IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 20, 2021, accepted March 4, 2021, date of publication March 9, 2021, date of current version March 17, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3064684*

# A Hybrid Method for Remaining Useful Life Prediction of Proton Exchange Membrane Fuel Cell Stack

**FU-KWUN WANG[1,2], ZEMENU ENDALAMAW AMOGNE[1], AND JIA-HONG CHOU[1]**
[1]Department of Industrial Management, National Taiwan University of Science and Technology, Taipei 10607, Taiwan
[2]Department of Business Administration, Asia University, Taichung 41354, Taiwan

Corresponding author: Fu-Kwun Wang (fukwun@mail.ntust.edu.tw)

This work was supported by the Ministry of Science and Technology, Taiwan under Grant MOST-107-2221-E011-100-MY3.

**ABSTRACT** Proton exchange membrane fuel cell (PEMFC) is a clean and efficient alternative technology for transport applications. The degradation analysis of the PEFMC stack plays a vital role in electric vehicles. We propose a hybrid method based on a deep neural network model, which uses the Monte Carlo dropout approach called MC-DNN and a sparse autoencoder model to analyze the power degradation trend of the PEMFC stack. The sparse autoencoder can map high-dimensional data space to low-dimensional latent space and significantly reduce noise data. Under static and dynamic operating conditions, using two experimental PEMFC stack datasets the predictive performance of our proposed model is compared with some published models. The results show that the MC-DNN model is better than other models. Regarding the remaining useful life (RUL) prediction, the proposed model can obtain more accurate results under different training lengths, and the relative error between 0.19% and 1.82%. In addition, the prediction interval of the predicted RUL is derived by using the MC dropout approach.

**INDEX TERMS** Deep neural network model, Monte Carlo dropout approach, remaining useful life prediction, sparse autoencoder model.

## I. INTRODUCTION

A proton exchange membrane fuel cell (PEMFC) is an electrochemical method that rapidly converts chemical energy stored in a hydrogen tank into electrical energy. The main application is focused on transportation, which has high power density and excellent dynamic characteristics compared to other types of fuel cells. The two major prognostics jobs in prognostic and health management are to predict degradation trends and estimate the remaining useful life (RUL). The stack aging indexes such as total voltage, power, polarization curves, and electrochemical impedance spectroscopy (EIS) can be used to characterize the degradation trend of PEMFCs (see the first column of Table 1). Also, many degradation model parameters based on the measurement data can be used as the degradation indexes. The stack voltage and the stack power are currently the most commonly used measurement-based degradation indexes for PEMFCs.

PEMFC prognostics methods are classified as model-based methods, model-free methods, and hybrid methods [1]. Since the power generation system of the PEMFC is a very complex nonlinear multi-dimensional dynamic strongly-coupled machine, the principle of fuel cell generator and internal chemical reaction process has not been thoroughly studied yet, so model-based PEMFC aging prediction is complicated.

Model-based methods, such as the mechanism degradation model, the empirical degradation model [2]–[4], the semi-mechanism degradation model, or the semi-empirical degradation model [5]–[7], have been used to obtain the prognostics of PEMFCs. Chen *et al.* [8] present an aging prediction method based on a voltage aging model and an extended Kalman filter algorithm. The experimental data of postal fuel cell electric vehicles verified the feasibility of the method. When the training data set exceeds 45 hours, the average mean relative error in the prediction phase will be reduced to 0.68%.

The model-free methods such as data-based method, machine learning such as support vector regression (SVR), deep learning, and signal processing algorithm can be used to implement the degradation trend prediction and the RUL estimation for PEMFCs. Liu *et al.* [9] present a data-based

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao-Sheng Si.

**TABLE 1. Previous works on PEFMC prognostic based on measurement degradation indexes.**

| Degradation index | Method | Type of study | Reference |
|---|---|---|---|
| Voltage | W-GDMH | Short-term prognostics on FC1 & FC2 | [9] |
| | ARIMA-LSTM | Degradation model on 2 groups of FCs | [14] |
| | LSSVM-RPF | RUL prediction on FC1 | [16] |
| | AutoML-AUKF | Long-term degradation trend prediction on FC1 &FC2 | [17] |
| | PF | RUL prediction on FC1 & FC2 | [3] |
| | EKF | RUL prediction framework on FC1 | [25] |
| | BG-PF | SOH and RUL prediction on FC1 & FC2 | [26] |
| Power | DWT | Online power prediction for FC1 & FC2 | [13] |
| | RSVAR | RUL prediction framework based on higher-order regime | [27] |
| | DAE-LS-SVMR | dynamics on FC1 & FC2 | [28] |
| | Improved PF | PEMFC power prediction on FC1 & FC2 | [29] |
| | CIDPF | RUL prognostic tool on FC1 & FC2 | [30] |
| Polarization curve | Linear regression | Pattern-recognition-based diagnosis approach on FC1 & FC2 | [31-33] |
| and EIS such as fuel | PF | RUL prediction model on FC1 & FC2 | [3] [34-35] |
| cell ohmic losses, | GPSS | A model for inferring the RUL of PEMFC on FC1 | [5] |
| reaction activity | EKF | RUL prediction framework on FC1 & FC2 | [25] [36] |
| losses, and reactants | BG-PF | SOH and RUL prediction on FC1 & FC2 | [26] |
| mass transfer losses | UPF | Novel HI for prognostic on FC1 | [37] |
| | EKF-INFORM | Observer-based prognostic algorithm on variable load | [38] |
| | AUKF | SOH and RUL prediction model on FC1 | [39] |

Note: W-GDMH = wavelet and group method of data handling; LSSVM-RPF = least square support vector machine-regularized particle filter; AUKF = adaptive unscented Kalman filter; BG-PF = bond graph-particle filter; DWT = discrete wavelet transform; RSVAR = regime switching vector autoregressive; DAE-LS-SVMR = deep auto encoder-least square-support vector machine regression; UPF = unscented particle filter; CIDPF = characterization induced disturbances in particle filter; GPSS = Gaussian process state space; UPF = unscented particle filter; EKF-IFORM = extended Kalman filter-inverse first order reliability method; AUKF = adaptive unscented Kalman filter; FC1 and FC2 are two experimental PEMFC stacks [40].

method based on the group method of data handling and the wavelet analysis for a short-term PEMFC prognostic. Liu *et al.* [10] report that the adaptive neuro-fuzzy inference system (ANFIS) with fuzzy c-means (ANFIS-FCM) strategy outperforms several methods such as the Elman neural network, the group method of data handling, and the wavelet decomposition approach in the short-term prognostics. Ma *et al.* [11] use a grid long short-term memory (G-LSTM) recurrent neural network (RNN) to study the PEFMC degradation. A method based-on sparse auto-encoder (SAE) and deep neural network (DNN) is applied for the RUL prediction of a PEMFC system [12]. When the training set length is set to 500 hours, the prediction accuracy can reach 99.68%. A discrete wavelet transform (DWT) approach is used as an online prognostic [13]. Using power signals of two PEMFC stacks under static and dynamic operating conditions for RUL prediction, the results show that prediction error is less than 3%. A data fusion method based on LSTM and auto-regressive integrated moving average (ARIMA) is used to predict PEMFC degradation in transportation applications with a minimum root mean square error (RMSE) value of 0.0039 [14].

Hybrid methods combining the model-based and model-free methods with different hybrid strategies can be found in [15]–[23]. Liu *et al.* [23] proposed a multi-scale hybrid degradation index that combines the degradation of membranes and electrodes to estimate the aging state of PEMFCs. They reported that when the prediction length is less than 400 h, the automated machine learning algorithm can provide an estimation error of less than 3.5%.

Some measurement-based degradation indexes have been used in the RUL prediction, such as a particle filter (PF) [24], an extended Kalman filter (EKF) [25], a PF-based hybrid method [26]. The summary of previous studies on PEMFC using the measurement-based degradation indexes is presented in Table 1.

However, most researches in the RUL analysis of PEMFCs are not dealing with their prediction uncertainty. Deep neural networks with the Monte-Carlo (MC) dropout method or variational inference can avoid overfitting and offer uncertainty estimates. In this study, we propose a hybrid method based on a deep neural network model with an MC dropout approach called MC-DNN and an SAE model to analyze the degradation trend of the PEMFC stack, where power is used as an aging index. In addition, we apply the rolling prediction method to the MC-DNN model to obtain RUL prediction. Using the SVR model, a multistep-ahead prediction of input features such as total voltage and current can be derived. The main contributions of our proposed method include:

1) Extracting high relevant features by compressing the input dataset to low dimensions;

2) To obtain the estimated value, our model also provides uncertainty of the PEMFC stack RUL; and

3) The differential evolution (DE) algorithm used in the MC-DNN model improves prediction performance by adjusting hyperparameters.

The overall organization of the paper is as follows. Section 2 introduces our proposed method. Section 3 describes the experimental data. Section 4 discusses the analysis results. Finally, Section 5 concludes the paper and provides directions for future work.

## II. DEEP NEURAL NETWORK WITH MC DROPOUT

### A. DEEP NEURAL NETWORK

Deep learning based on the artificial neural network is run under a deep architecture or hierarchy, consisting of multiple hidden layers, capturing the high-level abstraction behind data and characterizing the complicated nonlinear relationship between inputs and targets. Such a deep neural network has been successfully applied for several regression and classification systems, including speech recognition, natural language processing, and music information retrieval [41]–[44]. Shrestha and Mahmood [45] provide a thorough overview of the neural networks and deep neural networks. The DNN receives data at the input layer, merges the input data with the weights in each node, and transforms it in a nonlinear manner. The final output acts accordingly at the output layer by calculating the average gradient and adjusting the weights and activation.

The output of the $k^{\text{th}}$ layer is given by

$$\mathrm{A}_k = \mathrm{g}(\mathrm{W}_k \mathrm{A}_{k-1} + \mathrm{b}_k) \tag{1}$$

where $k = 1, 2, \ldots, K$; $\mathrm{A}_{k-1}$ is the input datum of the $k^{\text{th}}$ layer; $\mathrm{W}_k$ is the connection weight matrix between the $(k-1)$ layer and the $k$ layer; $\mathrm{b_k}$ is the bias of the $k^{\text{th}}$ layer; g represents the activation function of the $k^{\text{th}}$ layer. Rectified linear unit (ReLU) is usually employed as an activation function of neurons in DNNs, where $\mathrm{ReLU(A)} = \max\{0,\mathrm{A}\}$. Figure 1 depicts a DNN with two hidden layers.

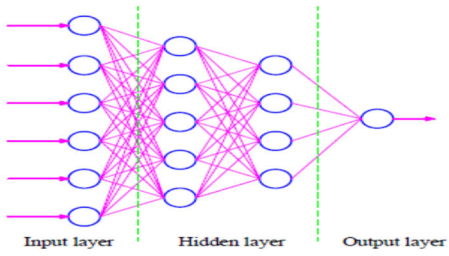


**FIGURE 1.** The architecture of a deep neural network with two hidden layers.

### B. MC DROPOUT

Besides letting the network predict future data better (reduce overfitting), the dropout also enables us to obtain the model uncertainty [46], [47]. While predicting new data, instead of using all neurons (disabling dropout layers), we can keep dropout enabled and predict multiple times (see Figure 2). In the MC dropout approach, the dropout is applied at both training and test time. Given the same data point, our model could predict different values depending on which nodes are chosen to be kept. During test time, the prediction is not deterministic, as in the case of ordinary dropout. MC dropout generates random predictions and interprets them as samples from the probabilistic distribution, and it is called Bayesian interpretation. By running multiple forward passes through the model with a different dropout optimization method mask every time, we can model the uncertainty with the MC dropout approach.

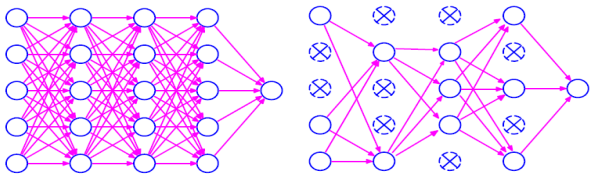


**FIGURE 2.** Neural network with and without dropout.

A sample of possible model output for sample t with a given trained neural network model with dropout $f_h$ can be obtained as:

$$f_h^{d_0}(t), f_h^{d_1}(t), \ldots f_h^{d_M}(t) \tag{2}$$

where $d_i$ is the dropout mask and $M$ is the number of inferences. Predictive posterior mean $(p)$ and an estimate of the uncertainty $(u)$ of the model regarding $t$ can be computed by using the following two formulas [46], [47].

$$p = \frac{1}{M}\sum_{i=0}^{M}\left(f_h^{d_i}(t)\right) \tag{3}$$

$$u = \frac{1}{M}\sum_{i=0}^{M}\left[f_h^{d_i}(t) - p\right]^2 \tag{4}$$

### C. PROPOSED MODEL

The proposed method is based on the DNN model with MC dropout for PEMFC aging analysis, as shown in Figure 3.

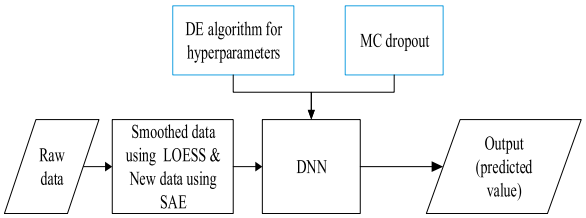


**FIGURE 3.** Architecture of the proposed method.

The details of the analysis steps are given as follows:

*Step 1:* Locally estimated scatterplot smoothing (LOESS) combines much of the simplicity of linear least squares regression with the flexibility of nonlinear regression [48]. LOESS does this by fitting simple models to localized subsets of the data to build up a function that describes the deterministic part of the data's variation, point by point. LOESS is essentially used to visually assess the relationship between two variables and is especially useful for large datasets, where trends can be hard to visualize.

**TABLE 2.** Parameters used for SAE model.

| Structural parameters | Input layer size | Hidden size | Output layer size | Activation function |
|---|---|---|---|---|
| | 25 | 10 | 25 | ReLu |
| Learning parameters | Batch size | Learning rate | Optimizer | Sparse parameter |
| | 100 | 0.01 | Adam | 0.001 |

*Step 2:* Sparse auto-encoder is an extension of the auto-encoder, which can learn relatively sparse features by introducing a sparse penalty (see Figure 4) [49].

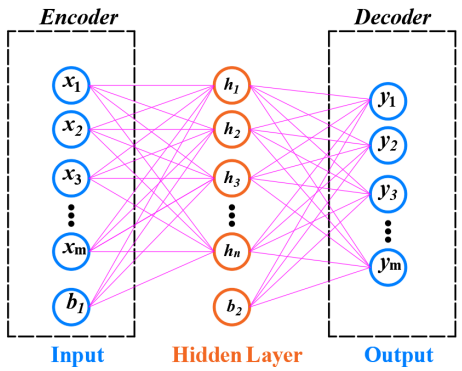


**FIGURE 4.** Structure of sparse auto-encoder.

The difference between sparse autoencoder and autoencoder is the sparsity penalty. The loss function in SAE most involves penalized activations in the hidden layer. The activation can reduce the overfitting problem to reach the Sparsity. There are different ways to construct the sparse penalty for SAE, such as L1 regularization and the KL-divergence-based approach. In this paper, we use L1 regularization, which is widely used in deep learning. The SAE improves the efficiency of traditional autoencoder algorithms. The parameters of the SAE in this study are shown in Table 2.

The input layer, which is the encoder layer, can compress the input data points between the input and hidden layers. The original data points will be extracted by the linear layer with activation function $H_j$, as shown in Equation (5).

$$H_j = ReLU\left(\sum_{i=1}^{k} W_{ij}x_i + b_j\right) \tag{5}$$

where $i$ is the number of neurons of the input layer, $j$ is the number of neurons in the hidden layer, $x$ is the input data points, $H_i$ is the hidden layer's output data, $W_{ij}$ is defined as weight and $b_j$ is the bias. The output data of the hidden layer with the ReLU activation function. Moreover, L1 regularization is used to compute the weight in the hidden layer, where L1 regularization is defined by

$$L1 = \lambda \left\| W_{ij} \right\|, \quad \frac{dL1}{dw} = \frac{\left\| W_{ij} \right\|}{dw} \tag{6}$$

where $\lambda$ is penalty parameter or sparse parameter (see Figure 5) [50].

The input data go through the hidden layer with L1 regularization. The weight, which is $W_{ij}$ is controlled by L1 regulation. Through forward and backpropagation, the output of the hidden layer would be changed to close 1 or $-1$. The output layer, which is the decoder layer, uses a linear layer with an ReLU activation function where $O_i = ReLU\left(\sum_{j=1}^{k} W_{ji}H_j + b_i\right)$. The decoder layer can reconstruct the data points to the original features generating similar data points with input data. The features can be reduced by SAE, which means the output data have a higher correlation.

*Step 3:* The normalized SAE output data is used to train the DNN model with two hidden layers to predict the degradation trend.

*Step 4:* The optimal hyperparameters are found from the DE algorithm, and the MC dropout approach is used to obtain the prediction intervals.

The procedure of the DE algorithm [51] for obtaining the optimal hyperparameters is given as follows.

*Step 1:* Set up the hyperparameters range, the number of parameter vectors, the differential weighting factor, and the crossover probability.

*Step 2:* Define the fitness function as the mean absolute percentage error (MAPE):

$$MAPE = \frac{\sum_{t=1}^{n} |(A_t - F_t)/A_t|}{n} \times 100\%, \tag{7}$$

where $A_i$ is the actual value at the period $i$, $F_i$ is the predicted value at the period $i$, and $n$ is the number of periods used in the calculation.

*Step 3:* Perform the proposed model on each individual in the population and calculate the MAPE.

*Step 4:* If the convergence rule or the maximum iteration number (=100) is reached, the estimated values of hyperparameters such as the number of neurons and the dropout rate are output.

When the optimal hyperparameters are obtained, these values will be used as input parameters in the proposed model for the training and test data.

Regarding the RUL prediction, RUL is defined as the time span from the current time to the end of life (EOL), where a certain amount of power losses gives EOL. That is, it can be expressed as $RUL_t = t_{EOL} - t_{sp}$, where $RUL_t$ is the remaining life at $t_{sp}$, $t_{EOL}$ is the end of life, and $t_{sp}$ is the time at the starting point for RUL prediction. We consider three input features such as time, voltage, and current to predict a PEMFC stack's power. Because input features after the starting time are unknown, a rolling forecasting approach based on the SVR model is used to obtain the multistep-ahead prediction for the input features such as voltage and current. The predicted voltage and current values are then used for power prediction

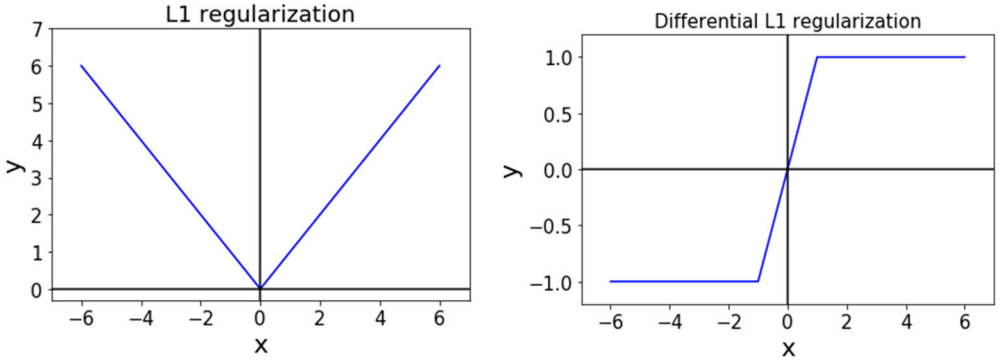


**FIGURE 5.** L1 regularization and differential L1 regularization.

using the MC-DNN model. Figure 6 provides the structure of the RUL prediction using the MC-DNN model with a rolling forecasting approach. When the predicted power reaches the EOL value, the predicted RUL is obtained. Note the number of multistep-ahead forecasts affects the prediction performance.

## III. EXPERIMENTAL EVALUATION

### A. DATA DESCRIPTION

We verify the proposed method using data from the two experimental PEMFC stacks of the IEEE PHM 2014 Data Challenge [40]. The maximum setting of the test bench for full cells is 1 kW. Each stack has five fuel cells. Two aging tests were performed at a constant current of 70A and a dynamic current of 70A with 7A oscillation. Several parameters were collected, including single-cell and stack voltages, temperatures, gas flows, air, and hydrogen hygrometry rates. Other characteristics and experimental operating parameters of the PEMFC stack are provided in Table 3.

**TABLE 3.** PEFMC stack characteristics and experimental operating parameters.

| Parameter | Value |
|---|---|
| Each cell active area | 100 $cm^2$ |
| Cell number | 5 |
| Nominal voltage per cell | 0.6 V |
| Cooling temperature | $20^0$C ~ $80^0$C |
| Cooling flow | 0 ~ 10 l/min |
| Gas temperature | $20^0$C ~ $80^0$C |
| Gas humidification | 0 ~ 100 % RH |
| Gas pressure | 0 ~ 2 bars |
| Air flow | 0 ~ 100 l/min |
| H2 flow | 0 ~ 30 l/min |
| Current | 0 ~ 300 A |

The total of 143862 sets called FC1 under the constant current testing condition are recorded from 0 h to 1054 h. For the dynamic current testing condition, a total of 127370 sets are recorded from 0 h to 1020 h. In this study, we reconstructed the raw data by using the mean of each variable per hour. Thus, 1055 sets and 1021 sets with 25 variables are used for further analysis. The plot of power for FC1 is shown in Figure 7(a). The plots of power for FC2 with raw data shows five outliers (see Figure 7(b). These outliers are replaced by a smoothing method called locally estimated scatterplot smoothing (LOESS).

### B. HYPERPARAMETERS SETTING

The DNN model with two hidden layers is configured as follows: the number of neurons in the first and second hidden layers are 15 and 12, respectively, batch size = 10, dropout rate = 0.002, the number of epochs = 5000, activation function = ReLU, the loss function is mean square error (MSE), and the optimizer is Adam, in which $MSE = \frac{\sum_{i=1}^{n}(A_i - F_i)^2}{n}$, where $A_i$ is the actual value at the period $i$, $F_i$ is the predicted value at the period $i$, and $n$ is the number of periods. This setting is used for RUL prediction when the training length = 550. The other analysis under different training lengths will be provided in the following sections. The system runs on the Visual Studio Code 1.43.0 environment with the Torch library. The computer configuration is as follows: Memory: 8.00 GiB; Processor: Intel Core i7-7700HQ CPU with Intel HD Graphics 630 and NVIDIA GeForce GX 1050; OS: Windows NT x64; Disk: 1050 GB.

## IV. ANALYSIS RESULTS

### A. POWER DEGRADATION ANALYSIS

The PEMFC state of health can be monitored by different parameters like voltage, air and hydrogen stoichiometry rates, and temperature. Under these health monitoring parameters, the PEMFC stack under dynamic or constant operating conditions goes through aging in the experiment. Figure 8 depicts the behavior of some monitoring variables for the PEMFC stack under the experiment. From the figure, it can be seen that, as time increases, the power drops, which indicates the aging of the PEMFC stack.

The training lengths are set to 500 hours, 600 hours, and 700 hours, respectively, and the remaining time is used as the test data set. The three test lengths are 655 hours, 555 hours, and 455 hours for FC1 and 520 hours, 420 hours, and 320 hours for FC2. Three measures such as MSE, root mean square error (RMSE), and MAE are used to compare the prediction performance of different models, in which

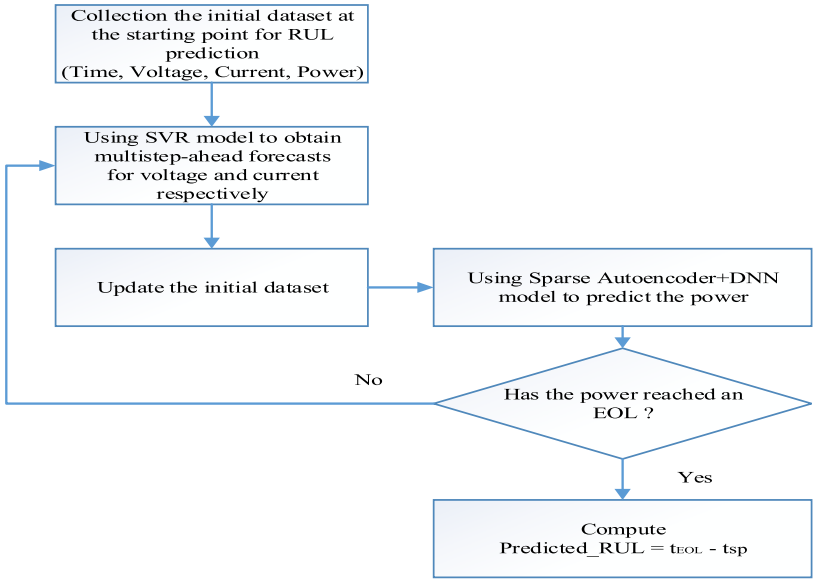


**FIGURE 6.** Structure of the RUL prediction.

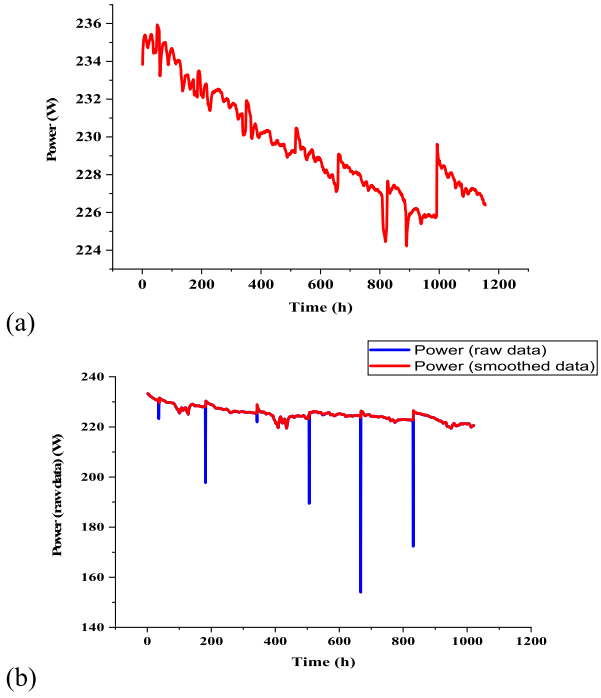


**FIGURE 7.** Power degradation of the PEMFC stack: (a) FC1 (raw data) and (b) FC2 (raw data & smoothed data).

$MAE = \frac{\sum_{i=1}^{n} |A_i - F_i|}{n}$, where $A_i$ is the actual value at the period $i$, $F_i$ is the predicted value at the period $i$, and $n$ is the number of periods.

The three models' prediction results under the three different training lengths with their batch size and dropout value for FC1 are shown in Table 4. The DNN model is based on raw data. The SAE-DNN model is with the optimal parameter of batch size. The MC-DNN model is with two hidden layers and the optimal parameters of batch size and dropout. When the training length is set to 600 h, the MC-DNN model with a batch size of 27 and a dropout value of 0.04 can provide an MSE value of 0.0015. Compared with the DNN model and the SAE-DNN with a batch size of 27, the MSE values are 0.0114 and 0.0234. The MC-DNN model with the optimal parameters outperforms the DNN model and SAE-DNN model in MSE, RMSE, and MAE in all training lengths. Besides, we found that the prediction performance is affected by the training length.

We compare our proposed model with other models using the FC2 data for the stack's dynamic operating condition. The MC-DNN model with two hidden layers and the batch sizes of 25, 24, and 10 for 500 h, 600 h, and 700 h training lengths respectively, results in an improved accuracy compared to the SAE-DNN model with a dropout value of 0.4 [12]. It should be noted that Liu *et al.* [12] do not provide the number of hidden layers and the batch size in the DNN model. Moreover, the dropout parameter is derived from the DE algorithm. The dropout values of 0.1, 0.04, and 0.006 are found for 500 h, 600 h, and 700 h training lengths, respectively. In Table 5, we found that our proposed model outperforms the other two models in terms of MSE, RMSE, and MAE.

When the training length is at 700 h, our model with a batch size of 10 and a dropout value of 0.006 provides the best MSE, RMSE, and MAE values of 0.0001, 0.0120, and 0.0082, respectively (see Table 5). The result shows that the prediction performance of the MC-DNN model with the optimized parameters provides a better result compared to the other two models.

### B. RUL PREDICTION

For the PEMFC stack, the US Department of Energy (DoE) pointed out that PEMFC cannot achieve its function when the initial power is reduced by 10% [52]. However, the testing time of FC2 is limited. Here, 5 power drops were considered: 3.5 %, 4 %, 4.5 %, 5 %, and 5.5 % of its initial power for FC2.

**TABLE 4.** Prediction results of different model on FC1.

| Method | Training length [0,500h] | | | Training length [0,600h] | | | Training length [0,700h] | | |
|---|---|---|---|---|---|---|---|---|---|
| | MSE | RMSE | MAE | MSE | RMSE | MAE | MSE | RMSE | MAE |
| DNN (batch_size) | 0.0327 (16) | 0.1807 | 0.1548 | 0.0114 (27) | 0.1067 | 0.0879 | 0.0408 (17) | 0.2020 | 0.1723 |
| SAE-DNN (batch_size) | 0.0093 (16) | 0.0965 | 0.0863 | 0.0274 (27) | 0.1655 | 0.1604 | 0.0234 (17) | 0.1531 | 0.1421 |
| MC-DNN (batch_size, dropout) | 0.0017 (16,0.01) | 0.0415 | 0.0399 | 0.0015 (27,0.04) | 0.0382 | 0.0338 | 0.0024 (17,0.03) | 0.0485 | 0.0463 |

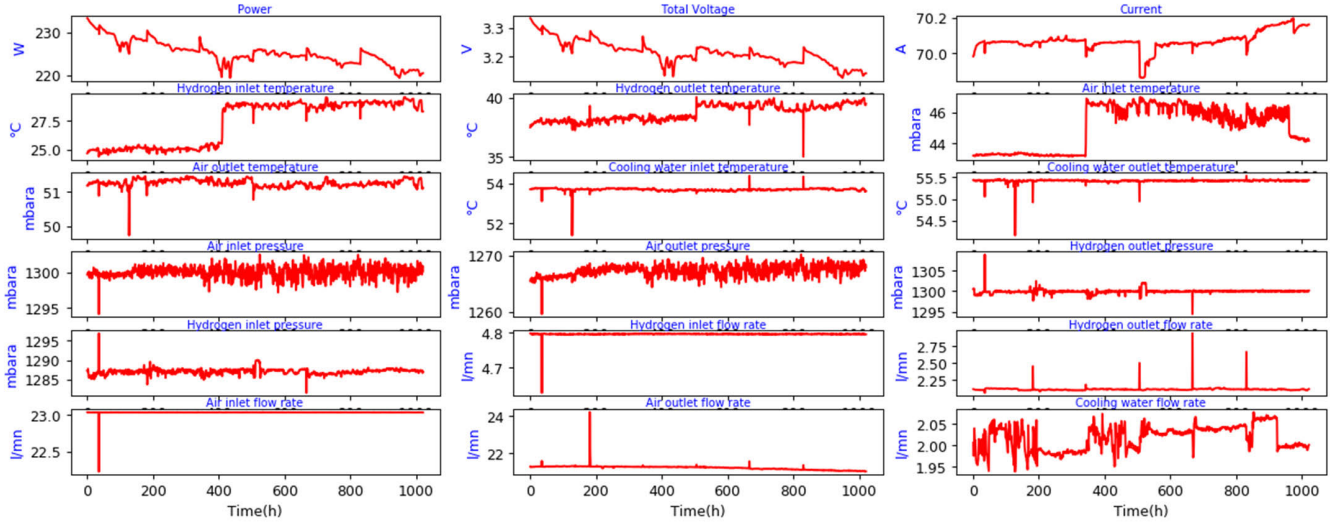


**FIGURE 8.** The behavior of some monitoring parameters for the PEMFC stack.

**TABLE 5.** Prediction results of different model on FC2.

| | Training length [0,500h] | | | Training length [0,600h] | | | Training length [0,700h] | | |
|---|---|---|---|---|---|---|---|---|---|
| Method | MSE | RMSE | MAE | MSE | RMSE | MAE | MSE | RMSE | MAE |
| SAE-DNN [13] | 0.1121 | 0.3348 | 0.2035 | 1.2355 | 1.1115 | 0.5526 | 0.2487 | 0.4987 | 0.3619 |
| SAE-DNN (batch_size) | 0.0822 (25) | 0.2867 | 0.2376 | 0.0244 (24) | 0.1563 | 0.0426 | 0.0181 (10) | 0.1346 | 0.1064 |
| MC-DNN (batch_size, dropout) | 0.0202 (25,0.1) | 0.1422 | 0.0370 | 0.0238 (24,0.04) | 0.1543 | 0.0368 | 0.0001 (10,0.006) | 0.0120 | 0.0082 |

Thus, the remaining useful life prediction before those power losses are starting after 550 h. Using the MC-DNN model with a multistep-ahead prediction (=3), the predictive power and its prediction interval at 95% confidence level are shown in Figure 9. The prediction MSE, RMSE, and MAE values are obtained as 0.2216, 0.1474, and 0.3839, respectively.

To evaluate the performance of different methods, the final Score of 5 RUL estimates is derived by

$$\text{Score} = \frac{\sum_{i=1}^{5} A_i}{5} \tag{8}$$

where $A_i = \begin{cases} \exp(-\ln(0.5) \times (Er_i/5)) \text{ if } Er_i \le 0 \\ \exp(\ln(0.5) \times (Er_i/20)) \text{ if } Er_i > 0 \end{cases}$ and $Er_i = \frac{\text{Acutal\_RUL}_i - \text{Predicted\_RUL}_i}{\text{Acutal\_RUL}_i} \times 100\%$. The positive value of Er means the prediction RUL signaling early than the end of life. The larger the value of Score, the higher the prediction accuracy.

The predicted RUL values and percent errors of the proposed method, a model based on PF [53], and a power model [54] for FC2 under five different failure threshold values are provided in Table 6. Note: failure threshold (FT) is defined as the percentage value of the initial power drop. The average and standard deviation of the percent errors are shown in the last two rows of Table 6, where the percent error values of the PF power, and MC-DNN models are obtained as −9.437, −8.479, and −0.396, respectively. Moreover, the standard deviations of percent error using the PF, power and MC-DNN models are obtained as 19.798, 18.635,

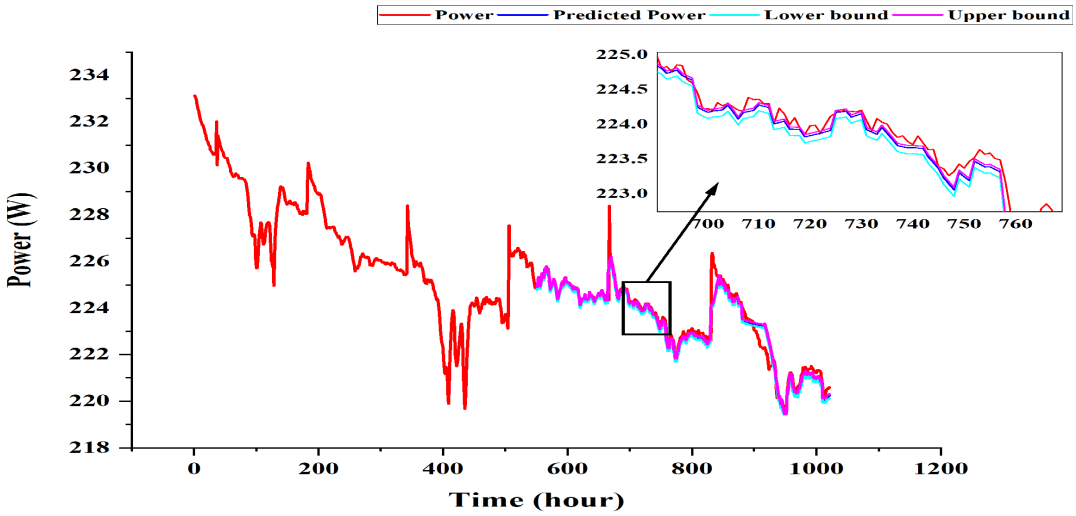


**FIGURE 9.** The prediction intervals of the power at the training length = 550.

**TABLE 6.** Comparison results of different models.

| FT | Actual RUL(h) | Predicted RUL(h) | | | %Err | | |
|---|---|---|---|---|---|---|---|
| | | PF | Power | MC-DNN | PF | Power | MC-DNN |
| 3.5% | 21.4442 | 31.0495 | 30.4107 | 21.8288 | -44.792 | -41.813 | -1.793 |
| 4.0% | 194.1917 | 193.6087 | 194.2174 | 191.9308 | 0.300 | -0.013 | 1.164 |
| 4.5% | 209.7127 | 209.7391 | 209.9835 | 209.7825 | -0.013 | -0.129 | -0.033 |
| 5.0% | 384.3280 | 384.6244 | 384.7621 | 384.1197 | -0.077 | -0.113 | 0.054 |
| 5.5% | 386.7023 | 396.7766 | 387.968 | 392.0059 | -2.605 | -0.327 | -1.371 |
| | Average of Ers | | | | -9.437 | -0.973 | -0.396 |
| | Standard deviation of Ers | | | | 19.798 | 3.722 | 1.191 |

and 1.191, respectively. The results indicate that the proposed method performs better than the PF model in terms of the RUL prediction's accuracy and stability

The PF is based on the MC method and recursive Bayesian estimation method, which employs random sample particles and associated weights to represent posterior probability density. The following two equations formulate the PF method:

$$x_k = f_k(x_{k-1}, w_{k-1}) \tag{9}$$

$$z_k = h_k(x_k, v_k) \tag{10}$$

where $x_k$ is the system state, $z_k$ is the observation, $f_k$ is the state transition function, $h_k$ is the measurement function, $w$ and $v$ are the system's Gaussian noises and the measurements, respectively. In a particle filtering framework, the Bayesian update is processed sequentially with particles having probability information of unknown parameters. The details of the PF algorithm can be found in [53].

The degradation trend model used in [53] is given by

$$x_k = a \cdot (t_k - t_{k-1})^2 + b \cdot (t_k - t_{k-1}) + x_{k-1} \tag{11}$$

where $x_k$ is the current state, $a$ and $b$ are the $2^{nd}$ order polynomial degradation model parameters, $t_k$ and $t_{k-1}$ are the current time step and previous time step, respectively.

The power model used in [54] is given by

$$x(t) = \lambda \cdot t^{\beta} + \sigma_B B(t) \tag{12}$$

where $x(t)$ is the current state, $\lambda$, $\beta$ and $\sigma_B$ are the model parameters, $t$ is the current tme, and $B(t)$ is a standard Brownian motion.

The final scores of PF, power and MC-DNN models are obtained as 0.7379, 0.7847, and 0.9121, respectively (see Table 7). The results show that the MC-DNN model outperforms the other two models in terms of Score. Therefore, the proposed method can offer a more reliable RUL prediction for the PEMFC stack. Note: MC-DNN parameters are the number of neurons = 15 and 12 for the first and second hidden layers, batch_size = 10, dropout = 0.002, epochs = 5000.

**TABLE 7.** The final scores of different models.

| FT | PF | Power | MC-DNN |
|---|---|---|---|
| 3.5% | 0.0020 | 0.0030 | 0.7798 |
| 4.0% | 0.9897 | 0.9982 | 0.9605 |
| 4.5% | 0.9982 | 0.9823 | 0.9954 |
| 5.0% | 0.9973 | 0.9845 | 0.9981 |
| 5.5% | 0.7023 | 0.9556 | 0.8269 |
| Final score | 0.7379 | 0.7847 | 0.9121 |

### C. ROBUSTNESS OF PROPOSED METHOD

To demonstrate the robustness of the proposed MC-DNN model on the RUL prediction of a PEMFC stack, we consider six training lengths, and the failure threshold is set at 935 h. Note: different training lengths for predicting RUL with the

**TABLE 8.** RUL prediction results and errors under different training lengths.

| Training length (h) | Actual RUL (h) | Predicted RUL (h) | Error | Relative error (%) | Prediction interval |
|---|---|---|---|---|---|
| 550 | 385 | 392.0059 | 7.01 | 1.82 | [390.37,394.37] |
| 600 | 335 | 336.8737 | 1.87 | 0.56 | [336.45,337.27] |
| 650 | 285 | 286.6364 | 1.64 | 0.57 | [284.43,287.15] |
| 700 | 235 | 236.3693 | 1.36 | 0.58 | [236.08,236.69] |
| 750 | 185 | 186.2191 | 1.22 | 0.65 | [185.57,186.57] |
| 800 | 135 | 135.2503 | 0.25 | 0.19 | [134.44,135.65] |

batch size in parentheses: 550 (12); 600 (17); 650 (19); 700 (15); 750 (9); 800 (6). The predicted RUL values and their prediction interval at the 95% confidence level are shown in Table 8. For example, when the training length is set at 700 h, the RUL prediction error is 1.36. The results show that the proposed model can provide more accurate RUL prediction under a larger training length.

## V. CONCLUSION

The power degradation analysis of the PEFMC stack plays an important role in electric vehicles. This paper proposes a hybrid method based on a deep neural network model, which has a dropout approach called MC-DNN and a sparse autoencoder model to analyze the power degradation trend of the PEMFC stack and its RUL prediction. We use LOESS to smooth the raw data from the PEMFC stack. The sparse autoencoder model is used to generate new data as input in the DNN model. The DE algorithm is used to obtain the optimal hyperparameters of the DDN model, and the MC dropout approach is used to obtain the prediction interval.

The experimental PEMFC stack is used to verify our proposed model. The results show that the prediction performance of the MC-DNN model with two hidden layers and optimized parameters provides better results compared with the other two models such as DNN and SAE-DNN. In addition, the data of the first 550 hours was used to obtain the trained model, and the five power drop values at 3.5 %, 4 %, 4.5 %, 5 %, and 5.5 % of the initial power value were selected as the end of life. The RUL prediction performance of our proposed model is compared with the score value of the PF model. The results show that our model is better than the PF model with a score value of 0.9121 while the Score values of PF and power model are 0.7379 and 0.7847, respectively. Furthermore, our proposed model can achieve more accurate RUL prediction, and the relative error is between 0.19% and 1.82% under different training lengths of the PEMFC stack. Based on the above analysis results, we can conclude that the proposed model is suitable for the RUL prediction of the PEMFC stack.

The proposed method could be applied to other fuel cells and help practitioners make more reliable RUL prediction.

## REFERENCES

[1] H. Liu, J. Chen, D. Hissel, J. Lu, M. Hou, and Z. Shao, "Prognostics methods and degradation indexes of proton exchange membrane fuel cells: A review," *Renew. Sustain. Energy Rev.*, vol. 123, May 2020, Art. no. 109721.

[2] X. Zhang, D. Yang, M. Luo, and Z. Dong, "Load profile based empirical model for the lifetime prediction of an automotive PEM fuel cell," *Int. J. Hydrogen Energy*, vol. 42, no. 16, pp. 11868–11878, Apr. 2017.

[3] D. Zhou, Y. Wu, F. Gao, E. Breaz, A. Ravey, and A. Miraoui, "Degradation prediction of PEM fuel cell stack based on multiphysical aging model with particle filter approach," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 4041–4052, Jul. 2017.

[4] Y. Wang, Y. Hu, and C. Sun, "Remaining useful life prediction for proton exchange membrane fuel cell using stochastic fusion filtering," *IFAC-PapersOnLine*, vol. 51, no. 21, pp. 158–162, 2018.

[5] Z. Hu, L. Xu, J. Li, M. Ouyang, Z. Song, and H. Huang, "A reconstructed fuel cell life-prediction model for a fuel cell hybrid city bus," *Energy Convers. Manage.*, vol. 156, pp. 723–732, Jan. 2018.

[6] L. Zhu and J. Chen, "Prognostics of PEM fuel cells based on Gaussian process state space models," *Energy*, vol. 149, pp. 63–73, Apr. 2018.

[7] M. Mayur, M. Gerard, P. Schott, and W. Bessler, "Lifetime prediction of a polymer electrolyte membrane fuel cell under automotive load cycling using a physically-based catalyst degradation model," *Energies*, vol. 11, no. 8, p. 2054, Aug. 2018.

[8] K. Chen, S. Laghrouche, and A. Djerdir, "Fuel cell health prognosis using unscented Kalman filter: Postal fuel cell electric vehicles case study," *Int. J. Hydrogen Energy*, vol. 44, no. 3, pp. 1930–1939, Jan. 2019.

[9] H. Liu, J. Chen, M. Hou, Z. Shao, and H. Su, "Data-based short-term prognostics for proton exchange membrane fuel cells," *Int. J. Hydrogen Energy*, vol. 42, no. 32, pp. 20791–20808, Aug. 2017.

[10] H. Liu, J. Chen, D. Hissel, and H. Su, "Short-term prognostics of PEM fuel cells: A comparative and improvement study," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6077–6086, Aug. 2019.

[11] R. Ma, T. Yang, E. Breaz, Z. Li, P. Briois, and F. Gao, "Data-driven proton exchange membrane fuel cell degradation predication through deep learning method," *Appl. Energy*, vol. 231, pp. 102–115, Dec. 2018.

[12] J. Liu, Q. Li, Y. Han, G. Zhang, X. Meng, J. Yu, and W. Chen, "PEMFC residual life prediction using sparse autoencoder-based deep neural network," *IEEE Trans. Transport. Electrific.*, vol. 5, no. 4, pp. 1279–1293, Dec. 2019.

[13] M. Ibrahim, N. Y. Steiner, S. Jemei, and D. Hissel, "Wavelet-based approach for online fuel cell remaining useful lifetime prediction," *IEEE Trans. Ind. Electron.*, vol. 63, no. 8, pp. 5057–5068, Aug. 2016.

[14] R. Ma, Z. Li, E. Breaz, C. Liu, H. Bai, P. Briois, and F. Gao, "Data-fusion prognostics of proton exchange membrane fuel cell degradation," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 4321–4331, Jul. 2019.

[15] D. Zhou, F. Gao, E. Breaz, A. Ravey, and A. Miraoui, "Degradation prediction of PEM fuel cell using a moving window based hybrid prognostic approach," *Energy*, vol. 138, pp. 1175–1186, Nov. 2017.

[16] Y. Cheng, N. Zerhouni, and C. Lu, "A hybrid remaining useful life prognostic method for proton exchange membrane fuel cell," *Int. J. Hydrogen Energy*, vol. 43, no. 27, pp. 12314–12327, Jul. 2018.

[17] D. Zhou, A. Al-Durra, K. Zhang, A. Ravey, and F. Gao, "Online remaining useful lifetime prediction of proton exchange membrane fuel cells using a novel robust methodology," *J. Power Sources*, vol. 399, pp. 314–328, Sep. 2018.

[18] J. Liu and E. Zio, "Prognostics of a multistack PEMFC system with multiagent modeling," *Energy Sci. Eng.*, vol. 7, no. 1, pp. 76–87, Feb. 2019.

[19] H. Liu, J. Chen, D. Hissel, and H. Su, "Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method," *Appl. Energy*, vol. 237, pp. 910–919, Mar. 2019.

[20] D. Zhang, P. Baraldi, C. Cadet, N. Yousfi-Steiner, C. Bérenguer, and E. Zio, "An ensemble of models for integrating dependent sources of information for the prognosis of the remaining useful life of proton exchange membrane fuel cells," *Mech. Syst. Signal Process.*, vol. 124, pp. 479–501, Jun. 2019.

[21] K. Chen, S. Laghrouche, and A. Djerdir, "Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization," *Energy Convers. Manage.*, vol. 195, pp. 810–818, Sep. 2019.
[22] K. Chen, S. Laghrouche, and A. Djerdir, "Degradation model of proton exchange membrane fuel cell based on a novel hybrid method," *Appl. Energy*, vol. 252, Oct. 2019, Art. no. 113439.
[23] H. Liu, J. Chen, D. Hissel, M. Hou, and Z. Shao, "A multi-scale hybrid degradation index for proton exchange membrane fuel cells," *J. Power Sources*, vol. 437, Oct. 2019, Art. no. 226916.
[24] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Prognostics of PEM fuel cell in a particle filtering framework," *Int. J. Hydrogen Energy*, vol. 39, no. 1, pp. 481–494, Jan. 2014.
[25] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, "Extended Kalman filter for prognostic of proton exchange membrane fuel cell," *Appl. Energy*, vol. 164, pp. 220–227, Feb. 2016.
[26] M. S. Jha, M. Bressel, B. Ould-Bouamama, and G. Dauphin-Tanguy, "Particle filter based hybrid prognostics of proton exchange membrane fuel cell in bond graph framework," *Comput. Chem. Eng.*, vol. 95, pp. 216–230, Dec. 2016.
[27] A. Hochstein, H.-I. Ahn, Y. Tat Leung, and M. Denesuk, "Switching vector autoregressive models with higher-order regime dynamics application to prognostics and health management," in *Proc. Int. Conf. Prognostics Health Manage.*, Jun. 2014, pp. 1–10.
[28] S. Hong, L. Sun, J. Yin, T. Y. China, Y. Wang, and W. Zhu, "PEMFC power prediction based on deep auto-encoder and LS-SVMR," in *Proc. IEEE 3rd Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2018, pp. 391–396.
[29] D. Zhang, C. Cadet, C. Bérenguer, and N. Yousfi-Steiner, "Some improvements of particle filtering based prognosis for PEM fuel cells," *IFAC-PapersOnLine*, vol. 49, no. 28, pp. 162–167, 2016.
[30] M. Jouin, R. Gouriveau, D. Hissel, M. C. Pera, and N. Zerhouni, "Remaining useful life estimates of a PEM fuel cell stack by including characterization induced disturbances in a particle filter model," in *Proc. Int. Disc. Hydrogen. Energy Appl.*, May 2014, pp. 1–10.
[31] R. Onanena, L. Oukhellou, D. Candusso, F. Harel, D. Hissel, and P. AkninŸ, "Fuel cells static and dynamic characterizations as tools for the estimation of their ageing time," *Int. J. Hydrogen Energy*, vol. 36, no. 2, pp. 1730–1739, Jan. 2011.
[32] R. Onanena, L. Oukhellou, D. Candusso, A. Same, D. Hissel, and P. Aknin, "Estimation of fuel cell operating time for predictive maintenance strategies," *Int. J. Hydrogen Energy*, vol. 35, no. 15, pp. 8022–8029, Aug. 2010.
[33] W. O. L. Vianna, I. P. de Medeiros, B. S. Aflalo, L. R. Rodrigues, and J. P. P. Malere, "Proton exchange membrane fuel cells (PEMFC) impedance estimation using regression analysis," in *Proc. Int. Conf. Prognostics Health Manage.*, Jun. 2014, pp. 1–8.
[34] M. S. Jha, M. Bressel, B. Ould-Bouamama, G. D. Tanguy, M. Hilairet, and D. Hissel, "Particle filter based prognostics of PEM fuel cell under constant load," *Int. J. Renew. Energy Res.*, vol. 6, no. 2, pp. 644–657, Jan. 2016.
[35] L. Mao, L. Jackson, and T. Jackson, "Investigation of polymer electrolyte membrane fuel cell internal behaviour during long term operation and its use in prognostics," *J. Power Sources*, vol. 362, pp. 39–49, Sep. 2017.
[36] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Fuel cell remaining useful life prediction and uncertainty quantification under an automotive profile," in *Proc. IECON - 42nd Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2016, pp. 5477–5482.
[37] J. Chen, D. Zhou, C. Lyu, and C. Lu, "A novel health indicator for PEMFC state of health estimation and remaining useful life prediction," *Int. J. Hydrogen Energy*, vol. 42, no. 31, pp. 20230–20238, Aug. 2017.
[38] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, "Remaining useful life prediction and uncertainty quantification of proton exchange membrane fuel cell under variable load," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2569–2577, Apr. 2016.
[39] H. Liu, J. Chen, C. Zhu, H. Su, and M. Hou, "Prognostics of proton exchange membrane fuel cells using a model-based method," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 4757–4762, Jul. 2017.
[40] *IEEE PHM 2014 Data Challenge*, La Fédération FCLAB, Belfort, France, 2014.
[41] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017.
[42] K. Nagpal, D. Foote, Y. Liu, P.-H.-C. Chen, E. Wulczyn, F. Tan, N. Olson, J. L. Smith, A. Mohtashamian, J. H. Wren, G. S. Corrado, R. MacDonald, L. H. Peng, M. B. Amin, A. J. Evans, A. R. Sangoi, C. H. Mermel, J. D. Hipp, and M. C. Stumpe, "Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer," *NPJ Digit. Med.*, vol. 2, no. 1, p. 48, Jun. 2019.
[43] S. Nevo, V. Anisimov, G. Elidan, R. El-Yaniv, P. Giencke, Y. Gigi, A. Hassidim, Z. Moshe, M. Schlesinger, G. Shalev, A. Tirumali, A. Wiesel, O. Zlydenko, and Y. Matias, "ML for flood forecasting at scale," 2019, *arXiv:1901.09583*. [Online]. Available: http://arxiv.org/abs/1901.09583
[44] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.
[45] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, Apr. 2019.
[46] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," 2015, *arXiv:1506.02142*. [Online]. Available: http://arxiv.org/abs/1506.02142
[47] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Stat. Assoc.*, vol. 74, no. 368, pp. 829–836, Dec. 1979.
[48] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: An approach to regression analysis by local fitting," *J. Amer. Stat. Assoc.*, vol. 83, no. 403, pp. 596–610, Sep. 1988.
[49] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, vol. 89, pp. 171–178, Jul. 2016.
[50] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent neural network regularization," 2014, *arXiv:1409.2329*. [Online]. Available: http://arxiv.org/abs/1409.2329
[51] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.
[52] US DOE. (2011). *The Department of Energy Hydrogen and Fuel Cells Program Plan, Technical Report*. [Online]. Available: https://www.energy.gov/sites/prod/files/2014/03/f12/program_plan2011.pdf
[53] J. K. Kimotho, T. Meyer, and W. Sextro, "PEM fuel cell prognostics using particle filter with model parameter adaptation," in *Proc. Int. Conf. Prognostics Health Manage.*, Jun. 2014, pp. 1–6.
[54] X.-S. Si, "An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 5082–5096, Aug. 2015.

**FU-KWUN WANG** received the Ph.D. degree in industrial engineering from Arizona State University, Tempe, USA. He is currently a Distinguished Professor with the Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan. He has published more than 150 articles in these fields. His research interests include reliability engineering, quality control, and predictive analytics.

**ZEMENU ENDALAMAW AMOGNE** is currently pursuing the Ph.D. degree with the Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan. His research interests include predictive maintenance and quality management.

**JIA-HONG CHOU** is currently pursuing the degree with the Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan. His research interests include predictive maintenance and deep learning model.