

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



Modeling the Behavior of Amphiphilic Aqueous Solutions

*Gonzalo Astray Dopazo, Cecilia Martínez-Castillo,
Manuel Alonso-Ferrer and Juan Carlos Mejuto*

Abstract

Two types of predictive models based on artificial neural networks (ANN) and quadratic regression model developed in our laboratory will be summarized in this book chapter. Both models were developed to predict the density, speed of sound, kinematic viscosity and surface tension of amphiphilic aqueous solutions. These models were developed taking into account the concentration, the number of carbons and the molecular weight values. The experimental data were compiled from literature and included different surfactants: i) hexyl, ii) octyl, iii) decyl, iv) tetradecyl and v) octadecyl trimethyl ammonium bromide. Neural models present better adjustment values, with R^2 values above 0.902 and AAPD values under 2.93% (for all data), than the quadratic regression models. Finally, it is concluded that the quadratic regression and the neural models can be powerful prediction tools for the physical properties of surfactants aqueous solutions.

Keywords: amphiphiles, surfactants, physical properties, modeling, artificial neural network

1. Introduction

Amphiphilic compounds have a well-defined structure; two parts clearly differentiated that will determine the behavior in aqueous systems [1] and is the key factor to their relationship with the internal and the external interfaces in aqueous systems [1]. One part of the amphiphilic compound is hydrophilic and the other part is hydrophobic [1, 2] and both are linked by a covalent bond [2].

In aqueous systems, the most important application of surfactants (in volume and economic impact terms), generally a long-chain hydrocarbon group is used as the hydrophobic group (although i) fluorinated, ii) oxygenated hydrocarbon or iii) siloxane chains can also be used) and an ionic or highly polar group as a head or hydrophilic group [3]. The different types of amphiphilic molecules can be differentiated according to the bonds between their two parts, hydrophilic and hydrophobic [2]. For example i) a hydrophilic head can be covalently bound to hydrophobic alkyl chain, whether single, double, or triple, also, ii) an amphiphilic bolaform is formed by two hydrophilic heads covalently linked with a hydrophobic alkyl chain and iii) a Gemini amphiphile is two surfactants covalently linked by their charged heads [2]. These compounds can be also classified based on the chemical nature of their hydrophilic group with subgroups according to the tail, so that, four basic categories can be defined: i) anionic, ii) cationic, iii) nonionic and iv) amphoteric (and zwitterionic) [3].

The property of amphiphiles to self-assemble in aqueous solution to design well-defined structures makes them become interesting molecules that can be applied in different fields [2] such as:

- i. Pharmaceutical to overcome: i) the important manufacturing costs, ii) the poor pharmacokinetic characteristics and iii) the low bacteriological efficiency of the natural cationic antimicrobial peptides (AMPs), using novel and diverse cationic amphiphiles that can mimic the AMPs amphiphilic topology [4], or even as anti-cancer drug delivery vehicles using block copolymer micelles (poly(ethylene oxide) and poly(L-amino acid)) [5],
- ii. in the cleaning sector, where they were used to clean oily deposits from solid surfaces using mixed solutions of fatty acid sulfonated methyl esters and using as cosurfactant dodecyldimethylamine oxide [6]. Yavrukova et al. [6] study the cleaning process of porcelain and stainless steel and concluded that the SME mixtures can be a hopeful system for formulations in household detergency,
- iii. in Chemistry, where this kind of molecules are studied as a developer of supramolecular nanotubes architectures [7],
- iv. in Medical Science to accelerate wound healing using antioxidant shape amphiphiles [8], or
- v. in Food Chemistry using amphiphiles to modulate organoleptic properties in foods post harvested technology or for potential food applications [9, 10], among others.

As previously said, these kinds of molecules can form different types of aggregates. These structures are formed when a certain concentration, called critical micelle concentration (cmc), is reached. This parameter can be defined as the specific concentration for a particular surfactant at which determinate solution properties change strongly [3]. According to Myers [3], different authors showed that the aggregated structure type depends on what is known as critical packing parameter. This parameter ($CPP = v/a_o l_c$) establishes the relationship between the volume of the hydrophobic part of the molecule (v), the optimal area of the head group (a_o) and the critical length of the hydrophobic tail (l_c), and it controls the packing in aggregate structures [3]. The structures that can be formed are i) spherical micelles (when the value of CPP is less than 1/3), ii) cylindrical micelles ($1/3 < CPP < 1/2$), iii) bilayer vesicles ($1/2 < CPP < 1$), iv) lamellar phases ($CPP \approx 1$) and, finally, v) inverted micelles ($CPP > 1$) [11]. Some of these structures are shown in **Figure 1**.

According to Gómez-Díaz et al. [1, 12], different physical properties have been used to characterize the aggregation processes by means of measured different experimental values. These authors have been demonstrate that density and kinematic viscosity do not alter when the micellization point is reached so that they are not utilized to determine knowledge about the behavior of the colloidal aggregate [1, 12]. On the other hand, the variation of the rest of the measured properties, speed of sound and surface tension, can be used to determine the cmc value. The property variation can give rise to the existence of two trend lines which intersection can be used to determine the cmc [1, 12]. As claimed by Gómez-Díaz et al. [1, 12], the cmc value, using the surface tension and the speed of sound was similar. Nevertheless, the cmc value using the surface tension was, for the hexyl, octyl and

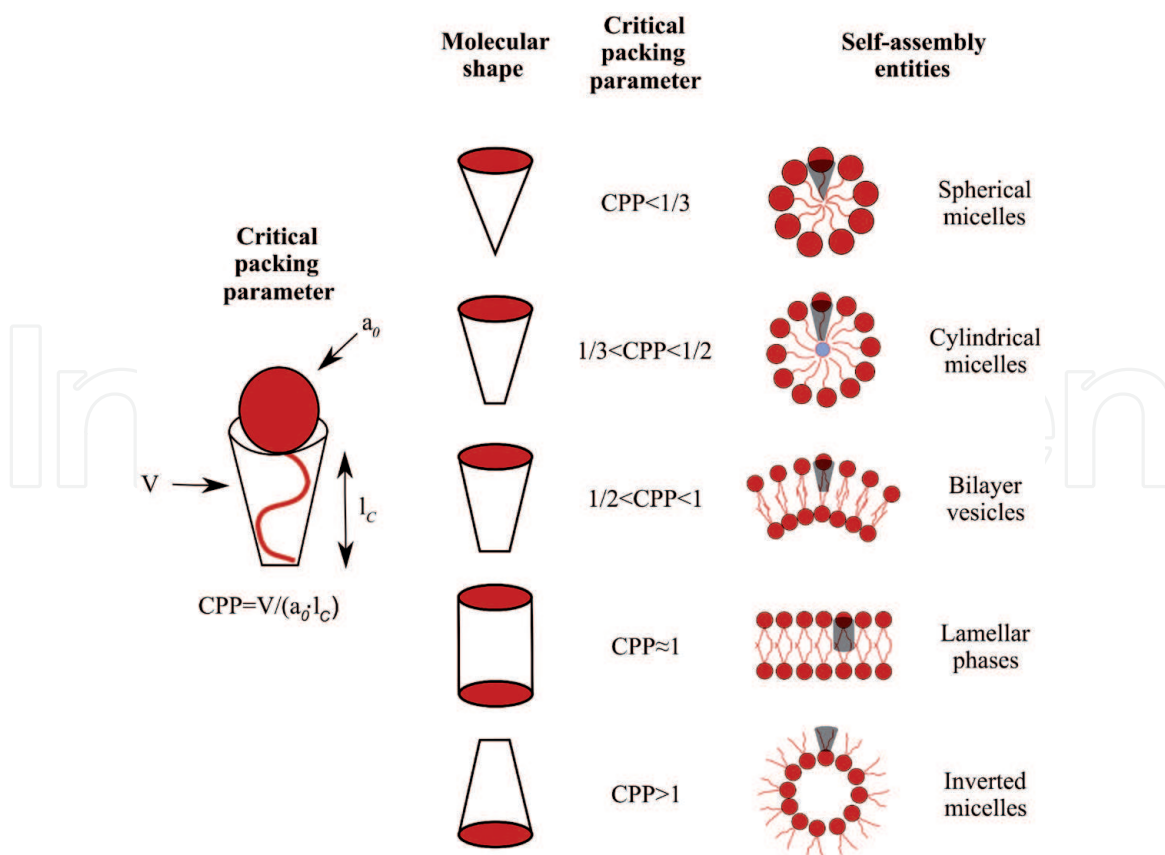


Figure 1. Critical packing parameter (where: v is the lipophilic chain volume; a_0 is the hydrophilic core cross-section area and l_c is the lipophilic chain length), molecular shape and self-assembly entity formed by each different amphiphile. Redraw from the original figure of Wang et al. [11].

decyl trimethyl ammonium bromide, a bit lower than when the speed of sound is used [1, 12] (which can be attributed to the effect of small impurities amounts upon the surface tension value) [1].

The study of solutions behavior to know its properties, and to be able to calculate the cmc, required a lot of work, time-consuming and material cost. Due to these facts, modeling the physical properties of these solutions could help to reduce material and time costs. Thus, the study of methodologies such as artificial neural networks (ANN) and response surface (RS) are interesting and due to this in our research group, a study about this possibility were carried out by Astray & Mejuto [13].

On the one hand, and regarding response surface methodology, it was firstly described by Box and Wilson in 1951 [13, 14]. The RSM is used as a tool for optimization tasks by relating the variables of the process and its response [15, 16]. The experimental data could be fit to a polynomial equation which must describe the data behavior to achieve statistical previsions [17], therefore, this methodology is based on the development of empirical mathematical models to describe the system under study [18]. These models can be used when the response, or responses, are influenced by different variables [17]. An RSM model can work with a reduced amount of experimental trials and can be used to develop, improve and optimize different process [19]. The RSM can use a set of mathematical and statistical tools to fit the experimental data to an Equation [17], usually, linear or square polynomial functions [17, 18]. Different experimental designs could be used which randomizes the experimental error and equals the experimental points distribution, for de independent variables, in the range investigated [20]. RSM models can be applied in different areas such as:

- i. Chemical Engineering to extract alumina from coal fly ash optimizing different variables involved in the process ($K_2S_2O_7/Al_2O_3$ molar ratio, calcining temperature and calcining time) [21],
- ii. in Environmental Science to study the biodegradation of the strobilurin fungicide Pyraclostrobin using bacteria from orange cultivation plots to develop a bioremediation method [22],
- iii. in Biomedical applications to extract anthocyanins from blueberry optimizing the ultrasonic time, ultrasonic temperature, freezing time and liquid–solid ratio [23] or in
- iv. in Biotechnology to optimize the culture media and reduce the production cost of urease bacteria to achieve an eco-friendly process controlling different parameters (yeast extract, whey and heating temperature) [24], inter alia.

On the other hand, artificial neural networks are computational modeling tool that consists of a set of simple processing elements (neurons), massively interconnected capable to process data [20]. This kind of models can try to simulate the path in which the human brain process the information, that it is, ANN are inspired in the biological system [25]. ANN is made up of different neurons layers: an input layer to receive the information, one or more intermediate (or hidden) layers where the information is processed, and an output layer, with one or more neurons, where the predicted value is generated (**Figure 2**). Each neural network is characterized by a specific topology or architecture. To facilitate your identification each neural model implemented can be named such as i-h-o, using the number of neurons presented in the input (i), hidden (h) and output (o) layer [13].

These models present different advantages such as: are non-linearity systems that allow better data fit, are non-sensitivity to noise (uncertain data and measurement errors), present high parallelism (fast processing and failure-tolerance), among others [26]. According to Baş and Boyacı [20], ANNs represent non-linearities better than RS, although ANNs cannot produce a similar model equation to RS models. This kind of approach can be used in a multitude of fields such as:

- i. Engineering to diagnose and classify of bearing faults [27] or to model hot deformation in titanium alloys [28],
- ii. in Food Technology to determine the botanical origin of honey using different parameters (ashes content, electrical conductivity, among others) [29] or food authenticity [30] (carried out in our laboratory),
- iii. in Renewable Energy to predict three components of solar irradiation in Odeillo (France) [31], or
- iv. or in diverse fields (also carried out in our laboratory) such as Palynology [32] or Hydrology [33], among others.

This book chapter summary the quadratic regression and neural models developed in our research group [13] to predict, for amphiphilic aqueous solutions, the i) density (ρ in $g \cdot cm^{-3}$), ii) speed of sound (u in $m \cdot s^{-1}$), iii) kinematic viscosity (ν in $mm^2 \cdot s^{-1}$), and iv) surface tension (σ in $mN \cdot m^{-1}$) taking into account i) concentration (C), ii) carbons number (n °C) and iii) molecular weight (M_w).

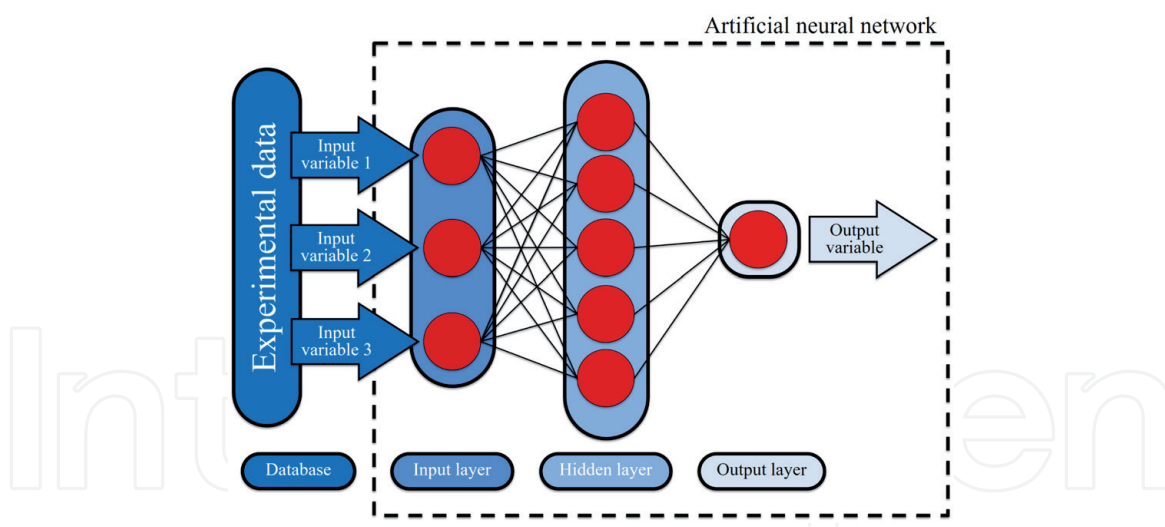


Figure 2.
 Representation of a typical neural network with 3 neurons in the input layer, 5 neurons in one hidden layer and one neuron in the output layer (topology 3–5–1). Redraw from the original figure of Astray & Mejuto [13].

2. Material and methods

2.1 Artificial neural networks as an approximation approach

Artificial Intelligence models based on artificial neural networks have been widely used in the area of chemistry to model and predict processes related to physical properties. This type of model has shown great reliability to model and predict density, dynamic viscosity, and surface tension, among others.

A good example of the use of artificial neural networks to determine properties of interest in micellar systems is the research carried out by Katritzky et al. [34] who developed a model to predict the critical micellar concentration of non-ionic surfactants based on different parameters related to its molecular structure. According to the authors, the models developed could be used for prediction or analysis of new non-ionic surfactants similar to those used in this research. On the other hand, Fatemi et al. [35] developed a model based on artificial neural networks to predict the critical micellar concentration of different anionic and cationic compounds. The selected input variables included the Balaban index, the heat of formation, among others. The results obtained were compared with the predictions of a multiple linear regression model and it was shown that the neural network is superior to multiple linear regression model to predict the log CMC of anionic and cationic surfactants. Along the same line, Kardanpour et al. [36] reported a wavelet neural network (WNN) to predict the critical micellar concentration of Gemini surfactants. The developed model used twelve different descriptors from the molecular structure. According to the authors, the results reveal the ability of the model to determine CMC and demonstrate, in comparison with MLR models, that the models based on neural networks are superior to the MLR approach (due to the ability of the WNN model to work with nonlinearities between the input variables and the CMC).

The researchs listed above demonstrate the ability of artificial neural networks to predict the critical micellar concentration of different surfactants. But to predict the value of this CMC, it is necessary to carry out different experimental studies to determine any particular property that allows determining the CMC value as a function of some abrupt variation of that property. Two of these properties are surface tension and speed of sound whose experimental work requires a great deal of work time and expense in labour and reagents. The different experiments carried out for each variable would determine the CMC as a function of the intersection

of the two trend lines (as mentioned above [1, 12]). Due to these facts, an ANN approach could be very useful to lower costs and be able to make approximations easier, so designing models that are capable to predict this variable depending on the different mixtures could be a very recommendable tool. The claim that artificial neural networks are useful tools because they can minimize the time of experimental treatment and operating costs can be contrasted in different studies reported in the bibliography. An example of this, is the study carried out by Belhaj et al. [37] in which they use artificial neural networks to predict absorption values for alkyl ether carboxylate (AEC) and alkyl polyglucoside (APG). Thus, this book chapter summarizes the research carried out in our research group to predict density, speed of sound, kinematic viscosity and surface tension of amphiphilic aqueous solutions [13].

In addition to our work, surface tension modeling was also carried out by different authors. Khazaei et al. [38] developed an ANN to predict the surface tension of multicomponent mixtures at different temperatures was employed. The input variables were: reduced temperature, critical pressure and volume, and an acentric factor of the mixture. The obtained average absolute relative deviations were low and the ANN model, compared with well-known models (Brook-Bird equation, Flory theory and group contribution theory) has proved a high prediction capacity. The authors concluded that ANN can be helpful for engineering calculations and they emphasized that the ANN model can be a robust approach to predict complex input-output systems. Other interesting research was carried out by Gharagheizi et al. [39] developed neural models to determine the surface tension of pure compounds at different temperatures and atmospheric pressure. The authors investigated compounds belonging to 78 different chemical families and the results were satisfactory (according to different statistical parameters) with an absolute average deviation of 1.7% and a squared correlation coefficient of 0.997. On the other hand, Bakeri et al. [40] used 20 hydrocarbons mixtures to determine the surface tension. The model developed by the authors showed the best accuracy when they are compared with other four well-known classical models. On the other hand, density and kinematic viscosity, in this case, for different systems of biofuels and their blends with diesel fuel, can be predicted using ANNs [41]. In this case, two artificial neural networks were developed to predict kinematic density and viscosity. The models developed used 6 input variables, temperature, volume fractions, among others. The results reported by the authors indicate that the models obtained good correlations. Density and speed of sound of binary ionic liquid and ketone mixtures can also be predicted by ANNs [42]. In this case, the artificial neural network models used as input variables, the temperature and the mole fraction, among others, to determine these two variables. The models developed presented an overall average percentage error lower than 2.5%, so the authors concluded that this model was applicable for the prediction of these variables in binary ionic liquid and ketone mixtures.

Nevertheless, the use of artificial neural networks is not only limited to the prediction of the previous properties, ANNs can also be used to tensammetric analysis of different nonionic surfactants (Brij 30, 35, 56 and 96) [43]. Authors concluded that ANNs can be a possible candidate to determine nonionic surfactants. Another interesting study is the one developed by Jha et al. [44] that developed a feedforward artificial neural network with three layers to predict the diffusion coefficient of a micellar system with sodium dodecyl sulfate (SDS). The model uses the temperature and NaCl and SDS concentrations as input variables. The ANN is capable to model the experimental behavior (correlation coefficient upper than 0.99) and it is concluded that the model is usable to calculate this property. ANN models can also be used to investigate the different factors that affect particle size

in a Nanoemulsion System (Virgin Coconut Oil) that contain copper peptide [45]. The model used, to predict the particle size, four input variables composed of the amount of virgin coconut oil, Tween 80:Pluronic F68, xanthan gum and water. The ANN demonstrated its ability to model the particle size according to the four input variables and showed good determination coefficients upper than 0.97. Finally, another interesting research is that carried out by Rocabruno-Valdés et al. [46] in which the authors develop artificial neural models to predict different properties (dynamic viscosity, density and cetane number of biodiesel) using as input variables the temperature, the number of carbon and hydrogen atoms and methyl esters composition. The correlation coefficients obtained were upper than 0.91. According to the authors, the ANN models provide an adequate prediction and can be interesting for their inclusion in simulators.

2.2 Database

To carry out this work, the experimental data obtained by Gomez et al. [1, 12] were used. The used surfactants were: hexyl trimethyl ammonium bromide (HTABr), octyl trimethyl ammonium bromide (OTABr) and decyl trimethyl ammonium bromide (DTABr) from [1] and tetradecyl trimethyl ammonium bromide (TDTABr) and octadecyl trimethyl ammonium bromide (ODTABr) from [12]. All these reagents were supplied by Fluka with a purity $\geq 98\%$. [1, 12]. The authors prepared the aqueous solutions by mass using an analytical balance Kern 770 (precision 10^{-4} g) [1, 12].

The output variables for each aqueous solution were determined (at 298 K) with different instruments: i-ii) the density and speed of sound using an Anton Paar DSA 5000 vibrating-tube densimeter and sound analyzer, iii) the kinematic viscosity by means the transit time for liquid meniscus through a capillary viscosimeter (supplied by Schott) and iv) the surface tension using a tensiometer Krüss K-11 using the Wilhelmy plate method [1, 12].

2.3 Modeling procedure for predictive models

The surface model, which is used to evaluate the influence of each input variable on the physical properties (density, speed of sound, kinematic viscosity and surface tension) used the combination of input variables linearly, quadratically and cross-correlated [13]. That is, experimental data can be approximate using a generalized second-order polynomial model (Eq. (1) [47]). In this sense, the model was used to correlated each dependent variable (y_{pred}) using the input variables (x), the regression coefficients (b) and the random error value (ε).

$$y_{pred} = b_0 + \sum_{j=1}^k b_j x_j + \sum_{j=1}^k \sum_{i=1}^k b_{ji} x_j x_i + \sum_{j=1}^k b_{jj} x_j^2 + \varepsilon \quad (1)$$

The response surface methodology was created to carry out the experiments with previous analysis of the relationship between the variables (generally standardized), with a homogeneous distribution of the experiments [13]. Nevertheless, in this case, the experimental data used are not homogeneously distributed and the data have not been standardized [13].

The other predictive model used is based on artificial neural networks. The ANN require to split the data into at least two different groups -training (T) and validation (V)-, which has been carried out by the authors [13] randomly. The set of training data was used to train the ANN model, while the validation data set is used to check the good training of the model [13]. An important aspect of this

methodology is that it based on the trial-error procedure [13] to find the optimal combination of parameters for prediction. Once the database is presented to the input layer, the training can start, the data are propagated to the first intermediate layer and the information is treated by the propagation function (Eq. (2)) to obtain a single value (S_i), being: x_f the input data (in the input neuron f), w_{fi} the weight (importance among the neurons f and i) and b_i the bias value associated with the intermediate neuron i . [13]. The single response is processed by the activation function (Eq. (3)) being: S_{res} the output value of the neuron [13]. This process is repeated throughout all the neurons in the intermediate and output layer where the predicted value is generated.

$$S_i = \sum_{f=1}^N w_{fi} x_f + b_i \quad (2)$$

$$S_{res} = \frac{1}{1 + e^{-S_i}} \quad (3)$$

2.4 ANN's parameters

The authors [13] used a total of 80 cases to develop different prediction models (RS and ANN). In this case, the database was divided into two groups. A first group, with 75% of the cases (60), to train the model and a second group, with the remaining 25%, to validate the model (20) [13].

The learning rate and momentum values were set at 0.7 and 0.8, respectively. The models were developed at different training cycles in order to locate the point from which could be overtrained.

2.5 Adjustments parameters

The results were analyzed by the authors [13] using different statistics to determine the adjustment power, such as the coefficient of determination (R^2), the root mean square error (RMSE) (Eq. (4)) or the average absolute percentage deviation (AAPD) (Eq. (5)) for the training and the validation phases. Individual percentage deviations (IPD) is also used.

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (y_{pred} - y_{experimental})^2}{n}} \quad (4)$$

$$AAPD = \frac{\sum_{j=1}^n \left(\left| \frac{y_{pred} - y_{experimental}}{y_{experimental}} \right| \right) \cdot 100}{n} \quad (5)$$

2.6 Computer equipment and software

The input variables, necessary to determine the desired variables, were obtained from the Sigma Aldrich and Chemdraw Professional 15 trial (PerkinElmer) [13]. Microsoft Excel Professional Plus 2013 (Microsoft) was used for RS modeling, and the software EasyNN plus v14.0d (Neural Planner Software Ltd.) was used to ANN modeling [13]. A computer server with an Intel® processor Core™ i7 processor with 16 GB of RAM was used to develop the models [13].

The figures of this book chapter were made with Inkscape 0.92 and Microsoft PowerPoint Professional Plus 2016 (Microsoft).

3. Results and discussion

The adjustments for the models developed [13] are shown in **Table 1**. It can be seen heterogeneous results for the surface and neural models.

Response surface models present good determination coefficients in the training phase, varying between the value obtained for the density model (0.994) and the value obtained for the kinematic viscosity model (0.906). These good values contrast with the value obtained for the surface tension model which reports a low determination coefficient value (0.505).

For the first three models (density, speed of sound and kinematic viscosity) the values of determination coefficient obtained in the validation phase are similar (with a minimal descent to the obtained R^2 values in the training phase) varying between the value obtained for the density model (0.985) and the R^2 value obtained for the kinematic viscosity model (0.885). The response surface model, with the worst-performing behavior for the training phase, the model developed to predict surface tension, showed, for the validation phase, a determination coefficient of 0.503, similar to that obtained in the training phase (0.505).

Regarding the root mean square error values obtained by the response surface models developed by Astray & Mejuto [13], it can be seen that the density model present an RMSE value around $0.001 \text{ g}\cdot\text{cm}^{-3}$, in both phases, the speed of sound models around $7.1837 \text{ m}\cdot\text{s}^{-1}$ and $6.3226 \text{ m}\cdot\text{s}^{-1}$ in training and validation phase, respectively. The model developed to predict kinematic viscosity presents an RMSE value around $0.1003 \text{ mm}^2\cdot\text{s}^{-1}$ for the training phase and $0.0569 \text{ mm}^2\cdot\text{s}^{-1}$ for validation phase, and the worst model developed, the surface tension model, $8.3304 \text{ mN}\cdot\text{m}^{-1}$ and $8.1307 \text{ mN}\cdot\text{m}^{-1}$, for training and validation phase, respectively. The size of these errors can best be understood if they are given in terms of average absolute percentage deviation. The AAPD values reported for each phase are very similar. In this case, the errors obtained for each model (for all data) were: 0.08%, 0.31%,

Model	Training phase		Validation phase	
	R^2	RMSE	R^2	RMSE
RS_ρ	0.994	0.0012	0.985	0.0011
RS_u	0.976	7.1837	0.972	6.3226
RS_ν	0.906	0.1003	0.885	0.0569
RS_σ	0.505	8.3304	0.503	8.1307
ANN_ρ	0.999	0.0004	0.999	0.0003
ANN_u	0.998	1.9393	0.998	1.7093
ANN_ν	0.999	0.0108	0.994	0.0104
ANN_σ	0.449	9.6859	0.457	9.6827
ANN'_σ	0.985	1.4956	0.940	2.8593

Table 1.
Adjustments for training and validation phase for the RS and ANN models selected by the authors [13]. Determination coefficient (R^2) and root mean square error (RMSE) for the models developed by surface (RS) and neural models (ANN). The subscript ρ corresponds to the variable density, u to the speed of sound, ν to the kinematic viscosity and σ is the surface tension. Table adapted from data reported by Astray & Mejuto [13].

5.18% and 14.73%, for density, speed of sound, kinematic viscosity and surface tension model, respectively. It can be seen how the AAPD value for the density and speed of sound prediction models are very low, the error of the kinematic viscosity model presents an error of 5.18% that can be considered feasible. In these cases, the error that is not acceptable is the one reported by the surface tension model (14.73%) since it is clearly much higher than the rest, and above the 10% which is considered, in our laboratory, as an acceptable error.

With all this, it can be said that the models designed to determine the density, the speed of sound and the kinematic viscosity are useful models for the prediction of these properties. The model to predict the surface tension should not be used due to its high AAPD.

The adjustments for the ANN models developed [13] can be shown in **Table 1**. ANN models were developed based on the trial-error method to obtain the best models for each predict output variable (more than 400 neural networks were developed) [13]. All models developed by the authors [13] presented a different topology: i) 3–7–1 for the density model, ii) 3–5–1 for the speed of sound model, iii) 3–6–1 to the kinematic viscosity model and iv) 3–1–1 for the surface tension model. Thus, each model presents, in the input layer, three variables: concentration, number of carbons and molecular weight and intermediate layer of each model varies from a single neuron, to predict the surface tension, to seven in the density model, in addition to that, each selected model has a different number of training cycles [13].

It can be observed (**Table 1**) that, in general, the ANN provided by the authors [13] adjust, properly, the desired variables, both in the training and in the validation phase. The model to predict the density value is the model with the best adjustments, in fact, and take into account the adjustments in terms of determination coefficient and root mean square error, this model presents values of 0.999 and $0.0004 \text{ g}\cdot\text{cm}^{-3}$, respectively, for the training phase and values of 0.999 and $0.0003 \text{ g}\cdot\text{cm}^{-3}$, respectively, for validation phase. Once again, as was RS models case, the model to predict the density is the model with the best adjustments, in fact, the AAPD values reported for both phases were 0.02%.

The behavior of the rest of the models follows the pattern of the RS models, that is, the models to predict the speed of sound and the kinematic viscosity are, in this order, the second and the third-best model [13].

The model destined to predict the speed of sound presents adjustments, in terms of coefficient of determination, very close to the model destined to predict density (0.998 in both phases), presenting relatively low RMSE values ($1.9393 \text{ m}\cdot\text{s}^{-1}$ and $1.7093 \text{ m}\cdot\text{s}^{-1}$).

The kinematic viscosity model has slightly lower adjustment than the previous two models. In this sense, and always in terms of the determination coefficient, the value for the training phase remains similar to the two previous models, however, for the validation phase, this value falls slightly to 0.994. Even so, the model seems to be predicting the kinematic viscosity values correctly, especially if it be taking into account the low RMSE values ($0.0108 \text{ mm}^2\cdot\text{s}^{-1}$ and $0.0104 \text{ mm}^2\cdot\text{s}^{-1}$, for training and validation, respectively) [13].

Finally, the worst model developed using artificial neural networks is the model designed to determine surface tension [13]. It can be seen in **Table 1** show the values obtained fall significantly, in fact, the determination coefficient value falls to 0.449 and 0.457 for the training and validation phase, respectively. It seems clear that this low value of determination coefficient indicates the impossibility of the model to make correct predictions. This fact is demonstrated with the high RMSE values obtained for the training and validation phase ($9.6859 \text{ mN}\cdot\text{m}^{-1}$ and $9.6827 \text{ mN}\cdot\text{m}^{-1}$, respectively).

As stated above, the size of the errors made by the different ANN models can best be understood in terms of AAPD. In this case, the errors obtained (for all data) by density, speed of sound, kinematic viscosity and surface tension model were: 0.02%, 0.10%, 0.62% and 18.13%, respectively.

In the same way that occurs with the surface models, the ANN surface tension model should not be used to predict surface tension (AAPD above 10%). The other three models can be used for prediction.

3.1 Comparison of response surface and neural models

Once the models have been analyzed separately, it is necessary to make a comparison between them.

As previously stated, the models to predict density are the best models according to the adjustments. This means that this model is useful to predict physical properties of surfactants aqueous solutions (at least with the surfactants studied).

On the one hand, although in general, the AAPD in the RS_ρ model is around 0.08%, according to the authors [13], some cases present a bigger IPD value (0.25–0.49%). Even so, these values are very low. Despite the good performance of this RS model, the ANN model seems to work a little better, improving each adjustment parameter (see **Table 1**). In fact, the AAPD values in the case of the ANN_ρ are below to the value obtained by the RS_ρ model. This improvement is observable in terms of RMSE being, for both phases together (with all the data), very significant ($0.0012 \text{ g}\cdot\text{cm}^{-3}$ vs. $0.0004 \text{ g}\cdot\text{cm}^{-3}$) which represents an important improvement. For both models, it seems clear that the most important variable to determine the density is the concentration with an importance value around 59.00% for the ANN_ρ model and around 89.25% for ANN_ρ model [13].

The second-best models, based on their adjustments, are the models to predict the speed of sound. The RS_u and ANN_u model developed by Astray & Mejuto [13], present good results, in fact, the RS_u model presents, for all data, an R^2 value of 0.974 (with some cases presenting an IPD >1%), while for the ANN_u presents a better value of determination coefficient (0.998), representing a slight improvement of 2.46%. The same behavior occurs regarding the RMSE, where the ANN_u model improved this parameter by around 73.00%. The authors [13] reported that the ANN model has an AAPD value of around 0.10% and a highest IPD value around 0.45%. In both cases, very similar values are obtained for the training and validation phases. Concerning the importance of the variables, in the same way, that the models developed to predict the density, the most important variable to determine the speed of sound is the concentration with an importance value of 89.31% for the RS_u model and 63.30% for the ANN_u model [13].

The third-best model according to its results is the model to predict the kinematic viscosity. In this case, the behavior of the RS_v model is slightly different from the one presented by the ANN_v model. Thus, it is observed that the RS_v model cannot predict with accuracy the kinematic viscosity and showed a slight dispersion of the data (predicted vs. experimental) that can be seen in the figure presented by Astray & Mejuto [13]. This dispersion is reflected in the adjustment parameters of the RS_v model that presents, for all cases, a determination coefficient of 0.903 and an AAPD value of 5.18%). It is noteworthy that, according to Astray & Mejuto [13], there are more than 30 cases with an IPD value in the range of 5.34% to 26.51%. On the other hand, the model based on ANN, predicts with accuracy for the training and the validation phase, showed an R^2 upper than 0.993. According to the AAPD values provided, for all data, the AAPD value obtained by the ANN_v model (0.62%) compared to the AAPD value of the RS_v model (5.18%) represents a decrease around 88.05% [13]. Regarding the input variables, in the same way,

that the models developed to predict the density and the speed of sound, the most important variable to determine the kinematic viscosity is the concentration [13].

Finally, from all models developed, both surface tension models were the worst models according to their adjustments. These models are the models with the highest dispersion (RS_{σ} model can be seen in Figure 3.G [13]). The adjustment parameters for all data are low, in fact, the determination coefficient for the RS_{σ} model is around 0.503 (being even lower in the case of ANN_{σ} model -0.451). Taking into account the AAPD values for all data, it can be seen how neither the RS_{σ} model (14.73%) and the ANN_{σ} model (18.13%) are capable to predict with accuracy the surface tension value. Regarding the input variables, once again, the most important variable to determine the surface tension is the concentration [13].

Due to these poor results, the authors [13] proposed an alternative ANN model (ANN'_{σ}) to improve the prediction of surface tension. In this new ANN model, the input variables were increased using the predictions of the ANN models of density, speed of sound and kinematic viscosity, that is, this new ANN model presents an input layer with six variables. The ANN'_{σ} model needs more cycles for its training (12800 cycles) compared to the three input variables ANN model (1200 cycles). It can be shown a notable improvement in the adjustments for the training and the validation phases. The determination coefficient increases to 0.985 for the training phase and 0.940 for the validation phase, while the RMSE values decrease from the $9.6851 \text{ mN}\cdot\text{m}^{-1}$ (for all case) to $1.9291 \text{ mN}\cdot\text{m}^{-1}$. In the same way, an improvement in the AAPD values is observed, which, in the new model, is below 3.86%. This new model takes the predicted speed of sound as the most important variable, unlike the previous ones, where the most important variable was focused on the concentration. The predicted speed of sound is followed by the predicted kinematic viscosity, the predicted density and the concentration (number of carbons and molecular weight present lower importance).

Given the results obtained by the surface models and the neural models [13], it can be concluded that the models developed to determine density, sound speed and kinematic viscosity are models suitable for their use in the laboratory due to the low AAPD values that presented (between 0.02% and 5.18%, for all the data cases). Regarding the models for surface tension prediction, as previously mentioned, these cannot be used for laboratory use, because they present errors upper than 10%. The alternative ANN model developed by the authors [13], appears to offer acceptable results in terms of determination coefficient and AAPD value. This alternative model improves the original RS and ANN model.

All the models developed [13] can be improved in different ways. The response surface models could be improved by adapting the experimental cases to an experimental design before the experimental measurements, allowing on the one hand to save economic costs and time, and on the other, favoring the development of an RS model based on a precise experimental design. It would also be very convenient to develop a response surface model trying to find surfactants that allow the variables to vary constantly in a range. All these improvements could favor the improvement of the models destined to predict the density, the speed of the sound and the kinematic viscosity.

Likewise, and given the ANN model that uses six input variables [13], it would be interesting to develop an RS model that includes the predictions of density, speed of sound and kinematic viscosity as input variables of the model (although it would be necessary to see how to treat the different variation of the values within the range under study).

Neural network models could be improved by including different input variables that are capable of better identification of the different surfactants. Another interesting approach could be the increase the database for their modeling.

4. Conclusion

The development of models based on response surfaces and neural networks to predict different physical properties of surfactants aqueous solutions (i) density, ii) speed of sound, iii) kinematic viscosity and iv) surface tension) can be a good alternative to save money and time in the laboratory.

In general terms, this kind of models can adjust, with accuracy, the density, the kinematic viscosity and the speed of sound with determination coefficient upper than 0.902 and lower APPD values than 5.20% (for all data). In contrast to these good adjustments, surface tension models do not work properly and presented (for all data) low determination coefficients (0.503 and 0.451 for RS and ANN model, respectively) and high APPD values (14.73% and 18.13% for RS and ANN model, respectively). It seems that this problem can be solved, in the case of models based on neural networks, with the inclusion of new variables from the predictions of the previous models. With this modification, the new neural model improves (for all data) each adjustment parameter (0.974 and 2.92% for determination coefficient and AAPD value, respectively).

In conclusion, RS and ANN models can be powerful prediction tools for the properties (density, speed of sound, kinematic viscosity or surface tension) of surfactants aqueous solutions. These models could therefore facilitate daily laboratory work, saving time and money. However, it would be interesting to improve the models using other development alternatives or, even, improve these model using different approaches such as support vector machines or random forests, among others.

Acknowledgements

Gonzalo Astray thanks to the University of Vigo for his contract supported by “Programa de retención de talento investigador da Universidade de Vigo para o 2018” budget application 0000 131H TAL 641. Cecilia Martínez-Castillo and Manuel Alonso-Ferrer thanks to the University of Vigo for her contract supported by FEADER 2018/002B project (Xunta de Galicia, Consellería de Medio Rural, Project “Desarrollo de modelos de predicción de origen en vinos de denominaciones de origen gallegas”).

IntechOpen

Author details

Gonzalo Astray Dopazo^{1,2*}, Cecilia Martínez-Castillo³, Manuel Alonso-Ferrer³ and Juan Carlos Mejuto¹

1 Departamento de Química Física, Facultade de Ciencias, Universidade de Vigo, Ourense, España

2 CITACA, Universidade de Vigo Campus Auga, Ourense, España

3 Grupo de Nutrición y Bromatología, Departamento de Química Analítica y Alimentaria, Facultade de Ciencias, Universidade de Vigo, Ourense, España

*Address all correspondence to: gastray@uvigo.es

IntechOpen

© 2021 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] Gómez-Díaz D, Navaza JM, Sanjurjo B. Density, kinematic viscosity, speed of sound, and surface tension of hexyl, octyl, and decyl trimethyl ammonium bromide aqueous solutions. *J Chem Eng Data*. 2007;52(3): 889-91.
- [2] Zhang X, Wang C. Supramolecular amphiphiles. *Chem Soc Rev*. 2011;40(1):94-101.
- [3] Myers D. Surfactant Science and Technology: Third Edition. Surfactant Science and Technology: Third Edition. 2006. 380 p.
- [4] Findlay B, Zhanel GG, Schweizer F. Cationic amphiphiles, a new generation of antimicrobials inspired by the natural antimicrobial peptide scaffold. *Antimicrob Agents Chemother*. 2010;4049-58.
- [5] Kwon GS, Kataoka K. Block copolymer micelles as long-circulating drug vehicles. *Adv Drug Deliv Rev*. 1995;16(2-3):295-309.
- [6] Yavrukova VI, Shandurkov DN, Marinova KG, Kralchevsky PA, Ung YW, Petkov JT. Cleaning Ability of Mixed Solutions of Sulfonated Fatty Acid Methyl Esters. *J Surfactants Deterg*. 2020;23(3):617-27.
- [7] Shimizu T, Masuda M, Minamikawa H. Supramolecular nanotube architectures based on amphiphilic molecules. *Chem Rev*. 2005;105(4):1401-43.
- [8] Li Z, Zhang J, Fu Y, Yang L, Zhu F, Liu X, et al. Antioxidant shape amphiphiles for accelerated wound healing. *J Mater Chem B*. 2020 Aug;8(31):7018-23.
- [9] Cid A, Morales J, Mejuto JC, Briz-Cid N, Rial-Otero R, Simal-Gándara J. Thermodynamics of sodium dodecyl sulphate-salicylic acid based micellar systems and their potential use in fruits postharvest. *Food Chem*. 2014 May;151:358-63.
- [10] Cid A, Mejuto JC, Orellana PG, López-Fernández O, Rial-Otero R, Simal-Gandara J. Effects of ascorbic acid on the microstructure and properties of SDS micellar aggregates for potential food applications. *Food Res Int* [Internet]. 2013;50(1):143-8. Available from: <http://www.sciencedirect.com/science/article/pii/S0963996912004206>
- [11] Wang N, Chen M, Wang T. Liposomes used as a vaccine adjuvant-delivery system: From basics to clinical immunization. *J Control Release*. 2019;303:130-50.
- [12] Gómez-Díaz D, Navaza JM, Sanjurjo B. Density, kinematic viscosity, speed of sound, and surface tension of tetradecyl and octadecyl trimethyl ammonium bromide aqueous solutions. *J Chem Eng Data*. 2007;52(5):2091-3.
- [13] Astray G, Mejuto JC. Approach of different properties of alkylammonium surfactants using artificial intelligence and response surface methodology. *Tenside, Surfactants, Deterg*. 2017;54(2):132-40.
- [14] Box GEP, Wilson KB. On the Experimental Attainment of Optimum Conditions. Johnson NL, editor. *J R Stat Soc Ser b*. 1951;13(1):1-45.
- [15] Yang QQ, Gan RY, Zhang D, Ge YY, Cheng LZ, Corke H. Optimization of kidney bean antioxidants using RSM & ANN and characterization of antioxidant profile by UPLC-QTOF-MS. *LWT*. 2019;114:108321.
- [16] Natabirwa H, Nakimbugwe D, Lung'aho M, Muyonga JH. Optimization of Roba1 extrusion conditions and bean extrudate properties using response surface methodology and multi-response desirability function. *LWT*. 2018;96:411-8.

- [17] Bezerra MA, Santelli RE, Oliveira EP, Villar LS, Escalera LA. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta*. 2008;76(5):965-77.
- [18] Teófilo RF, Ferreira MMC. Quimiometria II: planilhas eletrônicas para cálculos de planejamentos experimentais, um tutorial. *Quim Nova*. 2006;29(2):338-50.
- [19] Sarkar M, Majumdar P. Application of response surface methodology for optimization of heavy metal biosorption using surfactant modified chitosan bead. *Chem Eng J*. 2011;175:376-87.
- [20] Baş D, Boyacı IH. Modeling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. *J Food Eng*. 2007;78(3):846-54.
- [21] Guo C, Zhao L, Yang J, Wang K, Zou J. A novel perspective process for alumina extraction from coal fly ash via potassium pyrosulfate calcination activation method. *J Clean Prod*. 2020;271:122703.
- [22] Birolli WG, da Silva BF, Rodrigues-Filho E. Biodegradation of the fungicide Pyraclostrobin by bacteria from orange cultivation plots. *Sci Total Environ* [Internet]. 2020;746:140968. Available from: <http://www.sciencedirect.com/science/article/pii/S0048969720344971>
- [23] Yuan J, Li H, Tao W, Han Q, Dong H, Zhang J, et al. An effective method for extracting anthocyanins from blueberry based on freeze-ultrasonic thawing technology. *Ultrason Sonochem*. 2020;68:105192.
- [24] Kahani M, Kalantary F, Soudi MR, Pakdel L, Aghaalizadeh S. Optimization of cost effective culture medium for *Sporosarcina pasteurii* as biocementing agent using response surface methodology: Up cycling dairy waste and seawater. *J Clean Prod*. 2020;253:120022.
- [25] Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J Pharm Biomed Anal*. 2000;22:717-27.
- [26] Basheer IA, Hajmeer M. Artificial neural networks: Fundamentals, computing, design, and application. *J Microbiol Methods*. 2000;43:3-31.
- [27] Agrawal P, Jayaswal P. Diagnosis and Classifications of Bearing Faults Using Artificial Neural Network and Support Vector Machine. *J Inst Eng Ser C* [Internet]. 2020 Feb 7;101(1):61-72. Available from: <http://link.springer.com/10.1007/s40032-019-00519-9>
- [28] Zhao J, Ding H, Zhao W, Huang M, Wei D, Jiang Z. Modelling of the hot deformation behaviour of a titanium alloy using constitutive equations and artificial neural network. *Comput Mater Sci*. 2014;92:47-56.
- [29] Anjos O, Iglesias C, Peres F, Martínez J, García Á, Taboada J. Neural networks applied to discriminate botanical origin of honeys. *Food Chem*. 2015 May;175:128-36.
- [30] Martinez-Castillo C, Astray G, Mejuto JC, Simal-Gandara J. Random Forest, Artificial Neural Network, and Support Vector Machine Models for Honey Classification. *eFood*. 2020;1(1):69-76.
- [31] Benali L, Notton G, Fouilloy A, Voyant C, Dizene R. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renew Energy*. 2019;132:871-84.

- [32] Iglesias-Otero MA, Fernández-González M, Rodríguez-Caride D, Astray G, Mejuto JC, Rodríguez-Rajo FJ. A model to forecast the risk periods of *Plantago* pollen allergy by using the ANN methodology. *Aerobiologia* (Bologna). 2015;31(2):201-11.
- [33] Araujo P, Astray G, Ferrerio-Lage JA, Mejuto JC, Rodríguez-Suarez JA, Soto B. Multilayer perceptron neural network for flow prediction. *J Environ Monit*. 2011 Jan;13(1):35-41.
- [34] Katritzky AR, Pacureanu LM, Slavov SH, Dobchev DA, Karelson M. QSPR study of critical micelle concentrations of nonionic surfactants. *Ind Eng Chem Res*. 2008;47(23):9687-95.
- [35] Fatemi MH, Konuze E, Jalali-Heravi M. Prediction of critical micelle concentration of some anionic and cationic surfactants using an artificial neural network. *Asian J Chem*. 2007;19(4):2479-89.
- [36] Kardanpour Z, Hemmateenejad B, Khayamian T. Wavelet neural network-based QSPR for prediction of critical micelle concentration of Gemini surfactants. *Anal Chim Acta*. 2005;531(2):285-91.
- [37] Belhaj AF, Elraies KA, Alnarabiji MS, Abdul Kareem FA, Shuhli JA, Mahmood SM, et al. Experimental investigation, binary modelling and artificial neural network prediction of surfactant adsorption for enhanced oil recovery application. *Chem Eng J*. 2021;406:127081.
- [38] Khazaei A, Parhizgar H, Dehghani MR. The Prediction of Surface Tension of Ternary Mixtures at Different Temperatures Using Artificial Neural Networks. *Iran J Oil Gas Sci Technol*. 2014;3(3):47-61.
- [39] Gharagheizi F, Eslamimanesh A, Mohammadi AH, Richon D. Use of artificial neural network-group contribution method to determine surface tension of pure compounds. *J Chem Eng Data*. 2011;56(5):2587-601.
- [40] Bakeri G, Delavar M, Lashkenari MS. Surface Tension Prediction of Hydrocarbon Mixtures Using Artificial Neural Network. *J Oil, Gas Petrochemical Technol*. 2015;2(1):14-26.
- [41] Belmadani S, Hanini S, Laidi M, Si-Moussa C, Hamadache M. Artificial Neural Network Models for Prediction of Density and Kinematic Viscosity of Different Systems of Biofuels and Their Blends with Diesel Fuel. Comparative Analysis. *Kem u Ind*. 2020;69(7-8):355-64.
- [42] Rivera RR, Soriano A. Prediction of Density and Speed of Sound of Binary Ionic Liquid and Ketone Mixtures Using Artificial Neural Network. *E3S Web Conf*. 2019;120:01003.
- [43] Safavi A, Sedaghatpour F, Shahbaazi HR. Tensammetric analysis of nonionic surfactant mixtures by artificial neural network. *Electroanalysis*. 2005;17(12):1112-8.
- [44] Jha BK, Tambe SS, Kulkarni BD. Estimating Diffusion Coefficients of a Micellar System Using an Artificial Neural Network. *J Colloid Interface Sci*. 1995;170(2):392-8.
- [45] Samson S, Basri M, Fard Masoumi HR, Abdul Malek E, Abedi Karjiban R. An artificial neural network based analysis of factors controlling particle size in a virgin coconut oil-based nanoemulsion system containing copper peptide. *PLoS One*. 2016;11(7):e0157737.
- [46] Rocabrúno-Valdés CI, Ramírez-Verduzco LF, Hernández JA. Artificial neural network models to predict density, dynamic viscosity,

and cetane number of biodiesel. Fuel.
2015;147:9-17.

[47] Asfaram A, Ghaedi M, Goudarzi A,
Rajabi M. Response surface
methodology approach for optimization
of simultaneous dye and metal ion
ultrasound-assisted adsorption onto Mn
doped Fe₃O₄-NPs loaded on AC: Kinetic
and isothermal studies. Dalt Trans.
2015;44(33):14707-23.