Data-driven diagnosis of PEM fuel cell: A comparative study

Zhongliang Li, Rachid Outbib, Daniel Hissel, Stefan Giurgea

Data-driven diagnosis of PEM fuel cell: A comparative study

Zhongliang Li[a,b,c,*], Rachid Outbib[a], Daniel Hissel[b,c], Stefan Giurgea[b,c]

[a] *Laboratoire des Sciences de l'Information et des Systemes (LSIS), University of Aix-Marseille, France*
[b] *FEMTO-ST (UMR CNRS 6174), ENERGY Department, University of Franche-Comte, France*
[c] *FCLAB (Fuel Cell Lab) Research Federation, FR CNRS 3539, rue Thierry Mieg, 90010 Belfort Cedex, France*

Abstract

This paper is dedicated to data-driven diagnosis for Polymer Electrolyte Membrane Fuel Cell (PEMFC). More precisely, it deals with water related faults (flooding and membrane drying) by using pattern classification methodologies. Firstly, a method based on physical considerations is defined to label the training data. Secondly, a feature extraction procedure is carried out to pick up the significant features from cell voltages constructed vectors. Finally, classification is adopted in the feature space to realize the fault diagnosis. Various feature extraction and classification methodologies are employed on a 20-cell PEMFC stack. The performances of these methodologies are compared.

Keywords: Fault diagnosis, PEMFC, Water management, Classification, Feature extraction

1. Introduction

PEMFC is a promising alternative power generator, thanks to its high power density, high efficiency and environmental friendly property. However, reliability and durability are still two barriers which block its wide application. Thus, fault diagnosis is an efficient solution to guarantee the safe operation of the fuel cell. Indeed, more serious faults can be avoided with an inchoate fault diagnosis. Furthermore, the diagnosis results can be supplied to the control unit, thus help adjust the control commands to make the fuel cell operate efficiently and safely.

The faults that can occur on PEMFC are of two types: First, the irreversible ones that lead to degeneration of the system (for instance, tearing of membrane). Second, those that can be corrected and supervised. Among the important faults there are those related to water (i.e. flooding and membrane drying). More precisely, these faults have been considered as a major cause of power decay and consequently drawn much attention during the last decades. Hence, the problem of PEMFC has received intensive study and several results have been proposed.

Among the proposed results, a first category is based on the use of analytic models (see for instance [1], [2]). This approach consists in analyzing the residuals that are obtained by comparing measured inputs and outputs with analytical relationships.

A second category is based on expert knowledge, mainly, fuzzy logic (see [3]), Bayesian networks (see [4], [5]), and neural networks (see [6] and references therein) methodologies.

A third category concerns results that are achieved using signal processing. Electrochemical Impedance Spectroscopy (EIS) is considered as a powerful tool in analyzing the behavior of fuel cells and some results were established by using the fact that the impedances of fuel cell stacks in certain frequencies are sensitive to faults ([7] [8]). Besides, and recently, new results direct voltage signal analysis thanks to wavelet were proposed (see [9]).

Other approaches have also been proposed, for instance, those using some novel micro sensors [10], or those using multivariate statistical methods (see for instance [11]).

Generally speaking, and independently of PEMFC, data-driven diagnosis has attracted the interest of many authors. This approach offers the advantage of being more relevant for industrial applications. In this framework, various pattern classification techniques have been widely applied to the diagnosis domain [12] [13]. The common classification procedure usually proceeds two steps: firstly an empirical classifier is established from prior knowledge and history data, which is considered as training process. Then the real time data are processed by the classifier in order to determine whether and which faults occur. Notice that, in addition to the classification, some feature extraction procedures are usually carried out as a previous step to get relevant features from the raw data [14].

This paper is a contribution to the PEMFC diagnosis problem. The proposed approach is data-driven and combines feature extraction and classification. In our strategy, features are firstly extracted from data related to the individual cell voltages. Then, a classifier is used in order to distinguish the features of the possible states (i.e. "normal state", "flooding state" and "membrane drying state"). To do so, we consider for automatic feature extraction four methodologies that are Principal Component Analysis (PCA), Fisher Discrimination Analysis (FDA), Kernel Principal Component Analysis (KPCA), Kernel Fisher Discrimination Analysis (KFDA). On the other hand, we employ three typical classification methodologies which are Gaussian Mixture Model (GMM), k-Nearest Neighbor (kNN),

*Corresponding author. Tel.: +33 (0)3 84 58 36 28, Fax: +33 (0)3 84 58 36 36. E-mail address: zhongliang.li@lsis.org(Z. LI), rachid.outbib@lsis.org(R. OUTBIB), daniel.hissel@univ-fcomte.fr(D. HISSEL), stefan.giurgea@utbm.fr(S. GIURGEA)

Support Vector Machine (SVM). Our aim is to compare the performances of these methodologies, from the point of view of diagnosis precision and computation cost, so as to get a relevant tool for online diagnosis of PEMFC stack.

The paper is organized as follows. In section 2, the concerned fuel cell system and faults are presented. Section 3 is devoted to present methodologies used in this work, including methods for labeling training data, feature extraction, and classification. In section 4, diagnosis results of the concerned PEMFC system are given. The performances of different methods are also compared and discussed in this section. Finally, section 5 gives a conclusion.

2. PEMFC operation

2.1. PEMFC system

A running PEMFC is usually fed continuously with hydrogen on the anode side and with air on the cathode side. On the anode side, hydrogen is oxidized:

$$H_2 \rightarrow 2H^+ + 2e^-$$

With the protons transferred through membrane and electrons through the external circuit, oxygen is reduced on the cathode side:

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$$

With the conversion of chemical energy to electrical energy, the by-product water is generated and expelled mostly with the unreacted air from the cathode side.

To produce a useful voltage or power, many cells have to be connected in series, which is known as a fuel cell stack. In addition to the stack, practical fuel cell systems also contain several other ancillary subsystems: hydrogen supply subsystem, air supply subsystem, temperature management subsystem. As Fig. 1 shows, in the hydrogen supply subsystem and air supply subsystem, gas flows and pressures at stack inlets can be regulated respectively by using mass flow controllers and back-pressure valves. The relative humidity of the input air is regulated by a humidifier. Temperature management subsystem can make the fuel cell stack operate at an appropriate temperature.

Kinds of faults may occur in the fuel cell system. The faults may happen either inside the fuel cell stack or in the ancillary subsystems. The fuel cell stack is the heart of the whole system and the faults in the fuel cell stack can somehow reflect the faults in the ancillary components. Consequently, the diagnosis, especially online diagnosis of fuel cell stack is crucial and thus the focus of this paper.

2.2. Concerned faults

Water management has been considered as one of the most important issues in PEMFC. Flooding and membrane drying are the two main degradation mechanisms that occur when water management is not adequate [15]. The diagnosis of these two typical faults are therefore dedicated to the verification and the comparison of the methodologies used in the paper.

As Fig. 2 shows, a typical PEMFC consists of bipolar plates (BPs), gas diffusion layers (GDLs), catalyst layers (CLs), and a membrane. On both sides of BPs, gas channels (GCs) are grooved for gas flow. In a proper functioning PEMFC, the membrane should keep a certain water content to make the protons transfer through it effectively with low ohmic resistance. Hence, air is usually humidified before being fed into the fuel cells. At the same time, liquid water is generated in the cathode and expelled from the fuel cells by unreacted air and hydrogen or purged at regular intervals. Inside the fuel cell, water travels among different layers and moves between anode and cathode. The water content at the channels and different layers can be influenced by a number of factors, such as gas pressures, gas humidities, gas flow rates, stack temperature, and load current. Inadequate water management may cause the degradation of the PEMFC. On the one hand, dry membrane increases the ohmic losses, thus induces "membrane drying" fault. On the other hand, accumulation of liquid water in the GCs and/or gas porosities of GDLs and CLs, results in "flooding" fault. Excessive liquid water will block the reactant pathways, thus making the fuel cell stack degraded. As water is generated and expelled mostly in the cathode side, flooding happens generally in the cathode side [15].

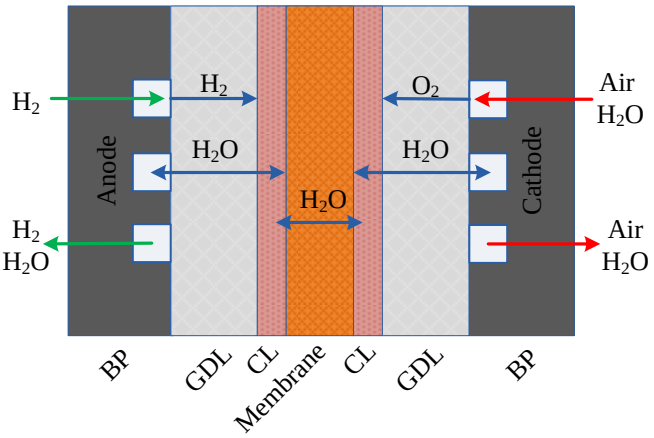

Figure 2: Schematic picture of water movement inside a PEMFC

3. Descriptions of diagnosis methodologies

3.1. Approach principle

The proposed approach in the paper is based on data and contains three steps (see Fig. 3) : data labeling process, model training process, and diagnosis process. The first two processes are off-line, while the third process is on-line.

The employed pattern classification methods belong to supervised ones. In supervised learning, data points for training should be provided with their category labels before the training procedure. Consequently, it is needed to define the classes and label the training data. In the labeling stage, a two-phase pressure drop model combined with statistical analysis is used to achieve this [16] [17]. After data labeling process, the training data can be labeled with three class labels: normal state, flooding fault state, and membrane drying state.

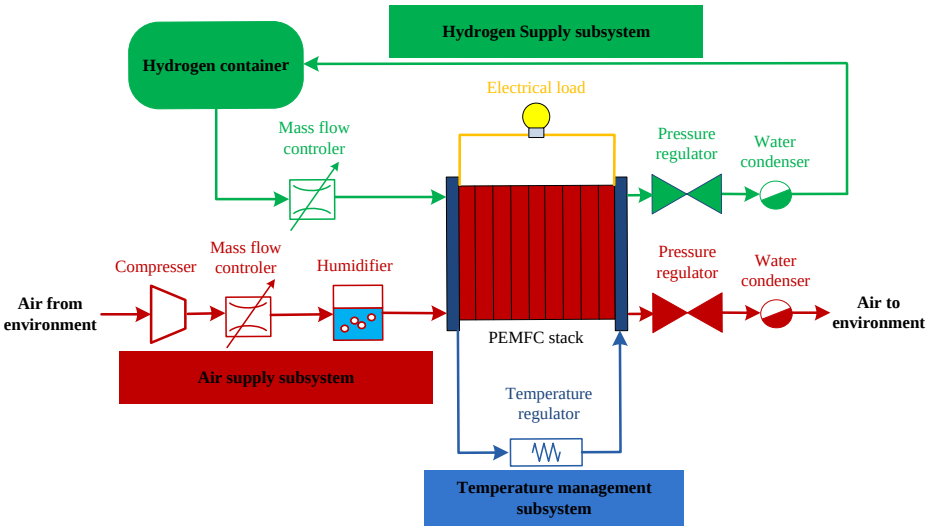

Figure 1: The schematic of PEMFC system

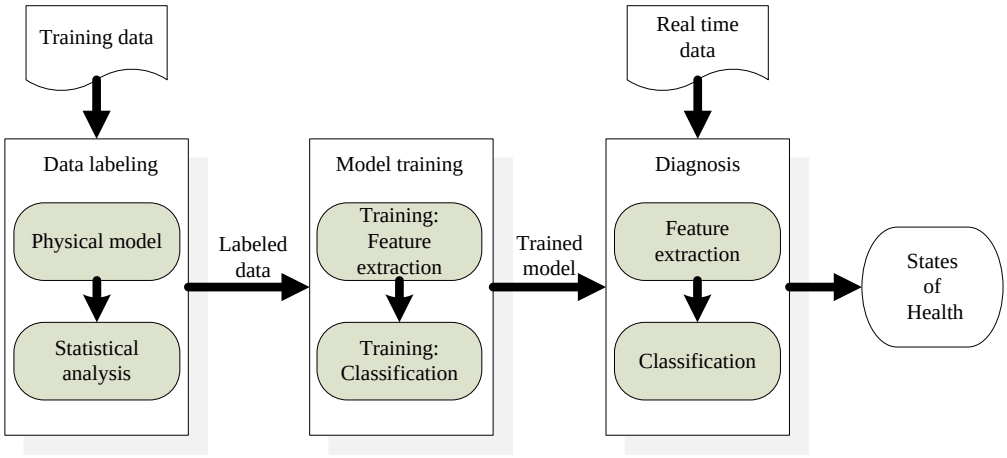

Figure 3: The framework of the diagnosis approach

In the model training and diagnosis stages, the individual cell voltages are the used variables. Cell voltages are chosen here on account of several points: first, fuel cell voltage signals are highly dependent to the current, electrochemical characteristics, temperature, and aging effect. The voltage variances can be used to determine the magnitude of the parameters for the fuel cell model [18]. In other words, the individual cell voltage can be seen as sensors inside the fuel cell stack. Second, the importance of monitoring individual cell voltages is stressed, since the cell with the lowest cell voltage in the stack restricts the maximum power output of the stack [19], it is necessary to monitor every single (or several) cell voltage(s) to some degree. Third, it is observed that the water management faults change the flow distributions of the gases and then make the distribution cell voltages varied [20]. Moreover, the cost of measurement of cell voltages is relatively low. Based on the above points, vectors constructed by individual cell voltages are considered as the original data for diagnosis.

In the model training process, two kinds of model need to be trained, which are feature extraction and classification. In feature extraction, the goal is to find certain projecting vectors to map the original high-dimension vectors to feature space, which is of low-dimension. Notice that in the classification, the classifiers will be trained in the feature space. By the trained classifiers, a new point in feature space can be fixed to one of the three predefined classes. In the literature, there are kinds of methodologies for feature extraction and classification. Hence, in order to make a comparison of the different methods, several relevant and representative methodologies are utilized in the paper.

In the diagnosis process, the real-time cell voltages are sampled and represented by vectors. Then, the feature extraction and classification procedures are respectively carried out by using the models obtained in the training process. Real-time data can be classified into three classes, and the diagnosis is realized accordingly after these two procedures.

3.2. Data labeling

In order to label the training data to the three classes: "normal", "flooding", and "membrane drying", the normal range of liquid water inside the fuel cell must be evaluated. To do so, an approach that combines pressure drop model with statistical analysis is proposed.

3.2.1. Pressure drop model

The pressure drop between inlet and outlet channels is significant of the gases removal out of the fuel cell, and it is relevant to the content of liquid water in the flow fields [15]. Since only the cathode inlet gas is humidified in the studied PEMFC system, and the generated water is mostly expelled from the cathode side, the cathode side is more relevant to the water management issues. Hence, the pressure drop model in the cathode side will be considered.

GCs are grooved on both sides of BPs for gas flow and different structures can be used. Fig. 4 depicts three classic structures: parallel, serpentine and interdigitated flow fields. Notice that the pressure drop model is dependent on the considered structures. For the parallel and serpentine flow fields, the air passes from the GCs, and the pressure drops throughout the GCs. The major part of pressure loss is associated with the frictional losses along the channel pipe [21]. In this case, the pressure drop model based on Darcy law is given by [16]

$$\Delta P = \frac{\mu}{K_{C_0}(1-s)^{n_k}} v_D L_{GC} \tag{1}$$

where v_D is the air flow rate, K_{C_0} is permeability, which is impacted by the sizes and the structures of flow fields. μ denotes viscosity of air and L_{GC} is the length of the channel. $s \in [0, 1)$ is defined as the volume fraction of GC occupied by liquid water, which is a key parameter characterizing the water quantity. n_k is a constant between 4.5 and 5.0 (see [16]).

For the interdigitated flow field, the pressure drop mainly occurs in the GDL. In this case, the pressure drop can be denoted as [17]

$$\Delta P = \frac{150(1-\varepsilon(1-s))^2\mu}{\varepsilon^3(1-s)^3 d_0^2} v_D L_{GDL} \tag{2}$$

where d_0 is the representative diameter of pore in GDL, L_{GDL} is the rib length of the BP. ε is constant that reflects the porosity of GDL. $s \in [0, 1)$ quantifies the portion of pores (in GDL) occupied by liquid water. Undoubtedly, in (1) and (2), s can be considered as a criteria to quantify the water content inside the fuel cells.

3.2.2. A statistical method

From (1) and (2), the quantity $\Delta P/v_D$ can be considered as a function depending on s:

$$W(s) = \frac{\Delta P}{v_D}$$

where

$$W^{(1)}(s) = \frac{\alpha}{(1-s)^{n_k}} \quad \text{and} \quad W^{(2)}(s) = \beta\frac{(1-\varepsilon(1-s))^2}{(1-s)^3} \tag{3}$$

with

$$\alpha = \frac{\mu L_{GC}}{K_0} \quad \text{and} \quad \beta = 150\frac{\mu L_{GDL}}{\varepsilon^3 d_0^2}$$

Clearly, $W^{(1)}$ and $W^{(2)}$ defined by (3) are increasing function for $s \in [0, 1)$. Thus, W can replace s to express water quantity. In normal state, it is considered that the fuel cell can operate in a range of s, so the values of W also distribute in a normal range. If the values of W follow a normal distribution, which is the case in our study (see section 5), a common statistical method "3-sigma" can be used in order to evaluate the limits of W:

$$W_{max} = \bar{W} + 3\sigma_W, W_{min} = \bar{W} - 3\sigma_W \tag{4}$$

where $\bar{W}$ and σ_W are respectively the average value and standard deviation of W samples in normal state.

In the labeling process, W_{max} is the threshold for flooding diagnosis while W_{min} is the one for drying fault diagnosis. The flow chart of the training data labeling procedure can be depicted as Fig. 5.

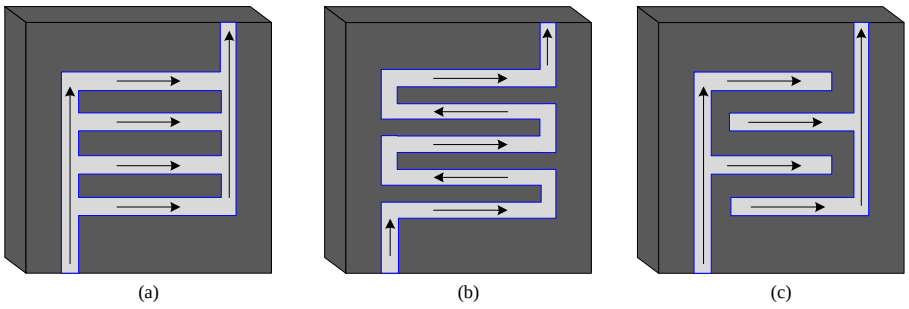

Figure 4: Three kinds of flow field structures. (a) Parallel flow field. (b) Serpentine flow field. (c) interdigitated flow field

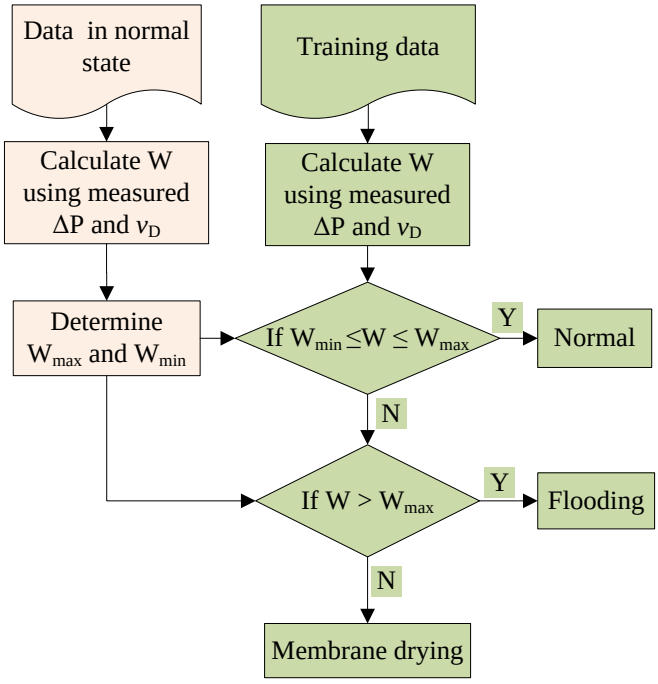

Figure 5: Training data labeling process

3.3. Feature extraction

For high power applications, for instance vehicles, many fuel cells are stacked in series to meet the power requirement. Hence, large dimension data has to be handled when individual cell voltages serve as the variables for diagnosis. In order to reduce the complexities of computations, it is necessary to lower the dimension of cell voltages constructed vectors by some means of feature extraction. At the same time, the feature extraction procedure is motivated to draw useful features for diagnosis. Based on the above considerations, four representative feature extraction methodologies, which can meet these two needs, are presented in this subsection. More precisely, two typical unsupervised and supervised methodologies: PCA, FDA, and their nonlinear forms KPCA and KFDA are considered.

The feature extraction problem can be described as follows: Collect N training samples $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N \in \mathbb{R}^M$, which are distributed in C classes $\{\omega_1, \omega_2, \ldots, \omega_C\}$. The sample number in ω_i is N_i, which satisfies $\sum_{i=1}^{C} N_i = N$. The class index of $\boldsymbol{x}_n$ is denoted by h_n, $h_n \in \{1, 2, \ldots, C\}$. The objective of the training process is to find L ($L < M$) M-dimension unit projecting vectors: $\{\boldsymbol{w}_1, \boldsymbol{w}_2, ..., \boldsymbol{w}_L\}$. With these vectors, a real time sample $\boldsymbol{x}$ can be projected to a L-dimension feature space, the projected vector $\boldsymbol{z}$ is expressed:

$$\boldsymbol{z} = [\boldsymbol{w}_1^T \boldsymbol{x}, \boldsymbol{w}_2^T \boldsymbol{x}, \ldots, \boldsymbol{w}_L^T \boldsymbol{x}]^T \tag{5}$$

3.3.1. PCA

PCA is an unsupervised dimensionality reduction and feature extraction technique that preserves the significant variability information in the original data set. It changes more relevant variables into seldom uncorrelated variables according to the lowest data missing rule [11]. The general PCA procedure is summarized by algorithm 1.

Algorithm 1 PCA

Training:

1: Collect samples $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$.

2: Perform singular value decomposition on covariance matrix:

$$\frac{1}{N} \sum_{n=1}^{N} (\boldsymbol{x}_n - \bar{\boldsymbol{x}})(\boldsymbol{x}_n - \bar{\boldsymbol{x}})^T = \boldsymbol{P} \Lambda \boldsymbol{P}^T \tag{6}$$

where $\boldsymbol{P} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_M]$, $\Lambda = diag(\lambda_1, \ldots, \lambda_M), \lambda_1 \geq \lambda_2, \geq , \ldots, \geq \lambda_M$.

3: Determine the number of principal components L by respecting criteria

$$\frac{\sum_{i=1}^{L} \lambda_i}{\sum_{i=1}^{M} \lambda_i} \geq Th, \frac{\sum_{i=1}^{L-1} \lambda_i}{\sum_{i=1}^{M} \lambda_i} < Th \tag{7}$$

where Th is a pre-set threshold, whose value is near but less than 1 (see [22]).

4: Save vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_L$.

Performing:

Calculate the projected vector of a new sample $\boldsymbol{x}$ as (5).

3.3.2. FDA

FDA is a supervised technique developed for reducing the dimension of the data in hope of obtaining a more manageable classification problem. The objective of FDA is to find mapping vectors that make the data in the same class concentrated

while the data in varied classes separated [23] [24]. FDA can be briefly formulated by algorithm 2.

Algorithm 2 FDA

Training:

1: Collect labeled samples: $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$.

2: Calculate within-class-scatter matrix $\boldsymbol{S}_w$ and between-class-scatter matrix $\boldsymbol{S}_b$.

$$\boldsymbol{S}_w = \sum_{i=1}^{C} \sum_{\boldsymbol{x}_n \in \omega_i} (\boldsymbol{x}_n - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_n - \bar{\boldsymbol{x}}_i)^T$$

$$\boldsymbol{S}_b = \sum_{i=1}^{C} N_i (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})^T$$

where $\bar{\boldsymbol{x}} = \sum_{n=1}^{N} \boldsymbol{x}_n / N$, and $\bar{\boldsymbol{x}}_i = \sum_{\boldsymbol{x}_n \in \omega_i} \boldsymbol{x}_n / N_i$.

3: Find the L eigenvectors of $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ with non-zero eigenvalues: $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_L$.

Performing:

Calculate the projected vector of a new sample as (5).

3.3.3. KPCA

KPCA is an extension of PCA, which aims to solve nonlinear PCA. The key idea of KPCA is intuitive and generic. In general, the nonlinear correlated data can always be mapped to a higher-dimensional space in which they vary linearly via a nonlinear mapping [25]. After that, PCA procedure can be carried out in the new space. Actually, this two-step process can be realized by introducing kernel functions and playing "kernel trick" [26]. KPCA is summarized by algorithm 3.

Algorithm 3 KPCA

Training:

1: Collect $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$.

2: Get $\boldsymbol{K} \in \mathbb{R}^{N \times N} : \boldsymbol{K}_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is a kernal function.

3: Modify $\boldsymbol{K}$

$$\tilde{\boldsymbol{K}} = \boldsymbol{K} - \boldsymbol{1}_N \boldsymbol{K} - \boldsymbol{K}\boldsymbol{1}_N + \boldsymbol{1}_N \boldsymbol{K} \boldsymbol{1}_N \tag{8}$$

where $\boldsymbol{1}_N \in \mathbb{R}^{N \times N}$ with terms all equal to $1/N$.

4: Find L eigenvectors of $\tilde{\boldsymbol{K}}$ with the largest eigenvalues, which are denoted as $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_L \in \mathbb{R}^N$.

Performing:

Calculate the projected vector of a new sample $\boldsymbol{x}$. The l^{th} ($l \in 1, 2, \ldots, L$) element of the projected vector $\boldsymbol{y}$ can be calculated as:

$$y_l = \sum_{n=1}^{N} \alpha_{ln} k(\boldsymbol{x}_n, \boldsymbol{x}) \tag{9}$$

where α_{kn} is the n^{th} element of eigenvector $\boldsymbol{\alpha}_l$.

The kernel functions corresponds to nonlinear mappings. In various kernel functions, Gaussian kernel function is usually the first choice for its high performance in most cases [27]. Hence, this popular kernel function is involved in the paper:

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = exp(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{\sigma}) \tag{10}$$

where σ is a constant that needs to be initialized.

3.3.4. KFDA

As KPCA, the key idea of KFDA is also to map the data to a new space by nonlinear mapping firstly, and then carry out FDA procedure in the new space. Kernel trick helps to realize the KFDA process in the same way as KPCA [28].

Algorithm 4 KFDA

Training:

1: Collect labeled samples: $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$.

2: Get kernel matrix $\boldsymbol{K}$.

3: Modify $\boldsymbol{K}$ as (8)

4: Get matrix $\boldsymbol{U}$:

$$\boldsymbol{U} = diag(\boldsymbol{U}_1, \boldsymbol{U}_2 \ldots, \boldsymbol{U}_C) \tag{11}$$

where $\boldsymbol{U}_i \in \mathbb{R}^{N_i \times N_i}$ with terms all equal to $1/N_i$.

5: Find L eigenvectors of $(\tilde{\boldsymbol{K}}\tilde{\boldsymbol{K}})^{-1}\tilde{\boldsymbol{K}}\boldsymbol{U}\tilde{\boldsymbol{K}}$ with the largest eigenvalues, which are denoted as $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_L \in \mathbb{R}^N$

Performing:

Calculate the projected vector of a new sample $\boldsymbol{x}$. The l^{th} ($l \in 1, 2, \ldots, L$) element of the projected vector $\boldsymbol{z}$ can be calculated as (9).

3.3.5. Remarks on feature extraction methods

1. It was verified in [29] that KFDA is equivalent to KPCA plus FDA. That is, KPCA is performed first then FDA is carried out in the feature space obtained by KPCA.
2. PCA and FDA can be seen as the special situation of KPCA and KFDA using the linear kernel function $k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T \boldsymbol{x}_j$. So FDA can be seen as the procedure of FDA in the PCA mapped space.
3. The performances of both KPCA and KFDA are highly related to the choice of kernel function and the parameters in the kernel function.
4. Both KPCA and KFDA have two drawbacks that the computation time may increase with the number of training samples, and the data patterns in the feature space are rather hard to interpret in the input data space [30].

3.4. Classification

The classification proceeds after feature extraction step. In this step, classifiers are trained in the feature space. The classification procedure can be described as follows: Given N samples $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N \in \mathbb{R}^L$, which are used for training, the samples are distributed in C classes: $\omega_1, \omega_2, \ldots, \omega_C$. The sample number in ω_i is N_i, which satisfies $\sum_{i=1}^{C} N_i = N$. The class index of $\boldsymbol{z}_n$ is denoted by h_n, $h_n \in \{1, 2, \ldots, C\}$. The objective of the training process is to get a classifier. With the classifier, the class index h of a real time sample $\boldsymbol{z}$ can be obtained. In order

to make a comparison among different classifiers, three representative classifiers: GMM, kNN, and SVM are under consideration. Without loss of generality, GMM is a preferable parametric classification method, while kNN and SVM are two typical non-parametric ones. kNN is a widely used method due to its simplicity and flexibility. The remarkable characteristics of SVM, such as good generalization performance, the absence of local minima and the sparse representation of solution, attract much attention in recent years [14].

3.4.1. GMM

GMM is a parametric classification methodology based on Bayes decision theory [31]. The classification is realized by calculating and comparing the class-conditional probabilities $p(z|\omega_i), i = 1, \ldots, C$. In GMM, this density is represented as a weighted sum of R_i component Gaussian densities in the following equation:

$$p(z|\omega_i) = \sum_{j=1}^{R_i} p(c_j|\omega_i) p(z|c_j, \omega_i) \tag{12}$$

where $p(c_j|\omega_i), j = 1, \ldots, R_i$ are the mixture weights, which satisfies $\sum_{j=1}^{R_i} p(c_j|\omega_i) = 1$, $p(z|c_j, \omega_i)$ are the component Gaussian densities. Each component density is a L-variate Gaussian function of the form,

$$p(z|c_j, \omega_i) = \frac{1}{(2\pi)^{M/2}|\Sigma_j|^{1/2}} exp\left\{-\frac{1}{2}(z - \boldsymbol{\mu}_j)^T \Sigma_i (z - \boldsymbol{\mu}_j)\right\} \tag{13}$$

with mean vector $\boldsymbol{\mu}_j$ and covariance matrix Σ_j. Parameters $\boldsymbol{\mu}_j$, Σ_j and $p(c_j)$ are collectively represented by the notation ζ_i:

$$\zeta_i = \{p(c_j), \boldsymbol{\mu}_j, \Sigma_j\} \quad i = 1, \ldots, R_i$$

The configuration of R_i is often determined by the complexity of the data distribution. A complex distribution could be described by a choice of a large R_i. Expectation-Maximization (EM) algorithm, whose details can be found in [32], is adopted to estimate ζ_i. The GMM classification method described by algorithm 5.

Algorithm 5 GMM

Training:

1: Collect labeled samples $z_1, z_2, \ldots, z_N$
2: Initial R_i for $i = 1, \ldots, C$.
3: **for** $i = 1$ to C
Estimate and save ζ_i by using EM algorithm.
end for

Performing:

1: For a new sample z, calculate $p(z|\omega_i, \zeta_i)$.
2: Class index h is assigned:

$$h = arg\left\{\max_{i\in\{1,\ldots,C\}} N_i p(z|\omega_i, \zeta_i)\right\} \tag{14}$$

3.4.2. kNN

kNN is a widely-used nonparametric classifier [24]. In the kNN procedure, the classification decision is based on the N training samples. The training step is needless, and the procedure is given by algorithm 6.

Algorithm 6 kNN

Training:

1: Collect and save labeled samples: $z_1, z_2, \ldots, z_N$.

Performing:

1: For a new sample z, calculate its Euclidean distances to $z_1, z_2, \ldots, z_N$.
2: Find the nearest k neighbors of z that are at the minimum Euclidean distances: $z_1^k, z_2^k, \ldots, z_k^k$, whose class indexes are $h_1^k, h_2^k, \ldots, h_k^k$.
3: z is assigned to a class to which most of the neighbors belong:

$$h = arg\left\{\max_{j\in\{1,\ldots,C\}} \sum_{i=1}^{k} \delta(h_i^k, j)\right\} \tag{15}$$

3.4.3. SVM

SVM is a classification method developed by V. Vapnik [33] and has been widely applied the last two decades. The basic theory comes from binary classification problem. As Fig. 6 shows, there are data samples distributed in two classes, suppose we have some hyperplane which separates the points. Then, SVM looks for the optimal hyperplane with the maximum distance from the nearest training samples. A subset of training samples that lie on the margin are called support vectors.

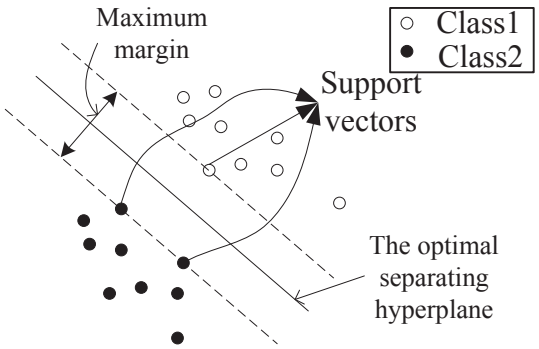

Figure 6: SVM schematic diagram

Algorithm 7 SVM

Training:

1: Collect labeled samples $z_1, z_2, \ldots, z_N$. $g_n \in \{-1, 1\}$, is the class label of sample z_n. Initial D.

2: Solve the quadratic problem:

$$minJ(\boldsymbol{a}) = \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m g_n g_m k(z_n, z_m) - \sum_{n=1}^{N} a_n$$
$$\text{s.t.} \sum_{n=1}^{N} a_n g_n = 0, \quad 0 \le a_n \le D, \quad \text{for} \quad n = 1, 2, \ldots, N \tag{16}$$

where $\boldsymbol{a} = [a_1, a_2, ..., a_N]^T$, $k(z_n, z_m)$ is a kernel function.

3: Save support vectors: $z_1^s, z_2^s, \ldots, z_S^s$ and corresponding g_n and a_n. Samples for which $a_n > 0$ are support vectors.

Performing:

For a new sample z, its class label is determined:

$$g = sign\left\{\sum_{n=1}^{S} a_n^s g_n^s k(z_n^s, z) + b\right\} \tag{17}$$

where

$$b = \frac{1}{S}\sum_{j=1}^{S}\left(g_j^s - \sum_{n=1}^{S} a_n^s g_n^s k(z_n^s, z_j^s)\right)$$

There are several points that need to be emphasized: first, notice that the real computation process is just correlated to the support vectors. This property is central to the practical applicability of SVM. Second, for solving the quadratic programming problem, a practical approach, Sequential Minimal Optimization (SMO) is used in our study [33]. Third, the basic SVM is a binary classifier. To extend the binary classifier to multi-classification situations, a method named "One-Against-One" is adopted in the paper. The technique classifies the classes in pairs by using binary SVM. The final classification is obtained by voting all binary classification results. The details can be found in [34].

4. Results and discussion

4.1. The PEMFC under considerations

A 1 kW test bench is used to test a 20-cell PEMFC stack. Table 1 summarizes the parameters of the investigated fuel cell stack. In the test bench, a number of physical parameters impacting or expressing stack performances can be controlled and monitored. Stack temperature (T_{fc}), stoichiometries of hydrogen and air (S_h, S_a), relative humidity of the inlet air (RH), load current (I) can be set. Inlet and outlet pressures of hydrogen and air (P_h, P_a), stack temperatures, current, stack voltage (V_s) and single cell voltages ($v(1), v(2), \ldots, v(20)$) can be monitored. The sample time of the test bench is 150ms.

The fuel cell temperature is considered to be the value of the temperature at cooling water (demineralized) outlet. The control of the relative humidity of input air is by the means of regulating air dew point temperature (The input air is set the same temperature as fuel cell). Additional details about the test bench and the test protocol have been previously published in [35].

Table 1: The paramenters of the investigated fuel cell stack

Cell area	100 cm^2
Cell number	20
Flow field structure	serpentine
Nominal output power	500 W
Nominal operating temperature	40 °C
Operating temperature region	20-65 °C
Maximum operating pressures	1.5 bar
Anode stoichiometry	2
Cathode stoichiometry	4

4.2. The results of data labeling

Based on the experimental test bench, experiments in normal conditions were firstly carried out. In the experiments, stack temperature is set at T_{fc} = 40°C, the stoichiometries of hydrogen and air were set at the nominal values as table 1, relative humidity RH was been situated between 75% and 98%, which is considered as a normal region. The output current was configured from 40A to 2A. As Fig. 7 shows, current was set at 20 discrete points, for every current point, 100 samples were collected to one group.

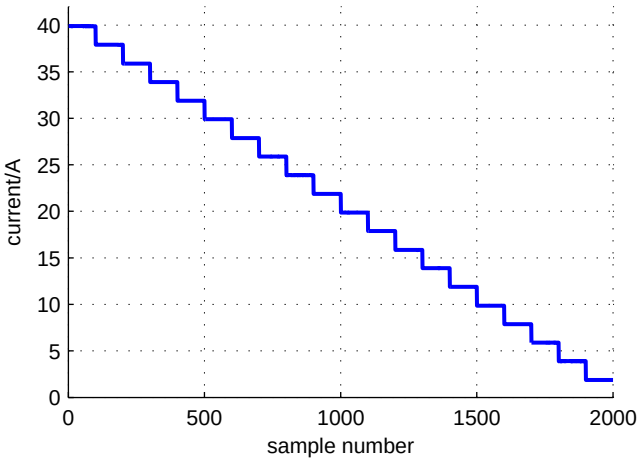

Figure 7: Current evolution in normal experiment

In the normal state experiments, water indicator W is calculated for every sample. The Lilliefors tests are used to test the null hypothesis that W follows a normally distributed population for each group [36]. The null hypotheses are not rejected with significance level 0.05 for all the 20 groups. Hence, it is reasonable to define the up and down limits as (4).

The values W of normal condition in different current points are as in Fig. 8. The up limit W_{max} and down limit W_{min} are also shown in this figure. Since the generated water increases with the rising of current, it can be seen that the bounders of the normal W increase globally with current.

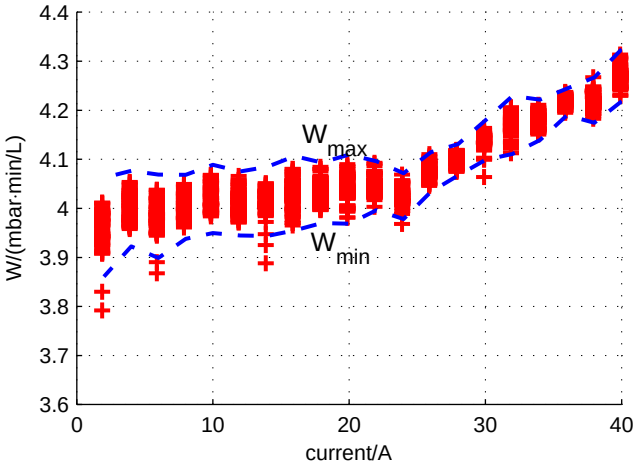

Figure 8: Values of parameter W in normal condition and their limits

In order to obtain the training data, fault experiments were carried out afterwards. In the fault experiments, the output current was fixed at 40A; stack temperature is set at T_{fc} = 40°C; the stoichiometries of hydrogen and air were set at the nominal values as table 1. Relative humidity RH was been situated between 85% and 110%, which would induce the flooding by importing some liquid water with the inlet air. The data acquired from a fault experiment was used for training procedure. Several independent fault experiments were done to further test and verify the approach.

The values of W in the fault experiments are compared with the corresponding limits W_{max} and W_{min}. The parameter W in a fault experiment is as Fig. 9. From the figure, it can be observed that the data points can be labeled with three class labels and the stack went through three successive states: membrane drying state, normal state, and flooding state. Actually, it is considered in our study that a certain amount liquid water exists in the air paths of a normal operating fuel cell stack. At the beginning of the experiment, some time was needed to construct the necessary water environment. Hence, the stack showed membrane drying state in this period. After this period, with the help of high humidified inlet air, the liquid water accumulated in the air paths. The flooding is therefore induced after a period of normal state.

Although the variable W deduced from the pressure drop model can be used for monitoring the flooding and membrane drying faults, pressure sensors and instruments for air flow measurement must be settled in both sides of air circuit. These sensors will increase the cost of the fuel cell system. Based on this consideration, an effort is made to realize the fault diagnosis by analyzing cell voltages.

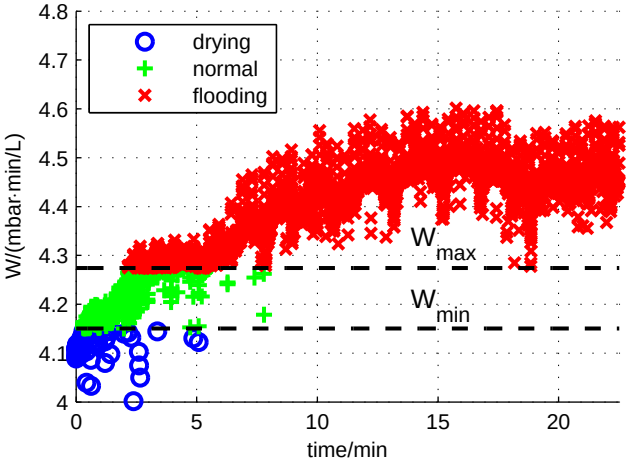

Figure 9: Parameter W of data in fault process

4.3. The results of feature extraction and classification

After the samples in fault experiments are labeled, we arrived to the model training step. As aforementioned, the analytic targets are individual cell voltages in this step. The individual cell voltages in normal experiment and fault experiment are respectively as Fig. 10 and 11.

From the figures, it can be observed that the amplitudes of cell voltages in a normal experiment are more homogeneous than in a fault experiment from an overall point of view.

In order to explore the information sufficiently, the vectors by cell voltages are constructed. The vectors in fault experiment can be expressed by a matrix $\boldsymbol{V}$:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_N \end{bmatrix}$$

where $\boldsymbol{v}_n = [v_n(1), v_n(2), \ldots, v_n(20)]$ is the cell voltages constructed vector of time sequence n. Each $\boldsymbol{v}_n$ is labeled. As aforementioned in section 3, the number of training samples N=9000.

With the matrix $\boldsymbol{V}$, the feature extraction models and classification models are trained successively. The kernel functions used in KPCA, KFDA and SVM are Gaussian kernels. Considering the distributions of data in feature spaces are not complex, the number of Gaussian components (R_i in (12)) in GMM is set to 1 for all the feature extraction methods. The parameters (σ in (10)) of the kernel functions and the parameter for SVM (D in (16)) are determined after several attempts to get a relatively higher classification accuracy. In kNN, k is set respectively for varied feature extraction methods with respect to the leave-one-out error [23].

Through feature extraction, the original 20-dimension data is projected to a 2-dimension space. Fig. 12, 13, 14 and 15 show respectively the data of a fault experiment in feature spaces generated by PCA, FDA, KPCA, and KFDA.

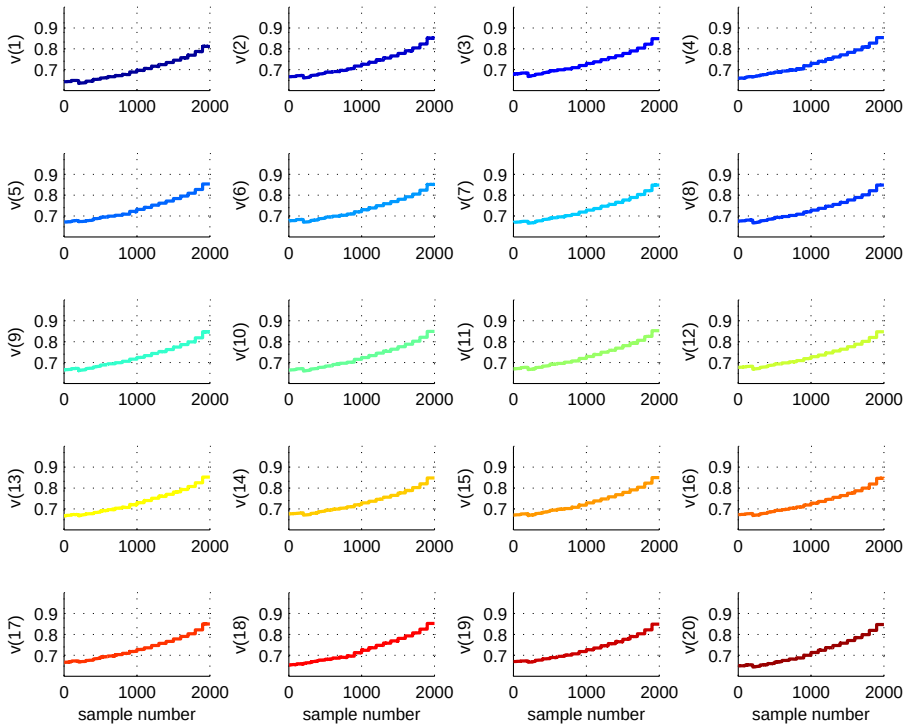

Figure 10: Cell voltages in a normal experiment. $v(n)$ denotes the voltage behavior of n^{th} cell counting from air entrance to air exit.

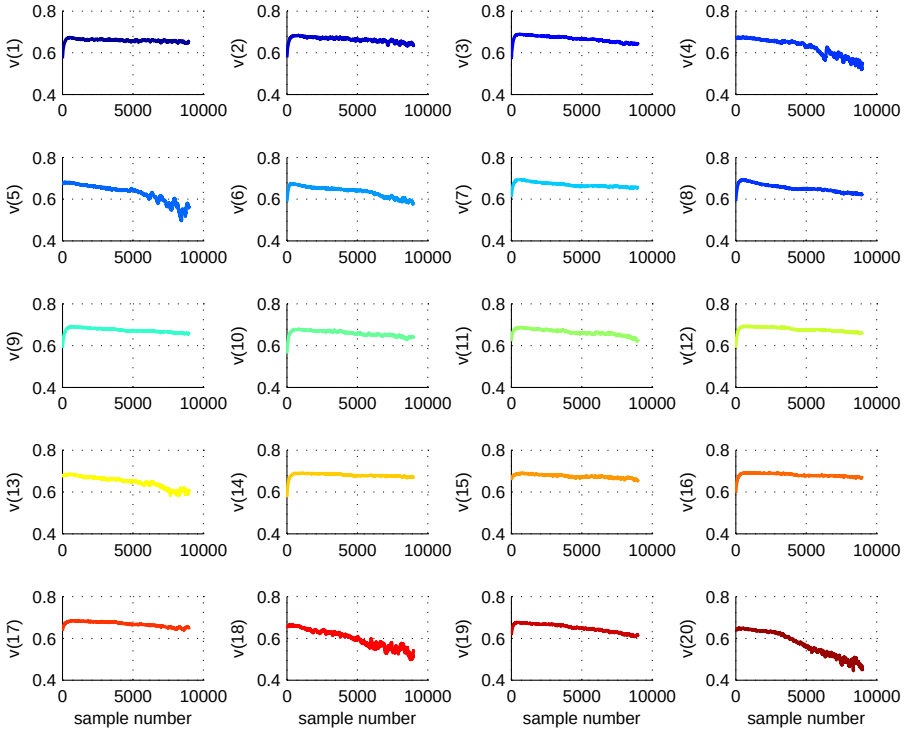

Figure 11: Cell voltages in a fault experiment

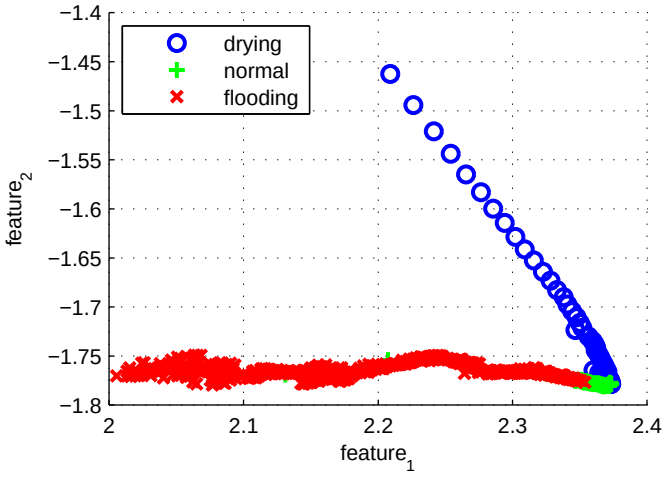

Figure 12: The features obtained by adopting PCA

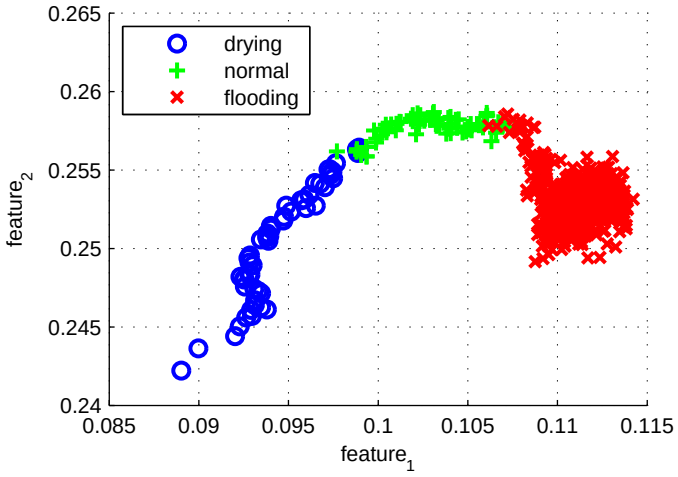

Figure 13: The features obtained by adopting FDA

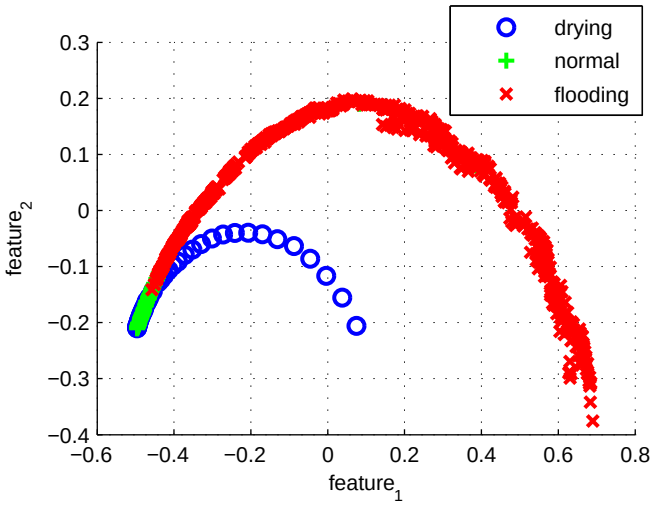

Figure 14: The features obtained by adopting KPCA

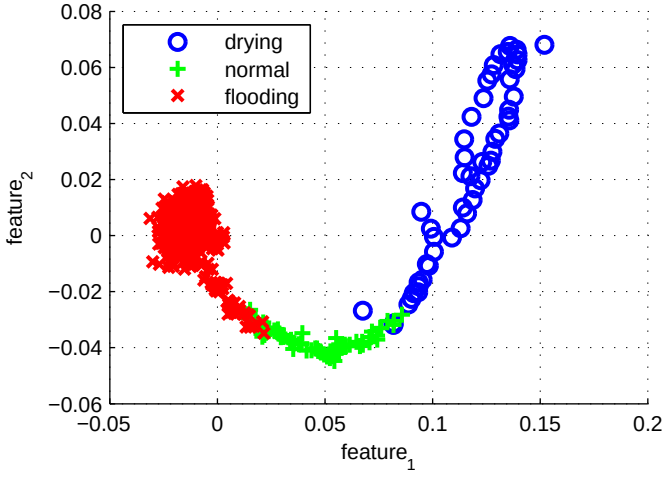

Figure 15: The features obtained by adopting KFDA

It can be seen that the data points in Fig. 12 and 14 disperse in the whole scale, whereas the overlap regions between data in the normal state and the other two fault states are large. In contrast, from Fig. 13 and 15, the points in the same class are more concentrated, the amounts of overlapping points are small, which means points in different classes are decentralized.

Classification methods GMM, kNN, and SVM are carried out in the different feature spaces. For instance, Fig. 16, 17 and 18 show respectively the visualization results of GMM, kNN, SVM classification in FDA feature space. It can be seen that the feature space is divided to three zones which denotes different states, and the boundaries determined by varied classifiers are different.

To evaluate the performance of the classifiers and feature extraction methods, error diagnosis rate (EDR) is defined. The error diagnosis points refer to the ones which are wrongly diagnosed. EDR means the proportion of error diagnosis points in total data points. Table 2 shows EDRs for different combinations of the feature extraction methods and the classification methods. In order to evaluate the robustness of approaches, the data, which is acquired from the other fault experiments than the one for training, was handled as test data. The test data was firstly labeled. Then, the trained feature extraction and classification models were used to process the test data. The EDRs of test data were thus obtained by comparing the diagnostic results with the labeling results.

From Table 2, the performances of the feature extraction methods can be compared. For each classification methodology, the error rates by using FDA and KFDA as feature extraction tools are generally lower than by using PCA and KPCA. The reason is that PCA and KPCA are unsupervised methodologies, the training data is treated equally without considering the label of each point, while FDA and KFDA are supervised methodologies, the labelling information is utilised sufficiently. Hence, we can consider that FDA and KFDA are more suitable for classification problems such as fault diagnosis.

Table 2: Results of varied classifications in different feature spaces

Feature extraction	Classification	EDR of training data	EDR of test data
	GMM	0.051	0.110
	kNN	0.016	0.110
PCA	SVM	0.015	0.129
	GMM	0.032	0.087
	kNN	0.013	0.070
FDA	SVM	0.014	0.070
	GMM	0.089	0.058
	kNN	0.052	0.121
KPCA	SVM	0.058	0.113
	GMM	0.034	0.085
	kNN	0.014	0.082
KFDA	SVM	0.016	0.075

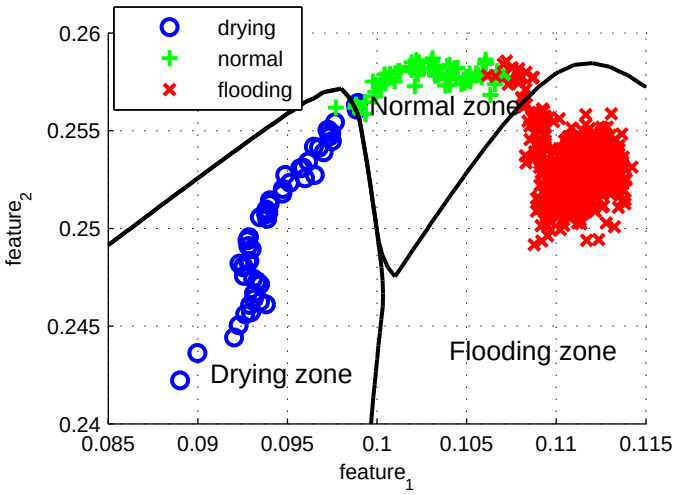

Figure 16: Classification results in FDA feature space by GMM

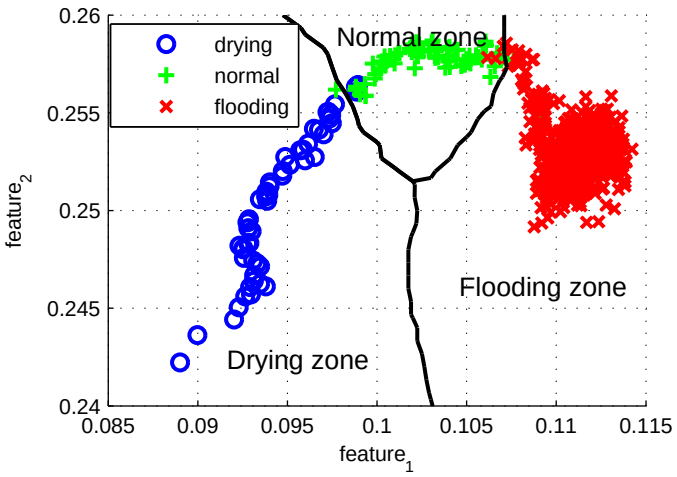

Figure 17: Classification results in FDA feature space by kNN

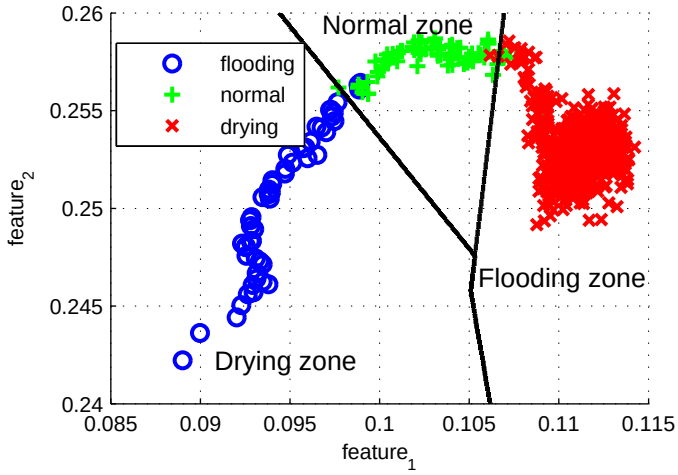

Figure 18: Classification results in FDA feature space by SVM

In practice, it is difficult to choose the highest performance classifier. Various problems may have different suitable classification solutions. Concerning our case, the choice of the proper classifier can be achieved firstly by comparing the EDRs of classifiers with FDA and KFDA as foregoing procedures. It can be observed that the error rates obtained by using kNN and SVM are always lower than GMM.

4.4. Discussion about computation costs

Apart from EDR, computation cost is a crucial factor that needs to be taken into account for real time implementation. In our approach, the training process is usually out of consideration, since it is completed off-line. Concerning the online diagnosis process, feature extraction methodologies and classification methodologies are considered respectively. The order notation $O()$ is used here to describe the computation cost. $f(x) = O(h(x))$ denotes that there exist x_0 and c_0, such that $|f(x)| \leq c_0|h(x)|$ for all $x > x_0$. The order notation is used to

give a bound on the limiting behavior of a function.

As table 3 and 4, different methodologies are evaluated from the perspectives of occupied memory and computation time (the notations are described in section 3). From this table, it can be seen that among different feature extraction methods, needed memory and computation time of KPCA and KFDA are in proportion to the number of training samples, thus are usually large. While for classifiers, the needed memory and computation time of kNN are in proportion of the number of training samples. These methodologies are less suitable than others for online diagnosis. Hence, considering synthetically the performances of diagnosis accuracy and feasibility of online implementation, FDA combined by SVM can be chosen as final solution in our case.

5. Conclusion

This paper presents an approach for diagnosis of water management faults for the PEMFC stacks. The procedure is realized by classifying the features that are extracted from the cell voltages constructed vectors.

In this approach, water indicator W is defined to describe the quantity of water inside the fuel cell stack and label data with normal and fault labels. Taking into account of the uneven character of cell voltage amplitudes, individual cell voltages are chosen as original variables for diagnosis. Different feature extraction and classification methods are employed and compared. The test results for a 20-cell stack show that FDA and SVM have higher performance and less computation costs comparing with other methods in our case. The EDR of diagnosis by using such an approach is always below 10%. It is therefore a very promising diagnosis proposal to diagnose the water management relevant faults for PEMFC.

It is possible to extend this approach to the diagnosis of other faults by increasing the number of classes and the training data related to the corresponding faults. This extension is under study.

[1] T. Escobet, D. Feroldi, S. de Lira, V. Puig, J. Quevedo, J. Riera, M. Serra, Model-based fault diagnosis in PEM fuel cell systems, Journal of Power Sources 192 (1) (2009) 216 – 223. doi:http://dx.doi.org/10.1016/j.jpowsour.2008.12.014.

[2] A. Hernandez, D. Hissel, R. Outbib, Modeling and Fault Diagnosis of a Polymer Electrolyte Fuel Cell Using Electrical Equivalent Analysis, IEEE Transaction on Energy Conversion 25 (1) (2010) 148–160. doi:10.1109/TEC.2009.2016121.

[3] D. Hissel, M. C. Péra, J. M. Kauffmann, Diagnosis of automotive fuel cell power generators, Journal of Power Sources 128 (2) (2004) 239–246. doi:http://dx.doi.org/10.1016/j.jpowsour.2003.10.001.

[4] S. Wasterlain, D. Candusso, F. Harel, X. François, D. Hissel, Diagnosis of a fuel cell stack using electrochemical impedance spectroscopy and bayesian networks, in: Vehicle Power and Propulsion Conference (VPPC), 2010 IEEE, 2010, pp. 1–6. doi:10.1109/VPPC.2010.5729184.

[5] L. A. M. Riascos, M. G. Simoes, P. E. Miyagi, A Bayesian network fault diagnostic system for proton exchange membrane fuel cells, Journal of Power Sources 165 (1) (2007) 267 – 278. doi:http://dx.doi.org/10.1016/j.jpowsour.2006.12.003.

[6] N. Yousfi-Steiner, D. Hissel, P. Moçotéguy, D. Candusso, Diagnosis of polymer electrolyte fuel cells failure modes (flooding & drying out) by neural networks modeling, International Journal of Hydrogen Energy 36 (4) (2011) 3067–3075. doi:http://dx.doi.org/10.1016/j.ijhydene.2010.10.077.

[7] W. Mérida, D. Harrington, J. L. Canut, G. McLean, Characterisation of proton exchange membrane fuel cell (PEMFC) failures via electrochemical impedance spectroscopy , Journal of Power Sources 161 (1) (2006) 264 – 274. doi:http://dx.doi.org/10.1016/j.jpowsour.2006.03.067.

[8] S. Asghari, A. Mokmeli, M. Samavati, Study of PEM fuel cell performance by electrochemical impedance spectroscopy, International Journal of Hydrogen Energy 35 (17) (2010) 9283–9290. doi:http://dx.doi.org/10.1016/j.ijhydene.2010.03.069.

[9] N. Yousfi-Steiner, D. Hissel, P. Moçotéguy, D. Candusso, Non intrusive diagnosis of polymer electrolyte fuel cells by wavelet packet transform, International Journal of Hydrogen Energy 36 (1) (2011) 740–746. doi:http://dx.doi.org/10.1016/j.ijhydene.2010.10.033.

[10] C.-Y. Lee, Y.-M. Lee, In-situ diagnosis of local fuel cell performance using novel micro sensors , International Journal of Hydrogen Energy 37 (5) (2012) 4448 – 4456. doi:http://dx.doi.org/10.1016/j.ijhydene.2011.11.098.

[11] J. Hua, L. Lu, M. Ouyang, J. Li, L. Xu, Proton exchange membrane fuel cell system diagnosis based on the signed directed graph method , Journal of Power Sources 196 (14) (2011) 5881 – 5888. doi:http://dx.doi.org/10.1016/j.jpowsour.2011.03.008.

[12] S. Abbasion, A. Rafsanjani, A. Farshidianfar, N. Irani, Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine, Mechanical Systems and Signal Processing 21 (7) (2007) 2933 – 2945. doi:http://dx.doi.org/10.1016/j.ymssp.2007.02.003.

[13] H. Li, D. Xiao, Fault diagnosis using pattern classification based on one-dimensional adaptive rank-order morphological filter, Journal of Process Control 22 (2) (2012) 436 – 449. doi:http://dx.doi.org/10.1016/j.jprocont.2011.12.005.

[14] L. Cao, K. Chua, W. Chong, H. Lee, Q. Gu, A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine, Neurocomputing 55 (1) (2003) 321 – 336. doi:http://dx.doi.org/10.1016/S0925-2312(03)00433-8.

[15] N. Yousfi-Steiner, P. Moçotéguy, D. Candusso, D. Hissel, A. Hernandez, A. Aslanides, A review on PEM voltage degradation associated with water management: Impacts, influent factors and characterization, Journal of Power Sources 183 (1) (2008) 260 – 274. doi:http://dx.doi.org/10.1016/j.jpowsour.2008.04.037.

[16] Y. Wang, S. Basu, C.-Y. Wang, Modeling two-phase flow in PEM fuel cell channels , Journal of Power Sources 179 (2) (2008) 603 – 617. doi:http://dx.doi.org/10.1016/j.jpowsour.2008.01.047.

[17] K. Ito, K. Ashikaga, H. Masuda, T. Oshima, Y. Kakimoto, K. Sasaki, Estimation of flooding in PEMFC gas diffusion layer by differential pressure measurement, Journal of Power Sources 175 (2) (2008) 732 – 738. doi:http://dx.doi.org/10.1016/j.jpowsour.2007.10.019.

[18] J. Kim, I. Lee, Y. Tak, B. Cho, State-of-health diagnosis based on hamming neural network using output voltage pattern recognition for a PEM fuel cell , International Journal of Hydrogen Energy 37 (5) (2012) 4280 – 4289. doi:http://dx.doi.org/10.1016/j.ijhydene.2011.11.092.

[19] P. Rodatz, F. Bchi, C. Onder, L. Guzzella, Operational aspects of a large PEFC stack under practical conditions , Journal of Power Sources 128 (2) (2004) 208 – 217. doi:http://dx.doi.org/10.1016/j.jpowsour.2003.09.060.

[20] A. Hernandez, D. Hissel, R. Outbib, Fuel cell fault diagnosis: A stochastic approach, in: Industrial Electronics, 2006 IEEE International Symposium on, Vol. 3, 2006, pp. 1984–1989. doi:10.1109/ISIE.2006.295877.

[21] R. Tirnovan, S. Giurgea, Efficiency improvement of a PEMFC power source by optimization of the air management , International Journal of Hydrogen Energy 37 (9) (2012) 7745 – 7756. doi:http://dx.doi.org/10.1016/j.ijhydene.2012.02.029.

[22] X. Yu, Multivariate statistics analysis, 1st Edition, Beijing: China Statistics Press, 1991.

[23] R. Duda, P. Hart, D. Stork, Pattern Classification, Wiley, 2001.

[24] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.

[25] T. Cover, Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, Electronic Computers, IEEE Transactions on EC-14 (3) (1965) 326–334. doi:10.1109/PGEC.1965.264137.

[26] C. Campbell, Kernel methods: a survey of current techniques , Neurocomputing 48 (14) (2002) 63 – 84. doi:http://dx.doi.org/10.1016/S0925-2312(01)00643-9.

[27] K.-P. Wu, S.-D. Wang, Choosing the kernel parameters for sup-

Table 3: Computation costs of varied methodologies

Methodologies	Feature extraction			
	PCA	FDA	KPCA	KFDA
Occupied memory	$O(ML)$	$O(ML)$	$O(MN + NL)$	$O(MN + NL)$
Computation time	$O(ML)$	$O(ML)$	$O(MN + NL)$	$O(MN + NL)$

Table 4: Computation costs of varied methodologies

Methodologies	Classification		
	GMM	kNN	SVM
Occupied memory	$O(CR(\frac{(1+M)(M+2)}{2}))$	$O(MN)$	$O(MS + S)$
Computation time	$O(CR(\frac{(1+M)(M+2)}{2}))$	$O(MN)$	$O(MS + S)$

port vector machines by the inter-cluster distance in the feature space , Pattern Recognition 42 (5) (2009) 710 – 717. doi:http://dx.doi.org/10.1016/j.patcog.2008.08.030.

[28] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach, Neural computation 12 (10) (2000) 2385–2404. doi:10.1162/089976600300014980.

[29] J. Yang, Z. Jin, J. yu Yang, D. Zhang, A. F. Frangi, Essence of kernel Fisher discriminant: KPCA plus LDA, Pattern Recognition 37 (10) (2004) 2097 – 2100. doi:http://dx.doi.org/10.1016/j.patcog.2003.10.015.

[30] Z.-B. Zhu, Z.-H. Song, Fault diagnosis based on imbalance modified kernel Fisher discriminant analysis , Chemical Engineering Research and Design 88 (8) (2010) 936 – 951. doi:http://dx.doi.org/10.1016/j.cherd.2010.01.005.

[31] T. Koski, J. M. Noble, Bayesian Networks: An Introduction, Wiley, 2009.

[32] G. McLachlan, D. Peel, Finite Mixture Models, Wiley, New York, 2004.

[33] J. C. Platt, Sequential Minimal Optimization : A Fast Algorithm for Training Support Vector Machines, Technical Report MSR-TR-98-14, Microsoft Research (1998) 1–21.

[34] C.-W. Hsu, C.-J. Lin, A comparison of methods for multiclass support vector machines, Neural Networks, IEEE Transactions on 13 (2) (2002) 415–425. doi:10.1109/72.991427.

[35] D. Candusso, A. De Bernardinis, M.-C. Péra, F. Harel, X. François, D. Hissel, G. Coquery, J.-M. Kauffmann, Fuel cell operation under degraded working modes and study of diode by-pass circuit dedicated to multi-stack association, Energy Conversion and Management 49 (4) (2008) 880–895. doi:http://dx.doi.org/10.1016/j.enconman.2007.10.007.

[36] H. Lilliefors, On the kolmogorov-smirnov test for normality with mean and variance unknown, Journal of the American Statistical Association 62 (318) (1967) 399–402. doi:10.1080/01621459.1967.10482916.