algorithms

MDPI

Article

An Optimized Data Analysis on a Real-Time Application of PEM Fuel Cell Design by Using Machine Learning Algorithms

Arun Saco [1], P. Shanmuga Sundari [2,*], Karthikeyan J [3] and Anand Paul [4]

[1] Department of Mechanical Engineering, Sri Venkateswara College of Engineering and Technology, Chitoor 517127, India
[2] Department of Computer Science and Engineering, Sri Venkateswara College of Engineering and Technology, Chittoor 517127, India
[3] Professor of Humanities and Sciences, Sri Venkateswara College of Engineering and Technology, Chittoor 517127, India
[4] School of Computer Science and Engineering, Kyungpook National University, Daegu 37224, Korea
* Correspondence: sigashanmu@gmail.com

Abstract: In recent years, machine learning algorithms have been applied in many real-time applications. Crises in the energy sector are the primary challenges experienced today among all countries across the globe, regardless of their economic status. There is a huge demand to acquire and produce environmentally friendly renewable energy and to distribute and utilize it efficiently because of its huge production cost. PEMFC are known for their energy efficiency and comparatively low cost, and can be an alternative energy source. The efficiency of these PEMFC can still be enhanced with the help of advanced technologies like machine learning and artificial intelligence, as they provide an optimal solution to explore the hidden knowledge from the generated data. The proposed model attempts to compare several design techniques with varied humidity levels. To enhance the performance of PEMFC, the various humidification processes were considered during the experimental study. The humidification reduces the heat during energy generation and increases the performance of PEM fuel cell. The humidity levels such as 100%, 50%, and 10% were considered to be tested with the machine learning models. The SVMR, LR, and KNN algorithms were tested and observed with the RMSE value as the evaluation parameters. The observed results show that SVMR has an RMSE rate of 0.0046, the LR method has an RMSE rate of 0.0034, and KNN has an RMSE rate of 0.004. The analysis shows that the LR model provides better accuracy than other models. The LR model enhances the PEMFC performance.

Keywords: PEM fuel cells; machine learning; SVMR; LR; KNN

Citation: Saco, A.; Sundari, P.S.; J, K.; Paul, A. An Optimized Data Analysis on a Real-Time Application of PEM Fuel Cell Design by Using Machine Learning Algorithms. *Algorithms* **2022**, *15*, 346. https://doi.org/10.3390/a15100346

Academic Editor: Frank Werner

Received: 12 July 2022
Accepted: 20 September 2022
Published: 25 September 2022

1. Introduction

The proton exchange membrane fuel cell (PEMFC) is an alternative source of energy for the automotive industry. The fuel cell works on the principle of converting the chemical energy of the fuel into electrical energy through electrochemical reactions. The hydrogen and oxygen are considered as the reactant and oxidant for the PEMFC. During the electrochemical reaction, hydrogen oxidation reaction and oxygen reduction reaction will take place. The electrons will be formed during the electrochemical reaction. The working principle of PEMFC is based on the electrochemical reaction of the reactant and oxidant in the electrode of the fuel cell. During this electrochemical reaction, the fuel cell responds to the exothermic reaction, thereby producing heat in the system. The temperature of the PEMFC during the electrochemical reaction varies from 30 °C to 80 °C depending upon the operating parameters, namely flow rate, pressure, and humidification of the reactant and oxidant, respectively [1].

A number of studies have been carried out in the field of PEMFC as expressed earlier [2,3]. The majority of these studies have been carried out by using the methodology of computational fluid dynamics (CFD). A very small amount of deviation was reported in most of the previous works between experimental and numerical (CFD) analysis; hence, it can be inferred that CFD is a very viable method of predicting and analyzing the performance of a PEMFC having a high accuracy percentage [4].

The main factor to be considered when designing a fuel cell is to ensure proper humidification without causing flooding as expressed in previous literature [5]. Flooding factors are considered when designing the flow channels and the operating conditions of the fuel cells. The PEMFC needs to be humidified for its proper functioning and prevent it from drying out. This humidification needs to be provided through an anode as well as a cathode inlet. However, higher humidification, especially for the cathode can result in flooding conditions, which in turn will lead to the deterioration of the performance of the fuel cell.

Fuel cells are gaining worldwide importance because of their clean nature, fuel, and reliability; they are finding increased applications in almost every industry. Proton exchange membrane fuel cells (PEMFCs) are preferred on a larger scale due to their lower operating conditions and high power output. Increased manufacturing costs for fuel cells have been a major concern for researchers around the world. Scaling up PEMFC can help reduce this very cost to a great extent. As found in the literature, an enormous number of studies have been done in the field of PEMFCs with a typical active area of 25 to 50 cm^2. The authors have found from the literature that there are not many studies that deal with scaled-up PEMFC with active areas of 100 cm^2. The power output of a single cell is increased as the PEMFC is scaled up. On the other hand, if the PEMFC is scaled up, it must be ensured that the flow channel is properly designed for effective water and species distribution; otherwise the power density would drop as compared to the standard PEMFC with an active area of 25 cm^2. Hence, this present work will contribute significantly to the existing literature by giving an in-depth analysis of scaled-up PEMFC and all other factors which affect its performance by evaluating some novel flow-field designs that numerically use CFD. A real-time numerical data analysis was implemented by [6] to enhance the performance of PEMFC with various designs.

To enhance the performance of the PEMFC, this paper adopts machine learning techniques which will give better performance in power generation. The numerical analysis methods is a time-consuming method that leads to less accuracy. However, the proposed methods try to reduce the operational cost by implementing prediction methods. The main objective of this method is to enhance the performance of the PEMFC.

1. The presented method exploits machine learning algorithms to optimize the power generation.
2. Various PEMFC design models were considered to prepare the data for analysis. In addition, the method considered various humidity levels.
3. The method experimented and analyzed three different regression algorithms to select optimized model for prediction.
4. Finally, the methods validated with existing numerical analysis methods.

Based on the observed results, the linear regression methods yield better results.

2. Literature Survey

2.1. PEMFC and Various Design

Neutron imaging techniques have been used by [7] for determining the distribution of water in the membrane electrolytes of the PEMFC and have been used as benchmark data for validating numerically predicted results by using a CFD code. The local distribution of liquid water as predicted by the model are compared to those obtained by using the techniques of neutron imaging at various different operating parameters. The numerical CFD experiment has been validated with experimental data by using polarization curves.

The 3D CFD simulations for the flow fields in the bipolar plate for hydrogen flow in a PEMFC was performed by [8]. The author has considered different flow field models for the bipolar plates by varying the geometry. OPEN FOAM software has been used to analyze the flow simulations and to obtain the velocity and pressure distributions. The variable parameters considered for these straight parallel flow fields are the depth, shape, and width of the channels. It is reported that the geometric changes in the structure of the parallel flow field could affect the flow and pressure distribution.

The numerical simulation on a wavy flow field structure and a vertical flow field structure using CFD code on PEM fuel cell performance is carried out by [9]. They reported that due to the wavy flow fields, vertical fluid flows were established, which helps in improving the effective diffusion of oxygen and delay concentration losses for high current density. The wavy structure helps to reduce concentration losses even at lower stoichiometry. Moreover, the wavy flow field design also helps to maintain a uniform current density and gas flow. A numerical investigation on the two-phase flow occurring in a PEMFC with tapered channel flow using air and water [10]. They reported more effective water removal for tapered channels due to increased air velocity. The channel tapering effect on fuel cell performance was significant at higher temperatures, low voltages, and high current density. When compared with normal rectangular channels, the water is removed more effectively with the tapered flow channel.

A 3D CFD model of a new type of flow field for a PEMFC which has a structure like the branches of a tree [11]. They studied three different tree types of flow field configuration. They were compared to the conventional flow field types like the parallel and the serpentine. They reported that the tree-shaped designs ensure a much more uniform distribution of the reactants and lower pressure drop when compared to the conventional patterns. In the tree-shaped design, it was found that as the bifurcations increase, and there is an increase in the active area, which increases the cell performance.

The numerical investigation on nature-inspired flow field design for a PEMFC was conducted by [12]. They also stated the role of CFD models in analyzing these nature-inspired flow field designs. Various nature-inspired designs have been considered, such as fractal designs, heuristic, biologically inspired designs, and formal, biologically inspired designs. A comparison of these designs to the conventional flow fields has also been presented, and the various challenges regarding the development of these nature-inspired flow designs. A 3D CFD investigation of the liquid water dynamics on the performance of PEMFC was conducted by [1]. The volume of the fluid model has been used for the simulation purpose. It was shown that droplets from inner and outer pores tend to move along the lower edge of the gas channel. They reported that there is an increase in GDL water surface coverage and a decrease in water volume fraction. The proper balance between the water volume fraction and the GDL surface water coverage ratio will optimize the performance of the fuel cell. A numerical simulation on a new compound flow-field design using a 3D CFD model was performed. They compared the performance characteristics between conventional parallel and serpentine flow fields [13]. The contours and the polarization curves for the flow fields have been used for comparison. The output numerical results of the models stated that the parallel flow field performance is lesser than the other two designs due to insufficient reactant distribution. They reported that the compound flow field is better for controlling and reducing flooding. However, the performance of the compound flow field was similar to that of the serpentine flow field.

A volume of fluid (VOF) model which has been coupled with a one-dimensional MEA model to study the effect of low misdistribution in parallel channels of a PEMFC [14]. The results state that the slug flow patterns increase the surface water coverage in the gas diffusion layer, decreasing cell performance. The performance of the fuel cell can be improved by optimizing the flow field resistance without causing much loss of pressure in the flow.

A detailed review on the application and development of stainless steel bipolar plates (BPPs) for PEMFC was performed by [15]. Various assemblies and processes required for manufacturing and optimizing the performance of these steel BPPs have been discussed. A detailed discussion about the various processes involved in stainless steel BPPs, such as laser welding, micro-stamping, rubber pad forming, and hydro-forming have taken place. They reported the effects of anti-corrosive and conducting coatings on stainless steel materials. The effects of various errors in the shape, size, and other parameters of the BPPs on the performance have also been reported.

A numerical and CFD simulation to determine the optimum channel width to rib height ratio for serpentine flow fields to get the optimum performance of the fuel cell [16]. Seven different flow fields have been analyzed, and its effects on water distribution, current density, flooding, and reactant gases mass fraction were also studied numerically. Pressure, stoichiometry, and temperature effects have also been evaluated in this study [17].

The numerical analysis on the anode and the cathode gas channel of a PEMFC with rectangular-shaped obstacles in the flow field was conducted by [18]. By using a 3D CFD model, the numerical analysis of the flow field with obstacles at different operating parameters of relative humidity, stoichiometry, and temperature was studied. The authors reported that the flow field with rectangular obstacles, and the current density values are higher than ordinary flow fields. The fuel cell performance is studied by using the I-V polarization curves [19]. An effective cooling system for a PEMFC was considered by [20], the authors of which performed a numerical investigation of new and different cooling flow fields. The parameters like thermal behavior, pressure drop, and coolant flow distribution of various thermal designs have been studied. They stated an increase in the mass flow rate of the cooling field as well as a minimization of the maximum temperature difference between flow field surface.

An investigation on water dynamics inside the PEMFC was performed by [21], the authors of which investigated the water content at the cathode of the PEM cell due to the dynamic behavior of liquid water. The surface coverage of the GDL by water can be greatly reduced, and this can be achieved by increasing the inter-pore distance, and decreasing the pore diameter. The three-dimensional PEMFC with different operational parameters and geometric inputs [22]. They have compared the simulated model with experimental results, and there is a good match between them. These parameters include species concentration, water content in the PEM, over-potentials, and the current densities. The effect of these factors on cell performance has also been evaluated.

2.2. *PEMFC Using the Machine Learning Model*

Machine learning and artificial intelligence (AI), have been proven to give better performance in data analysis, system control, and design optimization performance and energy development [5]. Advances in computational power, simulation, and machine learning enables researchers to explore large amounts of data, to provide inspiration and tools for designing new systems. This study [23] experiments with modeling and data analysis tools to build a framework for the study and development of high-temperature polymer electrolyte membrane fuel cells (HT-PEMFC). In [23], the machine learning technique is used to identify the two-phase flow pressure drop in a flow channel of a PEMFC. Three machine learning models: logistic regression, support vector machine, and artificial neural networks (ANN) are used to classify the liquid–gas two-phase flow pressure drop images into three pressure classes. Machine learning and AI, which are effective tools for data analysis/classification, system control/monitoring, and design/performance optimization, are gaining power in the material- and energy-development industries [23]. The machine learning-based modeling and analysis approach proposed here allows for quick identification of material qualities and device operating parameters that improve PEMFC performance [24]. Catalyst layers have been intensively investigated for not only PEMFC, but also many other systems, such as electrolyzes and sensors with Pt-catalyst electrodes. Machine learning and AI are immensely helpful, but also challenging, for catalyst layer development when such catalyst layers have been thoroughly studied not only for PEMFC, but also many other systems, such as electrolyzes and sensors with Pt-catalyst electrodes [25]. The parametric identification of a polymer electrolyte membrane (PEM) can be effectively performed by using the machine learning model [26].

3. Proposed Methodology

The proposed method is broadly classified into preprocessing, model selection, and evaluation. Figure 1 illustrated the overall architecture of the proposed method.

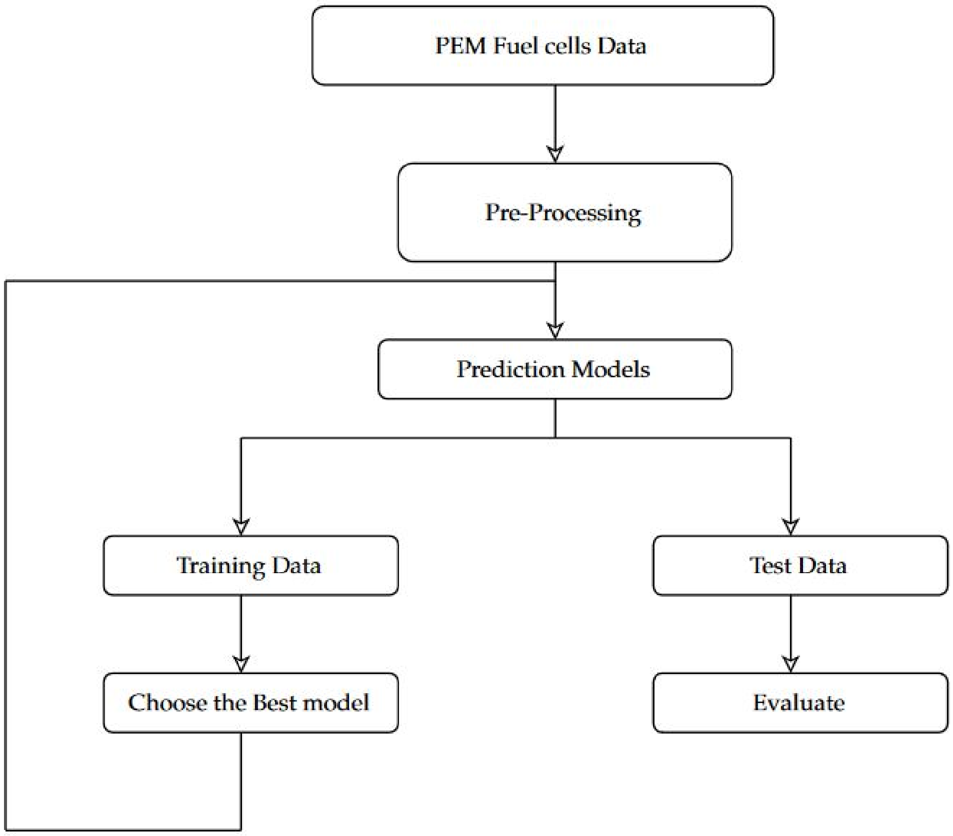

Figure 1. Architecture of proposed method.

3.1. Numerical Study

Numerical modeling consists of three stages, namely pre-processing, processing, and post-processing. In the pre- processing stage, the three-dimensional geometry is modeled by using Solid Works 10.0 software. Flow channel designs are illustrated in Figure 2, and the design parameter is listed in Table 1. Finite volume discretization of the PEMFC is modeled by using the ANSYS ICEMCFD 15.0 code. In the processing stage, governing equations, which are used to capture the flow physics and electrochemical reactions of PEMFC are solved by using the ANSYS Fluent 15.0. CFD code. The modeling of the proton exchange membrane fuel cell is created by using solidworks 10.0, which consists of nine parts, namely membrane electrolyte, catalyst layers, gas diffusion layers, flow channels, and the current collectors in both anode and cathode side. The proton exchange membrane fuel cell is discretized through a finite volume approach by using ANSYS ICEMCFD 15.0. The PEMFC parts are meshed individually by adopting the blocking technology to generate hexahedral elements. The angle and determinant value mesh lies in the range of 85° and 0.9, respectively.

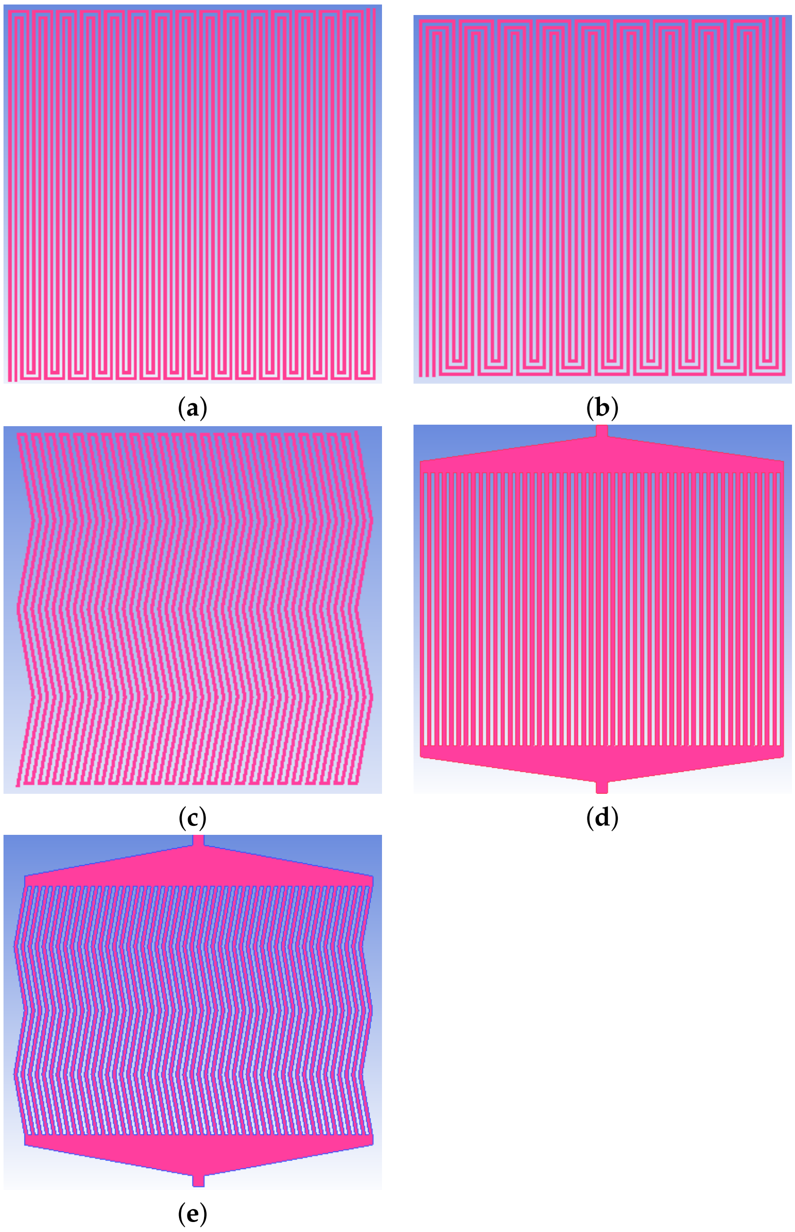

Figure 2. Flow channel design. (**a**) 2-Serpentine. (**b**) 3-Serpentine. (**c**) Serpentine zigzag. (**d**) Straight zigzag. (**e**) Straight zigzag.

Table 1. Design parameters.

Parameter	Value
PEMFC Cross-sectional Area	100 cm^2
Width of the flow Channel	1 mm
Width of the Landing (Rib)	1 mm
Depth of the Channel	1 mm
Thickness of the Gas Diffusion Layer	0.33 mm
Thickness of the Catalyst Layer	0.01 mm
Thickness of the Membrane	0.051 mm

3.2. Surrogate Models Development

The surrogate models were implemented with the help of machine learning algorithms. The machine learning algorithms provided the performance growth in many energies' efficacy models. The machine learning algorithms are classified into supervised, unsupervised learning, and predictive models: classification, clustering, and regression. Each method has different algorithms to process the data. The data analysis has been made based on the amount of data available, and the processing resources must all be considered while choosing the best algorithm. The regression analysis has a greater advantage over classification and clustering. The regression algorithm predicts the independent variable data based on dependent data. Adopting a predictive model provides better accuracy as per previous studies [27]. The proposed method considered three machine learning algorithms such as support vector machines, linear regression and k-nearest neighbor method to analyze the performance of PEMFC with various humidity levels.

3.3. Machine Learning

Machine learning and artificial intelligence play an important role in various scientific applications. The data analysis, classification, regression, and clustering techniques are widely applied in many applications like material and energy development. Machine learning and artificial intelligence have simplified the numerical analysis of PEMFC in recent decades, allowing for the discovery of essential knowledge of hidden patterns in energy development, design optimization, and energy efficiency [28].

The research objective is focused on three-dimensional, two-phase computational fluid dynamics to scale up PEMFC. However, existing studies show that the size of fuel cells are stacked together with other fuel cells, which is not adequate for agility purposes [29]. The present study focuses on different design methods to scale up an active area of 100 cm^2 with various flow field configurations namely two-serpentine, three-serpentine, serpentine zigzag, straight parallel, and straight zigzag. Numerical study has been conducted and validated with different design methods. The various humidification of hydrogen and oxygen is considered for better performance. Machine learning algorithms are introduced to analyze the data's hidden knowledge. The various algorithms that include classification, regression, and clustering reveal the relationship between the input and output parameters under appropriate training to test the data.

3.4. Support Vector Machine

In recent years, a massive amount of data analysis has been carried out in support vector machines, in a multidisciplinary environment [30–32]. SVM is powerful and robust in classification and regression methods in multidisciplinary research (Algorithm 1). It plays a significant role in pattern recognition which is helpful in research. The support vector machine for regression (SVMR) problem is based on structural risk minimization which provides good performance, with an epsilon (ϵ) intensive loss function obtained, and a symmetric performance of reduced training errors and reduced model complexities. These models work well even with small training data. Hence, SVM for regression has chosen to build a surrogate model. SVMR is used to find the hidden relationship between the input variables like voltage, humidity level, and design structure of the PEM cell and the output variable of the current.

Algorithm 1 An algorithm for support vector machine

1: Split the data into training and testing
2: Set random number for learning parameters
3: Repeat the process to optimal value
4: Fit the SVMR model
5: Find error (mse) value
6: Continue the same until no more changes in the error value
7: Find the best value for regression
8: End

3.5. Linear Regression

The linear regression model is a basic and simple regression technique [33,34]. Finding a linear relationship between one or more variables is done by using linear regression. There are two types of linear regression, simple and multiple linear regression. This model is used to estimate the relationship between the input variable (X) and output variable (Y) which work under the principle of statistical and machine learning concepts. The LR model predicts and calculates changes in output variable (Y) when the input variable changes (X). The input and output variables can be exchanged by correlation coefficient. The LR model predicts the target value (Y) based on the input variable (X). The simple linear regression is given in Equation (1),

$$y = mX + c => Y = s_0 X + s_1, \tag{1}$$

where from (1), y is the target (dependent) variable, X is the input (independent) variable, s_0 is the slope, and s_1 is the regression coefficient. The complexities of the regression model are used to evaluate the number of coefficients used in this model. The model coefficient or parameters are obtained together as m and c; after utilizing the training data to get the model coefficient estimates of $\hat{s_0}$ and $\hat{s_1}$. We have

$$\hat{y} = \hat{s_0} X + \hat{s_1}, \tag{2}$$

where from (2), $\hat{y}$ indicates the prediction. We also have

$$e_j = y_j - \hat{y}_j, \tag{3}$$

where from (3), e_j is the error difference between the actual and predicted value of j^{th} representation. The sum of square error is defined as

$$SSE = e_1^2 + e_2^2 .....e_n^2. \tag{4}$$

In regression, there is a concept of best-fit line, which is the line that best matches the provided data (Algorithm 2). SSE is called loss function or cost function, and minimizing the error would result in good fit or accuracy. This strategy is called the least squares method. By using some calculus, the least-squares method chooses 0 and 1 to minimize the SSE. Then, a new set of coefficients is generated, and it needs some metrics to verify that the estimated coefficients are accurate.

Algorithm 2 An algorithm for linear regression

1: Start
2: Load data
3: Split the data into training and testing
4: Set random number for learning parameters
5: Repeat the process to optimal value
6: Fit the LR model
7: Find error (mse) value
8: Continue the same until no more changes in the error value
9: Find the best value for regression
10: Predict the value for regression model
11: End

3.6. KNN for Regression

The KNN model is used for both classification and regression techniques (Algorithm 3). Based on the similarity among the features, the new value is predicted. With the closeness learned from the features in the training data, the model predicts the new value. At the initial step, the model calculates the distance between the new value and each data point in the training dataset. Based on the distance, the K value is chosen. There are various distance measures which include Euclidean distance, Manhattan distance, and Hamming distance, which are used to find the similarity distance between the training data points and new data points [35]. Choosing optimal K values influences the performance of the KNN model. The optimal value has been chosen in the training dataset based on error.

Algorithm 3 An algorithm for K-nearest neighbor for regression

1: Start
2: Load data
3: Split the data into training and testing
4: Initialize the K value
5: Find distance between each data points for data sample
6: Append the value in the index list
7: Sort the list and take average value of k-indexes in the list
8: End

4. Experimental Results

4.1. Data Generation

A real-time PEM heap data simulation is configured by containing cells and an active area of 100 cm^2. A real-time numerical simulation is carried out with an active area of 100 cm^2 PEMFC. The numerical work was carried out at various cell voltages ranging from 0 V to 0.95 V under different hydrogen and oxygen humidification [36]. The dataset was considered with 10%, 50%, and 100% humidity levels with different flow channel designs likes two serpentine, three serpentine, serpentine zigzag, straight parallel, and straight zigzag [6]. The experimental study was conducted based on humidity levels.

4.2. Evaluation Metric

The surrogate model is evaluated by using the following metrics [37]. The root means squared error (RMSE) and the squared correlation coefficient (R^2). The root means square error is the average of mean error between the actual and predicted value.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)} \tag{5}$$

$$R^2 = 1 - \sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{6}$$

4.3. Learning Parameter

The dataset is randomly split into training and testing. The k-fold cross-validation technique is applied to obtain the optimized parameter in the training dataset. For the experimental purpose, the five-fold validation is conducted and the radial basis function (RBF) kernel function is chosen to optimize the parameters like gamma and epsilon for SVM regression. The model learns the hyperparameter valuesy performed by a five cross-fold validation method. The optimal value for the support vector machine for regression was chosen based on different values tested with k-fold validation.

4.4. Result and Discussion

The machine learning-based modeling and analysis approach proposed here allows for flow channel design and operating parameters of hydrogen and oxygen that improve PEMFC performance. The performance of the presented method was evaluated with different models, such as SVMR, KNN, and LR models. The mentioned models predict the cell voltage is directly based on regression.

4.4.1. SVMR Method

The support vector machine regression method was tested with various learning parameters. An epsilon is the loss function for SVMR. The model was tested with various values, which range from [0.11, 0.12 ... 0.16]. The result illustrated in Figure 3. shows that 0.12 is the optimal value for predicting cell voltage by using SVMR (Table 2). The experiment was stopped at 0.16 (Table 3). as there is an increase in error value. In addition to this test, the experiment was conducted with various kernel functions to identify the optimal kernel function (RBF, poly, linear, sigmoid) for the data. The kernel functions help to solve the complexity of the data [38]. It is necessary to select optimized function for data. It is observed that the kernel function RBF gives better result as listed in Table 4 when compared to other functions. Figures 4–6 illustrate the comparison of various kernel functions with humidity levels of 10%, 50%, and 100% (Table 5). The training of the SVM regression model turns out to be the most expensive, as it necessitates the cross-validation tweaking of three hyperparameters. The results show that the 100% humidity level gives better results, as depicted in Figure 7.

Table 2. Results of SVMR.

			100% Humidity		
RMSE	0.0046	0.004	0.0094	0.024	0.0437
R^2	0.9149	0.9444	0.9456	0.5545	0.611
			50% Humidity		
RMSE	0.0031	0.0025	0.005	0.0131	0.015
R^2	0.9605	0.9644	0.9221	0.7492	0.8623
			10% Humidity		
RMSE	0.0039	0.0029	0.0047	0.0173	0.0172
R^2	0.9467	0.9563	0.9261	0.5919	0.8311

Table 3. Kernel function comparison (100% humidity).

epsilon	0.11	0.12	0.13	0.14	0.15	0.16
Kernel Function				**rbf**		
RMSE	0.0052	0.0055	0.0071	0.007	0.0082	0.0085
R^2	0.8699	0.8619	0.8205	0.8243	0.831	0.8247
Kernel Function				**poly**		
RMSE	0.0085	0.0098	0.0113	0.0115	0.0123	0.0129
R^2	0.7865	0.7538	0.7161	0.7118	0.746	0.7343
Kernel Function				**Linear**		
RMSE	0.0048	0.006	0.0073	0.0083	0.009	0.0098
R^2	0.8784	0.8502	0.8169	0.7924	0.7751	0.7545
Kernel Function				**sigmoid**		
RMSE	0.0058	0.0069	0.0083	0.0093	0.0109	0.0127
R^2	0.8537	0.8256	0.7925	0.7656	0.7262	0.6821

Table 4. Kernel function comparison (10% humidity).

Kernel Function	Rbf					
RMSE	0.0033	0.0026	0.0029	0.0035	0.0043	0.0054
R^2	0.9574	0.9603	0.956	0.9469	0.9344	0.9184
Kernel Function	**poly**					
RMSE	0.0123	0.0128	0.0131	0.0136	0.0146	0.0158
R^2	0.8325	0.8252	0.8208	0.8139	0.80007	0.784
Kernel Function	**Linear**					
RMSE	0.0043	0.0044	0.0059	0.0085	0.0111	0.0134
R^2	0.9405	0.9391	0.9191	0.8839	0.0848	0.8176
Kernel Function	**sigmoid**					
RMSE	0.0057	0.0062	0.0072	0.0084	0.0103	0.0115
R^2	0.9139	0.9066	0.8908	0.8734	0.8449	0.8261

Table 5. Kernel function comparison (50% humidity).

epsilon	0.11	0.12	0.13	0.14	0.15	0.16
Kernel Function	**rbf**					
RMSE	0.0011	0.0015	0.002	0.0027	0.003	0.0036
R^2	0.97	0.9606	0.9478	0.9307	0.9199	0.9079
Kernel Function	**poly**					
RMSE	0.0049	0.0055	0.0062	0.0071	0.0068	0.0074
R^2	0.8745	0.8596	0.8404	0.8184	0.825	0.8088
Kernel Function	**Linear**					
RMSE	0.0025	0.0036	0.0036	0.0037	0.0039	0.0039
R^2	0.934	0.9068	0.9068	0.904	0.8991	0.8991
Kernel Function	**sigmoid**					
RMSE	0.0049	0.0048	0.0048	0.006	0.0073	0.0079
R^2	0.8742	0.875	0.8758	0.8459	0.8119	0.7983

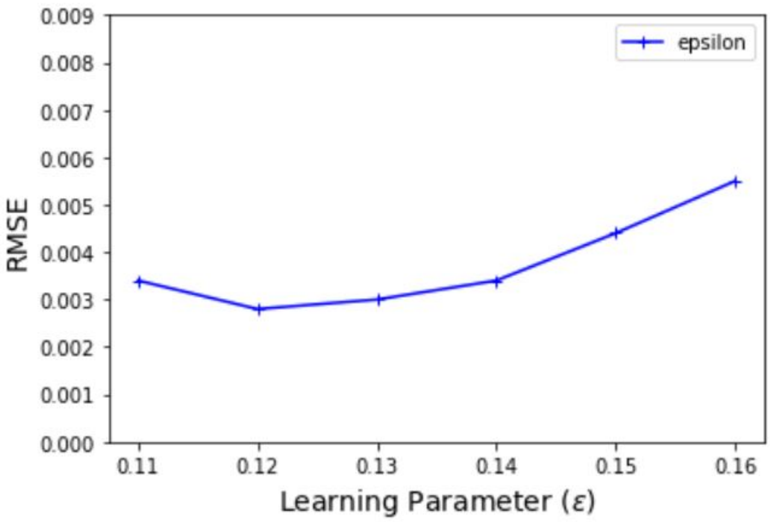

Figure 3. Learning parameter epsilon.

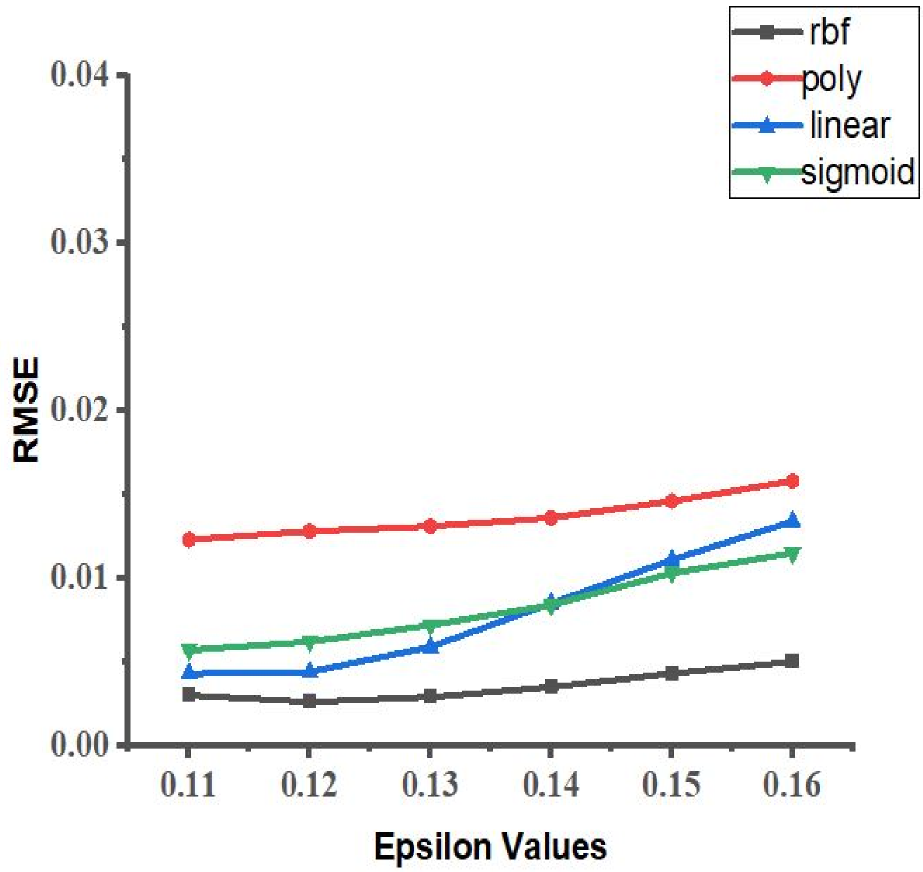

Figure 4. Kernel function comparison for 10% humidity.

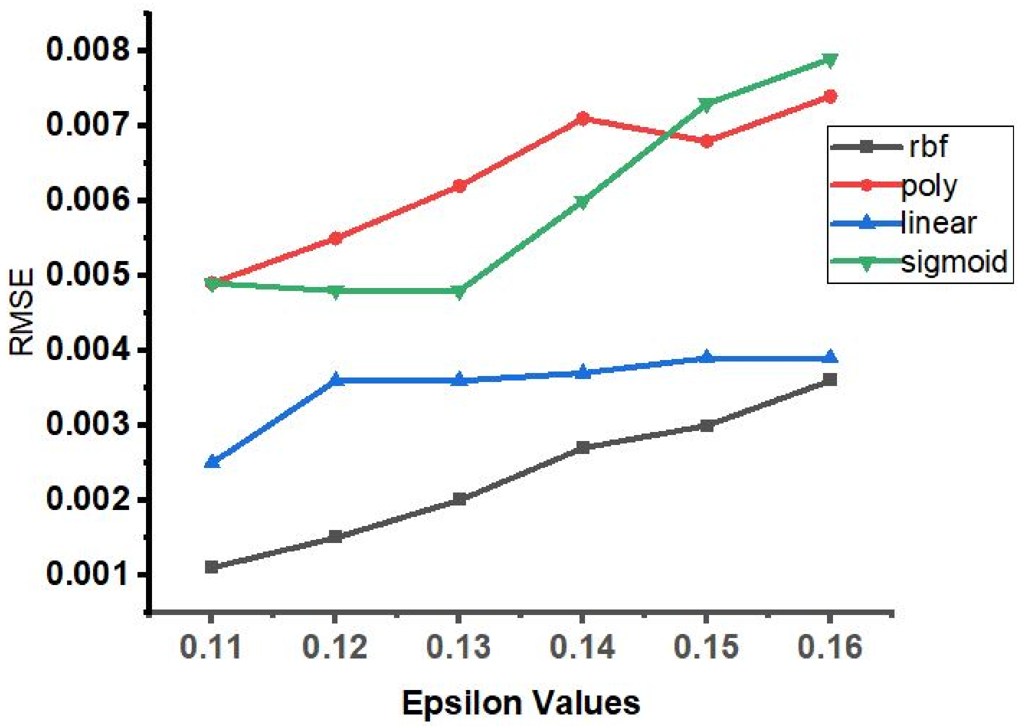

Figure 5. Kernel function comparison for 50% humidity.

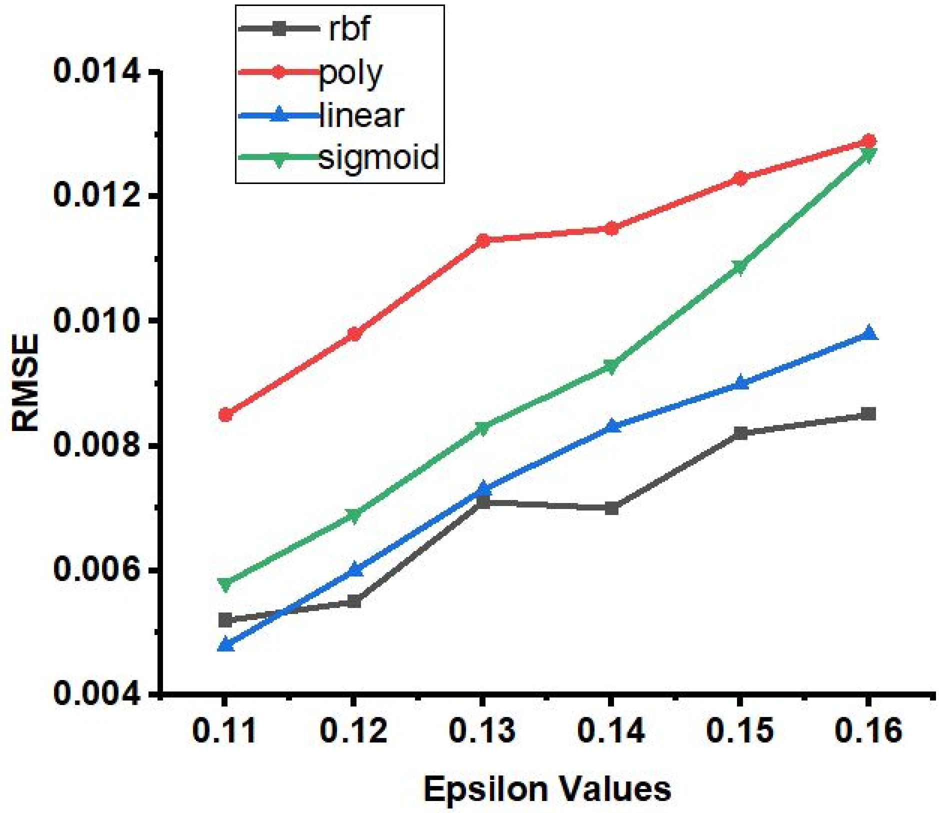

Figure 6. Kernel function comparison for 100% humidity.

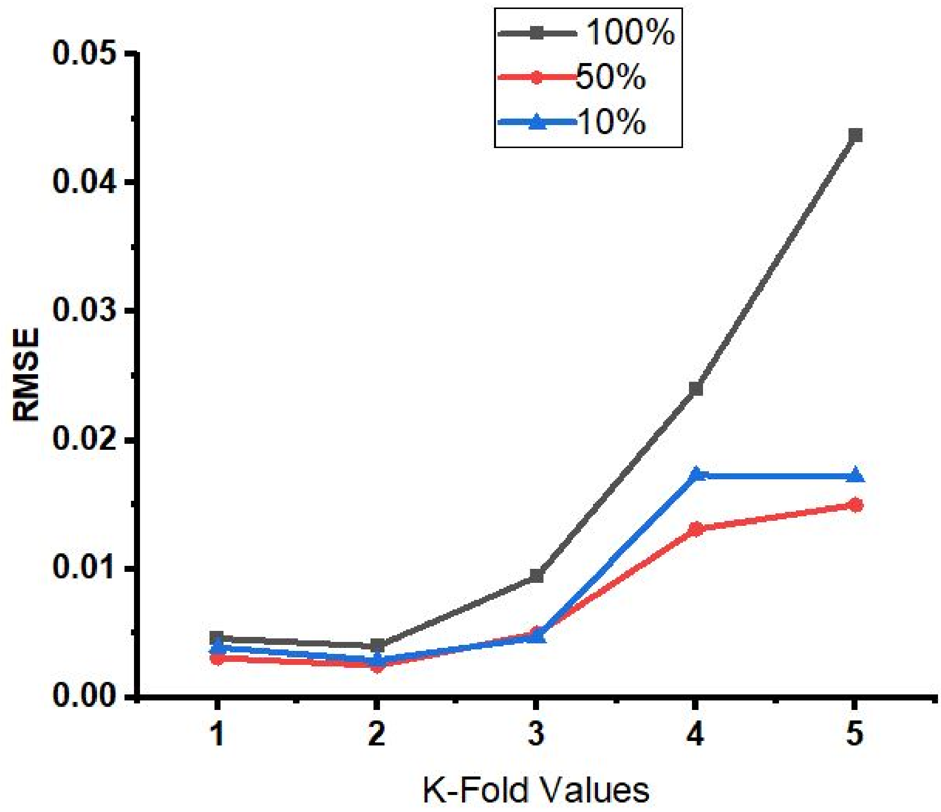

Figure 7. SVMR model with various humidity levels.

4.4.2. Linear Regression Method

The linear regression model was tested with various humidity levels at 10%, 50%, and 100%, and their results are listed in Table 6. It was observed from the Figure 8 that the LR method gives better results at the 50% humidity level. Based on the humidity level, the performance differs from machine learning models.

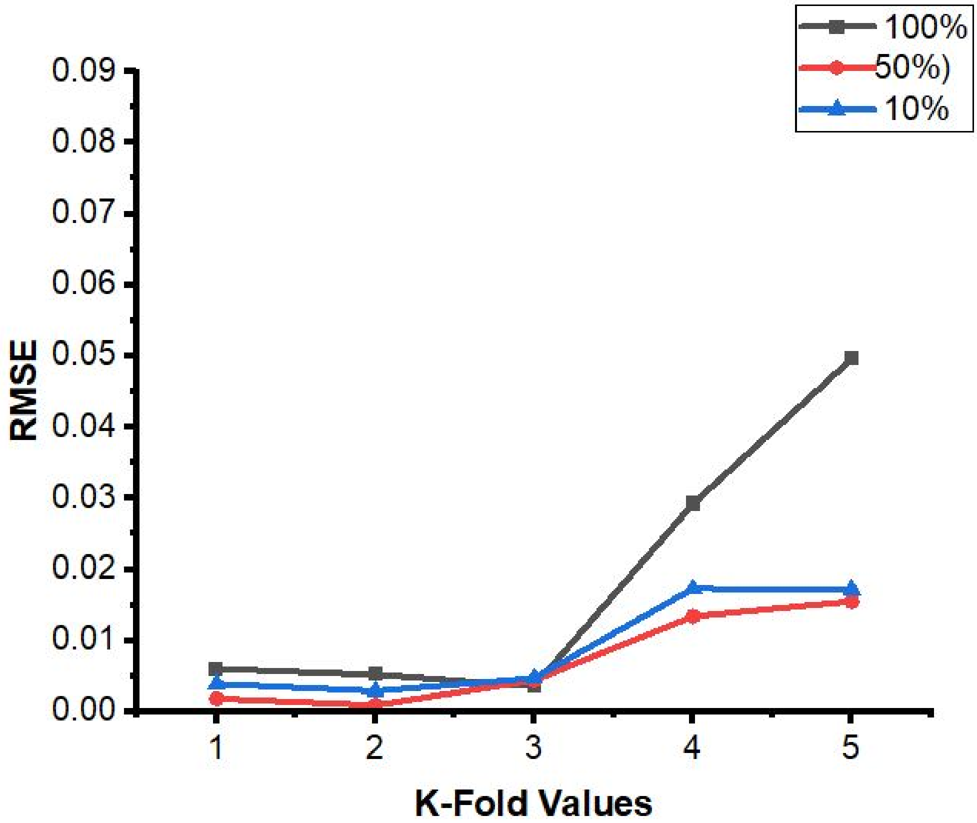

Figure 8. LR model with various humidity levels.

Table 6. Various humidity comparison of LR model.

			100% Humidity		
RMSE	0.0060	0.0052	0.0036	0.0292	0.0496
R^2	0.8489	0.8917	0.9716	0.0035	0.3107
			50% Humidity		
RMSE	0.0018	0.0009	0.0044	0.0134	0.0155
R^2	0.962	0.9762	0.8851	0.591	0.7851
			10% Humidity		
RMSE	0.0039	0.0029	0.0047	0.0173	0.0172
R^2	0.9467	0.9563	0.9261	0.5919	0.0172

4.4.3. KNN Model

The KNN model for regression was tested with various humidity levels, and Table 7 depicts the results. It was observed from Figure 9 that all the humidity level falls at certain 0.013 then it starts to deviate.

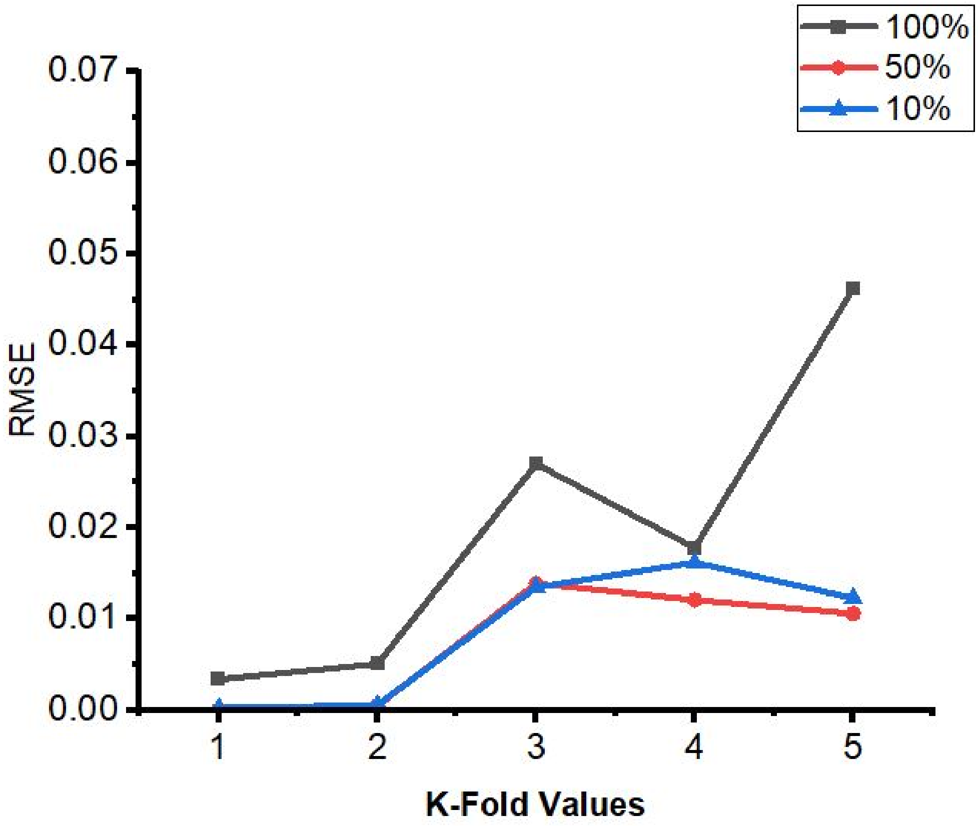

Figure 9. KNN model with various humidity levels.

Table 7. Results of KNN method.

			100% Humidity		
RMSE	0.0034	0.0051	0.027	0.0178	0.0462
R^2	0.9354	0.9281	0.8431	0.6681	0.5879
			50% Humidity		
RMSE	0.0001	0.0005	0.0139	0.0121	0.0106
R^2	0.9982	0.9923	0.7843	0.7698	0.9031
			10% Humidity		
RMSE	0.0003	0.0006	0.0135	0.0162	0.0123
R^2	0.9963	0.9914	0.7876	0.6186	0.8789

The study on comparative analysis of various humidity levels using various machine learning model was analyzed, and the results obtained from the various models show that 100% humidification works well when compared to other levels, as listed in Table 8. The results observed in Figure 10 show that LR methods perform well when compared to other models.

Table 8. Comparison of various models.

Metrics	SVMR	LR	KNN
RMSE	0.0046	0.0034	0.004
R^2	0.9343	0.9354	0.9444

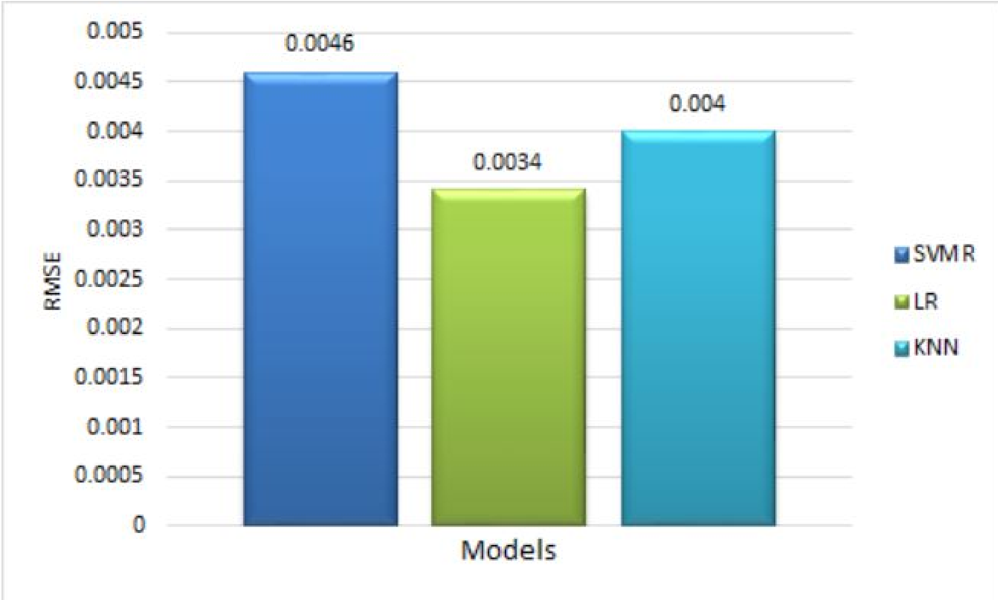

Figure 10. Different models' RMSE value.

4.4.4. Flow Channel Design

Furthermore, the investigation was extended by using the flow channel designs, namely two serpentine, three serpentine, serpentine zigzag, straight parallel, and straight zigzag with machine learning models. Here, the datasets were considered based on the flow channel designs. All the humidification levels are considered with respect to flow channel designs.The tested results illustrated in the Figure 11 show the algorithm gives different results based on the nature of designs.

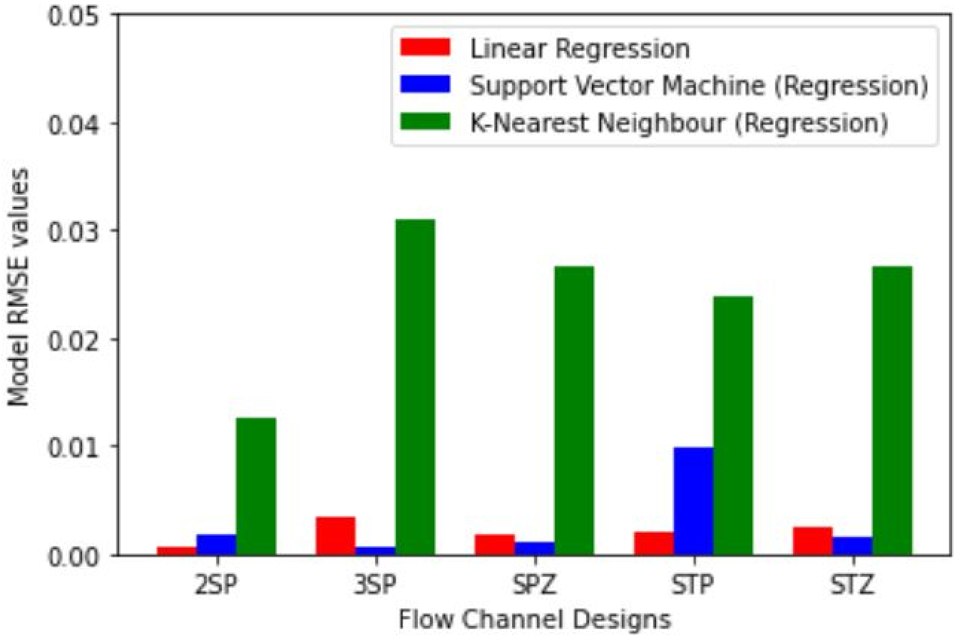

Figure 11. Flow channel design comparison.

4.4.5. Comparison with Numerical Study

The numerical study was conducted on [6] humidification of reactant to enhance the performance of PEMFC. The study considered various fuel cell designs at different levels of humidification. Under the numerical investigation, different cell designs yield different error rates, which increase the complexity of finding the optimal values. In order to find the optimal value, a machine learning algorithm was introduced to enhance the performance of PEMFC. The machine learning-based modeling and analysis approach proposed here allows swift identification of material qualities and device operating parameters that improve PEMFC performance.

The results observed that the machine learning model predicts an output voltage of 0.88 at the greatest power density point, which closely matches with the prediction by numerical models, as illustrated in Figure 12.

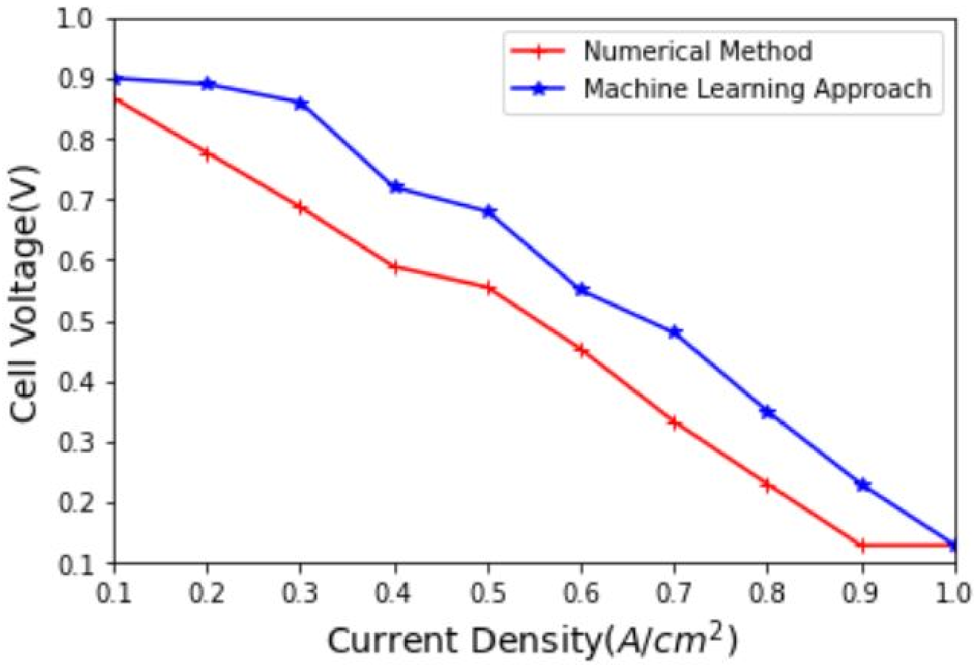

Figure 12. Comparison with numerical study.

5. Conclusions and Future Enhancement

This research experimented with various PEMFC designs with various humidity levels. Machine learning models, such as SVMR, LR, and KNN models, were tested with various humidity levels. The performance is compared by using the error rate of each model. The model which gives the lesser error is considered to be the best algorithm. The experimental analysis model yields the SVMR error rate of 0.0046, LR error rate of 0.0034, and the KNN error rate of 0.004. It is observed that the LR model gives better result. In the future, the research will extend to enhance the performance of PEMFC by using a deep learning approach while considering various parameters like oxygen level, water, humidity, and various cell designs [39].

Author Contributions: Conceptualization, P.S.S.; methodology, P.S.S.; validation, K.J.; formal analysis, P.S.S.; investigation, A.S.; resources, A.P.; writing—original draft preparation, P.S.S.; writing—review and editing, K.J.; project administration, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANN	Artificial Neural Networks
BPPs	Bipolar Plates
CFD	Computational fluid dynamics
GDL	Gas Diffusion Layer
KNN	K-Nearest Neighbor
LR	Linear Regression
ML	Machine Learning
MSE	Mean Square Error
PEMFC	Proton Exchange Membrane Fuel Cell
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SVMR	Support Vector Machine for Regression
VOF	Volume of Fluid
Symbols	
ϵ	Epsilon

References

1. Jo, J.H.; Kim, W.T. Numerical simulation of water droplet dynamics in a right angle gas channel of a polymer electrolyte membrane fuel cell. *Int. J. Hydrogen Energy* **2015**, *40*, 8368–8383. [CrossRef]
2. Jiao, K.; Xuan, J.; Du, Q.; Bao, Z.; Xie, B.; Wang, B.; Zhao, Y.; Fan, L.; Wang, H.; Hou, Z.; et al. Designing the next generation of proton-exchange membrane fuel cells. *Nature* **2021**, *595*, 361–369. [CrossRef] [PubMed]
3. Haider, R.; Wen, Y.; Ma, Z.F.; Wilkinson, D.P.; Zhang, L.; Yuan, X.; Song, S.; Zhang, J. High temperature proton exchange membrane fuel cells: Progress in advanced materials and key technologies. *Chem. Soc. Rev.* **2021**, *50*, 1138–1187. [CrossRef] [PubMed]
4. Zhang, S.; Liu, S.; Xu, H.; Mao, Y.; Wang, K. Numerical Investigation on the Performance of Proton Exchange Membrane Fuel Cell With Zigzag Flow Channels. *Front. Therm. Eng.* **2022**, *1*, 5. [CrossRef]
5. Wang, Y.; Seo, B.; Wang, B.; Zamel, N.; Jiao, K.; Adroher, X.C. Fundamentals, materials, and machine learning of polymer electrolyte membrane fuel cell technology. *Energy AI* **2020**, *1*, 100014. [CrossRef]
6. Selvaraj, A.S.; Rajagopal, T.K.R. Effect of flow fields and humidification of reactant and oxidant on the performance of scaled-up PEM-FC using CFD code. *Int. J. Energy Res.* **2019**, *43*, 7254–7274. [CrossRef]
7. Zhang, G.; Jiao, K. Multi-phase models for water and thermal management of proton exchange membrane fuel cell: A review. *J. Power Sources* **2018**, *391*, 120–133. [CrossRef]
8. Wilberforce, T.; El-Hassan, Z.; Khatib, F.; Al Makky, A.; Mooney, J.; Barouaji, A.; Carton, J.G.; Olabi, A.G. Development of Bi-polar plate design of PEM fuel cell using CFD techniques. *Int. J. Hydrogen Energy* **2017**, *42*, 25663–25685. [CrossRef]
9. Han, S.H.; Choi, N.H.; Choi, Y.D. Simulation and experimental analysis on the performance of PEM fuel cell by the wave-like surface design at the cathode channel. *Int. J. Hydrogen Energy* **2014**, *39*, 2628–2638. [CrossRef]
10. Mancusi, E.; Fontana, É.; de Souza, A.A.U.; de Souza, S.M.G.U. Numerical study of phase flow patterns in the gas channel of PEM fuel cells with tapered flow field design. *Int. J. Hydrogen Energy* **2014**, *39*, 2261–2273. [CrossRef]
11. Lorenzini-Gutierrez, D.; Hernandez-Guerrero, A.; Ramos-Alvarado, B.; Perez-Raya, I.; Alatorre-Ordaz, A. Performance analysis of a proton exchange membrane fuel cell using tree-shaped designs for flow distribution. *Int. J. Hydrogen Energy* **2013**, *38*, 14750–14763. [CrossRef]
12. Arvay, A.; French, J.; Wang, J.C.; Peng, X.H.; Kannan, A.M. Nature inspired flow field designs for proton exchange membrane fuel cell. *Int. J. Hydrogen Energy* **2013**, *38*, 3717–3726. [CrossRef]
13. Vazifeshenas, Y.; Sedighi, K.; Shakeri, M. Numerical investigation of a novel compound flow-field for PEMFC performance improvement. *Int. J. Hydrogen Energy* **2015**, *40*, 15032–15039. [CrossRef]
14. Ding, Y.; Bi, X.; Wilkinson, D. Numerical investigation of the impact of two-phase flow maldistribution on PEM fuel cell performance. *Int. J. Hydrogen Energy* **2014**, *39*, 469–480. [CrossRef]
15. Peng, L.; Yi, P.; Lai, X. Design and manufacturing of stainless steel bipolar plates for proton exchange membrane fuel cells. *Int. J. Hydrogen Energy* **2014**, *39*, 21127–21153. [CrossRef]
16. Rahimi-Esbo, M.; Ranjbar, A.; Ramiar, A.; Alizadeh, E.; Aghaee, M. Improving PEM fuel cell performance and effective water removal by using a novel gas flow field. *Int. J. Hydrogen Energy* **2016**, *41*, 3023–3037. [CrossRef]
17. Mariani, M.; Basso Peressut, A.; Latorrata, S.; Balzarotti, R.; Sansotera, M.; Dotelli, G. The Role of Fluorinated Polymers in the Water Management of Proton Exchange Membrane Fuel Cells: A Review. *Energies* **2021**, *14*, 8387. [CrossRef]
18. Bilgili, M.; Bosomoiu, M.; Tsotridis, G. Gas flow field with obstacles for PEM fuel cells at different operating conditions. *Int. J. Hydrogen Energy* **2015**, *40*, 2303–2311. [CrossRef]

19. Arif, M.; Cheung, S.C.; Andrews, J. A systematic approach for matching simulated and experimental polarization curves for a PEM fuel cell. *Int. J. Hydrogen Energy* **2020**, *45*, 2206–2223. [CrossRef]
20. Alizadeh, E.; Rahgoshay, S.; Rahimi-Esbo, M.; Khorshidian, M.; Saadat, S. A novel cooling flow field design for polymer electrolyte membrane fuel cell stack. *Int. J. Hydrogen Energy* **2016**, *41*, 8525–8532. [CrossRef]
21. Hossain, M.; Islam, S.Z.; Colley-Davies, A.; Adom, E. Water dynamics inside a cathode channel of a polymer electrolyte membrane fuel cell. *Renew. Energy* **2013**, *50*, 763–779. [CrossRef]
22. Yu, L.j.; Ren, G.p.; Qin, M.j.; Jiang, X.m. Transport mechanisms and performance simulations of a PEM fuel cell with interdigitated flow field. *Renew. Energy* **2009**, *34*, 530–543. [CrossRef]
23. Briceno-Mena, L.A.; Venugopalan, G.; Romagnoli, J.A.; Arges, C.G. Machine learning for guiding high-temperature PEM fuel cells with greater power density. *Patterns* **2021**, *2*, 100187. [CrossRef] [PubMed]
24. Wang, B.; Xie, B.; Xuan, J.; Jiao, K. AI-based optimization of PEM fuel cell catalyst layers for maximum power density via data-driven surrogate modeling. *Energy Convers. Manag.* **2020**, *205*, 112460. [CrossRef]
25. Wang, C.; Waje, M.; Wang, X.; Tang, J.M.; Haddon, R.C.; Yan, Y. Proton exchange membrane fuel cells with carbon nanotube based electrodes. *Nano Lett.* **2004**, *4*, 345–348. [CrossRef]
26. Sultan, H.M.; Menesy, A.S.; Kamel, S.; Selim, A.; Jurado, F. Parameter identification of proton exchange membrane fuel cells using an improved salp swarm algorithm. *Energy Convers. Manag.* **2020**, *224*, 113341. [CrossRef]
27. Gao, T.; Lu, W. Machine learning toward advanced energy storage devices and systems. *Iscience* **2021**, *24*, 101936. [CrossRef] [PubMed]
28. Hong, T.; Wang, Z.; Luo, X.; Zhang, W. State-of-the-art on research and applications of machine learning in the building life cycle. *Energy Build.* **2020**, *212*, 109831. [CrossRef]
29. Daud, W.; Rosli, R.; Majlan, E.; Hamid, S.; Mohamed, R.; Husaini, T. PEM fuel cell system control: A review. *Renew. Energy* **2017**, *113*, 620–638. [CrossRef]
30. Han, I.S.; Chung, C.B. Performance prediction and analysis of a PEM fuel cell operating on pure oxygen using data-driven models: A comparison of artificial neural network and support vector machine. *Int. J. Hydrogen Energy* **2016**, *41*, 10202–10211. [CrossRef]
31. Han, I.S.; Chung, C.B. A hybrid model combining a support vector machine with an empirical equation for predicting polarization curves of PEM fuel cells. *Int. J. Hydrogen Energy* **2017**, *42*, 7023–7028. [CrossRef]
32. Prabu, M. Canopy Removal on Satellite Images Using Classification and Contrast Enhancement. Available online: https://www.researchgate.net/publication/338606745_Canopy_Removal_On_Satellite_Images_Using_Classification_And_Contrast_Enhancement (accessed on 5 July 2022).
33. Placca, L.; Kouta, R.; Candusso, D.; Blachot, J.F.; Charon, W. Analysis of PEM fuel cell experimental data using principal component analysis and multi linear regression. *Int. J. Hydrogen Energy* **2010**, *35*, 4582–4591. [CrossRef]
34. Yu, C.; Yao, W. Robust linear regression: A review and comparison. *Commun.-Stat.-Simul. Comput.* **2017**, *46*, 6261–6282. [CrossRef]
35. Deng, Z.; Zhu, X.; Cheng, D.; Zong, M.; Zhang, S. Efficient kNN classification algorithm for big data. *Neurocomputing* **2016**, *195*, 143–148. [CrossRef]
36. Iranzo, A.; Boillat, P.; Rosa, F. Validation of a three dimensional PEM fuel cell CFD model using local liquid water distributions measured with neutron imaging. *Int. J. Hydrogen Energy* **2014**, *39*, 7089–7099. [CrossRef]
37. Lan, H.; Yang, L.; Zheng, F.; Zong, C.; Wu, S.; Song, X. Analysis and optimization of high temperature proton exchange membrane (HT-PEM) fuel cell based on surrogate model. *Int. J. Hydrogen Energy* **2020**, *45*, 12501–12513. [CrossRef]
38. Sánchez A.V.D. Advanced support vector machines and kernel methods. *Neurocomputing* **2003**, *55*, 5–20. [CrossRef]
39. Legala, A.; Zhao, J.; Li, X. Machine learning modeling for proton exchange membrane fuel cell performance. *Energy AI* **2022**, *10*, 100183. [CrossRef]