

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



UDLAP®

Universidad de las Américas Puebla

*Technology, Science and Culture: A Global
Vision, Volume IV*

IntechOpen

Technology, Science and Culture: A Global Vision, Volume IV 2021

Editors

Luis Ricardo Hernández
Martín Alejandro Serrano Meneses

Knowledge area co-editors

Aura Matilde Jiménez Garduño
Nelly Ramírez Corona
José Luis Sánchez Salas
Enrique Ajuria Ibarra
Roberto Rosas Romero

We continue with this series by discussing research topics related to Food Science, Intelligent Systems, Molecular Biomedicine, and Water Science. We aim to discuss the newest issues, theories, and research methods in each field, promote debates among top researchers and graduate students and generate collaborative works among them.

The interactions of recognized specialists in each field and graduate students, through different meetings, generated exciting discussions, which are presented in this book. Thus, Dr. Eusebio Juaristi, from the Department of Chemistry of the Center for Research and Advanced Studies of the Instituto Politécnico Nacional, contributes with the article titled “The 2021 Nobel Prize in Chemistry. Relevance of asymmetric organocatalysis and green chemistry”. Dr. Toby P. Breckon, from the Department of Computer Science, Durham University, Durham, UK, with the article “Evaluating Convolutional Neural Networks for Prohibited Item Detection Using Real and Synthetically Compositated X-ray Imagery”. From the School of Chemical Sciences and DCU Water Institute, Dr. Fiona Regan wrote the article “Water quality monitoring using innovative technologies”. Dr. Encarna Gómez-Plaza, from the University of Murcia, contributes with the article “ANTHOCYANINS and wine color: A complex story”. Finally, graduate students of the Universidad de las Américas Puebla further present their key findings in a series of articles.

We believe that interactions between students and high-level researchers of different areas contribute to creating multidisciplinary points of view generating the advancement of science.



United Nations
Educational, Scientific and
Cultural Organization



• UNESCO Chair on
• Hydrometeorological Risks,
• Universidad de las Américas Puebla

The number and impact of water-related natural disasters has increased since the middle of last century. As result of increased climate variability and the effects of global warming, the hydrometeorological risk has increased and spread, while the resilience of societies, in many cases, is not adequate. Consequently, the risk has increased. Floods and droughts, particularly in a changing climate, require greater understanding to generate better forecasts and proper management of these phenomena. Mexico, like other countries in the world, and of course in Latin America and the Caribbean region, suffers from both weather extremes.

The UNESCO Chair on Hydrometeorological Risks, held at the University of Americas Puebla, is devoted to the analysis, measurement, modelling and management of extreme hydro-meteorological events in the context of a more urbanized world, climate change and further vulnerable regions. Focused on the development of basic and applied research for the design of adaptation and mitigation measures, dissemination and preparation of decision makers as well as the public. In its activities keeps a gender focus, directed in particular to reduce the vulnerability of women to hydrometeorological disasters.

The Chair acts in the following fields:

1. Hydrometeorological risks and climate change.
2. Modelling and forecasting of hydrometeorological risks.
3. Integrated management of hydrometeorological risks.
4. Gender and hydrometeorological risks.

A detailed description of the UNESCO Chair on Hydrometeorological Risks, members and publications, can be obtained at its Website <https://www.udlap.mx/catedraunesco/>

The Chair publish a quarterly Newsletter, in Spanish and English, that can be consulted at <https://www.udlap.mx/catedraunesco/newsletters.aspx>



UDLAP®

HIGHLY SPECIALIZED DOCTORAL PROGRAMS

Doctoral Programs in: (Doesn't require MA degree for entry)

- Molecular Biomedicine
- Food Science
- Water Sciences
- Intelligent Systems

Doctoral Program in: (Requires MA degree for entry)

- Creation and Culture Theory

**All accepted students have a 100% tuition scholarship
and a monthly maintenance for the entire duration of the
program.**

For more information contact:

Research and Graduate Studies Office

Tel: +52 (222) 2 29 20 00 exts. 2725 y 6005

informes.doctorados@udlap.mx

www.udlap.mx

ACADEMIC EXCELLENCE

CHOOSE UDLAP

Contents

The 2021 Nobel Prize in Chemistry: Relevance of Asymmetric Organocatalysis and Green Chemistry <i>Eusebio Juaristi</i>	1
Evaluating Convolutional Neural Networks for Prohibited Item Detection Using Real and Synthetically Composited X-Ray Imagery <i>Neelanjana Bhowmik, Qian Wang, Yona Falinie A. Gaus, Marcin Szarek and Toby P. Breckon</i>	15
Water Quality Monitoring Using Innovative Technologies <i>Fiona Regan</i>	29
Anthocyanins and Wine Color: A Complex Story <i>Encarna Gómez-Plaza</i>	47
Neurodegeneration Triggered by Inhaled Nanopollutants <i>Cristina Hermosillo-Abundis, Miguel A. Méndez-Rojas and Oscar Arias-Carrión</i>	65
A Systematic Review of the Pharmacological Potential of Disintegrins on Breast Cancer <i>Araliz López-Pintor and Irene Vergara-Bahena</i>	75
Shedding Light on Four Selected Flavonoids with Anti-non-small Cell Lung Cancer Properties <i>Jorge L. Mejía-Méndez, Horacio Bach, Luis Ricardo Hernández and Eugenio Sánchez-Arreola</i>	87
Na _v 1.7 Sodium Channel: A Potential Analgesic Target <i>Mildred López-Vázquez</i>	109
Detection and Classification of Burnt Skin on Images with Sparse Representation of Image Patches and Dictionaries <i>Brenda Rangel-Olvera and Roberto Rosas-Romero</i>	121
Random Wavelet Coefficients Pooling for Convolutional Neural Networks <i>Daniel Trevino-Sanchez and Vicente Alarcón-Aquino</i>	135
Autoethnographical Contributions in the Construction of Feminist Knowledge since the Lockdown <i>Alejandra Acevedo Placeres</i>	153
Sueño en Otro Idioma: Queer Identities in Contemporary Mexican Melodrama <i>Ilia Lizbeth Espinosa Pacheco</i>	163
New Masculinities, Are They Truly New? <i>Jeaqueline Flores Alvarez</i>	175

A Hundred Stitches Make a Canvas: How the Practice of Embroidery Relates to Forms of Life in Poverty in Atempa Coyomeapan <i>José Mariano Amador</i>	187
Microplastics Environmental Risk Assessment: A Review <i>Andrea Arredondo-Navarro, Estefanía Martínez-Tavera and Deborah Xanat Cervantes Flores</i>	199
The Roll of Different Kind of Fungi to Eliminate Lignin and Organochlorines: A Review <i>Anaid Bautista-Guerrero and Jose Luis Sanchez-Salas</i>	217
The Usefulness of System Dynamics for Groundwater Management <i>David Eduardo Guevara-Polo and Carlos Patiño-Gómez</i>	229
Influences of ENSO, AMO, and PDO over Dry Periods in Mexico: A Systematic Review (1984–2020) <i>Regina Mijares-Fajardo and Polioptro F. Martínez-Austria</i>	251
Reviewing Composite Water and Ecosystem Indexes to Envision Indicators with an Ecohydrology Framework <i>Maria Pina and Carlos Patino-Gomez</i>	275
Technologies and Materials for Sulfate Removal in Water: A Review <i>Ariel Antonio Quintana-Baquedano, Jose Luis Sanchez-Salas and Deborah Xanat Flores-Cervantes</i>	289
Weibull Modeling of Microorganisms' Survival Kinetics by High Intensity Light Pulses on Different Food Systems <i>Abril E. García-Santiesteban, Enrique Palou and María Teresa Jiménez-Munguía</i>	303
White Maize Tortillas Fortified with Brown Algae <i>Macrocystis pyrifera</i> <i>Alexa Pérez-Alva, Diana K. Baigts-Allende and Milena M. Ramírez-Rodrigues</i>	315
Artificial Intelligence for Sensory Acceptability Prediction <i>Jorge Metri, Nelly Ramírez, Milena Ramírez and Diana Baigts</i>	329
Modeling of Drying Kinetics of Fresh and Osmodehydrated Apples during Convective Drying <i>Julio Emmanuel González-Pérez, Nelly Ramírez-Corona and Aurelio López-Malo</i>	345
Deep Eutectic Solvents: A Promising Technique to Anthocyanin Extraction for Food Coloring Agents <i>Oscar Jiménez-González, Nelly Ramirez-Corona and José Á. Guerrero-Beltrán</i>	361
Methods of Application of Essential Oils in Foods and Their Effects on Sensory Attributes <i>Scarlette L. Recio-Cázares, Aurelio López-Malo and Enrique Palou</i>	379

The 2021 Nobel Prize in Chemistry: Relevance of Asymmetric Organocatalysis and Green Chemistry

Eusebio Juaristi

Abstract

Chemistry is not a stagnant scientific discipline; on the contrary, chemistry is a continuously advancing activity. Indeed, the XXI century is registering the emergence of several highly relevant new chemical developments that are already having an impact in the way both academic institutions and chemical enterprise are working. Two innovative concepts are particularly relevant, (1) asymmetric organocatalysis and (2) green chemistry. Indeed, the 2021 Nobel Prize in Chemistry was bestowed on German chemist Benjamin List and Scottish chemist David MacMillan for their demonstration that small organic molecules such as chiral amino acids and heterocycles are capable to catalyze the enantioselective formation of carbon-carbon bonds in the absence of potentially toxic transition metals. The other field that is presently most appreciated is that of the so-called green chemistry, which concerns the elimination of chemical substances and procedures that could be harmful to animals, humans and/or the environment. Former Nobel Prize in Chemistry winner Ryōji Noyori has stated that the survival of humankind depends on the ability of chemists to master the principles of green chemistry. In this regard, during the past 15 years the author's own research has contributed to the development of more sustainable asymmetric organocatalysis.

Keywords: asymmetric organocatalysis, green chemistry, sustainable chemistry, (S)-Proline, stereoselective reactions

1. Introduction

The 2021 Nobel Prize in Chemistry was awarded to Drs. Benjamin List and David MacMillan in recognition of their relevant contributions in the field of asymmetric organocatalysis. **Figure 1** shows photographs of Benjamin List and David MacMillan when they were the age at which they carried out the research work that eventually led to the award of the Nobel Prize, in the year 2000.

Benjamin List was born and raised in Frankfurt, Germany, and studied organic chemistry at the Free University of Berlin, in the group of Professor Johann Mulzer,



Figure 1.

Benjamin List (left) (https://es.wikipedia.org/wiki/Benjamin_List) and David W. C. MacMillan (right) (https://es.wikipedia.org/wiki/David_MacMillan).

with whom he worked for the conformation of his doctoral thesis on the synthesis of vitamin B12, a molecule with the minimum formula $C_{72}H_{100}CoN_{18}O_{17}P$. This ambitious project requires the construction – piece by piece—of the target molecule, which involves a large consumption of time, chemical reagents and solvents. List then realized that for the realization of the proposed goal it would be much more efficient to carry out the reactions involved with the help of catalysts similar to enzymes that very selectively catalyze the chemical processes that occur in nature.

In a fortunate coincidence, in the late 1980s the research group of Dr. Richard Lerner, a leading researcher at the Scripps Institute in California, United States, was already exploring the possibility of using antibodies produced by the immune system of humans and animals as stereoselective catalysts. Thus, with a grant from the Alexander von Humboldt Foundation, List had the opportunity to collaborate with Lerner, who later promoted his hiring as an independent researcher at the Scripps Institute. List immediately set about examining small organic molecules that could function as enantioselective catalysts.

For his part, David W. C. MacMillan is of Scottish origin, having been born in the town of Bellshill, near Glasgow, where he carried out his professional studies in the field of chemistry, to later carry out doctoral studies at the University of California—Irvine, under the tutelage of Larry Overman. During his doctoral studies MacMillan devoted himself to the asymmetric synthesis of natural products. His interest in asymmetric organic synthesis took him to Harvard in 1996, where he was a postdoctoral research associate in the group of Professor David A. Evans.

MacMillan began his career as an independent researcher at the Department of Chemistry at the University of California—Berkeley in 1998. Two years later, in 2000, he published the seminal work where he demonstrated the potential of chiral organic molecules to perform asymmetric catalysis. At that time, he was offered tenure at Berkeley, but MacMillan decided to accept another offer that was presented to him at the time and moved to the California Institute of Technology (Caltech). In 2004 he moved with his research group to Princeton University, where he has been director of the Institute of Catalysis ever since.

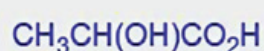
To understand the relevance of the projects addressed by List and MacMillan, it is convenient to review the concept of *chirality* of molecules that contain asymmetric tetravalent carbon. Thus, in the mid-nineteenth century, examples were known of molecules with the same *composition* (that is, the same type and number of constituent atoms) and the same *connectivity* (that is, the same way in which such atoms are linked together), and yet they are not the same. An illustrative example is α -hydroxypropionic acid (lactic acid) isolated by Scheele from milk in

Configuration (chirality)

Dilemma: molecules with the same composition and connectivity but different!



≠



$$[\alpha] = -26^\circ$$

$$[\alpha] = +26^\circ$$

(Scheele, 1780, from milk)

(Berzelius, 1807, from muscular tissues)

And with different biological effect !

Figure 2.

α-Hydroxypropionic acids as an example of molecules that have the same composition and connectivity and are nevertheless different, as can be seen from their different physical and biological activity: lactic acid with negative optical rotation is salty while α-hydroxypropionic acid isolated from muscle tissues such as the frog's leg is tasteless.

1780, which has the same $\text{C}_3\text{H}_6\text{O}_3$ composition and the same connectivity as the α-hydroxypropionic acid isolated from a frog's leg in 1807 by Berzelius, and which are different as it is shown by the values of the optical rotation of these compounds, which is of the same magnitude but different sign (**Figure 2**).

It would take several decades, until the year 1874, for this dilemma to be solved. Independently, chemists Joseph Achille Le Bel and Jacobus Henricus van't Hoff realized that it is very important to visualize molecules in three dimensions. Most relevantly, when four different substituents on the tetravalent carbon atom are oriented towards the vertices of a tetrahedron instead of being located on one shared plane, then the molecule will be asymmetric, as in the case of α-hydroxypropionic acid (**Figure 3**).

An analogy is the right and left hands, which have the same composition (the same pinkie, index, thumb, etc.) and the same connectivity (that is, the way they are attached to the palm of the hand), but actually they are not superimposable mirror images (**Figure 3**). Indeed, a right hand “fits” into the right glove but not into the left glove. Likewise, the left hand only fits into the complementary left glove. The fact that hand in Greek is called “*quir*” gives rise to the concepts “chirality”, “chiral molecule” and “center of chirality” [1].

Thus, molecules that are non-superimposable mirror images are isomers called *enantiomers*, which have the same physical (e.g., melting point) and chemical (i.e., reactivity) properties except when interacting with other chiral agents. In particular, the different taste of enantiomeric α-hydroxypropionic acids is a consequence of their different interaction with the receptor for the sense of taste, made up of chiral biomolecules containing α-amino acids, which are also chiral.

With the proposed three-dimensional model advanced by Le Bel and van't Hoff originates the area of *stereochemistry*, which refers to the study of *chemistry in three dimensions*. The contribution of Le Bel and van't Hoff may seem trivial now, but this is not so since the essential change in the way one should consider molecules, taking

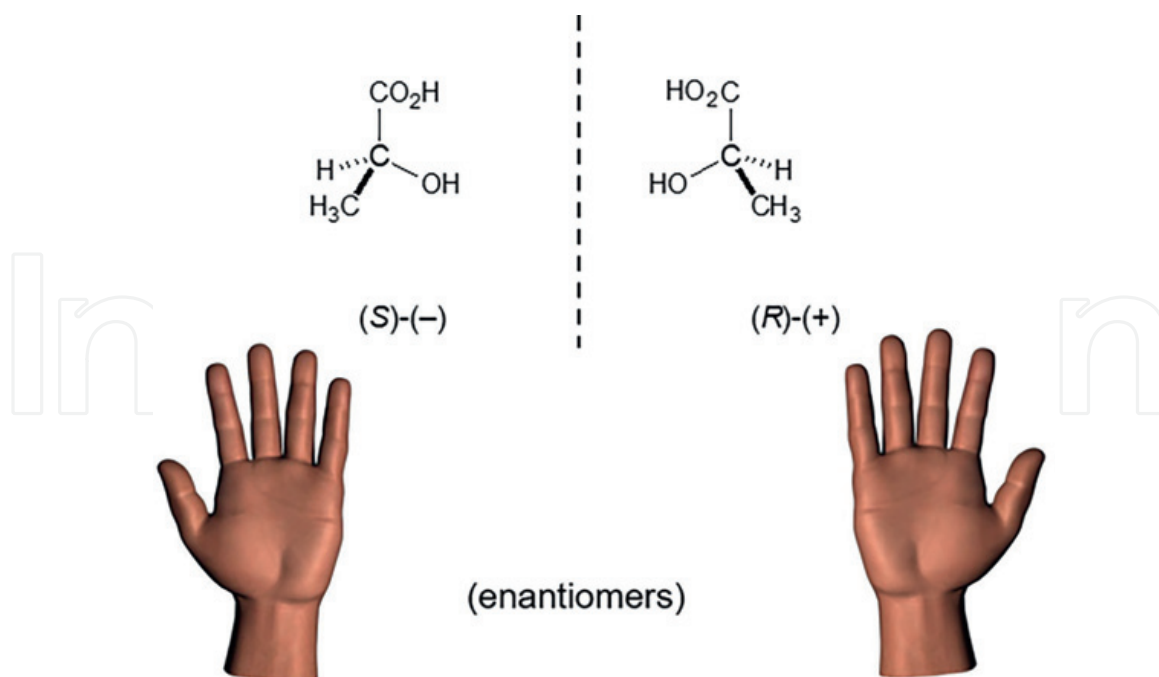


Figure 3.

Molecules that have a carbon atom with four different substituents, as is the case of α -hydroxypropionic acid, are asymmetric so that their mirror image is not the same, that is, they are not superimposable. Thus, the molecule and its mirror image are isomeric, and are called enantiomeric.

into account the spatial arrangement of their constituent atoms, required of inspiration and genius!

Another term that we must introduce here is “*stereogenic center*”, that is, a carbon atom that due to its asymmetry gives rise to the stereochemistry present in chiral molecules. A more complete discussion of these concepts can be found in reference 1.

The fact that enantiomeric molecules present different reactivity to reagents or biological receptors that are also chiral can have dramatic consequences in our daily lives. **Figure 4** shows two examples of chiral molecules that give rise to a desirable pharmacological effect in one specific configuration but provoke an undesirable effect in the opposite configuration. Thus, the (*S*) configured 1-chloro-2,3-propanediol is a valuable therapeutic agent while its (*R*) enantiomer is in fact a poison (*distomer*, from ‘*dis*’ = bad; that is, the toxic enantiomer).

In one illustrative case that gave rise to the strongest arguments supporting the relevance of the enantioselective synthesis of chiral drugs, that is of *asymmetric synthesis* in which the objective is to synthesize exclusively the enantiomer with the properties of interest (*eutomer*, from ‘*eus*’ = good, that is, the beneficial enantiomer) Thalidomide was administered in racemic form (that is, as a 50:50 mixture of its two enantiomers) in the 1960s for its effectiveness as a tranquilizing agent, without suspecting that the (*S*) enantiomer is a teratogenic compound (**Figure 4**). Indeed, even today, 60 years after the tragedy, there live affected people who were born without arms or legs as a result of the teratogenic activity of the bad enantiomer, the *distomer*.

In summary, in chiral bioactive molecules, as consequence of their stereospecific interaction with chiral bioreceptors, the biological activity is determined by their configuration. Thus, the control of enantiomeric purity is essential in the pharmaceutical, agrochemical, food and perfume industries, among others.

In this context, **Table 1** brings together a series of chiral drugs that were initially marketed as racemic mixtures but which, in the face of observations such as those

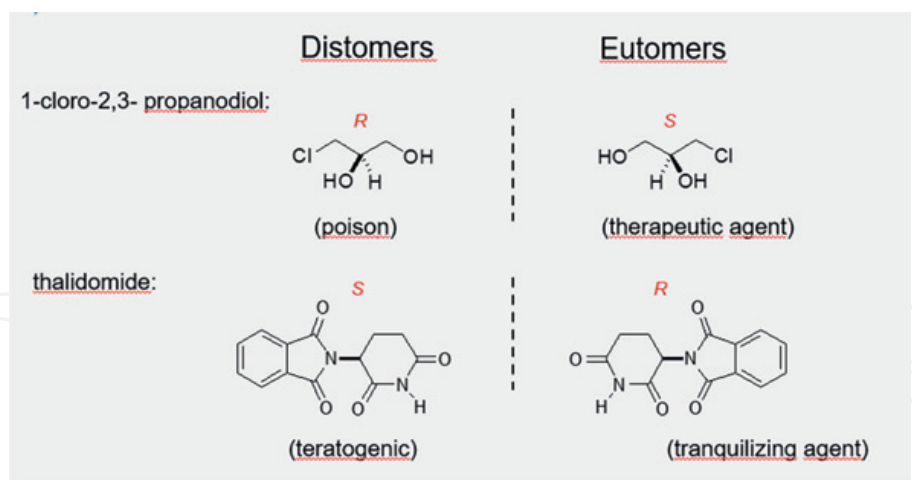


Figure 4.
Importance of asymmetric synthesis. Two illustrative examples of chiral molecules with very different pharmacological properties depending on their configuration.

mentioned above, required the development of asymmetric syntheses to produce exclusively the desired enantiomer (the eutomer) at an industrial level. This kind of development in the pharmaceutical industry is known as a “chiral switch” and it has enormous commercial value. For example, for decades the Mexican company Syntex distributed the anti-inflammatory agent Naproxen with great commercial success, but when the patent that protected it expired, Syntex was not able to develop the most efficient process for its asymmetric synthesis, which other pharmaceutical companies did. As consequence, the company’s market shares lost value and it eventually had to be sold to the Swiss company H. La Roche, which dismantled the existing research laboratories in Mexico. Unfortunately, most of the chemists who worked in these laboratories had to move to the United States, which was a significant loss of talent for Mexico.

Asymmetric synthesis emerged as a consolidated discipline at the end of the twentieth century. There are various strategies to carry out an asymmetric synthesis,

Drug	Therapeutic activity
Amoxicillin	antibiotic
Enalapril	antihypertensive
Ampicillin	antibiotic
Captopril	antihypertensive
Pravastatin	antihypercholesterolemic
Diltiazem	antihypertensive
Ibuprofen	anti-inflammatory
Lovastatin	antihypercholesterolemic
Naproxen	anti-inflammatory
Fluoxetine	antidepressant

Table 1.
Some chiral drugs that were initially marketed as racemic mixture but which later were produced as pure enantiomers via asymmetric synthesis.

such as, for example, via the so-called “*chiral pool*” or using chiral auxiliaries, which are incorporated into “prochiral” molecules to induce the stereoselective formation of new chiral centers before removing (and if possible reusing) the chiral auxiliary [1].

On the other hand, at the end of the twentieth century the first examples of the application of chiral catalysts in asymmetric synthesis were reported. The great advantage of these catalytic processes is that a chiral molecule of the catalyst can generate thousands of new chiral molecules, that is, they afford a multiplying effect of chirality. Among the most successful chiral catalysts are those reported by William S. Knowles, Ryoji Noyori, and K. Barry Sharpless, who received the 2001 Nobel Prize in Chemistry for their contributions in catalytic asymmetric synthesis (**Figure 5**) [2].

A disadvantage of the catalysts developed by Knowles, Noyori and Sharpless is that they use transition metals such as rhodium, ruthenium and titanium, which pose a health risk. This is a risk that cannot be taken by pharmaceutical and food companies, in which the presence of potentially toxic metals cannot be tolerated. As we will see in later paragraphs, the 2021 Nobel Prize in Chemistry was awarded to Benjamin List and David MacMillan precisely for their fundamental contributions in the area of sustainable asymmetric organocatalysis, which proceeds in the absence of metals (**Figure 6**).

In this context, Nature has developed extremely selective and efficient catalysts; however, the complex and exquisite architecture exhibited by these biomolecules involves the presence of thousands of components such as the constituent amino acids in the proteins of the corresponding enzymes (**Figure 7**).

In contrast to the metal catalysts developed by Knowles, Noyori, and Sharpless, List (while at the Scripps Institute in California), and MacMillan (at that time in the University of California at Berkeley) set about to design small organic molecules that could act as catalysts. In particular, knowing that aldolase enzymes (that is, the enzymes that catalyze enantioselective aldol reactions like the one shown in **Figure 7**) contain amino acid residues in the active site, List recalled that the enantioselective annealing reaction that Hajos, Parrish, Eder, Sauer and Wiechert had reported in the

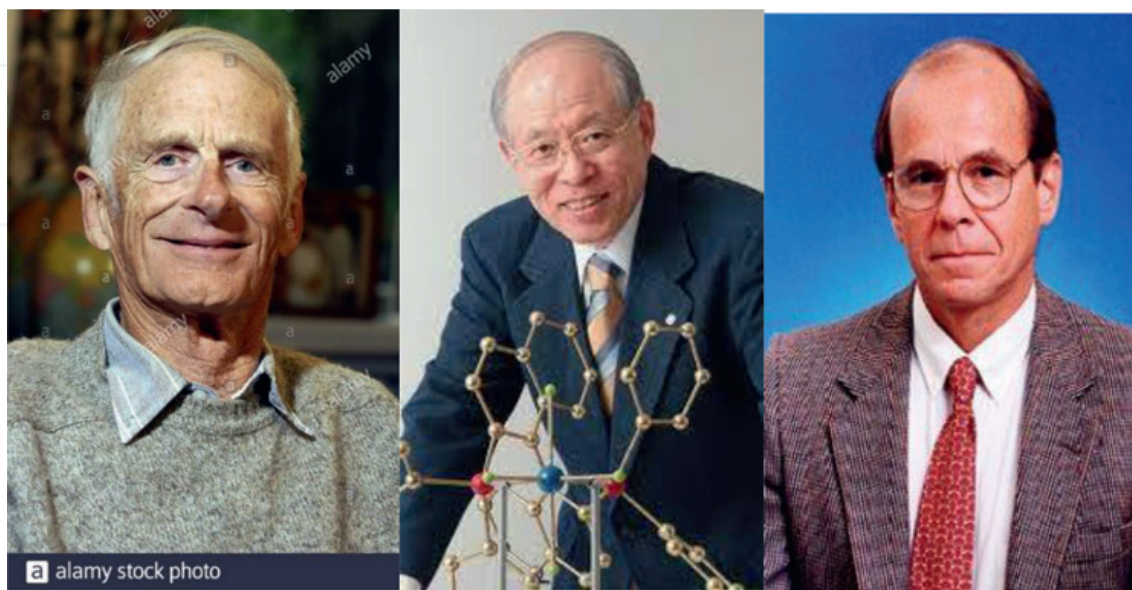


Figure 5.

The 2001 Nobel Prize in Chemistry was awarded (corresponding photographs arranged left to right) to William S. Knowles, Ryoji Noyori, and K. Barry Sharpless for developing asymmetric catalysis using metals coordinated to chiral organic ligands. (<https://www.nobelprize.org/prizes/chemistry/2001/summary/>).

Asymmetric organocatalysis



Figure 6.

Asymmetric organocatalysis is based on the use of small organic molecules, in substoichiometric quantities, for the activation of prochiral substrates in the absence of metals.

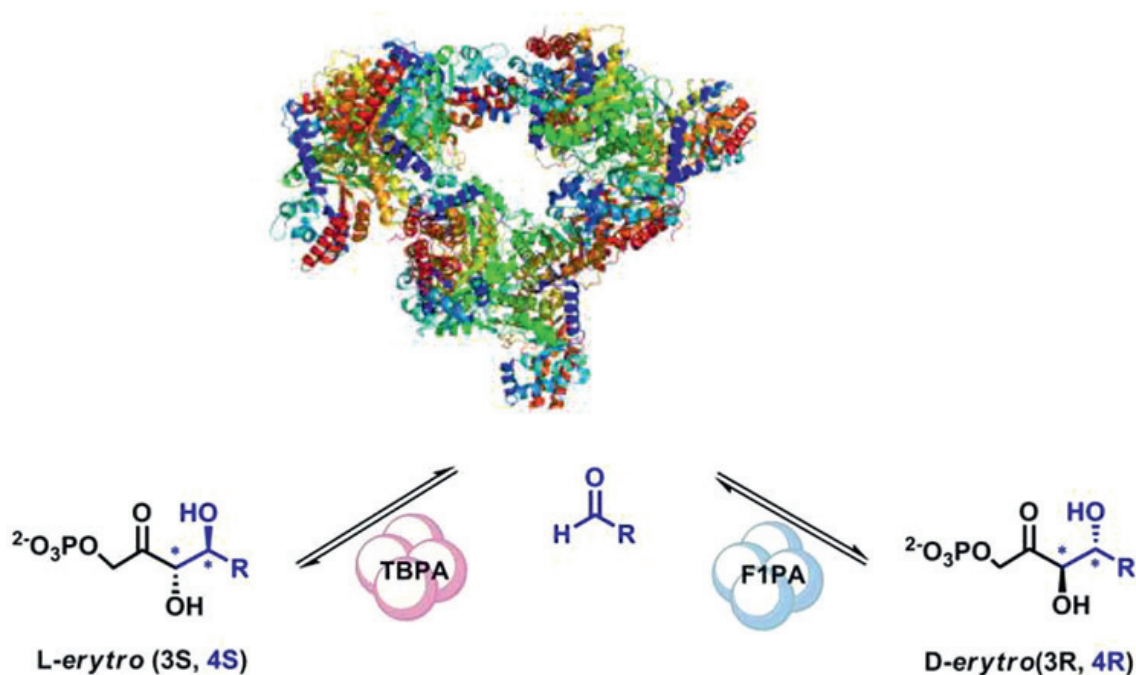


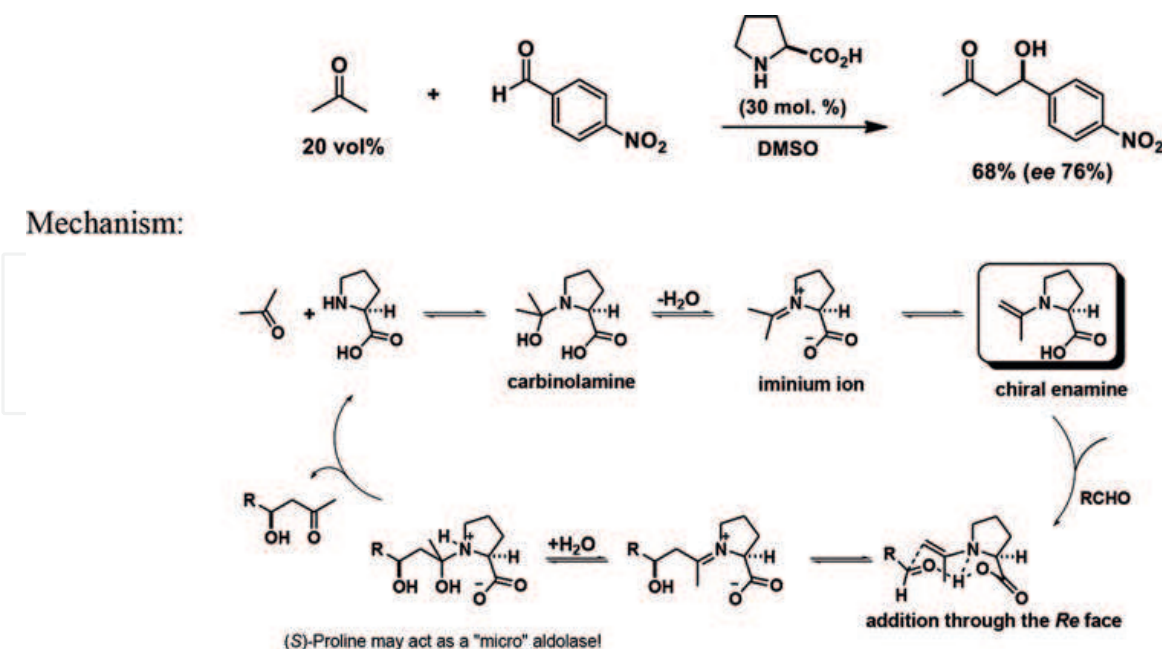
Figure 7.

Enzymes such as aldolases are capable of catalyzing enantioselective reactions with great efficiency.

1970s employed the amino acid (*S*)-proline as a catalyst. List then wondered whether (*S*)-proline can function as an enzyme and catalyze reactions by mechanisms similar to those used by aldolase enzymes.

And indeed, Benjamin List, Richard Lerner and Carlos Barbas demonstrated in the year 2000 that (*S*)-proline catalyzes the aldol condensation between acetone and benzaldehyde with 69% yield and an enantioselectivity of 76%, that is to say that the major enantiomer is obtained in 88% yield and the minor enantiomer in 12% yield (**Figure 8**) [3].

It should be noted that a phrase in the publication by List, Lerner and Barbas [3] “(*S*)-Proline acts as a micro-aldolase!” had a huge impact on the chemical community. Indeed, for the vast majority of chemists dedicated to organic synthesis, it is very attractive to be able to “compete” with Nature; specifically, the possibility of employing as enantioselective catalysts small chiral organic molecules that can exceed

**Figure 8.**

Asymmetric aldol reaction catalyzed by (*S*)-Proline. (*S*)-Proline acts as a “micro-aldolase”!.

the efficiency of natural biocatalysts (developed over millions of years of evolution) containing thousands of amino acids is very tempting.

Figure 8 also includes the accepted mechanism for the organocatalyzed enantioselective aldol reaction catalyzed by (*S*)-proline. First, the condensation of acetone with (*S*)-proline generates an iminium ion that rearranges itself to give a chiral enamine, which is the key intermediate in the process, since it can differentiate the prochiral faces of the aldehyde. Indeed, the addition of the enamine occurs predominantly on one of the carbonyl faces to give, after the hydrolysis of a new iminium ion, the major enantiomeric aldol product as well as the (*S*)-proline that can restart another catalytic cycle.

In the case of the work carried out by David MacMillan et al, a small chiral organic molecule was also used, a heterocyclic molecule called imidazolidinone. This chiral organic molecule readily condenses with unsaturated prochiral aldehydes to afford reactive chiral dienophiles that can then differentiate the prochiral faces of cyclopentadiene to give chiral bicyclic products in an enantioselective manner (**Figure 9**) [4].

Another topic of great relevance in the XXI century is the so-called “Green Chemistry”, which refers to the design of chemical products and processes that minimize or eliminate the production and handling of substances that are dangerous for the environment or for human health [5]. Indeed, it has been asserted that the survival of humanity depends on the ability of chemistry professionals to correctly apply the principles and goals of Green Chemistry [6].

In this context, my research group has been dedicated for the last 12 years to the development of methodologies that increase the sustainability of asymmetric organocatalysis [7, 8]. In particular, we have synthesized chiral dipeptides derived from (*S*)-proline that catalyze enantioselective aldol reactions in the absence of solvent by means of mechanochemical activation.

Figure 10 shows various types of mills used for grinding solid substances in the absence of solvent, from the mortar and pestle used since ancient times to the vibrational or planetary type mills used in chemistry laboratories.

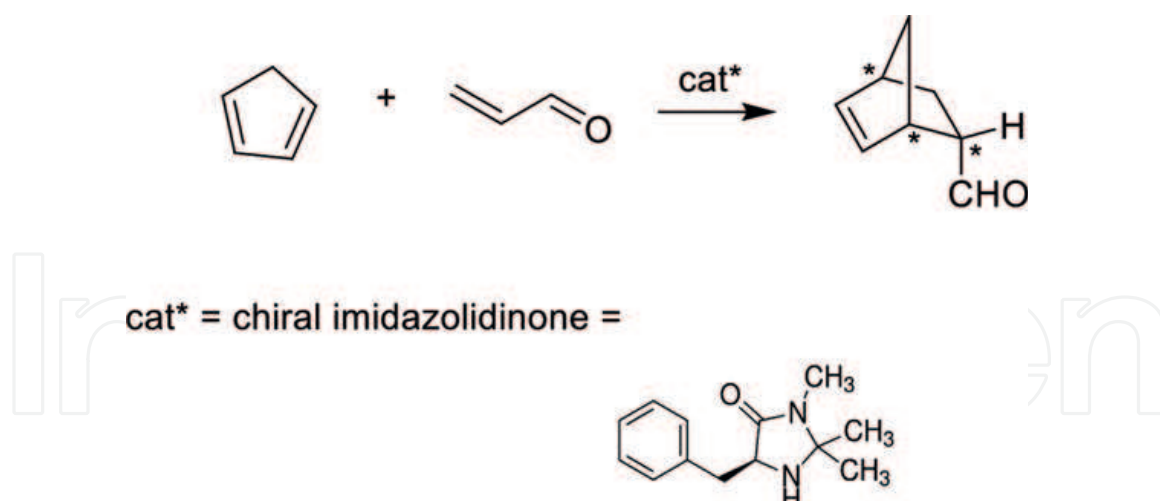


Figure 9.
David MacMillan et al. demonstrated that a chiral imidazolidinone can catalyze the enantioselective Diels-Alder reaction between α , β -unsaturated aldehydes and dienes [4].

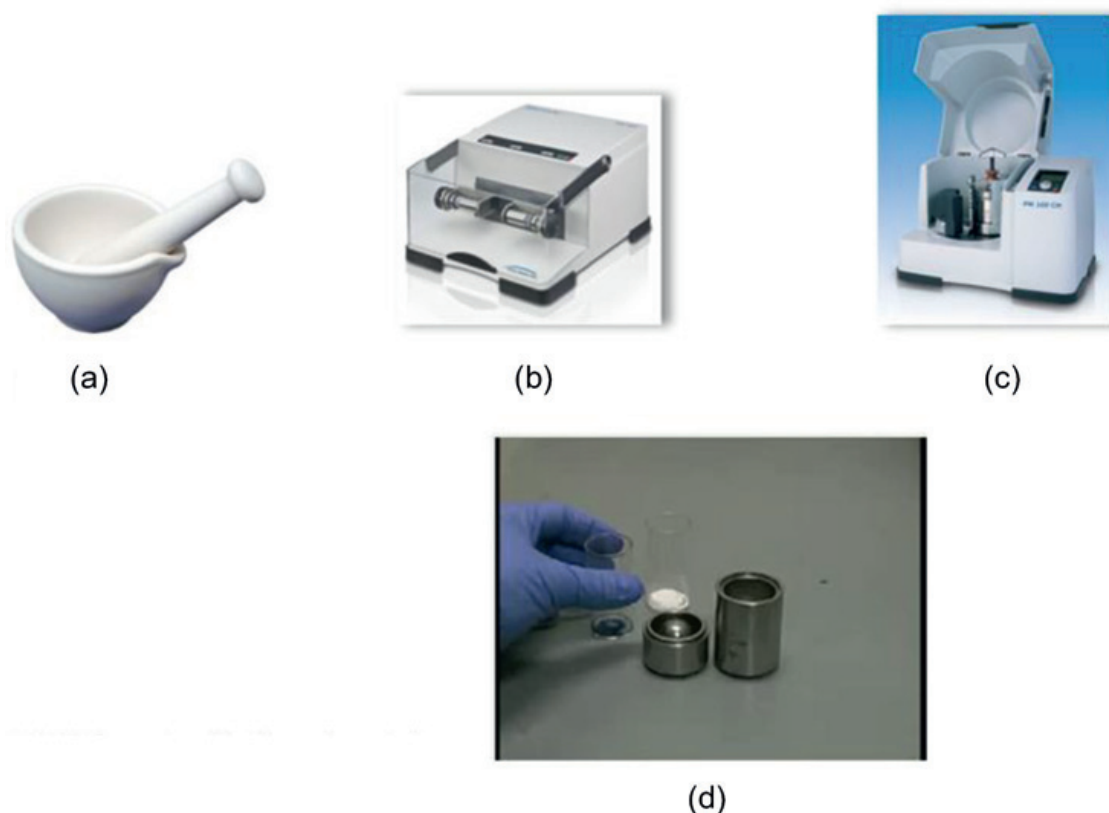


Figure 10.
(a–c): Some milling equipment used for the mechanochemical activation of laboratory-scale reactions. (d): Example of container used for grinding in the vibratory mill. (Photographs taken from reference [8].) (a) Mortar and pestle, (b) vibratory mill, (c) planetary mill, (d) stainless steel ball and container.

Figure 11 shows the aldol reaction between cyclohexanone and *p*-nitrobenzaldehyde catalyzed by the chiral dipeptide containing a (*S*)-proline residue and a phenylalanine residue. This reaction was carried out with mechanochemical activation (ball milling) in the absence of solvent. The reactions proceed with excellent yield and good diastereo- and enantioselectivity (95% enantiomeric excess, that is, the 1*S*,1'*R* enantiomeric product predominates over the 1*R*,1'*S* enantiomer in a ratio of 97.5 to 2.5) [9, 10].

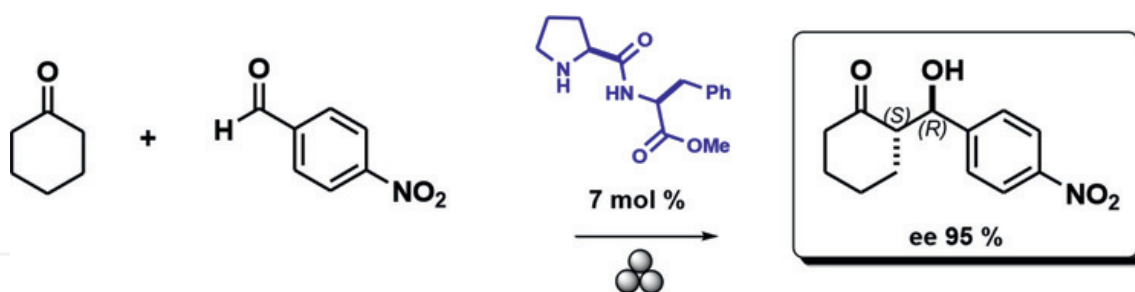


Figure 11. Asymmetric aldol reaction catalyzed by chiral dipeptides (*S*)-Proline-(*S*)-Phe in the absence of solvent [9, 10]. The symbol with three pellets below the arrow indicates that the reaction is activated by ball milling.

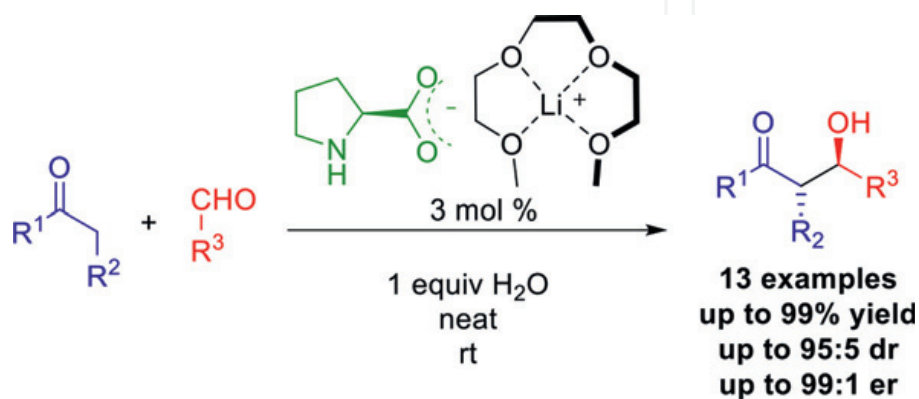


Figure 12. Improving the catalytic performance of (*S*)-Proline as an organocatalyst in asymmetric aldol reactions in the presence of Solvate Ionic Liquids: participation of a supramolecular aggregate [11].

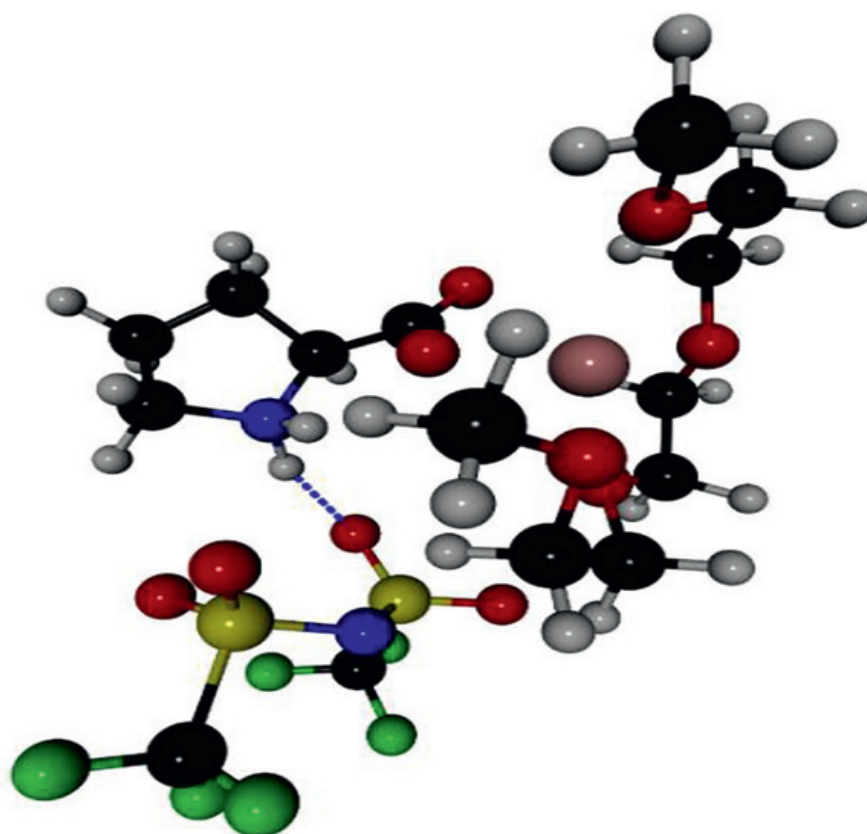


Figure 13. Supramolecular aggregate between (*S*)-proline and solvate ionic liquid.

More recently we found that the efficacy of (*S*)-proline as a chiral organocatalyst in asymmetric syntheses is markedly increased in the presence of so-called solvate ionic liquids, which are complexes of glymes with lithium salts (**Figure 12**) [11].

The evidence collected through theoretical calculations indicates the formation of an aggregate between (*S*)-proline and the solvate ionic liquid, that apparently leads to a sufficiently robust transition state that explains the observed high stereoselectivity (**Figure 13**).

2. Final comments

Like other scientific disciplines, chemistry is in full growth. The twenty-first century already registers a variety of relevant topics that present an extraordinary development. In this sense, the International Union of Pure and Applied Chemistry recently carried out an analysis of the most promising emerging fields of today (**Table 2**) [12]. Asymmetric organocatalysis and mechanochemistry stand out in this list, of which the first has just been recognized with the Nobel Prize in Chemistry 2021 awarded to Benjamin List and David MacMillan. During the last two decades diverse experimental protocols have made organocatalysis an even “greener” alternative by the use of friendlier reaction conditions, or via the application of solvent-free methodologies activated by mechanochemistry. Alternative strategies include the design and synthesis of more selective catalysts, or via the development of multicomponent one-pot organocatalytic reactions, or by the recycling and reuse of organocatalysts, or by means of the application of more energy-efficient activation techniques, among other approaches. It is reasonable to expect that Green Chemistry, including the field of mechanochemistry, will also be recognized by the Nobel Prize committee in the near future.

3D-bioprinting
Directed evolution of selective enzymes
Enantioselective organocatalysis
Flow chemistry
Metal Organic Frameworks for water harvesting
Nanopesticides
Plastics to monomers
Mechanochemistry
Reversible-deactivation of radical polymerization
Solid-state batteries

Table 2.
IUPAC announces the top ten emerging technologies in chemistry [12].

Acknowledgements

I am grateful to Dr. Claudia Gabriela Ávila-Ortiz for her expert assistance in the preparation of this document. I am also indebted to Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico) for financial support via grant A1-S-44097.

Conflict of interest

The author declares no conflict of interest.

Author details


Eusebio Juaristi^{1,2}

1 Departamento de Química, Centro de Investigación y de Estudios Avanzados,
Mexico City, Mexico

2 El Colegio Nacional, Mexico City, Mexico

*Address all correspondence to: ejarist@cinvestav.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Juaristi E. Introducción a la Estereoquímica y al Análisis Conformacional. Ciudad de México, México: El Colegio Nacional; 2007
- [2] Juaristi E. Premio Nobel de Química 2001: La Importancia de la Síntesis Asimétrica. Educación Química. 2002;**13**:6-7
- [3] List B, Lerner AR, Barbas CF III. Proline-Catalyzed Direct Asymmetric Aldol Reactions. Journal of the American Chemical Society. 2000;**122**:2395-2396
- [4] MacMillan DWC et al. New strategies for organic catalysis: The first highly enantioselective organocatalytic Diels–Alder reaction. Journal of the American Chemical Society. 2000;**122**:4243-4244
- [5] Anastas P, Eghbali N. Green chemistry: Principles and practice. Chemical Society Reviews. 2010;**39**:301-312
- [6] Noyori R. Green chemistry: The key to our future. Tetrahedron. 2010;**66**:1028
- [7] Hernández JG, Juaristi E. Recent efforts directed to the development of more sustainable asymmetric organocatalysis. Chemical Communications. 2012;**48**:5396-5409
- [8] Pérez-Venegas M, Juaristi E. Mechanoenzymology: State of the art and challenges towards highly sustainable biocatalysis. ChemSusChem. 2021;**14**:2682-2688
- [9] Hernández JG, Juaristi E. Asymmetric Aldol reaction organocatalyzed by (S)-proline-containing dipeptides: Improved stereoinduction under solvent-free conditions. The Journal of Organic Chemistry. 2011;**76**:1464-1467
- [10] Hernández JG, García-López V, Juaristi E. Solvent-free asymmetric Aldol reaction organocatalyzed by (S)-proline-containing thiodipeptides under ball-milling conditions. Tetrahedron. 2012;**68**:92-97
- [11] Obregón-Zúñiga A, Milán M, Juaristi E. Improving the catalytic performance of (S)-proline as organocatalyst in asymmetric Aldol reactions in the presence of solvate ionic liquids. involvement of a supramolecular aggregate. Organic Letters. 2017;**19**:1108-1111
- [12] Gomollón-Bel F. Ten chemical innovations that will change our world. Chemistry International. 2019;**41**:12-17

Evaluating Convolutional Neural Networks for Prohibited Item Detection Using Real and Synthetically Composited X-Ray Imagery

*Neelanjan Bhowmik, Qian Wang, Yona Falinie A. Gaus,
Marcin Szarek and Toby P. Breckon*

Abstract

Detecting prohibited items in X-ray security imagery is pivotal in maintaining border and transport security against a wide range of threat profiles. Convolutional neural networks (CNN) with the support of a significant volume of data have brought advancement in such automated prohibited object detection and classification. However, collating such large volumes of X-ray security imagery remains a significant challenge. This work opens up the possibility of using synthetically composed imagery, avoiding the need to collate such large volumes of hand-annotated real-world imagery. Here, we investigate the difference in detection performance achieved using real and synthetic X-ray training imagery for CNN architecture detecting three exemplar prohibited items, {*Firearm*, *Firearm Parts*, *Knives*}, within cluttered and complex X-ray security baggage imagery. We achieve 0.88 of mean average precision (mAP) with a Faster R-CNN and ResNet101 CNN architecture for this three-class object detection using real X-ray imagery. While the performance is comparable with synthetically composited X-ray imagery (0.78 mAP), our extended evaluation demonstrates both challenge and promise of using synthetically composed images to diversify the X-ray security training imagery for automated detection algorithm training.

Keywords: object detection, convolutional neural networks, X-ray images, airport security

1. Introduction

To ensure transport and border security, X-ray security screening is commonplace within public transport and border security installations such as airports, railway and metro stations. However, due to the nature of cluttered and complex X-ray imagery

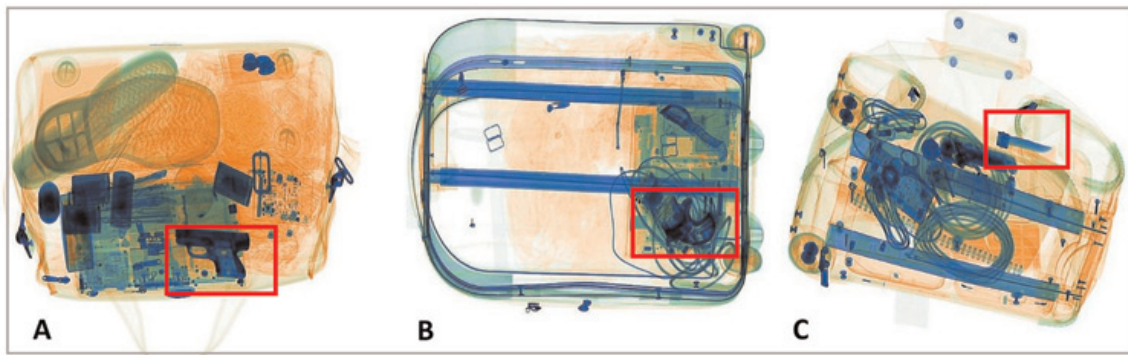


Figure 1. Exemplar X-ray security baggage images with prohibited objects—red box: (A) Firearm (B) Firearm Parts and (C) Knife.

(**Figure 1**), the process of X-ray screening is complicated by tightly packed items within baggage making it challenging and time-consuming to identify the presence of prohibited items. With the natural occurrence of such prohibited items being rare, previous studies cite time constraints as a major factor in the performance limitations of human operators for this screening task [1, 2].

Whilst challenging for a human, a reliable automatic prohibited item detection system may assist in improving the performance and throughput of such screening processes [3]. To date, contemporary X-ray security scanners already implement material discrimination via dual-energy multiple view X-ray imagery to enable threat material detection [4]. This use of dual-energy X-ray gives rise to the false-colour mapped appearance of X-ray security imagery (e.g., metals, alloy or hard plastic are shown in blue while less dense objects are shown in green/orange—see **Figure 1**).

Convolutional Neural Network (CNN) based methods have proven effective in detecting a wide range of object classes within this context [5–8]. However, the performance of such object detection approaches is heavily reliant on the availability of a substantial volume of labelled X-ray imagery. Unfortunately, the availability of such X-ray imagery datasets suitable for training CNN architectures is limited and also restricted in size and item coverage (e.g. GDXray [9], SIXray [10]).

Commonly, it is challenging to collect sufficient X-ray imagery containing example of prohibited items with large variations in pose, scale and item construction. To overcome this challenge, contemporary data augmentation schemes such as image translation, rotation, flipping and re-scaling are applied to enlarge the availability of otherwise limited training datasets [5]. However, such methods suffer from the fact that the resulting augmented dataset still lacks diversity in terms of prohibited item variation and inter-occlusion emplacement within complex and cluttered X-ray security imagery. This motivates the use of synthetically composed imagery, where such imagery readily enables the introduction of more variability in pose, scale and prohibited item usage in an efficient and readily available way.

In this work, we devise a Synthetically Composed (SC) data augmentation approach via the use of Threat Image Projection (TIP). TIP is an established process within operational aviation security for the monitoring of human operators which uses a smaller collection of X-ray imagery comprising of isolated prohibited objects (only), which are subsequently superimposed onto more readily available benign X-ray security imagery. Here this approach additionally facilitates the generation of synthetic, yet realistic prohibited X-ray security imagery for the purpose of CNN training. Our key contributions are the following: (a) the synthesis of high quality

prohibited images from benign X-ray imagery using a documented TIP approach and (b) an extended comparative evaluation on how real and synthetically generated X-ray imagery impacts the performance for prohibited object detection and classification using CNN architectures.

2. Related work

Traditional computer vision methods that rely on handcrafted features have been applied to prohibited item detection in X-ray security imagery such as Bag of Visual Words (BoVW) [3, 11, 12] and sparse representations [13]. However, the recent advancement in CNN have drawn more attention to prohibited item detection due to significant performance gains within X-ray security imagery [14–16]. The works of [14, 16] compare handcrafted features with a BoVW based sparse representation to CNN features. These shows that such deep CNN features achieve superior performance with more than 95% accuracy for prohibited item detection. The study of [14] exhaustively compares various CNN architectures to evaluate the impact of network complexity on overall performance. Fine tuning the entire network architecture for this problem domain yields 0.99% true positive, 0.01% false positive and 0.994% accuracy for generalized prohibited item detection [14].

Further work on prohibited item under X-ray security imagery is undertaken by Mery et al. [11], where regions of interest detection is performed across multiple views of the object. Subsequently, the candidate region obtained from an earlier segmentation step is then matched based on their similarity. This achieves 94.3% true positive and 5.6% false positive across multiple view X-ray security imagery. The work of [15] examines the relative performance of traditional sliding window driven CNN detection model based on [14] against contemporary region-based and single forward-pass based CNN variants such as Faster R-CNN [8], R-FCN [17], and YOLOv2 [18], achieving a maximal 0.88 and 0.97 mAP over 6-class object detection and 2-class firearm detection problems respectively.

To investigate the generalised applicability of CNN within X-ray security imagery, large X-ray imagery datasets are required. Existing public domain datasets such as GDXray [9] contains three major categories of prohibited items, {*Guns, Shurikens, Razor blades*}. However, images in GDXray are provided with lesser clutter and overlap making object detection less challenging than in typical operational conditions. By contrast, the SIXray dataset [10] contains six classes, {*Guns, Knives, Wrenches, Pliers, Scissors, Hammers*}, from cluttered operational imagery. This provides more inter-occluding imagery examples but at the same time provides significantly fewer prohibited item than benign samples akin to an operational (real-world) scenario, where the presence of prohibited items is low within stream-of-commerce (largely benign) X-ray security imagery.

To overcome the limited dataset availability, data augmentation has been used to increase overall dataset diversity. Whilst simple image data augmentation strategies such as translations, flipping and scaling do increase geometric diversity of the imagery they do not increase the appearance or content diversity of the dataset itself [19]. The work of [20] alternatively attempts data augmentation based on an Generative Adversarial Network (GAN) approach but generates synthetic prohibited items in isolation rather than within a full cluttered X-ray security image. By contrast, the work of [21] utilises an approach, similar to the concept of TIP, whereby a prohibited item is superimposed into X-ray security imagery. Therefore, in this work, we explore

the feasibility of TIP as a data augmentation strategy to support performance enhancement and evaluation of contemporary deep CNN architectures within the context of prohibited item detection in security X-ray imagery.

3. Proposed approach

We investigate the use of a full TIP pipeline, based on prior work in the field [22–24], to generate a range of appearance and contents based dataset variation (Section 3.1). Subsequently, CNN object detection architecture is used to evaluate the TIP based data augmentation approach and compare the performance with real X-ray security imagery (Section 3.2)

3.1 Synthetic X-ray security imagery via TIP

Our TIP pipeline consists of three components: *threat image transformation*, *insertion position determination* and *image compositing* as illustrated in **Figure 2**.

We use threat (prohibited item) images containing clean, isolated object signatures which can be easily segmented from their plain background via simple thresholding. To diversify the resultant synthetic images, we apply *threat image transformation* via rotating the threat signature by a random angle θ . Although other threat image transformation strategies (e.g., noise, illumination, magnification, etc.) have been explored in [25], our work focuses on the pure combination of our segmented threat signature and a benign X-ray security image, isolating the effects of other data augmentation techniques. We denote this transformed threat image as I_s and the i -th row, j -th column pixel as $I_s(i, j)$.

A valid insertion position within the bag image is determined based on the bag region and the shape of threat signature.

Given a bag image I_t , we use morphological operations to extract the bag region. Specifically, the original bag image is firstly binarised by thresholding (**Figure 3b**) to extract the foreground (target) region for insertion. Due to noise, a simple thresholding process cannot ideally separate background and foreground. We sequentially apply a series of appropriately parameterised morphological operations including dilation (**Figure 3c**), hole filling (**Figure 3d**) and erosion (**Figure 3e**) to identify the largest connected image region as the target for insertion (see **Figure 3f**). Obviously, a valid insertion of the threat signature has to guarantee the threat

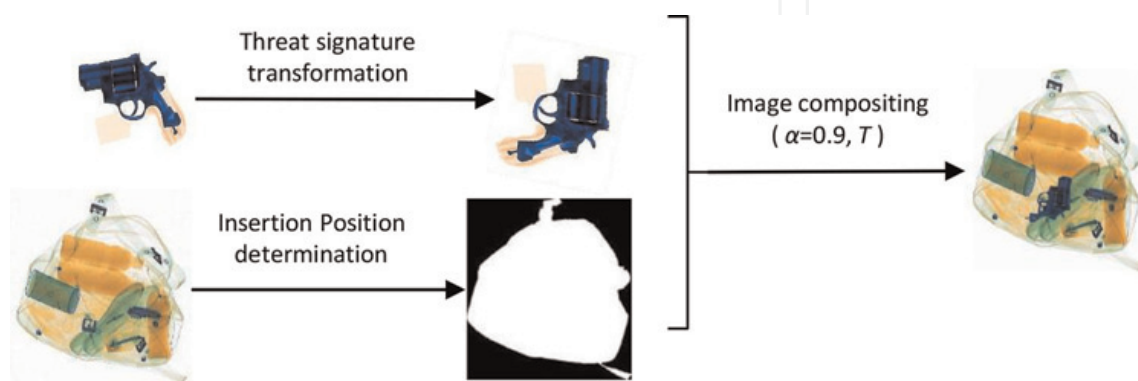


Figure 2.
Threat image projection (TIP) pipeline for synthetically composited image generation.

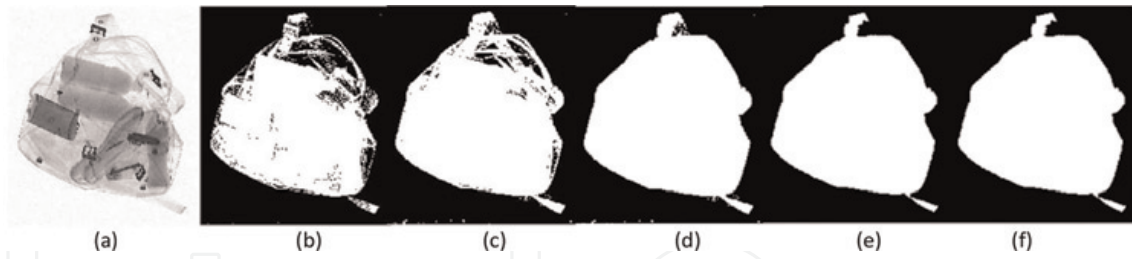


Figure 3.
 Image segmentation using morphological operations for insertion position determination. (a) Grayscale (b) Binarisation (c) Dilation (d) Hole filling (e) Erosion (f) Biggest region

signature is completely located inside this target region. To this end, we use a loop to generate a random insertion position until it is a valid one. The selected valid insertion position can be denoted by a binary mask matrix M of the same size as the target baggage image with elements of ones indicating the insertion region.

Finally, a threat signature I_s is superimposed onto the target bag image I_t in the selected valid position (denoted by M) to generate a synthetically composited image I_{TIP} . To ensure the plausibility of the composited TIP image, we consider two factors in image blend. Parameter α controls the transparency of the source image I_s ($\alpha = 0.9$). The other parameter is the *threat threshold* T ensuring the consistency of source image with the target image in terms of image contrast. The use of *threat threshold* T aims to remove the high-value pixels of the threat signature so that the inserted threat signature is not visually too bright comparing against the target region where it is superimposed. To calculate the value of T , we first transform the target image I_t to a grayscale image G_t . The threat threshold T can be empirically calculated by:

$$T = \min(\exp(\hat{g}^5) - 0.5, 0.95) \quad (1)$$

where \hat{g} is the normalised average intensity of the insertion region within G_t , calculated as:

$$\hat{g} = \frac{\sum_{i,j} G_t(i,j) * M(i,j)}{\sum_{i,j} 255 * M(i,j)} \in [0, 1] \quad (2)$$

The image compositing can be formulated as follows:

$$I_{TIP}(i,j) = \begin{cases} (1 - \alpha) I_t(i,j) + \alpha I_s(i',j'), & M(i,j) = 1 \text{ and } I_s(i',j') < 255 * T \\ I_t(i,j), & \text{otherwise} \end{cases} \quad (3)$$

where $I(i,j)$ denotes the value of pixel in i -th row and j -th column of the image I ; $I_s(i',j')$ denotes the pixel in source image corresponding to the pixel of $I(i,j)$. Since the value of T computed by Eq. (1) is in the range of 0.5–0.95, any pixel of the higher value than $T * 255$ in the source image will be ignored during image compositing process.

The proposed TIP approach is able to generate a large number of diverse synthetic X-ray baggage images containing prohibited items whose locations are accessible without any extra cost for training a supervised learning detection model.

3.2 Detection strategies

We use two representative CNN object detection model, Faster R-CNN [8] and RetinaNet [26], for the purposes of our evaluation.

Faster R-CNN [8] is an object detection algorithm which is the combination of its predecessor Fast R-CNN [27] and Region Proposal Network (RPN). Unlike Fast R-CNN [27], which utilises external region proposal, this architecture has its own region proposal network, which consists of convolutional layers that generate object proposals and two fully connected layers that predict coordinates of bounding boxes. The corresponding locations and bounding boxes are then fed into objectness classification and bounding box regression layers. Finally the objectness classification layer classify whether a given region proposal is an object or a background region while a bounding box regression layer predicts object localisation, at the end of the overall detection process.

RetinaNet [26] is an object detector where the key idea is to solve the extreme class imbalance between foreground and background classes. To improve the performance, RetinaNet employs a novel loss function called Focal Loss, where it modifies the cross-entropy loss such that it down-weights the loss in easy negative samples so that the loss is focusing on the sparse set of hard samples. Unlike Faster R-CNN [8] which apply two-stage approach, RetinaNet only apply one-stage approach, potentially to be faster and simpler.

4. Experimental setup

Our experimental setup comprises of real X-ray security imagery dataset and one constructed using the TIP based synthetic compositing approach outlined in Section 3.1. These are evaluated within a common CNN training environment using the CNN architecture outlined in Section 3.2.

Dbf3_{Real} dataset: The Durham Dataset Full Three-class (Dbf3) images are generated using a Smith Detection dual-energy X-ray scanner (**Figure 1**). It consists of total 7603 images, which is divided into three classes of prohibited item. In this experiment we uses subsets of the datasets, which consists of three types metallic prohibited item, {Firearm, Firearm Parts, Knives}. Out of these three classes, we incorporate 3192 images of firearms, 1204 images of firearms parts, and 3207 images of knives, within cluttered and complex X-ray security dataset.

Dbf3_{SC} dataset: The Synthetically Composited (SC) dataset is generated using TIP approach of Section 3.1. We use 3366 benign X-ray security images, generated by a Smith Detection X-ray scanner, and 123 individual prohibited objects of three classes {*Firearm*, *Firearm Parts*, *Knives*}. The prohibited item are composed into the benign images to create synthetically composited X-ray security imagery dataset. We use the same number of images as *Dbf3_{Real}* in the synthetically composited dataset. Exemplar images from *Dbf3_{Real}* (**Figure 4A**) and *Dbf3_{SC}* (**Figure 4B**) are visually realistic and challenging to distinguish from the real images.

Dbf3_{Real+SC} dataset: A subset of *Dbf3_{Real}* and subset of *Dbf3_{SC}* images are combined to create this dataset, where the numbers of synthetic and real images are used in equal number to present a data set with 50% of each which is itself the same size as *Dbf3_{Real}*.

The CNN architecture (Section 3.2) are trained on a GTX 1080Ti GPU, optimised by Stochastic Gradient Descent (SGD) with a weight decay of 0.0001, learning rate of

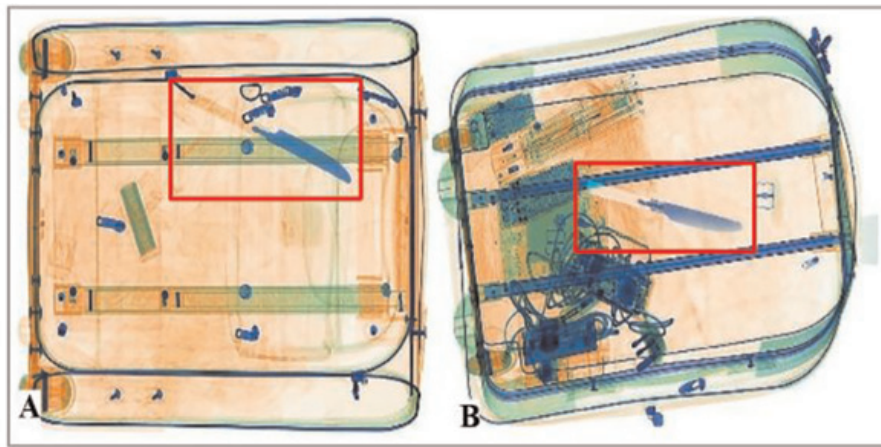


Figure 4.
 Visual comparison of real (A) and SC (B) X-ray security imagery of prohibited items.

0.01 and termination at maximum of 180k epochs. ResNet₅₀ and ResNet₁₀₁ are chosen as network backbone to operate within the detection framework of [28]. We split each dataset into training (60%), validation (20%) and test sets (20%) so that each split has similar class distribution. All CNN architecture are initialised with ImageNet [5] pre-trained weights for their respective model [29].

5. Evaluation

Our evaluation considers the comparative performance of CNN architecture to detect prohibited items using real X-ray imagery against prohibited items under synthetic X-ray imagery. We consider mean Average Precision (mAP) as our evaluation criteria following [15].

5.1 Prohibited item detection—Quantitative results

In the first set of experiments (**Table 1**, upper), prohibited items in X-ray security imagery are detected using the CNN architectures set out in Section 3.2. We use the *Dbf3* dataset consisting of three types of prohibited items {*Firearm*, *Firearm Parts*, *Knives*}. To provide performance benchmark, our CNN architectures are firstly trained and evaluated on real images of *Dbf3* ($Dbf3_{Real} \rightarrow Dbf3_{Real}$). The AP/mAP highlighted in **Table 1** (upper) denotes the maximal performance achieved. **Table 1** shows statistical results of prohibited item detection for Faster R-CNN [8] and RetinaNet [26] architecture using ResNet₅₀ and ResNet₁₀₁. Inline with the overall complexity of the network, we observe maximal mAP performance from ResNet₁₀₁ for all three prohibited item classes. In this performance benchmark, we observe that the best performance (mAP = 0.88) is achieved on $Dbf3_{real}$ by Faster R-CNN with ResNet₁₀₁ configuration, as presented in **Table 1** (upper).

In second set of experiments (**Table 1**, middle), the CNN architecture are trained on the synthetic X-ray imagery ($Dbf3_{SC}$) achieve 0.78 mAP when tested on same set of real X-ray imagery ($Dbf3_{Real}$) of **Table 1** (upper). Even though the performance is lesser when compared with former results (**Table 1**, upper), this experimental setting does not require any manual image labelling (as TIP insertion positions are known) and yet achieves surprisingly good performance on a standard benchmark. The performance gap between CNN architecture trained on real and synthetically composited

Train \Rightarrow Evaluation	Model	Network	Average precision			mAP
			Firearm	Firearm Parts	Knives	
$Dbf3_{Real} \Rightarrow$	Faster	ResNet ₅₀	0.87	0.84	0.76	0.82
	R-CNN [8]	ResNet ₁₀₁	0.91	0.88	0.85	0.88
$Dbf3_{Real}$	RetinaNet [26]	ResNet ₅₀	0.88	0.86	0.73	0.82
		ResNet ₁₀₁	0.89	0.86	0.73	0.83
$Dbf3_{SC} \Rightarrow$	Faster	ResNet ₅₀	0.82	0.77	0.55	0.71
	R-CNN [8]	ResNet ₁₀₁	0.86	0.80	0.66	0.78
$Dbf3_{Real}$	RetinaNet [26]	ResNet ₅₀	0.84	0.77	0.53	0.71
		ResNet ₁₀₁	0.84	0.76	0.54	0.72
$Dbf3_{Real+SC} \Rightarrow$	Faster	ResNet ₅₀	0.85	0.79	0.65	0.76
	R-CNN [8]	ResNet ₁₀₁	0.87	0.81	0.74	0.81
$Dbf3_{Real}$	RetinaNet [26]	ResNet ₅₀	0.85	0.81	0.64	0.76
		ResNet ₁₀₁	0.86	0.80	0.63	0.76

Table 1.

Detection results of varying CNN architecture trained on: upper \rightarrow $Dbf3_{Real}$, middle \rightarrow $Dbf3_{SC}$ and lower \rightarrow $Dbf3_{Real}$. All models are evaluated on set of real X-ray security imagery.

X-ray imagery is attributable to the domain shift problem whereby the distribution of training and test data differ. In the first experiment (**Table 1**, upper), the training and test data are from the same distribution since they are created by randomly dividing data captured under the same experimental conditions. By contrast, in this second experimental setup (**Table 1**, middle), the prohibited items used for the synthetic X-ray imagery ($Dbf3_{SC}$) data are different from those in the test X-ray imagery ($Dbf3_{Real}$) data. It is also noteworthy that prohibited images used for generating synthetic X-ray imagery ($Dbf3_{SC}$) data is a smaller set of prohibited item instances than in the real training images. As a result, CNN architecture trained on synthetic data have larger generalisation errors than those trained on real data. However, when tested on synthetic X-ray imagery ($Dbf3_{SC}$) data (**Table 2**), however, CNN architecture trained with real or synthetic CNN architecture have comparable performance. These experimental results show that it is essential to have diverse prohibited item signatures in the training data to improve the generalisation. It also largely explains why overall performance in **Table 2** (showing evaluation on the synthetic dataset, $Dbf3_{SC}$) is significantly higher than overall performance in **Table 1** (evaluation is on the real dataset, $Dbf3_{Real}$).

In the third set of experiments (**Table 1**, bottom), we evaluate the effectiveness of synthetic X-ray imagery by combining it with real images of $Dbf3$ to create $Dbf3_{Real+SC}$ dataset, as explained in the Section 4. We evaluate the testing sets of images from real $Dbf3$ (**Table 1**) and synthetically composite (**Table 2**) datasets. Surprisingly, the combination of real and synthetic imagery data does not improve the results (e.g. 0.81 vs. 0.88 on $Dbf3_{Real}$ and 0.89 vs. 0.91 on $Dbf3_{SC}$ with Faster R-CNN and ResNet101). This can also be explained by the domain shift problem mentioned previously. Possibly this data combination can perform well if we apply domain adaptation techniques [30] explicitly. In addition, we may also need to evaluate the quality of the TIP solution that underpins our work further.

Train \Rightarrow Evaluation	Model	Network	Average precision			mAP
			Firearm	Firearm Parts	Knives	
$Dbf3_{Real} \Rightarrow$	Faster	ResNet ₅₀	0.88	0.87	0.84	0.87
	R-CNN [8]	ResNet ₁₀₁	0.92	0.92	0.89	0.91
$Dbf3_{SC}$	RetinaNet [26]	ResNet ₅₀	0.89	0.87	0.83	0.86
		ResNet ₁₀₁	0.90	0.88	0.85	0.88
$Dbf3_{SC} \Rightarrow$	Faster	ResNet ₅₀	0.90	0.88	0.83	0.87
	R-CNN [8]	ResNet ₁₀₁	0.93	0.92	0.86	0.91
$Dbf3_{SC}$	RetinaNet [26]	ResNet ₅₀	0.91	0.89	0.84	0.88
		ResNet ₁₀₁	0.91	0.89	0.83	0.86
$Dbf3_{Real+SC} \Rightarrow$	Faster	ResNet ₅₀	0.89	0.86	0.83	0.86
	R-CNN [8]	ResNet ₁₀₁	0.91	0.89	0.87	0.89
$Dbf3_{SC}$	RetinaNet [26]	ResNet ₅₀	0.90	0.87	0.83	0.87
		ResNet ₁₀₁	0.90	0.88	0.84	0.87

Table 2.
Detection results of different CNN architecture trained on: Upper \rightarrow $Dbfg3_{Real}$ middle \rightarrow $Dbf3_{SC}$ and lower \rightarrow $Dbfg3_{Real}$. ALL models are evaluated on set of SC dataset.



Figure 5.
Exemplar detection of prohibited items in red box using Faster R-CNN [8] and trained on (A) $Dbfg3_{Real}$ and (B) $Dbf3_{SC}$ images.

5.2 Prohibited Item detection—Quantitative examples

Exemplar prohibited items detection results from Faster R-CNN [8] with ResNet₁₀₁ are depicted in **Figure 5**, using real (top row) and synthetic (bottom row) training imagery. These results illustrate that the synthetically composited imagery using TIP techniques can be effective in training detection architectures for prohibited item detection in cluttered X-ray security imagery.

We also visually inspect the detection results to investigate the performance difference when training the models using real and synthetic data. By comparing the results depicted in **Figure 6A1** and **B1**, the model trained with synthetic data fails to detect the knives since such type of knives have very different appearance from the

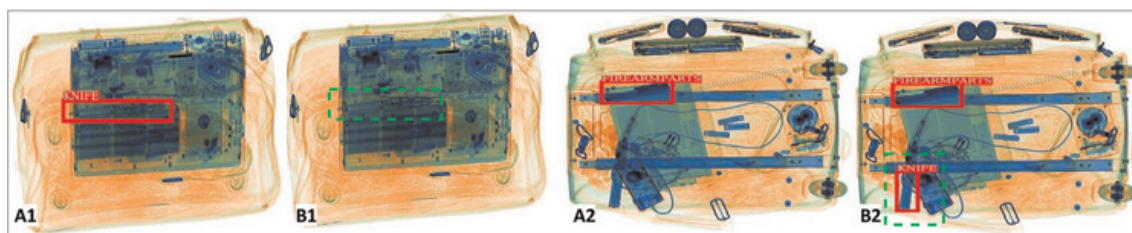


Figure 6. Exemplar prohibited item detection (by Faster R-CNN [8]) using $Dbfg3_{Real}$ (A1, A2) and $Dbf3_{SC}$ (B1, B2) training datasets. Green dashed box in B1 fails to detect, where in B2 wrongly detects as knife.

ones we used to generate the synthetic imagery. On the other hand, from **Figure 6A2** and **B2** we can see that the model trained on synthetic imagery has mistakenly detected something benign as a knife. These results account for the low performance for knife detection observed in **Table 1**. As a result, we need to either use more diverse threat signatures for data synthesis or particular domain adaptation techniques to tackle the potential domain shift problem identified previously.

6. Conclusion

This work explores the possibility of generating synthetically composite X-ray security imagery for training of CNN architecture to bypass the collecting a large amount of hand-annotated real-world X-ray baggage imagery. We synthesise high-quality synthetically composited X-ray images using TIP approach and we present an extensive comparison on how real and synthetic X-ray security imagery affects the performance of CNN architecture for prohibited object detection in cluttered X-ray baggage images. Our experimental comparison demonstrates Faster R-CNN achieves the highest performance with mAP: 0.88 when trained on *Real* data (the good), followed by *Real+Synthetic* (the bad) and *Synthetic* (the ugly) over a three-class, $\{\text{Firearms}, \text{Firearm parts}, \text{Knives}\}$, prohibited item detection problem. This demonstrates a strong insight into the benefits of using real X-ray training data, also challenge and promise of using synthetic X-ray imagery.

In our future work, it is worth further investigating how to improve the effectiveness of synthetically composited imagery for training CNN architecture. Based on other work [20], a potential direction is to generate more diverse prohibited items images using generative adversarial networks (GAN). The generated prohibited item images then could be used for generating synthetic baggage images using TIP or similar.

Acknowledgements

The authors would like to thank the UK Home Office for partially funding this work. Views contained within this paper are not necessarily those of the UK Home Office.

Conflict of interest

The author declares there is no conflict of interest.

IntechOpen

Author details


Neelanjan Bhowmik¹, Qian Wang¹, Yona Falinie A. Gaus¹, Marcin Szarek²
and Toby P. Breckon^{1*}

1 Department of Computer Science, Durham University, Durham, UK

2 School of Engineering, Cranfield University, Cranfield, UK

*Address all correspondence to: toby.breckon@durham.ac.uk

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Schwaninger A, Bolfig A, Halbherr T, Helman S, Belyavin A, Hay L. The impact of image based factors and training on threat detection performance in X-ray screening. In: International Conference on Research in Air Transportation (ICRAT). Virginia, USA: Fairfax; 2008. pp. 317-324
- [2] Blalock G, Kadiyali V, Simon DH. The impact of post-9/11 airport security measures on the demand for air travel. *The Journal of Law and Economics*. 2007;**50**(4):731-755
- [3] Turcsany D, Mouton A, Breckon TP. Improving feature-based object recognition for X-ray baggage security screening using primed visual words. In: International Conference on Industrial Technology. New York: IEEE; 2013. pp. 1140-1145
- [4] Mouton A, Breckon TP. A review of automated image understanding within 3D baggage computed tomography security screening. *Journal of X-ray Science and Technology*. 2015;**23**(5): 531-555
- [5] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. MIT Press. 2012. pp. 1097-1105
- [6] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA. 2016
- [7] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco, California, USA. 2017. pp. 4278-4284
- [8] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Transactions on Pattern Analysis and Machine Intelligence*. 2017;**39**(6): 1137-1149
- [9] Mery D, Riffo V, Zscherpel U, Mondragón G, Lillo I, Zuccar I, et al. Gdxray: The database of X-ray images for nondestructive testing. *Journal of Nondestructive Evaluation*. 2015; **34**(4):42
- [10] Miao C, Xie L, Wan F, Su C, Liu H, Jiao J, et al. Sixray: A large-scale security inspection X-ray benchmark for prohibited item discovery in overlapping images. In: *Conference on Computer Vision and Pattern Recognition*. New York: IEEE; 2019. pp. 2119-2128
- [11] Mery D, Riffo V, Zuccar I, Pieringer C. Automated X-ray object recognition using an efficient search algorithm in multiple views. In: *Conference on Computer Vision and Pattern Recognition Workshops*, Portland, Oregon, USA. 2013. pp. 368-374
- [12] Kundegorski ME, Akcay S, Devereux M, Mouton A, Breckon TP. On using feature descriptors as visual words for object detection within X-ray baggage security screening. In: *International Conference on Imaging for Crime Detection and Prevention*, Greater Noida, India. 2016. pp. 1-6
- [13] Mery D, Svec E, Arias M. Object recognition in baggage inspection using adaptive sparse representations of X-ray

images. In: Image and Video Technology, Auckland, New Zealand. 2015. pp. 709-720

[14] Akcay S, Kundegorski ME, Devereux M, Breckon TP. Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery. In: IEEE International Conference on Image Processing, Phoenix, Arizona, USA. 2016. pp. 1057-1061

[15] Akcay S, Kundegorski ME, Willcocks CG, Breckon TP. Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery. IEEE Transactions on Information Forensics and Security. 2018;**13**(9):2203-2215

[16] Mery D, Svec E, Arias M, Rizzo V, Saavedra JM, Banerjee S. Modern computer vision techniques for X-ray testing in baggage inspection. Transactions on Systems, Man, and Cybernetics: Systems. 2016;**47**(4): 682-692

[17] Dai J, Li Y, He K, Sun J. R-FCN: Object detection via region-based fully convolutional networks. In: Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain. 2016. pp. 379-387

[18] Redmon J, Farhadi A. Yolo9000: Better, faster, stronger. In: Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA. 2017. pp. 7263-7271

[19] Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using gan for improved liver lesion classification. In: International Symposium on Biomedical Imaging, Washington, DC. 2018. pp. 289-293

[20] Yang J, Zhao Z, Zhang H, Shi Y. Data augmentation for X-ray prohibited item images using generative adversarial networks. IEEE Access. 2019;**7**:894-902

[21] Jain D et al. An evaluation of deep learning based object detection strategies for threat object detection in baggage security imagery. Pattern Recognition Letters. 2019;**120**:112-119

[22] Neiderman EC, Fobes JL. Threat image projection system. 2005. US Patent 6,899,540

[23] Schwaninger A, Hardmeier D, Hofer F. Measuring visual abilities and visual knowledge of aviation security screeners. In: International Carnahan Conference on Security Technology, Albuquerque, NM, USA. 2004. pp. 258-264

[24] Schwaninger A, Michel S, Bolting A. A statistical approach for image difficulty estimation in X-ray screening using image measurements. In: Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, New York, NY, United States. 2007. pp. 123-130

[25] Rogers TW, Jaccard N, Protonotarios ED, Ollier J, Morton EJ, Griffin LD. Threat image projection (TIP) into X-ray images of cargo containers for training humans and machines. In: International Carnahan Conference on Security Technology, Orlando, Florida, USA. 2016. pp. 1-7

[26] Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: International Conference on Computer Vision, Venice, Italy. 2017. pp. 2980-2988

[27] Girshick R. Fast R-CNN. In: International Conference on Computer

Vision, Santiago, Chile. 2015.
pp. 1440-1448

[28] Girshick R, Radosavovic I, Gkioxari G, Dollár P, He K. Detectron. 2018. Available from: <https://github.com/facebookresearch/detectron>

[29] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA. 2017. pp. 1492-1500

[30] Wang Q, Bu P, Breckon T. Unifying unsupervised domain adaptation and zero-shot visual recognition. In: International Joint Conference on Neural Networks, Budapest, Hungria. 2019

Water Quality Monitoring Using Innovative Technologies

Fiona Regan

Abstract

Monitoring of water quality is controlled by policy and legislation and programmes of measures defining the approach. However, temporal and spatial frequency of monitoring is sometimes insufficient to meet the needs of water quality management. Therefore, there is a need for new and affordable technologies that can collect data on water quality in real time or near real time at high spatial frequency. These approaches will not replace the need to take a spot sample and report a result, but they can help with the decision-making process. This chapter looks at several novel sensing developments, the challenges that they are designed to address and the mechanism of sensing. The technologies represent a hierarchy of approaches from low-cost nonspecific sensing to very selective biosensor technologies. These systems are in different stages of development and application, but this shows the potential is bright for addressing the challenges related to water quality monitoring. These technologies when commercially available can add to the toolset for water monitoring. However, there are certain data gaps in small stream monitoring, nutrient monitoring in surface waters, and so on. Therefore, a disruptive approach involves citizen science in addressing these data gaps to help reach the SDG 6 goals.

Keywords: sensing, biosensors, microfluidics, multimodal data, water quality

1. Introduction

1.1 Water monitoring

Today water is the most precious and valuable resource. Water quality monitoring can be expensive to carry out and challenging in hard-to-reach locations. Typically, it still requires a grab or spot water sample to be manually taken. This is both costly, time-consuming, and prone to error. It is not possible to achieve good spatial and temporal data for the sites of interest using this approach. The rationale of water monitoring comes from the relevant policy or legislation in each country. The function of water monitoring is to ensure good ecosystem health or provide a water supply suitable for use. Considerations for any water monitoring programme are summarized in **Figure 1**.

Drinking water utilities face new challenges in day-to-day operation due to growing population, ageing infrastructure and climate-related weather events. Therefore there is a need to address more suitable approaches to monitoring water quality. Although the

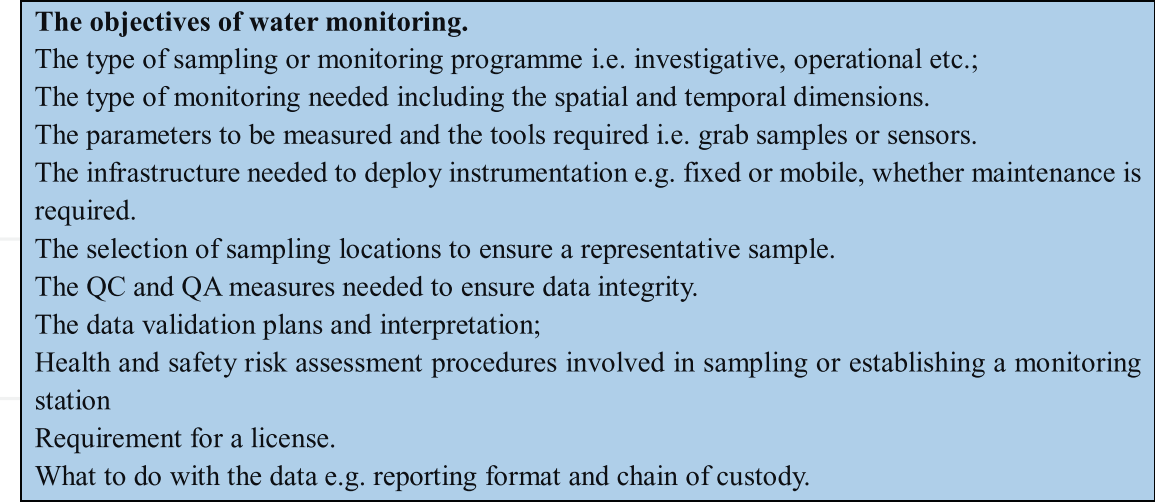


Figure 1.
Considerations when planning a water monitoring programme.

current monitoring approaches include the determination of physical, chemical and biological species, there are drawbacks: a) poor temporal and spatial information b) laborious and costly approach and c) absence of real-time water quality information to. This highlights the opportunity for continuous water quality monitoring. A drawback of course to using in-situ sensing is the level of maintenance required. In-situ sensing is commonly used for the measurement of common water quality parameters that contribute to and affect the environment. *In-situ* sensors allow for high-resolution monitoring, they can be used alone or deployed as part of an observation system [1, 2]. However, a major limitation currently is the lack of available sensors that can measure priority parameters such as chemicals of emerging concern, bacteria, or natural toxins in water. The benefit of being able to measure such parameters in real-time could mean that effective decision supports could be implemented, and appropriate measures applied.

1.2 New modes of sensing

For water monitoring many forms of data gathering are emerging including satellites with sensors, aerial vehicles that carry imaging systems and in-situ sensors for direct water quality measurements.

1.2.1 Remote sensing

Remote monitoring, via satellite assets and distributed in-situ sensors, has the potential to meet the challenges of data gaps in many areas including developing countries. Existing and new satellites are proving useful in coastal waters, estuaries, lakes, and reservoirs, which are relevant to water quality managers. The information from satellites can assist the policy design and implementation.

Chlorophyll a (*Chl-a*) determination in water bodies is a common application of remote sensing for water quality monitoring as an indicator for harmful algal blooms (HAB). Although remote sensing techniques contribute to very promising outputs and offers a replacement method for field sampling, enabling fast, temporal, spatial and frequent observations [3, 4], it suffers from some limitations concerning the accuracy of the products obtained, data continuity, excess tools, and software for atmospheric correction (scattering and absorption effects, cloud cover etc.) and the

precision of the results [4, 5]. Hence the added value of combining remote sensing with another monitoring tool such as *in-situ* sensing devices.

In-situ instrumentation, including flow meters and water quality sensors installed in remote and hard to access locations, can complement satellite information. Remote sensing data is generally used to monitor parameters such as secchi disk depth, temperature, total organic carbon, total suspended matter, turbidity, conductivity, colour dissolved organic matter, *Chl-a*, suspended solids and sea surface salinity [2]. Geographic Information System (GIS) is commonly used alongside remote sensing techniques, it is a tool that is used to process and analyse the changes observed in satellite imagery.

1.2.2 Unmanned aerial vehicles (UAVs)

Unmanned aerial vehicles (UAVs, or drones), are being used by researchers to monitor catchments. Structure from Motion (SfM) three-dimensional point clouds can be derived from digital images collected on board UAVs, similar to the type of data produced by Light Detection and Ranging (LiDAR) technology [6]. The technological development of UAVs and their ability to derive high-resolution 3D information has made UAV image acquisition appealing in several applications. The UAV systems are mounted with red, green, blue (RGB), colour-infrared (CIR) and multispectral (MS) cameras [7]. They allow for near real-time monitoring of an area and comparably low operational costs, with the RGB cameras being the most economical [8]. Images from airborne sensors can provide a unique method of monitoring [9]. Drone systems are used to get a better visual representation of the area being studied. These images are generally taken from a height between 500 and 1,000 meters. The drawback to this method of monitoring is that the payload for most of these systems is relatively low, therefore the instruments choice is important.

1.2.3 Autonomous underwater vehicles (AUVs)

The most recent developments in remote sensing are the growth of autonomous underwater vehicles which are also known as unmanned underwater vehicles (UUV) or remotely operated vehicles (ROV) [10]. AUV's are commonly used in oceanographic monitoring and are becoming more popular for industrial applications and environmental monitoring. These new machines have demonstrated to be cost-effective, efficient, and safe alternatives [11] for water quality assessment and monitoring. They employ the same "deploy and forget" ethos as autonomous sensors. It is possible to mount multiple water quality sensors on board the vehicle which can then be deployed in remote areas. AUV's require tether connections for them to be powered and controlled, this is done via fibre microwave transmissions, fibre optics, remote controllers, communication media and satellites so that they can receive and send data [11]. The hulls of these devices are becoming smaller as the need to minimize manufacturing resources increases, cutting down on power, fabrication costs and time [12] are all important aspects for the new wave of AUV's.

2. New water quality sensors are helping to democratize access to data

A variety of water quality sensors are commonly used for both in situ continuous monitoring as well as sample-based testing. Off-the-shelf sensors are readily available to measure pH, dissolved oxygen (DO), conductivity (often used to measure salinity),

turbidity, temperature, chlorine, various dissolved ions (including fluoride, ammonia, silver, nitrate, and nitrite that are often by-products of fertilizers), and total organic carbon (TOC). Water quality assessments include the identification of coloured dissolved organic matter (CDOM), Secchi disk depth (SDD), turbidity, total suspended sediments (TSS), water temperature (WT), total phosphorus (TP), sea surface salinity (SSS), dissolved oxygen (DO), biochemical oxygen demand (BOD), and chemical oxygen demand (COD). State-of-the-art fluorimeter solutions are not ideally designed for in situ long-term use. This is primarily because biofouling and baseline drift attenuate and confuse the signal [13]. At present, there are no viable *in-situ* electronic sensors for monitoring the microbial contamination of drinking water for example.

This chapter briefly outlines the rationale behind and development of five innovative sensors for water quality parameters—for monitoring catchment (watershed) level to the sea and addressing a range of current and emerging challenges in terms of target analytes.

Those innovations briefly addressed in this chapter are as follows:

1. Marine sensing using centrifugal microfluidics;
2. Hand-held phosphate analyser for freshwater monitoring;
3. *E. coli* sensing for bathing water;
4. eDNA for fish detection;
5. Disruptive innovation by citizen science.

2.1 Microfluidic sensing platform for marine sensing

2.1.1 The challenge being addressed

With the rapid increase of population growth, the global demand and pressure for clean water supplies have never been more apparent. The UN Ocean Decade seeks to improve marine water quality through science. Legislation such as the Marine Strategy Framework Directive in Europe (<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32008L0056>), has led to improved monitoring of marine waters. However, with a growing list of emerging contaminants of concern, the challenge of monitoring is increasing. The ideal criteria for environmental monitoring tools include autonomous deployment, robust, quantitative, and qualitative output, and cost-effectiveness. It is important to have a highly flexible monitoring platform that can process a large variety of analytes under different conditions. The development of microfluidic sensors as cost-effective strategies for environmental monitoring has been a primary goal for the technology since the 1980s. It has played a pivotal role in the downsizing of sample ranges and complex assay protocols used in analyte detection. One such platform capable of this flexibility is the lab-on-a-disc centrifugal platform (LOAD) [14, 15], a derivative of the lab-on-a-chip (LOAC) [16–19] microfluidic platform.

2.1.2 The sensing device

The Lab On A Disc (LOAD) centrifugal microfluidic platform that uses a centrifugal driving force. To date, on-chip water quality assessment systems have been

primarily developed using Lab On A Chip (LOAC) systems, with only a few examples reported on centrifugal disc (CD) platforms [20, 21]. The EU Framework 7 project, MariaBox (www.mariabox.com) set out to develop an autonomous sensor for marine application with the ambitious aim of autonomous operation for several months in the field. The biosensor disc was developed at DCU. The original biosensor concept was to enable the determination of eight analytes (including marine toxins domoic acid, saxitoxin; poly aromatic hydrocarbons and metals) on-chip simultaneously, with each requiring a different biological or chemical assay method. The final 8-analyte sensor disc evolved through development and testing and emerged as an effective design enabling different bio- and chemical assays to be performed on the same chip by using antibodies, valving, and optimal micro channel design.

The complexities in developing a multi-analyte centrifugal microfluidic platform are many. These technical iterations are summarized in **Figure 2a** as well as the incremental optimization practices that were employed. The process starts with an initial concept, biosensor assay integration with material surface modifications, microvalve development for synchronised assay execution, and the final deployable outputs.

There are multiple advantages associated with LOAD platforms including; autonomous pumping using centrifugal forces, precise liquid handling, control of samples using valves, ability to multiplex assays using identical test conditions and a myriad of detection techniques compatible on disc, making it an ideal technique for *in-situ* environmental monitoring. The LOAD consists of a multi-layer disc that when combined, holds the channels, valves and reservoirs (**Figure 2b**).

To develop the centrifugal microfluidic platform (**Figure 2a**), a number of factors were considered; sample/reagent type/size, analyte characteristics, sample matrix, assay protocols, transduction method, targeted style of use (autonomous, handheld), usage (single, multiuse). The detection strategy on centrifugal microfluidics, offers a myriad of transducing solutions for biosensing [14, 22]. The key to developing a successful, multi-analyte environmental processing unit lies within the integration and adaptation of new and current microfluidic features and transducing elements in microfluidic platforms [23]. With further development, this type of biosensor for multiple analyte determination can transform how data can be gathered from hard to reach or remote marine or surface water environments in real-time or near real-time.

LOAD devices allow for rapid on-site measurements, with miniaturisation and automation of laboratory-based analytical protocols, towards the development of inexpensive, portable, and compact devices [17]. The use of microlitre volumes results in a reduction in reagent consumption and waste production compared to laboratory protocols. This, coupled with reduced cost and lowered sample contamination risk compared with standard methods for water analysis makes disposable microfluidic devices an attractive option [16]. The ease with which discs can be modified and built means that bioassays can be readily transferred to sensor systems.

2.2 Hand-held phosphate analyser for freshwater monitoring

2.2.1 The challenge being addressed

Water quality monitoring within catchments is desirable to maintain the structure and functioning of the land, water bodies and aquatic ecosystem that inhabit the catchment. Problems such as nutrient enrichment, sediment influx and dissolved

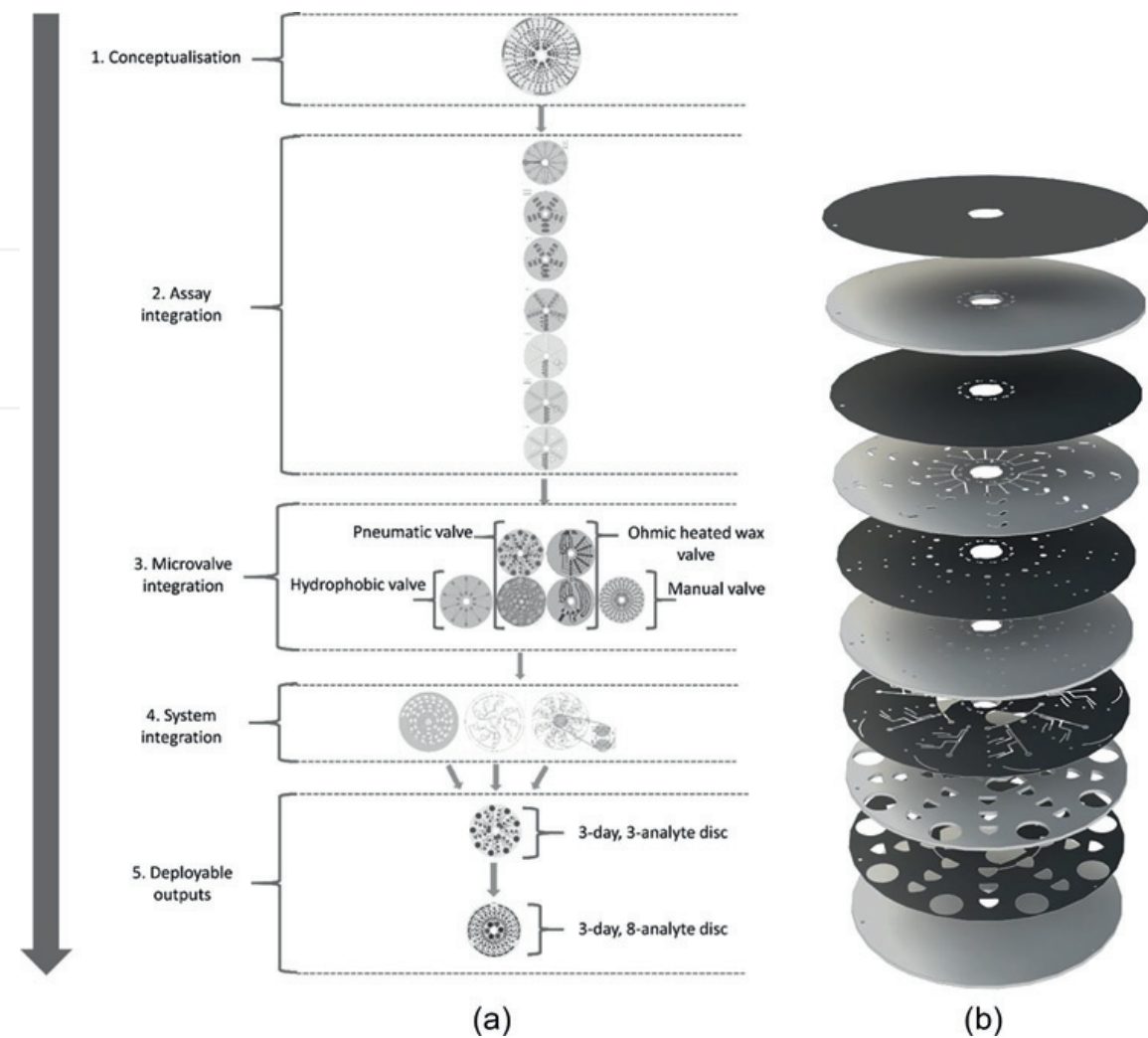


Figure 2. (a) (Left): The design tree for the MariaBox centrifugal processing unit (mCPU). In order of development, assay integration was prioritised to generate a proof of concept for testing assay integration. Next, microvalve integration strategies were investigated to automate assay protocol. This was then followed by the development and distribution of calibration and test discs for compatibility troubleshooting with the MariaBox platform. Finally, any compatibility issues were rectified, with two final deployable mCPU variants for the MariaBox platform produced; 3-day, 3-analyte disc (Biological analytes only) and 3-day, 8-analyte disc (all analytes). (b) (Right): Layer-by-layer disc design approach using CAD software for ease of compatibility and rapid modification. Poly(methyl methacrylate) (PMMA) Layers (white), ranging from 0.5–2 mm in thickness, are interlocked with pressure sensitive adhesive (PSA) layers (black), ranging from 80–150 μm in thickness. The thickness dimensions of both the PMMA and PSA make them ideal for μl liquid storage, and capillary action transportation, respectively.

oxygen depletion are characteristics of stressed water bodies [24]. Excessive riverine nutrient concentrations can have harmful effects on the aquatic ecosystem structure and functioning of a catchment [25].

Phosphorus (P) is an essential nutrient for life. It is a growth-limiting nutrient, which makes it an important parameter to monitor in water [26]. Elevated levels of growth-limiting nutrients lead to algal blooms [27]. Phosphorus exists in many different chemical forms in water [28]. The simplest method for estimating bioavailable phosphorus in water is to analyse for soluble reactive phosphate (SRP). Orthophosphates are the most abundant forms of SRP at pH levels typically encountered in natural waters [29].

The demand for rapid detection methods that provide real-time or near real-time quantification of phosphate levels in freshwater is recognised by both government

and legislative bodies [30]. Different methods and technologies for monitoring and detecting phosphate levels in freshwater has been reviewed by Warwick *et al.*, these technologies include: biometric receptors (synthetic receptors and molecular imprinted polymers (MIPs)), electrochemical detection (potentiometric, amperometric, voltametric and conductometric analysis) and optical detection methodologies (colorimetric absorbance and luminescence/fluorescence) [31]. Handheld sensors can be developed for rapid, robust and reliable environmental monitoring. They can provide near real or real-time analysis of environmental water pollution parameters. There is an increase in the demand for cheap, reliable and robust sensing devices that can be used out in the field daily to collect real-time or near real-time data on water quality parameters.

2.2.2 The phosphate sensor

Commonly used assays for optical based detection of phosphate include the Molybdenum blue [32], Vanadomolybdophosphoric acid [33] and Stannous Chloride [34]. These assays have some limitations related to their working complexity, chemical requirements (lifetime and stability) and monitoring duration.

A centrifugal microfluidic disc for this device was fabricated from poly(methyl methacrylate) (PMMA) and pressure sensitive adhesive (PSA) as in **Figure 2b**. Fluid movement for the assay is enabled by rotating the disc at ~8 Hz to generate a centrifugal force that acts from the centre of the disc, radially outwards. [24] The on-board microfluidic architecture, including an air ventilation system for performance-enhanced mixing of samples and reagents, also enables precise fluidic manipulation. In this device all chemical reagents needed in the assay are dried onto the centrifugal disc. This has transformed how the sensor can operate in that the only intervention needed is the addition of the water sample—thus making it an ideal device for taking into the field for rapid and sensitive analysis. A long optical path length of 75 mm was included on the disc design to maximize absorbance signal, achieving the desired sensitivity and limit of detection.

Such a portable sensor that can be transported around the catchment and detect the levels of specific nutrients in real-time will give more information about the quality of water in the overall catchment, not just one point (as in a spot sample or a single *in-situ* sensor), improving the temporal and spatial resolution of the data is obtained. These devices are easily transported around a catchment, therefore can increase the temporal and spatial data collection, filling in data gaps from the use of standard *in-situ* monitoring and remote sensing. **Figure 3** illustrates the sensor unit, the measurement disc, optical detection unit and control panel. This system has been tested in the field and demonstrates the potential for low-cost potable nutrient sensors in water quality monitoring.

2.3 Bacterial *E. coli* sensing for bathing water

2.3.1 The challenge being addressed

The biological contamination of water is a major concern for public health reasons. The COVID-19 pandemic has meant that more people are turning to sea swimming as a form of exercise and for general wellbeing. The current approach to monitoring bathing water quality requires a sample to be collected and returned to the laboratory for analysis. The time to result is generally 24-72 hours after sample collection.

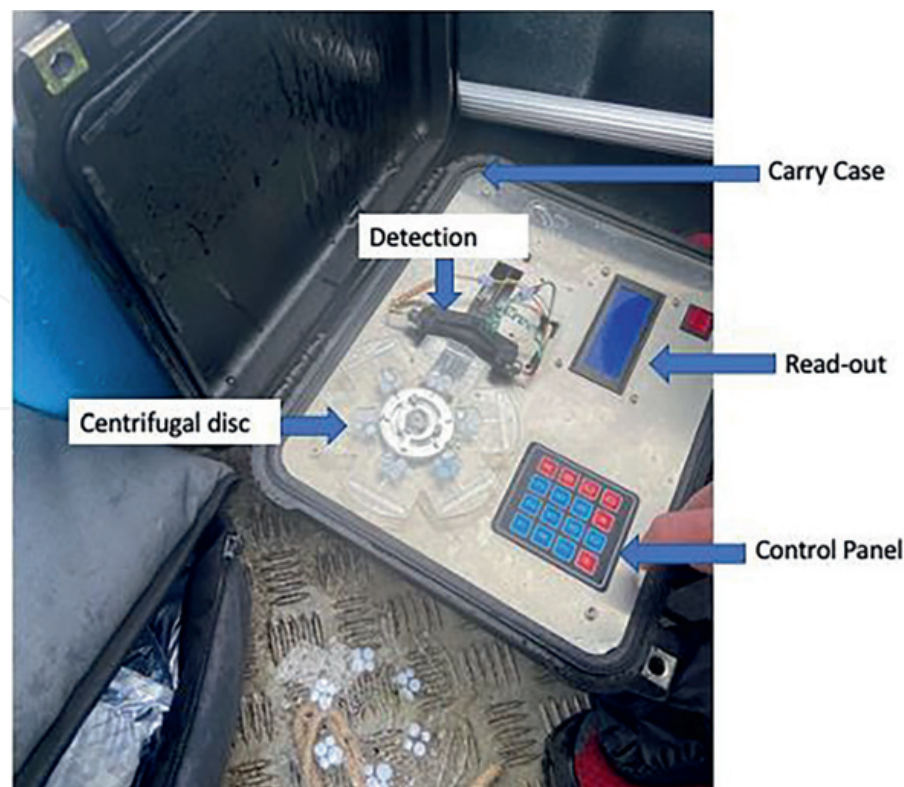


Figure 3.
Image showing the use of the phosphate analyser onboard a boat—showing the keypad below the centrifugal platform and the read-out to the right.

The determination of microbial water quality is a key requirement for water safety management. The microbial contamination of water is associated mainly with faecal coliforms (FC) such as mainly *Escherichia coli* (*E. coli*) and *Enterococcus*. Although FC has traditionally been regarded as good indicators of faecal contamination of waters, recent reviews suggest *E. coli* be a better, more specific indicator. [35] Both *E. coli* and *Enterococcus* can be classified as indicator organisms for faecal contamination and can be used to assess the hygienic quality of water. The EU Water Framework Directive's (WFD) bathing water directive is based on the classification and quality of bathing waters as 'poor', 'sufficient', 'good' or 'excellent' based on the presence of intestinal *Enterococcus* and *E. coli*. The EU Bathing Water Directive (<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31976L0160>) provides specific limits, in Colony Forming Units (CFU), of *E. coli* and *Enterococcus*, as a guide for acceptable water quality for both inland and bathing waters, reported in **Figure 4** [37, 38].

2.3.2 The sensing technology

The high contribution of faecal coliforms, particularly *E. coli*, to the microbiological contamination of water is of growing concern. The standard method for *E. coli* detection is time consuming as it focuses on culture-based methods. A method for the rapid identification of *E. coli* is the optical detection of β -Glucuronidase (GUD) activity using a soluble fluorescent molecule as a marker. GUD is a fluorometrically detectable enzyme that cleaves the glycosidic bond of glucuronides in living organisms. GUD is used widely as a marker enzyme for *E. coli* and can be exploited by incorporating chromogenic and fluorogenic substrates into growing mediums. Another method is the direct extraction of GUD from *E. coli* cells and measuring its activity. These methods eliminate any

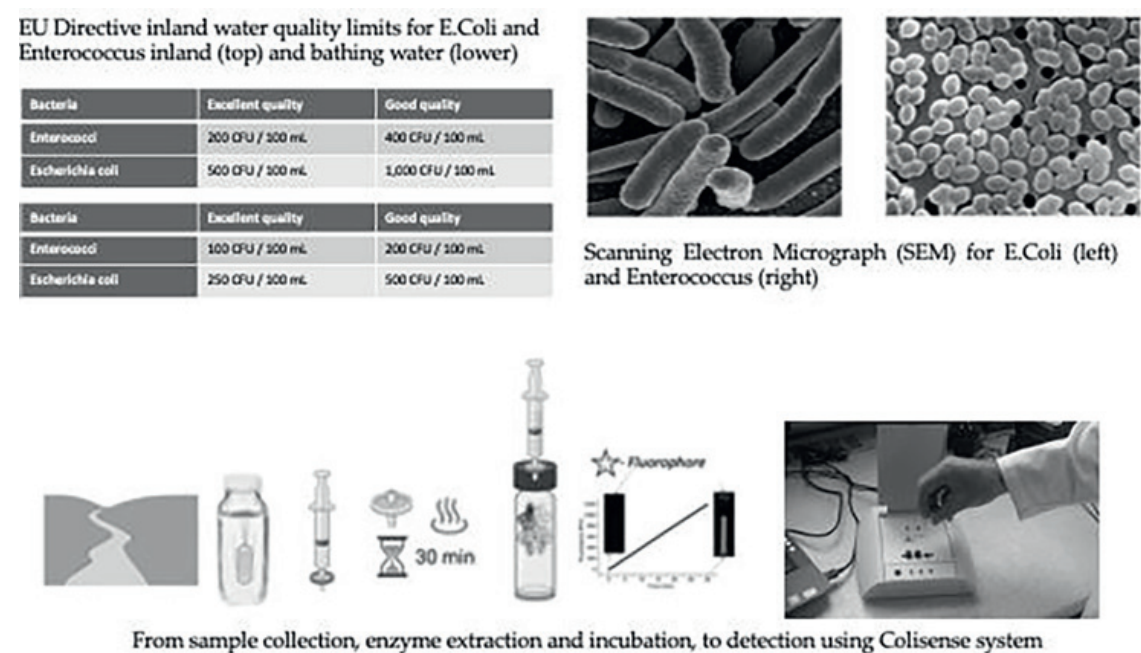


Figure 4. Schematic illustrating the elements of a bathing water monitoring sensor (i) the legislation driving the development; (ii) bacterial target species; and (iii) sample collection to measurement with the Colisense fluorimeter [36].

time-consuming culturing steps, providing a rapid monitoring system for the detection of *E. coli*. **Figure 4** shows the steps involved in the Colisense measurement system. This sensor developed in DCU provides for a sample collection to answer in 75 minutes [39]. The process involves sample collection, filtering to trap the bacteria, lysis to release the GUD enzyme and then addition of a fluorogenic substrate for detection using a mini fluorimeter as shown in **Figure 4**. The slope of the fluorescence response relates to the concentration of enzyme present and thus the level of bacterial contamination. This sensor methodology has been demonstrated during bathing seasons and it has been shown to correlate well with the standard culture-based method. The robust methodology can be further developed and commercialized to provide an early warning device for bathing water applications. This type of simple, robust technology has the potential to transform how information can be provided to sea swimmers and policy makers. The challenge in scaling this technology relates to the current legislation which relies on standard culture methods. A mind-set change is needed if low-cost sensor technology is to be adopted for routine environmental monitoring to protect human health.

2.4 eDNA for fish detection

2.4.1 The challenge being addressed

Environmental DNA (eDNA) is defined as the genetic material obtained directly from an environmental sample such as water, soil and sediment. The use of eDNA assessment is well established for detecting the presence or absence of rare or invasive species from the organic material they leave behind [40]. An organism provides a rich source of eDNA through the cells and waste they shed and excrete including faeces, mucus, gametes, hair and skin [41]. eDNA has been used to monitor a wide variety of species in a diverse range of environments.

There are two main approaches to eDNA monitoring: targeting specific species or detection of multiple species from a single sample, termed eDNA metabarcoding [42]. Monitoring of specific species, particularly endangered or invasive species, is vital for conservation purposes. The sensitivity, and importantly specificity, of molecular techniques has enabled the non-invasive monitoring of individual species from a variety of environmental samples [43]. Due to the low concentration of eDNA, these techniques require specific amplification of target DNA to a detectable limit [43]. For this reason, quantitative PCR (qPCR) is most commonly used and through robust primer design, has been shown to enable monitoring of several target species [44–47]. Moreover, qPCR has the potential to be adapted for any target of interest. Although the gold standard, other nucleic acid amplification techniques also offer potential for use in single species detection [48, 49]. For example, isothermal amplification techniques

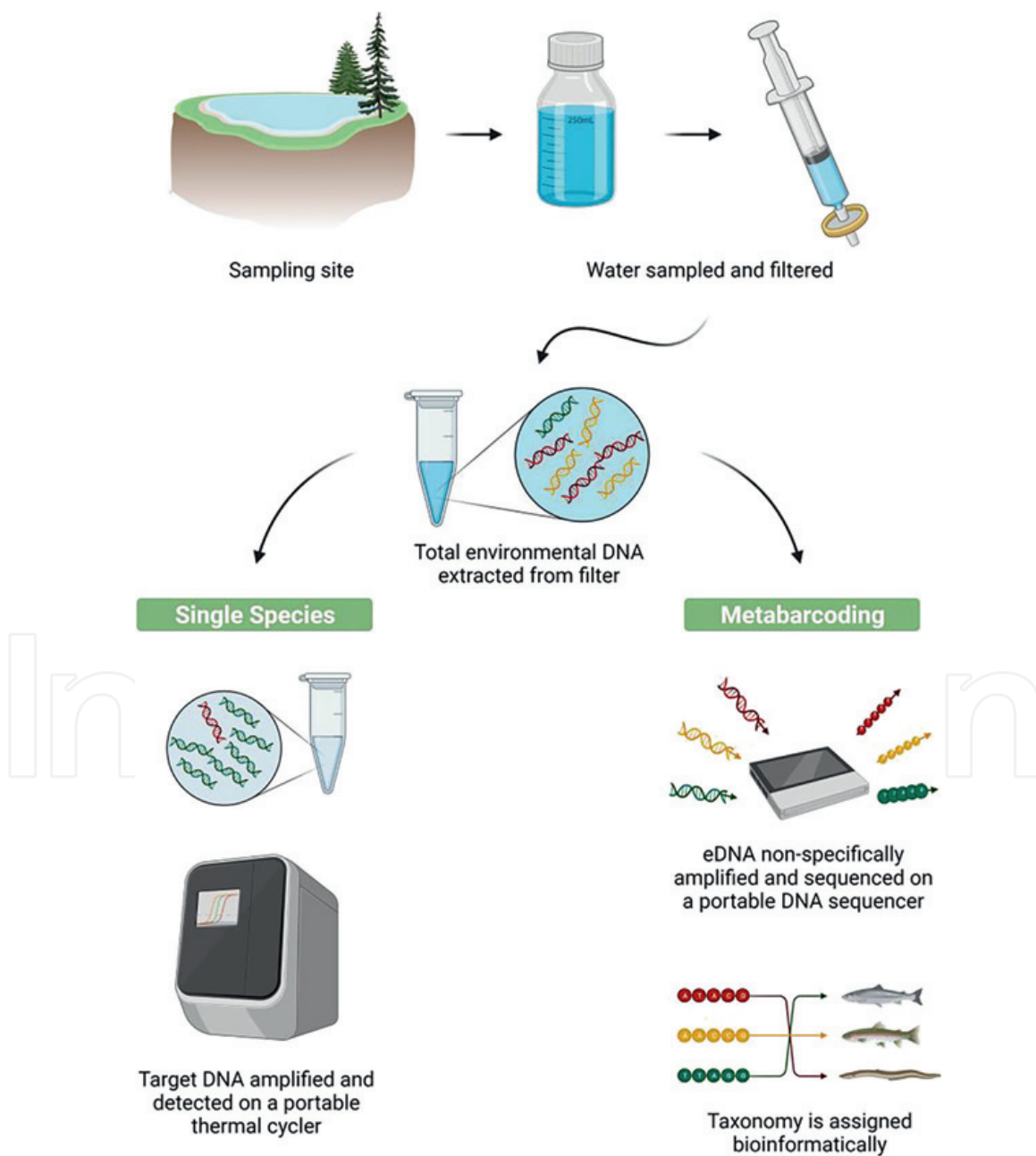


Figure 5.
Scheme showing the methods used to collect and analyse eDNA.

such as Recombinase Polymerase Amplification (RPA) and Loop-Mediated Isothermal Amplification (LAMP) have been used for diagnostics in the medical field [50] highlighting their specificity and sensitivity. These methods rely on enzymatic activity to maintain DNA amplification efficiency without the need for thermal cycling. On the other hand, eDNA metabarcoding allows identification of multiple species across all taxa from microbes to higher vertebrates as illustrated in **Figure 5**.

2.4.2 The sensing approach

The aim of this work is the development of a biosensor device for single species monitoring that can enable eDNA technology to be taken out of the laboratory and into the field. Conventional methods for eDNA detection pose a logistical challenge for on-site monitoring and adaptation to biosensor devices, due to the need for high temperatures and thermal cycling. Isothermal amplification methods have evolved as an alternative to PCR, enabling nucleic acid amplification at a constant temperature [51]. These methods can be coupled to fluorescence-based detection to enable a highly sensitive and specific analysis of target sequences [52, 53]. An alternative molecular eDNA approach, coupling isothermal RPA to a CRISPR-Cas12a detection system, has been developed.

In this work, three salmonid species were targeted, *S. salar*, *S. trutta* and *S. alpinus*, due to their ecological, economic and cultural importance in Ireland. The developed assays enabled detection of these species at a single temperature of 37°C. This switch from thermal cycling to isothermal detection enhances the potential for adaptation to a biosensor device. This approach has been a transformation in terms of how eDNA sensing in the field can be delivered. The CRISPR-Cas system, more commonly known for its role in genome editing, enables highly specific sequence recognition [53], which can be adapted to detect any species from eDNA samples from a variety of sources. This improves the ability to differentiate closely related species and enhances the capabilities of eDNA as a tool for biodiversity monitoring [54]. This work has for the first time demonstrated the detection of Atlantic salmon eDNA in the presence of other potentially competing species. This exciting development using the CRISPR-Cas system has meant that a sensor can be developed for eDNA target species and used directly in the field providing real-time or near real-time occurrence data.

3. Disruptive innovation with citizen science (CS)

3.1 The challenge being addressed

The United Nations Sustainable Development Goals (SDGs), announced in September 2015, has a vision of achieving a higher level of human health and well-being worldwide by the year 2030. The SDG targets specific to water and sanitation call for more detailed monitoring and response to understand the coverage and quality of safely managed sources. This provides challenges to monitoring but also opportunities to use innovative and disruptive approaches to gathering water quality data.

In order to fill gaps in data for water bodies, it is necessary to find ways of gathering information at higher spatial and temporal frequency than is feasible with current monitoring programmes.

3.2 What is citizen science

Citizen science relates to the active participation of the public in scientific research projects. This is a rapidly expanding field in open innovation. CS provides an integrated model of co-creation leading to public knowledge production with science. CS provides a valuable tool for citizens to play a more active role in sustainable development, and it can benefit the scientific community in the gathering of data that would otherwise be difficult and costly to obtain. Because any citizen can go out to collect data, sample sizes are therefore potentially very large [55]. This means that we can improve environmental event detection having more people collecting good quality data. Data gathered can inform on accelerated human impact on ecosystems, where increased and on-going monitoring is required to understand and potentially remedy environmental functions and discussed by Loisell *et. al.*, [56].

3.3 What success looks like

Earthwatch’s FreshWater Watch (FWW) programme is a global CS project that began in 2012 to investigate water quality and ecosystem degradation within freshwater bodies [57]. The programme is based on continuous monitoring of freshwater bodies. The basis of the approach is that it works with local groups and scientists who have a scientific question, to provide answers about local water quality and ecosystems [56].

In June 2019, in a partnership between Earthwatch, DCU Water Institute, a monitoring programme began on the river Liffey as it flows through Dublin city and inner suburbs. In September 2019, a WaterBlitz was carried out using the FWW methodology, in which over 350 samples were taken at sites chosen by citizens and not confined to the urban river Liffey [57]. Analysis of the data indicated a correlation of higher nitrates with agricultural land and also indicated elevated phosphate levels after heavy rainfall at locations which were sampled pre- and post-precipitation. Simple



Figure 6.
A citizen scientist compares the colour of the sample tube with a colour card when measuring nutrients in surface water.

colorimetric test kits were used to measure nitrate and phosphate levels in local water bodies (**Figure 6**) and results revealed interesting and valuable data relating to sources and levels of pollution around the country. The scale of data gathered over a short timeframe reflected the real potential of CS in water quality monitoring and in delivering real change in water management. This CS activity can be truly disruptive in how water can be monitored and how we can address SDG 6 targets.

4. Conclusion

The growing pressures on the environment arising from population growth, pollution control and climate change present a challenge for the current water monitoring methods. There are new chemical pollutants to monitor, events to track and trends to observe so that effective decisions can be made based on the information. Currently, the types of sensors available are limited to physical parameters or some basic water quality parameters. This chapter showcases the development of a number of low-cost or biosensor innovations that meet specific needs and address emerging analytical challenges. The work herein demonstrates the potential of microfluidic technology, integration of optical sensing with bioassays and portable systems to address current and emerging water quality challenges. The information provides a snapshot of the technologies and their performance, as a means to demonstrate the need and potential for innovation in water quality management.

The disruptive innovation that can be achieved is clear when disciplines of chemistry, biology, engineering and environmental sciences are merged. However, real disruption can also be achieved by mobilising the citizen scientist who has a personal interest as well as the opportunity to monitor their local water body and provide data that would otherwise not be gathered. We must combine the innovations of science and technology with the new wave of citizen engagement to address the future challenges of water quality monitoring.

Acknowledgements

The author acknowledges multidisciplinary collaborative work done on developing these technologies by researchers in the Water Institute including Ciprian Briciu-Burghina, Molly Williams, Gillian Duffy, Joyce O'Grady, Brendan Heery, Anne Parle McDermott, Nigel Kent, Caroline Murphy, Jennifer Fitzgerald and Ivan Maguire.

Conflict of interest

The author declares there is no conflict of interest.

Thanks

Thanks to Molly Williams who provided **Figure 5**.

IntechOpen


IntechOpen

Author details

Fiona Regan
School of Chemical Sciences, and DCU Water Institute, Dublin City University,
Dublin, Ireland

*Address all correspondence to: fiona.regan@dcu.ie

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Daniel A, Laes-Huon A, Barus C, Beaton A, Blandfort D, Guigues N, et al. Toward a harmonization for using in situ nutrient sensors in the marine environment. *Frontier in Marine Science*. 2020;**6**:1-22
- [2] Gholizadeh M, Melesse AM, Reddi L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors (Switzerland)*. 2016;**16**:8
- [3] Anderson C, Berdalet E, Kudela R, Cusack C, Silke J, O'Rourke E, et al. Scaling up from regional case studies to a global harmful algal bloom observing system. *Frontiers in Marine Science*. 2019;**6**(250):1-24
- [4] Japitana M, Demetillo AT, Burce MEC, Taboada EB. Catchment characterization to support water monitoring and management decisions using remote sensing. *Sustainable Environmental Research*. 2019;**29**:1-10
- [5] Wang K, Franklin SE, Guo X, He Y, McDermid GJ. Problems in remote sensing of landscapes and habitats. *Progress in Physical Geography*. 2009;**33**:747-768
- [6] Menezes-Blackburn D, Giles C, Darch T, George T, Blackwell M, Stutter M, et al. Opportunities for mobilizing recalcitrant phosphorus from agricultural soils: A review. *Plant Soil*. 2018;**427**:5-16
- [7] Zheng H, Cheng T, Li D, Zhou X, Yao X, Tian Y, et al. Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. *Remote Sensing*. 2018;**10**:6
- [8] Concha A, Civera J. RGBDTAM: A cost-effective and accurate RGB-D tracking and mapping system. 2017. DOI: 10.48550/arXiv.1703.00754
- [9] Lawley V, Lewis M, Clarke K, Ostendorf B. Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. *Ecological Indicators*. 2016;**60**:1273-1283
- [10] He Y, Wang D, Ali Z. A review of different designs and control models of remotely operated underwater vehicle. 2020;**53**:1561-1570
- [11] Li Y, Liu R, Liu S. The design of an autonomous underwater vehicle for water quality monitoring. *IOP Conference Series Material Science Engineering*. 2018;**301**:1
- [12] Walling DE, Collins AL, Stroud RW. Tracing suspended sediment and particulate phosphorus sources in catchments. *Journal of Hydrology*. 2008;**350**:274-289
- [13] EPA. Water quality in 2020, an indicators report. 2021. ISBN: 978-1-80009-009-5
- [14] Ducrée J, Haeberle S, Lutz S, Pausch S, Zengerle R, Stetten F, V, et al. The centrifugal microfluidic Bio-Disk platform. *Journal of Micromechanics Microengineering*. 2007;**17**:S103-S115
- [15] Kong LX, Perebikovskiy A, Moebius J, Kulinsky L, Madou M. Lab- on-a-CD: A fully integrated molecular diagnostic system. *Journal of Laboratory Automation*. 2016;**21**(3):323-355
- [16] Stone HA, Stroock AD, Ajdari A. Engineering flows in small devices:

Microfluidics toward a Lab-on-a-Chip. *Annual Review in Fluid Mechanics*. 2004;**36**:381-411

[17] Haeberle S, Zengerle R, Mark D, Von Stetten F, Zengerle R. Microfluidic platforms for lab-on-a-chip applications. *Lab Chip*. 2007;**7**:1094-1110

[18] Liu Y, Sun Y, Sun K, Song L, Jiang X. Recent developments employing new materials for readout in lab-on-a-chip. *Journal of Material Chemistry*. 2010;**20**:7305-7311

[19] Mark D, Haeberle S, Roth G, von Stetten F, Zengerle R. Microfluidic lab-on-a-chip platforms: Requirements, characteristics and applications. *Chemical Society Review*. 2010;**39**:1153-1182

[20] Maguire I, O’Kennedy R, Ducr  e J, Regan F. A review of centrifugal microfluidics in environmental monitoring. *Analytical Methods*. 2018;**10**:1497-1515

[21] Regan F, Fitzgerald J, Murphy C, Maguire I, O’Kennedy R. Convenient “one-step” extraction method for autonomous sensing of marine algal toxins. In: *Proceedings of the OCEANS 2017*. Aberdeen. 2017

[22] Maguire I, Fitzgerald J, McPartlin D, Heery B, Murphy C, Nwankire C, et al. A centrifugal microfluidic-based approach for multi-toxin detection for real-time marine water-quality monitoring. In: *Proceedings of the OCEANS 2017—Aberdeen*. 2017

[23] Maguire I, Fitzgerald J, Heery B, Nwankire C, O’Kennedy R, Ducr  e J, et al. A novel microfluidic analytical sensing platform for the simultaneous detection of three algal toxins in water. *ACS Omega*. 2018;**3**:6624-6634

[24] OECD. Agriculture’s impact on aquaculture: Hypoxia and Eutrophication. 2012;1-46

[25] Robertson GW. *Marine Ecosystem Management*, vol. 1. 1980

[26] Fabricius K, De’ath G, McCook L, Turak E, Williams D. Changes in algal, coral and fish assemblages along water quality gradients on the inshore Great Barrier Reef. *Marine Pollution Bulletin*. 2005;**51**:384-398

[27] Guide to the Water Framework Directive—Catchments.ie—Catchments.ie.” [Online]. Available from: <https://www.catchments.ie/guide-water-framework-directive/>

[28] DHPLG. River Basin Management Plan for Ireland 2018-2021 Executive Summary. 2017

[29] Chapman DV, Bradley C, Gettel G, Hatvani I, Hein T, Kov  cs J, et al. Developments in water quality monitoring and management in large river catchments using the Danube River as an example. *Environmental Science Policy*. 2016;**64**:141-154

[30] Tierney D, O’Boyle S. Water quality in 2016: An indicators report. 2018

[31] Warwick C, Guerreiro A, Soares A. Sensing and analysis of soluble phosphates in environmental samples: A review. *Biosensors and Bioelectronics*. 2013;**41**:1-11

[32] Yeh P, Yeh N, Lee C, Ding TJ. Applications of LEDs in optical sensors and chemical sensing device for detection of biochemicals, heavy metals, and environmental nutrients. *Renewable and Sustainable Energy Reviews*. 2017;**75**:461-468

[33] Neves M, Souto MRS, T  th IV, Victal SMA, Drumond MC, Rangel AOSS. Spectrophotometric flow system using vanadomolybdophosphate detection chemistry and a liquid

waveguide capillary cell for the determination of phosphate with improved sensitivity in surface and ground water samples. *Talanta*. 2008;**2**:527-532

[34] Gaikwad R, Kumbhar D, Ganeshkhind C. Development of portable optical sensor based system for detection of phosphate in wastewater. *International Journal of Chemical and and Physical Science*. 2018;**7**:470-475

[35] Briciu-Burghina C, Heery B, Duffy G, Brabazon D, Regan F. Demonstration of an optical biosensor for the detection of faecal indicator bacteria in freshwater and coastal bathing areas. *Analytical and Bioanalytical Chemistry*. 2019;**411**:7637-7643

[36] EU. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality

[37] Edberg S, Rice E, Karlin R, Allen M. *Escherichia coli*: The best biological drinking water indicator for public health protection. *Journal of Applied Microbiology*. 2000;**88**(S1):106S-116S

[38] Paruch AM, Mæhlum T. Specific features of *Escherichia coli* that distinguish it from coliform and thermotolerant coliform bacteria and define it as the most accurate indicator of faecal contamination in the environment. *Ecological Industries*. 2012;**23**:140-142

[39] Briciu-Burghina Heery CB, Regan F. Continuous fluorometric method for measuring β -glucuronidase activity: Comparative analysis of three fluorogenic substrates. *Analyst*. 2015;**140**(17):5953-5964

[40] Goldberg CS, Strickler KM, Pilliod DS. Moving environmental DNA methods from concept to practice for

monitoring aquatic macroorganisms. *Biological Conservation*. 2015;**183**:1-3

[41] Thomsen PF, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, et al. Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*. 2012;**21**(11):2565-2573

[42] Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, et al. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*. 2017;**26**(21):5872-5895

[43] Thomsen PF, Willerslev E. Environmental DNA—An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*. 2015;**183**:4-18

[44] Beans C. Core concept: Environmental DNA helps researchers track pythons and other stealthy creatures. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;**115**(36):8843-8845

[45] Cai W, Ma Z, Yang C, Wang L, Wang W, Zhao G, et al. Using eDNA to detect the distribution and density of invasive crayfish in the Honghe-Hani rice terrace World Heritage site. *PLoS ONE*. 2017;**12**(5):e0177724

[46] Carlsson JEL, Egan D, Collins PC, Farrell ED, Igoe F, Carlsson J. A qPCR MGB probe based eDNA assay for European freshwater pearl mussel (*Margaritifera margaritifera* L.). *Aquatic Conservation: Marine and Freshwater Ecosystems*. 2017;**27**(6):1341-1344

[47] Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P, et al. Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*. 2011;**6**(8):e23398

- [48] Williams KE, Huyvaert KP, Piaggio AJ. Clearing muddied waters: Capture of environmental DNA from turbid waters. PLOS ONE. 2017;**12**(7):e0179282
- [49] Williams MA. The development of a novel biosensor for single species detection using environmental DNA. PhD thesis; 2022
- [50] Obande GA, Singh KKB. Current and future perspectives on isothermal nucleic acid amplification technologies for diagnosing infections. Infection and Drug Resistance. 2020;**13**:455-483
- [51] Bodulev OL, Sakharov IY. Isothermal nucleic acid amplification techniques and their use in bioanalysis, biochemistry. Biokhimiia. 2020;**85**(2):147
- [52] Li Y, Li L, Fan X, Zou Y, Zhang Y, Wang Q, et al. Development of real-time reverse transcription recombinase polymerase amplification (RPA) for rapid detection of peste des petits ruminants virus in clinical samples and its comparison with real-time PCR test. Scientific Reports. 2018;**8**(1):1-9. DOI: 10.1038/s41598-018-35636-5
- [53] Kaminski MM, Abudayyeh OO, Gootenberg JS, Zhang F, Collins JJ. CRISPR- based diagnostics. Nature Biomedical Engineering. 2021;**8**(7):643-656
- [54] Williams MA, O'Grady J, Ball B, Carlsson J, de Eyto E, McGinnity P, et al. The application of CRISPR-Cas for single species identification from environmental DNA. Molecular Ecology Resources. 2019;**19**(5):1755-0998. DOI: 10.1111/1755-0998.13045
- [55] Cunha DGF, Casali SP, de Falco PB, Thornhill I, Loiselle SA. The contribution of volunteer-based monitoring data to the assessment of harmful phytoplankton blooms in Brazilian urban streams. Science Total Environment. 2017;**584**:586-594
- [56] Loiselle SA, Frost PC, Turak E, Thornhill I. Citizen scientists supporting environmental research priorities. Science Total Environment. 2017;**598**:937
- [57] Hegarty S, Hayes A, Regan F, Bishop I, Clinton R. Using citizen science to understand river water quality while filling data gaps to meet United Nations Sustainable Development Goal 6 objectives. Science Total Environment. 2021;**783**:146953

Anthocyanins and Wine Color: A Complex Story

Encarna Gómez-Plaza

Abstract

Color is one of the most important sensory attributes of a red wine, directly related to the visual acceptance and capable of modulating other sensory attributes such as odor and taste. Anthocyanins are the main compounds responsible for the color of red wines. These are flavonoid compounds that are present in the skins of the grapes and are transferred to must during the winemaking process, especially during crushing and maceration. The color of quality red wines lasts in time; however, anthocyanins are unstable compounds. The reactions and mechanisms behind the maintenance of red wine color during long aging periods are reviewed as well as the different technologies used to increase the amount of anthocyanins in young wines.

Keywords: wine, color, anthocyanins, polymeric compounds, copigmentation, pyranoanthocyanins

1. Introduction

Red wines are a very complex hydroalcoholic matrix, with a wide variety of compounds extracted from grapes and with a high quantity of metabolites released by yeast during the fermentation process. Alcohol and water represent most of the wine composition, however, small quantities of other compounds are the main responsible for the wine color, aroma and flavor.

The color of red wines is a very important feature, we cannot forget that it is the first attribute to be perceived by wine consumers and is directly associated with its quality. Moreover, as stated by Peynaud [1], color also provides information about variety, type of elaboration, defects, the type of aging and the conservation during storage.

The typical normal sequence in wine tasting is to view, smell, and taste. The visual appearance, and color in particular, may have such an impact that it may also influence on smell, which, in turn, may impact on taste. As an example, Morrot et al. [2] conducted a tasting session where three wines were presented to a tasting panel, a white wine, a red wine and the white wine colored red by an odorless and tasteless colorant. The white wine was described using normal terms associated to white wines while the same wine colored red was described using typical red wine descriptors. Also, other experience showed that the aroma description of a wine changed if the tasting was conducted in opaque glasses. The most important conclusion is that the color of the wine plays an important role not only in appearance but also in defining wine attributes and quality [3, 4].

Given the importance of color on wine quality and consumer appreciation, viticulturists, winemakers and scientists dedicate a lot of attention to the grapes (with the enormous polyphenolic variability that may be present in grapes, influenced by ripening, genetic, or environmental factors) and the winemaking process (alcoholic and malolactic fermentation), since wine color starts in the grapes but develops in the winery.

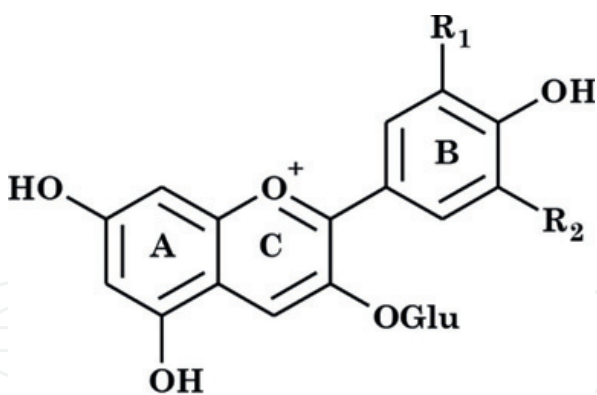
Two types of phenolic compounds are paramount in wine color. The most obvious significant compounds are the anthocyanins, that provide the red color of wines, but also the flavan-3-ols are very important compounds, since they assure the long-term wine color.

2. The colored world of anthocyanins

The anthocyanins are the main pigments in many flowers and fruits, e.g. grapes. They are not only responsible for red color of grapes but for a great diversity of the colors (red, violet and blue) found in nature.

Five anthocyanins are particularly abundant in grapes, cyanidin, delphinidin, malvidin, peonidin, and petunidin. Individually, the color of each anthocyanin is determined by its structure, as an example, those anthocyanins with two substituents in the B ring present orange hues, while those with three substituents present more of a red-purple color. Also, the existence of methoxyl groups in the B-ring affects color: the higher the number of methoxyl groups, the higher the displacement toward purple hues (Figure 1) [5].

In aqueous solution, the color of anthocyanins is strongly dependent on the pH [6]. At very acidic pH (pH~1), anthocyanins are present in their flavylium cation form which has a red color. When the pH is raised to 3–4, the flavylium cation is



Aglicone	R1	R2
Cyanidin	OH	H
Peonidin	OCH ₃	H
Delphinidin		OH OH
Petunidin	OCH ₃	OH
Malvidin	OCH ₃	OCH ₃

Figure 1.
Chemical structure of the main anthocyanins found in grapes and wines.

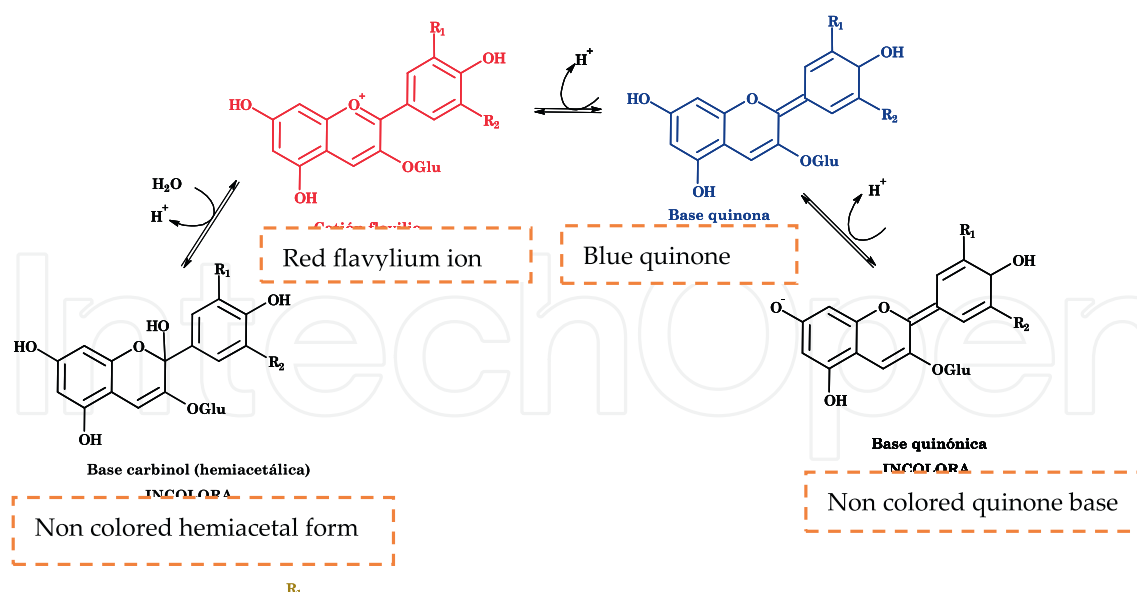


Figure 2.
 pH related structural changes in anthocyanins.

involved in two parallel reactions in equilibrium: deprotonation to form the violet quinonoidal base and hydration at the C-2 position yielding a non-colored hemiketal form. These forms are also in equilibrium with the cis- and trans-chalcone forms which present a yellow color. With a pH rise to values up to 6, the quinonoidal base can be ionized to form the respective blue anionic quinoidal form (**Figure 2**) [6, 7].

Since the flavylium anthocyanin structure is highly unstable due to its deficiency on electrons the free forms of anthocyanins are usually found as glycosylated forms in nature, as these are more stable. Most anthocyanins are 3-O-glycosides and the most common sugar moiety is D-glucose. It has been reported that the stability of anthocyanins increases with the number of methoxyl groups linked to the B-ring, and decreases with the number of hydroxyl groups linked to this ring. Thus, the most stable grape anthocyanidin is malvidin, whereas delphinidin has been shown to be the least stable [8]. Furthermore, the presence of acylation (esterification of the sugar moiety by different organic acids) in the anthocyanin structure also increases its stability and acylated anthocyanins are more resistant against pH changes than the corresponding monoglucosides.

2.1 Localization of anthocyanins in the grape berry and evolution during ripening

These compounds are located in the grape skins, in the first three or four cell layers of the hypodermis and exceptionally in the pulp, in the “teinturier” grapes [9, 10]. At the subcellular level, these water-soluble molecules are found normally within the vacuoles where they can be accumulated in spherical vesicles called “anthocyanoplasts” or “vacuolar anthocyanin inclusions” [11].

Anthocyanins appear in grapes at veraison and accumulate during maturation, generally reaching their highest value around technological maturity, and then decreasing. This decrease usually coincides with the overripening of the grape [12–14]. This maximum accumulation depends on the environment, climatic conditions and the variety.

The variability in grape anthocyanin composition, due to their structural diversity allows to use anthocyanins for discriminating varieties and species. The composition of anthocyanins has led to grape and wine classifications established according

to their content in acylated and non-acylated anthocyanins. Various authors have statistically established the classification of varieties according to their anthocyanin profile, applying multivariate analysis to verify varietal origin [15–21] or try to objectively distinguish the origin of the wines. Within a variety, anthocyanin levels may vary depending on the temperature, on the insolation conditions, and also on the terroir [21]. Also viticultural practices such as the system of fertilization or irrigation can modify the composition anthocyanin [22–25] but certain features of their anthocyanin profile can be tightly linked to variety.

2.2 Transference of anthocyanins from grape to must-wine

During the normal fermentation of red grapes, the grapes are destemmed and crushed without breaking the seeds. The must, along with the skins and seeds is inoculated with an active dried yeast culture and left to ferment. At some point (decided by the enologist) the wine is racked off the skins and the skins are pressed to recover the wine captured in the skins.

The anthocyanins that accumulate in the vacuoles are released from the cells when these grape cell walls are mechanically broken during crushing. Anthocyanins are not stable compounds. Right after these compounds are extracted, they start to change, due to the participation of these compounds in various reactions.

The question is, why the aged wines are still red if anthocyanins are such unstable compounds? Which reactions are occurring in the must and wine that assure the red wine color in the long term?

3. Assuring wine color in the long term

3.1 Copigmentation

Considering the anthocyanin equilibria in aqueous solution and the wine pH (pH 3–4), anthocyanin should be mainly present in a colorless form. Nevertheless, there are some mechanisms, known under the generic name of copigmentation, that contribute to stabilize the colored forms.

The phenomenon of copigmentation is due to non-covalent hydrophobic interactions between the aromatic nuclei of anthocyanins, with other molecules called copigments, normally colorless. It is a spontaneous and exothermic process consisting of the stacking of the copigment on the planar flavylum ion or quinoidal forms of anthocyanins [26]. Structurally, it manifests itself in a vertical anthocyanin-copigment stacking giving rise to a sandwich-like structure, establishing weak Van der Waals-type bonds and hydrophobic interactions between the compounds (**Figure 3**).

These sandwich-like groups generate a hydrophobic environment, due to the glucose molecules of the stacked anthocyanins, and in this way the nucleophilic attack of water molecules is prevented, and therefore the hydration of the anthocyanins is slowed, shifting the balance towards the colored form and therefore increasing the percentage of colored anthocyanins [27, 28].

Copigmentation also changes the absorbance spectrum of the anthocyanin chromophore, leading to an increase in the absorbance in the visible range (hyperchromic effect) and a displacement of the visible wavelength of maximum absorption (bathochromic effect).

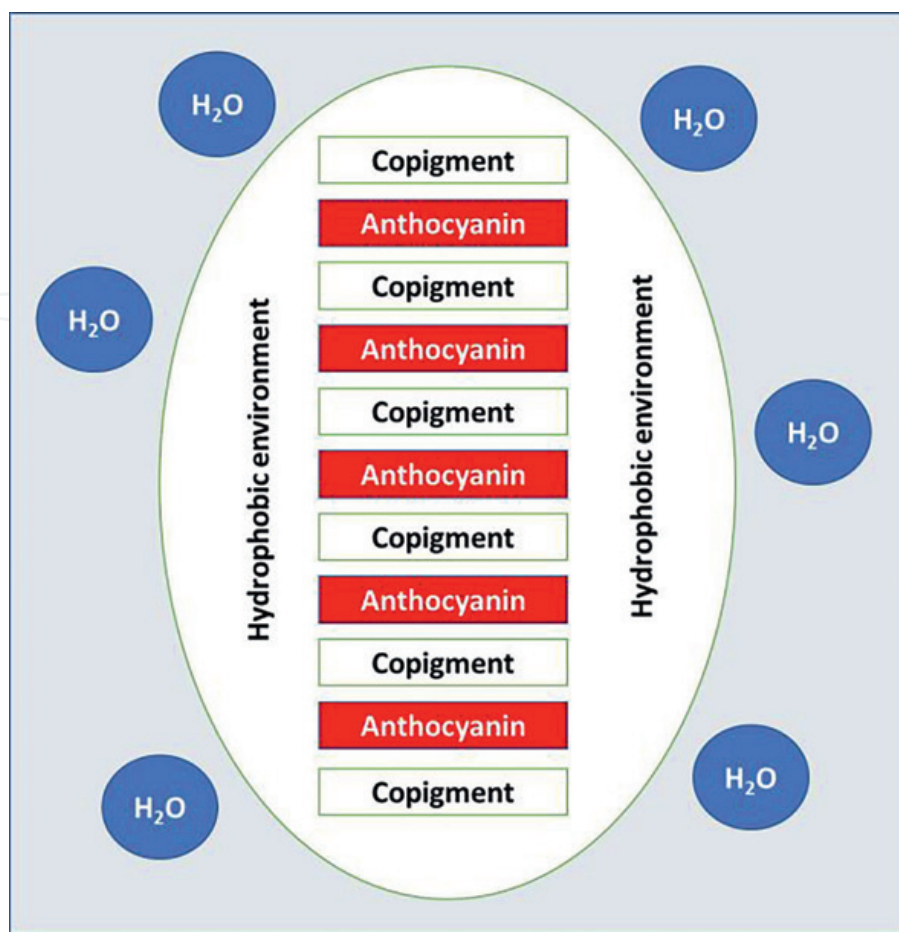


Figure 3.
 Schematic representation of the anthocyanin copigmentation.

There are different types of compounds that can act as anthocyanin copigments, such as amino acids, nucleotides, carbohydrates or phenolic compounds. Flavonoids, in particular flavanols and flavonols, and hydroxycinnamic derivatives are the compounds that could most act as co-pigments due to their larger presence in wines.

Some authors show that the contribution of copigmentation in the color of young red wines can reach 30–50%, since at the pH of the wine, about 80% of the anthocyanins are in a colorless hydrated form [29].

The stability of the co-pigmentation complex in wines is limited and can be affected by several parameters:

Storage temperature: the increase in temperature produces a decrease in the copigmentation process and therefore affects the color [29–31].

pH: by increasing the pH there is a lower contribution of the anthocyanins in the form of the flavillium cation, therefore the co-pigmentation complex is also lowered [29, 32].

Ethanol content: the increase in ethanol favors the rupture of the copigmentation complex, although in the range of usual ethanol concentrations in wine, this influence is limited [29, 33].

Nature and structure of the cofactor: In wine, several compounds can act as cofactors such as hydroxycinnamic acids (caffeic, p-coumaric and ferulic acids), flavonols and flavanols. Flavonols are very good cofactors, although they are found in small quantities in wines [34, 35]. Flavan-3-ols, both monomers and oligomers intervene in co-pigmentation, although polymeric flavan-3-ols are less effective

as co-pigments. [30, 36, 37]. Among monomers, epicatechin is considered a better co-pigment than catechin [27, 38], dimer procyanidins with C4–C6 bonds appear to behave as better co-pigments than their respective dimers with C4–C8 bonds, possibly due to their more open and flexible conformation and the esterification of flavanols with gallic acid, as in the case of (–) - epicatechin-3-O-gallate, improves their capacity as copigments [27, 36].

Molar cofactor / anthocyanin ratio: increasing this ratio produces an increase in the effects produced by copigmentation [29, 30, 34, 39]. For this reason, copigmentation can be improved in wines by the prefermentative addition of substances that could act as cofactors. Some authors reported that the prefermentative addition of (+)-catechin produces increases in wine color from 9 to 13% and from 36 to 60% with caffeic acid [34, 35]. However this addition is not a regulated practice in enology. As an example of authorised practices, the addition of enological tannins also led to an improvement of wine chromatic characteristics and its stability [40]. Also, a covinification of some grape varieties could be an interesting practice for obtaining a richer phenolic composition. In this regard, covinification of Tempranillo and Graciano varieties results in higher copigmentation values than in the corresponding monovarietal wines [41]. Also, the addition of wood derivatives during the maceration process has been evaluated as an alternative technological application to modulate the phenolic composition and color characteristics of red wines, especially in warm climate regions [42].

3.2 Formation of anthocyanin-derived and polymeric compounds

During aging, the color due to copigmentation decreases [34, 43]. This is also accompanied by a decrease in monomeric anthocyanin concentration, since they can be involved in degradation reactions, such as oxidation, but color not always decrease, even when copigmentation diminish. Studies have shown that the reaction of anthocyanins with other wine components leads to the formation of red anthocyanin-derived compounds and polymerized compounds.

Pyranoanthocyanins: They are anthocyanin-derived compounds, formed by nucleophilic cycloaddition of a polarizable double bond of a compound with an anthocyanin. The formation of the pyran ring in the anthocyanin structure makes it less sensitive to pH variations than the anthocyanins from which they originate, due to the fact that the C4 position is blocked. Moreover, these compounds have shown to express more color than the corresponding anthocyanin [44].

Figure 4 shows the different types of pyranoanthocyanins that appear in wines, whose structure will depend on the compound that incorporates to the anthocyanin:

- Carboxypyrananthocyanins, generated by the addition of pyruvic acid or acetaldehyde.
- Phenylpyrananthocyanins, generated by the addition of vinylphenol derivatives from hydroxycinnamic acids.
- Flavanylpyrananthocyanins, whose precursors are vinylflavanols.

Their concentration in wines will depend on the initial wine anthocyanin composition, the concentration of precursors, as well as the presence of small amounts of oxygen, since an oxidation stage intervenes in their formation [44].

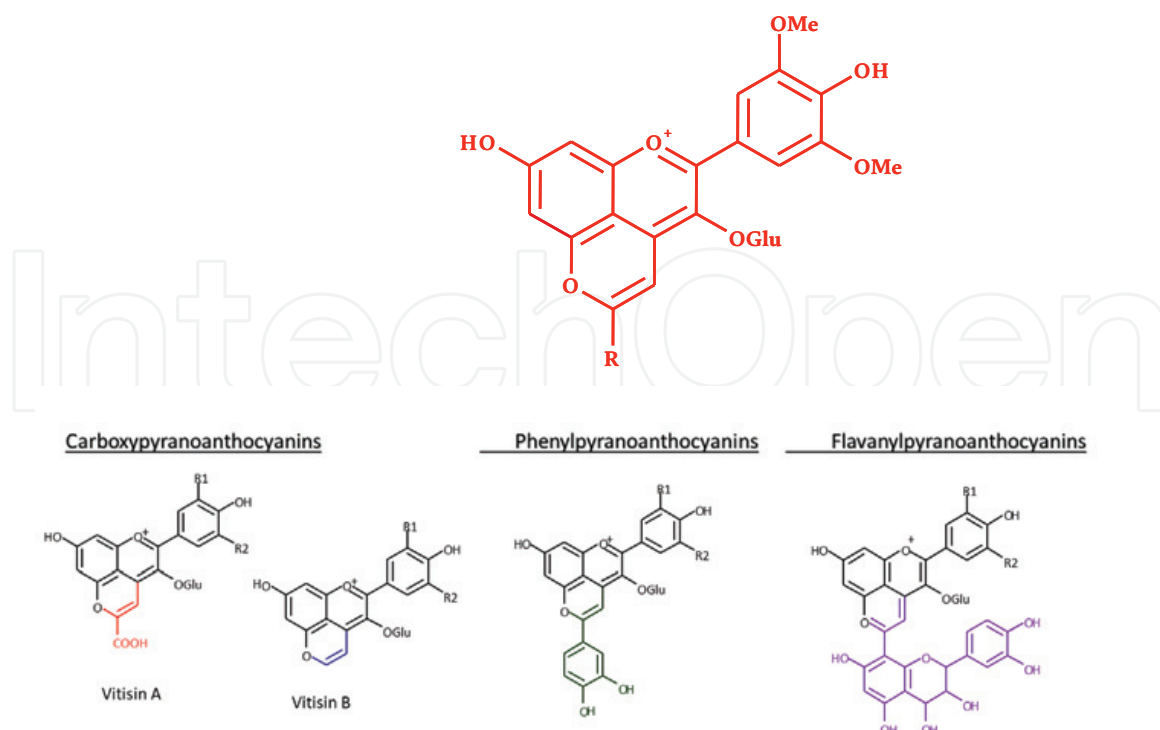


Figure 4.
Pyranoanthocyanidin structure and main types of pyranoanthocyanins.

The pyranoanthocyanins generated by the cycloaddition of acetaldehyde or pyruvic acid to the monomeric anthocyanin make up the so-called vitisins or carboxypyrananthocyanins.

The carboxypyrananthocyan obtained by the cycloaddition of pyruvic acid and malvidin-3-glucoside is called vitisin A. Furthermore, pyruvic derivatives of cyanidin, delphinidin, petunidin and peonidin have been identified in wine, both from monoglycosides and their acylated forms [44–47]. Vitisin A has a λ_{max} of 511 nm.

The structure of vitisin B is the decarboxylated structure of vitisin A, generated by the cycloaddition of acetaldehyde to malvidin-3-glucoside. Later, coumarylvitisin B was also detected in wine [44, 46–50]. Vitisin B has a maximum wavelength ~ 20 nm shorter than vitisin A.

The formation of vitisin A and B depends on the anthocyanin composition, pH, concentration of pyruvic acid and acetaldehyde. Pyruvic acid and to a lesser extent acetaldehyde appear during alcoholic fermentation by the action of yeasts. Vitisin A is the most rapidly formed carboxypyrananthocyanin and the first to be detected in young wines [51], On the other hand, vitisin B begins to be synthesized at the end of alcoholic fermentation, since acetaldehyde production is proportional to the amount of fermented sugar, being higher at the end of alcoholic fermentation.

Vitisins are more stable than monomeric anthocyanins and present a slower degradation rate than monomeric anthocyanins. After one year of aging, the amount of monoglycoside anthocyanins decreases between 80 and 90%, while vitisin A decreases between 15 and 25% and after 30 months, the monomeric anthocyanins in the wine have practically disappeared, while there are still between 20 and 35% of the initial amount of vitisins [52], even increases have been found in oxidative aging increases significantly [47, 53].

The phenylpyrananthocyanins are formed by the addition of hydroxycinnamic acids (p-coumaric, caffeic, ferulic and synapic acid) or their decarboxylation

products (4-vinylphenol, 4-vinylcatechol, 4-vinylguayacol and 4-vinyl syringol) to a monomeric anthocyanin.

The flavanlypyranoanthocyanins were initially synthesized in a model solution from malvidin-3-glucoside, acetaldehyde and flavanols (both monomeric such as (+) - catechin and (–) - epicatechin, as dimeric such as procyanidin B2 and B3), obtaining small amounts of malvidin- 3-glucoside-vinyl (epi) catechin and malvidin-3-glucoside-vinyldicatechin. Later they were identified in Port wines and red wines [44–47, 50, 54]. Like the previous pyranoanthocyanins (carboxypyrananthocyanins and phenylpyranoanthocyanins), these pigments are more resistant to changes in pH or discoloration by hydration than free anthocyanins and do not suffer discoloration by SO₂.

The age of the wines influences the concentration of flavanlypyranoanthocyanin with a flavanol unit, which increased with the years of aging of the wine, being the most important pigments in 6-year-old wines, possibly due to the degradation of catechin-dimer flavanlypyrantocyanins whose concentrations they were lower in older wines [54].

3.3 Polymeric compounds

Other reactions that may occur with anthocyanins are the formation of polymeric pigments, that are, in general, more stable to pH variations, and temperatures than monomeric anthocyanins, providing stability to the color of wine. These polymerization reactions mainly involve anthocyanins and polymeric flavan-3-ols, commonly known as wine tannins [55].

Direct condensations between phenolic compounds: The condensations between anthocyanins and tannins are nucleophilic addition reactions, since the phenolic compound (anthocyanin or tannin) can act as an electrophile (compound with a poor electron density) and be attacked by nucleophilic agents (compounds with an excess in electron density). Anthocyanins behave as electrophiles when they are in the form of the flavillium cation, on the other hand, their hydrated forms act as nucleophiles. Also, tannins can intervene as nucleophiles since their monomers have a negative charge density in their C6 and C8 positions, while the breaking of the interflavanol bond (CC) between the monomers that constitute tannin, an intermediate adduct is generated which operates as an electrophile [56]. Therefore, the condensation between anthocyanins and tannins presents two types of products; tannin-anthocyanin (T-A) and anthocyanin-tannin (A-T). They present a color similar to free anthocyanins and their existence has been confirmed in wine [56, 57].

Tannin-anthocyanin condensation (T-A): Procyanidin or tannin in an acid medium such as wine can undergo hydrolysis, a phenomenon that occurs spontaneously, generating an intermediate adduct, a carbocation or active catechin. This is an electrophilic structure that reacts with anthocyanin in the hemiacetal form (nucleophile), giving rise to a colorless compound, which after dehydration presents an orange color. T-A condensation occurs in the absence of oxygen, thus explaining its appearance in wines subjected to reducing environments, such as a bottle or tank [57].

Anthocyanin-tannin condensation (A-T): The mechanism consists of the union of the C4 position of the anthocyanin in the form of a flavillium cation (electrophile) with the negative positions C6 or C8 of the tannins (nucleophile), forming a colorless AT dimer which after its oxidation, facilitated by the presence of oxygen, it takes on a red color (A+T) [56]. It has been shown in wine that these compounds have greater stability than monomeric anthocyanins, and their percentage increases during the maturation time [57].

Acetaldehyde-mediated condensation between phenolic compounds: The formation mechanism of this polymerization consists of the activation of the acetaldehyde in the form of carbocation, which reacts with tannin in positions C6 or C8, giving a cationic aldehyde-flavanol adduct. Subsequently, an attack occurs on an anthocyanin in hydrated or hemiacetalic form at its C8 position [56]. Upon deprotonation, the violet colored quinoidal base is formed. Yeasts produce acetaldehyde under anaerobic conditions, at low pH, high sugar concentration, presence of SO₂ and at low temperatures. Acetaldehyde is also a by-product of the metabolism of acetic and lactic bacteria [58] and it can also be generated in wine by the oxidation of ethanol. The presence of this condensation compounds in wines produces an increase in color [59–61] and they are more resistant to SO₂ discoloration or hydration than free anthocyanins, not because the C2 and C4 positions of the anthocyanin are occupied, but because of steric hindrances produced by the monomers surrounding the anthocyanin [59]. Its formation is faster than direct condensation between anthocyanin-tannins, but these compounds are less stable. The rupture of the ethyl bridge leads again to the acetaldehyde-flavanol cation adduct (ethylflavanol) that can subsequently intervene in new condensations or generate, after dehydration, pyranoanthocyanins (flavanyl-pyranoanthocyanins) [59]. For this reason, the presence of this compounds can play an important role as more stable pyranoanthocyanic precursors.

4. Technologies to increase the extraction of anthocyanins and the wine color

As stated before, an essential step in the red winemaking process is the extraction of the colored compounds, along with tannins and aromatic molecules, from skins to the must during maceration. This means that it is necessary to keep skins and seeds in contact with the must in the tank during fermentation to extract and diffuse these phenolic compounds in the must and to obtain the end color of the wine and this extraction will be affected by several technological factors such as fermentation temperature, cap management, SO₂, ethanol content. There are several techniques that may favor the extraction and increase de concentration of anthocyanins in musts and wine, assuring enough substrate for long-term highly colored wines.

Cold maceration: The extraction kinetics can also be improved with enological techniques such as the use of cold soak (cold maceration), dry ice, or enzymes affecting the extraction of anthocyanins and proanthocyanidins. Cold soak is based on the use of low maceration temperatures, 4–15°C [62]. It applies refrigeration to the must in order to delay the beginning of fermentation and that means that maceration is done with the absence of ethanol, so extraction of anthocyanins is more favored than tannin extraction.

Enzymes: They are a commercial mixture of different enzymes with different pectolytic activities, mainly endo-polygalacturonase, pectin-methyl esterase and pectin-lyase activities, that are added to the crushed grapes with the purpose is to disassemble the cell wall structure, thus facilitating extraction of the compounds located inside the grape cells [63]. These enzymes have been shown to have a positive effect in several studies [64–66].

The use of high temperatures (thermovinification): Although generally heating of the grapes before fermentation is called “thermovinification”, different pre-fermentation heating processes are currently being applied in wineries. These techniques can be classified into two groups, depending on whether the cooling of the grapes,

similarly to heating, is conducted using heat exchangers, or whether the cooling is conducted into a vacuum chamber.

The first kind of process is designated as “thermovinification”. The second group involves the technique called thermo-flash, flash détente or flash-release [67].

During thermovinification the grape mash is heated up to 85°C, and both heating and cooling are conducted in heat exchangers [68]. Heating is conducted for a period of time of less than one hour, after which the grape mash is pressed to separate the solid parts and perform fermentation as for white wine. If heating at the same temperature is extended for a longer period of time (up to 24 h), and the fermentation is conducted in the presence or absence of the solid phase, the process is called pre-fermentation hot maceration.

The process called “flash release” or “flash détente” consists in rapidly heating the grapes at temperatures between 85–95°C by a direct injection of steam. Grapes are then introduced into a vacuum that instantly vaporizes the water, thereby cooling the treated grapes and weakening their skin cell envelopes by boiling the water inside the cells [69, 70]. This effect on the skin cells enhances extractability in subsequent fermentation process that may be conducted with or without the solid parts of the grapes. Flash technology differs from traditional thermo-vinification because the traditional method does not involve a vacuum and there is no flash water waste produced.

Emerging technologies: There are other emerging technologies such as high hydrostatic pressure (HHP), pulsed electric fields (PEFs), ultrasounds (USs), irradiation, pulsed light (PL), ozone and electrolyzed water are opening up new possibilities in wine technology. Among them, US and PEF are already used technologies and they can improve the extraction of anthocyanins from skins to must.

Pulsed Electric Fields: PEFs increase the rate and yield in the extraction of phenolic compounds in red grape processing [71]. PEF processing consists in the intermittent application of short duration pulses (ms-s) of high voltage (kV) to the crushed grapes located between two electrodes. The applied external voltage generates an electric field whose strength depends not only on voltage intensity, but also on the distance between the electrodes. When exposed to a sufficiently strong electric field, the cell membrane undergoes a phenomenon called electroporation, consisting in the increment of cell envelope permeability as a consequence of the formation of pores in the cytoplasmic membrane [72]. By using PEFs, the extraction of total phenolic compounds can be increased by 10–50% [73] and maceration time can be reduced by 40–50% using PEF intensities of 5–10 kV/cm [74]. Under these conditions, the temperature increment is very low, so it can be considered a nonthermal technology.

Ultrasounds: USs are sonic waves in the range of 20–100 kHz producing cavitation, which is the formation of gas bubbles that are expanded and compressed according to the variation of the wave intensity. Through this phenomenon, local pressure can reach values of 50 MPa and temperatures at specific points can also exceed 1000°C [75]. USs help to destroy cell walls of skins, facilitating the extraction of phenolic compounds. This technology can be used in enology to carry out a fast and continuous extraction, extremely reducing maceration times, and the extracted juice can later be fermented at low temperature in the absence of solids (skins and seeds). The best extraction results can be obtained at low frequencies (30 kHz) inside the USs range. Using this technology, it is possible to reduce maceration time from 7 to 2–3 days with the same degree of extraction in terms of tannins and pigments [76].

The emergence of non-thermal technologies can lead to high quality products while saving energy. Most of these technologies are locally clean processes and

therefore appear to be more environment-friendly, with less environmental impact than traditional ones.

As a conclusion, the importance of anthocyanins and their reactivity in must and wines has been reviewed, indicating the importance of color in wine appreciation and how these reactions are paramount for a long-term wine color stability.

Conflict of interest


The author declares there is no conflict of interest.

Author details

Encarna Gómez-Plaza
University of Murcia, Murcia, Spain

*Address all correspondence to: encarna.gomez@um.es

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Peynaud E. The Taste of Wine. 2nd ed. New York: Wiley; 1996
- [2] Morrot G, Brochet F, Dubourdieu D. The color of odors. *Brain Lang.* 2001;**79**:309-320
- [3] Gonzalez-Miret ML, Wei J, Luo R, Hutchings J, Heredia F. Measuring colour appearance of red wines. *Food Quality and Preference.* 2007;**18**:79-101
- [4] Parr W, White K, Heatherbell D. The nose knows: Influence of colour on perception of wine aroma. *Journal of Wine Research.* 2003;**14**:79-101
- [5] Khoo H, Azlan A, Tang S, Meng LS. Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutrition Research.* 2017;**61**:1361779
- [6] Brouillard R, Chassaing S, Fougèrouse A. Why are grape/fresh wine anthocyanins so simple and why is it that red wine color lasts so long? *Phytochemistry.* 2003;**64**:1179-1186
- [7] Santos-Buelga C, Francia-Aricha E, Rivas-Gonzalo JC. Role of flavan-3-ol structure on direct condensation with anthocyanins. In: Vercauteren J, Chèze C, Dumon MC, Weber JF, editors. *Polyphenols 96.* Paris: XVIIIe Journées Internationales Groupe Polyphénols; 1996. pp. 9-10
- [8] Yang P, Yuan C, Wang H, Han F, Liu Y, Wang L, et al. Stability of anthocyanins and their degradation products from cabernet sauvignon red wine under gastrointestinal pH and temperature conditions. *Molecules.* 2018;**23**:354
- [9] Moskowitz A, Hrazdina G. Vacuolar contents of fruit subepidermal cells from *Vitis* species. *Plant Physiology.* 1981;**68**:686-692
- [10] Souquet J, Cheynier V, Broussaud F, Moutounet M. Polymeric proanthocyanidins from grape skins. *Phytochemistry.* 1996;**43**:509-512
- [11] Markham K, Gould K, Winefield S, Mitchell K, Bloor S, Boase MR. Anthocyanic vacuolar inclusions—Their nature and significance in flower colouration. *Phytochemistry.* 2000;**55**:327-336
- [12] Ribereau-Gayon P. Evolution des composés phénoliques au cours de la maturation du raisin. *Connaissance Vigne Vin.* 1972;**2**:161-175
- [13] Pirie A, Mullins MG. Concentration of phenolics in the skin of grape berries during fruit development and ripening. *American Journal of Enology and Viticulture.* 1980;**31**:34-36
- [14] Glories Y. La maturità fenolica delle uve: primo parametro da controllare per una corretta vinificazione in rosso. *Vignevini.* 1999;**3**:46-50
- [15] González-Sanjosé ML, Santa-Maria G, Díez C. Anthocyanins as parameters for differentiating wines by grape variety, wine-growing region, and wine-making methods. *Journal of Food Composition and Analysis.* 1990;**3**:54-66
- [16] Katalinic V, Males P. Compositional changes in grape polyphenols throughout maturation. *Journal of Wine Research.* 1997;**8**:169-177
- [17] Holbach B, Marx R, Ackermann M. Determination of anthocyanins composition of red wine by HPLC. *Lebensmittelchemie.* 1997;**51**:78-80

- [18] Hesford F, Schneider K. Anthocyanins, the natural red coloring matter of the wine. *Schweizerische Zeitschrift für Obst- und Weinbau*. 1997;**133**:559-561
- [19] Kallithraka S, Arvanitoyannis IS, Kefalas P, El-Zajouli A, Soufleros E, Psarra E. Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin. *Food Chemistry*. 2001;**73**:501-514
- [20] García-Beneytez E, Revilla E, Cabello C. Anthocyanin pattern of several red grape cultivars and wines made from them. *European Food Research and Technology*. 2002;**215**:32-37
- [21] Ortega-Regules A, Romero-Cascales I, López-Roca JM, Ros-Garcia JM, Gómez-Plaza E. Anthocyanin fingerprint of grapes: Environmental and genetic variations. *Journal of the Science of Food and Agriculture*. 2006;**86**:1460-1467
- [22] Delgado R, González MR, Martín P. Effects d'interaction entre fertilisation azotée et fertilisation potassique sur la composition anthocyanique et les caractéristiques chromatiques du raisin cv. Tempranillo. *Journal International des Sciences de la Vigne et du Vin*. 2006;**40**:141-150
- [23] Wu L, Li P, Jia H, Phillip F, Bao X, Zhao F, et al. The effect of foliar application of K₂SO₄ or KH₂PO₄ on skin color of the 'Kyoho' Grape. *Agronomy*. 2021;**11**:2361
- [24] Kyrleou M, Koundouras S, Kallithraka S, Theodorou N, Proxenia N, Kotseridis Y. Effect of irrigation regime on anthocyanin content and antioxidant activity of *Vitis vinifera* L. cv. Syrah grapes under semiarid conditions. *Journal of the Science of Food and Agriculture*. 2016;**96**:988-996
- [25] Ju Y, Yang B, He S, Tu T, Min Z, Fang Y, et al. Anthocyanin accumulation and biosynthesis are modulated by regulated deficit irrigation in Cabernet Sauvignon (*Vitis vinifera* L.) grapes and wines. *Plant Physiology and Biochemistry*. 2019;**135**:469-479
- [26] Escribano T, Buelga S, Celestino. Anthocyanin copigmentation—Evaluation, mechanisms and implications for the colour of red wines. *Current Organic Chemistry*. 2012;**16**:715-723
- [27] Liao H, Cal Y, Haslam E. Polyphenols interactions. Anthocyanins: Copigmentation and colour changes in young red wines. *Journal of the Science of Food and Agriculture*. 1992;**59**:299-305
- [28] Mirabel M, Saucier C, Guerra C, Glories Y. Copigmentation in model wine solutions. Occurrence and relation to wine aging. *American Journal of Enology and Viticulture*. 1999;**50**:211-218
- [29] Boulton R. The copigmentation of anthocyanins and its role in the color of red wine: A critical review. *American Journal of Enology and Viticulture*. 2001;**52**:67-87
- [30] Bakowska A, Kucharska A, Oszmianski J. The effect of heating, UV irradiation and storage on stability of the anthocyanin-polyphenol copigment complex. *Food Chemistry*. 2003;**81**:349-355
- [31] Mazza G, Brouillard R. The mechanism of co-pigmentation of anthocyanins in aqueous solutions. *Phytochemistry*. 1990;**29**:1097-1102
- [32] Gris E, Ferreira E, Falcao L, Bordignon-Luiz MT. Caffeic acid copigmentation of anthocyanins from Cabernet Sauvignon grape extracts in model systems. *Food Chemistry*. 2007;**100**:1289-1296

- [33] Hermosín-Gutiérrez I. Influence of ethanol content on the extent of copigmentation in a Cencibel young red wine. *Journal of Agriculture and Food Chemistry*. 2003;**51**:4079-4083
- [34] Darias-Martín J, Carrillo M, Díaz E, Boulton R. Enhancement of red wine colour by pre-fermentation addition of copigments. *Food Chemistry*. 2001;**73**:217-220
- [35] Fernández-Iglesias CM, López-Roca JM, Gómez-Plaza E. La copigmentación en vinos tintos. Efecto de la adición de cofactores. *Enólogos*. 2004;**29**:34-38
- [36] Berké B, Freitas AP. Influence of procyanidin structures on their ability to complex with oenin. *Food Chemistry*. 2005;**90**:453-460
- [37] Kunsági-Máte S, Szabó K, Nikfardjam MP, Kollár L. Determination of the thermodynamic parameters of the complex formation between malvidin-3-O-glucoside and polyphenols. Copigmentation effect in red wines. *Journal of Biochemical and Biophysical Methods*. 2006;**69**:113-119
- [38] Brouillard R, Wigand MC, Dangles O, Cheminat A. pH and solvent effects on the copigmentation reaction of malvin with polyphenols, purine and pyrimidine-derivatives. *Journal of the Chemical Society-Perkin Transactions*. 1991;**2**:1235-1241
- [39] González-Manzano S, Mateus N, Freitas V, Santos-Buelga C. Influence of the degree of polymerisation in the ability of catechin to act as anthocyanin copigments. *European Food Research and Technology*. 2008;**227**:83-92
- [40] Alcalde-Eon C, García-Estévez I, Ferreras-Charro R, Rivas-Gonzalo JC, Escribano-Bailón MT. Adding oenological tannin vs. overripe grapes: Effect on the phenolic composition of red wines. *Food Composition and Analysis*. 2014;**34**:99-133
- [41] García-Marino M, Escudero-Gilete ML, Heredia F, Escribano T, Rivas-Gonzalo J. Color-copigmentation study by tristimulus colorimetry (CIELAB) in red wines obtained from Tempranillo and Graciano varieties. *Food Research International*. 2013;**51**:123-131
- [42] Gordillo B, Baca-Bocanegra B, Rodríguez-Pulido F, González-Miret M, García-Estévez I, Quijada-Morín N, et al. Optimization of an oak chips-grape mix maceration process. Influence of chip dose and maceration time. *Food Chemistry*. 2016;**206**:249-259
- [43] Talcott S, Brenes C, Pires D, Del Pozo-Infran D. Phytochemical stability and color retention of copigmented and processed muscadine grape juice. *Journal of Agriculture and Food Chemistry*. 2003;**51**:957-963
- [44] Atanasova V, Fulcrand H, Cheynier V, Moutounet M. Effect of oxygenation on polyphenol changes occurring in the course of wine-making. *Analytica Chimica Acta*. 2002;**458**:15-27
- [45] Wang H, Race E, Shirinhand J. Anthocyanin transformation in Cabernet Sauvignon wine during aging. *Journal of Agriculture and Food Chemistry*. 2003;**51**:7989-7994
- [46] Alcalde-Eon C, Escribano-Bailón MT, Santos-Buelga C, Rivas-Gonzalo JC. Separation of pyranoanthocyanins from red wine by column chromatography. *Analytica Chimica Acta*. 2004;**513**:305-318
- [47] Alcalde-Eon C, Boido E, Carrau F, Dellacassa E, Rivas-Gonzalo JC. Pigment profiles in monovarietal wines produced

in Uruguay. *American Journal of Enology and Viticulture*. 2006;**57**:449-459

[48] Revilla I, Pérez-Magariño S, González-Sanjosé ML, Beltran S. Identification of anthocyanin derivatives in grape skin extracts and red wines by liqued chromatography with diode array and mass spectrometric detection. *Journal of Chromatography A*. 1999;**847**:83-90

[49] Morata A, Gómez-Cordoves C, Colomo B, Suárez JA. Pyruvic acid and acetaldehyde by different strains of *Saccharomyces cerevisiae*: Relationship with vitisin A and B formation in red wines. *Journal of Agriculture and Food Chemistry*. 2003;**51**:7402-7409

[50] Boido E, Alcalde-Eon C, Carrau F, Dellacassa E, Rivas-Gonzalo JC. Aging effect of the pigment composition and color of *Vitis vinifera* L. c.v. Tannat wines. Contribution of the main pigment families to wine color. *Journal of Agriculture and Food Chemistry*. 2006;**54**:6692-6704

[51] Revilla I, González-Sanjosé ML. Evolution during the storage of red wines treated with pectolytic enzymes: New anthocyanin pigment formation. *Journal of Wine Research*. 2001;**12**:183-197

[52] Mateus N, de Freitas V. Evolution and stability of anthocyanin-derived pigments during Port wine aging. *Journal of Agriculture and Food Chemistry*. 2001;**49**:5217-5222

[53] Del Álamo-Sanza M, Nevares I. Wine aging in bottle from artificial systems (staves and chips) and oak woods. Anthocyanin composition. *Analytica Chimica Acta*. 2006;**513**:229-237

[54] He J, Santos-Buelga C, Mateus N, De Freitas V. Isolation and quantification of oligomeric pyranoanthocyanin-flavanol

pigments from red wines by combination of column chromatographic techniques. *Journal of Chromatography A*. 2006;**1134**:215-225

[55] Pati S, Losito I, Gambacorta G, La Notte E, Palmisano F, Zambonin GP. Simultaneous separation and identification of oligomeric procyanindins and anthocyanin-derived pigments in raw red wine by HPLC-UV-ESI-MS. *Journal of Mass Spectrometry*. 2006;**41**:861-871

[56] Salas E, Fulcrand H, Meudec E, Cheynier V. Reactions of Anthocyanins and Tannins in Model Solutions. *Journal of Agricultural and Food Chemistry*. 2004;**51**:7951-7961

[57] Remy S, Fulcrand H, Labarbe B, Cheynier V, Moutonet M, Boido E, et al. First confirmation in red wine of products resulting from direct anthocyanin-tannin reactions. *Journal of the Science of Food and Agriculture*. 2000;**80**:745-751

[58] Liu SQ, Pilone GJ. An overview of formation and roles of acetaldehyde in winemaking with emphasis on microbiological implications. *International Journal of Food Science and Technology*. 2000;**35**:49-61

[59] Escribano-Bailón T, Álvarez-García M, Rivas-Gonzalo JC, Heredia FJ, Santos-Buelga C. Color and stability of pigments derived from the acetaldehyde-mediated condensation between malvidin-3-O-glucoside and (+)-catechin. *Journal of Agriculture and Food Chemistry*. 2001;**49**:1213-1217

[60] Es-Safi N, Cheynier V, Moutounet M. Role of aldehydic derivatives in the condensation of phenolic compounds with emphasis on the sensorial properties of fruti-derived foods. *Journal of Agriculture and Food Chemistry*. 2002;**50**:5571-5585

- [61] Lee DF, Swinny EE, Jones GP. NMR identification of ethyl-linked anthocyanin-flavanol pigments formed in model wine ferments. *Tetrahedron Letters*. 2004;**45**:1671-1674
- [62] Panprivech S, Lerno L, Brenneman C, Block D, Oberholster A. Investigating the effect of cold soak duration on phenolic extraction during Cabernet Sauvignon fermentation. *Molecules*. 2015;**20**:7974-7989
- [63] Pinelo M, Arnous A, Meyer AS. Upgrading of grape skins: Significance of plant cell-wall structural components and extraction techniques for phenol release. *Trends in Food Science & Technology*. 2006;**17**:579-590
- [64] Romero I, Ros JM, López JM, Gómez E. The effect of a commercial pectolytic enzyme on grape skin cell wall degradation and color evolution during the maceration process. *Food Chemistry*. 2012;**130**:626-631
- [65] Gil R, Moreno A, Vila R, Fernández JI, Martínez A, Gómez E. Influence of low temperature prefermentative techniques on chromatic and phenolic characteristics of Syrah and Cabernet Sauvignon Blanc. *European Food Research and Technology*. 2009;**228**:777-788
- [66] Belda I, Conchillo L, Ruiz J, Navascués E, Marquina D, Santos A. Selection and use of pectinolytic yeasts for improving clarification and phenolic extraction in winemaking. *International Journal of Food Microbiology*. 2016;**223**:1-8
- [67] Maza M, Alvarez I, Raso J. Thermal and non-thermal physical methods for improving polyphenol extraction in red winemaking. *Beverages*. 2019;**5**:47
- [68] Nordestgaard S. Fermentation: Pre-fermentation heating of red grapes: A useful tool to manage compressed vintages? *Australian and New Zealand Grapegrower and Winemaker*. 2017;**637**:54-61
- [69] Ageron D, Escudier JL, Abbal P, Moutounet M. Prétraitement des raisins par flash détente sous vide poussé. *Revue Française d'Oenologie*. 1995;**35**:50-53
- [70] Doco T, Williams P, Cheynier V. Effect of flash release and pectinolytic enzyme treatments on wine polysaccharide composition. *Journal of Agricultural and Food Chemistry*. 2007;**55**:6643-6649
- [71] Garde-Cerdán T, Martí-Belloso O, Ancín-Azpilicueta C. Pulsed electric field and fermentation. In: Shikha Ojha K, Tiwari B, editors. *Novel Food Fermentation Technologies*. New York: Springer International Publishing; 2016. pp. 85-123
- [72] Tsong TY. Electroporation of cell membranes. In: Neumann E, Sowers AE, Jordan CA, editors. *Electroporation and Electrofusion in Cell Biology*. Boston, MA, USA: Springer; 1989. pp. 149-163
- [73] Yang N, Huang K, Lyu C, Wang J. Pulsed electric field technology in the manufacturing processes of wine, beer, and rice wine: A review. *Food Control*. 2016;**61**:28-38
- [74] López-Alfaro I, González-Arenzana L, López N, Santamaría P, López R, Garde-Cerdán T. Pulsed electric field treatment enhanced stilbene content in Graciano, Tempranillo and Grenache grape varieties. *Food Chemistry*. 2013;**141**:3759-3765
- [75] Chemat F, Khan MK. Applications of ultrasound in food technology: Processing, preservation and extraction. *Ultrasonic Sonochemistry*. 2011;**18**:813-835

[76] Bautista-Ortín A,
Jiménez-Martínez MD, Jurado R,
Iniesta JA, Terrades S, Andrés A, et al.
Application of high-power ultrasounds
during red wine vinification.
International Journal of Food Science
and Technology. 2017;52:1314-1323

Neurodegeneration Triggered by Inhaled Nanopollutants

Cristina Hermosillo-Abundis, Miguel A. Méndez-Rojas and Oscar Arias-Carrión

Abstract

Nanoparticles, also known as ultrafine particles, present in polluted air have shown a detrimental effect on human health, especially on respiratory and cardiovascular functions. They are considered nanopollutants and, given their small size and unique physicochemical properties, it has been suggested that they could reach the brain through the olfactory nerve pathway. Neurodegeneration is on the rise, especially in highly polluted cities, and a variety of nanoparticles present in the air have been related. Recent findings on airborne nanoparticles translocation routes and neurodegenerative processes may allow us to better understand the impact of nanopollutants in human health. The outcomes may have major implications in the development of public policies and regulations required to protect human health from exposition to nanomaterials, produced either incidentally or by design.

Keywords: nanopollutants, nose-brain translocation, neurodegeneration, inhaled, ultrafine particles

1. Introduction

Environmental pollution is silently responsible for millions of deaths worldwide, by the year 2060 it is expected to cause between 6 and 9 million deaths; surpassing even the deaths related to tobacco [1]. Exposure to air pollution has been linked to harmful effects on health such as cardiovascular diseases, ischemia, heart attacks, lung cancer and an may increase the risk and trigger neurodegenerative diseases such as multiple sclerosis, Parkinson's disease, and Alzheimer's disease. It is also related to the onset of childhood neurodevelopmental disorders such as attention deficit, autism and the onset of schizophrenia [1–5].

A diverse mixture of suspended matter or particulate matter (PM) including man-made nanomaterials, gases, and organic and inorganic compounds is found in environmental pollution; suspended matter can be classified according to aerodynamic diameter or particle size. Particles bigger than $2.5\ \mu\text{m}$ are considered coarse particles, those smaller than $2.5\ \mu\text{m}$ fine particles, and those smaller than $0.1\ \mu\text{m}$ are classified as ultrafine particulate matter or ultrafine particles (UFP or UFPM for short) [3, 6]. The latter are of special interest due to their potential toxic effects since their small size and physicochemical properties give them the possibility of interacting with

different organs to which no other particle has access and, in addition, these interactions are different from those that the same material of a larger size could have. For example, their elimination can be considerably slowed down, which is linked to bioaccumulation for long periods exerting potentially harmful effects on the organism [1]. Specifically, their translocation to organs such as the heart, kidney, liver, or brain makes them a latent threat to public health; as it has already been recorded that they manage to cross biological barriers by various routes. In the case of the lungs and brain, the absorption of such particles can occur through the nasal mucosa causing systemic inflammation and oxidative stress on the central nervous system (CNS).

The most common entry routes of UFPs into the body are inhalation, dermal absorption, and ingestion, although the skin under proper conditions is an efficient barrier making the entry routes of greatest risk to the respiratory and gastric systems. The gastrointestinal tract is about 200 m² of surface area lined with mucosa, represents a not very efficient barrier, despite the existence of cellular junctions that limit the passage of potentially harmful substances; since it allows the passage of practically any material below 500 nm (0.5 µm) [7]. It has been described that the junctions between epithelial cells can undergo modifications that facilitate the passage of ultrafine materials smaller than 15 nm allowing their arrival to the bloodstream and their consequent arrival to different organs including the brain [3]. The inhalation route, the respiratory system is divided into upper and lower airways, which represent two distinct pathways for particles that may be inhaled. The upper airways include the nose, nasal cavity, pharynx, and larynx; while the lower airways include the trachea and lungs including bronchi, bronchioles, and alveoli. The lung surface area is approximately 150 m² and particles have the potential to cross the air-blood barrier into the bloodstream and disperse in the direction of other organs including the brain where again the blood-brain barrier (BBB) limits the free distribution of any potentially harmful substances [8]. As for the particles that remain attached to the nasal mucosa, there are two entry routes: particles can cross the respiratory epithelium reaching the blood vessels which gives them systemic access, or they can be absorbed by the epithelial cells and reach the brain in different forms of transport. It has been proposed that they may enter via the olfactory or trigeminal pathway through axonal transport, by paracellular transport by entering through the spaces between neuronal axons, or even that they could enter by transcellular or paracellular transport to the sustentacular olfactory epithelium [9].

An important problem is that these types of nanomaterials are not only found suspended in the air as part of the components of environmental pollution, but some of them are also similar in composition and characteristics to others used in biomedical areas for therapeutic purposes. Specifically, nanoparticles are materials that due to their characteristics have stood out in the development of applications in biomedicine due to their benefits and physical properties that are different from those of larger particles. Their dimensions are attractive for biomedical applications as they may show better clinical performance compared to conventional ways to deliver drugs or imaging tissues and organs [10]. However, these same characteristics give them the particularity of crossing barriers intended to protect us from potential toxic effects, such as the epithelial barrier of the skin, the gastrointestinal tract (GIT), the air-blood barrier of the lungs and even organ-specific barriers such as the blood-testicular barrier, the placental barrier, and the BBB [11]. The latter has the function of preserving the integrity of the CNS and cautiously selecting what can enter the CNS. When these barriers are breached by nanoparticles, they can be transported through the organism by systemic circulation, distributing and accumulating in different organs and tissues

without being correctly metabolized or eliminated by the organism, causing damage that does not necessarily depend on the dose but on their chronic permanence where they have been deposited [12].

Some of the most widely used nanoparticles for biomedical applications are magnetite (Fe_3O_4), carbon-based nanostructures (carbon nanotubes, graphene, carbon nanoparticles), silica (SiO_2), as well as some metallic nanoparticles, in particular those containing silver (Ag) and gold (Au). Among the benefits derived from their use are improvements in imaging, creation of biosensors for disease detection (especially neurodegenerative diseases and cancer), targeted drug delivery, delivery of drugs to the brain and other relevant biomedical applications such as antiviral nanotherapeutics that could provide a rapid and efficient response to viruses causing pandemics such as the current one [13]. But to take advantage of the advances that nanomaterials offer and transfer them efficiently and safely to clinical application, it is necessary to identify those that present the best benefit-toxicity ratio, especially when it comes to materials with the ability to translocate and accumulate in the brain and potentially may cause neurodegenerative alterations.

Synthetic nanomaterials are manufactured for several technological applications in sectors such as agriculture, cosmetics, pharmaceuticals, engineering, the automotive industry, construction, among others. They can also be released to the environment after disposal and degradation of consumer items that contain them. Their potential environmental effects, caused by their diverse physical properties, are mostly unknown; one of the important problems of pollutant nanomaterials is their fate as they may end up being incorporated into bodies of water, food chains, and even the atmosphere where they can accumulate and be breathed [6, 14, 15].

There are also airborne nanoparticles coming from natural processes, such as volcanic ash, forest fires carbon particles, dust from natural wastes, fecal matter, decomposition of biological structures and several other sources. However, multiple human activities such as fires for land use, logging, mining, construction, overcrowding, etc., alter their natural cycles, incorporating large amounts of nanoparticles into the atmosphere in a short time, generating a considerable increase in their presence as part of environmental pollution [15].

As previously described, nanoparticles can cross the BBB and once inside can generate diverse toxic effects, such as inducing oxidative activity through the production of reactive oxygen species (ROS) or inducing the abnormal folding of proteins setting the conditions for the development of CNS degenerative diseases. The composition, size, and shape of nanoparticles, as well as the nature of their biocorona (protein corona), influence their toxicity, therefore it is required to identify which nanomaterials have a lower risk of inducing harmful reactions and how they are biodistributed once they enter the organism [12, 16].

In this regard, the specific factors that determine which nanoparticles enter the brain are mostly unknown. Size, shape, stiffness, and composition of nanoparticles are considered important and, under physiological conditions, the nature of the adsorbed biomolecule corona (proteins, lipids, etc.) determines the biological responses.

2. Nanoparticles, types, and characteristics

Nanoparticles (or UFPs) are those that measure between 1 to 100 nm in all dimensions of their surface; they can be composed of different components, either natural or

synthesized for some purpose. As previously mentioned, the occurrence of nanoparticles in the environment can be the result of anthropogenic activities and in these cases, they are usually part of the pollutants to which the population of urban areas is exposed. In contrast, anthropogenic (synthetic) nanoparticles have very diverse applications in areas such as medicine, engineering, agriculture, cosmetics, clothing and more [3, 17, 18]. The synthesis of nanoparticles can be done physically, chemically, or biologically, the latter is gaining popularity because there is no risk of generating toxic materials during the process unlike chemical synthesis and various microorganisms can be used for their syntheses such as fungi or bacteria [17]. Different types of synthesized nanoparticles can be described according to their physicochemical characteristics and chemical composition: carbon-based nanoparticles, ceramic, metallic, semiconductor, polymeric, or lipid-based materials, among others.

3. Nanoparticles as contaminants

One characteristic that differentiates nanoparticles from other materials is that their small size and large surface to volume ratio give them unique properties in terms of their interaction with organisms, affecting how these materials translocate through the biological barriers (membranes). This property can be very useful for biomedical applications; however, it has been described [6, 19] that significant toxic effects can be observed due to the passage of nanoparticles to compartments unreachable by materials of larger dimensions. This is a consequence not only of their small size but also of their permeability, their large surface area and their tendency to form agglomerates [3]. This is a major problem given that inhaled nanoparticles not only enter the lungs but also there is evidence of their passage through the nasal mucosa, which may imply access to the central nervous system (CNS) with all the consequences that this entails.

One of the materials that have been studied concerning CNS toxicity are magnetite nanoparticles (Fe_3O_4); it has been observed that they can affect microvascular endothelial cells that form the BBB and can induce cell damage even at the mitochondrial level [19]. Damage at the mitochondrial level in the CNS could affect the cell cycle and regulation of apoptosis, making the organism more prone to cancer, neuronal dysfunction, and demyelination.

Another issue regarding nanoparticles, either from environmental pollution or derived from anthropogenic activities, is that they can damage the nasal epithelium after constant exposure, generating high rates of dysplasia in nasal tissues and inducing the accumulation of P53 protein, which in turn leaves the brain unprotected since the membranes lose their integrity and increase their permeability facilitating their passage through the nasal epithelium [20–22]. Exposure to environmental pollutants can also lead to BBB alteration, degeneration of cortical neurons, apoptotic glial cells, apolipoprotein E deposition in muscle cells and in pericytes, non-neuritic plaques and neurotropic tangles [23]. Therefore, exposure to this type of nanoparticles may initiate damage that leads to the development of neurodegenerative pathologies even at an early age [20, 21].

In this sense, it is relevant that oxidative stress damage, DNA damage and the presence of amyloid plaques in the olfactory bulb, frontal cortex and hippocampus, as a consequence of particulate matter and ozone present in environmental pollution, have already been clinically evidenced; microglia activation, inflammation, oxidative stress and changes in the BBB are the main mechanisms that lead to CNS pathology

and therefore neurodegenerative disorders such as ischemia, Parkinson's disease and Alzheimer's disease could be associated with such exposure [23].

It is evident that nanoparticles, both natural and derived from anthropogenic actions, can have a negative impact on the human organism. This impact can represent a high cost for the global health system, and that research aimed at understanding the mechanisms must be facilitated, in order to design strategies that allow us to avoid damage, but also to take advantage of the properties to design nanomaterials that can provide benefits, reduce costs for the health sector and increase the quality of life of people, thus becoming applied for the benefit of society.

4. Nanoparticles with biomedical applications

Nanomedicine deals with the biomedical applications of nanomaterials; in this area, nanomaterials are used creatively for the design and preparation of new forms of drug delivery, improving their precision and effectiveness due to their unique properties, or for the creation of novel formulations, sensors for early diagnosis and medical imaging contrast agents for the early detection of diseases in a minimally invasive way, with higher resolution, sensitivity, and specificity, among other uses.

Quantum dots are an example: they have unique physical characteristics, especially their size which can be controlled to tune their light emission properties; they have been used to solve problems related to the modulation of the intensity and duration of fluorescence when staining cell or tissue samples for microscopic analysis, in addition to the fact that by coating their surface they can target specific sites and allow, for example, the detection of tumors [24].

Magnetic nanoparticles are being successfully used in medical magnetic resonance imaging (MRI) as contrast agents to visualize biological structures that cannot be identified with conventional MRI, which has a major impact on diagnosis, allowing patients to receive more appropriate care in the best possible time [25].

Perhaps the greatest potential application of nanomaterials is in the design of novel drug delivery systems, in which polymeric and lipid-based nanoparticles seem to lead, although examples can be found with other nanomaterials usually PEGylated (coated with polyethylene glycol) to improve their biocompatibility; in the case of lipid nanoparticles, as they form a vesicle with a double membrane, the drug can be loaded inside the vesicle or between the lipid bilayer and be directed into a specific target by functionalizing their surface [18]. Extensive research has been done with functionalized nanoparticles to act as virucidal agents for Zika, rabies, influenza, HIV, hepatitis, papilloma, among other viral diseases. [13]. This is a promising area considering the current pandemic and the need for vaccines in short times not only to respond to SARS-Cov-2 but for emerging viral diseases to come, so having candidate nanomaterials ready to generate such vaccines is pressing.

Another area of opportunity is the development of nanoparticle-based biosensors. These are being used not only for the detection of diseases but also for the stratification of such pathologies, and even for the early detection of asymptomatic patients who would benefit from early diagnosis, for example for cancer. Among the advantages of these biosensors are their very high specificity, sometimes better than conventional biopsies, and that they are non-invasive so that sensors can be generated for routine use for screening and monitoring of drug resistance or function [26].

5. Nanoparticles and the brain

One of the most promising potential applications is the design and synthesis of nanoparticles with characteristics that make them capable of crossing the BBB, thus allowing drugs (or even genes) to be taken directly to the CNS, exerting their therapeutic action at that site; this has the potential to offer alternatives to improve the quality of life of millions of people suffering from diseases that do not respond to treatment with conventional drugs and who usually receive doses that generate undesirable side effects derived from the toxicity of the drug itself [18, 27]. Another interesting application related to the brain is that given the ease of exosomes to diffuse freely through the BBB and circulate in different body fluids, and the fact that exosomes reflect the environment of the cell from which they originate, biosensors can take advantage to detect exosomes derived from cells with various CNS pathologies either cancer or neurodegenerative diseases [26, 28].

Unfortunately, the passage of nanoparticles into the brain implies that they can interact with neural cells, with the positive and negative consequences that this can have; although pollutant nanoparticles were previously described as possible causative agents of neurodegenerative diseases, it is a reality that any nanomaterial, depending on its characteristics and their biocorona could have the capacity to form ROS or promote inappropriate protein folding. Even more serious is that the harmful effects may be greater in the long term and related to the dose increase, since many of these materials will not be able to be eliminated from the site where they are deposited, exerting their negative effects indefinitely [8].

Of equal concern is that nanomaterials can induce changes in miRNAs, histones, and DNA methylation, producing changes in gene expression patterns that will not necessarily be beneficial and studies that contemplate identifying toxicity of nanomaterials for biomedical use should include an analysis of changes in gene expression to identify possible mechanisms of damage and/or cause of various diseases among which will, of course, include neurodegenerative diseases [8, 29].

6. BBB and internalization pathways to the brain

Although nanoparticles can enter different compartments of the organism and generate local toxicity, their presence in the brain has been the subject of multiple studies to elucidate the mechanisms for brain internalization as toxicity in the CNS can lead to the onset of neurodegenerative diseases.

Previous studies have identified a gradient of damage and accumulation of metals in the following direction, olfactory mucosa > olfactory bulb > frontal cortex [23], and the possible routes through which environmental pollutants reach the brain have been described [21].

Recent works [30] have mentioned the existence of distinct routes of post-exposure entry of nanoparticles into the nasal cavity, the trigeminal route, systemic route, and olfactory route. These entry routes present nanoparticles with compartments that differ in their cellular composition and structure, representing barriers for drug transport and delivery. Nanoparticles can reach the brain either bypassing the BBB or by axonal transport along the olfactory nerve and once there they can become toxic especially due to their increased surface area which increases chemical reactivity and biological activity [12].

Of the different types of nanoparticles that are frequently used in brain-related biomedical applications, titanium oxide, iron oxide, silver, gold, silica, and carbon nanoparticles have been described to exhibit neurotoxic effects ranging from oxidative stress to alterations of brain development in embryonic stages. However, changes in their surface area, size or morphology may have impacts on both the degree of toxicity and their beneficial effects; more extensive studies on their mechanisms of action including modification of gene expression are required to identify more suitable materials, for biomedical use [12, 31–33].

Recapitulating the various biomedical applications and specifically those related to the CNS, it is important to identify which nanomaterials are currently under study and to verify their toxicity which may arise from the accumulation of these materials in the brain, otherwise, they could not be considered for the clinical setting and their potential advantages could not be translated into concrete benefits for society.

7. Conclusions

Exposure to nanoparticles of natural and anthropogenic origin is practically inevitable, either because they are present in environmental pollution or because they are used as components in many products for biomedical, cosmetic, structural, or other applications. Their small size and composition give them unique physicochemical features that influence the way how they interact with the organism; some of these nanoparticles can pass easily through membranes intended to protect our cells, others cannot be excreted and will bioaccumulate in tissues and organs with all the consequences that may derive not only from their localization, but also from their reactivity, aggregation, or even their potential to modify gene expression patterns. Therefore, their translocation to the brain continues to be studied both to identify entry routes of toxic agents as well as to exploit these mechanisms for drug delivery, internalization of contrast media, and in various other applications.

From these data, it is important to know whether different nanomaterials present in environmental pollution manage to translocate to the brain when they are aerosolized, how they are biodistributed, whether they have effects associated with neurodegeneration, etc. This would make it possible to issue recommendations on the strategies required to deal with environmental pollution, but also to regulate the fate of nanomaterials from different industries and to implement protocols for their manufacture, manipulation, and disposal. This may also have an impact on the identification of better drug delivery vehicles to the brain, which has been extremely complex due to the limitations on mobility that the BBB represents to the passage of drugs entering the organism through the different routes of administration.

For that matter, it is important to find an answer to whether there is a neurodegenerative effect derived from chronic exposure to aerosolized silica, magnetite, or carbon black nanoparticles in animal models, to deepen in our understanding of neurodegeneration processes and identify potential entry routes of nanoparticles into the brain. Approaches through *in vivo* models can lead the way for clinical studies with a broader result of possible nanoparticle damaging effects. However, knowing that inhaled nanopollutants can translocate to the brain and promote neurodegenerative effects has implications in the way we approach disease, but also in prevention and, in the public policies created to regulate exposition to nanomaterials independently of their source. Since prevention is the best health approach possible, it is where our

efforts should be concentrated. It can save lives, suffering, and money to a greater extent than solutions brought once a health problem has exploded. For that matter, studying the effects of nanopollutants present in the air that we are breathing every day at higher rates, is a fundamental step to gather the necessary information needed to create preventive strategies concerning pollution exposition to prevent health risks.

Acknowledgements

ACHA is thankful to CONACYT (México) for a PhD scholarship.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author details


Cristina Hermosillo-Abundis¹, Miguel A. Méndez-Rojas^{1*} and Oscar Arias-Carrión²

¹ Chemical-Biological Sciences Department, UDLAP, Puebla, México

² Movement and Sleep Disorders Unit, Dr. Manuel Gea González General Hospital, México City, México

*Address all correspondence to: miguela.mendez@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Landrigan PJ. Air pollution and health. *The Lancet*. 2017;2(1):e4-e5
- [2] Boyes WK, van Thriel C. Neurotoxicology of nanomaterials. *Chemical Research Toxicology*. 2020;24:1121-1144
- [3] Flood-Garibay JA, Angulo-Molina A, Méndez-Rojas MÁ. Particulate matter, urban air pollution and its effect on the Nervous 2 System. 48. 2022 [Manuscript in preparation]
- [4] Grandjean P, Landrigan PJ. Neurobehavioural effects of developmental toxicity. *The Lancet Neurology*. 2014;13(3):330-338
- [5] Kim H, Kim W-H, Kim Y-Y, Park H-Y. Air pollution and central nervous system disease: A review of the impact of fine particulate matter on neurological disorders. *Front Public Health*. 2020;8:575330
- [6] Thompson JE. Airborne particulate matter: Human exposure and health effects. *Journal of Occupational and Environmental Medicine*. 2018;60(5):392-423
- [7] Salvo-Romero E, Alonso-Cotoner C, Pardo-Camacho C, Casado-Bedmar M. Función barrera intestinal y su implicación en enfermedades digestivas. *Revista Espanola Enfermedades Digestivas*. 2015;107:11
- [8] Gehr P. Interaction of nanoparticles with biological systems. *Colloids and Surfaces B: Biointerfaces*. 2018;172:395-399
- [9] Newman L, Rodrigues AF, Jasim DA, Vacchi IA, Ménard-Moyon C, Bianco A, et al. Nose-to-brain translocation and cerebral biodegradation of thin graphene oxide nanosheets. *Cell Reports Physical Science*. 2020;1(9):100176
- [10] McNamara K, Tofail SAM. Nanoparticles in biomedical applications. *Advances in Physics: X*. 2017;2(1):54-88
- [11] Meng H, Leong W, Leong KW, Chen C, Zhao Y. Walking the line: The fate of nanomaterials at biological barriers. *Biomaterials*. 2018;174:41-53
- [12] Teleanu D, Chircov C, Grumezescu A, Volceanov A, Teleanu R. Impact of nanoparticles on brain health: An up to date overview. *JCM*. 2018;7(12):490
- [13] Mukherjee S, Mazumder P, Joshi M, Joshi C, Dalvi SV, Kumar M. Biomedical application, drug delivery and metabolic pathway of antiviral nanotherapeutics for combating viral pandemic: A review. *Environmental Research*. 2020;191:110119
- [14] Gómez-Hernández R, Panecatí-Bernal Y, Méndez-Rojas MÁ. High yield and simple one-step production of carbon black nanoparticles from waste tires. *Heliyon*. 2019;5(7):e02139
- [15] Lespes G, Faucher S, Slaveykova VI. Natural nanoparticles, anthropogenic nanoparticles, where is the frontier? *Front Environmental Science*. 2020;8:71
- [16] Shannahan J. The biocorona: A challenge for the biomedical application of nanoparticles. *Nanotechnology Reviews*. 2017;6(4):345-353
- [17] Hasan S. A review on nanoparticles: Their synthesis and types. *Research Journal of Recent Sciences*. 2015;4:1-3
- [18] Murthy SK. Nanoparticles in modern medicine: State of the art and future challenges. *International Journal of Nanomedicine*. 2007;2(2):129-141

- [19] Gárate-Vélez L, Escudero-Lourdes C, Salado-Leza D, González-Sánchez A, Alvarado-Morales I, Bahena D, et al. Anthropogenic iron oxide nanoparticles induce damage to brain microvascular endothelial cells forming the blood-brain barrier editor. JAD. 2020;76(4):1527-1539
- [20] Calderon-Garciduenas L, Rodriguez-Alcaraz A, Valencia-Salazar G, Mora-Tascareno A, Garcia R, Osnaya N, et al. Nasal biopsies of children exposed to air pollutants. Toxicological Pathology. 2001;29(5):558-564
- [21] Calderon-Garciduenas L, Maronpot RR, Torres-Jardon R, Henriquez-Roldan C, Schoonhoven R, Acuna-Ayala H, et al. DNA damage in nasal and brain tissues of canines exposed to air pollutants is associated with evidence of chronic brain inflammation and neurodegeneration. Toxicological Pathology. 2003;31(5):524-538
- [22] Calderon-Garciduenas L, Rodriguez-Alcaraz A, Janszen D, Morgan KT. Nasal epithelium as a sentinel for airborne environmental pollution. Toxicological Sciences. 1998;46:352-364
- [23] Calderón-Garcidueñas L, Azzarelli B, Acuna H, Garcia R, Gambling TM, Osnaya N, et al. Air pollution and brain damage. Toxicological Pathology. 2002;30(3):373-389
- [24] Gao X, Cui Y, Levenson RM, Chung LWK, Nie S. In vivo cancer targeting and imaging with semiconductor quantum dots. Nature Biotechnology. 2004;22(8):969-976
- [25] Harisinghani MG, Deserno WM. Noninvasive detection of clinically occult lymph-node metastases in prostate cancer. The New England Journal of Medicine. 2003;348:2491-2499
- [26] Xu L, Shoaie N, Jahanpeyma F, Zhao J, Azimzadeh M, Al-Jamal KT. Optical, electrochemical and electrical (nano)biosensors for detection of exosomes: A comprehensive overview. Biosensors and Bioelectronics. 2020;161:112222
- [27] Opris I, Lebedev MA, Pulgar VM, Vidu R, Enachescu M, Casanova MF. Editorial: Nanotechnologies in neuroscience and neuroengineering. Frontier in Neuroscience. 2020;12(14):33
- [28] Lorencova L, Bertok T, Bertokova A, Gajdosova V, Hroncekova S, Vikartovska A, et al. Exosomes as a source of cancer biomarkers: Advances in electrochemical biosensing of exosomes. ChemElectroChem. 2020;7(9):1956-1973
- [29] Sierra MI, Valdés A, Fernandez AF, Torrecillas R, Fraga MF. The effect of exposure to nanoparticles and nanomaterials on the mammalian epigenome. IJN. 2016;11:6297-6306
- [30] Selvaraj K, Gowthamarajan K, Karri VVSR. Nose to brain transport pathways an overview: Potential of nanostructured lipid carriers in nose to brain targeting. Artificial Cells, Nanomedicine, and Biotechnology. 2017;28:1-8
- [31] Ways TM, Ng KW, Lau WM, Khutoryanskiy VV. Silica nanoparticles in transmucosal drug delivery. Pharmaceutics. 2020;12:751
- [32] Malhotra N, Lee J-S, Liman RAD, Ruallo JMS, Villaflores OB, Ger T-R, et al. Potential toxicity of iron oxide magnetic nanoparticles: A review. Molecules. 2020;25(14):3159
- [33] Uo M, Akasaka T, Watari F, Sato Y, Tohji K. Toxicity evaluations of various carbon nanomaterials. Dental Materials Journal. 2011;30(3):245-263

A Systematic Review of the Pharmacological Potential of Disintegrins on Breast Cancer

Araliz López-Pintor and Irene Vergara-Bahena

Abstract

Breast cancer is the most common type of cancer and causes the highest mortality in women worldwide. The triple-negative breast cancer (TNBC) subtype is present in 15–20% of women with breast cancer, and it does not have estrogen receptors, progesterone receptor or HER2. It occurs in young, Hispanic, African-American women, characterized by worse prognosis, high recurrence and metastasis, and has limited treatment options. Integrins are heterodimeric transmembrane receptors involved in cancer by regulating angiogenesis, proliferation, migration, cell invasion and metastasis. Disintegrins are small proteins (4–14 kDa) localized in snake venom capable of interacting and interfering with the activity of integrins involved in cancer. There is an urgent need to provide therapeutic alternatives for breast cancer. In this review, we present the *in vitro* and *in vivo* effects of disintegrins that have proven their effect in breast cancer.

Keywords: breast cancer, disintegrins, integrins, triple-negative breast cancer, venom snakes

1. Introduction

Cancer is a condition characterized by uncontrolled cell proliferation, avoiding growth suppressive factors, supporting cell death, allowing replicative immortality, driving angiogenesis, causing invasion and metastasis [1]. Breast cancer (BC) is the neoplasm with the highest frequency and mortality in women [2]. Prognosis and treatment depend on the type and stage of BC [3].

According to the treatment response, there are three subtypes of BC; hormone receptor-positive, human epidermal growth factor 2 (HER 2) positive, and triple-negative. Triple-negative breast cancer (TNBC) is a heterogeneous subtype characterized by the absence of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER-2). This cancer subtype accounts for 15–20% of women diagnosed with breast cancer [4]. It usually occurs in young, Hispanic and African-American patients [5]. Patients with TNBC have BRCA1/2 mutations, have a worse prognosis than any other cancer subtype due to high recurrence rates (being higher in the first 1–4 years of follow-up) and limited therapeutic

options. Currently, the only systemic treatment for TNBC is chemotherapy [5]; however, this therapeutic option is not standardized [6]. Additionally, TNBC is usually more aggressive and more prone to metastasis [7–9]. Unfortunately, metastasis contributes to 90% of cancer-related deaths [10]. Owing to, it is necessary to look for alternative treatments that offer patients diagnosed with BC. This work presents disintegrins that have proven their effect on BC both *in vitro* and *in vivo*.

2. Integrins

Integrins are heterodimeric transmembrane proteins responsible for mediating cell-cell and cell-extracellular matrix (ECM) interactions. These proteins bind to ECM ligands such as fibronectin, collagen, laminin, nidogen, vitronectin and other adhesion molecules such as ICAM-1. In healthy tissues, integrins play a key role in hemostasis, cell survival, migration, and proliferation; while in pathological processes, these receptors are associated with angiogenesis, inflammatory activity, osteoporosis and metastasis [11–13].

Integrins are homo- or hetero-dimeric proteins of α and β subunits combination. So far, 18 α -subunits and 8 non-covalently bound β -subunits combined constitute 24 types of integrins [13, 14]. The majority of integrins recognize the Arg-Gly-Asp (RGD) sequence present in essential ECM proteins. It has been shown that the modulation of residues flanking the RGD sequence may be the mechanism that regulates the specificity of integrin-ligand interactions [15, 16]. Each integrin binds to one or more ligands [17].

Some examples are integrins like $\alpha 5\beta 1$, $\alpha 4\beta 1$ and $\alpha v\beta 3$ that have an affinity to fibronectin; $\alpha 1\beta 1$ and $\alpha 2\beta 1$ recognize collagen; $\alpha v\beta 3$ and $\alpha v\beta 5$ can bind to vitronectin and fibrinogen and, $\alpha 2\beta 1$, $\alpha 3\beta 1$ and $\alpha 6\beta 1$ exerts activity on laminin [18]. Besides, $\alpha IIb\beta 3$ expressed in platelets bind fibrinogen or von Willebrand factor that contribute to platelet aggregation [19]. Additionally, some integrins interact with cell receptors such as $\alpha 4\beta 1$ and $\alpha 9\beta 1$ that binds to VCAM-1 (Vascular cell adhesion molecule-1) and $\alpha 4\beta 1$ to mucosal vascular addressin cell adhesion molecule-1 (MadCAM-1) [18, 20].

Integrins transduce signals (from the outside to the inside of the cell and vice versa) mediated by the extracellular environment; ligand binding leads to conformational changes allowing the integrin to produce signals for cells to migrate, proliferate or undergo apoptosis [21, 22]. Integrins have three conformational states that depend on the regulation of the adaptor proteins and the traction force of the cytoskeleton at the time the integrins bind their ligands. The open extended form indicates a high affinity for the ligand, when ligand binds triggers activation of the signaling pathway [23], whereas the closed folded or closed extended form has a low affinity for the ligand [14].

In normal tissues, integrins have checkpoints over cell proliferation, however, integrin expression is altered in cancer. Such is the case of $\alpha v\beta 3$, integrin overexpressed in tumor cells which can promote cell migration and angiogenesis. In concert with $\alpha v\beta 5$, these integrins contribute to drug resistance and radiotherapy. Invasion and metastasis occur when a cancer cell escapes from the primary tumor, invades, survives in the bloodstream, and colonizes distant sites. Integrins influence these processes by modulating the colonization of metastatic sites and leading to anchorage-independent survival of tumor cells. Integrins have been associated with high rates of metastasis and poor prognosis in different types of cancers [23].

3. Animal venoms

Venoms are molecular cocktails constituted by mixtures of different families of proteins, peptides, amines, salts, polypeptides and enzymes, [24, 25] designed by natural selection to act on the vital systems of the prey or victim. Components of venoms can act on receptors, channels, or through their enzymatic activity, causing the dysfunction of the central nervous system (CNS) and peripheral nervous system (PNS), the cardiovascular system, blood coagulation and homeostasis of the prey. Venoms are a rich source of pharmacological bioactive compounds, but research on them is scarce, considering that there are 700 species of cone snails, 725 snakes, 1500 scorpions, and 37,000 spiders identified [26, 27].

3.1 Snake venoms

Since the discovery in 1980 of captopril, a drug derived from the venom of the *Bothrops jararaca* snake, it has been demonstrated that components of venoms are beneficial in the medical area to treat different diseases including cancer [25]. Among the snake venom components, phospholipases A₂ (PLA₂s) of *Bothrops jararacussu* showed antitumor activity on cell lines of breast cancer, acute T-cell leukemia and Ehrlich ascitic tumor [28]. A metalloproteinase of the *Bothrops jararaca* venom induced apoptosis and decreased the proliferation and metastasis in melanoma [29]. The L-amino acid oxidase (LAAO) isolated from *Daboia russelii russelii* induced apoptosis on breast cancer [30].

4. Disintegrins

Disintegrins are low molecular weight (4–14 kDa), nonenzymatic, disulfide-rich proteins derived from the proteolysis of snake venom metalloproteases (SVMPs). These molecules are phylogenetically related to human ADAMs (a disintegrin and a metalloprotease) [31]. SVMPs are divided according to their multidomain metalloprotease structure into P-I, P-II and P-III classes. Members of the P-I class are composed of a metalloprotease domain (MD), P-II class proteins have a MD and a disintegrin-like domain, and P-III have a cysteine-rich or disintegrin-like domain, and sometimes a lectin domain. Both P-II and P-III classes are subdivided according to their proteolytic processing and the ability to organize dimers [11, 31].

Depending on the type and processing, SVMPs yield different disintegrins structurally classified into monomers, homodimers and heterodimers. Monomeric disintegrins come from the processing by the P-IIa class of SVMPs, homodimers are derived from P-IIb class and heterodimers from the P-IIc class. P-III SVMPs are derived from disintegrin-like proteins (not containing the RGD motif), which consist of a cysteine-rich “disintegrin-like” domain, covalently linked and with a molecular weight of 30 kDa [31]. Another classification names disintegrins as short, medium, long, homo and heterodimers, based on the number of amino acid residues and disulfide bonds [12, 26, 32].

4.1 Disintegrins binding motifs

Disintegrins interact with specific integrins depending on the presence of a particular integrin-binding motif located in the hairpin loop, modulating the

selectivity toward cellular responses. There are three common motifs identified in disintegrins: RGD, KGD and KTS [33–35]. Other disintegrins contain different tripeptide sequences but conserve aspartic acid such as the MVD, KGD and WGD motifs [31].

RGD disintegrins correspond to the largest and most investigated family, they can inhibit physiological functions of $\alpha 3\beta 1$, $\alpha 4\beta 1$, $\alpha 5\beta 1$, $\alpha 6\beta 1$, $\alpha 7\beta 1$, $\alpha 8\beta 1$, $\alpha v\beta 1$, $\alpha v\beta 3$, $\alpha IIb\beta 3$ and $\alpha v\beta 5$ integrins. There are some dimeric disintegrins with motif MLD which binding to $\alpha 3\beta 1$ integrins, $\alpha 4\beta 1$, $\alpha 5\beta 1$, $\alpha 6\beta 1$, $\alpha 7\beta 1$, $\alpha 9\beta 1$, $\alpha IIb\beta 3$ and $\alpha 4\beta 7$, whereas others with KTS motifs in their active site have shown selectivity toward the $\alpha 1\beta 1$ collagen receptor [31].

4.2 Disintegrins and cancer

The main activity of disintegrins is to interfere with integrin-ligand interactions [35], i.e., they block binding integrins with ECM proteins such as collagen, fibronectin, laminin, and fibrinogen, preventing intracellular signals, cellular adhesion, migration, and metastasis [36]. The biological activity of disintegrins has prompted the research with these peptides as inhibitors of cancer progression [31]. For example, eristostatin binds and blocks $\alpha 4\beta 1$, an integrin associated with metastasis in human melanoma cells [37], jerdoni inhibits tumor growth and improves survival in a mouse model of melanoma [12], and rhodocetin, which prevents fibrosarcoma cell invasion by blocking $\alpha 4\beta 1$ to collagen [38].

Isolation and purification involve arduous processes to extract components from snake venom having small yielding. The production of recombinant disintegrins is a strategy to obtain large quantities of these proteins, suitable for *in vitro* and *in vivo* testing. Different disintegrins have been produced thanks to recombinant technology como albolatin [39], obtustatin [40], rubistatin [22] and salmosin [41]. Hosts such as *Escherichia coli* and *Pichia pastoris* are widely used for recombinant protein production with high yielding and easy genetic manipulation [11].

5. Methodology

5.1 Data search strategy

The database used for this review was PubMed. The Mesh search terms were “Neoplasms and Disintegrins”. The articles excluded were unassociated with breast cancer. The systematic review was performed under the guidelines of Preferred Reporting Items for Systematic reviews and Meta-Analyses PRISMA [42].

5.2 Inclusion and exclusion criteria

Inclusion criteria were based on the following: 1. Locating studies of disintegrins with activity in breast cancer; 2. that used the MCF-7, MDA-MB-231 cell line and human cell primary culture; 3. that were written in English. The exclusion criteria were: 1. that the disintegrins were not ADAM; 2. that they evaluated only the MDA-MB-435 or murine breast cancer cell line (66.3p); 3. that they did not evaluate the anticancer activity; 4. that the publications were not written in English due to not having suitable tools to perform a faithful translation.

5.3 Data extraction and analysis

The data were analyzed by two reviewers (ALP and IVB). The descriptive variables taken were type of disintegrin, cell line, the effect *in vitro* and effects *in vivo*.

5.4 Selection of studies

The papers identified with the term Mesh: “Neoplasms Disintegrins” yielded 152 publications. We decided to add one article that included human breast cancer primary cultures. The final analysis was 9 articles after applying the inclusion and exclusion criteria described above.

6. Results

Disintegrin	Cell line	Effects <i>in vitro</i>	Effects <i>in vivo</i>	References
Contortrostatin	Cell primary culture (ductal breast carcinoma)	Cytoskeleton disruption and apoptosis.	N.D.	[43]
DisBa-01	MDA-MB-231	Adhesion and VEGF signaling impairment.	N.D.	[44]
Lansbermin-I	MCF-7 MDA-MB-231	Adhesion, migration, and cell invasion.	N.D.	[45]
r-viridistatin 2	MDA-MB-231	Invasion cell	N.D.	[46]
Rhodostomin Trigramin Triflavin	MCF-7 MDA-MB-231	Cell adhesion, migration, and invasion.	Osteolytic and tumor growth in BALB/c-nu/nu mice.	[47]
Totonacin	MCF-7 MDA-MB-231	Migration and adhesion cell.	N.D.	[48]
Tzabcanin	MCF-7	Cell adhesion	N.D.	[49]
Vaa-Dis	MDA-MB-231	Viability and textto migration cell.	N.D.	[50]
Vicrostatin	MDA-MB-231	Cell invasion. Cytoskeleton disruption. Angiogenic, apoptosis, migration, proliferation.	Tumor growth and survival in nude mice.	[51, 52]

Abbreviations. N.D.: Not determined.

Table 1.
Types of snake venom disintegrins that evaluated their effects in breast cancer.

7. Discussion

BC ranks first in terms of frequency and mortality in the world [2]. Metastasis is responsible for 90% of deaths [53] and some integrins such as $\alpha v \beta 3$, $\alpha 5 \beta 1$ and $\alpha v \beta 6$ are involved in this process [31, 47, 54]. It has been observed that disintegrins may be an approach for future treatments by inhibiting the activity of integrins in cancer [11, 18, 55]. **Table 1** lists the disintegrins with effects on BC (cell line MCF-7) and TNBC (cell line MDA-MB-231). Most of them cause alterations in cell migration and adhesion.

We also find recombinant disintegrins such as DisBa-01 [44], r-viridistatin [46], and vicrostatin with promising pharmacological activity on cancer progression [51, 52]. Vicrostatin, is a chimeric recombinant disintegrin, produced as a genetic fusion between the C-terminal tail of echistatin and contortrostatin, it has shown the greatest anti-cancer potential [51] and several studies have evaluated its effect. Vicrostatin encapsulated in liposomes, was tested *in vivo* showing >80% of decrease in tumor growth and increased survival, similar results to those obtained in the group treated with the monoclonal antibody bevacizumab, used as the gold standard in antiangiogenic therapy [52]. Vicrostatin has also shown promising results in ovarian cancer [56], prostate cancer [57] as well as a diagnostic marker in glioma [58].

Contortrostatin is noteworthy for being the first work using human cell primary culture extracted from BC tissue. These primary cells maintain the original cellular phenotype, which is desirable for cell binding and motility experiments, effects present in metastasis [43].

In this review we only considered papers about human BC, we decided to discard the effect of crotoxin 2 because its effect was evaluated in murine breast cancer lines (66.3p). The effects of crotoxin 2 were the same as those disintegrins which cause inhibition of cell migration. In contrast to other disintegrins, to achieve a decrease in lung colonization crotoxin 2 required 1000 µg/kg, high doses to avoid the metastatic effect [59]. It would now need to be tested in human cancer lines to demonstrate if it has the same effects.

Other papers not considered were those naming cell lines as belonging to breast cancer [60–64], but recently, several studies have confirmed that they are melanoma lines [65–67]. We also do not include the study of the disintegrin DisBa-01 evaluated on the MDA-MB-231 (TNBC cell line) because it had no effect on the binding nor cell proliferation [68].

It has been demonstrated that disintegrins have high selectivity in the binding and inhibition of integrins, an activity highlighted for their therapeutic potential [31]; thus, this review shows the panorama and potential of disintegrins in BC, but there are still more studies and problems to be solved in order to reach their clinical application.

8. Conclusion

BC continues to be cancer with the highest number of cases and deaths worldwide. While the TNBC subtype is characterized by a worse prognosis, due to high rates of recurrences and metastasis, there is no standardized treatment so far, reasons that call the attention of researchers to find new therapeutic proposals. Thus, noting that several integrins are overexpressed during the progression of BC; disintegrins have pharmacological potential as a new alternative in cancer.

Acknowledgements

ALP is thankful to CONACYT (Mexico) for a Ph.D. scholarship.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Notes/thanks/other declarations

Thanks to the UDLAP Research and Postgraduate Department for supporting this publication.

IntechOpen


IntechOpen

Author details

Araliz López-Pintor and Irene Vergara-Bahena*
Chemical-Biological Sciences Department, UDLAP, Puebla, México

*Address all correspondence to: irene.vergara@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011;**144**(5):646-674
- [2] Globocan. Estimated number of new cases in 2020, worldwide, both sexes, all ages [Internet]. 2020. Available in: https://gco.iarc.fr/today/online-analysis-pie?v=2020&mode=cancer&mode_population=continents&population=900&populations=900&key=total&sex=0&cancer=39&type=0&statistic=1&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&nb_items=7&group_cancer=1&include_nmsc=0&include_nmsc_other=1&half_pie=0&donut=0
- [3] Waks AG, Winer EP. Breast cancer treatment. *Journal of the American Medical Association*. 2019;**321**(3):316-316
- [4] Gerratana L, Basile D, Buono G, De Placido S, Giuliano M, Minichillo S, et al. Androgen receptor in triple negative breast cancer: A potential target for the targetless subtype. *Cancer Treatment Reviews*. 2018 Jul;**68**:102-110
- [5] Chen L, Li CI, Tang M-TC, Porter P, Hill DA, Wiggins CL, et al. Reproductive factors and risk of luminal, HER2-overexpressing, and triple-negative breast cancer among multiethnic women. *Cancer Epidemiology, Biomarkers & Prevention*. 2016;**25**(9):1297-1304
- [6] Mahtani R, Kittaneh M, Kalinsky K, Mamounas E, Badve S, Vogel C, et al. Advances in therapeutic approaches for triple-negative breast cancer. *Clinical Breast Cancer*. 2021;**21**(5):383-390
- [7] Shen M, Jiang Y-Z, Wei Y, Ell B, Sheng X, Esposito M, et al. Tinagl1 suppresses triple-negative breast cancer progression and metastasis by simultaneously inhibiting integrin/FAK and EGFR signaling. *Cancer Cell*. 2019;**35**(1):64-80.e7
- [8] Xiao H, Yang W. MiR-205 dysregulations in breast cancer: The complexity and opportunities. *Non-Coding RNA*. 2019;**5**(4):53
- [9] Merino Bonilla JA, Torres Tabanera M, Ros Mendoza LH. Breast cancer in the 21st century: From early detection to new therapies. *Radiology English Edition*. 2017;**59**(5):368-379
- [10] Arrechea Irigoyen MA, Vicente García F, Córdoba Iturriagoitia A, Ibáñez Beroiz B, Santamaría Martínez M, Guillén GF. Subtipos moleculares del cáncer de mama: Implicaciones pronósticas y características clínicas e inmunohistoquímicas. *Anales del Sistema Sanitario de Navarra*. 2011;**34**(2):219-233
- [11] David V, Succar BB, de Moraes JA, Saldanha-Gama RFG, Barja-Fidalgo C, Zingali RB. Recombinant and chimeric disintegrins in preclinical research. *Toxins*. 2018;**10**(8):321
- [12] Zhou X-D, Jin Y, Chen R-Q, Lu Q-M, Wu J-B, Wang W-Y, et al. Purification, cloning and biological characterization of a novel disintegrin from *Trimeresurus jerdonii* venom. *Toxicon*. 2004;**43**(1):69-75
- [13] Della-Casa MS, Junqueira-de-Azevedo I, Butera D, Clissa PB, Lopes DS, Serrano SMT, et al. Insularin, a disintegrin from *Bothrops insularis* venom: Inhibition of platelet aggregation and endothelial cell adhesion by the native and recombinant GST-insularin proteins. *Toxicon*. 2011;**57**(1):125-133

- [14] Wang J, Dong X, Zhao B, Li J, Lu C, Springer TA. Atypical interactions of integrin $\alpha_v \beta_8$ with pro-TGF- β_1 . *Proceedings of the National Academy of Sciences*. 2017;**114**(21):E4168-E4174
- [15] Rahman S, Aitken A, Flynn G, Formstone C, Savidge GF. Modulation of RGD sequence motifs regulates disintegrin recognition of $\alpha_{IIb}\beta_3$ and $\alpha_5\beta_1$ integrin complexes. *The Biochemical Journal*. 1998;**335**(2):247-257
- [16] Huang TF, Holt JC, Lukasiewicz H, Trigramin NS. A low molecular weight peptide inhibiting fibrinogen interaction with platelet receptors expressed on glycoprotein IIb-IIIa complex. *The Journal of Biological Chemistry*. 1987;**262**(33):16157-16163
- [17] Li Z-H, Zhou Y, Ding Y-X, Guo Q-L, Zhao L. Roles of integrin in tumor development and the target inhibitors. *Chinese Journal of Natural Medicines*. 2019;**17**(4):241-251
- [18] Saviola AJ, Burns PD, Mukherjee AK, Mackessy SP. The disintegrin tzabcanin inhibits adhesion and migration in melanoma and lung cancer cells. *International Journal of Biological Macromolecules*. 2016;**88**:457-464
- [19] Fujii Y, Okuda D, Fujimoto Z, Horii K, Morita T, Mizuno H. Crystal structure of trimestatin, a disintegrin containing a cell adhesion recognition Motif RGD. *Journal of Molecular Biology*. 2003;**332**(5):1115-1122
- [20] Juarez P, Comas I, Gonzalez-Candelas F, Calvete JJ. Evolution of snake venom disintegrins by positive darwinian selection. *Molecular Biology and Evolution*. 2008;**25**(11):2391-2407
- [21] Eble JA, Bruckner P, Mayer U. Vipera lebetina venom contains two disintegrins inhibiting laminin-binding β_1 integrins. *The Journal of Biological Chemistry*. 2003;**278**(29):26488-26496
- [22] Carey CM, Bueno R, Gutierrez DA, Petro C, Lucena SE, Sanchez EE, et al. Recombinant rubistatin (r-Rub), an MVD disintegrin, inhibits cell migration and proliferation, and is a strong apoptotic inducer of the human melanoma cell line SK-Mel-28. *Toxicon*. 2012;**59**(2):241-248
- [23] Yousefi H, Vatanmakanian M, Mahdiannasser M, Mashouri L, Alahari NV, Monjezi MR, et al. Understanding the role of integrins in breast cancer invasion, metastasis, angiogenesis, and drug resistance. *Oncogene*. 2021;**40**(6):1043-1063
- [24] Tasoulis T, Isbister G. A review and database of snake venom proteomes. *Toxins*. 2017;**9**(9):290
- [25] Minutti-Zanella C, Gil-Leyva EJ, Vergara I. Immunomodulatory properties of molecules from animal venoms. *Toxicon*. 2021;**191**:54-68
- [26] Kini RM, Clemetson KJ, Markland FS, McLane MA, Morita T. *Toxins and Hemostasis*. Netherlands: Dordrecht; 2011
- [27] Calvete JJ. Proteomic tools against the neglected pathology of snake bite envenoming. *Expert Review of Proteomics*. 2011;**8**(6):739-758
- [28] Roberto PG, Kashima S, Marcussi S, Pereira JO, Astolfi-Filho S, Nomizo A, et al. Cloning and identification of a complete cDNA coding for a bactericidal and antitumoral acidic phospholipase A₂ from Bothrops jararacussu Venom. *The Protein Journal*. 2004;**23**(4):273-285
- [29] Maria DA, da Silva MGL, Correia MC, Ruiz IRG. Antiproliferative effect of the jararhagin toxin on

B16F10 murine melanoma. BMC Complementary and Alternative Medicine. 2014;**14**(1):446

[30] Mukherjee AK, Saviola AJ, Burns PD, Mackessy SP. Apoptosis induction in human breast cancer (MCF-7) cells by a novel venom l-amino acid oxidase (Rusvinoxidase) is independent of its enzymatic activity and is accompanied by caspase-7 activation and reactive oxygen species production. Apoptosis. 2015;**20**(10):1358-1372

[31] Arruda Macedo J, Fox J, Souza CM. Disintegrins from snake venoms and their applications in cancer research and therapy. Current Protein & Peptide Science. 2015;**16**(6):532-548

[32] Sun M-Z, Cui Y, Guo C, Zhao B, Liu S. rAdinbitor, a disintegrin from Agkistrodon halys brevicaudus stejneger, inhibits tumorigenicity of hepatocarcinoma via enhanced anti-angiogenesis and immunocompetence. Biochimie. 2015;**116**:34-42

[33] Lima-dos-Santos I, Della-Casa MS, Portes-Junior JA, Calabria PAL, Magalhães GS, Moura-da-Silva AM. Characterization of Neuwiedin, a new disintegrin from Bothrops neuwiedi venom gland with distinct cysteine pattern. Toxicon. 2015;**104**:57-64

[34] Da Silva M, Lucena S, Aguilar I, Rodríguez-Acosta A, Salazar AM, Sánchez EE, et al. Anti-platelet effect of cumastatin 1, a disintegrin isolated from venom of South American Crotalus rattlesnake. Thrombosis Research. 2009;**123**(5):731-739

[35] Calvete JJ, Moreno-Murciano MP, Theakston RDG, Kisiel DG, Marcinkiewicz C. Snake venom disintegrins: Novel dimeric disintegrins and structural diversification by disulphide bond engineering. The Biochemist. 2003;**372**(3):725-734

[36] Oliva IB, Coelho RM, Barcellos GG, Saldanha-Gama R, Wermelinger LS, Marcinkiewicz C, et al. Effect of RGD-disintegrins on melanoma cell growth and metastasis: Involvement of the actin cytoskeleton, FAK and c-Fos. Toxicon. 2007;**50**(8):1053-1063

[37] Danen EHJ, Marcinkiewicz C, Cornelissen IM, van Kraats AA, Pachter JA, Ruiter DJ, et al. The disintegrin eristostatin interferes with integrin $\alpha 4 \beta 1$ function and with experimental metastasis of human melanoma cells. Experimental Cell Research. 1998;**238**(1):188-196

[38] Eble JA, Niland S, Dennes A, Schmidt-Hederich A, Bruckner P, Brunner G. Rhodocetin antagonizes stromal tumor invasion in vitro and other $\alpha 2 \beta 1$ integrin-mediated cell functions. Matrix Biology. 2002;**21**(7):547-558

[39] Singhamatr P, Rojnuckarin P. Molecular cloning of albolatin, a novel snake venom metalloprotease from green pit viper (Trimeresurus albolabris), and expression of its disintegrin domain. Toxicon. 2007;**50**(8):1192-1200

[40] Brown MC, Eble JA, Calvete JJ, Marcinkiewicz C. Structural requirements of KTS-disintegrins for inhibition of $\alpha 1 \beta 1$ integrin. The Biochemical Journal. 2009;**417**(1):95-101

[41] Hong S-Y, Lee H, You W-K, Chung K-H, Kim D-S, Song K. The snake venom disintegrin salmosin induces apoptosis by disassembly of focal adhesions in bovine capillary endothelial cells. Biochemical and Biophysical Research Communications. 2003;**302**(3):502-508

[42] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies

that evaluate health care interventions: Explanation and elaboration. PLoS Medicine. 2009;**6**(7):e1000100

[43] Marinou I, Havaki S, Goutas N, Vlachodimitropoulos D, Baltatzis G, Konstantakou EG, et al. Action of the disintegrin contortrostatin on breast cancer cell primary cultures. *Advanced Breast Cancer Research*. 2013;**02**(04):161-169

[44] Montenegro CF, Salla-Pontes CL, Ribeiro JU, Machado AZ, Ramos RF, Figueiredo CC, et al. Blocking $\alpha v \beta 3$ integrin by a recombinant RGD disintegrin impairs VEGF signaling in endothelial cells. *Biochimie*. 2012;**94**(8):1812-1820

[45] Montealegre-Sánchez L, Gimenes SNC, Lopes DS, Teixeira SC, Solano-Redondo L, de Melo RV, et al. Antitumoral potential of lansbermin-I, a novel disintegrin from *Porthidium lansbergii lansbergii* Venom on breast cancer cells. *Current Topics in Medicinal Chemistry*. 2019;**19**(22):2069-2078

[46] Lucena SE, Jia Y, Soto JG, Parral J, Cantu E, Brannon J, et al. Anti-invasive and anti-adhesive activities of a recombinant disintegrin, r-viridistatin 2, derived from the Prairie rattlesnake (*Crotalus viridis viridis*). *Toxicon*. 2012;**60**(1):31-39

[47] Yang R-S, Tang C-H, Chuang W-J, Huang T-H, Peng H-C, Huang T-F, et al. Inhibition of tumor formation by snake venom disintegrin. *Toxicon*. 2005;**45**(5):661-669

[48] Rivas Mercado E, Neri Castro E, Bénard Valle M, Rucavado-Romero A, Olvera Rodríguez A, Zamudio Zuñiga F, et al. Disintegrins extracted from totonacan rattlesnake (*Crotalus totonacus*) venom and their anti-adhesive and anti-migration effects

on MDA-MB-231 and HMEC-1 cells. *Toxicology In Vitro*. 2020;**65**:104809

[49] Saviola AJ, Modahl CM, Mackessy SP. Disintegrins of *Crotalus simus tzabcan* venom: Isolation, characterization and evaluation of the cytotoxic and anti-adhesion activities of tzabcanin, a new RGD disintegrin. *Biochimie*. 2015;**116**:92-102

[50] Latinović Z, Leonardi A, Petan T, Žlajpah M, Križaj I. Disintegrins from the Venom of *Vipera ammodytes ammodytes* efficiently inhibit migration of breast cancer cells. *Acta Chimica Slovenica*. 2017;**2017**:555-559

[51] Minea RO, Helchowski CM, Zidovetzki SJ, Costa FK, Swenson SD, Markland FS. Vicrostatin – An anti-invasive multi-integrin targeting chimeric disintegrin with tumor anti-angiogenic and pro-apoptotic activities. *PLoS ONE*. 2010;**5**:e10929

[52] Minea R, Helchowski C, Rubino B, Brodmann K, Swenson S, Markland F. Development of a chimeric recombinant disintegrin as a cost-effective anti-cancer agent with promising translational potential. *Toxicon*. 2012;**59**(4):472-486

[53] Medeiros B, Allan AL. Molecular mechanisms of breast cancer metastasis to the lung: Clinical and experimental perspectives. *International Journal of Molecular Sciences*. 2019;**20**(9):2272

[54] Petitclerc E, Strömblad S, Cheres DA, Brooks PC. Integrin $\alpha V \beta 3$ promotes M21 melanoma growth in human skin by regulating tumor cell survival. *Cancer Research*. 1999;**59**(11):2724-2730

[55] Akhtar B, Muhammad F, Sharif A, Anwar MI. Mechanistic insights of snake venom disintegrins in cancer treatment. *European Journal of Pharmacology*. 2021;**899**:174022

- [56] Swenson SD, Markland FS, Minea R. A novel, non-cytotoxic, anti-invasive therapeutic agent for ovarian cancer. In: Proceedings of the 2016 International Conference on Biomedical and Biological Engineering, Shanghai, China. 2016
- [57] Markland FS, Swenson S, Minea R, Pinski J. Abstract PO030: A novel integrin-targeted therapeutic agent for prostate cancer. En: Systemic Tumor Microenvironment. American Association for Cancer Research. 2021;2021:PO030-PO030
- [58] Swenson S, Minea R, Tuan C, Thein T-Z, Chen T, Markland F. A novel venom-derived peptide for brachytherapy of glioblastoma: Preclinical studies in mice. *Molecules*. 2018;23(11):2918
- [59] Galán JA, Sánchez EE, Rodríguez-Acosta A, Soto JG, Bashir S, McLane MA, et al. Inhibition of lung tumor colonization and cell migration with the disintegrin crotatroxin 2 isolated from the venom of *Crotalus atrox*. *Toxicon*. 2008;51(7):1186-1196
- [60] Zhou Q, Sherwin RP, Parrish C, Richters V, Groshen SG, Tsao-Wei D, et al. Contortrostatin, a dimeric disintegrin from contortrix contortrix, inhibits breast cancer progression. *Breast Cancer Research and Treatment*. 2000;61(3):249-259
- [61] Dong D, Ko B, Baumeister P, Swenson S, Costa F, Markland F, et al. Vascular targeting and antiangiogenesis agents induce drug resistance effector GRP78 within the tumor microenvironment. *Cancer Research*. 2005;65(13):5785-5791
- [62] Swenson S, Costa F, Ernst W, Fujii G, Markland FS. Contortrostatin, a snake venom disintegrin with anti-angiogenic and anti-tumor activity. *Pathophysiology of Haemostasis and Thrombosis*. 2005;34(4-5):169-176
- [63] Swenson S, Costa F, Minea R, Sherwin RP, Ernst W, Fujii G, et al. Intravenous liposomal delivery of the snake venom disintegrin contortrostatin limits breast cancer progression. *Molecular Cancer Therapeutics*. 2004;3(4):499-511
- [64] Minea R, Swenson S, Costa F, Chen TC, Markland FS. Development of a novel recombinant disintegrin, contortrostatin, as an effective anti-tumor and anti-angiogenic agent. *Pathophysiology of Haemostasis and Thrombosis*. 2005;34(4-5):177-183
- [65] Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genetics*. 2000;24(3):227-235
- [66] Rae JM, Creighton CJ, Meck JM, Haddad BR, Johnson MD. MDA-MB-435 cells are derived from M14 Melanoma cells—a loss for breast cancer, but a boon for melanoma research. *Breast Cancer Research and Treatment*. 2007;104(1):13-19
- [67] Korch C, Hall EM, Dirks WG, Ewing M, Faries M, Varella-Garcia M, et al. Authentication of M14 melanoma cell line proves misidentification of MDA-MB-435 breast cancer cell line: M14 and MDA-MB-435 cell lines. *International Journal of Cancer*. 2018;142(3):561-572
- [68] Ramos OHP, Kauskot A, Cominetti MR, Bechyne I, Salla Pontes CL, Chareyre F, et al. A novel $\alpha v \beta 3$ -blocking disintegrin containing the RGD motive, DisBa-01, inhibits bFGF-induced angiogenesis and melanoma metastasis. *Clinical & Experimental Metastasis*. 2008;25(1):53-64

Chapter

Shedding Light on Four Selected Flavonoids with Anti-non-small Cell Lung Cancer Properties

Jorge L. Mejía-Méndez, Horacio Bach, Luis Ricardo Hernández and Eugenio Sánchez-Arreola

Abstract

Flavonoids are polyphenolic secondary metabolites classified as anthocyanins, chalcones, flavones, flavonols, flavanones, and isoflavonoids. The antibacterial, antifungal, anti-inflammatory, and anticancer properties of flavonoids are attractive for biomedical purposes. The anticancer potential of many flavonoids is well-known. However, the effect of some of them against human health concerns, such as lung cancer (LC), remains elusive. LC is the leading cause of cancer mortality worldwide. LC is classified into small-cell lung cancer (SCLC; 15%) and non-small cell lung cancer (NSCLC; 85%). Through PubMed, Google Scholar, and SciFinder databases, we addressed the distribution, chemical features, and therapeutic properties of flavonoids. Furthermore, we focused on the chemical characteristics, sources, biosynthesis, pharmacokinetics, anticancer, and anti-NSCLC activities of four selected flavonoids: luteolin, naringenin, kaempferol, and baicalein. The four selected flavonoids interfere with metastatic processes, promote apoptotic events, and diminish the proliferation of NSCLC cells, although further studies are required to extend our knowledge about the anti-NSCLC properties, efficacy, safety, and toxicity of the selected flavonoids. In the foreseeable future, the four selected flavonoids (alone or in combination) may be prospective candidates for drug design against NSCLC.

Keywords: lung cancer, non-small cell lung cancer, natural products, secondary metabolites, flavonoids

1. Introduction

Cancer is a generic term for a broad group of diseases caused by dysregulation of cell growth and exacerbated cellular proliferation. Contrary to the orderly and self-regulating manner of normal cells in the human organism, cancer cells divide and colonize tissues usually reserved for other cells. Cancer incidence is caused by many factors, such as unhealthy diet and poor physical activity, excessive alcohol intake, exposure to oncogenic viruses (e.g., human papillomavirus; HPV), immoderate sun exposure, tobacco consumption, among others [1]. Cancers with the highest mortality

globally are lung cancer (18.2%), colorectal cancer (9.5%), liver cancer (8.4%), stomach cancer (7.8%), and breast cancer (6.9%) [2].

Lung cancer (LC) is a complex type of cancer that remains the leading cause of death worldwide. By 2021, an estimated 235,760 persons will be diagnosed with LC, whereas 131,880 persons will probably die of LC [3]. The incidence of LC is subjected to exposure to physical and chemical components, family history, genetic alterations, and comorbidities. Considering that the estimated survival rate of LC is five years, several modifications have been made to diagnostic tools and treatment schemes to improve a patient's prognosis. The histological analysis and molecular scanning of LC lead to its categorization into small-cell lung cancer (SCLC, 15% of all lung cancers) and non-small cell lung cancer (NSCLC, 85% of all lung cancers) [4].

NSCLC is one of the two major categories of LC; its subtypes include adenocarcinoma, large-cell carcinoma, and squamous cell carcinoma. Over the past decades, there has been an increasing understanding of the biology, diagnosis, and treatment of NSCLC. Nowadays, it is known that the incidence of NSCLC is related to family history, genetic susceptibility, and exposure to physical and chemical carcinogens. NSCLC is diagnosed by testing genetic mutations, imaging techniques, and histologic features. This type of cancer is treated with various medical and pharmacological approaches, such as surgery, radiotherapy, targeted therapy, and immunotherapy [5].

The treatment of NSCLC presents diverse limitations, such as drug cytotoxicity, incompatible pharmacokinetic profiles, low accumulation rate in tumor tissues, side effects, and poor solubility. Therefore, the study of relatively new therapeutic molecules with anticancer potential such as flavonoids has unveiled promising results to treat NSCLC *in vitro* and *in vivo*.

Flavonoids are one of the largest classes of secondary metabolites produced in plants and foods. Flavonoids are classified into various subclasses, such as anthocyanins, flavones, flavonols, and flavanones. Despite their ecophysiological relevance, multiple studies have assessed many flavonoids' anticancer properties and synergistic action [6]. However, the anti-NSCLC properties of some flavonoids persist elusively.

We consulted PubMed, Google Scholar, and SciFinder to compile contemporary literature about flavonoids' distribution, chemical features, and therapeutic properties. In addition, this work is focused on the chemical characteristics, sources, biosynthesis, pharmacokinetics, anticancer, and anti-NSCLC activities of four selected flavonoids: luteolin, naringenin, kaempferol, and baicalein. A general scenario of LC and NSCLC is also addressed. In the literature, the four selected flavonoids have been reported to interfere with metastatic processes, arrest the cell cycle, restrain oncogenic signaling pathways activation, inhibit the invasion of NSCLC, improve the efficacy of current chemotherapy drugs, and diminish NSCLC cells' proliferation *in vitro* and *in vivo*. Further studies are required to extend our knowledge about the anti-NSCLC properties, efficacy, safety, and toxicity of luteolin, naringenin, kaempferol, and baicalein.

2. A general scenario of lung cancer and non-small cell lung cancer

LC or bronchogenic carcinoma refers to tumors originating in the pulmonary parenchyma or bronchi. LC is the leading cause of cancer death globally and is the second most prevalent type of cancer in both sexes. About 1.6 million people die of LC each year, and the overall 5-year survival rate is only 15% [7].

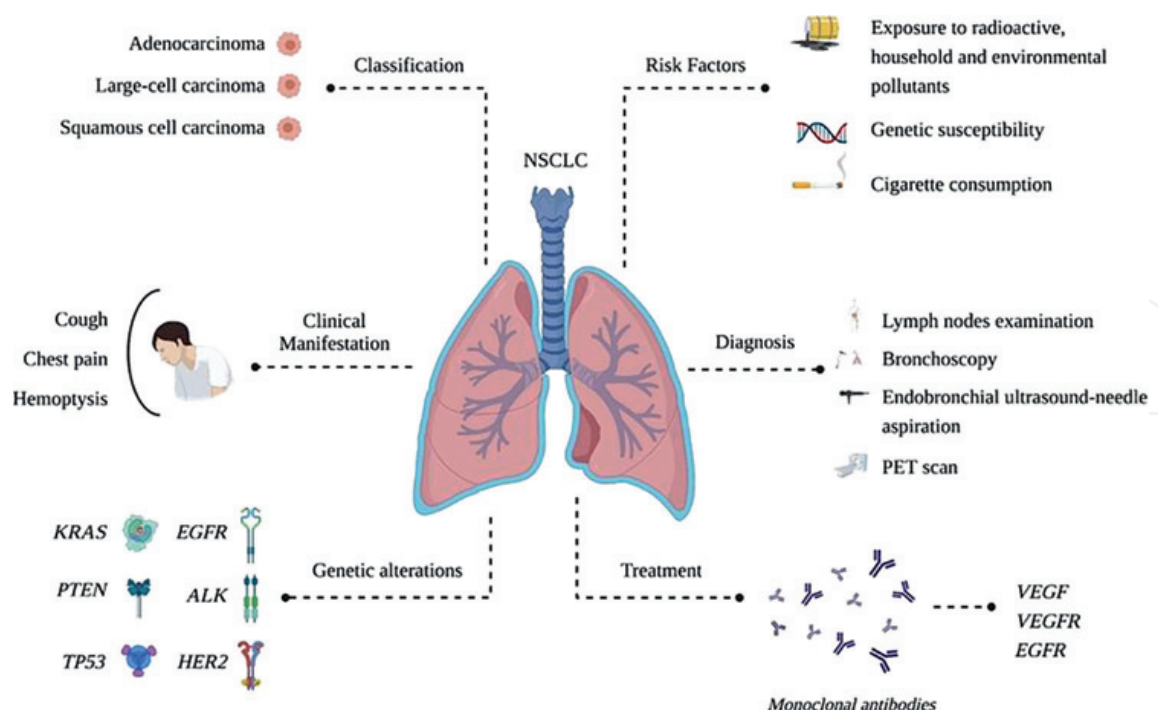


Figure 1.
 Non-small cell lung cancer (NSCLC) characteristics: classification, risk factors, genetic alterations, clinical manifestation, diagnosis, and treatment.

Even though the incidence of LC is associated with several factors, such as geographic location, sex, race, lifestyle, exposure to physical, chemical, or biological carcinogens, genetic polymorphisms, and comorbidities [8]; cigarette smoking is directly related to 82% of LC deaths and its major histological subtypes [9]. According to its histological characterization and by international convention, LC is divided into two large groups: small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC).

NSCLC is diagnosed in 80–85% of cases. Based on cell origin and histological features, the subtypes of NSCLC are adenocarcinoma, large-cell carcinoma, and squamous cell carcinoma [10]. The incidence of NSCLC is attributed to various elements, such as advanced age, current smoking, molecular features, genetic heterogeneity, and exposure to chemical carcinogens (e.g., radioactive radon) (**Figure 1**) [11]. However, bacterial and viral infections also represent a major concern associated with the incidence, morbidity, and mortality of NSCLC [12]. In addition, NSCLC risk factors increase mutation rates in the cancer cell genome and contribute to the progression of NSCLC.

The cancer cell genome is characterized by various somatic mutations and epigenetic changes that confer a growth advantage to the cancer cells and contribute to cancer development and progression [13]. Generally, 69% of patients with advanced NSCLC exhibit aberrations in the following proteins: anaplastic lymphoma kinase (ALK), EGFR, c-ros oncogene 1 (ROS1), human epidermal growth factor receptor (HER2), hepatocyte growth factor receptor (HGRF), Kirsten rat sarcoma virus (KRAS), V-raf murine sarcoma oncogene homolog B1 (BRAF), among others [14]. NSCLC patients also manifest alterations on TP53 (40–60%), phosphatase and tensin homolog (PTEN, 50%), liver kinase B1 (LKB1, 34%), and p16INK4A (17–58%) (**Figure 1**) [15]. Such mutations' epidemiological, geographical, and sociological distribution could vary worldwide [16].

NSCLC patients manifest cough, weight loss, dyspnea, chest pain, and hemoptysis [17]. The diagnosis of NSCLC consists of the evaluation of family history and physical (e.g., lymph nodes examination), histopathological (e.g., bronchoscopy), and cytological examinations (e.g., endobronchial ultrasound-needle aspiration) [18]. Given the necessity to establish appropriate therapy, management, and prognosis in recently diagnosed NSCLC patients, imaging techniques such as positron emission tomography (PET) scan, fluorodeoxyglucose (FDG)-PET scanning, thoracic computed tomography (CT), and magnetic resonance imaging (MRI), are broadly implemented (**Figure 1**) [19, 20]. The implementation of diagnostic tools leads to cancer staging.

Cancer staging determines how the spread and localization of a particular type of cancer. For example, the tumor-nodes-metastasis (TNM) classification is a system for categorizing a malignancy. According to the TNM staging system, the stages of NSCLC are local (IA, IB, and IIA), locally advanced (IIB, IIA, and IIIB), and advanced (IIIB, and IV) [21].

Depending on the stage of NSCLC, treatment includes the administration of chemotherapy (e.g., platinum-based doublet therapy) and monoclonal antibodies that target immune system T cells, ligands on the tumor cells, or aberrant proteins, such as vascular endothelial growth factor (VEGF) (e.g., bevacizumab, and docetaxel), vascular endothelial growth factor (VEGF) receptor-2 (e.g., docetaxel with or without ramucirumab), and epidermal growth factor receptor (EGFR) (e.g., necitumumab, gefitinib, and erlotinib) (**Figure 1**) [22].

Additionally, NSCLC treatment includes radiation technologies and immunotherapy. Radiation therapy (e.g., stereotactic and hadron therapy) has been used for stage III NSCLC treatment [23], whereas immunotherapy for NSCLC uses monoclonal antibodies (e.g., atezolizumab, nivolumab, pembrolizumab, durvalumab, and avelumab) [22]. Immunotherapy has been used for stage III NSCLC and in clinical trials utilizing immunotherapy to treat NSCLC [24]. Surgical resection procedures have been recommended for stages I through IIIA of NSCLC [21]. However, NSCLC treatment is prone to failure due to variations in the cell regulatory mechanisms and the drug targets.

The therapeutic failure in NSCLC arises from cell-intrinsic mechanisms that promote cancer cells' survival, proliferation, and metastasis. These processes comprise of changes in the expression of drug transporters, activation of pro-survival and anti-apoptotic pathways, the influence of the tumor microenvironment, drug resistance, and severities of the drug metabolism [25]. Although the NSCLC therapy depends on the dosing regimens, protocols, and differences in genetic predisposition, it can cause serious adverse effects (AEs), such as anemia, nausea, vomiting, diarrhea, pneumonitis, psoriasis, and transient thyroid malfunction [26, 27]. Flavonoids such as diosmetin, genistein, fisetin, and epigallocatechin gallate (EGCG) are molecules that potentiate the effect of current chemotherapeutic drugs in cancer models [28].

3. Flavonoids

Drug discovery is a pivotal field that has demonstrated the pharmacological properties of a broad library of natural products (NPs), such as avermectin, artemisinin, cyclosporine, irinotecan, lovastatin, morphine, penicillin, and paclitaxel [29]. NPs are compounds derived from animals, microorganisms, and plants. Their structural diversity is classified into alkaloids, cardiac glycosides, cyanogenic glycosides, tannins, terpenes, steroids, saponins, phenylpropanoids, polyketides, polysaccharides, and flavonoids.

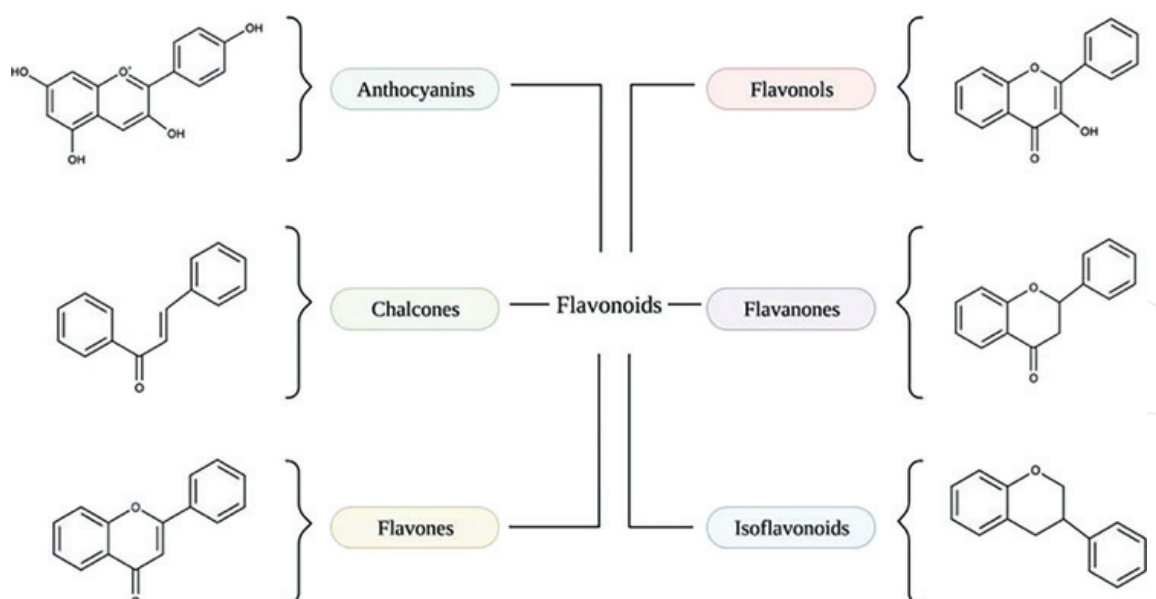


Figure 2.
 Flavonoid's classification: anthocyanins, chalcones, flavones, flavanones, and isoflavonoids.

Flavonoids are a large family of structurally and biosynthetically related polyphenolic compounds (~6000 members). Flavonoids are distributed in higher plants, and their general chemical structure is the flavan nucleus which consists of fifteen carbon atoms organized into three rings (C6-C3-C6). Two of those rings (A and B) are aromatic, whereas a three-carbon-atom heterocyclic ring connects them. The last ring, termed C, is an oxygen-containing pyran ring. According to their level of oxidation, conjugation, glycosidic residues, number of substituents, and enzymatic modification [30], flavonoids are classified into anthocyanins, chalcones, flavones, flavonols, flavanones, and isoflavonoids (**Figure 2**) [31]. Other classes of flavonoids include aurons and biflavones [32].

Although several factors influence the process of flavonoid biosynthesis (e.g., light, water availability, temperature, hormones, physical injuries, etc.) [33], the structural characteristics of these categories are determined by spectroscopy, preparation and extraction methods, chromatography, and spectrometry (e.g., mass spectrometry).

4. Properties of flavonoids: an emphasis in their anticancer potential

In plants, flavonoids are specialized metabolites that display several functions: defense against predators, protection against elements of biotic and abiotic stress, signaling molecules, detoxifying agents, growth regulators, inhibitors of protein components of pathogens, among others [34]. In human health, their antioxidant effects and vast pharmacological properties have driven their application in treating and preventing diseases, such as cancer, cardiovascular diseases, gastrointestinal disorders, and microbial infections. In addition, their anti-allergic, anti-inflammatory, antidiabetic, neuro, hepatic, and gastroprotective are attractive properties to the scientific community [35].

Flavonoids are convenient in cancer therapy because they can induce cytotoxicity on multidrug-resistant cancer cell lines, inhibit the transcription of oncogenes, ATP-dependent efflux pumps (i.e., P-glycoprotein), and glucose transporters [36].

The anticancer effects of flavonoids are due to their capacity to target major signaling pathways such as nuclear factor kappa B (NF- κ B), mitochondrial, and phosphatidylinositol-3-kinase (PI3K) pathways.

Besides, flavonoids inhibit cytochrome P450 (CYPs) enzymes (e.g., CYP1B1 and CYP1A) entailed in the metabolism of anticancer drugs [37]. Flavonoids such as fisetin, robinetin, myricetin, eupafolin, and hispidulin, are widely used to delay carcinogenesis, promote cytoprotection, genoprotection, and interfere with inflammatory and oxidative mechanisms *in vitro* and *in vivo* [38].

Inflammation plays a significant role in the evolution and potentiation of cancer invasion by increasing the production of reactive oxygen species (ROS), leading to oxidative DNA damage and reducing DNA repair. The abundance of other components such as chemokines or pro-inflammatory cytokines (e.g., tumor necrosis factor- α ; TNF- α) leads to cancer progression, angiogenesis, and neoplastic growth.

Overproduction of immune cells such as macrophages, mast cells, neutrophils, and T lymphocytes also contributes to cancer malignancy by releasing extracellular proteases, pro-angiogenic factors, inflammatory mediators, and favoring tumor microenvironment architecture [39].

Flavonoids regulate these events owing to their antioxidative and radical scavenging activities. Such NPs also modulate the production and gene expression of pro-inflammatory molecules, inflammation-related cells, and arachidonic acid metabolism enzymes activities [40]. In LC cells, many categories of flavonoids promote the activation of apoptosis and autophagy, cell adhesion, inhibit cell proliferation, retard invasion, impede metastasis, down-regulate epithelial-mesenchymal transition (EMT) in the tumor microenvironment, arrest cell cycle, and target cell signaling pathways. For this reason, it is possible to find the anti-lung cancer properties of various flavonoids in contemporary literature, such as apigenin, diosmetin, quercetin, hesperidin, catechin, genistein, and malvidin [39, 40]. In addition, luteolin, naringenin, kaempferol, and baicalein are flavonoids distributed on accessible natural sources, such as fruits, beverages, and plants.

5. Selected flavonoids

Luteolin, naringenin, kaempferol, and baicalein are flavonoids distributed on accessible natural sources such as fruits, beverages, and plants. Like other flavonoids, their biosynthesis occurs via the phenylpropanoid pathway, except that it can be enhanced with heterologous systems. Such natural products' pharmacological properties and mechanisms against human health threats have been documented over the last decade. However, there is a need to expand our knowledge about their capacity to face NSCLC treatment limitations and develop novel active research fields. Therefore, in the following sections, we present the sources, chemical characteristics, biosynthesis, pharmacokinetics, anticancer properties, and anti-NSCLC effects of our four selected flavonoids: luteolin, naringenin, kaempferol, and baicalein.

5.1 Luteolin

Luteolin (3',4',5,7-tetrahydroxyflavone) is a glycosylated flavone present in various plants such as celery, sweet bell peppers, carrots, etc. In addition, it has been isolated from fruits, tea, wine, and seeds. Luteolin has two benzene rings (A, B), an oxygen-containing ring (C), and a 2–3 carbon double bond; the hydroxyl groups are

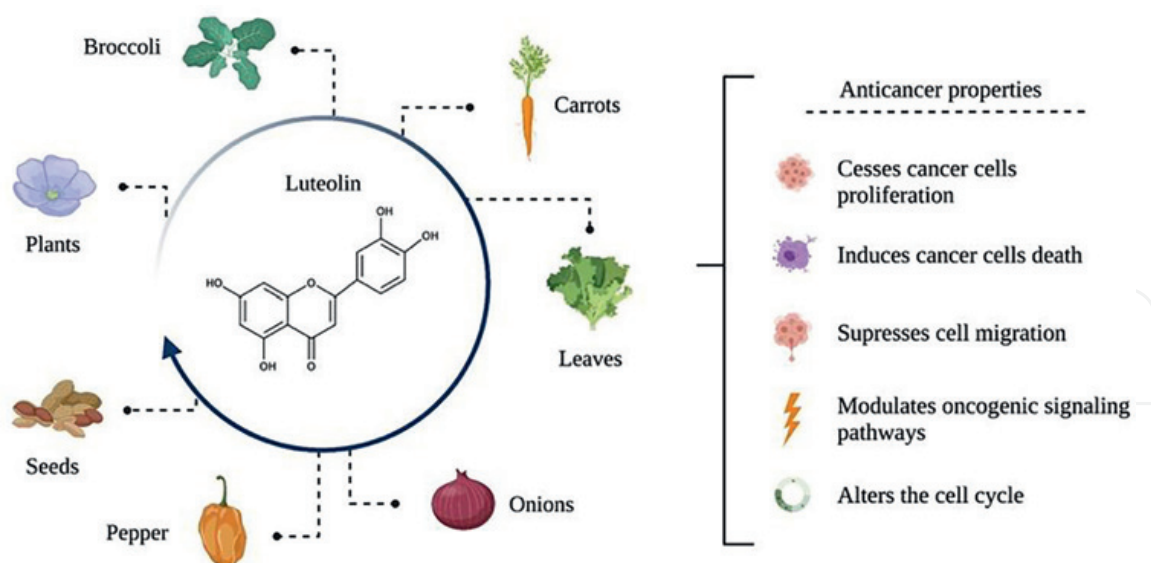


Figure 3.
 Chemical structure, sources, and anticancer properties of luteolin.

placed at carbons 5,7,3' and 4'; like other members of this category of NP, its condensed form is C6-C3-C6 (**Figure 3**) [41].

In plants, the synthesis of luteolin comprises of two different pathways, depending on the ring synthesis. The ring B is synthesized by the phenylpropanoid pathway by a sequential enzymatic activity, including the conversion of the amino acid phenylalanine into 4-coumaroyl-CoA by phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase, and 4-coumaroyl CoA ligase catalysis [42]. The ring A is then synthesized when a molecule of 4-coumaroyl-CoA is condensed with three molecules of malonyl-CoA by a chalcone synthase [43]. This process generates the intermediate naringenin, which by the activity of a chalcone isomerase, apigenin is transiently synthesized. This last intermediate is converted into luteolin by a flavonoid 3'-hydroxylase activity. This process has been replicated *de novo* in heterologous systems, such as *Streptomyces albus* [44].

Even though flavonoids metabolism and pharmacokinetic parameters have been areas of active research over the last decade, little is known about the bioavailability of flavones. Luteolin, like other flavonoids, is mainly absorbed in the small intestine and is biotransformed by liver and kidney enzymes (e.g., catechol-O-methyltransferase) [45]; its monoglucuronide and aglycone forms have been detected in human plasma [46]. In rat models, luteolin and its metabolites (e.g., luteolin 3'-O- β -D-glucuronide and luteolin-7-O- β -D-glucuronide) are distributed in the gastrointestinal tract, liver, kidney and lung, and the biliary excretion is the dominant elimination pathway of conjugated luteolin and its derivatives [47]. Other analogs, such as luteolin-7-O-glucoside, are excreted through urine and feces [48].

Luteolin acts as an anticancer agent against NSCLC by inhibition of proliferation of tumor cells, protection from carcinogenic stimuli, activation of cell cycle arrest, induction of apoptosis through different signaling pathways (e.g., c-Jun N-terminal kinase (JNK), phosphatidylinositol 3-kinase (PI3K)/protein kinase B (Akt), protein kinase (PKC), etc.), induction of the epithelial biomarker E-cadherin expression, reversion of the EMT process, and by modulation of microRNAs (miRNAs) expression [49, 50].

For instance, Jiang and coworkers [51] revealed the tumor-suppressive activity of luteolin on NSCLC cancer cell lines A549 and NCI-H460. In this study, the effect of

luteolin was observed in a dose-dependent manner with a half-maximal inhibitory concentration (IC_{50}) of $40\ \mu\text{M}$ [51]. The tumor inhibitory effect of 50, 100, and 200 mg/kg/day luteolin was investigated using the NCI-H460 xenograft nude mice model [51]. Moreover luteolin treatment upregulated the expression of potential tumor suppressors such as miR-34a-5p, and increased the activation of cysteine proteases (i.e., caspase-3 and -9) involved in cell death by apoptosis [51]. In another study, the effect of luteolin on A549 cell migration was assessed. Treatment with 10, 20, and $40\ \mu\text{M}$ of luteolin reduced the number of invading cells across the matrix and transwell membrane in a concentration-dependent manner [52]. In the same concentration range, luteolin reduced the number of filopodia, suppressed the expression of activated focal-adhesion kinase (FAK), and activated proto-oncogene tyrosine-protein kinase (SRC). In another study, treatment with 5, 10, 20, or $40\ \mu\text{M}$ luteolin arrested the cell cycle of NCI-H1975 and NCI-H1650 cells at G1 phase with a concomitant decrease in the expression of cell cycle regulators, such as cyclin D1 and D3 [53].

5.2 Naringenin

Naringenin (2,3-dihydro-5,7-dihydroxy-2-(4-hydroxyphenyl)-4H-1-benzopyran-4-one) and its derivatives (naringenin-7-rhamnoglucoside and naringenin-7-glucoside) are naturally occurring flavanones found in many edible fruits (e.g., lemon peels, pomelo, grapefruits, oranges, and tomatoes) and vegetables [54–56]. This flavanone is derived from the hydrolysis of its glycone forms naringenin and narirutin (**Figure 4**) [57].

The biosynthesis of naringenin occurs via shikimic acid and acylpolymalonate metabolic pathways. Naringenin possesses 2 aromatic rings joined by a linear 3-carbon chain (C6-C3-C6) with a saturated 3-carbon chain and an oxygen atom at carbon 4. Depending on the arrangement of related functional groups in its structure, naringenin displays a wide array of therapeutic properties such as neuroprotective, antimicrobial, antiviral, antifibrotic, antidiabetic, and anti-inflammatory [58, 59]. In this context, the high reactivity of hydroxyl groups and the presence of carbohydrate moieties confers antioxidant advantages against a variety of radical species [60].

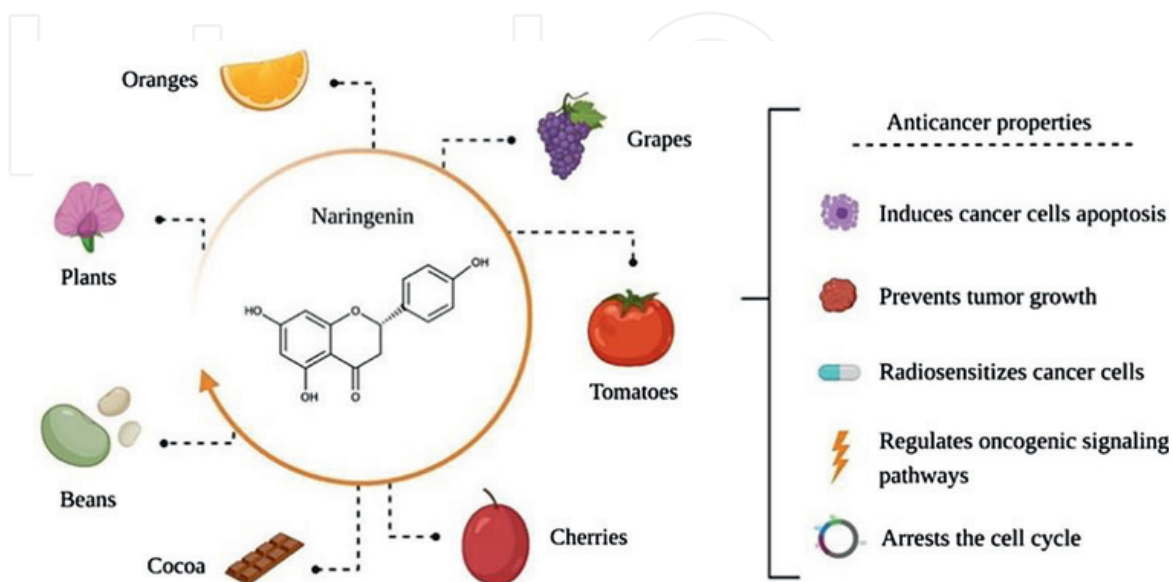


Figure 4.
Chemical structure, sources, and anticancer properties of naringenin.

Regarding its pharmacokinetic properties, the absorption of naringenin takes place through both passive diffusion and active transport. In rats, naringenin and its derivative naringin, have been identified in the lung, trachea, kidney, liver, spleen, heart, muscle, fat, and brain tissues [61]. In order to be metabolized, naringenin undergoes glucuronidation and sulfation with its major metabolites naringenin-*o*- β -D-glucuronide, *p*-hydroxybenzoic acid, *p*-hydroxyphenylpropionic acid, and *p*-coumaric acid, are found in liver, kidney, heart, brain, plasma, and urine. Naringenin excretion occurs through the biliary (naringenin 7-glucuronide 4'-sulfate and naringenin 7-glucuronide) and urinary (naringenin 7,4'-disulfate and naringenin 4'-glucuronide) pathways [62].

Naringenin exhibits antidiabetic, anti-inflammatory, anti-atherosclerotic, hepatoprotective, nephroprotective, and neuroprotective properties [63]. Recently, the *O*-ethyl, *O*-dodecyl, *O*-methyl, and *O*-pentyl derivatives of naringenin have been recognized for its anticancer, and antimicrobial properties [64]. The anticancer potential of naringenin has showed that it can induce apoptosis and cell cycle arrest in a variety of cancer cell lines (e.g., breast tumor, glioblastoma, lung tumor) [65], act as an anti-angiogenic chemopreventive agent [66], suppress mutations, modulate oxidative stress, regulate signaling factors involved in the inflammatory response (e.g., interleukin-8, interleukin-6, inducible nitric oxide synthase, etc.), inhibit carcinogenesis, induce major pathways of apoptosis, and promote caspase activation [67, 68]. Similar effects have been observed for other naringenin derivatives such as 5,4'-dihydroxy flavanone-7-yl [69].

In NSCLC, the effect of naringenin has been evaluated in a few models. For example, in A549 cells, the molecular mechanism of the naringenin-induced apoptotic event has been evaluated [70]. In this study, the growth inhibition effect was observed with 100 and 200 μ M of naringenin, and an apoptosis activation with a concentration of 800 μ M through the activation of caspase-3 and -9 [71]. In another study, the effect of current therapies such as radiation therapy has been improved with naringenin treatment. For example, naringenin promoted radiosensitization of NCI-H23 cells [71]. In such study, the cancer cells' survival was lowered by 50% with a supplementation of 100 μ M naringenin before treatment with radiation. After radiation treatment, naringenin acted as a radiosensitizer and radioprotector at 6 Gy. On the other hand, naringenin treatment decreased the expression of protease enzymes such as metalloproteinase-2 (MMP-2) and -9 (MMP-9) and increased the expression of caspase-3 [71]. Similarly, another study reported that 10, 100, and 200 μ mol/L naringenin treatment increased 10–140% mRNA expression of caspase-3 and decreased 6–55% MMP-2 and 5–60% MMP-9 expression levels [72].

5.3 Kaempferol

Kaempferol (3,5,7-trihydroxy-2-(4-hydroxyphenyl)-4H-chromene-4-one) and its derivatives (e.g., kaempferol-3-*O*-rhamnoside, kaempferol-3-*O*-arabinofuranoside, kaempferol-3-*O*-(6-*p*-coumaroyl)-glucoside, kaempferol-3-*O*-(2,6-di-*p*-coumaroyl)-glucoside, etc.) are polyphenols widely distributed in foods (e.g., green chili, asparagus, beans, blueberries, etc.) fruits (e.g., strawberry and gooseberry), vegetables (e.g., cauliflower), beverages (e.g., grapefruit juice, red raspberry juice, teas, etc.) [73], and plants belonging to the families *Aspidiaceae*, *Aspleniaceae*, *Polypodiaceae*, and *Iridaceae* (**Figure 5**) [74].

Structurally, kaempferol is synthesized from diphenyl propane via condensation of three molecules of malonyl-CoA and one molecule of 4-coumaroyl-CoA.

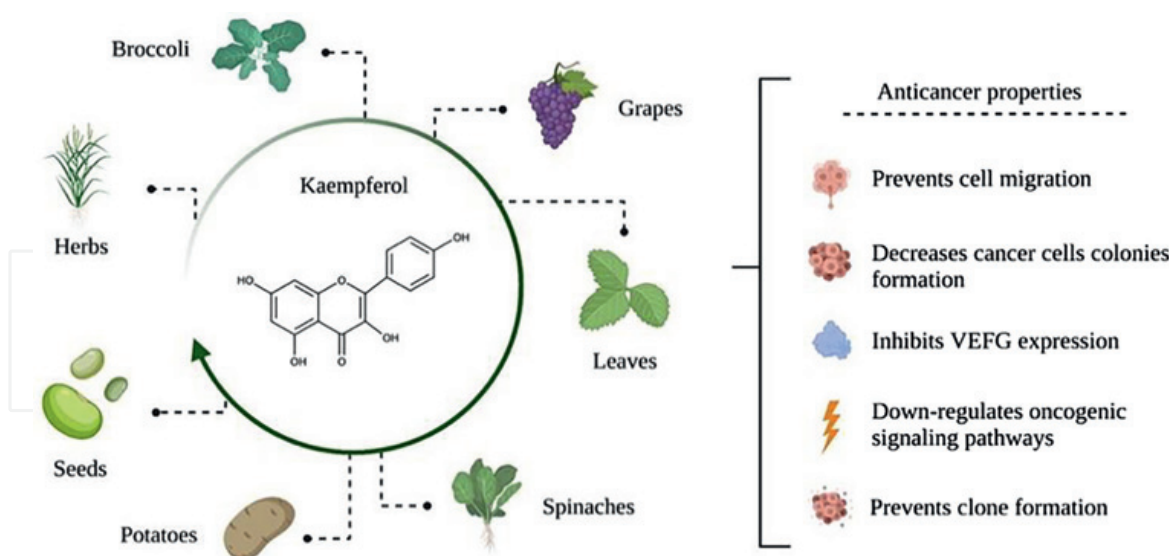


Figure 5.
Chemical structure, sources, and anticancer properties of kaempferol.

Kaempferol is absorbed from the small intestine by passive diffusion and metabolized into 4-methyl phenol, 4-hydroxyphenyl acetic acid, and phloroglucinol by the normal flora of the colon [75]. Kaempferol metabolites (e.g., kaempferol-3-glucuronide and kaempferol mono- and di-sulfates) are distributed in body tissues, blood, and plasma. Their excretion occurs through feces and urine [76, 77].

Kaempferol exhibits a handful of pharmacological properties, such as anticancer, antidiabetic, anti-obesity, antimicrobial, anti-inflammatory, antiallergic, antiplatelet aggregation, anti-bone disorders, and cardiovascular protection [78]. The anti-cancer potential of kaempferol is primarily achieved by inhibition of the proliferation of cancer cells by triggering apoptosis (e.g., activation of caspases-3, -7, and -9), cell cycle arrest at G2/M phase, downregulation of signaling pathways and expression of EMT-related markers, prevention of the accumulation of reactive oxygen species (ROS), inhibition of angiogenesis [79], sensitization of cancer cell growth to chemotherapeutic pharmaceuticals (e.g., ovarian cancer cells to cisplatin) [80], among other effects.

During the study of therapeutic molecules against NSCLC, kaempferol treatment (50 μ M) inhibited the proliferation of A549 (51%) and NCI-H460 (57%) cells, increased the expression of cleaved caspase-3, -9, and Bax, and arrested the cell cycle at the G1 phase [79]. The same dose of kaempferol down-regulated the transcription of major cellular oxidative and inflammatory regulators such as nuclear factor erythroid 2-related factor 2 (Nrf2) [79]. Finally, 25 μ M of kaempferol inhibited the key metastatic inducers tumor growth factor β 1 (TGF- β 1) of epithelial-mesenchymal transition (EMT) and herein, interfering with the activation of downstream effectors involved in NSCLC progression [80].

5.4 Baicalein

Baicalein (5,6,7-trihydroxyflavone) is the major component of the roots of *Scutellaria baicalensis* Georgi and *Scutellaria lateriflora*. Its glycone form (5,6-dihydroxy-7-O-glucuronide) is also an important flavone constituent of *S. baicalensis*. Lower concentrations of baicalein can be found in other sources, such as foods

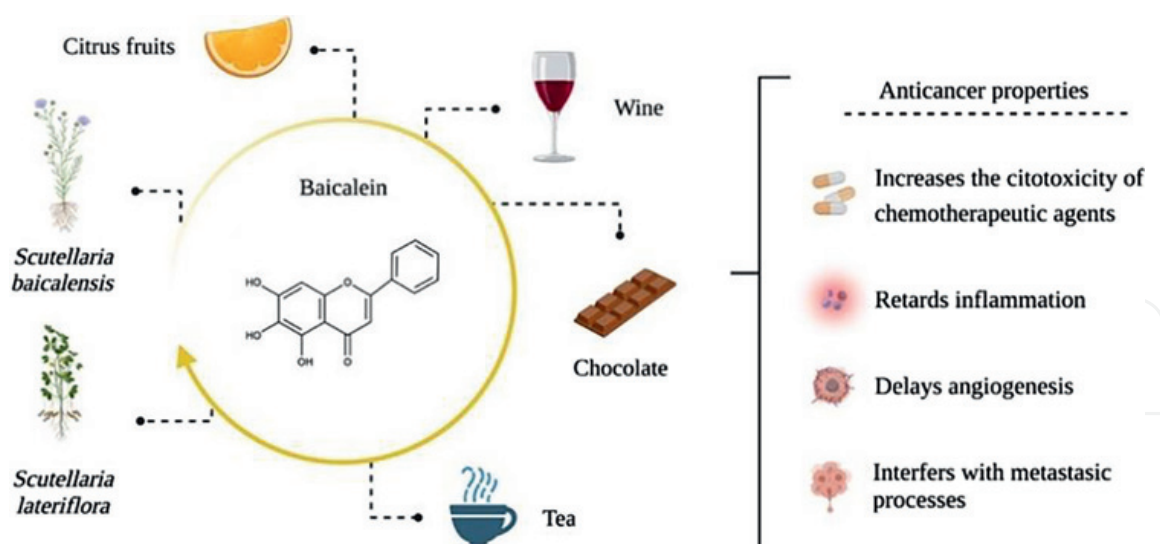


Figure 6.
 Chemical structure, sources, and anticancer properties of baicalein.

(e.g., citrus fruits and chocolate) and beverages (e.g., wine and tea) [81]. The structure of baicalein lacks a 4'-hydroxyl group on its ring B compared to other flavones (**Figure 6**).

Although baicalein is synthesized by the flavonoid pathway (which is part of phenylpropanoid metabolism), recent evidence suggests that because of the lack of a 4'-hydroxyl group, an alternative path recruits cinnamic acid to form cinnamoyl-coenzyme A (CoA) through a CoA ligase. This modification will allow a condensation with malonyl-CoA by a chalcone synthase to form a chalcone, and then isomerized by chalcone isomerase to form pinocembrin. This intermediate could be converted by a flavone synthase to form chrysin, and then modified by hydroxylases, methyltransferases, and glycosyltransferases to form baicalein [82].

Comparably, baicalin is synthesized via the phenylpropanoid pathway by many enzymes (e.g., phenylalanine ammonia-lyase, cinnamate 4-hydroxylase, chalcone synthase, etc.); baicalin can be catalyzed back to baicalein by UDP-glucuronate by the baicalein 7-O-glucuronosyl transferase [83].

Even with the complex metabolic process of most flavones, baicalein was found to be ample in the small and large intestine of mice, especially its glucuronide and glycosylated forms, whereas its dehydroxylated, methylated, and sulfated metabolites have been found in the entire intestine [84, 85].

The safety, tolerability, and pharmacokinetics of oral baicalein tablets (200, 400, and 600 mg) in a phase I single-center study of healthy Chinese subjects were assessed [86]. Results showed that baicalein tablets are safe and well-tolerated with mild adverse events (e.g., proteinuria). The excretion of baicalein and its metabolites (e.g., baicalein-6-O-glucuronide, baicalein-6-O-sulfate, baicalein-7-O-glucuronide, etc.) was determined mainly in the urine.

Baicalein has multiple biological properties, including cardioprotective, neuroprotective, hepatoprotective, attenuates renal dysfunction, increases the cytotoxicity of chemotherapeutic agents (e.g., cisplatin) [87], antibacterial [88], anti-inflammatory, antiviral (e.g., against dengue virus) [89], and antioxidant [90]. Furthermore, in cancer, baicalein has the potential to decrease cancer growth and retard cancer-promoting processes such as angiogenesis, inflammation, and metastasis [91]. Hence, baicalein properties have been studied in distinct cancers like breast cancer, cervical

cancer, gastric cancer, hepatocellular carcinoma, multiple myeloma, pancreatic cancer, prostate cancer, and LC [92].

In another study where 10 or 40 $\mu\text{mol/L}$ baicalein treatment was used, an inhibition of A549 and NCI-H1299 cell invasion and metastasis was observed. Also, a down-regulation of the membrane-cytoskeleton linker proteins (i.e., ezrin) was measured [93].

The combination of baicalein and the chemotherapeutic docetaxel was investigated for a potential synergism against the proliferation of A549 and LCC cells [94]. Results showed an increase in the percentage of apoptotic cells compared with baicalein or docetaxel individually tested. In addition, the combination of baicalein plus docetaxel arrested the cell cycle of A549 cells at the G1 phase, whereas for LLC cells, it was blocked in the G2/M phase. These events are related to the suppression of cyclin D1, cyclin-dependent kinase 4 (CDK4), cyclin B1, cyclin-dependent kinase 6 (CDK6), and cell division cycle 25C (CDC25C) expression in both cell lines.

6. Discussion

Flavonoids constitute a broad group of secondary metabolites with fascinating pharmacological and nutraceutical properties. Since their discovery in the 1930s by Nobel Prize Dr. Albert Szent-Gyorgyi, many analytical methods have been settled to enlighten their chemical characteristics, differences with other subclasses of natural NP, quality, pharmacokinetics, and medicinal properties.

The anticancer and antioxidant activities of flavonoids are subjected to their structural organization, functional groups, and the total number and configuration of hydroxyl groups. For example, while the ortho-dihydroxy structure in the B-ring and the 2,3-double bond in conjugation with the 4-oxo function of the C-ring of luteolin is responsible for its antioxidant capacity [95]. In addition, the hydroxyl group at C3 in the C ring of kaempferol is important for its biological activity [96]. The hydroxyl group and moieties on the C7 atom in the structure of naringenin influences its antioxidant activity [97]. The antioxidant and anticancer activity of baicalein is attributed to seven double bonds in the two aromatic rings and three hydroxyls on ring A [97].

Screening therapeutic molecules against NSCLC is developed through *in vitro* and *in vivo* models. *In vitro* models include cancer cell lines and organoids; the former is a 3D classical helpful approach to expand our knowledge about the molecular and cell biology of various types of cancer in a limitless replicative manner [98]. This study alluded to several NSCLC cell lines, such as A549, LLC, NCI-460, NCI-H1299, NCI-H1975, NCI-H23, and NCI-H1650. However, other NSCLC cell lines include NCI-H358, TL-1, HCC4006, HCC827, and NCI-226 [99].

AEs, low solubility, cytotoxicity, incompatible pharmacokinetics, and drug resistance are significant drawbacks of NSCLC treatment. AEs are negative outcomes caused by medical interventions, some examples include prolongation of the hospital stay, injuries, side effects, or death [100]. On the other hand, drug resistance is an intrinsic or acquired serious concern that contributes to the growth, survival, chemoresistance, and evolution of NSCLC. This event occurs due to overexpression of receptors, increased cell signaling, mutations, and protein fusion. For example, the overexpression of EGFR and the fusion of ALK with echinoderm microtubule-associated protein-like 4 (EML4) have been correlated with poor response to antibodies treatment, poor chemo-sensitivity, lower survival, lower response rate, NSCLC progression, and metastasis [101]. Therefore, a strategy to combat drug resistance is

the administration of two or more drugs. We found that the synergism of baicalein plus docetaxel is an effective, safe, non-toxic approach to induce cell cycle arrest at the G1 and G2/M phase, apoptotic events, and inhibit tumor growth angiogenesis in A549 and LLC cells.

On the other hand, the reports mentioned in this work identified that luteolin suppressed essential cell cycle regulators, such as cyclin D1 and cyclin D3. Such results are important because the overexpression of cyclin D1 and D3 is commonly associated with the evolution and progression of several types of cancer, including LC and NSCLC [102]. CDK4 and CDK6 are necessary proteins for the regulatory activity of cyclin D1 and cyclin D3. We described baicalein as a powerful flavone that, in combination with docetaxel, inhibited CDK4 and CDK6 expression. Inhibition of both proteins translated into the arrest of the cell cycle, down-regulation of CDC25C, cyclin B1, and β -catenin on NSCLC models. We commented on the influence of naringenin on members of the caspase cascade (caspase-3 and -9) and the effect of kaempferol and its analogs on the inhibition of cellular regulators such as Nrf2.

7. Conclusions

This work presented a general scenario of the classification, risk factors, genetic alterations, diagnosis, and treatment of NSCLC. We highlighted the anticancer potential of four selected flavonoids against NSCLC: luteolin, naringenin, kaempferol, and baicalein. We documented their sources, derivatives, biosynthesis, pharmacokinetics, and their capacity to inhibit NSCLC-promoting processes. We compiled recent evidence about their ability to interfere with major signaling pathways, transcription factors, membrane-cytoskeleton proteins, and intrinsic apoptotic pathways on NSCLC cell lines.

We also addressed the possibility of using them with current chemotherapy medications or as radiosensitizers and radioprotectors. Finally, we mentioned differences in their structure-activity relationships and summarized their effect on fundamental cellular processes, such as cell proliferation, death, metabolism, migration, and fate. Even though this review presented luteolin, naringenin, kaempferol, and baicalein as prospective drug candidates for NSCLC treatment, further studies are required to assess the efficacy safety and improve their bioavailability.

Acknowledgements

Jorge L. Mejía-Méndez thanks Consejo Nacional de Ciencia y Tecnología (CONACyT) for his doctoral fellowship.

Conflict of interest

The authors declare no conflict of interests.

Notes/thanks/other declarations

Authors do not have other declarations.

IntechOpen

Author details


Jorge L. Mejía-Méndez¹, Horacio Bach², Luis Ricardo Hernández¹
and Eugenio Sánchez-Arreola^{1*}

1 Department of Chemical Biological Sciences, Universidad de las Américas Puebla,
Puebla, Mexico

2 Department of Medicine, Division of Infectious Diseases, University of British
Columbia, Vancouver, BC, Canada

*Address all correspondence to: eugenio.sanchez@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Stein CJ, Colditz GA. Modifiable risk factors for cancer. *British Journal of Cancer*. 2004;**90**:299-303. DOI: 10.1038/sj.bjc.6601509
- [2] Global Cancer Observatory. Estimated Number of Deaths in 2020, Worldwide, both Sexes, All Ages [Internet]. 2020. Available from: <https://gco.iarc.fr/> [Accessed: February 20, 2022]
- [3] National Cancer Institute. Cancer Stat Facts: Lung and Bronchus Cancer [Internet]. 2021. Available from: <https://seer.cancer.gov/statfacts/html/lungb.html> [Accessed: February 20, 2022]
- [4] Zappa C, Mousa SA. Non-small cell lung cancer: Current treatment and future advances. *Translational Lung Cancer Research*. 2016;**5**(3):288-300. DOI: 10.21037/tlcr.2016.06.07
- [5] Reck M, Rabe KF. Precision diagnosis and treatment for advanced non-small-cell lung cancer. *The New England Journal of Medicine*. 2017;**377**:849-861. DOI: 10.1056/NEJMra1703413
- [6] Patel B, Das S, Prakash R, Yasir M. Natural bioactive compound with anticancer potential. *International Journal of Advances in Pharmaceutical Sciences*. 2010;**2010**:1
- [7] Romaszko AM, Doboszyńska A. Multiple primary lung cancer: A literature review. *Advances in Clinical and Experimental Medicine*. 2018;**27**: 725-730. DOI: 10.17219/acem/68631
- [8] Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. *The European Respiratory Journal*. 2016;**48**:889-902. DOI: 10.1183/13993003.00359-2016
- [9] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA: A Cancer Journal for Clinicians*. 2021;**71**:7-33. DOI: 10.3322/caac.21654
- [10] Yuan M, Huang LL, Chen JH, Wu J, Xu Q. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy*. 2019;**4**:61. DOI: 10.1038/s41392-019-0099-9
- [11] Christopoulos P, Budczies J, Kirchner M, Dietz S, Sülthmann H, Thomas M, et al. Defining molecular risk in ALK⁺ NSCLC. *Oncotarget*. 2019;**10**:3093-3103. DOI: 10.18632/oncotarget.26886
- [12] Sarihan S, Ercan I, Saran A, Cetintas SK, Akalin H, Engin K. Evaluation of infections in non-small cell lung cancer patients treated with radiotherapy. *Cancer Detection and Prevention*. 2005;**29**:181-188. DOI: 10.1016/j.cdp.2004.11.001
- [13] Chakravarthi BV, Nepal S, Varambally S. Genomic and epigenomic alterations in cancer. *The American Journal of Pathology*. 2016;**186**(7):1724-1735. DOI: 10.1016/j.ajpath.2016.02.023
- [14] Wheeler DA, Wang L. From human genome to cancer genome: The first decade. *Genome Research*. 2013;**23**: 1054-1062. DOI: 10.1101/gr.157602.113
- [15] Johnson JL, Pillai S, Chellappan SP. Genetic and biochemical alterations in non-small cell lung cancer. *Biochemistry Research International*. 2012;**2012**:940405. DOI: 10.1155/2012/940405
- [16] Fois SS, Paliogiannis P, Zinellu A, Fois AG, Cossu A, Palmieri G. Molecular

epidemiology of the main druggable genetic alterations in non-small cell lung cancer. *International Journal of Molecular Sciences*. 2021;**22**:612. DOI: 10.3390/ijms22020612

[17] Thammakumpee K. Clinical manifestation and survival of patients with non-small cell lung cancer. *Journal of the Medical Association of Thailand*. 2004;**87**(5):503-507

[18] Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, et al. Methods for staging non-small cell lung cancer: Diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;**143**:e211S-e250S. DOI: 10.1378/chest.12-2355

[19] Schrevels L, Lorent N, Doooms C, Vansteenkiste J. The role of PET scan in diagnosis, staging, and management of non-small cell lung cancer. *The Oncologist*. 2004;**9**:633-643. DOI: 10.1634/theoncologist.9-6-633

[20] Quint LE. Staging non-small cell lung cancer. *Cancer Imaging*. 2007;**7**(1):148-159. DOI: 10.1102/1470-7330.2007.0026

[21] Collins LG, Haines C, Perkel R, Enck RE. Lung cancer: Diagnosis and management. *American Family Physician*. 2007;**75**:56-63

[22] Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;**553**:446-454. DOI: 10.1038/nature25183

[23] Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clinic Proceedings*. 2008;**83**:584-594. DOI: 10.4065/83.5.584

[24] Nasser NJ, Gorenberg M, Agbarya A. First line immunotherapy for non-small cell lung cancer. *Pharmaceuticals (Basel)*. 2020;**13**:373. DOI: 10.3390/ph13110373

[25] Sosa Iglesias V, Giuranno L, Dubois LJ, Theys J, Vooijs M. Drug resistance in non-small cell lung cancer: A potential for NOTCH targeting? *Frontiers in Oncology*. 2018;**8**:267. DOI: 10.3389/fonc.2018.00267

[26] Muthu V, Myllemngap B, Prasad KT, Behera D, Singh N. Adverse effects observed in lung cancer patients undergoing first-line chemotherapy and effectiveness of supportive care drugs in a resource-limited setting. *Lung India*. 2019;**36**:32-37. DOI: 10.4103/lungindia.lungindia_321_17

[27] Zarogoulidis P, Chinelis P, Athanasiadou A, Tsiouda T, Trakada G, Kallianos A, et al. Possible adverse effects of immunotherapy in non-small cell lung cancer; treatment and follow-up of three cases. *Respiratory Medicine Case Reports*. 2017;**22**:101-105. DOI: 10.1016/j.rmcr.2017.07.004

[28] Zanoaga O, Braicu C, Jurj A, Rusu A, Buiga R, Berindan-Neagoe I. Progress in research on the role of flavonoids in lung cancer. *International Journal of Molecular Sciences*. 2019;**20**:4291. DOI: 10.3390/ijms20174291

[29] Katiyar C, Gupta A, Kanjilal S, Katiyar S. Drug discovery from plant sources: An integrated approach. *Ayu*. 2012;**33**:10-19. DOI: 10.4103/0974-8520.100295

[30] Kushwaha PP, Vardhan PS, Kumari P, Mtewa AG, Kumar S. Bioactive lead compounds and targets for the development of antimalarial drugs. In: Egbuna C, Kumar S, Ifemeje JC, Ezzat SM, Kaliyaperumal S, editors. *Phytochemicals as Lead*

Compounds for New Drug Discovery. Amsterdam, Netherlands: Elsevier; 2020. pp. 305-316. DOI: 10.1016/B978-0-12-817890-4.00020-2

[31] Hinderer W, Seitz HU. Flavonoids. In: Constabel F, Vasil IK, editors. *Phytochemicals in Plant Cell Cultures*. Amsterdam, Netherlands: Elsevier; 1988. pp. 23-48. DOI: 10.1016/B978-0-12-715005-5.50009-1

[32] Pietta PG. Flavonoids as antioxidants. *Journal of Natural Products*. 2000;**63**:1035-1042. DOI: 10.1021/np9904509

[33] Dias MC, Pinto DCGA, Silva AMS. Plant flavonoids: Chemical characteristics and biological activity. *Molecules*. 2021;**26**:5377. DOI: 10.3390/molecules26175377

[34] Havsteen BH. The biochemistry and medical significance of the flavonoids. *Pharmacology & Therapeutics*. 2002;**96**(2-3):67-202. DOI: 10.1016/S0163-7258(02)00298-X

[35] Kavitha RV, Kumar JR, Egbuna C, Ifemeje JC. Phytochemicals as therapeutic interventions in neurodegenerative diseases. In: Egbuna C, Kumar S, Ifemeje JC, Ezzat SM, Kaliyaperumal S, editors. *Phytochemicals as Lead Compounds for New Drug Discovery*. Amsterdam, Netherlands: Elsevier; 2020. pp. 161-178. DOI: 10.1016/B978-0-12-817890-4.00010-X

[36] Chambers CS, Viktorová J, Řehořová K, Biedermann D, Turková L, Macek T, et al. Defying multidrug resistance! Modulation of related transporters by flavonoids and flavonolignans. *Journal of Agricultural and Food Chemistry*. 2020;**68**(7):1763-1779. DOI: 10.1021/acs.jafc.9b00694

[37] Lou JS, Yao P, Tsim KWK. Cancer treatment by using traditional Chinese

medicine: Probing active compounds in anti-multidrug resistance during drug therapy. *Current Medicinal Chemistry*. 2018;**25**:5128-5141. DOI: 10.2174/0929867324666170920161922

[38] Kicinska A, Jarmuszkiewicz W. Flavonoids and mitochondria: Activation of cytoprotective pathways? *Molecules*. 2020;**25**(13):3060. DOI: 10.3390/molecules25133060

[39] Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;**420**:860-867. DOI: 10.1038/nature01322

[40] García-Lafuente A, Guillaumon E, Villares A, Rostagno MA, Martínez JA. Flavonoids as anti-inflammatory agents: Implications in cancer and cardiovascular disease. *Inflammation Research*. 2009;**58**(9):537-552. DOI: 10.1007/s00011-009-0037-3

[41] Lin Y, Shi R, Wang X, Shen HM. Luteolin, a flavonoid with potential for cancer prevention and therapy. *Current Cancer Drug Targets*. 2008;**8**(7):634-646. DOI: 10.2174/156800908786241050

[42] Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: Biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science*. 2012;**3**:222. DOI: 10.3389/fpls.2012.00222

[43] Tsao R. Chemistry and biochemistry of dietary polyphenols. *Nutrients*. 2010;**2**(12):1231-1246. DOI: 10.3390/nu2121231

[44] Marín L, Gutiérrez-Del-Río I, Yagüe P, Manteca Á, Villar CJ, Lombó F. *De Novo* biosynthesis of Apigenin, Luteolin, and Eriodictyol in the Actinomycete *Streptomyces albus* and production improvement by feeding and spore conditioning. *Frontiers in*

Microbiology. 2017;**8**:921. DOI: 10.3389/fmicb.2017.00921

[45] Hollman PCH. Absorption, bioavailability, and metabolism of flavonoids. *Pharmaceutical Biology*. 2004;**42**:74-83. DOI: 10.3109/13880200490893492

[46] Manach C, Donovan JL. Pharmacokinetics and metabolism of dietary flavonoids in humans. *Free Radical Research*. 2004;**38**:771-785. DOI: 10.1080/10715760410001727858

[47] Deng C, Gao C, Tian X, Chao B, Wang F, Zhang Y, et al. Pharmacokinetics, tissue distribution and excretion of luteolin and its major metabolites in rats: Metabolites predominate in blood, tissues and are mainly excreted via bile. *Journal of Functional Foods*. 2017;**35**:332-340. DOI: 10.1016/j.jff.2017.05.056

[48] Liu H, Huo X, Ding L, Feng X, Jiang M, Pan G, et al. Metabolic profiling of luteolin-7-O-glucoside in rat urine, plasma, bile and feces after oral administration using ultra-high-performance liquid chromatography/quadrupole time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*. 2016;**30**:447-459. DOI: 10.1002/rcm.7456

[49] Imran M, Rauf A, Abu-Izneid T, Nadeem M, Shariati MA, Khan IA, et al. Luteolin, a flavonoid, as an anticancer agent: A review. *Biomedicine & Pharmacotherapy*. 2019;**112**:108612. DOI: 10.1016/j.biopha.2019.108612.

[50] Han K, Meng W, Zhang JJ, Zhou Y, Wang YL, Yang S, et al. Luteolin inhibited proliferation and induced apoptosis of prostate cancer cells through miR-301. *Oncotargets and Therapy*. 2016;**9**:3085-3094. DOI: 10.2147/OTT.S102862

[51] Jiang ZQ, Li MH, Qin YM, Jiang HY, Zhang X, Wu MH. Luteolin inhibits tumorigenesis and induces apoptosis of non-small cell lung cancer cells via regulation of MicroRNA-34a-5p. *International Journal of Molecular Sciences*. 2018;**19**:447. DOI: 10.3390/ijms19020447

[52] MasraksaW, TanasawetS, HutamekalinP, Wongtawatchai T, Sukketsiri W. Luteolin attenuates migration and invasion of lung cancer cells via suppressing focal adhesion kinase and non-receptor tyrosine kinase signaling pathway. *Nutrition Research and Practice*. 2020;**14**(2):127-133. DOI: 10.4162/nrp.2020.14.2.127

[53] Zhang M, Wang R, Tian J, Song M, Zhao R, Liu K, et al. Targeting LIMK1 with luteolin inhibits the growth of lung cancer in vitro and in vivo. *Journal of Cellular and Molecular Medicine*. 2021;**25**:5560-5571. DOI: 10.1111/jcmm.16568

[54] Mbaveng AT, Zhao Q, Kuete V. Harmful and protective effects of phenolic compounds from African medicinal plants. In: Kuete V, editor. *Toxicological Survey of African Medicinal Plants*. Amsterdam, Netherlands: Elsevier; 2014. pp. 577-609. DOI: 10.1016/B978-0-12-800018-2.00020-0

[55] Ni H, Zhang SF, Gao QF, Hu Y, Jiang ZD, Chen F. Development and evaluation of simultaneous quantification of naringin, prunin, naringenin, and limonin in citrus juice. *Food Science and Biotechnology*. 2015;**24**:1239-1247. DOI: 10.1007/s10068-015-0159-z

[56] Venkateswara Rao P, Kiran SDVS, Rohini P, Bhagyasree P. Flavonoid: A review on Naringenin. *Journal of Pharmacognosy and Phytochemistry*. 2017;**6**:2778-2783

- [57] Nouri Z, Fakhri S, El-Senduny FF, Sanadgol N, Abd-ElGhani GE, Farzaei MH, et al. On the neuroprotective effects of naringenin: Pharmacological targets, signaling pathways, molecular mechanisms, and clinical perspective. *Biomolecules*. 2019;**9**:690. DOI: 10.3390/biom9110690
- [58] Jadeja RN, Devkar RV. Polyphenols and flavonoids in controlling non-alcoholic steatohepatitis. In: Watson RR, Preedy RV, Zibadi S, editors. *Polyphenols in Human Health and Disease*. Amsterdam, Netherlands: Academic Press; 2014. pp. 615-623. DOI: 10.1016/B978-0-12-398456-2.00047-5
- [59] Ramos-Tovar E, Muriel P. Phytotherapy for the liver. In: Watson R, Preedy V, editors. *Dietary Interventions in Liver Disease*. Amsterdam, Netherlands: Academic Press; 2019. pp. 101-121. DOI: 10.1016/B978-0-12-814466-4.00009-4
- [60] Heim KE, Tagliaferro AR, Bobilya DJ. Flavonoid antioxidants: Chemistry, metabolism and structure-activity relationships. *The Journal of Nutritional Biochemistry*. 2002;**13**(10):572-584. DOI: 10.1016/S0955-2863(02)00208-5
- [61] Zeng X, Su W, Zheng Y, He Y, He Y, Rao H, et al. Pharmacokinetics, tissue distribution, metabolism, and excretion of naringin in aged rats. *Frontiers in Pharmacology*. 2019;**10**:34. DOI: 10.3389/fphar.2019.00034
- [62] Joshi R, Kulkarni YA, Wairkar S. Pharmacokinetic, pharmacodynamic and formulations aspects of Naringenin: An update. *Life Sciences*. 2018;**215**:43-56. DOI: 10.1016/j.lfs.2018.10.066
- [63] Den Hartogh DJ, Tsiani E. Antidiabetic properties of naringenin: A citrus fruit polyphenol. *Biomolecules*. 2019;**9**:99. DOI: 10.3390/biom9030099
- [64] Sargazi ML, Juybari KB, Tarzi ME, Amirkhosravi A, Nematollahi M, Mirzamohammadi S, et al. Naringenin attenuates cell viability and migration of C6 glioblastoma cell line: A possible role of hedgehog signaling pathway. *Molecular Biology Reports*. 2021;**48**:6413-6421. DOI: 10.1007/s11033-021-06641-1
- [65] Choi J, Lee DH, Jang H, Park SY, Seol JW. Naringenin exerts anticancer effects by inducing tumor cell death and inhibiting angiogenesis in malignant melanoma. *International Journal of Medical Sciences*. 2020;**17**:3049-3057. DOI: 10.7150/ijms.44804
- [66] Ghanbari-Movahed M, Jackson G, Farzaei MH, Bishayee A. A systematic review of the preventive and therapeutic effects of naringin against human malignancies. *Frontiers in Pharmacology*. 2021;**12**:639840. DOI: 10.3389/fphar.2021.639840
- [67] Bak Y, Kim H, Kang JW, Lee DH, Kim MS, Park YS, et al. A synthetic naringenin derivative, 5-hydroxy-7,4'-diacetyloxyflavanone-N-phenyl hydrazone (N101-43), induces apoptosis through up-regulation of Fas/FasL expression and inhibition of PI3K/Akt signaling pathways in non-small-cell lung cancer cells. *Journal of Agricultural and Food Chemistry*. 2011;**59**:10286-10297. DOI: 10.1021/jf2017594
- [68] Lu WL, Yu CR, Lien HM, Sheu GT, Cherng SH. Cytotoxicity of naringenin induces Bax-mediated mitochondrial apoptosis in human lung adenocarcinoma A549 cells. *Environmental Toxicology*. 2020;**35**(12):1386-1394. DOI: 10.1002/tox.23003
- [69] Baruah TJ, Hauneihkim K, Kma L. Naringenin sensitizes lung cancer NCI-H23 cells to radiation by downregulation of akt expression and

metastasis while promoting apoptosis. Pharmacognosy Magazine. 2020;**16**:229-235. DOI: 10.4103/pm.pm_535_19

[70] Shi X, Luo X, Chen T, Guo W, Liang C, Tang S, et al. Naringenin inhibits migration, invasion, induces apoptosis in human lung cancer cells and arrests tumour progression in vitro. Journal of Cellular and Molecular Medicine. 2021;**25**(5):2563-2571. DOI: 10.1111/jcmm.16226

[71] Lee YJ, Wu TD. Total synthesis of kaempferol and methylated kaempferol derivatives. Journal of the Chinese Chemical Society. 2001;**48**:201-206. DOI: 10.1002/jccs.200100033

[72] Singab ANB, Ayoub IM, El-Shazly M, Korinek M, Wu TY, Cheng YB, et al. Shedding the light on Iridaceae: Ethnobotany, phytochemistry and biological activity. Industrial Crops and Products. 2016;**92**:308-335. DOI: 10.1016/j.indcrop.2016.07.040

[73] Alam W, Khan H, Shah MA, Cauli O, Saso L. Kaempferol as a dietary anti-inflammatory agent: Current therapeutic standing. Molecules. 2020;**25**(18):4073. DOI: 10.3390/molecules25184073

[74] Dabeek WM, Marra MV. Dietary quercetin and kaempferol: Bioavailability and potential cardiovascular-related bioactivity in humans. Nutrients. 2019;**11**:2288. DOI: 10.3390/nu11102288

[75] Labib S, Hummel S, Richling E, Humpf HU, Schreier P. Use of the pig caecum model to mimic the human intestinal metabolism of hispidulin and related compounds. Molecular Nutrition & Food Research. 2006;**50**:78-86. DOI: 10.1002/mnfr.200500144

[76] Imran M, Rauf A, Shah ZA, Saeed F, Imran A, Arshad MU, et al. Chemo-preventive and therapeutic effect of

the dietary flavonoid kaempferol: A comprehensive review. Phytotherapy Research. 2019;**33**:263-275. DOI: 10.1002/ptr.6227

[77] Imran M, Salehi B, Sharifi-Rad J, Gondal TA, Saeed F, Imran A, et al. Kaempferol: A key emphasis to its anticancer potential. Molecules. 2019;**24**:2277. DOI: 10.3390/molecules24122277

[78] Luo H, Daddysman MK, Rankin GO, Jiang BH, Chen YC. Kaempferol enhances cisplatin's effect on ovarian cancer cells through promoting apoptosis caused by down regulation of cMyc. Cancer Cell International. 2010;**10**:16. DOI: 10.1186/1475-2867-10-16

[79] Fouzder C, Mukhuty A, Kundu R. Kaempferol inhibits Nrf2 signalling pathway via downregulation of Nrf2 mRNA and induces apoptosis in NSCLC cells. Archives of Biochemistry and Biophysics. 2021;**697**:108700. DOI: 10.1016/j.abb.2020.108700

[80] Jo E, Park SJ, Choi YS, Jeon WK, Kim BC. Kaempferol suppresses transforming growth factor- β 1-induced epithelial-to-mesenchymal transition and migration of A549 lung cancer cells by inhibiting Akt1-mediated phosphorylation of Smad3 at threonine-179. Neoplasia. 2015;**17**:525-537. DOI: 10.1016/j.neo.2015.06.004

[81] Donald G, Hertzner K, Eibl G. Baicalein: An intriguing therapeutic phytochemical in pancreatic cancer. Current Drug Targets. 2012;**13**:1772-1776. DOI: 10.2174/138945012804545470

[82] Yin B, Li W, Qin H, Yun J, Sun X. The use of Chinese Skullcap (*Scutellaria baicalensis*) and its extracts for sustainable animal production. Animals (Basel). 2021;**11**:1039. DOI: 10.3390/ani11041039

- [83] Zhao Q, Zhang Y, Wang G, Hill L, Weng JK, Chen XY, et al. A specialized flavone biosynthetic pathway has evolved in the medicinal plant, *Scutellaria baicalensis*. *Science Advances*. 2016;**2**:e1501780. DOI: 10.1126/sciadv.1501780
- [84] Yuan Y, Wu C, Liu Y, Yang J, Huang L. The *Scutellaria baicalensis* R2R3-MYB transcription factors modulates flavonoid biosynthesis by regulating GA metabolism in transgenic tobacco plants. *PLoS One*. 2013;**8**:e77275. DOI: 10.1371/journal.pone.0077275
- [85] Zhang L, Lin G, Chang Q, Zuo Z. Role of intestinal first-pass metabolism of baicalein in its absorption process. *Pharmaceutical Research*. 2005;**22**: 1050-1058. DOI: 10.1007/s11095-005-5303-7
- [86] Li L, Gao H, Lou K, Luo H, Hao S, Yuan J, et al. Safety, tolerability, and pharmacokinetics of oral baicalein tablets in healthy Chinese subjects: A single-center, randomized, double-blind, placebo-controlled multiple-ascending-dose study. *Clinical and Translational Science*. 2021;**14**:2017-2024. DOI: 10.1111/cts.13063
- [87] Sahu BD, Mahesh Kumar J, Sistla R. Baicalein, a bioflavonoid, prevents cisplatin-induced acute kidney injury by up-regulating antioxidant defenses and down-regulating the MAPKs and NF- κ B pathways. *PLoS One*. 2015;**10**:e0134139. DOI: 10.1371/journal.pone.0134139
- [88] Liang S, Deng X, Lei L, Zheng Y, Ai J, Chen L, et al. The comparative study of the therapeutic effects and mechanism of baicalin, baicalein, and their combination on ulcerative colitis rat. *Frontiers in Pharmacology*. 2019;**10**:1466. DOI: 10.3389/fphar.2019.01466
- [89] Zandi K, Teoh BT, Sam SS, Wong PF, Mustafa MR, Abubakar S. Novel antiviral activity of baicalein against dengue virus. *BMC Complementary and Alternative Medicine*. 2012;**12**:214. DOI: 10.1186/1472-6882-12-214
- [90] Gao Y, Snyder SA, Smith JN, Chen YC. Anticancer properties of baicalein: A review. *Medicinal Chemistry Research*. 2016;**25**:1515-1523. DOI: 10.1007/s00044-016-1607-x
- [91] Liu H, Dong Y, Gao Y, Du Z, Wang Y, Cheng P, et al. The fascinating effects of baicalein on cancer: A review. *International Journal of Molecular Sciences*. 2016;**17**(10):1681. DOI: 10.3390/ijms17101681
- [92] Lu L, Zhang M, Wang X, Zhang Y, Chai Z, Ying M, et al. Baicalein enhances the antitumor efficacy of docetaxel on nonsmall cell lung cancer in a β -catenin-dependent manner. *Phytotherapy Research*. 2020;**34**(1):104-117. DOI: 10.1002/ptr.6501
- [93] Zhang X, Ruan Q, Zhai Y, Lu D, Li C, Fu Y, et al. Baicalein inhibits non-small-cell lung cancer invasion and metastasis by reducing ezrin tension in inflammation microenvironment. *Cancer Science*. 2020;**111**:3802-3812. DOI: 10.1111/cas.14577
- [94] Seelinger G, Merfort I, Wölfl U, Schempp CM. Anti-carcinogenic effects of the flavonoid luteolin. *Molecules*. 2008;**13**(10):2628-2651. DOI: 10.3390/molecules13102628
- [95] Wu T, He M, Zang X, et al. A structure-activity relationship study of flavonoids as inhibitors of *E. coli* by membrane interaction effect. *Biochimica et Biophysica Acta* 2013;**1828**:2751-2756. <https://doi.org/10.1016/j.bbamem.2013.07.029>.

- [96] Lee ER, Kang YJ, Kim HJ, et al. Regulation of apoptosis by modified naringenin derivatives in human colorectal carcinoma RKO cells. *Journal of Cellular Biochemistry* 2008;**104**:259-273. <https://doi.org/10.1002/jcb.21622>.
- [97] Li K, Fan H, Yin P, Yang L, Xue Q, Li X, Sun L, Liu Y. Structure-activity relationship of eight high content flavonoids analyzed with a preliminary assign-score method and their contribution to antioxidant ability of flavonoids-rich extract from *Scutellaria baicalensis* shoots. *Arabian Journal of Chemistry* 2018;**11**:159-170. <https://doi.org/10.1016/j.arabjc.2017.08.002>.
- [98] Kim J, Koo BK, Knoblich JA. Human organoids: Model systems for human biology and medicine. *Nature Reviews. Molecular Cell Biology* 2020;**21**:571-584. <https://doi.org/10.1038/s41580-020-0259-3>.
- [99] Kellar A, Egan C, Morris D. Preclinical murine models for lung cancer: Clinical trial applications. *BioMed Research International* 2015;**2015**:621324. <https://doi.org/10.1155/2015/621324>.
- [100] Garrouste-Orgeas M, Philippart F, Bruel C, Max A, Lau N, Misset B. Overview of medical errors and adverse events. *Annals of Intensive Care*. 2012;**2**:2. DOI: 10.1186/2110-5820-2-2
- [101] Wangari-Talbot J, Hopper-Borge E. Drug resistance mechanisms in non-small cell lung carcinoma. *Journal of Cancer Research Updates*. 2013;**2**:265-282. DOI: 10.6000/1929-2279.2013.02.04.5
- [102] Alao JP. The regulation of cyclin D1 degradation: Roles in cancer development and the potential for therapeutic invention. *Molecular Cancer*. 2007;**6**:24. DOI: 10.1186/1476-4598-6-24

Chapter

Na_v1.7 Sodium Channel: A Potential Analgesic Target

Mildred López-Vázquez

Abstract

Chronic pain is one of the most complex clinical conditions to manage, it can affect approximately 20% of the population and has a high economic cost. In recent years, the voltage-gated sodium channel Na_v1.7 has been proposed as a therapeutic target against pain due to the evidence that certain mutations in the gene that encodes it can cause chronic pain or the absence of this dependent nature. Several molecules have been discovered - both synthetic and natural - that seek to selectively inhibit Na_v1.7 in the search of an effective treatment without side effects against chronic pain.

Keywords: sodium channel, Na_v1.7, pain, analgesic

1. Introduction

Pain is an essential sense that has evolved in complex organisms to minimize cell and tissue damage and thus increase survival. The onset of pain results in the adoption of behaviors that apart the body from a “dangerous environment” and allow tissue repair. Pain also protects us from our environment, by teaching us what situations and behaviors are likely to lead to injury. Pain pathways operate at numerous levels in the nervous system and are under voluntary and involuntary control. The neural process through which organisms encode noxious stimuli is called nociception, and is mediated by the activation of sensory terminals called nociceptors. Blocking these processes with analgesics has been a major pharmaceutical achievement [1].

Humans have nine subtypes of voltage-gated sodium channels (Na_v), denoted from Na_v 1.1 to Na_v 1.9 and which are expressed in various organs; Na_v 1.1, Na_v 1.2, Na_v 1.3 and Na_v 1.6 channels predominate in the central nervous system, while Na_v1.7, Na_v 1.8 and Na_v 1.9 are mainly expressed in the peripheral nervous system, Na_v 1.4 channels are expressed in skeletal muscle and Na_v 1.5 in the heart. Of all of them, the Na_v1.7, Na_v 1.8 and Na_v 1.9 channels play an important role in pain signaling [2, 3].

The Na_v1.7 channel has been shown to play a crucial role in pain sensation and there is interesting genetic evidence linking Na_v1.7 and the gene encoding it SCN9A to pain disorders in humans [4].

2. Na_v1.7

Voltage gated sodium channels are transmembrane proteins that are formed by an alpha (α) subunit composed of approximately 2000 amino acid residues, this subunit

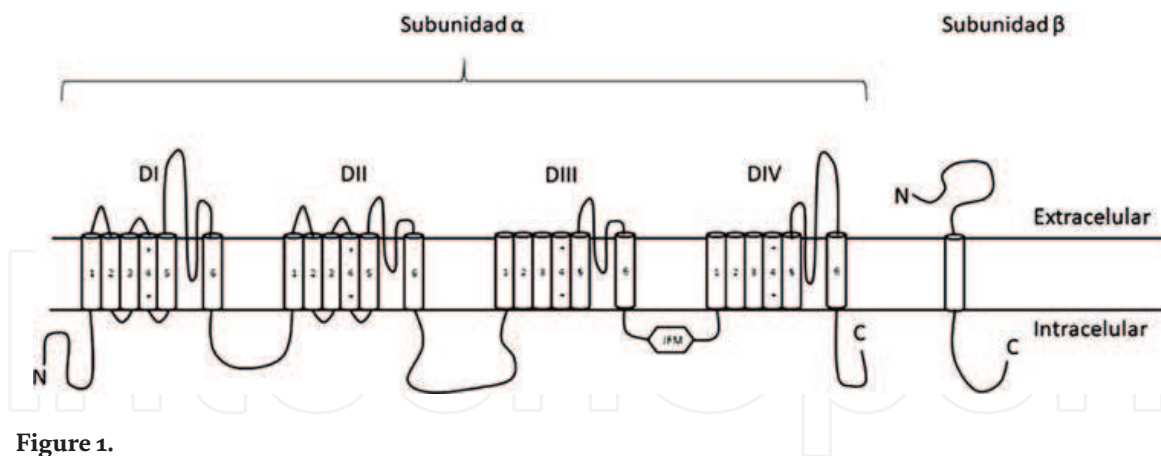


Figure 1.
Voltage-gated sodium channel structure.

is organized in four homologous domains that make up the Na^+ channel pore. Each domain is made up of 6 transmembrane segments; 1 to 4 form the voltage sensing domains (VSDs) and segments 5 and 6 of each domain form the central module of the pore. The β -subunits of sodium channels are composed of an N-terminal extracellular immunoglobulin-like fold, a single transmembrane segment, and a short intracellular segment as seen in **Figure 1** [5].

All sodium channel α -subunits consist of 4 homologous domains that form a single, voltage-gated aqueous pore. The α -subunits are greater than 75% identical over the amino acid sequences comprising the transmembrane and extracellular domains. The α -subunits show distinct patterns of expression, and are associated with accessory β -subunits which modify channel properties and interact with cytoskeletal and extracellular matrix proteins. Despite the broadly similar properties of voltage-gated sodium channels, there is good evidence for a specialised functional role of the various isoforms. Voltage-gated sodium channels provide the inward current that generates the upswing of an action potential in response to supra-threshold depolarisations of the membrane potential. At present 9 homologous channel-forming α -subunits ($\text{Na}_v1.1$ to $\text{Na}_v1.9$), and four accessory β subunits ($\beta1$ to $\beta4$) have been cloned from a variety of tissues and classified as belonging to the voltage-gated sodium channel family according to their amino acid sequence. Although the α -subunit alone is sufficient for the formation of a functional channel, the accessory β -subunits increase the efficiency of channel expression and are required for normal kinetics and voltage dependence of channel gating. In addition, β -subunits have an important role in the localization of α -subunits. The $\beta4$ -subunit has even been shown to play a role in gating channels through a positively charged C-terminal region [6].

Mutations in the nine Na_v channel subtypes ($\text{Na}_v1.1$ – $\text{Na}_v1.9$) are associated with migraine, epilepsy, pain, and cardiac and muscle paralysis syndromes.

Specifically, the $\text{Na}_v1.7$ channel is highly expressed in the soma, axons and peripheral endings of the nociceptive neurons of the spinal and trigeminal ganglia, in the olfactory neurons and in the neurons of the sympathetic ganglia. And to a lesser degree in the nervous system, liver, heart muscle and spinal cord.

$\text{Na}_v1.7$ is designated as a threshold channel and is expressed on the surface of peripheral pain-sensitive neurons or nociceptors (in about 85%), where it conducts Na^+ currents in response to membrane depolarizations that are generated by potentially damaging events to the tissues, activating the action potential and sending pain signals. It can be said that it acts as a volume knob that sets the level of pain gain in the

neurons. The Na_v1.7 channel could also be seen as an amplifier of pain signals generated by receptors in nociceptor terminals.

Both gain and loss-of-function mutations in the SCN9A gene have been reported to result in a variety of conditions. When there is a gain of function, there is evidence of painful conditions such as hereditary erythromelalgia, paroxysmal extreme pain disorder, and idiopathic small fiber neuropathies. In contrast, loss-of-function mutations in the SCN9A gene have been reported to cause a rare disorder called congenital insensitivity to pain, characterized by the complete loss of the ability to feel pain stimuli.

The important role of Na_v 1.7 in pain generation has created immense interest in this channel as an analgesic target.

2.1 A little history

The first hint that Na_v1.7 might have an important role in pain signaling was given by a research group in 2004, who showed that three-generation patients from a Chinese family with a persistent pain syndrome called primary erythromelgia had missense mutations in the SCN9A gene, which codes for Na_v1.7. Primary erythromelgia is a rare disease characterized by intermittent burning pain with redness and heat in the feet and hands in response to moderate exercise. In patients, standing, exercise, or local exposure to heat can induce the symptoms, and keeping the involved extremities at an icy cold temperature could be the only way to relieve the pain. The mutations in SCN9A that the group identified are located in the II/S5 segment (L858H) and the loop region between II/S4 and II/S5 (I848T) of Na_v1.7 [7].

The next year, Dib-Hajj et al. discovered a third mutation in Na_v1.7 in a large family with primary erythromelgia, they described a single substitution of phenylalanine by valine (F1449V) at codon 1449 in the sodium channel Na_v1.7. This single amino acid substitution alters the biophysical properties of Na_v1.7 and reduces the threshold for action potential firing and bursting of dorsal root ganglion neurons. This is what causes gain-of-function alterations in the channel, which causes hypersensitivity to pain [8].

Around the same time, researchers at University College London examined the role of Na_v1.7 in pain pathways using knockout mice, which were found to be insensitive to inflammatory pain, showing that loss-of-function mutations abolish pain perception [9].

The first case of a patient with a congenital inability to perceive pain was said to have been reported in the early twentieth century. Only very few patients have since been described and are categorized as having a congenital indifference to pain. In 2006 researchers at the University of Cambridge described individuals from three families with the extraordinary phenotype of a congenital inability to perceive any form of pain, in whom all other sensory modalities were preserved and the peripheral and central nervous systems were apparently otherwise intact. The index case for the study was a ten-year-old child, who acted in a “street theater” putting knives through his arms and later walking on hot coals without feeling any pain, reasons why he was well known to the medical service. He died at the age of 14, after jumping off a house roof, before being able to be part of the study. Subsequently, they studied three further consanguineous families in which there were individuals with similar histories of a lack of pain appreciation. The study included 6 children (between 6 and 12 years old), who had never felt any pain in any part of their body. Even as babies they had shown no evidence of pain appreciation. None knew what pain felt like, although the older individuals realized what actions should elicit pain (including acting as if in pain after football tackles). All had injuries to their lips (some requiring later plastic

surgery) and/or tongue (with loss of the distal third in two cases), caused by biting themselves in the first 4 years of life, fractures and even osteomyelitis; despite this, all children were found to have normal vision, hearing, and appearance. Detailed neurological examinations revealed that each could correctly perceive the sensations of touch, warm and cold temperature, proprioception, tickle and pressure, but not painful stimuli. Genome sequencing of these individuals showed disruption of one gene, SCN9A, causing loss-of-function mutations in Na_v1.7 and complete loss of nociceptive input [1].

It is important to note that apart from the inability to feel pain, anosmia (loss of smell) is the only other sensory impairment in individuals with this channelopathy. Weiss et al examined human patients with loss-of-function mutations in SCN9A and show that they were unable to sense odours. To establish the essential role of Na_v1.7 in odour perception, they generated conditional null mice in which Na_v1.7 was removed from all olfactory sensory neurons. In the absence of Na_v1.7, these neurons still produce odour-evoked action potentials but fail to initiate synaptic signalling from their axon terminals at the first synapse in the olfactory system. The mutant mice no longer display vital, odour-guided behaviours such as innate odour recognition and avoidance, short-term odour learning, and maternal pup retrieval. The no significant side effects in people lacking Na_v1.7 such as cognitive, motor or non-nociceptive sensory impairments other than anosmia give support to the concept of Na_v1.7 antagonists as analgesics [10].

These studies opened the way to propose and study the Na_v1.7 channel as a therapeutic target to develop drugs that could serve as a new and potentially safer analgesic by selectively blocking it.

2.2 Why does the Na_v1.7 channel represent a therapeutic target against pain?

If the different subtypes of sodium channels are related to particular mechanisms of pain, then specific antagonists of those subtypes could, in theory, produce pain treatment without side effects.

Specifically, there are three binding sites that appear to offer the highest potential for the discovery and optimization of selective Na_v1.7 inhibitors: 1) the extracellular vestibule of the pore, tetrodotoxin (TTX) and saxitoxin (STX) binding sites; 2) the extracellular loops of voltage-sensing domain II (VSD2) AND 3) the extracellular loops of voltage-sensing domain IV (VSD4).

3. Channel inhibitory molecules

Findings about Na_v1.7 gain-of-function mutations causing intense pain and loss-of-function mutations of the same channel causing painlessness have put Na_v1.7 in the therapeutic spotlight as an important new target for development of analgesics, for which a channel blocking antagonist is required.

Despite the peculiar and promising nature of the Na_v1.7 channel, the use of ion channel blockers for the treatment of pain is not new. For years, blockers of sodium (Na⁺), calcium (Ca⁺⁺) and potassium (K⁺) channels have been used, such as local anesthetics, antiarrhythmics or antiepileptics, as adjuvants in the treatment of neuropathic pain. The problem lies in the poor selectivity of these drugs, in the adverse cardiac, neurological, hematological, and digestive effects that limit their use, and in the inter-individual variability that, together with the particular susceptibility to develop chronic pain, limits their efficacy; therefore, the challenge facing Na_v1.7

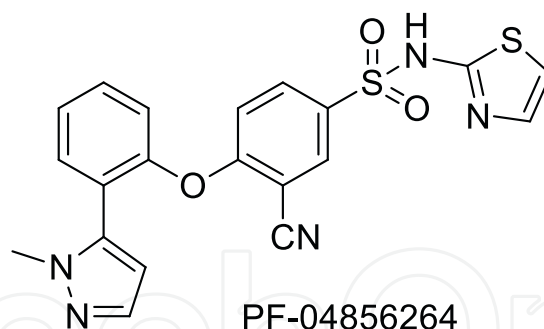


Figure 2.
Nav1.7 inhibitory arylsulfonamide.

inhibitory molecules is that they must be extremely selective, for which the three binding sites described above have been proposed. Some of the compounds studied so far and proposed as selective inhibitors of Nav1.7 are described below.

3.1 Small molecules as sodium channel blockers

In recent years, small molecules have been developed that seek to selectively inhibit domain IV of the voltage sensor, arylsulfonamides are some of them. In 2012, it was shown that the arylsulfonamide named PF-04856264 and shown in **Figure 2** interacts with amino acid residues on an extracellular facing region of the homologous Domain 4 voltage sensor of Nav1.7, selectively inhibiting it (IC₅₀, 28 nM) over Nav1.5; this is important due block of the Nav1.5 channel may lead to arrhythmia and thus limit the therapeutic potential of nonselective Nav1.7 inhibitors [11].

In 2016, Focken et al. reported a series of aryl sulfonamides that act as nanomolar potent, isoform-selective inhibitors of the human sodium channel hNav1.7. In this study, the optimization of these inhibitors is described. The main goal of this research was to improve potency against hNav1.7 while minimizing off-target safety concerns, so, they generated the compound of the **Figure 3**. This agent displayed significant effects in rodent models of acute and inflammatory pain and demonstrated that binding to the voltage sensor domain 4 site of Nav1.7 leads to an analgesic effect *in vivo* and corroborate the importance of hNav1.7 as a drug target for the treatment of pain [4].

Vixotrigine (**Figure 4**) is another Nav1.7 channel blocker under clinical investigation to treat peripheral neuropathic pain conditions, including trigeminal neuralgia, which is characterized by brief episodes of severe pain from one or more branches of the trigeminal nerve. Clinical trial results suggest that vixotrigine 150 mg three

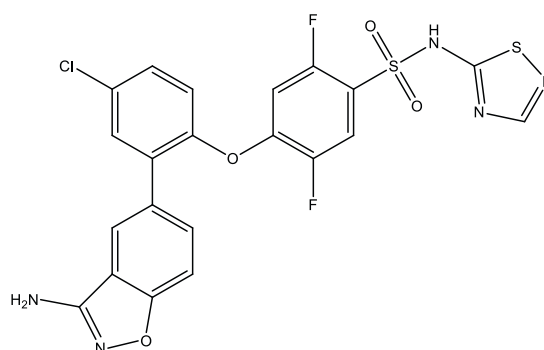


Figure 3.
Sulfonamide that acts as an inhibitor of the human sodium channel hNav1.7.

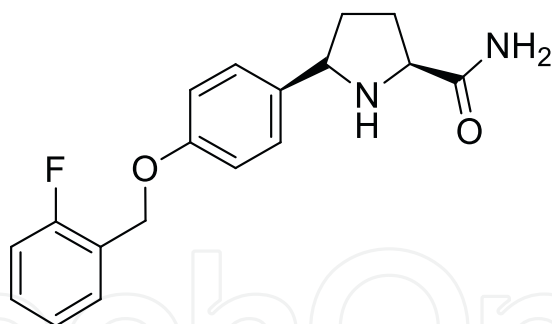


Figure 4.
Vixotrigine is a voltage-dependent Nav blocker.

times a day may be an effective and safe treatment for trigeminal neuralgia pain with only dizziness and headache as side effects, and even a phase III study suggests that a higher dose (250 mg three times a day) may provide additional benefit in those who do not respond adequately to the first dose [12, 13].

In 2017 Swain et al. described a series of acidic diaryl ether heterocyclic sulfonamides that are potent and sub-type selective $\text{Na}_v1.7$ inhibitors. Their design strategies and results from pre-clinical PK and Clinical human microdose PK data are described leading to the discovery of the first sub-type selective $\text{Na}_v1.7$ inhibitor clinical candidate: 4-[2-(5-Amino-1H-pyrazol-4-yl)-4-chlorophenoxy]-5-chloro-2-fluoro-N-1,3-thiazol-4-ylbenzenesulfonamide (PF-05089771) which binds to a site in the voltage sensing domain. (**Figure 5**). This was the first potent and selective molecule which binds to the domain IV voltage sensor region of the sodium channel to progress into the clinic. The authors demonstrated that the development of selective agents such as PF-05089771 offers new potential therapeutic modelities for acute and chronic pain [14].

Recently Wang et al. designed and synthesized a series of compounds with ethanoanthracene and aryl sulfonamide moieties. They detected the inhibitory activity of this compounds on sodium channels with electrophysiological techniques. They found that a compound of theirs potently inhibited $\text{Na}_v1.7$ channels stably expressed in HEK293 cells ($\text{IC}_{50} = 0.64 \pm 0.30 \text{ nmol/L}$) and displayed a high $\text{Na}_v1.7/\text{Na}_v1.5$ selectivity [15].

3.2 Toxins as sodium channel blockers

On the other hand, studies of different toxins have been carried out for therapeutic benefit, finding that they bind to DSV2, a different domain than small molecules.

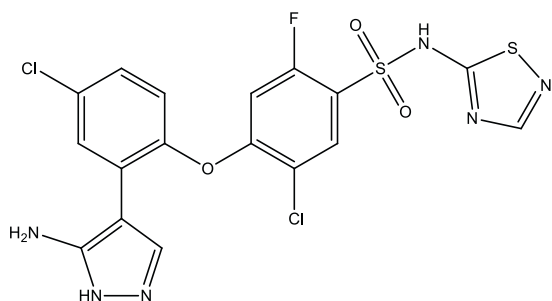


Figure 5.
PF-05089771.

For example, ProTx-II, a peptide from spider venom was the first inhibitor of Na_v1.7 reported to have greater than 50-fold selectivity over the other Nav isoforms.

In 2010, Xiao et al. discovered that ProTx-II and huwentoxin-IV (HWTX-IV), cystine knot peptides from tarantula venoms, preferentially block hNa_v1.7. This group described that the mutation E818C increases ProTx-II's and HWTX-IV's IC₅₀ for block of hNa_v1.7 currents by 4- and 400-fold, respectively. Also, they reported that ProTx-II, but not HWTX-IV, preferentially interacts with hNa_v 1.7 to impede fast inactivation by trapping the domain IV voltage-sensor in the resting configuration [16].

Knowing that Protoxin-II (ProTx2) from the Peruvian green velvet tarantula (**Figure 6**) is an inhibitor cystine-knot peptide and selective antagonist of the human Na_v1.7 channel, in 2019, Xu et al. visualize ProTx2 in complex with voltage-sensor domain II (VSD2) from Na_v1.7 using X-ray crystallography and cryoelectron microscopy. They described that membrane partitioning orients ProTx2 for unfettered access to VSD2, where ProTx2 interrogates distinct features of the Na_v1.7 receptor site. ProTx2 positions two basic residues into the extracellular vestibule to antagonize S4 gating-charge movement through an electrostatic mechanism. ProTx2 has trapped activated and deactivated states of VSD2, revealing a remarkable ~10 Å translation of the S4 helix, providing a structural framework for activation gating in voltage-gated ion channels. This work delivered key templates to design selective Nav channel antagonists. Here, the researchers verified that ProTx2 shows 30–100 times more selectivity on the Na_v1.7 channel than on the other Nav isoforms. The authors also demonstrated that this toxin is the most potent inhibitor available, testing its structural bases in the quest to accelerate the design of new modulators [17].

Finally, in 2020, Pajouhesh et al. described the discovery and characterization of ST-2262, a Na_v1.7 inhibitor that blocks the extracellular vestibule of the channel with an IC₅₀ of 72 nM and greater than 200-fold selectivity over off-target sodium channel isoforms, Na_v1.1–1.6 and Na_v1.8. ST-2262 (**Figure 7**) was discovered through a rational design strategy aimed at identifying derivatives of natural bis-guanidinium toxins that preferentially inhibit hNa_v1.7 over other off-target hNa_v isoforms. The potency of ST-2262 against hNa_v1.7 stably expressed in HEK293 cells was assessed by manual patch clamp electrophysiology with a voltage-protocol that favors the resting state of the channel. Using a stimulation protocol with a holding potential of –110 mV and a stimulus frequency of 0.33 Hz, the IC₅₀ of ST-2262 against hNa_v1.7 was measured



Figure 6.
Peruvian Green Velvet Tarantula.

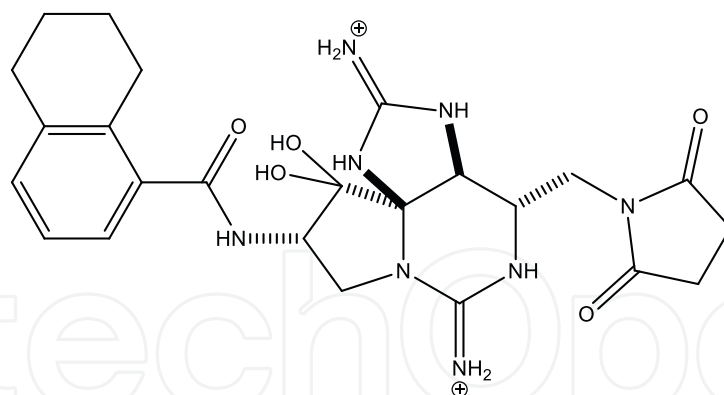


Figure 7.
ST-2262 is a potent and selective inhibitor of $hNa_v1.7$ and it is a derivative of natural bis-guanidinium toxins.

at $0.072 \mu\text{M}$ (95% confidence interval (CI) $0.064\text{--}0.082$). The authors find that, in contrast to other $Na_v1.7$ inhibitors that preferentially inhibit the inactivated state of the channel, ST-2262 is equipotent in a protocol that favors the resting state of the channel, a protocol that favors the inactivated state, and a high frequency protocol. In a non-human primate study, animals treated with ST-2262 exhibited reduced sensitivity to noxious heat. These findings establish the extracellular vestibule of the sodium channel as a viable receptor site for the design of selective ligands targeting $Na_v1.7$ [18].

4. Conclusions

The voltage-gated sodium channel $Na_v1.7$ has been shown to play an important role in human pain signaling through neuronal excitability. Studies of different rare conditions point to the SCN9A gene, which codes for said channel as the one responsible for its gain or loss of function, which has demonstrated its intrinsic relationship with pain, this has given rise to the Na_v channel 1.7 is an interesting therapeutic target to develop analgesics, despite the fact that a large catalog of channel antagonists have been discovered, there is still no certainty for their medical prescription, however, there is extensive and promising research in clinical trials to exploit the characteristics of $Na_v1.7$ and there are, in the future, specific drugs that relieve patients of severe pain, thus improving their quality of life.

Acknowledgements

The author thanks Universidad de las Americas Puebla and CONACyT for the PhD fellowship.

Conflict of interest

The author confirm that this article content has no conflicts of interest.

IntechOpen


IntechOpen

Author details

Mildred López-Vázquez
Departamento de Ciencias Químico-Biológicas, Universidad de las Américas Puebla,
San Andrés Cholula, Pue, México

*Address all correspondence to: mildred.lopezvz@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Cox J, Reimann F, Nicholas A, Thornton G, Roberts E, Springell K, et al. An SCN9A channelopathy causes congenital inability to experience pain. *Nature*. 2006;**444**:894-898. DOI: 10.1038/nature05413
- [2] King GF, Vetter I, Gain N. No pain: Nav1.7 as an analgesic target. *ACS Chemical Neuroscience*. 2014;**5**(9):749-751
- [3] Pérez CN, Martínez TM, Díaz Mato I. Canalopatías, un novedoso enfoque en la fisiopatología del dolor. *Revista Cubana de Pediatría*. 2017;**89**(3):381-394
- [4] Focken T, Liu S, Chahal N, Dauphinais M, Grimwood ME, Chowdhury S, et al. Discovery of aryl sulfonamides as isoform-selective inhibitors of Nav1.7 with efficacy in rodent pain models. *ACS Medical Chemistry Letters*. 2016;**7**(3):277-282
- [5] Catterall WA. Forty years of sodium channels: Structure, function, pharmacology, and epilepsy. *Neurochemical Research*. 2017;**42**(9):2495-2504. DOI: 10.1007/s11064-017-2314-9
- [6] Wood JN, James Boorman BSP. Voltage-gated sodium channel blockers; Target validation and therapeutic potential. *Current Topics in Medicinal Chemistry*. 2005;**5**(6):529-537. DOI: 10.2174/1568026054367584
- [7] Yang Y, Wang Y, Li S, Xu Z, Li H, Ma L, et al. Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythromelgia. *Journal of Medical Genetics*. 2004;**41**:171-174. DOI: 10.1136/jmg.2003.012153
- [8] Dib-Hajj SD, Rush AM, Cummins TR, Hisama FM, Novella S, Tyrrell L, et al. Gain-of-function mutation in Nav1.7 in familial erythromelgia induces bursting of sensory neurons. *Brain*. 2005;**128**(8):1847-1854
- [9] Nassar M, Stirling LC, Forlani G, Baker MD, Matthews EA, Dickenson AH, et al. Nociceptor-specific gene deletion reveals a major role for Nav1.7 (PN1) in acute and inflammatory pain. *Proceedings of the National Academy of Sciences*. 2004;**101**(34):12706-12711
- [10] Weiss J, Pyrski M, Jacobi E, Bufer B, Willnecker V, Schick B, et al. Loss-of-function mutations in sodium channel Nav1.7 cause anosmia. *Nature*. 2011;**472**(7342):186-190
- [11] McCormack K, Santos S, Chapman ML, Krafte DS, Marron BE, Wes CW, et al. Voltage sensor interaction site for selective small molecule inhibitors of voltage-gated sodium channels. *Proceedings of the National Academy of Sciences*. 2013;**110**(29):E2724-E2732
- [12] Zakrzewska JM, Palmer J, Morisset V, for the Study Investigators, et al. Safety and efficacy of a Nav1.7 selective sodium channel blocker in patients with trigeminal neuralgia: A double-blind, placebo-controlled, randomised withdrawal phase 2a trial. *Lancet Neurology*. 2017;**16**(4):291-300
- [13] Kotecha M, Cheshire WP, Finnigan H, Giblin K, Naik H, Palmer J, et al. Design of phase 3 studies evaluating vixotrigine for treatment of trigeminal neuralgia. *Journal of Pain Research*. 2020;**13**:1601-1609
- [14] Swain NA, Batchelor D, Beaudoin S, Bechle BM, Bradley PA, Brown AD,

et al. Discovery of clinical candidate 4-[2-(5-Amino-1H-pyrazol-4-yl)-4-chlorophenoxy]-5-chloro-2-fluoro-N-1,3-thiazol-4-ylbenzenesulfonamide (PF-05089771): Design and optimization of diaryl ether aryl sulfonamides as selective inhibitors of Nav1.7. *Journal of Medicinal Chemistry*. 2017;**60**(16):7029-7042

[15] Wang JT, Zheng YM, Chen YT, Gu M, Gao ZB, Nan FJ. Discovery of aryl sulfonamide-selective Nav1.7 inhibitors with a highly hydrophobic ethanoanthracene core. *Acta Pharmacology Sinica*. 2020;**41**(3):293-302

[16] Xiao Y, Blumenthal K, Jackson JO, Liang S, Cummins TR. The Tarantula Toxins ProTx-II and Huwentoxin-IV differentially interact with human Nav1.7 Voltage sensors to inhibit channel activation and inactivation. *Molecular Pharmacology*. 2010;**78**(6):1124-1134. DOI: 10.1124/mol.110.066332

[17] Xu H, Li T, Rohou A, Arthur CP, Tzakoniati F, Wong E, et al. Structural basis of Nav1.7 inhibition by a gating-modifier spider toxin. *Cell*. 2019;**176**(4):702-715

[18] Pajouhesh H, Beckley JT, Delwig A, et al. Discovery of a selective, state-independent inhibitor of Nav1.7 by modification of guanidinium toxins. *Scientific Reports*. 2020;**10**:14791. DOI: 10.1038/s41598-020-71135-2

Detection and Classification of Burnt Skin on Images with Sparse Representation of Image Patches and Dictionaries

Brenda Rangel-Olvera and Roberto Rosas-Romero

Abstract

Burn injuries are the fourth leading cause of death among accidental deaths, which is the third leading cause of death in humans. Early treatments are required to stop the degeneration of the muscles, which could lead to the complications of the patient, even death. A proper and in time, treatment can lead to a full recovery of the person. Professionals often mislead these early treatments due to the stress and the lack of sleep during their job. Thus, an artificial medical assisting method was developed to detect and classify burns. Skin burns in color images were accurately detected and classified according to burn degree to assist clinicians during diagnosis and early treatment. Detection and classification of burnt areas are performed using the sparse representation of feature vectors with over-redundant dictionaries.

Keywords: burn injuries, sparse representation, over-redundant dictionaries, burn classification and detection

1. Introduction

The human body is covered by skin and this muscle protects all the others. This muscle is essential for the human senses. It not only provides sensation but also helps with the skin regulation and synthesis of vitamin D. Skin can get damage by burn accidents. In a burn injury, some skin layers are destroyed, and first treatments are required as soon as possible. These treatments are based on intensity and severity. Burn area, depth, and location are the determining factors for the gravity. While partial-thickness burns can heal immediately with minimal scarring, deep partial-thickness and full-thickness burns require more than three weeks to close. They are often associated with significant scarring and functional limitations unless removed and grafted within the first days of the injury [1]. Thus, the correct diagnosis of burn depth is necessary to achieve optimal results. It is also known that burn injury may extend to grow over the first days heading to the conversion of superficial burns to deep burns. A better perception of the mechanisms leading to burn wound regeneration is probable to lead to new treatments that can result in the limitation of burn

wound progression, which leads to better healing. Therefore, it is vital to identify the type of skin and the position of the injury [2]. Despite its importance, verified methods to measure wound closure are lacking, making any comparisons of innovative therapies difficult. Wound healing and long-term consequences are determined by burn depth. The distinction between each degree lies in how many layers of skin are damaged [3].

Research interest in systems for automatic diagnosis of skin burns is motivated to assist health care personnel in certain situations. Clinics in small towns may rely on the work of interns or professional apprentices. Additionally, quantitative analysis of important burnt skin features can be used as contextual information for clinical interpretation, for recommendation of treatment. Classification of skin burns is also useful in assisting healing since the treatment depends on the burn degree and size.

Previous work in this area has focused on the classification of burn wounds by applying various artificial vision approaches. Cognitive systems are machine models whose function is inspired by the way that the human brain works. The use of cognitive systems is not intended to replace humans but rather to augment human capabilities through automatic assistance tools [4]. Although there is much research done, there are many opportunities in the area. Segmentation and image processing are fields widely used for artificial vision and medical imaging, especially for burn detection applications [5–9].

Şevik et al. [10] proposed a model to classify burns with texture-based features to identify between burnt skin, healthy skin, and background. The authors combined various algorithms to obtain a model that can find burn wounds with a small error. They also developed different models to compare them all and choose the best results as the most accurate method. The goal was to classify the images into three different objects: skin, burn, and background regions. The best process was by applying the fuzzy c-means algorithm at the stage of the segmentation. In the next step of the algorithm, a multilayer feed-forward artificial neural network trained with the back-propagation algorithm was used for classification. The authors created their database in collaboration with Turkish hospitals. It contains 105 images from various patients. The photos had to be pre-processed to standardize the size of the pictures. The authors used texture patterns from Haralick features from the Gray Level Co-occurrence Matrix (GLCM) for the feature extraction process. From a different perspective, Rehman et al. [11] proposed a computer-aided diagnosis (CAD) for the classification of the burnt skin. Their focus lies at the identification of the depth of the injury. Their objective was to identify the thickness of injured human skin and then be able to classify among first, second, and third-degree burns. The authors used the Otsu method of thresholding for the segmentation step and then applied the statistical method GLCM to obtain the feature vector for the input of the classifier. They used contrast and correlation instead of texture. The author created their database with the help of the Burn Center of Allied Hospital Faisalabad, Pakistan. It contains around 173 images from the different types of burn depth. Yadav et al. [12] suggested a classification model to diagnose burns based on support vector machine (SVM). The training is performed by classifying images into two classes, those that need grafts and those that are non-graft. They skipped the segmentation step and converted the pictures to CIELAB space since the range of colors on that scale is more extensive than RGB. They obtained Hue, Kurtosis, Skewness, and Chroma to determine the area and depth of the wound and, based on the analysis, classify the injury in graft or non-graft with the help of the SVM.

The rest of the manuscript is organized as follows. The following section describes the methodology for feature extraction and the image segmentation and classification methods. The detection and classification results are reported in the Results section. Finally, conclusions are presented in the last section.

2. Methodology

An overview of the proposed method is shown in **Figure 1** [13]. Each image was segmented into two regions: no burns and burns. Burns were classified as first-degree burns, second-degree burns, or third-degree burns. Image segmentation was performed by classifying each image patch. Each image patch (black square in **Figure 1**) was classified by (1) extracting a feature vector (red square in **Figure 1**), and (2) classifying the feature vector based on its closest reconstruction over six dictionaries. The dimensions of the image patch to be classified were smaller than the dimensions of the patch for feature extraction. The extracted features were color and texture. The class assigned to a patch depended on the dictionary with the best reconstruction of the extracted feature vector. The best sparse reconstruction corresponded to the minimal Euclidean distance between the feature vector and its reconstruction.

To obtain color features, three histograms were generated, P_R , P_G , and P_B . One histogram was obtained from each RGB channel. Five statistical attributes were computed for each channel: the *average value*

$\mu = \sum_{I=0}^{255} I P(I)$, *variance* $\sigma^2 = \sum_{I=0}^{255} (I - \mu)^2 P(I)$, *skewness* $m_3 = \sum_{I=0}^{255} (I - \mu)^3 P(I)$, *kurtosis* $m_4 = \sum_{I=0}^{255} (I - \mu)^4 P(I)$, and *entropy* $H = -\sum_{I=0}^{255} P(I) \log [P(I)]$. To conduct texture analysis, a color image was transformed into a grayscale image by using the luminosity model, which consists of a weighted average of the RGB channels: $I = 0.21 R + 0.72 G + 0.07 B$. According to the luminosity model, the green channel is the one that contributes the most, which agrees with the fact that green is the dominant color in skin. Color was extracted from three channels within each patch: green

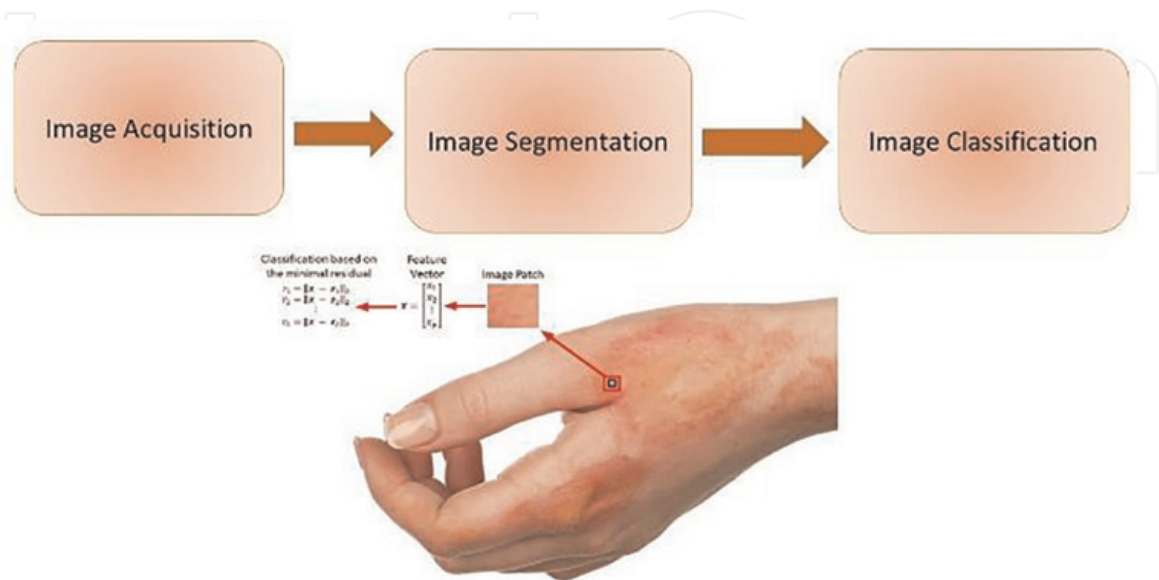


Figure 1.
Skin burn detection and classification model.

channel, blue channel, and a linear combination of the three RGB channels (luminosity model) as it is shown in **Figure 2**. To convert color information to a single grayscale value for texture analysis, the luminosity model was used.

Textural features convey statistical information about the relative positions of the pixel intensity values within the region of interest. *Textural features* are obtained from the gray level co-occurrence matrix (GLCM), where these two concepts were introduced by Robert Haralick [14–16]. The GLCM specifies the distribution of pairs of pixel intensities according to (1) the distance between these two pixels and (2) the angle of the line segment that joins them. There are four possible angles: 0° (horizontal), 45° (diagonal), 90° (vertical), and 135° (anti-diagonal). The GLCM of a pair of pixel intensities (I_m, I_n) at distance d and angle φ is defined as

$$P(I_m, I_n, d, \varphi) = \frac{\text{Number of pairs } (I_m, I_n) \text{ at distance } d \text{ and angle } \varphi}{\text{Total number of pairs}} \quad (1)$$

where $\{P(I_m, I_n, d, \varphi); m = 1, 2, \dots, N; n = 1, 2, \dots, N\}$ is a second-order histogram defined as the probability of occurrence of a pair of pixel intensities in terms of their relative position and angle, and N is the number of pixel intensities. Seven textural features were computed from the GLCM: (1) the angular second moment $ASM = \sum_{m=1}^N \sum_{n=1}^N P(I_m, I_n)^2$, (2) the contrast $CON = \sum_{\ell=1}^N \ell^2 \sum_{m=1}^N \sum_{n=1, |I_m - I_n| < \ell}^N P(I_m, I_n)$, (3) the inverse difference moment $IDF = \sum_{m=1}^N \sum_{n=1}^N \frac{P(I_m, I_n)}{1 + (I_m - I_n)^2}$, (4) the correlation $Corr = \frac{\sum_{m=1}^N \sum_{n=1}^N I_m I_n P(I_m, I_n) - \mu^2}{\sigma^2}$, (5) the co-occurrence matrix variance $Var = \sum_{m=1}^N \sum_{n=1}^N (I_m - \mu)^2 P(I_m, I_n)$, (6) the difference

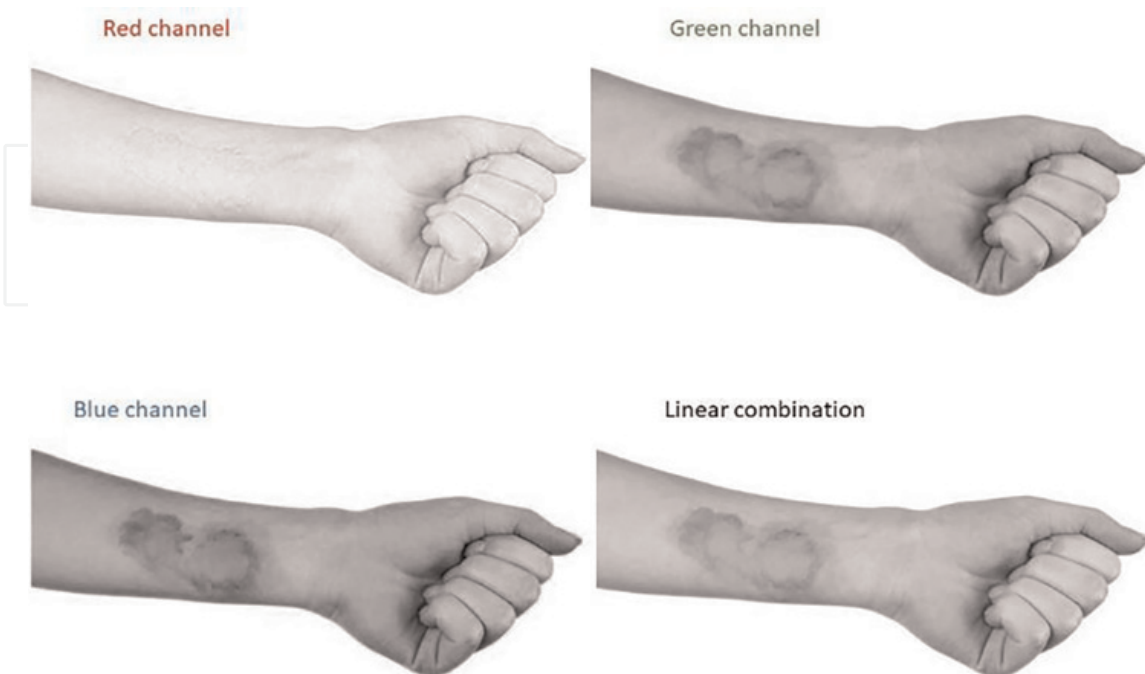


Figure 2.
Image channels and luminosity model.

average $DA = \sum_{m=1}^N \sum_{\ell=1}^N I_m P(I_m, I_n \pm \ell)$, and (7) the co-occurrence matrix entropy $H = \sum_{m=1}^N \sum_{n=1}^N P(I_m, I_n) \log [P(I_m, I_n)]$.

In this work, the GLCM was computed for three different values of angle φ : 0° (horizontal), 45° (diagonal), and 90° (vertical). The GLCM was generated for two pixel intensity values (I_m, I_n) at each of three distance values $d = 1, 2$, and 3 . Thus, nine GLCMs were generated for each image patch. Seven texture-based features were computed from each GLCM. Thus, 45 texture-based features were used to build each feature vector. The total number of entries in an image patch's feature vectors was thus nine color features and 405 texture-based features.

The classification of a feature vector \mathbf{x} extracted from an image patch consisted of (1) sparsely reconstructing the feature vector over a dictionary for each class D_i (healthy skin, first-degree-burnt skin, second-degree-burnt skin, third-degree-burnt skin, shadowed skin, background), followed by (2) determining the dictionary or class that gave the sparse reconstruction s_i closest to the feature vector \mathbf{x} under analysis. The closest sparse reconstruction corresponds to the smallest distance between the feature vector and its corresponding approximation according to the minimization problem $class = \min_i \|\mathbf{x} - s_i\|^2$, where s_i is the sparse reconstruction of \mathbf{x} with dictionary D_i as it is shown in Figure 3.

3. Results

We used color images from various individuals and conditions of luminosity, skin color, background and presence of shadows. These images were obtained from public websites BurnVictims [17], howzak [18], ISCN [19], Healthline [14], MedicalNewsToday [15], Brouhard [16], UrgentCare [20]. The images were studied and tagged with the assistance of an experienced plastic surgeon with experience in

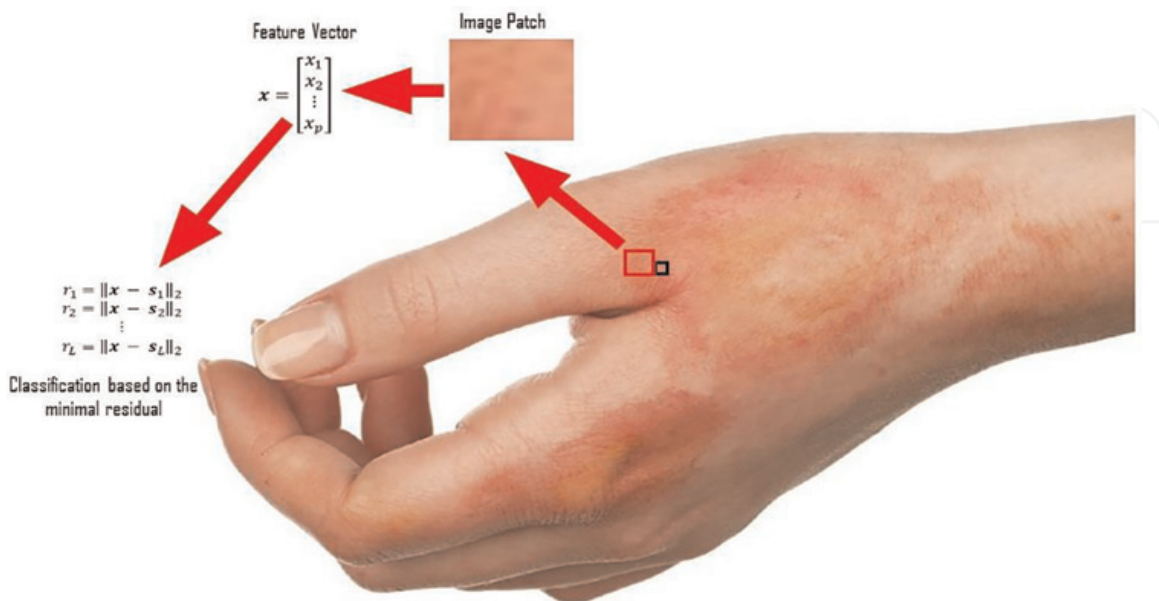


Figure 3.
Classification of image patches.

treating burnt patients. In general, first-degree burns were superficial (epidermal) lesions that are red and dry. Second-degree burns were white or yellow and were mostly wet and involved blisters. Third-degree burns were bloody and leathery lesions revealing subcutaneous structures and were red, brown or black. The collection includes 20 color images that show 33 superficial burns, 20 images with 26 partial thickness burns and 20 images with 56 third-degree burns. These images were characterized by a diversity of backgrounds. **Figure 4** shows four cases of burns of first-, second-, and third- degrees. The image formats were JPG and PNG. The image resolution ranged from $12.13 \frac{\text{pixels}}{\text{in}}$ to $89.55 \frac{\text{pixels}}{\text{in}}$. The variation in image resolution was due to the variation in distance between the camera and the subject. The image dataset's average image resolution was $54.67 \frac{\text{pixels}}{\text{in}}$.

Image segmentation was performed using six dictionaries for each class: *healthy skin*, *first-degree-burnt skin*, *second-degree-burnt skin*, *third-degree-burnt skin*, *shadowed skin*, and *background*. Each dictionary contained 3000 normalized atoms (unit vectors) with $54 \frac{\text{features}}{\text{atoms}}$. Color and texture were used as features for atom extraction and patches were obtained from images at various locations. **Figure 5** shows six sets of patches, each corresponding to a different class. Patches were used for extraction of dictionary atoms. Color and texture played a significant role as features for discrimination. Identification of shadowed skin is not necessary for clinical analysis; however, these dictionaries were helpful to discard these regions; otherwise, shadowed skin might be classified as a first- or third-degree burn. The occurrence of shadowed skin usually takes place around finger joints and between fingers. The introduction of a dictionary for shadowed skin allowed a reduction in the number of false positives.

Two sets of experiments were conducted using either dictionaries that were directly built and learned dictionaries that were trained with *K - singular value decomposition*. *Untrained dictionaries* required of collecting atoms from multiple images to generate each dictionary *D*, a matrix where the number of rows is equal to the number of features (54 features), and the number of columns is equal to the number of atoms

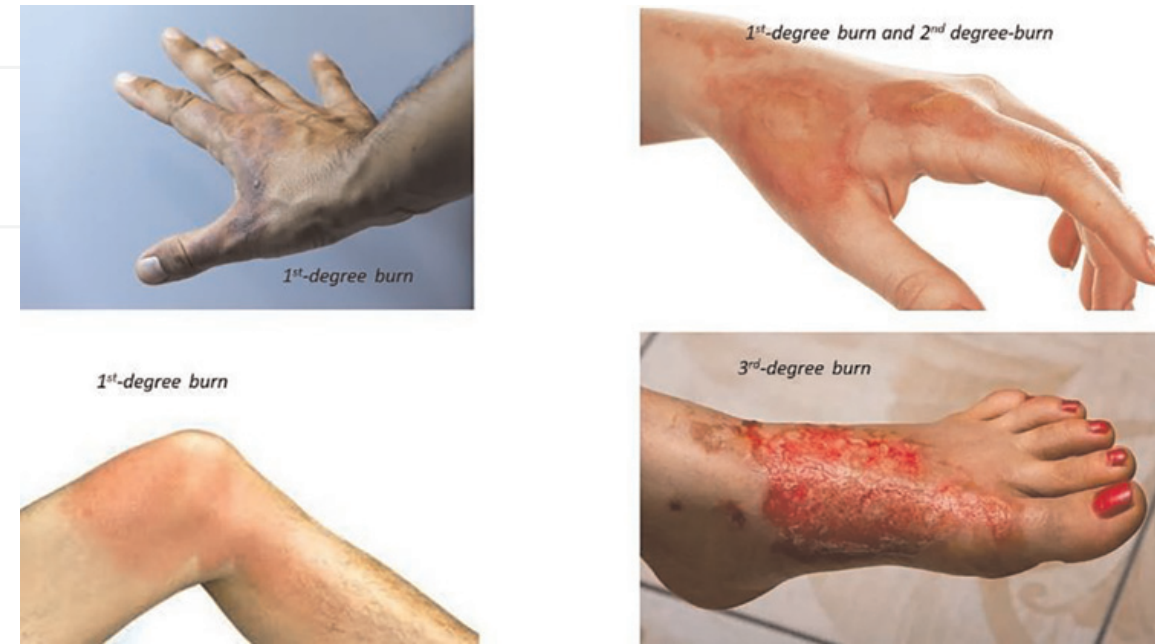


Figure 4.
Examples of images with first-, second-, and third-degree burns.

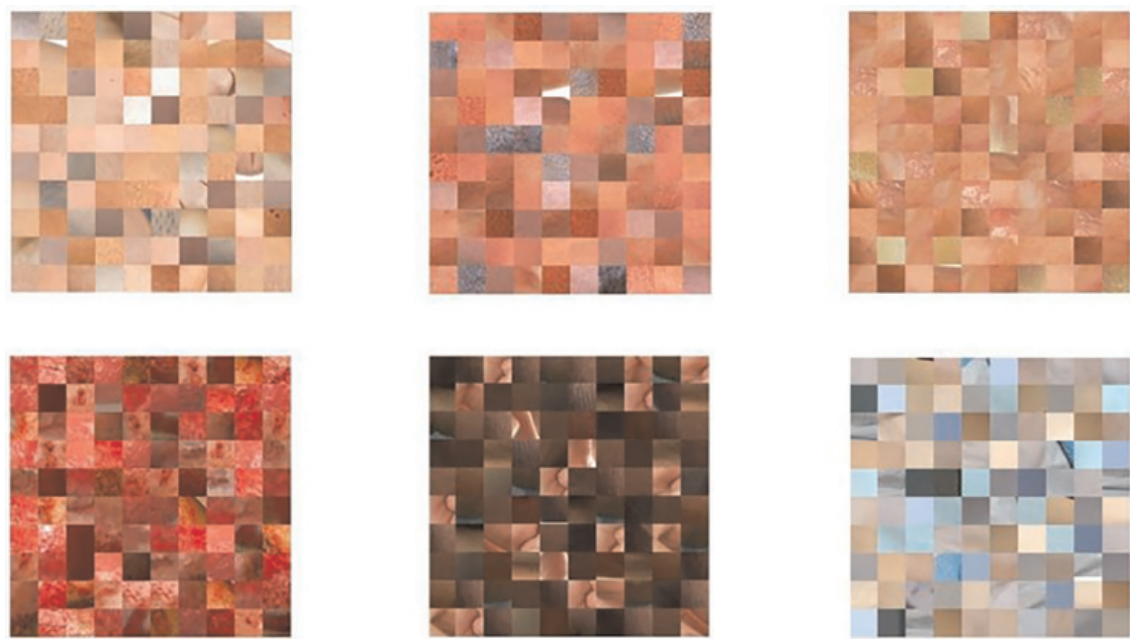


Figure 5.
Collections of patches from first-, second-, and third-degree burns, healthy skin, shadowed skin, and background.

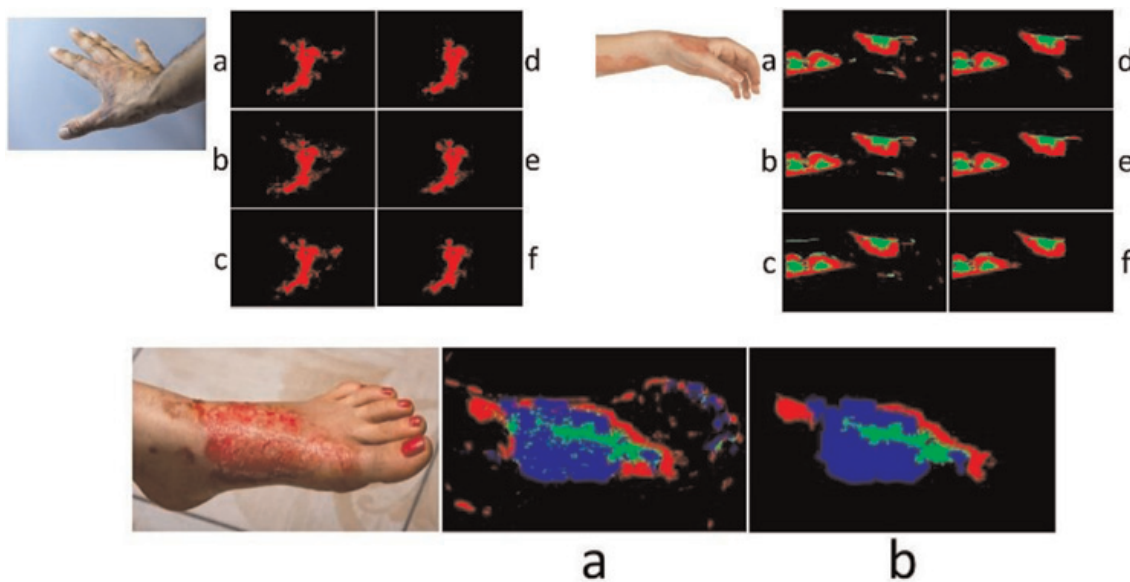


Figure 6.
Three examples of detection and classification of burnt skin. First-degree burn on dark skin (upper left panel), first- and second-degree burns (upper right panel), third-degree burn (lower panel).

(3000 atoms). An additional set of 500 observations X was necessary to train each learned dictionary with K – singular value decomposition. These observations were collected from various images at various positions to train the dictionaries. Feature values were mostly positive, while atom entries after K -SVD were both positive and negative. During dictionary learning, K – *singular value decomposition* was applied to generate dictionaries for each of the six classes.

Figure 6 shows the results of skin burn segmentation and classification. The images under analysis shows regions with first-degree and second-degree burns. Second-degree burns affect deeper skin layers (epidermis and dermis). Image patches detected as healthy skin or background are painted in black. A second-degree burn is

usually surrounded by a first-degree burns. First-degree burns are painted in red, second-degree burns are painted in green. The image under analysis presents shadowed skin around the fingers, which might be incorrectly classified as burns. Thus, extraction of connected components was used to find the largest objects and eliminate the smallest objects. An object was suppressed if its size was less than one fifth the size of the largest object. Feature vectors were extracted from 23×23 patches. The first two rows in each panel show the segmentation results obtained with untrained dictionaries. The third row of each panel shows the result of segmentation based on trained dictionaries. The best results were obtained with untrained dictionaries. The second and third columns of each panel show the results before and after extraction of connected components, respectively. After all the patches were classified, each image patch was re-classified by analyzing the class of the nearest patches and taking the majority vote. The results of re-classification depended on the dimensions of the patch to be classified. The third panel presents a challenging case since this image is characterized by the presence of shadowed skin. Furthermore, toenails are painted in red so that these regions ended up being classified as third-degree burns (false positives) before extraction of connected components. Third-degree burns destroy the epidermis and dermis, and their physical appearance is typically red. For this case, the best results were obtained by using untrained dictionaries.

Table 1 shows the number of true positives, false positives, and false negatives for each segmentation experiment in **Figure 6** and for each burn degree. When optimal experimental settings (second row) were used for this image, the accuracy and sensitivity were both 100%.

The classification performance of the proposed method for identification and classification is shown in **Table 2**, in which $Sensitivity = \frac{TP}{TP+FN}$ and $Precision = \frac{TP}{TP+FP}$ values are presented for each burn degree. The detection performance is shown in the last column, where TP accounts for all the detected burns (all degrees), FN accounts for all the undetected burns (all degrees), and FP correspond to all the incorrectly detected burns (all degrees). Performance metrics were obtained by using 4-fold cross-validation in which 75% of the images (15 images) were randomly selected to build dictionaries, and the other 25% (5 images) were used to test burn identification and classification. This process was repeated four times so that all images were tested. Sets of 3000 patches were obtained from 15 images to build dictionaries. The performance metrics were obtained by using two strategies for the building of dictionaries, extraction of atoms from images and trained dictionaries. Various sparsity factor values ($L = 1, 2, 3$) were tried during segmentation, and $L = 1$ provided the best performance. The size for patch labeling was 1 in some cases and 3 in other cases. The $0 < T < 1$ for suppression of the smallest connected components was experimentally established as $0.25 < T < 0.3$. Since the purpose of this research is to assist physicians in the identification and classification of burnt skin, the segmentation process was aimed at burn regions. According to **Figure 6**, the segmentation process identified multiple burns, some of which were not true burns, i. e., false positives. *False positives* corresponded to healthy skin or shadowed skin or background. False positives were characterized as considerably smaller in absolute area than TP ; this is the reason why extraction of connected components was used to detect and suppress the smallest components. The results of suppression of the smallest connected components are shown in the rightmost column in panels of **Figure 6**. The introduction of three classes, shadowed skin, healthy skin, and background, reduced the prevalence of false positives while improving the skin burn degree classification performance.

	1 st degree TP	1 st degree FP	1 st degree FN	2 nd degree TP	2 nd degree FP	2 nd degree FN
Untrained dictionaries $L = 1$	3	2	0	3	0	0
Untrained dictionaries $L = 3$	3	0	0	3	0	0
Trained dictionaries $L = 1$	3	0	0	3	1	0

Table 1.
Number of true positives, false positives and false negatives for the 1st and 2nd degree burns shown in the right upper panel in Figure 6.

	1 st degree burns	2 nd degree burns	3 rd degree burns
Sensitivity untrained dictionaries	0.9394	0.9615	0.9565
Sensitivity trained dictionaries	0.9091	0.9231	0.9217
Precision untrained dictionaries	0.9394	0.8929	0.9402
Precision trained dictionaries	0.8824	0.8889	0.9138

Table 2.
Sensitivity and precision for the classification and detection of skin burns.

In addition to higher performance, another advantage of untrained dictionaries over trained dictionaries is low computational complexity, since the collection of observations and the training stage are not required. In order to explain why untrained dictionaries provide better detection performance, there are two observations: (1) Atoms are normalized feature vectors, and most features (average value, variance, contrast, angular second moment, inverse difference moment) are positive. (2) After training, atoms are adjusted and most of their entries are no longer positive. Untrained atoms use raw features without adjustments introduced with training. Untrained features lie on a constrained sub-space given that they are non-negative. Thus, they are more densely concentrated within the same sub-space where the signal to be reconstructed is located. This is why the error introduced by the sparse reconstruction of a signal is smaller over untrained atoms.

4. Conclusions

The proposed method for detection and classification of burns on images is found to be adequate as an assisting tool in the diagnosis of skin burns and it offers non-intrusiveness and low economic cost. The tools and analyses in the current work contribute to skin burn detection and classification: (1) Achieving high classification performance (sensitivity and precision above 90\%); (2) sparse reconstruction of image patches with dictionaries was shown to be adequate for this application; (3) Analyzing images under diverse conditions of skin color (ethnicity) and illumination. It was found that untrained dictionaries outperformed trained dictionaries for this application. The two groups of features most commonly used in the imaging-based classification of burns, namely color and texture, provided the reported method with good classification performance.

Acknowledgements

The authors would like to acknowledge the Mexican National Council on Science and Technology (CONACyT) and the Universidad de las Américas Puebla (UDLAP) for their support through the doctoral scholarship program.

Conflict of interest

Authors declare no conflict of interests.


IntechOpen

Author details

Brenda Rangel-Olvera* and Roberto Rosas-Romero*
Department of Computing, Electronics and Mechatronics, Universidad de las
Américas Puebla, Puebla, Mexico

*Address all correspondence to: brenda.rangel@udlap.mx and
roberto.rosas@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Saraswathi D, Naimisha U, Yukthashree PJ, Sahana G. Human burn diagnosis using machine learning. *IJCST*. 2020;8(3):8
- [2] Hoffman M. The skin (Human Anatomy): Picture, definition, function, and skin conditions. WebMD. 2019. <https://www.webmd.com/skin-problems-and-treatments/picture-of-the-skin> [Accessed: April 12, 2021]
- [3] Mehta M, Tudor GJ. *Parkland Formula*. Treasure Island (FL): StatPearls Publishing; 2020
- [4] Amershi S, Cakmak M, Knox WB, and Kulesza T. "Power to the people: The role of humans in interactive machine learning." *AI Magazine*. Dec 2014;35(4):105–120. DOI: 10.1609/aimag.v35i4.2513
- [5] Abubakar A, Ugail H. Discrimination of human skin burns using machine learning. In: Arai K, Bhatia R, Kapoor S, editors. *Intelligent Computing*. Vol. 997. Cham: Springer International Publishing; 2019. pp. 641-647. DOI: 10.1007/978-3-030-22871-2_43
- [6] Rowland R et al. Burn wound classification model using spatial frequency-domain imaging and machine learning. *Journal of Biomedical Optics*. 2019;24(5):1. DOI: 10.1117/1.JBO.24.5.056007
- [7] Bhansali R, Kumar R. BurnNet: An efficient deep learning framework for accurate dermal burn classification. *medRxiv*. 2021. DOI: 10.1101/2021.01.30.21250727
- [8] Jiao Y, Qian C, Fei S. Mask convolution for filtering on irregular-shaped image. In: 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Wuxi, China. 2018. pp. 115-118. DOI: 10.1109/DCABES.2018.00039
- [9] Khan FA et al. Computer-aided diagnosis for burnt skin images using deep convolutional neural network. *Multimed Tools Application*. 2022;3(2): 512-525. DOI: 10.1007/s11042-020-08768-y
- [10] Şevik U, Karakullukçu E, Berber T, Akbaş Y, Türkyılmaz S. "Automatic classification of skin burn colour images using texture-based feature extraction," *IET Image Processing*, Sep 2019;13(11): 2018–2028. DOI: 10.1049/iet-ipr.2018.5899.
- [11] Rehman Butt AU, Ahmad W, Ashraf R, Asif M, Cheema SA. Computer aided diagnosis (CAD) for segmentation and classification of burnt human skin. In: 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Swat, Pakistan. 2019. pp. 1-5. DOI: 10.1109/ICECCE47252.2019. 8940758
- [12] Yadav DP, Sharma A, Singh M, Goyal A. Feature extraction based machine learning for human burn diagnosis from burn images. *IEEE Journal of Translational Engineering and Health Medicine*. 2019; 7:1-7. DOI: 10.1109/JTEHM.2019.2923628
- [13] Rangel-Olvera B, Rosas-Romero R. Detection and classification of burnt skin via sparse representation of signals by over-redundant dictionaries. *Computers in Biology and Medicine*. 2021;132:104310. DOI: 10.1016/j.combiomed.2021.104310

[14] Khan A. Burns: Types, Treatments, and More. <https://www.healthline.com/health/burns>

[15] Elaine KL. Second-Degree Burn: Everything You Need to Know. <https://www.medicalnewstoday.com/articles/325189>

[16] Brouhard R. Burn Pictures: A Close Look At First, Second, and Third Degree. <https://www.verywellhealth.com/burn-pictures-4020409>

[17] Burn Victims Resource Center. Phases of Burn Care Download. <https://www.burnvictimsresource.org/first-degree-burns-look-like/>

[18] Howzak. Best Home Remedies for Burns and How to Treat Them. <https://howzak.com/best-home-remedies-for-burns-and-how-to-treat-them-2018/>

[19] "Burns," ISCN. <https://skincare.network/burns/> [Accessed: February 3, 2022]

[20] Family First Urgent Care. How Are First, Second, and Third Degree Burns Treated Differently? <https://familyfirst-urgentcare.com/how-are-first-second-and-third-degree-burns-treated-differently/>

Random Wavelet Coefficients Pooling for Convolutional Neural Networks

Daniel Trevino-Sanchez and Vicente Alarcon-Aquino

Abstract

The number of applications requiring effective object detection and classification are constantly increasing. In this regard, one of the most utilized solutions are Convolutional Neural Networks (CNN), which are also constantly pushed to improve their accuracy. When improving CNN, the key layers are the ones related to feature extraction and size reduction. In recent years, one way to improve the layers behavior is to combine them with different techniques, like the Multi-Resolution Analysis (MRA). The aim of the study is to use the Wavelet Transform Coefficients randomly incorporating the Lifting Scheme, which is a second-generation of MRA in a formal model as a pooling layer inside the CNN. This can emulate the Max-pooling dynamic-filter effect that prevents the CNN from overfitting but independently from the signal's shape, achieving a measurable increment of its accuracy using a benchmark dataset. To validate this study, seven pooling methods, including the proposed model, are tested using three different wavelet functions and one benchmark dataset. The results are compared with five of the most used pooling methods, which are also considered as the state-of-the-art.

Keywords: convolutional neural network, feature extraction, lifting scheme, multi-resolution analysis, random pooling, wavelet coefficients

1. Introduction

Detect and classify objects correctly is challenging, especially if they are cropped or occluded. Even hard weather conditions or poorly lit scenes constitute an obstacle to achieve the task. The Convolutional Neural Network (CNN) is one of the techniques with highest performance in this research field. However, the drawbacks previously mentioned are some of the reasons to constantly seeking to improve CNNs.

There is constant research exploring new technologies to improve the CNN performance. Since deep learning was first proposed there have been plenty of recommendations to increase CNN accuracy. They go from enlarging training datasets and reducing batches [1], to pretrained networks, skip connections [2], deeper architectures [3], and complex connections [4]. CNN can obtain huge amounts of features this is the main reason they are being used in many applications. The training stage is

where all key patterns are learned. This capability is constructed updating CNN parameters values, and the network accuracy improves by extracting more relevant patterns during training.

Nevertheless, increasing the effectiveness of CNNs architectures is not only about learning patterns, but extracting high quality features and moreover keeping the right ones. Therefore, feature extraction is still an open field to research. To address this challenge some approaches have focused on different feature extraction techniques to identify key patterns. For example, Farah Malik et al. [5] and Li et al. [6] utilized spectral analysis to analyze heart sounds, the first one uses wavelet transform to produce a scalograms to trains a CNN, while the second one uses Denoising Autoencoders to extract features from the spectrograms to do the same. In Haque and Mishu [7], Principal Component Analysis is used to extract a feature vector and reduce its dimensionality. Then the images are classified in a Multi-scale CNN. Statistical methods are also common, the work in Rao et al. [8], focused on detecting objects with camouflage. First, the image is divided into pieces of equal size named subblocks, each subblock is decomposed into two levels employing Discrete Wavelet Transform and Daubechies wavelet Db2. Then the coefficients obtained are used to extract features statistically. These features feed the final stage of the algorithm to detect and locate the objects. In general, the main idea is to do more with less.

Within the CNN there are several kinds of layers, and where most of the size reduction takes place is in the pooling layers [9, 10]. And according to Liu et al. [11, 12], the pooling process enlarges receptive field and decreases computational cost. But it is also in the dimension reduction stage where some attributes are lost. Traditional pooling methods like average, maximum, mixed, and stochastic are known to lose important details, specifically high-frequency patterns. Moreover, according to Fujieda et al. [13, 14], conventional CNN lost most of the spectral information. And they can be considered as a deficient form of Multi-Resolution Analysis (MRA). This problem can be overcome by incorporating a well-structured MRA like the Wavelet Transform (WT). The WT is particularly useful preserving details by capturing frequency and location features, according to Daubechies [15, 16]. CNN feature extraction process can be enriched directly from the inside to get more useful set of patterns, allowing less relevant features while reducing the map dimensionality (see, e.g., [17–21]). However, contrarily to the traditional methods it requires more computational resources.

One of the most popular methods used in CNN for size reduction is the max-pooling, but it presents some unsolved issues that need to be addressed, like lack of formality. Even though an equation describes how to apply the Max-pooling layer, it is an empirical solution within the CNN. Contrarily, the proposed random wavelet pooling can be presented as a formal mathematical model. It uses the Lifting Scheme, which is a second-generation wavelet transform, meets all these requirements, and fits inside the CNN as another layer. Formalizing the pooling method leads to a better understanding of its internal behavior, it helps to design improved models that optimize the process. In this sense, new models may reproduce and enhance the advantages of the process without repeating the disadvantages of the pooling methods.

Additionally, max-pooling behaves like a dynamic filter, but also like a random amplifier, both controlled indirectly by the signal. In many cases, this property helps to prevent the network from overfitting and increase its accuracy. But in some others, it reduces its effectiveness and prevents CNN from reaching high accuracy levels. Some studies have been focusing on this problem creating new pooling techniques. They focus on improving these layers by blending MRA techniques within the CNN layers to get useful patterns [22]. Following that idea, MRA can be seen as high and

low-pass filters that decompose the signal into separated components that can be used to duplicate the max-pooling behavior.

The main contribution of this study is: a computational model that functions as a pooling layer within a CNN selecting signal components randomly as a dynamic filter independent from the signal features preventing overfitting and improving accuracy.

Other contributions are:

- Max-pooling frequency response analysis.
- Lifting Scheme model that selects coefficients randomly.
- Implementation of Lifting Scheme for wavelet functions Daubechies 4 and 6.
- Embedded of the previous designs as CNN layers.
- Train and test the CNN using a selected benchmark dataset to test the accuracy.
- Assessment of 5 standard methods and 7 wavelet methods using 3 different wavelet functions.
- State-of-the-art accuracy achieved by the proposed method.

The organization of this paper is as follows. Section 2 presents some related work using wavelet transform with CNN models. Section 3 briefly explains the lifting scheme and the most common pooling methods. Section 4 describes the methodology and the proposed model. Section 5 reports the characteristics of the experiments and their results. Finally, Section 6 draws some conclusions.

2. Related work

CNNs are becoming more capable for image analysis application, due to the increasing number of layers and complex connections. However, adding more layers to increases performance produce more intricate models and higher computational cost. Eventually, deeper networks may start losing accuracy. On the other hand, MRA keeps all the information contained and distributed into its coefficients through the different resolution levels. Each one contains specific features; higher resolutions show more details and lower resolutions shows only the strongest features. In recent years, more studies have applied MRA to the CNN. Some of these studies are presented in **Table 1**, where the first two columns indicate the publication year and the application.

All the studies shown are applications embedded within the CNN to replace the pooling layers. The last three columns describe the kind of wavelet solution implemented. For example, the third column shows that Discrete Wavelet Transform (DWT), Lifting Scheme Wavelet Transform (LSWT) and Wavelet Package Transform (WPT) are the only wavelet transforms utilized in the analyzed studies. At the same time, the last three columns complete each model information by adding the wavelet family selected, the number of wavelet coefficients applied and the levels of decomposition. The wavelets used are families like Daubechies or Coiflet. The 2D wavelet transforms produce 4 coefficients and all analyzed models require the approximation coefficients. However, not all studies utilize the details coefficients.

Research	Description	Wavelet transform	Wavelet function	No. coeff.	Levels of dec.
Proposed model	Handwritten digits classification	Lifting scheme	Haar, Db4 and Db6	4	1
2021 [22]	Handwritten digits classification	Lifting scheme	Haar	1	1
2019 [10]	Handwritten digits classification	DWT	Haar, Db, Coiflet and mixes	1	3
2019 [23]	Detect lung tumors	DWT	Haar and Shannon	1	3
2019 [12]	Image restoration	PWT	Haar, Db2, Haar + Db2	4	3

Note: This table presents some of the most recent research as examples and does not intent to be a survey study.

Table 1.
Wavelets Transforms used to improve CNN pooling layers.

In Ma et al. [24], the detail coefficients are used partially. In the last column indicates the number of levels of decomposition applied in each model. These studies introduce new frameworks with interesting results, like Bastidas-Rodriguez et al. [25], this model is embedded in a larger CNN to detect scenes and classify textures or compress images [24]. Other approaches also detect scenes and classify handwriting digits [26]. But, replacing pooling layers is the most common uses of wavelets within CNN like in Refs. [9–12, 23], as well as activation functions as reported in Refs. [27–31]. It is used for most of the previous applications, as well as for image restoration, lung tumors detection, among many more.

The aim of the study, initially inspired by Williams and Li [9], is to use the Wavelet Transform Coefficients randomly in a formal model as a pooling layer, to emulate the Max-pooling effect but independently from the signal’s shape. Achieving a measurable increment of its accuracy using a benchmark dataset.

3. Background

3.1 The lifting scheme

MRA like the lifting scheme is used to design different wavelets. This wavelet transform is named as the second-generation wavelet transform and developed in Sweldens [32]. It uses convolution operations instead of filters, simplifying mathematical operations [33, 34]. The lifting scheme splits the signal in even and odd samples to apply them simple operations, as shown in **Figure 1**. The general version of the lifting scheme was presented in Sole and Salembier [36]. The Lifting Scheme steps are detailed in Trevino-Sánchez and Alarcón-Aquino [22]. The Wavelet Lifting Scheme decomposes the 1D Signal into two coefficients: approximations (L: Low Frequency) and details (H: High Frequency).

Figures 1–3 show three wavelet functions for a 1D lifting scheme: Haar, Daubechies 4 (Db4) and Daubechies 6 (Db6) accordingly, where ($\downarrow 2$) indicates that the signal is downsampled by half. Using Z means that a previous sample is used, while Z^{-1} is for the next one. Additionally, Z^{-1} also states that the signal is splitted in even and odd components [35].

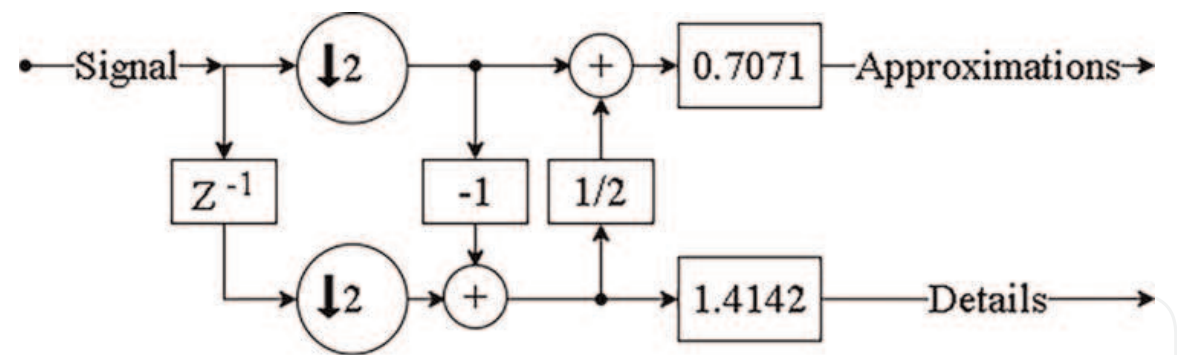


Figure 1.
The Haar Wavelet Lifting Scheme for a 1D Signal [35].

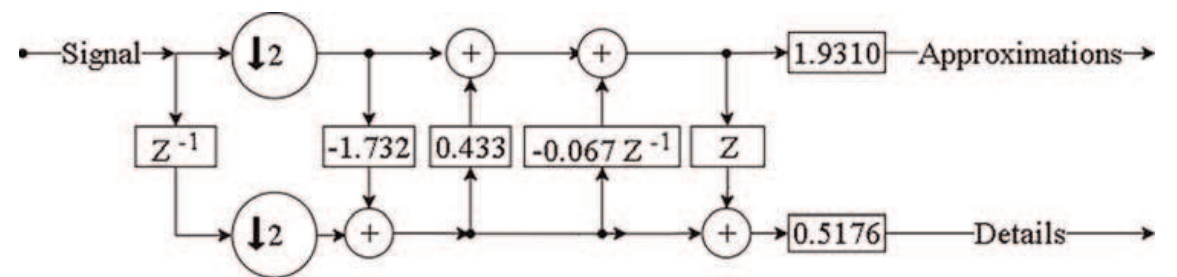


Figure 2.
The Daubechies 4 Wavelet Lifting Scheme for a 1D Signal [35].

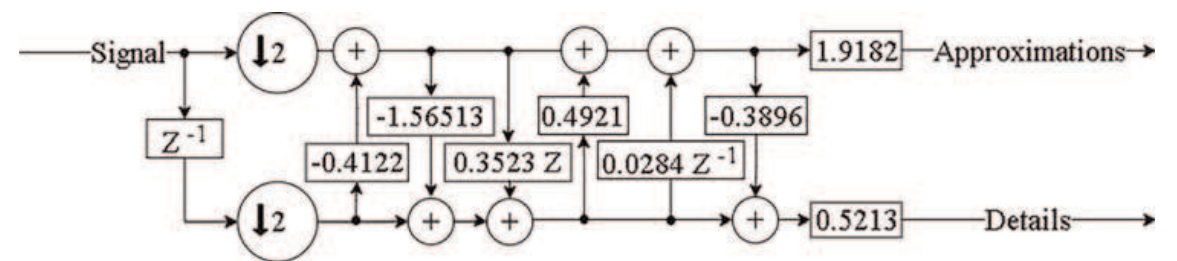


Figure 3.
The Daubechies 6 Wavelet Lifting Scheme for a 1D Signal [35].

When the lifting scheme needs to be applied to an image a 2D version is required, as shown in **Figure 4**. In this example a Haar wavelet function is used but it can be applied to any other wavelet function. It simply repeats the same block two more times, one for the low and one for high frequency coefficients. The first block splits the signal in one direction (horizontal), while the other two blocks divide the signal in the other direction (vertical) [37]. At the end of the decomposition process, four coefficients are obtained, they are also known as approximations (LL: low-low pass filters), verticals (HL: high-low pass filters), horizontals (LH: low-high pass filters) and diagonals (HH: high-high pass filters).

3.2 Pooling

The pooling layer is a subsampling technique used to reduce the size of the information. This method shirks a predefined region into one single value. The most used pooling methods are max-pooling [38] and average pooling [39]. Max-pooling takes the highest value of each region R of size of $r \times r$ using a stride of $S \times S$ and no padding,

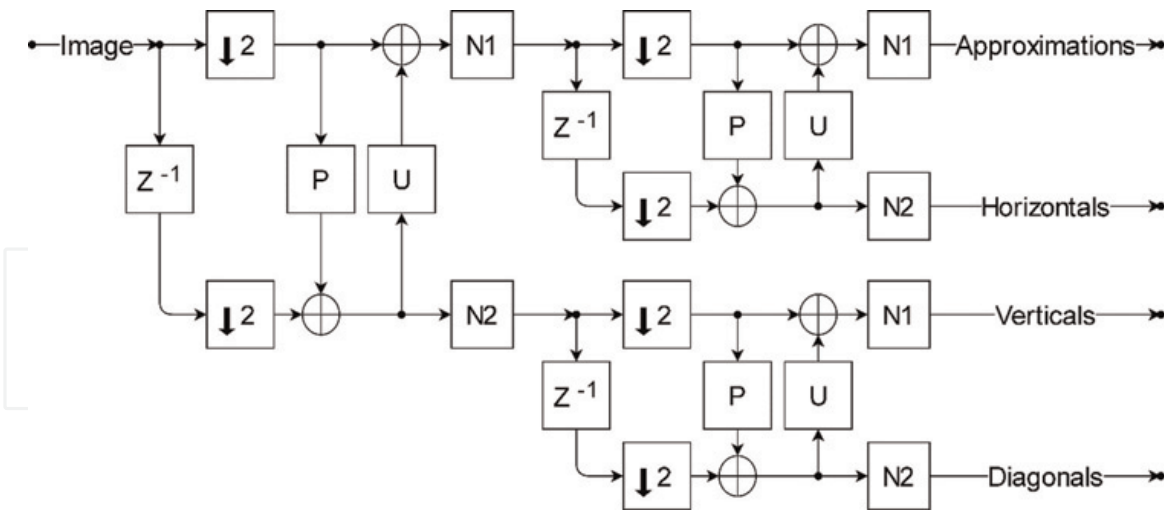


Figure 4.

Haar Wavelet Lifting scheme for 2D decomposes an image into four subbands. The subbands names are approximations (LL), horizontals (LH), verticals (HL), and diagonals (HH).

until the entire image is reduced. In this study, both parameters r and S take the value of 2, these parameters need to be equal to avoid skipping any value, and thus ignore that piece of information.

The equation for max-pooling is described in Eq. (1).

$$F_{max}(x', y') = \max_{\substack{k=a \dots a+(r-1) \\ l=b \dots b+(r-1)}} \{I(k, l)\} \quad (1)$$

$$\left\{ \begin{array}{l} i = \{0, 1, 2, \dots, n\}, \quad x' = i + 1, \\ j = \{0, 1, 2, \dots, m\}, \quad y' = j + 1, \\ I(k, l) \mid a = ((S * i) + 1), n \mid a \leq (x - (r - 1)), \\ b = ((S * j) + 1), m \mid b \leq (y - (r - 1)), \\ \forall (k, l) \exists (a, b), \quad \forall I(k, l) \in I(x, y) \end{array} \right\}$$

where I is the input image and (x, y) its dimensions. F_{max} is the obtained output and (x', y') are indexes that point out to each element of the reduced information map. $I(k, l)$ is the pooling region R and the indexes (k, l) indicate each value of the region. Additionally, (a, b) are variables used to slide the region indexes through the entire images at a S pace, while (n, m) prevents from overpassing the image size considering both S and r . Finally, (i, j) are the iterations required to cover all possible regions in the image.

Similarly, to max-pooling, average pooling reduces the feature map by calculating the average value of each region. The equation for average pooling is shown in Eq. (2).

$$F_{avg}(x', y') = \frac{1}{|I(k, l)|} \sum_l \sum_k I(k, l) \quad (2)$$

where $F_{avg}(x', y')$ is the output, and $|I(k, l)|$ is the magnitude or number of elements in the pooling region. The elements of the pooling region are added by row and by column until all have been added, then the result is divided by the number of elements to obtain the average value of the entire region.

There are also two probabilistic pooling methods: mix pooling [40] and stochastic pooling [41]. Mix pooling takes either the maximum or the average value randomly during the training process. Mix pooling can select values the region or between channels. Mix pooling is described in Eq. (3).

$$F_{mix}(x', y') = \lambda F_{max}(x', y') + (1 - \lambda) F_{avg}(x', y') \quad (3)$$

where λ is a random value from the interval $[0, 1]$.

For the stochastic pooling, each value within the region is assigned with a probability, the higher the value, the higher the probability it gets. All probability values from the same region add up 100%. Then a value is selected based on the probabilities of the region. The Eq. (4) normalizes the value of each element within a pooling region to calculate their probabilities $p(k, l)$.

$$p(k, l) = \frac{I(k, l)}{\sum' \sum^k I(k, l)} \quad (4)$$

$$F_{stoch}(x', y') = P(p(k, l)) \quad (5)$$

where $P()$ is a function that selects a sample from the multinomial distribution created with the probabilities obtained by the Eq. (4).

4. Methodology

All experiments used MATLAB® R2020b & Deep Learning Toolbox, and the CNN utilized was the MatConvNet, which is also reported in Williams and Li [9]. Only one change was made to the CNN, the local response normalization (LRN) layer is used instead of the batch normalization (BN) layer for practicality of the applications. All training used stochastic gradient decent, initial learning rate of 0.001, minibatch of 64, and 2 iterations, the first one for 20 epochs and the second one for 40 epochs. Besides, all tests were run on a 64-bit operating system (Windows 10). The core is AMD RYZEN 7 4800H series with Radeon Graphics @ 2.90 GHZ, with 16.0 GB RAM and the GPU utilized is NVIDIA GEFORCE GTX 1050. This study used three wavelet functions implemented for the lifting scheme: the Haar, Daubechies 4 and 6.

4.1 Max-Pooling frequency analyzes procedure

In CNN, feature loss in pooling layers is more relevant than in convolution layers. Thus, the pooling process is where most of the size reduction of the information takes place. The main idea of the pooling method is to select just one value and eliminate the rest. Depending on the pooling method that value can be the maximum, the average, or a randomly selected one. For example, the average pooling loses relevant information by acting like a low-pass filter, generalizing features when obtaining its mean value. In contrast, max pooling focus on the highest magnitude producing an irregular effect that allows the CNN to improve its accuracy and prevent overfitting, which is the reason to be selected for this study over other methods.

The pooling methods used in this study are divided into two groups, the single or standard methods and the wavelet methods. The objective it to perform a complete comparison among all these methods. The first group consists of five methods:

Average (Avg), Maximum (Max.), Mixed (Mix.), Mix. by Region and Stochastic (Stoch.), which are the most referred methods for CNNs. The second group is made of seven methods based on The Lifting Scheme: The Lifting Scheme only for the LL Coefficient (Lift), hybrid combination of Lifting Scheme LL coefficient and max-pooling (Lift. + Max), The Lifting Scheme for all the Coefficient (Lift.+ Coeff.), hybrid combination of all the Lifting Scheme Coefficient and Max-pooling (Lift. + Max. + Coeff.), Lifting LL coefficient with one other coefficient selected randomly (LH, HL and HH) for the entire channel (Lift. + Rand. Channel), same as before but for a 2x2 region (Lift. + Rand. Region) and finally, the last one is also as the previous one but adding the Max-pooling (Lift. + Max. + Rand. Region) for a 2x2 region.

In a previous study [22], the Max-pooling is considered as some kind of high-pass filter that retains some high frequency features. However, the results suggested otherwise. In consequence, the max-pooling effect was analyzed in the frequency domain. To keep it as simple as possible, we reduce the pooling process for a 1D signal. To analyze the max-pooling frequency response several simulated signals are generated randomly to circumscribe as much spectrum values and combinations as possible, as shown in **Figure 5**.

Once the max-pooling frequency response is analyzed, its behavior may be emulated by a new model. In this sense, the proposed model looks to reproduce the effects of max-pooling but eliminating the need of an input signal with specific characteristics, to trigger them at some point.

4.2 Proposed model

The proposed model in this study is a pooling method that incorporates the 2D lifting scheme. The lifting scheme is used as a base of the model because it is very suitable within CNN architecture, as in Ma et al. [24] and Bastidas-Rodriguez et al. [25].

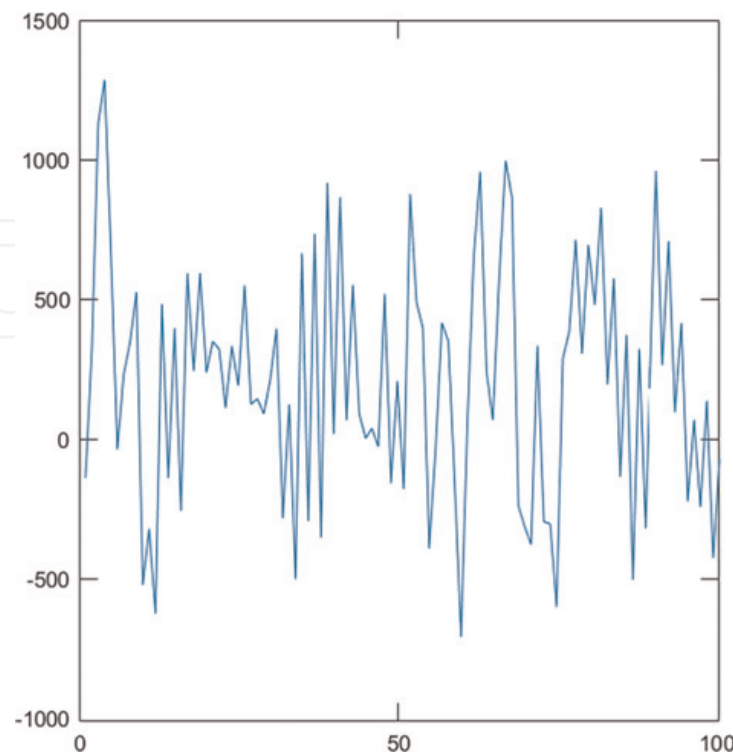


Figure 5.
Example of a random input signal used to produce max-pooling frequency responses.

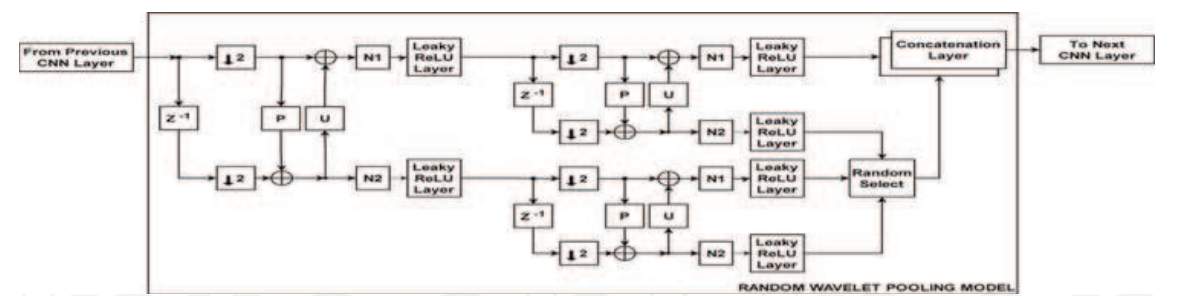


Figure 6. The proposed model is a pooling method that uses the lifting scheme. It extracts the LL coefficient and concatenates it with a randomly selected coefficient (LH, HL or HH). The model shown is for the Haar case, but it may be also constructed for the Db4 and Db6.

It is also a MRA technique that preserves frequency and space information when reducing dimensionality without any loss. Even if only the approximation coefficients were used, the lifting scheme could extract relevant features from images to improve CNN performance.

The model is shown in **Figure 6**, although it is for the Haar case, it can be also constructed for other wavelet functions such as Db4 and Db6. First, it extracts the four wavelet coefficients (LL, LH, HL or HH) and selects randomly one of the three coefficients with the highest frequencies. Finally, the selected coefficients are concatenated with the lowest frequency coefficient (LL), according to Williams and Li [26], this subband has most of the image energy and structure information and it is the most alike to the original image. The proposed model is then embedded within the CNN instead of every pooling layer. This pooling layer performs 2x2 region based downsampling, and 2x2 stride with no padding.

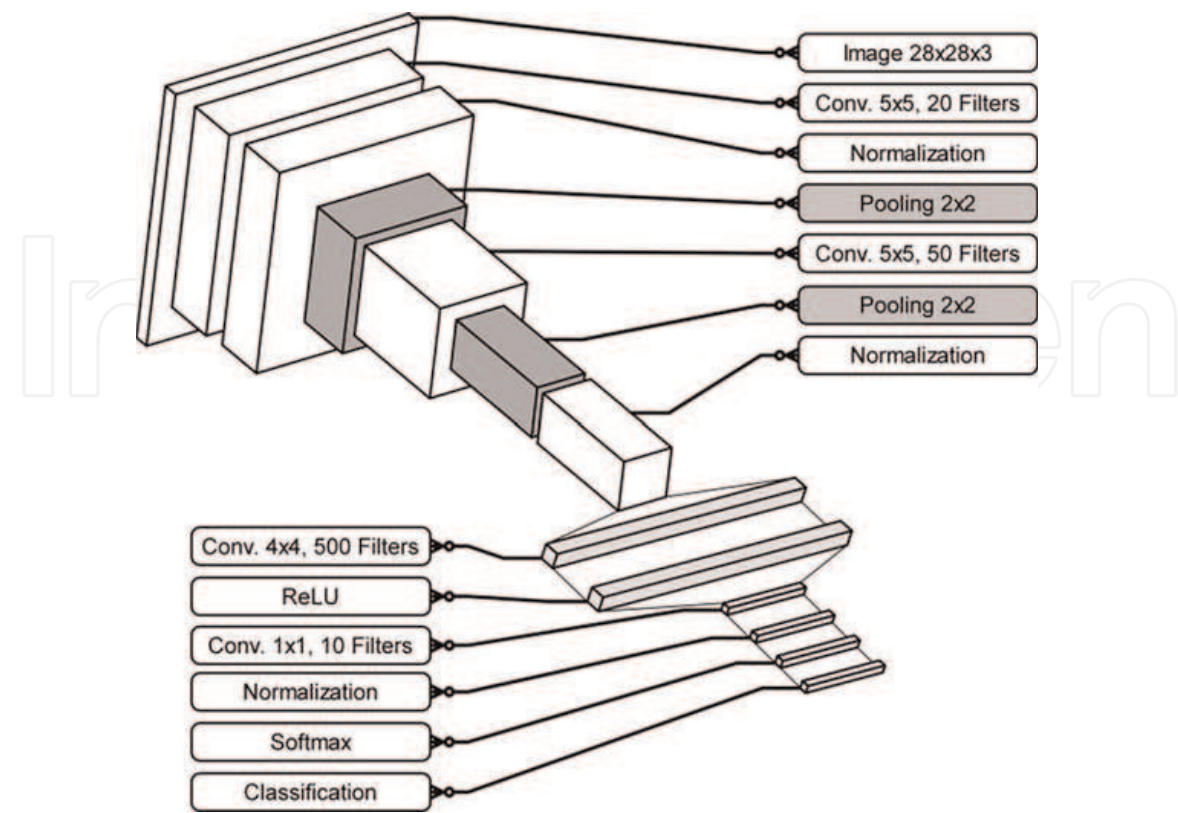


Figure 7. Architecture used for the CNN. The highlighted blocks indicate the pooling layers.

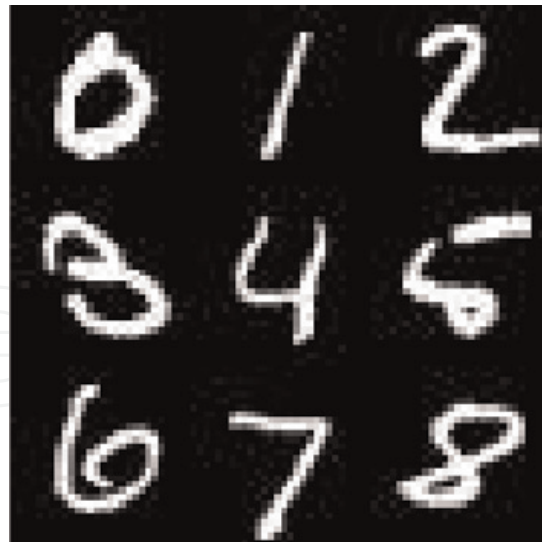


Figure 8.
The benchmark dataset selected for this study was the MNIST.

4.3 CNN architecture and benchmark dataset

The network structure used for this study is made of the following layers: four convolutions, three normalizations, one ReLU, one softmax, one classification, and two pooling. The highlighted blocks indicated in **Figure 7** are the pooling layers. Each pooling method has been implemented in those layers, utilizing one single method for the entire network. All methods have been tested for the dataset considering a 2×2 region, 2×2 stride and no padding, except for the Lift. Random Channel which uses the entire image channel as a region.

The benchmark dataset used for training and testing is the MNIST, as shown in **Figure 8**. The MNIST dataset [42] contains a large number of color images of hand-written single digits divided into 10 classes, and their size is 28×28 pixels. The training set is made of 60,000 images and the testing set of 10,000 images.

5. Results and discussion

The results presented first contain the max-pooling frequency response, and then the accuracy measured for every tested model.

5.1 Max-pooling frequency response

The first thing to be analyzed was the max-pooling frequency response, as explained earlier, the max-pooling was reduced to a 1D signal to facilitate the results interpretation. Besides, multiple random signals were tested to look for a pattern behavior. **Figure 9** shows some of the frequency responses obtained during these tests. Item (a) presents an irregular response, it has different peaks and valleys without any pattern, while response (b) resembles a low-pass filter, (c) a multi-pass band filter and (d) a stop band filter. However, most of the values are above 0 dB, which means that max-pooling functions more like an amplifier rather than a filter. In summary, the obtained responses show a dynamic behavior dependent of the input signal, different signals triggers different frequency responses. Additionally, although

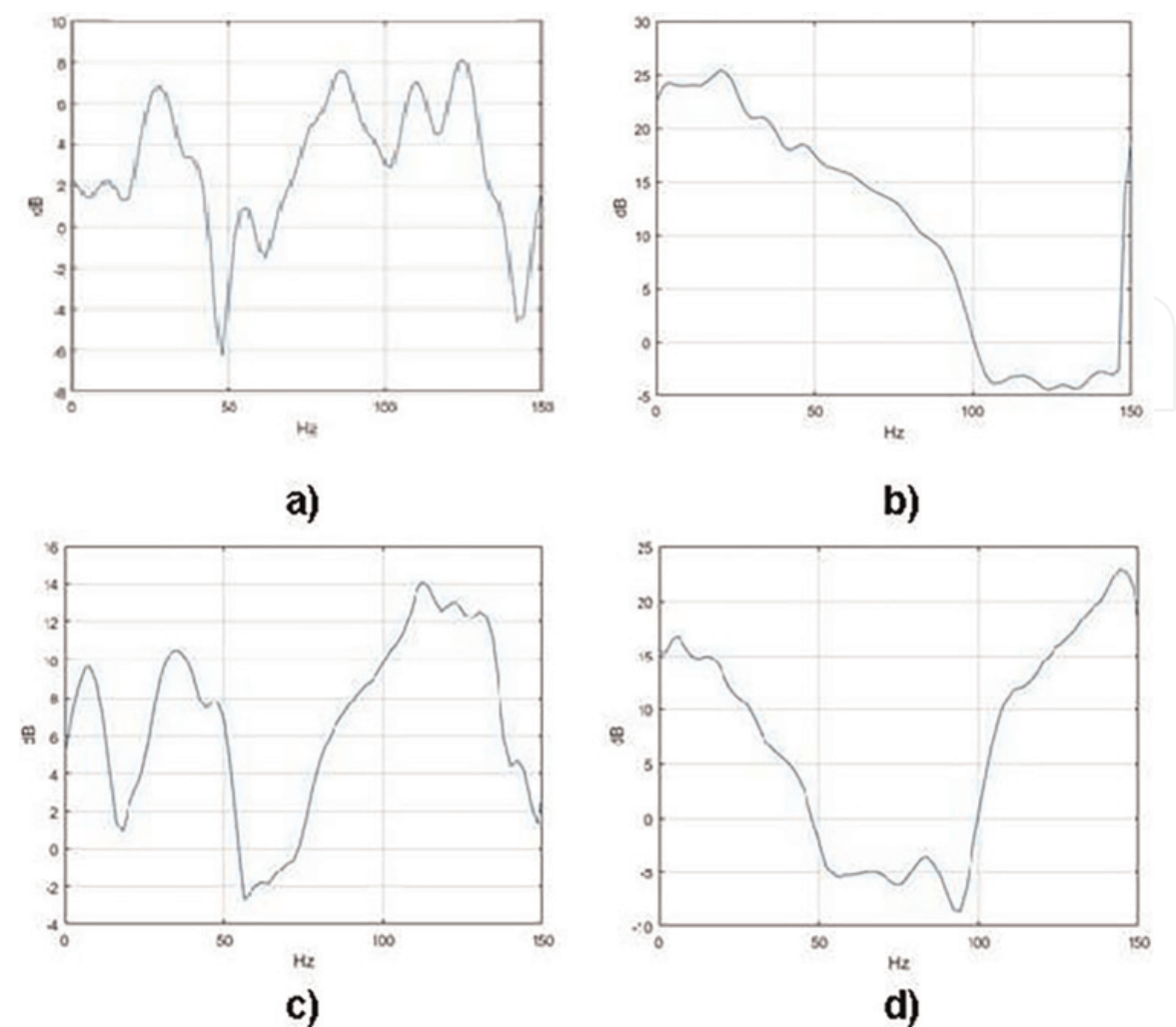


Figure 9.
Different max-pooling frequency responses behaves like (a) an irregular filter/amplifier, (b) low-pass filter/amplifier, (c) multiband filter/amplifier, (d) stopband filter/amplifier.

max-pooling minimizes some frequency values, its main action is to maximizes some signal harmonics. This behavior consistent but the response is different for every signal, there are however, signals that produces poor responses too and that is a disadvantage.

5.2 Models accuracy

The test results are concentrated in **Tables 2** and **3**. The standard pooling methods are located in the first table and the wavelet methods in the second table. They are organized as follows. There is one method per column and both tables are divided into two sections, both were performed during the training process. The first section groups 20 iteration-experiments, and the second is for the 40 iteration-experiments.

The first part of the results is shown in **Table 2**. The most relevant but expected observation in the first table is that max-pooling has the highest accuracy values. It is known that among the standard methods max-pooling is the state-of-the-art, as previous studies suggest [22].

On the other hand, the mix-pooling method by channels has the worst result on the table. However, when the mixing method is made by 2x2 region, it reaches the second

Number of iterations	Average (%)	Maximum (%)	Mixed (%)	Mixed by region (%)	Stochastic (%)
20	98.3	99.2	98.2	99.1	98.2
40	98.7	99.0	97.9	99.0	98.4

Table 2.
Accuracy results for conventional pooling methods.

Wavelet	Lift. (%)	Lift. + Max. (%)	Lift. + Coeff. (%)	Lift. + Max. + Coeff. (%)	Lift. + Rand. Channel (%)	Lift. + Rand. Region * (%)	Lift. + Max. + Rand. Region (%)
20 Iterations							
Haar	99.2	99.3	99.1	99.1	99.0	99.1	99.3
Daubechies 4	99.1	99.2	99.2	99.1	98.9	99.2	99.2
Daubechies 6	99.3	99.2	99.0	99.1	98.8	99.3	99.2
40 Iterations							
Haar	99.2	99.4	99.0	99.2	99.1	99.3	99.4
Daubechies 4	99.3	99.4	99.1	99.3	98.9	99.3	99.2
Daubechies 6	99.2	99.2	99.1	99.2	99.2	99.3	99.3
*Proposed model.							

Table 3.
Accuracy results for the wavelet pooling methods.

highest accuracy levels, just after max-pooling. This points out that processing small sections of an image keeps details better than taking it as a whole.

In **Table 3**, each section uses three wavelet functions with different vanishing moments: Haar, Daubechies 4 and Daubechies 6. In addition, there is one wavelet functions per row and the proposed model is indicated by an asterisk.

From this second part of the results, an improvement from standard to wavelet methods can be noticed. Even for max-pooling, it goes from 99.2% as stand-alone version to 99.4% as a hybrid method. If we observe wavelets methods without max-pooling combination, the proposed method using random selection by 2x2 region is the most consistent and it reaches the second highest accuracy (99.3%). In a previous study [22] this dataset had the best results for max-pooling methods; therefore, it is expected to have two hybrid version with max-pooling with the highest accuracy (99.4%). As explained before, this may change drastically using different datasets since max-pooling is signal dependable, on the contrary the proposed method reaches the second highest value, but consistently and invariant to signal effects.

It is also interesting that the wavelet model with the lowest accuracy is the Lifting Scheme with all the coefficients fixed. Even most of the standard methods get better results than this approach. In these models all the filters are always present and there is no changing effect at all. All the model behavior is expected and predictable, and therefore there is not any non-linearity that helps to prevent overfitting and selects different attributes in every iteration.

6. Conclusion

In this study, max-pooling frequency response is analyzed, and instead of a sliding or changing filter, the results shown a dynamic harmonic amplifier behavior. This phenomenon helps max-pooling to prevent the CNN from overfitting. This keeps the CNN learning rather than memorizing the dataset values, and thus achieving higher accuracy results.

Most of the wavelet models achieve high results, which point out that MRA can improve CNN pooling performance, especially when using hybrid models. However, by analyzing the results of the wavelet model that uses all the coefficients permanently, it is evident that having a nonlinear selecting mode, like a random model, to extract features in every iteration is essential.

In summary, contrarily to max-pooling, not been dependable on the signal shape gives more consistent results. In addition, the mix pooling methods exhibit that processing small regions produced better results than the opposite. All these factors are important to improve CNN performance. The model proposed takes the best part of both issues and moreover, it also takes advantages of having wavelet functions with different vanishing moments.

Finally, implementing two additional wavelet functions based on the lifting scheme shows its feasibility and enrich the model for future applications. Future works include the use of this model in harder imaging like infrared, ultraviolet or satellite imaging.

Acknowledgements

This work was financially supported by the Mexican National Council of Science and Technology (CONACyT) and the Universidad de las Americas Puebla (UDLAP), Mexico.

Conflict of interest

The author confirm that this article content has no conflicts of interest.


IntechOpen

Author details

Daniel Trevino-Sanchez* and Vicente Alarcon-Aquino
Department of Computing, Electronics and Mechatronics, Universidad de las
Americas Puebla, Puebla, México

*Address all correspondence to: daniel.trevinosz@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. Distributed under the terms of the Creative Commons Attribution - NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited. 

References

- [1] Acremont A, Fablet R, Baussard A, Quin G. CNN-based target recognition and identification for infrared imaging in defense systems. *Sensors*. 2019;**19**(9): 2040
- [2] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2016. pp. 770-778
- [3] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2015. pp. 1-9
- [4] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Re-thinking the inception architecture for computer vision. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2016. pp. 2818-2826
- [5] Farah Malik AE, Barin S, Yüksel ME. Accurate classification of heart sound signals for cardiovascular disease diagnosis by wavelet analysis and convolutional neural network: Preliminary Results. In: *28th Signal Processing and Communications Applications Conference*. 2020. pp. 1-4
- [6] Li F, Liu M, Zhao Y, Kong L, Dong L, Liu X, et al. Feature extraction and classification of heart sound using 1D convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*. 2019;**59**:1-11
- [7] Haque MR, Mishu SZ. Spectral-spatial feature extraction using PCA and multi-scale deep convolutional neural network for hyperspectral image classification. In: *22nd International Conference on Computer and Information Technology*. 2019. pp. 1-6
- [8] Rao CP, Reddy AG, Rama Rao CB. Target detection using multi-resolution analysis for camouflaged images. *International Journal on Cybernetics & Informatics*. 2016;**5**(4):135-147
- [9] Williams T, Li R. Wavelet pooling for convolutional neural networks. In: *International Conference on Learning Representations*. 2018
- [10] Ferrà A, Aguilar E, Radeva P. Multiple wavelet pooling for CNNs. In: *European Conference on Computer Vision 2018 Workshops Lecture Notes in Computer Science*. 2019. pp. 671-675
- [11] Liu P, Zhang H, Zhang K, Lin L, Zuo W. Multi-level wavelet-CNN for image restoration. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2018. pp. 886-895
- [12] Liu P, Zhang H, Lian W, Zuo W. Multi-level wavelet convolutional neural networks. *IEEE Access*. 2019;**7**: 74973-74985
- [13] Fujieda S, Takayama K, Hachisuka T. Wavelet convolutional neural networks for texture classification. In: *Computing Research Repository*. 2017
- [14] Fujieda S, Takayama K, Hachisuka T. Wavelet convolutional neural networks. In: *Computing Research Repository*. 2018
- [15] Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*. 1990;**36**(5): 961-1005
- [16] Daubechies I. Ten lectures on wavelets. In: *Society for Industrial and Applied Mathematics*. 1992

- [17] Oyallon E, Belilovsky E, Zagoruyko S. Scaling the scattering transform: Deep hybrid networks. In: IEEE International Conference on Computer Vision. 2017. pp. 5619-5628
- [18] Aslam A, Hayat K, Iqbal-Umar A, Zohuri B, Zarkesh-Ha P, Modisette D, et al. Wavelet-based convolutional neural networks for gender classification. *Journal of Electronic Imaging*. 2019;**28**(1):01301201
- [19] Williams T, Li R. An ensemble of convolutional neural networks using wavelets for image classification. *Journal of Software Engineering and Applications*. 2018;**11**:69-88
- [20] Huang H, He R, Sun Z, Tan T. Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution. In: IEEE International Conference on Computer Vision. 2017. pp. 1698-1706
- [21] Ma W, Pan Z, Guo J, Lei B. Achieving super-resolution remote sensing images via the wavelet transform combined with the recursive Res-Net. *IEEE Transactions on Geoscience and Remote Sensing*. 2019;**57**(6):3512-3527
- [22] Trevino-Sánchez D, Alarcón-Aquino V. Hybrid pooling with wavelets for convolutional neural networks. *Journal of Intelligent & Fuzzy Systems*. 2021; **2021**:1-10
- [23] Rossetto AM, Zhou W. Improving classification with CNNs using wavelet pooling with nesterov-accelerated Adam. In: Proceedings of 11th International Conference on Bioinformatics and Computational Biology. Vol. 60. 2019. pp. 84-93
- [24] Ma H, Liu D, Xiong R, Wu F. iWave: CNN-based wavelet-like transform for image compression. *IEEE Transactions on Multimedia*. 2020;**22**(7):1667-1679
- [25] Bastidas-Rodriguez MX, Gruson A, Polanía LF, Fujieda S, Prieto-Ortiz F, Takayama K, et al. Deep adaptive wavelet network. In: IEEE Winter Conference on Applications of Computer Vision. 2020. pp. 3100-3108
- [26] Williams T, Li R. Advanced image classification using wavelets and convolutional neural networks. In: 15th IEEE International Conference on Machine Learning and Applications. 2016. pp. 233-239
- [27] Liu JW, Zuo FL, Guo YX, Li TY, Chen JM. Research on improved wavelet convolutional wavelet neural networks. *Applied Intelligence*. 2020;**51**(6): 4106-4126
- [28] Stepanov AB. Construction of activation functions for wavelet neural networks. In: XX IEEE International Conference on Soft Computing and Measurements. 2017. pp. 397-399
- [29] Azeem MF, Banakar A, Kumar V. Comparative study of different types of wavelet functions in neural network. In: IEEE International Joint Conference on Neural Network Proceedings. 2006. pp. 1061-1066
- [30] Sharan RV. Spoken digit recognition using wavelet scalogram and convolutional neural networks. In: IEEE Recent Advances in Intelligent Computational Systems. 2020. pp. 101-105
- [31] Xu B, Wang N, Chen T, Li M. Empirical evaluation of rectified activations in convolutional network. In: Computing Research Repository. 2015
- [32] Sweldens W. The lifting scheme: A construction of second-generation wavelets. *SIAM Journal on Mathematical Analysis*. 1998;**29**(2):511-546

- [33] Mallat SG. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989;**11**(7):674-693
- [34] Akansu AN, Haddad RA. *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*. 2nd ed. San Diego, USA: Academic Press; 2001. p. 499
- [35] la Cour-Harbo A, Jensen A. *Wavelets and the Lifting Scheme*. Aalborg University; 2007. pp. 1-44
- [36] Sole J, Salembier P. Generalized lifting prediction optimization applied to lossless image compression. *IEEE Signal Processing Letters*. 2007;**14**(10): 695-698
- [37] Daubechies I, Sweldens W. Factoring wavelet transforms into lifting steps. *The Journal of Fourier Analysis and Applications*. 1998;**4**(3):247-269
- [38] Nielsen MA. *Neural Networks and Deep Learning*. Determination Press; 2015
- [39] Lee CY, Gallagher PW, Tu Z. Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2018;**40**(4):863-875
- [40] Yu D, Wang H, Chen P, Wei Z. Mixed pooling for convolutional neural networks, rough sets and knowledge technology. In: *Lecture Notes in Computer Science*. 2014. pp. 364-375
- [41] Zeiler M, Fergus R. Stochastic pooling for regularization of deep convolutional neural networks. In: *1st International Conference on Learning Representations*. 2013
- [42] LeCun Y, Cortes C. *MNIST handwritten digit database*. 2010

Chapter

Autoethnographical Contributions in the Construction of Feminist Knowledge since the Lockdown

Alejandra Acevedo Placeres

Abstract

This document presents the autoethnographic work and the data collection techniques used in my doctoral thesis as tools that have served to explore the connection between the individual and the social, as well as to validate the evocative narrative as a methodological tool in the construction of a new feminist knowledge from prison and shows my experiences lived in the San Miguel prison (CERESO) in the city of Puebla, Mexico.

Keywords: autoethnography, prisons, confinement, qualitative research, evocative autoethnography, methodology

1. Introduction

Autoethnography has been used as a qualitative research method through which the researcher describes his or her own experience for a sociological understanding of the phenomena studied. In recent years autoethnography has begun to be considered not only as a research method, but also as a “genre of writing, research, history and method that connects the autobiographical and personal with the cultural, social and political” ([1], p. xix), as Carolyn Ellis points out. This text focuses on the description of the use of autoethnography in my experience as an intern, social manager and researcher at the CERESO of San Miguel. The purpose is to present the autoethnographic work and the data collection techniques used in the doctoral research that have served to validate the evocative narrative [2] as a methodological tool in the construction of a new feminist knowledge confinement in prison and also to explore the connection between the individual and the collective.

2. Methodology

The doctoral thesis entitled “Subjectivities and sorority in the CERESO de San Miguel. An autoethnographic story” is narrated in my own way of embodying the experience of confinement, in turn it is also told from the bodies and experiences of other women with similar experiences in the same context of confinement. The

analysis that is presented is directed from phenomenology to show the ways in which women have been subjectivated and the processes of subjectivation that have turned us into prisoners and in this specific context into Persons Deprived of Liberty (PPL¹ -acronym in Spanish-), what is sought with these stories is that they contribute to revealing the existence and how the sorority relationships between the internal women within the CERESO of San Miguel take place.

As this research focuses on personal experiences, one of the interpretative and referential frameworks through which the approach to the social experience is made is phenomenology. The analysis is aimed at showing that subjective reality and social reality are intimately related. Sorority relationships of the women inside CERESO are also inscribed, in order to study how the inmates experience the processes of subjectivation from their own bodies and the sorority relationships that emanate from them.

The work analyzes the experience from four key elements of phenomenology pointed out by Juan Luis Álvarez-Gayou ([3]: 85): temporality, spatiality, corporeality and relationality. In this sense, we explore how the inmates feel the “world” and not how they think about it, that is, how the time lived in confinement/temporality is interrelated with spaces/spatiality and bodies/corporeality and how this affects or impacts the creation of sorority/relationality relations. Thus, the lived experience in seclusion is contextualized by a relationship with objects, people, events and situations.

3. Autoethnographical contributions

I began my research work without thinking of autoethnography as a methodology or interpretative frame of reference. I was in the first months of my doctoral studies and I was not sure how to talk about my experience. Nor did I know how my story could contribute “scientifically valid” knowledge to science. What was clear to me was the need to interpret my experience and the life stories of the women interns I had met, stories to which I was also a witness. I wanted the thesis to be a dialog, a conversation with the women inmates to unveil how sorority relationships are built inside ([4], cit. Bénard 20-25). Thus, while planning and designing the methodology and data collection methods, I began to feel an uneasiness and uncertainty about the role I played in this research as part of the social phenomenon I was analyzing. In analytical terms, an important element that gives important nuance to the thesis is my participation in the phenomemo, which brought me closer to a different epistemological perspective to understand this lived experience, and the way in which we can make sense of the experience from introspection and writing as qualitative research tools ([5], pp. 3-4).

The thesis starts off with my own testimony, narrating my own experience. This translates, in the words of Elizabeth Aguirre - Armendariz [6] to “telling the thesis”. In turn, the becomes the starting point to make a first approach to the field of study from the own introspection. The incorporation of autoethnography as a methodology has also allowed me to position myself as an “active agent with narrative authority” within the research:

Performing autoethnography has allowed me to position myself as active agent with narrative authority over many hegemonizing dominant cultural myths that

¹ PPL is the acronym in spanish for Person Deprived of Liberty, and it is the way in which the inmates are named in this penitentiary

restricted my social freedom and personal development, also causing me to realize how my Whiteness and class membership can restrict the social freedom and personal development of others [7].

As Tami Spry puts it, sharing an experience from one's position as a researcher inspires readers to reflect critically on their own life experience. Under this assumption, throughout the thesis I describe some of my personal experiences in the CERESO of San Miguel as an intern, to produce an autoethnographic work that consists of the construction of the narrative between the lived, thoughts and feelings. Tami Spry says: "Autoethnographic performance makes us acutely aware of how we 'I-witness' our own reality constructions." ([7], p. 707). In this sense, the autoethnographic work, from my position as an inmate, helps me to develop a reconstruction of some practices that are generated in the CERESO that serves as a bridge and encounter between those who read us and the women who speak here about part of our life and daily life in this context of confinement.

Between the experiences and the theoretical plane that was presented to me, it was complicated for me, not only to establish a starting point, but also to define the position from which I was going to develop the research work ([8], pp. 115-119). As I progressed with my fieldwork, the review of information I had collected and its sources allowed me to find the path to follow. It was evident that I was not only elaborating an academic work; at this point I was already telling my own life story. In this way, I began to outline a multi-positionality in the development of the research.

This multiple position is presented throughout the work from three different perspectives: as a researcher, as a leader of a social management project and as an intern. This multipositionality in which I find myself is distinguished throughout my research. Is not just reflected in the stories and descriptions, but at different times it is presented as a distinct writing. It has not been easy. The writing exercise required a mental and emotional effort to allow me to find the balance between academics and my own passions.

In the role of researcher, I elaborated qualitative research of an inductive nature. This approach does not intend to develop concepts and understandings from the collection of data as a way to evaluate models or test hypotheses. On the contrary, the purpose is to understand the lived experience and at the same time to understand how other people live it, which allows flexibility both in data collection and in data analysis ([9], pp. 217-238). Thus, the questions about what is happening in the women's section, what subjectivities "know" in order to transform their relationships into sorority strategies? It was not possible to remain only in the descriptive part of the interviews, it was necessary to ask questions from a different position, to ask much more concrete and direct questions. That is, to go deeper with additional questions about the meaning of their experiences for inmates, through a prolonged group study in which, in addition to participant observation, inmates had other sources of information. This is why I designed, as part of the procedure to obtain qualitative data, four sessions of group memory workshops where interns participated as part of my focus group.

Additionally, another source of data collection that enriches the research and data collection was my interaction in the penitentiary center as a social manager. Since 2005, together with my sister Ana, we founded a civil association with the purpose of promoting projects based on playful methodologies as a tool to develop socioemotional skills. This contributed in the formation of a positive personality that provides children with maximum development potential in health, growth and learning capacities to transform not only the lives of these children, but also of the community,

our state and the country. This program has given me access to the female section of San Miguel CERESO on a permanent basis. The actions and activities that we have carried out allow me to have a much more intimate approach with the inmates and their children. My constant presence in the female section has facilitated the elements to carry out a participant observation. This interaction and participation undoubtedly links me to the situation I am observing. It can be said that as a researcher I have become a member of the group to be studied. From this positionality, the challenge has been to set aside my own beliefs, perspectives and predispositions so that I do not try to take some circumstances for granted. At the same time, the information gathered through participant observation and interaction with inmates, custodial staff, family members and volunteers, is part of the corpus that is being analyzed in the doctoral research.

Unfortunately, this entire work plan was interrupted in 2020. The first work session with the focus group had barely been carried out, when the governor of the state of Puebla decreed that, starting in April, visits to prisons would be suspended as a control measure to prevent the spread of COVID-19 among inmates. As a society, we were facing a pandemic of unimaginable dimensions that would obviously have implications in my research and fieldwork.

While I was organizing my ideas and rethinking my thesis work in the face of the control measures adopted by the pandemic, I discovered that I was part of the experience, and that, if my interest was to contribute new perspectives to feminist knowledge, the multipositional nature in which I found myself was already a source of knowledge. In addition, although I could not enter the prison as a researcher, I was allowed to enter as a social manager. In this way, the methodology of the thesis started to shape up more towards an autoethnography, but the thread of the narrative itself was still missing. The question at this point was: what was part of my story and experience that I had yet to narrate? It is here where my position as an inmate at CERESO appeared as the missing piece to complete the narrative I was elaborating. This provided me with the opportunity to transcend from the individual to the social.

Autoethnography has been a methodological tool and the method of analysis of social practices and dynamics within CERESO because “autoethnographies are highly personalized, revealing texts in which authors tell stories about their own lived experience, relating the personal to the cultural” [10], then the story of my life as an autoethnographic narration becomes a revealing text that tells us this:

In March 2015, I was arrested for an alleged crime of extortion committed by a public servant. I was detained and consigned to a prison as a precautionary measure. I was deprived of my liberty in the female section of CERESO San Miguel during 3 months and 3 weeks. There, I lived the prison experience as an inmate. The situation was in the midst of a political scenario. I was accused and imprisoned for a crime I did not commit. My stay transgressed each of the spheres of my life: professional, family, labor, social and economic. It disturbed my economic and emotional stability and, why deny it, even my mental health. In this context, talking about my experiences in confinement was not only painful and confusing, but it also seemed a bit self-centered and narcissistic because I felt that in this place there were other life stories to tell. On the other hand, perhaps another reason for not writing about it was not wanting to acknowledge my own history as an inmate. Avoiding it and not writing about it was in one way or another the way to run away from my memories. When I finally decided to speak about my experience and use it as a thread of conversation with the interns, I found not only meaning, but the richness of lived experience and a way to tell our stories.

The pandemic and the suspension of my fieldwork were important events because I realized that my research did not depend entirely on ethnographic sessions. Instead, I understood that it was work, which had started the day I entered the CERESO as an intern. It continued constantly until today, encouraged by the social projects that I had managed recently. I have always been doing fieldwork, and I was also part of the social phenomenon that I was exploring. Thus, all personal experiences are valuable data that helped me to write the following chapters.

4. Results

The richness of autoethnography allows me to speak from my own experience, under the understanding that Carolyn Ellis, Tony E. Adams and Arthur P. Bochner (cit. in Bénard 23) have of it. These authors claim that autoethnography produces texts that are artistic and evocative of personal experiences, which is why these texts are often criticized for being too artistic and not scientific, or vice versa. Autoethnography is sometimes still rejected for not being rigorous, theoretical and analytical enough or for being too emotional, esthetic or therapeutic (quoted in Bénard 23). Yet, in confinement environments, and in the specific case of my personal experience, it becomes a useful tool to better understand the processes of subjectification and sorority relations that are being explored as social phenomena. In this context, and in an exercise of full awareness of the implications it would have on the process and results of the research (Alvarez-Gayou 23–28), I decided not only to incorporate autoethnography as a methodological tool, but also to give it a specific weight and recognize the richness of my own experience as part of the process and results of the research.

As was to be expected, the appearance of the pandemic and the resulting confinement and social distancing that we experienced had implications for my research work. These became relevant challenges in the research process. One of the challenges was facing the suspension of activities in my fieldwork. The governor's decree that prevented access to CERESO caused a considerable delay in my thesis, which forced me to focus on the search for alternative sources of information and to go deeper in reviewing the field journal entries. During this process, I discovered that fieldwork was also writing. Writing about my experience, writing about the lives of other inmates and writing about our processes is part of the fieldwork when it comes to working with autoethnography as a source of information. Although I was going through autoethnography, ethnography was also present, since the voice of other inmates telling their own stories often intervened in the narrative. In the midst of the health crisis that impacted my research, I was faced with another challenge: how to make autoethnography and ethnography coexist at the same time, without losing myself in their limits? In this back-and-forth of ideas, I elaborated what would be the first sketches of an autoethnographic account, and thus I understood that autoethnography is a process and a product.

Another challenge I faced was to find my own writing style, using autoethnography as a methodology was a real personal and academic challenge. From the moment I started the research I asked myself: How was it possible that talking about my personal life could contribute something to scientific knowledge? How would I talk about myself, about my own experience at CERESO San Miguel? Would an autoethnographic thesis have only an exploratory scope? I had been adamant about not wanting to talk about myself. I wanted to talk about the sorority relationships that take place in the female section of the CERESO of San Miguel. I wanted the inmates to

talk. I wanted to have a conversation with them. At some point, even the name of the thesis was related to the topic of giving them a voice.

Faced with this situation and the question of how to make sense of my own experiences and at the same time include those of others? I had to do a lot of further reading and with that I got closer to several works about women in prisons. Of course, I read the life stories of the inmates that have been compiled in books and anthologies such as the book *Bajo la sombra de un Hamuchil* (2017) the result of a writing workshop at the CERESO of Morelos and the books formed by the compilation of texts that DEMAC and the Ministry of Public Security have published under the name of *Bajo Condena* (2003). But my purpose was not to tell one or more life stories. My purpose now was to tell my story, my experience, and in turn to dialog with other women about our experiences embodied in our bodies and reflected in sorority relationships.

Faced with this situation and the question of how to make sense of my own experiences and at the same time include those of others, I had to do further reading to know more about women in prisons. Of course, I read the life stories of the inmates that have been compiled in books and anthologies, such as the book *Bajo la sombra de un Hamuchil* (2017), the result of a writing workshop at the CERESO in the state of Morelos, and the compilations edited and published by Documentation and Women's Studies A.C. and the Ministry of Public Security of México under the title: *Bajo Condena* (2003). But my purpose was not to tell one or more life stories. My purpose now was to tell my story, my experience, and in turn to engage in dialogs with other women about our experiences embodied in our bodies and reflected in sorority relationships.

To employ a writing style that incorporated a genre of writing and an uncommon research method such as autoethnography, also met the expectations of an academic paper. I was still trapped in my own questions: what to tell, how to tell it? But above all, how to write it? The most difficult thing for me during my studies has been writing. I had to find my own style of writing that the academy would accept and understand; a writing that gave me access, as a kind of authorization, to write from my experiences and emotions. I had to write an autoethnographic account with academic validity.

By reviewing Elizabeth Aguirre-Armendáriz's [6] story about her thesis work I felt much more confident that I too "could transit through autoethnography and with autoethnography" (47). I found in her a true ally for my research work. Aguirre-Armendáriz's research approach began with the question: from where and how far should I go back, both in my exploration and in my narrative, to understand and show its scope, its possibilities and its limits? Through these readings I discovered that autoethnography can be used as a research method, as a method of analysis and as a genre of writing. This helped me to better structure my thesis work, but, above all, it made better sense of my ideas. Now, I have firmly decided to use autoethnography as a methodology, method and way of writing.

Autoethnography is used in the thesis to express the results in a more natural language, with a freer style of writing, to raise my voice, to make myself heard. Finally, I speak with my fellow interns and with you, the readers. After all this process, I feel satisfied and convinced of the methodological contribution of my research work. The road has been long, but discovering that through autoethnography it is possible to generate meaningful, accessible and evocative knowledge to transform and provoke social change has been a revelation [2]. Transforming my memories into notes and using writing as an experiential method of inquiry to analyze social components of my experience gives shape and congruence to my research work.

5. Conclusions

I am convinced that in my research, autoethnography is not only a useful, but also a challenging tool with which I systematically describe and analyze personal experience to understand cultural experience. I am confident that my autoethnographic account, going beyond a descriptive narrative, has become a critical reflection through which instead of looking only at “them”, I can look at myself and thereby make knowledge from the place where I dwell. Like Tamy Spry [7] did with her autoethnographic work, I hope that my research clearly expresses the interactive textures that occur between my experience and that of the other inmates in the context of institutionalized confinement in a penitentiary. At the same time, this approach fulfills the objective of unveiling an unknown reality, finding the theoretical link to reflect on the process of creating sorority relationships between these women.

I have developed this research from a favored position that allows me to construct and reconstruct discourses and narratives from different perspectives. Firstly, from my experiences lived as an inmate in the San Miguel penitentiary in Puebla, Puebla in 2015. Then, as a leader of a project of management and support to inmates (2015–2020) and now as a graduate research student since 2018. These experiences and perspectives have yielded copious data and information that has been used to produce a vast autoethnographic and ethnographic work that constitutes the innovative part of the research. It generates a dialog between the narratives of the lived and my thoughts and feelings with data and findings collected with ethnographic techniques, such as life histories and participatory workshops.

Finally, I can conclude that an autoethnographic narrative is a genuine companion method that builds a dialog with other qualitative methodologies. It also habilitates spaces for critical and theoretical reflection. The analysis of my own reality can accommodate the intersubjective interpretation of all the realities that converge in my story. My research becomes a space to make ourselves visible not from inequality, but from our own wisdom derived from our learning experiences, our strengths and struggles. I am creating a document in the form of a conversation with our own personal baggage to question and speak about women in confinement, their social constitution and, probably, to break down all the preconceptions we have about them. As a result, life stories and experiences in confinement become visible, generating modes of subjectification and what kind of relationships they put together to generate and strengthen pacts and alliances that translate into sorority [11].

Conflict of interest

The author confirm that this article content has no conflicts of interest.

IntechOpen

Author details


Alejandra Acevedo Placeres^{1,2}

1 Doctorado en Creación y Teorías de la Cultura. Ex Hacienda Sta, Puebla, México

2 Universidad de la Américas Puebla, Ex Hacienda Sta, Puebla, México

*Address all correspondence to: alejandra.acevedops@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Ellis C. The Ethnographic I. A Methodological Novel about Autoethnography. United States of America: Altamira Press; 2004. p. xix
- [2] Ellis C, Adams TE, Bochner AP. Autoethnography: An Overview. Alemania: Forum: Qualitative Social Research. 2010. Available from: <http://nbnresolving.de/urn:nbn:de:0114-fqs1101108>
- [3] Álvarez-Gayou J. Cómo hacer investigación cualitativa. Paidós: Fundamentos y metodología; 2003
- [4] Bérnard S. Autoetnografía: Una metodología cualitativa. Universidad Autónoma de Aguascalientes; 2019
- [5] Ellis C. Final Negotiations: A Story of Love, Loss, and Chronic Illness. Temple University Press; 2018
- [6] Aguirre-Armendáriz E. Cuando contar la tesis es hacer la tesis: investigación y escritura autoetnográfica. Centro Latinoamericano de Pensamiento Crítico; 2015
- [7] Spy T. Performing autoethnography: An embodied methodological praxis. Qualitative Inquiry. 2001;7(6):706-732
- [8] Araiza A. Ciencia, subjetividad y poder: Claves feministas para la construcción de conocimiento. Universidad Autónoma del Estado de Hidalgo; 2017
- [9] Castañeda M. Etnografía Feminista. In: Blázquez N, Flores F, Ríos M, editors. Investigación feminista: epistemología, metodología y representaciones sociales. UNAM, Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades; 2012. pp. 217-238
- [10] Richardson L. Evaluar la etnografía en Autoetnografía. In: Calva B, Autoetnografía SM, editors. Una metodología cualitativa. Mexico: Universidad Autónoma de Aguascalientes; 2019. pp. 183-186
- [11] Toussaint M. Cuéntame tu vida, Historia oral: historias de vida. Secuencia. 2016;95:259-273

Sueño en Otro Idioma: Queer Identities in Contemporary Mexican Melodrama

Ilia Lizbeth Espinosa Pacheco

Abstract

This article examines the configuration of queer identity in the Mexican film *Sueño en otro idioma* (Dream in another language, 2017), directed by Ernesto Contreras, arguing that contemporary Mexican melodrama presents new dynamics related to gender, class, and ethnic identities and is more interested in exploring the personal crisis that modernity has arisen in the country, instead of a fixed idea of mexicaness. When melodrama appeared in Mexican cinema, it reproduced the stereotypes of the heteronormative and patriarchal hegemony that dominated the configuration of the Mexican post-revolutionary nation. Unlike previous films, *Sueño en otro idioma* shows that these dynamics are evolving within Mexican culture, breaking with old representations about indigenous people, homosexuality, and violent masculinity and thus proposing a new, reconfigured notion of diverse Mexican identities. This paper analyzes the foundational characteristics of melodrama still present in the Mexican culture and those related to queer identities that the film has subverted through their representation.

Keywords: identity, Mexican cinema, melodrama, queer identity, gender

1. Introduction

Historically, melodrama and romantic comedy have been the most watched film genres in Mexico. In the beginning, the film industry contributed to the post-revolutionary national project to create an idea of a united, developing country. The melodrama and ranchera comedy produced in the Golden Age era (1936–1956) displayed a series of stereotypes related to an inherited European patriarchal culture and, according to Koper [1], to the notion of *mestizaje* that followed an official cultural policy depicting a “raceless” nation where the indigenous was apparently valued. These movie genres showed hegemonic stereotypes such as the macho, the pure and virginal woman, and the *Indio*, which were seen as “idealized identities” (p. 100) that reinforced the inequitable power relations and the discriminatory effects of this policy. Later, with a national transformation driven by modernity, these stereotypes evolved accordingly with changing economic, political, and social contexts, but still guided by a heteronormed society where class and gender inequalities remained, and indigenous people were reduced or made invisible.

Therefore, when indigenous people appear on the screen, they do not have a leading role or agency; their identity conflicts are not represented. They are conceived in a one dimensional fashion: precisely as “indigenous”, “poor people” only ([2], p. 102), or simply as “heterosexuals”. They are generally not portrayed as human beings with multiple dimensions, including various sexual identities. The practices and behavior of indigenous people represented on screen are defined by the religious syncretism that is inherited from the Spanish Conquest, and in which the cosmogony of their origins has become veiled. They are often presented as victims, forced to assimilate *mestizaje* or related to a class condition defined by a rurality that is perceived as backward. Most are minor characters, like farmers, domestic workers, healers and midwives; confident and hardworking people, but not very intelligent ([3], p. 284). Fiction films like *La Llovizna* (1977), *Ley de Herodes* (1999), *Piedras Verdes* (2001), *Roma* (2018), *El ombligo de Guie'dani* (2018), and *Nuevo orden* (2020) are good examples of this trend.

In contrast, *Sueño en otro idioma* (2017) is a melodrama that presents different characteristics. The film breaks away from class, race, gender, and ethnicity stereotypes that represent an established series of ideas about being Mexican. The story takes place in a small town defined by the encounter between the ancient culture of *Zikril* and worldwide neoliberalism. The film depicts a contemporary indigenous homosexual couple, like the metonymy of a nation where pre-Hispanic cultures, *mestizaje*, the social transformations of the Mexican Revolution, and the current crisis of modernity meet, showing another facet of Mexican identity. The characters act and make decisions, they are the protagonists, they are not victims or reduced to class differences, their inner world and their sexuality is explored, and although syncretism is present, their pre-Hispanic past is not veiled. The protagonists fight to be part of the national imaginary, in order to transform the stereotypical system that has long organized the idea of mexicaness.

2. The construction of a National Mexican Identity

For Stuart Hall, identity is the sense of a common world that is shared between the members of a group through “systems of representation” ([4], p. 4). These systems of representation are cognitive mechanisms that let people situate in the world and react to it. Hall explains that these are “signifying practices”, signs and symbols (sounds, language, written words, electronically produced images, musical notes, objects) “that stand for or represent, in our mental life, things which are or may be out there in the world” (p. 4) and that people use to signify concepts, ideas, and feelings and to communicate and exchange meanings (p. 5). According to him, the meanings shared by a group help organize and regulate daily social practices, what is accepted and what must be rejected, since “they help to set the rules, norms, and conventions by which social life is ordered and governed” (p. 5). This is how their culture is shaped. The exchange of meanings in a community creates a sense of identity, “a sense of who we are and with whom we belong” (p. 3), a sense of fitting based on the common.

However, identity is not static or uniform, it changes and evolves with the vicissitudes of societies and the exchanges that exist with the “others”. For Hall and Du Gay [5], identity is related with the sense of identification, with “a construction, a never-finished process: always in ‘process’. It is not determined [because it] is always possible to ‘win’ or ‘lose’ it, sustain it or abandon it” (p. 15, my translation). So, identity “is an articulation process”, where identification occurs “to a greater or lesser extent”

but “never in totality” (p. 15, my translation). For the authors, it works by means of difference. It points out the limits that separate us from others, and, although it recognizes “stable” elements that unify a group, identities react to context and otherness, so that they “are fragmented and are never singular” (p. 17). Hall and Du Gay establish that identities are “constructed in many ways through different discourses, practices and positions, often crossed and antagonistic. They are subject to radical historicization, and in a constant process of change and transformation” ([5], pp. 16–17, my translation). It means that we can find a historical time that has defined our identity with an idea close to reality but, at the same time, has experienced different transformations related to all the systems of representation that configure cultures in time, just like it has been revealed in cinema.

In Mexico, cinema has contributed to forging the idea of Nation, a system of representation that has established a series of conceptions about what is mexicaness. Most of them had been determined by patriarchal, heteronormative, and racist codes of behavior inherited from a sociopolitical process of national transformation through the Conquest, the Revolution, and the arrival of modern capitalism. Those struggles have also been fought in the realm of representation, where all the social sectors (men, women, indigenous people, *mestizos*, rich, poor, queer) and their subjectivities have been looking forward to being part of the national imaginary, in order to transform the stereotypical system that has long organized the idea of being Mexican. De la Garza [6] explores the changes that national Mexican identity across time: “In 1810, the main external other was ‘Spain’”, “in 1847 this other became ‘the United States’” (p. 414). With the revolution, “Mexicans were all *mestizos*” (p. 416) and the others became the *criollos*, remaining “indigenous minority” (p. 415). Later, with the “regional economic integration” (p. 419) promoted by the North American Free Trade Area, narratives were then related to the idea of “partnership and complementarity” (p. 419). Today, most conflicts are focused on the city, economic and political crises, drug trafficking and understanding new queer identities.

For Koper [1], through siècle XX, the idea of Mexican national identity was grounded “in racial and class-based notions of ‘Mexicanity’” (p. 98). She explains that the public policies implemented by the state after the Revolution “reinforcing cultural and racial whitening of the Mexican nation” (p. 98) in a kind of “internal colonialism” (p. 98) that intended to make all indigenous communities uniform under the concept of *mestizaje*. The result of these policies was the apparent incorporation of the indigenous with society, but at the same time, their exclusion. What was valued about indigenous people was the past: narratives that praised an imaginary pre-Columbian culture, but that excluded its legal and current conditions of existence. The concept of *indigenismo* was a myth, an attempt to build a homogeneous nation where “their concern over the pluri-ethnic composition of the fractured nation, reshaped the image of the indigenous in order to forge one national identity” ([1], p. 100). The ideological idea of *mestizaje* hid racism at the discursive level, and *indigenista* policies categorize indigenous people as “passive recipients of governmental aid” (p. 102), associating them with poverty and in need of education to become better integrated with a developed society.

Film productions portraying indigenous people supported government policies that “glorify pre-Columbian heritage [...] historical epics depicting the Spanish conquest and *mestizaje*” ([1], p. 104). Also, the great number of movies of the *comedia ranchera* genre produced during the Golden Age of Mexican cinema embraced religious syncretism, *mestizo* identity, and “valorized Mexican pueblo as joyful yet coarse drunks” (p. 110). In contrast, Koper claims that the cinema of Emilio “El

Indio” Fernández exalted an indigenous heritage and showed the marginal side of government policies. Nevertheless, his stories reproduced the patriarchal regime, the importance of the Mexican man and the submission of women. Besides, the film director was criticized because his casting choices reproduced *mestizo* identity over “indigenous somatic features” ([1], p. 109). The characters in Fernández’s films work in oppositional concepts: “nature vs civilization” (p. 108), victims vs. perpetrators, naïve vs. abusive, pure vs. impure, corrupted vs. uncorrupted; they all represent the stereotypical identities discussed so far.

For Hall and Du Gay [5], the basis of identities is “difference”. They claim, based under the readings of Derrida, Laclau and Butler, that “the radically disturbing admission that the ‘positive’ meaning of any term -and with it its ‘identity’- can only be constructed through the relationship with the other, the relationship with what he is not, with precisely what he lacks, with what has been called its constitutive outside” (p. 18, my translation). Hall argues that this is related to cultural meanings and how these are frequently organized as binaries or opposites, when one of them is considered abnormal, wrong, or unacceptable. Once the “normal” has been “established or fixed”, all that falls around it is dismissed or rejected. This “difference” is represented as “the other”, and to be reduced into this relation of opposed meanings, gives rise to stereotypes. People who are in any way significantly different from the majority are frequently exposed to this binary-opposite form of representation ([4], pp. 225–229). Thus, racial, ethnic, sexual, class and disabilities foreground the dimension of difference. Hall also claims that “the whole repertoire of imaginary and visual effects through which ‘difference’ is represented at one historical moment is a ‘regime of representation’” ([4], p. 232). In it, difference is significant in a very simplistic way, losing the richness that this difference itself entails, failing to capture the diversity of the world. There is always a power relationship in signifying practices because it is the dominant representation system which imposes a classificatory system over the “others”.

However, the meanings that we share in culture are not “straightforward, rational or instrumental” or finally fixed because “they mobilize powerful feelings and emotions of both a positive and negative kind” ([4], p. 10). Interpretation depends on the context and on personal history in a community. Concepts and categories that classify the elements of reality for a mutual understanding are changing (Types), and just as types change, stereotypes do too, despite the power relationships that try to maintain a fixed meaning that allows classifying, evaluating, and discriminating difference. Regimes of representation are constantly undermined, “as representations interact with one another, substituting for each other, displacing one another along unending chain” ([4], p. 10). Meanings can be struggled over, challenging, negotiating and transforming the very cultural identities in resonance with new contexts and situations. This is the reason why identity allows a general cultural recognition, but, at the same time, accepts changes with time.

Sueño en otro idioma presents, in Mexican culture, a fragmented identity. The first part of this identity are the indigenous *Zikril* people. Only three of them are still alive: Jacinta (Mónica Miguel), Evaristo (Eligio Meléndez/Juan Pablo de Santiago), and Isauro (José Manuel Poncelis/Hoze Meléndez). They still speak their ancient language. In their cosmogony, *zikriles* can speak with animals and understand the world. The second fragment of identity in the film are the *mestizos*, descendants of both *zikriles* and Spaniards who established themselves in this region long ago. All of them had forgotten their indigenous roots and, facing an economic crisis, are trying to learn English and travel to the United States in order to have a better life. This group

is represented by Lluvia (Fátima Molina) and her friends. The third identity are the *criollos*, who represent the purest heritage of the Spaniards, hold power and symbolize the progress of the country. The *criollos* in the film are Martín (Fernando Álvarez Reveil), a young linguist that has come to the village with the intention of preserving the *Zikril* language, and María (Nicolasa Ortiz Monasterio), a woman that forces the main characters to confront their past and their present.

The film represents both types and stereotypes in relation to indigenous people, gender roles and even professional practices. The western, educated, modern, and superior white man is represented by Martín and his honorable aspiration of saving the language with an apparently moral superiority. Jacinta is the wise, old, indigenous woman dressed in traditional attire who embodies *Zikril* cosmogony. María is the young religious woman of European origins who considers herself superior to the indigenous people and whose accomplishment is to marry a man to fulfill her marital obligations. Lluvia is to the opposite of María: a *Malinche* figure who betrays her past because she wants to be part of the global world. Conversely, the protagonists, Evaristo and Isauro, break their stereotype. They are two indigenous homosexual men that in their youth fall in love with each other. They represent a different idea about indigenous subjectivities. Their emotional world embodies the syncretism of their identity, while their beliefs and cultural traditions enter in conflict.

The previous descriptions adhere to cinema as a “system of representation” through which people make sense of the world. It feeds and configures representations, offers sources of knowledge and provides behavior guidelines. Meanings shown on the screen share values, behavior patterns and attitudes over types and stereotypes, working as devices that project social conventions to recognize, build and rebuild our identities. But, just like identity is an ongoing process, its representation in cinema can also demonstrate nuances and challenges. Throughout the years, while some types and stereotypes have remained stable, others have evolved, evidencing different sociopolitical and socioeconomic crises that Mexico has faced, as well as the changing processes that shape identity.

3. Mexican Melodrama and Otherness: masculinity and Queer Identities

Andrea Noble [7] establishes that it is necessary to understand Mexican melodrama in terms of the social, the representation of masculinity and patriarchal values. This is paired with a specific cultural context that is determined by three master narratives: religion, nationalism and modernization. For her, “the kinds of narrative structure that define melodrama are similarly to be found as central components in the rituals of the Catholic religion that arrived in the ‘New World’ with the Spanish invasion and revolve around sin, suffering, abnegation and punishment” (p.100). In the same way, Carlos Monsiváis links Catholic guilt with violence and suggests that “from the second half of the 19th century, the ecclesiastical power has been one of the guides of the content of the melodrama to warn people about moral indecency and to turn personal experiences into a deterministic view of events” (p. 1, my translation). He further argues that:

Violence is our cross, and thanks to the lack of rights, we are Christ to scale: mugged, beaten, murdered for the sins of society or our lack of foresight or our reluctance to exercise complaint. On a rudimentary level, the formulation of melodramatic violence is deeply indebted to theology; on another, the Catholic melodrama influences

the styles of movies and soap operas, and on a third level, melodrama is the “exorcism” that turns violence, unfortunately real, in the discharge of woes and resignations that disables or mediates the will to act. ([8], p.2, my translation).

Evaristo's dream sequence in *Sueño en otro idioma* shows these Mexican melodrama features. Evaristo arrives to the monastery looking for Isauro, who is taking his Spanish class with the priest. Through the chapel window, he sees Isauro and calls him. Suddenly, he looks at the big figure of Christ on the cross high above the chapel's altar and feels guilty. He steps back from the figure and notices a river of blood flowing under his feet. He screams and wakes up all flustered. The protagonist experiences a powerful feeling of guilt for loving Isauro because he knows that by doing so, he is breaking the moral codes of Christian society and, consequently, the laws of God. His private feelings are symbolically revealed, as well as his awareness of the heteronormative discourse defended by the church. Evaristo is fighting an inner battle between his personal desire and his social obligations. His love for Isauro becomes an evil that he must resist. Soon after having this dream, he proposes to María and shuns Isauro. This decision will mark him for life, as he will carry his forbidden love until they arrive at *El Encanto*. The modest chair that he carries everywhere is the metaphor for this forbidden love that will be with him beyond death.

These aspects are reinforced by the esthetics of the scene. The emotional state of Evaristo is shown through heightened expressions in close-ups during the dream sequence and also after he wakes up. During the dream, the church is presented with low-key lighting, high contrast and gentle dolly-in movements. The colors on screen are very saturated, giving the blood a surreal texture. The rhythm of the scene speeds up when Evaristo sees blood Christ's face too. Music and fast-cut editing add tension to the sequence.

In Mexican melodrama, masculinity and patriarchal values are directly related to the opposition between the female and male figures and the construction of a macho identity. Machillot establishes that the term “macho” arises between 1910 and 1915 (in [9], p. 254) and was used to turn vulgar *mestizos* into heroes of the revolution, placing them as subjects that contributed to the building of the Nation. For De la Mora [10], the macho is the quintessential virile image (p. 2) associated with being “violent, rude, irritable, dangerous, impulsive, boastful, superficial, distrustful, unstable and false” (Machillot in [9], p. 255, my translation) and was seen as the means by which the Mexican male “recovered” his manhood and revitalized the nation (Domínguez in [10], p. 5). De la Mora [10] suggests a macho is aggressive and fearless, are capable of inflicting violence on others and “must limit their sexual desire to women, at least in public” (p. 122). At the same time, this stereotype generates its opposite, the homosexual, without which the former cannot function.

De la Mora [10] determines that “homosexuality is thus the other side of the coin shaping and complementing the social construction of Mexican male heterosexuality, the other side without which, Mexican masculinities would lose their cultural distinctiveness” (p. 18). For him, homosexuality fights desires and anxieties about masculinity. For Michael Schuessler [11], the homosexual was always depicted as “the other”, a stereotypical image that endured for more than a century. The origins of this discriminated figure can be traced back to the Spanish conquest, where these two worlds, pre-Hispanic and European, were “considered diametrically opposed: civilization vs. barbarism, polytheism vs monotheism” (p. 135), culture vs. ignorance. Even the chronicles of Bernal Díaz del Castillo and Fray Bernadino de Sahagún claimed that the sexuality of the Mesoamerican people displayed the abhorrent nature of the effeminate male and the persecution of the “*pecado nefando*” (p. 135).

However, Gerardo González [12] demonstrates that different ethnic groups in Mexico believed in a dual deity in which the principles of the masculine and the feminine functioned in correspondence and complementarity. Sexuality was a source of joy and creative potential. Homosexuality was visibly present in different rituals and festivities. Thus, the relationship with the feminine principle of the cosmos was fundamental. The Spanish Conquest imposed the Christian doctrine where sexuality is perceived as sin and homosexuality as guilt. The subsequent control of indigenous identity laid the foundations for the construction of modern Mexican identity. Mexico is a virile nation that struggles with the legacy colonialism and the development of a predominantly macho identity.

The systems of representations that configure a culture across time make it difficult to recognize “sexuality in different indigenous groups and the perception of homosexuality within their communities” ([2], p. 100, my translation). Just like stereotypes and their relationship with otherness, patriarchal gender prescriptions are deeply entrenched and reject anything that is not “the natural”, “the normal”, and “the moral” (p. 101). Accordingly, Bautista claims that there is a tendency to infantilize indigenous people because it is assumed that “they do not know what they want or what is good for them” ([2], p. 101, my translation), associating them with simplistic characteristics. For the author, it is very important to conceive them as sexual subjects, and more importantly, as queer sexual subjects (p. 101). For Benshoff [13], “human sexuality is multiple, varying and diverse”, and he further establishes that “[...] nor is it choice between two singular things (either heterosexuality or homosexuality)” (p. 29.) In fact, he acknowledges that there are a multiplicity of sexualities, a continuum from straight to gay, covering the vast areas between them “as well as the sexualities that might lie beyond them” with the term queer (p. 2). He also recognizes the importance of film representation so “films are cultural artifacts that are intricately connected to our understanding of among other things) gender, sexuality, history, and identity” (p. 2). The conception of sexuality has been presented as a dual model and the representations of it in cinema have responded to dominant social norms. However, the recognition of queer identities is increasingly more common, the sexual imaginary is being transformed and reconfiguring the “fixed” identities.

According to Schuessler [11], the characterization of homosexuality in Mexican cinema has evolved from “the caricaturized and very secondary ‘sissy’, into a character with truly human dimensions” (p. 133). The first Mexican film with a homosexual character was *La casa del ogro* (1938), but it was until *El lugar sin límites* (1978) that the figure of the homosexual acquires human dimensions. The film “explores and deconstructs the myth of the Mexican macho within the context of rural Jalisco, home of the charro, tequila and the mariachi” ([11], p. 140). The movie is a critique of traditional heterosexuality and deconstructs the impervious macho ideal that was prevalent until then. With this film, director Arturo Ripstein challenges the ideological conventions of national melodrama and of the macho Nation too:

He uses the structures of melodrama to confront audiences with everything that should not be said or shown, explicitly question and expose family values and religious morals, gender expectations, sexual norms, as well as vernacular forms and practices intimately associated with Mexican national identity. They desecrate holy spaces and institutions and thereby often alienate and enrage Mexican audiences. His films do not have happy endings. Ripstein sees the family as a “nucleus of destruction and horror” [...] Demolish the basic values of certain bourgeoisie who believe that religion, family and country are the most important factors one has. His films take an enterally different route in terms of narrative, visual style, and ideology. ([11], p. 133).

In *Sueño en otro idioma*, the main conflict also addresses these two issues: homosexuality and an indigenous identity confronted by the *mestizaje* process in which theology is the guide for morality. Evaristo and Isauro fell in love, but Catholic prescriptions imposed by the Church make Evaristo feel guilty and therefore reject Isauro. In order to follow the religious doctrine, Evaristo married María, against his desire. For more than fifty years, he stayed away from Isauro, until Martín forced them to speak with each other again. This encounter makes it clear that they still love each other, but for Evaristo this only represents regret and suffering. He always depicts a macho identity in order to neglect his real desire: he is aggressive, controlling, and overprotecting the integrity of his daughter Lluvia.

One of the scenes that better addresses this inner repression is when, after meeting again, Evaristo and Isauro go to the beach they used to go when they were young. In the same place, Isauro touches Evaristo's knee saying a few words to him in *Zikril*, to which Evaristo reacts violently, the same way he did years ago. Non-diegetic music adds tension to the scene, and the montage sequence cross-cuts between the present and the past, revealing the fight they had fifty years ago. The shots are parallel with close-ups disclosing Evaristo's repression and anger, as well as Isauro's grief. Evaristo hits Isauro with the chair he always carries with him, as a metaphor for the guilt, the anger and the sadness that he has carried all this time.

For De la Mora [11], "the hold that the stereotype of the Mexican macho exerts does not mean that notions of manhood and manliness are ahistorical, unchanging, monolithic and uncontested" (p. 9). In the 1970s, the films start to depict another idea of mexicaness in relation with gender and sexuality, where the Mexican man was less authoritarian and more sensitive and women were more active and independent. De la Mora [11] claims that the films from that period frequently functioned as allegories about the crisis of the patriarchal state, just as the ruling Institutional Revolutionary Party (or PRI, its acronym in Spanish) was unable to manage the economy and dominate national politics. These larger social changes were related with the politically uncertain years leading up and following the student massacre at Tlatelolco Plaza in 1968. These changes were "shaped by growing social disparities, the government authorities' lack of respect for civil liberties, the impact of Mexican feminism and the women's movement, and the homosexual liberation movement" (p. 109).

In the 1990s, the film industry was affected by the economic crisis in the country; however, this was one of the most productive decades and several melodramatic films received national and international awards: *La mujer de Benjamin* (1991) and *Como agua para chocolate* (1992). These movies showed stories related with identity issues, such as violence, family crisis, gender identities, migration. Also, these films demonstrated a novel production management. This cinema offered alternative inscriptions of sexual/cultural identity during the consolidation of neoliberalism in Mexico.

4. Neoliberalism, subjectivity and melodrama

In Mexico, the new conditions of life and work under the changes of modernity have transformed the social, economic and cultural spheres. The welfare state has diminished, and family dynamics have changed. Also, technology has had a very important role in the organization of daily life and the sharing of new meanings about gender, sexuality and race. These new conditions have undermined the patriarchal structure of the family: "accelerated insertion of women in the world of productive

work, drastic reduction in the number of children, the separation between sex and reproduction, transformation in couple relationships, in the roles of father and male, and in the perception that women have of themselves" ([14], p. 14, my translation). Nowadays, women have a more active role, while the influence of the family and the Church has decreased. In the same sense, new legislation about queer identities has helped develop more respectful and humane forms of social inclusion. Stereotypes are still present, but they are challenged by new systems of representation that endorse all these changes.

Globalization is transforming the national and the local, not only the economy but culture and social relations too. For Salinas [15], the government's development in this globalized economy has assumed a neoliberal character, in which entrepreneurship is rewarded yet "collective history" and communities are ignored (p. 118). Systems of representation are depicting new experiences and narratives, in response to a hybridization process defined by more heterogeneous contexts. In the media, new esthetic explorations are revealing the uncertainties of the "contemporary soul" and new narratives structures are displaying mundane experiences and contemporary crises.

Laura Podalsky [16] states that these esthetic and narrative changes disclose a reconfiguration of masculinity in contemporary Mexican cinema through the exploration of male subjectivities. There are new films that feature male protagonists in liminal periods of their lives, on the cusp of their transformative possibilities or at the end of their livesage, old men experiencing the implications of their disabilities, and men on the verge of adulthood experiencing an incredible loss (p. 162). By doing so, these films rework temporal and spatial markers in radical ways (p. 162). As she explains: "space does not function as a backdrop for the unfolding of dramatic conflicts or as an allegory delineating national identity. Instead, rural and urban landscapes function as the material register of male subjectivity. Time is not registered in terms of history, but rather through movement" (p. 163). For the author, these explorations reflect the global transformation that reshapes the role of men in Mexican society.

The main storyline in *Sueño en otro idioma* explores the subjectivity of Evaristo. The spectator travels in time to the past and into his mind during the narrative. With the use of flashbacks, the spectator is witness to his youth and his encounters with Isauro. Most of them take place at the beach, a kind of paradise where they can be whatever they want. The forest is another element of the *mise-en-scène* that explores the subjectivity of the characters. It protects the millenary knowledge of the *Zikril* people. The cave *El Encanto* is situated here too. This is the portal to an eternal life of enjoyment, according to *Zikril* mythology. One of the most important elements that reflect the subjectivity of the characters is the *Zikril* language. This indigenous language makes it possible to revive the love between the characters and, at the same time, travel to the world they belong to. *Zikril* is the key element that represents the love between them, and with it, the indigenous world that has been marginalized by the modern process of *mestizaje*.

One of the scenes that show this subjectivity mediated by language is when Evaristo e Isauro meet in order to have a conversation in *Zikril* that Martin wants to record. Evaristo speaks with Isauro in *Zikril* about Lluvia's and Martín's relationship, which he disapproves of. This is the first time that Evaristo and Isauro laugh together and Evaristo discovers that he is still in love with Isauro. The scene is shot with close-ups, but also strategic two-shots reveal the intimacy between the two characters.

5. Conclusion

Cinema as a “system of representation” is a critical arena of contestation and struggle. Although it has served to reproduce ideologies, it has also allowed the reconfiguration of identities following political, social and economic changes. In this sense, identity as meaning is not fixed. It is an ongoing process. Even the attempt to “fix” it is the result of representational practices that privilege one meaning over others. Nevertheless, this can be transformed, enriched, or subverted, just as the extensive journey of melodrama has shown.

Melodrama has been one of the most significant genre categories that represent Mexican identity. With the first movies, identities were represented following the moral and social codes of the patriarchy, where agency was established by heteronormative male power. With the implementation of a global neoliberal system, representations in melodrama evolved into broader categories and new issues, like loneliness, family crisis, the politics of the body, gender relations and subjectivity. As Patricia Torres [17] points out, we can then recognize that the expansion of the features that define the melodramatic expressions and contemporary melodrama have been re-signified through new discourses and narrative and esthetic resources, delivering a new approach to gender identities (p. 213). Although stereotypes are still present, melodrama continues to question its audience and is still an important narrative structure in the media. Therefore, it is fundamental to understand contemporary Mexican culture.

Sueño en otro idioma is a film that reinforces some stereotypes from the Mexican cinema of the Golden Age, such as violence as a characteristic of virility and masculinity, the paternal figure as responsible for feminine decency and the morality imposed by religion. However, the film also challenges the conventions of the genre by addressing gender and sexuality in relation to cultural identity. The characterization of Evaristo and Isauro as indigenous men differs from other characterizations marked by abuse and helplessness. They have agency and are able to confront their inner passions in a hybrid context where the pre-Hispanic and the globalized worlds can coexist. Likewise, Lluvia, Evaristo's granddaughter is a woman with agency. She decides what she wants to be and chooses who she wants to share her life with. Although she takes care of her grandfather, she can accept his homosexuality and support his search for happiness.

Mexican cinema operates as a device for social projection and as a mechanism in which our realities are constructed and reworked. *Sueño en otro idioma* shows that the love experience of Evaristo and Isauro reflects a mixed world where modernity and the cosmogony of their origin can coexist. The exploration of Evaristo's subjectivity is an example of an identity in transition in this globalized world. At the same time, the film shows how the old stereotypes that have shaped Mexican culture are no longer useful, particularly when they face the challenge to accept new possible forms of existence in contemporary times.

Conflict of interest

The author confirms that this article content has no conflicts of interest.

Filmography

- Arau, A. (Director/Productor). (1992). *Como agua para chocolate*, [Film], CONACULTA, IMCINE.

- Carrera, C. (Director). (1991). *La mujer de Benjamín*, [Film], Centro de Capacitación Cinematográfica, IMCINE/FOPROCINE.
- Contreras, E. (Director). (2017). *Sueño en otro idioma*, [Film], Agencia Sha, Alebrije cine y video, Revolver Amsterdam, FOPROCINE/IMCINE, Estudios Churubusco Azteca, S. A.
- Cuarón, A. (Director). (2018). *Roma*, [Film], Participant Media, Esperanto Filmoj
- De Fuentes, F. (Director). (1938). *La casa del Ogro*, [Film], Compañía Mexicana de Películas s. de R.L.
- Estrada, L. (Director). (1999). *La ley de Herodes*, [Film], Bandidos Films.
- Flores, A. (Director). (2001). *Piedras Verdes*, [Film], Videocine, De cuernos al abismo.
- Franco, M. (Director). (2020). *Nuevo orden*, [Film], The Match Factory, Teorema, Les Films d'Ici.
- Olhovich, S. (Director). (1977). *La llovizna*, [Film], Conacina, Dasa Films.
- Ripstein, A. (Director). (1978). *El lugar sin límites*, [Film], Conacite 2.
- Sala, X. (Director). (2018). *El ombligo de Guie'dani*, [Film], Xavi Sala p.c.

Author details

Ilia Lizbeth Espinosa Pacheco
Arts and Humanities, Universidad de las Américas Puebla, Puebla, Mexico

*Address all correspondence to: ilial.espinosapo@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Koper N. The unmaking of Indigeneity. Indigenous representations in the Golden Age of Mexican Cinema. *Poliarchia*. 2017;**9**:97-118
- [2] Bautista R. Reflexiones acerca de la diversidad sexual entre jóvenes indígenas en México. *Revista de Estudios Sociales*. 2018;**63**:100-109
- [3] Muñiz, C., Marañón, F. y Saldierna A. 2014. Retratando la realidad? Análisis de los estereotipos de los indígenas presentes en los programas de ficción de la televisión mexicana, *Palabra Clave* 2, pp. 263-293
- [4] Hall S. Representation. London: Sage Publications; 1998
- [5] Hall S, du Gay P. Cuestiones de identidad cultural. Buenos Aires: Amorrortu; 2003
- [6] De la Garza A. Diversity, difference and nation: Indigenous peoples on Mexican screen. *National Identities*. 2010;**4**:413-424
- [7] Noble A. National Mexican Cinema. New York: Routledge; 2012
- [8] Monsiváis C. La política del melodrama. (nd). Seminario Educar la Mirada. Flacso. México. Recuperado el 8 octubre de 2020
- [9] Rodríguez Z. Machillot, Didier. Machos y Machistas. Historia de los estereotipos mexicanos. México: Paidós; 2014
- [10] De la Mora S. Cinemachismo: Masculinities and Sexuality in Mexican Film. United States of America: University of Texas Press; 2006
- [11] Schuessler M. “Vestidas, Locas, Mayates” and “Machos”: History and Homosexuality in Mexican Cinema. *Chasqui*. 2005;**34**:132-144
- [12] González G. Entre Cuilonimiquiztlan y Sodoma, homosexualidad, cultura y ley en el México Colonial [tesis]. Facultad de Filosofía y Letras. México: UNAM; 2013
- [13] Beshoff G. Queer Images: A History of Gay and Lesbian Film in America. United States of America: Rowman & Littlefield Publishers; 2006
- [14] Barbero M. El melodrama en televisión o los avatares de la identidad industrializada. In: Herlinghaus, editor. Narraciones anacrónicas de la modernidad. Melodrama e intermedialidad en América Latina. Chile: Editorial Cuarto Propio; 2002. pp. 171-198
- [15] Salinas C. Melodrama, Identities and Modernity in Latin American Cinema: From Amores Perros to Sábado. *Aisthesis*. 2010;**48**:112-127
- [16] Podalsky L. Landscapes of subjectivity in Contemporary Mexican Cinema. *New Cinemas: Journal of Contemporary Film*. 2011;**9**:2-3
- [17] Torres P. El universo de las emociones: la reinención del melodrama en el cine mexicano contemporáneo. In: Schmidt-Welle F, Wher C, editors. 2015, Nationbuilding en el cine mexicano desde la poca de Oro hasta el presente. México: Bonilla Artigas Editores; 2015

New Masculinities, Are They Truly New?

Jeaqueline Flores Alvarez

Abstract

This article analyzes the concept of “new masculinity”, a term that has been used excessively to boast about a series of changes concerning men’s discussions, habits, practices, and behaviors that point towards more just gender relations. However, in the digital era, it is crucial to observe the role the culture industry and pseudo-intellectuals, or “prosumers” play in the construction and reconstruction of masculinity, in the emergence of a media configuration that provides great benefits to capitalism and helps preserve patriarchal privileges. Recognizing the trajectories and mutations that male chauvinism, sexism, and the patriarchy have experienced with the goal of adapting to the neoliberal and culturalist current of the times, makes it possible to demonstrate the media contraptions that contribute to perpetuating models that, celebrated as “new”, maintain strong links to traditional masculinity founded upon gender inequality. This will allow us to explain the “neo-chauvinist” and “retrosexist” trends that emerge in responses to viral feminism, as a strategy to continue perpetuating gender violence, both physical and symbolic.

Keywords: masculinities, sexism, hispter, neo-chauvinism, retrosexism, feminism, prosumerism

1. Introduction

The 21st century is celebrated as the potential era of a new gender order, an ideal moment to subvert prejudices and social inequalities: male chauvinism, misogyny, classism, racism, and homophobia. In this historic moment, dismantling the model of traditional or hegemonic masculinity¹, questioning the supposed masculine norm, is presented as an urgent undertaking in our society. However, this era of feminist awakening too finds aggressive resistance and reactionary attacks coming from those who feel themselves threatened by such claims, complaints, and charges. Men have been called to evaluate themselves, to reflect on their customs, their emotions, their relationship with women, and that for many can be uncomfortable and ultimately a waste of time.

¹ Hegemonic Masculinity can be defined as “the configuration of gender practices which embodies the currently accepted answer to the problem of the legitimacy of the patriarchy, which guarantees (or is taken to guarantee) the dominant position of men and the subordination of women” ([1], p. 39).

It is evident that many men have interpreted the call for equality and the questioning of their privilege as a provocation, as a loss of power and “manhood”. Stopping to reflect on the forms of masculinity carries with it great benefits for men, but also carries all the risks associated with removing something external that has become a part of one’s being [2]. As far back as three decades ago talk of the possibility of constructing new masculinities began. In light of this, means of communication, literature, and movements by men searching to recover the “essence of masculinity” all produced ideas of what this “new man” might resemble.

Faced with the idea of renewed masculinity, today we ask ourselves if there exists an authentic change in masculinities or if at their core, they remain unaltered. What are “new masculinities”? Are they truly less chauvinistic? Less homophobic? Egalitarian? Less violent? This idea of the appearance of “new masculinities” has been exploited in recent years to designate a series of discourses, representations, behaviors, habits, and attitudes adopted by men that are proclaimed to be “new”, healthier, alternative, egalitarian, responsible, empathetic, positive, that supposedly run counter to traditional or toxic masculinity violent and chauvinistic masculinity based upon gender inequality. In this manner, the term has been widely applauded and commercialized in Latin America, by means of the ideal of the “new metrosexual man” [3] and by the literature around new forms of “being a man” [4].

Theoretically, this new masculinity being spread promises a change that paints and image of a healthier man, more emotional, affectionate, sensible, not competitive, not homophobic, nonviolent, without addictions, responsible and respectful. A man who takes part in domestic work and caring for children, that is, a kind of masculinity that reconciles the tension between public and domestic life, between work life and family life. Additionally, new masculinity assumes the recognition of different forms of being a man, and of diverse sexual orientations. In short, does this concept lead us to an essentially deceptive or hopeful idea that male chauvinism and the patriarchy are being left behind as things of the past?

On this point I consider it important to stop and analyze what lies behind this so-called “new” masculinity, that which is hidden, that which remains, that which is untouched and continues reproducing itself. The idea of the emergence of “new masculinity” has accommodated discussions about change or deconstruction experienced by the various forms of masculinity as well as important advances for “gender parity”² that wish to depict a harmonic and uplifting scene. However, the first thing we must ask ourselves in regards to this is whether “new masculinity” looks beyond individual shifts or deconstructions and transcends to the collective and political spheres with actions that threaten hegemonic masculinity, the fraternal pact between men and the structures of the patriarchy.

2. Gender Prosumerism: “Neo-chauvinism” and “Retrosexism” in the Digital era

Communication media are spaces crucial for initiating and analyzing debates surrounding the form in which masculinity and femininity are shaped in contemporary society. As with any product of culture, gender roles are historic, they change, they

² Nancy Fraser [5] proposes using the term parity, as a concept distinct from equity and equality, which encompasses the spheres of the justice system, redistribution of wealth, and recognition, that is to say, the economic sphere as well as the cultural sphere.

are modified, they are negotiated, and they adapt to social, political, economic, and cultural transformations. Based on this, Cultural Studies give an account of the influence the culture industry has over the construction of biases and social meanings, of the central role that technology plays in the creation and propagation of popular culture [6], as with the unbreakable interconnection between human beings and technology [7, 8], the overlap of which spans from the biological to the socio-cultural and political.

In the digital age, the consumer has ceased to be passive, becoming instead a “prosumer” [9], which is to say, one who assumes the dual roles of consumer and producer, one who meddles in the development of new products or services. “Prosumerism” allows us to understand consumption as a static, permanent process wherein beyond being simply passive consumers, we are emotionally and personally involved, and whereby we ourselves become agents who negotiate, reinvent, resist, challenge, and transform existing cultural patterns [10]. Immersed in a “culture of participation” [11], the idea of the passive recipient is substituted for that of the consumer subject who is, in addition, a collaborator and co-creator [12]. As a consequence, the consumer subjects interact and ally themselves with networks of people and virtual communities to produce and distribute content of interest to themselves as well as to the group, seizing whatever the market has to offer.

Based on the previous, with the goal of ascertaining and examining the gender configurations shared and popularized by means of digital media, the concept of “gender prosumers” [13] arises in the academic literature, allowing us to analyze the ways in which men and women actively participate in the modeling and remodeling of masculinity and femininity, as with the role of intellectuals or pseudo-intellectuals within this process. Thus, on full display are the contemporary trends of “neo-chauvinism” and “retrosexism” which function as “disciplinary devices of symbolic violence” ([13], p. 120, my translation) and which, supported by the culture industry and based on prosumerism, contribute to the legitimization of violence against women and the reinforcement of gender inequality.

Along this line, the work of Alissa Quart [14] y Sayak Valencia [13] give an account of the means via which discussion about hegemonic masculinity, tied to the guidelines of the market and habits of consumption, has been costumed with a progressive veil of that tries to conceal sexism and male chauvinism, decorating these with facades of “education” and being in the “avant-garde”. Their investigations give an account of the characteristic circulation of discussions, images, and media representations about gender and sexuality that are supported by and reproduced by specific groups that actively participate in reinforcing existing gender dichotomies, producing, reproducing, and consuming the symbolic violence that circulates in the media.

Both authors refer to “hipster sexism”³, led by young middleclass university students who, masking themselves with intellectualism, art, culture, and being in style, acquire a material measure of influence within movements and on social medial, as well as within the academic sphere. As Quart [14] explains, these groups hold sway with men of varying ages and social standings by means of their wide publicity and the sale of cultural merchandise: t-shirts, films, magazines, and television programs. Given their university education, hipsters proclaim themselves “men of culture” and laud themselves for being “cultural leaders” capable of being both critical and analytical, by which means their products and statements acquire relevance within

³ Sayak Valencia [13] refers to arts and humanities students as being among the principal producers/reproducers of “hipster sexism”.

the media and social networks, marking and influencing trends. Among other things, the characteristic manner in which this group utilizes humor and sarcasm to appeal to freedom of expression must be highlighted, for it functions as a way to avoid criticism for expressing markedly misogynistic, homophobic, classist, and racist statements.

In this manner, the cultural and economic capital wielded, together with the “humorous” tone of their assertions, generally serve to soften their violent and cynical messages. For example, in so called “anti-feminist” Facebook groups these pseudo-intellectuals make use of garish phrases and cite prestigious authors to make jokes and to satirize and denigrate women, attempting to make known their cultural capital. The sexist hipster is one who tries to depoliticize and minimize the relevance of feminism, who calls feminists “feminazis”⁴, one who makes videos explaining how to patriarchy does not exist, who claims feminism is equivalent to male chauvinism, who claims that violence cannot be gender-specific, one who is certain that feminism inverted the reality of the situation and that it is actually men who suffer discrimination, assertions that go viral among groups of men who identify with this avant-garde sexism or misogyny.

In compliance with the ideals of hegemonic masculinity, by means of misogynistic discussions that pass as “cool”, neo-chauvinism employs feminist rhetoric, manufactures confusing discourse and representations of masculinity that disguise male chauvinism and sexism [15] which perpetuate “the objectification of women, but in a manner that uses mockery, quotation marks, and paradox” ([14]: 1). Upon declaring, for example, that women have achieved the same rights as men, men would now be the victims of feminism, and these groups apparently take it upon themselves to “defend their sons from the dangers of feminism” ([15], p. 10, my translation).

“Hipster sexism” and other contra-feminist trends are based on the idea that gender equality has been achieved and from this assumption, shielding themselves at all times behind their “humorous” and “only joking” character, denigrating women and sexual minorities becomes masked or “hidden” in the joke. As explained by Isabel Menéndez [15], this avant-garde sexism is based on:

The use of accounts characterized by two elements: a humorous script or outline, and the use of the retro esthetic. The underlying message, however, is based on the same familiar sexism despite incorporating more elevated insults, tolerated today due to their supposedly entertaining character, behind which it is easy to perceive modern misogyny. It demands reflection upon sexism itself - “you-know-I-only-do this-in-jest” - so the objectification of women is inscribed in the action itself, despite starting from the assumption that sexism is something old-fashioned. Its radical novelty is being at once explicit but banal, violent but ironic ([15], p. 18, my translation).

Converging with hipster sexism, as contemporary strategies seeking to depoliticize feminism, Menéndez [15] refers to “retrosexism” and “neo-chauvinism” to reflect upon gender paradigms that, promoted by the culture industry, point towards a regression in terms of gender equality. According to the author, retrosexism is a

⁴ The term “feminazi” was first used by Rush Limbaugh, an American conservative broadcaster, in his popular talk shows during the nineties. It equates feminism, an academic and political movement, with nazism, a totalitarian political doctrine whose prime representative is Adolf Hitler (1889–1945), as a way to underline the “dangerous” nature of the former and discredit the female liberation movement. For Limbaugh (1992: 194), “feminazis” are “women who are obsessed with perpetuating a modern-day holocaust: abortion”, as he describes in his book *The Way Things Ought to Be*.

narrative that cocoons itself within a bygone era in order to attenuate its true sexist cargo, which takes advantage of humor, cheekiness, and jest as stratagems for perpetuating gender injustices and “from there are reinforced classic attitudes of gender distinctions as parody and irony impede the ability to determine whether these distinctions are being supported or subverted” ([15], p. 18, my translation). Viewed in this light, hipster sexism is presented as a variant of retrosexism.

Under this logic, neo-chauvinism – another reactionary line of discourse from the avant-garde – is based on the fear of losing uniquely male privileges, a loss which gender parity could possibly bring about. This trend makes boastful claims of the injustices, abuses, and dangers of feminism, posing men as “victims” and women as “opportunists” seeking to take advantage of current conditions, which take assume “equality”, as a means to legitimize their male chauvinist burdens and complaints [15, 16]. Among the principal characteristics of neo-chauvinism Rubiales [16] cites: 1) perceiving gender equality as a threat, 2) equating feminism with male chauvinism, 3) denial of the existence of gender violence against women and, 4) questioning and satirizing the rights and autonomy of women. This is to say, neo-chauvinism is a veiled form of male chauvinism, but considering themselves oppressed and maligned by feminism within a contest where male chauvinism does not enter into the politically correct discourse as concerns masculinity. As Menéndez explains:

Neo-chauvinism is not new. It is the same male chauvinism form yesteryear but dressed in new robes: it is fed by a series of practices and rhetoric - post-chauvinism, post-feminism, hipster sexism – each one complex, their principal characteristic being ambiguity, the dynamic between power and control, which does not always reveal itself as such. Insistence upon messages that delegitimize emancipatory philosophies like feminism can be effective if, in addition, they are accompanied by a complete discourse that insists upon the obsolescence of complaints for equality and the prudishness of criticism of the use of sex and women's bodies ([15], p. 25, my translation).

In this way, in addition to positioning men as victims of feminism, these reactionary groups and movements defend the hyper-sexualization of women, for prostitution and sexual work by women as inevitable necessities; in sum, for the subordination and sexual servility of women before men [17]. Along this same course is aligned the trend of post-feminism, that is to say, associations that strongly criticize and deflect the concerns of “second wave of feminism” [18] which is blamed for being “victim feminism” ([19], p. 136).

Post-feminism, founded upon neoliberal philosophy, is based on individualism and the “ethics of personal choice”, defends traditional ideals of femininity, sexist beauty stereotypes and suggests looking at the “empowering” aspects of prostitution and pornography ([15]: 13). That is to say, it is a segregationist and depoliticized feminism that leaves aside the questioning of hegemonic masculinity, criticisms of patriarchal capitalism, while at the same time that it denies the existence of gender injustice. Such points of criticism formed the axis of the grievances of “second wave” feminism of the seventies [5]. Therefore, rooted in the idea that we live in a just society wherein gender equality has been achieved, post-feminism merged easily with hegemonic neoliberalism which contributes to ingraining “a visual and narrative rhetoric that emanates from neoliberal media for the purpose of discrediting feminism as a gender ideology belonging to the past” ([20]: 4, my translation).

As explained by Raewyn Connell [1], the patriarchy is maintained thanks to structural and institutional complicity and the agreement reached between men and

women. Along this same line, Rita Segato [21] emphasizes that the patriarchy is not embodied solely by men but by women as well who, in conjunction with their male counterparts, contribute to maintaining and supporting the patriarchal order. In this sense it is important to point out the participation of those women who partake in anti-feminist trends, who contribute to propagating hate, jokes and their own denigration, as well as that of other women. The above comes about in “a reactionary and patriarchal discourse that is reproduced and magnified in the media and social networks; an ambiguous and confusing discourse that even achieves the support and backing of many women who believe that feminism and its entire agenda is detrimental to them” ([15]: 4, my translation).

3. Gender Parity: advance or regression?

Therefore, do we speak of progress or regression? Isabel Menéndez [15] underlines the prevalence of a “regressive metadiscourse” that, promoted by the culture industry, has a strong resonance with the new/information media, cyberspace, and popular culture. A deceptive discourse that pretends to be liberating and feminist, makes use of feminist symbols and narratives to distort, confuse, and in this way creates an anti-feminist and hateful discourse against women which requires the agreement and consent of, and is received and reproduced with pleasure by, broad sectors of the population. Such that, by means of an intentional misrepresentation of feminism, that which at one point was understood to be oppressive has now acquired a character of liberation.

But this regression and reactionary organization goes beyond the symbolic plane and is not confined to the media, materializing itself in the form of an increase in violence against women as well as acts of terrorism. Over the last number of years, as a reactionary response to feminism and the “crisis of masculinity”⁵, we have witnessed the emergence of groups referring to themselves as “ultra-chauvinist”⁶ as

⁵ Since the early 2000s the “crisis of masculinity” has positioned itself as a hot topic with as strong resonance within academia as within social media. For researchers like Ángels Caribí [22], this “crisis” manifested itself as much in men’s need to remedy physical deficiencies, as with the use of Viagra and esthetic plastic surgery, as in the increase in instances of domestic violence. For the author, domestic violence is intimately connected with employment insecurity and material loss that men experience; but, in addition, with their inability to deal with their emotions in a self-confident manner, without harming others. Given this, exercising violence allows those perpetrators to “feel like men” and experience the feelings of power and control that they consider having been lost.

⁶ An emblematic case is that of Roosh Valisadeh, a US blogger and activist who is known for his openly misogynist statements, for creating guides that help tourists have sex with women from various countries and for opting in favor of legalizing rape “if it takes place on private property”. Through his statements on social media, books, and conferences, Roosh has promoted ideas about “neo-masculinity”, an ideology, founded upon masculine superiority, which emphasizes discussions that make appeals to biology, that urge the importance of maintaining traditional gender roles as regards sex, and that promote the idea of the ineffectiveness of feminism. In this way, in 2016, the blogger called upon men from 44 countries to combat “feminist oppression”, exhorting them to form “tribes of men” to fight for their rights. Considering the fact of his being a dangerous individual legitimizing gender violence, his events and his influence were able to be halted by local governments in several cities thanks to pressure from feminist organizations.

well as terrorist trends embodied by the “incel community”⁷ or separatist antifeminist groups like the “Men Going Their Own Way” (MGTOW)⁸ movement, which aims to counteract the social pressure and questioning of their privileges by feminism. Thus, at the same time there sprout groups of men who seek to critically reflect upon and deconstruct their masculinity, other movements like that of the mythopoetics⁹ and those of neomascularity or antifeminism seek to recover the supposed “essence” of masculinity and strengthen the traditional model based on the gender hierarchy, inequality and masculine supremacy.

Within the context of advanced capitalism, the theory of markets together with the emergence of prosumers is useful for analyzing gender politics as a strategy that gets us closer to an analysis of the models of masculinity that men and women produce, consume, update, and renew [13]. By these means it becomes possible to analyze the repositioning, the masks, the camouflage, and maneuvers via which is articulated a gender paradigm that, boasting of being “new”, solidifies the masculine heteropatriarchal hegemony, violence, and gender inequality. Additionally, it is crucial to insist that hipster sexism, neo-chauvinism and retrosexism as reactionary trends masked by pseudo-intellectualism, culture, and humor, far from justifiable by claims that they are “simply memes”, opinions, cartoons, jokes or simple publicity campaigns promoting some product, have social effects that materialize in daily life with severe consequences for the lives and well-being of women.

4. How should we position ourselves towards “New Masculinities”?

Presently, the work of teasing out new forms of masculinity presents itself as a political-ethical wager that requires there be criticism of the patriarchy and the existing gender order, and, as a consequence, criticism of masculinity and power relationships [3, 23]. We must recognize that that which today takes precedence in the political sphere, among non-governmental organizations (NGOs) and the media, is a discussion of gender equality and rejection of male chauvinism, a discussion

⁷ The “incel” (involuntarily celibate) movement is an online community that has brought together thousands of heterosexual men who consider themselves nonconformists and who are frustrated by their romantic and sexual propositions constantly being rejected by women. As a means of taking revenge, this group incites hate, rape, and violence towards women. Some of their more well-known members are Elliot Rodger, a 22-year-old American who in 2014 murdered six people on a university campus in Isla Vista, California before then killing himself. And Alek Minassian of Canada who in 2018, bearing the standard of Elliot Rodger, carried out a massacre in Toronto, Canada that took the lives of 10 people and injured 15 others. More information on these cases can be found at <https://www.thecut.com/2018/04/incel-meaning-rebellion-alex-minassian-elliott-rodger-reddit.html>

⁸ Using the slogan “Women are your enemy, men are under attack”, the online community Men Going Their Own Way is a separatist organization headed by heterosexual men who consider themselves victims of feminism and seek to defend masculine/male supremacy by breaking affective/emotional ties with women, without, however, excluding the eventuality of sexual intercourse. This anti-feminist movement has a strong presence in countries such as Canada, the United States, Australia, and the United Kingdom. To learn more about the basis of the movement consult the book *Men Going Their Own Way: The Red Pill Anti-Feminism Survival Guide* (2019), by Charles River.

⁹ The mythopoetic movement is inspired by Robert Bly’s book *Iron John* (1992), which serves as the ideological basis for this spiritual trend that seeks to recover the essential nucleus of masculinity by means of therapy, retreats, and male circles. This movement proliferated in the United States during the eighties and nineties.

appropriated not solely by institutions but rather by a great quantity of people who assume conditions are ever improving. However, it is an incoherent discussion that masks power inequalities and leaves intact patriarchal privileges and concessions thereto.

It is important to note that politically correct discourse on gender equality, of a neoliberal slant that promotes certain institutions, which entangles itself with discussions of renewed masculinity, lacks a political foundation if it does not examine and contend with the power structure, privileges, complicity, dominance, and pacts of/by hegemonic masculinity [1]. Therefore, it is important to expose sexism's new disguises and the strategies of "recycling cultural patriarchy" ([9], p. 102), the inequitable covering-up behind talk of gender equity or equality, as illustrated by contemporary violent and male chauvinist practices.

We can no longer afford to think that certain actions on the part of men like crying, involving themselves in caring for children, cooking, changing diapers, doing the dishes, doing laundry, using deodorant, or going to the doctor are indicative of some kind of "new masculinity". By assuming that any novel practices on the part of men server as guides for proclaiming that masculinity has been renewed for the better, we run the risk of depoliticizing the criticism of masculinity. Therefore, changes in certain habits and patterns of behaviors on the part of men can be deceiving in they do not involve redistribution of power, given that these changes can exist harmoniously with traditional and detrimental forms of masculinity.

This becomes clear if we stop to analyze the dearth of support for "neo", "retro", "light" and "metrosexual" forms of masculinity. According to Ivan García's [3] explanation in an interview:

There is a very light, very superficial version of new masculinities, for example, a man who simply cries is labeled neo-masculine, it seems to me that exempting personal conduct from asking for the democratization of power, certainly does not lend support to the claim that one is a "new man". A man who today, in contrast to the past, changes diapers, but is the same sexist as ever towards his female colleagues, or is now more vain and moisturizes his skin, hence for some metrosexuality is a new masculinity (García, 2012, as cited in [3], p. 101, my translation).

Interpreting actions and personal behavior, or patterns of consumption, as markers of a weakening of the structures, or as some radical and fundamental change within the masculinities, impedes continued work regarding matters invaluable to the feminist agenda. In this sense, as regards the transformations undergone by masculinity, questioning men's power at the individual as well as at the collective level becomes central in diverse spheres: individual, the family, the community, institutional, organizational, and structural [3]. Viewed in this light, it is impossible to laud these "new masculinities" that are devoid of anything truly new, but rather they have adjusted and restructured from current conditions by means of discussions and practices that are friendly towards hegemonic masculinity. Thus, much work remains pending.

Conflict of interest

The author confirm that this article content has no conflicts of interest.

IntechOpen


IntechOpen

Author details

Jeaqueline Flores Alvarez
Doctorado en Creación y Teorías de la Cultura, Escuela de Artes y Humanidades,
Universidad de las Américas Puebla, Mexico

*Address all correspondence to: yaqyoga@gmail.com

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Connell R. La organización de la masculinidad. In: Valdés T, Olavarria J, editors. *Masculinidades. Poder y crisis*. Isis Internacional y FLACSO Chile; 1997. pp. 31-48
- [2] Rodríguez AJ. *La nueva masculinidad de siempre: Capitalismo, deseo y falofobias*. Anagrama; 2020. 213 p
- [3] García LF. *Nuevas masculinidades: Discursos y prácticas de resistencia al patriarcado*. FLACSO; 2015. 229 p
- [4] Viveros M. Masculinidades alternativas y masculinidades relacionales. In: Ramírez P, editor. *Memorias Cátedra Abierta Hernán Henao*. Medellín: Instituto de Estudios Regionales; 2007. pp. 106-116
- [5] Fraser N. *Fortunas del feminismo. Traficantes de sueños*; 2015. 280 p
- [6] Du Gay P, Hall S, Janes L, Koed Madsen A, Mackay H, Negus K. *Doing Cultural Studies: The Story of the Sony Walkman*. 2nd ed. London: SAGE; 2013. 175 p
- [7] Ferrando F. *El Posthumanismo Filosófico*. Brazil: Universidad de Curitiba; 2019 Available from: <https://www.youtube.com/watch?v=HrcZ75x6mYc&t=1214s>
- [8] Kember S, Zyilinska J. *Life after New Media: Mediation as a Vital Process*. Cambridge: MIT Press; 2012. 288 p
- [9] Potra S. What defines a prosumer? An insight in participative consumer behavior. In: 5th International Conference on Managment, Leadership and Governance (ICMLG); 2017. pp. 380-385
- [10] Storey J. *Cultural Studies and the Study of Popular Culture*. Edinburgh: Edinburgh: Edinburgh University Press; 2010. 694 p
- [11] Jenkins H. *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press; 2006. 336 p
- [12] Xie C, Baggozi R, Troye S. Trying to prosume: Toward a theory of consumers as co-creators of value. *Journal of the Academy of Marketing Science*. 2008;**36**:109-122
- [13] Valencia S. Nuevas masculinidades? Sexismo hipster y machismo light. In: Bercovich S, Cruz S, editors. *Topografías de las violencias. Alteridades e impasses sociales*. El Colegio de la Frontera Norte; 2015. pp. 107-124
- [14] Quart A. The Age of Hipster Sexism. *The Cut*. 2012. Available from: https://www.thecut.com/2012/10/age-of-hipster-sexism.html#_ga=2.128345740.1840053351.1616615524-655423978.1616615524
- [15] Menéndez MI. Entre machismo y retrosexismo: Antifeminismo en industrias culturales. *Prisma Social*. 2017;**2**:1-30. Available from: <https://revistaprismasocial.es/article/view/1544>
- [16] Rubiales A. El neomachismo. *El País*. 2010. Available from: https://elpais.com/diario/2010/01/15/opinion/1263510005_850215.html
- [17] Lorente M. Tú haz la comida, que yo cuelgo los cuadros. *Trampas y tramposos en la cultura de la desigualdad*. Barcelona: Ares y Mares; 2014. 288 p
- [18] Heywood L, Drake J. *Third Wave Agenda: Being Feminist, Doing Feminism*. Minneapolis: University of Minnesota Press; 1993. 280 p

[19] Ferguson A. Moral responsibility and social change: A new theory of the self. *Hypatia*. 1997;12(3):116-141

[20] Moreno I. Postfeminismos: representaciones de género en la cultura popular neoliberal. Madrid: Universidad Complutense de Madrid; 2012. 94 p

[21] Segato R. Las nuevas formas de guerra y el cuerpo de las mujeres. México: *Pez en el Árbol*; 2014. 120 p

[22] Connell RW, Messerschmidt JW. Hegemonic masculinity: Rethinking the concept. *Gender & Society*. 2005;19(6):829-859

[23] Faur E. Masculinidades y desarrollo social: Las relaciones de género desde la perspectiva de los hombres. Bogotá: Arango Editores Ltda; 2004. 315 p

Chapter

A Hundred Stitches Make a Canvas: How the Practice of Embroidery Relates to Forms of Life in Poverty in Atempa Coyomeapan

José Mariano Amador

Abstract

Embroidery making is one of the practices carried out in the indigenous town of Atempa, Coyomeapan, in the Sierra Negra mountain range of the state of Puebla. The temporal, physical, social, and geographic milieu that it occupies in people's lives leads us to consider it as an expression that sheds light on the community's form of life. This study analyzes the intersections among the scaled levels of capital that configure embroidery making with other activities and places. The case, from the theoretical lens of lifestyles and unequal development, raises a series of questions over the interaction of these practices and places, which constitute, in the collective imagination, the precarious and vulnerable lives in this town. In this sense, this article posits that the practice of embroidery in Atempa reveals the way in which this town connects to the world via the different levels of domination-subordination produced by its commercial, social and spatial relations, and, among other factors, how it shapes their ways of life in poverty. This perspective allows the studying of its practices and contexts as part of the phenomenon of poverty from a local level, without disassociating it from its global conditioning.

Keywords: embroidery, lifestyles, unequal development, practices, poverty, Mexico, space, scale, place

1. Introduction

They bring it over from Chilac. A lady comes and distributes it among us. All the women know how to embroider; no other way. There is not any work here like in Tehuacan... They pay us little... \$20 to \$25 pesos, depending on what we are stitching.

They make anything from dresses to blouses that they resell in Cancun or Matamoros¹. They come with the design already made. We just stitch the pattern. We put up our own string and material. We do two a week, because we each have our own chores to do. My aunts taught me. My mother taught me. We grew up with it because there aren't any jobs. We learned very young. (A. and L. Campos, personal conversation, 14 of February 2020).

On the various occasions that I visited the Campos² family in Atempa, Coyomeapan, Puebla, I found a similar scene: three women weaving, sitting in their patio, under the shade of the extending roofs over stone floors. Veronica, Leticia and Antonia (grandmother, daughter and granddaughter) belong to different generations of women from the same family that, among their diverse activities, are also embroiderers. Sitting next to each other, heads down, their gaze focused on the next stitch to color the petals of a flower. The three women barely say anything and only look up to put a new string through the needle. They stitch in a hurry, as they are committed to deliver a shipment with multiple pieces the following day. This performance is replicated in most of the houses in this community of 205 inhabitants [1], since embroidery making is a primary source of income for its residents, in addition to their work in the fields, their trade of products derived from agriculture and livestock, and their paid domestic work.

Atempa is one of forty-two towns in the municipality of Coyomeapan. It is situated in a depression among various hills in the heart of Puebla's Sierra Negra mountain range. Its denizens are indigenous, Nahuatl-speaking first, Spanish-speaking second. The temporal, physical, social, and geographic space that embroidery occupies in these people's lives in Atempa makes me consider how an expression can shed light on their form of life [2]. The field of self-expression within embroidery alludes to an area of content revealing how this practice implicates diverse precepts of precarity, vulnerability, and precarious work dynamics operating from unequal relations in scaled spatial and social capital.

The conditions of Precarity and Marginalization of this town have been studied by academia and institutions of social development, interpreting its findings from public policy or by way of diverse economic variables within a wider geography: regional or municipal, in the best of cases. Of these investigations, we know that the municipality of Coyomeapan, in which the town of Atempa is located, is one of the rural indigenous territories with the most extreme poverty in the state of Puebla [3]. Nevertheless, the majority of these findings had been conducted without considering the local, social and material transformations in concomitance and synchronicity with its actors' practice, who are the protagonists in the phenomena of this investigation. Analyzing the case from this perspective raises questions over the contradictions, the particularities and conditionings that arise among microgeographies when expressing specific lifestyles by these practices [2].

¹ Tehuacan is the second most important populated commercial hub in the state of Puebla, after its state capital. It is situated 120 minutes from Coyomeapan, a hillside municipality bordering the state of Oaxaca, and 30 minutes from the peri-urban municipality of San Gabriel Chilac. Cancun is a tourist hub well known for its domestic and foreign vacationers located in the state of Quintana Roo. Matamoros, Tamaulipas, is located on the border with the United States, and is an important commercial cross-border city.

² The names of the participants have been changed to protect their identity for the sake of this investigation.



Figure 1.
Embroidery in action.

In this text, by analyzing the testimonies of embroiderers in Atempa, compiled in a series of interviews and data obtained by way of participatory observation, the ways in which this town connects with other territories through differing levels of domination-subordination in its commercial, social, and spatial relations are revealed. This, among other phenomena, produces ways of life whose constitutive elements are defined by the poverty registry. Not only does this just concern the practice of embroidery, which we can recognize among the dynamics of accumulation, but also the meaning behind its making. A way of its making is defined by time and particular durations; by constitutive spaces through its act of social relations; by specific actors that implement certain techniques, that make available their body, the capacity and their knowledge to complete certain tasks. It is also defined by utilities and objects such as tools and supports, and scenographic objects that bear witness to the acts, whose contribution resides in its capacity to set the stage and contextualize. As a result, the work of textualizing is complicated. Putting into words what appears both simultaneously and on its own rhythm makes impossible the endeavor to deliver the intersubjective experience as it is perceived and expressed by its practitioners ([4], pp. 101–130).

The following lines present, firstly, the context of the investigation in empirical terms, with references of these connections between the concrete practice and other practices, and how their implications are scaled up and down from the local level. In the next segment, spaces for this practice, its materiality and correlation with its capacity to reproduce actions that make embroidering possible are analyzed. Lastly, the connection between these scaled spaces and the meaning of conformity in poverty lifestyles is discussed (**Figure 1**).

2. The practice of embroidery in Atempa

Some of the inhabitants of the municipality of Coyomeapan that serve as managers rely on the textile *maquiladoras* in the municipality of San Gabriel Chilac in the valley

of Tehuacan to bring work to the communities of their own region. They distribute assignments among the localities, which consist of making traditional patterns of flowers and animals on pieces of fabric. These patterns are then used to put together dresses in the workshops at Chilac. The work provider gives instructions for the product's creation, including techniques and materials. They give them the design or illustration of the patterns to be replicated in large quantities. However, they do not provide them with even the supplies (needles, threads, hoops and fabric) or technical knowledge, so each person invests in their own resources and goes with a member of their family or community for help with the task in case they were not already a master of it.

With a big assignment, the members of the family get together in a group effort. On some occasions, they also include the men. Although uncommon for men to do embroidery work, they still participate. Ramon, a member of the Campos family, Leticia's brother, assures that "if you're not in the fields, whether boy or girl, you learn to sew. One must help the family...otherwise we won't have anything to eat. Here everyone must work with cloth". (R. Campos, personal conversation, 24 of April 2021). Family support, in this sense, is essential to achieve the production goals promised to their employers.

Compared with other economic and productive activities, embroidery in Atempa seems to be a practice that bestows certain autonomy and comfort for the worker. They decide if they take on the job, the time they'll dedicate to each piece, and where they will work on it. However, they are not backed by any formal contract that establishes a relationship with their employer. The commitment is made orally, trusting in the word of both the employer and employee in terms of remuneration and how the product is to be delivered. There are no work benefits, nor is there any legal fiscal obligation with any party. As with other work that is made from home, the opposite of the supposed benefits found in this kind of work, there exists a system of control that defines the practices through determining deadlines, quality control, creative process, designs, technique and economic value for handiwork [5]. This explains the means of production: hurried, in family, within inadequate spaces, simultaneous with other duties, with low-quality materials and tools, and without active participation in the design development. The embroiderers are aware of the possibility that their work provider may stop visiting them for a time, depriving them of the income of said activity. Nevertheless, they know that the product they work on is profitable enough to have their labor requested for over 30 years, though intermittently, as part of their production chain.

Sewing is a practice that is transferred from generation to generation. Antonia learned to sew from her aunt Leticia; Leticia was taught by her mother. Ramon shares his memories about his childhood in the company of his mother rooted in embroidering. He recalls that, at seven years old, she taught him to work sewing fabric and embroidering tablecloth. He boasts about the evolution of his technique that younger generations, like those of Antonia at 19 years old, do not learn so young anymore. This is attributed to the economic support provided by government social programs, whose conditions stipulate that children study and not work. Likewise, the parents of these families have prioritized they attend school until they are no longer beneficiaries of these programs. The young people that still preserve these practices can now access the web through their mobile phones and *Google* search embroidery techniques and design patterns that are different from those of San Gabriel Chilac. There are those that assure that they dedicate enough time to start their own businesses with these newly found designs.

Every two or three weeks, the manager collects the finished cloths from the different communities. They review the quality of the work and pay according to their agreement, be it \$20 to \$60 pesos per pattern varying by size, material, and complexity. They deliver the product to the workshops in Chilac, where the dressmaking families are then in charge of assembling the complete garment: blouses, *huipiles* or chanel-style dresses. Sometimes, the manager or other members of the same family are also dressmakers, reducing assignment product costs that would go towards the person in charge of commissioning the patterns in the localities. The complete garment is sold to small retail dealers from \$150 to \$300 pesos. The public resells them on the internet or in tourist areas across the country, and in these places, their selling price can start as low as \$150 pesos. The Campos family members assume their garments possibly end up as *artesanias* or souvenirs in the major cities throughout the country. They know that their target consumer is primarily national tourists and foreigners.

The description of the experience of embroidering by Leticia and Antonia is full of digressions. This evidences the practice's dialectical relationship with other practices. For example, Antonia, when sharing her personal experience about her embroidering, shies away from any direct descriptions that constitute her actions and explains the conditions of her personal health, which are made clear the moment she is there sitting while sewing. She narrates that she left school in her second year of middle school due to illness. She is not able to walk as an outcome. The school that she attended is situated in the auxiliary *junta* of San Juan Cuautla, a town about an hour away through dirt roads and pathways.

The doctor told me that I could not overexert myself, so the teachers came. They insisted that I go back to studying. I told them that I could do the work from home, but they did not want it, because according to them, supervisory boards were coming to review them. And since I would not be there, they would mark me as absent. Well they did not give me a chance to study from home, and I told the teacher that were I permitted, I would go into school a couple of days per week, but they still would not allow it (A. Campos, personal communication, 14 of February 2020).



Figure 2.
The practitioner's space.

As a result of this situation, Antonia finds in her embroidery a way to collaborate with the needs of the family, adapting within the limits of her health. She can work from home, while the rest work the farms and fields. Additionally, she thinks that her embroidery allows her to continue her education in a more open format. She knows that, as opposed to the scholastic system, the practice can be done from anywhere (**Figure 2**).

3. The spaces for practicing

The house patio is the primary space that the Campos family uses for their trade, first for its openness and secondly for its tranquility. It has a rectangular shape, balanced by the surrounding construction that breaks from the uneven rhythm of vegetation everywhere else one looks. The walls of the rooms that make up the building that the family calls home flank the area. From the front, one can find the little shop that they attend to from a window, and next to it, to the left, is a place where the elderly rest. Nearby are the dining area and kitchen. Perpendicularly, on the far left, in the corner of the patio you find Leticia's room and a place for prayer. Adjacent, another room is used for meal prepping, with a makeshift ventilation for smoke from the fireplace, and next to this is a walkway to enter and exit the complex. A building with two floors that protrudes at the back on the right side where Ramon resides. Since Ramon is also a traditional medic and inspector for the town, this building also houses an office and a waiting area. This building is separated from the embroidery scene of the patio by large flowerpots of different sizes, some steps and by a translucent roof that filters light, providing shade and a cooler temperature than that of the patio. Some walls are made of block and others of adobe. They reveal a history in each room. The care of the construction work, different in each wall, as well as the selective application of paint and ironwork, help us assume that each investment reveals the hierarchy of every family member and the spaces they inhabit.

It is evident that the patio I refer to here sets the scene for many acts. It participates actively in various practices that can sometimes unfold in parallel, simultaneously, consecutively and even confusedly. On that instance, for example, during the inter-generational embroidery scene, next to the basket of threads that were being used at that time for stitching, another basket rested with the grains of the cobs that, perhaps, days before, had been shelled in that very same place. Or even then, when we focus our attention on the center of the patio and look up, we notice overhead the hanging lines with recently washed clothes dripping onto the floor. One can infer that some of the family members, if not all of them together, were washing them there that same morning.

The practice of embroidering also takes place in other scenarios as other activities are taking place. Rosa assures us that some of the older family members have developed the ability to sew and stroll at the same time. They take advantage of the time it takes to get from one town to another to catch up on their embroidering. In addition to dominating their technique, these people, who know their routes well, have developed a kinesthetic sense to the point of not tripping or falling, all while not missing a single stitch despite all the ups and downs of the irregular paths. It is also common to observe in community meetings, like school meetups or religious celebrations, how mainly women can both attend the event while also maintaining their focus down on the sewing hoop upon which they are working.

Although a large part of these actions take place with one's hands, good vision is imperative. For that, the Campos family recognizes the importance of light to be

able to sew. In this sense, during the day, the natural daylight, free and abundant in the patio, gives an ideal spot for this job. During the night, the practice changes, requiring more effort to coordinate hands and eyes. Additionally, it becomes a more intimate affair, since it is done inside single rooms with or without company, and the surrounding darkness hinders the objects and fades the details of their scenery. The embroiderers thus depend on electricity and find the best way to get hold of a light source. Those who lived twenty years ago remember how there was no electricity in Atempa, and that diminished their productivity:

For some, light is very important, and for others, it is saddening. Because those that already had some money began buying their electrical apparatuses like their televisions or their tape recorders. Before, it was with cassette tape. They began buying what they liked. They listened to music. But for others it was sad because they did not have the resources to buy. For some came a storm, and for others, happiness. One could work during the night with electricity because one could see more. Because then, one survived embroidering cloths. Some cloths you'd stitch flowers, and you start with a needle and sewing hoop. Before you could only see with daylight, because there was no light at night, and by candlelight it was too difficult to do those things. (R. Campos, personal conversation, 24 of April 2021).

The testimony of the family makes evident that the arrival of electricity deepened social differences within the towns. Those that have electricity can dedicate more work hours and thereby obtain better remuneration. Those that cannot are unlikely to improve their productive capacity in quantity, time, and quality, and are often rejected by employers who seek to make their processes and investments more efficient.



Figure 3.
The practiced space.

The space for this practice is not just one. We could discuss the ideal conditions for production in these spaces that are favorable to the purposes of action and facilitate work being done. We can think about space from the perspective that it is shaped by a person's own actions and objects, from their movement and displacement, rather than by space as a container [6], for it helps us perceive the overlaps between practices and their transformative capacity. De Certau affirms that "space is a practiced place" and that it is not univocal or stable ([7], pp. 117). This is similar to the Campos family's patio, where, throughout the morning hours, it functions as the space to hang clothes dry and as a sewing area, as a corral for the hens all day long, as a storage for corn and as a place to use the motor mill for preparing *nixtamal*. Just like this patio, with its various amounts of uses and applications, space is produced and transformed at every moment (**Figure 3**) [8].

4. Forms of life threaded through levels of meaning and scales

Forms of life are languages ([2], pp. 26). They are made up of elements of significance that deliver meaning to the world we inhabit both individually and collectively. These elements integrate and correlate among different levels and scales.

It is in these practices that elements interact with time and space. Practical scenes require subjects and a material context, whether social or symbolic, to operate their reality and to be able to transform it. The places of practice become enriched with the presence or absence of other participants, their tools and other objects that configure that space. Each one of them – in their structure, shape and use – bears significant elements that engage in dialog within a particular language system. The semiotics of practice, like theory and methodology, identifies fields of expression at different levels: "a practice integrates signs and texts, and even an isolated sign" ([9], pp. 18). The level of immanence of these practices allows us to move at generative scales of expression according to the relevance of that movement in its analysis. Thus, it has the same capacity for signification; in this case, they are the patio, the means of work of the embroiderers and even the designs embroidered on a fabric.

The practical scene and the articulated strategies adjust to the act and then are re-signified. In other words, for example, the practice changes if one sews in the patio, or during the night inside a room, or while walking. In addition, other practices that are coupled with embroidering, like sewing garments or its commercialization. When a practical scene is adjusted, it also moves all the levels that constitute it.

If we think about the space of this system, we can consider that the scenes of practices can be created from scaled spaces such as rooms, houses or neighborhoods, and also from the personal, familial or communal interactions at a local level on a spatial scale. We can speak about regional, national or global scales, insofar as there is a tendency to generalize and typify forms of life. The distinction of practices in this proportion is blurred. Just as in the scaled system of meaning, in terms of space, "what happens at one scale cannot be understood outside of the nested relationships that exist across a hierarchy of scales" ([10], pp. 95).

Thus, a practice cannot be understood as exempt from other practices, nor as a form of life independent from other forms of life. Fontanille points out that "a form of life exists...in the confrontation and comparison with others" ([2], pp. 71). This is the result of the conflict generated by being and acting upon a world of counterpoints with other existences and performances that adjust their forms to different extents. It has, at the level of practice, the transforming potential of its form to change, profoundly, the course of its life [2].

5. Atempa in the geography of poverty

Although embroidering by itself is not associated with forms of life in poverty, in this case the spatial and temporal context of such practice is found within a social, political, economic and symbolic geography that permeates the significance of each element that constitutes it and its relationship with precarity. The geography of poverty defines its territories on a global scale characterized by the impossibility of its inhabitants to make their own decisions on how to direct their lives, as a result for having denied them access to social, economic and political means to exercise the power of choice, according to Amartya Sen [11]. The geography of poverty, from a global scale, generalizes ways of life and the actions to confront their effects. From a local scale, a single lifestyle position is distinguished in its locality in relation to other spatial, social and economic positions relative to other spaces and forms of life.

In Atempa, the way to access economic, social, and political means are extremely limited. Embroidering is one of the few options a person has for attaining economic resources, one that does not demand so much from the body as those that work the fields, or for those that do not require commuting like farm or domestic workers. However, embroidering experiences the same workdays as those in a textile workshop: a systematized production of pieces that will be assembled in a later industrial process. Unlike other towns, despite the time dedicated for sewing, is that there is no technical and/or artistic tradition that accounts for a worldview rich in symbols regarding its past origins.

As described above, these practices respond with shortages to the commercial demands other territories and scales they relate to. It competes with other towns where conditions and means of making undervalue manpower in order to keep better commissions with higher earnings. This prevents the possibility to establish their own prices in the market. Producers are also aware that if the materials become cheaper, so will the quality of the products. Nonetheless, they need to compensate the costs of shipping at the expense of a competitive price for their work. Also, they rival mechanized processes, where production time is significantly reduced by programmed machines that emulate artisanal stitching.

Even though this market product refers to an ancestral, indigenous cosmovision expressed by shapes embodied through the application of needle onto fabric, the commercial intention of this practice has uprooted the communities that originally decorated their own dresses with these techniques. With this I mean that the traditional expression whose origins are in the indigenous communities of San Gabriel Chilac has been displaced to communities of neighboring regions in the Tehuacan valley and throughout the Sierra Negra, where these represented elements in the design have a symbolic, expressive and representative value of an identity that is not their own. In this sense, the value for these communities shifts from day-to-day or symbolic use to exchange or trade use. Furthermore, those that are involved in the production process understand that the practice responds to the demands of the market.

6. Conclusions

The practice of embroidering in Atempa, Coyomeapan, has integrated itself into the rules and language of poverty, displaying precarious and vulnerable ways of life in fields of expression and content. However, the spatial coordinates of the area, geographically situated as part of global poverty, make experiencing this way of life unique, resisting the typification of its practices and its development.



Figure 4.
The patio, the practice's place and its overlaps.

The expression of commercial, social, and spatial relationships across different levels allows us to analyze this practice and its local contexts as part of the phenomenon of poverty and its connection worldwide. It is necessary to develop a meticulous effort, pausing to take in every field of the practice's expression, identifying exhaustively the way in which its elements relate significantly with poverty, precarity and experienced vulnerability. This way, we can envision a realization of the details concerning the processes of existence. For example, what is revealed by the techniques, the designs and compositions, the materials, the ways of doing, the negotiations with suppliers, the internal roles within the familial organization, the geography, the infrastructure and a series of elements that particularly occur in the spatial environment of Atempa.

The theory on forms of life offers a variety of critical resources that, when focused primarily on the field of analyzing practices, allows one to visualize the significant elements of these forms in action, time and space. The 'action' element during these practices reveal how forms of life have a power to transform other imposed or unbalanced forms, like a power of persistence or conciliation. Thus, it is evident in the case of forms of life associated with poverty, the field of analyzing practice is indispensable.

It seems that the phenomenon of poverty settles in physical and social spaces, penetrating all layers of existence where one looks, reproducing itself in every practice. On a deeper level, the questioning arises from whether a remedy exists, or if it is possible to transform said way of life. That questioning is not new, and attempting to answer it is even less new. The importance resides not in abandoning this questioning while these forms of life in poverty endure; rather, it resides in trying to reclaim the capacity to make sense of it, and then, disassociating from it (**Figure 4**).

Conflict of interest

The author confirm that this article content has no conflicts of interest.

IntechOpen


IntechOpen

Author details

José Mariano Amador
Escuela de Artes y Humanidades, Universidad de las Américas Puebla, Puebla, Mexico

*Address all correspondence to: josem.amador@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Pueblos de México en Internet. (s.f.). Atempa (Coyomeapan, Puebla). 2021. Recovered the 9th of November 2021 from <https://mexico.pueblosamerica.com/i/atempa-8/>
- [2] Fontanille J. Formas de vida. Fondo Editorial de la Universidad de Lima; 2018
- [3] Diario Oficial de la Federación. Decreto por el que se formula la Declaratoria de las Zonas de Atención Prioritaria para el año. 2019. Available from: https://www.dof.gob.mx/nota_detalle.php?codigo=5547481&fecha=28/12/2018, 2018
- [4] Dondero M. “Semiótica de la acción: textualización y notación”. La inmanencia en cuestión III. Tópicos del Seminario. 2015;33:101-130
- [5] Allen S, Wolkowitz C. The control of Women’s Labour: The Case of Homeworking. Feminist Review. 1986;22:25-51
- [6] Pandya V. Movement and space: Andamanese cartography. American Ethnologist. 1990;17(4):775-797
- [7] De Certeau M. The Practice of Everyday Life. University of California Press; 1988
- [8] Vergara A. Etnografía de los lugares: Una guía antropológica para estudiar su concreta complejidad. Ediciones Navarra; 2015
- [9] Fontanille J. Prácticas Semióticas. Fondo Editorial de la Universidad de Lima; 2016
- [10] Harvey D. Espacios de esperanza. Ediciones Akal; 2018
- [11] Sen A. Amartya Sen: Collective Choice and Social Welfare. First Harvard University Press Edition; 2017

Microplastics Environmental Risk Assessment: A Review

*Andrea Arredondo-Navarro, Estefanía Martínez-Tavera
and Deborah Xanat Cervantes Flores*

Abstract

Plastics have been critical in human development due to their unique properties. However, when discarded, they accumulate in the natural environment due to their lack of biodegradability and when exposed to environmental conditions, these polymers become microplastics (MPs). Many experimental studies have detected MPs in the environment, but the identification and separation methods used, work for certain types of MPs while overlooking or damaging others, which can affect the development and application of systematic risk evaluations such as the ecological risk assessments (ERAs). A literature review was conducted of papers published up to 2021 looking for ERAs made for MPs on experimental studies. The eight selected articles were analyzed and scored based on concentration, toxicology, and quality assurance (QA) criteria. The hazard index (HI), the pollutant load index (PLI), the potential ecological risk index (RI), and the risk quotient (RQ) based on species sensitivity distribution (SSD) were compared and assessed. To compare the indexes correctly, modifications were made to previous studies.

Keywords: microplastics, polymer, environmental risk assessment, risk, comparison, review

1. Introduction

Plastics have been critical in human development. Because of their unique properties including a wide range of temperature use, their resistance to chemicals and light, and their mechanical strength, plastics have brought important advances in health, housing, transportation, science, and education. However, like other man-made materials, when discarded, they accumulate permanently in the natural environment due to their lack of biodegradability. As time goes by and they are exposed to the drag of currents other environmental factors (such as temperature changes, mechanical abrasion, solar irradiation), these polymers break down into microplastics (MPs; 1–5000 m in their longest dimension). The MPs generated in this way are called secondary MPs and are the most abundant type in water bodies [1]. Some MPs are intentionally manufactured for specific applications such as microbeads used in exfoliants or detergents and are called primary MPs [1].

The study of MPs as contaminants began in 2007, and since then many experimental studies have been conducted in which MPs are detected and/or identified. The methodology for the characterization of MPs in water includes sampling, separation from the environmental matrix, quantification, and identification. However, some identification and separation methods work for certain types of MPs while overlooking or damaging others, generating studies that are not comparable with each other and hindering possible solutions such as the estimation of ecological risk assessment (ERA), which can provide the basis for future legislation regarding the use and discharge of MPs.

The ERA is an exposure and dose-response study based on the comparison between the concentration of a chemical of interest found in an ecosystem of interest (in this case water bodies) with the data on the expected effects of this compound for a specific group (flora, fauna, or human). The exposure studies are affected by the lack of standard methods, as described before, as different types and sizes of MPs are collected and identified depending on the used methodology. The dose-response relationship is studied with animal or plant toxicology and then extrapolated for humans with some standard values (see [2]) to obtain a safe concentration level. Quantifying this value is basic for proper risk analysis, and it can be observed linearly and clearly in chemicals since only the species studied and the concentration varies. But for MPs, it is not so simple because there are many other variables that can generate the toxicity of the polymer, such as additives, size, shape, color, and type of the MP. For this reason, toxicological studies are also not standardized and are performed with different types, forms, concentrations, and sizes of MPs as well as different organisms and toxicological endpoints.

Because of the reasons presented above, Gouin et al. [3] identified that the adequate development of an ERA should include the following:

- A common definition for MP categorization and exposure metrics.
- A consensus on reporting requirements aligned with the physicochemical properties of the MPs under test and the availability of analytical methods to characterize the polymer material. As well as, the properties of the test system that may influence particle aggregation, agglomeration, sedimentation, dissolution, etc.
- The inclusion of chemical controls to differentiate effects related to the physical aspects of the polymer by itself (form, color, and size) from the ones produced by the associated additives, monomers, or adsorbed pollutants.
- The development of methodological protocols to strengthen reproducibility and interpretability.
- The identification and prioritization of environmentally relevant exposures.
- The identification of sentinel test species based on sensitivity distributions (SSD) of species based on mechanistic understanding of physiological and behavioral traits.
- A consensus on appropriate effect endpoints based on relevant chronic exposure scenarios.

These key points indicate that standardization and information reporting are essentially the basis of developing an ERA. Therefore, the present review aims to help researchers on plastic pollution in water to observe how the ERAs for MPs have been carried out and improve what other authors have made. Although only eight experimental articles will be analyzed, throughout the paper references to authors proposing a framework for conducting risk analysis for MPs will be found [3–5].

The purpose of the review is to analyze and compare experimental articles that have calculated an environmental risk assessment (ERA) of MPs in a water body. The aspects to be compared are: (a) the criteria used to define a water body as safe or not (PNEC, HI, RI, etc.); (b) which aspects are considered by these criteria; (c) the results of each ERA; (d) the concentration data used in the risk estimation; and finally (e) if the experiments performed comply with the necessary quality in terms of sampling, separation, and identification of MPs in water bodies according to Koelmans et al. [6].

2. Methods

A literature review was conducted of papers published up to 2021, which were identified by the Scopus search engine. The search terms were: Risk assessment AND Water OR Waters OR Estuarine OR Surface OR River OR Ocean OR Marine OR Lake AND Microplastics OR Microplastic on the article title. A total of 14 articles matched this search strategy, from which four were excluded because even though they identified MPs in water bodies, they carried out the ERA only for other pollutants (e.g., heavy metals), performed MPs measurements in sediments that were excluded from this review or were Reviews of toxicological articles. The eight selected articles were analyzed and scored based on the criteria observed in **Table 1**.

Environmental risk analysis estimation		Quality assurance (QA) of the experimental article based on Koelmans et al. [6] criteria
Toxicology	Concentration	
Based on chemical composition (monomer, additives, others)	Size range specified	Sampling methods
Based on morphology	Chemical composition specified	Sample size
Based on size	Percentage of each identified polymer type is included in the ERA	Sample processing and storage
Based on other (specify)	Color specified	Lab preparation
	Morphology specified	Clean air conditions
	Site characterization (were MPs found (surface, sediments, mid-level)	Negative controls
	Site characterization (main site activities; population, weather)	Positive controls
		Sample treatment
		Polymer ID
		Limit of detection (LOD) specified.

Table 1.
Criteria to score the environmental risk analysis for MPs selected.

The ERA criteria observed on the first two columns of **Table 1** were based on the articles made by Refs. [3, 4, 7], and the quality assurance column was based on the article made by Koelmans et al. [6]. Both the ERA criteria and the quality controls carried out in the experimental part of the article are considered for this review, as quality is the basis of good data collection and reporting how the experiment is performed is key to reproducibility and future standardization of methods.

The scores given to each criterion are:

- Toxicology: for each criterion used as a basis for the study, a point will be assigned.
- Concentration: for each specified criterion in the article, one point will be assigned. Except for site characterization (main site activities, population, and weather) on which a point is assigned per description; if the study area is broadly described, only one point will be assigned.
- QA: the points will be assigned based on Koelmans et al. [6] ranking scores.
 - One point will be assigned to each specified detection limit (one for size limit, one for shape limit, and one for polymer type limit) and one for each section of the experiment (one for sampling limit, one for separation limit, and one for identification limit).
 - In addition to the criteria established by Koelmans et al. [6], the detection limits must be specified to be able to interpret the results more accurately.

3. Results and discussion

3.1 Toxicology: method used to determine if a site is safe or not

According to the main quoted authors in the compared articles, it can be stated that Lithner et al. [8] and Everaert et al. [9] laid the foundations for the current estimation of ERA related to MPs. The first authors [8] established a hazard level for each polymer based on the toxicity of their monomers, while the second ones [9] proposed to compare the predicted no-effect concentration (PNEC) (to which the RQ was added later), calculated with a species sensitivity distribution (SSD) analysis (a statistical model that is used for other chemical compounds and nanomaterials [10]) with the predicted environmental concentration (PEC).

Some improvements have been made to the method proposed by Everaert et al. [9]. For example, Besseling et al. [11] limited the articles on which they based their SSD to those that studied endpoints assumed to affect population size; Burns and Boxall [12] used ecotoxicity data limited to 10–5000 μm particle size exposures; and Jung et al. [13] screened the toxicity data used for the SSD to fragment and fiber MPs in the size range of 20–300 μm while measuring (instead of predicting) the MPs concentration in a marine environment.

On the other hand, there have been no important improvements to the path taken by Lithner et al. [8] probably because it is difficult to integrate toxicity information

for all monomers or additives associated with polymers. However, many articles are based on the principles they implemented. The methods used in the reviewed articles will be explained below.

3.1.1 Hazard index (HI) and pollutant load index (PLI)

The hazard index (HI) (Eq. (1)) was proposed by Xu et al. [14].

$$HI = \sum P_n * S_n \quad (1)$$

Where HI is the total polymer risk index caused by MPs, P_n is the specific MP polymer type percentage at each sampling site, and S_n is the polymer hazard score established by Lithner et al. [8], that is based on the toxicity of its monomers. Eq. (1) does not consider the concentration of the polymer; therefore, it would not be correct to use it without a complementary index. For this reason, Xu et al. [14] proposed to incorporate the pollutant load index (PLI), an equation proposed by Tomlinson et al. [15] to observe the pollution difference between zones on the same waterbody.

$$PLI = \sqrt{C_f^i} \quad (2)$$

$$C_f^i = C_i / C_{oi} \quad (3)$$

$$PLI_{zone} = \sqrt[n]{PLI_1 PLI_2 PLI_n} \quad (4)$$

Where C_f^i is the MP concentration factor, C_i is the MP abundance of the sampling site and C_{oi} is the minimum average concentration that may be on the available literature or the minimum concentration found on the study area [15].

Eqs. (1) and (4) were used to develop an Ecological Risk Assessment by 62.3% of the reviewed articles [14, 16–18], but Picó et al. [16] made a modification to Eq. (2); $PLI = C_f^i$. These studies could have been comparable easily if all the authors would have used the same C_{oi} (in particles per liter p/L) value or the same equation. This makes the use of data difficult or impossible to compare [14], but some modifications will be made to their calculations, so a comparison is made in the present review.

3.1.2 Potential ecological risk index (RI)

This factor was introduced to the microplastic risk analysis by Pan et al. [17].

$$E_r^i = T_r^i * C_f^i \quad (5)$$

$$T_r^i = P_i / C_i * S_n \quad (6)$$

$$RI = \sum_{i=1}^n E_r^i \quad (7)$$

Where E_r^i is the potential ecological risk index that represents the potential ecological hazard from single MPs polymer, T_r^i is the ecotoxicity response factor, P_i is the proportion of each individual MPs polymer type, S_n is the hazard score of MPs polymer type (i) given by Lithner et al. [8], n is the number of types of MPs polymers contained in the sample, and RI is the potential ecological risk from combined MPs

polymers [17, 18]. It is important to highlight that the calculation of RI allows combining the toxic effects of the polymers and the MPs concentration of the study area. Eqs. (5)–(7) were used by 25% of the reviewed articles [17, 18], and provide additional information for future studies to complement HI and PLI (1)–(4).

Eqs. (1)–(7) are based only on the toxicological information of the polymers [8], without considering the additive, size, form, and other aspects that could generate toxicity in an organism. Therefore, these articles will only have a score of 1 in the toxicology column displayed in the next section (**Table 2**).

The interpretation of the described indexes can be observed in **Table 3**.

The hazard level for the HI calculation observed in **Table 3** is based on Lithner et al. [8], which classifies these levels with the toxicity data of the monomers (S_n). It is important to note that a polymer is not easily depolymerized to the monomeric state during use and even less so when discarded in nature [23]. Additionally, this is an issue that the plastic industry has been improving for many years along with increasingly stringent regulations especially for polymers used in the food industry (packaging and others) which are among the most discarded and found in the environment [24]. Still, the authors [8] argue that polymerization reactions during production are rarely complete, therefore, unreacted residual monomers can be found in the polymeric material.

The PLI hazard level was originally developed by Tomlinson et al. [15] for estuarine environments specifically for bioaccumulation of heavy metals on intertidal algae or fauna. Later, Angulo [25] assessed the articles that calculated the PLI for different organisms and assigned them a risk category which is now used for MP concentrations in the water column. This value must be used carefully as it was originally developed for other purposes but using it in a standardized manner (establishing an international C_{oi} or giving it specific thresholds according to the pollution level of the studied waterbody) could yield comparable studies around the world.

The potential ecological risk index and risk (E_i and RI) are a combination of the chemical toxicity coefficient for each polymer type introduced by Lithner et al. [8] and the C_f^i (MP concentration factor) introduced by Tomlinson et al. [15] and can be taken as the most complete approach until now. As it is a new perspective only estimated by Refs. [17, 18] the risk categories must be calibrated with studies made around the world in more and less polluted areas.

3.1.3 Risk quotient (RQ) and predicted-no-effect-concentration (PNEC) based on species sensitivity distribution (SSD)

This method to estimate the environmental risk is completely different from those discussed in the previous sections and it has been used by 25% of the articles reviewed [13, 21]. It is an estimate of the hazardous concentration for protecting 95% of the species (HC_5), the predicted-no-effect concentration (PNEC, Ec. 8), and a risk quotient (RQ) (relationship between the measured environmental concentration (MEC) and the PNEC; Ec. 9), based on a species sensitivity distribution (SSD), a tool to establish safe limits on chemical concentrations on surface waters.

The general procedure to perform an SSD is: first to compile the results of many toxicity tests performed separately on a given chemical using several species; second, a statistical distribution that best fits the data and; third, the fitted distribution is used to infer a concentration that protects the desired proportion of similar species (HC_5) [26].

Ref.	Study's objective	Equations/method used	Toxicology score	ERA results
[18]	MPs in two natural reserves: Chagan Lake/Xianghai Lake, and ERA	HI/PLI/RI C_{oi}^- 6.65 p/L based on Everaert et al. [9]	1	<i>Chagan Lake</i> HI = 27; PLI = 3.21; RI = 15.03 <i>Xianghai Nature Reserve</i> HI = 15.38; PLI = 2.16; RI = 0.83
[16]	MPs in artificial channels/ponds (mixed surface/treated waste-water) influenced by two Saudi Arabia cities, Riyadh and Al-Jubail.	(modified Eq. (2)) $PLI = \frac{C_{oi}^i}{C_{oi}^-}$ 0.1 (Al Jubail)/1.7 (Riyadh) p/L	1	Same HI (polymer type similar on the two sites); HI = 881 PLI significant differences; Riyadh PLI = 20; Al-Jubai PLI = 2
[19]	MP characterization in surface water of the Manas River Basin	HI/PLI C_{oi}^- 9 (April)/6 (July) p/L	1	<i>April</i> : HI = 82.28; PLI = 1.37 <i>July</i> : HI = 65.08; PLI = 1.52
[17]	MPs ERA in Dongshan Bay	HI/PLI/RI C_{oi}^- 1×10^{-5} p/L (min. published concentration) [20]	1	HI = 12.94; PLI = 14.2; RI = 21.5
[14]	MPs ERA on the surface waters of the Changjiang Estuary and the adjacent East China Sea.	HI/PLI C_{oi}^- 5×10^{-4} p/L	1	<i>East China Sea</i> PLI = 20.4; HI (no reported value); Hazard Level of IV <i>Changjiang Estuary</i> PLI = 18.4; HI (no reported values); Hazard Level of III
[13]	ERA posed by non-spherical MPs in surface/subsurface in seawaters in the coastal, continental shelf, and deep-sea areas of South Korea.	(1) SSD, (2) HC_5 , (3) PNEC	2	HC_5 = 59.8 n/L; PNEC = 12 n/L Mean Mec Mean = 0.09–1.57 n/L Maximum Mec = 7.88 n/L
[21]	ERA of MPs in the Yongjiang River (drinking water source for Nanning City)	(1) SSD, (2) HC_5 , (3) PNEC, (4) RQ	0	RQ = 0.10–1.56 (near tributaries). RQupstream = 0.26; RQmid stream = 0.75; RQ downstream = 0.33
[22]	Examine the effects of marine regions, polymer types, and aging effects on MPs bound DLCs; (2) assess the DLC health risks based on bioassays and predicted dioxin-like effects to marine organisms	(1) Comparison of already estimated EC_{10} (similar k_{ow} compounds) with DLC (PHAs/PCBs) concentrations desorbed from MPs	4.5	The concentration of DLCs on MPs was influenced by regions, polymer types, and aging, but not plastic size. <i>DLCs on unaged pellets</i> Concentration << EC_{10} <i>DLCs on aged pellets and styrofoams</i> Concentration >> EC_{10}

Those colored in gray represent the studies that were performed with Eqs. (1)–(7) and those that are not colored represent those that used a probabilistic approach.

Table 2.
Reviewed articles; studied areas, equations/method used to develop the ERA, toxicology score, and ERA results.

HI (range)	0–1	1–10	10–100	100–1000	>1000
Hazard level	I	II	III	IV	V
PLI (range) [14]	<10		10–20	20–30	>30
Hazard level	I		II	III	IV
Potential ecological risk assessment					
Ei (range)	<40	40–80	80–160	160–320	>320
RI (range)	<150	150–300	300–600	600–1200	>1200
Risk category	Minor	Medium	High	Danger	Extreme danger
Index interpretation in terms of hazard level and risk category.					

Table 3.
Risk assessment of MPs pollution—MPs-induced risk index (HI) and pollution load index (PLI).

$$PNEC = HC_5 / \text{Assessment Factor (AF)} \tag{8}$$

$$RQ = MEC / PNEC \tag{9}$$

To estimate a PNEC for a single chemical compound using an SSD correctly, it is necessary to include a minimum of 10 chronic toxicity data and to cover eight taxonomic classes. If it does not accomplish these requirements and others (see ECHA [27]), a value of five is given to the AF. The usual requirements needed for chemical compounds are not always met by toxicological studies for MPs, in addition to uncertainties in these studies not covered by the AF. These can include; differences in the polymers used, MP morphologies, and MP sizes and colors used in the selected toxicological articles that may cause different toxic effects on the studied organisms [11].

As these articles are based on toxicological studies, the authors of this type of ERA may include articles in which the toxic effects are generated by different aspects of the MP. For example, by choosing studies where only MPs of a certain type or a certain size were used, the calculated PNEC value can be said to include type and size toxicology because both may trigger the toxic effects on which the PNEC was based. Therefore, the score given to the Refs. [13, 21] articles on the toxicology column displayed in the next section (**Table 2**), will range from 0 (when they did not screen any toxicological criteria) to 2 (when they screened specific toxicological criteria).

There are many loose ends in the development of ERAs with the SSD method. Some authors suggest addressing them with a probabilistic approach. For example, Koelmans et al. [28] proposed and tested a rescaling method to improve the alignment of SSD. Their objective was to englobe the diversity of MPs found in the environment via continuous probability density functions. They corrected the size ranges differences, developed a method to convert number to volume and mass concentration (or vice versa), and one to correct for differences in the MPs sizes, shapes, and densities. However, the authors conclude that their study will be more enriching when future experimental studies are developed with a higher level of quality in terms of sampling, laboratory control, and identification of MPs.

3.1.4 Toxicity based on adsorbed pollutants

One of the main worries of researchers around the world is the polymer’s high sorption capacity toward hydrophobic organic contaminants (HOCs). MPs can

enhance HOCs transportation resulting in an extensive distribution of MP-associated chemicals entering every level of the food chain in the marine environment [19, 29]. Therefore, the study made by Chen et al. [22] is included in the present review, in order to consider this associated environmental risk.

The mentioned authors compared concentrations of dioxin-like polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyl (PCBs) extracted from micro/meso plastic debris from two plastic accumulation zones in marine environments with the effect concentration at which 10% of the species are affected EC_{10} , or EC_{10} from chemicals with similar octanol-water partitioning constants (K_{ow} s) when the values were not available. They also reported variations in the concentration of these toxic compounds according to descriptive parameters including the type of polymer, form, size, adsorption capacity, and weathering. Therefore, their score on the toxicology column (where a point is given for each descriptive parameter) is the highest.

This is an unexplored way of performing risk analysis for MPs, but should be explored further, since as mentioned throughout the paper, MPs have different ways of being toxic.

3.2 Toxicology: results

The following table shows the articles reviewed, the studied area, the methods each one used to perform their environmental risk analysis, their toxicology score explained before, and the results they obtained on their ERA.

The first thing to observe from **Table 2** is that the authors [14, 16–19] used the same basis (Eqs. (1)–(7)) or small variations of them (e.g., Picó et al. [16]). Therefore, they will be compared first.

The HI values vary a lot from study to study (as observed in **Figure 1**) and are difficult to compare since the polymer types detected in the study vary according to the sampling depth (surface, mid-level, or bottom), the waterbody, the degradation compound used, and the identification method. Though, Picó et al. [16] study has the highest HI value (see **Figure 1** and **Table 2**), which matches with also having the highest PLI value for Riyadh, it is important to mention that these are the result of a more conservative approach by the authors, intended by modifying Eq. (2).

The RI was only calculated in two articles, one of freshwater environment [18], and the other one on a marine environment [17]. Considering these differences, direct comparison (by recalculation) will not be performed for this index as it varies according to the type of polymers found in the area and the concentration (it will be recalculated for the PLI comparison below).

Another aspect to notice from **Table 2** is that [14, 16–19] studies could have been comparable if they have used the same C_{oi} . Fortunately, most of them included sufficient data to calculate the indices with a modified and equal C_{oi} value and be able to compare the results in the present review. The C_{oi} that will be used to compare between surface or inland waters [16, 17, 19] will be 0.15 p/L since it is the lowest value measured in the three articles, and the modification made by Picó et al. [16] on Eq. (2) will be recalculated with the original equation to get an equivalent PLI. The C_{oi} that will be used to compare marine environments will be 1×10^{-5} p/L as it is the lowest concentration measured in the two articles [14, 17]. The reason that the lowest concentration values were chosen is that this approach intends to be a conservative risk estimation.

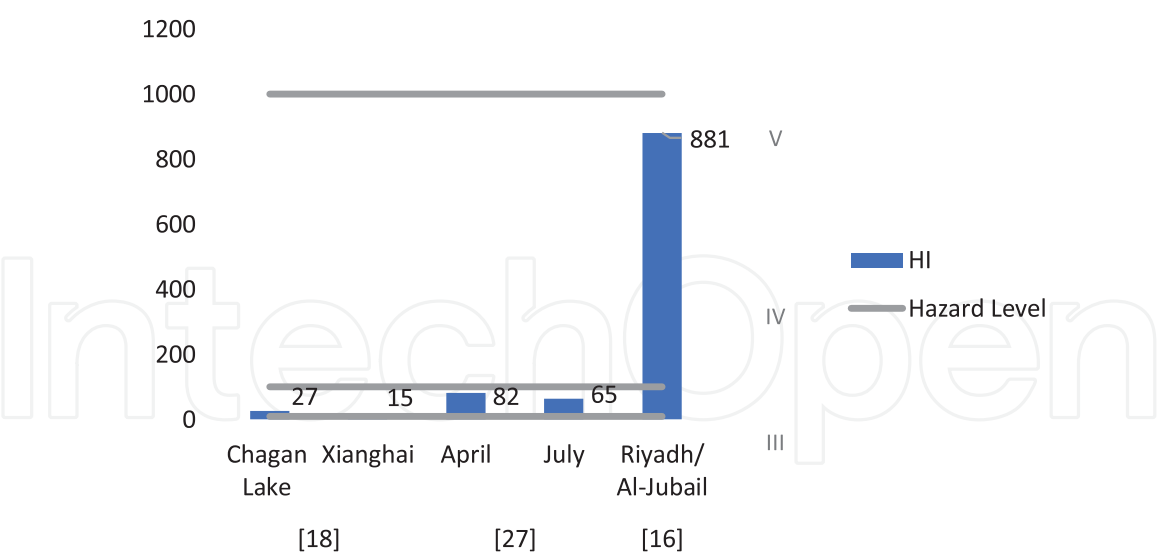


Figure 1.
Hazard Indexes results from five different study areas.

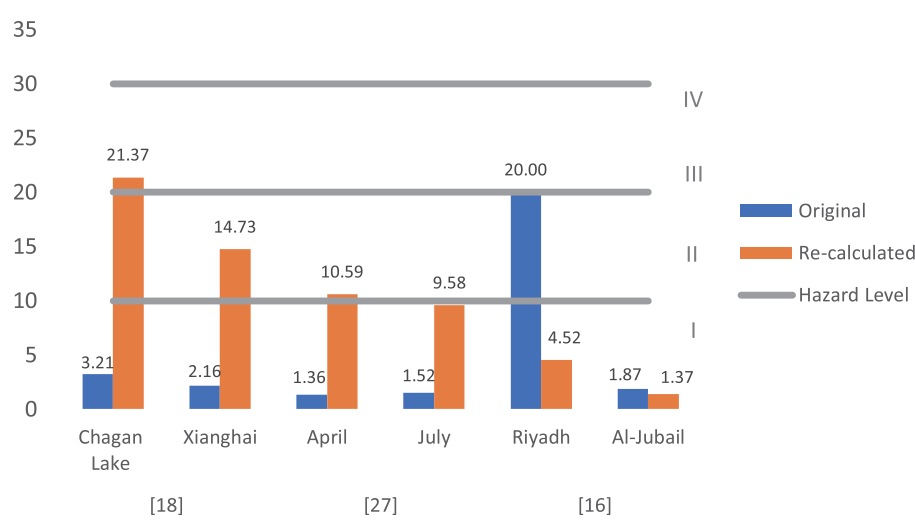


Figure 2.
Original (blue) and modified (orange) PLI results for surface/inland water environments.

Figure 2 and 3 exhibit the original PLI values and the results of the modifications for surface or inland waters and marine environments respectively.

Figure 2 indicates that with the modified C_{oi} in the first two articles observed from left to right, the PLI value increased at least six times. While in the last article [16] (from where the C_{oi} value was chosen) and the modification made to Eq. (2) the value decreased six times, indicating that their original approach was much more conservative than the others. The hazard level (HL) increased from I to III for Yin et al. [18], I to II for Refs. [18, 19], and from I to almost II for Wang et al. [19], while decreasing from II to I in the Riyadh artificial surface channel and remaining in I in the Al-Jubail artificial water pond [16] (consult Lithner et al. [8] for each HL definition).

The three study areas are Asian waterbodies [18, 19], and were both conducted in China and Picó et al. [16] in Saudi Arabia.

The highest PLI values are found on the Chagan Lake and the Xianghai reserve. This may be the result of them both being lentic ecosystems (a lake and a wetland) with no significant water flow, being accumulation systems in which the

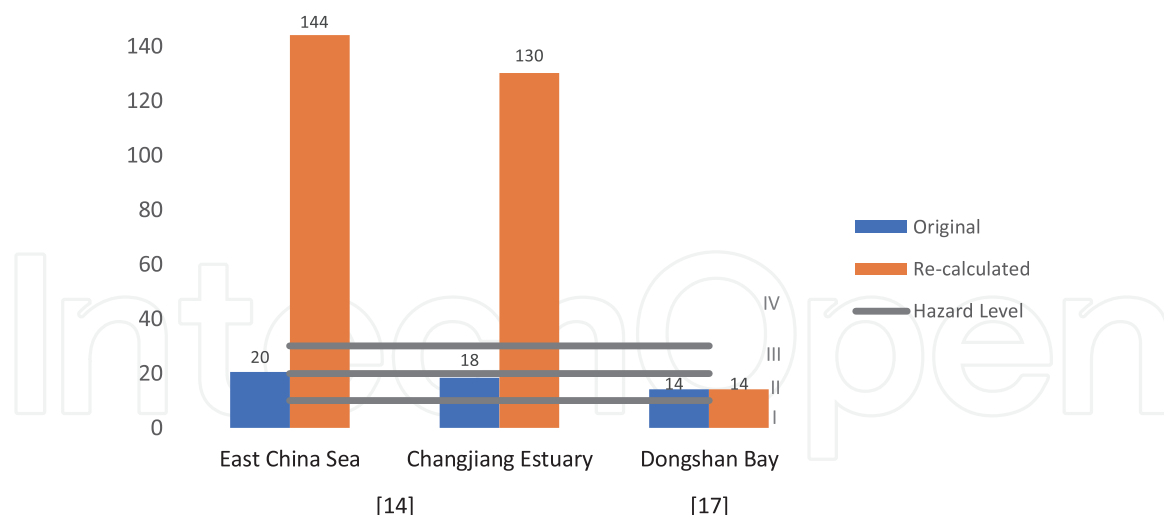


Figure 3.
 Original (blue) and modified (orange) PLI results for marine environments.

concentration of perennial pollutants such as MPs will increase over time. The difference in the PLI values between the Chagan Lake and the Xianghai Nature Reserve can be explained by the fact that in the lake, tourism, and fishing (activities that generate MP pollution) are carried out, while in the natural reserve these activities are less common but occur. Overall, the results are consistent with the findings of previous papers that establish that the concentrations of MPs in China are the highest [29]. Sampling methods may influence the results, as the smallest MPs that they were able to sample are 1, 20, and 38 μm . However, there is no relationship between the lowest detection limit (1 μm) and the highest PLI value.

PLI hazard levels for freshwater (**Figure 2**) are much lower (maximum of III) compared to PLI values for marine environments (**Figure 3**). This is attributed to the low C_{oi} used in the modified PLIs and the high concentrations of MPs found in Chinese estuarine or coastal environments.

Although Dongshan Bay has several sources of MPs, such as wastewater discharge, seawater farming, and a special semi-closed terrain, it has the lowest PLI modified value [17], which is also the same as the original value. This may be due to the sampling method they used (Manta Trawl with 330 μm of pore opening) versus the sampling method used by Xu et al. [14] (pump with a coupled mesh of 70 μm of pore opening), which results in smaller MPs being collected, thus a larger MP concentration. The modified C_{oi} value may be sub-estimated too, as it is based on the Dongshan Bay study but may be used as a base for future ERAs.

In two of the remaining articles [13, 21] the authors performed the risk analysis based on an SSD explained previously in the document. The authors [21] were the first ones on estimating a risk analysis based on this statistical approach proposed by Everaert et al. [9] and comparing it with real measured data (MEC). While Jung et al. [13] used the same approach but limited their study to certain toxicological MP forms (fragment and fiber) and in the size range of 20–300 μm .

The article performed by Zhang et al. [21] (one of the first articles to perform SSD), did not include the confidence intervals (CI). Later, Jung et al. [13] argued that the screening they performed, accomplished a confidence interval reduction from 52 to 19, when they included the toxicity data for all types of MP for the SSD derivation, regardless of shape and size. They also mentioned that the articles made by Refs. [9, 11, 12] reported IC values greater than 100, proving that the limitation of certain

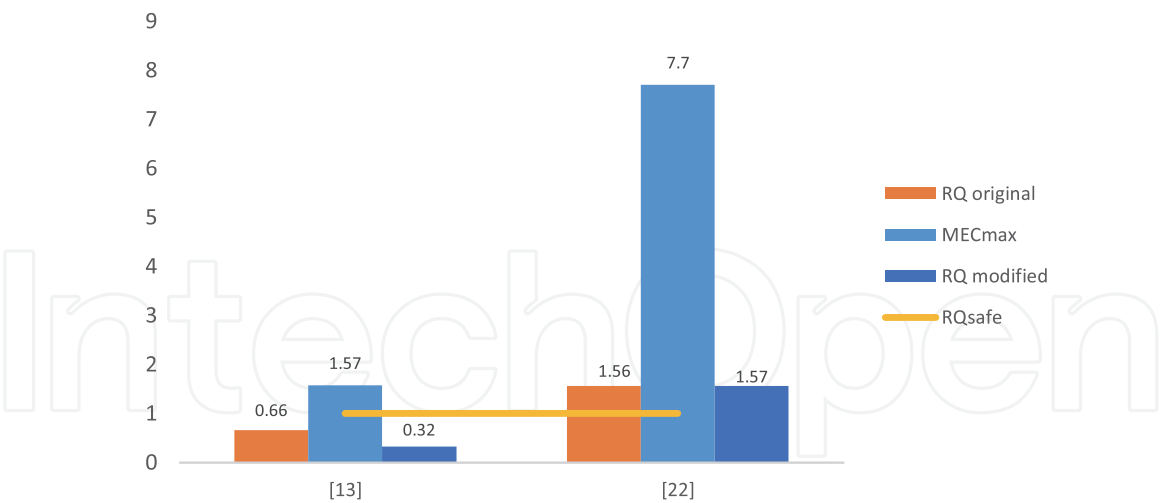


Figure 4.
The measured environmental concentration (MEC) and the risk quotient (RQ).

shapes, colors, sizes, and types may improve the statistical reliability when deriving a PNEC value.

In **Figure 4** the values of the obtained MEC, the minimum and maximum RQ values measured on the study area can be observed, as well as the RQ safe limit (that is one), can be observed.

The observed differences between studies are related to differences in sampling sites and methodologies. The risk analysis by Jung et al. [13] was made on a marine environment, they based their SSD on chronic toxicity data of marine and freshwater species and limited their study to certain form and size of MPs. On the other side, the study by Zhang et al. [21] was performed on a river, based their statistical analysis on toxicity data of freshwater organisms, and did not screen any MP feature. Also although, Jung et al. [13] used a sampling method with a pore opening of 20 m and Zhang et al. [21] of 50 m, their MEC and RQ values are lower, which agrees with other studies that indicate that freshwater systems have higher MPs concentrations [30].

The approach taken by Jung et al. [13] is a good start to delimit statistical analyses and make their results more reliable. The selected PNEC to compare the RQ value (Eq. (9)) was the one obtained by Zhang et al. [21] in their SSD analysis. It was chosen since it was the lower of the two analyses (Jung et al. [13] obtained a value of 12) and as in previous comparisons, the lower concentrations are chosen to have a more conservative approach to this pollution problem.

The last study Chen et al. [29], used a very different approach than the others. Therefore, there will not be a comparison as such, but the results will be commented on.

The authors discovered that the DLCs adsorption on plastic particles is influenced by the coastal and oceanic regions, polymer types, aging effects but did not find a relationship between the size and the sorbed pollutants. For pollutants on unaged pellets, the concentration is lower than the EC₁₀ thresholds, and dioxin-like effects may not occur (EC₁₀ can replace the no observed effect concentration (NOEC)). However, for aged pellets and styrofoam PAHs and PCBs concentrations are higher than the EC₁₀ on 11 out of 11 scenarios for aged pellets and on 2 out of 11 scenarios for styrofoam. In addition, dioxin-like effects are likely to occur once these MPs are exposed to marine organisms [29].

The toxicological score given to these three articles is 2 for Jung et al. [13], on which the toxicological studies were screened (form and size), therefore the ERA is

developed from two distinct approaches; 0 for Zhang et al. [21] as they did the SSD without screening the toxicological studies, and 4.5 for Chen et al. [29] because they compared the toxic effects of the plastic particle type (0.5 because they only compared three types), weathering, size, form, and adsorbed pollutants.

When conducting ERAs, it is recommended that a standard C_{oi} is used in addition to the baseline found in the study zone to be able to compare them with other risk analyses in different areas. The SSD can be improved by screening information, as mentioned above. Sampling, separation, and identification methods should also be reported to understand differences in results when these do not make sense to the naked eye as in the case of Dongsham Bay. For this reason, a score will be given to the experimental studies compared in this review.

3.3 Concentration and quality assurance

The scores observed in **Table 4** are based on the criteria explained in **Table 1**.

It is important to highlight that in **Table 2**, Wang et al. [19] has the highest scores, while one of the lowest in **Table 4**. In the quality assurance section, specifically in the sampling section no date, coordinates, surface depth, sample size, processing, and storage data are specified. In the laboratory preparation, no clean air specifications, sample treatment, positive nor negative controls are specified. It is essential to meet all the criteria listed in **Table 1** or at least report whether they were considered or why they could not be met.

The highest score is accomplished by Chen et al. [22], they reached the highest punctuation of the concentration section, as they reported and did every criterion. Unlike the quality section, in which they did not achieve the points for the minimum sample size criterion of 500 L, they did perform a negative control but without triplicate, and like all the other articles they did not report a positive control (see Koelmans et al. [6]).

Only Pan et al. [17] had the two points that represented a sampling size higher than 500 L, as they used a special-made manta trawl net, and other authors used sampling containers or pumps. This can be also related to their low MP concentration reported, as the nets pore opening is of 300 μm , which might allow from MPs below 300 μm to not be accounted for.

The LOD criterion was one of the least accomplished, although each step of the experimentation should have detection limits in terms of size, type, and shape of the

Ref.	Toxicology score	Concentration score	QA score
[18]	1	7	11.5
[16]	1	9	12.5
[19]	1	7	9
[17]	1	8	11
[14]	1	7	11
[13]	2	4	12
[21]	0	7	11
[22]	4.5	4.5	4

Table 4.
Toxicology score, concentration score, and quality assurance (QA) score for the reviewed articles.

polymer. For example, if the surface of a waterbody is sampled, it is likely that only polymers with a certain density are being recovered because more dense polymers are usually found at the bottom. The problems arising from density difference also apply to density separation. In addition, in the stereoscopic visualization step, certain colors, and sizes may not be clearly visible. Finally, in the detection step, the micro-Raman is known to be sensitive to fluorescence so it may miss certain colors of MPs [30].

In any case, it is important to use reviews and articles already published to identify these detection limits and report them in any experimental work. Especially if a risk analysis is to be done, in which the study will likely be compared with others where different methods are used.

4. Conclusions

The present review highlights the need to adequately create an environmental risk analysis for MPs, as it allows for better and simpler comparisons between studies, a way to communicate the hazards of MPs to society, and a basis for future regulations. The approaches used by the authors of the reviewed articles can serve as a foundation for future risk analysis. Still, they have some limitations, such as not including all the ways in which a MP can be toxic, not including all the uncertainties in the statistical analysis of SSD, not indicating the differences in their experimental methodologies (that generate difficult-to-compare articles), and others. Some corrections were made to the reviewed articles to determine more comparable parameters. In every case, significant variations were obtained due to the difference in the detection methods used or the studied waterbody.

The most integral approach is the one in which the RQ or PNEC values are estimated based on an SSD, because the toxicological studies included many types, sizes, shapes, and concentrations of MPs, as well as several control species. In the future, this approach may narrow the toxicological studies used as mentioned above and with the aid of statistics will probably become more and more accurate overtime. The least complete approach is the one based on the toxicity of the monomers as it does not consider any other toxicological aspects of the MP.

The article also highlights the importance on the reporting of the MPs detection methodology and the importance of QA in experimental articles.

Other authors may use this article to identify some key parameters to make comparisons between experimental articles and risk analysis of MPs.

Acknowledgements

AAN would like to thank UDLAP and CONACYT for making her work possible.

Conflict of interest

The author confirm that this article content has no conflicts of interest.

IntechOpen

Author details


Andrea Arredondo-Navarro^{1*}, Estefanía Martínez-Tavera²
and Deborah Xanat Cervantes Flores¹

1 University of the Américas Puebla (UDLAP), Puebla, Mexico

2 Popular Autonomous University of the State of Puebla (UPAEP), Puebla, Mexico

*Address all correspondence to: andrea.arredondono@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Lambert S, Wagner M. Environmental performance of bio-based and biodegradable plastics: The road ahead. *Chemical Society Reviews*. 2017;**46**:6855-6871
- [2] Fowle J III, Dearfield K. *Risk Characterization Handbook*. Washington, DC: Science Policy Council; 2000
- [3] Gouin T, Becker R, Collot A, Davis J, Howard B, Inawaka K, et al. Toward the development and application of an environmental risk assessment framework for microplastic. *Environmental Toxicology and Chemistry*. 2019;**38**:2087-2100
- [4] Hung C, Klasios N, Zhu X, Sedlak M, Sutton R. Methods matter: Methods for sampling microplastic and other anthropogenic particles and their implications for monitoring and ecological risk assessment. *Integrated Environmental Assessment and Management*. 2021;**17**:282-291
- [5] De Ruijter V, Redondo-Hasselerharm P, Gouin T, Koelmans A. Quality criteria for microplastic effect studies in the context of risk assessment: A critical review. *Environmental Science and Technology*. 2020;**54**:11692-11705
- [6] Koelmans A, Mohamed Nor N, Hermsen E, Kooi M, Mintenig S, De France J. Microplastics in freshwaters and drinking water: Critical review and assessment of data quality. *Water Research*. 2019;**155**:410-422
- [7] Brander S, Renick V, Foley M, Steele C, Woo M, Lusher A, et al. Sampling and quality assurance and quality control: A guide for scientists investigating the occurrence of microplastics across matrices. *Applied Spectroscopy*. 2020; **74**:1099-1125
- [8] Lithner D, Larsson A, Dave G. Environmental and health hazard ranking and assessment of plastic polymers based on chemical composition. *Science of the Total Environment*. 2011;**409**:3309-3324
- [9] Everaert G, Van Cauwenberghe L, De Rijcke M, Koelmans A, Mess J, Vandegehuchte M, et al. Risk assessment of microplastics in the ocean: Modelling approach and first conclusions. *Environmental Pollution*. 2018;**242b**: 1930-1938
- [10] USEPA. Species Sensitivity Distribution (SSD) Toolbox [En línea]. 2021. Available from: <https://www.epa.gov/chemical-research/species-sensitivity-distribution-ssd-toolbox> [Último acceso: December 2021]
- [11] Besseling E, Redondo-Hasselerharm P, Foekema E, Koelmans A. Quantifying ecological risks of aquatic micro- and nanoplastic. *Critical Reviews in Environmental Science and Technology*. 2019; **49**:32-80
- [12] Burns E, Boxall A. Microplastics in the aquatic environment: Evidence for or against adverse impacts and major knowledge gaps. *Environmental Toxicology and Chemistry*. 2018;**37**:2776-2796
- [13] Jung J, Park J, Eo S, Choi J, Song Y, Cho Y, et al. Ecological risk assessment of microplastics in coastal, shelf, and deep sea waters with a consideration of environmentally relevant size and shape. *Environmental Pollution*. 2021;**270**: 116217
- [14] Xu P, Peng G, Su L, Gao Y, Gao L, Li D. Microplastic risk assessment in surface waters: A case study in the Changjiang Estuary, China. *Marine Pollution Bulletin*. 2018;**133**:647-654

- [15] Tomlinson D, Wilson J, Harris C, Jeffrey D. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer Wissenschaftliche Meeresuntersuchungen*. 1980;**33**:566–575
- [16] Picó Y, Soursou V, Alfarhan A, El-Sheikh M, Barceló D. First evidence of microplastics occurrence in mixed surface and treated wastewater from two major Saudi Arabian cities and assessment of their ecological risk. *Journal of Hazardous Materials*. 2021; **416**:125747
- [17] Pan Z, Liu Q, Jiang R, Li W, Sun X, Lin H, et al. Microplastic pollution and ecological risk assessment in an estuarine environment: The Dongshan Bay of China. *Chemosphere*. 2021;**262**:127876
- [18] Yin K, Wang D, Zhao H, Wang Y, Guo M, Liu Y, et al. Microplastics pollution and risk assessment in water bodies of two nature reserves in Jilin Province: Correlation analysis with the degree of human activity. *Science of the Total Environment*. 2021;**799**:149390
- [19] Wang G, Lu J, Li W, Ning J, Zhou L, Tong Y, et al. Seasonal variation and risk assessment of microplastics in surface water of the Manas River Basin, China. *Ecotoxicology and Environmental Safety*. 2021
- [20] Pan Z, Guo H, Chen H, Wang S, Sun X, Zou Q, et al. Microplastics in the Northwestern Pacific: Abundance, distribution, and characteristics. *Science of the Total Environment*. 2019;**12**:141–151
- [21] Zhang X, Leng Y, Liu X, Huang K, Wang J. Microplastics' Pollution and Risk Assessment in an Urban River: A Case Study in the Yongjiang River. Nanning City, South China: Exposure and Health; 2020
- [22] Chen Q, Zhang H, Allgeier A, Zhou Q, Ouellet J, Crawford S, et al. Marine microplastics bound dioxin-like chemicals: Model explanation and risk assessment. *Journal of Hazardous Materials*. 2019
- [23] Han M. 5-Depolymerization of PET Bottle via Methanolysis and Hydrolysis, de Recycling of Polyethylene Terephthalate Bottles. William Andrew Publishing; 2019. pp. 85-108
- [24] FDA, 21 CFR 177. 2021
- [25] Angulo E. The Tomlinson Pollution Load Index applied to heavy metal, Mussel-Watch data: A useful index to assess coastal pollution. *The Science of the Total Environment*. 1996;**43**:1277
- [26] Center for Computational Toxicology and Exposure, EPA's, Species Sensitivity Distribution (SSD) Toolbox, The United States Environmental Protection Agency's Center for Computational Toxicology and Exposure. 2020
- [27] ECHA. Chapter R.10: Characterisation of dose [concentration]-response for environment, de Guidance on information requirements and chemical safety assessment. 2008
- [28] Koelmans A, Redondo-Hasselerharm P, Nor N, Kooi M. Solving the nonalignment of methods and approaches used in microplastic research to consistently characterize risk. *Environmental Science & Technology*. 2021;**54**: 12307–12315
- [29] Chen H, Qin Y, Huang H, Xu W. A regional difference analysis of microplastic pollution in global freshwater bodies based on a regression model. *Water*. 2020;**650**-2: 1913-1922

[30] Strady E, Dang T, Dao T, Dinh H, Do T, Duong T, et al. Baseline assessment of microplastic concentrations in marine and freshwater environments of a developing Southeast Asian country, Viet Nam. *Marine Pollution Bulletin*. 2021;**162**:111870

The Roll of Different Kind of Fungi to Eliminate Lignin and Organochlorines: A Review

Anaid Bautista-Guerrero and Jose Luis Sanchez-Salas

Abstract

The lignin biodegradation initiate mainly by the fungi organisms and particular interest is focus on the white-rot fungi because apart to use cellulose as carbon source and energy, also can degrade lignin by ligninolytic enzymes such as lignin peroxidase (LiP), manganese peroxidase (MnP), laccases (Lac), and versatile peroxidase (VP). Interesting results, are reported that these enzymes are involve in the degradation of some persistent organic pollutants such as organochlorines. The mechanism to degrade the organochlorines, is thanks apparently by laccase enzymes and other mono and dioxygenases like MnP and VP which can transform chlorophenols to innocuous by-products facilitating the uptake by the fungus or other microorganisms. The use of these different fungus or their enzymes can help to reduce the pollution on the environment (water and soil systems) reducing the negative effects on aquatic animals, plants, and animals including the human being.

Keywords: lignin, biodegradation, laccase, enzymes, fungi

1. Introduction

Lignin is one of the most abundant aromatic heteropolymer on earth, it is a component of plant cell walls, along with cellulose and hemicelluloses [1, 2]. The structure of lignin varies depending on the type of plant, however, three main types are recognized: G-type lignin (guaiacil), found in hardwoods; GS type lignin (syringil) present in softwoods; and GSH lignins (p-hydroxyphenyl) in grasses [3, 4]. Lignin forms a non-reactive and insoluble material around the layer rich in cellulose, which stiffens the cell wall and protects it from microbial attack. The value of lignin in the industry becomes more promising, as it has great potential in the production of chemicals with higher value, requiring its depolymerization into oligo and monomeric aromatic compounds as well as uniform sets of desired aromatic products [4]. However, the natural degradation process is not completely clear. Some detailed molecular studies of lignin degradation are limited by the lack of methods to analyze the decomposition of it. C-labeled lignin has been used to identify several degraders [5]. Generally, tissues with a higher content of lignin, polyphenol, and wax, decompose more slowly [6]. The resistance of plant biomass biodegradation is directly

related to the presence of lignin. According to some authors, in nature, lignin mineralization is an enzyme-dependent process, which is catalyzed by complex ligninolytic enzyme systems composed of extracellular oxidoreductases, such as laccase (Lac) and peroxidases [7].

2. Persistent organic pollutants

Persistent organic pollutants (POPs) are chemicals that are resistant to degradation in the environment and biota, bioaccumulate and are toxic [8]. These contaminants include various lipophilic compounds, which accumulate primarily in lipid-containing tissues such as adipose tissue and move within the body attached to lipids and accumulate during life become a source of chronic internal exposure as they are released continuously from the adipose tissue to the circulation and vital organs with lipid content. POPs cover a variety of lipophilic chemicals resistant to environmental degradation that bioaccumulate in food webs and living organisms including humans [9]. Biomagnification is produced in humans when several orders of magnitude being higher than in the environment. Organochlorine pesticides such as DDT, lindane, chlordane, and hexachlorobenzene are typical examples of POPs. Other persistent organic pollutants are produced as industrial chemicals or by-products, including polychlorinated biphenyls, polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), and polybrominated diphenyl ethers [9].

On the other hand, polycyclic aromatic hydrocarbons (PAHs) are one of the main types of environmental pollutants that often occur during fuel combustion, and are widely distributed in our environment. The effects that these pollutants cause on human health, like POPs, are cause for concern since they can damage the nervous and endocrine systems, they can even become carcinogenic even in low concentrations [10, 11].

3. Generation of organochlorine compounds by some industries

The intensive development of agriculture and industry produces large quantities of plant raw materials, as well as the presence of pesticides, toxic xenobiotics, polycyclic aromatic hydrocarbons, chlorophenols, metals, organochlorine compounds, among other pollutants [7]. Once released into the environment, they decompose very slowly in air, water, soil, and in living organisms [11]. Particularly, these compounds, which are environmental pollutants, mainly used as biocides of wide spectrum in industry and agriculture, their formation during the bleaching of the pulp on papermaking process and the direct discharge of industrial waste [12]. A classic example is the dichlorodiphenyltrichloroethane (DDT), which has been widely used in the last decades as an insecticide for the protection of crops and for the control of vector-borne diseases such as typhus and malaria [13].

The most common sources of organochlorine pollutants, apart from the paper industry, are the combustion of organic matter, partial transformation of phenoxy pesticides such as 2,4-dichlorophenoxyacetic acid and 2,4,6-trichlorophenoxyacetic acid, treatment of wood against fungi and insects and preservation of raw hides in leather tanning industries [12].

The toxicity of these compounds tends to increase with their degree of chlorination. Among chlorinated phenols, pentachlorophenol (PCP) has been widely used

as a preservative in wood and leather due to its toxicity to bacteria, mold, algae, and fungi in these materials [12].

On the other hand, inorganic chloride (Clinorg) supplied to soil through atmospheric deposition can undergo plant absorption, leaching, and react to form organochlorines (Clorg). Natural chlorination can occur through biotic and abiotic processes under environmental conditions. Also, the Clinorg becomes Clorg during the decomposition and humification of the vegetal material in the surface of soil [14].

4. Organochlorinated contaminants and their toxic effects

Organochlorine pesticides (OCP) are among the most abundant global pollutants with high concern. According to the Stockholm Convention, POPs included nine internationally produced and extensively used in agricultural crops or vector control. These compounds are highly toxic and potentially carcinogenic according to the US Environmental Protection Agency (USEPA) stated since 1980 [15]. One of these compounds is pentachlorophenol, which is toxic to any living organism, as it is an inhibitor of oxidative phosphorylation. It is also a pollutant recalcitrant to degradation due to its stable aromatic ring and high chloride content, thus persisting in the environment [12].

The regulation of exposure to these persistent chlorinated pollutants by the population continues, mainly through the consumption of fatty foods of animal origin that previously were fed with contaminated fodder. The resistance of persistent organic compounds to chemical and metabolic degradation implies that they concentrate more as they move through food networks [9]. Despite the prohibitions imposed on most of these pollutants for several years, numerous investigations have reported the continuing and ubiquitous presence of organochlorines in the global atmosphere [15].

Exposure to these pollutants is associated with effects on human health, including cancer, reproductive defects and behavioral changes, which are related to the alteration of the functions of some hormones, growth factors, enzymes, and neurotransmitters [11]. Organochlorine pesticides such as dichlorodiphenyltrichloroethane (DDT) was widely used before recognizing their toxicity and persistence. Exposure to pesticides has been associated with arthritis, breast cancer, and diabetes [8]. Actually, these organochlorine compounds are bio-magnified through the food chain. In the case of humans, when consuming foods of animal origin such as meat, fish, milk, among others, they are exposed to high levels of these pollutants because their biodegradation is slow and they accumulate in the human body as mentioned above. An example is DDT, which can remain for 50 years in the body, mainly in fatty tissue, its distribution is through the bloodstream in plasma or lipids. Also, these contaminants cross the placental barrier reaching the fetus and is secreted in breast milk [12]. This compound has been detected at high levels in liver, brain, along the food chain [13]. Other examples are dichlorodiphenyldichlorethylene (DDE) and methoxychlor, which, although they are weak as estrogens and antiandrogens. They have been found to accumulate in fatty tissues in organisms and are associated with disease, including cancer [16].

Another problem of the organochlorine compounds due to the low biodegradation, is their sediment deposition, which straight to substantial increases in the residence time of these compounds in aquatic ecosystems, thus facilitating bioaccumulation. These pollutants are known to bind principally with organic matter as the humic acids present in the sediments due to their hydrophobicity [15].

5. Ligninolytic fungi

The methods of chemical and physical lignin degradation have several disadvantages, including their high costs and the possibility to produce secondary contamination as mentioned above [12]. In contrast, biological treatment or biodegradation is an attractive option as it involves energy saving and is environmentally friendly [17].

Bioremediation involves the use of living organisms, usually bacteria or fungi, to remove contaminants from soil and water. Thus, there is an intense interest in white rot fungi that has the ability to degrade some extremely persistent or toxic environmental pollutant [18]. Microorganisms have developed sophisticated metabolic and enzymatic systems for the degradation and conversion of lignin in nontoxic monomers, it is believed that white rot fungi are the main tool for the depolymerization of lignin, since they can secrete a collection of peroxidases and ligninolytic laccases that degrade oxidatively lignin to produce small aromatic molecules [4]. The use of fungi for biodegradation has already been investigated for several decades, one of the best studied is *Phanerochaete chrysosporium* [17]. Most of the white rot fungi belong to basidiomycetes and are responsible for the complete mineralization of the woody components including lignin [19]. Other ligninolytic enzyme-producing fungi are species of ascomycetes such as *Trichoderma reesei*. However, in nature, the degradation of lignin during the process of decomposition of the wood is possible mainly by basidiomycetes [20].

A large variety of white rot fungi simultaneously attacks lignin, hemicellulose, and cellulose, while some others only attack mainly lignin. For example, *Ceriporiopsis subvermispora*, *Phlebia* spp., *Physisporinus rivulosus*, and *Dichomitus squalens* selectively attack lignin, since, the purpose of these fungi is to break down the lignin barrier and make cellulose and hemicellulose accessible to these microorganisms [21]. *Trametes versicolor*, *Heterobasidium annosum*, *P. chrysosporium*, and *Irpex lacteus* simultaneously degrade all components of the cell wall [20].

Another fungus belonging to white rot is *Pleurotus ostreatus*, which produces laccase and manganese peroxidase, but does not secrete lignin peroxidase [22]. *Pleurotus ostreatus* expresses multiple laccase genes encoding isoenzymes with different properties, the physiological meaning of this multiplicity being as yet unknown [23].

Experiments carried out with *Myrothecium verrucaria*, a fungus belonging to ascomycetes, showed that this fungus has the ability to secrete enzymes such as lignin peroxidase (LiP), manganese peroxidase (MnP), and laccase (Lac), and has been used as a biological pretreatment to reduce the lignin content in corn stover [24]. Other important enzyme for lignin degradation is the versatile peroxidase (VP), it appears to be produced from different fungus genera as *Pleurotus*, *Bjerkandera*, *Lipista*, *Panus*, and *Trametes* species [18].

Particularly, the degradation of lignin by *Ceriporiopsis subvermispora*, is produced by a mechanism of a single electron-oxidation closely dependent on the action of MnP. During the decomposition of the wood by *C. subvermispora*, it releases a series of unsaturated fatty acids including linoleic acid, which, according to some studies, was found during the first week of culture before lignin degradation. This reaction seems to depend on the presence of unsaturated fatty acids that are peroxidized by MnP generating organoperoxy radicals capable of oxidizing non-phenolic lignin structures. Therefore, it is suggested that *C. subvermispora* uses the peroxidation reactions of fatty acids initiated by the enzyme MnP to initiate the degradation of lignin in the cell walls of wood [25].

On the other hand, it has been found that *P. chrysosporium*, when producing LiP and MnP, could degrade anthracene and phenanthrene, which are polycyclic

aromatic hydrocarbons (PAHs) evidencing the role of these enzymes to degrade these types of compounds. Anthracene is used in the production of artificial colorants, insecticides, or coating materials and is listed as one of the priority pollutants of the United States Environmental Protection Agency (USEPA) [26]. Furthermore, previous studies, has been found that white rot fungi such as *Phlebia lindtneri* and *Phlebia brevispora* could hydroxylate polychlorinated dibenzo-p-dioxins, 2,7-dichlorodibenzo-p-dioxin, 2,3,7-trichlorodibenzo-p-dioxin, 1,2,8,9-tetrachlorodibenzopyrodioxine, and 1,2,6,7-tetrachlorodibenzo-p-dioxin [13]. It has been shown that hydroxylation of an aromatic ring is important as the first step for degradation of dioxins and hydrocarbons [13]. Recent studies have reported that *Ganoderma lucidum* could break down trichlorethylene and drugs such as ibuprofen and carbamazepine, since this fungus has the potential to produce all three types of ligninolytic enzymes, although enzymatic activities vary between different strains and culture conditions [27]. All these reports, shows the potential of different fungus strains to be used to eliminate, lignin and POPs.

6. Ligninolytic enzymes and their characteristics

So far, the white rot fungi are the main eukaryotic microorganisms reported that produce these enzymes and play a crucial role in the degradation of plant raw materials, as well as numerous phenolic contaminants in soil bioremediation and in industrial waters [7]. The four major lignin modifying enzymes are lignin peroxidase (LiP), manganese peroxidase (MnP), laccase (Lac), and the versatile peroxidase (VP). White rot fungi contain the four enzymes and therefore can decompose and mineralize several environmental contaminants in non-toxic forms [18].

It has been observed that white rot basidiomycetes such as *Coriolus versicolor*, *P. chrysosporium*, and *T. versicolor*, degrade lignin more efficiently. However, electron microscopy studies of the early stages of fungal degradation have shown that oxidative ligninolytic enzymes are too large to penetrate the micropores of the cell wall of wood. Therefore, it has been suggested that before enzymatic attack, low molecular weight reactive oxidant compounds should initiate changes in lignin structure [20].

Manganese peroxidase (MnP) was first isolated from *P. chrysosporium* [28]. This enzyme is produced by almost all basidiomycete fungi that cause white wood rot and contain a heme group as prosthetic group, which is always extracellular [29]. The production of MnP in *P. chrysosporium* is enhanced by the limitation of nutrients such as C, N, and also by Mn^{2+} . It is also shown that the fungus *Irpex lacteus* has a great potential in the pretreatment of lignocellulose as well as in the biodegradation of xenobiotics, producing MnP as the main enzyme in many controlled conditions tested [30]. This peroxidase, requires Mn^{2+} as its reducing substrate, oxidizing it to Mn^{3+} , which is then chelated by several organic acids as oxalate, which are secreted by fungi [18]. Mn^{3+} chelating complexes in turn act as low molecular weight diffusible oxidants of the relatively labile phenolic substructures in the lignin. However, Mn^{3+} chelating complexes act as low molecular weight oxidants of the phenolic substructures of lignin. Although in natural lignin contains other non-phenolic substructures and these are abundant and recalcitrant to degradation by this mechanism. The efficiency of MnP is dependent on the peroxidation of unsaturated fatty acids, generating in the process, species that are capable of oxidizing and excising non-phenolic structures in synthetic lignins. It has also been shown that unsaturated fatty acids increase the ability of MnP to bleach (i.e., delignify) wood pulps [29].

The discovery of the enzyme lignin peroxidase (LiP), was confirmed in 1983 from the fungus *Phanerochaete chrysosporium*. LiP oxidizes aromatic lignin rings and a wide range of dangerous contaminants such as phenols, dyes, xenobiotics, and this enzyme has a redox potential ($E_0 \sim 1.2$ V at pH 3.0–4.5) [28, 31]. LiP is a biotechnologically important enzyme that has broad applications such as the delignification of lignocellulosic materials, conversion of coal into low molecular weight fractions, which could be used for the production of basic chemicals, in bio-pulp and bio-bleaching in industries of paper, and on the elimination of recalcitrant organic pollutants and for enzymatic polymerization in the polymer industries [18]. *P. chrysosporium* can produce LiP since it depends only on the limitation of carbon, nitrogen, or sulfur. It is important to mention that this enzyme uses a monomeric compound, veratryl alcohol as an inducer for the catalysis of the reaction [32].

In case of Laccase (Lac), it was first discovered by Yoshida in the Japanese lacquer tree *Rhus vernicifera* in 1883 [28]. This enzyme is capable of oxidizing phenols and aromatic amines. These are typical of basidiomycetes of white rot [33]. Lac is a multi-copper oxidase in its structure, the availability of copper in the medium could allow the synthesis of the enzyme. In addition, the presence of copper in *Pleurotus ostreatus* cultures decreases the activity of extracellular proteases that can degrade laccase [34]. A study with *P. ostreatus* in solid culture, using sugar cane bagasse as a substrate, showed that laccases are induced by copper sulfate and ferulic acid and two concentration levels of organic nitrogen in the form of yeast extract and its regulation of their expression may be substantially diverse among fungal species [22]. Lac is capable of catalyzing the direct oxidation of ortho and para-diphenols, aminophenols, polyphenols, polyamines, and aryl diamines, as well as some inorganic ions. It couples the four single-electron oxidations of the reducing substrate to the four-electron reducing division of the dioxygen, using four Cu atoms distributed against three sites, defined according to their spectroscopic properties [23].

The other important ligninolytic enzyme is versatile peroxidase (VP), which was first discovered approximately 20 years ago in the fungus *Pleurotus eryngii*, and a few years later in the *Bjerkandera* species [35]. VP has recently been described as a new ligninolytic peroxidase family, together with LiP and MnP, both reported by *Phanerochaete chrysosporium* [18]. In general, VP enzymes, share the same characteristics to most peroxidases, but the VP enzyme is unique with respect to the substrate that can oxidize. VP oxidizes high potential redox dyes [18]. Therefore, VP is an oxidoreductase that degrades lignin and in natural settings this enzyme cooperates with LiP, MnP, and Lac to break down lignocellulose. VP is considered as a hybrid enzyme because it combines catalytic features of MnP, oxidizing Mn^{2+} to Mn^{3+} , and can oxidize aromatic compounds when complexed with organic acids [36]. This enzyme is distinguished from LiP and MnP in that it has multiple active sites to directly oxidize Mn^{2+} and some phenolic and non-phenolic substrates (such as Orange II, Reactive Black 5, and Direct Blue 199) in the absence of Mn^{2+} . In addition, VP contains a conserved manganese binding site from MnP and a conserved tryptophan residue from LiP [37].

7. Effects of some factors for the growth of ligninolytic fungi

Although there are microorganisms capable of secreting this type of extracellular enzymes, the efficiency of complex degradation processes depends on the potential of the degrading organism, its oxidative mechanisms and culture conditions. The

problem that persists today is the slow cultivation of fungus species in wood and the sensitivity to growing conditions [17]. These extracellular oxidoreductase enzymes are generally secreted in low amounts. The degradation of fungal lignin is considered as a secondary metabolic event that is triggered when the indispensable nutrients like nitrogen are depleted [38].

Has been seek the way to increase the enzyme production by the different fungus strains and some reports that such increase can be achieved by modifying the source of carbon and nitrogen in the medium. Since the cost of enzymes is very high, this is a limitation in industrial use, however, the use of agricultural waste not only lowers cost, but also solves an environmental problem [34].

Therefore, the production of lignin-modifying peroxidase has traditionally been studied in media with low nitrogen concentration. Phenol rich but also nitrogen rich crops often increase laccase activity in white rot fungi and a complex nitrogen source such as peptone is generally required for efficient laccase production [38]. Copper is a known inducer of fungal lacquer activity. Its promoter effect on Lac production can be explained at the folding and post-translational level of proteins since the fungal laccases contain four copper atoms at two active redox sites. Transcriptional activation of the Lac genes by copper with several white rot fungi has also been demonstrated [38].

In general, and as happen with other enzymes, the enzymatic activities and biodegradation rates are influenced by pH, temperature, substrate concentration, presence of mediators, for example, 2,2'-azino-bis (3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) and veratryl alcohol, cofactors such as Cu^{2+} and Mg^{2+} , inhibitors such as organic acids, for example, citric, oxalic, and tartaric acid [28]. Molecular cloning studies of ligninolytic enzyme genes, such as the *Ganoderma lucidum* laccase gene, expressed in *Pichia pastoris*, have been carried out, taking into account the effects of temperature, pH, and nitrogen source on the expression of the enzyme in *P. pastoris*, obtaining that at a pH of 7.0 and a temperature of 20–30°C the Lac was stable [33].

It is now considered that extracellular lignin-modifying enzymes (Lac, LiP, MnP, and VP) together with new peroxidases containing a hemo group secreted by white rot fungi may be crucial for the biodegradation of lignocelluloses and lignocelluloses peroxidases secreted by hemo-containing fungi are oxidizing and biocatalyst proteins for the conversion of lignin-like compounds and aromatic compounds, and show specificity against a wide range of organic and inorganic compounds. Lignin modifying peroxidases (LiP, MnP, and VP) are structurally related to hemo-glycoproteins belonging to the family of fungal hemo-peroxidases [38].

8. Conclusion

This review shows the importance of degrading lignin, specifically in the bleaching process in the paper industry, since highly harmful pollutants are generated, thus affecting the environment and human health. Furthermore, other persistent organic pollutants (POPs), which are related to pesticides, drugs, dyes, etc., produced and wasted in other industrial process, are also on big concern. The toxicity of all of these POPs particularly organochlorines tends to increase according to the degree of chlorination turning them more resilient and produce biomagnification, because accumulate in adipose tissue on different animals including human beings from where is distributed through blood at different organs and tissues giving concentrations higher than those of the environment with the potential to cause diseases such as cancer and reproductive defects. However, exist different strategies to remove and reduce them in the environment but need to be

improved. Within these strategies, exist the biological methods particularly filamentous fungi, bacteria, and yeasts, with potential to degrade lignin and several POPs thanks to the synthesis and secretion of extracellular enzymes as LiP, MnP, Lac, and VP than depolymerize and transform POPs in nontoxic compounds. These methods are friendly to the environment, they are less expensive, compared to the physical-chemical ones.

Acknowledgements

This work was supported partially by the “Foundation Universidad de las Americas Puebla” for the student’s scholarship and a complementary scholarship by CONACYT.

Conflict of interest

All authors declared that they have no conflicts of interest to this work.

Author details

Anaid Bautista-Guerrero and Jose Luis Sanchez-Salas*
Universidad de las Americas Puebla, Ex Hacienda de Sta. Catarina Martir. Cholula,
Puebla, Mexico

*Address all correspondence to: jluis.sanchez@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Kamimura N, Sakamoto S, Mitsuda N, Masai E, Kajita S. Advances in microbial lignin degradation and its applications. *Current Opinion in Biotechnology*. 2019;**56**:179-186. DOI: 10.1016/j.copbio.2018.11.011
- [2] Longe L, Couvreur J, Grandchamp ML, Garnier G, Allais F, Saito K. Importance of mediators for lignin degradation by fungal laccase. *ACS Sustainable Chemistry & Engineering*. 2018;**6**:10097-10107. DOI: 10.1021/acssuschemeng.8b01426
- [3] Zeng Y, Zhao S, Yang S, Ding S. Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Current Opinion in Biotechnology*. 2014;**27**:38-45. DOI: 10.1016/j.copbio.2013.09.008
- [4] Li C, Chen C, Wu X, Tsang CW, Mou J, Yan J, et al. Recent advancement in lignin biorefinery: With special focus on enzymatic degradation and valorization. *Bioresource Technology*. 2019;**291**:121898. DOI: 10.1016/j.biortech.2019.121898
- [5] Ahmad M, Taylor CR, Pink D, Burton K, Eastwood D, Bending GD, et al. Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders. *Molecular BioSystems*. 2010;**6**:815-821. DOI: 10.1039/b908966g
- [6] Kleber M. What is recalcitrant soil organic matter? *Environment and Chemistry*. 2010;**7**:320-332. DOI: 10.1071/EN10006
- [7] Knezevic A, Milovanovic I, Stajic M, Loncar N, Brceski I, Vukojevik J, et al. Lignin degradation by selected fungal species. *Bioresource Technology*. 2013;**138**:117-123. DOI: 10.1016/j.biortech.2013.03.182
- [8] Schecter A, Colacino J, Haffner D, Patel K, Opel M, Papke O, et al. Perfluorinated compounds, polychlorinated biphenyls, and organochlorine pesticide contamination in composite food samples from Dallas, Texas, USA. *Environmental Health Perspectives*. 2010;**118**:796-802. DOI: 10.1289/ehp.0901347
- [9] Lee DH, Porta M, Jacobs DR Jr, Vandenberg LN. Chlorinated persistent organic pollutants, obesity, and type 2 diabetes. *Endocrine Reviews*. 2014;**35**:557-601. DOI: 10.1210/er.2013-1084
- [10] Herzig R, Lohmann N, Meier R. Temporal change of the accumulation of persistent organic pollutants (POPs) and polycyclic aromatic hydrocarbons (PAHs) in lichens in Switzerland between 1995 and 2014. *Environmental Science and Pollution Research*. 2019;**26**:10562-10575. DOI: 10.1007/s11356-019-04236-9
- [11] Mrema EJ, Rubino FM, Brambilla G, Moretto A, Tsatsakis AM, Colosio C. Persistent organochlorinated pesticides and mechanisms of their toxicity. *Toxicology*. 2012;**307**:74-88. DOI: 10.1016/j.tox.2012.11.015
- [12] Karn SK, Chakrabarty SK, Reddy MS. Characterization of pentachlorophenol degrading *Bacillus* strains from secondary pulp-and-paper-industry sludge. *International Biodeterioration & Biodegradation*. 2010;**64**:609-613. DOI: 10.1016/j.ibiod.2010.05.017
- [13] Xiao P, Mori T, Kamei I, Kondo R. A novel metabolic pathway for

biodegradation of DDT by the white rot fungi, *Phlebia lindtneri* and *Phlebia brevispora*. Biodegradation. 2011;22:859-867. DOI: 10.1007/s10532-010-9443-z

[14] Leri AC, Mynen SCB. Organochlorine turnover in forest ecosystems: The missing link in the terrestrial chlorine cycle. Global Biogeochemical Cycles. 2010;24:1-8. DOI: 10.1029/2010GB003882

[15] Parween M, Ramanathan AL, Khillare PS. Persistence, variance and toxic levels of organochlorine pesticides in fluvial sediments and the role of black carbon in their retention. Environmental Science and Pollution Research. 2014;21:6525-6546. DOI: 10.1007/s11356-014-2531-6

[16] Martyniuk CJ, Mehinto AC, Denslow ND. Organochlorine pesticides: Agrochemicals with potent endocrine-disrupting properties in fish. Molecular and Cellular Endocrinology. 2020;507:110764. DOI: 10.1016/j.mce.2020.110764

[17] Chang AJ, Fan J, Wen X. Screening of fungi capable of highly selective degradation of lignin in rice straw. International Biodeterioration & Biodegradation. 2012;72:26-30. DOI: 10.1016/j.ibiod.2012.04.013

[18] Yadav M, Yadav HS. Applications of ligninolytic enzymes to pollutants, wastewater, dyes, soil, coal, paper and polymers. Environmental Chemistry Letters. 2015;13:309-318. DOI: 10.1007/s10311-015-0516-4

[19] Hirosue S, Tazaki M, Hiratsuka N, Yanai S, Kabumoto H, Shinkyo R, et al. Insight into functional diversity of cytochrome P450 in the white-rot basidiomycete *Phanerochaete chrysosporium*: Involvement of versatile monooxygenase. Biochemical and

Biophysical Research Communications. 2011;407:118-123. DOI: 10.1016/j.bbrc.2011.02.121

[20] Dashtban M, Schraft H, Syed TA, Qin W. Fungal biodegradation and enzymatic modification of lignin. International Journal of Biochemistry and Molecular Biology. 2010;1:36-50

[21] Nayan N, Sonnenberg ASM, Hendriks WH, Cone JW. Screening of white-rot fungi for bioprocessing of wheat straw into ruminant feed. Applied Microbiology. 2018;125:468-479. DOI: 10.1111/jam.13894

[22] Karp SG, Faraco V, Amore A, Birolo L, Giangrande C, Soccol VT, et al. Characterization of laccase isoforms produced by *Pleurotus ostreatus* in solid state fermentation of sugarcane bagasse. Bioresource Technology. 2012;114:735-739. DOI: 10.1016/j.biortech.2012.03.058

[23] Piscitelli A, Del Vecchio C, Faraco V, Giardina P, Macellaro G, Miele A, et al. Fungal laccases: Versatile tools for lignocellulose transformation. Comptes Rendus Biologies. 2011;334:789-794. DOI: 10.1016/j.crvi.2011.06.007

[24] Su Y, Yu X, Sun Y, Wang G, Chen H, Chen G. Evaluation of screened lignin-degrading fungi for the biological pretreatment of corn stover. Scientific Reports. 2018;8:1-11. DOI: 10.1038/s41598-018-23626-6

[25] Cunha GGS, Masarin F, Norambuena M, Freer J, Ferraz A. Linoleic acid peroxidation and lignin degradation by enzymes produced by *Ceriporiopsis subvermispota* grown on wood or in submerged liquid cultures. Enzyme and Microbial Technology. 2010;46:262-267. DOI: 10.1016/j.enzmictec.2009.11.006

[26] Wu YR, Luo ZH, Vrijmoed LLP. Biodegradation of anthracene and

benz[a]anthracene by two *Fusarium solani* strains isolated from mangrove sediments. *Bioresource Technology*. 2010;**101**:9666-9672. DOI: 10.1016/j.biortech.2010.07.049

[27] Ting WTE, Yuan SY, Wu SD, Chang BV. Biodegradation of phenanthrene and pyrene by *Ganoderma lucidum*. *International Biodeterioration & Biodegradation*. 2011;**65**:238-242. DOI: 10.1016/j.ibiod.2010.11.007

[28] Shi K, Liu Y, Chen P, Li Y. Contribution of lignin peroxidase, manganese peroxidase, and Laccase in lignite degradation by mixed white-rot fungi. *Waste and Biomass Valorization*. 2020;**12**:3753-3763. DOI: 10.1007/s12649-020-01275-z

[29] Kapich AN, Korneichik TV, Hatakka A, Hammel KE. Oxidizability of unsaturated fatty acids and of a non-phenolic lignin structure in the manganese peroxidase-dependent lipid peroxidation system. *Enzyme and Microbial Technology*. 2010;**46**:136-140. DOI: 10.1016/j.enzmictec.2009.09.014

[30] Chowdhary P, Shukla G, Raj G, Ferreira LFR, Bharagava RN. Microbial manganese peroxidase: A ligninolytic enzyme and its ample opportunities in research. *SN Applied Sciences*. 2019;**1**:1-12. DOI: 10.1007/s42452-018-0046-3

[31] Singh AK, Bilal M, Iqbal MNH, Raj A. Lignin peroxidase in focus for catalytic elimination of contaminants—A critical review on recent progress and perspectives. *International Journal of Biological Macromolecules*. 2021;**177**:58-82. DOI: 10.1016/j.ijbiomac.2021.02.032

[32] Singh D, Chen S. The white-rot fungus *Phanerochaete chrysosporium*: Conditions for the production of lignin-degrading enzymes. *Applied Microbiology and Biotechnology*.

2008;**81**:399-417. DOI: 10.1007/s00253-008-1706-9

[33] You LF, Liu ZM, Lin JF, Guo LQ, Huang XL. Molecular cloning of a laccase gene from *Ganoderma lucidum* and heterologous expression in *Pichia pastoris*. *Journal of Basic Microbiology*. 2014;**54**:S134-S141. DOI: 10.1002/jobm.201200808

[34] El-Batal AI, ElKenawy NM, Yassin AS, Amin MA. Laccase production by *Pleurotus ostreatus* and its application in synthesis of gold nanoparticles. *Biotechnology Reports*. 2015;**5**:31-39. DOI: 10.1016/j.btre.2014.11.001

[35] Gonzalez-Perez D, Alcalde M. The making of versatile peroxidase by directed evolution. *Biocatalysis and Biotransformation*. 2018;**36**:1-11. DOI: 10.1080/10242422.2017.1363190

[36] Đurđić KI, Ostafe R, Delmas AD, Popovic N, Schillberg S, Fischer R, et al. Saturation mutagenesis to improve the degradation of azo dyes by versatile peroxidase and application in form of VP-coated yeast cell walls. *Enzyme and Microbial Technology*. 2020;**136**:109509. DOI: 10.1016/j.enzmictec.2020.109509

[37] Liu J, Zhang S, Shi Q, Wang L, Kong W, Yu H, et al. Highly efficient oxidation of synthetic and natural lignin-related compounds by *Physisporinus vitreus* versatile peroxidase. *International Biodeterioration and Biodegradation*. 2019;**136**:41-48. DOI: 10.1016/j.ibiod.2018.10.009

[38] Makela MR, Lundell T, Hatakka A, Hildén K. Effect of copper, nutrient nitrogen, and wood-supplement on the production of lignin-modifying enzymes by the white-rot fungus *Phlebia radiata*. *British Mycological Society*. 2013;**117**:62-70. DOI: 10.1016/j.funbio.2012.11.006

The Usefulness of System Dynamics for Groundwater Management

David Eduardo Guevara-Polo and Carlos Patiño-Gómez

Abstract

The water crisis is global, and it is manifested in groundwater stored in aquifers; also, the climate crisis is worsening it. In this context, there is an ongoing paradigm shift in water management where complexity of water problems is being addressed. For that matter, there are several interdisciplinary approximations and modeling techniques available in the literature. Integrated Water Resources Management for the achievement of water security can be modeled using system dynamics. Dynamic modeling is multidisciplinary, cross-scalar, flexible, modular, adaptable, and allows stakeholders to be involved, therefore, it is appealing to IWRM and groundwater management. In the latter case, aquifers are usually modeled in a broader scale of a watershed and their interrelationships with human activities and climate change have been modeled successfully. In the future, there are opportunities for these modeling approach, which are joining different water subsystems, including water systems in broader human systems, and combining system dynamics with other techniques that can increase its suitability for decision making. System dynamics modeling is useful to inform public policies and projects that can address the groundwater crisis, especially in countries where the impacts are already evident, like Mexico.

Keywords: system dynamics, water management, groundwater, water security, water crisis, aquifer

1. Introduction

During the period 2014–2021, “Groundwater in a changing environment” was the second theme of the eighth phase of the Strategic Plan of the International Hydrologic Programme of the United Nations Science Culture and Education Organization (UNESCO). In 2022, “Groundwater: Making the invisible visible” is the theme of World Water Day and consequently, of the World Water Development Report, both sponsored by the United Nations. Moreover, groundwater is the source of resilience during the occurrence of droughts and therefore, is an essential component of water supply and climate action, therefore, it contributes directly to Sustainable Development Goals 6 and 13. Groundwater is a relevant component of the international water agenda. The dependence of human and ecological communities on

groundwater varies substantially across the globe, but in no location is groundwater not utilized [1].

Groundwater is stored in aquifers. These water bodies are distinguished from other elements of the hydrological cycle, because of their features: flux, storage, and residence time [2]. However, it must not be forgotten that the groundwater subsystem is encompassed by the water system and therefore, its management cannot be made in isolation. The current multiplicity of water uses, and the soaring of stakeholders involved in water issues has forced a paradigm shift for water management, for which system dynamics turns out to be an interesting approach.

System dynamics represent not only modeling but also a radical change of thought [3]. In that sense, the current chapter argues the usefulness of system dynamics for its application in groundwater management. As shall be seen, the simulation of aquifers using this approach allows modelers to include them in broader water systems simulation and describe its behavior on time. Groundwater, surface water, humans, and ecosystems are all interconnected in ways that necessitate an integrated approach to management [4]. The chapter is structured in four sections. The first one elaborates on how the global water crisis belongs to the global risks of the twenty-first century, how it is interrelated with other problems, and how current technology has shifted the paradigm of water management. This section end with a comment on the global relevance of groundwater. On a lower scale, the second section discusses the current situation of groundwater resources in Mexico and describes its consequences. Next, there is a description of the historical development of system dynamics theory and its application to water management, as well as its features. Finally, we undertake a systematic review of the literature available on its applications to groundwater management and mention some opportunities for future research on this topic.

2. Context of the global groundwater crisis

Problems of the twenty-first century must be seen as different facets of one single crisis, which is largely a crisis of perception [5]. In Ref. [6], the World Economic Forum reported the identified risks for 2020. They are classified into five categories: economic, environmental, geopolitical, technological, and social but, above all, it is fundamental to underscore the interrelation that exists between them. Certainly, given the systemic nature of Earth, there are expectations that environmental risks are interrelated. However, the relationship between these and those of the other categories is not obvious and, as suggested in Capra and Luisi [5], the incapacity to identify and understand these relationships has not allowed us to manage these risks. In this framework, the water crisis emerges.

In this context, water management can no longer be postponed. The modern approach of water management is characterized by two fundamental concepts: water security and integrated water resources management (IWRM), where the former is the strategic objective, and the latter is the method for reaching it [7]. It is important to underscore that IWRM requires a multidisciplinary perspective to be conducted; it must consider biophysical aspects provided by science and engineering but in addition, got to be enriched with the perspective of social and economic sciences given the wide range of stakeholders involved in water issues. This last premise, along with the increasing computation power and the increase in the complexity of water problems, constitutes the paradigm of complexity of water management [8].

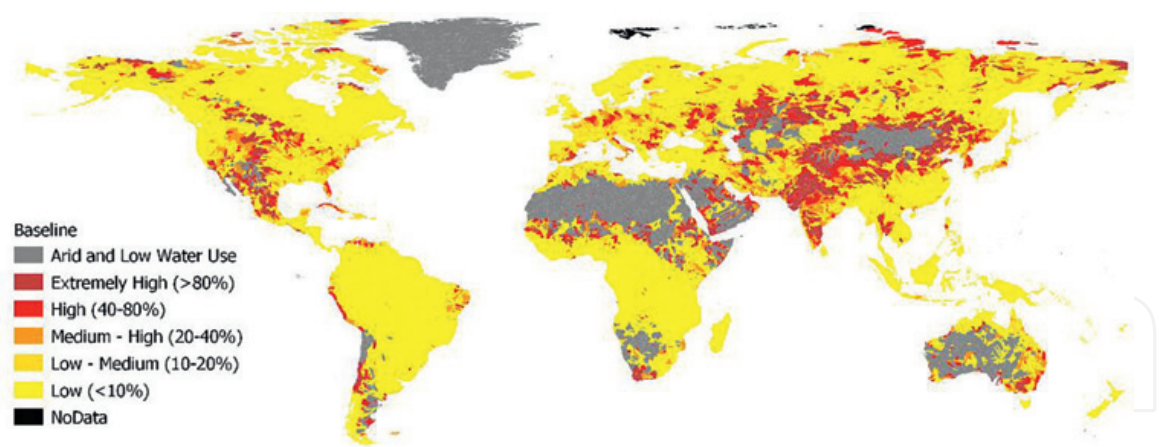


Figure 1.
Water stress in the world. Adapted from Hofste et al. [9].

The water crisis has a global dimension. In **Figure 1**, the water stress in the base scenario, which considers data for 1960–2014. It is evident that in the subtropical region, especially in the northern hemisphere, the water stress is predominantly high and very high. The cases of Mexico and the United States in North America in the west and India and Saudi Arabia in the east are worth highlighting. Analogously, in the subtropical region of the southern hemisphere, the outlook is not favorable in northern Chile and some regions of Argentina, in South America and, in the region of the Sahel, South Africa, and Australia, as well. Likewise, in **Figure 2**, the projection of change in water stress concerning base scenario towards 2040 under a business-as-usual scenario (combination of socioeconomic scenario SSP2 and climate change scenario RCP8.5) is shown. In previously mentioned regions, there will be an increase of water stress in a higher proportion.

In this global context, groundwater performs an essential role. In principle, it is important in sustaining streams, lakes, wetlands, and aquatic communities [10], therefore it is relevant for sustaining biodiversity. On the other hand, groundwater represents almost half of freshwater supply worldwide, around 40% of irrigation water, and a third of water required for industry [11]. Nevertheless, despite its importance, groundwater remains a minor player in water resources management [1] and

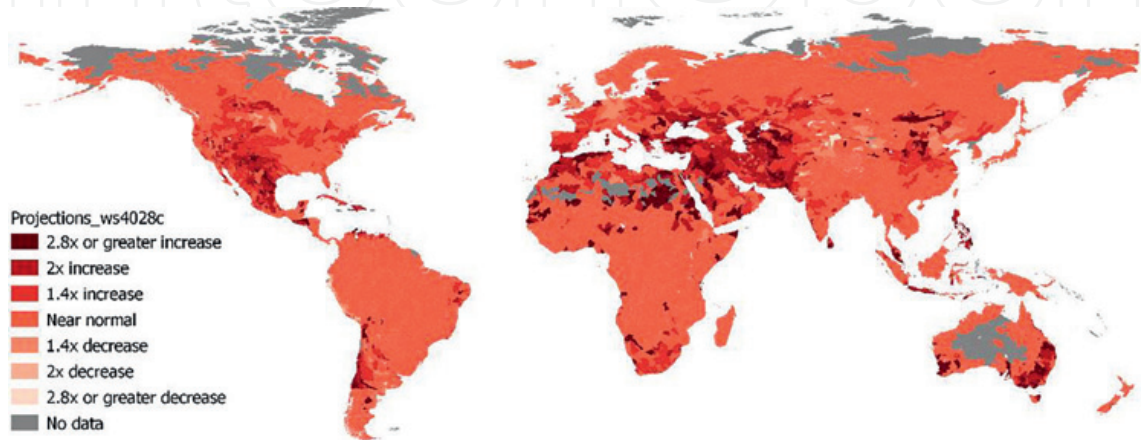


Figure 2.
Water stress projection regarding baseline scenario. Adapted from Hofste et al. [9].

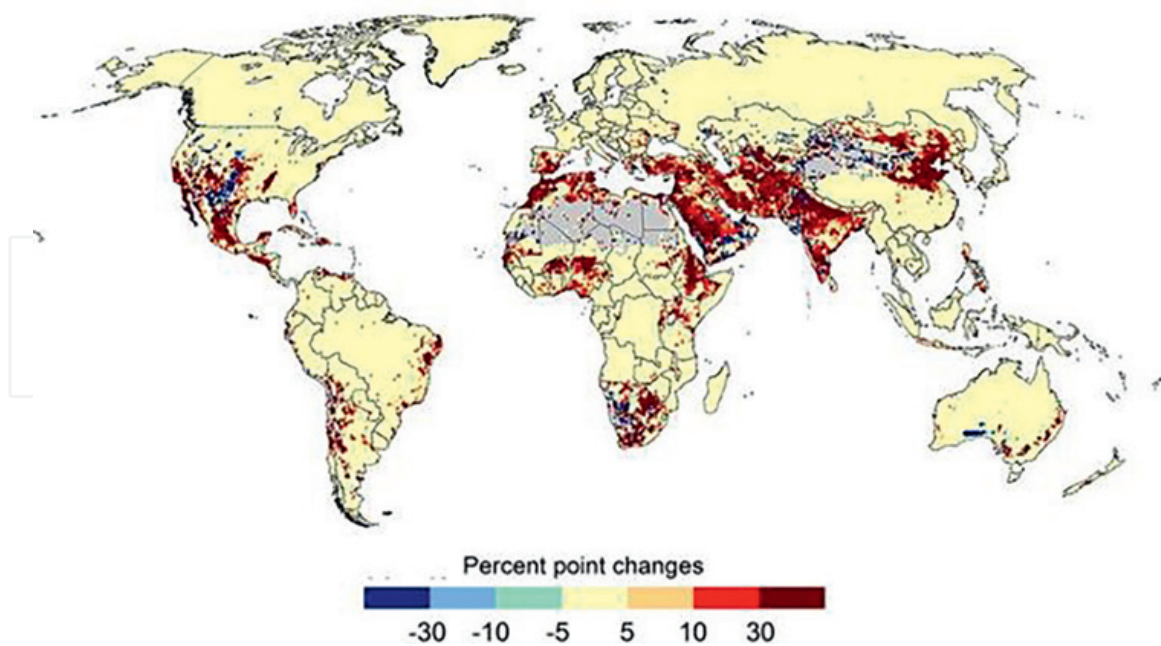


Figure 3.
Groundwater stress projection for 2041–2070. Adapted from Herbert and Döll [12].

the evidence is that, although the impact of groundwater extraction is most acute and obvious at local scales, right now, groundwater depletion is a global problem [2].

In that respect, in Herbert and Döll [12] the scenario shown in **Figure 3** is suggested. In the latter, the projection represents the mean of 10 scenarios, which considers five General Circulation Models (GCMs), forced by a climate change scenario RCP8.5 and socioeconomic SSP2, and two irrigation scenarios. It is noted that subtropical regions of the northern hemisphere are currently the most affected and, and if that was not enough, affectations may increase in the future. The case of Mexico, the United States, the Middle East, South Africa, and the Sahel region are worth highlighting [12].

3. Context of the groundwater crisis in Mexico

In this global context, Mexico will face several challenges given the simultaneous pressures of population growth, urbanization, climate variability, and climate change [13]. Mexico's population will reach near 148 million people in 2050, which means an increase of 23% concerning 2015 [14] while the urban population will increase by 35.7 million in 2050 [15]. These increases will escalate the demand for food, energy, products, and services, with the consequent augmentation of water demand. Additionally, population growth has an unfavorable distribution concerning water availability because 77% of the population is established where 33% of available water is found [16].

Moreover, Mexico's climate has been classified as semi-arid to arid, and as a result, and particularly in the northern half of the country, we are continuously experiencing droughts of different magnitudes [17] and with climate change, an increase in their severity and frequency can be expected [18]. Besides, because of the latter, a decrease in cumulative precipitation is expected [19] and positive anomalies of temperature [20] in Mexican territory. Temperature changes will boost water demand, mainly for the environment and food production, while changes in precipitation and runoff

point to a decrease in water availability [18]. Actually, in the RCP 6.0 scenario, a reduction between 16 and 25% in precipitation is expected in the northern region of the country, while in the northeast a reduction up to 40% is anticipated [21]. Climate change impacts at the basin scale have been documented previously. For example, in Herrera-Pantoja and Hiscock [22] it was suggested that surface runoff in the Rio Queretaro basin can decrease 9.16%, while in Molina-Navarro et al. [23] it was reported that reductions of 45% in the short term can be anticipated in runoff of the river Guadalupe basin, in Baja California.

In this context, groundwater must be managed strategically because it looks more resilient to rainfall variability caused by climate change [21], and besides, it may mitigate drought effects [24]. In fact, in the National Program Against Drought (Pronacose, in Spanish) of the Mexican water authority Conagua, the exploitation degree of aquifers was selected as an indicator of adaptation capacity to compute climate vulnerability [25] and for example, in the watershed-scale program against drought of the Rio Balsas basin council, three of the five strategic measures related to water supply involve groundwater [26].

It is worth mentioning that, in general, groundwater has advantageous properties concerning other water sources: lower evaporation losses, protection of contamination caused by human activities, flexible management of the resource, and wide spatial distribution [27]. However, despite these features and their relevance in the current climate context, Mexican aquifers are severely affected. Aquifers' overexploitation and overallocation characterize current water management [28] and if this trend prevails, the resilience of human systems to water shortages will be compromised [29]. Nationally, groundwater supplies 35.5, 58.6, and 56.3% of water demand for irrigation, public urban, and industrial use, respectively [30]. By consequence, in Mexico, groundwater has a critical role in water security and its depletion can compromise food security, environmental security, health, and economic development.

There are several definitions of overexploitation [27]. Conagua defines it as the condition where the ratio of extraction to mean annual recharge is greater than 1.10. In Mexico, there are 105 overexploited aquifers [25], which represent 16.07% of all aquifers and are illustrated in **Figure 4**. It might sound like a small fraction, however, these overexploited aquifers are supply sources for the most important urban and industrial centers in Mexico, and around 42 million inhabitants depend partially on them [31], as shown in **Figure 5**. By consequence, despite that the national balance of groundwater is positive (i.e., 92.5 km^3 of recharge and 29.9 km^3 of extraction in a year), some local balances are not [25].

Additionally, overallocation, which refers to the assignation of groundwater volume assignation beyond the resource availability, is a product of the miscalculation of the mean annual availability of aquifers [28]. This indicator is computed by subtracting the compromised natural discharge and the recharge from the allocations; the former is defined as compromised water for surface allocations or environmental purposes [32]. In 2020, it was reported that nationally 92.4 km^3 of recharge occurred, and 42.9 km^3 were compromised as natural discharge; hence, there were 49.5 km^3 of availability, and 41.5 km^3 were allocated. That means that 83.1% of mean annual availability was allocated [33]. However, it is worth highlighting that the mean annual availability of 275 aquifers is negative, which reflects the phenomena of overallocation. Is it interesting to underscore that 7 of the 10 aquifers with less availability are found in the state of Chihuahua: Los Juncos, Laguna de Santa María, Laguna de Tarabillas, Laguna de Hormigas, Lagunas la Vieja, Jiménez-Camargo, and Meoqui-Delicias; their availability is found between -697.79 and -165.04 km^3 .



Figure 4.
Overexploited aquifers in Mexico. Reproduced from Conagua [25].

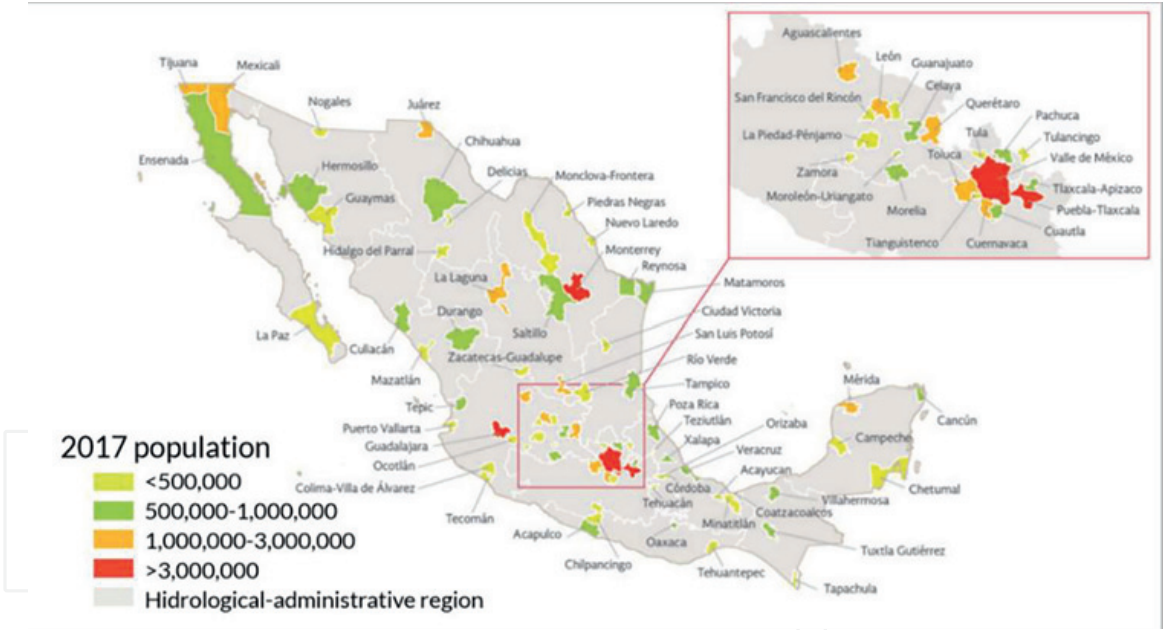


Figure 5.
Metropolitan zones in Mexico. Reproduced from Conagua [25].

It is worth mentioning that the compromise of the drinking water supply is not the only consequence of aquifers' overexploitation. This has also caused spring discharge depletion, disappearance of lakes and wetlands, decrease of base flow in rivers, elimination of native vegetation, and ecosystems loss [27]. This represents serious affectations. Firstly, intensive extraction of groundwater can alter the quality through the upwell of thermal and/or mineralized water of underlying aquifers and the infiltration of organic contaminants from sewage and rainfall percolation [34]. Similar contamination is widely documented in Mexico. For example, in Camacho et al. [35] there was a report of arsenic occurrence in Chihuahua and Coahuila states in

concentrations higher than permissible. And if that was not enough, besides meaning a public health risk, overexploitation and pollution of aquifers also have an impact in groundwater dependent ecosystems. For example, in Esteller and Diaz-Delgado [36] it was reported that the ecosystem of the Almoloya del Río wetland had been affected by the overexploitation of the Valle de Toluca aquifer. Finally, the overexploitation of aquifers can explain subsidence in urban areas, which has important implications for infrastructure and buildings operation and maintenance. Possibly the best known case is that of Valle de Mexico, but there is also evidence for Aguascalientes [37] and the cities of Tepic, Nayarit [38]; Celaya, Guanajuato; Morelia, Michoacán; and Queretaro, Queretaro [39].

4. System dynamics and its application to water management

The water crisis is not an exception to the crisis of perception stated in Section 2. In this context, system dynamics represents an interesting method, as argued next. The approach of system dynamics was proposed by Forrester in 1956 for industrial applications. Forrester combined applied control theory, information theory, decision theory, and other relevant theories and methods relevant to this purpose [40] and argued that an event-oriented vision and a lineal-causal thought are not useful to address complex problems [41]. His books *Industrial Dynamics*, *Urban Dynamics*, and *Principles of Systems* were the starting point of system dynamics.

Some years later, Forrester and the Club of Rome, headed by Aurelio Peccei, collaborated to bring the system dynamics approach to a global scale and to a time horizon of more than 100 years. In effect, in Meadows et al. [42] there is an assessment of which would be the behavior of flow and storage of natural resources in the future and describe how they may mean limits to growth. The title of this report, *Limits to Growth*, became the motto of the global environmental movement [43]. Likewise, Forrester published individually his book *World Dynamics*. Later, Forrester and his collaborators developed a programming language called DYNAMO which made the *Limits to Growth* model possible. The textbook of Richardson and Paugh published in 1981 is based on this language.

Next, the ideas of system dynamics were brought to corporate and citizen practice in the book *The Fifth Discipline* [44]. The five disciplines described are personal mastery, mental models, shared vision, team learning, and systems thinking. This book created a new era in business and management because it inspiringly presented a framework for the learning organization [45]. During this decade, many companies, consultations firms, and governmental organizations began to use system dynamics to address critical issues [46]. In effect, Sterman published his book *Business Dynamics* in 2000, which collects concepts, techniques, and tools for dynamic simulation. Eventually, these developments permeated in IWRM, as described next.

Currently, the four interdisciplinary approximations to water management are IWRM, Water-Energy-Food Nexus (WEF Nexus), Nature-Based Solutions (NbS), and socio-hydrology [47]. In an extensive bibliometric and network analysis, the authors identified the main academic communities dedicated to studying the WEF nexus in the urban environment. There is no doubt that the current paradigm of water management is interdisciplinary. Given the complexity of water problems, the adequate approach to address them is a flexible one, that includes “hard” and “soft” approximations, that is, that it considers the construction of hydraulic infrastructures like dams, aqueducts, pipelines, and centralized treatment plants but at the same

time, looks to increase water use productivity using economic tools and public policy [48]. For that matter, assessment approaches are crucial.

Assessment approaches are classified into four categories: water urban systems, which encompasses integrated modeling of water systems; Urban Metabolism, which reunites territorial and mass balance; consumption methods like Life Cycle Analysis, water footprint, and input-output analysis; and complex systems (ecological network analysis and system dynamics) [49]. Similarly, in Hassan and Garg [48], there is a description of the different modeling techniques that can boost the flexible approach and conclude that system dynamics allow to model it while in Ghalehkhondabi et al. [50] there is a review on the soft computing techniques that allow to include non-linearity, stochasticity, and imprecision of water demand data and mentions that the potential of these techniques and especially that of system dynamics has not been availed.

The history of dynamic simulation for water management has been classified previously in Winz et al. [51]. For stages are distinguished: limitation in adaptation, during the 60s and 70s decade; the surge of integrated water resources management, at the end of the 80s decade; the 90s decade; and the recent developments, from 2000 onwards. Next, the advances in the application of dynamic modeling to water management are described through their historical development.

A few years after Forrester's developments, the first application of system dynamics to hydrology can be found by Crawford and Linsley in 1966. In effect, in Crawford and Burges [52] it is documented that the model was unique in its times because it reunited knowledge and theory of the different hydrologic processes but considered their interdependences and incorporate them in a model whose objective was to describe hydrologic balance continuously on time. In Winz et al. [51] it is mentioned that later Hamilton (1969) developed a model that besides flow and storage, included socioeconomic factors that looked to explain the systemic causes of the crisis in the American watershed of Susquehanna river. The perspective of these authors preceded the surge of integrated water resources management concept, so in their time it was visionary.

However, during the 70s decade, there was a limit in the adaptation of these models, attributable to two factors pointed out by Winz et al. [51]: the indifference for complexity and uncertainty in the daily operation of water systems and negative publicity of the findings of Forrester and Meadows. The adoption of the concept of IWRM and its inherent interdisciplinary feature, along with the soaring of computing power and the increase in the complexity of water problems constitute the paradigm of complexity in water management [8], under which proliferated a high quantity of studies using system dynamics approach. There have been several review articles published that will be discussed further (see [45, 46, 53–56]).

Dynamic simulation aids our capacity to make predictions of future states and to improve our understanding of the problem as a necessary step towards affecting sustainable and effective change [51]. The former activity is critical to inform public policy or projects related to water management but also to involve relevant stakeholders (including, of course, policy-makers and authorities that materialize these policies and projects) because it allows describing the behavior of the system under different scenarios.

Dynamic simulations models have been classified previously. The first classification of dynamic simulation models for water management is reported in Winz et al. [51] and it was realized based on the problem that they address: regional analysis and river basin planning, urban water, flooding, irrigation, and pure process models. A second

classification, proposed in Chen and Wei [40] divides the models into three categories: flooding control and disaster mitigation, water resources security, and water environment security. The first one encompasses management and forecast of floodings; the second reunites water resources management, water resources carrying capacity, water resources planning, and sustainability; the third one consists of pollution control, water quality management, pollution early warning systems, and water ecology.

However, the classification that has been complemented further in literature can be found in Mirchi et al. [41], which is based on the model's objective. This way, models are classified as predictive simulation models, descriptive integrated models, and participatory and shared vision models. Combined system dynamics models can be added to this classification [57]. In the first category, the processes governing particular subsystems within a broader water resources system are simulated with a quantitative approach to defining tactics, while in the second, the approach is holistic and strive to identify and characterize the main feedback loops among two or more disparate subsystems and finally, the third involves stakeholders and decision-makers to foster the understanding of the system [41]. In Zomorodian et al. [58], the three kinds of descriptive integrated models are reported: hydro-economic, socio-hydro-logic, and integrated water resources management models.

Previously, the modeling process in the classic literature of system dynamics was reported in Luna-Reyes and Andersen [59]. A comprehensive explanation of the methodology for dynamic simulation in the context of water management can be found in Simonovic [8]. In **Table 1**, there is a summary of the methodology for dynamic simulation in the context of water management and it is possible to appreciate that essentially, there are six processes to be conducted. Firstly, the problem articulation shapes the entire modeling [60] because involves identifying the key variables and their past and possible future behavior, as well as defining the time horizon [46].

Process	[51]	[40]	[60]	[57]
Process 1	Problem definition	Define objectives	Articulation	Objective determination
		Determine system boundary	Dynamic hypothesis	
Process 2	Qualitative/conceptual modeling	Design graphical structure		Polarity analysis
Process 3	Quantitative/numerical modeling	Develop stock-flow diagrams	Formulation	Model structure
		Formulate mathematical model		Simulation
Process 4	Model evaluation/testing	Calibrate and validate	Model testing	Theory building
			Calibration	
Process 5	Policy analysis		Policy formulation and evaluation	
Process 6	Implementation	Implement		

Table 1.
Processes for dynamic simulation in water management.

Next, there is conceptual modeling where qualitative tools can be used, such as causal relationships, where positive or negative relationships between the variables are identified; causal loop diagrams (CLD) and feedback loops; stock and flow diagrams (SFD); reference modes [4] and system archetypes [61]. To accomplish a successful system dynamics application, extensive computer simulations should be performed only after a clear picture of the integrated water resources system has been established through reasonably simplified conceptual models [61]. A result of this preliminary sketch is the development of developing a dynamic hypothesis, which plays a significant role in complexity reduction [60] and describes the problematic behavior of the system [46]. This process is appealing to involve stakeholders to identify the key variables that affect the system, as well as to propose and refine the dynamic hypothesis that describe the problematic behavior. It is important to note that relevant stakeholders are to be mapped for particular local conditions.

Following, the simulation model is formulated. The simulation model is the refinement and closure of a set of dynamic hypotheses to an explicit set of mathematical relationships, where the experimentation process is iterative and flexible [8]. It is important to underscore that the connection between qualitative and quantitative modeling is the use of the dynamic hypothesis and how it can be modified and adjusted, according to the findings made during the simulation process. The mathematical equations used in the simulation models describe detail complexity, and frequently the modeler will be tempted to include more math into the model to increase accuracy. However, it must be considered that, at the strategic level, emphasis should be placed on trend identification and pattern recognition rather than exact quantitative predictions of dynamic variables [61]. This claim leads to the following step of the modeling process.

It is worth recalling that a model is useful when addresses the right problem at the right scale and scope, and represents the system response correctly [51]. In that sense, for models to be useful as decision support tools for water resources planning and management, it is necessary to verify the model structure to ensure that mathematical equations and interrelationships between subsystems follow logical explanations and are not spurious or erroneous [61]. One method to do this is through calibration, which is concerned with adjusting the model variables in such a way that the model outputs can fit the collected data [60]; however, due to limited data or lack of appropriate methods to quantify particular subsystems [61] often makes this process not possible, which gives way to the concept of model validation [3]. This process is conducted using structural tests, sensitivity analysis [60], behavior tests (which encompasses calibration and comparison with reference modes), and implications tests [51].

Finally, depending on the models objective it can either be used for policy formulation and evaluation, where it is investigated how specific change in a parameter in a system dynamics model affects the system's response [60], or theory building, where the outcome is a theory that is a structured, explanatory, abstract and coherent set of statement about a partial reality [62]. In addition, in Randers [63] it is argued that besides some tests, implementation deals with the identification of potential users, translation of study insights to an accessible form, and diffusion of study insights. This can be done through the building of a decision support system (DSS) (see for example, Martínez-Valderrama et al. [53]), the organization of technology transfer workshops with decision-makers, and/or divulgation activities with the community.

Dynamic models are useful because they are multidisciplinary, cross-scalar, modular object-oriented, transparent, adaptable; in addition, they support a variety

of goals, allow users to quickly become familiar, and have methods for consensus building and team learning [51]. When the complexity of water management issues is considered, these are valuable features. In effect, water management problems are multiscalar because besides involving the biophysical borders of watersheds, they happen in political-administrative divisions. Also, they can be described using different subsystems according to water users, as described in the next section. Finally, they require the involvement of different stakeholders: several authorities, local communities, cities, and different kinds of users and each one of them has a distinct mental model of what are the root causes of problems, as well as their approach to solutions. These arguments support why system dynamics are useful for water management. Nevertheless, there are several opportunities for system dynamics application to water management, and they have been identified by the authors that have undertaken reviews of the literature on the topic (see [45, 46, 53–56]).

In Chen and Wei [40], a synthesis of the insights of 35 studies is reported. Firstly, they indicate the necessity of integrating studies of different water subsystems: flooding control, disaster mitigation, security in water supply, and environmental supply, for example. Secondly, to consider climate change and the intensification of human activities. Thirdly, to advance in uncertainty theories that limit applications of system dynamics. Finally, to integrate system dynamics with other methods like geographic information systems (GIS), remote sensing (RS), hydrologic models, and water quality to broaden their spatial scope.

In Zarghami et al. [60], a revision focused on applications of dynamic simulation to water supply and distribution networks was made. The authors identify 32 records oriented towards water supply and four to sewage and conclude that this approach is useful to be applied in water distribution networks because their performance indicators have a dynamic behavior and are subject to internal and external variables of the system. However, they suggest that they must be complemented to reinforce their capacities, with reliability methods, for example.

After the revision and synthesis reported in Zomorodian et al. [58], the authors coincide with the ideas in Chen and Wei [40] and Mashaly and Fernald [57]: it is necessary to integrate water subsystems to have a better perspective of the behavior of the entire system and to combine system dynamics with other methods and tools to strengthen the modeling framework. A relevant contribution of these authors is the identification of the necessity of adding qualitative variables to modeling, especially in the moment of incorporating social and political processes to modeling.

Recently, in Mashaly and Fernald [57] it was reported that since 1996, 199 articles have been published using system dynamics methods to solve agricultural water management problems. Besides, they found that 64% are integrated descriptive models, 16% are predictive simulation models, 8% are shared vision models, and 12% are combined. The authors advocate for preventing the creation of simplistic models through the involvement of specialists that have worked with the system and the use of validation and/or calibration methods.

Finally, in Cerecedo Arroyo and Martínez Austria [3] the first systematic review of dynamic modeling reported in the literature can be found. It is distinguished from previous work because of its transparency in search protocols. The authors reported 27 studies applied to watersheds and aquifers and indicate that there are relevant opportunities for the validation of these models as mentioned in Mashaly and Fernald [57], as well as the incorporation of water quality aspects.

5. System dynamics for groundwater management

Possibly the first record of application of system dynamics to groundwater is found in Hong-gang [54]. Although the author does not develop formally a simulation model, provides an interesting analysis of the main feedback loops that characterize groundwater, which counteract each other. The first one is characterized by the lag that exists between water extraction and the population perception of the depletion, which is reinforced with government subsidies. The second is distinguished by the closure of water wells because of water quality deterioration. Both trigger more wells drilling to obtain more water in the short term, depleting aquifers' storage. Now, one of the first dynamic simulation models applied to aquifers is proposed in Assaf [55], which poses two modules: the physical and the economic. The former includes extractions, natural and artificial recharge, and geologic features of the aquifer, and the latter consists of productivity and the net present value of water resources. Therefore, although theoretical developments are traced back to 2001, it was until 2009 that simulation models started to thrive. Under this consideration, a systematic review was conducted.

The search was made on December 14th, 2021. The question to search was: What are the recent advances in the application of system dynamics to groundwater? One hundred and ninety seven records in Scopus were identified using the search string TITLE-ABS-KEY (system PRE/0 dynamics) AND TITLE-ABS-KEY (groundwater OR aquifer) and excluding the areas of biochemistry, business, physics, chemistry, materials science, and medicine. After reading the titles of the records, those who were not related to the question were excluded; this was the first exclusion criteria. The next exclusion criteria were the relevance of the abstract to the research question, after which 40 records were not retrieved. Next, 68 records were sought for retrieval during the entire timeframe 2001–2021. Twenty-two of these were published between 2017 and 2021 but three were not related to the research question. Hence, 19 studies were included in this systematic review, which are reported in **Table 2**.

It is interesting to note that 11 of these 19 studies were conducted in Iran. In addition, there are three studies in China, two in the Mediterranean region, one in the United States, one in Brazil, and one in Morocco. These regions are either arid zones or regions subjected to high water stress, therefore they rely entirely or partially on groundwater resources for their freshwater supply. In that sense, it is worth underscoring that groundwater is frequently modeled encompassed by the watershed or water system, which is aligned with the principles of IWRM. It is important to note that the most comprehensive study found is reported in Jiang et al. [67]: it is structured in four subsystems (socioeconomic, natural resources, water, and eco-environmental) and includes policies and impacts.

The classification purposed for the reviewed articles is strategic, tactical, and descriptive. They are distinguished by their time horizon and the objective they pursue. The first two look to project scenarios into the future, but with a different time horizon and objective: the former refers to strategic actions that look to increase groundwater sustainability, while the latter emphasizes short-term actions operation or management of infrastructure or irrigation. From the former kind, it is worth emphasizing that in Benabderrazik et al. [65] and Gómez Martín et al. [66], there was an involvement of stakeholders during the modeling process, which simultaneously validates the model and improves the understanding of the system's behavior. Regarding the latter case, in Barati et al. [70] the objective is to optimize cropping patterns, while in Silva and Teixeira [77] propose actions to improve institutional relationships in the context of water governance. Next, in Barati et al. [74], although

Reference	Study site	Focus	Approach (tactical, strategic, or descriptive)	Simulation period or time horizon	Time resolution
[64]	Aquifer of Hamedan-Bahar watershed, Iran	Aquifer	Strategic	2019–2050	0.0625 months
[56]	Rio Grande Basin, USA	Watershed	Strategic	2017–2099	1 year
[65]	Two sites in Morocco	Groundwater and farmers income	Strategic	2008–2050	Not stated explicitly
[66]	Medina del Campo, Spain	Watershed	Strategic	100 years	1 month
[67]	Yangtze Economic Belt, China	Integrated assessment model	Strategic	1990–2100	1 year
[68]	Kerman catchment, Iran	Watershed	Strategic	2012–2037	1 month
[69]	Najaf-Abad plain, Iran	Watershed	Descriptive	2001–2012	1 month
[70]	Kermanshah plain, Iran	Watershed	Tactical	10 years	1 month
[71]	Yongding watershed, China	Watershed	Strategic	2006–2035	Not stated
[53]	Coastal aquifers, mediterranean	Aquifer	Descriptive	50 years	Not stated
[72]	Hashtgerd Plain, Iran	Watershed	Strategic	2020–2050	Not stated
[73]	Iran	National	Strategic	2020–2050	1 year
[74]	Qazvin irrigation network, Iran	Irrigation network	Descriptive	2011–2021	1 year
[75]	Azarshahr watershed, Iran	Watershed	Descriptive	2006–2016	1 year
[76]	Sao Carlos, Sao Paulo, Brazil	Watershed	Strategic	2008–2028	1 year
[77]	Lenjanat subbasin, Iran	Watershed	Tactical	5 years	1 month
[78]	Shanshan county, China	County	Strategic	2006–2030	1 month
[79]	Tehran, Iran	City	Strategic	2011–2041	1 year
[80]	Firuzabad Plain aquifer, Iran	Aquifer	Strategic	1992–2042	1 year

Table 2.
Records from systematic review.

not in a predictive fashion, the authors focus on the management of irrigation infra-structure. Finally, the third classification looks to describe the past behavior of the system and identify which policies modified the indicators of the system.

Another element worth comparing between the studies is the time resolution. The time resolution is the time-step considered during the performance of the simulation. In effect, eight of the studies used a yearly time-step, while six used a monthly time step, four did not state this information nor it is visible on the plots, and one uses a time step of 0.0625 months (1.875 days). The time-step is relevant, especially in those regions where there is intra-annual variability, to make more accurate decisions based on the simulations.

There is no doubt that the most pressing impact of climate change is felt through water resources, but the specific effects must be assessed in local studies. In effect, four studies considered climate change scenarios in their analysis. Interestingly, in Afruzi et al. [64] it was reported that, in the aquifer of Hamedan-Bahar watershed in Iran, current water management practices pose a more serious threat to groundwater resources sustainability than climate change. Also, in the Yangtze Economic Belt, climate change impacts in the economy, energy and food systems are negligible as long as the scenario is not RCP 8.5 [67]. Conversely, in the Hashtgerd Plain, Iran, climate change will intensify the water shortage [72] and, in Medina del Campo, Spain, it can turn Nature-Based Solutions insufficient in a worst-case climate change scenario. These effects were assessed through climate change effects in precipitation and temperature, but other effects must be also considered like soil degradation, plant pathogens, pests, increase of ground-level ozone [66], land yield, and extreme hydrometeorological disasters [67]. Considering these findings, studies of the impact of climate change in groundwater resources should be encouraged to include further effects but also to observe which are the effects in other regions of the world. For that matter, the studies presented in this review prove that system dynamics is a useful approach.

There is presence of coupling of system dynamics with other techniques like Compromise Programming, where it was used to rank strategies based on indicators [68]; Nash bargaining method, a quantitative model for conflict resolution in water allocation [69]; System of Environmental and Economic Accounts for Water (SEEA-Water) [75], to diagnose local water security indicators; CA-Marko model, to simulate the prediction of land use [71]; optimization [72, 73] and Agent-based model, to simulate the interaction between stakeholders and agencies [77]. In the context of groundwater management, there is a diversity of combination methods that have been used in the last 5 years, to complement system dynamics modeling, something that was suggested in Refs. [46, 54, 55].

6. Conclusion: opportunities for the application of system dynamics in groundwater management

The water crisis is already happening and one of its manifestations, although silent and lagging, is on groundwater. The paradigm shift of water management must manifest in practice urgently. This implies a radical change of thought where water management is not addressed from a unidimensional perspective, but rather from a multidisciplinary one. In this context, there is a requirement of simulation models that allow this multidisciplinary to be included, along with the role of stakeholders in water issues and effects of climate change in the hydrological cycle. These simulation models must inform public policy and projects aimed at achieving water security.

Several approaches are available in the literature, among which system dynamics is found. System dynamics models can be predictive, descriptive, or participatory and may have multiple goals, like supporting public policy and/or increasing the

understanding of the system among stakeholders; also, they possess valuable features that allow addressing water management complexity.

There are several opportunities for system dynamics application to water management. Firstly, to integrate several water subsystems. This is already happening in the context of groundwater management with aquifers, which are usually modeled in the broader context of a watershed. However, there are opportunities to include them in disaster mitigation (droughts, for instance), water quality, groundwater-dependent ecosystems, and water supply. Also, it is necessary to include the effects of climate change and variability in water systems and in aquifers particularly.

In addition, water systems must be included within broader human systems: food security, economic security, energy security, and environmental security. For example, an interesting opportunity is to model land-use policies to analyze their effects on aquifer recharge and surface runoff. Another one deals with the analysis of the water-energy nexus in the context of unconventional hydrocarbons exploitation. Finally, the inclusion of water pollution control infrastructure to include reuse or managed aquifer recharge (MAR) in water policy can be assessed.

Certainly, to materialize these applications system dynamics would have to be integrated with other methods, such as GIS, RS, hydrologic models, water quality models, or reliability methods. In addition, to incorporate social and political processes there is a need to add qualitative variables, as well as to increase the interaction with stakeholders, not only during the technology transfer and divulgation but during the modeling process itself. Moreover, we must continue developing tools to validate the models. This is especially important for groundwater models because usually aquifer data is scarce and therefore, models cannot be calibrated.

System dynamics represents not only a modeling approach but a radical change of thought that is appealing to the interconnected crises that the world lives in; in particular, to the water crisis. Countries where the water crisis is already evident, like Mexico, must embrace system dynamics to inform public policies and projects that improve water management. This chapter elaborated on why this is necessary and why system dynamics represent an interesting opportunity to address these issues.

Conflict of interest

The authors declare that they have no competing interests.

Notes/thanks/other declarations

The first author would like to thank Universidad de las Americas Puebla (UDLAP); the scholarship and resources provided for this research is acknowledged.

The first author would like to thank Consejo Nacional de Ciencia y Tecnología (Conacyt); the scholarship granted for this research is acknowledged.

IntechOpen


IntechOpen

Author details

David Eduardo Guevara-Polo and Carlos Patiño-Gómez*
Department of Civil and Environmental Engineering, Universidad de las Américas
Puebla, San Andrés Cholula Puebla, Mexico

*Address all correspondence to: carlos.patino@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Jakeman AJ, Barreteau O, Hunt RJ, Rinaudo JD, Ross A. Integrated groundwater management: Concepts, approaches and challenges. In: *Integrated Groundwater Management: Concepts, Approaches and Challenges*. Berlin, Germany: Springer; 2016. pp. 1-762. DOI: 10.1007/978-3-319-23576-9
- [2] Aeschbach-Hertig W, Gleeson T. Regional strategies for the accelerating global problem of groundwater depletion. *Nature Geoscience*. 2012;5(12):853-861. DOI: 10.1038/ngeo1617
- [3] Cerecedo Arroyo ME, Martínez Austria PF. Dynamic water system modeling: A systematic review. *Water Practice Technology*. 2021;16(3):744-755. DOI: 10.2166/wpt.2021.051
- [4] Fienen MN, Arshad M. The international scale of the groundwater issue. In: *Integrated Groundwater Management*. Berlin, Germany: Springer; 2016
- [5] Capra F, Luisi PL. *The Systems View of Life*. Cambridge, United Kingdom: Cambridge University Press; 2014
- [6] WEF. *The Global Risks Report 2020*. Davos, Switzerland: World Economic Forum; 2019
- [7] Martínez-Austria PF. Modelos dinámicos para la gestión de la cuenca del río Bravo. In: *La cuenca del río Bravo y el cambio climático*. Puebla, Mexico: UDLAP; 2018. pp. 210-229
- [8] Simonovic SP. *Managing Water Resources: Methods and Tools for a Systems Approach*. Paris, France: UNESCO Publishing; 2009
- [9] Hofste RW, et al. Technical Note Aqueduct 3.0: Updated Decision-Relevant Global Water Risk Indicators [Online]. Technical Note. 2019. pp. 1-53. Available from: https://files.wri.org/s3fs-public/aqueduct-30-updated-decision-relevant-global-water-risk-indicators_1.pdf
- [10] Alley WM, Healy RW, LaBaugh JW, Reilly TE. Flow and storage in groundwater systems. *Science*. 2002;296(5575):1985-1990. DOI: 10.1126/science.1067123
- [11] UN Water, *Groundwater Overview: Making the Invisible Visible*. Geneva, Switzerland: UN Water; 2018
- [12] Herbert C, Döll P. Global assessment of current and future groundwater stress with a focus on transboundary aquifers. *Water Resources Research*. 2019;55(6):4760-4784. DOI: 10.1029/2018WR023321
- [13] Fonseca-Ortíz CR, Mastachi-Loza CA, Díaz-Delgado C, Esteller-Alberich MV. The water-energy-food nexus in Mexico. In: Raynal-Villasenor JÁ, editor. *Water Resources of Mexico*. Berlin, Germany: Springer; 2020
- [14] Conapo. Población a mitad de año. 2020. Available from: <https://datos.gob.mx/busca/dataset/proyecciones-de-la-poblacion-de-mexico-y-de-las-entidades-federativas-2016-2050/resource/ac0cad37-004b-45a4-831a-2538c13819cb>
- [15] Marengo-Mogollón H. Water resources in Mexico: Some proposals for the future. In: *Water Resources of Mexico*. Berlin, Germany: Springer; 2020
- [16] Arreguin Cortes FI, Cervantes Jaimes CE. *Water Security and Sustainability in Mexico*. Vol. 6. Berlin, Germany: Springer; 2020

- [17] Marin LE et al. Virtual capacity building in water resources management as a strategy to cope with hydrometeorological risks and climate change in Mexico. In: Hydrometeorological Risks and Climate Change 2014. Puebla, Mexico: UDLAP; 2015
- [18] Martínez-Austria PF. Climate change and water resources. In: Water Resources of Mexico. Berlin, Germany: Springer; 2020
- [19] Rivas-Acosta I. Efectos del cambio climático en el recurso hídrico de México (agua superficial). In: Arreguin Cortes FI, López-Pérez M, Rodríguez-López O, Montero Martínez MJ, editors. Atlas de vulnerabilidad hídrica en México ante el cambio climático. Morelos, Mexico: IMTA; 2015
- [20] Montero Martínez MJ, Martínez Jiménez J, Castillo Pérez NI, Espinoza Tamarindo BE. Escenarios climáticos en México proyectados para el siglo XXI: precipitación y temperaturas máxima y mínima. In: Martínez-Austria PF, Patiño-Gómez C, editors. Efectos del cambio climático en los recursos hídricos de México. Volumen III. Atlas de vulnerabilidad hídrica en México ante el cambio climático. Morelos, Mexico: IMTA; 2010
- [21] Martínez-Austria PF, Díaz-Delgado C, Moeller-Chavez G. Seguridad hídrica en México: diagnóstico general y desafíos principales. Ingeniería del Agua. 2019;23(2):107. DOI: 10.4995/ia.2019.10502
- [22] Herrera-Pantoja M, Hiscock KM. Projected impacts of climate change on water availability indicators in a semi-arid region of central Mexico. Environmental Science & Policy. 2015;54:81-89. DOI: 10.1016/j.envsci.2015.06.020
- [23] Molina-Navarro E, Hallack-Alegría M, Martínez-Pérez S, Ramírez-Hernández J, Mungaray-Moctezuma A, Sastre-Merlín A. Hydrological modeling and climate change impacts in an agricultural semiarid region. Case study: Guadalupe River basin, Mexico. Agricultural Water Management. 2016;175:29-42. DOI: 10.1016/j.agwat.2015.10.029
- [24] Escalante-Sandoval C, Núñez-García P. Effect of climate change on occurrence of meteorological drought in Zacatecas, Mexico. In: Raynal-Villasenor JÁ, editor. Hydrometeorological Risks and Climate Change 2014. Puebla, Mexico: UDLAP; 2015
- [25] Conagua. Estadísticas del agua en México. 2018
- [26] Ortega-Gaucin D, Velasco-Velasco I, López-Pérez M, Cardoso-García J. Program of preventive and mitigation drought measures of the Rio Balsas basin council (PPMDM-RBBC). In: Raynal-Villasenor JÁ, editor. Hydrometeorological Risks and Climate Change 2014. Puebla, Mexico: UDLAP; 2015
- [27] Alcocer-Yamanaka VH, Chávez-Guillén R. Aprovechamiento del agua subterránea. In: H2O Gestión del agua. Mexico City, Mexico: Sacmex; 2016
- [28] Hatch-Kuri G. Capítulo 7. Agua subterránea en México: retos y pendientes. In: Aguas Subterráneas en México. Mexico City, Mexico: FES Transformación; 2018. pp. 149-170
- [29] Amanambu AC et al. Groundwater system and climate change: Present status and future considerations. Journal of Hydrology. 2020;589:125163. DOI: 10.1016/j.jhydrol.2020.125163
- [30] Raynal-Gutierrez ME. Water use and consumption: Industrial and domestic. In: Raynal-Villasenor JÁ, editor. Water

Resources of Mexico. Berlin, Germany: Springer; 2020

[31] Gutiérrez-Ojeda C, Escolero-Fuentes OA. Groundwater resources of Mexico. In: Raynal-Villasenor JÁ, editor. *Water Resources of Mexico*. Berlin, Germany: Springer; 2020

[32] DOF. Norma Oficial Mexicana NOM-011-Conagua-2015. In: *Conservación del recurso agua-Que establece las especificaciones y el método para determinar la disponibilidad media anual de las aguas nacionales*. DOF—Diario Oficial de la Federación; 2015. p. 8

[33] DOF. Acuerdo por el que se actualiza la disponibilidad media anual de agua subterránea de los 653 acuíferos de los Estados Unidos Mexicanos. In: *mismos que forman parte de las regiones hidrológico-administrativas que se indican* [Online]. 2020. pp. 1-10. Available from: http://www.dof.gob.mx/nota_detalle.php?codigo=5600593&fecha=17/09/2020

[34] Ocampo-Astudillo A, Garrido-Hoyos SE, Salcedo-Sánchez ER, Martínez-Morales M. Alteration of groundwater hydrochemistry due to its intensive extraction in urban areas from Mexico. In: Otazo-Sánchez EM, Navarro-Frómata AE, Singh VP, editors. *Water Availability and Management in Mexico*. Berlin, Germany: Springer; 2020

[35] Camacho LM, Gutiérrez M, Alarcón-Herrera MT, de Villalba L, Deng S. Occurrence and treatment of arsenic in groundwater and soil in Northern Mexico and Southwestern USA. *Chemosphere*. 2011;**83**(3):211-225. DOI: 10.1016/j.chemosphere.2010.12.067

[36] Esteller MV, Diaz-Delgado C. Environmental effects of aquifer overexploitation: A case study in the highlands of Mexico. *Environmental*

Management. 2002;**29**(2):266-278. DOI: 10.1007/s00267-001-0024-0

[37] Acuña-Lara F, Pacheco-Martinez J, Luna-Villavicencio H, Hernaoández-Marioón M, Gonzaoólez-Cervantes N. Infiltration of surface water through subsidence failure assessment applying electric prospecting, case Aguascalientes Valley, Mexico. *Proceedings of the International Association of Hydrological Sciences*. 2020;**382**:5-9. DOI: 10.5194/piahs-382-5-2020

[38] León WHH, Martínez JP, Marín MH, Cenicerros RP. Land subsidence and its effects on the urban area of Tepic city, Mexico. *WIT Transactions on the Built Environment*. 2018;**179**:369-380. DOI: 10.2495/UG180341

[39] Castellazzi P et al. Land subsidence in major cities of Central Mexico: Interpreting InSAR-derived land subsidence mapping with hydrogeological data. *International Journal of Applied Earth Observation and Geoinformation*. 2016;**47**:102-111. DOI: 10.1016/j.jag.2015.12.002

[40] Chen Z, Wei S. Application of system dynamics to water security research. *Water Resources Management*. 2014;**28**(2):287-300. DOI: 10.1007/s11269-013-0496-8

[41] Mirchi A, Madani K, Watkins D, Ahmad S. Synthesis of system dynamics tools for holistic conceptualization of water resources problems. *Water Resources Management*. 2012;**26**(9):2421-2442. DOI: 10.1007/s11269-012-0024-2

[42] Meadows DH, Meadows DL, Randers J, Behrens WW. *Limits to Growth*. Virginia, United States: Potomac Associates; 1972

[43] Sachs W. *Planet Dialectics, Explorations in Environment and*

- Development. London, United Kingdom: Zed Books; 1999
- [44] Senge PM. *The Fifth Discipline, the Art and Practice of the Learning Organization*. New York: Currency; 1990
- [45] Bui HTM. From the fifth discipline to the new revolution: What we have learnt from Senge's ideas over the last three decades. *The Learning Organization*. 2020;**27**(6):489-498. DOI: 10.1108/TLO-04-2020-0062
- [46] Sterman JD. *Business Dynamics*. New York, United States: McGraw-Hill; 2000
- [47] Gain AK, Hossain S, Benson D, Di Baldassarre G, Giupponi C, Huq N. Social-ecological system approaches for water resources management. *International Journal of Sustainable Development and World Ecology*. 2021;**28**(2):109-124. DOI: 10.1080/13504509.2020.1780647
- [48] Hassan Q, Garg NK. Systems approach for water resources development. *Global Journal of Flexible Systems Management*. 2007;**8**(4):29-43. DOI: 10.1007/BF03396531
- [49] Renouf MA, Kenway SJ. Evaluation approaches for advancing urban water goals. *Journal of Industrial Ecology*. 2017;**21**(4):995-1009. DOI: 10.1111/jiec.12456
- [50] Ghalekhondabi I, Ardjmand E, Young WA, Weckman GR. Water demand forecasting: Review of soft computing methods. *Environmental Monitoring and Assessment*. 2017;**189**(7):313. DOI: 10.1007/s10661-017-6030-3
- [51] Winz I, Brierley G, Trowsdale S. The use of system dynamics simulation in water resources management. *Water Resources Management*. 2009;**23**(7):1301-1323. DOI: 10.1007/s11269-008-9328-7
- [52] Crawford NH, Burges SJ. History of the Stanford Watershed Model. *Water Resources IMPACT*. 2004;**6**:2
- [53] Martínez-Valderrama J, Ibáñez J, Alcalá FJ. AQUACOAST: A simulation tool to explore coastal groundwater and irrigation farming interactions. *Scientific Programming*. 2020;**2020**:1-20. DOI: 10.1155/2020/9092829
- [54] Hong-Gang X. Exploring effective policies for underground water management in artificial oasis: A system dynamics analysis of a case study of Yaoba Oasis. *Journal of Environmental Sciences*. 2001;**13**(4):476-480
- [55] Assaf H. A hydro-economic model for managing groundwater resources in semi-arid regions. *WIT Transactions on Ecology and the Environment*. 2009;**125**:85-96. DOI: 10.2495/WRM090091
- [56] Bai Y, Langarudi SP, Fernald AG. System dynamics modeling for evaluating regional hydrologic and economic effects of irrigation efficiency policy. *Hydrology*. 2021;**8**:2
- [57] Mashaly AF, Fernald AG. Identifying capabilities and potentials of system dynamics in hydrology and water resources as a promising modeling approach for water management. *Water (Switzerland)*. 2020;**12**:294-304
- [58] Zomorodian M, Lai SH, Homayounfar M, Ibrahim S, Fatemi SE, El-Shafie A. The state-of-the-art system dynamics application in integrated water resources modeling. *Journal of Environmental Management*. 2018;**27**:294-304
- [59] Luna-Reyes LF, Andersen DL. Collecting and analyzing qualitative

data for system dynamics: Methods and models. *System Dynamics Review*. 2003;**19**(4):271-296. DOI: 10.1002/sdr.280

[60] Zarghami SA, Gunawan I, Schultmann F. System dynamics modelling process in water sector: A review of research literature. *Systems Research and Behavioral Science*. 2018;**35**(6):776-790. DOI: 10.1002/sres.2518

[61] Mirchi A. System dynamics modeling as a quantitative-qualitative framework for sustainable water resources management: Insights for water quality policy in the Great Lakes Region. In: Submitted in Partial Fulfillment of the Requirements for the Degree in Civil Engineering, Michigan Technological University, USA. 2013

[62] Schwaninger M, Grösser S. System dynamics as model-based theory building. *Systems Research and Behavioral Science*. 2008;**25**(4):447-465. DOI: 10.1002/sres.914

[63] Randers. Guidelines for model conceptualization. In: *Elements of the System Dynamics Method*. Massachusetts, United States: MIT Press; 1980. pp. 117-139

[64] Afruzi A, Zare Abyaneh H, Abdolabadi H. Local strategies to manage groundwater depletion under climate change scenarios—A case study: Hamedan-Bahar Plain (Iran). *Arabian Journal of Geosciences*. 2021;**14**:1548

[65] Benabderrazik K, Kopainsky B, Tazi L, Joerin J, Six J. Agricultural intensification can no longer ignore water conservation—A systemic modelling approach to the case of tomato producers in Morocco. *Agricultural Water Management*. 2021;**256**:107082. DOI: 10.1016/j.agwat.2021.107082

[66] Gómez Martín E, Máñez Costa M, Egerer S, Schneider UA. Assessing the

long-term effectiveness of Nature-Based Solutions under different climate change scenarios. *Science of the Total Environment*. 2021;**794**:148515

[67] Jiang H, Simonovic SP, Yu Z, Wang W. What are the main challenges facing the sustainable development of China's Yangtze economic belt in the future? An integrated view. *Environmental Research Communications*. 2021;**3**(11):115005

[68] Momeni M, Behzadian K, Yousefi H, Zahedi S. A scenario-based management of water resources and supply systems using a combined system dynamics and compromise programming approach. *Water Resources Management*. 2021;**35**(12):4233-4250. DOI: 10.1007/s11269-021-02942-z

[69] Naghdi S, Bozorg-Haddad O, Khorsandi M, Chu X. Multi-objective optimization for allocation of surface water and groundwater resources. *Science of the Total Environment*. 2021;**776**:146026. DOI: 10.1016/j.scitotenv.2021.146026

[70] Barati K, Abedi Koupai J, Darvishi E, Azari A, Yousefi A. Cropping pattern optimization using system dynamics approach and multi-objective mathematical programming. *Journal of Agricultural Science and Technology*. 2020;**22**(5):1397-1412

[71] Dai D, Sun M, Lv X, Lei K. Evaluating water resource sustainability from the perspective of water resource carrying capacity, a case study of the Yongding River watershed in Beijing-Tianjin-Hebei Region, China. *Environmental Science and Pollution Research*. 2020;**27**(17):21590-21603. DOI: 10.1007/s11356-020-08259-5

[72] Mehrazar A, Massah Bavani AR, Gohari A, Mashal M, Rahimikhoob H.

Adaptation of water resources system to water scarcity and climate change in the suburb area of megacities. *Water Resources Management*. 2020;**34**(12):3855-3877. DOI: 10.1007/s11269-020-02648-8

[73] Barati AAAA, Azadi H, Scheffran J. A system dynamics model of smart groundwater governance. *Agricultural Water Management*. 2019;**221**:502-518. DOI: 10.1016/j.agwat.2019.03.047

[74] Hosseinzadeh Ghazichaki Z, Monem MJ. Development of quantified model for application of control systems in irrigation networks by system dynamic approach. *Irrigation and Drainage*. 2019;**68**(3):433-442. DOI: 10.1002/ird.2331

[75] Mahdavi T, Bagheri A, Hosseini SAA. Applying the System of Environmental and Economic Accounts for Water (SEEA-Water) for integrated assessment of water security in an aquifer scale—Case study: Azarshahr Aquifer, Iran. *Groundwater for Sustainable Development*. 2019;**9**:100261

[76] Silva SP, Teixeira BAN. Modelling system dynamics to evaluate urban water supply management and production of future scenarios. *Journal of Urban and Environmental Engineering*. 2019;**13**(2):317-328

[77] Ohab-Yazdi SA, Ahmadi A. Using the agent-based model to simulate and evaluate the interaction effects of agent behaviors on groundwater resources, a case study of a sub-basin in the Zayandehroud River basin. *Simulation Modelling Practice and Theory*. 2018;**87**:274-292. DOI: 10.1016/j.simpat.2018.07.003

[78] Chen C, Ahmad S, Kalra A, Xu ZX. A dynamic model for exploring water-resource management scenarios in an

inland arid area: Shanshan County, Northwestern China. *Journal of Mountain Science*. 2017;**14**(6):1039-1057. DOI: 10.1007/s11629-016-4210-1

[79] Ghasemi A, Saghafian B, Golian S. System dynamics approach for simulating water resources of an urban water system with emphasis on sustainability of groundwater. *Environment and Earth Science*. 2017;**76**(18)

[80] Mokhtar A, Aram S. Systemic insights into agricultural groundwater management: Case of Firuzabad Plain, Iran. *Water Policy*. 2017;**19**(5):867-885. DOI: 10.2166/wp.2017.159

Chapter

Influences of ENSO, AMO, and PDO over Dry Periods in Mexico: A Systematic Review (1984–2020)

Regina Mijares-Fajardo and Polioptro F. Martínez-Austria

Abstract

Forecasting climatic variables through teleconnections has become more common for different planning activities. In Mexico, different sea surface temperature (SST) patterns have shown to force droughts, consequently, precipitation anomalies could be anticipated with the study of these patterns. To our knowledge, this is the first study that reviews the individual and coupled influences of ENSO, AMO, and PDO over dry periods in Mexico. Overall, through a systematic review using the PRISMA methodology, 75 documents were analyzed, and their conclusions were compared. In general, this study concludes that the majority of the authors coincided in the following teleconnections: La Niña (El Niño) forces dry (wet) conditions in northern Mexico in both winter and summer, likewise, in this region dry periods are associated with AMO+ during summer and PDO– during winter. On the other hand, for central and southern Mexico, for both winter and summer, El Niño generally drives drought conditions while La Niña has been linked with greater than average humidity, in this region, AMO– (AMO+) and PDO+ (limited information) induce dry (humid) circumstances. Thus, understanding drought trends in México driven by climatic indices is imperative to predict situations that threaten the natural resources and economic activities. With the results obtained, it is possible to deepen with greater certainty in the investigation of the influence of these teleconnections.

Keywords: teleconnections, drought, Mexico, Atlantic Multidecadal Oscillation, Pacific Decadal Oscillation, El Niño-Southern Oscillation

1. Introduction

Droughts are one of the most threatening natural disasters for human civilization and the ecosystems. Even though there is no universally accepted definition for these phenomena in the scientific literature [1], droughts are characterized by the prolonged absence or pronounced deficiency of precipitation [2]. This negative anomaly in precipitation could be driven by numerous factors, however, changes in SST (sea surface temperature) have shown to be one of the main forces that induce droughts [3]. Moreover, in the late Holocene, three main SST patterns: the Pacific

Decadal Oscillation (PDO), El Niño Southern Oscillation (ENSO), and Atlantic Multidecadal Oscillation (AMO), have been associated to significant changes in precipitation regimes over Mexico [4, 5].

On one side, PDO reflects the monthly variations of Pacific SST north of the 20°N with a periodicity of 15–20 years [6]. The main climatic effects of the PDO are concentrated in the North Pacific and southwestern North America, primarily during the winter season, and can be responsible for up to 50% of the annual precipitation variability [7]. Some authors suggest that during its positive phase, the monsoonal precipitation is reduced in central Mexico [5]. The AMO, on the other side, measures medium-term cycles (20–40 years) of positive and negative detrended anomalies in North Atlantic SST [6, 8]. The AMO variability has been linked with changes of climate and extreme events [8]. During boreal summer, AMO has been associated with intensification of the hurricane season in the Atlantic and droughts in North America [9].

Lastly, ENSO is the most studied climatic oscillation because it affects atmospheric circulation throughout the world [6, 7, 10, 11]. Moreover, several authors consider that ENSO is the most important factor for interannual climatic variation in Mexico [7, 12–15]. ENSO is made up of two phenomena; its oceanic component, El Niño, refers to the positive anomaly of at least 0.5°C in the equatorial Pacific SST for six consecutive months or more [13]. La Niña reflects the opposite. On the other hand, its atmospheric component, the Southern Oscillation Index (SOI), is an air mass equilibrium movement between the Pacific and Indo-Australian areas [10]. Although El Niño and SOI are not perfectly correlated in terms of minor variations, large negative SOI values are associated with warm ONI (Oceanic Niño Index) events [10]. Additionally, the PDO has the potential to act simultaneously with ENSO, and depending on the phasing, it can enhance its effects [16].

Teleconnections are statistical associations between climatic variables perceptible in spatially separated areas [17]. The studies reviewed in this document analyzed the teleconnections between dry periods in México and SST patterns. For this purpose, the majority of authors measured dry periods with precipitation anomalies, which refer to a deviation from the average registered in a determined location. These precipitation anomalies were estimated by numerous methods; some authors used standardized anomalies of monthly, daily, or yearly data; others used the average annual or monthly accumulated precipitation. To define dry periods, other authors used the Standardized Precipitation Index (SPI), whose use was suggested by the World Meteorological Organization [18] as a general recommendation to characterize meteorological drought. Additional indices used were the Palmer Drought Severity Index (PDSI), Tropical Rainfall Index (TRI), and the Outgoing Longwave Radiation Precipitation Index (OPI). Proxy data such as tree ring width and information from ancient sediments was used to estimate dry conditions as well.

Droughts, unlike most natural disasters but like a disease, generally begin before the symptoms appear [19]. Therefore, understanding their causes and expected variations is critical [20]. The potential of forecasting droughts through the study of SST patterns relies in providing time to take mitigation actions. This review paper aims to synthesize the main findings regarding the effects that AMO, ENSO, and PDO can have over dry periods in Mexico. To achieve this goal, the following research questions were addressed.

- RQ1: What are the main conclusions achieved by the majority of studies on the topic?

- RQ2: How rare or common are divergent conclusions among different authors?
- RQ3: What are the climatic oscillations/seasons/regions that have received the most attention among studies in Mexico?

Moreover, this is the first study that reviews not only the individual but the coupled influences of ENSO, AMO, and PDO over droughts in Mexico. Findings in this review will be useful for local authorities, decision-makers, academics, among others. Furthermore, by identifying the least explored research directions, this review will also help researchers when selecting potential areas of study in the field.

2. Methods

This study was conducted and reported with the aid of one of the most prevalent systematic literature guides [21]. First, the database Elsevier's Scopus was selected because it is the largest of peer-reviewed literature. Second, the search string was defined as ((drought) AND (ENSO) OR (El AND Niño) OR (PDO) OR (Pacific AND Decadal AND Oscillation) OR (Atlantic AND Multidecadal AND Oscillation) OR (AMO) AND ((review) OR (Mexico))) as to include the different indices in all their written forms. The reason to include review or Mexico was to find documents that comprised both, studies focused in Mexico and documents that reviewed the main global effects or effects in a broader region, which contained national information. The search string was applied to article titles, abstracts, and keywords. After the search string was applied, 295 documents that were published until the end of 2020 were obtained. Manual search was conducted to include relevant content from other sources, 11 documents were selected for screening. Later on, the PRISMA method [22] was employed for selecting the documents to be included in this review.

After excluding the studies based on areas that were not in the field, excluding the one document written in Chinese, and excluding the least relevant documents, 221 documents were left. After the titles and reviews were screened, 137 documents remained for full-text reading, in this phase, 56 full-text articles were excluded because neither the effects of ENSO, AMO, or PDO over precipitation in México were addressed. In the last phase, six documents were eliminated because they focused on the physical mechanisms involved in teleconnections but did not mention their specific effects on precipitation over Mexico. This information is presented in **Figure 1**. The 75 documents that were comprised in this systematic review include those in which their primary objective was to study the teleconnections of interest and those which broach the subject as a partial topic. These 75 documents also include papers of primary research (self-conducted) as well as secondary research (conducted by others). Lastly, even though the search string only involved the word "drought," most of the documents studied humid periods as well, therefore, wet conditions are also included in this systematic review.

3. Results

For the study of climatic oscillations, the ENSO phenomenon was studied through the use of different indices such as El Niño 3/3.4 index, SOI, ONI, and Extended Multivariate ENSO Index (MEI). Only a minority of authors referred to this oscillation

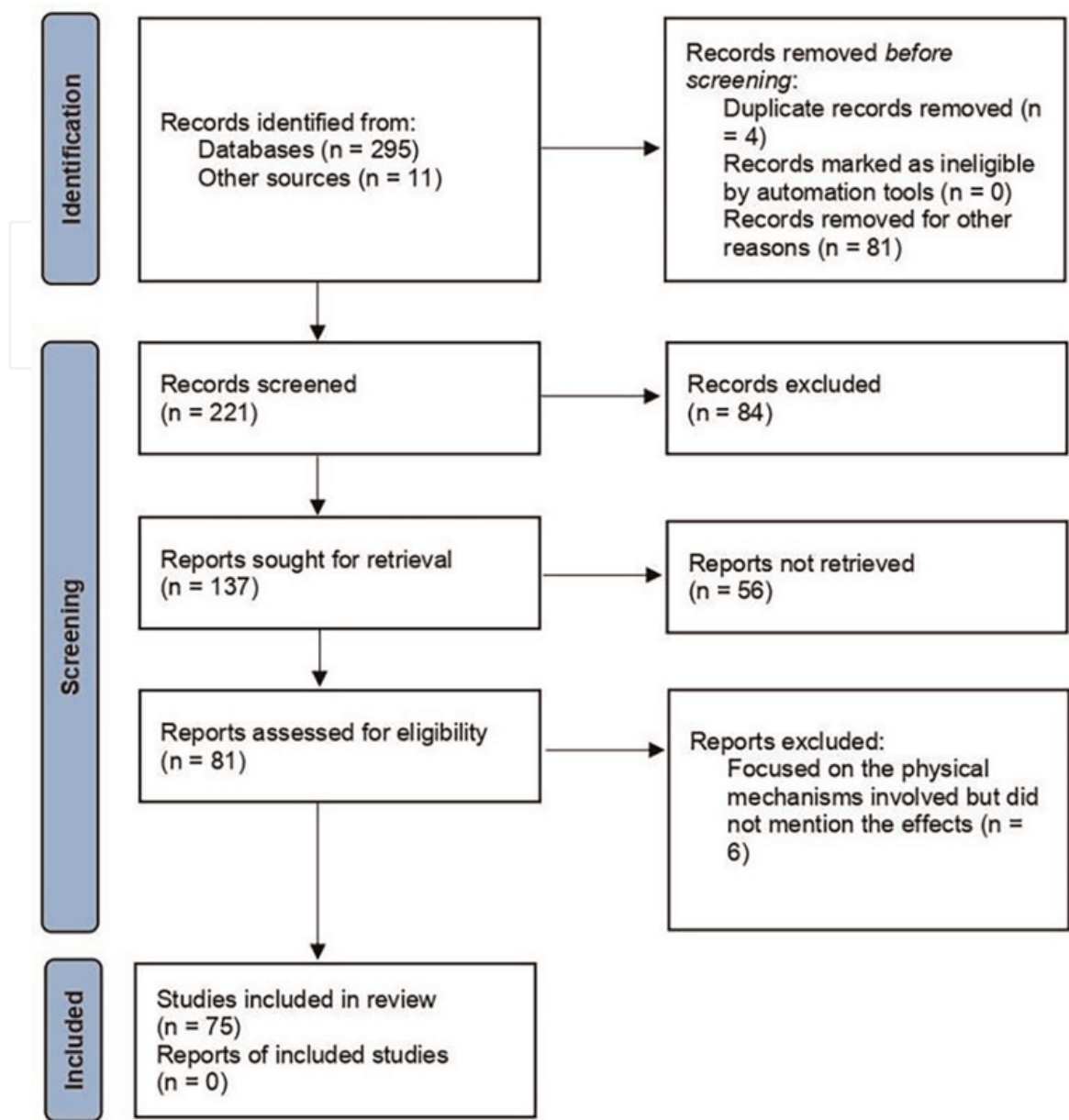


Figure 1.
PRISMA flowchart for document selection.

by means of anomalies in the equatorial eastern Pacific SST [23–26]. On the other hand, the PDO events were measured directly by the original index, except for Mo and Schemm [3] whose study was considered, despite the fact that they contemplate North Pacific SST anomalies. Finally, the AMO oscillation was referred to, in less than 30% of the cases, as North Atlantic SST anomalies. The other 70% of the authors used the AMO index to express their conclusions, and only Ruprich-Robert et al. [27] used the Atlantic Multidecadal Variability (AMV) Index to identify teleconnections.

Table 1 summarizes the number of articles divided by area of study. What stands out in the table is that the majority of studies reviewed, published from 1984 to 2020, focused in northwestern Mexico or in a region that comprises part of the country or the country as a whole (i.e., subtropical North America or Tex-Mex region). On the other side, northeastern Mexico has been the area with less focused investigations.

Turning now to **Table 2**, an overview of all conclusions reached by the authors is presented by region of the country and by the season of the year. The category “most

Region of study			# Articles
Mexico	North	East	2
		West	19
	Center		9
	Center and South		6
	South		4
	Whole country		9
Region that includes Mexico			17
Whole world			9

Table 1.
Studies per region.

of the territory” includes those authors who referred to Mexico as a whole region, without distinguishing the north, center or south. All 75 articles analyzed in this document, published between 1984 and 2020, are gathered in this table. Also, the cells colored in blue denote that at least one conclusion contained is divergent from the majority of the authors, with majority meaning over 50%. Likewise, cells colored in green signify that two or more conclusions there contained, are divergent from the majority of the authors. Interestingly, although the modes of climate variability cannot fully explain the physical mechanisms that drive extreme events in precipitation patterns [77], most of the authors agree on general conclusions, thus, only a minority of cells is colored.

Furthermore, **Table 3** shows the summary statistics for the percentage of matching and mismatching conclusions. Cells colored in green denote the conclusions repeated by two or less authors, these are referred to as limited information given the lack of sufficient studies to support the conclusion. Cells colored in magenta highlight the conclusions with more than 50% of coincidence. This percentage is obtained by dividing the number of articles that agreed on the condition caused by a specific index, during a specific season, in a specific region, over the total number of articles that addressed the effects of that climatic index for that season and in that region.

From the chart, it can be seen that by far, the most studied SST pattern is ENSO with a total of 158 conclusions driven by different authors. This number is more than twice the sum of the conclusions driven for AMO [34] and PDO [39]. Closer inspection of the table shows that the northern part of the country has received the most attention among different authors; in total, 107 conclusions were reached for this region with the most teleconnections studied during winter [37] and in unspecified season [53], which in most cases means annual conclusions. Conversely, the central and southern regions of Mexico received less attention, with 87 conclusions that were mostly reached for unspecified season or the whole year [44], and for the summer season [28]. Finally, for articles that mentioned the country in general, 39 conclusions were found that were distributed quasi homogeneously in all seasons.

Lastly, **Table 4** provides the results obtained from the systematic review of 75 articles that focused primarily or partially on the study of the teleconnections of interest. Cells colored in magenta show the general conclusions in which more than 50% of the articles agree but at least three authors studied the corresponding phase of the climatic oscillation, in the corresponding region, during the corresponding season.

Precipitation anomalies over Mexico									
North			Center and south			Most of the territory			
Ref.	Summer	dns	Winter	Summer	dns	Winter	Summer	dns	Winter
[28]			H (EN)						
[29]			H (EN)						
[30]	H (EN)	D (LN) H (EN)		H (EN)					
[31]				H (EN)					
[32]								D (EN)	H (LN)
[33]			H (EN)						
[34]								D (EN)	
[35]		H (EN)			D (EN) Ctr				
[36]			H (EN) Nw						
[37]			D (LN) H (EN)						
[38]	D (LN) H (EN)								
[39]				D (EN) Sth					
[14]			D (LN) W-S						
[40]								D (EN)	
[41]		D (LN) H (EN) Nw							
[42]				D (EN) Ctr		D (EN) Ctr	D (EN)		D (EN)
[43]			D (LN) Nw						
[44]			D (LN) H (EN) Nw			D (EN) H (LN) Sth	D (EN)		D (EN) Os
[45]		D (LN, -PDO) H (EN, +PDO) Nw							

Precipitation anomalies over Mexico									
North				Center and south			Most of the territory		
Ref.	Summer	dns	Winter	Summer	dns	Winter	Summer	dns	Winter
[46]			H (EN)						
[47]					D (EN) Ctr				
[48]			H (EN)	H (EN)	H (EN)	D (EN)	D (EN)		
[49]		D (EN, +PDO) Nw			D (EN, +PDO, –SOI, –AMO) Sth				
[50]			H (EN) Ne	D (EN) H (LN) TSt					
[51]			H (EN)						
[9]	D (+AMO) ASO								
[3]		D (+SST in NP, NA) H (–SST in NP, NA) Nw	H (EN)						
[1]			H (EN)						
[52]	H (EN)			H (EN) Ctr, NS			D (EN)		
[53]				D (–AMO, SAWP) H (+AMO, LAWP) Sth					
[54]				D (EN) Ctr	D (EN) Ctr				
[55]	H (EN)	D (LN, +SST in NA)		D (EN)	D (EN, –SST in NA) H (+SST in NA) Sth				H (EN)
[7]			D (–PDO) H (+PDO)	D (+PDO) H (–PDO) Sth		D (–PDO) H (+PDO) Ctr			H (+PDO)
[56]		D (–PDO, +AMO)		D (EN) H (LN)	D (+PDO, –AMO)		D (EN) H (LN)		

Precipitation anomalies over Mexico									
North			Center and south				Most of the territory		
Ref.	Summer	dns	Winter	Summer	dns	Winter	Summer	dns	Winter
[57]								ID (LN) D (EN, +SST in TNA)	D (LN) H (EN)
[58]					D (EN) Ctr				
[8]		D (LN)							
[59]					D (EN)				
[60]	H (EN)			D (EN)					
[15]						D (LN) H (EN) W-S			
[61]							D (EN)		
[62]			D (LN) H (EN)						
[10]				D (EN)	D (EN) Sth				
[6]							D (EN)		H (EN)
[63]		D (+AMO, −PDO, +SOI) Nw							
[64]			D (LN) W-S, Nw, H (EN) S						
[65]			D (LN) H (EN) W- S						
[5]				D (EN, +PDO, −AMO) H (LN, −PDO, +AMO)	D (+PDO, −AMO)				
[66]					D (EN) H (LN) Sth				

Precipitation anomalies over Mexico									
North				Center and south			Most of the territory		
Ref.	Summer	dns	Winter	Summer	dns	Winter	Summer	dns	Winter
[67]					D (EN) H (LN)				
[23]		D (LN, −PDO, +SST in EEP) H (EN, +PDO, −SST in EEP)							
[68]				D (EN) H (LN)	D (EN) H (LN)	D (LN) H (EN)			
[69]					D (EN) H (LN)				
[24]	D (EN, +SST in ETP)	D (EN, +PDO, −AMO)	D (LN) H (EN) W-S, Nw	D (EN)	D (EN, +PDO, −AMO)				
[27]	D (+SST in NA)	D (+SST in NA)						D (LN)	
[20]		D (LN)			D (EN) Ctr			D (EN)	
[70]					D (EN) H (LN) Sth				
[71]		D (LN) H (EN) Nw							
[72]					D (EN)				
[25]				D (EN) Sth	D (EN, +SST in EEP, −NA) H (LN, −SST in EEP, +NA) Sth				
[73]		D (−PDO) Nw	D (−PDO) Nw						
[26]					H (−SST in EP) Sth		D (EN) H (LN)		
[74]			D (LN) H (EN) Ne						
[75]	D (LN, −PDO, +AMO) Nw								

Precipitation anomalies over Mexico									
North			Center and south				Most of the territory		
Ref.	Summer	dns	Winter	Summer	dns	Winter	Summer	dns	Winter
[76]	H (–AMO) Nw	D (LN, –PDO, +AMO) H (EN, +PDO) Nw			D (EN, –AMO) H (LN, +AMO) Sth				D (LN, –PDO) H (EN, +PDO)
[77]		D (+AMO)							
[16]		H (+AMO, –PDO)	H (EN) Nw	D (EN)		D (EN) H (LN)	D (+PDO) H (EN, –PDO)	D (–AMO, +PDO)	D (–PDO) H (EN, +PDO)
[78]			D (LN, +AMO, –PDO) H (EN, +PDO) W-S, Nw						D (LN) Os
[79]					D (LN) H (EN) Ctr				
[80]		D (LN, –PDO, –SST in EEP) Nw, AA							
[81]		D (LN, –PDO)	H (EN)						
[82]					D (EN)				
[83]	D (EN) H (LN) Ne	D (+AMO) H (–AMO) Ne, AA							
[84]		H (EN)			H (EN)				
[85]		D (+AMO) H (–AMO)	D (LN) H (EN)			H (LN)			

AA, annual accumulated; AMO, Atlantic Multidecadal Oscillation; ASO, August September–October; Ctr, center; D, dry periods; dns, does not specify; EEP, equatorial and eastern Pacific; EN, El Niño; EP, Eastern Pacific; ETP, Eastern tropical Pacific; H, humid periods; ID, intense drought; LAWP, Large Atlantic Warm Pool; LN, La Niña; NA, North Atlantic; Ne, northeast; NP, North Pacific; NS, not significant; Nw, northwest; Os, only some; PDO, Pacific Decadal Oscillation; S, spring; SAWP, Small Atlantic Warm Pool; SOI, Southern Oscillation Index; SST, sea surface temperature; Sth, south; TNA, Tropical North Atlantic; TSth, tropical south; W-S, winter-spring.

Table 2.
Summarized precipitation anomalies driven by climatic indices.

Region	North						Center and south						Most of the territory					
	Summer		dns		Winter		Summer		dns		Winter		Summer		dns		Winter	
	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN
SST pattern	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN
Condition	Humid	Dry	Dry	Humid	Humid	Dry	Humid	Dry	Humid	Dry	Dry	Humid	Dry	Humid	Dry	Humid	Dry	Humid
# Articles	5	2	3	1	21	11	4	0	3	1	4	3	8	2	5	0	2	1
% Match	71%	67%	27%	8%	100%	100%	24%	0%	14%	11%	67%	60%	89%	100%	100%	0%	29%	33%
Condition	Dry	Humid	Humid	Dry	Dry	Humid	Dry	Humid	Dry	Humid	Humid	Dry	Humid	Dry	Humid	Dry	Humid	Dry
# Articles	2	1	8	12	0	0	13	4	19	8	2	2	1	0	0	3	5	2
% Match	29%	33%	73%	92%	0%	0%	76%	100%	86%	89%	33%	40%	11%	0%	0%	100%	71%	67%
Total	7	3	11	13	21	11	17	4	22	9	6	5	9	2	5	3	7	3
SST pattern	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−	AMO+	AMO−
Condition	Dry	Humid	Dry	Humid	Dry	Humid	Humid	Dry	Humid	Dry	Humid	Dry	Humid	Dry	Dry	Humid	Humid	Dry
# Articles	3	1	9	3	1	0	2	2	3	7	0	0	0	0	1	0	0	0
% Match	100%	100%	90%	75%	100%	0%	100%	100%	100%	100%	0%	0%	0%	0%	100%	0%	0%	0%
Condition	Humid	Dry	Humid	Dry	Humid	Dry	Dry	Humid	Dry	Humid	Dry	Humid	Dry	Humid	Humid	Dry	Dry	Humid
# Articles	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
% Match	0%	0%	10%	25%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%
Total	3	1	10	4	1	0	2	2	3	7	0	0	0	0	1	1	0	0
SST pattern	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−	PDO+	PDO−
Condition	Humid	Dry	Dry	Humid	Humid	Dry	Dry	Humid	Dry	Humid	Humid	Dry	Dry	Humid	Dry	Humid	Humid	Dry
# Articles	0	1	2	1	2	3	2	2	4	0	1	1	1	1	1	0	3	2
% Match	0%	100%	33%	10%	100%	100%	100%	100%	100%	0%	100%	100%	100%	100%	100%	0%	100%	100%
Condition	Dry	Humid	Humid	Dry	Dry	Humid	Humid	Dry	Humid	Dry	Dry	Humid	Humid	Dry	Humid	Dry	Dry	Humid
# Articles	0	0	4	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
% Match	0%	0%	67%	90%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Total	0	1	6	10	2	3	2	2	4	0	1	1	1	1	1	0	3	2

Table 3.
Percentage of matching conclusions.

Precipitation anomalies in Mexico forced by ENSO					Forced by AMO			Forced by PDO		
Region	Season	ENSO phase	Condition	References	AMO phase	Condition	References	PDO phase	Condition	References
North	Summer	El Niño	Humid	[30, 38, 52, 55, 60]	AMO+	Humid	—	PDO+	<u>Humid</u>	—
			Dry	[24, 83]		Dry	[9, 27, 75]		<u>Dry</u>	—
		La Niña	Humid	[83]	AMO—	<u>Humid</u>	[76]	PDO—	<u>Humid</u>	—
			Dry	[38, 75]		<u>Dry</u>	—		<u>Dry</u>	[75]
		dns	Humid	[23, 30, 35, 41, 45, 71, 76, 84]	AMO+	Humid	[16]	PDO+	Humid	[3, 23, 45, 76]
			Dry	[23, 24, 49]		Dry	[3, 27, 55, 56, 63, 76, 77, 83, 85]		Dry	[24, 49]
	Winter	El Niño	Humid	[1, 3, 16, 24, 28, 29, 33, 36, 37, 44, 46, 48, 50, 51, 62, 64, 65, 74, 78, 81, 85]	AMO+	<u>Humid</u>	—	PDO+	<u>Humid</u>	[7, 78]
			Dry	—		<u>Dry</u>	[78]		<u>Dry</u>	—
		La Niña	Humid	—	AMO—	<u>Humid</u>	—	PDO—	Humid	—
			Dry	[14, 24, 37, 43, 44, 62, 64, 65, 74, 78, 85]		<u>Dry</u>	—		Dry	[7, 73, 78]
		El Niño	Humid	[30, 31, 48, 52]	AMO+	<u>Humid</u>	[5, 53]	PDO+	<u>Humid</u>	—
			Dry	[5, 10, 16, 24, 25, 39, 42, 50, 54–56, 60, 68]		<u>Dry</u>	—		<u>Dry</u>	[5, 7]
Center and south	Summer	El Niño	Humid	[30, 31, 48, 52]	AMO+	<u>Humid</u>	[5, 53]	PDO+	<u>Humid</u>	—
			Dry	[5, 10, 16, 24, 25, 39, 42, 50, 54–56, 60, 68]		<u>Dry</u>	—		<u>Dry</u>	[5, 7]
		La Niña	Humid	[5, 50, 56, 68]	AMO—	<u>Humid</u>	—	PDO—	<u>Humid</u>	[5, 7]
			Dry	—		<u>Dry</u>	[5, 53]		<u>Dry</u>	—

Precipitation anomalies in Mexico forced by ENSO					Forced by AMO			Forced by PDO		
Region	Season	ENSO phase	Condition	References	AMO phase	Condition	References	PDO phase	Condition	References
	dns	El Niño	Humid	[48, 79, 84]	AMO+	Humid	[25, 55, 76]	PDO+	Humid	—
			Dry	[10, 20, 24, 25, 35, 47, 49, 54, 55, 58, 59, 66–70, 72, 76, 82]		Dry	—		Dry	[5, 24, 49, 56]
		La Niña	Humid	[25, 26, 66–70, 76]	AMO–	Humid	—	PDO–	<u>Humid</u>	—
			Dry	[79]		Dry	[5, 24, 25, 49, 55, 56, 76]		<u>Dry</u>	—
	Winter	El Niño	Humid	[15, 68]	AMO+	<u>Humid</u>	—	PDO+	<u>Humid</u>	[7]
			Dry	[16, 42, 44, 48]		<u>Dry</u>	—		<u>Dry</u>	—
		La Niña	Humid	[16, 44, 85]	AMO–	<u>Humid</u>	—	PDO–	<u>Humid</u>	—
			Dry	[15, 68]		<u>Dry</u>	—		<u>Dry</u>	[7]
	Most of the territory	Summer	El Niño	Humid	AMO+	<u>Humid</u>	—	PDO+	<u>Humid</u>	—
			Dry	[6, 26, 42, 44, 48, 52, 56, 61]		<u>Dry</u>	—		<u>Dry</u>	[16]
			La Niña	<u>Humid</u>		<u>Humid</u>	—		<u>Humid</u>	[16]
			<u>Dry</u>	—		<u>Dry</u>	—		<u>Dry</u>	—
		dns	El Niño	Humid	AMO+	<u>Humid</u>	—	PDO+	<u>Humid</u>	—
			Dry	[20, 32, 34, 40, 57]		<u>Dry</u>	[57]		<u>Dry</u>	[16]
			La Niña	Humid		<u>Humid</u>	—		<u>Humid</u>	—
			Dry	[27, 57, 78]		<u>Dry</u>	[16]		<u>Dry</u>	—
	Winter	El Niño	Humid	[6, 16, 55, 57, 76]	AMO+	<u>Humid</u>	—	PDO+	Humid	[7, 16, 76]
			Dry	[42, 44]		<u>Dry</u>	—		Dry	—
		La Niña	Humid	[32]	AMO–	<u>Humid</u>	—	PDO–	<u>Humid</u>	—
			Dry	[57, 76]		<u>Dry</u>	—		<u>Dry</u>	[16, 76]

Table 4.
 General conclusions reached by most authors.

Cells underlined in italics represent the areas that have received the less attention and therefore, constitute interesting opportunities for future research.

4. Discussion

Other relevant but less repeated conclusions that were found but are not shown in **Table 4** are listed below:

1. The precipitation-El Niño teleconnection in Mexico is more significant when the climatic phenomenon is classified as intense, strong, or moderate [47, 48, 69].
2. The precipitation-AMO teleconnection has the greatest influence in northern Mexico [60, 76, 77], especially in the monsoon region [9].
3. The precipitation-ENSO teleconnection has the greatest (least) influence in the center and south (north) of Mexico [10, 23, 47, 49, 59, 61].
4. Even in areas where the effect of climatic indices is more noticeable, only 30–40% of variation in precipitation can be attributed to these modes [44, 49, 50].

Furthermore, the effects of the co-occurrence of AMO, ENSO, and PDO may be different from those when presented individually. For example, it has been noticed that El Niño events occur more frequently during positive phases of PDO [47, 49] and the most important climate impacts in the country could occur when ENSO and PDO are in phase [7, 16, 77, 78]. Besides, changes in the strength and spatial scale of the precipitation-ENSO teleconnection could be related to SST variability in the North Atlantic [74]. Some authors suggest that the strongest impacts arise when the North Atlantic and the tropical Pacific exhibit opposite anomalies [27, 77]. Additionally, Stahle et al. [60] observed that the response of the June PDSI to the combined ENSO-AMO index is stronger and more widespread in northern Mexico than for the ENSO or AMO indices alone. Hence, contemplating the synergistic effects of ocean-atmospheric circulations opens the possibility of making better predictions based on the effect of these circulation systems on terrestrial precipitation [50].

Among possible limitations that the validity of this study could concern, it is important to mention that the selection of only one database could restrict the articles screened, for this reason, the largest digital library of peer-reviewed literature was selected to conduct the review. In addition, to enlarge the examination, manual search was conducted to include some of the most relevant documents that were not encompassed in the Elsevier's Scopus search. Also, when conducting a systematic review, there is always a risk that there would be some papers unavailable in digital databases or documents that miss important keywords in their metadata [86].

Lastly, further research might focus on the teleconnections that have the most divergent conclusions, for instance, those in which $\frac{3}{4}$ or less authors match conclusions. Examples include summer ENSO conditions in the north, for both phases (El Niño and La Niña), and El Niño conditions in the north during the whole year. Both phases of ENSO might also be studied for the winter season in the center, south, and most of the territory. In addition, PDO+ and AMO– conditions in the north could be suggested for further research given the existence of opposite conclusions. Among these studies with divergent findings, the uneven extension of the study area was

found as a commonality. Attention must be paid to the spatial extent considered because interpretation of anomalies in precipitation driven by climatic indices can be interfered by small-scale processes such as localized convection.

5. Conclusions

This document presents a summary of research carried out in the period 1984–2020, concerning the individual and coupled effects of climatic oscillations ENSO, AMO, and PDO, over dry periods in México. After reviewing 75 articles, the main conclusions reached are summarized below:

- Positive AMO phases, together with negative PDO periods and La Niña conditions, generally force dry periods in northern Mexico. Conversely, negative AMO phases, together with positive PDO periods and El Niño conditions, generally force humid periods in northern Mexico.
- In central and southern Mexico, dry (humid) conditions have been associated with negative (positive) AMO phases, together with positive (limited information) PDO phases and El Niño (La Niña) conditions.
- For most of the territory, both El Niño and La Niña can drive negative precipitation anomalies. In general, La Niña (El Niño) induces dry (humid) periods during the winter season, while El Niño is related to dry conditions during summer. Additionally, positive PDO phases are related to greater than average humidity during winter.
- El Niño events occur more frequently during positive PDO phases and when both phenomena are in phase, their effects are intensified.
- Strength and spatial scale of the precipitation-ENSO teleconnection could be related to North Atlantic SST.

Given the potential predictability of these climatic indices, studies of this type open the possibility to roughly predict future droughts. Knowing the effects of climatic indices in advance, will aid in preventing and reducing adverse conditions in productive activities and society. These improved seasonal climate predictions play an essential role in climate change adaptation and water resources management.

Acknowledgements

Support provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) with a PNPC grant is appreciated. Additional support was received from the Universidad de las Américas Puebla (UDLAP) through a graduate fellowship.

Conflict of interest

The authors declare no conflict of interest.


IntechOpen

Author details

Regina Mijares-Fajardo* and Polioptro F. Martínez-Austria
Department of Environmental and Civil Engineering, University of the Americas
Puebla, Puebla, Mexico

*Address all correspondence to: regina.mijaresfo@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Nicholas RE, Battisti DS. Drought recurrence and seasonal rainfall prediction in the Río Yaqui Basin, Mexico. *Journal of Applied Meteorology and Climatology*. 2008; **47**:991-1005. DOI: 10.1175/2007JAMC1575.1
- [2] WHO. World Meteorological Organization International Meteorological Vocabulary. 2nd ed. Geneva: World Meteorological Organization (WMO); 1992
- [3] Mo KC, Schemm JE. Droughts and persistent wet spells over the United States and Mexico. *Journal of Climate*. 2008; **21**:980-994. DOI: 10.1175/2007JCLI1616.1
- [4] Metcalfe SE, Barron JA, Davies SJ. The Holocene history of the North American Monsoon: 'Known knowns' and 'known unknowns' in understanding its spatial and temporal complexity. *Quaternary Science Reviews*. 2015; **120**:1-27. DOI: 10.1016/j.quascirev.2015.04.004
- [5] Park J, Byrne R, Böhnelt H. The combined influence of Pacific decadal oscillation and Atlantic multidecadal oscillation on central Mexico since the early 1600s. *Earth and Planetary Science Letters*. 2017; **464**:1-9. DOI: 10.1016/j.epsl.2017.02.013
- [6] del Jiménez M. Indicadores Climáticos: Una manera para identificar la variabilidad Climática a escala global. In: Anexo del Informe Técnico: Elaboración de un boletín con información hidroclimática de los mares de México. 2015
- [7] Mendez Gonzalez J, Ramirez Leyva A, Cornejo Oviedo E, Zárate Lupercio A, Cavazos Perez T. Teleconexiones de la Oscilación Decadal del Pacífico (PDO) a la precipitación y temperatura en México. *Investigation Geography*. 2010; **73**:57-70
- [8] Arndt DS, Blunden J. State of the climate in 2011. *Bulletin of the American Meteorological Society*. 2012; **2012**:93. DOI: 10.1175/2012BAMSStateoftheClimate.1
- [9] Curtis S. The Atlantic multidecadal oscillation and extreme daily precipitation over the US and Mexico during the hurricane season. *Climate Dynamics*. 2008; **30**:343-351. DOI: 10.1007/s00382-007-0295-0
- [10] Rojas O, Li Y, Cumani R. Understanding the Drought Impact of El Niño on the Global Agricultural Areas: An Assessment Using FAO's Agricultural Stress Index (ASI). Food and Agriculture Organization of the United Nations; 2014. ISBN: 9789251086711
- [11] Van Loon AF. Hydrological drought explained. *Wiley Interdisciplinary Reviews Water*. 2015; **2**:359-392. DOI: 10.1002/wat2.1085
- [12] Stahle DW, Cleaveland MK. Southern oscillation extremes reconstructed from tree-rings of the Sierra Madre Occidental and Southern Great Plains. *Journal of Climate*. 1993; **6**: 129-140. DOI: 10.1175/1520-0442(1993)006<0129:SOERFT>2.0.CO;2
- [13] Trenberth KE, Hoar TJ. El Niño and climate change. *Geophysical Research Letters*. 1997; **24**:3057-3060. DOI: 10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2
- [14] Díaz SC, Therrell MD, Stahle DW, Cleaveland MK. Chihuahua (Mexico)

winter-spring precipitation reconstructed from tree-rings, 1647–1992. *Climate Research*. 2002;**22**: 237-244. DOI: 10.3354/cr022237

[15] Cerano-Paredes J, Méndez-González J, Amaro-Sánchez A, Villanueva-Díaz J, Cervantes-Martínez R, Rubio-Camacho EA. Reconstrucción de precipitación invierno-primavera con anillos anuales de pinus douglasiana en la reserva de la biosfera sierra de manantlán, Jalisco. *Revista Chapingo Serie Ciencias Forestales y del Ambiente*. 2013;**19**: 413-423. DOI: 10.5154/r.rchscfa.2013.02.007

[16] Durán-Quesada AM, Sorí R, Ordoñez P, Gimeno L. Climate perspectives in the Intra-Americas seas. *Atmosphere*. 2020;**11**:1-32. DOI: 10.3390/ATMOS11090959

[17] Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, et al., editors. *IPCC Climate Change 2013: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2013

[18] World Meteorological Organization. Experts agree on a universal drought index to cope with climate risk. WMO Press Release No. 872: Geneva. 2009

[19] Ault TR. On the essentials of drought in a changing climate. *Science* 80. 2020; **368**:256-260. DOI: 10.1126/SCIENCE.ABC4034

[20] Stevenson S, Overpeck JT, Fasullo J, Coats S, Parsons L, Otto-Bliesner B, et al. Climate variability, volcanic forcing, and last millennium hydroclimate extremes. *Journal of Climate*. 2018;**31**:4309-4327. DOI: 10.1175/JCLI-D-17-0407.1

[21] Pati D, Lorusso LN. How to write a systematic review of the literature.

HERD: Health Environments Research & Design Journal. 2018;**11**:15-30. DOI: 10.1177/1937586717747384

[22] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;**2021**:372. DOI: 10.1136/bmj.n71

[23] Llanes-Cárdenas O, Gaxiola-Hernández A, Estrella-Gastelum RD, Norzagaray-Campos M, Troyo-Diéguez E, Pérez-González E, et al. Variability and factors of influence of extremewet and dry events in Northern Mexico. *Atmosphere*. 2018;**9**:1-16. DOI: 10.3390/atmos9040122

[24] Quiroz-Jiménez JD, Roy PD, Beramendi-Orosco LE, Lozano-García S, Vázquez-Selem L. Orbital-scale droughts in central-northern Mexico during the late quaternary and comparison with other subtropical and tropical records. *Geological Journal*. 2018;**53**:230-242. DOI: 10.1002/gj.2888

[25] Hidalgo HG, Alfaro EJ, Amador JA, Bastidas Á. Precursors of quasi-decadal dry-spells in the Central America Dry Corridor. *Climate Dynamics*. 2019;**53**: 1307-1322. DOI: 10.1007/s00382-019-04638-y

[26] Rodríguez-Vera G, Romero-Centeno R, Castro CL, Castro VM. Coupled interannual variability of wind and sea surface temperature in the Caribbean Sea and the Gulf of Mexico. *Journal of Climate*. 2019;**32**:4263-4280. DOI: 10.1175/JCLI-D-18-0573.1

[27] Ruprich-Robert Y, Delworth T, Msadek R, Castruccio F, Yeager S, Danabasoglu G. Impacts of the Atlantic multidecadal variability on North American summer climate and heat waves. *Journal of Climate*. 2018;**31**:

3679-3700. DOI: 10.1175/JCLI-D-17-0270.1

[28] Kousky VE, Kagano MT, Cavalcanti IFA. A review of the Southern Oscillation: Oceanic-atmospheric circulation changes and related rainfall anomalies. *Tellus*. 1984;**35**:490-504. DOI: 10.1111/j.1600-0870.1984.tb00264.x

[29] Cleaveland MK, Cook ER, Stahle DW. Secular variability of the Southern Oscillation detected in tree-ring data from Mexico and the southern United States. In: Henry F, Vera M, editors. *El Niño: Historical and Paleoclimate Aspects of the Southern Oscillation*. Cambridge: Cambridge University Press; 1992. pp. 271-291

[30] Diaz HF, Markgraf V. *Historical and Paleoclimatic Aspects of the Southern Oscillation*. Cambridge: Cambridge University Press; 1992

[31] Pereyra Diaz D, Angulo Cordova Q, Palma Grayeb BE. Effect of ENSO on the mid-summer drought in Veracruz State, Mexico. *Atmosfera*. 1994;**7**:211-219

[32] Bell GD, Halpert MS, Ropelewski CF, Kousky VE, Douglas AV, Schnell RC, et al. Climate Assessment for 1998. *Bulletin of the American Meteorological Society*. 1999;**80**. DOI: 10.1175/1520-0477-80.5.s1

[33] Kane RP. Some characteristics and precipitation effects of the El Niño of 1997–1998. *Journal of Atmospheric and Solar - Terrestrial Physics*. 1999;**61**: 1325-1346. DOI: 10.1016/S1364-6826(99)00087-5

[34] Adem J, Mendoza VM, Ruiz A, Villanueva EE, Garduño R. Recent numerical experiments on three-months extended and seasonal weather prediction with a thermodynamic model. *Atmosfera*. 2000;**13**:53-83

[35] Filonov A, Tereshchenko I. Niño 1997–98 monitoring in mixed layer at the Pacific Ocean near Mexico's West Coast. *Geophysical Research Letters*. 2000;**27**: 705-707. DOI: 10.1029/1999GL002347

[36] Magaña V, Conde C. Climate and freshwater resources in northern Mexico: Sonora, a case study. *Environmental Monitoring and Assessment*. 2000;**61**:167-185. DOI: 10.1023/A:1006399025537

[37] Minnich RA, Vizcaino EF, Dezzani RJ. The El Niño/Southern Oscillation and precipitation variability in Baja California, Mexico. *Atmosfera*. 2000;**13**:1-20

[38] Díaz SC, Touchan R, Swetnam TW. A Tree-Ring reconstruction of past precipitation for Baja California Sur, Mexico. *International Journal of Climatology*. 2001;**21**:1007-1019. DOI: 10.1002/joc.664

[39] Curtis SS. Interannual variability of the bimodal distribution of summertime rainfall over Central America and tropical storm activity in the far-eastern pacific. *Climate Research*. 2002;**22**: 141-146. DOI: 10.3354/cr022141

[40] Hunt BG, Elliott TI. Mexican megadrought. *Climate Dynamics*. 2002; **20**:1-12. DOI: 10.1007/s00382-002-0265-5

[41] Salinas-Zavala CA, Douglas AV, Diaz HF. Interannual variability of NDVI in northwest Mexico. Associated climatic mechanisms and ecological implications. *Remote Sensing of Environment*. 2002; **82**:417-430. DOI: 10.1016/S0034-4257(02)00057-3

[42] Adams RM, Houston LL, McCarl BA, Tiscareño ML, Matus JG, Weiher RF. The benefits to Mexican agriculture of an El Niño-southern oscillation (ENSO)

early warning system. *Agricultural and Forest Meteorology*. 2003;**115**:183-194. DOI: 10.1016/S0168-1923(02)00201-0

[43] Cleaveland MK, Stahle DW, Therrell MD, Villanueva-Diaz J, Burns BT. Tree-ring reconstructed winter precipitation and tropical teleconnections in Durango, Mexico. *Climate Change*. 2003;**59**:369-388. DOI: 10.1023/A:1024835630188

[44] Magana VO, Vázquez JL, Pérez JL, Pérez JB. Impact of El Niño on precipitation in Mexico. *Geofisica International*. 2003;**42**:313-330

[45] Reyes-Coca S, Troncoso-Gaytán R. Modulación multidecenal de la lluvia invernal en el noroeste de Baja California. *Ciencias Marinas*. 2004;**30**: 99-108. DOI: 10.7773/cm.v30i11.117

[46] González-Elizondo M, Jurado E, Nívar J, González-Elizondo MS, Villanueva J, Aguirre O, et al. Tree-rings and climate relationships for Douglas-fir chronologies from the Sierra Madre Occidental, Mexico: A 1681–2001 rain reconstruction. *Forest Ecology and Management*. 2005;**213**:39-53. DOI: 10.1016/j.foreco.2005.03.012

[47] Mendoza B, Jáuregui E, Diaz-Sandoval R, García-Acosta V, Velasco V, Cordero G. Historical droughts in central Mexico and their relation with El Niño. *Journal of Applied Meteorology*. 2005;**44**:709-716. DOI: 10.1175/JAM2210.1

[48] Bravo JL, Gay C, Conde C, Estrada F. Probabilistic description of rains and ENSO phenomenon in a coffee farm area in Veracruz, México. *Atmosfera*. 2006;**19**:49-74

[49] Mendoza B, Velasco V, Jáuregui E. A study of historical droughts in southeastern Mexico. *Journal of Climate*.

2006;**19**:2916-2934. DOI: 10.1175/JCLI3726.1

[50] Caso M, González-Abraham C, Ezcurra E. Divergent ecological effects of oceanographic anomalies on terrestrial ecosystems of the Mexican Pacific coast. *Proceedings of the National Academy of Sciences of the United States of America*. 2007;**104**:10530-10535. DOI: 10.1073/pnas.0701862104

[51] Villanueva-Diaz J, Stahle DW, Luckman BH, Cerano-Paredes J, Therrell MD, Cleaveland MK, et al. Winter-spring precipitation reconstructions from tree rings for northeast Mexico. *Climatic Change*. 2007;**83**:117-131. DOI: 10.1007/s10584-006-9144-0

[52] Peralta-Hernández AR, Barba-Martínez LR, Magaña-Rueda VO, Matthias AD, Luna-Ruiz JJ. Temporal and spatial behavior of temperature and precipitation during the canícula (midsummer drought) under El Niño conditions in central México. *Atmosfera*. 2008;**21**:265-280

[53] Wang C, Lee SK, Enfield DB. Climate response to anomalously large and small Atlantic warm pools during the summer. *Journal of Climate*. 2008;**21**:2437-2450. DOI: 10.1175/2007JCLI2029.1

[54] Kienel U, Bowen SW, Byrne R, Park J, Böhnelt H, Dulski P, et al. First lacustrine varve chronologies from Mexico: Impact of droughts, ENSO and human activity since AD 1840 as recorded in maar sediments from Valle de Santiago. *Journal of Paleolimnology*. 2009;**42**:587-609. DOI: 10.1007/s10933-009-9307-x

[55] Seager R, Ting M, Davis M, Cane M, Naik N, Nakamura J, et al. Mexican drought: An observational modeling and

tree ring study of variability and climate change. *Atmosfera*. 2009;**22**:1-31

[56] Méndez M, Magaña V. Regional aspects of prolonged meteorological droughts over Mexico and central America. *Journal of Climate*. 2010;**23**: 1175-1188. DOI: 10.1175/2009JCLI3080.1

[57] Seager R, Vecchi GA. Greenhouse warming and the 21st century hydroclimate of southwestern North America. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;**107**:21277-21282. DOI: 10.1073/pnas.0910856107

[58] Sosa Nájera S, Lozano García S, Roy PD, Caballero M. Registro de sequías históricas en el occidente de México con base en el análisis elemental de sedimentos lacustres: El caso del lago de Santa María del Oro. *Boletín de la Sociedad Geológica Mexicana*. 2010;**62**: 437-451

[59] Lachniet MS, Bernal JP, Asmerom Y, Polyak V, Piperno D. A 2400 yr Mesoamerican rainfall reconstruction links climate and cultural change. *Geology*. 2012;**40**:259-262. DOI: 10.1130/G32471.1

[60] Stahle DW, Burnette DJ, Diaz JV, Heim RR, Fye FK, Paredes JC, et al. Pacific and Atlantic influences on Mesoamerican climate over the past millennium. *Climate Dynamics*. 2012;**39**: 1431-1446. DOI: 10.1007/s00382-011-1205-z

[61] Bhattacharya T, Chiang JCH. Spatial variability and mechanisms underlying El Niño-induced droughts in Mexico. *Climate Dynamics*. 2014;**43**:3309-3326. DOI: 10.1007/s00382-014-2106-8

[62] Pompa-García M, Jurado E. Seasonal precipitation reconstruction and teleconnections with ENSO based on tree

ring analysis of *Pinus cooperi*. *Theoretical and Applied Climatology*. 2014;**117**: 495-500. DOI: 10.1007/s00704-013-1018-6

[63] Návar J. Hydro-climatic variability and perturbations in Mexico's north-western temperate forests. *Ecohydrology*. 2015;**8**:1065-1072. DOI: 10.1002/eco.1564

[64] Parazoo NC, Barnes E, Worden J, Harper AB, Bowman KB, Frankenberg C, et al. Influence of ENSO and the NAO on terrestrial carbon uptake in the Texas-northern Mexico region. *Global Biogeochemical Cycles*. 2015;**29**:1247-1265. DOI: 10.1002/2015GB005125

[65] Chávez-Gándara MP, Cerano-Paredes J, Nájera-Luna JA, Pereda-Breceda V, Esquivel-Arriaga G, Cervantes-Martínez R, et al. Reconstrucción de la precipitación invierno-primavera con base en anillos de crecimiento de árboles para la región de San Dimas, Durango, México. *Bosque*. 2017;**38**:387-399. DOI: 10.4067/S0717-92002017000200016

[66] Anderson W, Seager R, Baethgen W, Cane M. Trans-Pacific ENSO teleconnections pose a correlated risk to agriculture. *Agricultural and Forest Meteorology*. 2018;**262**:298-309. DOI: 10.1016/j.agrformet.2018.07.023

[67] Díaz-Esteban Y, Raga GB. Weather regimes associated with summer rainfall variability over southern Mexico. *International Journal of Climatology*. 2018;**38**:169-186. DOI: 10.1002/joc.5168

[68] Maass M, Ahedo-Hernández R, Araiza S, Verduzco A, Martínez-Yrizar A, Jaramillo VJ, et al. Long-term (33 years) rainfall and runoff dynamics in a tropical dry forest ecosystem in western Mexico: Management

implications under extreme hydrometeorological events. *Forest Ecology and Management*. 2018; **426**:7-17. DOI: 10.1016/j.foreco.2017.09.040

[69] Martínez-Austria PF, Díaz-Jiménez D. Tendencias de la precipitación y su relación con el Índice Oceánico El Niño. El caso de la Región Mixteca, México. *Ingeniería del Agua*. 2018;**22**:1-14. DOI: 10.4995/ia.2018.7779

[70] Arellano-Monterrosas JL, Ruiz-Meza LE. Climate variability and extreme events in the Zanatenco river basin, Chiapas. *Tecnología y Ciencias del Agua*. 2019;**10**:249-274. DOI: 10.24850/j-tyca-2019-03-10

[71] Esquivel-Arriaga G, Cerano-Paredes J, Sánchez-Cohen I, Velásquez-Valle MA, Flores-López F, Bueno-Hurtado P. Temporal analysis of droughts (1922–2016) in the upper Nazas River Basin using SPI and its relationship with ENSO. *Tecnología y Ciencias del Agua*. 2019;**10**:126-151. DOI: 10.24850/j-tyca-2019-05-05

[72] Fichez R, Linares C, Chifflet S, Conan P, Esparza ACR, Denis L, et al. Spatiotemporal variability in Terminos Lagoon (Mexico) waters during the 2009–2010 drought reveals upcoming trophic status shift in response to climate change. *Regional Environmental Change*. 2019;**19**: 1787-1799. DOI: 10.1007/s10113-019-01519-2

[73] O'Mara NA, Cheung AH, Kelly CS, Sandwick S, Herbert TD, Russell JM, et al. Subtropical Pacific ocean temperature fluctuations in the common era: Multidecadal variability and its relationship with southwestern North American megadroughts. *Geophysical Research Letters*. 2019;**46**:14662-14673. DOI: 10.1029/2019GL084828

[74] Torbenson MCA, Stahle DW, Howard IM, Burnette DJ, Villanueva-Díaz J, Cook ER, et al. Multidecadal modulation of the ENSO teleconnection to precipitation and tree growth over subtropical North America. *Paleoceanography and Paleoclimatology*. 2019;**34**:886-900. DOI: 10.1029/2018PA003510

[75] Campos MN, Cárdenas OL, Gaxiola A, González GEG. Meteorological interaction between drought/oceanic indicators and rainfed maize yield in an arid agricultural zone in northwest Mexico. *Arabian Journal of Geosciences*. 2020;**13**

[76] Cavazos T, Luna-Niño R, Cerezo-Mota R, Fuentes-Franco R, Méndez M, Pineda Martínez LF, et al. Climatic trends and regional climate models intercomparison over the CORDEX-CAM (Central America, Caribbean, and Mexico) domain. *International Journal of Climatology*. 2020;**40**:1396-1420. DOI: 10.1002/joc.6276

[77] De Luca P, Messori G, Wilby RL, Mazzoleni M, Di Baldassarre G. Concurrent wet and dry hydrological extremes at the global scale. *Earth System Dynamics*. 2020;**11**:251-266. DOI: 10.5194/esd-11-251-2020

[78] Gutierrez-Garcia G, Leavitt SW, Trouet V, Carriquiry JD. Tree ring-based historic hydroclimatic variability. *Baja California Peninsula*. 2020;**125**

[79] Li Y, Hernandez JH, Aviles M, Knappett PSK, Giardino JR, Miranda R, et al. Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato, México. *Journal of Hydrology*. 2020;**582**:124517. DOI: 10.1016/j.jhydrol.2019.124517

- [80] Llanes-Cárdenas O, Norzagaray-Campos M, Pérez-González E, Gaxiola A, López-Rocha JS, González-González GE. Trend analysis and historical and recent return periods of erosivity indicators in the state of Sinaloa, Mexico. *Arabian Journal of Geosciences*. 2020;**13**:1-13. DOI: 10.1007/s12517-020-5153-y
- [81] Martínez-Sifuentes AR, Villanueva-Díaz J, Estrada-ávalos J. Runoff reconstruction and climatic influence with tree rings, in the mayo river basin, Sonora, Mexico. *iForest - Biogeosciences and Forestry*. 2020;**13**:98-106. DOI: 10.3832/ifor3190-013
- [82] Roy PD, García-Arriola OA, Garza-Tarazon S, Vargas-Martínez IG, Muthusankar G, Giron-García P, et al. Late Holocene depositional environments of Lake Coatetelco in Central-Southern Mexico and comparison with cultural transitions at Xochicalco. *Palaeogeography Palaeoclimatology Palaeoecology*. 2020; **560**:110050. DOI: 10.1016/j.palaeo.2020.110050
- [83] Roy PD, Vera-Vera G, Sánchez-Zavala JL, Shanahan TM, Quiroz-Jiménez JD, Curtis JH, et al. Depositional histories of vegetation and rainfall intensity in Sierra Madre Oriental Mountains (northeast Mexico) since the late Last Glacial. *Global and Planetary Change*. 2020;**187**:103136. DOI: 10.1016/j.gloplacha.2020.103136
- [84] Ruiz-alvarez O, Singh VP, Enciso-medina J, Ontiveros-Capurata RE, Corrales-Suastegui A. Spatio-temporal trends of monthly and annual precipitation in Aguascalientes, Mexico. *Atmosphere*. 2020;**11**:437. DOI: doi.org/10.3390/atmos11050437
- [85] Stahle DW, Cook ER, Burnette DJ, Torbenson MCA, Howard IM, Griffin D, et al. Dynamics, variability, and change in seasonal precipitation reconstructions for North America. *Journal of Climate*. 2020;**33**:3173-3195. DOI: 10.1175/JCLI-D-19-0270.1
- [86] Akdur D, Demirörs O, Bilgisi Öz M. Systematic reviews in model-driven engineering: A tertiary study model. *Journal of Aeronautical Special Technology*. 2020;**13**:57-68

Reviewing Composite Water and Ecosystem Indexes to Envision Indicators with an Ecohydrology Framework

Maria Pina and Carlos Patino-Gomez

Abstract

The main objective of this review is to identify and gather most common indicators used to conceive an ecohydrologic index. The research was made by means of a systematic review of different available literature, related to water and ecological indexes. The articles' eligibility was based on studies focused on basins as geographic scale and published in English or Spanish. Additionally, the indexes reviewed must have shown clearly the indicators utilized in their methodology. The selected bibliographic dataset for this study comes from the search strings: water sustainability index, water health index, water risk index, water vulnerability index, water quality index, water sustainability indicators, ecological index, and ecosystem index. The main outcomes resulted in a combination of hydrologic, ecologic, economic, institutional, and political indicators lists. The most repeated indicators on the lists are presented in this document and their ecohydrological relationship is analyzed; as well as the future directions for the realization of an ecohydrological index.

Keywords: ecohydrology, composite indexes, indicators, sustainable ecosystems, basin hydrology, water demands, watershed assessment

1. Introduction

For humanity, access to freshwater is a right and a necessity; however, the contamination and the increasing scarcity of this precious resource have alarmed the entire world. Currently, investigations executed around water knowledge are focusing on multiple subjects. One of them, the water cycle, is without objection a complex topic, and it has been studied for its undeniable connection with life. The water cycle is related to the atmosphere and the ground; the atmosphere is tied with the carbon cycle, the ground with the biota, and the system continues with multiple relationships that happen at the same time.

Moreover, added to the difficulty of modeling the water cycle and its relationship with biodiversity and climate; political, social, and economic issues need to be linked. Hydrological processes such as interception, evapotranspiration, and the effects of irrigation need a combined investigation of impacts such as the extraction

of groundwater, changes in land use [1]; as well as to vapor pressure deficits, temperature changes, and extreme events (floods and droughts) [2]. Nonetheless, the impacts mentioned earlier have their own functioning, which have been aggravated seriously in recent years with the overexploitation of ecosystem resources and the anthropogenic activities that have ushered to irregular climate change.

Attempting to fathom and handle the problems related to the water cycle and ecosystems, the science of ecohydrology allows researchers to address the principle that water is a critical factor for the biosphere's functioning, bio productivity, and biodiversity-ecosystem services for society [3]. Ecohydrology aspires to integrate the knowledge and understanding of both hydrological and ecological processes that co-occur at different scales, considering ecosystems as the primary tool to increase water sustainability in basins [3]. "A key challenge faced by science in the 21st century is not only understanding the causal relationships between abiotic factors and biotic factors in all types of ecosystems, but also linking them to the dynamics of society" ([4] p. 396). Sometimes humans are narrow-minded, this further complicates noticing these connections, therefore the reality is oversimplified [5]. It exists a great necessity to change that idea and start to think systematically.

Since ecohydrology studies variables of its interest as a multidimensional phenomenon, composite indexes methodology could be a proper technique to be applied in this field. A composite index is a measure that determines and expresses the degree or state of a variable [6]. The variable is measured by employing a single score obtained by combining other values, sometimes through a direct sum or by more complex manners. The indexes and their distinct aspects are a bridge between specialists' work and their scope being understood by the common people.

Environmental, water or ecological indexes embody several variables for their calculation, and they are called indicators. Quantitative or qualitative indicators can represent the variables since they categorize a large amount of data. Indicators have been used as an instrument to estimate changes and risks, but lately, they have also taken importance in water and resources management [7] for government policies. Indicators need constant upgrades to generate better quality and trustworthiness of their results. One path for indicators' improvement is creating a scheme of criteria for selecting and assessing them along with designing the mediums to do it [8]. In this work, the authors believe that indicators are a viable method of assessing the complexity and depth interrelation of the freshwater system with the ecological medium; that is the main reason to review the published literature related to the subject. From examining the available indexes works with a systematic review, it is possible to find the right ecohydrological indicators to be the baseline for a new composite index.

2. Material and methods

2.1 Data collection

The bibliographic dataset for this study was selected from two sources, the first one was the data information from Vollmer et al. [7] review paper, and the second was extracting a body of data (December 2021) from the Scopus (<https://www.scopus.com>) search engine.

The methodology applied in Vollmer et al. [7] was utilized for the systematic review of this document. The authors' investigation resulted in an assortment of 95 indexes in their article. However, for this analysis, only indexes with basin as the geographic

scale were selected, deriving the usage of 30 indexes of their research; unfortunately, information of one of them was not found. Since the review paper [7] was published six years ago, an update of the new indexes was necessary. Having this in mind, a search of the latest literature was made, using the Scopus portal, and following Vollmer et al.'s approach [7]. The procedure for the search was using the criteria shown below:

- I. Only consider indexes where water is the focus [7].
- II. Solely review indexes with documented and transparent methods [7].
- III. The literature review covers the years 2016 to 2021.
- IV. Selection of indexes exclusively with basin as geographic scale.
- V. Contemplate articles written in English and Spanish.
- VI. The authors must have created a new index, or at least have modified an existing one they also must show clearly the indicators used for their index.
- VII. The elective subject areas for the searching are environmental science, social sciences, agricultural and biological sciences, decision sciences, engineering, energy, earth and planetary sciences, mathematics, and multidisciplinary.

Given that the searching strings used by Vollmer et al. [7] do not cover completely the ecosystems and ecologic outlook, two more search terms were aggregated (**Figure 1**). The procedure for searching with the new terms was almost the same as the criteria described above, except for numbers I and III (years span: 1980–2021).

The update work altogether with the new subjects connected to ecology, it was made following the PRISMA methodology. The flow diagram (**Figure 2**) displays the identification, screening, and inclusion steps of this process.

2.2 Data analysis

The election of the 76 articles and gathered their respective indexes, resulted in a compilation of data of more than 1000 different indicators. Some of the indicators



Figure 1.
Search terms used in the Scopus search engine.

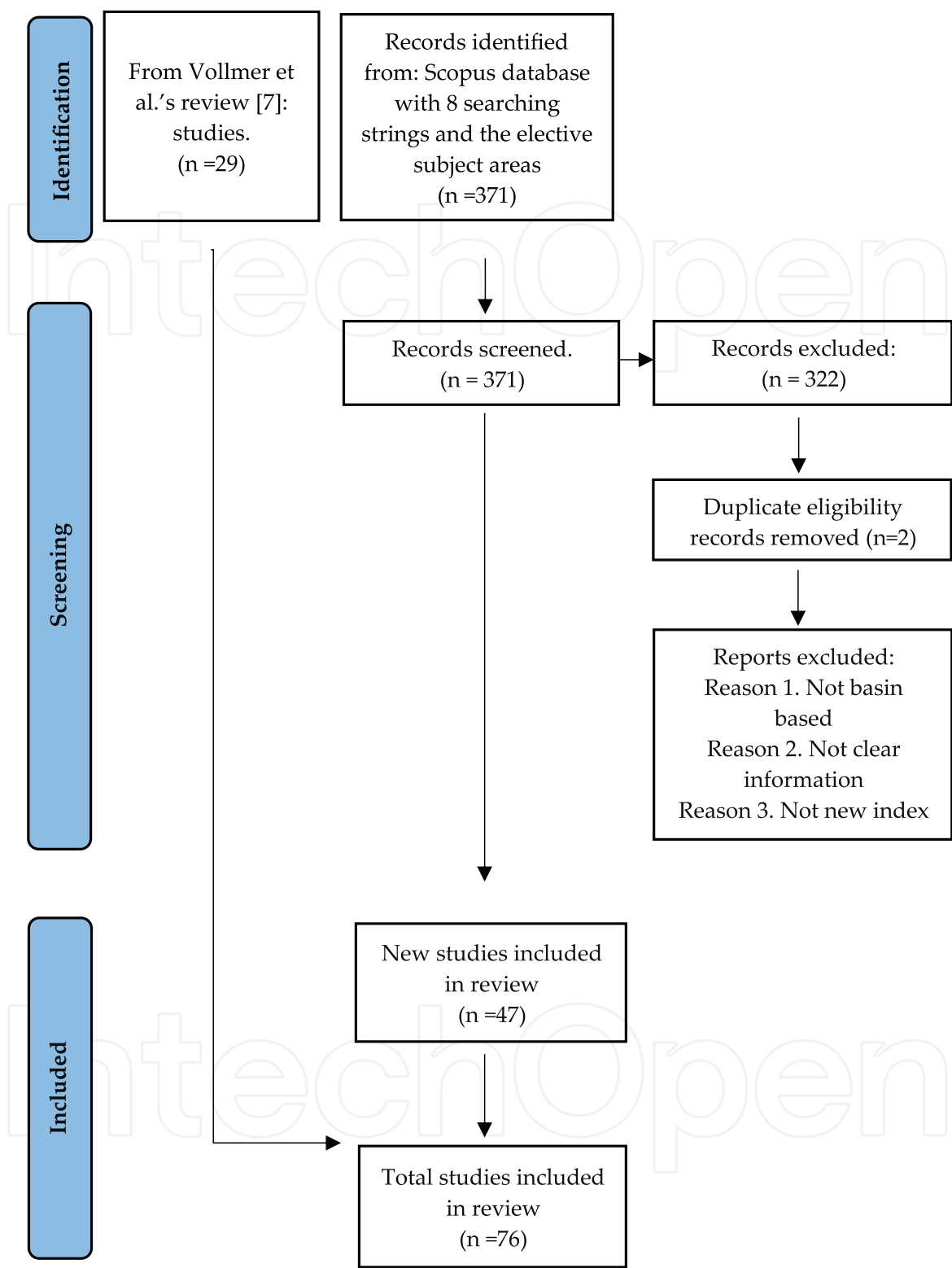


Figure 2. PRISMA 2020 flow diagram for updated systematic review, which included the new search of the database. Own elaboration with information from [7, 9].

share a similar context, for example, Pires et al. [10] measure the “burden of water-associated diseases (expressed in DALYs) with comparative risk assessment”; similarly, Cervantes-Jiménez et al. [11] assess the “average disability-adjusted life years (DALYs) due to intestinal disease”. Both evaluate problems related to population health due to poor water quality and lack of its treatment, which avoids “access

to safe water”, another indicator, presented by Perez-Foguet and Garriga [12], and Chidammodzi and Muhandiki [13].

Alternative to this, other authors present a short version of the indicators suggested by an older index, this is the case of Gallego-Ayala and Juárez [14] who took 24 indicators from Hopper [15]. Nonetheless, it also exists the circumstance in which more than one article mentions the same indicator for their index (e.g., channel alteration [16–18]).

An indicator can also be divided into other indicators, so while one author considers a set of aspects, another handles them individually. For instance, heavy metals as one indicator [10, 19–21]; or split it in indicator for chromium, aluminum [22, 23], arsenic [23], mercury [23], lead [23], zinc [23, 24], and iron [23, 25].

Regarding the categories used in the articles to classify the indicators, it is possible to see that some indicators can be put into different classes; to illustrate, population density is grouping in economic factors [26], pressure [27], water demand vulnerability [28], social [11], social development [29], and social and urban activity [30].

Also, it is important to recognize that indicators have neutral, negative, or positive dimensions. It means that some of them measure the state, the impact, or the alternatives to improve the situation. As a final remark in this section, it is to notice that some indicators employed to build an index are index by themselves, namely the environment pressure index (EPI) [22, 31, 32].

3. Results and discussion

3.1 Gathered data

From the indicators collected data, a summary of 39 indicators with at least four mentions among the 76 articles selected are illustrated in **Table 1**.

Indicator/parameter	No. of occurrences	Sources
Dissolved oxygen	14	[19–25, 27, 33–38]
Nitrite and nitrate nitrogen	11	[10, 22–25, 30, 33–35, 37, 38]
Ammonia nitrogen/ammonium	10	[10, 19, 22, 23, 25, 33–35, 37, 38]
pH	9	[19, 20, 23, 25, 34, 36–39]
Water/stream temperature	9	[22, 23, 28, 34, 36–40]
Total phosphorus	9	[18, 20, 22–24, 27, 28, 33, 35]
Population density	9	[10, 11, 26–30, 41, 42]
Water quality (surface and groundwater) throughout the catchment	8	[10, 40, 43–48]
Human development index related (HDI)	7	[11, 12, 29, 31, 32, 49, 50]
Biochemical oxygen demand	7	[10, 11, 23, 27, 32, 34, 35]
Total suspended solids	7	[18, 22–25, 34, 35]
Organic nitrogen/total nitrogen (in surface and groundwater)	7	[18, 20, 22, 23, 27, 28, 33]

Indicator/parameter	No. of occurrences	Sources
Environmental water demand/flow requirements	7	[10, 21, 51–55]
Use of soil/ catchment land use	7	[30, 35, 36, 40, 52, 56, 57]
(Integrate) Water resource management	7	[13, 15, 18, 32, 46, 56, 58]
Water per capita (basin availability)	6	[10, 12, 26, 31, 32, 59]
Groundwater exploitation potential/groundwater depletion	6	[10, 11, 18, 41, 42, 60]
Shannon-Wiener family diversity	5	[25, 30, 34, 37, 61]
Chloride (Cl-) (total or dissolved)	5	[22, 23, 25, 33, 37]
Sulfate/ sulphate (total or dissolved)	5	[20, 22, 23, 25, 33]
Biological contaminants (<i>E. coli</i> / Fecal coliforms/ Total coliforms)	5	[10, 22, 23, 25, 37]
Occurrence of diseases related to water resources	5	[12, 18, 56, 60, 62]
Proportion of urban land use per basin/ urbanization of catchment	5	[11, 28, 35, 41, 63]
Forest cover	5	[26, 27, 49, 62, 64]
Vegetation cover as percentage of basin area/land cover naturalness	5	[17, 18, 21, 32, 59]
Water-use efficiency	4	[32, 44, 45, 65]
Water stress	4	[50, 58, 66, 67]
Water quantity available	4	[43–45, 56]
Total water storage capacity	4	[10, 27, 56, 65]
Total annual water demand	4	[27, 41, 51, 52]
Flow variations/ alteration	4	[21, 40, 50, 68]
Chemical oxygen demand	4	[19, 23, 25, 34]
Total dissolved solids	4	[23, 25, 36, 51]
Salinity	4	[10, 20, 25, 51]
Heavy metals (surface and groundwater)	4	[10, 19–21]
Total riparian vegetation cover	4	[16, 17, 40, 69]
Management/impact of invasive species	4	[13, 18, 70, 71]
Protected natural area	4	[10, 12, 30, 62]
Education level	4	[11, 46, 62, 64]

Table 1.
Most mentioned indicators or parameters.

3.2 Result remarks related to ecohydrology framework

As it was alluded to in the introduction, water available for the ecological and anthropological demands are subjected substantially to basin characteristics, the climate, and land cover naturalness; together with what occurs in the basin as consequence, mostly by the human impact [4]. Taking this consideration, ecohydrology science introduces to its principles the five multi-dimensional parameters within river basins: water, biodiversity, ecosystem services for society, resilience to climatic changes, and cultural heritage and education (WBSR-CE) [4, 72].

From the displayed indicators examined in **Table 1**, there is no doubt that all of them can be part of an ecohydrological index. These indicators enter the concept of ecohydrology described above. However, the issue at hand is if the information to collect and measure the data is available; and if the indicators can be normalized in such a way that when correctly weighted, they are to be added to the composite index which will face calibration.

4. Conclusion and future directions

With respect to this study, the authors identified that more and more research related to composite indexes include parameters of culture, economics, politics, institutions, governance, and social issues. It also was notable that there is not a perfect number of indicators to compose an index, while some investigations use five indicators, others could use 115, all depending on the detailed framework wanted. On the other hand, the absence of data available, affordable, or accessible is summed to the reasons to avoid utilizing specific indicators; nevertheless, the use of indexes is still creating scientific publications.

Considering all the points exposed in this manuscript, the final purpose to elaborate this systematic review was envisioning the creation of an ecohydrological composite index for basins, which currently does not exist. The work pretends to be a framework to strengthen ecosystems' sustainability and to ensure water demands. An end-user tool that processes the indicators' values and calculations of the ecohydrological composite index is also in plans to develop.

Acknowledgements

The author would like to thank the Mexican National Council on Science and Technology (CONACyT) and the Universidad de las Americas Puebla for the scholarship granted.

Conflict of interest

The authors declare that they have no conflict of interest and all authors read and agree with the manuscript.

IntechOpen


IntechOpen

Author details

Maria Pina* and Carlos Patino-Gomez
Civil and Environmental Engineering Department, Universidad de las Americas
Puebla, Puebla, Mexico

*Address all correspondence to: maria.pinars@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Potter NJ, Chiew FHS. An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray–Darling Basin using rainfall- runoff models. *Water Resources Research*. 2011;**47**:1-12. DOI: 10.1029/2010WR010333
- [2] Sangines de Carcer P, Vitasse Y, Penuelas J, Jassey VE, Buttler A, Signarbieux C. Vapor-pressure deficit and extreme climatic variables limit tree growth. *Global Change Biology*. 2017;**24**(3):1108-1122. DOI: 10.1111/gcb.13973
- [3] Zalewski M. Ecohydrology in the Anthropocene - necessity to change paradigm from mechanistic to evolutionary/ecosystemic [Video]. 2020. Available from: <https://hidroinformatica.itaipu.gov.py/aulavirtual/mod/book/view.php?id=1056>
- [4] Zalewski M, Kiedrzyńska E, Wagner I, Izydorczyk K, Boczek JM, Jurczak T, et al. Ecohydrology and adaptation to global change. *Ecohydrology & Hydrobiology*. 2021;**21**(3):393-410. DOI: 10.1016/j.ecohyd.2021.08.001
- [5] Matija M. 10 pasos para alinear la cabeza y el corazón y salvar el planeta. Ciudad de Mexico: Editorial Planeta Mexicana; 2020
- [6] Mazziotta M, Pareto A. Methods for Constructing Composite Indices: One for All or All for One? *Rivista italiana di economia, demografia e statistica*. 2013;**LXVII**(2):67-80
- [7] Vollmer D, Regan HM, Andelman SJ. Assessing the sustainability of freshwater systems: A critical review of composite indicators. *Ambio*. 2016;**45**:765-780. DOI: 10.1007/s13280-016-0792-7
- [8] Pires A, Morato J, Peixoto H, Bradley S, Muller A. Synthesizing and standardizing criteria for the evaluation of sustainability indicators in the water sector. *Environmental Development Sustainable*. 2020;**22**:6671-6689. DOI: 10.1007/s10668-019-00508-z
- [9] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ*. 2021;**29**:71
- [10] Pires A, Morato J, Peixoto H, Botero V, Zuluaga L, Figueroa A. Sustainability Assessment of indicators for integrated water resources management. *Science of The Total Environment*. 2017;**578**:139-147. DOI: 10.1016/j.scitotenv.2016.10.217
- [11] Cervantes-Jiménez M, Díaz-Delgado C, González-Sosa E, Ángel Gómez-Albores M, Mastachi-Loza CA. Proposal of a water management sustainability index for the 969 sub-basins of Mexico. *Journal of Maps*. 2020;**16**(2):432-444. DOI: 10.1080/17445647.2020.1763486
- [12] Pérez-Foguet A, Giné GR. Analyzing water poverty in basins. *Water Resources Management*. 2011;**25**(14):3595-3612. DOI: 10.1007/s11269-011-9872-4
- [13] Chidammodzi CL, Muhandiki VS. Development of indicators for assessment of Lake Malawi Basin in an Integrated Lake Basin Management (ILBM) framework. *International Journal of the Commons*. 2015;**9**(1):209
- [14] Gallego-Ayala J, Juárez D. Performance evaluation of River Basin Organizations

to implement integrated water resources management using composite indexes. *Physics and Chemistry of the Earth, Parts A/B/C*. 2012;**50-52**:205-216

[15] Hooper B. River basin organization performance indicators: Application to the Delaware River basin commission. *Water Policy*. 2010;**12**(4):461-478. DOI: 10.2166/wp.2010.111

[16] Sirombra MG, Mesa LM. A method for assessing the ecological quality of riparian forests in subtropical Andean streams: QBRy index. *Ecological Indicators*. 2012;**20**:324-331. DOI: 10.1016/j.ecolind.2012.02.021

[17] Munné A, Prat N, Solà C, Bonada N, Rieradevall M. A simple field method for assessing the ecological quality of riparian habitat in rivers and streams: QBR index. *Aquatic Conservation: Marine and Freshwater Ecosystems*. 2002;**13**(2):147-163. DOI: 10.1002/aqc.529

[18] Vollmer D, Shaad K, Souter NJ, Farrell T, Dudgeon D, Sullivan CA, et al. Integrating the social, hydrological and ecological dimensions of freshwater health: The Freshwater Health Index. *Science of The Total Environment*. 2018;**627**:304-313. DOI: 10.1016/j.scitotenv.2018.01.040

[19] Sakai N, Mohamad ZF, Nasaruddin A, Abd Kadir SN, Mohd Salleh MSA, Sulaiman AH. Eco-Heart Index as a tool for community-based water quality monitoring and assessment. *Ecological Indicators*. 2018;**91**:38-46. DOI: 10.1016/j.ecolind.2018.03.079

[20] Flint N, Rolfe J, Jones CE, Sellens C, Johnston ND, Ukkola L. An Ecosystem Health Index for a large and variable river basin: Methodology, challenges and continuous improvement in Queensland's

Fitzroy Basin. *Ecological Indicators*. 2017;**73**:626-636. DOI: 10.1016/j.ecolind.2016.10.007

[21] Xu X. Comprehensive assessment of the water ecological security of the Xiangjiang river basin based on physico-chemistry and organism indices. *Applied Ecology and Environmental Research*. 2019;**17**(2):4547-4574

[22] Kerans BL, Karr JR. A Benthic Index of Biotic Integrity (B-IBI) for Rivers of the Tennessee Valley. *Ecological Applications*. 1994;**4**(4):768-785. DOI: 10.2307/1942007

[23] Casillas-García LF, de Anda J, Yebra-Montes C, Shear H, Díaz-Vázquez D, Gradilla-Hernández MS. Development of a specific water quality index for the protection of aquatic life of a highly polluted urban river. *Ecological Indicators*. 2021;**129**:107899. DOI: 10.1016/j.ecolind.2021.107899

[24] Smajgl A, Larson S, Hug B, De Freitas DM. Water use benefit index as a tool for community-based monitoring of water related trends in the Great Barrier Reef region. *Journal of Hydrology*. 2010;**395**(1-2):1-9. DOI: 10.1016/j.jhydrol.2010.09.007

[25] Tampo L, Lazar IM, Kaboré I, Oueda A, Akpataku KV, Djaneye-Boundjou G, et al. A multimetric index for assessment of aquatic ecosystem health based on macroinvertebrates for the Zio river basin in Togo. *Limnologica*. 2020;**83**:125783. DOI: 10.1016/j.limno.2020.125783

[26] Han M, Qingwang R, Wang Y, Du J, Hao Z, Sun F, et al. Integrated Approach to Water Allocation in River Basins. *Journal of Water Resources Planning and Management*. 2013;**139**(2):159-165. DOI: 10.1061/(asce)wr.1943-5452.0000255

- [27] Jun KS, Chung E-S, Sung J-Y, Lee KS. Development of spatial water resources vulnerability index considering climate change impacts. *Science of The Total Environment*. 2011;**409**(24):5228-5242. DOI: 10.1016/j.scitotenv.2011.08.027
- [28] Chang H, Jung I-W, Strecker A, Wise D, Lafrenz M, Shandas V, et al. Water supply, demand, and quality indicators for assessing the spatial distribution of water resource vulnerability in the Columbia River Basin. *Atmosphere-Ocean*. 2013;**51**(4):339-356. DOI: 10.1080/07055900.2013.777896
- [29] Zamanzad-Ghavidel S, Sobhani R, Etaei S, Hosseini Z, Montaseri M. Development of hydro-social-economic-environmental sustainability index (HSEESI) in integrated water resources management. *Environmental Monitoring and Assessment*. 2021;**193**(8):1-29. DOI: 10.1007/s10661-021-09129-4
- [30] Ortega M, Velasco J, Millan A, Guerrero C. An ecological integrity index for Littoral Wetlands in Agricultural Catchments of Semiarid Mediterranean Regions. *Environmental Management*. 2004;**33**(3):412-430. DOI: 10.1007/s00267-003-3059-6
- [31] Silva J da, Fernandes V, Limont M, Dziedzic M, Andreoli CV, Rauen WB. Water sustainability assessment from the perspective of sustainable development capitals: Conceptual model and index based on literature review. *Journal of Environmental Management*. 2020;**254**:109750. DOI: 10.1016/j.jenvman.2019.109750
- [32] Chaves HML, Alipaz S. An integrated indicator based on basin hydrology, environment, life, and policy: The Watershed Sustainability Index. *Water Resources Management*. 2007;**21**(5):883-895. DOI: 10.1007/s11269-006-9107-2
- [33] Boyacıoğlu & Boyacıoğlu. Ecological water quality index associated with factor analysis to classify surface waters. *Water Science and Technology: Water Supply*. 2020
- [34] Etemi FZ, Bytyçi P, Ismaili M, Fetoshi O, Ymeri P, Shala-Abazi A, et al. The use of macroinvertebrate based biotic indices and diversity indices to evaluate the water quality of Lepenci river basin in Kosovo. *Journal of Environmental Science and Health, Part A*. 2020;**55**(6):748-758. DOI: 10.1080/10934529.2020.1738172
- [35] Villeneuve B, Souchon Y, Usseglio-Polatera P, Ferréol M, Valette L. Can we predict biological condition of stream ecosystems? A multi-stressors approach linking three biological indices to physico-chemistry, hydromorphology and land use. *Ecological Indicators*. 2015;**48**:88-98. DOI: 10.1016/j.ecolind.2014.07.016
- [36] Oberholster PJ, McMillan P, Durgapersad K, Botha AM, de Klerk AR. The Development of a Wetland Classification and Risk Assessment Index (WCRAI) for Non-Wetland Specialists for the Management of Natural Freshwater Wetland Ecosystems. *Water, Air, & Soil Pollution*. 2014;**225**(2)
- [37] Baptista DF, Souza RSG de, Vieira CA, Mugnai R, Souza AS, Oliveira RBS de. Multimetric index for assessing ecological condition of running waters in the upper reaches of the Piabanha-Paquequer-Preto Basin, Rio de Janeiro, Brazil. *Zoologia (Curitiba)*. 2011;**28**(5):619-628. DOI:10.1590/S1984-46702011000500010
- [38] Salinas-Camarillo VH, Carmona-Jiménez J, Lobo EA. Development of the Diatom Ecological Quality Index

- (DEQI) for peri-urban mountain streams in the Basin of Mexico. *Environmental Science and Pollution Research*. 2020;**28**(12):14555-14575. DOI: 10.1007/s11356-020-11604-3
- [39] Veras DS, Castro ER, Lustosa GS, de Azevêdo CA, Juen L. Evaluating the habitat integrity index as a potential surrogate for monitoring the water quality of streams in the cerrado-caatinga ecotone in northern Brazil. *Environmental Monitoring and Assessment*. 2019;**191**(9):1-9. DOI: 10.1007/s10661-019-7667-x
- [40] Tipa G, Teirney LD, New Zealand. Ministry for the Environment. A Cultural Health Index for Streams and Waterways: Indicators for Recognising and Expressing Māori Values. Wellington, N.Z: Ministry for the Environment; 2003
- [41] Sullivan CA. Quantifying water vulnerability: A multi-dimensional approach. *Stochastic Environmental Research and Risk Assessment*. 2010;**25**(4):627-640. DOI: 10.1007/s00477-010-0426-8
- [42] UNESCO-IHP, IGRAC, WWAP. GEF Transboundary Waters Assessment Programme (TWAP): Transboundary Aquifers and SIDS Groundwater Systems. 2012
- [43] Abel N. Natural Values: Exploring Options for Enhancing Ecosystem Services in the Goulburn Broken Catchment. Csiro: Erscheinungsort Nicht Ermittellbar; 2003
- [44] Ioris AAR, Hunter C, Walker S. The development and application of water management sustainability indicators in Brazil and Scotland. *Journal of Environmental Management*. 2008;**88**(4):1190-1201. DOI: 10.1016/j.jenvman.2007.06.007
- [45] Schneider F, Bonriposi M, Graefe O, Herweg K, Homewood C, Huss M, et al. Assessing the sustainability of water governance systems: The sustainability wheel. *Journal of Environmental Planning and Management*. 2014;**58**(9):1577-1600. DOI: 10.1080/09640568.2014.938804
- [46] Kang M-G, Lee G-M. Multicriteria evaluation of water resources sustainability in the context of watershed management. *JAWRA Journal of the American Water Resources Association*. 2011;**47**(4):813-827. DOI: 10.1111/j.1752-1688.2011.00559.x
- [47] Sun X, Xiong S, Zhu X, Zhu X, Li Y, Li LB. A new indices system for evaluating ecological-economic-social performances of wetland restorations and its application to Taihu Lake Basin. *China Ecological Modelling*. 2015;**295**:216-226
- [48] Picone C, Henke R, Ruberto M, Calligaris E, Zucaro R. A synthetic indicator for sustainability standards of water resources in agriculture. *Sustainability*. 2021;**13**(15):8221. DOI: 10.3390/su13158221
- [49] Ferreira SCG, de Lima AMM, Corrêa JAM. Indicators of hydrological sustainability, governance and water resource regulation in the Moju river basin (PA) – Eastern Amazonia. *Journal of Environmental Management*. 2020;**263**:110354. DOI: 10.1016/j.jenvman.2020.110354
- [50] ILEC. Methodology for the GEF Transboundary Waters Assessment Programme. Vol. 3. Methodology for the Assessment of Transboundary Lake Basins, UNEP; 2011. p. 69
- [51] Abdi-Dehkordi M, Bozorg-Haddad O, Chu X. Development of a combined index to evaluate sustainability of water

resources systems. *Water Resources Management*. 2021;**35**(9):2965-2985. DOI: 10.1007/s11269-021-02880-w

[52] Pellicer-Martínez F, Martínez-Paz JM. The Water Footprint as an indicator of environmental sustainability in water use at the river basin level. *Science of The Total Environment*. 2016;**571**:561-574. DOI: 10.1016/j.scitotenv.2016.07.022

[53] Yano S, Yamaguchi M, Yokoi E, Kanayama T, Kubota A, Ogawada D, et al. Using the sectoral and statistical demand to availability index to assess freshwater scarcity risk and effect of water resource management. *Journal of Hydrology X*. 2020;**8**:100058. DOI: 10.1016/j.hydroa.2020.100058

[54] Bao K, Liu J, You X, Shi X, Meng B. A new comprehensive ecological risk index for risk assessment on Luanhe River, China. *Environmental Geochemistry and Health*. 2018;**40**(5):1965-1978. DOI: 10.1007/s10653-017-9978-6

[55] Smakhtin V, Revenga C, Döll P. Taking into account environmental water requirements in global-scale water resources assessments. Colombo, Sri Lanka: International Water Management Institute (IWMI); 2004. p. 24

[56] Crispim DL, Pimentel Da Silva GD, Fernandes LL. Rural water sustainability index (RWSI): An innovative multicriteria and participative approach for rural communities. *Impact Assessment and Project Appraisal*. 2021;**39**(4):320-334. DOI: 10.1080/14615517.2021.1911752

[57] Peruchi Trevisan D, da Conceição BP, Almeida D, Imani M, Balzter H, Eduardo ML. Environmental vulnerability index: An evaluation of the water and the vegetation quality in a Brazilian Savanna and Seasonal Forest biome. *Ecological Indicators*.

2020;**112**:106163. DOI: 10.1016/j.ecolind.2020.106163

[58] Cai J, Zhao D, Varis O. Match words with deeds: Curbing water risk with the Sustainable Development Goal 6 index. *Journal of Cleaner Production*. 2021;**318**:128509. DOI: 10.1016/j.jclepro.2021.128509

[59] Babel MS, Pandey VP, Rivas AA, Wahid SM. Indicator-based approach for assessing the vulnerability of freshwater resources in the Bagmati River Basin, Nepal. *Environmental Management*. 2011;**48**(5):1044-1059. DOI: 10.1007/s00267-011-9744-y

[60] Corrêa MA, Teixeira BA. do N. Developing sustainability indicators for water resources management in Tietê-Jacaré basin, Brazil. *Journal of Urban and Environmental Engineering*. 2013:8-14. DOI: 10.4090/juee.2013.v7n1.008014

[61] Cai W, Xia J, Yang M, Wang W, Dou C, Zeng Z, et al. Cross-basin analysis of freshwater ecosystem health based on a zooplankton-based Index of Biotic Integrity: Models and application. *Ecological Indicators*. 2020;**114**:106333. DOI: 10.1016/j.ecolind.2020.106333

[62] Fraser Basin Council. Measuring & Reporting on Sustainability: A Report on Lessons Learned. FBC; 2011

[63] Kuhar U, Germ M, Gaberščik A, Urbanič G. Development of a River Macrophyte Index (RMI) for assessing river ecological status. *Limnologica*. 2011;**41**(3):235-243. DOI: 10.1016/j.limno.2010.11.001

[64] Aura CM, Nyamweya CS, Owiti H, Odoli C, Musa S, Njiru JM, et al. Citizen science for bio-indication: Development of a community-based index of ecosystem integrity for assessing the status of afrotropical riverine

- ecosystems. *Frontiers in Water*. 2021;**2**:1-13. DOI: 10.3389/frwa.2020.609215
- [65] Cai X, McKinney DC, Lasdon LS. A framework for sustainability analysis in water resources management and application to the Syr Darya Basin. *Water Resources Research*. 2002;**38**(6):21
- [66] Kefayati M, Saghafian B, Ahmadi A, Babazadeh H. Empirical evaluation of river basin sustainability affected by inter-basin water transfer using composite indicators. *Water and Environment Journal*. 2017;**32**(1): 104-111. DOI: 10.1111/wej.12304
- [67] UNEP-DHI, UNEP. Transboundary River Basins: Status and Trends, Summary for Policy Makers. Nairobi: United Nations Environment Programme (UNEP); 2016
- [68] Davies PE, Harris JH, Hillman TJ, Walker KF. The Sustainable Rivers Audit: Assessing river ecosystem health in the Murray - Darling Basin, Australia. *Marine and Freshwater Research*. 2010;**61**(7):764. DOI: 10.1071/MF09043
- [69] van Oosterhout MP, van der Velde G. An advanced Index of Biotic Integrity for use in tropical shallow lowland streams in Costa Rica: Fish assemblages as indicators of stream ecosystem health. *Ecological Indicators*. 2015;**48**:687-698. DOI: 10.1016/j.ecolind.2014.09.029
- [70] Brooks R, McKenney-Easterling M, Brinson M, Rheinhardt R, Havens K, O'Brien D, et al. A Stream–Wetland–Riparian (SWR) index for assessing condition of aquatic ecosystems in small watersheds along the Atlantic slope of the eastern U.S. *Environmental Monitoring and Assessment*. 2008;**150**(1-4):101-117. DOI: 10.1007/s10661-008-0673-z
- [71] Magee TK, Ringold PL, Bollman MA, Ernst TL. Index of Alien Impact: A method for evaluating potential ecological impact of Alien plant species. *Environmental Management*. 2010;**45**(4):759-778. DOI: 10.1007/s00267-010-9426-1
- [72] UNESCO. Ecohydrology as an integrative science from molecular to basin scale: Historical evolution, advancements, and implementation activities. 2016. Available from: <https://unesdoc.unesco.org/ark:/48223/pf0000245512?1=null&queryId=4661711d-1493-4c5a-8228-6b252a68eb0e>. [Accessed: 4 January 2022]

Technologies and Materials for Sulfate Removal in Water: A Review

Ariel Antonio Quintana-Baqueda, Jose Luis Sanchez-Salas and Deborah Xanat Flores-Cervantes

Abstract

With the constant increase in world population, water pollution, water stress, and the shortage of drinking water sources for human use and consumption, it is mandatory for technology and research to focus on alternatives that provide options to counter or stop these problems. That is why a new and important alternative is proposed as an additional source of water supply for human consumption: sulfur removal in water. Sulfur-containing water effluents exist throughout the world. However, these waters are normally only used for recreational uses or are discarded, since they are rejected by humans for consumption due to their bad smell and taste, despite the fact that these waters do not represent a significant risk to human health. This review provides information for users to identify the best possible technology available based on temperature, pH, maximum sulfate concentration, economic availability, human effects, and land/space resources. This work includes a) the use of natural samples and/or synthetic water with possible interferences; b) the use of real world concentrations above the established limits by WHO and EPA; c) the use environmental conditions of pH and temperature to reduce costs; d) the regeneration of the material/technology and possible uses for the recovered sulfur material; e) the use of nanotechnology and more environmentally friendly technologies.

Keywords: sulfate, sulfur, hydrogen sulfide, removal sulfate, sulfurous waters, sulfur in water, sulfur-containing

1. Introduction

Today, the world is experiencing water stress and it is estimated that by 2030, the demand for water will increase by 50% [1]. This vital liquid must supply more than 7,900 million people around the world [2], a population that continues to increase as well as its economic activity. Added to this, the growing contamination of surface waters, stresses the water available for human consumption, among others. That is why there is an urgent need to develop new technological options for the purification and supply of water. Water containing sulfur, could prove to be an additional source of water for human use and consumption. However, it is known that sulfur can occur in different forms in nature, mainly sulfate and hydrogen sulfide. Sulfur speciation in the liquid medium can be visualized as a function of pH and its exposure to oxygen.

Higher concentrations of sulfate will be found where higher oxygen concentrations are present, mainly surface waters, while hydrogen sulfide will be important in ground water where oxygen is usually depleted. In addition, the water pH will depend on the geographical area and the presence of other chemical or microbial forms in the reservoir or water source [3].

Sulfur, particularly in sulfate form, is abundant in natural effluents, in groundwater where the soil is rich in gypsum, as well as in wastewater from industrial sources [4]. Sulfur-containing water effluents exist throughout the world. However, these waters are normally only used for recreational uses or are discarded, and not considered for human consumption because of their bad smell and taste [4]. However, these waters do not represent a significant risk to human health. An example is the city of Puebla Capital in Mexico, where sulfur-containing water could contribute an additional 7.4% in the supply of drinking water, which is urgently needed, since there are communities that do not receive continuous water service due to limited volumes of water available [5].

Sulfate is not carcinogenic, and its side reactions are limited to stomach problems at concentrations above 300 mg of sulfate per liter in children and above 500 mg L⁻¹ in adults [4]. The World Health Organization (WHO) and the Environmental Protection Agency (EPA), in the United States specify a limit of 250 mg L⁻¹ [6, 7]. In Mexico, sulfate concentration limits are regulated by SEMARNAT (Ministry of the Environment and Natural Resources), with a limit of 400 mg L⁻¹ [8]. These agencies suggest nanofiltration and ion exchange as recommended methodologies for the treatment of sulfates present in water; however, there exist other alternative technologies and materials that could remove sulfate present in water and be used in different countries.

In this review, are describe the existing technologies for sulfate removal, and the information contained in each report will be evaluated. The analysis will be done considering the following factors: availability, economy, speed, and efficiency. Due to the scarce information in the literature on the research topic, no filter is used for the year of publication, everything studied so far had to be considered to maintain a broad and sustained panorama. Finally, are identified and proposed different opportunity areas for sulfate removal.

2. Technologies and materials

According to the literature, the methodologies that have been studied for sulfur removal in water are membrane filtration, adsorption, ion exchange, bioreactors, and constructed wetlands. **Table 1** specifies the main characteristics of each technology or material, grouped together in three categories: physical, chemical, and biological procedures.

The Technology Readiness Level (TRL) is an index that helps measure or qualify the maturity of the technological components of a system. Here, this analysis is used to determine the level of application in which each investigation was carried out, what does it consist of, if it was at the laboratory level, pilot scale or even at the industrial level. The scale is established from level 1 (passage from scientific research to applied research) to level 9 (technology in its final form). In **Figure 1**, the level of maturity based on the information presented in each study is shown. As can be seen, most of

Methodology	Description
Physical	
Nano-filtration	Description: Process by which an effluent pass through a membrane. Membranes with pore sizes between 0.001 and 0.01 mm. Common use: to reduce total salinity (reduction from 50 to 70%), eliminate a significant part of divalent ions (up to 95%; mainly sulfates) [9], proteins, molecules, among others [10]. Special aspects to consider: Calcium sulfate causes high turbidity in the solution, which causes membrane fouling. In general, acidification is recommended prior to treatment with membranes, to avoid fouling caused by the presence of minerals in the water [11].
Reverse Osmosis	Description: Process by which an external hydraulic pressure (hydraulic pump) is applied to force the water through the membrane against the osmotic pressure. Common use: For the reduction of high levels of nitrate, sulfate, sodium, and total dissolved solids [12]. Special aspects to consider: selection and / or design of a membrane, pretreatments, tendency to clog, and operating conditions [9]. For the treatment of sulfates with reverse osmosis it is important that the input concentration is less than 700 mg L ⁻¹ [13].
Adsorption ^a	Description: The principle of operation is the adhesion of contaminants through a porous medium. This porous medium in industry, due to its low costs is activated carbon, but it can be zeolites, silica, pumice stone, among others. Common use: This is one of the most common and used methods in the treatment of water for the removal of a wide range of contaminants [14]. Special aspects to consider: Over time, different ways of increasing the surface area have been investigated, from changing the raw material starting from organic materials, to activating with acids or bases to improve the elimination affinity with the contaminant to be eliminated. One of the well known adsorbent is activated carbon which is made from many carbonaceous sources including coal, coke, pest, wood, and coconut shell [14]. Organic matter and turbidity are important factors present in the water that affects the performance of treatments in drinking water specifically on activated carbon [15].
Chemical	
Ion exchange	Description: The resins are composed of high concentrations of polar, acidic, or basic groups and can be of two types; anionic and cationic and each one of them is prepared with acids or bases, weak or strong according to the application. The size of the ionic resin beads ranges from 0.3 to 1.2 mm [11]. Common use: It is mainly used for the removal of heavy metals. Especially in the removal and recovery of Pb ²⁺ , Ni ²⁺ , Cu ²⁺ , Cd ²⁺ and Fe ³⁺ ions from both natural and industrial water sewage water [16]. Also used for sulfate removal [11]. Special aspects to consider: In general, resins are operated in columns to promote exchange and their main advantage is that their exchange capacity can be regenerated and recovered [16]. As with reverse osmosis, scaling of CaSO ₄ is a common problem [13].
Biological	
Bioreactor	Description: Aerobic or anaerobic, and it works continuously or in batches [17]. In bioelectrochemical systems, anodic, cathodic, or both redox reactions are catalyzed by microorganisms. In circumstances where the potential of the cathode-catalyzed redox reaction is greater than the anodic one, as in the case of oxidation of organic compounds and presence of oxidant at the cathode, electricity is generated in microbial fuel cells [18]. Common use: You can have various designs and configurations of reactors, depending on the application, for example: Membrane bioreactor, biofilm formation reactors, rotating packed bed biofilms Reactors, Monolithic Biofilm Reactor, Fixed Bed Reactors, Bubble Column Reactor, among others [18]. Special aspects to consider: Alternative to treat waters with high levels of sulfate (700–1400 mg L ⁻¹), due to the reduction of microorganisms with affinity for the substrate [19]. High concentrations of molecular H ₂ S (g) can inhibit microbial activity, especially at a reactor pH of 6 to 7 [20].

Methodology	Description
Artificial wetland	Description: Constructed wetlands are wastewater phytopurification systems. It consists of the development of a macrophyte culture rooted in a waterproofed gravel bed. The action of macrophytes enables a series of complex physical, chemical, and biological interactions through which the wastewater is gradually and slowly purified [21]. Common use: Removal of heavy metals, mainly zinc and lead, although not necessarily due to biogeochemical stabilization, maturation of wetlands or seasonal effects, but by the variable load of metal in the influent [22]. Special aspects to consider: They are natural systems of easy application, low costs, and operational requirements [21]. Long term treatment of sulfate in these systems can be compromised as substantial metal cations are required to precipitate high levels of sulfide anions [22].

^aAdsorption is both a chemical and a physical process, however, for purposes of the criteria for ordering the table, it is included in physical.

Table 1.
Summary of analyzed methods for sulfate removal.

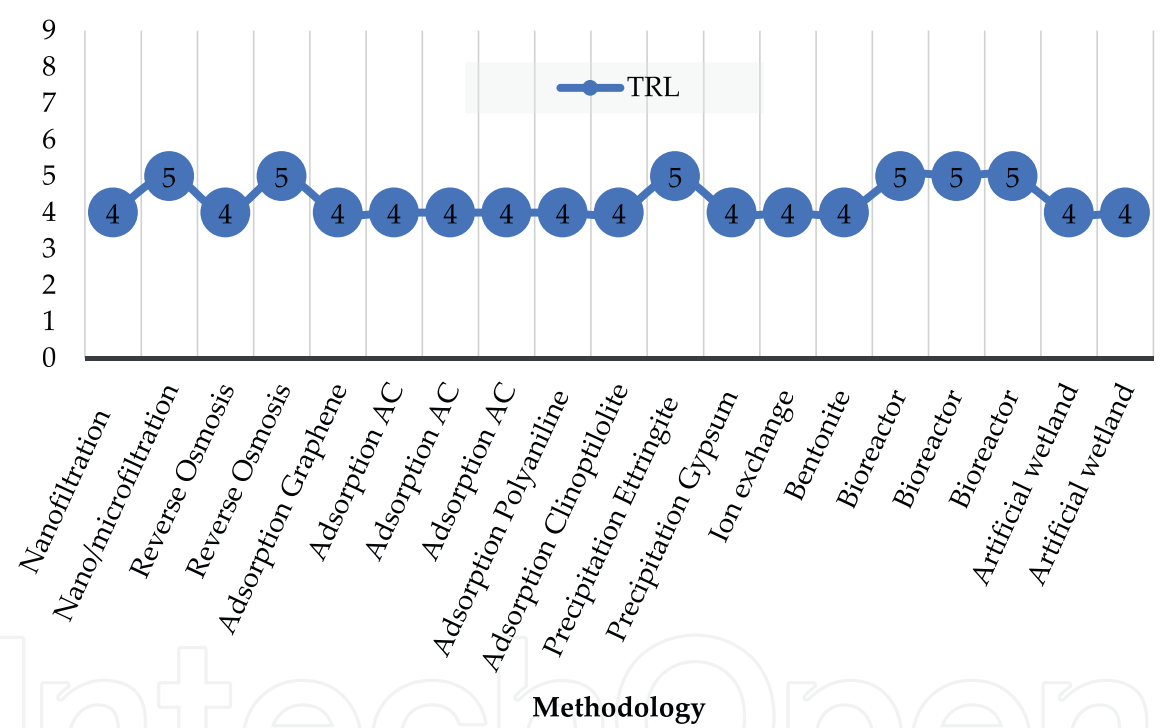


Figure 1.
TRL of each methodology analyzed. * AC (Activated Carbon).

the studies fall in level 4 (validation in laboratory environment) or 5 (validation in relevant environment, in this case, prototype) [23].

In order to understand how the experiments were made and to compare between studies, it is important to determine the characteristics and parameters used. **Table 2** lists the reports found related to the topic where at least two of the following parameters are available: maximum input concentration, % removal, pH, and/or T.

- mg L⁻¹ Maximum input concentration
- % Maximum percentage of removal

	#	Methodology	Article	Reference	mg L ⁻¹	%	pH	T (°C)
Physical	1	Nanofiltration	Sulfate removal from water	[11]	5,393	97	7.5	RT
	2	Nanofiltration and microfiltration	Nanofiltration and microfiltration for the removal of chromium, total dissolved solids, and sulfate from water	[10]	500	98	6.3	25
	3	Reverse osmosis	Reduction of sulphate content in aqueous solutions by reverse osmosis using cellulose acetate membranes	[24]	25,455	NR	6	25
	4	Reverse osmosis	A review of sulfate removal options for mine waters	[13]	600	98	6	NR
	5	Adsorption graphene	Thermodynamic, kinetic and isotherm studies of sulfate removal from aqueous solutions by graphene and graphite nanoparticles	[25]	75	NR	3	RT
	6	Adsorption Activated carbon	Sulfate ion removal from water using Activated carbon Powder prepared by ziziphus Spina-Christi Lotus Leaf	[26]	20	84.5	6	45
	7	Adsorption Activated carbon	Sulfate removal from acid mine drainage using polypyrrole-grafted granular activated carbon	[27]	773	NR	8	RT
	8	Adsorption Activated carbon	Removal of sulfate from wastewater by activated carbon	[14]	520	43	8	25
	9	Adsorption Polyaniline and composites	Removal of sulfate from water using polyaniline and its composites	[28]	66.21	91	7	RT
	10	Adsorption Nano-Clinoptilolite	Removal of sulfate from Gamasiab river water samples by using natural nano-Clinoptilolite	[29]	45	95	9.5	RT
	11	Adsorption Biochar	Kinetic and isotherm studies on adsorptive removal of sulfate by cotton shell derived biochar: Recovery of sulfates from marcasite soil.	[30]	100	91	9.8	RT
	12	Adsorption Zirconium oxide-modified biochar	Adsorption of sulfate ion from water by zirconium oxide-modified biochar derived from pomelo peel.	[31]	300	NR	2	45
	13	Adsorption Zea mays Stems	Efficient Sulfate adsorption on modified adsorbents prepared from Zea Mays Stems	[32]	100	NR	4	25

	#	Methodology	Article	Reference	mg L ⁻¹	%	pH	T (°C)
Chemical	1	CESR (Cost Effective Sulfate Removal)	A review of sulfate removal options for mine waters	[13]	800	75	11	NR
	2	CESR (Cost effective sulfate removal)	A new process for sulfate removal from industrial waters	[33]	8000	80	12	25
	3	Ion exchange	Sulfate removal from water	[11]	3,665	90	7.7	NR
	4	Bentonite	Sulfate removal from water	[11]	25	NR	7.5	NR
Biological	1	Bioreactor	A review of sulfate removal options for mine waters	[13]	3000	82	7	37
	2	Bioreactor	Improved sulphate removal rates at increased sulphide concentration in the sulphidogenic bioreactor	[19]	700	NR	8.5	25
	3	Bioreactor	Removal of sulphate, COD and Cr (VI) in simulated and real wastewater bysulphate reducing bacteria enrichment in small bioreactor and FTIR study	[20]	50	81	8.5	25
	4	Artificial wetland	Sulfate and metal removal in bioreactors treating acid mine drainage dominated with iron and aluminum.	[34]	608	94	NR	NR
	5	Artificial wetland	Removal of sulfate, zinc, and lead from alkaline mine wastewater using pilot-scale surface-flow wetlands at tara mines, Ireland.	[22]	900	72	NR	NR

NR. No Record
RT. Room Temperature

Table 2.
Summary of relevant parameters depicted in each study for sulfate removal reviewed.

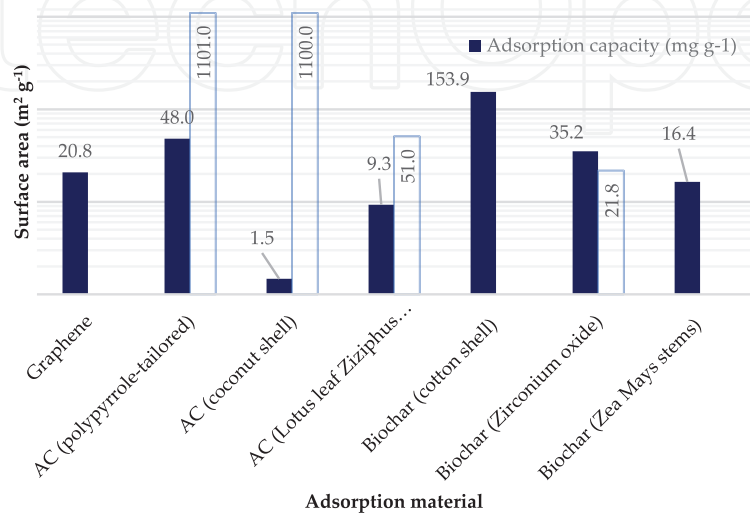


Figure 2.
Maximum adsorption capacity and surface area reported for different tested materials. * AC (Activated Carbon).

- pH Hydrogen potential
- T Temperature

When not available, “N/R” is reported.

As can be seen, within the physical treatments, nine of them are by adsorption, which requires an independent analysis of the information. The analysis between different adsorption materials and a comparison between the sulfate adsorption capacity for each material tested can be seen in **Figure 2**.

3. Analysis of results

As show in **Figure 1**, the information found, indicates that the evaluated technologies still need to mature, as no evidence of their use at higher TRLs were found. The evaluated studies only included laboratory or prototype applications. However, there is indication that reverse osmosis and ion exchange resins have higher TRLs, as water treatment and purification companies already offer the products at a commercial scale [35].

Table 2 shows the big diversity in the information retrieved from the research papers, with important variations in the experimental conditions, especially for maximum concentrations and pH ranges. However, variations in efficiencies, and temperature, were lower. Considering commercial or water treatment plant applications, the removal efficiencies presented are adequate; however, some pHs and temperatures are not at environmental conditions, requiring additional expenses.

The range of maximum initial sulfate concentrations in the analyzed samples vary widely ranging from 25,000 mg L⁻¹ in a case using reverse osmosis, 8,000 mg L⁻¹ with CESR, and 5,000 mg L⁻¹ for nanofiltration. On the other end, there were also cases with concentrations below the limits established by international standards (250 mg L⁻¹). Regarding this last point, the studies carried out by adsorption, which are the majority, work with concentrations as low as 20 mg L⁻¹ reporting a removal above 90%. The problem of working with such low concentrations is that they already are below the limit of permitted concentrations. In addition, when up scaling the process to higher concentrations and full-scale applications, the margin of error would increase considerably and probably the system will be saturated easily. Studies should consider sulfate concentrations similar to those found in natural effluents, for example, 750 mg L⁻¹ obtained in the city of Puebla, Mexico [36] or in the rivers in western Canada with records up to 3040 mg L⁻¹, or the sea water that contains around 2700 mg L⁻¹ [7, 11].

Most treatments report removal rates of more than 72%, except for Hong and collaborators [27], where the removal efficiency was 43% using activated carbon (polypyrrole-tailored). Nanofiltration and reverse osmosis had the highest removal rates, with values above 97%. However, it is important to consider some advantages and disadvantages of these systems. For example, for nanofiltration, which removes up to 97% of the sulfates present in water, it is necessary to include pretreatments, and increasing the pressure to values between 40 and 80 psi depending on the studies conditions. This pumping process increases the cost, since it implies higher energy consumption, and at the same time reduces the half-life of the membrane [11]. Similarly, reverse osmosis removes up to 98% (with samples that exceed 6,000 mg L⁻¹); however, it needs pretreatment such as chlorination, softening and/or ion exchange, to avoid flakes formation by reactions on the walls of the membranes [24].

Although not all the reports indicate a required pretreatment, adsorption might also require the modification of the pH to increase the removal efficiency, depending on the pollutant to be removed. According to the methodology and reagents used, adsorption has been used in acidic or alkaline conditions, giving better results with acidic pH. Similarly, temperature is a parameter that directly affects the efficiency of removal by adsorption. In both cases, when pH and temperature modifications are required, removal efficiency costs increase [26].

For biological processes, although removal percentages are above 72%, response time and environmental conditions must be considered. For example, O'Sullivan and collaborators reported a retention time of more than 30 days [22]. However, these processes are among the most economical, as they show, not require temperature changes, work naturally, and typically operate at a neutral pH in the case of constructed wetlands. Regarding the bioprocess, constructed wetlands differs from bioreactors, because the former requires a large land and therefore a high initial investment.

In **Table 2**, the pH and temperature are reported whenever they are present in the study evaluated, as well as the concentration of the samples studied and the percentages of sulfate removal. In addition, it is also important to identify the presence of other components in the water samples, such as competitive ions, that might affect the removal efficiency of sulfates from water. Although Bowel and collaborators mention an extra advantage of their method, because nitrate can be removed together with sulfates during the ettringite precipitation, unfortunately they do not show any results about it [13].

Before introducing a particular method to remove sulphates from water, it is necessary to have all the information previously mentioned for the correct characterization of the samples to be treated, as well as defining their origin. It is important to differentiate the analysis of water from natural sources than the studies using synthetic water, because the natural water contains different factors, which cannot be controlled as with synthetic water. For example, the presence of other anions will compete with sulfates during the adsorption process; also, when using membranes, the presence of calcium sulfate directly affects the generation of flakes, clogging the pores of the membrane [11].

Finally, there are notable differences in the adsorption capacity of evaluated adsorption studies. For example, Rumjit and collaborators [30], with Biochar, reported a maximum adsorption capacity of 153.85 mg g^{-1} , while Rahmati and collaborators [26] reported a value of 9 mg g^{-1} with activated carbon (see **Figure 2**). Rumjit and collaborators [30], however, reported a pH of 9.8 at room temperature, but no record of the surface area of the biochar tested. On the other hand, Rahmati et al, reported the surface area of their material of $51 \text{ m}^2 \text{ g}^{-1}$ working at pH 6 and temperature of 45°C but has less removal capacity [26]. Furthermore, Rahmati, with a lower adsorption capacity, reports a higher surface area in the material tested than Ao and collaborators [31] ($21.7 \text{ m}^2 \text{ g}^{-1}$), which has the second highest removal capacity from the studies reviewed with 35.2 mg g^{-1} [31].

Only one of the three biochar studies reports its surface area, which is much lower than the one reported for the activated carbon, where surface areas of up to $1100 \text{ m}^2 \text{ g}^{-1}$ are reached in at least two research articles. Of the nine articles reported with adsorption, 55.6% do not report surface area data, 22% do not record the data of maximum adsorption capacity, and those that do, present values at different conditions.

4. Conclusions

The objective of this document was to carry out a methodological analysis of the technologies available for the removal of sulfur in water for human consumption. This review provides a guideline to possible users or possible research in the area of sulfate removal. A user, based on the characteristics of the water, concentration, pH, T, can choose a method that is fit for a given application, considering the available resources (monetary, human, space), and characteristics of the water to be treated.

Given the complexity and diversity of the technologies and materials evaluated, choosing the best methodology or material depends on several factors, including equipment availability, technology, operation, maintenance, economy, response speed, efficiency. These factors become more relevant for developing countries like Mexico, where there is a lack of trained technicians, or places with the greatest need for water have limited or no access to replacement parts.

All necessary pretreatment, pH change, use of reagents before or after treatment, installation or initial investment, operation, and maintenance requirements, must be considered when evaluating which methodology is better, especially if there are economic restrictions. When considering biological treatment, it is important to keep in mind the high retention times needed for relevant removal efficiencies. Finally, it is also important to keep in mind that a 98% efficiency removal with a sample below 66 mg L^{-1} (as is the case of Adsorption Polyaniline) is not the same as a 90% efficiency with concentrations of $3,665 \text{ mg L}^{-1}$ (as in the case of ion exchange).

This review also helps identify the following research gaps and opportunities:

1. Future research should include samples with concentrations above the limits established by WHO and EPA.
2. For the technologies and materials to have to potential to be implemented at a large scale, experimental conditions (T and pH) should be similar to those found in the environment.
3. Synthetic samples should include possible cations or anions found in natural waters to identify changes in the removal processes or efficiencies. This in addition to including natural samples.
4. The studies should include regeneration to identify the potential possibilities for commercial use and to recover sulfur to be used as raw material for other products or applications.
5. Attention should be given to technologies and materials that consume less energy and are more environmentally friendly. For example, to consider biochar from different sources of biomass, instead of fossil fuel-based AC.
6. There is room for the development of nanomaterials that could resemble some of the studied adsorbants but at a nanoscale, enhancing the surface area available for sorption, the presence of functional groups, and/or the inclusion of other desirable characteristics.

Sulfur water has not been considered for human consumption due to social rejection caused by the characteristic taste and smell of sulfur present in the water. However, this water has the potential to become an important source of drinking water after treatment. This review provides the basis to identify technologies that can be effective at removing sulfate from potential drinking water sources. In particular, for countries like Mexico, there is still room for improvement in developing a methodology that is effective and economic, and that can be used at a large scale in water purification plants. Currently, the best options are physical treatments, such as the use of membranes or adsorbents, but other materials or a combination of them could provide better alternatives.

Acknowledgements

This review was possible thanks to the Fundación Universidad de las Américas Puebla.

Funding

This work was supported by a CONACYT grant (Registration number 443326), which allocated funds for the development of research and a scholarship for AAQB.

Authors' contributions

All authors contributed with ideas for the generation of the article. AAQB wrote the general document. JLA and DXFC reviewed, corrected, and provided feedback. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that there is no conflict of interests.

IntechOpen

Author details

Ariel Antonio Quintana-Baquedano¹, Jose Luis Sanchez-Salas²
and Deborah Xanat Flores-Cervantes^{3*}

1 Departamento de Ingenieria Civil y Ambiental, Universidad de las Americas Puebla, Puebla, Mexico

2 Departamento de Ciencias Quimico-Biologicas, Universidad de las Americas Puebla, Puebla, Mexico

3 Departamento de Ingenieria Quimica, Alimentos y Ambiental, Universidad de las Americas Puebla, Puebla, Mexico

*Address all correspondence to: deborah.flores@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] UNESCO. The United Nations world water development report, 2017: Wastewater: The untapped resource – UNESCO Biblioteca Digital [Internet]. 2017. Available from: <https://unesdoc.unesco.org/ark:/48223/pf0000247153>
- [2] Worldometers.info. Población Mundial: 7.8 Billones de Personas (2022) – Worldometer. 2022. [Internet] Available from: <[https://www.worldometers.info/es/poblacion-mundial/#:~:text=Poblaci%C3%B3n%20Mundial%3A%207.8%20Billones%20de%20Personas%20\(2022\)%20%2D%20Worldometer](https://www.worldometers.info/es/poblacion-mundial/#:~:text=Poblaci%C3%B3n%20Mundial%3A%207.8%20Billones%20de%20Personas%20(2022)%20%2D%20Worldometer)> [Accessed: 1 February 2022]
- [3] Arslan F. Electrooxidation as a pretreatment process before cyanidation. *Noble Metals*. 2012;**2012**:370-388
- [4] Severiche CA, González H. Evaluación analítica para la determinación de sulfatos en aguas por método turbidimétrico modificado. *Ingenierías USBMed*. 2012;**3**(2):6-11
- [5] Gárfias J, Arroyo N, Aravena R. Hydrochemistry and origins of mineralized waters in the Puebla aquifer system, Mexico. *Environmental Earth Sciences*. 2010;**59**(8):1789-1805
- [6] U.S.EPA. USEPA, 2002 Edition of Drinking Water Standards and Health Advisories. EPA 822-R-02-038, Washington DC.pdf. 2002. p. 10-2
- [7] WHO. Sulfate in Drinking-water. Guidelines for drinking water quality [Internet]. 2004; Available from: WHO/SDE/WSH/03.04/114
- [8] NOM-127-SSA1-1994. DOF – Diario Oficial de la Federación. Vol. 5002338. 2007. p. 1-14
- [9] Guizard C. TECNICAS MEMBRANARIAS de FILTRACION de LIQUIDOS Microfiltracion – Ultrafiltracion – Nanofiltracion – Osmosis inversa. Módulo de enseñanza en fenómenos interfaciales. 1999;**56**:1-52
- [10] Zolfaghari G, Kargar M. Nanofiltration and microfiltration for the removal of chromium, total dissolved solids, and sulfate from water. *MethodsX*. 2019;**6**:549-557
- [11] Darbi A, Viraraghavan T, Jin Y-C, Braul L, Corkal D. Sulfate Removal from Water. *Water Quality Research Journal of Canada*. 2003;**38**
- [12] Bergsrud F. Treatment Systems for Household Water Supplies Reverse Osmosis. *Water Quality Specialist. Water Quality Coordinator*; 1992. pp. 1-4
- [13] Howell RJ. A review of sulfate removal options for mine waters: A review of sulphate removal options for mine waters. Vol. 2. 2004
- [14] Sadeq SM. Removal of sulfate from waste water by activated carbon. *Journal Al-Khwarizmi Engineering Journal*. 2009;**5**:72-76
- [15] Torres-Lozada P, Amezquita-Marroquín CP, Agudelo-Martínez KD, Ortiz-Benítez N, Martínez-Ducura DS. Evaluation of turbidity and dissolved organic matter removal through double filtration technology with activated carbon. *DYNA (Colombia)*. 2018;**85**(205):234-239
- [16] Raeissi AS, Shahadat M, Bushra R, Nabi SA. Development of titanium-supported ion-exchange adsorbent for removal of metal pollutants. *Arabian Journal for Science and Engineering*. 2018;**43**(7):3601-3609

- [17] Flores A, Flores M, Reyes J, Astorga C. Diseño y modelado de un bioreactor tipo batch y continuo para aplicaciones de control automático. Aplicaciones a la microbiología industrial. 2013;**2013**:86-92
- [18] Ayol A, Peixoto L, Keskin T, Abubackar HN. Reactor designs and configurations for biological and bioelectrochemical c1 gas conversion: A review. International Journal of Environmental Research and Public Health. MDPI. 2021;**18**:2-37
- [19] Greben HA, Maree JP, Eloff E, Murray K. Improved sulphate removal rates at increased sulphide concentration in the sulphidogenic bioreactor. 2005. Available from: <http://www.wrc.org.za>
- [20] Singh R, Kumar A, Kirrolia A, Kumar R, Yadav N, Bishnoi NR, et al. Removal of sulphate, COD and Cr (VI) in simulated and real wastewater by sulphate reducing bacteria enrichment in small bioreactor and FTIR study. Bioresource Technology. 2011;**102**(2):677-682
- [21] Hoffmann H, Platzer C, Winker M, von Muench E. Technology Review of Constructed Wetlands Subsurface flow constructed wetlands for greywater and domestic wastewater treatment. Deutsche Gesellschaft fü. 2011;**1**:27-35
- [22] O'sullivan AD, Murray DA, Otte ML. Removal of Sulfate, Zinc, and Lead from Alkaline Mine Wastewater Using Pilot-scale Surface-Flow Wetlands at Tara Mines, Ireland. Vol. 23. 2004
- [23] Conrow Edmund H. Estimating Technology Readiness Level Coefficients. Spacecraft and Rockets. 2011;48
- [24] Bódalo A, Gómez J-L, Gómez E, Gerardo León MT. Reduction of sulphate content in aqueous solutions by reverse osmosis using cellulose acetate membranes. Desalination. 2004;**162**:55-60
- [25] Naghizadeh A, Ghasemi F, Derakhshani E, Shahabi H. Thermodynamic, kinetic and isotherm studies of sulfate removal from aqueous solutions by graphene and graphite nanoparticles. Desalination and Water Treatment. 2017;**80**:247-254
- [26] Rahmati M, Yeganeh G, Esmaeili H. Sulfate ion removal from water using activated carbon powder prepared by Ziziphus spina-christi lotus leaf. Acta Chimica Slovenica. 2019;**66**(4): 888-898
- [27] Hong S, Cannon FS, Hou P, Byrne T, Nieto-Delgado C. Sulfate removal from acid mine drainage using polypyrrole-grafted granular activated carbon. Carbon. 2014;**73**:51-60
- [28] Vahid Babae HE. Removal of Sulfate from Water Using Polyaniline and Its Composites. Asian Journal of Research in Chemistry. 2011;**4**:827-833
- [29] Salimi AH, Mousavi SF, Farzin S. Removal of sulfate from Gamasiab river water samples by using natural nano-Clinoptilolite. Journal of Applied Research in Water and Wastewater [Internet]. 2019;**6**(1):39-44
- [30] Rumjit NP, Samsudin NA, Low FW, Thomas P, Lai CW, Velayudhaperumal Chellam P, et al. Kinetic and isotherm studies on adsorptive removal of sulfates by cotton shell derived biochar: Recovery of sulfates from marcasite soil. Sustainable Chemistry and Pharmacy. 2021;**20**:100361
- [31] Ao H, Cao W, Hong Y, Wu J, Wei L. Adsorption of sulfate ion from water by zirconium oxide-modified biochar derived from pomelo peel. Science of the Total Environment. 2020;**708**:135092

[32] Tejada-Tovar C, Villabona-Ortíz Á, Gonzalez-Delgado AD, Herrera A, Viera De la Voz A. Efficient sulfate adsorption on modified adsorbents prepared from *Zea mays* stems. *Applied Sciences*. 2021;**11**(4):1596

[33] Reinsel MA. A new process for sulfate removal from industrial waters. *Journal American Society of Mining and Reclamation*. 1999;**1999**(1):546-550

[34] McCauley CA, O'Sullivan AD, Milke MW, Weber PA, Trumm DA. Sulfate and metal removal in bioreactors treating acid mine drainage dominated with iron and aluminum. *Water Research*. 2009;**43**(4):961-970

[35] JHUESA Water Technology. 2022. Available from: <https://jhuesa.com/tecnologias/osmosis-inversa>

[36] Hernández-García F. Combinación de Electrodialisis, Intercambio Iónico y Ósmosis Inversa para la Desnitrificación de Aguas Potencialmente Potables. 2019;120. Available from: https://cideteq.repositorioinstitucional.mx/jspui/bitstream/1021/217/1/Combinacióndeelectrodialisis%2Cintercambioiónicoyósmosisinversaparadesnitrificacióndeaguaspotencialmentepotables_rees.pdf

Weibull Modeling of Microorganisms' Survival Kinetics by High Intensity Light Pulses on Different Food Systems

*Abril E. García-Santiesteban, Enrique Palou
and María Teresa Jiménez-Munguía*

Abstract

The practical application of any non-thermal technology (NTT) including high intensity light pulses (HILP) requires detailed basic studies on various microbial reduction effects and target microbial reduction parameters such as the inactivation rate (k), decimal reduction time (D) and temperature resistance coefficient (z). Particularly, for HILP technology such parameters have not yet been widely reported for different types of food, mostly because in NTT inactivation curves display non-linear concave shapes and do not follow first order kinetics. A simple alternative to the use of linear models is the Weibull distribution model. This work presents the Weibull modeling of the inactivation by HILP of *Pseudomonas fluorescens*, *Listeria innocua*, *Aspergillus niger*, and *Aspergillus flavus* in different food systems, previously reported by other authors. A D-value like parameter (D_v) is also calculated in aims to contribute to the industrial development of HILP technology. It was found that, in most cases, the Weibull distribution model represents a better alternative to linear inactivation models; and that it can give information related to the biological factors affecting the inactivation efficiency of HILP.

Keywords: HILP, microorganism inactivation, Weibull model, food safety

1. Introduction

One of the most important aspects to consider as part of quality in food processing is food safety. The ability to produce food free from pathogens is critical in the food processing industry [1, 2]. Different food technologies have been studied with the aim to inhibit or inactivate deteriorative microorganisms and pathogens. Thermal processing is highly effective and commonly used to eliminate or reduce pathogens; however it is not always suitable to use since it can induce chemical and physical changes leading to the degradation of nutritional content and organoleptic properties [2–4]. This has led researchers to investigate nonthermal technologies (NTT), or emerging technologies in food processing, such as High Intensity Light Pulses (HILP).

HILP treatments apply short light pulses of broad spectrum (200–1100 nm) that are rich in UVC light to decontaminate food and surfaces in short periods of time [5, 6]. Its main inactivation mechanism is DNA denaturation and dimer formation, preventing bacterial cells to replicate and eventually causing cell death. There is also a photothermal effect due to the mild increase of temperature and a photophysical effect due to microorganisms' cell membrane damage and consequent elution of proteins [3, 5].

The effects of HILP on inactivating different microorganisms have been studied on solid foods such as vegetables [7, 8], fruits [9], seeds [3, 10], meat, and fish [11]. The wide range of the log reduction on microbial population reported in different studies show that it depends on: processing parameters (number of pulses, discharge voltage and distance from the light source), food chemical composition, food type (liquid, semisolid or solid), and resistance of the microorganisms being inactivated [5].

Thermal treatments in food processing require detailed information on target microbial reduction parameters such as inactivation rate constant (k ; min^{-1}) for first order kinetics, decimal reduction time (D ; min), and temperature resistance coefficient (z ; °C) [1, 12]. In thermal processing, D is the time required to inactivate 90% of the initial population at a given temperature, while z is the increase in temperature required to achieve a logarithmic reduction. Both values, in thermal processing, can be found in the literature for different food-microorganism systems. Nevertheless, the practical application of any NTT including HILP, requires detailed basic studies on various microbial reduction effects [1]. Particularly, for HILP technology such parameters have not yet been widely reported for different types of food, in part because in NTT, inactivation curves do not follow first order kinetics.

Most inactivation curves in NTT display non-linear concave shapes and therefore have low values of correlation coefficient (R^2) when modeling with first order kinetic equations. The reason for this is that in complex biochemical systems like foods, the microbial population and processing conditions lack uniformity which is required for a better fit of a linear kinetic model [12].

A simple reliable alternative to linear models is the cumulative form of the Weibull distribution, simply known as the Weibull model in food microbiology [1, 12]. This model has been extensively applied to model lifetime data in medical, biological, and engineering sciences [13]. The Weibull survival function predicts the survival ratio of a microbial population as the result of the cumulative exposure of individual microorganisms with different resistances to a given lethal agent [12, 14]. On this model, parameter b is equivalent to the inactivation rate constant, k (min^{-1}), while the exponent n indicates the shape of the curve.

Forcing the fit of NTT inactivation curves, including those of HILP, to linear models corresponding to first order kinetics produces a lack of accuracy and might result in either consumers' health risk due to under-processing or in undesired losses in nutritional and sensorial quality of foods due to over-processing [14]. Therefore, the use of process design parameters that are equivalent to D and z -value, derived from non-linear kinetics, such as the Weibull model, might contribute substantially to NTT development and guidelines for their industrial implementation [12].

But, although there have been several studies offering information on the microbial reduction on foods by HILP, further mathematical modeling is missing in most cases to describe the inactivation effects of this technology and the factors affecting it.

For that reason, the aim of this work is to present the Weibull modeling of different food-microorganism systems previously investigated by other authors; and to analyze the obtained information to further understand HILP inactivation and give

useful information in the form of *t* *D*-value thermal processing like parameters for the implementation of this technology at industrial scale.

2. Methodology

2.1 Data obtention, selection and homogenization

The data used for the present study was obtained from a search of different previously published information by other authors. A total of 11 studies were pre-selected and a final group of only 4 studies (8 data sets) was used to obtain data for the food systems analyzed (**Table 1**). The selection was based on the type of food (solid foods), number of points in the curves (at least six points), number of repetitions (at least three repetitions with mean and standard deviation) and inactivation as a function of Total Dose (*T.D.*, J cm⁻²).

In some cases, the inactivation curves were given as a function of treatment times; therefore, to obtain the total dose (*TD*), Eq. (1) was used [18]. Where *e*₁ (Jcm⁻²) is the energy of a single pulse, *t* (s) is the treatment time and *f* (Hz) is the frequency of the light pulses.

$$TD = e_1 \cdot t \cdot f \tag{1}$$

2.2 Data modeling using the Weibull model

The logarithmic reduction values (log (*N/N*₀)) were obtained from the interpretation of the inactivation curves reported on each article, where *N* is the survival population after HILP treatment and *N*₀ is the initial number of microorganisms.

The inactivation of microorganisms was fitted using the Weibull model (Eq. (2)), where log (*N/N*₀) is the survival ratio after HILP treatment, *TD* is the total dose (Jcm⁻²) required to achieve a certain amount of microbial log reduction, *b* (min⁻¹) is equivalent

Food	Microorganism	Reference
Mozzarella cheese	<i>Pseudomonas fluorescens</i>	[15]
	<i>Enterobacteriaceae</i>	
Spinach surface	<i>Listeria innocua</i>	[13]
	<i>Escherichia coli</i>	
Low moisture malting barley	<i>Aspergillus carbonarius</i>	[16]
	<i>Aspergillus flavus</i>	
Cheddar cheese	<i>Escherichia coli</i> ATCC 25922	[17]
	<i>Pseudomonas fluorescens</i>	
	<i>Listeria innocua</i>	
White american singles	<i>Escherichia coli</i> ATCC 25922	
	<i>Pseudomonas fluorescens</i>	
	<i>Listeria innocua</i>	

Table 1.
Summary of food data sources and corresponding authors, used for the HILP modeling analysis.

to the inactivation rate constant while the exponent n indicates the shape of the curve ($n < 1$, upward-concave curve; $n > 1$, downward-concave curve) [12, 19, 20].

$$\log \left(\frac{N}{N_0} \right) = -bTD^n \tag{2}$$

2.3 D_v values calculation for HILP in food systems

The D_v value represents the equivalent of D-value in thermal processes [20]. In this study, D_v is defined as the total fluence (Jcm^{-2}) required to inactivate 90% of the viable cells [1]. In other words, it is the energy required to achieve the reduction of one logarithmic cycle. The D_v value was calculated from the substitution of $\log (N/N_0) = -1$ and the obtained b and n values from Eq. (2) for each set of data.

3. Results and discussion

Eight different data sets of inactivation kinetics, from the collected data, were organized to compare 4 different microorganisms and 5 different types of food. The data sets were chosen to study their inactivation curves by Weibull modeling. The inactivation rate constant (b , min^{-1}) and the shape of the curve parameter n , as well as the theoretical decimal reduction doses (D_v , Jcm^{-2}) for one logarithmic reduction, are presented in **Table 2**.

In all cases, $n < 1$ indicates upward concave curves as can be observed in the corresponding figures. Even though the Weibull model is of empirical nature, some associations with biological effects can be made. For instance, some authors have pointed out that $n < 1$ indicates that the remaining cells have the ability to adapt to the applied stress. This indicates that the process is becoming progressively less efficient [21]. Also, the fact that all of the obtained n values are different from 1, indicates that for these systems, inactivation does not follow a classical first order kinetic and that, therefore a linear model is not adequate for these data sets [19].

For the analyzed systems, the inactivation rate, b , shows no significative difference in relation to the microorganism, nor to the type of food. However, for the calculated D_v value of *P. fluorescens* in mozzarella cheese, there can be observed a deviation from the

Microorganism	Sample	b (min^{-1})	n	D_v (J cm^{-2})
<i>P. fluorescens</i>	Mozzarella cheese surface [15]	0.65811	0.10565	52.45
	Cheddar cheese [17]	0.90888	0.44733	1.24
	White American singles [17]	1.02013	0.38366	0.95
<i>L. innocua</i>	Spinach surface [13]	1.79700	0.10197	0.0032
	Cheddar cheese [17]	1.17223	0.36229	0.645
	White American singles [17]	1.2430	0.3290	0.516
<i>A. carbonarius</i>	Low moisture malting barley grains [16]	0.81668	0.12950	4.77
<i>A. flavus</i>		1.05498	0.12441	0.65

Table 2.
Values of b and n parameters, and lethal dose value (D_v) obtained by Weibull modeling inactivation kinetics data sets.

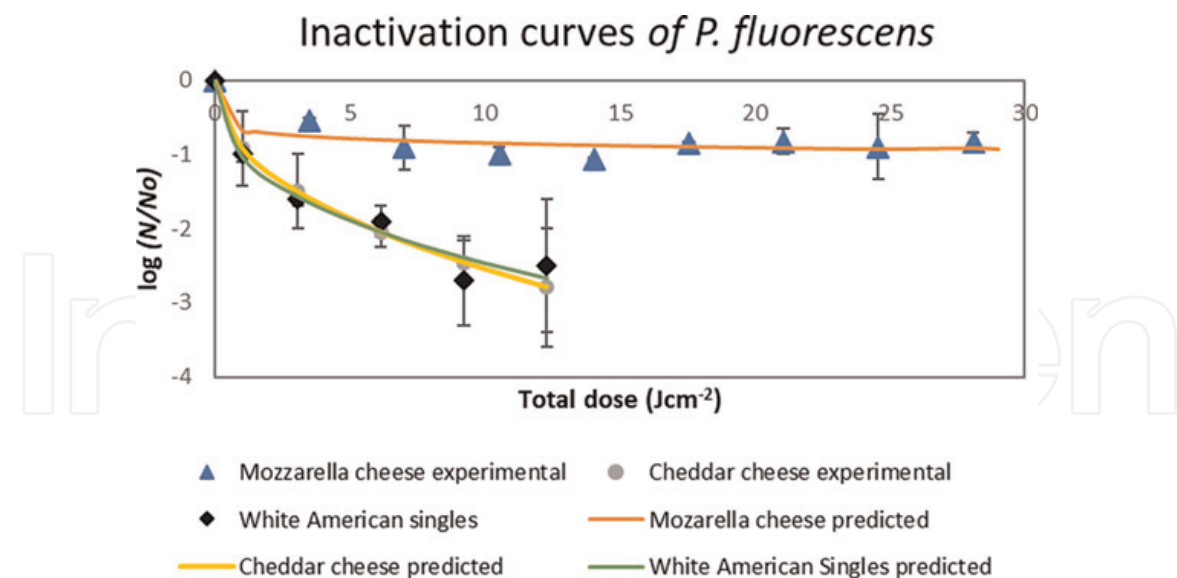


Figure 1.
Inactivation curves of *P. fluorescens* on: (a) mozzarella cheese (experimental (▲) and predicted (—) values) [15]; (b) Cheddar cheese (experimental (●) and predicted (—) values) and (c) white American singles (experimental (◆) and predicted (—) values) [17]. Note: Experimental values are presented as mean and s.d.

obtained experimental data. For instance, when calculating D_v for the inactivation of *P. fluorescens* in mozzarella cheese (Figure 1) [15], the obtained value is almost five times the experimental value ($\sim 10.53 \text{ J cm}^{-2}$) applied to achieve the inactivation of one logarithmic cycle of *P. fluorescens*. In this case, the experimental data shows that most of the inactivated population was achieved at a small dose (equivalent to a 2 s treatment [15]) and that after this time the final population remained approximately constant (Figure 1). Therefore, the values of n and b are small and D_v calculation is inaccurate in comparison to the experimental values. This supports the relevance of choosing an adequate model that fits the data accurately. This means that even when the Weibull model has been extensively applied, this particular system might not be a good fit [13].

In cheddar cheese and American white singles (Figure 1), it is observed that the total inactivation as well as the inactivation rate, b , are higher in comparison to the ones obtained for mozzarella cheese. These results are due to the lower D_v values calculated for cheddar and American white singles cheese (1.24 and 0.95 J cm^{-2} respectively). This inactivation kinetics show that after one log of the microorganisms' population of inactivation was achieved on cheddar and white American cheese, the inactivation rate decreased gradually as the treatment time increased. This difference in behavior can be attributed to the higher b and n values obtained for both kinetics in comparison to the values obtained for mozzarella cheese. For instance, lower inactivation levels on the same microorganism in different food could be attributed to the non-effective access of the radiation due to the irregular surfaces. The surface of some cheeses might be rough and have pores or crevices where microorganisms are able to hide from the light which is known as the 'shadow effect' [15, 22]. The 'shadow effect' might be caused as well by the microbial cell distribution over the food surface. It is known that at higher initial contamination levels, the inactivation efficiency of pulsed light decreases, given the fact that dead microorganisms might create a protective layer to other microbial cells; making it difficult for the light beams to reach these sheltered cells [22]. Other surface characteristics affect the inactivation efficacy, such as the inoculation of microorganisms in very smooth surfaces, as the mozzarella cheese surface, which might lead to cell clustering and therefore shadowing [1].

Differences in inactivation levels could also be related to the location of the microorganisms and the presence of protective substances in some cases. Therefore, the composition also plays an important role, given the fact that some components might have protective effects vs. other types of foods [13].

In addition, it is also important to consider other possible sources of variation in the inactivation kinetics such as errors during the experimental assay. It is noticeable that in mozzarella cheese (Figure 1) after a dose equivalent to $\sim 14.04 \text{ J cm}^{-2}$, total inactivation decreases as the standard deviation seems to increase. Therefore, it is important to consider the repeatability of each experiment in an aim to reduce non-precise data. Moreover, environmental conditions, such as pH, temperature, salt or sugar concentration, water activity, among others, should be carefully controlled since they play an important role in growth kinetics [14].

The inactivation of *L. innocua* on spinach leaves, cheddar cheese and American singles cheese, is presented in Figure 2; where it is shown that, at similar irradiation doses ($12\text{--}13 \text{ J cm}^{-2}$), slightly higher total inactivation levels were achieved on cheddar cheese and white American singles (Figure 2, black triangles, and green rhomboids, respectively) than the one achieved on spinach surface (Figure 2, blue dots). However, for this microorganism, D_v in spinach surface is significantly lower than in cheddar cheese and American white singles. This is explained by the fact that, as it is observed in Figure 2, a rapid inactivation occurs at very low energy levels, achieving an inactivation of ~ 1.5 logarithmic cycles at fluences lower than 1 J cm^{-2} . After this point, the inactivation rate decays as the slope for this portion of the graph tends to 0 with higher energy doses, achieving a final inactivation of ~ 2.4 log cycles. This indicates that the increase in the applied energy did not produce proportional increases in the total inactivation of the microorganism [13].

In Figure 2, it can be observed that *L. innocua* inactivation kinetic, shows slightly different behavior for cheddar and white American cheese in the sense that initial

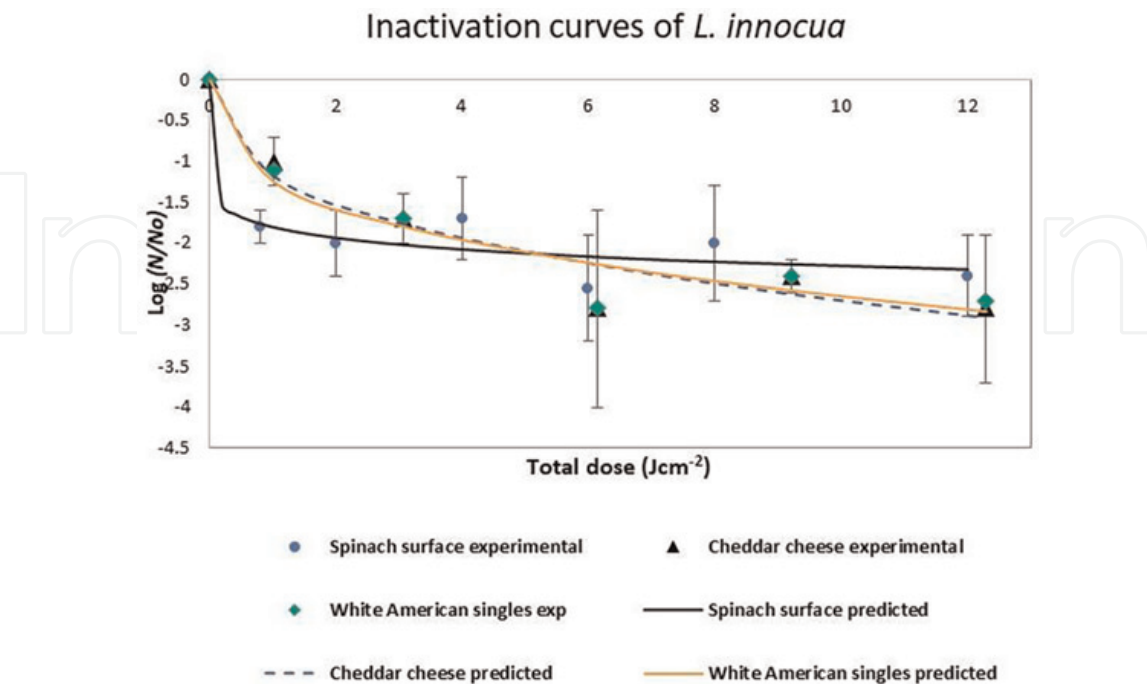


Figure 2. *Listeria innocua* inactivation curves on: a) spinach surface (experimental (●) and predicted (—) values) [13]; (b) cheddar cheese (experimental (▲) and predicted (---) values) and (c) white American singles (experimental (◆) and predicted (—) values) [17]. Note: Experimental values are presented as mean and s.d.

microbial inactivation occurs at higher inactivation doses, observing not only calculated D_v values higher than that obtained for spinach but also significantly higher n values, related to the less pronounced upward concavity of the curves 2.b and 2.c in comparison to 2.a. This behavior indicates that the rate of inactivation decreases as fluence increases affecting also the calculated D_v value [13, 19]. In fact, this behavior is noticed in all the systems analyzed in the present study, indicating that for pulsed light inactivation, linear models are not accurate. Instead, non-linear models, such as the Weibull model, should be applied [19].

In comparison to bacteria, fungi show generally higher resistance to light pulses [23]. This is shown in **Figure 3**, where it can be observed that higher energy doses were used to achieve inactivation levels of ~ 1.4 and ~ 1.75 log cycles for *A. carbonarius* and *A. flavus*, respectively, than the energy doses used for *P. fluorescens* and *L. innocua*. **Figure 3** shows a similar inactivation behavior for *A. carbonarius* and *A. flavus*, where a rapid inactivation is achieved at very low doses (~ 4 and ~ 3 J cm for *A. carbonarius* and *A. flavus*, respectively). The inactivation rate, b , decreases as the energy doses increases, achieving a tailing effect indicating that conidia inactivation rate dramatically fell down after energy doses equivalent to 18 J cm^{-2} [16]. However, parameter b is significantly higher for *A. flavus* than for *A. carbonarius* because the initial microbial population decay is higher in *A. flavus*. This behavior is also related to the small n values (~ 0.12) corresponding to a pronounced upward concave curve and might be the cause of the significantly higher D_v value obtained for *A. carbonarius* in comparison to the D_v value obtained for *A. flavus* since the inactivation rate for the latter one is significantly higher, as mentioned before.

The persistent residual population, tailing effect, showed for *L. innocua* inactivation on spinach surface (**Figure 2A**) and for *A. carbonarius* and *A. flavus* inactivation on malting barley (**Figure 3**) might be attributed to factors such as genotypic differences within the population providing enhanced survival in a minority of cells, the microorganisms were not inoculated homogeneously and were not evenly distributed on the surface or the light treatments were not applied homogeneously, due to the

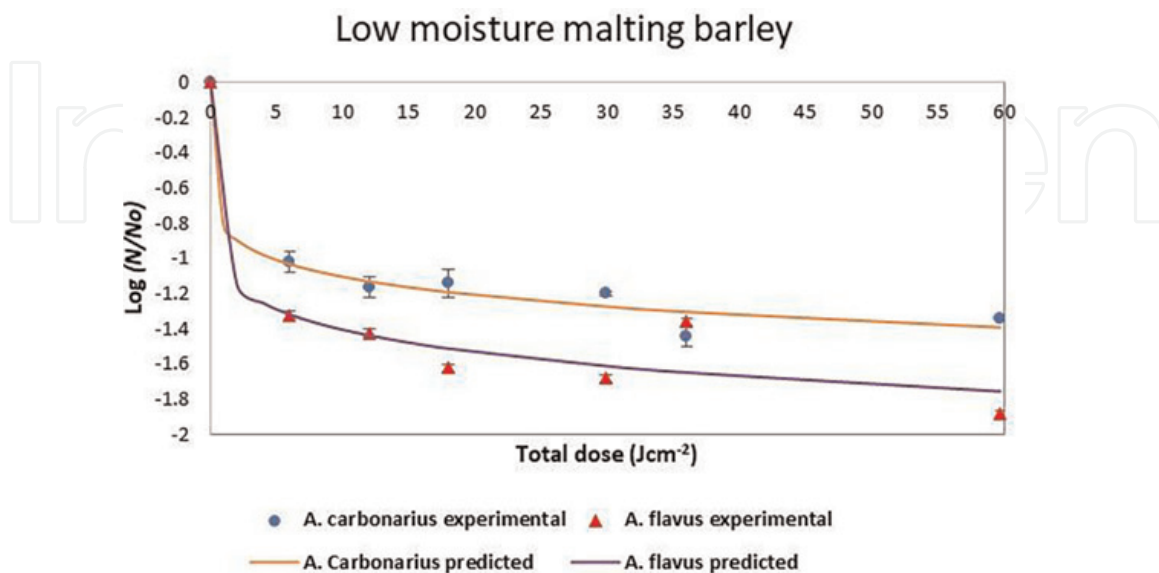


Figure 3. Inactivation of (a) *A. carbonarius* (experimental (●) and predicted (—) values) and (b) *A. flavus* (experimental (▲) and predicted (—) values) in low moisture malting barley [16]. Note: Experimental values are presented as mean and s.d.

shape of the food causing the sample to receive the light only on one side (which is likely the case for barley seeds). It should also be taken into account that some microorganisms might grow to deepen into the food tissue, thus being inaccessible to light radiation [16, 24].

4. Conclusions

The analysis of the HILP inactivation curves of the different data sets demonstrated that the application of a Weibull model generally offers a better fit to experimental data than linear model kinetics. Nevertheless, there are some exceptions where neither a first order kinetic nor the Weibull model offers a good fit. On those cases, important deviations from the experimental data might be caused, leading to inaccurate calculation of process parameters, like D_v parameter. That is the case of *P. fluorescens* inactivation on mozzarella cheese, where the calculates D_v was five times higher than the experimental value.

One advantage of the Weibull model is that despite its empirical nature, its parameters can be correlated to biological phenomena. For instance, the curve shape parameter n , that showed in all the kinetics analyzed, the efficiency of the treatment was reduced as the total energy increased, due to different biological factors.

In conclusion, the Weibull model represents a better alternative, in comparison to first order kinetics models, to describe the inactivation of microorganism by HILP. It also provides further information about the biological phenomena occurring during HILP inactivation process and can be used to calculate other process parameters such as the D_v or decimal dose.

Acknowledgements

This work was supported by Universidad de las Américas Puebla (UDLAP) and Consejo Nacional de Ciencia y Tecnología (CONACYT). Author García-Santiesteban acknowledges the financial support for her PhD studies in Food Science given by UDLAP and CONACYT.

Conflict of interest

The authors of the present work declare that there are not known conflict of interests.

Nomenclature list

b	inactivation rate in Weibull model, given in min^{-1}
D_v value	total fluence (Jcm^{-2}) required to inactivate 90% of the viable cells
HILP	high intensity light pulses
k	inactivation rate in first order kinetics, given in min^{-1}
n	inactivation curve shape parameter in Weibull model
NTT	non thermal technologies

IntechOpen


IntechOpen

Author details

Abril E. García-Santiesteban*, Enrique Palou and María Teresa Jiménez-Munguía
Department of Chemistry, Environmental and Food Engineering, Universidad de las
Américas Puebla, Puebla, Mexico

*Address all correspondence to: abril.garciasn@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Hwang H-J, Seo J-H, Jeong C, Cheigh C-I, Chung M-S. Analysis of bacterial inactivation by intense pulsed light using a double-Weibull survival model. *Innovative Food Science and Emerging Technology*. 2019;**56**:102185
- [2] Yang Y, Meier F, Ann Lo J, Yuan W, Lee Pei Sze V, Chung H-J, et al. Overview of recent events in the microbiological safety of sprouts and new intervention technologies: Sprout safety and control. *Comprehensive Reviews in Food Science and Food Safety*. 2013;**12**(3):265-280
- [3] Kim W-I, Choi SY, Han I, Cho SK, Lee Y, Kim S, et al. Inhibition of *Salmonella enterica* growth by competitive exclusion during early alfalfa sprout development using a seed-dwelling *Erwinia persicina* strain EUS78. *International Journal of Food Microbiology*. 2020;**312**:108374
- [4] Hb O, Song KY, Joung KY, Shin SY, Kim YS. Effects of chia (*Salvia hispanica* L.) seed roasting conditions on quality of cookies. *Italian Journal of Food Science*. 2018;**26**:54-66
- [5] Hwang H-J, Cheigh C-I, Chung M-S. Construction of a pilot-scale continuous-flow intense pulsed light system and its efficacy in sterilizing sesame seeds. *Innovative Food Science and Emerging Technolog*. 2017;**39**:1-6
- [6] Rowan NJ. Pulsed light as an emerging technology to cause disruption for food and adjacent industries – Quo vadis? *Trends in Food Science and Technology*. 2019;**88**:316-332
- [7] Kaack K, Lyager B. Treatment of slices from carrot (*Daucus carota*) using high intensity white pulsed light. *European Food Research Technology*. 2007;**224**:561-566
- [8] Huang R, Chen H. Evaluation of inactivating *Salmonella* on iceberg lettuce shreds with washing process in combination with pulsed light, ultrasound and chlorine. *International Journal of Food Microbiology*. 2018;**285**: 144-151
- [9] Huang Y, Chen H. Inactivation of *Escherichia coli* O157:H7, *Salmonella* and human norovirus surrogate on artificially contaminated strawberries and raspberries by water-assisted pulsed light treatment. *Food Research International*. 2015;**72**:1-7
- [10] Aron Maftai N, Ramos-Villaruel AY, Nicolau AI, Martín-Belloso O, Soliva-Fortuny R. Pulsed light inactivation of naturally occurring moulds on wheat grain: Pulsed light inactivation of naturally occurring moulds on wheat grain. *Journal of Science Food and Agriculture*. 2014;**94**:721-726
- [11] Ozer NP, Demirci A. Inactivation of *Escherichia coli* O157:H7 and *Listeria monocytogenes* inoculated on raw salmon fillets by pulsed UV-light treatment. *International Journal of Food Science and Technology*. 2006;**41**(4): 354-360
- [12] Serment-Moreno V. Microbial modeling needs for the nonthermal processing of foods. *Food Engineering Review*. 2020;**13**:465-489
- [13] Agüero MV, Jagus RJ, Martín-Belloso O, Soliva-Fortuny R. Surface decontamination of spinach by intense pulsed light treatments: Impact on quality attributes. *Postharvest Biology and Technology*. 2016;**121**:118-125

- [14] Peleg M, Corradini MG. Microbial growth curves: What the models tell us and what they cannot. *Critical Review in Food Science and Nutrition*. 2011;**51**(10): 917-945
- [15] Lacivita V, Conte A, Lyng JG, Arroyo C, Zambrini VA, Del Nobile MA. High intensity light pulses to reduce microbial load in fresh cheese. *Journal of Dairy Research*. 2018;**85**(2):232-237
- [16] Zenklusen MH, Coronel MB, Castro MÁ, Alzamora SM, González HHL. Inactivation of *Aspergillus carbonarius* and *Aspergillus flavus* in malting barley by pulsed light and impact on germination capacity and microstructure. *Innovative Food Science Emerging Technology*. 2018;**45**:161-168
- [17] Proulx J, Sullivan G, Marostegan LF, VanWees S, Hsu LC, Moraru CI. Pulsed light and antimicrobial combination treatments for surface decontamination of cheese: Favorable and antagonistic effects. *Journal of Dairy Science*. 2017; **100**(3):1664-1673
- [18] Luksiene Z, Gudelis V, Buchovec I, Raudeliuniene J. Advanced high-power pulsed light device to decontaminate food from pathogens: Effects on *Salmonella typhimurium* viability in vitro: In vitro decontamination of food using high-power pulsed light device. *Journal of Applied Microbiology*. 2007; **103**(5):1545-1552
- [19] van Boekel M. On the use of the Weibull model to describe thermal inactivation of microbial vegetative cells. *International Journal of Food Microbiology*. 2002;**74**(1-2):139-159
- [20] John D, Ramaswamy HS. Comparison of pulsed light inactivation kinetics and modeling of *Escherichia coli* (ATCC-29055), *Clostridium sporogenes* (ATCC-7955) and *Geobacillus stearothermophilus* (ATCC-10149). *Current Research Food Science*. 2020;**3**: 82-91
- [21] Izquier A, Gómez-López VM. Modeling the pulsed light inactivation of microorganisms naturally occurring on vegetable substrates. *Food Microbiology*. 2011;**28**(6):1170-1174
- [22] Cudemos E, Izquier A, Medina-Martínez MS, Gómez-López VM. Effects of shading and growth phase on the microbial inactivation by pulsed light. *Czech Journal of Food Science*. 2013; **31**(2):189-193
- [23] Heinrich V, Zunabovic M, Varzakas T, Bergmair J, Kneifel W. Pulsed light treatment of different food types with a special focus on meat: A critical review. *Critical Review Food Science and Nutrition*. 2016;**56**(4): 591-613
- [24] Alzamora SM, Guerrero SN, Raffelini S, Ferrario M, Schenk M. Hurdle technology in food processing. In: Pareek S, editor. *Fresh-cut Fruits and Vegetables: Technology, Physiology and Safety*. Boca Raton, FL: CRC Press; 2016. pp. 101-138

White Maize Tortillas Fortified with Brown Algae *Macrocystis pyrifera*

Alexa Pérez-Alva, Diana K. Baigts-Allende
and Milena M. Ramírez-Rodrigues

Abstract

This study aimed to analyze the effect of incorporating different concentrations (10, 20 and 30% (w/w)) of the brown seaweed *Macrocystis pyrifera* into white maize tortillas in terms of chemical composition, color, texture, phenolic compounds, and antioxidant capacity. The moisture and lipid content decreased with the addition of seaweed, while the ash, protein and fiber content increased. The changes in carbohydrates were not significant ($p > 0.05$). The concentration of Na, K, Mg, and P increased significantly ($p < 0.05$) with the incorporation of the alga. Higher percentage of seaweed led to *masas* and tortillas with greener tones as it affected L^* , a^* , b^* parameters, hue and chroma and to bigger differences between the control samples and the samples with seaweed. Total phenolic content increased linearly with the incorporation of seaweed, as well as FRAP values. While the *masa* with 20% of seaweed showed the highest percentage of ABTS inhibition (86%). Finally, the control sample of *masa* presented the lowest values of adhesiveness and hardness, while the control sample of tortilla showed the highest hardness value.

Keywords: seaweed, *Macrocystis pyrifera*, white maize, tortilla, mineral profile, texture profile analysis, antioxidant capacity

1. Introduction

Cereals, from the Latin word '*cerealis*' which means 'grain', comprise starchy grains from the grass family known as *Gramineae* [1]. They are a key component of the human diet and are the most widely used crop [2]. Some grains from this family are wheat, rice, maize, sorghum, millet, barley, and rye. Rice, wheat, and maize provide at least 30% of the food calories to more than 4.5 billion people in 94 developing countries [3]. Among them, maize (*Zea mays* L.) was the most produced crop in 2019 with 1,148,487,291 tons, compared to wheat and rice, with 765,769,635 and 755,473,800 tons, respectively [4]. Maize was domesticated in what is now Mexico approximately 7000 years ago and was spread across the world shortly after the European discovery of the Americas. Currently, the United States, Brazil, Mexico, Argentina, India, France, Indonesia, South Africa, and Italy produce 79% of the world's maize production [5].

In Mexico, maize is considered the most representative crop for its economic, social, and cultural importance [6, 7], with a production of 27,228,242 tons in 2019 [8], which places it as the eighth producer in the world [9]. Unlike the United States, where “yellow” corn is industrially produced to be used as animal feed or to make industrial corn derivatives ranging from sweeteners to plastics, in Mexico the vast majority of maize grown is “white” and it is destined for human consumption [10]. Most of the maize in Mexico is consumed as tortillas, being the basis of the daily diet [7].

There are different ways in which tortillas can be elaborated, traditionally nixtamalized maize is used [11]. The nixtamalization process consists on cooking the grains in an alkaline solution, followed by overnight steeping [6]. Afterwards, the maize is washed and stone ground to obtain a wet dough, known as *masa* [6, 12]. The *masa* can be used to prepare tortillas or to produce instant maize flour [12]. During this process, calcium is incorporated into the maize, the bioavailability of niacin (B3) improves, and mycotoxins’ presence is reduced [12, 13].

The nutritional value of maize is affected by the variety, environment, and sowing conditions [14], and can also be modified after nixtamalization [15]. In general, maize is considered to be a source of vitamins (B-complex, A, C, and K), β -carotene and selenium [2], but its protein content is considered to be of low nutritional quality as the concentration of lysine and tryptophan is low [14]. When compared to other cereals as wheat, oats, and rice, maize has a higher antioxidant capacity, attributed to a higher total phenolic content [13, 16]. The main component of tortillas are carbohydrates, mainly starch and fiber. While the protein, lipid and niacin (B3) content has been reported to increase after nixtamalization, phytochemicals like carotenoids, polyphenols, and anthocyanins can decrease [15].

Tortillas are an excellent vehicle to enhance nutrition options for maize consumers [17]. In fact, tortillas have been fortified with grasshopper (*Sphenarium purpuracens*) [18], nopal (*Opuntia ficus-indica*) [19], faba-bean (*Vicia faba*), white-bean (*Phaseolus vulgaris*), [14] and muicle (*Justicia spicigera* Schechtendal) [6], among others. With this additions, the mineral [18, 19], protein, [14, 18] and dietary fiber [14, 19] contents increased, as well as the antioxidant capacity [6].

In the last decades, seaweeds have gained attention as potential sources for the elaboration of food and feed products [20, 21]. Seaweeds are marine, photosynthetic algae that are prolific in every ocean. They are divided into three main classes, or phyla: *Phaeophyta* (brown algae), *Rhodophyta* (red algae), and *Chlorophyta* (green algae) [22]. Brown algae are a source of unsaturated fatty acids, alginate, biopolymers, and phlorotannins, which are polymeric structures of the monomer phloroglucinol [23]. *M. pyrifera* is a brown seaweed found in the Pacific Ocean, mainly used as feed for abalone and hydrocolloid production [20]. In this study, the seaweed *M. pyrifera* was used to partially substitute nixtamalized corn *masa* to elaborate functional tortillas, with an increased phenolic content and antioxidant activity. Therefore, the effect of incorporating different concentrations of seaweed into white maize tortillas in terms of chemical composition, color, texture, phenolic compounds, and antioxidant capacity was analyzed.

2. Materials and methods

2.1 Materials

White nixtamalized *masa* was purchased in a local store in Puebla, Mexico. The seaweed *M. pyrifera* was bought from ALGAS PACIFIC (Baja California, Mexico).

The reagents used for the extraction and quantification of the phenolic compounds and antioxidant activity were of analytic grade, and were purchased from Sigma-Aldrich Co (St. Louis, MO, USA).

2.2 Tortilla elaboration

The seaweed was first dried at 60°C for 4 h using a food dehydrator (Excalibur 3948CDB, Sacramento, CA, USA.), milled and then sieved using a 270 µm mesh. Next, the nixtamalized *masa* was divided into four parts and 10 (M10), 20 (M20) or 30 (M30) % (w/w) of seaweed powder was incorporated, while the remaining section was used as a control (M0). Then tortillas, T0, T10, T20 and T30 (depending on the concentration of seaweed) were produced. Each *masa* was divided into 30 g spherical portions, which were then pressed into flat discs and heated at 350°C for 1 min on one side, then turned over and heated for 30 s on the other side and finally 15 s on the initial side [6]. After analyzing color and texture, the remaining tortillas were cooled down and milled in order to obtain a homogeneous product that was frozen until further analysis.

2.3 Chemical composition

The chemical composition of the tortillas was determined following the official methods for moisture (925.09), ash (923.03), protein (954.01), lipids (920.39), and fiber (962.09) [24]. While carbohydrates were calculated by difference. The mineral content (Na, K, Ca, Mg) of the tortillas was determined by Inductively Coupled Plasma-Optical Emission Spectrometry (ICP-OES) based on AOAC Method 2011.4, as in Ref. [25]. The sample preparation consisted of a microwave digestion on a CEM Mars6 microwave system (Charlotte, NC, USA), and the determination was carried out in triplicate using an Agilent 5110 ICP-OES (Santa Clara, Ca, USA). The analytical measurements were made with the ICP Expert 7.5.1 software.

2.4 Color

The surface color was measured on randomly selected tortilla and *masa* spots and was measured in triplicate using a Konica Minolta model CR-400 colorimeter (Konica Minolta Holdings, Inc., Japan). The color was expressed as L*, a* and b* colorimetric coordinates in the CIELab scale, as in Ref. [26]. The hue angle (H), chroma (C) and change in color (ΔE) were calculated according to eqs. (1)–(3), respectively [27].

$$H = \tan^{-1}(b^* / a^*) \quad (1)$$

$$C = \sqrt{(a^*)^2 + (b^*)^2} \quad (2)$$

$$\Delta E = \sqrt{(\Delta L^2 + \Delta a^2 + \Delta b^2)} \quad (3)$$

2.5 Texture

The texture of tortillas was measured as a perforation test [28]. The test conditions were 50 mm distance from the puncture base, and the applied force was the necessary to cause the tortilla to break. The texture profile analysis (TPA) of *masas* was

done following the procedure described in [29] to determine hardness, adhesiveness, cohesiveness, and springiness. Both texture analysis were done in triplicate, using a texturometer Shimadzu EZ-SX Texture Analyzer (Shimadzu Co., Japan) and the software used for data analysis was Trapezium X (Shimadzu Co., Japan).

2.6 Total phenolic content (TPC) and antioxidant capacity

The extracts used for both TPC and antioxidant capacity were prepared by mixing 1 g of *masa* with 10 ml of distilled water and let to stand for 12 h. Next, samples were centrifuged at 1325 x g for 10 min. The three assays were measured using a UV-Vis spectrophotometer model Multiskan Sky Microplate (Thermo Fischer Scientific, USA). Total phenolic content (TPC) was measured using the Folin–Ciocalteu assay as in Ref. [30]. A gallic acid calibration curve (0–1000 mg/L) was prepared to calculate the TPC and the results were expressed as gallic acid equivalents (mg GAE/100 g of *masa*).

In order to quantify the antioxidant capacity, 2,2'-azinobis-(3-ethylbenzothiazoline-6-sulfonic acid (ABTS) was used. The methodology for ABTS was done following the method in [31] and results were reported as % of inhibition and were calculated using the following eq. (4).

$$\% \text{inhibition} = ((\text{initial absorbance} - \text{final absorbance}) / (\text{initial absorbance})) \times 100 \quad (4)$$

The methodology used for Ferric Reducing Antioxidant Power Assay (FRAP) was the one described in [32]. The results were expressed as μmol Trolox equivalents (TE)/100 g of *masa*.

2.7 Statistical analysis

All tests were performed in triplicate and the results had a coefficient of variation (CV) < 5%. The results were analyzed using one-way ANOVA ($\alpha = 0.05$) with Minitab 19 (Minitab, LLC, State College, PA, USA). The mean differences were determined using a Tukey's multi-comparison test ($\alpha = 0.05$).

3. Results and discussion

The addition of seaweed promoted a significant ($p < 0.05$) increase in ash, protein, and fiber, while the moisture and lipid decreased and no change was observed in the carbohydrates (**Table 1**). The ash content increased 10 times when comparing the control and the tortillas with 30% of seaweed, while the fiber almost triplicated when the same samples were compared (7 and 17%, respectively). The moisture values from this study were similar to the ones reported in Ref. [28], but lower than the ones reported in Ref. [6], where values ranging between 48 and 50% were obtained. However, the ash content reported in this study (4.63–11.60%) was higher than the content reported for tortillas fortified with different concentrations of *Jatropha curcas* L. flour [28] with values between 0.7 and 2%. It has been reported that the mineral fraction of seaweeds can reach up to 40% of dry matter, and in some cases can be higher than the ones of land plants and animal products [33]. Despite seaweeds not having a notoriously high protein content when compared to traditional protein sources [34], the addition of more seaweed resulted in a significant augment ($p < 0.05$) of protein content in tortillas. However, while the protein content of T0

	Moisture (%)	Ash (%)	Protein (%)	Lipids (%)	Carbohydrates (%)	Fiber (%)
T0	44.1 ± 0.17 ^a	1.16 ± 0.02 ^d	4.49 ± 0.02 ^d	0.370 ± 0.01 ^a	49.9 ± 0.20 ^a	6.77 ± 0.29 ^d
T10	40.5 ± 0.16 ^b	4.63 ± 0.04 ^c	4.75 ± 0.02 ^c	0.345 ± 0.01 ^b	49.8 ± 0.20 ^a	10.2 ± 0.44 ^c
T20	36.9 ± 0.15 ^c	8.11 ± 0.07 ^b	5.01 ± 0.02 ^b	0.327 ± 0.01 ^c	49.7 ± 0.18 ^a	13.4 ± 0.40 ^b
T30	33.3 ± 0.14 ^d	11.6 ± 0.11 ^a	5.27 ± 0.03 ^a	0.311 ± 0.01 ^d	49.6 ± 0.16 ^a	16.6 ± 0.38 ^a

Different letters in the same row represents significant differences ($p < 0.05$). Results are expressed as mean ($n = 3$) and standard deviation (mean ± SD). T10, T20, T30 = tortilla with 10, 20 or 30% of seaweed, T0 = control.

Table 1.
Chemical composition of tortillas with different concentrations of seaweed.

(4.49%) was similar to the control sample from reference [28] (4.65%), the addition of seaweed resulted in lower protein values (4.75–5.27%) than the addition of *J. curcas* flour (6.52–10.85%) [28]. In other, the protein content from hand-made white tortillas has been reported to be between 7.5 and 9.6% [35], the protein content of the tortillas from this study was lower. As seaweeds are not considered to have an elevated lipid content [33], it was not expected to observe a high increase in the lipid fraction. The lipid content in all tortillas from this study was lower than the one reported in other studies [28, 35], with lipid values ranging between 1.87 and 3.95%. The differences in the lipid content can be related to the variety of the used maize as well as to the nixtamalization conditions, as lipids from the kernel can be solubilized during the alkaline treatment or during the steeping time [35]. The fiber content of all samples was higher than the reported in [28] (0.8–1.98%) while the fiber content of T30 was similar to the one reported in Ref. [35] (15.72%).

With an increment in ash content, the increase in minerals would be expected, specially as seaweeds can contain between 10 and 20 times the mineral amount when compared to land plants since seaweeds gather them from the seawater [36]. All the analyzed minerals (Na, Ca, P, K, and Mg) increased significantly ($p < 0.05$) with an increment in seaweed (**Table 2**). The most notorious increase in minerals was Na, which increased almost 100 times when T0 and T30 were compared, followed by K, which augmented 16 times, Mg went from 85 to 380 mg/100 g, Ca was almost triplicated when comparing the same samples (T0 and T30), and P barely increased from 315 (T0) to 360 mg/100 g (T30), however it was also significant. The mineral composition of seaweeds is influenced by the environmental conditions, their age and their capability to absorb inorganic substances from the environment. Due to the polysaccharides in their cell walls, brown seaweeds (like *M. pyrifera*) present higher affinity to absorb calcium, magnesium, sodium, and potassium salts due to the presence of alginic acid, alginate, and alginic acid salt [36]. The Ca content of all samples was higher than the content of white maize tortillas (155.5 mg/100 g) [35]. The recommended dietary allowance (RDA) for women and men between 18 and 50 years old for Ca is 1000 mg [37], amount that could be covered with 130 g of T30. The RDA of P is 700 mg [37] value that could be covered with 200 g of T30, with 50–70 g of T30 covers the RDA of K (2600–3400 mg [37]) while the RDA of Mg (310–420 mg [37]) could be covered with 80–110 g of T30. While the Chronic Disease Risk Reduction Level (CDRR) of Na is 2300 [37], dose that is not overpassed with 200 g of T30, which is the amount of tortilla needed to cover all the minerals.

The increase in seaweed concentration in *masas* led to a significant increase in the differences (ΔE) between T0 and T10, T20, and T30. Additionally, the seaweed

	Sodium (Na)	Calcium (Ca)	Phosphorus (P)	Potassium (K)	Magnesium (Mg)
T0	10.4 ± 0.26 ^d	363 ± 3.11 ^d	315 ± 3.60 ^d	331 ± 0.11 ^d	84.7 ± 0.01 ^d
T10	335 ± 0.52 ^c	500 ± 2.91 ^c	330 ± 3.24 ^c	1852 ± 1.94 ^c	184 ± 0.70 ^c
T20	661 ± 0.94 ^b	636 ± 2.93 ^b	344 ± 2.88 ^b	3374 ± 3.87 ^b	284 ± 1.39 ^b
T30	986 ± 1.39 ^a	773 ± 3.19 ^a	359 ± 2.52 ^a	4895 ± 5.81 ^a	384 ± 2.09 ^a

All results are expressed as mg/100 g. Different letters in the same row represents significant differences ($p < 0.05$). Results are expressed as mean ($n = 3$) and standard deviation (mean ± SD). T10, T20, T30 = tortilla with 10, 20 or 30% of seaweed. T0 = control.

Table 2.
Mineral content of tortillas with different concentrations of seaweed.

	L*	a*	b*	Hue	Chroma	ΔE
M0	81.9 ± 0.46 ^a	−0.21 ± 0.01 ^a	23.9 ± 0.42 ^a	90.5 ± 0.01 ^d	23.9 ± 0.42 ^a	
M10	62.1 ± 0.92 ^b	−1.86 ± 0.02 ^b	19.8 ± 0.35 ^b	95.4 ± 0.13 ^c	19.9 ± 0.35 ^b	20.4 ± 0.91 ^c
M20	54.3 ± 1.04 ^c	−2.54 ± 0.03 ^c	19.5 ± 0.39 ^{b,c}	97.4 ± 0.14 ^b	19.6 ± 0.38 ^b	28.1 ± 0.99 ^b
M30	48.4 ± 0.90 ^d	−3.28 ± 0.05 ^d	19.2 ± 0.24 ^c	99.7 ± 0.22 ^a	19.5 ± 0.23 ^b	34.0 ± 0.87 ^a

Different letters in the same row represents significant differences ($p < 0.05$). Results are expressed as mean ($n = 3$) and standard deviation (mean ± SD). M10, M20, M30 = masa with 10, 20 or 30% of seaweed. M0 = control.

Table 3.
Color of masas with different concentrations of seaweed.

affected all the color parameters analyzed (**Table 3**) on *masas*. The lightness (L*) of T0 decreased to half its value in T30. The negative values of a* are indicative of greener tonalities [6], it could be seen that the increase in seaweed concentration led to an increase of this tonalities. The b* values also decreased when the concentration of algae was augmented, which means that the *masa* was less yellow and started to have more blue tonalities. On the contrary, the seaweed lead to a significant increase in hue values, changing from 90.5 to 99.7, while chroma was higher in the control sample and showed no differences between samples T10, T20 and T30. Similarly, the addition of muicle led to a variation in L*, a* and b* parameters [6]. The authors reported a decrease in lightness and in b* values, indicative that color tended to display increasingly blue tonalities, while an increase in a*, related to an increase in red tonalities, was also observed [6]. While the addition of *J. curcas* flour to *masa* led to increase in L* and a*, while there were no changes in b* [28].

The analyzed color parameters of tortillas were also significantly modified with the incorporation of seaweed (**Table 4**). Similarly, to the *masas*, there was a bigger difference (ΔE) between T0 and T30. Additionally, the lightness (L*) presented a similar behavior to the one observed in *masas*, with the value from T0 decreasing to almost half its value in T30. With the progressive incorporation of seaweed, the a* changed from positive values to negative ones, meaning that the samples became greener as a result of incorporation the algae. Despite the addition of seaweed, the values of b* remained positive, and while there was no significant difference between T20 and T30, this values were significantly smaller that T0 and T10. Chroma decreased with the increase of seaweed and went from 23.8 (T0) to 18.2 (T30), while T30 presented the highest hue values (95.4), followed by T20, and T0 showed the lowest values (84.3).

	L*	a*	b*	Hue	Chroma	ΔE
T0	70.6 ± 2.38 ^a	2.37 ± 0.08 ^a	23.6 ± 0.70 ^a	84.3 ± 0.19 ^d	23.8 ± 0.70 ^a	
T10	51.5 ± 1.29 ^b	-0.49 ± 0.02 ^b	20.2 ± 0.58 ^b	91.4 ± 0.02 ^c	20.2 ± 0.58 ^b	30.7 ± 1.33 ^b
T20	48.6 ± 1.60 ^c	-1.14 ± 0.02 ^c	18.6 ± 0.72 ^c	93.5 ± 0.14 ^b	18.7 ± 0.71 ^c	33.8 ± 1.62 ^a
T30	46.5 ± 1.67 ^c	-1.73 ± 0.05 ^d	18.1 ± 0.63 ^c	95.4 ± 0.22 ^a	18.2 ± 0.63 ^c	36.0 ± 1.68 ^a

Different letters in the same row represents significant differences ($p < 0.05$). Results are expressed as mean ($n = 3$) and standard deviation ($SD = \pm$). T10, 20, 30 = tortilla with 10, 20 or 30% of seaweed. TC = control.

Table 4.
Color of tortillas with different concentrations of seaweed.

However all the hue values could be described as yellowish-green. Tortillas fortified with white bean flour showed reddish color, with the value of a^* in the positive range, while b^* values had a clear tendency to yellow [14].

The antioxidant capacity and TPC of *masas* were significantly affected by the seaweed concentration (**Figures 1 and 2**). M0 presented the lowest values in the three assays, similarly to the effect of incorporating muicle extracts [6]. The highest values of both FRAP and TPC were obtained for M30. The TPC showed a linear trend with higher concentrations of seaweed. The TPC increased 5 times between M0 and M30, while the FRAP value from M30 increased more than twice when compared with M0. The TPC of all samples was higher than the value reported for hand-made tortillas from white maize (~75 mg GAE/100 g) [35]; while the TPC of T0 was similar to the one reported in [38] (~160 mg GAE/100 g). The highest percentage of inhibition of ABTS was obtained for M20 (90%), while the lowest value was 70% for M0. The incorporation of different varieties of beans to wheat tortillas also led to the increase in antioxidant capacity when compared to their control samples [39].

The texture of *masas* and tortillas was significantly affected by the incorporation of seaweed (**Figures 3 and 4**, respectively). M0 showed the lowest values for hardness and adhesiveness when compared to M10, M20 and M30. The adhesiveness and hardness of the *masa* showed a similar trend, the highest value was obtained for M10,

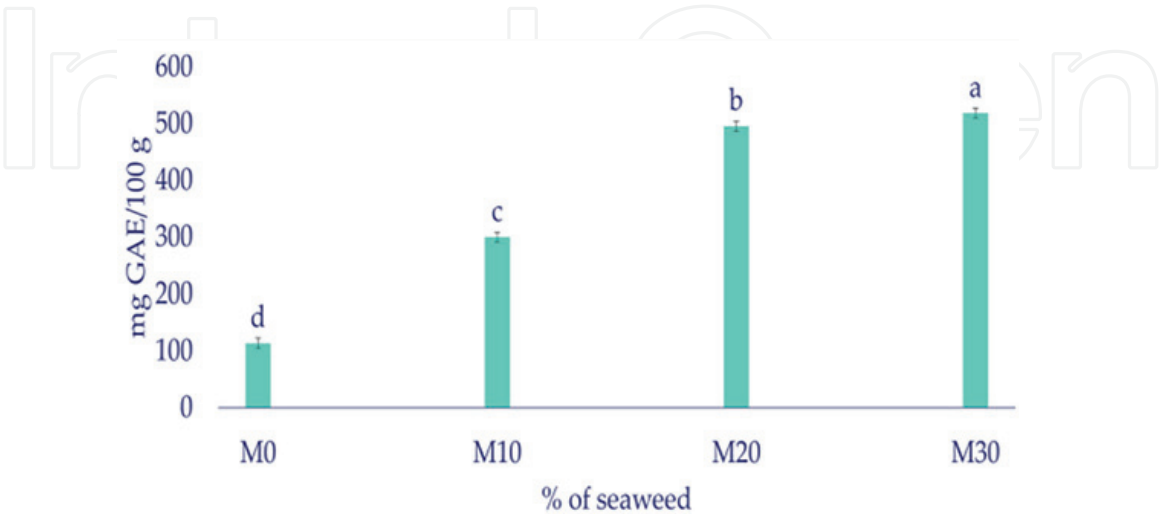


Figure 1.
Total phenolic content of masas with different concentrations of seaweed. Each value represent the results from mean and standard deviation. Different letters indicate significant differences.

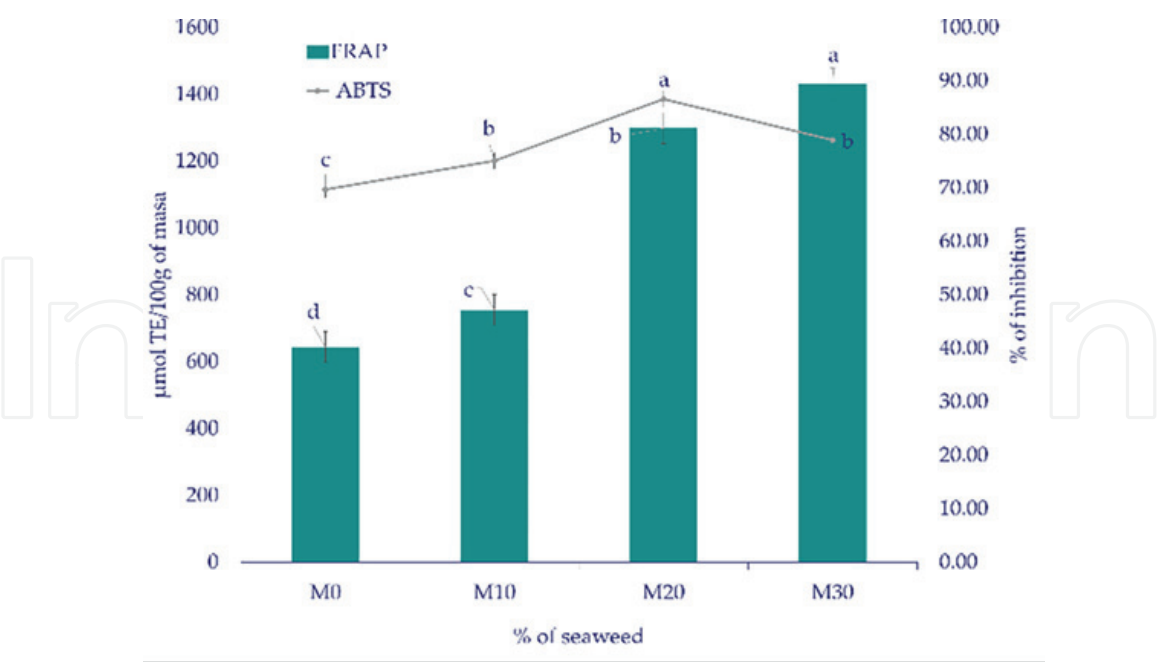


Figure 2. Antioxidant capacity (FRAP and ABTS) of masas with different concentrations of seaweed. Each value represent the results from mean and standard deviation. Different letters indicate significant differences.

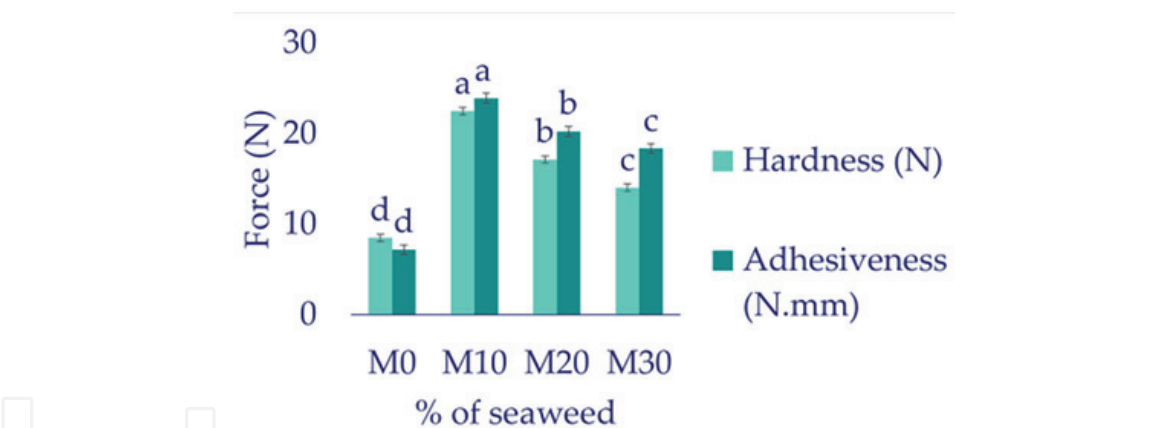


Figure 3. Texture profile analysis (TPA) of masas with different concentrations of seaweed. Each value represent the results from mean and standard deviation. Different letters indicate significant differences.

and decreased with a higher concentration of seaweed. There was not a linear trend depending on the concentration of seaweed, which was similar to the effect of incorporating grasshopper (*S. purpuracens*) flour. However, the adhesiveness of tortillas with seaweed was higher than the ones with grasshopper flour ($\sim 0.10 - \sim 0.20$ N) [18]. In the same study, the control sample showed higher hardness than *masa* with grasshopper flour [18], while in this study M0 showed the lowest values for both adhesiveness and hardness. The opposite behavior was observed in tortillas, were T0 required more applied force to puncture the tortilla. In the tortillas this trend was not observed, as T10 required the lowest force to get punctured while T20 and T30 showed no significant difference among them. Different concentrations of muicle did not affect the hardness of tortillas, which was lower (2.48–2.81 N) [6] than the one reported in this study (3.5–7.5 N).

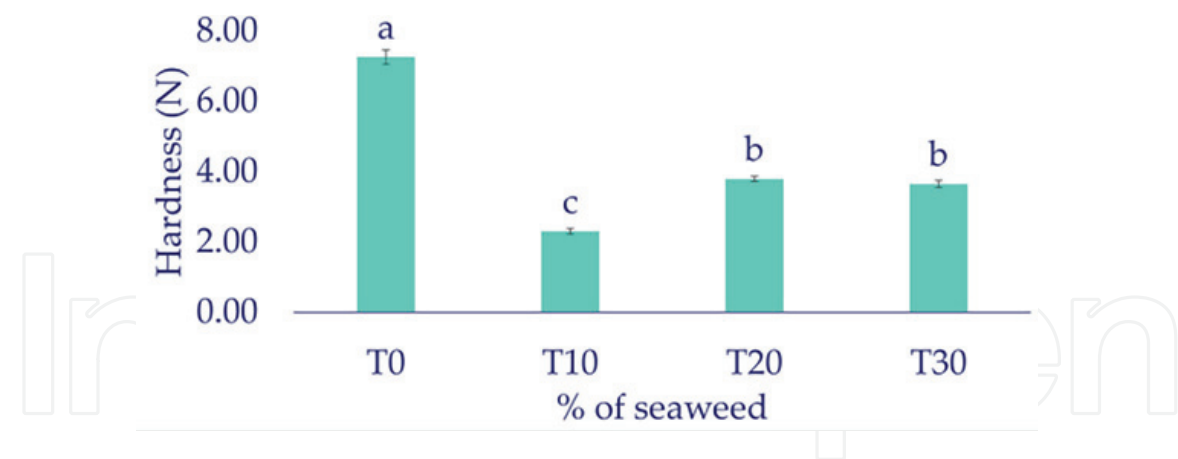


Figure 4. Texture of tortillas with different concentrations of seaweed. Each value represent the results from mean and standard deviation. Different letters indicate significant differences.

4. Conclusions

The addition of the seaweed *Macrocystis pyrifera* on white maize *masa* and tortillas modified significantly all the analyzed parameters. Except for texture of both *masa* and tortillas and ABTS of *masa*, the effect of the seaweed incorporation showed a linear trend. The results showed that fortifying tortillas with *M. pyrifera* could lead to a functional product with increased phenolic compounds and antioxidant capacity, which also covers the RDA of several minerals (Ca, Mg, P and K). However, it would be important to further study the acceptability of consumers, focusing mainly on color and flavor.

Acknowledgements

The author A. Pérez-Alva thank Universidad de las Américas Puebla (UDLAP) and Consejo Nacional de Ciencia y Tecnología (CONACyT) for the financial aid for her doctoral studies.

Conflict of interest

The authors declare no conflict of interest.

IntechOpen


IntechOpen

Author details

Alexa Pérez-Alva, Diana K. Baigts-Allende and Milena M. Ramírez-Rodrigues*
Department of Chemical, Food and Environmental Engineering, Universidad de las
Americas Puebla, Cholula, Puebla, Mexico

*Address all correspondence to: milena.ramirez@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Sarwar MH, Sarwar MF, Sarwar M, Qadri NA, Moghal S. The importance of cereals (Poaceae: Gramineae) nutrition in human health: A review. *Journal of Cereals Oilseeds*. 2013;**4**(3):32-35
- [2] Yaqoob S, Cai D, Liu M, Zheng M, Zhao C-B, Liu J-S. Characterization of microstructure, physicochemical and functional properties of corn varieties using different analytical techniques. *International Journal of Food Properties*. 2019;**22**(1):572-582
- [3] Shiferaw B, Prasanna BM, Hellin J, Bänziger M. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Security*. 2011;**3**(3):307
- [4] FAO. Crops and Livestock Products [Internet]. Available from: <https://www.fao.org/faostat/en/#data/QCL>. [Accessed: October 20, 2021]
- [5] Gwirtz JA, Garcia-Casal MN. Processing maize flour and corn meal food products. *Annals of the New York Academy of Sciences*. 2014;**1312**(1):66-75
- [6] Alvarez-Poblano L, Roman-Guerrero A, Vernon-Carter EJ, Alvarez-Ramirez J. Exogenous addition of muicle (*Justicia spicigera* Schechtendal) extract to white maize tortillas affects the antioxidant activity, texture, color, and in vitro starch digestibility. *LWT*. 2020;**133**:110120
- [7] Escalante-Araiza F, Gutiérrez-Salmeán G. Traditional Mexican foods as functional agents in the treatment of cardiometabolic risk factors. *Critical Reviews in Food Science and Nutrition*. 2021;**61**(8):1353-1364
- [8] SIAP. Anuario Estadístico de la Producción Agrícola [Internet]. Available from: <https://nube.siap.gob.mx/cierreagricola/>. [Accessed: October 20, 2021]
- [9] SIAP. Panorama Agroalimentario 2020 [Internet]. Available from: https://nube.siap.gob.mx/gobmx_publicaciones_siap/pag/2020/Atlas-Agroalimentario-2020. [Accessed: October 20, 2021]
- [10] Sweeney S, Steigerwald DG, Davenport F, Eakin H. Mexican maize production: Evolving organizational and spatial structures since 1980. *Applied Geography*. 2013;**39**:78-92
- [11] Sánchez-Vega LP, Espinoza-Ortega A, Thomé-Ortiz H, Moctezuma-Pérez S. Perception of traditional foods in societies in transition: The maize tortilla in Mexico. *Journal of Sensory Studies*. 2021;**36**(2):e12635
- [12] de los Ángeles Cornejo-Villegas M, Gutiérrez-Cortez E, Rojas-Molina I, Del Real-López A, de la Luz Zambrano-Zaragoza M, Martínez-Vega V, et al. Physicochemical, morphological, and pasting properties of nixtamalized flours from quality protein maize and its particle distribution. *LWT – Food Science and Technology*. 2013;**53**(1):81-87
- [13] de la Parra C, Saldivar SOS, Liu RH. Effect of processing on the phytochemical profiles and antioxidant activity of corn for production of *masa*, tortillas, and tortilla chips. *Journal of Agricultural and Food Chemistry*. 2007;**55**(10):4177-4183
- [14] Salazar D, Rodas M, Arancibia M. Production of tortillas from nixtamalized corn flour enriched with Andean crops flours: Faba-bean (*Vicia faba*) and white-bean (*Phaseolus vulgaris*). *Emirates Journal of Food and Agriculture*. 2020;**32**:731-738

- [15] Escalante-Aburto A, Mariscal-Moreno RM, Santiago-Ramos D, Ponce-García N. An update of different Nixtamalization technologies, and its effects on chemical composition and nutritional value of corn tortillas. *Food Review International*. 2020;**36**(5):456-498
- [16] Adom KK, Liu RH. Antioxidant activity of grains. *Journal of Agricultural and Food Chemistry*. 2002;**50**(21):6182-6187
- [17] León-Murillo JR, Gutiérrez-Dorado R, Reynoso-Camacho R, Milán-Carrillo J, Perales-Sánchez JXX, Cuevas-Rodríguez EO, et al. Tortillas made with extruded flours of blue maize and Chía seeds as an nutritious and nutraceutical food option: Tortillas preparadas con harinas extruidas de maíz Azul y Semillas de chía Como una opción de alimento nutritivo y nutraceutico. *Agrociencia*. 2021;**55**(6):487-506
- [18] Contreras Jiménez B, Oseguera Toledo ME, Garcia Mier L, Martínez Bravo R, González Gutiérrez CA, Curiel Ayala F, et al. Physicochemical study of nixtamalized corn *masa* and tortillas fortified with “chapulin” (grasshopper, *Sphenarium purpurascens*) flour. *CyTA – Journal of Food*. 2020;**18**(1):527-534
- [19] Cornejo-Villegas MA, Acosta-Osorio AA, Rojas-Molina I, Gutiérrez-Cortéz E, Quiroga MA, Gaytán M, et al. Study of the physicochemical and pasting properties of instant corn flour added with calcium and fibers from nopal powder. *Journal of Food Engineering*. 2010;**96**(3):401-409
- [20] Leyton A, Flores L, Mäki-Arvela P, Lienqueo ME, Shene C. Macrocytis pyrifera source of nutrients for the production of carotenoids by a marine yeast *Rhodotorula mucilaginosa*. *Journal of Applied Microbiology*. 2019;**127**(4):1069-1079
- [21] MaS A-E, Mansilla A, Ojeda J, Marambio J, Rosenfeld S, Mendez F, et al. Nutritional properties of dishes prepared with sub-Antarctic macroalgae—An opportunity for healthy eating. *Journal of Applied Phycology*. 2017;**29**(5):2399-2406
- [22] Shannon E, Abu-Ghannam N. Seaweeds as nutraceuticals for health and nutrition. *Phycologia*. 2019;**58**(5):563-577
- [23] Ford L, Theodoridou K, Sheldrake GN, Walsh PJ. A critical review of analytical methods used for the chemical characterisation and quantification of phlorotannin compounds in brown seaweeds. *Phytochemical Analysis*. 2019;**30**(6):587-599
- [24] AOAC. Official Methods of Analysis. 17th ed. Washington, DC: Association of Official Analytical Chemists; 1997
- [25] Kumaravel S, Alagusundaram K. Determination of mineral content in Indian spices by ICP-OES. *Oriental Journal of Chemistry*. 2014;**30**(2):631-636
- [26] Méndez-Lagunas LL, Cruz-Gracida M, Barriada-Bernal LG, Rodríguez-Méndez LI. Profile of phenolic acids, antioxidant activity and total phenolic compounds during blue corn tortilla processing and its bioaccessibility. *Journal of Food Science and Technology*. 2020;**57**(12):4688-4696
- [27] Wrolstad RE, Smith DE. Color analysis. In: Nielsen SS, editor. *Food Analysis*. 4th ed. Boston: Springer; 2010. pp. 573-585
- [28] Argüello-García E, Martínez-Herrera J, Córdova-Téllez L, Sánchez-Sánchez O, Corona-Torres T. Textural, chemical and sensorial properties of maize tortillas fortified

with nontoxic *Jatropha curcas* L. flour. *CyTA – Journal of Food*. 2017;**15**(2):301-306

[29] Montemayor-Mora G, Hernández-Reyes KE, Heredia-Olea E, Pérez-Carrillo E, Chew-Guevara AA, Serna-Saldívar SO. Rheology, acceptability and texture of wheat flour tortillas supplemented with soybean residue. *Journal of Food Science and Technology*. 2018;**55**(12):4964-4972

[30] Xiang L, Xiao L, Wang Y, Li H, Huang Z, He X. Health benefits of wine: Don't expect resveratrol too much. *Food Chemistry*. 2014;**156**:258-263

[31] Re R, Pellegrini N, Proteggente A, Pannala A, Yang M, Rice-Evans C. Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radical Biology & Medicine*. 1999;**26**(9-10):1231-1237

[32] Loayza FE, Brecht JK, Simonne AH, Plotto A, Baldwin EA, Bai J, et al. Enhancement of the antioxidant capacity of ripe tomatoes by the application of a hot water treatment at the mature-green stage. *Postharvest Biology and Technology*. 2020;**161**:111054

[33] Cian RE, Fajardo MA, Alaiz M, Vioque J, González RJ, Drago SR. Chemical composition, nutritional and antioxidant properties of the red edible seaweed *Porphyra columbina*. *International Journal of Food Sciences and Nutrition*. 2014;**65**(3):299-305

[34] Angell AR, Angell SF, de Nys R, Paul NA. Seaweed as a protein source for mono-gastric livestock. *Trends in Food Science and Technology*. 2016;**54**:74-84

[35] Colín-Chávez C, Virgen-Ortiz JJ, Serrano-Rubio LE, Martínez-Téllez MA, Astier M. Comparison of nutritional properties and bioactive compounds

between industrial and artisan fresh tortillas from maize landraces. *Current Research Food Science*. 2020;**3**:189-194

[36] Lozano Muñoz I, Díaz NF. Minerals in edible seaweed: Health benefits and food safety issues. *Critical Reviews in Food Science and Nutrition*. 2020;**62**:1-16

[37] U.S. Department of Agriculture and U.S. Department of Health and Human Services. Dietary Guidelines for Americans, 2020-2025 [Internet]. Available from: https://www.dietaryguidelines.gov/sites/default/files/2020-12/Dietary_Guidelines_for_Americans_2020-2025.pdf. [Accessed: October 20, 2021]

[38] Aguayo-Rojas J, Mora-Rochín S, Cuevas-Rodríguez EO, Serna-Saldivar SO, Gutierrez-Urbe JA, Reyes-Moreno C, et al. Phytochemicals and antioxidant capacity of tortillas obtained after lime-cooking extrusion process of whole pigmented Mexican maize. *Plant Foods for Human Nutrition*. 2012;**67**(2):178-185

[39] Anton AA, Ross KA, Lukow OM, Fulcher RG, Arntfield SD. Influence of added bean flour (*Phaseolus vulgaris* L.) on some physical and nutritional properties of wheat flour tortillas. *Food Chemistry*. 2008;**109**(1):33-41

Artificial Intelligence for Sensory Acceptability Prediction

Jorge Metri, Nelly Ramírez, Milena Ramírez and Diana Baigts

Abstract

Sensory evaluations are usually expensive and time consuming but are indispensable to ensure the success of a food product in the market. Artificial intelligence, such as machine learning models, is emerging as a powerful tool in food science to predict or classify sensory attributes. This study aimed to evaluate two machine learning models, a Multivariate Linear Regression Model (MLRM), and an Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict sensory acceptability of mayonnaise. A dataset of 64 data points was used, the analysis of the data indicates that oil content (%), viscosity (Pa.s), taste score, and texture score were the most pertinent data for the models (correlation index of 0.25–0.75). MLRM showed an adjustment of $r^2 = 0.864$ in the prediction with a positive bias. ANFIS model demonstrated a slightly higher adjustment ($r^2 = 0.910$) than the MLRM; additionally, ANFIS fit better to the data as it does not show any bias in the results. The obtained results suggest that tested algorithms can predict sensory acceptability, being a useful approach that can be used for reducing time and number of food sensory evaluations.

Keywords: machine learning, mayonnaise, ANFIS, multivariate regression, sensory evaluation

1. Introduction

Food production has faced important challenges in recent years, the demand for high quality foods has arisen the demands for methods to obtain rapid food manufacturing, low-level of uncertainty and error in the food formulation, safety, sustainability, and sensory satisfaction for the consumers [1]. Sensory analysis comprises a crucial step for product development because it determines in a great extent the overall acceptability, thus, the success of the product in the market. Due to the subjectivity and variability of the sensory attributes, these methods imply a large quantity of evaluations (usually >100) of trained or non-trained consumers. Therefore, a large quantity of samples is needed, which is time consuming and results in several waste of food products [2].

Artificial intelligence (AI) are computational systems, which can learn from experience and predict, classify, cluster, or even determine abstract patterns in the data provided (deep learning models). These AI methods have the advantage to manipulate a

high volume of data, detect non-evident characteristics among the features, and thus, producing a projective result (prediction) in different scenarios. The application of artificial intelligence in the food area is vast, ranging from quality control (detection and determination of pesticides, proximate composition estimation, authentication of food origin, and classification of oil blends), microbiological analysis [3–6], and sensory quality (aroma profile, texture quality, among others) [1, 7].

In the sensory science field, the most used AI algorithms (models) are the partial least squares (PLS), support vector machines (SVM), random forest (RF), and deep learning approaches (artificial neural networks). These models have been successfully used for estimating sensory quality using the chemical composition [8], volatile profile of the product [9], and image analysis [10] to obtain the features for the models. However, predicting specifically the overall acceptability is challenging due to the variability among the products' formulation and characteristics.

Some regression models such as multiple linear regression (MLR), partial least squares (PLS), and regression trees have been used in the forecast of sensory quality of food products [11]. For instance, a PLS-ANN hybrid model has been reported for predicting the acceptability of green tea beverages [12], a regression trees algorithm has been applied in fresh pineapple acceptability [13] and two regression models (multiple linear regression and PLS) were used for the sensory acceptance of beef bouillon [11]. However, the models based on linear regression have some limitations such as: strong correlation among independent and dependent variables, overfitting of the model, and ignoring non-linear associations [11] which can result in a low performance ($r^2 = 0.600\text{--}0.800$).

One way to surpass the limitations of linear models is the use of artificial neural networks. Adaptive Neuro-Fuzzy Inference System (ANFIS) are known as hybrid algorithms because these models combine the benefits of neural networks (fitting and finding non-linear associations among data) and the human if-then logical learning; hence, these models can learn from data with no clear relationship among the variables and the desired output. ANFIS models have been used for predicting different type of food attributes, especially for quality control [14], drying [15], physical stability [16], rheological properties [16], but their application in sensory acceptability [17] is limited, yet it demonstrated promising results for this purpose (>90% of predictability).

Implementing AI models for estimating the sensory acceptability as a rapid screening could be useful to evaluate and reformulate food products without compromising a large quantity of samples, reducing the experimental time, or involving expensive food sensory panels. Therefore, the aim of this study was to use the formulation, the physical properties, and sensory scores to predict the acceptability of mayonnaises throughout a MLRM and an ANFIS model.

2. Methodology

2.1 Case study

The mayonnaise is one of the most common seasoning products, it is used in a wide variety of food dishes (such as salads, sandwiches, hamburgers, among others). The traditional formulation consists of oil (up to 80% of the product), vinegar or lime juice, egg yolk, sugar, salt, and pepper. Nevertheless, there are some variations of this sort of products like low-fat mayonnaise, egg-free mayonnaise, virgin oil mayonnaise, among others [18]. These varieties have different ingredients such as stabilizers,

emulsifiers, and thickening agents [19] which makes challenging the prediction of the sensory attributes.

The present case of study is to estimate the acceptability of mayonnaises using the formulation, physical attributes, and sensory evaluation scores. To achieve the objective, a dataset of different features of the mayonnaises was generated and two different algorithms were used: a Multivariate Linear Regression Model (MLRM) and an Adaptive Neuro-Fuzzy Inference System (ANFIS).

2.2 Software and equipment

The database processing, the models' development, and the data visualization were done with Python 3.8.1 using Spyder (v 5.1.1) in Anaconda Navigator for Windows (v. 2.0.4). The libraries used for the database processing were Pandas (v. 1.2.5) and Statsmodels (v. 0.12.2). The MLRM was constructed using Scikit-Learn library (v. 0.24.2) and the ANFIS model was developed with Keras (v. 2.4.3) and Tensorflow (v. 2.5). The data visualization (Figures) was done with Matplotlib (v. 3.3.4).

The computational equipment was a laptop with a processor Intel core i5 10th gen (Quad-core at 2.5 GHz), 16 GB RAM at 2933 MHz, and a NVIDIA GeForce GTX 1660ti GPU (6 GB of VRAM, 1770 MHz and 1536 CUDA cores).

2.3 Database development

The product formulation and some physical characteristics were selected to develop the database considering that these features have an impact in the sensory properties, and they could be used for the acceptability prediction. The data was retrieved from published research articles, around 34 studies were found with the desired characteristics, which included all formulation of the mayonnaise, viscosity or texture analysis, and sensory evaluations. The research papers with incomplete data (i.e.: incomplete texture analysis, incomplete rheological analysis, without acceptability determination) were discarded. After the information filtering, the experimental data of 22 research papers were included.

The database obtained included the formulation (oil content, egg yolk, water content, flavorings, and vinegar), viscosity, taste score, texture score, and acceptability score; furthermore, a correlation matrix was generated to select the most pertinent data for the models. The whole database was comprised of 64 observations and 10 features (matrix 64 x 10), nevertheless, it was separated in the following sets:

- The training-testing set was comprised by the first 56 data points of the database. This dataset was used for training and testing the models.
- The validation set contained the last 8 points of the database. This dataset was used for the validation of the models because these data were not seen by the algorithms during the training.

2.4 Development of machine learning models

2.4.1 Multivariate linear regression model (MLRM)

Multivariate linear regression is a common mathematical approach to modeling data where a dependent variable, usually, has a direct association with the

independent variables. Then, the dependent variable is fitted with the independent variables in a mathematical function (Eq. (1)), which is used to predict the new data. The MLRM training and test split was done with 85/15% (respectively) of the data from the training-testing set.

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots \beta_nx_n \tag{1}$$

Where Y is the output variable, β_0 is the intercept, $\beta_n x_n$ are the coefficient and the independent variable, respectively.

2.4.2 Adaptive neuro-fuzzy inference system (ANFIS) model

The ANFIS is a method developed in 1993 by Jang [20], this model (ANFIS) combines the fuzzy logic learning (if-then rules) to convert the input data into a desired output through several interconnections in an artificial neural network. Along these interconnections, several if-then rules are set to establish a relationship between the inputs and the output [21]. The architecture of the model is complex, in **Figure 1** is shown a schematic representation of two inputs and the interconnections between the layers, usually, the ANFIS architecture contains 5 layers as follows.

In the first layer (Layer 1), known as the input layer, is where the membership functions are determined for the input nodes by the premise parameters. The membership functions (MF) help to define the degree of truth (e.g.: IF “ x ” is True and “ y ” is False THEN “ $z = f(x, y)$ ”) for a crisp value in a scale from 0 to 1, where 0 is False and 1 is True. Each membership function (μ) was calculated by the algorithm following the gaussian MF equation (Eq. (2)).

$$\mu_{Ai}(X) = e^{-\frac{1}{2}(\frac{x-c}{\sigma})^2} \tag{2}$$

Where A is the layer node, X is the input variable, c and σ represent the premise parameters.

Layer 2 is the fuzzifier or rule layer, where the firing strengths (w_i) for the membership functions are determined. The firing strengths represents the degree of

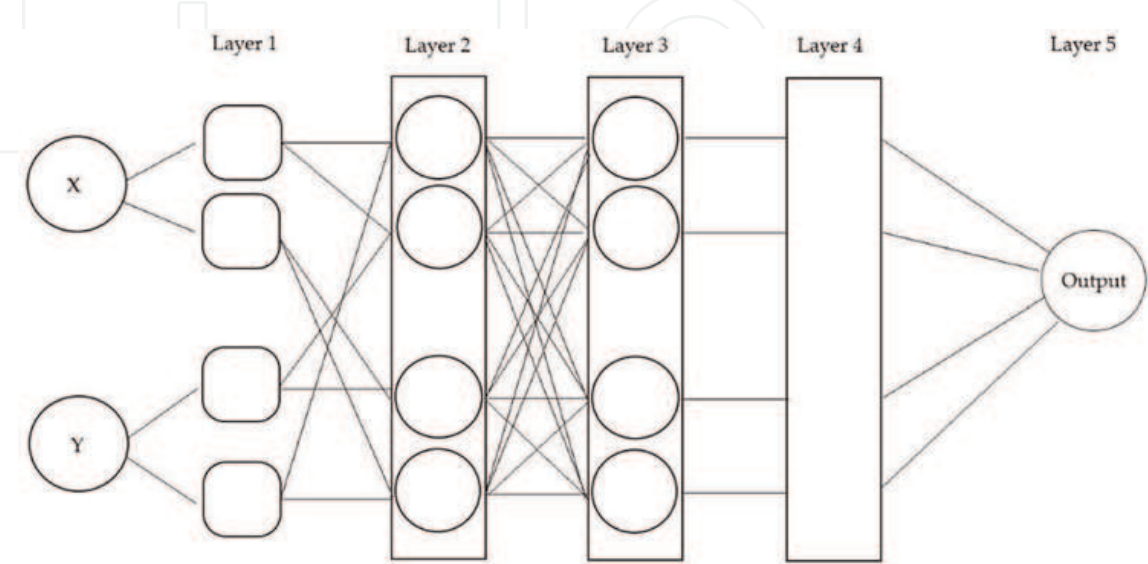


Figure 1. Schematic representation of the ANFIS architecture where X and Y represent two input variables.

the effect of each membership function to the output variable. The firing strengths were calculated as shown in Eq. (3).

$$w_i = \mu_{Ai}(X) * \mu_{Bi}(Y) \quad (3)$$

Where w is the firing strength and μ is the membership function. A and B represent the input nodes and X and Y the input variables.

In the third layer (Layer 3), the firing strengths are normalized (\overline{w}_i) as the ratio of the n^{th} fire strength to the total firing strengths. This could be expressed as follows (Eq. (4)).

$$\overline{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + w_4} \quad (4)$$

Where w_n corresponds to each of the firing strengths calculated in the second layer.

Layer 4 is known as the defuzzification layer, in this layer, the weighted values of rules are calculated. This is expressed by the following equation (Eq. (5)).

$$\overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \quad (5)$$

\overline{w}_i represents the output of the third layer and p, q, r are the consequence parameters.

Layer 5 is the output layer, the output is a summatory of all outputs of the fourth layer, as it is shown in Eq. (6).

$$\sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (6)$$

The source code for the ANFIS was adapted from a repository on Github [22]. All the data was scaled (from 0 to 1), and the results obtained were re-scaled to original values using the MinMaxScaler function of Scikit-Learn library (v. 0.24.2). The optimization of the algorithm was done using the Adam optimizer, and the mean squared error was used as loss value function. The parameters of the ANFIS model were set after varying the number of membership functions (4–16) and the training and test data split was set at the same conditions as the MLRM (train and test split at 85/15% of the training-tasting set).

2.5 Models evaluation

The models performance was evaluated with the validation set (8 unseen data points) [17], comparing the adjustment (r^2) and the mean squared error (mse) in the predicted values with Eqs. (7) and (9), respectively.

$$r^2 = 1 - \frac{\sum_{i=1}^N [O_i - T_i]^2}{\sum_{i=1}^N [O_i - T_m]^2} \quad (7)$$

Where O_i is the real value, T_i is the predicted value, N the number of data, and T_m is the ratio of the real values and N (Eq. (8)).

$$T_m = \frac{\sum_{i=1}^N O_i}{N} \tag{8}$$

$$mse = \frac{1}{N} \sum_{i=1}^N (O_i - T_i)^2 \tag{9}$$

3. Results and discussion

3.1 Database construction and cleaning

The data obtained from literature was standardized as follows. All components for the formulation (oil content, egg yolk, water content, flavorings, and vinegar) were recalculated to be expressed as percentage (%). Viscosity was standardized in Pa.s. All sensory scores were found in different scales (e.g.: 1–5, 1–9, 1–10, 1–100), therefore, the data was rescaled to a range of 1–10. The rescaling was done multiplying the score by 10 and dividing the result by the higher value of the original scale.

For selecting the appropriate features of the database, a correlation matrix was generated (**Figure 2**). It was observed that few characteristics (taste and texture scores) were closely correlated (~0.75) with the acceptability; meanwhile, the oil content, viscosity, and egg yolk content showed positive correlation but with lower impact on the desired variable (0.00 — 0.25). In the opposite way, the water, vinegar, and stabilizers content were negatively correlated with the acceptability in a range of –0.50 to –0.25. The features with better correlation (oil content, viscosity, taste score, and texture score) were selected for the acceptability prediction.

It is well known that all ingredients in the mayonnaises play a specific and important role [19, 23], however, most of the data such as water content, flavorings,

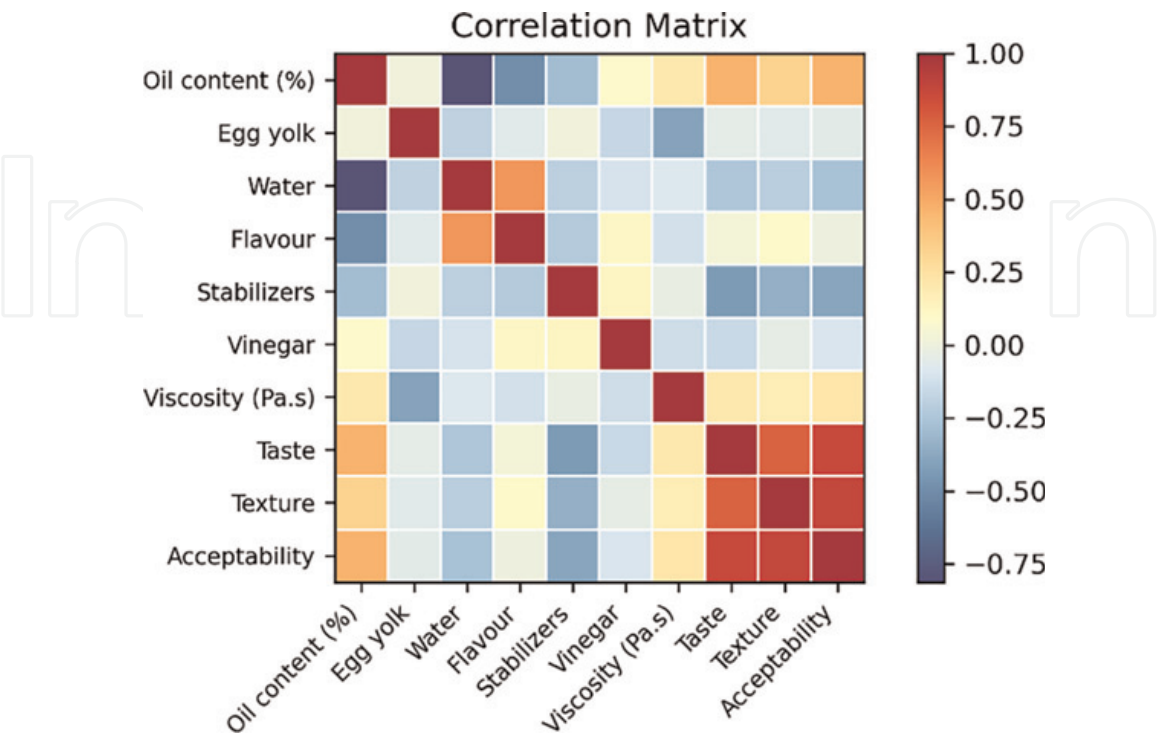


Figure 2. Correlation matrix for mayonnaise ingredients, viscosity, and sensory parameters.

stabilizers, and vinegar were not critical for the algorithms. Conversely, the oil content (%) showed a weak relationship with the acceptability (~ 0.25) but it represented an important input for the models. The high influence of this parameter could be related to its role in the product because it contributes to the viscosity, mouthfeel, appearance, and flavor of the product [23]. Regarding the correlations among acceptability and the two sensory attributes (taste and texture) found herein, it agrees with that reported in Caracciolo et al. [24], regarding the willingness to consume (comparable with acceptability) of microgreens (leafy vegetables). Among different features such as aroma, astringency, bitterness, texture, sourness, appearance, and flavor, the authors observed that flavor and texture were the features with more relationship with the willingness to consume [24].

3.2 Multivariate linear regression model (MLRM)

The training of the MLRM (**Figure 3A**) showed large variation between the predicted data and the real value, resulting in a low adjustment ($r^2 = 0.616$) and a mean squared error of 0.477. Although the residuals obtained from the training data (**Figure 3C**) showed a random distribution, which may indicate that the data fitted well for the model, the absolute value of these residuals was high (up to 1.5). The following equation was obtained from the MLRM for the prediction:

$$\begin{aligned} \text{Acceptability} = & -0.751 + 0.002(\text{Oil content}\%) - 3.09 \times 10^{-5}(\text{Viscosity}) \\ & + 0.509(\text{Flavor score}) + 0.586(\text{Texture score}). \end{aligned}$$

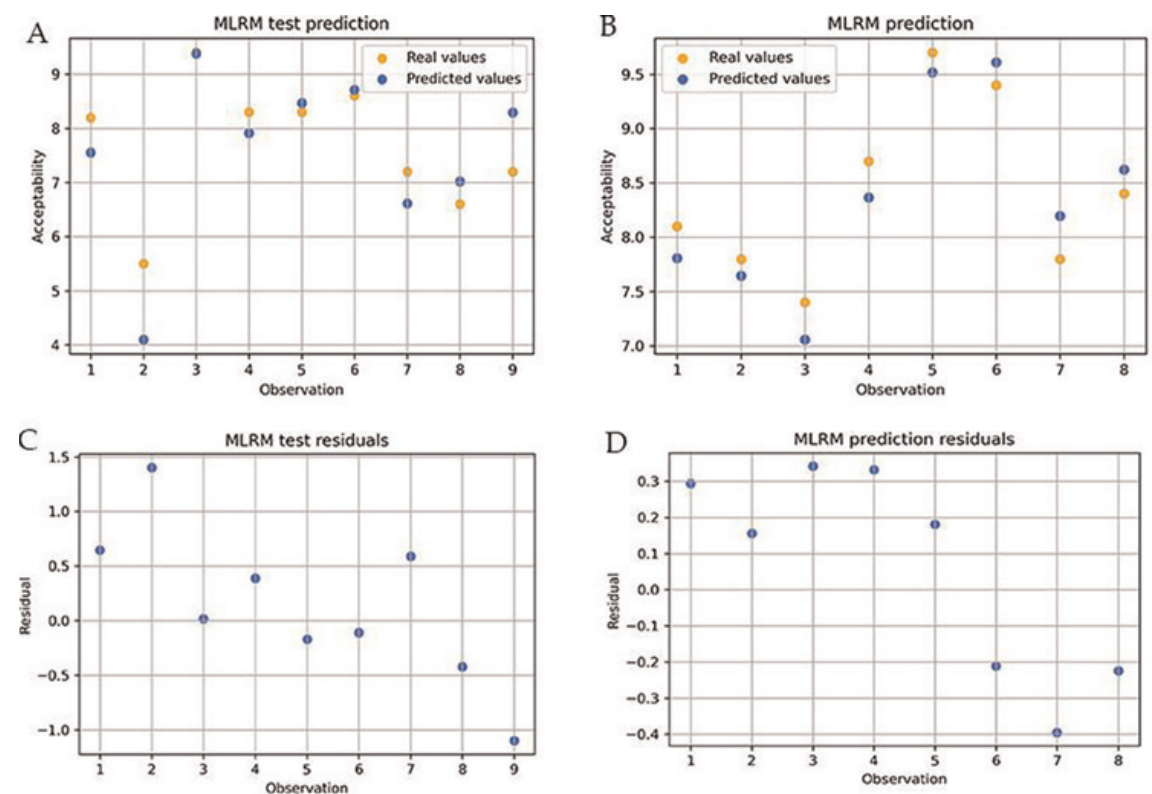


Figure 3. Prediction of testing data (A) and validation data (B) and their residuals (C, D) of sensory acceptability of mayonnaise using a multivariate linear regression model.

In the validation of the model (unseen data), which is shown in **Figure 3B**, the MLRM showed an increase of the adjustment ($r^2 = 0.864$) and a decrease of the mean squared error (0.077) in comparison to the training. The behavior of the residuals (**Figure 3D**) showed a positive bias in the results (the model tends to underestimate the data), additionally, the random distribution of the residual was slightly loss; however, this underestimation could be advantageous because it is possible that the real acceptability be better than the predicted by the model. Hence, these results showed that the MLRM could be applied to estimate the acceptability of a mayonnaise by using non-sensory and sensory parameters with considerably good reliability.

Some regression models reported for sensory data showed results comparable to the obtained in the present study, overall, the MLRM have low predictive performance ($r^2 = 0.600$ – 0.800). In a study of the application of a PLS (partial least squares) algorithm to predict the overall acceptability of green tea showed an adjustment of 0.776 in the prediction of unseen data [12]. Similarly, a multivariate linear regression models showed low predictive performance ($r^2 = 0.806$) for the oxidative stability of virgin olive oil [14] and the sensory quality of beer ($r^2 = 0.660$) [25]. Regarding the model of this work, it showed a better performance than the studies previously cited, probably the data selection and the strong correlation of sensory evaluations could improve the predictive performance of this model [11].

3.3 ANFIS model

The architecture for the ANFIS was selected by testing different conditions such as type of membership function, number of membership functions, and number of epochs (number of training cycles of the model) until obtain the best adjustment.

The selection of membership function types (sigmoid, gaussian, and generalized bell) were tested using a fixed number of membership functions (MF = 6) and 100 epochs. It was observed that sigmoid MF reduced the prediction accuracy at values around $r^2 = 0.700$, however, gaussian and generalized bell MF showed similar results ($r^2 \sim 0.800$). The gaussian type MF was selected because the use of only two premise parameters (c and σ) allows a faster calculation and usually gives better results compared to generalized bell and sigmoid type MF [26–28].

The number of membership functions also influenced the performance of the model, therefore, different MF number (4, 8, 12, and 16 MF) were tested using 100 epochs to obtain the best results. The lower adjustment ($r^2 = 0.79$) was obtained using 4 MF, which was increased gradually ($r^2 \sim 0.830$) using 8 and 12 MF. Compared to the previous conditions, the best performance was observed using 16 MF ($r^2 = 0.910$).

The learning curves of the model during the training are shown in **Figure 4**, it is observed that the training loss curve cross the validation loss near to the epoch number 30, then, both values (training and validation) converge close to the epoch number 80 epochs reaching an error value near to 0, this indicates that the model was well trained (by minimizing the error) and adjusted. Therefore, the number of epochs were kept at 100 to avoid an overfitting of the model; the final architecture of the ANFIS was: 4 inputs, 16 membership functions (MF), and 100 epochs.

Due to the batch size set for the ANFIS model (batch size = 8), one data point was missed in the training-testing set compared to the testing of the MLRM. Hence, the testing of ANFIS were done with 8 data points (**Figure 5A**). The results showed that the performance of ANFIS exceed the MLRM because the adjustment and error values ($r^2 = 0.941$ and a mean squared error of 0.057) were better. In addition, the residuals

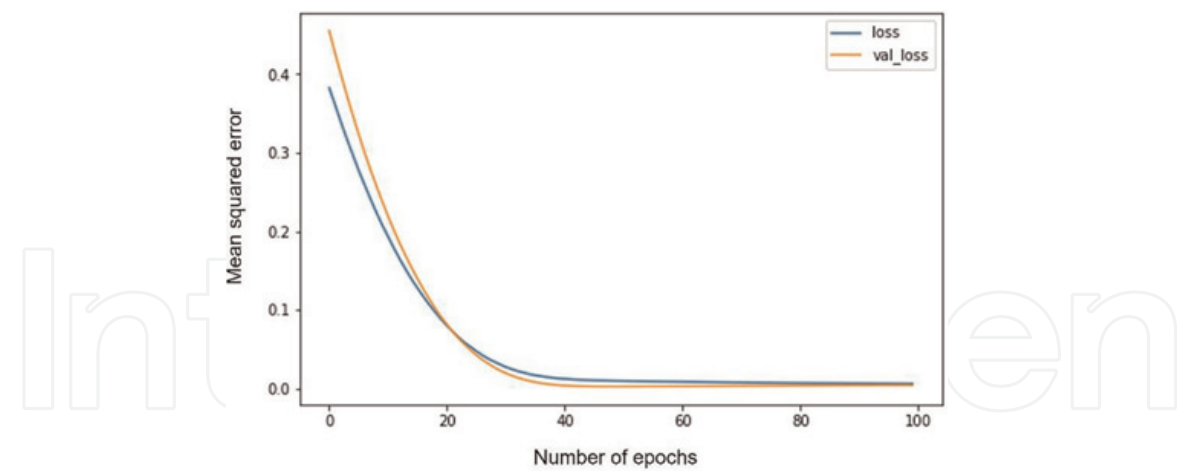


Figure 4.
ANFIS training curve. Training loss (*loss*) and validation loss (*val_loss*).

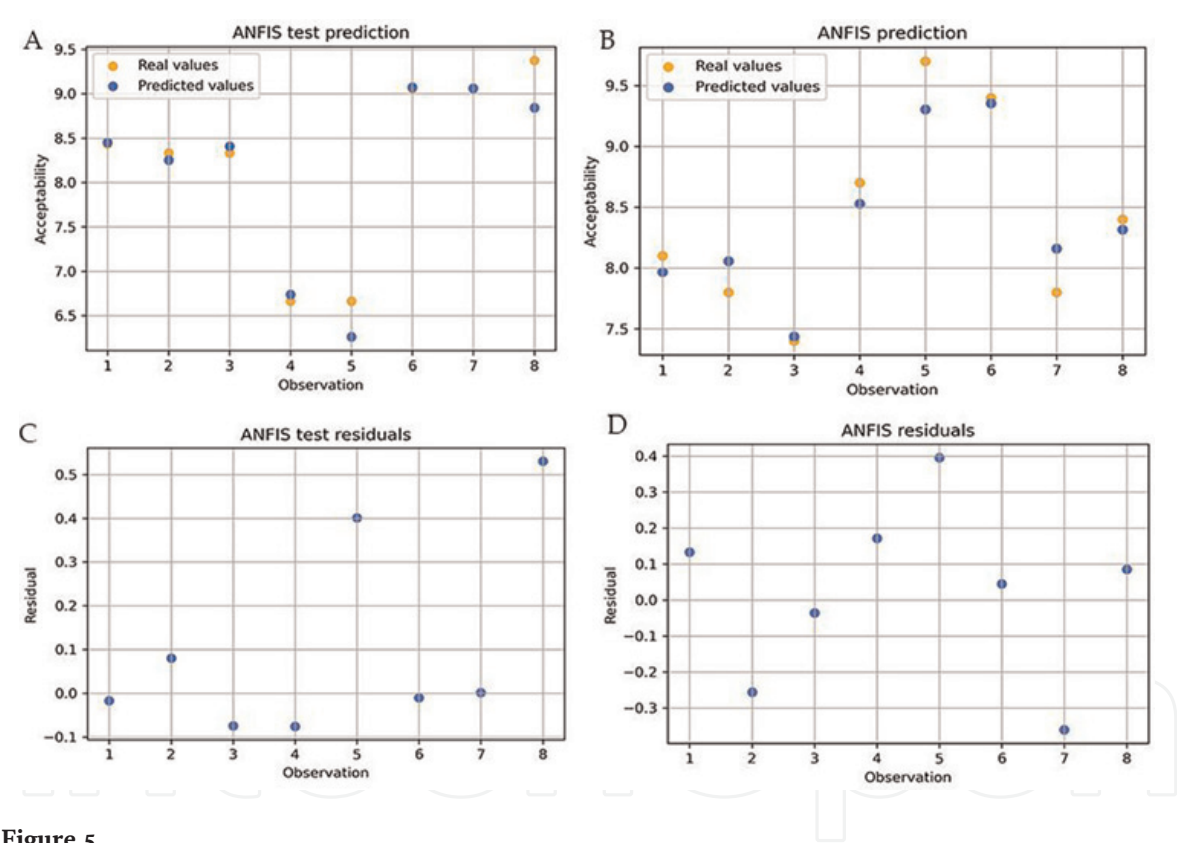


Figure 5.
Prediction of testing data (A) and validation data (B) and their residuals (C, D) of sensory acceptability of mayonnaise using an adaptive neuro-fuzzy inference system (ANFIS) model.

obtained (Figure 5C) in the model also demonstrate a positive bias but with a lower absolute value (0.1–0.5) compared to the MLRM.

The validation of the ANFIS is shown in Figure 5B. It was observed a slight reduction in the adjustment and the error ($r^2 = 0.910$ with a $mse = 0.051$) than the test prediction. The residuals (Figure 5D) indicate a good fitting of the data because these showed a random distribution. Despite the reduction of the adjustment in the validation, the ANFIS demonstrated to be a better option to predict the sensory acceptability of mayonnaise in comparison to the MLRM.

	Prediction summary							Models' performance		
	Oil content (%)	Viscosity (Pa·s)	Taste	Texture	Acceptability	MLRM prediction	ANFIS prediction		MLRM	ANFIS
1	75	160	6.9	8.30	8.10	7.81	7.97	r ²	0.864	0.910
2	75	700	7.10	7.82	7.80	7.64	8.06	mse	0.077	0.051
3	75	1000	6.30	7.50	7.40	7.06	7.44			
4	76.90	57.60	8.00	8.30	8.70	8.37	8.53			
5	53.80	57.20	9.10	9.40	9.70	9.52	9.30			
6	53.80	4.96	9.4	9.30	9.40	9.61	9.35			
7	53.80	3.35	8.00	8.10	7.80	8.20	8.16			
8	40	16.52	8.10	8.80	8.40	8.62	8.32			

Table 1.
Models' prediction and performance summary. Obs = observation. r² = adjustment. Mse = mean squared error. MLRM = multivariate linear regression model. ANFIS = adaptive neuro-fuzzy inference system.

The ANFIS model obtained in the present study showed a lower adjustment compared to other studies, for instance in the prediction of virgin olive oil stability ($r^2 = 0.988$), moisture ratio in papaya slices after drying ($r^2 = 0.996$), and the sensory acceptability in ice creams ($r^2 = 0.930$) [14, 15, 17]. Herein, it was needed a more complex architecture compared to the model reported by Bahram-Pavar et al. [17], which used 100 epochs, 2 MF, and 4 sensory parameters of quality-descriptive-analysis (QDA). The difference in both architectures could be due to the nature of the input variables, as it is shown in the correlation matrix, the sensory attributes are closely related to the overall acceptability; in contrast, the formulation, and the viscosity (used in our model) are not explained with a clear linear behavior. It is hypothesized that the model requires more MF to create a better relationship between the acceptability and the non-linear features, yet it is possible to use these variables.

In **Table 1** is summarized the data from the validation of the models, as well as the summary of their performance metrics. Overall, the ANFIS model showed best results, indicated by the higher adjustment than the obtained from the MLRM. It has been documented that hybrid models, such as ANFIS, are more suitable for predicting non-linear data [12, 14]. This is thanks to the training through back propagation and the multilayer architecture, that could provide a better adaptation to the non-linear parameters [12]. In this case, the viscosity and the oil content are not highly correlated with the acceptability, thus, it is probable that these data did not fit well for the prediction in the linear model (MRLM). Hence, it could be inferred that ANFIS model is superior to the MRLM if non-linear parameters are used, with possible research or industrial application; however, ANFIS model requires more computational skills in comparison to MRLM. This latter, despite the lower performance, could be considered as a simple model for non-industrial tasks, for instance, in academic practices.

Finally, it is important to point out some limitations of the present study, for instance, the high variability within the formulation of the products (use of gums, proteins, starches from different sources, different flavorings, among others), the low volume of the database, because most of the machine learning usually needs large datasets (>1000 data points) for being statistically significant.

However, the strength of this study is that the data used was not from a single study, this suggest that machine learning and deep learning algorithms could be implemented for different formulations in foods, therefore, it could be considered for testing novel ingredients or to evaluate and calibrate reformulations without the time consuming of conventional sensory analysis. Additionally, this type of works could impulse the generation of foods databases for developing industrial solutions using artificial intelligence models.

4. Conclusions

Multivariate linear regression (MLRM) and Adaptive Neuro-Fuzzy Inference System (ANFIS) could be used for predicting sensory data in mayonnaises. Both linear (MLRM) and non-linear (ANFIS) models seems to fit well for this task (good predictability), additionally they are moderately easy to implement, and present good versatility against the differences of the mayonnaises' formulation. Regarding the computational skills required for ANFIs model, it present a higher complexity, but due to its good performance, this model could be proposed for industrial applications or scientific approaches (research and development). In contrast, MLRM is easy to understand, easy for data interpretation, and rapid implementation; these features

could be exploited for academic purposes or to explore simple reformulations by new food manufacturers (i.e: start-ups). For further studies, it is considered as priority to expand the dataset for validating both models with more observations.

Acknowledgements

Thanks to the Universidad De Las Americas Puebla and CONACyT for the scholarship granted to the student Jorge Metri. Also, a special acknowledge for David Cantú and Gabriel Solana for their advice and encouragement in developing this study.

Conflict of interest

The authors declare that there is no conflict of interest and all authors read and agree with the manuscript.

Author details

Jorge Metri*, Nelly Ramírez, Milena Ramírez and Diana Baigts
Departamento de Ingeniería Química, Universidad De Las Américas Puebla,
Ambiental y Alimentos, Puebla, México

*Address all correspondence to: jorge.metrioa@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. Distributed under the terms of the Creative Commons Attribution - NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited. 

References

- [1] Mohammadi V, Minaei S. Artificial intelligence in the production process. *Engineering Tools in the Beverage Industry*. 2019;3:27-63
- [2] Yu P, Low MY, Zhou W. Design of experiments and regression modelling in food flavour and sensory analysis: A review. *Trends in Food Science and Technology*. 2018;71:202-215
- [3] Hiura S, Koseki S, Koyama K. Prediction of population behavior of *listeria monocytogenes* in food using machine learning and a microbial growth and survival database. *Science Reports*. 2021;11(1):1-11. [Internet] [cited 2021 Sep 20]. Available from: <https://www.nature.com/articles/s41598-021-90164-z>
- [4] Alghooneh A, Alizadeh Behbahani B, Noorbakhsh H, Tabatabaei YF. Application of intelligent modeling to predict the population dynamics of *Pseudomonas aeruginosa* in frankfurter sausage containing *Satureja bachtiarica* extracts. *Microbial Pathogenesis*. 2015;85:58-65
- [5] Lou W, Nakai S. Application of artificial neural networks for predicting the thermal inactivation of bacteria: A combined effect of temperature, pH and water activity. *Food Research International*. 2001;34(7):573-579
- [6] Oladunjoye AO, Oyewole SA, Singh S, Ijabadeniyi OA. Prediction of *listeria monocytogenes* ATCC 7644 growth on fresh-cut produce treated with bacteriophage and sucrose monolaurate by using artificial neural network. *LWT - Food Science and Technology*. 2017;76:9-17
- [7] Zhou L, Zhang C, Liu F, Qiu Z, He Y. Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety*. 2019;18(6):1793-1811. [Internet] [cited 2021 Sep 14]. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/1541-4337.12492>
- [8] Wood JE, Allaway D, Boulton E, Scott IM. Operationally realistic validation for prediction of cocoa sensory qualities by high-throughput mass spectrometry. *Analytical Chemistry*. 2010;82(14):6048-6055. [Internet] [cited 2021 Sep 20]. Available from: <https://pubs.acs.org/doi/abs/10.1021/ac1006393>
- [9] Vigneau E, Courcoux P, Symoneaux R, Guérin L, Villière A. Random forests: A machine learning methodology to highlight the volatile organic compounds involved in olfactory perception. *Food Quality and Preference*. 2018;68:135-145
- [10] Viejo CG, Torrico DD, Dunshea FR, Fuentes S. Development of artificial neural network models to assess beer acceptability based on sensory properties using a robotic pourer: A comparative model approach to achieve an artificial intelligence system. *Beverages*. 2019;5(2):33. [Internet] [cited 2021 Sep 20]. Available from: <https://www.mdpi.com/2306-5710/5/2/33/htm>
- [11] Krishnamurthy R, Srivastava AK, Paton JE, Bell GA, Levy DC. Prediction of consumer liking from trained sensory panel information: Evaluation of neural networks. *Food Quality and Preference*. 2007;18(2):275-285
- [12] Yu P, Low MY, Zhou W. Development of a partial least squares-artificial neural network (PLS-ANN) hybrid model for the prediction of consumer liking scores of ready-to-drink green tea beverages. *Food Research International*. 2018;103:68-75

- [13] Schulbach KF, Portier KM, Sims CA. Evaluation of overall acceptability of fresh pineapple using the regression tree approach. *Journal of Food Quality*. 2007; **30**(6):993-1008
- [14] Arabameri M, Nazari RR, Abdolshahi A, Abdollahzadeh M, Mirzamohammadi S, Shariatifar N, et al. Oxidative stability of virgin olive oil: Evaluation and prediction with an adaptive neuro-fuzzy inference system (ANFIS). *Journal of the Science of Food and Agriculture*. 2019; **99**(12):5358-5367. [Internet] [cited 2021 Sep 14]. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/jsfa.9777>
- [15] Yousefi A. Estimation of papaw (Carica papaw L.) moisture content using adaptive neuro-fuzzy inference system (ANFIS) and genetic algorithm-artificial neural network (GA-ANN). *Iran Food Science and Technology Research Journal*. 2017; **12**(6):767-779. [Internet] [cited 2021 Sep 14]. Available from: https://ifstrj.um.ac.ir/article_35803.html
- [16] Ghoush MA, Samhouri M, Al-Holy M, Herald T. Formulation and fuzzy modeling of emulsion stability and viscosity of a gum-protein emulsifier in a model mayonnaise system. *Journal of Food Engineering*. 2008; **84**(2): 348-357
- [17] Bahram-Parvar M, Salehi F, Razavi SMA. Adaptive neuro-fuzzy inference system (ANFIS) simulation for predicting overall acceptability of ice cream. *Engineering in agriculture, environment and food*. 2017; **10**(2):79-86
- [18] Saget S, Costa M, Styles D, Williams M. Does circular reuse of chickpea cooking water to produce vegan mayonnaise reduce environmental impact compared with egg mayonnaise? *Sustainability*. 2021; **13**(9):4726
- [19] Mirzanajafi-Zanjani M, Yousefi M, Ehsani A. Challenges and approaches for production of a healthy and functional mayonnaise sauce. *Food Science & Nutrition*. 2019; **7**(8):2471-2484
- [20] Jang JSR. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*. 1993; **23**(3):665-685
- [21] El-Hasnony IM, Barakat SI, Mostafa RR. Optimized ANFIS model using hybrid metaheuristic algorithms for Parkinson's disease prediction in IoT environment. *IEEE Access*. 2020; **8**: 119252-119270
- [22] Lenhard G. Adaptive-Network-Based Fuzzy Inference System (ANFIS) based on Keras on top of Tensorflow 2.0. [Internet]. 2020. [cited 2021 Sep 13]. Available from: <https://github.com/gregorLen/AnfisTensorflow2.0>
- [23] Taslikh M, Mollakhalili-Meybodi N, Alizadeh AM, Mousavi M-M, Nayebzadeh K, Mortazavian AM. Mayonnaise main ingredients influence on its structure as an emulsion. *Journal of Food Science and Technology*. 2021; **59**:1-9. [Internet] [cited 2021 Aug 17]. Available from: <https://link.springer.com/article/10.1007/s13197-021-05133-1>
- [24] Caracciolo F, El-Nakhel C, Raimondo M, Kyriacou MC, Cembalo L, De PS, et al. Sensory attributes and consumer acceptability of 12 microgreens species. *Agronomy*. 2020; **10**:7, 1043. [Internet] [cited 2021 Sep 15]. Available from: <https://www.mdpi.com/2073-4395/10/7/1043/htm>
- [25] Guido LF, Curto A, Boivin P, Benismail N, Gonçalves C, Barros AA. Predicting the organoleptic stability of beer from chemical data using multivariate analysis. *European Food Research and Technology*. 2006; **226**(1):

57-62. [Internet] [cited 2021 Sep 17].
Available from: <https://link.springer.com/article/10.1007/s00217-006-0508-5>

[26] Daher A, Hoblos G, Khalil M, Chetouani Y. Parzen window distribution as new membership function for ANFIS algorithm-application to a distillation column faults prediction. *Chemometrics and Intelligent Laboratory Systems*. 2018;**175**: 1-12

[27] Talpur N, MNM S, Hussain K. An investigation of membership functions on performance of ANFIS for solving classification problems. In: *MS&E*. Melaka, Malaysia: Institute of Physics Publishing; 2017. p. 012103. [Internet] [cited 2021 Sep 28]. Available from: <https://ui.adsabs.harvard.edu/abs/2017MS&E..226a2103T/abstract>

[28] Habibi E, Salehi M, Yadegarfar G, Taheri A. Optimization of the ANFIS using a genetic algorithm for physical work rate classification. *The International Journal of Occupational Safety and Ergonomics (JOSE)* is a refereed interdisciplinary quarterly journal. Published since 1995. 2018; **26**(3):436-443. DOI: 10.1080/1080354820181435445. [Internet] [cited 2021 Sep 28]. Available from: <https://www.tandfonline.com/doi/abs/10.1080/10803548.2018.1435445>

Chapter

Modeling of Drying Kinetics of Fresh and Osmodehydrated Apples during Convective Drying

Julio Emmanuel González-Pérez, Nelly Ramírez-Corona and Aurelio López-Malo

Abstract

This work used different mathematical models to describe the drying kinetics of osmodehydrated apple cubes (Granny Smith var.) with sucrose solution. The osmodehydration pre-treatment was carried out with different concentrations (40–60Brix) of sucrose solution, at 40°C, 10 g solution/g product and 1440 min. The dimensionless volume, water activity and drying kinetics of fresh (1.5 cm by side) and osmodehydrated apple cubes subjected to convective drying (50°C and 3.5 m/s) were investigated. Newton, Henderson and Pabis, Page, Weibull and Fick's second law models were used to describe drying curves. Using osmodehydration as a drying pre-treatment decreased the drying time and product shrinkage with respect to fresh product. Page's model was the best fit ($R^2 > 0.948$), followed by Weibull and Henderson and Pabis model ($R^2 > 0.907$). The Newton model was not adequate to describe drying processes with high mass transfer rate ($R^2 > 0.775$). For the diffusive model, it is required to characterize the real dimensions of the product during the process to achieve a $R^2 > 0.917$. Effective water diffusivity of fresh and osmodehydrated apple was 6.77×10^{-10} and 3.77×10^{-10} m²/s, respectively.

Keywords: convective drying, diffusivity, mass transfer, osmotic dehydration, sucrose solution

1. Introduction

Convective drying removes water from a wet product by the simultaneous transfer of heat and mass [1, 2]. The process involves using a stream of hot air (provided by an electrical resistance and a propeller fan) that distributes the temperature to one or more drying chambers containing trays with samples [3]. The hot air increases the energy transfer over the product surface by convection and can reach the product interior by conduction or diffusion depending on the product structure and dimensions [4, 5]. At the same time, product moisture is transferred from the product surface to the air by convection and from the product interior by diffusion, convection or capillarity [1]. Thus, the drying rate and the quality of the dried product depend on the external conditions (temperature, relative humidity, airflow speed and

direction) and internal conditions (geometry, thickness, porosity and chemical composition of the product) of the process [6].

The main advantages of the industrial use of convective drying are that it allows processing large quantities of foods, the heat distribution within the system is uniform, and the cost of the equipment is relatively low [7, 8]. However, drawbacks include low process speed, significant loss of product quality (shrinkage, deformation, water holding capacity, color and flavor) and high energy consumptions when working at high temperatures [4, 9–11]. Damage to product quality could be minimized by using low process temperatures, reducing detrimental reactions, and improving the quality of dried foods. Unfortunately, this increases the duration of treatment and the energy costs of the process [12]. On the other hand, osmotic dehydration, blanching or ultrasound are pre-treatments used to prevent enzymatic browning, morphometrical changes, or add solutes to modify flavor, color, texture, etc. [13–15].

Osmodehydration of fruits and vegetables is a mass transfer process in which food is immersed in a hypertonic solution (of one or more solutes) known as osmotic solution. During this process, the product undergoes partial dehydration, thereby increasing its solids content [16]. However, the mass transfer process stops when the difference between the water activity of the osmotically dehydrated sample and the osmotic solution is not significant [17, 18]. This pre-treatment of drying is widely used because it helps to reduce consumption of energy and drying time [19, 20].

The kinetics of the dehydration process is a complex heat and mass transfer phenomenon [21]. Models of dehydration of food materials provide simple representations of a complex geometric solid process, which have proven to be very useful for drying system's design, construction, and operation. To this end, these transport properties can be monitored through kinetic models of drying [22].

Theoretical, empirical, and semi-theoretical models have been reported to describe dehydration process kinetics for biological materials. The choice of the model is influenced by the behavior of experimental data of the process according to the type of material, product geometry, equipment technology and process conditions [23]. Theoretical models take into account the internal resistance to heat and energy transfer; in contrast, empirical and semi-theoretical models, only consider the external resistance to transfer at the interface of the food matrix and the drying system [24], also provide greater understanding of the transport processes by better fitting the experimental data [25].

Therefore, this work used different mathematical models to describe the drying kinetics of osmodehydrated apple cubes (Granny Smith var.) with sucrose solution.

2. Methodology

2.1 Materials and methods

Fresh apple (*Malus domestica* L.) of var. Granny Smith was locally purchased (San Andrés Cholula, Puebla, Mexico). Apples were washed, sanitized by immersion into peracetic acid solution (100 mg/L) for 10 min, peeled, and cut into cubes (1.2 cm × 1.2 cm × 1.2 cm) with an industrial vegetable cutter. Osmotic solutions were prepared with food-grade granulated refined sucrose and distilled water.

2.2 Osmodehydration experiments

Apple cubes (w_0) were immersed in the osmotic solution (40, 50 and 60°Brix sucrose) at a mass ratio of 1:10 (to avoid excessive dilution), and at a temperature of 40°C [25] until an equilibrium state of mass transfer (without significative differences between water activity of osmotic solution and product) was reached. The mass transfer equilibrium state was determinate by measuring water activity in a digital hydrometer (Aqua Lab, 4TEV, USA) [17]. Then, the osmodehydrated apple cubes were removed from the osmotic solution. The excess liquid was removed with absorbent paper, weighed (w_{OD}), and the moisture content (Y_0 and Y_{DO} , fresh and osmodehydrated product, respectively) was determined to calculate the water loss (W), Eq. (1), and solids gain (G) of the product, Eq. (2) [26]. All the osmodehydration experiments were carried out in triplication.

$$W = \frac{w_0 Y_0 - w_{OD} Y_{OD}}{w_0} \quad (1)$$

$$G = \frac{w_{OD}(1 - Y_{OD}) - w_0(1 - Y_0)}{w_0} \quad (2)$$

2.3 Drying experiments

2.3.1 Drying curves

Fresh and osmodehydrated apple cubes were subjected to convective drying in a dryer designed and built at the Universidad de las Américas Puebla. The dryer includes a fan (of axial flow), an electric device with two 1000-W electric resistances, and a drying chamber with three tray perforations (28 × 18 cm). The perforations allow the free passage of air, thus drying both sides of the samples. The samples were dried at 50°C and air velocity of 3.5 m/s [4, 5] until constant weight (± 0.05 g). A digital balance was used to record the evolution of sample weight during the process. Weight data during the process was used to estimate the kinetics of free moisture fraction (Ψ) using the Eq. (3). Experimental kinetics of drying were carried out in triplicated.

$$\Psi = \frac{mc_{pt} - mc_{p\infty}}{mc_{p0} - mc_{p\infty}} \quad (3)$$

Where mc_{p0} is initial moisture content, mc_{pt} is moisture content at time t based on recoded weight, and $mc_{p\infty}$ is equilibrium moisture content.

2.3.2 Evolution of dimensionless volume

Dimensionless volume was analyzed using image processing developed by González-Pérez et al. [4]. Briefly, a digital image of slices (1 mm-thick) of a group of 5 samples corresponding to certain free moisture fractions (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9) was acquired. The 4608 × 3456 pixels digital image was taken with a camera (Coolpix L810, Nikon Corp, Japan) using natural light and blue background (like contrast). ImageJ® (ver. 1.49 h, Dresden, Germany) was used to analyze the

bidimensional images. The measure of contour and surface (S) gave information about product shape [27]. Finally, the evolution of dimensionless estimated with Eq. (4) was fitted with Eq. (5), and characteristic length (L) was estimated with Eq. (6). Eq. (6) will be used in the next section to estimate the evolution of effective diffusion.

$$\frac{V}{V_0} = \frac{S^{3/2}}{S_0^{3/2}} \quad (4)$$

$$\frac{V}{V_0} = m + (1 - m)\Psi^n \quad (5)$$

$$L = (V_0(m + (1 - m)\Psi^n))^{1/3} \quad (6)$$

Where V and S are the volume (m) and sample surface at time t , V_0 and S_0 are the volume (m) and sample surface at initial time, m and n are dimensionless parameters of the volume model.

2.4 Modeling of drying curves

2.4.1 Mathematical models

Mathematical models have been developed to optimize the dryers and/or predict the quality of the dried product [28, 29]. To find out the best mathematical model to predict drying curves some thin-layer drying models were used, including Newton [30], Henderson and Pabis [31], Page [32] and Weibull [33] model according to Eqs. (7)–(10), respectively. Non-linear regression was used to estimate model parameters. The correlation coefficient (R^2) and root mean square error (RMSE) were estimated to evaluate the goodness of fit of selected models.

$$\Psi = e^{(-k_1 t)} \quad (7)$$

$$\Psi = ae^{(-k_2 t)} \quad (8)$$

$$\Psi = e^{(-k_3 t^b)} \quad (9)$$

$$\Psi = e^{-(t/k_4)^c} \quad (10)$$

Where k_1 , k_2 and k_3 are rate constants in 1/s for Newton, Henderson and Pabis, and Page model, respectively; k_4 is scale parameter in s; a and b are empirical parameters (dimensionless) of Henderson and Pabis model, respectively; c is shape-parameter of Weibull model.

2.4.2 Effective water diffusion

Effective water diffusion was determined with the Fick's second law for a cubic geometry according to Eqs. (11)–(12) [34, 35]. The solution of Eq. (11) considers an initial homogenous water distribution, a constant water diffusivity parameter, a solid sample with no volume change, and an isothermal process [34].

$$\Psi = \left(\frac{8^3}{\pi^2} \sum_{i=1}^{\infty} \frac{1}{(2i+1)^2} e^{\left(\frac{-(2i+1)^2 \pi^2}{4} \tau \right)} \right)^3 \quad (11)$$

$$\tau = \frac{Dt}{0.5L} \quad (12)$$

Where τ is the Fourier number (dimensionless), D is the effective water diffusivity (m^2/s), t is the process time (s), and L is the thickness of the sample (m).

The modified slope method was used to calculate the diffusivity of water in the product considering its shrinkage. Briefly, the τ term of Eq. (11) was replaced by a more general definition (θ), as shown in Eq. (13) [36].

$$\theta = \int_0^t \frac{D(t)}{(0.5L(t))^2} dt = \int_0^t \frac{D(\Psi)}{(0.5L(\Psi))^2} dt \quad (13)$$

Where $D(t)$ and $L(t)$ are the diffusivity and sample thickness as a function of time, respectively. $D(\Psi)$ and $L(\Psi)$ are the diffusivity and sample thickness as a function of free moisture fraction, respectively.

2.5 Modeling of drying curves

Fick's second law solutions were performed with the Matlab software and its Statistic Toolbox 7.3 (Matlab R2020b, MathWorks Inc., Natick, MA, USA). Experimental data were analyzed with an ANOVA using Minitab v.17 Statistical Software (Minitab Inc., State College, PA, USA) and the Tukey test was used to analyze the mean comparisons considering a 95% confidence.

3. Results

3.1 Pre-treatment

At 1440 min for all the experiments, the water activity of osmodehydrated samples and sucrose solutions had non-significant differences ($p > 0.05$), which imply an equilibrium state of mass transfer [17]. The initial water activity of apples decreased from 0.983 ± 0.004 to values between 0.914 and 0.958 (**Table 1**). The lowest water activity value was reached with the most concentrated osmotic solution, implying that increased sucrose concentration favors mass transfer [17, 37]. This behavior can be observed with the mass transfer parameters in which water loss values of 0.492 to 0.635 g/g were found with 40–60°Brix concentrations, respectively. In the case of solute gain, values of 0.130–0.203 g/g were reached using 40–60°Brix concentrations, respectively. Likewise, water loss was more significant than solute gain in all experiments, similar to found by others investigations during osmodehydration of foods with sucrose [37–39].

Osmodehydrated samples shrank in all the experiments ($p < 0.05$). The volume reduction of the osmodehydrated samples in the equilibrium state was 53–56%. The shrinkage of the samples increased ($p < 0.05$) with increasing concentration of the osmotic agent (**Table 1**), some authors associate this behavior with increased water

Sample ^a	Moisture content (g/g fresh product)		a_w ^b		$L \times 10^3$ (m) ^c	
	Initial	Final	Initial	Final	Initial	Final
Fresh	0.86 ± 0.15 ^A	0.05 ± 0.01 ^A	0.983 ± 0.004 ^A	0.259 ± 0.035 ^C	6.09 ± 0.44 ^A	4.05 ± 0.09 ^C
OD 40	0.68 ± 0.05 ^B	0.13 ± 0.01 ^B	0.958 ± 0.001 ^B	0.626 ± 0.020 ^B	4.74 ± 0.15 ^B	4.70 ± 0.15 ^A
OD 50	0.60 ± 0.01 ^C	0.13 ± 0.03 ^B	0.939 ± 0.003 ^C	0.631 ± 0.025 ^B	4.65 ± 0.05 ^B	4.40 ± 0.02 ^B
OD 60	0.52 ± 0.05 ^D	0.19 ± 0.03 ^A	0.914 ± 0.001 ^D	0.668 ± 0.012 ^A	4.63 ± 0.04 ^B	4.10 ± 0.04 ^C

^aOD, osmodehydrated samples (40 °C and 1440 min) with 40, 50 y 60 °Brix sucrose solutions, data expressed as mean ± standard deviation of three independent experiments
^bwater activity at 25.02 ± 0.40
^cL, characteristic length.

Table 1.
Estimated diffusion coefficients during convective drying of fresh and osmodehydrated apple cubes in sucrose solution.

loss, which reduce the volume of the sample [27]. Souraki et al. [40] found a 75–80% volume reduction of osmodehydrated apple slabs with sucrose solutions of 30–50°Brix at different temperatures (30–50°C). In addition, the authors observed a significative ($p < 0.05$) volume reduction due to the increase of temperature and concentration of osmotic agent.

3.2 Drying

Figure 1 shows the free moisture fraction curves of convective drying (at 50°C and a constant air velocity of 3.5 m/s) of fresh and osmodehydrated (OD) apple cubes with sucrose solutions. The drying time for all the drying experiments was different ($p < 0.05$). The convective drying process had a longer duration for the fresh samples (26,400 s), followed by the osmodehydrated samples with the 40°Brix (12,900 s), 50° Brix (8400 min) and 60°Brix (7500 min) sucrose solutions. The pre-treated samples had a lower ($p < 0.05$) initial moisture content compared with the fresh sample (**Table 1**). However, after the drying process, the osmodehydrated samples showed a higher ($p < 0.05$) moisture content than the untreated dried samples. This behavior is accentuated with pre-treatments with higher concentration of the osmotic solution. Due to that, the increasing of osmotic solutions concentration increases the solutes impregnation [13, 41]. Some authors have reported that the solutes present on the outer part of the product form a crust that decreases the water transfer rate [42]. Assis et al. [25] osmodehydrated apple cubes in sucrose and sorbitol solutions (60° Brix, considering an osmotic solution to product ratio of 4, 1, 60°C, 50 rpm agitation for 8 h) and reached $aw = 0.851$ – 0.949 . Untreated apple cubes took 21,600 s to dry, and osmodehydrated ones reduced from 20 to 22% drying time using sucrose and 30–39% using sorbitol. Sethi and Kaur [38] observed the same behavior during drying fresh and osmodehydrated pineapple, concluding that convective drying time with hot air decreases with increasing solids gain during osmotic dehydration.

The dimensionless volume (V/V_0) and average perimeter were determined at different free moisture fraction (**Figures 2 and 3**, respectively), to investigate the morphometric characteristics of the cubes during the drying process. According to **Figure 2**, regardless of the experiment conditions, all samples subjected to convective drying shrank from free moisture fractions of 0.9. However, untreated dry samples

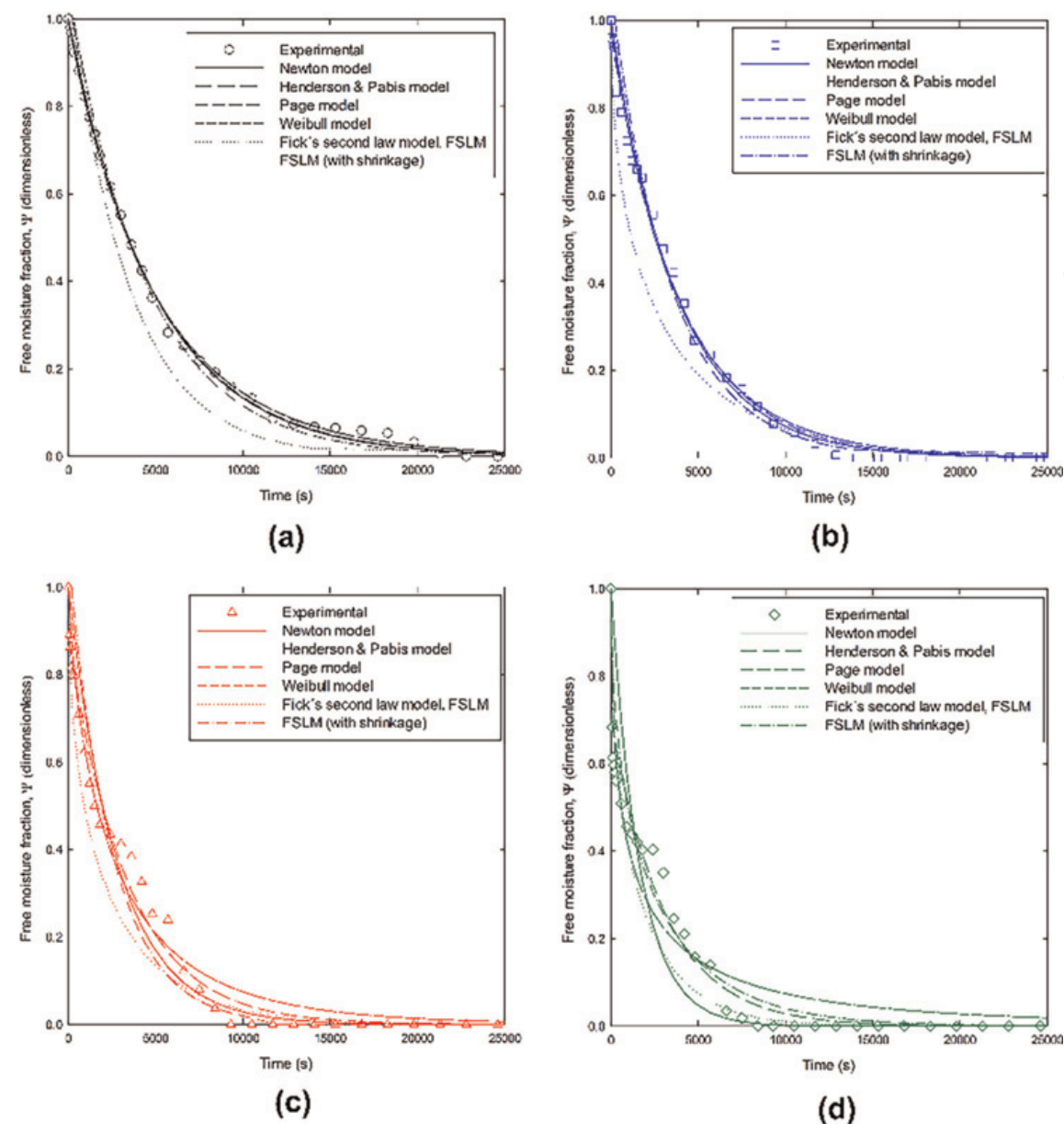


Figure 1. Experimental drying profiles and fitting numerical of drying models during convective drying of (a) fresh and osmodehydrated apple with (b) 40°brix (c) 50°brix and (d) 60°brix sucrose solutions.

(fresh) shrank more (71% compared to the initial volume) than treated dry samples (10–45% compared to initial volume). Osmodehydrated samples increased their V/V_0 reduction during the drying process at lower concentrations of the osmotic agent, mainly because they had a lower water loss and solute gain during the osmodehydration process. However, according to **Figure 3** and **Table 1**, the shrinkage of the dried samples previously treated by osmodehydration increased at high concentrations of the osmotic agent.

All the dry samples showed a shape change (**Figure 3**) known as the “corner effect”, consisting of an accentuated shrinkage at mid-length and a less pronounced at the edges. This behavior was similar to those obtained for pumpkin cylinder or chayote slices [27, 43]. In addition, the fresh dried samples showed higher deformation than the dried osmodehydrated samples (**Figure 3**). Osmotic dehydration can reduce or prevent the collapse of the structure of osmodehydrated samples due to solute incorporation [13]. In addition, a moderate rate of solute transfer is important to avoid deformation of

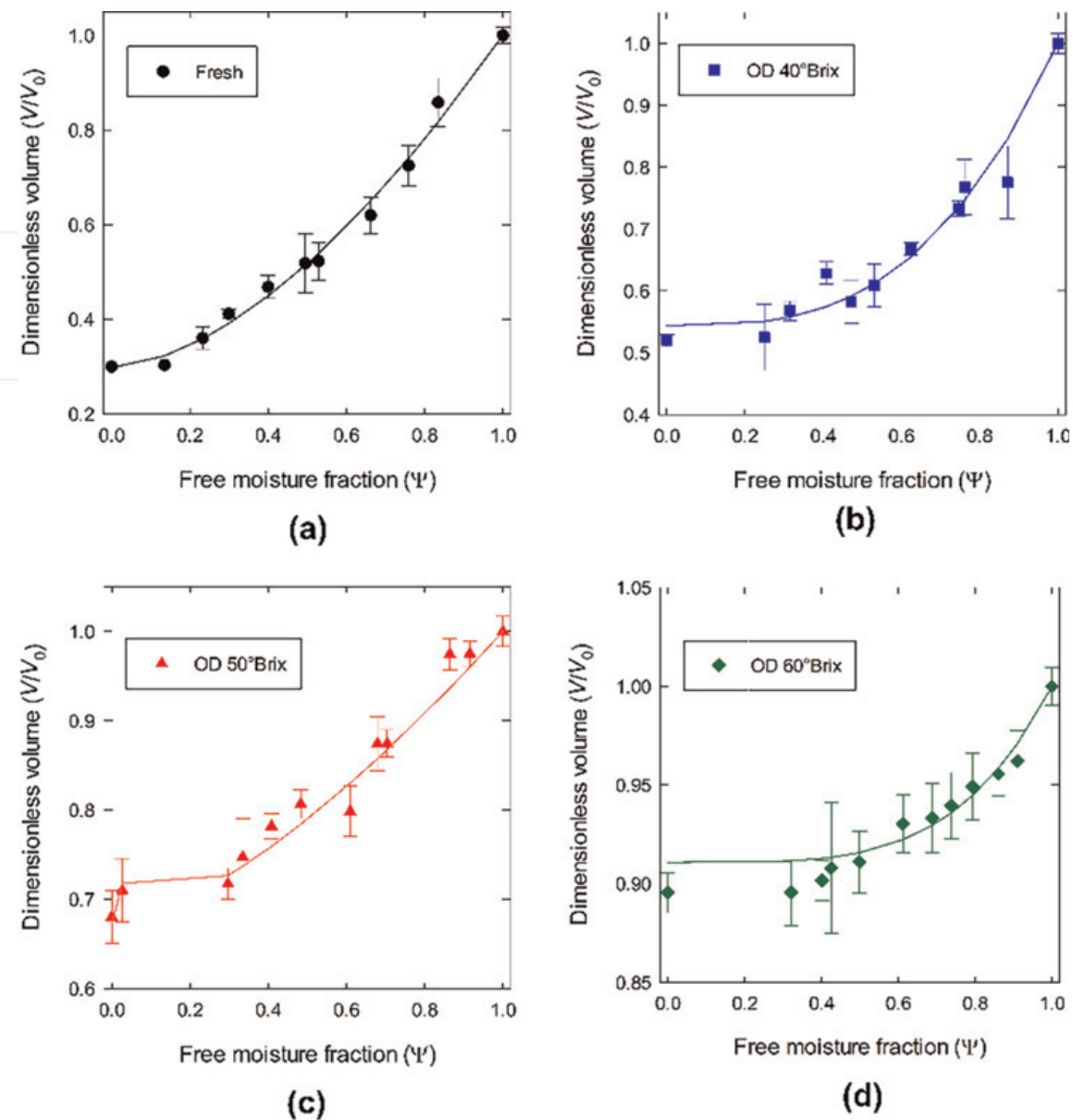


Figure 2. Experimental data (dots) and simulated results (line) of volumetric shrinkage of slabs in relation to dimensionless moisture on dry basis. (a) fresh, and osmodehydrated apple with (b) 40°brix (c) 50°brix and (d) 60°brix sucrose solutions.

osmodehydrated products before the convective drying process, as in the case of osmodehydrated sample at 60°Brix. This rate can be controlled by decreasing the osmotic agent's concentration, modifying the viscosity of the osmotic agent (adding pectins), or adjusting the process temperature [17]. Another alternative is to increase the impregnated solutes using vacuum impregnation, high hydrostatic pressures, among other non-thermal methods that allow modifying the rate of solute transfer [17, 44, 45].

The evolution of the dimensionless volume of the samples subjected to convective drying as a function of the free moisture fraction was described with the model $V/V_0 = m + (1-m) \Psi^n$. The fit allowed to adequately describe the experimental data with a correlation coefficient of $R^2 > 0.870$. The parameter m represents the fraction in which the studied response changes during drying. Low values of the parameter m indicate that the shrinkage of the sample during the process was fast. According to **Table 2**, m increased with the osmotic solution concentration in the case of the

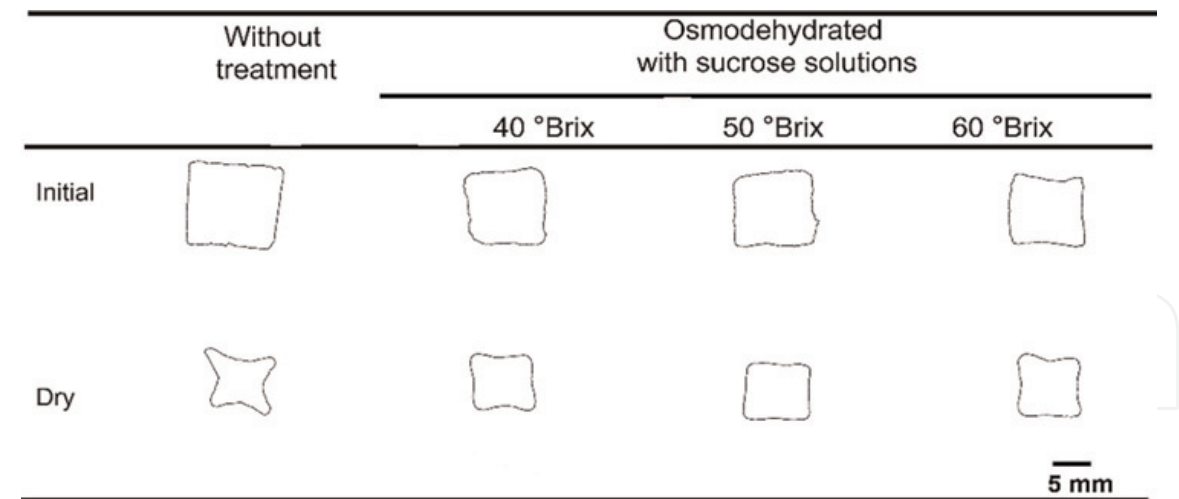


Figure 3. Morphometric variations of fresh and osmodehydrated apple cubes processed by convective. Average perimeter corresponding to five samples.

Model parameter*	Sample			
	Fresh	Osmodehydrated		
		40°Brix	50°Brix	60°Brix
m	0.30	0.54	0.67	0.91
n	1.67	3.00	1.47	4.17
RMSE	0.006	0.009	0.009	0.003
R ²	0.992	0.950	0.897	0.870

*V/V₀ = m + (1-m) Ψⁿ.

Table 2. Fit parameters of dimensionless volume model.

osmodehydrated samples, and this implied that the presence of solutes reduced shrinkage during the drying process [4].

3.3 Mathematical modeling of drying curves

Table 3 shows the kinetic parameters obtained for all the proposed models and their corresponding statistical values (R² and RMSE). Newton’s model adequately described the experimental data for drying kinetics of the fresh sample (R² = 0.997) and osmotic agent concentrations below 50°Brix (R² > 0.965). However, Newton’s model did not adequately describe (R² = 0.775) the kinetics of the samples pre-treated with 60°Brix; this is because the model does not allow to describe kinetics with high mass transfer rates [46]. The Henderson and Pabis, Page, and Weibull models had a good fit (R² > 0.90). In general, the model that best described the experimental data was the Page model with R² > 0.948. Parameter k₁, k₂ and k₃ correlate the ability of the drying process to exclude moisture from the food and moisture diffusivity [46]. Table 3 shows that the parameter values of k₁ and k₂ increase with samples pre-treated in solutions with higher solids content; this implies that the increase in drying

Model parameter	Sample			
	Fresh	Osmodehydrated		
		40°Brix	50°Brix	60°Brix
Newton model				
k ₁	2.02 × 10 ⁻⁴	2.64 × 10 ⁻⁴	3.57 × 10 ⁻⁴	6.13 × 10 ⁻⁴
R ²	0.997	0.995	0.965	0.775
RMSE	0.003	0.004	0.009	0.020
Henderson and Pabis model				
k ₂	2.00 × 10 ⁻⁴	2.52 × 10 ⁻⁴	2.94 × 10 ⁻⁴	3.22 × 10 ⁻⁴
a	0.99	0.97	0.88	0.70
R ²	0.997	0.996	0.985	0.946
RMSE	0.003	0.003	0.006	0.009
Page model				
k ₃	3.26 × 10 ⁻⁴	4.73 × 10 ⁻⁴	5.26 × 10 ⁻⁴	5.71 × 10 ⁻⁴
b	0.94	0.93	0.73	0.46
R ²	0.998	0.996	0.985	0.948
RMSE	0.003	1.48 × 10 ⁻⁴	0.004	0.009
Weibull model				
k ₄	4959.37	3770.50	2676.41	1256.72
c	0.95	0.93	0.72	0.46
R ²	0.907	0.996	0.985	0.948
RMSE	0.003	0.004	0.002	0.009
Fick's second law (without shrinkage)				
D _{eff}	14.1 × 10 ⁻¹⁰	6.37 × 10 ⁻¹⁰	8.11 × 10 ⁻¹⁰	11.5 × 10 ⁻¹⁰
R ²	0.648	0.707	0.778	0.879
RMSE	0.035	0.032	0.025	0.015
Fick's second law (with shrinkage)				
D _{eff}	6.77 × 10 ⁻¹⁰	3.77 × 10 ⁻¹⁰	4.63 × 10 ⁻¹⁰	7.03 × 10 ⁻¹⁰
R ²	0.960	0.972	0.981	0.917
RMSE	0.021	0.010	0.015	0.009
k ₁ , k ₂ and k ₃ are rate constant in 1/s for Newton, Henderson and Pabis and Page model, respectively; k ₄ , scale parameter in s; a and b are empirical parameters (dimensionless) of Henderson and Pabis model; c is shape-parameter of Weibull model; and D _{eff} is effective water diffusion (m ² /s).				

Table 3.
Fit parameters of models applied to drying curves of fresh and osmodehydrated apple cubes.

rate is favored by pre-treatment. This same behavior occurred with the diffusivity values, a tendency with increasing of concentration of osmotic agent for each of these parameters.

The diffusive model of convective drying solution allowed predicting the water distribution of fresh and osmodehydrated apple cubes with sucrose solutions. The

water diffusion coefficients of fresh cubes considering a constant volume were $14.1 \times 10^{-10} \text{ m}^2/\text{s}$ and for the pre-treated samples were found between $6.37\text{--}11.5 \times 10^{-10} \text{ m}^2/\text{s}$ (**Table 3**). However, according to the evolution of the dimensionless volume (**Figure 3**) there is shrinkage of product during the process, which explains why the correlation coefficients ($R^2 > 0.648$) were not good. Also, R^2 decreased with increasing samples shrinkage. It is clear that considering the evolution of the dimensionless volume of the product in the estimation of the diffusion improves the model's fit ($R^2 > 0.917$). The diffusion equation considers the product's dimensions, so that shrinkage during the process affects the estimation of diffusion coefficients [4, 47, 48]. The D_{eff} values of the drying process were $6.77 \times 10^{-10} \text{ m}^2/\text{s}$ and between $3.77\text{--}7.03 \times 10^{-10} \text{ m}^2/\text{s}$ for fresh and osmodehydrated samples, respectively. Similar to the values obtained for the drying process for yam flour, chayote slabs, apple, tomatoes, cucumbers or apricot [43, 46, 48–50]. In addition, D_{eff} values without consider real dimensions caused overestimation of 108.27, 68.97, 75.16 and 63.58%.

4. Conclusions

The application of osmodehydration as a drying pre-treatment decreased the drying time and shrinkage of the product compared with the fresh product. Increasing the osmotic agent (sucrose solution) concentration increased product shrinkage during the osmodehydration process.

The models used allowed describing the drying kinetics. Page's model was the best fit, followed by Weibull model and Henderson and Pabis model. These models are a great alternative to analyze the mass transfer rate. However, the Newton model was not adequate to describe drying processes with mass transfer rate. In the case of the diffusive model, it is required to characterize the real dimensions of the product during the whole process to achieve a better fit.

Acknowledgements

Author González-Pérez Julio Emmanuel (<https://orcid.org/0000-0002-0218-6158>) gratefully acknowledges financial support for his PhD studies from Universidad de las Americas Puebla (UDLAP) and the National Council for Science and Technology (CONACyT) of Mexico.

Conflict of interest

Authors declare no conflict of interest associated with this research.

Notes/thanks/other declarations

None.


IntechOpen

Author details

Julio Emmanuel González-Pérez, Nelly Ramírez-Corona and Aurelio López-Malo*
Department of Chemical, Food, and Environmental Engineering, Universidad de las
Américas Puebla, Ex Hacienda Santa Catarina Mártir, Puebla, Mexico

*Address all correspondence to: aurelio.lopezm@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Castro AM, Mayorga EY, Moreno FL. Mathematical modelling of convective drying of fruits: A review. *Journal of Food Engineering*. 2018;223:152-167
- [2] Dehghannya J, Bozorgi S, Heshmati MK. Low temperature hot air drying of potato cubes subjected to osmotic dehydration and intermittent microwave: Drying kinetics, energy consumption and product quality indexes. *Heat and Mass Transfer*. 2018; 54(4):929-954
- [3] Oliveira SM, Brandão TRS, Silva CLM. Influence of drying processes and pretreatments on nutritional and bioactive characteristics of dried vegetables: A review. *Food Engineering Reviews*. 2016;8(2):134-163
- [4] González-Pérez JE, López-Méndez EM, Ochoa-Velasco CE, Ruiz-López II. Mass transfer and morphometric characteristics of fresh and osmodehydrated white mushroom pilei during convective drying. *Journal of Food Engineering*. 2019;262:181-188
- [5] López-Méndez EM, Ortiz-García-Carrasco B, Ruiz-Espinosa H, Sampieri-Croda A, García-Alvarado MA, Ochoa-Velasco CE, et al. Effect of shape change and initial geometry on water diffusivity estimation during drying of gel model systems. *Journal of Food Engineering*. 2018;216:52-64
- [6] Sabarez HT. Thermal drying of foods. In: Rosenthal A, Deliza R, Welte-Chanes J, Barbosa-Cánovas GV, editors. *Fruit Preservation*. New York: Springer New York; 2018. pp. 181-210
- [7] Onwude DI, Hashim N, Janius R, Abdan K, Chen G, Oladejo AO. Non-thermal hybrid drying of fruits and vegetables: A review of current technologies. *Innovative Food Science and Emerging Technologies*. 2017;43: 223-238
- [8] Antal T, Kerekes B, Sikolya L, Tarek M. Quality and drying characteristics of apple cubes subjected to combined drying (FD pre-drying and HAD finish-drying): Apple drying. *Journal of Food Processing & Preservation*. 2015;39(6):994-1005
- [9] Zecchi B, Gerla P. Effective diffusion coefficients and mass flux ratio during osmotic dehydration considering real shape and shrinkage. *Journal of Food Engineering*. 2020;274:1-8
- [10] Estévez-Sánchez KH, González-Pérez JE, Ochoa-Velasco CE, García-Alvarado MA, Cruz-González D, Sampieri A, et al. Point set registration for reduced geometry mismatch during estimation of mass transfer properties in osmotic dehydration of complex-shaped foods. *Drying Technology*. 2020;38(4): 506-517
- [11] González-Pérez JE, Guerrero-Beltrán JÁ. Tomatillo or husk tomato (*Physalis philadelphica* and *Physalis ixocarpa*): A review. *Scientia Horticulturae*. 2021;288: 110306
- [12] de Lima AG, da Silva JV, Pereira EM, dos Santos IB, de Lima WM. Drying of bioproducts: Quality and energy aspects. In: Delgado JMPQ, Barbosa de Lima AG, editors. *Drying and Energy Technologies*. Cham: Springer International Publishing; 2016. pp. 1-18
- [13] González-Pérez JE, López-Méndez EM, Luna-Guevara JJ, Ruiz-Espinosa H, Ochoa-Velasco CE, Ruiz-López II. Analysis of mass transfer and morphometric characteristics of white

mushroom (*Agaricus bisporus*) pilei during osmotic dehydration. Journal of Food Engineering. 2019;**240**:120-132

[14] Pérez-Won M, Lemus-Mondaca R, Tabilo-Munizaga G, Pizarro S, Noma S, Igura N, et al. Modelling of red abalone (*Haliotis rufescens*) slices drying process: Effect of osmotic dehydration under high pressure as a pretreatment. Innovative Food Science and Emerging Technologies. 2016;**34**:127-134

[15] Nowacka M, Laghi L, Rybak K, Dalla Rosa M, Witrowa-Rajchert D, Tylewicz U. Water state and sugars in cranberry fruits subjected to combined treatments: Cutting, blanching and sonication. Food Chemistry. 2019;**299**: 125122

[16] Akharume F, Smith A, Sivanandan L, Singh K. Recent progress on osmo-convective dehydration of fruits. SDRP Journal of Food Science and Technology. 2019;**4**(9):956-969

[17] González-Pérez JE, Ramírez-Corona N, López-Malo A. Mass transfer during osmotic dehydration of fruits and vegetables: Process factors and non-thermal methods. Food Engineering Reviews. 2021;**13**(2):344-374

[18] Pacheco-Angulo H, Herman-Lara E, García-Alvarado MA, Ruiz-López II. Mass transfer modeling in osmotic dehydration: Equilibrium characteristics and process dynamics under variable solution concentration and convective boundary. Food and Bioproducts Processing. 2016;**97**:88-99

[19] da Costa Ribeiro AS, Aguiar-Oliveira E, Maldonado RR. Optimization of osmotic dehydration of pear followed by conventional drying and their sensory quality. LWT—Food Science and Technology. 2016;**72**:407-415

[20] Dermesonlouoglou EK, Bimpilas A, Andreou V, Katsaros GJ, Giannakourou MC, Taoukis PS. Process optimization and kinetic modeling of quality of fresh-cut strawberry cubes pretreated by high pressure and osmosis. Journal of Food Processing & Preservation. 2017;**41**(5):1-14

[21] Karathanos VT, Belessiotis VG. Application of a thin-layer equation to drying data of fresh and semi-dried fruits. Journal of Agricultural Engineering Research. 1999;**74**(4): 355-361

[22] de Lima AG, Delgado JMP, Neto SR, Franco CM. Intermittent drying: Fundamentals, modeling and applications. In: Delgado JMPQ, Barbosa de Lima AG, editors. Drying and Energy Technologies. Cham: Springer International Publishing; 2016. pp. 19-41

[23] Tayyab Rashid M, Ahmed Jatoti M, Safdar B, Wali A, Muhammad Aadil R, Sarpong F, et al. Modeling the drying of ultrasound and glucose pretreated sweet potatoes: The impact on phytochemical and functional groups. Ultrasonics Sonochemistry. 2020;**68**:105226

[24] Onwude DI, Hashim N, Janius RB, Nawi NM, Abdan K. Modeling the thin-layer drying of fruits and vegetables: A review: Thin-layer models of fruits and vegetables. Comprehensive Reviews in Food Science and Food Safety. 2016; **15**(3):599-618

[25] Assis FR, Morais RMSC, Morais AMMB. Mathematical modelling of osmotic dehydration kinetics of apple cubes. Journal of Food Processing & Preservation. 2017;**41**(3):1-16

[26] Cichowska J, Figiel A, Stasiak-Róžańska L, Witrowa-Rajchert D. Modeling of osmotic dehydration of

apples in sugar alcohols and dihydroxyacetone (DHA) solutions. Food. 2019;8(1):1-17

[27] Mayor L, Moreira R, Sereno AM. Shrinkage, density, porosity and shape changes during dehydration of pumpkin (*Cucurbita pepo* L.) fruits. Journal of Food Engineering. 2011;103(1):29-37

[28] Ayetigbo O, Latif S, Abass A, Müller J. Osmotic dehydration kinetics of biofortified yellow-flesh cassava in contrast to white-flesh cassava (*Manihot esculenta*). Journal of Food Science and Technology. 2019;56(9):4251-4265

[29] Manguiera ER, de Lima AGB, de Assis CJ, Costa NA, de Souza CC, de Abreu AKF, et al. Foam-mat drying process: Theory and applications. In: Delgado JMPQ, Barbosa de Lima AG, editors. Transport Processes and Separation Technologies. Cham: Springer International Publishing; 2021. pp. 61-87

[30] Ndukwu MC, Dirioha C, Abam FI, Ihediwa VE. Heat and mass transfer parameters in the drying of cocoyam slice. Case Studies in Thermal Engineering. 2017;9:62-71

[31] Ayala-Aponte AA, Molina-Cortés A, Serna-Cock L. Osmotic dehydration of green mango samples (*Mangifera indica* L., *Filipino* Var.) in ternary solutions. Vitae. 2018;25(1):8-16

[32] Chambi HNM, Lima WCV, Schmidt FL. Osmotic dehydration of yellow melon using red grape juice concentrate. Food Science and Technology. 2016;36(3):468-475

[33] Lemus-Mondaca R, Pizarro-Oteíza S, Perez-Won M, Tabilo-Munizaga G. Convective drying of osmo-treated abalone (*Haliotis rufescens*) slices:

Diffusion, modeling, and quality features. Journal of Food Quality. 2018; 2018:1-10

[34] Bezerra Pessoa TR, de Lima AGB, Martins PC, Pereira VC, Alves TCO, da Silva ES, et al. Osmo-convective dehydration of fresh foods: Theory and applications to cassava cubes. In: Delgado JMPQ, Barbosa de Lima AG, editors. Transport Processes and Separation Technologies. Cham: Springer International Publishing; 2021. pp. 151-183

[35] Srikiatden J, Roberts JS. Measuring moisture diffusivity of potato and carrot (core and cortex) during convective hot air and isothermal drying. Journal of Food Engineering. 2006;74(1):143-152

[36] Ruiz-López II, Ruiz-Espinosa H, Arellanes-Lozada P, Bárcenas-Pozos ME, García-Alvarado MA. Analytical model for variable moisture diffusivity estimation and drying simulation of shrinkable food products. Journal of Food Engineering. 2012;108(3):427-435

[37] Lech K, Michalska A, Wojdyło A, Nowicka P, Figiel A. The influence of the osmotic dehydration process on physicochemical properties of osmotic solution. Molecules. 2017;22(12):1

[38] Sethi K, Kaur M. Effect of osmotic dehydration on physicochemical properties of pineapple using honey, sucrose and honey-sucrose solutions. International Journal of Engineering and Advanced Technology. 2019;9(1): 6257-6262

[39] Bera D, Roy L. Osmotic dehydration of litchi using sucrose solution: Effect of mass transfer. Journal of Food Processing and Technology. 2015;6:7

[40] Souraki BA, Ghavami M, Tondro H. Correction of moisture and sucrose

effective diffusivities for shrinkage during osmotic dehydration of apple in sucrose solution. Food and Bioproducts Processing. 2014;**92**(1):1-8

[41] Cichowska J, Samborska K, Kowalska H. Influence of chokeberry juice concentrate used as osmotic solution on the quality of differently dried apples during storage. European Food Research and Technology. 2018; **244**(10):1773-1782

[42] Sette P, Franceschinis L, Schebor C, Salvatori D. Fruit snacks from raspberries: Influence of drying parameters on colour degradation and bioactive potential. International Journal of Food Science and Technology. 2017; **52**(2):313-328

[43] Ruiz-López II, Huerta-Mora IR, Vivar-Vera MA, Martínez-Sánchez CE, Herman-Lara E. Effect of osmotic dehydration on air-drying characteristics of chayote. Drying Technology. 2010; **28**(10):1201-1212

[44] Alolga RN, Osaé R, Essilfie G, Saalia FK, Akaba S, Chikari F. Sonication, osmosonication and vacuum-assisted osmosonication pretreatment of Ghanaian garlic slices: Effect on physicochemical properties and quality characteristics. Food Chemistry. 2021; **343**:128535

[45] Luo W, Tappi S, Wang C, Yu Y, Zhu S, Dalla Rosa M, et al. Effect of High Hydrostatic Pressure (HHP) on the antioxidant and volatile properties of candied wumei fruit (*Prunus mume*) during osmotic dehydration. Food and Bioprocess Technology. 2019;**12**(1): 98-109

[46] Kamal MM, Ali MR, Shishir MRI, Mondal SC. Thin-layer drying kinetics of yam slices, physicochemical, and functional attributes of yam flour.

Journal of Food Processing Engineering. 2020;**43**:8

[47] Balzarini MF, Reinheimer MA, Ciappini MC, Scenna NJ. Mathematical model, validation and analysis of the drying treatment on quality attributes of chicory root cubes considering variable properties and shrinkage. Food and Bioproducts Processing. 2018;**111**: 114-128

[48] Aires KLCDA, da Silva WP, de Farias Aires JE, da Silva Júnior AF, Silva CMDPDSE. Apple osmotic dehydration described by three-dimensional numerical solution of the diffusion equation. Drying Technology. 2018;**36**(16):1970-1981

[49] Dermesonlouoglou EK, Pantelaiaki K, Andreou V, Katsaros GJ, Taoukis PS. Osmotic pretreatment for the production of novel dehydrated tomatoes and cucumbers. Journal of Food Processing Preservation. 2019;**43**:7

[50] Dermesonlouoglou EK, Giannakourou MC. Modelling dehydration of apricot in a non-conventional multi-component osmotic solution: Effect on mass transfer kinetics and quality characteristics. Journal of Food Science and Technology. 2018; **55**(10):4079-4089

Deep Eutectic Solvents: A Promising Technique to Anthocyanin Extraction for Food Coloring Agents

*Oscar Jiménez-González, Nelly Ramirez-Corona
and José Á. Guerrero-Beltrán*

Abstract

The extraction of natural pigments for food applications have increased by consumers and the food industry due to multiple health benefits and applications. New generations impulse changes to implement and investigate new types of solvents as an alternative to conventional organic solvents. Recently, deep eutectic solvents (DES) gain popularity in many fields, because of their advantages over organic solvents. DES are biodegradables, produced with natural compounds, safe for use, ecologically friendly, and obtained relatively at low cost. For the extraction of anthocyanins, few works have been reported. This review is focused on describing how the physicochemical characteristics of different types of DES affect the extraction of anthocyanins as well as the variables involved during extraction processes affect the obtained yield and their stability.

Keywords: Deep eutectic solvents (DES), anthocyanins, physicochemical characteristics, process variables

1. Introduction

The types of foods the consumers consume have taken greater relevance during the last decade [1, 2]. The preference for consuming natural, fresh, and uncrossed food products is not new. Particularly when the term “functional food” appeared in the foods and nutritional areas. Functional foods claimed to provide health benefits or cure some diseases after the incorporation to a food or modification of the food [3]. The food industry has taken this into account and is interested in increasing the consumption of natural compounds through natural ingredients in the food matrix. Due to this, researchers have investigated different compounds such as polyphenols, mainly colored polyphenols of vegetables, and their residues [4–6] in terms of extraction, stability, conservation.

Colored polyphenols include anthocyanins, betalains, and some flavonoids [7, 8]. Their molecular structure poses double bonds and acts as antioxidants [9, 10] against cardiovascular illness, cancer, diabetes, and heart and neurological ailments; also, some have exhibited anti-inflammatory and antimicrobial effects [11–13].

Anthocyanins are unstable to some factors such as pH, temperature, heat, enzymes, oxygen, and other compounds (co-pigments or metals) [14–16]. One or more of these factors could be involved during the anthocyanins extraction. Therefore, extraction type and process conditions play an important role in the recovery and further stability of anthocyanins. There exist different types of extraction methods; they are divided into physical (shearing force, ultrasound, microwaves, pulsed electric fields), chemical (decoction, maceration), and biological methods (enzymatic). Another way of classification is as conventional and non-conventional technologies [17, 18]. However, most of them require the use of at least one type of solvent. Mohd Fuad et al. [19] have mentioned that the selection of the solvent is crucial to recover the compound of interest due to the coexistence of many compounds with different chemical characteristics, polarities, wherein each compound has its own solubility in a solvent. The main solvents used for anthocyanin extraction are polar solvents (water, ethanol, acetone, acetic acid) [17]. However, these solvents present the disadvantage of being highly volatile, flammable, and corrosive as well as toxic, carcinogenic, and non-biodegradable [19]. In some cases, the use of solvents is excessive, expensive, requires filtering, multiple extractions, and techniques to remove the solvent using pressurized evaporation [17].

Recently, the deep eutectic solvents (DES), a new green alternative for organic solvents, are under research since they exhibit a promising capacity for extracting anthocyanins. DES consists of a mixture of two or more organic compounds which in combination present a reducing melting point, lower than each individual component, and exist as liquid form at room temperature [20, 21]. DES shows the advantage of being easily synthesized, inexpensive, biodegradable, and with low toxicity [22]. In addition, they have the capacity of being modified to improve extraction yields and be selective [23]. DES have been used to extract bioactive compounds such as flavonoids, polyphenols, and pigments [21].

2. Brief overview of anthocyanin

Anthocyanins are water soluble secondary metabolites that are present in the vacuoles of flowers, leaves, cereals, vegetables, and fruits [24–26]. These components belong to flavonoids and are a kind of phenolic compounds (polyphenolic). The molecular structure of anthocyanins contains three aromatic rings and multiple double bonds which gives nutraceutical properties as they can inhibit free radicals [27] acting as natural antioxidants. Anthocyanins possess protective and promoting effects on human health against neurodegenerative diseases, cancer, metabolic syndrome [12, 13].

Anthocyanins are dimers of an aglycone anthocyanidin and some sugar (xylose, galactose, arabinose, rutinose) or other compounds such as co-pigments, proteins, among others [28]. The anthocyanidin form is unstable and less soluble which is why the dimer form is the most common [29].

There are more than 700 anthocyanins and 23 anthocyanidins. The differences are in their chemical structure (number of hydroxyl or methoxyl groups, type, number, and position of sugar group bonded to anthocyanidin, and the presence of aliphatic or

aromatic acids bonded) [30, 31]. Most common anthocyanins are cyanidin (orange-red), delphinidin (purple-blue), pelargonidin (orange tones), malvidin (purple), peonidin (purple) and petudine (dark red or purple) [15, 28]. The tone and intensity of color depend on the number of hydroxyl and methoxyl groups. Hydroxylation increases blueness and methylation induces redness. Also, copigmentation enhances the color by a hyperchromic effect [32]. A bathochromic effect is observed when changes in pH of the solution or in the presence of copigments [15, 32].

3. What are them and why are so important?

DES were described by Abbot et al. [33] when combined choline chloride and urea (1:2 molar ratio), observing a decrease in the melting point of the mixture compared to that of each individual component (302 and 133°C, respectively). The eutectic mixture obtained was liquid at room temperature (melting point 12°C). **Figure 1** depicts the eutectic phenomena of two compounds. Components involved in the mixture interact via intermolecular forces mainly hydrogen bonding (except covalent or ionic bonds) [34, 35].

DES consists of a mix of halide salt or another hydrogen bond acceptor (HBA) and hydrogen bond donor (HBD) [22]. The combination and molar ratio of HBA-HBD to form a DES led to specific thermophysical and extraction characteristics [36]. Commonly preparation methods are i) *heating and stirring* (component plus heat over 60°C and magnetically stirring), ii) *evaporation* (predissolved components in water plus evaporation at 50°C in a rotatory evaporator), iii) *freeze-drying* (aqueous solution of both components submitted to lyophilization) [22, 23]. No additional solvent or purification step is necessary to synthesize DES [23]. The main HBA compound is choline chloride, which is mixed with other inexpensive and safe HBD like urea, ethylene glycol, and glycerol. Other examples, can be natural components present in vegetables such as sugars, polyalcohols, amino acids, organic acids and organic bases [22, 37, 38] (other examples are shown in **Figure 2**). The use of these natural

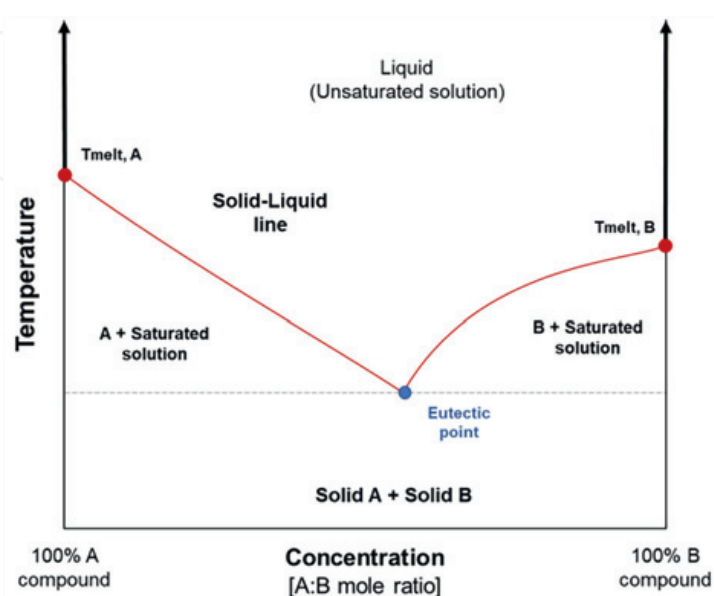


Figure 1.
Schematic representation of phase diagram of two components (A, B) of deep eutectic solvents.

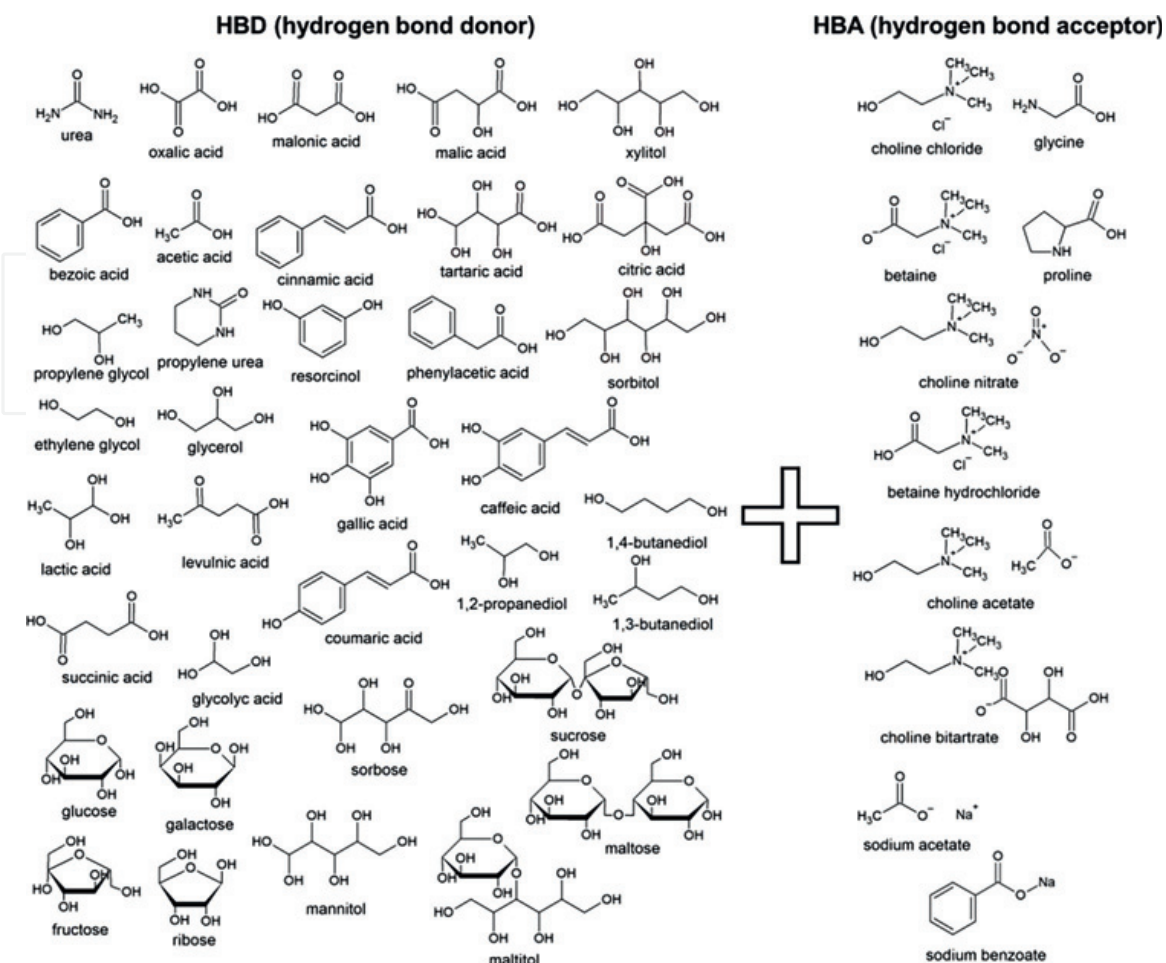


Figure 2.
Some HBD and HBA that can be used to form DES (NADES).

compounds introduces the concept of natural deep eutectic solvents (NADES), which are a classification of DES. NADES have the advantage of mimics the biological function of compounds mixtures in nature, playing specific roles in organisms [37]. For example, Vanda et al. [37] have mentioned that the transport of compounds, like flavonoids from certain tissues or cells, to another cell could be possible to fluids with similar composition to DES (and NADES).

Since eutectic phenomena were described, many authors have been used it for several purposes; DES has been used in fields such as chemistry (solvent for synthesis reactions), materials science (separation of azeotropes), physics, biology (in the isolation and fractionation of compounds), energy fuels (purification and processing in biodiesel production), agricultural (contaminant removal) and pharmacology [21, 22]. Furthermore, DES have the advantage of being green or ecologically friendly solvents [39]. DES are durable, biodegradable, have low toxicity, high solubility, and have relatively low cost [35].

Thus, when DES are used for the extraction, they have the ability to increase yield, selectivity, solubility, viscosity, and other important physicochemical properties [23, 40]. In addition, the use of DES in anthocyanin extraction could prevent thermal degradation that is common in extraction techniques. Besides, DES can improve the thermostability of anthocyanins once extracted [40]. Finally, DES can be easily eliminated, because during its formation/synthetization no chemical reaction occurs; then, the interactions of hydrogen bonding can be broken under appropriate

conditions [38] or by using microporous resins [20]. Furthermore, DES made with natural compounds can be considered safe for further application without further purification [41].

4. Application of DES in anthocyanins extraction

The tendency of use and research about anthocyanin extraction using DES or NADES is growing since 2015, as can be observed in **Figure 3**. Although, their use in the extractions of other nutraceutical compounds (phenolics, flavonoids, alkaloids, terpenoids) began some years before [21, 34, 42]. Now, DES have been studied in the extraction of anthocyanins of different parts of plants such as fruits, flowers and vegetables (**Table 1**). All examples cited in this work mention anthocyanins as primary compounds in the extraction), especially in the recovery of pigments from agro-industrial residues like the ones coming from the winemaking industry [69–72, 76] or juices industry [73].

Zuo et al. [21] have mentioned that due to the variety in molecule structure of the extracted compounds it has to take into account some considerations to make the extractions. For example, hydrophobicity or hydrophilicity, hydrogen-bonding capacity, polarity, viscosity, and the acidity tolerance range. One of the main advantages of DES is the feasibility of tuning their physical and chemical characteristics, in order to secure the desired extraction. In terms of anthocyanin extraction, a wide variety of DES has been used. The common composition of DES used is based on choline chloride (ChCl) as the hydrogen bond acceptor and polyalcohols, organic acids, sugars, or amides as hydrogen bond donors.

4.1 Influence of properties of DES due to composition

Considering that, there are multiples HBA and HBD combination to synthesize a DES and that each one has its own properties, it is difficult to know how it will behave during extraction. Therefore, some authors focused on screening different DES in the extraction of anthocyanins. Alañón et al. [43] evaluate eight different DES based

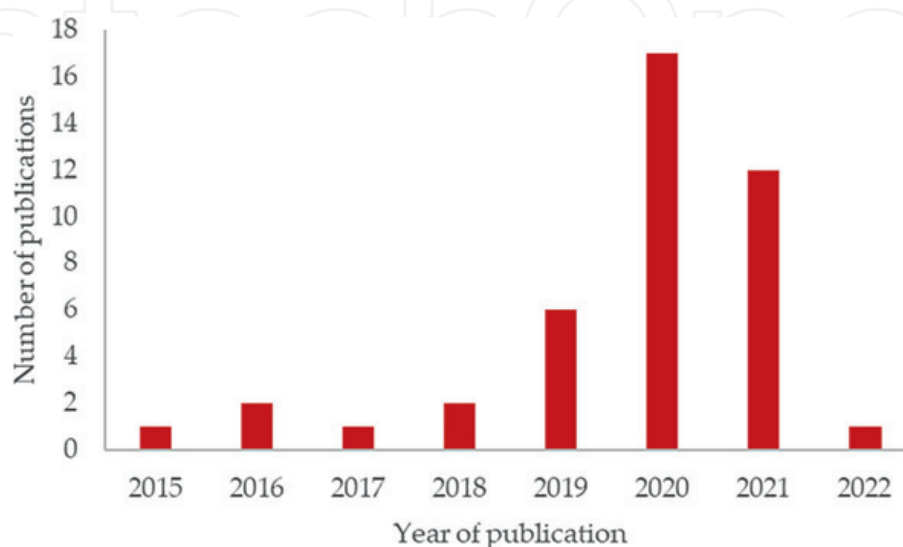


Figure 3.
Number of articles published on DES (NADES) for anthocyanin extraction.

Source	Best-performing solvent (yield)	Assisted technology	Ref.
Flowers			
<i>Hibiscus sabdariffa</i> L. calyces	ChCl-Ox (7.36 mg/g) CA-EtGly (17.87 mg/g) CA-EtGly (9.36 mg/g) SodAce-Foa (7.01 mg/g)	MAE UAE MAE UAE	[43–46]
<i>Myrothamnus flabellifolia</i> Welw	Suc-CA	MA	[47]
<i>Catharanthus roseus</i> var. Pacifica XP Apricot (orange) <i>C. roseus</i> var. Pacifica Orchid Halo (purple)	LA-Glu 1,2-Prop-ChCl Glu-Fru-Suc	UAE	[48]
Saffron (<i>Crocus sativus</i> L.)	LA-Glyc (8.06 mg/L)	MA	[49]
<i>Ixora javanica</i>	ChCl-EtGly (~13 mg/g) ChCl-PrGly (~12 mg/g)	UAE	[50]
Elderberry (<i>Sambucus nigra</i> L.) flowers	ChCl-Gly	MA	[51]
Fruits			
Grape skin (<i>Vitis vinifera</i> L.)	CA-Mal (36.14 mg/g)	UAE	[52]
	ChCl-Gly	MAE	[53]
	ChCl-Ox	MA	[54]
	ChCl-MaA		
	ChCl-Sor		
	ChCl-Pro-MaA ChCl-TA (3.33 mg/L)		
Mulberry puree (<i>Morus alba</i> L.)	ChCl-LA (4.25 mg/g)	UAE	[55]
	ChCl-Ca (~5 mg/g)	HSH-CBE	[56]
Blueberry (<i>Vaccinium corymbosum</i>)	ChCl-Gly-CA (362.3 mg/g)	UAE	[57]
Blueberry (<i>Vaccinium myrtillus</i> L.)	ChCl-LA	MA	[58]
	ChCl-MaA		
	ChCl-Ca		
<i>Ribes nigrum</i> L.	ChCl-LA	MAE	[59]
<i>Lycium ruthenicum</i> Murr.	ChCl-1,2-Pro	UAE	[60]
<i>Nitraria tangutorun</i> Bobr.	ChCl-1,2-Pro (~1.2 mg/g)	UAE	[61]
Black jamun (<i>Syzygium cumini</i>) pulp	ChCl-CA (~9.68 mg/g)	MAE	[62]
Blackberries (<i>Rubus fruticosus</i> L.)	ChCl-Ur-Gly (~325 mg/L)		[63]
Chequén (<i>Luma chequen</i> (Molina) A. Gray)	LA-Glu (3.30 mg/g)	UAE	[64]
Raspberries (<i>Rubus idaeus</i> L.)	ChCl-1,4-But (~1.2 mg/g)	UAE	[65]
Saskatoon berry (<i>Amelanchier alnifolia</i> Nutt.)	ChCl-LA (encapsulation)		[66]
Blue honeysuckle or Haskap (<i>Lonicera caerulea</i> L.)	ChCl-LA (5.82 mg/g)	NPCE	[67]
	CA-Mal (19.04 mg/g)	UAE	[68]
Fruit waste			
Wine waste (<i>V. vinifera</i> L.) cv. Merlot grape	ChCl-MaA (~5.5 mg/g)	UAE	[69]
	ChCl-TA (53.56%)	MA	[70]
Wine waste (<i>V. vinifera</i> L.) cv. Muscat of Hamburg	Gly-SodBen (~1.6 mg/g)	UAE	[71]
Wine waste (<i>V. vinifera</i> L.) cv. Tempranillo	ChCl-Ox (170.04 mg/L)	PHWE	[72]

Source	Best-performing solvent (yield)	Assisted technology	Ref.
Grape pomace (<i>V. vinifera</i> L.) cv. Plavac mali	ChCl-CA (~0.9 mg/g) ChCl-Pro-MA (~0.9 mg/g)	UAE	[41]
Cranberry juice pomaces	Glu-LA (1.54 mg/g)	UAE	[73]
Jaboticaba pomace (<i>Myrciaria cauliflora</i>)	ChCl-PrGly (279.45 mg/100 g)	MA	[74]
Blueberry pomace (<i>Vaccinium</i> spp.)	ChCl-Ox (24.1 mg/g)	MA	[75]
Blueberry wine waste (<i>Vaccinium</i> spp.)	ChCl-1,4-But (~8.5 mg/g)	UAE	[76]
Sour cherry pomace (<i>Prunus cerasus</i>)	ChCl-MaA (~2500 µg/g) (~2000 µg/g) (~2400 µg/g)	MA UAE MAE	[77]
Strawberry extrudate	ChCl-GlyA (1.38 mg/100 g)		[78]
Raspberry extrudate	ChCl-GlyA (0.38 mg/100 g)		[78]
Vegetable and vegetable waste			
Black carrots (<i>Daucus carota</i>)	Biodegradability test only		[79]
Onions solid waste (<i>Allium cepa</i> L.)	β-cyclodextrin methyl β-cyclodextrin 2-hydroxypropyl β-cyclodextrin	MA	[80]

ChCl: Choline chloride; Ox: Oxalic acid; CA: citric acid; EtGly: ethylene glycol; SodAce: Sodium acetate; FoA: Formic acid; LA: Lactic acid; Glu: Glucose; Suc: Sucrose; Fru: Fructose; 1,2-pro: 1,2-Propanediol; Glyc: Glycine; Gly: Glycerol; MaA: Malic acid; Pro: Proline; Sor: Sorbose; TA: Tartaric acid; 1,4-But: 1,4-Butanediol; Mal: Maltose; SodBen: Sodium benzoate; PrGly: Propyleneglycol; GlyA: Glycolic acid; MAE: microwave assisted extraction; UAE: Ultrasound assisted extraction; MA: Maceration; NPCE: Negative pressure cavitation extraction; HSH-CBE: high speed homogenization and cavitation burst extraction; and PHWE: Pressurized hot water extraction.

Table 1.
Source of anthocyanin extracted with DES reported in the literature.

in choline chloride for the extraction of anthocyanins of *Hibiscus sabdariffa* calyces. All DES tested were able to extract delphinidin-3-sambubioside (hibiscin) and cyanidin-3-sambubioside (gossypicyanin). Although DES that contains an organic acid (lactic and oxalic acid) reached the highest yields. Another example was reported by Bi et al. [55]; they evaluated the effect of ten DES based in choline chloride or choline bitartrate or choline acetate and organic acids, polyalcohols, and amide in the anthocyanin extraction of mulberry puree. The DES with urea (amide) were unable to extract anthocyanin, due to the basicity of the solvent. The organic acid-DES combinations were the ones with high extraction values, especially ChCl-lactic acid. In addition, these authors mentioned that there is not a clear relation between viscosity and extractability as other authors previously said. The organic acid relation with extractability was confirmed by Guo et al. [56] in mulberry, Kou et al. [67] in blue honeysuckle fruit (*Lonicera caerulea* L.), and Kou et al. [59] in *Ribes nigrum* L. On the contrary Dai et al. [48] and Alrugaibah et al. [73] mentioned that pH did not impact the extraction efficiency. However, the pH of DES is not the only parameter to consider. Other physicochemical properties such as solubility, viscosity, surface tension, and polarity are also very important [59].

The interaction of anthocyanins, which are polar molecules (depending on the number of hydroxyl groups, degree of methylation, and acylation pattern), with polar DES could increase the extractability. Anthocyanin compounds could be dissolved into a solvent with similar polarity by complying with the principle of similar compatibility [59]. This agrees with results reported by Loarce et al. [72] in grape pomace, Kou et al. [59] in *R. nigrum* L., Cvjetko Bubalo et al. [53] in grape skin, and Xue et al. [76]. In the former research, ten DES of different composition (based in ChCl with polyalcohol and sugars) and molar ratios were screened, finding that DES with similar polarities to that of anthocyanins yielded the higher values. These results differ with Dai et al. 2016 [48] and Sang et al. [60]. In the last-mentioned work, authors evaluate different ChCl and lactic acid, 1,2-Propanediol, 1,4-Butanediol. Organic acids had the lowest extraction yield compared to polyalcohol-based DES which are the less polar DES [81]. Authors also evaluated different ratios of ChCl, suggesting that anthocyanin extraction could be enhanced by making more hydrogen-hydrogen bonds between DES and anthocyanins. However, an excessive hydrogen-bound increase viscosity of the DES affecting the mass transfer parameters.

The high viscosity of the DES could affect the penetration of DES into the pores; therefore, a decrease in viscosity could increase diffusivity, and reduce surface tension [67]. Also, as it was mentioned above high viscosity affects the hydrogen-bonding between DES and anthocyanin compounds by making it more difficult to form [48, 50, 53, 60, 69]. Viscosity could be changed by adding water into the DES mixture. But the increase in water content could affect the polarity properties by weakening the interactions between DES and anthocyanin [56]. Cvjetko Bubalo et al. [53] evaluated the extraction of anthocyanin from grape skin. They observe that the eight anthocyanins founded in grape skin were better extracted when the water content was between 10 to 50% as well as the content/yield of the specific anthocyanins, increasing the polarity effect. For example, high polar anthocyanins (anthocyanin-3-*O*-monoglucosides) were better extracted when the water content was 50%. Meanwhile, less polar anthocyanins (malvidin-3-*O*-acetylmonoglucosides) increase when DES contain 25% of water. Those observations are in agreement with Bosiljkov et al. [69]; they found that certain levels of water into the eutectic mixture improve the extraction due to the polarity (by approaching the polarity of DES close to the water polarity, decreasing it in high polar DES like the organic acid-base or increase in polyalcohol or sugar DES with low polarity), favoring the mass transfer properties due to low viscosity.

4.2 Influence of DES due to process properties

Usually, the most critical factors that affect the extraction are the inherent properties of the solvents. However, extraction variables such as temperature and time also play an important role during the extraction. Furthermore, the combination of the solvent with other technologies (microwave-assisted extraction (MAE) or ultrasound-assisted extraction (USE), among others), and the variables of the process could improve the extractability.

The temperature is related to the change in physical characteristics of the DES, particularly, the reduction of viscosity and surface tension (similar to an increase in water content in DES) when temperature increase. Changes due to temperature could be attributed to the intermolecular forces by decreasing the strength of van der Waals forces and hydrogen bonding interaction, improving the flowability characteristics of DES [20, 34, 35]. Additionally, as previously mentioned, an improvement in mass

transfer parameters facilitates the penetration of DES in cells and the interaction with target compounds. Xue et al. [76] observed this behavior, but an excessive increase in temperature causes degradation of anthocyanins. Similar behavior was observed by Sang et al. [61] and Bozinou et al. [80]. Xue et al. [76] and Bozinou et al. [76, 80] focused their research in optimizing the extraction time [61, 76, 80], showing that shortest times do not allow the complete extraction, on the contrary, at longer time expose to the environmental conditions the quantity decrease [21]. Bozinou et al. [80] observed a fast extraction during the first minutes, reaching the equilibrium condition after 400 min, and a slight reduction after 1400 min. In another study Xue et al. [76] found that the reduction in anthocyanin content was notable after 30 min.

Both parameters, temperature and processing time are still involved when the extraction is assisted by some alternative technology. Microwave-assisted extraction (MAE) and ultrasound-assisted extraction (UAE) are the most used methods. The use of UAE and MAE reduces the temperature and time process by causing a disruption of cell structure (by cavitation or because of a rapid heat inside the cell, respectively) and releasing the internal compounds. The use of MAE and UAE reduce processing temperature and time, and the effect of both parameters is similar to the simple maceration method. It has been reported that increase the temperature and time at a certain point increases extraction yield and an excess of both can decrease the anthocyanin content. This effect was evaluated by Cvjetko Bubalo et al. [53], Fu et al. [75], Kou et al. [59], MacLean et al. [68], Grillo et al. [58], among others. Some of them optimize both parameters, among other parameters such as water content in DES (%), solid-liquid solvent ratio, power (W). However, the optimization parameters will depend on DES type, molar ratio, water content; type of material, state, and metabolite of interest.

4.3 Stability

In addition to the high extraction performance of DES and the physicochemical characteristics of the extraction method, some researchers have focused their studies in evaluating the stability of anthocyanins after being extracted with DES [41, 48, 55, 56, 67]. All observed that the stability of anthocyanins extracted by DES was higher compared to that obtained with a conventional solvent like ethanol. The stability could be attributed i) to the anthocyanin molecule interaction with DES mainly hydrogen bonding with the carboxyl and hydroxyl of cyanidin [48], ii) protection of the chromophore against nucleophilic attack [67], iii) substitution of core flavylum cation of anthocyanins [56], and iv) the reduction of the interaction of solutes with oxidative factors by reducing molecular movement, avoiding oxidative degradation [48, 55, 56].

5. Conclusions

DES can be used as green solvents substituting conventional solvents, which could be toxic. Eutectic mixtures have a great extractability of anthocyanin compounds due to differences in physicochemical characteristics. However, a systematical study is necessary to observe how each property (viscosity, polarity, solubility, pH, among others) or multiple properties of each DES mixture affects the extraction performance (exploring the potential link between properties) and including the external variables such as temperature and time. Besides, is important to know the nature of the target compound to select the ideal DES, which is why one unexplored area could

be the solvent design using computational techniques and further application/evaluation in the laboratory, based on different molecular properties. On the other side, multiple assisted methods could be used to improve extractions like ultrasound or microwave together with DES, the need to explore other technics such as pulsed electric fields, high hydrostatic pressures, among others could be also an interesting field of research. Furthermore, the investigation of conservation by different methods like encapsulation and application in food matrices is necessary in order to increase the interest in anthocyanin extraction with DES.

Acknowledgements

Author Jiménez-González thanks to the Universidad de las Americas Puebla (UDLAP) and the National Council for Science and Technology (CONACyT) of Mexico for the financial support for his PhD studies.

Conflict of interest


No conflict of interest is associated with this work.

Author details

Oscar Jiménez-González, Nelly Ramirez-Corona and José Á. Guerrero-Beltrán*
Department of Chemical, Food and Environmental Engineering, Universidad de las Americas Puebla, Puebla, Mexico

*Address all correspondence to: angel.guerrero@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Alzamora SM, López-Malo A, Tapia MS, Welti-Chanes J. Minimally Processed Foods. Encyclopedia of Food and Health. UK: Elsevier; 2016. pp. 767-771. DOI: 10.1016/B978-0-12-384947-2.00470-0
- [2] Bearth A, Cousin M-E, Siegrist M. The consumer's perception of artificial food additives: Influences on acceptance, risk and benefit perceptions. Food Quality and Preference. 2014;**38**:14-23. DOI: 10.1016/j.foodqual.2014.05.008
- [3] Topolska K, Florkiewicz A, Filipiak-Florkiewicz A. Functional food—Consumer motivations and expectations. IJERPH. 2021;**18**:5327. DOI: 10.3390/ijerph18105327
- [4] Md Salim NS, Singh A, Raghavan V. Potential utilization of fruit and vegetable wastes for food through drying or extraction techniques. NTNF. 2017;**1**:15-27. DOI: 10.31031/NTNF.2017.01.000506
- [5] Sagar NA, Pareek S, Sharma S, Yahia EM, Lobo MG. Fruit and vegetable waste: Bioactive compounds, their extraction, and possible utilization: Fruit and vegetable waste. Comprehensive Reviews in Food Science and Food Safety. 2018;**17**:512-531. DOI: 10.1111/1541-4337.12330
- [6] Wadhwa M, Bakshi MPS. Utilization of fruit and vegetable wastes as livestock feed and as substrates for generation of other value-added products 2013.
- [7] Naderi A, Rezaei S, Moussa A, Levers K, Earnest CP. Fruit for sport. Trends in Food Science & Technology. 2018;**74**:85-98. DOI: 10.1016/j.tifs.2018.02.013
- [8] Schreinemachers P, Simmons EB, Wopereis MCS. Tapping the economic and nutritional power of vegetables. Global Food Security. 2018;**16**:36-45. DOI: 10.1016/j.gfs.2017.09.005
- [9] Marova I, Carnecka M, Halienova A, Certik M, Dvorakova T, Haronikova A. Use of several waste substrates for carotenoid-rich yeast biomass production. Journal of Environmental Management. 2012;**95**:S338-S342. DOI: 10.1016/j.jenvman.2011.06.018
- [10] Neha K, Haider MR, Pathak A, Yar MS. Medicinal prospects of antioxidants: A review. European Journal of Medicinal Chemistry. 2019;**178**:687-704. DOI: 10.1016/j.ejmech.2019.06.010
- [11] Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, et al. Critical review: Vegetables and fruit in the prevention of chronic diseases. European Journal of Nutrition. 2012;**51**:637-663. DOI: 10.1007/s00394-012-0380-y
- [12] Sarı A, Şahin H, Özsoy N, Özbek ÇB. Phenolic compounds and in vitro antioxidant, anti-inflammatory, antimicrobial activities of *Scorzonera hieraciifolia* Hayek roots. South African Journal of Botany. 2019;**125**:116-119. DOI: 10.1016/j.sajb.2019.07.009
- [13] Zhang Q, Gonzalez de Mejia E, Luna-Vital D, Tao T, Chandrasekaran S, Chatham L, et al. Relationship of phenolic composition of selected purple maize (*Zea mays* L.) genotypes with their anti-inflammatory, anti-adipogenic and anti-diabetic potential. Food Chemistry. 2019;**289**:739-750. DOI: 10.1016/j.foodchem.2019.03.116
- [14] Bąkowska A, Kucharska AZ, Oszmiański J. The effects of heating,

UV irradiation, and storage on stability of the anthocyanin–polyphenol copigment complex. *Food Chemistry*. 2003;**81**:349-355. DOI: 10.1016/S0308-8146(02)00429-6

[15] Cortez R, Luna-Vital DA, Margulis D, Gonzalez de Mejia E. Natural pigments: Stabilization methods of anthocyanins for food applications: Stabilization of natural pigments. *Comprehensive Reviews in Food Science and Food Safety*. 2017;**16**:180-198. DOI: 10.1111/1541-4337.12244

[16] Kırca A, Özkan M, Cemeroğlu B. Effects of temperature, solid content and pH on the stability of black carrot anthocyanins. *Food Chemistry*. 2007;**101**:212-218. DOI: 10.1016/j.foodchem.2006.01.019

[17] Jiménez-González O, Guerrero-Beltrán JÁ. Extraction, microencapsulation, color properties, and experimental design of natural pigments obtained by spray drying. *Food Engineering Reviews*. 2021;**13**:769-811. DOI: 10.1007/s12393-021-09288-7

[18] Sridhar A, Ponnuchamy M, Kumar PS, Kapoor A, Vo D-VN, Prabhakar S. Techniques and modeling of polyphenol extraction from food: A review. *Environmental Chemistry Letters*. 2021;**19**:3409-3443. DOI: 10.1007/s10311-021-01217-8

[19] Mohd Fuad F, Mohd Nadzir M, Harun@Kamaruddin A. Hydrophilic natural deep eutectic solvent : A review on physicochemical properties and extractability of bioactive compounds. *Journal of Molecular Liquids*. 2021;**339**:116923. DOI: 10.1016/j.molliq.2021.116923

[20] Ali RA. Review on extraction of phenolic compounds from natural sources using green deep eutectic

solvents. *Journal of Agricultural and Food Chemistry*. 2021;**69**:878-912. DOI: 10.1021/acs.jafc.0c06641

[21] Zuo J, Geng S, Kong Y, Ma P, Fan Z, Zhang Y, et al. Current progress in natural deep eutectic solvents for the extraction of active components from plants. *Critical Reviews in Analytical Chemistry*. 2021;**0**:1-22. DOI: 10.1080/10408347.2021.1946659

[22] Ruesgas-Ramón M, Figueroa-Espinoza MC, Durand E. Application of deep eutectic solvents (DES) for phenolic compounds extraction: Overview, challenges, and opportunities. *Journal of Agricultural and Food Chemistry*. 2017;**65**:3591-3601. DOI: 10.1021/acs.jafc.7b01054

[23] Hansen BB, Spittle S, Chen B, Poe D, Zhang Y, Klein JM, et al. Deep eutectic solvents: A review of fundamentals and applications. *Chemical Reviews*. 2021;**121**:1232-1285. DOI: 10.1021/acs.chemrev.0c00385

[24] Horbowicz M, Kosson R, Grzesiuk A, Dębski H. Anthocyanins of fruits and vegetables—Their occurrence, analysis and role in human nutrition. *Vegetable Crops Research Bulletin*. 2008;**68**:5-22. DOI: 10.2478/v10032-008-0001-8

[25] Pietrini F, Iannelli MA, Massacci A. Anthocyanin accumulation in the illuminated surface of maize leaves enhances protection from photo-inhibitory risks at low temperature, without further limitation to photosynthesis. *Plant, Cell & Environment*. 2002;**25**:1251-1259. DOI: 10.1046/j.1365-3040.2002.00917.x

[26] Wang H, Cao G, Prior RL. Oxygen radical absorbing capacity of anthocyanins. *Journal of Agricultural and Food Chemistry*. 1997;**45**:304-309. DOI: 10.1021/jf960421t

- [27] Giuliani A, Cerretani L, Cichelli A. Colors: Properties and determination of natural pigments. Encyclopedia of Food and Health. UK: Elsevier; 2016. pp. 273-283. DOI: 10.1016/B978-0-12-384947-2.00189-6
- [28] Ruta F. Anthocyanins and anthocyanin-derived products in yeast-fermented beverages. Antioxidants. 2019;**8**:182. DOI: 10.3390/antiox8060182
- [29] Harborne JB. Anthocyanins as food colours. Phytochemistry. 1983;**22**:1067-1068. DOI: 10.1016/0031-9422(83)85072-9
- [30] Bueno JM, Sáez-Plaza P, Ramos-Escudero F, Jiménez AM, Fett R, Asuero AG. Analysis and antioxidant capacity of anthocyanin pigments. Part II: Chemical structure, color, and intake of anthocyanins. Critical Reviews in Analytical Chemistry. 2012;**42**:126-151. DOI: 10.1080/10408347.2011.632314
- [31] Arruda HS, Silva EK, Peixoto Araujo NM, Pereira GA, Pastore GM, Marostica Junior MR. Anthocyanins recovered from agri-food by-products using innovative processes: Trends, challenges, and perspectives for their application in food systems. Molecules. 2021;**26**:2632. DOI: 10.3390/molecules26092632
- [32] Li M, Zhao X, Sun Y, Yang Z, Han G, Yang X. Evaluation of anthocyanin profile and color in sweet cherry wine: Effect of sinapic acid and grape tannins during aging. Molecules. 2021;**26**:2923. DOI: 10.3390/molecules26102923
- [33] Abbott AP, Capper G, Davies DL, Rasheed RK, Tambyrajah V. Novel solvent properties of choline chloride/urea mixtures Electronic supplementary information (ESI) available: Spectroscopic data. Chemical Communication. 2003;**1**:70-71. DOI: 10.1039/b210714g
- [34] Huang J, Guo X, Xu T, Fan L, Zhou X, Wu S. Ionic deep eutectic solvents for the extraction and separation of natural products. Journal of Chromatography A. 2019;**1598**:1-19. DOI: 10.1016/j.chroma.2019.03.046
- [35] Alam MA, Muhammad G, Khan MN, Mofijur M, Lv Y, Xiong W, et al. Choline chloride-based deep eutectic solvents as green extractants for the isolation of phenolic compounds from biomass. Journal of Cleaner Production. 2021;**309**:127445. DOI: 10.1016/j.jclepro.2021.127445
- [36] Jeong KM, Lee MS, Nam MW, Zhao J, Jin Y, Lee D-K, et al. Tailoring and recycling of deep eutectic solvents as sustainable and efficient extraction media. Journal of Chromatography A. 2015;**1424**:10-17. DOI: 10.1016/j.chroma.2015.10.083
- [37] Vanda H, Verpoorte R, Klinkhamer PGL, Choi YH. Natural deep eutectic solvents: From their discovery to their applications. In: Ramón DJ, Guillena G, editors. Deep Eutectic Solvents. 1st ed. Wiley; 2019. pp. 61-81. DOI: 10.1002/9783527818488.ch4
- [38] Socas-Rodríguez B, Torres-Cornejo MV, Álvarez-Rivera G, Mendiola JA. Deep eutectic solvents for the extraction of bioactive compounds from natural sources and agricultural by-products. Applied Sciences. 2021;**11**:4897. DOI: 10.3390/app11114897
- [39] Marcus Y. Applications of Deep Eutectic Solvents. Deep Eutectic Solvents. Cham: Springer International Publishing; 2019. pp. 111-151. DOI: 10.1007/978-3-030-00608-2_4
- [40] Souza Mesquita LM, Martins M, Pisani LP, Ventura SPM, Rosso VV. Insights on the use of alternative solvents and technologies to recover bio-based

food pigments. Comprehensive Reviews in Food Science and Food Safety. 2021;**20**:787-818. DOI: 10.1111/1541-4337.12685

[41] Panić M, Gunjević V, Cravotto G, Radojčić RI. Enabling technologies for the extraction of grape-pomace anthocyanins using natural deep eutectic solvents in up-to-half-litre batches extraction of grape-pomace anthocyanins using NADES. Food Chemistry. 2019;**300**:125185. DOI: 10.1016/j.foodchem.2019.125185

[42] Balaraman H, Selvasembian R, Rangarajan V, Rathnasamy S. Sustainable and green engineering insights on deep eutectic solvents toward the extraction of nutraceuticals. ACS Sustainable Chemistry & Engineering. 2021;**9**:11290-11313. DOI: 10.1021/acssuschemeng.1c03034

[43] Alañón ME, Ivanović M, Pimentel-Mora S, Borrás-Linares I, Arráez-Román D, Segura-Carretero A. A novel sustainable approach for the extraction of value-added compounds from *Hibiscus sabdariffa* L. calyces by natural deep eutectic solvents. Food Research International. 2020;**137**:109646. DOI: 10.1016/j.foodres.2020.109646

[44] Şahin S, Pekel AG, Toprakçı İ. Sonication-assisted extraction of *Hibiscus sabdariffa* for the polyphenols recovery: Application of a specially designed deep eutectic solvent. Biomass Conversion Bioref. 2020. DOI: 10.1007/s13399-020-00837-4

[45] Kurtulbaş E, Pekel AG, Bilgin M, Makris DP, Şahin S. Citric acid-based deep eutectic solvent for the anthocyanin recovery from *Hibiscus sabdariffa* through microwave-assisted extraction. Biomass Conversion Bioref. 2022;**12**:351-360. DOI: 10.1007/s13399-020-00606-3

[46] Zannou O, Koca I, Aldawoud TMS, Galanakis CM. Recovery and stabilization of anthocyanins and phenolic antioxidants of Roselle (*Hibiscus sabdariffa* L.) with hydrophilic deep eutectic solvents. Molecules. 2020;**25**:3715. DOI: 10.3390/molecules25163715

[47] Bentley J, Olsen EK, Moore JP, Farrant JM. The phenolic profile extracted from the desiccation-tolerant medicinal shrub *Myrothamnus flabellifolia* using Natural Deep Eutectic Solvents varies according to the solvation conditions. Phytochemistry. 2020;**173**:112323. DOI: 10.1016/j.phytochem.2020.112323

[48] Dai Y, Rozema E, Verpoorte R, Choi YH. Application of natural deep eutectic solvents to the extraction of anthocyanins from *Catharanthus roseus* with high extractability and stability replacing conventional organic solvents. Journal of Chromatography A. 2016;**1434**:50-56. DOI: 10.1016/j.chroma.2016.01.037

[49] Lakka A, Grigorakis S, Karageorgou I, Batra G, Kaltsa O, Bozinou E, et al. Saffron processing wastes as a bioresource of high-value added compounds: Development of a green extraction process for polyphenol recovery using a natural deep eutectic solvent. Antioxidants. 2019;**8**:586. DOI: 10.3390/antiox8120586

[50] Oktaviyanti ND, Kartini, Mun'im A. Application and optimization of ultrasound-assisted deep eutectic solvent for the extraction of new skin-lightening cosmetic materials from *Ixora javanica* flower. Heliyon. 2019;**5**:e02950. DOI: 10.1016/j.heliyon.2019.e02950

[51] Vladimir-Knežević S, Perković M, Zagajski Kučan K, Mervić M, Rogošić M. Green extraction of flavonoids and phenolic acids from elderberry

(*Sambucus nigra* L.) and rosemary (*Rosmarinus officinalis* L.) using deep eutectic solvents. Chemical Papers. 2022;**76**:341-349. DOI: 10.1007/s11696-021-01862-x

[52] Jeong KM, Zhao J, Jin Y, Heo SR, Han SY, Yoo DE, et al. Highly efficient extraction of anthocyanins from grape skin using deep eutectic solvents as green and tunable media. Archives of Pharmacal Research. 2015;**38**:2143-2152. DOI: 10.1007/s12272-015-0678-4

[53] Cvjetko Bubalo M, Ćurko N, Tomašević M, Kovačević Ganić K, Radojčić RI. Green extraction of grape skin phenolics by using deep eutectic solvents. Food Chemistry. 2016;**200**:159-166. DOI: 10.1016/j.foodchem.2016.01.040

[54] Sapone V, Cicci A, Franceschi D, Vincenzi S, Bravi. Antioxidant extraction and bioactivity preservation from winery by-products by natural deep eutectic solvents (nades). Chemical Engineering Transactions. 2020;**79**:157-162. DOI: 10.3303/CET2079027

[55] Bi Y, Chi X, Zhang R, Lu Y, Wang Z, Dong Q, et al. Highly efficient extraction of mulberry anthocyanins in deep eutectic solvents: Insights of degradation kinetics and stability evaluation. Innovative Food Science & Emerging Technologies. 2020;**66**:102512. DOI: 10.1016/j.ifset.2020.102512

[56] Guo N, Ping-Kou JY-W, Wang L-T, Niu L-J, Liu Z-M, et al. Natural deep eutectic solvents couple with integrative extraction technique as an effective approach for mulberry anthocyanin extraction. Food Chemistry. 2019;**296**:78-85. DOI: 10.1016/j.foodchem.2019.05.196

[57] da Silva DT, Pauletto R, da Cavalheiro S, Bochi VC, Rodrigues E, Weber J, et al. Natural deep eutectic

solvents as a biocompatible tool for the extraction of blueberry anthocyanins. Journal of Food Composition and Analysis. 2020;**89**:103470. DOI: 10.1016/j.jfca.2020.103470

[58] Grillo G, Gunjević V, Radošević K, Redovniković IR, Cravotto G. Deep eutectic solvents and nonconventional technologies for blueberry-peel extraction: Kinetics, anthocyanin stability, and antiproliferative activity. Antioxidants. 2020;**9**:1069. DOI: 10.3390/antiox9111069

[59] Kou P, Kang Y-F, Wang L-T, Niu L-J, Xiao Y, Guo N, et al. An integrated strategy for production of four anthocyanin compounds from *Ribes nigrum* L. by deep eutectic solvents and flash chromatography. Journal of Industrial and Engineering Chemistry. 2019;**80**:614-625. DOI: 10.1016/j.jiec.2019.08.053

[60] Sang J, Li B, Huang Y, Ma Q, Liu K, Li C. Deep eutectic solvent-based extraction coupled with green two-dimensional HPLC-DAD-ESI-MS/MS for the determination of anthocyanins from *Lycium ruthenicum* Murr. fruit. Analytical Methods. 2018;**10**:1247-1257. DOI: 10.1039/C8AY00101D

[61] Sang J, Liu K, Ma Q, Li B, Li C. Combination of a deep eutectic solvent and macroporous resin for green recovery of anthocyanins from *Nitraria tangutorum* Bobr. fruit. Separation Science and Technology. 2019;**54**:3082-3090. DOI: 10.1080/01496395.2018.1559190

[62] Sharma M, Dash KK. Deep eutectic solvent-based microwave-assisted extraction of phytochemical compounds from black jamun pulp. Journal of Food Process Engineering. 2021;**44**. DOI: 10.1111/jfpe.13750

[63] Souza HKS, Mateus N, de Freitas V, Gonçalves MP, Cruz L. Chemical/color stability and rheological properties of

cyanidin-3-glucoside in deep eutectic solvents as a gateway to design task-specific bioactive compounds. *ACS Sustainable Chemistry & Engineering*. 2020;**8**:16184-16196. DOI: 10.1021/acssuschemeng.0c04839

[64] Velásquez P, Bustos D, Montenegro G, Giordano A. Ultrasound-assisted extraction of anthocyanins using natural deep eutectic solvents and their incorporation in edible films. *Molecules*. 2021;**26**:984. DOI: 10.3390/molecules26040984

[65] Xue H, Tan J, Li Q, Tang J, Cai X. Optimization ultrasound-assisted deep eutectic solvent extraction of anthocyanins from raspberry using response surface methodology coupled with genetic algorithm. *Food*. 2020;**9**:1409. DOI: 10.3390/foods9101409

[66] Zeng Y-J, Xu P, Yang H-R, Zong M-H, Lou W-Y. Purification of anthocyanins from saskatoon berries and their microencapsulation in deep eutectic solvents. *LWT*. 2018;**95**:316-325. DOI: 10.1016/j.lwt.2018.04.087

[67] Kou P, Wan N, Wang L-T, Pan H-Y, Jiao J, Zhao C-J, et al. A sustainable and efficient preparation process of anthocyanins from blue honeysuckle fruit and comprehensive bioactivity assessment. *Journal of the Taiwan Institute of Chemical Engineers*. 2020;**116**:3-10. DOI: 10.1016/j.jtice.2020.10.029

[68] MacLean AMG, Silva YPA, Jiao G, Brooks MS. Ultrasound-assisted extraction of anthocyanins from Haskap (*Lonicera caerulea* L.) berries using a deep eutectic solvent (DES). *Food Technology and Biotechnology*. 2021;**59**:56-62. DOI: 10.17113/ftb.59.01.21.6869

[69] Bosiljkov T, Dujmić F, Cvjetko Bubalo M, Hribar J, Vidrih R, Brnčić M,

et al. Natural deep eutectic solvents and ultrasound-assisted extraction: Green approaches for extraction of wine lees anthocyanins. *Food and Bioprocess Technology*. 2017;**102**:195-203. DOI: 10.1016/j.fbp.2016.12.005

[70] Iannone A, Sapone V, Di Paola L, Cicci A, Bravi M. Extraction of anthocyanins from grape (*Vitis Vinifera*) skins employing natural deep eutectic solvents (nades). *Chemical Engineering Transactions*. 2021;**87**:469-474. DOI: 10.3303/CET2187079

[71] Alibade A, Lakka A, Bozinou E, Lalas SI, Chatzilazarou A, Makris DP. Development of a green methodology for simultaneous extraction of polyphenols and pigments from red winemaking solid wastes (Pomace) using a novel glycerol-sodium benzoate deep eutectic solvent and ultrasonication pretreatment. *Environments*. 2021;**8**:90. DOI: 10.3390/environments8090090

[72] Loarce L, Oliver-Simancas R, Marchante L, Díaz-Maroto MC, Alañón ME. Modifiers based on natural deep eutectic mixtures to enhance anthocyanins isolation from grape pomace by pressurized hot water extraction. *LWT*. 2021;**149**:111889. DOI: 10.1016/j.lwt.2021.111889

[73] Alrugaibah M, Yagiz Y, Gu L. Use natural deep eutectic solvents as efficient green reagents to extract procyanidins and anthocyanins from cranberry pomace and predictive modeling by RSM and artificial neural networking. *Separation and Purification Technology*. 2021;**255**:117720. DOI: 10.1016/j.seppur.2020.117720

[74] Benvenutti L, del Sanchez-Camargo Z, Ferreira SRS. NADES as potential solvents for anthocyanin and pectin extraction from *Myrciaria cauliflora* fruit by-product: In silico and experimental

approaches for solvent selection. *Journal of Molecular Liquids*. 2020;**315**:113761. DOI: 10.1016/j.molliq.2020.113761

[75] Fu X, Wang D, Belwal T, Xie J, Xu Y, Li L, et al. Natural deep eutectic solvent enhanced pulse-ultrasonication assisted extraction as a multi-stability protective and efficient green strategy to extract anthocyanin from blueberry pomace. *LWT*. 2021;**144**:111220. DOI: 10.1016/j.lwt.2021.111220

[76] Xue H, Tan J, Li Q, Tang J, Cai X. Ultrasound-assisted deep eutectic solvent extraction of anthocyanins from blueberry wine residues: Optimization, identification, and HepG2 antitumor activity. *Molecules*. 2020;**25**:5456. DOI: 10.3390/molecules25225456

[77] Popovic BM, Micic N, Potkonjak A, Blagojevic B, Pavlovic K, Milanov D, et al. Novel extraction of polyphenols from sour cherry pomace using natural deep eutectic solvents—Ultrafast microwave-assisted NADES preparation and extraction. *Food Chemistry*. 2022;**366**:130562. DOI: 10.1016/j.foodchem.2021.130562

[78] Vázquez-González M, Fernández-Prior Á, Bermúdez Oria A, Rodríguez-Juan EM, Pérez-Rubio AG, Fernández-Bolaños J, et al. Utilization of strawberry and raspberry waste for the extraction of bioactive compounds by deep eutectic solvents. *LWT*. 2020;**130**:109645. DOI: 10.1016/j.lwt.2020.109645

[79] Aslan Türker D, Doğan M. Application of deep eutectic solvents as a green and biodegradable media for extraction of anthocyanin from black carrots. *LWT*. 2021;**138**:110775. DOI: 10.1016/j.lwt.2020.110775

[80] Bozinou E, Lakka A, Poulianiti K, Lalas S, Makris DP. Cyclodextrins as

high-performance green co-solvents in the aqueous extraction of polyphenols and anthocyanin pigments from solid onion waste. *European Food Research and Technology*. 2021;**247**:2831-2845. DOI: 10.1007/s00217-021-03839-2

[81] Dai Y, van Spronsen J, Witkamp G-J, Verpoorte R, Choi YH. Natural deep eutectic solvents as new potential media for green technology. *Analytica Chimica Acta*. 2013;**766**:61-68. DOI: 10.1016/j.aca.2012.12.019

Methods of Application of Essential Oils in Foods and Their Effects on Sensory Attributes

*Scarlette L. Recio-Cázares, Aurelio López-Malo
and Enrique Palou*

Abstract

In recent years, demand for more natural products by consumers has stimulated the use of essential oils (EOs) by the food industry as additives, food supplements, insecticides, and sanitizers. Different methods of applying EOs in foods have been studied; however, an important factor for their industrial use is the effect of EOs on sensory attributes. So currently, research is mainly focused on methods that allow the incorporation of EOs in foods while ensuring consumer acceptance. Therefore, the aim of this work is to compare current literature reporting several methods of application of EOs and especially their effect on sensory attributes in different food systems.

Keywords: essential oils, application methods, natural preservatives, sensory attributes, food applications

1. Introduction

The shelf life, nutritional properties, sensory attributes and microbial quality of food products are important aspects for consumers; therefore, food industries seek to make use of substances or compounds that ensure the safety of their products without compromising their quality [1].

Different preservatives are utilized during food processing to ensure food safety and quality; however, most of them are synthetic, and some have been associated with long-term health problems, such as kidney failure, hyperthyroidism, liver cirrhosis, and cancer, among others [2, 3]. As a result, there is a growing demand for natural products that can serve as preservatives [1].

Essential oils (EOs) are secondary metabolites obtained from different plants, which have been widely utilized as natural preservatives because they are, in many cases, safer than synthetic ones for both humans and the environment [4]. They also have several biological properties including antibacterial, antiparasitic, antifungal, antioxidant, disinfectant, and insecticide ones [5].

Due to their properties, EOs have been utilized in the agri-food industry. They have been applied directly in food systems, either in powder or in liquid form, by means of

edible films or coatings, in vapor phase, as well as in nanoemulsions and liposomes [6]. However, their application in the industry is still limited due to their impact on taste and aroma, especially at the concentrations needed for an effective use as antimicrobial. Therefore, the most acceptable concentration in terms of both, safety and food quality will depend on the type of EO, as well as its active compounds, the type of food in which it is incorporated, the method of application, and the food processing technique [7].

Therefore, the aim of this work is to compare current literature reporting several methods of application of EOs, describing their fundamentals, advantages and disadvantages, and especially their effect on sensory attributes in different food systems.

2. Methods of application of essential oils

2.1 Direct contact

This method of application consists of adding the EO either in powder or liquid form, directly into food systems. However, their application is limited, due to the strong aroma and taste that EOs impart to food products, especially at concentrations at which they can be considered antimicrobial agents. Therefore, very few EOs have the chemical (qualitative and quantitative composition, required concentration), physical (solubility, volatility), and sensory (strong flavors or aromas) characteristics required to be applied by direct contact [8].

In-vitro and *in-vivo* studies have reported that the antimicrobial effect of EOs is due to the presence of different active compounds that act synergistically [9, 10]. However, concentrations obtained under *in-vitro* conditions are not always effective when applied to foods [9], so the concentration of EO needed to observe the same effect could increase and thus affect acceptable sensory levels [8].

2.1.1 Application of essential oils in direct contact and their effect on sensory attributes

The application of EOs by direct contact has been tested on different foods such as meat products (meat burgers, ground meat, sausages, and fish, among others) and fruits (pineapple, mango, guava, and apple). Cinnamon (*Cinnamomum verum*), mint (*Mentha spicata*), lemon balm (*Melissa officinalis*), savory (*Satureja montana*), and juniper (*Juniperus communis*) EOs have exhibited good results against pathogenic bacteria such as *Pseudomonas aeruginosa*, *Salmonella typhimurium*, and *Escherichia coli* [10–14]. Rosemary (*Salvia rosmarinus*) and cinnamon EOs reduced and prevented the growth of Gram-positive bacteria such as *Staphylococcus aureus* and *Listeria monocytogenes* [11, 13] that are generally more sensitive to the presence of EOs [15].

The effect on the sensory attributes of EOs applied by direct contact in food systems is displayed in **Table 1**; some authors report their application for fish preservation [13]. [13] added 4 types of EOs: fennel (*Foeniculum vulgare*), cardamom (*Elettaria cardamomum*), cinnamon, or rosemary to fish fingers and ten untrained judges evaluated their acceptability (appearance, aroma, and texture) during storage at 4°C; the authors report that in addition to extending the shelf life (to 12 d while in the control without EOs it was 4 d) of the product during storage, they also improved their sensory characteristics. Rosemary EO had the best results since it was greatly accepted in the three studied attributes, and it can be used as antimicrobial.

EO	Concentration	Food	Sensory attribute	Results	Reference
Sage (<i>Sage officinalis</i>), rosemary (<i>Salvia rosmarinus</i>), thyme (<i>Thymus vulgaris</i>), and clove (<i>Syzygium aromaticum</i>)	600 ppm	Smoked rainbow trout	Appearance, taste, and aroma	Clove EO best accepted in all attributes	[16]
Hisopo (<i>Hyssopus officinalis</i>) and coriander (<i>Coriandrum sativum</i>)	0.02%	Ground beef	Aroma and taste	Coriander EO best accepted in both attributes evaluated	[17]
Mint (<i>Mentha spicata</i>) and toronjil (<i>Melissa officinalis</i>)	0.625 and 1.25 μ L/mL, respectively	Pineapple, mango, guava and apple cashew juices	Appearance, aroma, viscosity, taste, aftertaste, and general acceptability	Both EOs had negative effects on taste, aftertaste, and overall acceptability; however, mint EO was more accepted in other studied attributes	[10]
Combination of Chinese cinnamon (<i>Cinnamomun cassia</i>) and cinnamon bark (<i>Cinnamomum verum</i>)	0.05%	Ready-to-eat meat and lean ground meat	Taste and aroma in both types of meat	The EO showed better acceptability in both meats in the two evaluated attributes	[11]
Fennel (<i>Foeniculum vulgare</i>), cardamom (<i>Elettaria cardamomum</i>), cinnamon, and rosemary	10 mL/kg	Fish fingers	Appearance, aroma, and texture	Rosemary EO best accepted in all attributes until the 12th day of storage	[13]
Clove, basil (<i>Ocimum basilicum</i>), thyme, and Chinese cinnamon	0.25, 0.125, 0.25, and 0.125%	Chicken sausages	Appearance, taste, juiciness, texture, and overall acceptability	0.25% Chinese cinnamon EO was best accepted in all attributes	[12]
Clove and Curry (<i>Murraya koenigii</i>)	0.05, 0.15, and 0.25%	Burfi	Taste, texture, color, appearance and overall acceptability	Overall taste and acceptability were affected as the concentration of EO increased. Being clove EO at 0.15% the best accepted.	[18]

EO	Concentration	Food	Sensory attribute	Results	Reference
Savory (<i>Satureja montana</i>)	0.075 and 0.150 µL/g	Pork sausages	Taste, aroma and general acceptability	The EO obtained by extraction with supercritical fluid at a concentration of 0.075 µL/g was best accepted in all attributes	[19]
Juniper (<i>Juniperus communis</i>) and savory	0.0625, 0.125, 0.25, and 0.5%	Beef marinated with wine	Marinated taste and odor defects	0.25% best accepted juniper EO in both attributes. Taste and aroma defects show a spicy, sharp, and artificial taste with savory EO	[20]
Chinese pepper (<i>Zanthoxylum piperitum</i>)	0.3%	Sashimi	Color, aroma, texture and overall acceptability	Good acceptability of all attributes until the 15th day of storage	[21]
Combination of thyme, rosemary, and cinnamon	0.078, 0.625, and 0.156 g/L	Jujube	Texture, taste, aroma, and overall acceptability	The combination of EOs did not generate negative effects on studied attributes	[14]

Table 1.
Effect of selected essential oils applied by direct contact on the sensory attributes of various foods.

One of the most common uses of EOs by direct contact is in meat products as antimicrobials. It has been reported that the combination of Chinese cinnamon (*Cinnamomum cassia*) at a concentration of 0.25 and 0.05% cinnamon bark generate good sensory acceptability. Generate good sensory acceptability in ready-to-eat meat and chicken sausages, respectively [11, 12]. Also, Chinese pepper EO in sashimi, where the studied sensory attributes were acceptable [21].

In addition, EOs of rosemary, thyme (*Thymus vulgaris*), sage (*Salvia officinalis*), and clove (*Syzygium aromaticum*) were evaluated in smoked trout fillets stored at 4°C [16]. Fish fillets with rosemary EO had lower sensory acceptability compared to other studied EOs (thyme, sage, and clove), which could be because these trout fillets were evaluated smoked and not raw like fish fingers [13], affecting the EOs' sensory affinity with the food product.

It has been reported [4] that the negative sensory effect can be minimized by properly selecting the EO according to the type of food; for example, cinnamon has better affinity with various fruits and their products [22, 23]. However, it was observed in burfi (an Indian dessert) added with clove and curry (*Murraya koenigii*) EOs [18] as well as in guava, pineapple, mango, and cashew apple juices with mint and lemon balm EOs [10]. In both studies, inhibition of microorganisms was achieved; but sensory properties were negatively affected, as the taste of the EO decreased sensory acceptability (taste, aroma, residue, and overall acceptability).

When savory EO was applied to pork sausage [19] and beef marinated with wine [20]. In both cases, the taste of the EO negatively affected sensory attributes (aroma and taste). the sensory attributes aroma and taste were negatively affected by the presence of the EO. It is known that the composition of the EO can be different depending on several factors, including the extraction method. [19] were able to improve the sensory acceptability of savory EO by changing the extraction method from hydro-distillation to supercritical fluid extraction. This technology decreased the number of volatile compounds (terpenes) and improved sensory acceptability.

2.2 Vapor phase

The vapor phase consists of evaluating the vapor generated by the EOs on different food systems. This method has been reported to have a greater antimicrobial effect compared to EOs in liquid form applied by direct contact, because more volume of liquid EOs is needed to achieve the same biological effect as in gaseous form [24].

The volatility of EOs depends on factors such as the molecular weight of their components; since the smaller the molecule, the diffusivity is higher [25]. Similarly, increasing the temperature of the environment encourages migration of volatile compounds [8]. It is difficult to compare vapor phase antimicrobial activity results because there is no standard method and there is variation in the composition of EOs. However, one of the most promising methods for this purpose seems to be the disk volatilization test [26], which can be modified for food preservation by generating “active packages”, which allow volatilization and migration of the active components from the “disc in the package” into the food [8].

2.2.1 Application of essential oils in vapor phase and their effect on sensory attributes

The effect on different sensory attributes of a selection of EOs applied in vapor phase to different food systems is shown in **Table 2**. EOs have been used in fruits such as apples and grapes to prevent/reduce the growth of pathogenic molds such as *Aspergillus*, *Alternaria*, *Penicillium*, and *Botrytis*. In the case of grapes, a sensory analysis was performed through a triangular test to evaluate if the exposure time of the EOs affected sensory attributes; results indicated that lavender (*Lavandula*), rosemary (*Salvia rosmarinus*), and mint (*Mentha piperita*) EOs could be perceived even 24 h after treatment, due to the large number of aromatic compounds that they contain [29].

Apples treated with oregano (*Origanum vulgare*), cinnamon (*C. verum*), and clove (*S. aromaticum*) EOs [28] in vapor phase were evaluated for sensory acceptability using an unstructured scale. When the concentration of oregano and cinnamon EOs increased, they passed on different aromas to the apples. Although oregano EO showed better antifungal activity (MIC: 2 μ L of EO/L_{air}), it negatively affected sensory attributes when utilized at this concentration. This may be because oregano EO has been reported to provide an herbaceous flavor unlike to the sweet flavor imparted by cinnamon EO [22, 23]. This effect has also been reported in vegetables such as mushrooms [27]. [29] have proposed that reducing the time of exposure to EO vapors or extending the time between treatment and consumption could be strategies to minimize the sensory changes in the product.

Similarly, other authors [31, 32] evaluated the exposure of different EOs to bread for the inhibition of *Penicillium* (lemongrass -*Cymbopogon citratus*- EO) and *Aspergillus* [Thyme -*T. vulgaris*-, clove, cumin -*Cuminum cyminum*- and oregano EOs],

EO	Concentration	Food	Sensory attribute	Results	References
Cinnamon (<i>Cinnamomum verum</i>), thyme (<i>Thymus vulgaris</i>), and clove (<i>Syzygium aromaticum</i>)	5 µL/L _{air}	Mushrooms	Aroma and appearance	Cinnamon EO had better acceptance in both attributes for 20 d of storage	[27]
Oregano (<i>Origanum vulgare</i>), cinnamon, and clove	4 and 16 mL/L _{air}	Apples	Taste and aroma	Cinnamon EO was best accepted in both attributes until the 14th of storage	[28]
Lavender (<i>Lavandula</i>), rosemary (<i>Salvia rosmarinus</i>), and mint (<i>Mentha piperita</i>)	5 mL	Grapes	Flavor	Judges sensed the difference between control without and with EOs until 24 h. The most perceived EO was lavender	[29]
Allspice (<i>Pimenta dioica</i>), rosemary, and thyme	0.8–11.7 mL/L _{air}	Germinated alfalfa seeds	Degree of difference between seeds germinated with and without EOs	Moderate difference in seeds with allspice and thyme EO, and light difference with thyme EO. In addition, the germinates of seeds treated with allspice EO had better sensory acceptance	[30]
Thyme, cumin, clove, and oregano	125–500 µL	Slices of bread	General acceptability of aroma and taste	Good acceptability with thyme and cumin EOs for 4 d of storage.	[31]

Table 2.
Effect of selected essential oils applied by vapor phase on the sensory attributes of various foods.

respectively. In both papers they reported that sensory attributes of aroma, taste, and general acceptability were acceptable even after 21 d of storage with lemongrass EO and 14 d with thyme and cumin EOs.

The antimicrobial activity in vapor phase of allspice (*Pimenta dioica*), thyme and rosemary EOs against *L. monocytogenes* and *S. typhimurium* in alfalfa seeds was evaluated [30]. In addition, a sensory analysis was made using a triangular test. The

results showed that the difference among samples was moderate between allspice and thyme EOs, and slightly for germinated seeds treated with rosemary EO. The authors emphasize that most panelists preferred the sprouts from seeds treated with allspice and rosemary EOs to the control samples that did not contain EOs. This proves that some EOs in vapor phase at an adequate concentration can have a positive impact on the sensory attributes of different food products.

2.3 Nanoemulsions

Another alternative for EO application is encapsulation; it is defined as a technology that allows trapping sensitive components in a homogeneous or heterogeneous matrix, which provides protection, and increases its stability to adverse conditions; in addition, it helps to control its release [33]. There are different encapsulation methods, including simple or complex coacervation, spray drying, fluidized bed coating, or emulsions/nanoemulsions [34].

Emulsions have proven to be good carriers for the delivery of lipid substances such as EOs. Emulsions are composed of two phases, the oily (O) and the aqueous (W). The oil-in-water (O/W) emulsion consists of small drops of oil as a dispersed phase contained in water as a continuous phase while in the water-in-oil (W/O) emulsion the drops of the dispersed phase are water, and the continuous phase is oil. In the food industry, the most common emulsions are O/W.

Emulsions can also be classified into micro- or nano-emulsions depending on the size of the droplets. The droplet size of nanoemulsion ranges between 20 and 1000 nm; it can be prepared using different methods, such as high-pressure homogenization, microfluidization, high-speed mechanical homogenization, ultrasound, spontaneous emulsion, and phase inversion, among others [35, 36]. Its importance is that the nanometric particle size allows increasing bioavailability, bio-efficiency and stability [37].

2.3.1 Application of essential oils by nanoemulsions in foods and their effect on sensory attributes

The use and incorporation of EOs' nanoemulsions in foods is a technique that has become relevant, because increase of the contact surface improves antimicrobial activity since active components are more available [38]. Although nanoemulsions have had more application in foods than other methods, sensory studies are still scarce, only a few reports have studied the effect of nanoemulsions on the sensory properties of foods.

The application of nanoemulsions in foods has been tested mainly for fish conservation (**Table 3**). The antimicrobial effect of nanoemulsions of rosemary (*Salvia rosmarinus*), laurel (*Laurus nobilis*), thyme (*T. vulgaris*), and *aloe vera* EOs at 4% on the quality of rainbow trout were evaluated [40]. The results of the microbiological analysis indicated that EOs reduced the count of mesophilic anaerobic bacteria, psychotrophic bacteria (main microorganisms responsible for fresh fish deterioration) and enterobacteria by 3 Log CFU/g. The EOs also increased the shelf life by 3 d more (17 d in total while the control without EOs was 14 d) of the fish treated with the EOs. Similarly, [39] evaluated the antimicrobial effect of lemon EO (*Citrus × lemon*) at 0.3% against the same bacteria, resulting in the reduction of only 2 Log CFU/g. For sensory evaluation, both studies used a hedonic scale of 9 and 10 points respectively, for the evaluation of taste, aroma, and texture in fish. The EOs tested on both of the

EO	Concentration	Food	Sensory attribute	Results	Reference
Lemon (<i>Citrus × lemon</i>)	0.3% and 1%	Salted sardines	Aroma, texture, and taste	Good acceptance in all attributes with both concentrations of EO	[39]
Rosemary (<i>Salvia rosmarinus</i>), laurel, thyme (<i>Thymus vulgaris</i>), and aloe vera (<i>Aloe vera</i>)	4%	Rainbow trout	Taste, aroma, color, and overall acceptability	Thyme and laurel EO were the best accepted in all attributes	[40]
Oregano (<i>Origanum vulgare</i>), thyme, and star anise	0.1%	Chinese carp	Color, aroma, texture, and overall acceptability	Best-accepted oregano EO on all attributes for 10 d of storage	[41]
Cinnamon (<i>Cinnamomum verum</i>), rosemary, and oregano	10%	Celery	Aroma, color, texture, appearance, taste, and overall acceptability	Cinnamon EO showed better acceptance in all attributes	[42]
Oregano and rosemary	40 and 80 µL/mL, respectively	Fresh cheese	Aroma, appearance, taste, and texture	Both EOs had a positive effect on aroma and flavor until the 21st day of storage	[43]
Lemongrass (<i>Cymbopogon citratus</i>), garlic (<i>Allium sativum</i>), oregano, and thyme	0.2, 0.4, and 0.8%	Salmon and cod	Selected sensory attributes for each food.	0.2% best accepted oregano EO for salmon; EOs masked cod's flavor	[44]

Table 3.
Effect of selected essential oils applied by means of nanoemulsions on the sensory attributes of various foods.

studied fish had good acceptability for all attributes, except for the rosemary EO, which obtained the lowest rating, since it gave a bitter taste to the trout.

Oregano (*O. vulgare*), thyme, and star anise (*Illicium verum*) nanoemulsions have been evaluated EOs Chinese carp [41] and, oregano, thyme, and garlic in salmon and cod [44]. In these studies, oregano EO was found to have the best antimicrobial activity against *Pseudomonas* (reduction of 1 Log CFU/g) and *Listeria* (CMI 0.02%) while also displaying the best sensory acceptability in terms of aroma, color, and texture and extending the shelf life of Chinese carp by 2 d (4 d in total, the shelf life if the control sample without EOs was 2 d). However, the EOs did not have a positive effect on the sensory properties of all the fish. For the cod, the EOs tested masked the characteristic taste of the fish.

Dávila-Rodríguez et al. [42] evaluated the antimicrobial activity of cinnamon (*C. verum*), rosemary, and oregano EOs encapsulated by (O/W) nanoemulsions against *E. coli* and *L. monocytogenes* and compared them to that of the non- encapsulated EOs.

The latter were less effective than the nanoemulsions at inhibiting the same microbial load (5 Log CFU/g); 50% less EO was required when applied as a nanoemulsion. At the same time, the authors performed sensory analysis using a 9-point hedonic scale to evaluate the attributes of aroma, color, texture, appearance, taste, and overall acceptability by applying the EOs nanoemulsions to fresh celery. As a result, they found that celery with cinnamon EO nanoemulsions was the best accepted, as it masked the typical flavor of celery, which is not usually accepted by some consumers. Similarly, [40] found that rosemary EO was the least accepted, due to the herbaceous aroma and bitter taste that this EO imparts, it was not accepted in various types of foods.

2.4 Edible films

The use of biodegradable materials can be an important means for incorporation of EOs. To reduce the consumption of plastic packaging, biodegradable packaging, edible films and coatings have been developed for the storage of fresh or processed foods.

An edible film (EF) is a preformed layer made from edible materials that can be applied to foods to extend their shelf life. They are elaborated from biopolymers such as carbohydrates, proteins, and lipids [45]. Previous studies have shown that films made from carbohydrates are widely utilized because they have good mechanical properties and act as a barrier against low polarity compounds; however, they can present high permeability against moisture [46]. On the other hand, edible films made from proteins, such as gelatin, are excellent barriers to oxygen, carbon dioxide, and some aromatic compounds; but their mechanical properties are not as efficient [47].

In addition to the main polymer matrix, it is necessary to incorporate some additives as plasticizing agents. These are non-volatile substances of low molecular weight that when added to a polymeric material modify its physical and mechanical properties; for example, its flexibility, manageability, and extension ability [48].

2.4.1 Application of essential oils by means of edible films to foods and their effect on sensory attributes

The application of edible packaging as adjuvants for preservation (storage, quality, and safety) of foods, is one of the most effective solutions to ensure their safety [49].

Edible films can be applied as wrapping, packaging method, or layer spacing [8]. However, when the film directly covers the product, it can change its sensory characteristics. On the other hand, migration of the compounds present in the film can occur, similar to the phenomenon of mass transfer in vapor phase, resulting in the same alteration of aromas and flavors. Despite this, very few studies have carried out sensory analysis to evaluate the acceptability of different foods that have utilized edible films with EOs as a preservation method.

Table 4 shows several studies where EOs have been applied to edible films and their effect on the sensory characteristics of selected meat and fish products. In general, EOs modified the studied sensory attributes. Sodium alginate EF with oregano (*O. vulgare*) and rosemary (*Salvia rosmarinus*) EOs have been evaluated in beef fillets [55] and with ginger (*Zingiber officinale*) and oregano EOs in tilapia fillets [56]. Results indicated that EFs with oregano EOs in these two products negatively affected sensory properties (aroma, taste, and overall acceptability) and judges preferred beef and tilapia fillets with rosemary and ginger EOs, respectively. Nonetheless, it has been reported that oregano EO applied in an EF of chitosan in red snapper, maintained good sensory attributes (aroma and taste) even after 20 d of storage [50].

EF	EO	Concentration (%)	Food	Sensory attribute	Results	Reference
Chitosan	Oregano (<i>Origanum vulgare</i>)	0.1	Red pargo	Aroma and taste	Good acceptability for both attributes during 20 d of storage	[50]
Chitosan	Anise (<i>Pimpinella anisum</i>)	1.5 and 2	Chicken burgers	Aroma and taste	Good acceptability across all attributes during 12 d of storage	[51]
Soy	Thyme (<i>Thymus vulgaris</i>) and oregano	1, 2, and 3	Beef	Appearance, color, aroma, taste, texture, and general acceptability of cooked meat	There was not significant difference in appearance, color, aroma, and texture. 1% thyme EO was best accepted in overall taste and acceptability	[52]
Gum farsi	Clove (<i>Syzygium aromaticum</i>) and thyme	1 and 2	Rainbow trout	Texture, color, aroma, and general acceptability	The combination of 2% EOs showed better results in all attributes up to day 12 of storage	[53]
Chitosan	Rosemary (<i>Salvia rosmarinus</i>) and thyme	0.2–0.1	Sucuk	Aroma, color, taste, and overall acceptability of raw and cooked sausages	1% thyme EO was best accepted in all attributes for both sausages	[54]
Sodium alginate	Oregano and rosemary	0.1	Beef steaks	Taste, aroma, and overall acceptability	EF with rosemary EO best accepted in all attributes	[55]
Sodium alginate	Ginger (<i>Zingiber officinale</i>) and oregano	0.1	Tilapia fillets	Aroma, taste, texture, and overall acceptability	Aroma was the only attribute that showed differences. Ginger EO being the best accepted	[56]

EF	EO	Concentration (%)	Food	Sensory attribute	Results	Reference
Cellulose	Oregano, thyme, and cinnamon (<i>Cinnamomum verum</i>)	1	Minced beef	Color, aroma, and taste in cooked meat	Oregano and thyme EOs better accepted in all attributes than cinnamon EO until day 15 of storage	[57]
Starch	Zataria (<i>Zataria multiflora</i>)	6	Ground beef empanadas	Taste, color, and aroma of cooked meat.	Good acceptability in all attributes evaluated	[58]

Table 4.
Effect of selected essential oils applied by means of edible films (EFs) on the sensory attributes of various foods.

A less common EO is zataria (*Zataria multiflora*), its main component is thymol (spice and wood flavor); it was incorporated into a starch EF and conferred good sensory properties on ground meat pies [58]. This same compound is present in thyme (*T. vulgaris*) EO, and it has been evaluated in products such as ground meat (color, aroma, and taste) [57], beef (appearance, color, aroma, taste, and overall acceptability) [52], and sausages (aroma, color, taste, and overall acceptability [54]; all these authors highlight that the use of thyme EO generated good sensory attributes on their studied meat products.

2.5 Liposomes

Liposomes are another encapsulation method and are defined as vesicles with a spherical structure, consisting of a phospholipid bilayer (or two or more of these bilayers, separated by liquid regions) that enclose aqueous or lipid soluble materials. The size of liposomes can vary from a few nanometers to 1 μm in diameter, and they are often used as delivery systems for a wide variety of bioactive compounds such as EOs [59].

The conventional manufacturing process of liposomes is based on the use of organic solvents to dissolve liposome building units, removal of the solvent to form the lipid bilayer, and the subsequent high-energy dispersion of the bilayer in water to form liposomes. However, its low load capacity, the high cost of materials and the need for complex preparation procedures limits its application in the case of large-scale productions [60].

2.5.1 Application of liposome essential oils to foods and their effect on several sensory attributes

To overcome the volatility and instability of EOs, liposomes have been studied to encapsulate EOs and protect them against external environments that could accelerate their volatility [61]. It has been reported that liposomes could improve the chemical stability and bioavailability of hydrophobic antimicrobials such as EOs and

their bioactive compounds [62]. The antimicrobial effect of liposomes with different EOs has been studied, such as: curry, eucalyptus, cinnamon, rosemary, and nutmeg [63–67]; however, their application and study in foods is still scarce.

The few reported studies of liposomes with EOs applied in foods have been evaluated in different types of meats. The antimicrobial effect of different concentrations of chrysanthemum (*Chrysanthemum*) EO encapsulated in triple layer liposomes (chitosan and pectin) against *Campylobacter jejuni* in fresh chicken were evaluated [61]. It was demonstrated that liposomes have a high antimicrobial activity against *C. jejuni* in chicken, reducing 3.2 Log CFU/mL. Furthermore, they performed a sensory analysis with a 7-point hedonic scale and evaluated different attributes such as color, aroma, taste, and overall acceptance of the treated chicken. Results indicated that triple-layer liposomes with chrysanthemum EO could not only minimize chicken deterioration but also maintain its sensory quality.

Moreover, other authors [68] created liposomes to encapsulate laurel (*L. nobilis*) EO (50 µL) and silver nanoparticles, mixed with chitosan in order to coat polyethylene films for packaging pork. Good antimicrobial activity was demonstrated against *E. coli* and *S. aureus*, maintaining the quality of pork at 4°C for 15 d. At the same time, sensory analysis was performed through the descriptive analysis method. The parameters of sensory analysis of meat included color, viscosity, elasticity, and taste. The authors were evaluated and qualified with a scale from 1 to 10. Sensory evaluation showed that liposomes have a positive effect on preserving the quality of pork (color, flavor, viscosity, and elasticity) and could be used to extend the shelf life of fresh pork stored at 4°C.

2.6 Advantages and disadvantages of described methods of application of essential oils

Table 5 lists the advantages and disadvantages of application methods mentioned in this review. Among these, it is emphasized that the use of nanoemulsions or vapor phase require a lower amount of EO to achieve the same effects than those generated by direct contact [69], mainly due to the increase of the contact area [31, 70]. However, negative aspects include that there are no standardized methods for application EOs in vapor phase and that instability of nanoemulsions can generate phase separation and thus not achieve a correct incorporation of EO into foods [31, 70].

In the case of edible films, previous dispersion methods are required for preparation of the film; however, they are effective and replace the use of materials such as plastics [71] which are widely utilized by the food sector. On the other hand, the use of liposomes seems to be a promising method of application; but its use in foods is still scarce, representing a research opportunity for its possible application for food preservation.

Due to the inherent differences in application methods, an adequate comparison among them cannot be properly made; since both the concentration of EOs and their controlled release (or migration) of some of their compounds can alter sensory properties of treated foods. Based on the enumerated studies, the application method will influence sensory characteristics of food products, so selection of the method will depend on both the EO and the food matrix. In general, the use of EOs in foods is generally intended to provide favorable sensory attributes. It is important to emphasize that, according to the active compounds of each EO, they will provide different sensory characteristics for a particular food group, so it is important to discern which compounds are responsible for the aroma and/or flavor.

Method of application	Advantages	Disadvantages	Reference
Direct contact	Easy and fast application. Good antimicrobial activity against different pathogenic microorganisms.	Higher amount of EO is required which affects the sensory properties of food. Provide a strong aroma and taste to the food. High hydrophobicity and volatility.	[69]
Vapor phase	Less AE is required to achieve microbial inactivation. may not affect sensory properties.	It is necessary to create packaging or model systems for your application. The active component of the oils can migrate by diffusion to the surface of the food affecting sensory characteristics. The composition of the food may interact with the EOs, retaining some compounds.	[8, 31]
Nanoemulsions	Have a higher active surface/ volume ratio due to its droplet size, which improves the transport of active compounds through biological membranes. Reduction of EO concentration for microbial inactivation. Sensory acceptance due to low EO concentration.	The use of surfactants and stabilizers is required. Their stability has not been tested in some food systems.	[70]
Edible films	Biodegradable, prevent moisture loss, color fading, lipid oxidation and the formation of unpleasant odors. In addition, can improve the shelf life of some foods.	Investment cost for new production equipment, lack of materials with desired functionalities, lack of regulations. Difficult application in some types of food.	[71]
Liposomes	Good dispersibility in water, good biocompatibility, low diffusion, slow release and prolonged residual activity.	Low effectiveness during storage does not extend the shelf life. Low application in food.	[72]

Table 5.
Advantages and disadvantages of various methods of applying essential oils to food.

3. Final remarks

Chemical compounds derived from natural sources such as reported essential oils have proven their effectiveness as antimicrobials. However, it is important to emphasize that the method of application of these essential oils or their components will depend on factors such as the food to be treated and the type of target microorganisms. In addition, the current challenge lies in their use to increase food safety without

affecting the quality of the product, especially its sensory attributes. In this context, sensory analysis of foods treated with essential oils, or their components is scarce, and opens a field of research to further identify consumer preferences.

Acknowledgements

S. L. Recio-Cázares gratefully acknowledges *Universidad de las Américas Puebla* (UDLAP) and *Consejo Nacional de Ciencia y Tecnología* (CONACyT) for the scholarships granted for her Ph.D. studies.

Conflict of interest

No conflict of interest declared.

Author details

Scarlette L. Recio-Cázares, Aurelio López-Malo and Enrique Palou*
Department of Chemical, Food, and Environmental Engineering, Universidad De Las Américas Puebla, Puebla, Mexico

*Address all correspondence to: enrique.palou@udlap.mx

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Cueva C, Moreno-Arribas MV, Bartolomé B, Salazar Ó, Vicente MF, Bills GF. Antibiosis of vineyard ecosystem fungi against food-borne microorganisms. *Research in Microbiology*. 2011;**162**(10):1043-1051
- [2] Manning L. Categorizing food related illness: Have we got it right? *Critical Reviews in Food Science and Nutrition*. 2013;**57**(9):1938-1949
- [3] Singh P, Gandhi N. Milk preservatives and adulterants: Processing, regulatory and safety issues. *Food Reviews International*. 2015;**31**(3):236-261
- [4] Burt S. Essential oils: Their antibacterial properties and potential applications in foods—A review. *International Journal of Food Microbiology*. 2004;**94**(3):223-253
- [5] Solórzano-Santos F, Miranda-Novales MG. Essential oils from aromatic herbs as antimicrobial agents. *Current Opinion in Biotechnology*. 2012;**23**(2):136-141
- [6] Gyawali R, Ibrahim SA. Natural products as antimicrobial agents. *Food Control*. 2014;**46**:412-429
- [7] Goñi P, López P, Sánchez C, Gómez-Lus R, Becerril R, Nerín C. Antimicrobial activity in the vapour phase of a combination of cinnamon and clove essential oils. *Food Chemistry*. 2009;**116**(4):982-989
- [8] Ribeiro-Santos R, Andrade M, Melo NR, Sanches-Silva A. Use of essential oils in active food packaging: Recent advances and future trends. *Trends in Food Science & Technology*. 2017;**61**:132-140
- [9] Fisher K, Phillips CA. The effect of lemon, orange and bergamot essential oils and their components on the survival of *Campylobacter jejuni*, *Escherichia coli* O157, *Listeria monocytogenes*, *Bacillus cereus* and *Staphylococcus aureus* in vitro and in food systems. *Journal of Applied Microbiology*. 2006;**101**(6):1232-1240
- [10] de Sousa Guedes JP, da Costa Medeiros JA, de Souza Silva RS, de Sousa JM, da Conceição ML, de Souza EL. The efficacy of *Mentha arvensis* L. and *M. piperita* L. essential oils in reducing pathogenic bacteria and maintaining quality characteristics in cashew, guava, mango, and pineapple juices. *International Journal of Food Microbiology*. 2016;**238**:183-192
- [11] Ghabraie M, Vu KD, Tata L, Salmieri S, Lacroix M. Antimicrobial effect of essential oils in combinations against five bacteria and their effect on sensorial quality of ground meat. *LWT—Food Science and Technology*. marzo de. 2016;**66**:332-339
- [12] Sharma H, Mendiratta SK, Agarwal RK, Kumar S, Soni A. Evaluation of anti-oxidant and antimicrobial activity of various essential oils in fresh chicken sausages. *Journal of Food Science and Technology*. 2017;**54**(2):279-292
- [13] Abdeldaiem MHM, Ali HGM, Ramadan MF. Impact of different essential oils on the characteristics of refrigerated carp (*Cyprinus carpio*) fish fingers. *Food Measure*. 2017;**11**(3):1412-1420
- [14] Nikkhah M, Hashemi M. Boosting antifungal effect of essential oils using combination approach as an efficient strategy to control postharvest spoilage and preserving the jujube fruit quality. *Postharvest Biology and Technology*. 2020;**164**:111159

- [15] Bhavaniramy S, Vishnupriya S, Al-Aboody MS, Vijayakumar R, Baskaran D. Role of essential oils in food safety: Antimicrobial and antioxidant applications. *Grain & Oil Science and Technology*. 2019;**2**(2):49-55
- [16] Emir Çoban Ö, Patir B, Yilmaz Ö. Protective effect of essential oils on the shelf life of smoked and vacuum packed rainbow trout (*Oncorhynchus mykiss* W.1792) fillets. *Journal of Food Science and Technology*. 2014;**51**(10):2741-2747
- [17] Michalczyk M, Macura R, Tesarowicz I, Banaś J. Effect of adding essential oils of coriander (*Coriandrum sativum* L.) and hyssop (*Hyssopus officinalis* L.) on the shelf life of ground beef. *Meat Science*. 2012;**90**(3):842-850
- [18] Badola R, Panjagari NR, Singh RRB, Singh AK, Prasad WG. Effect of clove bud and curry leaf essential oils on the anti-oxidative and anti-microbial activity of burfi, a milk-based confection. *Journal of Food Science and Technology*. 2018;**55**(12):4802-4810
- [19] Šojić B, Pavlić B, Tomović V, Ikonić P, Zeković Z, Kocić-Tanackov S, et al. Essential oil versus supercritical fluid extracts of winter savory (*Satureja montana* L.)—Assessment of the oxidative, microbiological and sensory quality of fresh pork sausages. *Food Chemistry*. 2019;**287**:280-286
- [20] Vasilijević B, Mitić-Ćulafić D, Djekic I, Marković T, Knežević-Vukčević J, Tomasevic I, et al. Antibacterial effect of *Juniperus communis* and *Satureja montana* essential oils against *Listeria monocytogenes* in vitro and in wine marinated beef. *Food Control*. 2019;**100**:247-256
- [21] He Q, Li Z, Yang Z, Zhang Y, Liu J. A superchilling storage–ice glazing (SS-IG) of Atlantic salmon (*Salmo salar*) sashimi fillets using coating protective layers of *Zanthoxylum* essential oils (EOs). *Aquaculture*. 2020;**514**:734506
- [22] Pateiro M, Barba FJ, Domínguez R, Santna AS, Mousavi Khaneghah A, Gavahian M, et al. Essential oils as natural additives to prevent oxidation reactions in meat and meat products: A review. *Food Research International*. 2018;**113**:156-166
- [23] Perricone M, Arace E, Corbo MR, Sinigaglia M, Bevilacqua A. Bioactivity of essential oils: A review on their interaction with food components. *Frontiers in Microbiology*. 2015;**6**:1-7
- [24] Reyes-Jurado F, Navarro-Cruz AR, Ochoa-Velasco CE, Palou E, López-Malo A, Ávila-Sosa R. Essential oils in vapor phase as alternative antimicrobials: A review. *Critical Reviews in Food Science and Nutrition*. 2020;**60**:1641-1650
- [25] Saifullah M, Shishir MRI, Ferdowsi R, Tanver Rahman MR, Van Vuong Q. Micro and nano encapsulation, retention and controlled release of flavor and aroma compounds: A critical review. *Trends in Food Science & Technology*. 2019;**86**:230-251
- [26] Mejía-Garibay B, Palou E, López-Malo A. Composition, diffusion, and antifungal activity of black mustard (*Brassica nigra*) essential oil when applied by direct addition or vapor phase contact. *Journal of Food Protection*. 2015;**78**:843-848
- [27] Jiang T, Luo Z, Ying T. Fumigation with essential oils improves sensory quality and enhanced antioxidant ability of shiitake mushroom (*Lentinus edodes*). *Food Chemistry*. 2015;**172**:692-698
- [28] Frankova A, Smid J, Bernardos A, Finkousova A, Marsik P, Novotny D,

- et al. The antifungal activity of essential oils in combination with warm air flow against postharvest phytopathogenic fungi in apples. *Food Control*. 2016;**68**:62-68
- [29] Servili A, Feliziani E, Romanazzi G. Exposure to volatiles of essential oils alone or under hypobaric treatment to control postharvest gray mold of table grapes. *Postharvest Biology and Technology*. 2017;**133**:36-40
- [30] Lorenzo-Leal AC, Palou E, López-Malo A. Evaluation of the efficiency of allspice, thyme and rosemary essential oils on two foodborne pathogens in in-vitro and on alfalfa seeds, and their effect on sensory characteristics of the sprouts. *International Journal of Food Microbiology*. 2019;**295**:19-24
- [31] Císarová M, Hleba L, Medo J, Tančinová D, Mašková Z, Čuboň J, et al. The in vitro and in situ effect of selected essential oils in vapour phase against bread spoilage toxicogenic aspergilli. *Food Control*. 2020;**110**:107007
- [32] Mani López E, Valle Vargas GP, Palou E, López MA. Penicillium expansum Inhibition on Bread by Lemongrass Essential Oil in Vapor Phase. *Journal of Food Protection*. 2018;**81**:467-471
- [33] Patra JK, Das G, Fraceto LF, Campos EVR, Rodriguez-Torres M, Acosta-Torres LS, et al. Nano based drug delivery systems: Recent developments and future prospects. *Journal of Nanobiotechnology*. 2018;**16**:71
- [34] Velázquez-Contreras C, Osorio-Revilla G, Gallardo-Velázquez T. Encapsulation of orange essential oil in a spout-fluid bed dryer with a draft tube on a bed of inert solids. *Drying Technology*. 2014;**32**:1718-1726
- [35] Jin W, Xu W, Liang H, Li Y, Liu S, Li B. Nanoemulsions for food: Properties, production, characterization, and applications. En: *Emulsions*. 2016:1-36
- [36] Walker RM, Decker EA, McClements DJ. Physical and oxidative stability of fish oil nanoemulsions produced by spontaneous emulsification: Effect of surfactant concentration and particle size. *Journal of Food Engineering*. 2015;**164**:10-20
- [37] McClements DJ. Edible nanoemulsions: Fabrication, properties, and functional performance. *Soft Matter*. 2011;**7**(6):2297-2316
- [38] Otoni CG, de Moura MR, Aouada FA, Camilloto GP, Cruz RS, Lorevice MV, et al. Antimicrobial and physical-mechanical properties of pectin/papaya puree/cinnamaldehyde nanoemulsion edible composite films. *Food Hydrocolloids*. 2014;**41**:188-194
- [39] Alfonzo A, Martorana A, Guarrasi V, Barbera M, Gaglio R, Santulli A, et al. Effect of the lemon essential oils on the safety and sensory quality of salted sardines (*Sardina pilchardus* Walbaum 1792). *Food Control*. 2017;**73**:1265-1274
- [40] Ozogul Y, Yuvka İ, Ucar Y, Durmus M, Kösker AR, Öz M, et al. Evaluation of effects of nanoemulsion based on herb essential oils (rosemary, laurel, thyme and sage) on sensory, chemical and microbiological quality of rainbow trout (*Oncorhynchus mykiss*) fillets during ice storage. *LWT*. 2017;**75**:677-684
- [41] Huang Z, Liu X, Jia S, Zhang L, Luo Y. The effect of essential oils on microbial composition and quality of grass carp (*Ctenopharyngodon idellus*) fillets during chilled storage. *International Journal of Food Microbiology*. 2018;**266**:52-59

- [42] Dávila-Rodríguez M, López-Malo A, Palou E, Ramírez-Corona N, Jiménez-Munguía MT. Antimicrobial activity of nanoemulsions of cinnamon, rosemary, and oregano essential oils on fresh celery. *LWT*. 2019;**112**:108247
- [43] Diniz-Silva HT, de Sousa J, da Silva Guedes J, do Egypto R, de Queiroga RC, Madruga MS, et al. A synergistic mixture of *Origanum vulgare* L. and *Rosmarinus officinalis* L. essential oils to preserve overall quality and control *Escherichia coli* O157:H7 in fresh cheese during storage. *LWT*. 2019;**112**:107781
- [44] Pedrós-Garrido S, Clemente I, Calanche JB, Condón-Abanto S, Beltrán JA, Lyng JG, et al. Antimicrobial activity of natural compounds against *listeria* spp. and their effects on sensory attributes in salmon (*Salmo salar*) and cod (*Gadus morhua*). *Food Control*. 2020;**107**:106768
- [45] Sánchez Aldana D, Contreras-Esquível JC, Nevárez-Moorillón GV, Aguilar CN. Caracterización de películas comestibles a base de extractos pécticos y aceite esencial de limón Mexicano. *CyTA—Journal of Food*. 2015;**13**:17-25
- [46] Parra D, Tadini C, Ponce P, Lugao A. Mechanical properties and water vapor transmission in some blends of cassava starch edible films. *Carbohydrate Polymers*. 2004;**58**:475-481
- [47] Jang S-A, Shin Y-J, Song KB. Effect of rapeseed protein-gelatin film containing grapefruit seed extract on 'Maehyang' strawberry quality: Edible film packaging of strawberry. *International Journal of Food Science & Technology*. 2011;**46**(3):620-625
- [48] Osés J, Fernández-Pan I, Mendoza M, Maté JI. Stability of the mechanical properties of edible films based on whey protein isolate during storage at different relative humidity. *Food Hydrocolloids*. 2009;**23**(1):125-131
- [49] Gómez-Estaca J, López-de-Dicastillo C, Hernández-Muñoz P, Catalá R, Gavara R. Advances in antioxidant active food packaging. *Trends in Food Science & Technology*. 2014;**35**(1):42-51
- [50] Vatavali K, Karakosta L, Nathanailides C, Georgantelis D, Kontominas MG. Combined effect of chitosan and oregano essential oil dip on the microbiological, chemical, and sensory attributes of Red Porgy (*Pagrus pagrus*) stored in ice. *Food Bioprocessing and Technology*. 2013;**6**(12):3510-3521
- [51] Mahdavi V, Hosseini SE, Sharifan A. Effect of edible chitosan film enriched with anise (*Pimpinella anisum* L.) essential oil on shelf life and quality of the chicken burger. *Food Science and Nutrition*. 2018;**6**(2):269-279
- [52] Yemiş GP, Candoğan K. Antibacterial activity of soy edible coatings incorporated with thyme and oregano essential oils on beef against pathogenic bacteria. *Food Science and Biotechnology*. 2017;**26**(4):1113-1121
- [53] Dehghani P, Hosseini SMH, Golmakani M-T, Majdinasab M, Esteghlal S. Shelf-life extension of refrigerated rainbow trout fillets using total Farsi gum-based coatings containing clove and thyme essential oils emulsions. *Food Hydrocolloids*. 2018;**77**:677-688
- [54] Demirok Soncu E, Arslan B, Ertürk D, Küçükkaya S, Özdemir N, Soyer A. Microbiological, physicochemical and sensory characteristics of Turkish fermented sausages (sucuk) coated with chitosan-essential oils. *LWT*. 2018;**97**:198-204
- [55] Vital ACP, Guerrero A, Kempinski EMBC, de Monteschio J,

- Sary C, Ramos TR, et al. Consumer profile and acceptability of cooked beef steaks with edible and active coating containing oregano and rosemary essential oils. *Meat Science*. 2018;**143**:153-158
- [56] Vital ACP, Guerrero A, Ornaghi MG, Kempinski EMBC, Sary C, Monteschio JO, et al. Quality and sensory acceptability of fish fillet (*Oreochromis niloticus*) with alginate-based coating containing essential oils. *Journal of Food Science and Technology*. 2018;**55**(12):4945-4955
- [57] Agrimonti C, White JC, Tonetti S, Marmiroli N. Antimicrobial activity of cellulosic pads amended with emulsions of essential oils of oregano, thyme and cinnamon against microorganisms in minced beef meat. *International Journal of Food Microbiology*. 2019;**305**:108246
- [58] Amiri E, Aminzare M, Azar HH, Mehrasbi MR. Combined antioxidant and sensory effects of corn starch films with nanoemulsion of *Zataria multiflora* essential oil fortified with cinnamaldehyde on fresh ground beef patties. *Meat Science*. 2019;**153**:66-74
- [59] Khorasani S, Danaei M, Mozafari MR. Nanoliposome technology for the food and nutraceutical industries. *Trends in Food Science & Technology*. 2018;**79**:106-115
- [60] Fathi M, Vinceković M, Jurić S, Viskić M, Režek Jambrak A, Donsì F. Food-grade colloidal systems for the delivery of essential oils. *Food Reviews International*. 2021;**37**:1-45
- [61] Lin L, Zhu Y, Thangaraj B, Abdel-Samie MAS, Cui H. Improving the stability of thyme essential oil solid liposome by using β -cyclodextrin as a cryoprotectant. *Carbohydrate Polymers*. 2018;**188**:243-251
- [62] Shidhaye S, Vaidya R, Sutar S, Patwardhan A, Kadam V. Solid lipid nanoparticles and nanostructured lipid carriers—Innovative generations of solid lipid carriers. *CDD*. 2008;**5**:324-331
- [63] Valencia-Sullca C, Jiménez M, Jiménez A, Atarés L, Vargas M, Chiralt A. Influence of liposome encapsulated essential oils on properties of chitosan films: Liposome encapsulated essential oils in chitosan films. *Polymer International*. 2016;**65**(8):979-987
- [64] Cui H, Li W, Lin L. Antibacterial activity of liposome containing curry plant essential oil against *Bacillus cereus* in rice. *Journal of Food Safety*. 2017;**37**(2):e12302
- [65] Tang J, Ge Y. Development and evaluation of novel eucalyptus essential oil liposomes/chitosan composite sponges for medical use. *Fibers Polymers*. 2017;**18**(3):424-433
- [66] Maccioni A, Santo A, Falconieri D, Piras A, Manconi M, Maxia A, et al. Inhibitory effect of rosemary essential oil, loaded in liposomes, on seed germination of *Acacia saligna*, an invasive species in Mediterranean ecosystems. *Botany*. 2019;**97**(5):283-291
- [67] Zhu Y, Li C, Cui H, Lin L. Plasma enhanced-nutmeg essential oil solid liposome treatment on the gelling and storage properties of pork meat batters. *Journal of Food Engineering*. 2020;**266**:109696
- [68] Wu Z, Zhou W, Pang C, Deng W, Xu C, Wang X. Multifunctional chitosan-based coating with liposomes containing laurel essential oils and nanosilver for pork preservation. *Food Chemistry*. 2019;**295**:16-25
- [69] Tongnuanchan P, Benjakul S. Essential oils: Extraction, bioactivities,

and their uses for food preservation:
Bioactivities and applications of
essential oils. *Journal of Food Science*.
2014;**79**(7):R1231-R1249

[70] Artiga-Artigas M, Guerra-Rosas MI,
Morales-Castro J, Salvia-Trujillo L,
Martín-Belloso O. Influence of essential
oils and pectin on nanoemulsion
formulation: A ternary phase
experimental approach. *Food
Hydrocolloids*. 2018;**81**:209-219

[71] Umaraw P, Munekata PES,
Verma AK, Barba FJ, Singh VP, Kumar P,
et al. Edible films/coating with tailored
properties for active packaging of
meat, fish and derived products.
Trends in Food Science & Technology.
2020;**98**:10-24

[72] Chen W, Cheng F, Swing CJ, Xia S,
Zhang X. Modulation effect of core-wall
ratio on the stability and antibacterial
activity of cinnamaldehyde liposomes.
Chemistry and Physics of Lipids.
2019;**223**:104790