Chapter

Application of Artificial Neural Networks to Chemical and Process Engineering

Fabio Machado Cavalcanti, Camila Emilia Kozonoe, Kelvin André Pacheco and Rita Maria de Brito Alves

Abstract

The accelerated use of Artificial Neural Networks (ANNs) in Chemical and Process Engineering has drawn the attention of scientific and industrial communities, mainly due to the Big Data boom related to the analysis and interpretation of large data volumes required by Industry 4.0. ANNs are well-known nonlinear regression algorithms in the Machine Learning field for classification and prediction and are based on the human brain behavior, which learns tasks from experience through interconnected neurons. This empirical method can widely replace traditional complex phenomenological models based on nonlinear conservation equations, leading to a smaller computational effort – a very peculiar feature for its use in process optimization and control. Thereby, this chapter aims to exhibit several ANN modeling applications to different Chemical and Process Engineering areas, such as thermodynamics, kinetics and catalysis, process analysis and optimization, process safety and control, among others. This review study shows the increasing use of ANNs in the area, helping to understand and to explore process data aspects for future research.

Keywords: chemical and process engineering, Artificial Neural Networks, Big Data, Modeling

1. Introduction

Many chemical and process engineers are excited about the applications of Artificial Intelligence (AI) to their fields of expertise. AI can be defined as the ability of digital-computers to perform tasks at which people are better, at the moment [1]. In this context, Machine Learning (ML) is seen as one of the most relevant subareas, providing computers with the ability to learn without being programmed explicitly. It is essentially a form of applying statistics to estimate complex functions with less emphasis on obtaining the confidence intervals around them [2].

This current excitement was also stimulated by the Big Data boom related to the analysis and interpretation of large data volumes (of the order of several terabytes), which are generated at high rates and present various formats (numbers, text, multimedia, among others). Industry 4.0 requires this piece of knowledge from chemical and process engineers since process plants have large volumes of stored

historical data, obtained through sensors that measure thousands of variables in the order of seconds [3]. The analysis and exploitation of these data is a critical component for the operation of an industrial process.

In this framework, the so-called Artificial Neural Networks (ANNs) have numerous advantages and applications. They are universal nonlinear approximators based on the human brain-behavior through interconnected neurons that learn tasks from experience; in that case, from data [4]. Similarly to the nervous system, artificial neural networks are organized in the form of several simple individual elements – nodes or neurons – which interconnect with each other, forming networks capable of storing and transmitting information from/to the outside. Another relevant capacity of ANNs is their plasticity, which, through a learning process, allows changing in the interconnection pattern of its elements [5]. ANNs have been widely used in modeling or regression (linear and nonlinear) for one or several independent variables. It is worth noting that their use is not new in Chemical and Process Engineering, dating from the 1980s with some progress along the way, decisively contributing to the resurgence of the interest of the scientific community in this subject [3].

Modeling, simulation, and optimization are essential activities and competitive differentials among researchers to meet the challenges produced by environmental and commercial restrictions. In this context, ANNs are mostly used in process prediction and classification, as they are a robust nonlinear regression. In particular, this technique should be used when the solution of a problem is hampered by some of the following points: lack of physical or statistical understanding of the problem, statistical variations of the observable data, and the nonlinear mechanism responsible for the generation of the data [6].

In general terms, the use of neural networks consists of the following steps: 1- establishing the network architecture; 2- providing experimental data; 3- adjusting the network parameters – also known as their weights – until they learn the phenomenon (step called training); and 4- using the trained network with new input data for predicting the corresponding output data.

ANNs have been successfully applied to chemistry to correlate spectra of analytical methods and product properties [7]; in catalysis, to determine the relationships between the catalyst structure and its activity [8]; in process modeling, to predict product performance and operating conditions [9], and particularly in process control and fault diagnosis [10]. The main reasons for the growing popularity of the neural network approach are its lower computational cost compared to other methods and its ability to solve complex nonlinear problems [5]. Therefore, this review study demonstrates the increasing use of ANNs in Chemical and Process Engineering, helping to understand and to explore process data aspects for future research.

2. Most common activation functions used in chemical and process engineering applications

A neural network contains hyperparameters to be tuned prior to training in order to achieve the best configuration. Among them, the following can be mentioned: (i) number of hidden neurons, (ii) activation function, (iii) optimizer, and (iv) regularization and their dependencies (learning rate, optimizer specific, dropout rate, etc.).

Particularly, activation functions determine the output of the model, its accuracy, and the computational efficiency of training a model; therefore, they are an essential part of the structure of the neural networks. The Sigmoid function, Hyperbolic Tangent (TanH), and ReLU (Rectified Linear Unit) are the most

common in Chemical Engineering; however, recent studies improve these classical activation functions, defining new ones, such as Leaky ReLU, Swish, H-Swish [11].

In the sigmoid activation function, the output values are bounded between 0 and 1, normalizing each neuron output. However, there is a problem with the vanishing gradient, and outputs are not zero-centered. To make the modeling easier, the TanH was proposed, for which the outputs are zero-centered, i.e., when the inputs contain strongly negative, neutral, and strongly positive values.

In order to circumvent the computational expense, the ReLU was proposed. It is a computationally efficient linear activation function that will output the input directly if it is positive; otherwise, it will output zero. A further development is the Leaky ReLU, whereby the slope is changed to the left of $x = 0$, avoiding the dying ReLU problem, whereby some neurons can die for all inputs and remain inactive.

Therefore, the correct definition of the activation function is a fundamental part of the hyperparameter tuning to guarantee the best configuration of a neural network. In the course of the chapter, we will always mention which activation function each work used in the summary tables.

3. Applications to chemical and process engineering

In recent decades, there have been a large number of studies using ANNs in chemical engineering, from molecular property prediction [12], fault diagnosis [13], predictive control [14], and optimization [15, 16]. The use of first-principles knowledge must be integrated with the neural network in order to retain a more physical understanding of the system [14]. In the following subsections, we presented the principal papers of each area, with tables summarizing the characteristics of the ANNs used.

3.1 Thermodynamics and transport phenomena

Several data-driven models have been employed to predict phase equilibrium and transport phenomena coefficients for various chemical systems [17]. Indeed, these fields already have some empiricism in their standard mathematical formulations. For example, flash algorithms have some empiricism when using binary interaction parameters in subjective mixing rules [18], and the majority of transport phenomena coefficients are estimated from empirical correlations, sometimes questionable [19]. Therefore, the use of ANNs is a better way to find functional relationships between the model variables instead of first determining these constants [20].

Moreover, ANNs reveal a conceivably faster choice to those property prediction calculations in process simulations, limiting process control applications that require to be conducted in real-time. For this, Poort et al. [21] studied the replacement of conventional Equations of State (EoS) for property and phase stability calculations on a binary mixture of methanol–water. They trained ANNs with data generated through the Thermodynamics for Engineering Applications (TEA) to represent four kinds of flash algorithms, leading to an enhancement of 15 times for the predictions of properties and 35 times for classification of the phases.

Also noteworthy is that ANNs have also been used to predict if a particular mixture forms an azeotrope – essential information to design and to control a separation process. Alves et al. [22] successfully developed an ANN classification model to determine whether binary mixtures can exhibit (or not) azeotropy based solely on the properties of pure components as input variables. Therefore, it shows the power of ANNs for this type of thermodynamic evaluation since it does not take into account the non-ideality of the mixture.

They are also widely employed to predict thermal-physical properties of ionic liquids, such as density and viscosity [23]. The primary source of these values comes from experiments at the laboratory since ionic liquids do not present a universal description of their phase behavior. For example, using the definition of group contribution and the operating temperature, Valderrama et al. [24] successfully developed a three-layer FF-ANN to estimate the density of ionic liquids.

ANNs have also been employed in statistical thermodynamics techniques, which compute physicochemical properties from molecular simulations. One of these methods – the High-Throughput Force Field Simulation (HT-FFS) – can generate large volumes of data. ANNs can be trained with these data, thus building a gray-box model to improve the property predictions with a lower computational effort [25]. They have also been used in Density Functional Theory (DFT) calculations to replace some physical functionals with data-driven ones, finding the energy levels for electronic structures of different compounds with a balance between computational cost and accuracy [26].

Regarding their application to Transport Phenomena, it is well-known that ANNs – as an excellent universal approximator for any nonlinear function [27] – can be used for estimating convective heat- and mass-transfer coefficients [17]. Mainly in situations in which there is no mathematical correlation that can adjust them, as is the case of bubble columns. For this, Verma and Srivastava [19] successfully built an ANN model from literature data with eight inputs related to the system configuration of a bubble column (gas velocity, Prandtl number, number of holes, hole diameter, column diameter, surface tension, gas holdup, and bed height) and one output (heat coefficient).

Table 1 displays a summary of the current applications of neural networks to thermodynamics and transport phenomena discussed above. In the table, we specify the field, case study, class of neural network, activation function, topology and software used in each work.

3.2 Kinetics and catalysis

Neural networks have been successfully applied to catalysis to determine the relationship between the catalyst structure and its activity [8]. As heterogeneous catalysis has developed increasingly efficient experimentation techniques, the number of new data have increased exponentially [28], both from synthesis and from characterization and catalytic tests [29]. Thus, there is a need for more adequate tools to manage these large amounts of experimental data, to understand and to model it, and to generate a way to optimize the catalytic performance [30].

Two types of ANNs applications have been described so far in the frame of combinatorial catalysis: (i) ANN catalyst compositional models, correlating composition and synthesis variables with catalytic performance, and (ii) ANN kinetic models, correlating reaction conditions with the catalytic performance [31]. For example, those applications include the design of ammoxidation of propylene catalyst [32], design of methane oxidative decoupling catalyst [33], analysis and prediction of results of the decomposition of NO over zeolites [34], among other studies. Also, ANNs have been used combined with genetic algorithms for designing propane ammoxidation catalysts [35]. Another work successfully reported the viability of ANNs in the analysis and prediction of catalytic results within a collection of catalysts produced by combinatorial techniques [36]. Recently, an ANN was applied to estimate the rate of dehydration reaction of methanol in dimethyl ether synthesis [37]. The results showed that an ANN is a powerful tool for evaluating the reaction rate instead of using sophisticated kinetic model equations.

References	Field	Case Study	Class of Neural Network	Activation Function	Topology***	Software
[18]	Phase Equilibrium	Vapor–Liquid equilibrium of NH_3/H_2O and CH_4/C_2H_6 systems	FF-ANN*	Sigmoid	2–13-2	in-house software
[20]	Transport Phenomena	Determination of reduced boiling point from molecular weight and acentric factor	FF-ANN	Sigmoid	2–2–2-1	Matlab
[21]	Phase Equilibrium	Vapor–liquid flash calculations	FF-ANN	Linear/ Sigmoid	3–10-2	Keras-Python
[22]	Phase Equilibrium	Prediction of azeotrope formation	FF-ANN	Sigmoid	16–6-1	in-house software
[24]	Ionic Liquids	Estimation of physical properties of ionic liquids	FF-ANN	Tanh	10–15–15-1	Matlab
[25]	Molecular Thermodynamics	Enhancing the High-Throughput Force Field Simulation (HT-FFS)	FF-ANN	Linear/ ELU**	25–16–8-4-3	PyTorch
[26]	Molecular Thermodynamics	Correlation functionals of the electronic density	Fully connected neural networks	Sigmoid	4–8 neurons in each hidden layer	TensorFlow

**FN-ANN stands for Feed-Forward Artificial Neural Network.*
***ELU stands for Exponential Linear Unit.*
****The first and last elements in topology represent the number of neurons in the input and in the output layer, respectively. Among them, the number of neurons in the hidden layer(s).*

Table 1.
Current applications of ANNs to thermodynamics and transport phenomena.

The number of publications in this catalysis field has had an upward trend, especially in the last decade with the high demand for practical applications of the concepts of Big Data. The group of Turkish researchers led by Günay and Yildirim has excelled with work in the field, using not only ANNs for extracting knowledge from catalytic data, but also decision tree algorithms to determine the heuristic conditions and rules that lead to a high performance of the catalyst. For example, in work about carbon monoxide oxidation over Cu-based catalysts, they successfully used 1337 data points from 20 studies for evaluating catalyst performance using ANNs [38].

In the field of heterogeneous catalysis, ANNs can be used to select better possible catalysts – cheaper, less toxic, and composed of non-precious metals – for a given reaction, thus reducing the massive number of needed high-throughput experiments, peculiar conjuncture of combinatorial catalysis [39]. In this direction, Cavalcanti et al. [40] used a three-layer feedforward neural network to predict the

ideal composition of the catalyst in the water-gas-shift reaction and discover useful trends through sensitivity analysis. The input variables for ANN were several, while the only output variable considered was the conversion of CO. The model for the reaction was successfully developed, exhibiting the power of ANNs for predicting better catalysts and operating conditions for the process.

Recently, Cavalcanti et al. [8] showed that ANNs are able to predict the variables that most influence the conversion of CO in the water-gas-shift reaction, that is, temperature and surface area. The results can be used to conduct subsequent research in an optimized manner in this area, as it aims at the well-managed use of environmental resources, in the sense of selecting efficient catalysts for producing hydrogen - a clean energy source.

In the same topic, Garona et al. [41] presented an empiric model for the Fischer-Tropsch Synthesis (FTS) reaction using ANNs. A database of FTS to light olefins was assembled from the literature, and feedforward neural networks were used to build more complete models, which helped to predict optimal catalyst composition and operating conditions.

It is also noteworthy that ANNs were also used to model the sintering of a catalyst in a dry reformer [42]. In particular, the effects of temperature, pressure, and catalyst diameter on methane and CO_2 conversions, H_2/CO ratio, and molar percentage of solid carbon deposited on the catalyst (responsible for deactivation) have been studied. The ANN design activity was automated using a Genetic Algorithm (GA) search over the set of possible network topologies. The inclusion of the effective number of parameters in the GA objective function led to networks that performed well over testing data points.

Another application is in the determination of acidity in zeolites with data from FTIR spectroscopy [43]. FF-ANNs were used for analyzing multivariate base on the characteristic absorbance of 11 zeolite samples after metal substitution (Zn, Cu, Ga, and Ag) in the ~3612 cm^{-1} region. The developed regression method presented the same results of acid sites from other conventional and expensive methodologies.

Thus, in order to formulate a new kind of catalyst, it is essential to identify the catalysis past [44]. Therefore, by using ANNs, it is possible to convert historical data from past publications into valuable information, leading to a great acceleration in the development of new catalysts with better performances for a given process [8]. **Table 2** presents a summary of the current applications of neural networks to catalytic processes.

3.3 Process analysis and optimization

The applications of neural networks to the process analysis are increasing. Assidjo et al. [45] modeled the drying process of the production of coconut using a neural network. The goal is to predict the moisture of dried grated coconut whose dynamics are not well known. The authors used a feedforward fully connected neural network, whereby the selected architecture was 9–4-1, selected based on the minimum error in the test set. The results indicate that the neural network proposed, constructed using industrial plant data, can be used as a predicting method.

Fernandes and Lona [46] applied neural networks to the field of polymerization. The authors also highlighted some topologies, the number of data points needed, and the concept of stacked neural networks that can enhance the prediction of the final model.

Alves and Nascimento [47] used industrial plant data for constructing neural networks to detect gross errors; the case study was an isoprene unit facility.

Alves and Nascimento [4] studied the production of high purity isoprene from a C_5 cut arising from a pyrolysis gasoline unit. The first principle models were replaced by neural networks in the final grid search of the optimal parameters for

the process. The set of 10 neural networks were defined to represent the whole flowsheeting, whereby the number of hidden layers was defined by the minimum error in the test set. Lastly, the framework successfully optimized a chemical plant under study using neural networks with industrial data.

References	Field	Case Study	Class of Neural Network	Activation Function	Topology**	Software
[29]	Modeling of catalytic processes	Catalytic activity for n-paraffin isomerization	FF-ANN*	Sigmoid/ Tanh	4–8–6-3	SNNS neural networks simulator
[32]	Catalyst design	Design of catalyst for propane ammoxidation	FF-ANN	Sigmoid	6–20–12–2	in-house software
[33]	Catalyst design	Design of a catalyst for methane oxidative coupling	FF-ANN	Sigmoid	6–20–9-2	in-house software
[34]	Modeling of catalytic processes	Analysis of NO decomposition over Cu/ZSM-5 zeolite	FF-ANN	Sigmoid	4–32-1	in-house software
[36]	Combinatorial catalysis	Modeling of catalysts for oxidative dehydrogenation of ethane	FF-ANN	Not described	13–26–12-6	SNNS neural networks simulator
[37]	Modeling of catalytic processes	Estimation of the reaction rate in methanol dehydration	FF-ANN	Tanh/ Linear	3–6-1	Matlab
[38]	Modeling of catalytic processes	Selective CO Oxidation over Copper-Based Catalysts	FF-ANN	Tanh	14–7–7-1	Matlab
[8]	Catalyst selection	Catalyst selection for the WGS reaction	FF-ANN	Sigmoid	51–12-1	R - neuralnet
[41]	Modeling of catalytic processes	Fischer-Tropsch synthesis to lower-olefins	FF-ANN	Sigmoid	30–15-2	R - neuralnet
[42]	Catalyst deactivation	Dry reformer under catalyst sintering	FF-ANN	Tanh	3–12–5-6-1	in-house software
[43]	Determination of catalyst acidity	Determination of acidity in metal incorporated zeolites by FTRI	FF-ANN	Tanh	6–10-1	Matlab

**FN-ANN stands for Feed-Forward Artificial Neural Network.*
***The first and last elements in topology represent the number of neurons in the input and in the output layer, respectively. Among them, the number of neurons in the hidden layer(s).*

Table 2.
Current applications of ANNs to catalytic processes.

Khezri et al. [15] proposed a hybrid model for optimizing a large-scale gas to liquids process. The dataset was constructed using a simulation model of the GTL process. Different topologies were compared to select the most promising one; one and two hidden layers with different number of neurons were tested. The optimal configuration was two hidden layers with 7 and 15 hidden neurons each. The ANN was modeled using the information of the tail gas unpurged ratio, recycled tail gas to FT ratio, H_2O/C in the syngas section, and CO_2 removal percentage as input features; the outputting was wax production rate. The ANN model was then used for optimization purposes.

Wang et al. [16] proposed a framework for predicting the operating trend of an industrial process. The framework contains three major steps: (i) multivariate correlation analysis, to deal with the correlation between the historical industrial data, (ii) clustering, due to nonlinear dense data and unclear operating trend types and (iii) a convolutional neural network (CNN), formed by five parts (input layer, convolutional layer, ReLU layer, pooling layer, and fully connected layer).

The authors pointed out the importance of the convolutional networks to extract important features from the dataset. Moreover, the advantage of such a framework was compared with traditional convolutional neural networks and recurrent neural networks (RNNs) for a methanol production process.

Cai et al. [48] analyzed an industrial process using data-driven models. The case study was the industrial reverse osmosis concentrate (ROC) treatment with the fluidized bed reactor Fenton (FBR-Fenton) process. Prior to modeling, a statistical analysis was carried out to determine the most relevant features as input (Fe^{2+} dosage, H_2O_2 dosage, pH, and HRT). Two approaches were studied, ANN and linear regression. The former showed more accurate predictions, consisting in one input layer (4 neurons), 4 hidden layers (10 neurons each) and one output layer (2 neurons) using ReLU as an activation function, due to the least computationally dense mechanism and also a general approximation for most scenarios [11].

The crystallization process and the quality of the products was studied by Lin et al. [49]. The authors used a Raman spectrum as input for a two-layer back propagation neural network with four hidden neurons to predict the solution concentration and slurry density simultaneously. They also compared the output prediction of the neural network with other algorithm predictions (characteristic peaks regression, principal component regression, partial least-squares regression), and the results indicated the superior prediction characteristics of the neural network due to its inner nonlinear nature.

Chemical process synthesis is a complex scheme, which comprises process modeling and design, and combinatorial defiance. There are two major approaches: the traditional sequential form and the optimization-based synthesis using superstructure models. In the former category, the problem is solved in sequential scheme, by decomposition whereby there is a hierarchy of elements that can be depicted by an Onion Diagram (reactor, separation, heat recovery and utility) [50].

The latter category considers the full integration between decisions at the single step, i.e. determine the optimal structure and operating conditions simultaneously. Therefore, this approach contemplates all possible complex interactions between the engineering choices, including equipment (potentially selected in the optimized flowsheet), the interconnection and operating conditions formulated as an optimization problem [51–53].

There is a diversity of proposed methodologies to represent a general process superstructure [54–56]. However, due to the inner complexity of the superstructure (**Figure 1**), the large-scale non-convex Mixed-Integer Nonlinear Programs (MINLP) require effective approaches to solve them.

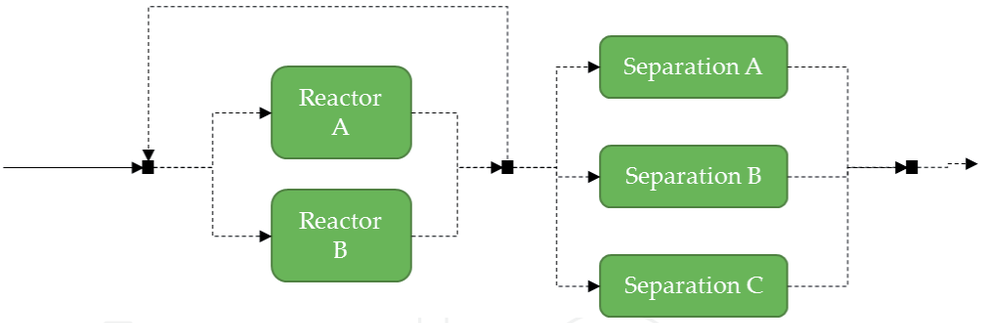

Figure 1.
Simple superstructure representation compared with different separation processes.

The use of simplified models or surrogates at the unit operation level is advantageous because they are present in any process simulator. Additionally, surrogates can be used to represent an entire subsystem consisting of a definite number of units. Artificial Neural Networks (ANNs) may be used to generate the surrogate models, due to their fitting characteristics [57].

In order to circumvent the solution problem of a superstructure, Henao and Maravelias [58] proposed a framework to replace complex unit models (based on first-principle) with surrogate models, developed using artificial neural networks. The authors proposed simpler surrogate models for pumps, compressors and flash vessels. The authors used two case studies (Absorption-based CO_2 capture system and maleic anhydride process superstructure) to validate the proposed framework. The results indicate the possibility of using neural networks embedded in a rigorous optimization procedure.

Savage et al. [59] proposed a hybrid machine learning-based framework to optimize the chemical process (the CryoMan Cascade cycle system was used as a case study). The authors compared different surrogate models algorithms (ANN and Kriging Partial Least Squares); the results indicated a reduction in the time needed for the optimization when compared with the rigorous model. Moreover, they found that a single large ANN was unable to capture the high nonlinearity of the process under study based on the final accuracy. Therefore, the authors broke the surrogate model into a series of parallel sub-models, revealing to have increased the final accuracy.

According to Klemes et al. [60], despite the substantial level of maturity of the process modeling, the nature of connections of the problem still allows improvements. Nascimento et al. [61] also presented alternatives for the optimization of industrial facilities using neural networks and compared them with industrial data.

Table 3 presents a summary of the current applications of neural networks to process analysis and optimization.

3.4 Process safety and control

One of the most common applications of ANNs to the area of process safety and control is in fault detection and diagnosis. These systems are built to identify habitual process behavior and recognize atypical variations in the chemical plant that can lead to an accident [64]. Generally, deep neural networks – ANNs that contain several hidden layers – are used to extract spatial and temporal aspects of the data for this purpose [65]. Their inputs are the sensors responsible for the variable measurement, and their outputs of the kind of faults (e.g., tube plugging, valve blockage, catalyst deactivation, among others) [66].

References	Field	Case Study	Class of Neural Network	Activation Function	Topology****	Software
[45]	Process Analysis	Grated coconut industry	FF-ANN*	Tanh	9–4-1	Matlab
[16]	Industrial Process Operating (Predictive Control)	Methanol production	CNN**	ReLu	5 convolution layers, 36 filters, and 3 pooling layers	Caffe
[62]	Process Analysis	Fluidized bed reactor Fenton process	FF-ANN	ReLu	4–10– 10 – 10 – 10 – 2	R - Keras
[14]	Predictive Control	non-isothermal continuous stirred tank reactors	RNN***	Tanh	2 hidden layers with 30 neurons in each layer	Python-Keras
[15]	Process Optimization	Large scale gas to liquids process	FF-ANN	Sigmoid	4–7–15-1	Matlab
[4]	Process Optimization	Isoprene Process	FF-ANN	Sigmoid	10 neural networks (all with one hidden layer)	in-house software
[58]	Process Synthesis	Absorption-based CO_2 capture and Maleic Anhydride process	FF-ANN	Tanh	Several neural networks (all with one hidden layer)	Matlab
[59]	Process Synthesis	CryoMan Cascade cycle system	FF-ANN	Not Described	Not Described	Python-PyTorch
[63]	Process Analysis	Thermo-catalytic methane decomposition	FF-ANN	Sigmoid	6–9-1	Matlab
[49]	Process Analysis	Crystallization process	FF-ANN	Not Described	two-layer neural network with four hidden neurons	Matlab

*FN-ANN stands for Feed-Forward Artificial Neural Network.
**CNN stands for Convolutional Neural Network.
***RNN stands for Recurrent Neural Network.
****The first and last elements in topology represent the number of neurons in the input and in the output layer, respectively. Among them, the number of neurons in the hidden layer(s).

Table 3.
Current applications of neural networks to process analysis and optimization.

However, determining the various hyperparameters of deep neural networks demands a considerable amount of time, which is not suitable for fast online process applications. Based on this, Peng et al. [67] applied a method to reduce the training time of these complex types of network architecture: the Broad Learning

System (BLS). It uses an incremental learning procedure and enlarges the network in width, making a quick training stage possible. They successfully employed this strategy in a batch fermentation process for fault detection utilizing the Affinity Propagation (AP) algorithm in a Long Short-Term Memory (LSTM) deep neural network to cluster distinct stage data.

Another use is in developing models to control the process quality through variables that do not have online sensors. On the one hand, variables such as pressure, temperature, and mass flow rate can be easily measured by manometers, thermocouples, and mass flow controllers, respectively. On the other hand, the online measurement of a variable such as pH in the process is a challenge since no large-scale equipment exists for this, depending on an offline laboratory analysis. Therefore, ANNs can be used to develop these so-called *soft-sensors* to predict quality parameters from a large volume of industrial data, improving the process control quality [68].

Finally, ANNs are also used to replace complex phenomenological models in Model Predictive Control (MPC) architectures and Real-Time Optimization (RTO) strategies [69]. Both applications depend on the model accuracy and the velocity of solving the model equations to drive the controlled variable to the desired set-point. The former is related to dynamic processes and the latter to steady-state operations [69]. Since ANNs have a lower computational response than first-principle models, they are a suitable alternative to make these control strategies possible and efficient.

A successful application of this kind of substitution can be found elsewhere [70], in which an ANN is used to replace a very detailed computational fluid dynamic (CFD) model that represents the synthesis of phthalic anhydride in a fixed-bed catalytic reactor for an MPC structure. Moreover, a hybrid model approach (first-principles combined with ANN) was employed in an MPC by Zhang et al. [69] to drive a reaction process in a continuous stirred tank reactor (CSTR) to optimal operating conditions. They represented the reaction rates by neural networks instead of using the nonlinear Arrhenius Law to describe the reaction phenomenon. Indeed, this well-known equation was used to generate the dataset for training the network under numerous variations in temperature and reactant concentrations. The MPC acted to stabilize the chemical process, driving it to the lowest total cost conditions.

Wu et al. [14] proposed a hybrid machine-learning model that incorporates first principles into a recurrent neural network. The authors studied two models, a partially-connected RNN model and a weight-constrained RNN model and applied them to a chemical process containing two well-mixed, non- isothermal continuous stirred tank reactors in series. The two proposed models outperformed a Lyapunov-based model predictive controller based on prediction accuracy, smoother state trajectories and economic advantages.

It is worth mentioning that ANNs are being used to build detectors to prevent cyber-attacks against process plants [71]. Nowadays, with highly automated systems for controlling chemical plants with real-time operation, breaches in cyber-secure failures can exist, which may cause accidents and economic losses. With this in mind, Chen et al. [71] developed a feedback-MPC control architecture with an ANN-detector that can identify the probabilities of cyber-attacks in networked sensors. Therefore, the applicability of ANNs in these safety and control strategies is very significant for the integrability of industrial plants.

Table 4 shows a summary of the current applications of neural networks to the area of process safety and control.

References	Field	Case Study	Class of Neural Network	Activation Function	Topology****	Software
[67]	Fault Detection	Penicillin fermentation process	LSTM***	Sigmoid	10–20–15-2	Matlab
[68]	Soft Sensors	pH control in a chemical process	RNN**	Tanh	5–14–1	Not described
[69]	Surrogate model in MPC and RTO	Reaction process in a CSTR	FF-ANN*	Tanh	3–10-1	Matlab
[70]	Surrogate model from CFD in MPC	Phthalic anhydride synthesis in a fixed-bed catalytic reactor	RNN	ReLu	3–64–64–1	Keras
[14]	Hybrid model in a MPC	Two-consecutive CSTRs	RNN	Tanh	2–30–30-4	IPOPT-Python
[71]	Cyber Security	MPC integrated with cyber-secure feedback controller	FF-ANN	Tanh	4–12–10-9	Matlab

*FN-ANN stands for Feed-Forward Artificial Neural Network.
**RNN stands for Recurrent Neural Network.
***LSTM stands for Long Short-Term Memory.
****The first and last elements in topology represent the number of neurons in the input and in the output layer, respectively. Among them, the number of neurons in the hidden layer(s).

Table 4.
Current applications of ANNs to process safety and control.

4. Future works

Today, ANNs are one of the most found subjects in the scientific literature of Chemical and Process Engineering; and their use tends to continue growing. This can be explained by the launch of Industry 4.0, in which these data-driven models play an essential role in the implementation of some type of intelligent systems in processes [72]. Thus, to remain relevant in this current scenario, companies need specialized professionals on their team. For this reason, this topic has been introduced into the curriculum of most Chemical Engineering degree programs [73].

Indeed, the continuous availability of large volumes of stored data in industrial processes will lead to the development of new ANN approaches for process modeling and data interpretation. These models will deliver more direct relationships between cause and effect variables for process optimization and control through MPC strategies. Therefore, the automation of entire plant units will conduct to intelligent processes, capable of making decisions for safer operation, and with a reliable protection system against cyber-attacks.

Another subarea worth mentioning for future developments is the design of new materials. The use of ANNs has led to a decrease in the number of lengthy and

costly laboratory experiments for analyzing the performance of polymers, ceramics, glasses, and mainly, catalysts. Therefore, it is possible to convert data from past publications and from high-throughput (HT) experiments into information, leading to a surprising acceleration in developing new materials with better performances for a given process.

5. Conclusions and perspectives

This chapter presented the ANNs and their Chemical and Process Engineering applications, showing how they have become a powerful tool for modeling chemical processes. This analysis also showed their increasing application, helping to understand and analyze process data features for future research in thermodynamics, transport phenomena, kinetics and catalysis, process analysis and optimization, and process safety and control.

The prospective availability of large volumes of data with good quality will make ANNs one of the most used methods to represent a process, estimate thermodynamic properties, develop new catalysts, replace complex phenomenological models, and improve control and safety strategies. Moreover, in real chemical processes, a particular part of the inputs affect only a section of the outputs. Therefore, the knowledge of first principles embedded in a data driven machine learning model is a challenge for the next studies.

Acknowledgements

We gratefully acknowledge the support of the RCGI – Research Centre for Gas Innovation, hosted by the Universidade de São Paulo (USP) and sponsored by FAPESP – The São Paulo Research Foundation (2014/50279-4) and Shell Brasil. In addition, the authors acknowledge the financial support provided by FAPESP for doctoral scholarships (Grant 2017/11940-5 and 2017/26683-8).

Conflict of interest

The authors declare no conflict of interest.

Author details

Fabio Machado Cavalcanti, Camila Emilia Kozonoe, Kelvin André Pacheco and Rita Maria de Brito Alves*
Escola Politécnica - Universidade de São Paulo, São Paulo, Brazil

*Address all correspondence to: rmbalves@usp.br

References

[1] Rich E. Artificial intelligence. 1st ed. McGraw-Hill, Inc.; 1983.

[2] Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. Vol. 1. MIT press Cambridge; 2016.

[3] Venkatasubramanian V. The promise of artificial intelligence in chemical engineering: Is it here, finally? AIChE J. 2019;65(2):466-478.

[4] Alves RMB, Nascimento CAO. Neural network based approach applied to for modeling and optimization an industrial isoprene unit production. AIChE Annu Meet Conf Proc. 2004;7663-7682.

[5] Himmelblau DM. Accounts of experiences in the aPplication of artificial neural networks in chemical engineering. Ind Eng Chem Res. 2008;47(16):5782-5796.

[6] Haykin S. Redes Neurais - Principios e Praticas. 2nd ed. Bookman; 2001.

[7] Gemperline PJ, Long JR, Gregoriou VG. Nonlinear Multivariate Calibration Using Principal Components Regression and Artificial Neural Networks. Vol. 63, Bricout, J.; Fontes, J. C. Ann. Falslf. Expert. Chlm. 1991.

[8] Cavalcanti FM, Schmal M, Giudici R, Brito Alves RM. A catalyst selection method for hydrogen production through Water-Gas Shift Reaction using artificial neural networks. J Environ Manage. 2019 May; 237:585-594

[9] Nascimento CAO, Giudici R. Neural network based approach for optimisation applied to an industrial nylon-6,6 polymerisation process. Comput Chem Eng. 1998;22:595-600.

[10] Guo L, Kang J. A hybrid process monitoring and fault diagnosis approach for chemical plants. Int J Chem Eng. 2015;2015.

[11] Nwankpa C, Ijomah W, Gachagan A, Marshall S. Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. arXiv Prepr. 2018;arXiv:1811.

[12] Hirschfeld L, Swanson K, Yang K, Barzilay R, Coley CW. Uncertainty Quantification Using Neural Networks for Molecular Property Prediction. J Chem Inf Model. 2020 Aug;60(8):3770-3780.

[13] Zhang S, Bi K, Qiu T. Bidirectional Recurrent Neural Network-Based Chemical Process Fault Diagnosis. Ind Eng Chem Res. 2020 Jan;59(2):824-834.

[14] Wu Z, Rincon D, Christofides PD. Process structure-based recurrent neural network modeling for model predictive control of nonlinear processes. J Process Control. 2020;89:74-84.

[15] Khezri V, Yasari E, Panahi M, Khosravi A. Hybrid Artificial Neural Network–Genetic Algorithm-Based Technique to Optimize a Steady-State Gas-to-Liquids Plant. Ind Eng Chem Res. 2020 May;59(18):8674-8687.

[16] Wang Y, Ren YM, Li H. Symbolic Multivariable Hierarchical Clustering Based Convolutional Neural Networks with Applications in Industrial Process Operating Trend Predictions. Ind Eng Chem Res. 2020 Aug;59(34):15133-15145.

[17] Verma AK. Process Modelling and Simulation in Chemical, Biochemical and Environmental Engineering. CRC Press; 2015.

[18] Vashishtha M. Application of artificial neural networks in prediction of vapour liquid equilibrium data.

Proc - 25th Eur Conf Model Simulation, ECMS 2011. 2011;6(Cd):142-5.

[19] Verma AK, Srivastava A. ANN based Model for Heat Transfer from Immersed Tubes in a Bubble Column: Effects of Immersed Surface and Sparger Geometry Conference on Thermal Systems. In: Proceedings of Fourth National Seminar on Thermal Systems. 2003. p. 135-9.

[20] Joss L, Müller EA. Machine Learning for Fluid Property Correlations: Classroom Examples with MATLAB. J Chem Educ. 2019;96(4):697-703.

[21] Poort JP, Ramdin M, van Kranendonk J, Vlugt TJH. Solving vapor-liquid flash problems using artificial neural networks. Fluid Phase Equilib. 2019;490:39-47.

[22] Brito Alves RM, Quina FH, Oller Nascimento CA. New approach for the prediction of azeotropy in binary systems. Comput Chem Eng. 2003;27(12):1755-1759.

[23] Yusuf F, Olayiwola T, Afagwu C. Application of Artificial Intelligence-based predictive methods in Ionic liquid studies: A review. Fluid Phase Equilib. 2021 Mar 1;531:112898.

[24] Valderrama JO, Reátegui A, Rojas RE. Density of ionic liquids using group contribution and artificial neural networks. Ind Eng Chem Res [Internet]. 2009 Mar 18 [cited 2021 Feb 9];48(6):3254-9. Available from: https://pubs.acs.org/sharingguidelines

[25] Gong Z, Wu Y, Wu L, Sun H. Predicting Thermodynamic Properties of Alkanes by High-Throughput Force Field Simulation and Machine Learning. J Chem Inf Model. 2018;58(12):2502-2516.

[26] Dick S, Fernandez-Serra M. Machine learning accurate exchange and correlation functionals of the electronic density. Nat Commun. 2020;11(1).

[27] Hornik K, Stinchcombe M, White H. Multilayer Feedforward Networks are Universal Approximators. Neural Networks. 1989;2(5):359-366.

[28] Erdem Günay M, Yıldırım R. Recent advances in knowledge discovery for heterogeneous catalysis using machine learning. Catal Rev - Sci Eng. 2020;

[29] Serra JM, Corma A, Chica A, Argente E, Botti V. Can artificial neural networks help the experimentation in catalysis? Catal Today. 2003;81(3):393-403.

[30] Senkan S. Combinatorial heterogeneous catalysis - A new path in an old field. Vol. 40, Angewandte Chemie - International Edition. John Wiley & Sons, Ltd; 2001. p. 312-29.

[31] Hattori T, Kito S. Neural network as a tool for catalyst development. Catal Today. 1995 Apr 7;23(4):347-355.

[32] Hou ZY, Dai Q, Wu XQ, Chen GT. Artificial neural network aided design of catalyst for propane ammoxidation. Appl Catal A Gen. 1997 Nov 4;161(1-2):183-190.

[33] Huang K, Chen FQ, Lü DW. Artificial neural network-aided design of a multi-component catalyst for methane oxidative coupling. Appl Catal A Gen. 2001 Oct 5;219(1-2):61-68.

[34] Sasaki M, Hamada H, Kintaichi Y, Ito T. Application of a neural network to the analysis of catalytic reactions Analysis of NO decomposition over Cu/ZSM-5 zeolite. Appl Catal A, Gen. 1995 Nov 23;132(2):261-270.

[35] Cundari TR, Deng J, Zhao Y. Design of a propane ammoxidation catalyst using artificial neural networks and genetic algorithms. In: Industrial and Engineering Chemistry Research.

American Chemical Society; 2001. p. 5475-5480.

[36] Corma A, Serra JM, Argente E, Botti V, Valero S. Application of artificial neural networks to combinatorial catalysis: Modeling and predicting ODHE catalysts. ChemPhysChem. 2002;3(11):939-945.

[37] Valeh-E-Sheyda P, Yaripour F, Moradi G, Saber M. Application of artificial neural networks for estimation of the reaction rate in methanol dehydration. Ind Eng Chem Res. 2010 May 19;49(10):4620-4626.

[38] Günay ME, Yildirim R. Neural network Analysis of Selective CO Oxidation over Copper-Based Catalysts for Knowledge Extraction from Published Data in the Literature. Ind Eng Chem Res. 2011;50(22):12488-12500.

[39] Baumes L, Farrusseng D, Lengliz M, Mirodatos C. Using artificial neural networks to boost high-throughput discovery in heterogeneous catalysis. QSAR Comb Sci. 2004;23(9):767-778.

[40] Cavalcanti FM, Schmal M, Giudici R, Brito Alves RM. A Catalyst Selection Method for the Water-Gas Shift Reaction using Artificial Neural Networks. In: 1st Latin American Conference on Sustainable Development of Energy, Water and Environment Systems - LA SDEWES. Rio de Janeiro; 2018. p. 1-11.

[41] Garona HA, Cavalcanti FM, Abreu TF, Schmal M, Brito Alves RM. Using Artificial Neural Networks for Fischer-Tropsch Synthesis to Lower-Olefins Production Optimization. In: 15th Conference on Sustainable Development of Energy, Water and Environment Systems - SDEWES. Cologne; 2020. p. 1-15.

[42] Azzam M, Aramouni NAK, Ahmad MN, Awad M, Kwapinski W, Zeaiter J. Dynamic optimization of dry reformer under catalyst sintering using neural networks. Energy Convers Manag. 2018 Feb 1;157:146-156.

[43] Juybar M, Khanmohammadi Khorrami M, Bagheri Garmarudi A, Zandbaaf S. Determination of acidity in metal incorporated zeolites by infrared spectrometry using artificial neural network as chemometric approach. Spectrochim Acta - Part A Mol Biomol Spectrosc. 2020 Mar 5;228:117539.

[44] Schmal M. Heterogeneous Catalysis and its Industrial Applications. 1st ed. Switzerland: Springer; 2016.

[45] Assidjo E, Yao B, Kisselmina K, Amané D. Modeling of an industrial drying process by artificial neural networks. Brazilian J Chem Eng. 2008;25(3):515-522.

[46] Fernandes FAN, Lona LMF. Neural network applications in polymerization processes. Brazilian J Chem Eng. 2005;22(3):401-418.

[47] Alves RMB, Nascimento CAO. Gross errors detection of industrial data by neural network and cluster techniques. Brazilian J Chem Eng. 2002;19(4):483-489.

[48] Cai Q, Lee BCY, Ong SL, Hu J. Application of a Multiobjective Artificial Neural Network (ANN) in Industrial Reverse Osmosis Concentrate Treatment with a Fluidized Bed Fenton Process: Performance Prediction and Process Optimization. ACS ES&T Water. 2021;0-11.

[49] Lin M, Wu Y, Rohani S. Simultaneous Measurement of Solution Concentration and Slurry Density by Raman Spectroscopy with Artificial Neural Network. Cryst Growth Des [Internet]. 2020 Mar 4 [cited 2021 Feb 8];20(3):1752-9. Available from: https://dx.doi.org/10.1021/acs.cgd.9b01482

[50] Douglas JM. Conceptual Design of Chemical Processes. Vol. 1. New York: McGraw-Hill; 1988. 1110 p.

[51] Yeomans H, Grossmann IE. A systematic modeling framework of superstructure optimization in process synthesis. Comput Chem Eng. 1999;23(6):709-731.

[52] Graciano JEA, Le Roux GAC. Improvements in surrogate models for process synthesis. Application to water network system design. Comput Chem Eng. 2013;59:197-210.

[53] Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. Comput Chem Eng. 2020;136:106808.

[54] Biegler LT, Grossmann IE, Westerberg AW. Systematic Methods of Chemical Process Design. Upper Saddle River, NJ: Prentice Hall PTR; 1999. 808 p.

[55] Grossmann IE, Daichendt MM. New trends in optimization-based approaches to process synthesis. Comput Chem Eng. 2003;20(6-7):665-683.

[56] Ryu J, Kong L, Pastore de Lima AE, Maravelias CT. A generalized superstructure-based framework for process synthesis. Comput Chem Eng. 2020;133:106653.

[57] Haykin S. Neural networks: a comprehensive foundation. Prentice Hall PTR; 1994.

[58] Henao CA, Maravelias CT. Surrogate-based process synthesis. Vol. 28, Computer Aided Chemical Engineering. Elsevier B.V.; 2010. 1129-1134 p.

[59] Savage T, Almeida-Trasvina HF, del Río-Chanona EA, Smith R, Zhang D. An adaptive data-driven modelling and optimization framework for complex chemical process design. Comput Aided Chem Eng. 2020;48:73-78.

[60] Klemeš J, Friedler F, Bulatov I, Varbanov P. Sustainability in the Process Industry: Integration and Optimization. New York, Chicago, San Francisco, Lisbon, London, Madrid, Mexico City, Milan, New Delhi, San Juan, Seoul, Singapore, Sydney, Toronto: McGRAW-HILL; 2011. 385 p.

[61] Nascimento CAO, Giudici R, Guardani R. Neural network based approach for optimization of industrial chemical processes. Comput Chem Eng. 2000 Oct 1;24(9-10):2303-2314.

[62] Cai QQ, Lee BCY, Ong SL, Hu JY. Fluidized-bed Fenton technologies for recalcitrant industrial wastewater treatment–Recent advances, challenges and perspective. Vol. 190, Water Research. Elsevier Ltd; 2021. p. 116692.

[63] Alsaffar MA, Ghany MARA, Ali JM, Ayodele BV, Mustapa SI. Artificial Neural Network Modeling of Thermo-catalytic Methane Decomposition for Hydrogen Production. Top Catal [Internet]. 2021 Jan 2 [cited 2021 Feb 8];1:3. Available from: https://doi.org/10.1007/s11244-020-01409-6

[64] Md Nor N, Che Hassan CR, Hussain MA. A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. Rev Chem Eng. 2020;36(4):513-553.

[65] Luo L, Xie L, Su H. Deep Learning with Tensor Factorization Layers for Sequential Fault Diagnosis and Industrial Process Monitoring. IEEE Access. 2020;8:105494-105506.

[66] Gao X, Yang F, Feng E. A process fault diagnosis method using multi-time scale dynamic feature extraction based on convolutional neural network. Can J Chem Eng. 2020;98(6):1280-1292.

[67] Peng C, Lu RW, Kang O, Kai W. Batch process fault detection for multi-stage broad learning system. Neural Networks. 2020;129:298-312.

[68] Kamat S, Madhavan KP. Developing ANN based virtual/soft sensors for industrial problems. IFAC-PapersOnLine. 2016;49(1):100-105.

[69] Zhang Z, Wu Z, Rincon D, Christofides PD. Real-time optimization and control of nonlinear processes using machine learning. Mathematics. 2019;7(10):1-25.

[70] Wu Z, Tran A, Ren YM, Barnes CS, Chen S, Christofides PD. Model predictive control of phthalic anhydride synthesis in a fixed-bed catalytic reactor via machine learning modeling. Chem Eng Res Des. 2019;145:173-183.

[71] Chen S, Wu Z, Christofides PD. A cyber-secure control-detector architecture for nonlinear processes. AIChE J. 2020;66(5):1-18.

[72] Hernavs J, Ficko M, Berus L, Rudolf R, Klančnik S. Deep Learning in Industry 4 . 0 – Brief Overview. J Prod Eng. 2018;21(2):1-5.

[73] Duever TA. Data science in the chemical engineering curriculum. Processes. 2019;7(11).