scientific reports

Check for updates

OPEN

Predictive modeling of membrane reactor efficiency using advanced artificial neural networks for green hydrogen production

Mehrdad Mahmoudi, Ahad Ghaemi✉, Ahmad Rahbar Kelishami & Salman Movahedirad

The imperative to decarbonize the energy sector has prompted substantial advancements in clean electricity generation, with hydrogen emerging as a promising low-carbon energy carrier. While hydrogen synthesis from renewable sources is crucial, challenges persist, necessitating innovative approaches for efficient and sustainable production. This study leverages diverse artificial neural network (ANN) models to assess and predict system efficiency based on key operational variables in membrane reactor systems. The multilayered perceptron (MLP) and radial basis function (RBF) methodologies are employed, with the MLP models optimized across twelve training algorithms and eight activation functions, exploring up to three hidden layers with variable neuron counts. The MLP model, utilizing the Levenberg-Marquard training algorithm and Tangent-Sigmoid activation function, achieved a high correlation coefficient (R^2) of 0.9975 for training and 0.9962 for testing, and a mean squared error (MSE) of 0.00425 for training and 0.23951 for testing, indicating precise and accurate efficiency predictions. The Log-Sigmoid activation function also performed well, with R^2 values of 0.9971 (training) and 0.9961 (testing), and MSE values of 0.004086 (training) and 0.17694 (testing). Optimization of the RBF network identified the best performance with a spread parameter of 1 and 35 neurons, although the MLP model demonstrated superior accuracy and reduced computational time. Statistical analysis, encompassing correlation coefficient, mean squared error, Root Mean Squared error, absolute average deviation, absolute average relative deviation, and runtime, confirms the network's consistent and accurate estimation of system efficiency across various input variables. The study highlights that applying tansig and logsig activation functions, configured with neuron counts of 20, 17, 6 and 23, 20, 2 at the first, second and third hidden layers, respectively, offers enhanced accuracy and reliability. The MLP model's high performance underscores its potential to identify optimal conditions for H_2 generation based on system efficiency, thereby advancing membrane reactor technology for hydrogen production.

The pivotal role of the energy sector's decarbonization cannot be overstated in the endeavor to mitigate greenhouse gas (GHG) emissions. Substantial strides have been made in the recent past, particularly in the generation of clean electricity from low-carbon sources[1]. In this regard, hydrogen stands out as a highly promising energy carrier with the potential to replace fossil fuels, offering the prospect of energy generation with significantly reduced or near-zero greenhouse gas emissions. Unfortunately, in contrast to fossil fuels, hydrogen is not naturally abundant. Nevertheless, it can be synthesized from various primary energy sources and subsequently utilized as a fuel through direct combustion in an internal combustion engine. The extensive adoption of hydrogen depends on reducing production costs and incorporating renewable sources[2–6]. In the context of addressing these challenges, low-carbon gaseous carriers, prominently represented by biogas, biomethane, and low-carbon hydrogen, assume a pivotal role. These entities are at the forefront of spearheading a paradigm shift towards sustainable energy solutions. The prospect of widespread adoption of hydrogen-based systems, exemplified by fuel cells and industrial boilers, is poised to catalyze an increased demand for green hydrogen[7].

A predominant approach in the quest for low-carbon hydrogen production involves harnessing fossil fuels with utilization, storage, and carbon capture[8]. Another avenue for hydrogen production from renewable

School of Chemical, Petroleum and Gas Engineering, Iran University of Science and Technology, Tehran, Iran. ✉email: aghaemi@iust.ac.ir

sources is biomass utilization, predominantly through pyrolysis and gasification[8]. An economical and promising strategy that has emerged involves the steam reforming of raw biogas[6], where the composition predominantly includes methane and carbon dioxide. While the conversion process bears similarities to that of natural gas, a crucial adaptation is required to align with the typical production scale of existing biogas plants, which is notably smaller, approximately 100 times so[9].

Exploring the technical and economic feasibility, prior research has demonstrated the potential of this method for hydrogen production[10–12]. Broadly, the synthesis of hydrogen is achievable through biogas steam reforming (BSR) within a wide temperature range of 600 to 1000 °C, encompassing endothermic and reversible reactions. This intricate process involves catalytic reactions, often intricately combined, and can be executed at lower pressures (typically under atmospheric pressure) employing either tubular fixed beds or fluidized reactors[13–15]. The gas stream generated by this conversion process is a hydrogen-rich mixture, requiring the separation of CO_2 and other significantly influencing the capital costs involved in this process.

The proposition of process intensification via membrane reactor stands as a compelling prospect to streamline expenses and amplify efficiency in small-scale hydrogen production. By incorporating selective membranes within the reactor's reactive zone, the synthesis of pure hydrogen occurs within a unified apparatus, eliminating the necessity for resource-intensive downstream processes. Notably, active proton-conductive ceramic membranes in proton ceramic reactors have demonstrated favorable results in this context[16]. Fernandez et al. developed and studied thin-film Pd-Ag ceramic-supported membranes for high-temperature fluidized bed membrane reactors. Despite achieving leak-tight sealing at 600 °C, the membranes exhibited pinhole defects after seven days, prompting further investigation into the cause and potential improvements in the membrane preparation process[17].

Within this framework, modeling plays a pivotal role in supporting the industrial development of membrane reactor technology. This is especially evident in fluidized-bed configurations, where complex fluid dynamics interact with the reactive and permeation processes. A multi-scale modeling approach is generally necessary to precisely describe the entire process[7].

In the realm of numerical simulations, there has been a notable surge in the application of mathematical modeling to Membrane Reactors in recent years, and several 1D and 2D models have been proposed for modeling such reactors in the literature[18]. Poto et al. recently detailed a membrane reactor for hydrogenation of CO_2 to DME[19] using a 1D nonisothermal model. Following optimization of the membrane flux, the authors described how an effective cooling strategy improved the performance of the reactor compared to an adiabatic reactor. Furthermore, Cruellas et al.[20] used a simplified non-isothermal model to simulate improvements in the efficiency of a packed bed membrane reactor for oxidative coupling of methane and to outline the limits of packed bed reactors. Nevertheless, the 1D model is unable to account for radial mass transport, which can have a major impact on the efficiency of membrane reactors in some cases. In many cases, resistance of bed-to-wall mass transfer is a key factor affecting reactor performance, especially with selective and highly permeable membranes and to obtain a more accurate evaluation of these limitations, it is desirable to utilize 2D models, as they take into account the radial fluctuations in momentum, heat, and mass[18]. In this case, Walter et al.[21] performed a numerical investigation on packed bed membrane reactors used for propane dehydrogenation processes, demonstrating this modeling approach. The authors utilized COMSOL to tackle the numerical simulation and determined that the extent of heat production within the reactor is significantly influenced by the radial concentration profiles of oxygen, which is due to the radial diffusion of oxygen across the membrane in the reactor.

With the widespread use and impressive performance of ANN technology across many scientific and engineering domains, an ANN-centric model stands out as a strong contender to circumvent the difficulties of direct numerical simulations. In this regard[22], explored the applicability of artificial neural networks in predicting and optimizing hydrogen production across various energy modes, including biomass, solar, and wind and discussed the advantages and limitations of different hydrogen production methods and highlighted how ANN-based simulations can significantly enhance the efficiency and accuracy of hydrogen production processes. In further work[24], they also investigated the use of support vector regression (SVR) and random forest (RF) models to predict hydrogen production from biomass. They focused on the impact of catalysts in the gasification process, which is critical for reducing tar formation and enhancing hydrogen yield. Zamaniyan et al.[23] devised a three-layer feedforward backpropagation neural network (FFBN) to model an industrial hydrogen production plant utilizing steam methane reforming (SMR). The architecture incorporated four input neurons, representing crucial production parameters temperature, pressure, steam to carbon ratio, and CO_2/CH_4 ratio. In the output layer, three neurons corresponded to primary process outcomes: temperature, CO mole fraction, and hydrogen mole fraction in the produced hydrogen. The study aimed to optimize the hidden layer's neuron count to minimize the MSE between predicted and actual outputs. After rigorous exploration, the optimal number of neurons in the hidden layer was identified as five, yielding the smallest MSE value of 0.00045.

Pardo et al.[24] developed seven MLP models for simulating a steam reforming (SR) plant. Utilizing a one-year dataset at 15-minute intervals, consisting of 31,874 samples with 22 input variables, the models were optimized through genetic and memetic algorithms. The training involved 70% of randomly shuffled data, and testing utilized the remaining samples. MSE served as the loss function, with sigmoid activation, stochastic Gradient Descent, a learning rate of 0.1, 100 epochs, and a single hidden layer. The 13-10-1 network structure was found suitable for predicting hydrogen production.

Bilgic et al. presented a novel approach to optimizing the hydrogen production system within water electrolysis under the influence of a magnetic field[25]. By integrating machine learning, their study predicted the effect of various input parameters, including magnetic field, electrode material, electrolyte type, and temperature, on hydrogen output. The ANN model demonstrated high predictive accuracy, with a correlation coefficient of 0.973 and MSE of 0.01125.

Alsaffar et al.[26] developed an MLP model for thermo-catalytic CH_4 decomposition and hydrogen production. Employing Bayesian regularization (BR) and Levenberg-Marquardt (LM) training methods, input parameters included reaction temperature, catalyst weight, flow rate, calcination temperature, calcination time, and specific volume, with hydrogen yield (HY) as the output. The dataset comprising 68 samples was partitioned into three segments: training, validation, and testing, constituting 70%, 15%, and 15% of the data, respectively. Achieving an MSE of 0.03 and R2 of 0.953, the LM-trained network outperformed the BR-trained counterpart, demonstrating a 7-16-1 model topology with a predicted HY of 86.56%. Vo et al.[27] employed a 3-3-4-2-2 structured Feedforward Backpropagation Network to optimize hydrogen production. The model, trained with 81 and tested with 32 samples using the gradient descent algorithm, achieved 98.91% accuracy, demonstrating the efficacy of ANN in modeling HP and analyzing parameter relationships.

While numerous studies focus on predicting the performance of various hydrogen production systems, a comprehensive examination of the utilization of Artificial Neural Network models, specifically in the context of a membrane reactor system for hydrogen production, is lacking. This study seeks to fill this void by doing a comparative analysis of two independent techniques (MLP and RBF), examining twelve different algorithms, investigating four distinct activation functions, and analyzing a wide range of hidden layers and neurons. Furthermore, the study highlights the importance of conducting additional research to explore the importance of data and randomizations in the division of training and testing data. This will lead to a better understanding and improvement of these systems.

Materials and methods
Data collection
The dataset used for the training and validation of the models consists of 206 data points, extracted from preexisting literature sources[7,28]. These studies utilized advanced simulation models, which have been validated against experimental results to ensure their accuracy. Specifically, they provide detailed insights into the behavior of fluidized bed auto-thermal membrane reactors. Although the data points themselves are based on simulations, the models employed have undergone experimental validation, making them a reliable proxy for real-world performance. Due to the current scarcity of comprehensive experimental datasets for such systems, these experimentally validated simulations serve as the most viable alternative for accurate modeling in this study. The input and output data necessary for this study were gathered from suitable sources, including details provided in the text, tables, and graphical figures of the referenced papers. Data points presented in graphical form were extracted using the "Plot Digitizer" software, which was calibrated to convert graph pixels into their corresponding x and y coordinates for further analysis[29]. The dataset included information on reforming reactions, bubble-emulsion hydrodynamics, and membrane diffusion. The comprehensive nature of this dataset enhances the robustness of the models developed for predicting reactor performance. In the conceptualization of these studies, the initial step involved mixing the compressed biogas/biomethane with compressed air, followed by preheating. Subsequently, just before entering the reactor inlet, the mixture underwent an additional blending with steam. The resulting permeate, comprising pure hydrogen, and the retentate streams exiting the reactor were then subjected to cooling in two distinct heat exchangers.

This study assessed reactor performance and the comparison of various operating conditions, with system efficiency (η_{system}) identified as the designated output for the ANN models. Nine pivotal parameters were selected from the existing body of research and employed as inputs for the ANN models. These input parameters encompassed membrane area (m²), Mass flow of Feed (g/s), LHV of feed fuel (KJ/kg), feed CH_4 mole fraction, feed CO_2 mole fraction, feed N_2 mole fraction, feed O_2 mole fraction, pressure (bar), and reactor diameter (m). The inputs and output ranges are presented in Table 1. The system efficiency was given as follows:

$$\eta_{system} = \frac{m_{H_2}.LHV_{H_2}}{m_F.LHV_F + \frac{W_{aux}}{\eta_{el,ref}}} \qquad (1)$$

Equation 1 defines system efficiency where the terms $m_{H_2}, m_F, LHV_{H_2}, LHV_F, W_{aux}$ and $\eta_{el,ref}$ are Mass flow of permeated hydrogen (kg/s), Mass flow of fuel (kg/s), Lower heating value of hydrogen (kJ/kg), Lower heating

Parameters	Units	Type	Symbol	Ranges
Membrane Area	m²	Input	A	1.384–8.582
Mass flow of Feed	g/s	Input	B	3.42–12.25
LHV of feed fuel	KJ/kg	Input	C	12.7–45.7
CH_4 Mole fraction	-	Input	D	44.2–96.0
CO_2 Mole fraction	-	Input	E	2–34
N_2 Mole fraction	-	Input	F	1.5–3.8
O_2 Mole fraction	-	Input	G	0.5–2.7
Pressure	bar	Input	H	10–14
Diameter	m	Input	I	0.37–0.60
System efficiency	%	Output	J	62.1–71.6

Table 1. Input and output parameters for the ANN models and their ranges[7,28].

value of fuel (kJ/kg), Auxiliaries consumption (kW) and Average electric efficiency of the power generating park respectively.

Evaluation and architecture of an artificial neural network

An architecture of linked mathematical "neurons" is used by ANN to describe complex processes. The dataset is structured and connections between the ANN's layers are established using these neurons[30]. The network uses different numbers of neurons at different points in the training phase. These Interconnected neurons in an ANN allow output data to be transferred across neurons. Network weights indicate the relevance of conveyed data. Neurons calculate by adding weighted input data and comparing it to their threshold or bias value. An activation function is applied to the input data sum when it exceeds the neuron's bias value to generate output data. ANNs learn by optimizing weights and biases[31]. The general form of computations in ANNs is expressed as Eq. 2

$$net = f\left(\left(\sum_{i=1}^{n} x_i w_i\right) + b\right) \tag{2}$$

The results produced by the ANN model and the actual responses are compared to evaluate the network's performance. Thus, to determine the appropriateness and performance of the network model, the analysis involves calculating the MSE, root mean square error (RMSE), coefficient of determination (R^2), absolute average deviation (AAD), and absolute average relative deviation (AARD). Table 2 provides the pertinent equations utilized by the statistical analysis tool. Lower values of MSE, RMSE, ADD, and %AARD imply improved model precision. On the other hand, a model is typically deemed effective when the R2 value approaches 1.

Where $Y_{predicted}$, Y_{actual} and Y_{mean} refer to predicted Y value by ANN, actual Y value, and average Y value, respectively. It's important to note that Normalization of the dataset is crucial prior to presenting it to the network and prediction methods in order to mitigate the impact of a wide range of values[35]. The data underwent normalization using the Eqs. (3–5) to align the input and output data within the range of -1 to 1, ensuring compatibility with the transfer or activation function adjustments.

$$x_{normalized} = s.x + o \tag{3}$$

$$s = \frac{d_{max} - d_{min}}{x_{max} - x_{min}} \tag{4}$$

$$o = \frac{x_{max} d_{min} - x_{min} d_{max}}{x_{max} - x_{min}} \tag{5}$$

In this scenario, x_{max} corresponds to the maximum value within the datasets, and x_{min} signifies the minimum value. d_{max} and d_{min} set the upper and lower boundaries for the acceptable output range (with dmax = 1 and dmin = − 1). It acts as the scaling parameter, and o serves as the offset, enabling adjustment for optimal alignment within the range of − 1 to 1.

Multilayer perceptron architectures

Since its introduction by Rosenblatt[36] in the late 1950s, the perceptron algorithm has gained widespread recognition and is used regularly[37]. This methodology employs a simple feedforward neural network architecture that is particularly effective for classification and regression tasks. The model comprises an input layer, multiple hidden layers, and an output layer. The input from each layer is propagated to the subsequent layer. During the training phase, it is crucial to optimize various weights and biases within the network. An excessive number of neurons can lead to overfitting and prolonged processing times, both of which are undesirable. The output of each perceptron is defined by Eq. 6.

Index	Equation
Mean Square Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_{predicted} - Y_{actual}\right)^2$
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_{predicted} - Y_{actual}\right)^2}{\sum_{i=1}^{n} \left(Y_{predicted} - Y_{mean}\right)^2}$
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_{predicted} - Y_{actual}\right)^2}$
Absolute Average Deviation (AAD)	$AAD = \frac{1}{n} \sum_{i=1}^{n} \left(|Y_{actual} - Y_{predicted}|\right)$
Absolute Average Relative Deviation (AARD)	$AARD = \frac{\sum_{i=1}^{n} \left(\frac{|Y_{actual} - Y_{predicted}|}{Y_{actual}}\right)}{n} \times 100$

Table 2. Evaluation Metrics for Model Assessment through statistical analysis[32–34].

$$\gamma_i\left(x^{(j)}\right) = \phi\left(\sum_{i=1}^{N} w_{ik} x_k^{(j)} + b_i\right) \tag{6}$$

In this context, $\varphi(x)$ activation function, γ_i is the output value of the ith neuron, $x^{(j)}$ is the input value of the jth layer, $x_k^{(j)}$ is the value of the kth neuron in the jth layer, w_{ik} is the weight connecting the ith neuron to the kth neuron, and b_i is the bias value associated with the ith neuron.

The topological configuration of the MLP artificial neural network model was presented in Fig. 1. Optimizing hyperparameters, such as the number of layers, neurons per layer, transfer functions, and training methods in an artificial neural network, is contingent on the intricacy of the problem and the size of the dataset. The primary objective is to construct a neural network with the minimum acceptable number of nodes (neurons). However, the MLP model's outcomes may exhibit variation based on the initial weights assigned to input parameters, representing one of its drawbacks. To fine-tune these hyperparameters, a systematic tuning approach was developed and implemented. This approach involved testing different training methods including Levenberg-Marquardt backpropagation 'trainlm', Bayesian regularization backpropagation 'trainbr', BFGS quasi-Newton backpropagation 'trainbfg', Resilient backpropagation 'trainrp', Scaled conjugate gradient backpropagation 'trainscg', Conjugate gradient backpropagation with Powell-Beale restarts 'traincgb', Conjugate gradient backpropagation with Fletcher-Reeves updates 'traincgf', Conjugate gradient backpropagation with Polak-Ribiére updates 'traincgp', One-step secant backpropagation 'trainoss', Gradient descent with momentum and adaptive learning rate backpropagation 'traingdx', Gradient descent with momentum backpropagation 'traingdm', Gradient descent backpropagation 'traingd', with various transfer functions such as tangent sigmoid 'tansig', Log-sigmoid 'logsig', Positive linear 'poslin', saturating linear 'satlin', Symmetric saturating linear 'satlins', Linear 'purelin', Symmetric hard-limit 'hardlims' and Hard-limit 'hardlim', exploring different hidden layers with different neuron combinations. Each configuration underwent five rounds of training to ensure independence from the random assignment of weights/biases, resulting in 300,000 trained models. Subsequently, optimized cases were selected from these models based on MSE and R^2 values for further analysis.

Radial basis function (RBF)
In the late 1980s, Moody and Darken presented the RBF neural network. A few of the many benefits of this neural network type include its capacity for self-learning, its efficient storage, its quick search speed, and its quick computation times[38]. In contrast to other neural network architectures, the RBF network features a more streamlined structure and employs a more rapid learning algorithm. It comprises three distinct layers: the input layer, the hidden layer, and the output layer. The input layer handles the processing of significant non-linear information, which is subsequently passed through the hidden layer. The output layer then generates the final result. The arrangement of the RBF network employed in the present study is depicted in Fig. 2.

The Gaussian function shown in Eq. 7 is commonly used as an activation function for radial-based neurons among the different options available.
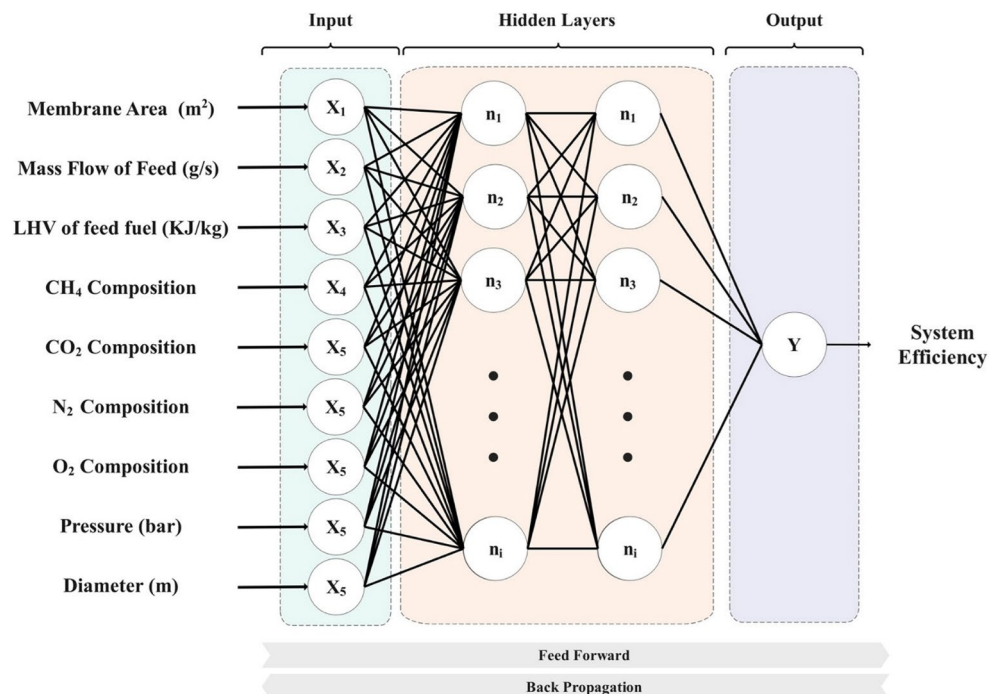

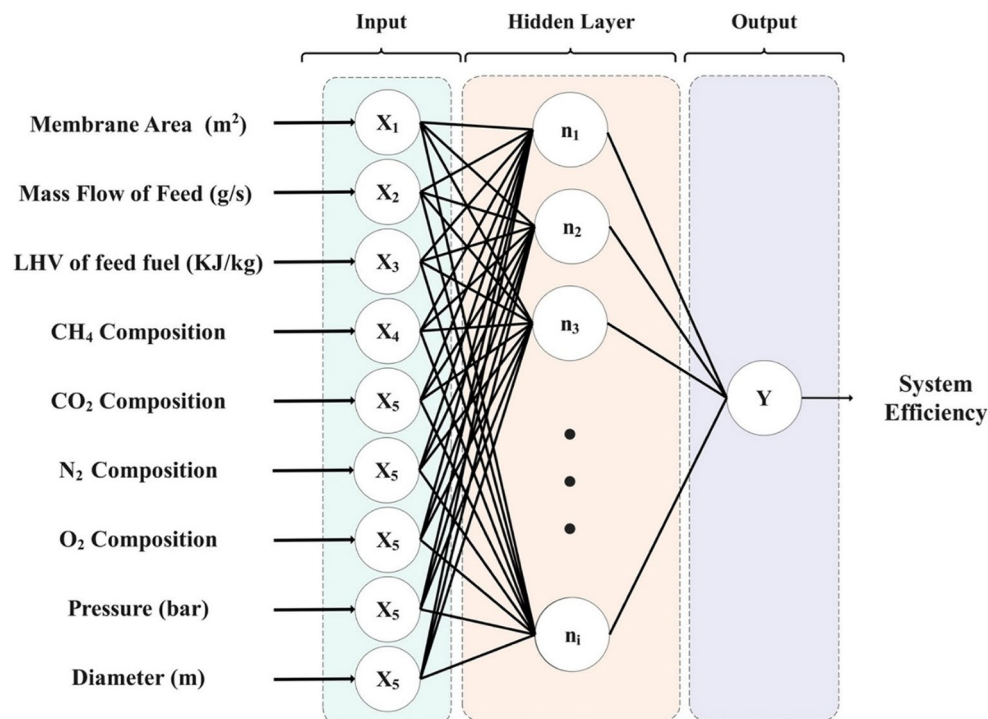**Figure 1**. The depicted topological configuration of the MLP artificial neural network model.

Figure 2. The depicted topological configuration of the RBF artificial neural network model.

$$\phi_i \left(\|x - c_i\| * b\right) = exp\left(-\frac{1}{2\sigma_i^2}(\|x - c_i\| * b)^2\right) \tag{7}$$

The parameters c_i and σ denote the center and spread of the Gaussian function, respectively, whereas x represents the input and φ_i the output in this context. Also, the bias term is represented by b. In a linear fashion, the activation function and the weight vector w of the output layer are combined to produce the network's output (y).

$$y = \sum_{i=0}^{n} \phi_i w_i \tag{8}$$

Where, " w_i " denotes the combined weighted value of the ith basis function.

Optimization of operating conditions using genetic algorithm

After optimizing the hyperparameters of the ANN models, a genetic algorithm (GA) was employed to determine the best operating conditions for the membrane reactor hydrogen generation plant based on the optimized network. The genetic algorithm, first proposed by John Holland[39], is a probabilistic global optimization technique that mimics the mechanism of natural biological evolution. The genetic algorithm is based on the fundamental darwinian concept of survival of the fittest. It works by iteratively improving a population of potential solutions to approach the best possible outcome. In contrast to conventional optimization methods, the genetic algorithm performs its search by considering a population of solutions instead of a single point. This approach helps prevent the algorithm from converging to poor answers.

In the GA, potential solutions to a specific problem are encoded in a chromosome-like data structure, where genes represent the parameters to be optimized. In this study, float-point coding is used for optimization of the operational parameters of the mentioned membrane reactor system once the optimal artificial neural network model is available. Each chromosome vector is coded as a vector of floating point numbers corresponding to the input parameters. The chromosome is defined as a real number vector $X = (x_1, x_2, \ldots, x_9)$, where $-1 < x_i < 1$.

The GA uses a fitness function to evaluate the adaptability of individuals without external information in the evolution search. The adaptability is expressed by the fitness value, where a higher fitness value indicates better adaptability subject to constraints and better viability of the individual. The fitness function, which is not constrained by definition domain, continuity, and differentiability, requires that the objective function is defined in the form of a non-negative maximum. In this optimization, the system efficiency based on the output of the ANN model is selected as the fitness function.

GA encompasses three fundamental operators: the selection operator, crossover operator, and mutation operator. The selection operator plays a pivotal role in identifying suitable parents for generating the subsequent generation of solutions. It achieves this by discerning individuals based on their fitness, with a higher likelihood

of selection bestowed upon those exhibiting superior fitness levels. Thus, heightened fitness correlates positively with the probability of selection, thereby promoting the propagation of favorable genetic traits through successive generations. The crossover operator stands as a foundational mechanism for generating novel chromosomes. It engenders offspring with a blend of genetic material derived from both parental sources. In this optimization scenario, the straightforward arithmetic crossover method is employed due to its uncomplicated nature, as elucidated below:

$$\begin{cases} c_1 = \alpha\, f_1 + (1 - \alpha)\, f_2 \\ c_2 = \alpha\, f_2 + (1 - \alpha)\, f_1 \end{cases} \tag{9}$$

In this context, α represents a pseudo-random number sampled uniformly from the interval [0, 1]. The parental individuals selected for crossover are designated as f_1 and f_2, while the resultant offspring, produced through the crossover process, are denoted as c_1 and c_2. Mutation serves a crucial purpose in the GA framework as it addresses potential shortcomings arising from the selection and crossover processes. While selection and crossover operations collaborate to explore novel solutions, they may inadvertently lead to premature convergence, risking the loss of valuable genetic diversity. The fundamental role of mutation within the GA paradigm is to reintroduce lost or underexplored genetic material into the population, thereby safeguarding against premature convergence and the emergence of suboptimal solutions.

Results and discussion

Comprehensive ANN models were constructed to showcase the prospective utility of ANN in predicting the system efficiency of a hydrogen production membrane reactor. These models were generated employing twelve distinct training functions, eight varied activation functions, and up to four hidden layers featuring different combinations of neurons in each layer. The evaluation of ANN models utilized the metrics outlined in Table 2. The prediction outcomes of these ANN models underwent thorough assessment, presented alongside statistical analyses and data interpretations.

Tuning the MLP network

To mitigate the risk of overfitting in our neural network model, a rigorous data partitioning strategy was implemented. Initially, 15 data points were randomly selected from the dataset and held out, ensuring they were never exposed to the model during training. Subsequently, the remaining dataset was stratified into three distinct subsets: 70% for network training, 15% for validation, and an additional 15% designated for network testing. This systematic approach aimed to enhance the robustness and generalization capability of the neural network by rigorously assessing its performance on diverse and unseen datasets. Furthermore, an ANN model was developed to predict system efficiency, utilizing a wide range of hidden layers, training algorithms, and activation functions. The relationship among hidden layers, training functions, and mean squared error, which represents the difference between the actual and predicted system efficiency, is depicted in Fig. 3. This representation is based on the outcomes of the ANN model's forecasts for the training dataset and the unseen validation set, which consists of entirely fresh data points. Each data point in this figure represents the optimal outcome derived from an extensive pool of over 3000 meticulously trained models. The models exhibited diversity in the number of hidden layers and neurons present in each layer, thereby encompassing a thorough investigation of the performance landscape of the neural network.

The observations obtained from Fig. 3 provide a detailed understanding of this work, highlighting a distinct pattern in the performance of specific algorithms and training functions. Although the Mean Squared Error reached an acceptable level during training, certain combinations resulted in notably higher MSE values for the validation dataset. This disparity indicates a possible problem of overtraining, where the model performs exceptionally well in delivering precise responses for the data it was trained on but faces difficulties in applying its knowledge to unfamiliar datasets. Nevertheless, a set of noteworthy exceptions, like lm-tansig, lm-logsig, and lm-poslin, exhibited excellent precision on both the training and validation datasets. Therefore, these specific models were chosen for additional thorough examination due to their strong performance and potential for dependable generalization. Within the category of sigmoidal activation functions, both tansig and logsig (depicted in Fig. 4) exhibit similar effects, to utilize the minimal number of neurons. Tansig, achieving the highest average coefficients of determination for both training and testing ($R^2 = 0.9979$ and 0.9962) with [20 17 6] neurons, outperformed other configurations. In Fig. 4, following closely is logsig ($R^2 = 0.9971$ and 0.9961) with 23, 20, 2 neurons at the first, second, and third layers, offering an acceptable prediction of system efficiency. Furthermore, Table 3 showcases some of the top-performing average correlation coefficients between the actual and predicted system efficiency, emphasizing superior activation functions (tansig, logsig, poslin, and satlins) along with the associated ANN models, algorithms, and hidden layers.

RBF training and optimization

Optimizing the parameters of the RBF network is crucial in comparing the RBF and MLP networks, and this optimization process is essential to ensure an in-depth inspection of the RBF network's performance compared to the MLP network. In our RBF model's training phase, we extensively explored various models by adjusting spreads and maximum neurons. The MSE was systematically compared between the predicted and actual system efficiency for the training and validation datasets. As depicted in Fig. 5, the outcomes reveal that the optimal MSE for the validation dataset is attainable when the spread is set to 1, and the maximum number of neurons is 35. This careful examination of different model configurations provides valuable insights into the parameter settings that yield the most accurate predictions for system efficiency.
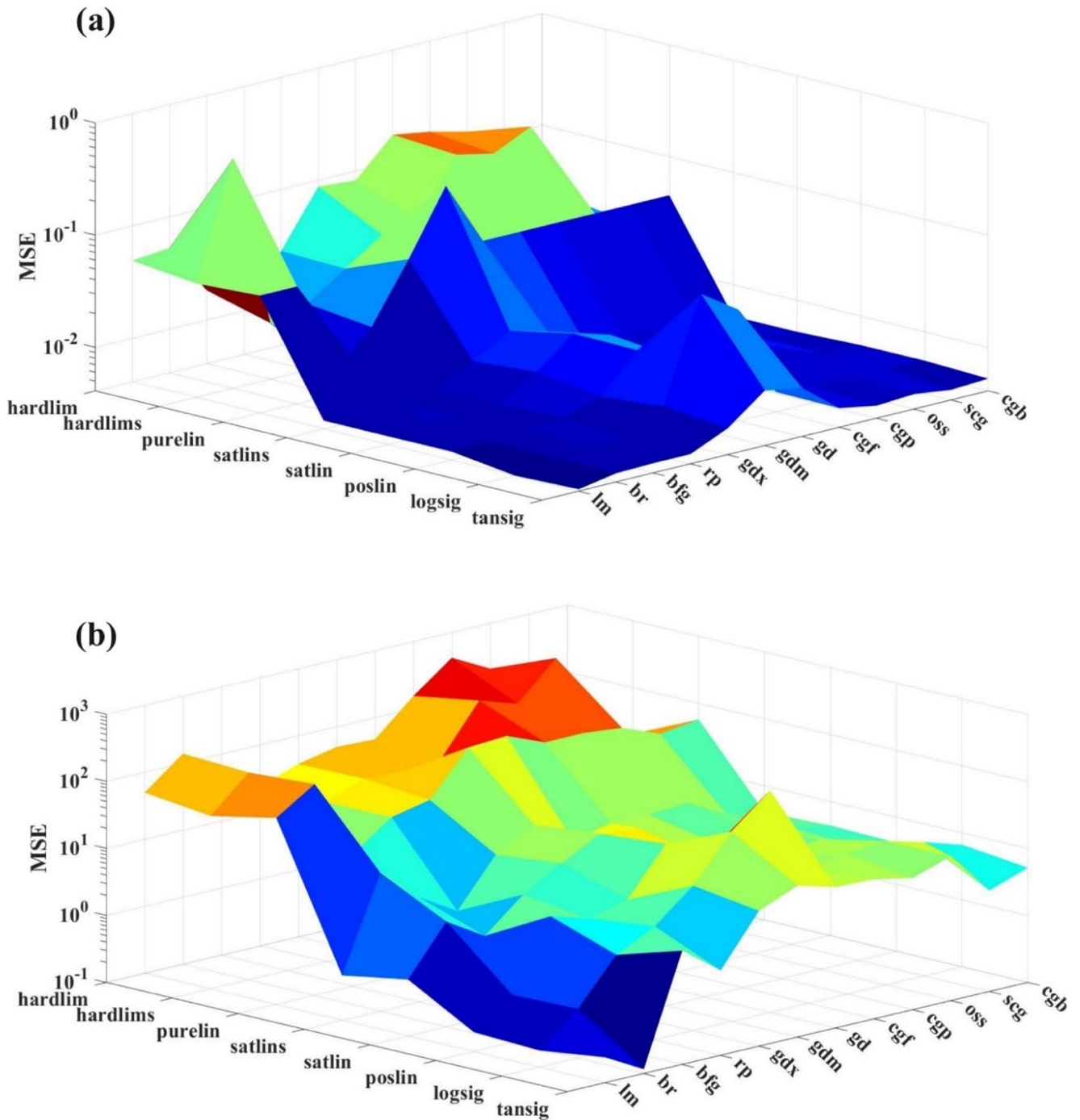
Figure 3. Prediction results of ANN models with different training algorithm and transfer functions for (**a**) trainng dataset (**b**) validation data set.

Upon attaining optimal values for the spread and maximum neurons, a comparative analysis between MLP and RBF training was conducted, as illustrated in Fig. 6. The results revealed that the best-case scenario for the MLP model exhibited superior accuracy with reduced computational time. This observation underscores the efficiency and effectiveness of the MLP model, making it a preferable choice in terms of accuracy and computational efficiency compared to the RBF model for this scenario.

Analysis of variable interrelationships

The Pearson correlation matrix is a pivotal tool in statistical analysis, showcasing the correlation coefficients among pairs of input features within a dataset (Fig. 7).

These coefficients measure the extent of linear relationships between variables, ranging from -1 to $+1$. A correlation of $+1$ indicates a perfect positive correlation, -1 signifies a perfect negative correlation, and 0 denotes
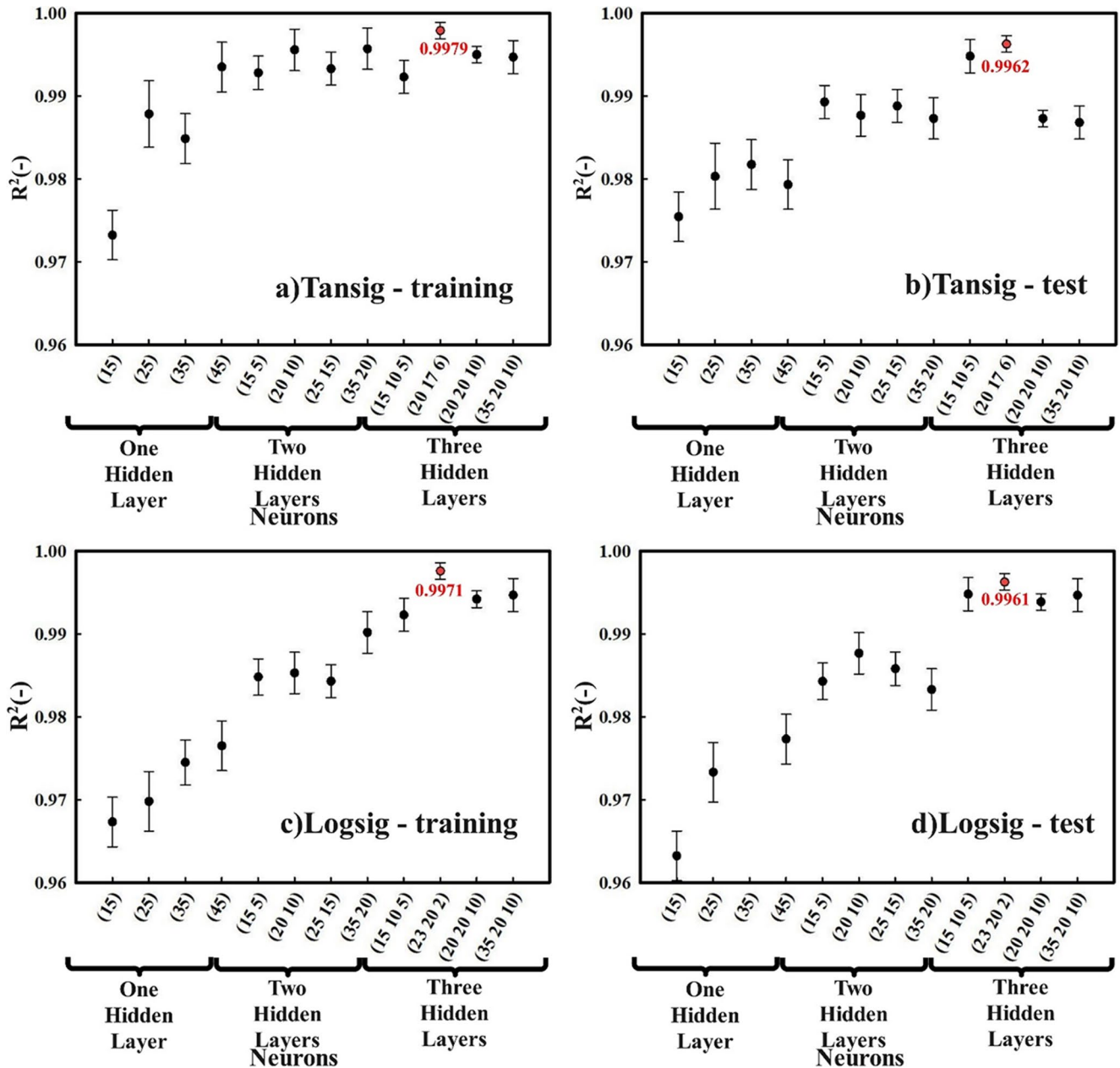
Figure 4. Impact of the sigmoidal activation function (applied to all layers with different combination of layers and neurons) on the prediction of System efficiency using ANN models: (**a**) Tansig-training, (**b**) Tansig-test, (**c**) Logsig-training, and (**d**) Logsig-test. Error bars represent the standard deviation over 20 iterations.

the absence of a linear correlation. The diagonal elements of the matrix invariably have a correlation coefficient of 1, reflecting each variable's self-correlation. Examining the correlation matrix yields critical insights into the dataset's underlying structure. The magnitude and direction of the correlation coefficients elucidate the strength and nature of relationships between variables. A coefficient approaching $+1$ suggests a strong positive correlation, meaning that as one variable increases, so does the other. Conversely, a coefficient nearing -1 indicates a strong negative correlation, where an increase in one variable corresponds to a decrease in the other[40,41]. Analyzing these correlations is invaluable for understanding the dataset's characteristics, guiding data preprocessing, and informing subsequent modeling steps. This analysis can help identify multicollinearity issues, select relevant features, and better understand the interplay between different variables in the context of system efficiency optimization[42]. In this study, we generated 1000 random data samples and utilized the optimized ANN model to derive the Pearson correlation matrix. This approach ensures that our correlation analysis reflects the underlying patterns learned by the model.

The correlation matrix (Fig. 7) reveals several significant relationships among the variables. Notably, system efficiency shows strong negative correlations with CH_4 (-0.34) and LHV (-0.31), highlighting the impact of these variables on overall system performance. Additionally, a significant positive correlation exists between system

A.F[1]	Alg.[2]	HL[3]	N.N.[4]	The results of the trained model			
				R^2(-)	MSE-train(-)	MSE-test(-)	RMSE
tansig	Lm	3	(20 17 6)	0.998	0.004255	0.23951	0.4894
logsig	Lm	3	(23 20 2)	0.997	0.004086	0.17694	0.4206
tansig	br	3	(17 13 7)	0.994	0.004901	0.11652	0.3413
logsig	br	3	(19 11 8)	0.994	0.004907	0.54319	0.7370
poslin	lm	3	(25 12 3)	0.995	0.004865	0.19453	0.4411
poslin	gdx	3	(24 18 5)	0.989	0.007072	5.29875	2.3019
satlins	rp	3	(25 14 3)	0.992	0.005052	1.67146	1.2929

Table 3. Correlation coefficients, corresponding MSE, and RMSE for the selected ANN models and activation functions. [1]Activation function. [2]Training Algorithm. [3]Hidden Layer. [4]Number of neurons.
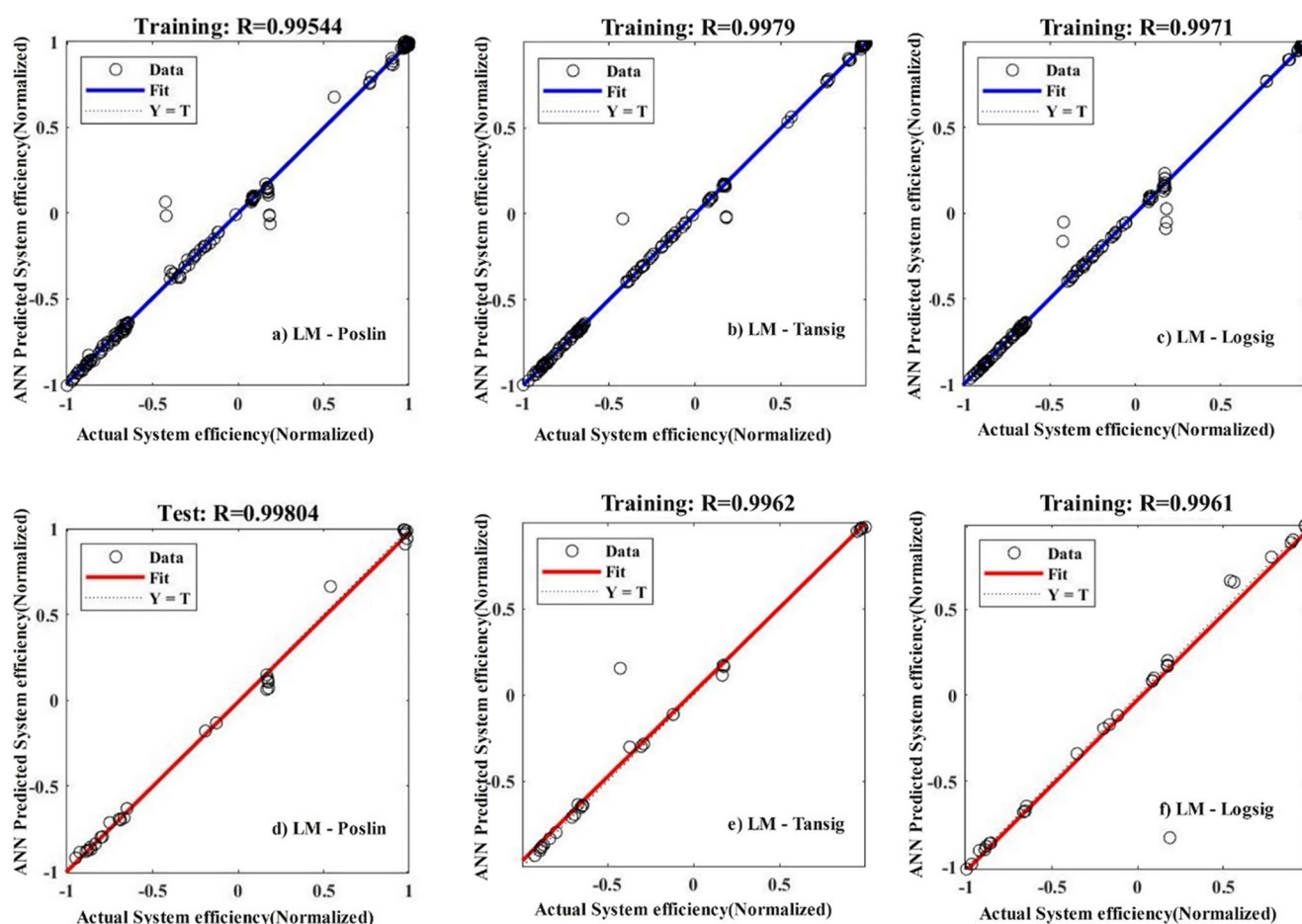
Figure 5. Regression plots for models with lm training function and tansig, logsig and posling transfer functions.

efficiency and reactor diameter (0.26), indicating that bigger reactor diameter corresponds with improved system efficiency. This detailed correlation analysis not only aids in feature selection and multicollinearity detection but also enhances our understanding of the system's dynamics. It enables us to identify which variables are most influential in determining system efficiency and to develop more accurate predictive models. We can optimize the system's performance more effectively by leveraging these insights.

The influence of various factors on system efficiency was also investigated using a deviation plot based on the optimized ANN model (Fig. 8). The optimal operating conditions were determined using a Genetic Algorithm, resulting in the following reference point values (Coded/Uncoded): membrane area (0.132/5.44), CH_4 (-0.999/44.2), CO_2 (0.226/21.62), N_2 (0.639/3.38), O_2 (0.243/1.86), mass flow (0.503/10.05), LHV (-0.999/12.7), pressure (0.997/13.99), and diameter (0.994/0.599). These values represent the ideal conditions under which the system operates at maximum efficiency (72.8). Furthermore, the calculated system efficiency was compared with those reported in the literature for conventional and membrane reactor systems. Specifically, the efficiency of
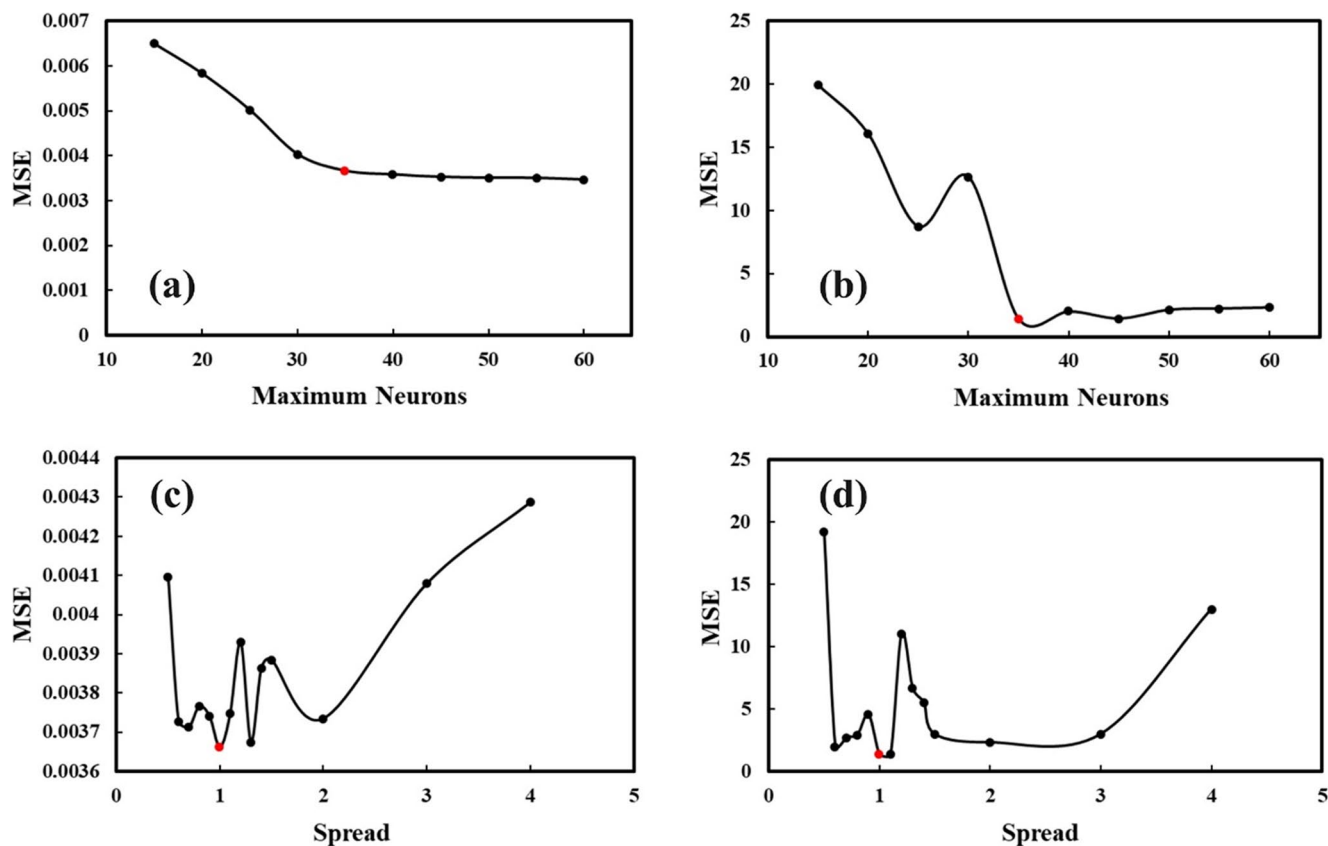
Figure 6. Optimization of the maximum neurons and spread parameter for the RBF neural network with **a**, **c**) Training and **b**, **d**)validation datasets.
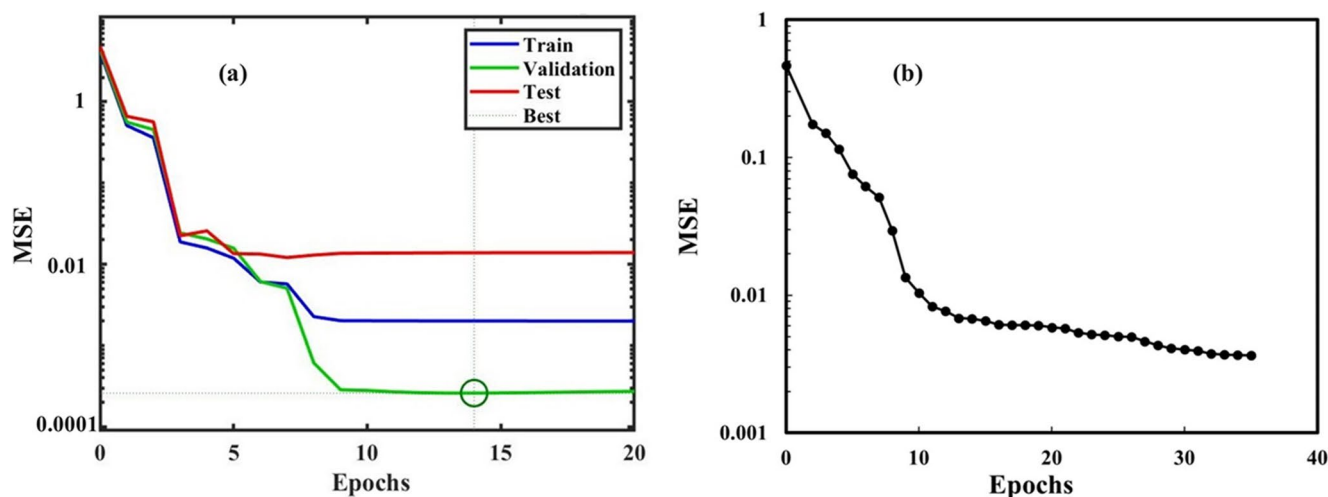

Figure 7. Comparison between best case scenarios for (**a**) MLP network and (**b**) RBF network.

biogas steam reforming (SR) coupled with pressure swing adsorption (PSA) and autothermal reforming (ATR) with PSA, as reported by Marcoberardino et al.[43], was considered. Comparisons were also made with other studies on Pd-based membrane reactors[7,28].

Figure 8 reveals that factor B (CH_4) significantly negatively influences system efficiency, as indicated by the steep decline in efficiency with increasing deviation from the reference point. Conversely, factors like a (membrane area) and G (mass flow) show relatively stable efficiency levels across the deviation range, suggesting a lesser impact on system performance. This detailed analysis highlights the critical system efficiency variables, enabling targeted optimization strategies to enhance overall performance.
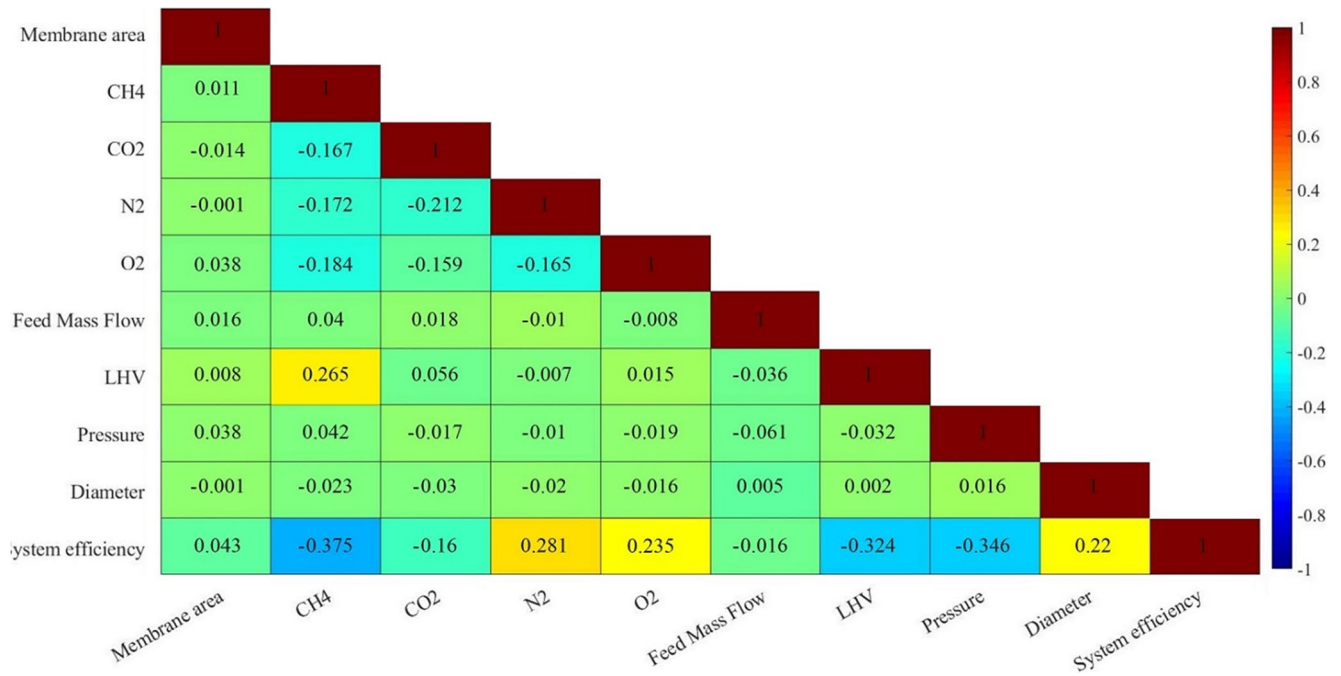
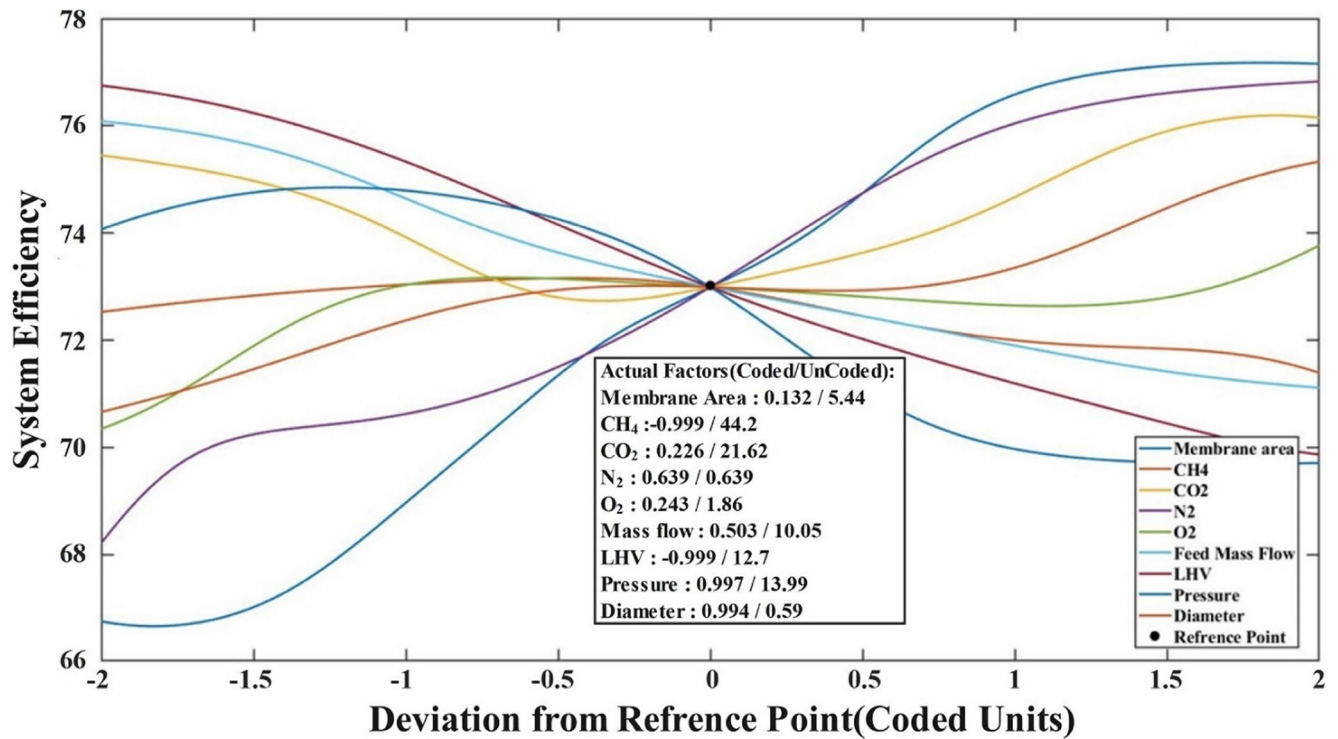**Figure 8.** Pearson correlation matrix based on the new data.

Figure 9. Deviation plot for system efficiency.

3D response surfaces of optimal network

In examining parameter effects on system efficiency via neural networks, a deliberate exploration was conducted on the impact of individual factors: pressure, membrane area, fuel LHV, reactor diameter, and feed mass flow (Fig. 9). This investigation involved the selective variation of two distinct parameters at a time while holding the remaining variables constant at values optimized through GA for achieving peak system efficiency (Fig. 10). Consequently, the neural network model calculated the system efficiency for these newly introduced data
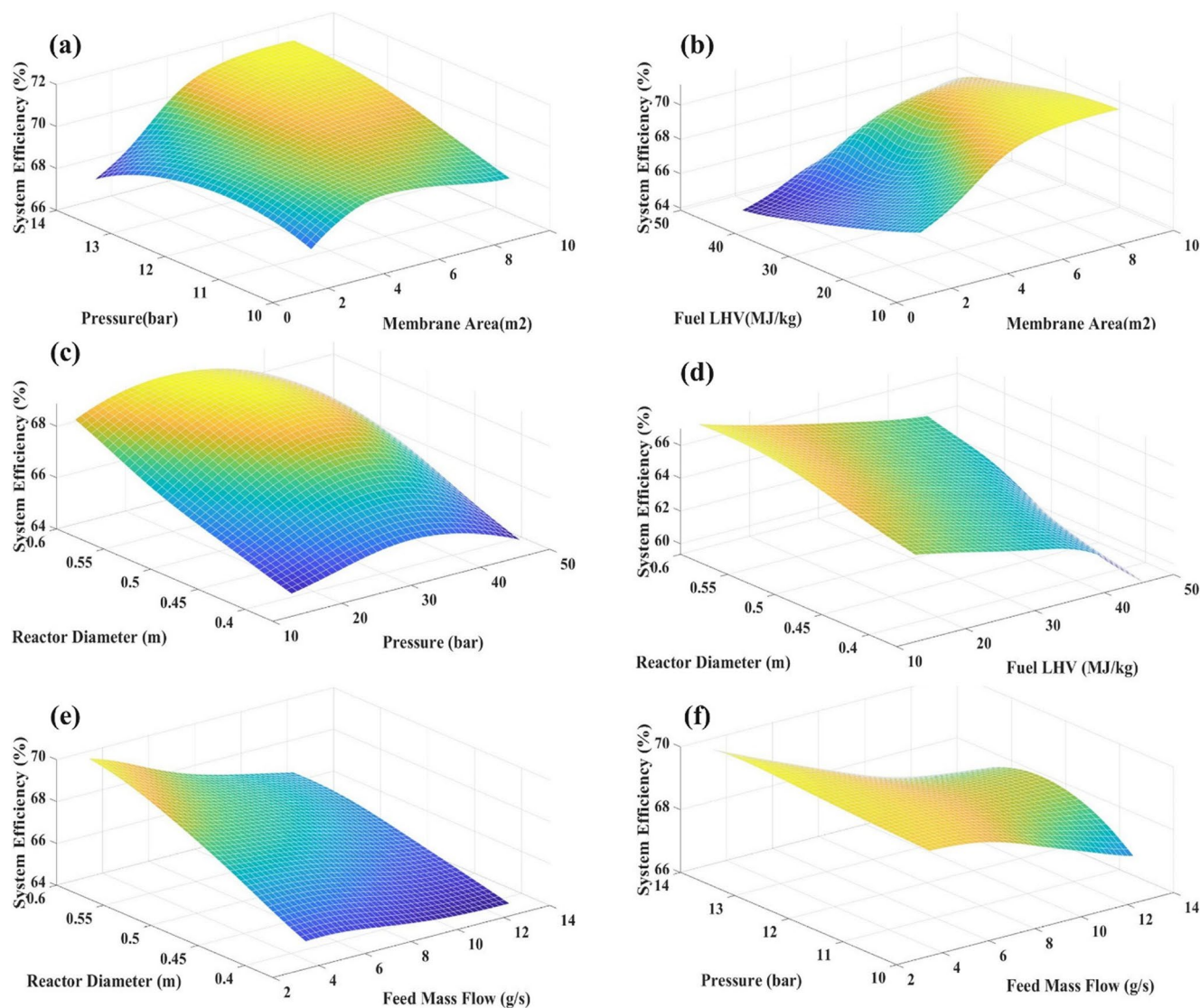
Figure 10. Three-dimensional response surface plots generated using MLP model for (**a**) pressure and membrane area, (**b**) Fuel LHV and membrane area, (**c**) Reactor Diameter and Pressure, (**d**) Reactor diameter and fuel LHV, (**e**) Reactor diameter and feed mass flow and (**f**) Pressure and feed mass flow.

points. The graphical representation of these calculations unfolds in the form of 3D response surfaces, elegantly depicted in Fig. 11.

The analysis of the Fig. 10 reveals several critical insights into the interactions between the parameters. In Fig. 10a, depicting pressure versus membrane area, it is evident that increasing membrane area consistently enhances system efficiency. However, the effect of pressure is more nuanced; higher pressures improve efficiency when membrane areas are large, whereas pressures above 12 bar diminish efficiency at smaller membrane areas. This suggests that an optimal pressure range exists that synergizes with larger membrane areas to maximize system performance. Similarly, in the graph of fuel LHV versus membrane area, higher fuel LHV generally reduces system efficiency, especially at lower membrane areas (Fig. 10b). Despite this negative trend, increasing the membrane area continues to improve efficiency, highlighting the importance of membrane dimensions in offsetting the adverse effects of fuel quality. Lastly, the interactions involving feed mass flow and pressure illustrate additional complexities (Fig. 10f). Increasing reactor diameter improves efficiency, but higher feed mass flow reduces efficiency, particularly in systems with larger reactor diameters. In contrast, the graph of pressure versus feed mass flow shows that increased pressure enhances system efficiency, whereas higher feed mass flow detracts from it. These observations indicate that while larger reactor dimensions and higher pressures are generally beneficial, they must be balanced with controlled feed mass flow rates to prevent efficiency losses. These observed trends could be contextualized using Eq. (1), where increasing parameters such as membrane area and reactor diameter likely enhance the mass flow of permeated hydrogen and its effective utilization, thereby boosting the numerator in the efficiency formula. Conversely, parameters like fuel LHV and feed mass flow may negatively impact the denominator and thus reducing overall system efficiency. Understanding these
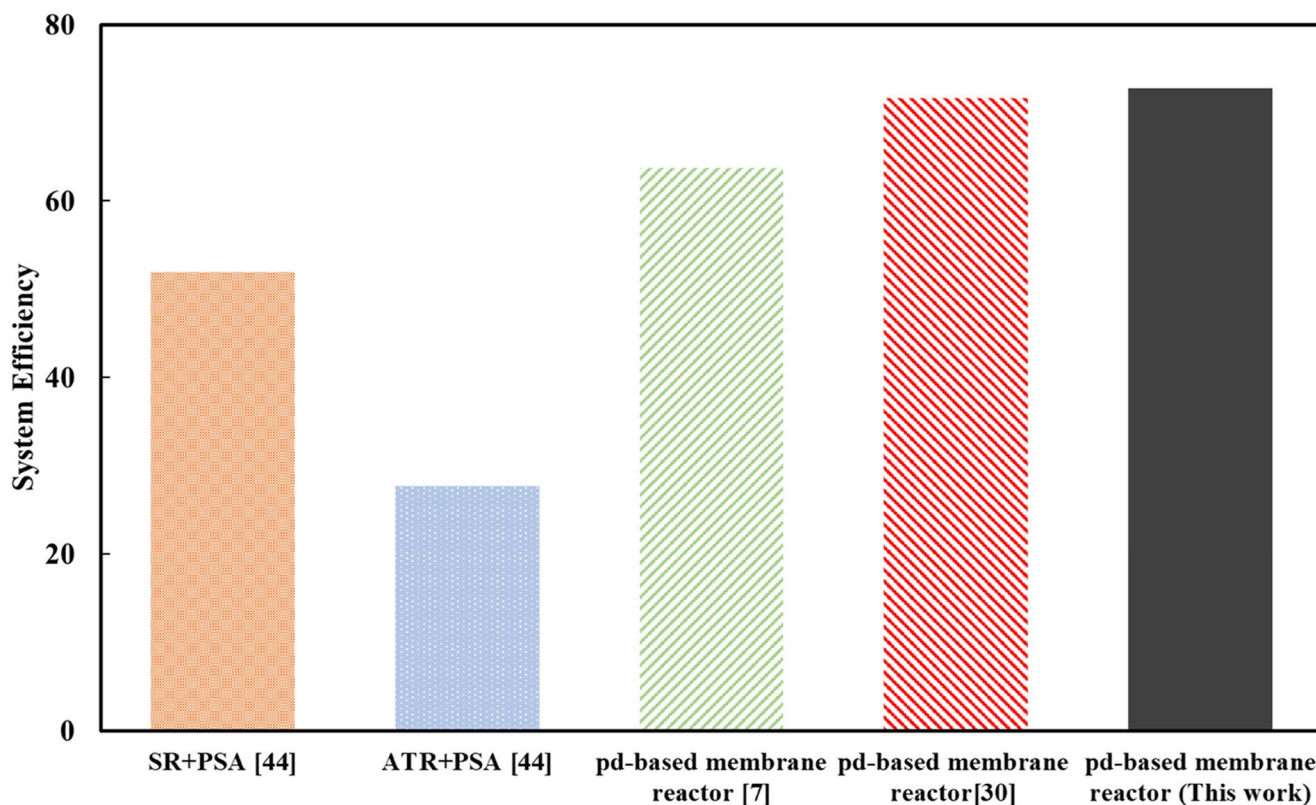
Figure 11. Comparison of best-case scenarios for system efficiencies in different biogas reforming studies.

interactions provides valuable insights into optimizing the operational parameters to maximize hydrogen generation efficiency.

Conclusion

This study employed various Artificial Neural Network models to scrutinize and forecast system efficiency concerning key operational variables in membrane reactor systems used for hydrogen production. The methodologies included MLP and RBF networks, optimized through twelve training algorithms and eight activation functions, with configurations exploring up to three hidden layers and variable neuron counts. The analysis revealed that the MLP models consistently outperformed the RBF models, particularly when using the Levenberg-Marquardt training algorithm and tansig activation function, which achieved a R^2 of 0.9975 during training and 0.9962 during testing, along with a mean MSE of 0.0042 and 0.2395, respectively. Increasing membrane area consistently enhanced system efficiency. Higher pressures improved efficiency when membrane areas were large; however, pressures above 12 bar reduced efficiency at smaller membrane areas, indicating an optimal pressure range for maximizing performance. Additionally, higher fuel with lower heating values generally reduced system efficiency, especially at lower membrane areas, but increasing the membrane area mitigated these adverse effects, underscoring the importance of optimizing membrane dimensions. Furthermore, increasing reactor diameter improved efficiency, while higher feed mass flow rates reduced efficiency, particularly in larger reactor systems. This highlights the necessity of balancing these parameters to prevent efficiency losses. Increased pressure was also found to enhance system efficiency, whereas higher feed mass flow rates detracted from it, emphasizing the need for controlled feed mass flow rates to maintain optimal efficiency. Quantitative analysis revealed that the MLP model using the logsig activation function, with an architecture of 23, 20, 2 neurons at the first, second and third layers, also performed well, achieving R^2 values of 0.9971 for training and 0.9961 for testing, with MSE values of 0.004086 for training and 0.17694 for testing. Optimization of the RBF network identified the best performance with a spread parameter of 1 and 35 neurons, although the MLP models demonstrated superior accuracy and reduced computational time. Genetic algorithms were utilized to find the optimal operating variables, enhancing the model's predictive capability and overall efficiency. After identifying the optimal operating conditions, the interrelationship between variables was analyzed using Pearson correlation matrices, deviation plots, and 3D response surface plots. These analyses provided deeper insights into the interactions among variables, revealing significant correlations and allowing for a more comprehensive understanding of how different factors influence system efficiency under optimal conditions. In conclusion, the study demonstrates that optimizing operational parameters such as membrane area, pressure, and reactor diameter is essential for maximizing hydrogen generation efficiency. The ANN models, particularly the MLP method, provide a robust framework for predicting and enhancing system performance, with the potential to identify optimal conditions for hydrogen production. Future research should focus on refining these models

and exploring additional parameters to achieve even greater efficiencies in hydrogen production systems, contributing to the advancement of membrane reactor technology for sustainable hydrogen production.

Data availability

References

1. Bouckaert, S. et al. *Net zero by 2050: a roadmap for the global energy sector* (2021).
2. Granovskii, M., Dincer, I. & Rosen, M. A. Economic and environmental comparison of conventional, hybrid, electric and hydrogen fuel cell vehicles. *J. Power Sources* **159** (2), 1186–1193 (2006).
3. Marbán, G. & Valdés-Solís, T. Towards the hydrogen economy? *Int. J. Hydrog. Energy*. **32** (12), 1625–1637 (2007).
4. Shinnar, R. The hydrogen economy, fuel cells, and electric cars. *Technol. Soc.* **25** (4), 455–476 (2003).
5. Subramani, V., Basile, A. & Veziroglu, T. N. *Compendium of hydrogen energy: hydrogen production and purification* (Woodhead Publishing, 2015).
6. Van Mierlo, J., Maggetto, G. & Lataire, P. Which energy source for road transport in the future? A comparison of battery, hybrid and fuel cell vehicles. *Energy. Conv. Manag.* **47** (17), 2748–2760 (2006).
7. Ongis, M. et al. Membrane reactors for green hydrogen production from biogas and biomethane: a techno-economic assessment. *Int. J. Hydrog. Energy*. **48** (51), 19580–19595 (2023).
8. Nikolaidis, P. & Poullikkas, A. A comparative overview of hydrogen production processes. *Renew. Sustain. Energy Rev.* **67**, 597–611 (2017).
9. Albrecht, U. et al. *Study on hydrogen from renewable resources in the EU. Final Report* 17 (2015).
10. Braga, L. B. et al. Hydrogen production by biogas steam reforming: a technical, economic and ecological analysis. *Renew. Sustain. Energy Rev.* **28**, 166–173 (2013).
11. Alves, H. J. et al. Overview of hydrogen production technologies from biogas and the applications in fuel cells. *Int. J. Hydrog. Energy*. **38** (13), 5215–5225 (2013).
12. Ohkubo, T., Hideshima, Y. & Shudo, Y. Estimation of hydrogen output from a full-scale plant for production of hydrogen from biogas. *Int. J. Hydrog. Energy*. **35** (23), 13021–13027 (2010).
13. Holladay, J. D. et al. An overview of hydrogen production technologies. *Catal. Today*. **139** (4), 244–260 (2009).
14. Ugarte, P. et al. Dry reforming of biogas in fluidized bed: process intensification. *Int. J. Hydrog. Energy*. **42** (19), 13589–13597 (2017).
15. Göransson, K. et al. Review of syngas production via biomass DFBGs. *Renew. Sustain. Energy Rev.* **15** (1), 482–492 (2011).
16. Clark, D. et al. Single-step hydrogen production from NH3, CH4, and biogas in stacked proton ceramic reactors. *Science*. **376** (6591), 390–393 (2022).
17. Fernandez, E. et al. Preparation and characterization of thin-film Pd–Ag supported membranes for high-temperature applications. *Int. J. Hydrog. Energy*. **40** (39), 13463–13478 (2015).
18. Gallucci, F. Modeling of membrane reactors. In *Current Trends and Future Developments on (Bio) Membranes* 315–335 (eds Basile, A. & Gallucci, F.) (Elsevier, 2023).
19. Poto, S., Gallucci, F. & d'Angelo, M. F. N. Direct conversion of CO2 to dimethyl ether in a fixed bed membrane reactor: influence of membrane properties and process conditions. *Fuel*. **302**, 121080 (2021).
20. Cruellas, A. et al. Oxidative coupling of methane: a comparison of different reactor configurations. *Energy Technol.* **8** (8), 1900148 (2020).
21. Walter, J. P. et al. Model-based analysis of fixed-bed and membrane reactors of various scale. *Chem. Ing. Tech.* **93** (5), 819–824 (2021).
22. Bilgiç, G. et al. Recent advances in artificial neural network research for modeling hydrogen production processes. *Int. J. Hydrog. Energy*. **48** (50), 18947–18977 (2023).
23. Zamaniyan, A. et al. Application of artificial neural networks (ANN) for modeling of industrial hydrogen plant. *Int. J. Hydrog. Energy*. **38** (15), 6289–6297 (2013).
24. Pardo, E. G. et al. Optimization of a steam reforming plant modeled with artificial neural networks. *Electronics* **9** (11), 1923 (2020).
25. Bilgiç, G. et al. Prediction of hydrogen production by magnetic field effect water electrolysis using artificial neural network predictive models. *Int. J. Hydrog. Energy*. **48** (53), 20164–20175 (2023).
26. Alsaffar, M. A. et al. Artificial neural network modeling of thermo-catalytic methane decomposition for hydrogen production. *Top. Catal.* **64**, 456–464 (2021).
27. Vo, N. D. et al. Combined approach using mathematical modelling and artificial neural network for chemical industries: steam methane reformer. *Appl. Energy*. **255**, 113809 (2019).
28. Di Marcoberardino, G. et al. Potentiality of a biogas membrane reformer for decentralized hydrogen production. *Chem. Eng. Processing-Process Intensif.* **129**, 131–141 (2018).
29. Shafizadeh, A. et al. Machine learning-enabled analysis of product distribution and composition in biomass-coal co-pyrolysis. *Fuel*. **355**, 129464 (2024).
30. Kolbadinejad, S. et al. Deep learning analysis of Ar, Xe, Kr, and O2 adsorption on activated Carbon and zeolites using ANN approach. *Chem. Eng. Processing-Process Intensif.* **170**, 108662 (2022).
31. Khoshraftar, Z. et al. Using halloysite nanotubes modified by tetraethylenepentamine for advanced carbon capture: experimental and modeling via RSM and ANNs. *Chem. Eng. J. Adv.* **16**, 100543 (2023).
32. Bakhtom, A. et al. Machine learning-guided prediction and optimization of precipitation efficiency in the bayer process. *Chem. Pap.* **77** (5), 2509–2524 (2023).
33. Naderi, K., Foroughi, A. & Ghaemi, A. Analysis of hydraulic performance in a structured packing column for air/water system: RSM and ANN modeling. *Chem. Eng. Processing-Process Intensif.* **193**, 109521 (2023).
34. Naderi, K. et al. Modeling based on machine learning to investigate flue gas desulfurization performance by calcium silicate absorbent in a sand bed reactor. *Sci. Rep.* **14** (1), 954 (2024).
35. Bahmanzadegan, F. & Ghaemi, A. Exploring the effect of Zeolite's structural parameters on the CO2 capture efficiency using RSM and ANN methodologies. *Case Stud. Chem. Environ. Eng.* **9**, 100595 (2024).
36. Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65** (6), 386 (1958).
37. Bariki, S. G. & Movahedirad, S. Comparative analysis of artificial neural network (ANN) models: CO2 loading in MDEA and blended MDEA/PZ solvents. *Fuel*. **357**, 129667 (2024).
38. Moody, J. & Darken, C. J. Fast learning in networks of locally-tuned processing units. *Neural Comput.* **1** (2), 281–294 (1989).

39. John, H. Holland. Genetic algorithms. *Sci. Am.* **267** (1), 44–50 (1992).
40. Khoshraftar, Z. & Ghaemi, A. *Enhanced carbon dioxide adsorption using lignin-derived and nitrogen-doped porous carbons: a machine learning approaches, RSM and isotherm modeling* 100668 (Case Studies in Chemical and Environmental Engineering, 2024).
41. Khoshraftar, Z., Ghaemi, A. & Hemmati, A. Comprehensive investigation of isotherm, RSM, and ANN modeling of CO2 capture by multi-walled carbon nanotube. *Sci. Rep.* **14** (1), 5130 (2024).
42. Mehrmohammadi, P. & Ghaemi, A. Investigating the effect of textural properties on CO2 adsorption in porous carbons via deep neural networks using various training algorithms. *Sci. Rep.* **13** (1), 21264 (2023).
43. Marcoberardino, G. D. et al. Green hydrogen production from raw biogas: a techno-economic investigation of conventional processes using pressure swing adsorption unit. *Processes.* **6** (3), 19 (2018).

Author contributions

Mehrdad Mahmoudi: Conceptualization, Methodology, Conceived and designed the experiments, Validation, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing. Ahad Ghaemi: Conceptualization, Methodology, Software, Conceived and designed the experiments, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Supervision Visualization, Project administrationAhmad Rahbar: Conceptualization, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - review & editing.Salman Movahedirad : Conceptualization, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - review & editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-75068-y.

Correspondence and requests for materials should be addressed to A.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.