JOURNAL OF RENEWABLE ENERGY AND SMART GRID TECHNOLOGY

Vol. 21, No. 1, January-June 2026



Roles of Activation Functions on the Polarization Modeling for the 1.2 kWp Proton Exchange Membrane Fuel Cells: A Focus on ReLU, Sigmoid, and Tanh

Chanin Khamwandee , Arnusorn Saengprajak*

Energy Research Unit, Department of Physics, Faculty of Science, Mahasarakham University, Maha Sarakham 44150, Thailand

***Corresponding author's email**: arnusorn.s@msu.ac.th

Article info:

Received: 28 July 2025
Revised: 20 February 2026
Accepted: 4 March 2026
DOI:
10.69650/rast.2026.263264

Keywords:

Fuel Cell
Modelling
ANN
Activation Function
Polarization Curve
ANNOVA
LabVIEW

ABSTRACT

Artificial neural network (ANN) analysis techniques have gained increasing popularity across diverse engineering applications, particularly in the multi-physics modeling of fuel cell systems. However, previous literature lacks a comprehensive comparison of the effectiveness of fundamental nonlinear activation functions in ANN architectures specifically tailored for the mid-sized proton exchange membrane fuel cell (PEMFC). This study investigates the influence of three primary activation functions—the rectified linear unit, logistic sigmoid, and hyperbolic tangent—on the predictive modeling of a 1.2 kWp PEMFC. The ANN model was developed within LabVIEW to simulate voltage across a dynamic operation, including variations in temperature, reactant pressure, and load current. The simulation outputs for each activation function were rigorously validated against empirical data using analysis of variance and comprehensive error analysis, including root mean square error. The results indicate that while the rectified linear unit and hyperbolic tangent models achieved performance within statistically acceptable margins, the logistic sigmoid function demonstrated superior accuracy in capturing the nonlinear voltage–current relationship. Specifically, the sigmoid function provided a more nuanced representation of the internal loss—namely, activation and concentration polarizations—characteristics associated with this scale of PEMFC systems. This study reveals that the superior sensitivity of the sigmoid function makes it the most suitable choice for digital twin diagnostics and real-time control of mid-sized PEMFC stacks.

1. Introduction

ANNs have become increasingly important tools in energy-system modeling because they provide a flexible, data-driven framework for representing complex nonlinear relationships that are difficult to capture using purely physics-based equations. In PEMFCs, the stack voltage is governed by strongly coupled electrochemical kinetics, ohmic resistance, thermal behavior, and mass-transport phenomena. These processes interact across multiple spatial and temporal scales, making analytical modeling challenging, particularly when operating conditions such as load demand, temperature, and reactant supply vary dynamically. An ANN approximates this behavior by learning the functional relationship between measured operating parameters—such as current density, temperature, gas pressure, and flow conditions—and the resulting voltage response. Through iterative training, the network adjusts internal weighting parameters to minimize the discrepancy between predicted and measured voltages, enabling it to infer system behavior directly from experimental data without requiring explicit formulation of electrochemical transport equations.

Because of this capability, ANN-based models are especially useful for fuel-cell applications where rapid prediction is needed, such as real-time performance estimation, fault detection, and system control. They also provide a practical alternative when detailed mechanistic models become computationally intensive or require parameters that are difficult to measure experimentally. Consequently, ANN approaches are increasingly adopted alongside conventional electrochemical models to support performance prediction, operational optimization, and digital monitoring of fuel-cell systems [1-2].

A critical factor influencing the performance of an ANN in fuel-cell modeling is the choice of activation function, as it determines how nonlinear relationships between operating conditions and voltage response are represented within the network. In electrochemical systems, the voltage–current relationship is governed by coupled activation, ohmic, and concentration losses, producing a smooth but highly nonlinear polarization curve. Among commonly used activation functions, the Rectified Linear Unit (ReLU), logistic sigmoid (Sigmoid), and hyperbolic tangent (Tanh) differ in how they approximate such behavior [3]. ReLU introduces a piecewise-linear response that can efficiently capture monotonic trends but may struggle to reproduce smooth electrochemical curvature, particularly in regions where voltage changes gradually with current. The Sigmoid function generates bounded, monotonic outputs, making it well suited for representing the gradual transitions typical of activation and ohmic loss regions. In contrast, the Tanh function produces symmetric outputs around zero, which can be advantageous for centered data but may be less compatible with the strictly positive and asymmetric nature of fuel-cell voltage behavior. Consequently, the selection of

activation function can directly influence how accurately the ANN reproduces the physical characteristics of fuel-cell performance.

Several recent studies have evaluated the performance of these activation functions in fuel cell modeling. For instance, Thi Hoa Le and colleagues [4] conducted a comparative study of ANN models using ReLU, Tanh, and Sigmoid to predict the voltage output of direct urea-hydrogen peroxide fuel cells. They found that the combination of the Sigmoid function and a specific training algorithm resulted in the most accurate voltage predictions. In another study, Pengjie Tian [5] examined the modeling of a 5.0 kWp PEMFC system using different activation functions and observed that the Tanh function yielded the lowest prediction error when evaluating voltage-current characteristics.

Vanja Subotić et al. [6] explored the use of artificial intelligence techniques for predicting and optimizing solid oxide fuel cell (SOFC) performance. Although their study included ReLU, Tanh, and Sigmoid activation functions, the rationale for choosing specific functions was not clearly described. Adithya Legala et al. [7] focused on proton exchange membrane fuel cells and used ReLU in their ANN model to predict cell voltage, achieving a very high correlation between predicted and actual values, indicating good model accuracy under the tested conditions.

While these studies provide useful insights, they primarily focus on either small single-cell models or large commercial-scale stacks, often with more than 350 individual cells. However, none of these studies systematically compares activation functions for mid-sized PEMFC stacks (fewer than 350 cells [8]) using experimental short-stack data. The physical behavior of a mid-sized PEMFC stack (often termed a short stack) differs substantially from both single-cell systems and large commercial stacks. In single cells, current distribution and temperature gradients are often assumed uniform, whereas large stacks exhibit strong manifold interactions, flow maldistribution, and thermal inertia. Mid-sized stacks represent a transitional regime in which these effects begin to emerge but remain weakly coupled. As a result, models developed for either extreme scale often perform poorly when applied to this intermediate configuration. This creates a clear need for dedicated modeling studies focused specifically on mid-sized PEMFC stacks. [9-11].

The present study provides a detailed comparative analysis of the ReLU, Sigmoid, and Tanh activation functions within ANN models for simulating the voltage output of the mid-sized stack. The models were developed using LabVIEW software and trained using experimental data that included variations in current, temperature, and gas pressure. The performance of each activation function was assessed using statistical techniques and evaluated using statistical performance metrics. The findings from this study aim to provide practical guidance for researchers and engineers in selecting suitable activation functions for ANN-based modeling of PEMFC systems, especially in mid-scale applications where data complexity and physical behavior must be accurately represented.

2. Materials and Methods

The study was focused on a 1.2 kWp PEMFC system, a mid-sized PEMFC with 46 cells. The basic information of PEMFC, details of the activation function equations used in ANNs, and the accuracy analysis methods are described as follows.

2.1 Proton Exchange Membrane

The 1.2 kWp Nexa™ power module proton exchange membrane fuel cell (PEMFC), manufactured by Ballard, was employed as the test system in this study (see Fig. 1). This PEMFC operates within a voltage range of approximately 22 to 50 volts and utilizes high-purity hydrogen fuel with a concentration of 99.99%. The fundamental working principle of the PEMFC is based on electrochemical reactions occurring within the cell, as schematically illustrated in Fig. 2. At the anode, hydrogen gas undergoes an oxidation reaction, releasing electrons that generate electrical current. Concurrently, at the cathode, oxygen undergoes a reduction reaction, combining with the electrons returning from the external circuit and positively charged hydrogen ions (protons) transported through the membrane. This overall reaction produces electrical energy and water as byproducts [12-13].

Despite the seemingly straightforward nature of this process, several types of voltage losses occur within the fuel cell, reducing its effective output voltage from the ideal theoretical value. These losses are generally categorized into three main types: activation loss, ohmic loss, and concentration loss. Activation loss refers to the voltage reduction caused by the slow kinetics of the electrochemical reactions at the electrodes, which require a certain amount of energy to overcome reaction barriers. Ohmic loss arises from the internal resistance of the cell components, including the membrane, electrodes, and electrical contacts, which impede the flow of ions and electrons, thereby causing voltage drops. Concentration loss occurs when the reactant gases—hydrogen and oxygen—are depleted near the electrode surfaces, leading to mass transport limitations that hinder the reaction rates [14-15].

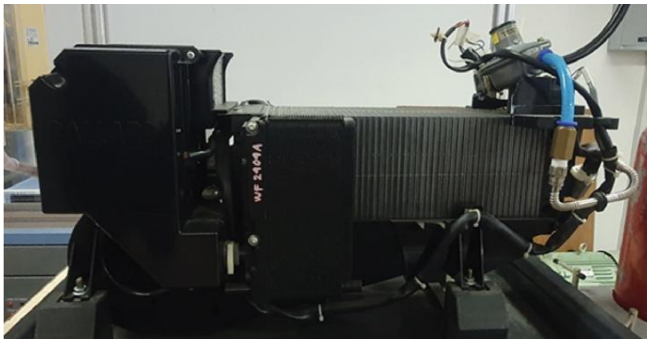


Fig. 1 The 1.2 kWp Nexa™ power module fuel cell.

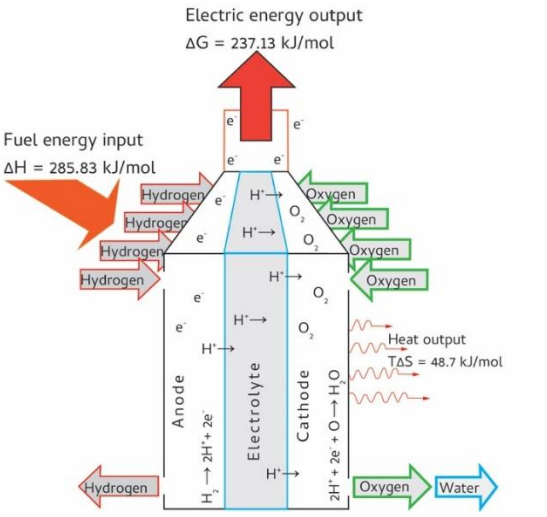


Fig. 2 Thermodynamic operation of fuel cells.

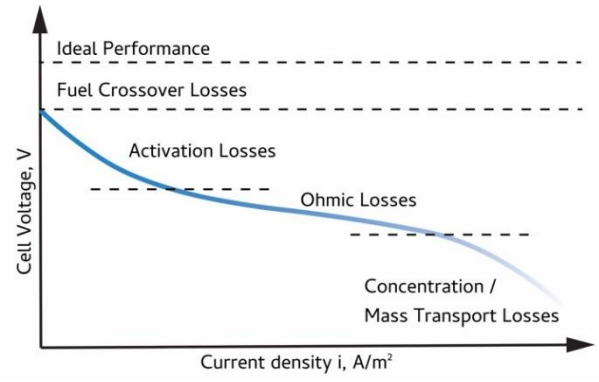


Fig. 3 Polarization curve of fuel cell.

These voltage loss mechanisms are commonly visualized through the polarization curve of the PEMFC, depicted in Fig. 3. This curve demonstrates how the cell voltage decreases with increasing current density due to the combined effects of the different losses. Understanding and modeling these losses are critical for improving PEMFC performance and efficiency, as they directly influence the usable voltage and power output of the system [16-17].

2.2 ANN Mathematical Model

In this study, the artificial neural network (ANN) is trained using a feedforward propagation algorithm with stochastic gradient descent (SGD), in which data flow from the input layer to the output layer through a single hidden layer [18]. The input variables include temperature, hydrogen pressure, and current (x_i, i = 1,…,n), as shown in Fig. 4 and Fig. 5. The learning rate was set to 0.1, and the number of training epochs was set to 10,000. These inputs are processed through the network (with three neurons in the input layer, five in the hidden layer, and one in the output layer, illustrated in Fig. 5), where each node combines the input values with associated weights and passes them through activation (transfer) functions, as shown in Equations (1) and (2). The resulting outputs at single hidden layer (z_j, j = 1,…,p) and the final output layer (y_k, k = 1,…,m), which voltages, are generated based on this layered computation. The network learns by adjusting the weights during training to minimize prediction error, enabling the model to simulate the complex voltage behavior of a PEMFC system.

The activations for the hidden node (z_j) are computed by:

$$z_j = f(h_j) = f\left(\sum_{i=1}^{n} v_{ji}\, z_i + v_{b,i}\right) \quad (1)$$

Here, x_i denotes the study variable of the i^{th} neuron in the input layer, and v_{ji} represents the weight connecting the i^{th} neuron in the input layer to the j^{th} neuron in the hidden layer. $v_{b,j}$ is the bias for the j^{th} hidden node, which initialized at 0.5 in this study. The h_j is total input to the j^{th} neuron.

The activation for the output node (y_k), in this study, it representing voltage, which is computed by:

$$y_k = f(h_k) = f\left(\sum_{j=1}^{p} w_{kj}\, z_j + w_{b,k}\right) \quad (2)$$

Where the h_k is the total input to the k^{th} neuron, the value is calculated as the weighted sum of the inputs from the previous hidden layer plus the bias term ($w_{b,k}$), initially set at 0.5. The z_j denotes the activation of the j^{th} neuron in the hidden layer, and the w^{kj} represents the weight connecting the j^{th} neuron in the hidden layer to the k^{th} neuron in the output layer (as shown in Fig. 6). The weights play a crucial role in adjusting the forecasted value to closely align with the target variable of interest (voltage) [19]. After the weight value is combined with the variable using Equations (1) and (2), the outputs (h_j and h_k) are adjusted to fit nonlinear data using an activation function (f) as shown in Equations (17).

In backpropagation, small initial weights and biases were assigned in LabVIEW, and activations were calculated layer by layer. The error was then used to adjust weights and biases, minimizing the difference between the predicted and target voltage, as illustrated in Fig. 6.

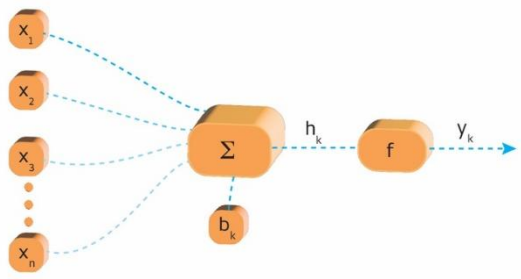


Fig. 4 Working principle of neural network.

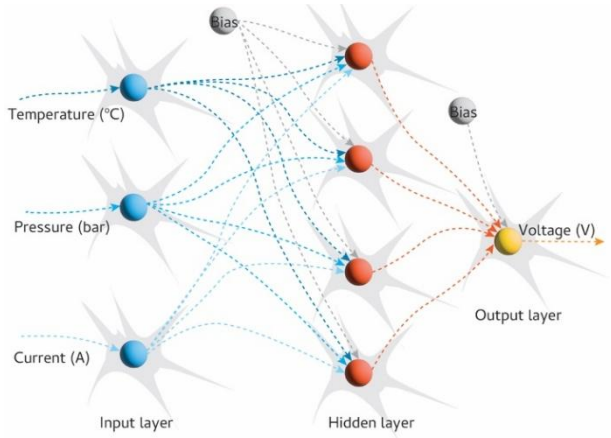


Fig. 5 Forward propagation operation of a (3-5-1) neural networks.

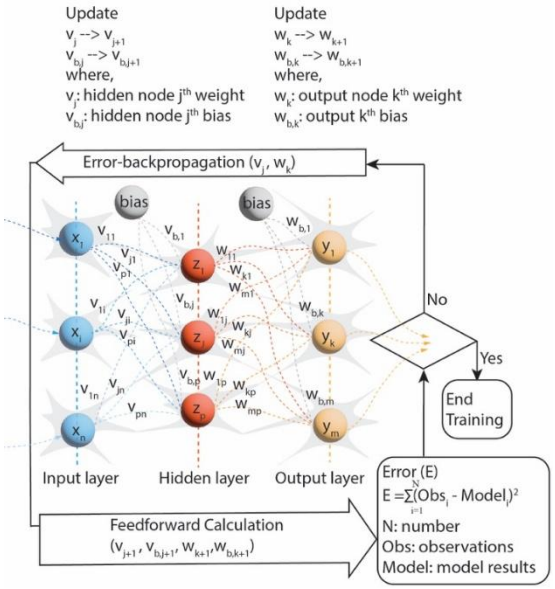


Fig. 6 Backpropagation process of ANN system.

In Figure 6, the error is calculated for both the output and hidden layer. Each node in the output layer (y_k, k=1,…,m) is compared with a target value (t_k) corresponding to the derivative of the activation function $f(h_k)$ with respect to the input of the k^{th} neuron. The error term (δ_k), the change in weight value (ΔW_{kj}), and the change in bias value ($\Delta W_{b,k}$) are then computed and propagated to all nodes in the preceding layer. This process is illustrated in Equations (3), (4), and (5).

$$\delta_k = (t_k - y_k) f(h_k) \tag{3}$$

$$\Delta W_{kj} = \alpha \delta_k z_j \tag{4}$$

$$\Delta W_{b,k} = \alpha \delta_k \tag{5}$$

Here, α represents the learning rate, z_j is the output from j^{th} neuron (hidden node). In the hidden layer, each node (z_j, j=1,…,p) receives the error input ($\delta_{in,j}$) from the output layer, as shown in Equation (6):

$$\delta_{in,j} = \sum_{j=1}^{p} \delta_k \, w_{kj} \tag{6}$$

Finally, Equation (7) calculates the error value (δ_j) for the j^{th} node in the hidden layer:

$$\delta_j = \delta_{in,j} f(h_j) \tag{7}$$

The change in weight (Δv_{ji}) and bias ($\Delta v_{b,j}$) value for the hidden layer are calculated using Equations (8) and (9):

$$\Delta v_{ji} = \alpha \delta_j x_i \tag{8}$$

$$\Delta v_{b,j} = \alpha \delta_j \tag{9}$$

When updating weights and biases, each output node (y_k, k = 1,…,m) adjusts its weight and bias using Equations (10) and (11) (with j = 0,…,p).

$$w_{kj+1} = w_{kj} + \Delta w_{kj} \tag{10}$$

$$w_{b,k+1} = w_{b,k} + \Delta w_{b,k} \tag{11}$$

In each hidden node (z_j, j=1,…,p), the weight (v_{ji}) and bias ($v_{b,j}$) are adjusted using Equation (12) and (13) (with i=0,…,n).

$$v_{ji+1} = v_{ji} + \Delta v_{ji} \tag{12}$$

$$v_{b,j+1} = v_{b,j} + \Delta v_{b,j} \tag{13}$$

Typically, the forecasted values have low tolerances, resulting in high accuracy for predicting the variables on which the study results are based [15].

2.3 Activation Function

The activation function in ANN has been used to generate the weighted sum from input signal nodes into an output signal. Adding the nonlinearity to the model results in the network discovering intricate linkages and patterns. The ANNs cannot represent the intricacy of real-world data without the nonlinearity activation function [20]. The common nonlinearity activation functions in ANNs are Sigmoid, Tanh, and ReLU, which are explained below.

2.3.1 Sigmoid Function

The Sigmoid function is a mathematical function often used in artificial neural networks (ANNs) for studying the nonlinearity behavior data [20]. It is an "S"-shaped curve and maps any input value to a value between 0 and 1, then converts to the voltage value at the final calculation, making it useful in PEMFC problems. The mathematical form of the Sigmoid function is:

$$f(h_{j,k}) = \frac{1}{1 + e^{-h_{j,k}}} \tag{14}$$

Where $h_{j,k}$ in Equation (14) is the input value at the position j,k (possibly a hidden layer activation or some input in a neural network), e is approximately equal to 2.71828 [21].

2.3.2 Tanh Function

The Tanh (hyperbolic tangent) function in Equation (15) is a common activation function that is used in neural networks. The formula for the Tanh function displays as:

$$f(h_{j,k}) = \tanh(h_{j,k}) \frac{(e^{h_{j,k}} - e^{-h_{j,k}})}{(e^{h_{j,k}} + e^{-h_{j,k}})} \tag{15}$$

The formula involves exponentials $e^{hj,k}$ and $e^{-hj,k}$, which represent the positive and negative exponential terms, respectively [22]. It maps input values to a range between -1 and 1, the function gives zero centered output, facilitating easier learning for subsequent layers. For this study, the final result has to be converted to the voltage value.

2.3.3 ReLU Function

ReLU is a nonlinearity activation function used in artificial neural networks, enabling it to model complex data patterns. ReLU is defined as:

$$f(h_{j,k}) = \max(0, h_{j,k}) = \begin{cases} 0 \; for \; h_{j,k} \leq 0 \\ h_{j,k} \; for \; h_{j,k} > 0 \end{cases} \tag{16}$$

If the value of $h_{j,k}$ in Equation (16) is less than or equal to 0, the output of the function is 0. If the value of $h_{j,k}$ is greater than 0, the output is simply $h_{j,k}$ itself, meaning that the function only outputs non-negative values. The output voltage obtains by converted the value of function [23]. ReLU is computationally more efficient than Tanh and Sigmoid, as it relies on simpler mathematical operations and avoids the exponential calculations required by the other functions. At a time only a few neurons are activated making the network sparse making it efficient and easy for computation.

2.4 Error Analysis

The three ANN models, incorporating Sigmoid, Tanh, and ReLU activation functions as described above, were systematically compared to determine which function provides the most accurate voltage predictions under varying operating conditions of the 1.2 kWp PEMFC. Voltage outputs at different gas pressures were plotted against changes in electrical current to visually assess each model's response across a wide range of realistic scenarios. To quantitatively evaluate model performance, residual error analysis and the Q-Q (Quantile-Quantile) plots were employed. These graphical tools helped assess how well the models fit the experimental data and whether their prediction errors were normally distributed, thereby indicating the validity of underlying assumptions.

A residual plot is a graphical tool used to evaluate the difference between observed (experimental) values and predicted values from a model. It displays residuals (calculated as the observed voltage value minus the predicted voltage value) against predicted values or an independent variable. Ideally, residuals should be randomly scattered around zero without any systematic patterns. Such randomness indicates that the model's errors are evenly distributed and that the model fits the data well. Patterns or trends in the residual plot may suggest issues like nonlinearity, heteroscedasticity (non-constant variance), or model misspecification.

A Q–Q plot is a graphical method to assess whether a dataset follows a particular theoretical distribution, most commonly the normal distribution. It plots the quantiles of the residuals against the quantiles of the reference distribution. If the residuals follow the specified distribution, the points will approximately lie along a straight reference line. Deviations from this line indicate departures from the assumed distribution, which can affect the validity of statistical tests and model assumptions. Together, residual and Q–Q plots provide complementary information on the adequacy of a model's fit and the assumptions underlying the statistical analysis, helping this research identify potential issues that may require model refinement.

In addition to visual analysis, standard error metrics (including mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) as shown in Equations (17) – (20) [24]) were calculated to provide objective measures of prediction accuracy. These metrics allowed a detailed comparison of the strengths and weaknesses of each activation function in capturing the complex nonlinear behavior of the PEMFC voltage response. The comprehensive evaluation revealed the specific conditions where each activation function excels or underperforms, offering valuable insights into their suitability for PEMFC modeling applications.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_t^s - Y_t^a) \tag{17}$$

RMSE provides the error in the same units as the predicted value (voltage, V, in this study) and is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_t^s - Y_t^a)} \tag{18}$$

MAE, calculated as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_t^s - Y_t^a| \tag{19}$$

and MAPE, which expresses the error as a percentage of actual values: were used alongside RMSE to evaluate model performance. Y_t^s, Y_t^a represent the predicted and actual values, respectively, and n is the number of data points.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|Y_t^s - Y_t^a| \times 100\% \tag{20}$$

R and R^2 (in Equation (21-22)) were also used to evaluate accuracy, where R is defined as:

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{21}$$

and R^2 as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{22}$$

where SS_{res} and SS_{tot} are defined by:

$$\mathrm{SS_{res}} = \sum_{\mathrm{i=1}}^{\mathrm{m}}(\mathrm{y_i} - \bar{y}_i)^2 \tag{23}$$

$$SS_{tot} = \sum_{i=1}^{m}(y_i - \bar{y}_i)^2 \tag{24}$$

where y_i and $\overline{y_i}$ in Equation (23-24) is observed voltage values and predicted voltage values by ANN with activation function, respectively. Furthermore, ANOVA was used to compare standard deviation, variance, and variation between experimental and model data, with significance determined at $p \leq 0.05$ using SPSS [25].

Analysis of Variance (ANOVA) is a statistical method used to compare the means of three or more groups to determine whether at least one group differs significantly from the others [26]. In this modeling study, ANOVA was applied to evaluate whether the predicted voltages generated by different activation functions (Sigmoid, Tanh, ReLU) significantly differ from each other and from the experimental data. The input dataset includes voltage predictions under varying operational conditions (current, temperature, pressure). ANOVA helps test the null hypothesis that the mean errors (e.g., residuals or absolute deviations) of the models are equal across all activation functions. If the ANOVA test reveals a statistically significant difference (typically $p < 0.05$), the Tukey's HSD post hoc analysis is conducted to identify which specific activation function pairs differ significantly.

3. PEMFC Experiment and Modelling

3.1 Experiment Data

The experimental operation of the 1.2 kWp PEMFC was carried out at the Department of Electrical Engineering, Faculty of Engineering, Rajamangala University of Technology Thanyaburi. Simulation and data analysis were subsequently performed at the Department of Physics, Faculty of Science, Mahasarakham University. The PEMFC was tested under varying hydrogen pressures ranging from 1 to 9 bar and electrical loads between 1 and 40 A. Voltage output (V_Exp) and cell temperature were continuously monitored using a built-in fuel cell controller. The experimental setup enabled the collection of diverse operating conditions to support model training and validation. The main components of the system, including the stack, control unit, and measurement devices, are illustrated in Fig. 7.

3.2 PEMFC Simulation

In the ANN modeling process developed using LabVIEW, the experimental input parameters include gas pressure (P), current (A), and temperature (T), while the model output is the predicted stack voltage (V). These variables are visually displayed on the front panel, enabling real-time monitoring and user interaction, whereas the computational operations of the artificial neural network are implemented within the block diagram, where data flow and algorithmic logic are executed.

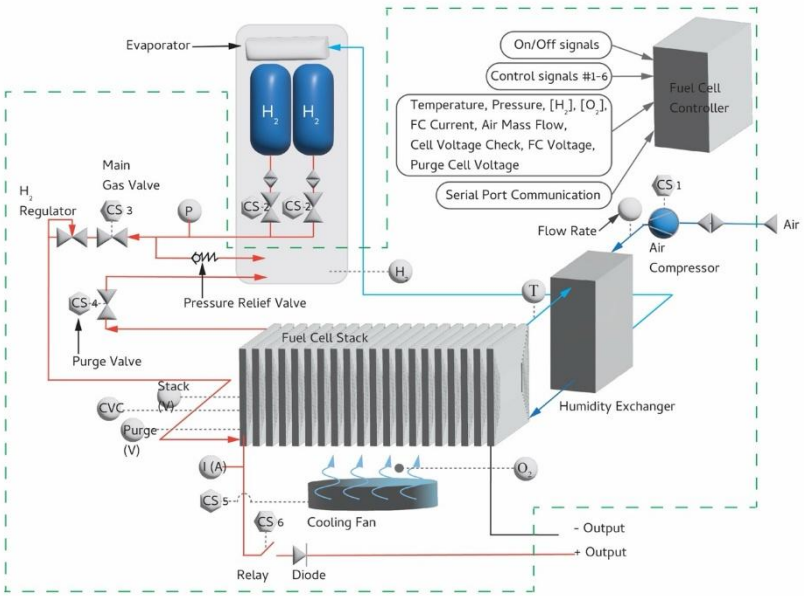


Fig. 7 Schematic diagram of 1.2 kWp Nexa™ power module PEMFC.

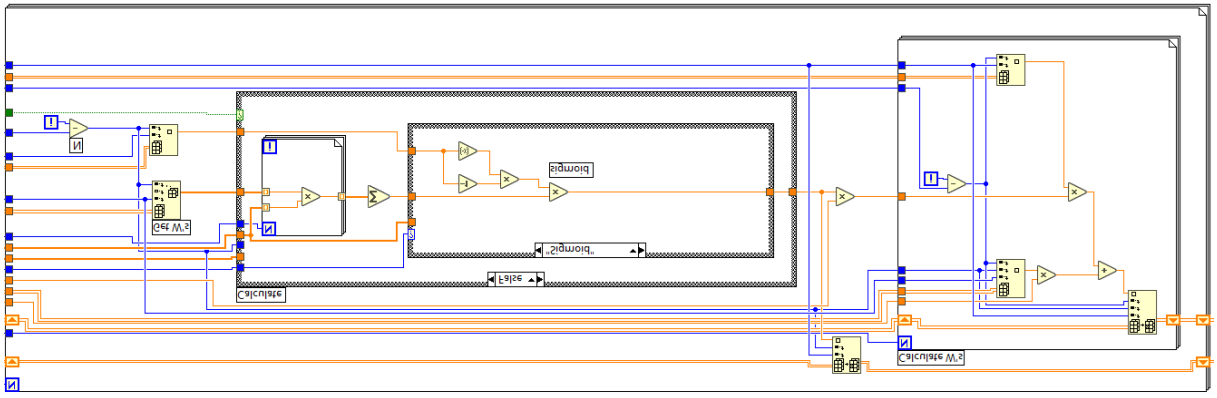


Fig. 8 The block diagram in the weighted adjustment process in LabVIEW.

Table 1 ANN architecture and training hyperparameters used in this study.

Category	Parameter	Symbol	Value
Model structure	ANN type	–	Feedforward ANN with backpropagation
	Number of input nodes	(n_{in})	3 (P, A, Tem)
	Number of hidden layers	(L_h)	1
	Neurons in hidden layer	(p)	5
	Number of output nodes	(n_{out})	1 (stack voltage, (V))
Activation function	Activation functions evaluated	–	Sigmoid, Tanh, ReLU
Initialization	Hidden-layer bias (initial)	($v_{b,j}$)	0.05
	Output-layer bias (initial)	($w_{b,k}$)	0.05
Training	Learning rate	(α)	0.1
	Momentum	(μ)	0.8
	Training cycles / epochs	(N_{epoch})	10,000
	Stopping criterion	–	Error tolerance = ± 1%
Preprocessing	Standardization	–	Mean–variance standardization (Batch Norm concept)

The ANN is trained using a feedforward propagation algorithm with stochastic gradient descent (SGD) and consists of a single hidden layer. The network architecture comprises three neurons in the input layer, five neurons in the hidden layer, and one neuron in the output layer. The available dataset is divided into training, validation, and testing subsets in a ratio of 60:20:20. A learning rate of 0.1 is employed, and the network is trained for 10,000 epochs. During training, the input dataset is multiplied by weighted connections and passed through selected activation functions—Sigmoid, Tanh, or ReLU—depending on the configuration under evaluation. For the Tanh activation function, which has a limited output range of −1 to 1, input and output data are scaled prior to training to prevent optimization difficulties and reduce prediction errors. The processed signals propagate to the output layer to generate the predicted stack voltage. A predefined tolerance level is used to evaluate the error between predicted and experimental values; if the error exceeds this threshold, backpropagation is applied to iteratively update the weight values. This training procedure is illustrated in Fig. 6, while Fig. 8 presents the detailed block diagram and weight update mechanism. The ANN architecture and training hyperparameters used in the study represents in Table 1.

4. Results and Discussion

Figure 9 shows the 2D plot of the polarization curve of the PEMFC at each pressure (1-9 bars), where the X and Y axes are current and voltage, respectively. The data reveal that there is a strong negative correlation between the PEMFC stack voltage and both temperature and current. The experiment voltage (V_Exp) variations are minimal during the activation loss phase, occurring at the low current, but become more prominent during the concentration loss phase, at higher current levels. The voltage in the activation losses region has high values (35-42 V) and low variation due to slow kinetics reaction, especially oxygen reduction at the cathode, at low current (1-10 A), as well as the system being under less stress of reactant supply, water management, and thermal condition are stable. In the Ohmic losses region, the voltage (29-36 V) has moderate fluctuation while operating in the medium current (10-30 A) due to the internal resistance caused by some thermal and hydration effects [27]. The high voltage fluctuations observed in the concentration loss region are primarily caused by increased reactant demand (particularly oxygen) when the cell operates under high current (30–40 A) and low voltage (25–32 V) conditions. At this stage, the system enters a mass transport-limited regime, where reactants cannot reach the catalyst sites quickly enough. As a result, voltage drops sharply due to factors such as oxygen starvation, pressure loss, water flooding, or membrane dry-out.

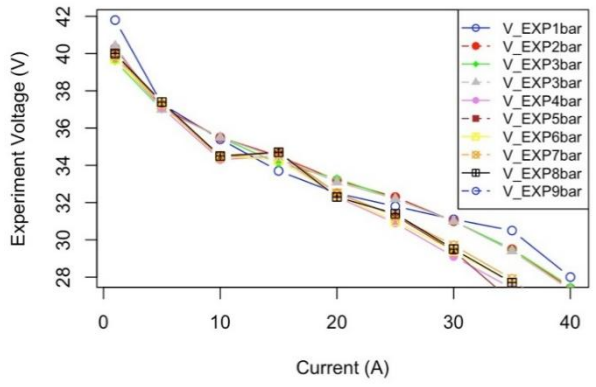


Fig. 9 The 2D plot of the polarization curve of the PEMFC at each pressure (1-9 bars).

Pressure has a relatively smaller effect on voltage, with slightly higher values observed at a group of lower pressures (1–3 bars) rather than at a group of higher pressures (5–9 bars), especially during the concentration losses region. Although increasing hydrogen pressure is generally expected to enhance PEMFC performance by improving reactant availability and reducing concentration losses, the observed voltage reduction at higher pressures (5–9 bar) in the Nexa 1.2 kWp stack suggests that secondary transport and water-management effects become dominant in this particular configuration. In compact, mid-sized stacks such as the Nexa system, higher inlet pressure can increase gas residence time and reduce effective purge efficiency, promoting liquid water accumulation within the gas diffusion layer and flow channels. This localized flooding restricts reactant access to the catalyst layer, thereby increasing mass-transport resistance despite the higher nominal reactant pressure. Additionally, elevated pressure may alter membrane hydration and compressive stress on the gas diffusion media, reducing porosity and limiting diffusive transport pathways. The combined effects of flow-field design, stack compression, and purge control strategy therefore appear to shift the system into a regime where excess pressure degrades rather than improves performance. This behavior is consistent with practical stack operation, where optimal pressure represents a balance between reactant supply and effective water removal rather than a simple monotonic performance increase. Pressure also affects the PEMFC's temperature distribution, especially during the ohmic losses. In that range, the maximum voltage is often found at 4 bars, the point at which PEMFC has reached maximum temperature for each current level.

Temperature variability within the PEMFC stack is more prominent at lower gas pressures (1–4 bar) but becomes more stable as pressure increases (5–9 bar). Temperature varies with pressure and is responsive to current and voltage, especially during activation loss and ohmic losses. In the activation losses, the temperature effect is very strong due to the temperature acceleration rates (Arrhenius behavior) [28]. The higher temperature reduces activation energy, increasing exchange current density. As a result, voltage increases noticeably with temperature.

In the Ohmic losses region, temperature has a moderate to strong effect. The membrane proton conductivity increases with temperature, causing reduced resistance. The voltage still increases with temperature, though the value is not as steep as in the activation loss region. In the concentration losses region, the temperature effect on PEMFC behavior is weak. The temperature does little to improve gas diffusion compared to how it improves kinetics or conductivity. The voltage drops sharply at high current due to oxygen starvation, regardless of temperature. So, the temperature dependence in this region is weak.

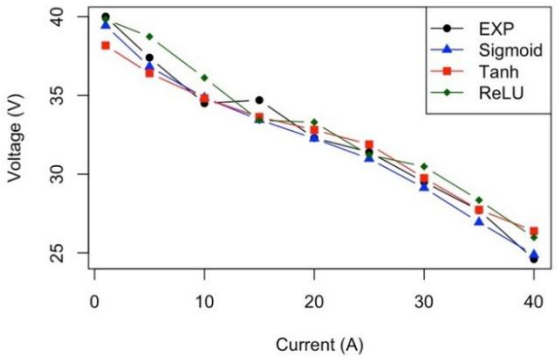


Fig. 10 The 2D plot of test data and experiment data at constant pressure (9 bar).

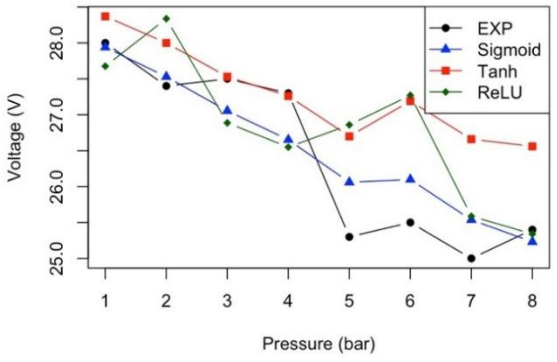


Fig. 11 The 2D plot of test data and experiment data at constant current (40 A).

4.1 Role of Sigmoid Function on PEMFC

In the Artificial Neural Network (ANN) model, the Sigmoid activation function produces voltage output data (V_sigmoid) ranging from approximately 24 to 40 volts, as illustrated in Fig. 10. The 2D plot demonstrates that the S-shaped curve inherent to the Sigmoid function effectively captures the nonlinear transitions between the different voltage loss phases observed in the experimental data. During the activation loss phase, where slow electrochemical kinetics dominate, the model exhibits minimal voltage variation, closely matching the experimental observations. This suggests that the Sigmoid function accurately models the sluggish voltage response typical of the early start-up conditions in a proton exchange membrane fuel cell (PEMFC).

Furthermore, the Sigmoid function performs well in representing the linearized voltage drop observed in the ohmic loss region. This region is characterized by internal resistance effects associated with the PEMFC's membrane, gas diffusion layers, humidifier components, and bipolar plate materials. The accurate prediction in this region indicates that the ANN model, when using the Sigmoid activation function, can effectively reflect the balance and interaction of water management, thermal conductivity, and material performance within the cell. Notably, the Sigmoid function successfully distinguishes between low-pressure and high-pressure data clusters and accurately represents the voltage deviations near the 4-bar critical point (as illustrated in Fig. 11). This capability underscores the Sigmoid function's strength in inferring detailed PEMFC behavior across all operational phases, making it the most suitable activation function among those tested for modeling the complex and interdependent electrochemical dynamics of a 1.2 kWp PEMFC stack.

4.2 Role of ReLU Function on PEMFC

Figure 10 illustrates the voltage output predicted by the artificial neural network (ANN) model using the ReLU activation function (V_ReLU), as a function of current and pressure. The scatter plot reveals that the ReLU-based model exhibits a greater degree of data dispersion compared to the Sigmoid function, particularly during the ohmic and concentration loss regions. While the ReLU model demonstrates reasonable predictive accuracy during the activation loss phase and at the onset of the concentration loss phase (comparable to the performance of the Sigmoid function), it fails to maintain this level of accuracy throughout the entire operational range of the fuel cell. This behavior is especially evident in the ohmic loss region, where accurate voltage prediction is critical for estimating internal resistance and membrane proton conductivity. The elevated errors in this region indicate that the ReLU-based ANN is less effective than the Sigmoid in capturing the underlying electrochemical and thermal dynamics of the PEMFC.

Moreover, the ReLU function appears to introduce a notable imbalance in voltage response under conditions where mass transport limitations dominate, particularly during the concentration loss phase. This imbalance may reflect the model's difficulty in accounting for variations in oxygen supply and reactant diffusion, which are critical for reliable voltage prediction under high-load or low-humidity conditions. As a result, although the ReLU activation function can offer computational advantages due to its simplicity and non-saturating nature, it demonstrates limitations in modeling the nonlinear and interdependent phenomena inherent to PEMFC operation. These findings underscore the importance of activation function selection in developing robust, physics-informed ANN models for fuel cell systems.

4.3 Role of Tanh Function on PEMFC

The results of applying the Tanh activation function within the ANN model (used to represent the relationship between output voltage (V_tanh) and key operational variables, i.e. pressure and current) are illustrated in Fig. 10-11. The plotted trends reveal that the Tanh function generates a negative correlation between voltage and current, which is consistent with the patterns observed in the Sigmoid-based model. However, despite this similarity, the voltage predictions derived from the Tanh function show poor alignment with the experimental data. Notably, the model exhibits the largest deviations and outliers among the three tested activation functions. These discrepancies persist across all critical loss regions of the fuel cell operation, namely activation loss and concentration loss phases. In addition, the Tanh function introduces significantly higher variability in voltage with respect to changes in pressure and temperature during the ohmic loss region (as illustrated in Fig. 11). This behavior deviates markedly not only from the experimental trends but also from the comparatively stable predictions generated by the Sigmoid and ReLU models. Such inconsistency suggests that the Tanh function lacks the robustness required to reliably capture the internal resistance characteristics of this 1.2 kWp PEMFC stack.

This limitation is especially prominent when considering the influence of thermal effects and membrane hydration dynamics, which are critical to accurate voltage prediction in proton exchange membrane fuel cells. Consequently, while the Tanh activation function may offer mathematical smoothness, its inability to represent the complex nonlinear interactions between variables in this application renders it less effective than its counterparts. These findings highlight the importance of selecting an activation function that can not only generalize well but also capture the physical realities of electrochemical processes within PEMFC systems.

4.4 Statistical Comparative

In the statistical evaluation of the ANN models developed for voltage prediction in a 1.2 kWp PEMFC system, two primary graphical diagnostic tools (Residuals vs Fitted plots and Quantile-Quantile (Q-Q) plots) were employed to assess model accuracy, distribution of errors, and reliability of predictions. These tools provide visual insights into how well the models perform under varying conditions and offer evidence of any systematic deviation from modeling assumptions.

The Residuals vs Fitted plots, presented in Fig. 13(a–c), were used to visualize the residual errors, which are calculated as the difference between experimentally measured voltages and those predicted by the ANN models. These plots were specifically generated for three different activation functions applied in the ANN framework: Sigmoid, Tanh, and ReLU. The fitted values (predicted voltages) are plotted along the x-axis, and the corresponding residuals are on the y-axis. A well-fitted model would ideally show residuals scattered randomly around the horizontal zero line, indicating homoscedasticity (constant variance of the residuals) and the absence of systematic patterns.

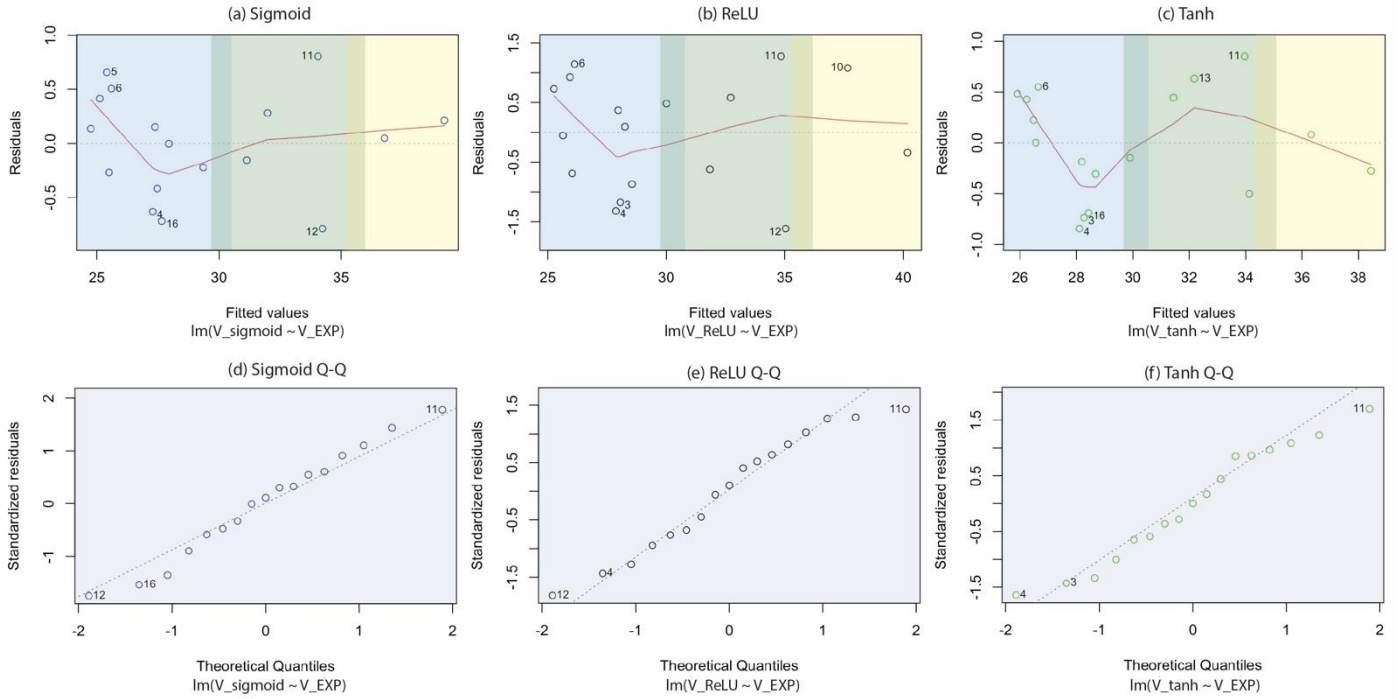


Fig. 13 The residuals plot against V_Exp: (a) V_sigmoid (b) V_ReLU (c) V_tanh and the Q-Q plot against V_Exp: (d) V_sigmoid (e) V_ReLU (f) V_tanh of the test data.

Figure 13a, which corresponds to the model using the Sigmoid activation function, shows a red line that remains nearly flat and closely aligned with the zero residual axis across the full voltage range. This horizontal pattern is particularly visible within the activation loss region (approximately 25–32 V), the ohmic loss region (29–36 V), and the concentration loss region (35–42 V). This behavior suggests that the Sigmoid-based ANN model achieves better prediction accuracy and consistency across all loss phases of the PEMFC operation. In contrast, Fig. 13b and 13c, which represent the ReLU and Tanh functions respectively, show residual lines that deviate more significantly from the horizontal axis. These deviations reflect potential systematic errors or inconsistencies in prediction accuracy, especially in specific voltage ranges corresponding to different physical loss mechanisms in the fuel cell. Diagnostic analysis of the ANN models shows that the sigmoid activation function provides the most reliable representation of the PEMFC voltage response. Its residuals are randomly distributed with near-normal behavior, indicating that the nonlinear electrochemical characteristics of the fuel cell are well captured. In contrast, the ReLU model exhibits systematic voltage-dependent deviations and non-normal residuals, reflecting limitations in representing the smooth polarization curve. The tanh model demonstrates intermediate performance but still presents structured bias. Overall, the sigmoid-based ANN offers the most robust predictive framework for the 1.2 kWp PEMFC system.

To further assess the distribution of residuals, Q-Q plots were used, as shown in Fig. 13(d–f). These plots compare the standardized residuals against the theoretical quantiles of a standard normal distribution. Ideally, the points should align along a 45-degree reference line if the residuals are normally distributed. The Q-Q plot for the Sigmoid-based model (Fig. 13d) shows residuals that align closely with the reference line, indicating that the residuals are approximately normally distributed. This alignment supports the statistical soundness and robustness of the Sigmoid-based ANN model for predicting PEMFC voltage. However, the plots for the ReLU (Fig. 13e) and Tanh (Fig. 13f) models demonstrate significant deviations from the reference line, particularly at the distribution tails. This suggests the presence of skewness, outliers, or non-normal distribution in those models, which could compromise the reliability of their predictions.

To statistically confirm the differences observed among the models, a one-way Analysis of Variance (ANOVA) was performed, followed by Tukey's post hoc test. Table 2 presents the ANOVA results using a significance threshold of $p \leq 0.05$. The analysis revealed no statistically significant difference between the experimental voltages and those predicted by the Sigmoid, Tanh, and ReLU models ($F = 0.083$, $p = 0.969$), indicating that all models reproduced the overall voltage level consistently.

Descriptive statistics in Table 3 further support this finding, showing that the mean predicted voltages were very close to the experimental values, ranging from 29.461 to 30.113 V. In addition, Tukey's post hoc indicates that the voltage simulated using the Sigmoid function, with a mean of 29.461 V, is the closest to the experimental result ($p = 0.999$). The Tanh function yields the second-closest result, also with a mean of 29.992 V (with $p = 0.995$). The largest deviation is observed with the ReLU function, which has a higher mean voltage of 30.113 V and a p-value of 0.988.

Although the post hoc comparison indicated that the Tanh predictions were statistically closer to the experimental mean voltage than the ReLU predictions (higher p-value), this does not imply superior predictive accuracy. Among the tested activation functions, the Sigmoid model consistently achieved the lowest prediction errors (MAE = 0.443, RMSE = 0.564 V, MAPE = 1.40%) and the highest correlation with experimental data ($R = 0.991$, $R^2 = 0.982$), as shown in Table 4. These results indicate that Sigmoid most effectively captured the nonlinear voltage response of the fuel cell. In contrast, Tanh produced the largest errors, possibly due to its symmetric output range, which is less compatible with strictly positive voltage data. ReLU showed intermediate performance, suggesting that its piecewise-linear structure may not fully represent the smooth electrochemical transitions of the polarization curve.

Table 2 Variance analysis results to compare differences in forecasted voltage values.

Voltage	Variance sources	SS	df	MS	F	P-value
Voltage 4 ways	Between groups	4.812	3	1.604	.083	.969
	Within a group	1238.695	64	19.355		
	combine	1243.507	67			

Table 3 Post Hoc analysis of voltage averages.

Voltage	$\overline{X}$	EXP	Sigmoid	Tanh	ReLU
$\overline{X}$		29.617	29.461	29.992	30.113
EXP	29.617		-.156 (0.999)	.375 (.995)	.495 (.988)
Sigmoid	29.461			.531 (.985)	-.651 (.973)
Tanh	29.992				-.120 (0.999)
ReLU	30.113				

Table 4 The results of the forecast accuracy analysis.

Statistical	Sigmoid	ReLU	Tanh
MAE	0.443	0.696	0.808
MSE	0.318	0.698	0.982
RMSE (V)	0.564	0.835	0.991
MAPE (%)	1.40	2.20	2.50
R	0.991	0.980	0.978
R-Square	0.982	0.961	0.957

Although Sigmoid and Tanh exhibit similar S-shaped nonlinearities, their performance differs due to their output ranges and interaction with fuel-cell voltage data. Stack voltage is strictly non-negative and varies within a limited operating window; thus, the Sigmoid function (0–1) is more naturally compatible with scaled voltage targets, resulting in more stable convergence [29-30]. In contrast, the zero-centered output of Tanh (−1 to 1) can introduce unnecessary sign changes in the error signal when predicting positive voltages, increasing optimization difficulty [31]. Despite normalization, Tanh is more prone to saturation near its bounds, leading to vanishing gradients and reduced learning efficiency under stochastic gradient descent with a fixed learning rate [32]. These effects collectively explain the higher MAE and MSE observed for Tanh compared with Sigmoid in this study.

The superior performance of the Sigmoid function can be attributed to its ability to effectively capture the nonlinear behavior associated with the three major voltage loss mechanisms in PEMFCs—activation, ohmic, and concentration losses. This capacity enables the model to produce voltage predictions that closely align with experimental observations. Residual error analysis supports this finding, demonstrating that the sigmoid-based ANN consistently delivers the most accurate results, with the smallest deviation from experimental values. Furthermore, Q–Q (quantile–quantile) plots of the residuals confirm the normality of the error distribution, particularly for the sigmoid-based model, whose residuals most closely follow the reference line. This indicates not only improved prediction accuracy but also a well-calibrated and unbiased error pattern.

Despite the performance differences observed, statistical analysis reveals no significant variation in predicted voltages among the three activation functions. With a p-value of 0.969, the results suggest that all models generate voltage outputs that are statistically comparable to the experimental data. Nonetheless, the sigmoid function demonstrates superior predictive accuracy, consistent with findings by Pengjie Tian [5], who reported similar results using a 1.5 kWp PEMFC system. It should be noted that humidity data were not included in this study. Given that humidity significantly influences both temperature and voltage, particularly during the concentration loss phase [22], and has a notable impact on overall PEMFC performance [23], future work should incorporate humidity as a critical input parameter to enhance model fidelity and generalizability.

Overall, the combination of residual analysis, Q-Q diagnostics, and statistical testing provides a comprehensive evaluation of model performance. The results clearly demonstrate that the ANN model using the Sigmoid activation function is more effective than the ReLU and Tanh function in modeling the complex voltage behavior of the 1.2 kWp PEMFC across all loss regions, offering a valuable predictive tool for the mid-size PEMFC system design and optimization.

5. Conclusion

This study evaluated the influence of activation functions in an artificial neural network (ANN) model for simulating the voltage of a 1.2 kWp proton exchange membrane fuel cell (PEMFC) stack using experimental data. Statistical analysis showed no significant differences in mean voltage level among the Sigmoid, Tanh, and ReLU models, indicating that all activation functions reproduced the overall stack voltage consistently. However, accuracy metrics demonstrated that the Sigmoid function provided the most reliable predictions, particularly under varying current, temperature, and gas pressure conditions. Its monotonic S-shaped response enabled the network to better capture the nonlinear transitions associated with activation, ohmic, and concentration losses, resulting in closer agreement with the experimental polarization behavior. In contrast, ReLU and Tanh showed reduced predictive accuracy, especially at higher current densities, suggesting limitations in representing the smooth electrochemical voltage curvature typical of PEMFC operation.

Although humidity is known to strongly influence PEMFC performance through its effects on membrane hydration, proton conductivity, and water management, the model remained robust because the primary operating inputs implicitly reflected hydration effects under steady-state conditions. Nevertheless, excluding humidity may reduce the model's ability to capture localized nonlinearities in the ohmic and concentration-loss regions, where water transport plays a dominant role. Incorporating humidity measurements in future studies would therefore likely improve predictive accuracy and help distinguish between limitations arising from input data and those related to network architecture. Overall, this work provides practical guidance for selecting activation functions in ANN-based modeling of mid-sized PEMFC stacks and underscores the importance of aligning neural-network design with the underlying electrochemical behavior governing fuel-cell performance.

6. Nomenclature

Notations	Description
x_i	Variable of i^{th} neuron in the input layer
y_k	Activations of k^{th} neuron at output layer
z_j	Activations of j^{th} neuron at hidden layer
v_{ji}	Weight connecting i^{th} input neuron to j^{th} hidden neuron
$v_{b,j}$	Bias for hidden node j^{th}
h_j	Total input to j^{th} neuron at hidden layer
h_k	Total input to k^{th} neuron at output layer
$h_{j,k}$	The input value at the position j,k
$w_{b,k}$	Bias for output layer k^{th}
w_{kj}	Weight for the hidden layer j^{th} to k^{th} neuron in the output layer
t_k	target value
δ_k	Error value k^{th} of output layer
$\delta_{in,j}$	Error input from higher (output) layer
δ_j	Error value for node j^{th} in hidden layer
Δv_{ji}	The change in weight value form hidden layer
$\Delta v_{b,j}$	The change in bias value form hidden layer
ΔW_{kj}	The change in weight value form output layer
$\Delta W_{b,k}$	The change in bias value form output layer
A	The learning rate
w_{kj+1}	The updating weights for output layer
$w_{b,k+1}$	The updating bias for output layer
v_{ji+1}	The updating weights for hidden layer
$v_{b,j+1}$	The updating bias for hidden layer
Y_t^s	The estimated value from the model
Y_t^a	The actual data value
N	Number of data points used to estimate model
m	Total number of data points in dataset
SS_{res}	The residual sum of squares
SS_{tot}	The total sum of squares
x_i, y_i	the i^{th} value of the independent variable x and y
$\overline{x}$, $\overline{y}$	the mean of the x values and y values

Acknowledgements

We gratefully acknowledge Assoc. Prof. Dr. Boonyang Plangklang, Renewable Energy System Research Center, Faculty of Engineering, RMUTT, for providing experimental data and valuable resources. We also sincerely thank Mahasarakham University for its financial and facility support under Grant No. 6604001, which significantly contributed to the successful completion of this research.

References

[1] Kumar, L. A., Renuka, D. K., Rose, S. L., Shunmuga priya, M. C. and Wartana, I. M., Deep learning based assistive technology on audio visual speech recognition for hearing impaired. *International Journal of Cognitive Computing in Engineering.* 3 (2022) 24-30, doi: https://doi.org/10.1016/j.ijcce.2022.01.003.

[2] Jackulin, C. and Murugavalli, S., A comprehensive review on detection of plant disease using machine learning and deep learning approaches. *Measurement: Sensors*. 24 (2022) 100441, doi: https://doi.org/10.1016/j.measen.2022.100441.

[3] Bouraya, S. and Belangour, A., A comparative analysis of activation functions in neural networks: unveiling categories. *Bulletin of Electrical Engineering and Informatics.* 13 (2024) 3301-3308, doi: https://doi.org/10.11591/eei.v13i5.7274.

[4] Le, T. H., Thakur, D. and Nguyen, P. K. T., Modeling and optimization of direct urea-hydrogen peroxide fuel cell using the integration of artificial neural network and bio-inspired algorithms. *Journal of Electroanalytical Chemistry*. 922 (2022) 116783, doi: https://doi.org/10.1016/j.jelechem.2022.116783.

[5] Tian, P. *Performance prediction of PEM fuel cell using artificial neural network machine learning.* Report No. 28022774, University of California, Irvine, (2020) 1-52.

[6] Subotić, V., Eibl, M. and Hochenauer, C., Artificial intelligence for time-efficient prediction and optimization of solid oxide fuel cell performances. *Energy Conversion and Management*. 230 (2021) 113764, doi: https://doi.org/10.1016/j.enconman.2020.113764.

[7] Legala, A., Zhao, J. and Li, X., Machine learning modeling for proton exchange membrane fuel cell performance. *Energy and AI*. 10 (2022) 100183, doi: https://doi.org/10.1016/j.egyai.2022.100183.

[8] Obermaier, M., Rauber, M., Bauer, A., Lochner, T., Du, F. and Scheu, C., Local Fuel Starvation Degradation of an Automotive PEMFC Full Size Stack. *Fuel Cells*. 20 (2020) 394-402, doi: https://doi.org/10.1002/fuce.201900180.

[9] Al-Baghdadi, M., Three-dimensional solid mechanics-CFD modeling of a PEM fuel cell stack. *International Journal of Energy and Environment*. 9 (2018) 1-26.

[10] Mastropasqua, L., Campanari, S. and Brouwer, J., Solid oxide fuel cell short stack performance testing - Part A: Experimental analysis and μ-combined heat and power unit comparison. *Journal of Power Sources.* 371 (2017) 225-237, doi: https://doi.org/10.1016/j.jpowsour.2017.10.028.

[11] Zhou, Y., Lin, G., Shih, A. J. and Hu, S. J., Multiphysics Modeling of Assembly Pressure Effects on Proton Exchange Membrane Fuel Cell Performance. *Journal of Fuel Cell Science and Technology*. 6 (2009) 1-7, doi: https://doi.org/10.1115/1.3081426.

[12] Tochigi, S. and Dowaki, K., Environmental Impact Assessment of PEM Fuel Cell Combined Heat and Power Generation System for Residential Application Considering Cathode Catalyst Layer Degradation. *Energies.* 16 (2023) 1985, doi: https://doi.org/10.3390/en16041985.

[13] Islam, M. S., Shudo, Y. and Hayami, S., Energy Conversion and Storage in Fuel Cells and Super-Capacitors from Chemical Modifications of Carbon Allotropes: State-of-Art and Prospect. *Bulletin of the Chemical Society of Japan*. 95 (2021) 1-25, doi: https://doi.org/10.1246/bcsj.20210297.

[14] Legala, A., Shahgaldi, S. and Li, X., Data-based modelling of proton exchange membrane fuel cell performance and degradation dynamics. *Energy Conversion and Management.* 296 (2023) 117668, doi: https://doi.org/10.1016/j.enconman.2023.117668.

[15] Pei, P., Fu, X., Zhu, Z., Ren, P. and Chen, D., Activation of polymer electrolyte membrane fuel cells: Mechanisms, procedures, and evaluation. *International Journal of Hydrogen Energy*. 47 (2022) 24897-24915, doi: https://doi.org/10.1016/j.ijhydene.2022.05.228.

[16] Du, F., Dao, T. A., Peitl, P. V. J., Bauer, A., Preuss, K., Bonastre, A. M., Sharman, J., Spikes, G., Perchthaler, M., Schmidt, T. J. and Orfanidi, A., Effects of PEMFC Operational History under Dry/Wet Conditions on Additional Voltage Losses due to

Ionomer Migration. *Journal of The Electrochemical Society.* 167 (2020) 144513, doi: https://doi.org/10.1149/1945-7111/abc83f.

[17] Kumar, R. and Subramanian, K. A., Analysis of activation, ohmic, and concentration losses in hydrogen fuelled PEM fuel cell. *Advances in Energy Research*. 8 (2022) 253–264, doi: https://doi.org/10.12989/eri.2022.8.4.253.

[18] Yang, B., Liang, B., Qian, Y., Zheng, R., Su, S., Guo, Z. and Jiang, L., Parameter identification of PEMFC via feedforward neural network-pelican optimization algorithm. *Applied Energy*. 361 (2024) 122857, doi: https://doi.org/10.1016/j.apenergy.2024.122857.

[19] Alsaeedi, A. H., Aljanabi, A. H., Manna, M. E. and Albukhnefis, A. L., A proactive metaheuristic model for optimizing weights of artificial neural network. *Indonesian Journal of Electrical Engineering and Computer Sci*ence. 20 (2020) 976-984, doi: http://doi.org/10.11591/ijeecs.v20.i2.pp976-984.

[20] Rasamoelina, A. D., Adjailia, F. and Sinčák, P. A Review of Activation Function for Artificial Neural Network. in *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI).* (2020), 281-286, doi: https://doi.org/10.1109/SAMI48414.2020.9108717.

[21] Sharma, S., Sharma, S. and Athaiya, A., Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology.* 4 (2020) 310–316, doi: https://doi.org/10.33564/IJEAST.2020.V04I12.054.

[22] Nanadegani, F. S., Lay, E. N., Iranzo, A., Salva, J. A. and Sunden, B., On neural network modeling to maximize the power output of PEMFCs. *Electrochimica Acta*. 348 (2020) 136345, doi: https://doi.org/10.1016/j.electacta.2020.136345.

[23] Zuo, B., Zhang, Z., Cheng, J., Huo, W., Zhong, Z. and Wang, M., Data-driven flooding fault diagnosis method for proton-exchange membrane fuel cells using deep learning technologies. *Energy Conversion and Management.* 251 (2022) 115004, doi: https://doi.org/10.1016/j.enconman.2021.115004.

[24] Deng, S., Zhang, J., Zhang, C., Luo, M., Ni, M., Li, Y. and Zeng, T., Prediction and optimization of gas distribution quality for high-temperature PEMFC based on data-driven surrogate model. *Applied Energy*. 327 (2022) 120000, doi: https://doi.org/10.1016/j.apenergy.2022.120000.

[25] Bayrak, Z. U. and Celik, N., Determining the Effects of Operating Conditions on Current Density of a PEMFC by Using Taguchi Method and ANOVA. *Arabian Journal for Science and Engineering*. 49 (2024) 10741-10752, doi: https://doi.org/10.1007/s13369-023-08470-9.

[26] Ariesanti, Y., Octavianus, P. G. M., Handayani, A. T. and Abbas, B., Characterization of Polyvinyl Alcohol–Collagen–Hydroxyapatite Composite Membrane from *Lates calcarifer* Scales for Guided Tissue and Bone Regeneration. *European Journal of Dentistry*. 17 (2023) 1153-1162, doi: https://doi.org/10.1055/s-0042-1759488.

[27] Sadeghi, H., Mirzaee, I., Khalilarya, S. and Ahmadi, N., Numerical investigation of gas channel geometry of proton exchange membrane fuel cell. *Renewable Energy Research and Applications.* 1 (2020) 93-114, doi: https://doi.org/10.22044/rera.2020.9182.1021.

[28] Mota, A., Gonzalez, E. R. and Eiswirth, M., Kinetic insights over a PEMFC operating on stationary and oscillatory states. *The Journal of Physical Chemistry A*. 115 (2011) 13773-13782, doi: https://doi.org/10.1021/jp205341w.

[29] Puranik, S. V., Keyhani, A. and Khorrami, F. Neural Network Modeling of Proton Exchange Membrane Fuel Cell. *IEEE Transactions on Energy Conversion*. 25 (2010) 474-483, doi: https://doi.org/10.1109/TEC.2009.2035691.

[30] Doumbia, M. L., PEM fuel cell modelling using artificial neural networks. *International Journal of Renewable Energy Research*. 4 (2014) 725-730.

[31] Lau, M. M. and Lim, K. H. Review of Adaptive Activation Function in Deep Neural Network. in *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES).* (2018), 686-690, doi: https://doi.org/10.1109/IECBES.2018.8626714.

[32] Wang, X., Qin, Y., Wang, Y., Xiang, S., and Chen, H., ReLTanh: An activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis. *Neurocomputing.* 363 (2019), 88-98, doi: https://doi.org/10.1016/j.neucom.2019.07.017.