Contribution to on-line diagnosis, fault classification and prognosis for PEMFC

Yunjin Ao

THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

PRÉPARÉE À UNIVERSITÉ DE TECHNOLOGIE DE BELFORT-MONTBÉLIARD

École Doctorale n° 37
Sciences Pour l'Ingénieur et Microtechniques

Doctorat de Énergie

par

Yunjin AO

Contribution to on-line diagnosis, fault classification and prognosis for PEMFC

Thèse présentée et soutenue à Belfort, le 13/12/2022

Composition du Jury :

Monsieur WIRA Patrice, *Professeur, Université de Haute Alsace*, Président
Monsieur OUTBIB Rachid, *Professeur, Aix-Marseille Université*, Rapporteur
Madame CADET Catherine, *Maître de conférences HDR, Université Grenoble Alpes*, Rapporteure
Monsieur PIERFEDERICI Serge, *Professeur, Université de Lorraine*, Examinateur
Monsieur CANDUSSO Denis, *Directeur de recherche, Université Gustave Eiffel*, Examinateur
Monsieur LAGHROUCHE Salah, *Professeur, UTBM*, Directeur de thèse
Monsieur DEPERNET Daniel, *Maître de Conferences HDR, UTBM*, Codirecteur de thèse

école doctorale **sciences pour l'ingénieur et microtechniques**

Title: Contribution to on-line diagnosis, fault classification and prognosis for PEMFC
Keywords: proton exchange membrane fuel cell (PEMFC), Prognostic Health Management (PHM), diagnosis, electrochemical impedance spectroscopy(EIS), equivalent circuit model (ECM)

Abstract: The proton exchange membrane fuel cell (PEMFC) is a promising energy source that offers several advantages such as no pollution, high efficiency, and low operating temperature. However, durability and reliability remain barriers to its large-scale commercialization. The development of tools for on-line diagnosis, fault classification and prognosis of PEMFC is a very important research topic to lift these barriers. The main objective of this thesis is to contribute to the development of these tools. Thus, we have proposed three diagnostic algorithms and one prognostic algorithm to address these challenges. First, a voltage fluctuation-based diagnostic method is proposed, and the faults resulting from different operating temperatures, stoichiometry and relative humidity are studied. The voltage fluctuation model is extracted by the autoregressive model (AR model), and the coefficients of the model are directly applied as diagnostic features. Four fault classification algorithms are proposed, applied, and compared under both single-fault and multiple-fault conditions. In the second stage, two electrochemical impedance spectroscopy (EIS)-based diagnostic methods are proposed and validated in real time, which can distinguish between flooding, drying-out and mass transport faults. The first method is based on an equivalent circuit model (ECM), in which the parameters of electrical elements can be identified and applied as diagnostic features. In addition, the adaptive neuro-fuzzy inference system (ANFIS) is proposed to perform the diagnosis, and the whole diagnosis process is implemented and validated in real time on a digital signal processor (DSP) system. The third proposed diagnostic method is based on zero-phase impedance and turning phase of EIS characterization. Experiments have shown that the two proposed features can represent the health status of the PEMFC in a practical way; therefore, they can be applied as features to perform fast and efficient diagnosis. The K-nearest neighbours (KNN) algorithm is applied to classify the different fault conditions, and the whole diagnostic method is also implemented on a DSP system, which validates its real-time applicability. Finally, a frequency domain Kalman filter (FDKF) based prognostic method is proposed. This method allows the prediction of the voltage evolution with different prediction horizons as well as a fast and accurate estimation of the remaining lifetime (RUL). This method is an advantageous alternative in terms of computation time compared to temporal approaches based on the same technique.

Titre: Contribution au diagnostic en ligne, à la classification des défauts et au pronostic des piles à combustible PEMFC
Mots clés: Pile à Combustible à Membrane Echangeuse de Proton (PEMFC), pronostic et gestion de la santé (PHM), diagnostic, spectroscopie d'impédance électrochimique (EIS), modèle de circuit équivalent (ECM)

Résumé: La pile à combustible à membrane échangeuse de protons (PEMFC) est une source d'énergie prometteuse qui présente plusieurs avantages tels que l'absence de pollution, un rendement élevé et une faible température de fonctionnement. Cependant, la durabilité et la fiabilité restent des obstacles à sa commercialisation à grande échelle. Le développement d'outils de diagnostic en ligne, de classification des défauts et de pronostic de la PEMFC sont des sujets de recherche très importants afin de soulever ces barrières. L'objectif principal de cette thèse est de contribuer au développement de ces outils. Ainsi, nous avons proposé trois algorithmes de diagnostic et un algorithme de pronostic pour relever ces défis. Tout d'abord, une méthode de diagnostic basée sur la fluctuation de tension est proposée, et les défauts résultant de différentes températures de fonctionnement, stoechiométries et humidités relatives sont étudiés. Le modèle de fluctuation de tension est extrait par le modèle autorégressif (modèle AR), et les coefficients du modèle sont directement appliqués comme caractéristiques de diagnostic. Quatre algorithmes de classification de défauts sont proposés, appliqués et comparés dans des conditions de défaut unique et de défaut multiple. Dans un second temps, deux méthodes de diagnostic basées sur la spectroscopie d'impédance électrochimique (EIS) sont proposées et validées en temps réel, et elles peuvent distinguer les défauts de noyage, d'assèchement et de transport de masse. La première est basée sur un modèle de circuit électrique équivalent (ECM), dans lequel les paramètres des éléments électriques peuvent être identifiés et appliqués comme caractéristiques de diagnostic. En outre, le système d'inférence neuro-flou adaptatif (ANFIS) est proposé pour réaliser le diagnostic, et l'ensemble du processus de diagnostic est mis en œuvre et validé en temps réel sur un processeur DSP dédié à l'embarqué. La troisième méthode de diagnostic proposée est basée sur les impédances à phase nulle et à inversion du sens de la phase caractérisant l'EIS. Les expérimentations ont montré que les deux caractéristiques proposées peuvent représenter l'état de santé de la PEMFC de manière pratique; elles peuvent donc être appliquées comme caractéristiques pour réaliser un diagnostic rapide et efficace. L'algorithme des K-voisins les plus proches (KNN) est appliqué à la classification des différentes conditions de défaut, et l'ensemble de la méthode de diagnostic est également mis en œuvre sur un système DSP, ce qui valide son applicabilité en temps réel. Enfin, une méthode de pronostic basée sur le filtre de Kalman dans le domaine fréquentiel (FDKF) est proposée. Cette méthode permet de prédire l'évolution de la tension avec différents horizons de prédiction ainsi qu'une estimation rapide et précise de la durée de vie utile restante (RUL). Cette méthode constitue une alternative avantageuse en termes de temps de calcul vis-vis des approches temporelles basées sur la même technique.

UBFC
UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ

Acknowledgements

It is never an easy task to achieve a doctorate thesis, and it is only possible with enough support. Therefore, I would like to thank all the people who have supported me.

First of all, I would like to express my deepest thanks to my thesis supervisors, Dr. Salah Laghrouche, and Dr. Daniel Depernet. They are always available when I have questions during the research, and give me a lot of precious advice. They also helped me a lot in the preparation of this thesis work. I would also like to thank Dr. Denis Candusso for providing me with the experimental data of the project "Decentralized Energy Production" managed by EFFICACITY institut and for helping and advising me in their exploitation.

Also, I want to thank my supervisors when I studied at Xi'an Jiaotong University in China, Dr. Qulan Zhou, and Dr. Na Li. They opened the door of scientific research for me, and I am very grateful for their help.

At the same time, I should say thanks to my friends, both in the laboratory and in life. Without their support and help, the thesis is impossible. There are too many names and I will not cite all of them. However, true friends always know how important they are.

Definitely, I can never done a thesis if my family has not raised me and supported me to pursue higher education. It is hard for them during the Covid period as I cannot go back often to see them. I am sorry and I am thankful.

Undoubtedly, I have to thank my fiancée for her infinite comprehension and support. It is a big challenge for both of us, but finally, we made it. Thank you very much and you have no idea how much happiness and support you have brought to me.

Finally, I want to especially thank my country, China, and the China Scholarship Council (CSC) for providing me with financial support during my Ph.D. Without it, it will not be possible for a poor countryside boy to have the chance to study in France. I wish that my country will be more and more prosperous.

Thanks to everyone who has helped me in my life, it is you who makes me what I am now. Thank you very much.

Contents

Abbreviations

AC	alternating current
AEC	acoustic echo cancellation
AFC	alkaline fuel cells
AI	artificial intelligent methods
AIC	Akaike information criterion
ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural network
AR	autoregressive
BIC	Bayesian information criterion
BPs	bipolar plates
BPNN	back propagation neural networks
CI	current interruption
CLs	catalyst layers
CNN	convolutional neural network
CPE	constant phase element
CV	cyclic voltammetry
DFA	discriminant function analysis
DFT	discrete Fourier transform
DoE	department of energy
DSP	digital signal processor
DWT	discrete wavelet transformation
ECM	equivalent circuit model
ECSA	electrochemical active surface area
EIS	electrochemical impedance spectroscopy
EKF	extended Kalman filter

ELM	extreme learning machine
EMD	empirical mode decomposition
EOL	end-of-life
EN	electrochemical noise
ESN	echo state network
FCEV	fuel cell electric vehicle
FCHJU	fuel cells and hydrogen joint undertaking
FCLAB	federation for fuel cell research
FDKF	frequency domain Kalman filter
FFT	fast Fourier transform
FT	Fourier transform
GA	genetic algorithm
GDL	gas diffusion layer
HFR	high-frequency resistance
IDFT	inverse discrete Fourier transform
KDFA	kernel discriminant function analysis
KNN	K-nearest neighbours
KPCA	kernel principal component analysis
LM	Levenberg–Marquardt algorithm
LSTM	long short-term memory
LSV	linear sweep voltammetry
MCFC	molten carbonate fuel cells
MEA	membrane electrode assembly
MID	mutual information difference
MIQ	mutual information quotient
MRA	multiple regression analysis
MRE	mean relative error
MSE	mean squared error
NANN	non-linear autoregressive neural network
NN	neural networks
PAFC	phosphoric acid fuel cells
PCA	principal components analysis
PEMFC	proton exchange membrane fuel cell

PF	particle filter
PFSA	perfluorosulfonic acid
PH	prediction horizon
PHM	prognostics and health management
PRBS	pseudo-random binary sequence
PSD	power spectrum density
PSO	particle swarm optimization
RF	random forest
RH	relative humidity
RMSE	root mean squared error
RUL	remaining useful life
RVM	relevance vector machine
SA	stoichiometry anode
SAA	simulated annealing algorithm
SC	stoichiometry cathode
SCI	serial communication interface
SOFC	solid oxide fuel cells
SSM-SVM	spherical-shaped multiple-class support vector
STFT	short-time Fourier transform
SVD	singular value decomposition
SVM	support vector machine
THD	total harmonic distortion
WT	wavelet transform
XGBoost	extreme gradient boosting

Introduction Générale

Les combustibles fossiles n'étant pas durables, et défavorables à l'environnement, l'utilisation des énergies renouvelables est considérée comme un élément important pour les futurs systèmes énergétiques [1]. L'hydrogène étant facilement accessible, stockable et propre, il s'agit d'un type d'énergie présentant un grand potentiel parmi les diverses énergies renouvelables. En outre, comme l'hydrogène n'émet pas de carbone lors de son utilisation, il est particulièrement important dans le contexte de la décarbonisation dans le monde entier. Par conséquent, la recherche scientifique sur l'hydrogène est de plus en plus importante.

La Pile à combustible à membrane échangeuse de protons (PEMFC) est l'une des plus importantes technologies permettant d'utiliser l'hydrogène, c'est-à-dire de convertir l'énergie de l'hydrogène en électricité. Elle a fait l'objet d'une grande attention et de nombreuses applications ont été développées sur la base de la PEMFC, comme les centrales électriques distribuées, les véhicules PEMFC et d'autres applications d'approvisionnement en énergie [2]. Cependant, comme là les performances de ces systèmes sont instables et se dégradent trop rapidement pendant le fonctionnement, la durée de vie de la PEMFC est toujours inférieure aux attentes. Par conséquent, la durabilité et la fiabilité sont toujours des obstacles critiques sur la voie de sa commercialisation à grande échelle [3].

Pour améliorer la durabilité et la fiabilité de la PEMFC, la gestion de sa santé est un sujet très important [4]. Il comprend toutes les actions qui peuvent maintenir un bon état de santé du système. Ainsi, il est important d'élaborer des méthodes pour évaluer l'état de santé des dispositifs PEMFC selon les informations collectées par toutes sortes de capteurs, afin de trouver les facteurs négatifs qui peuvent causer un défaut de fonctionnement ou une dégradation à long terme. Le diagnostic et le

pronostic sont des méthodes importantes pour la gestion de la santé. Le diagnostic permet de détecter les défauts pendant le fonctionnement de la PEMFC et de faire des suggestions sur l'élimination de ces défauts. En même temps, le pronostic permet d'avoir une prédiction sur le développement de l'état de santé à long terme. En outre, la dégradation à long terme peut être identifiée et les mesures correspondantes peuvent être prises pour réduire la dégradation et améliorer la durée de vie.

Plusieurs méthodes de diagnostic et de pronostic ont été proposées et appliquées dans la littérature pour les systèmes PEMFC. Cependant, la durée de vie ciblée n'est pas encore atteinte, en particulier pour les applications réelles (hors laboratoires), telles que les vibrations, les températures extrêmement élevées ou basses, la pollution par la poussière, etc. Par conséquent, il n'est pas facile de maintenir un fonctionnement normal et d'améliorer la durée de vie. Des stratégies de diagnostic et de pronostic plus pratiques sont nécessaires. Comme les défauts de la PEMFC peuvent se produire et évoluer très rapidement sans prévenir dans des conditions de fonctionnement complexes, la capacité de diagnostic en temps réel est une exigence critique pour la préservation du système. Ainsi, le défaut peut être détecté assez rapidement, afin de réagir immédiatement pour réduire les dommages. De plus, le coût du diagnostic et du pronostic doit être réduit, afin de pouvoir l'appliquer largement dans les applications commerciales.

Pour faire face à ces problèmes, le diagnostic en temps réel et le pronostic en ligne du système PEMFC sont étudiés dans cette thèse. L'objectif est de trouver des méthodes de diagnostic et de pronostic plus pratiques et rapides afin de contribuer au développement et à l'extension des applications PEMFC.

Plus précisément, trois algorithmes de diagnostic différents et une méthode de pronostic sont proposés. La première méthode de diagnostic est basée sur la fluctuation de tension et le modèle autorégressif. Ainsi, les défauts liés aux différentes stœchiométries, à la température du circuit de refroidissement et aux niveaux d'humidité relative peuvent être classés avec précision avec cette méthode. Les deuxième et troisième méthodes de diagnostic sont toutes basées sur la spectroscopie d'impédance électrochimique (SIE), et elles sont toutes validées en ligne sur un système de processeur de signaux numériques (DSP). L'une d'entre elles est réalisée à l'aide d'un modèle de circuit électrique équivalent (ECM) et d'un système d'inférence neuro-flou

adaptatif (ANFIS), tandis que l'autre est proposée en fonction des caractéristiques de détection rapide de l'EIS. Les deux méthodes peuvent reconnaître les défauts liés à l'inondation, au séchage et au transport de masse dans la PEMFC. Une méthode de pronostic est également proposée sur la base du filtre de Kalman dans le domaine fréquentiel (FDKF) qui, comparé à son équivalent dans le domaine temporel, offre la simplicité de mise en œuvre et une bien meilleure performance de calcul.

Toutes les méthodes proposées ont été validées en utilisant les résultats expérimentaux d'un banc d'essai dédié au diagnostic des PEMFC. Ce banc offre différentes conditions de fonctionnement et de test avec un accès aux mesures des principales variables du système. La pile à combustible considérée est une PEMFC à 12 cellules fonctionnant dans différentes conditions de stœchiométrie, de température du circuit de refroidissement et de taux d'humidité relative.

Le manuscrit de thèse est structuré comme suit :

- Dans le chapitre 1, nous fournissons un tour d'horizon sur les principes de la PEMFC et passons en revue les principales technologies de diagnostic. Tout d'abord, les principes de la PEMFC, les défauts courants de la PEMFC et le concept de diagnostic et de pronostic de la PEMFC sont présentés en détail. Ensuite, un état de l'art détaillé des méthodologies de diagnostic est donné, et les méthodes de diagnostic sont classifiées en deux groupes, à savoir les méthodes basées sur les modèles et les méthodes basées sur les données. En outre, en fonction des travaux de recherche précédents, les défis et les solutions du diagnostic des piles PEMFC en temps réel sont analysés, et les principaux objectifs et méthodes de la thèse sont présentés.

- Le chapitre 2 porte sur le diagnostic basé sur les données de fluctuation de tension à haute fréquence et un modèle autorégressif (modèle AR). Une nouvelle approche de diagnostic pour classer les différentes conditions de fonctionnement défectueuses du système PEMFC est proposée sur la base des modèles de fluctuation de tension de la pile, qui sont identifiés par le modèle autorégressif (modèle AR). Plutôt que d'utiliser les tensions de chaque cellule dans une pile pour le diagnostic, la méthode proposée se concentre sur la fluctuation de la tension de la pile dans le temps. Étant donné que seule la tension de la pile, et

non de chaque cellule, doit être collectée, cette méthode est moins complexe et plus pratique. Pour la première fois, le modèle AR est appliqué pour extraire les caractéristiques de fluctuation de la tension de la pile PEMFC, puis les conditions de défaut sont classées par plusieurs classificateurs. L'efficacité de la méthode de diagnostic proposée est démontrée expérimentalement dans 9 conditions de défaut unique et 8 conditions de défaut multiple. Ces défauts sont liés à la stœchiométrie de la cathode, la stœchiométrie de l'anode, la température du circuit de refroidissement et le niveau d'humidité relative. La précision du diagnostic est de 99% pour les conditions de défaut unique et de 93% pour les conditions de défaut multiple, ce qui permet de gagner du temps par rapport au diagnostic basé sur la méthode d'analyse des singularités. En outre, les effets de la fréquence d'échantillonnage et de la longueur de la fenêtre d'échantillonnage sur l'efficacité du diagnostic sont étudiés et discutés pour la première fois.

- Le chapitre 3 se concentre sur deux méthodes de diagnostic basées sur le EIS. La première est basée sur le modèle de circuit électrique équivalent (ECM) et le système d'inférence flou neuronal adaptatif (ANFIS), et une autre méthode de diagnostic est basée sur deux caractéristiques EIS facilement disponibles. Ces deux méthodes sont validées en temps réel sur un système embarqué de type industriel (DSP).

  Tout d'abord, une nouvelle méthode de diagnostic basée sur l'ECM et ANFIS est proposée. Une nouvelle méthode d'identification des paramètres combinant l'algorithme génétique (GA) et l'algorithme Levenberg-Marquardt (LM) est proposée pour identifier l'ECM d'ordre fractionnel, dans lequel l'impédance anodique, l'impédance cathodique et le transfert de masse sont tous considérés. Cette nouvelle méthode permet une meilleure exploitation des diagrammes EIS, et les relations internes entre les conditions de défaut et les paramètres ECM sont analysées en détail. Ensuite, sur la base des relations obtenues, un nouvel algorithme de diagnostic basé sur le clustering k-means et ANFIS est conçu pour identifier précisément plusieurs défauts qui peuvent se produire dans la PEMFC, tels que l'inondation de la membrane, le séchage et le défaut de transfert de masse. Enfin, l'efficacité de cette méthode est démontrée expérimentalement par l'exploitation des données EIS sous différents défauts et conditions de fonctionnement de la PEMFC. De plus, l'implémentation du

diagnostic en temps réel sur le système PEMFC est réalisée sur un processeur de signaux numériques (DSP) de type industriel basé sur l'algorithme de diagnostic proposé. Le cadre permettant de relier la mesure en ligne du EIS, le diagnostic et le contrôle de la PEMFC sont proposés, et la mise en œuvre du diagnostic en temps réel est mise en évidence. Dans le DSP, les paramètres ECM sont identifiés par une méthode d'optimisation efficace en termes de calcul et adaptée aux processeurs industriels, à savoir l'algorithme de Powell. Dans la deuxième phase, le système ANFIS à réglage adaptatif est appliqué au diagnostic des défauts. La démonstration expérimentale en temps réel prouve que la méthode de diagnostic est précise et réactive.

Ensuite, une méthode de diagnostic rapide basée sur le SIE est proposée et validée. Comme le diagnostic EIS traditionnel avec des mesures sur une large gamme de fréquences prend beaucoup de temps, une nouvelle méthode de diagnostic est proposée sur la base de deux caractéristiques SIE faciles à obtenir, à savoir l'impédance de phase zéro et la phase de rotation. L'impédance de phase zéro est la moyenne des points SIE avec un petit angle de phase dans la zone de haute fréquence, et elle peut représenter l'état de l'eau dans la membrane. La phase de retournement est la phase d'impédance qui relie deux demi-cercles dans la courbe EIS, et elle peut représenter la condition de flux de gaz dans la couche de diffusion de gaz (GDL). Il est démontré expérimentalement que les caractéristiques proposées peuvent représenter l'état de santé de la PEMFC. L'espace des caractéristiques peut être divisé en 7 zones de défaut par la classification KNN (K-nearest neighbours) et tous les cas sont assignés aux bonnes classes. Ces deux caractéristiques peuvent être extraites uniquement du SIE par quelques mesures, et l'algorithme de diagnostic est validé en temps réel sur un système DSP. Le temps total de diagnostic n'est que de 2,506 secondes, ce qui prouve que la méthode de diagnostic est rapide et réactive.

- Le chapitre 4 est consacré au développement d'une méthode de pronostic basée sur le filtre de Kalman dans le domaine fréquentiel (FDKF). Comme la dégradation affecte sérieusement la durabilité et le coût de la PEMFC, une nouvelle méthode basée sur le FDKF et un modèle de dégradation de la tension est proposée pour prédire la dégradation de la PEMFC dans le domaine fréquentiel. L'avantage de la méthode FDKF proposée est qu'elle peut

traiter les données par groupes, ce qui permet de réduire considérablement le temps de calcul avec une grande précision. Deux expériences de dégradation sous des courants constants et dynamiques ont été utilisées pour démontrer les performances de pronostic avec différentes conditions et différents temps d'apprentissage. Par rapport à la méthode traditionnelle du filtre de Kalman étendu (dans le domaine temporel), il a été démontré que la méthode proposée est plus précise et nécessite un temps de calcul beaucoup plus court.

- Enfin, une conclusion générale est donnée en fonction des études de cette thèse, et les améliorations possibles et les perspectives de travaux futurs sont proposées.

Les résultats de ce travail de thèse ont été publiés dans plusieurs revues internationales à comité de lecture et dans une conférence nationale.

- Y. Ao, S. Laghrouche, D. Depernet, and K. Chen, "Lifetime prediction for proton exchange membrane fuel cell under real driving cycles based on platinum particle dissolve model," International Journal of Hydrogen Energy, 2020.
- Y. Ao, K. Chen, S. Laghrouche, and D. Depernet, "Proton exchange membrane fuel cell degradation model based on catalyst transformation theory," Fuel Cells, vol. 21, no. 3, pp. 254-268, 2021.
- Y. Ao, S. Laghrouche, D. Depernet, and K. Chen, "Proton Exchange Membrane Fuel Cell Prognosis Based on Frequency-Domain Kalman Filter," IEEE Transactions on Transportation Electrification, vol. 7, no. 4, pp. 2332-2343, 2021.
- Y. Ao, S. Laghrouche, and D. Depernet, "Diagnosis of proton exchange membrane fuel cell system based on adaptive neural fuzzy inference system and electrochemical impedance spectroscopy," Energy Conversion and Management, vol. 256, pp. 115391, 2022.
- Y. Ao, S. Laghrouche, D. Depernet, and D. Candusso, "Diagnosis of proton exchange membrane fuel cell system based on autoregressive model," fédération nationale hydrogène (FRH2) du CNRS, Aussois France, 2022/06/01.

General Introduction

As fossil fuels are unsustainable and unfavourable for the environment, the utilization of renewable energy is considered an important part of the future energy system [1]. As hydrogen is readily accessible, storable, and clean, it is a kind of energy with great potential among various renewable energy. Also, as there is no carbon get involved during the utilization of hydrogen, it is especially important in the context of decarbonization all over the world. Therefore, the scientific research around hydrogen is more and more important.

The proton exchange membrane fuel cell (PEMFC) is one of the most important methods that can make use of hydrogen, i.e. convert hydrogen energy into electricity. It has received a lot of attention, and a lot of applications have been developed based on PEMFC, such as distributed power stations, PEMFC vehicles, and other energy supply applications [2]. However, as the performance is unstable and degrades too quickly during the operation, the PEMFC lifetime is still under expectation. Therefore, durability and reliability are still critical barriers on the road to large-scale commercialization of PEMFC [3].

To improve PEMFC durability and reliability, health management is a very important solution [4]. Health management includes all the actions that can maintain the health state of the system, and one major task is to check the health state of PEMFC according to the information collected by all kinds of sensors, so that to find out the negative factor that may cause operating fault or long-term degradation. Diagnosis and prognosis are important methods for health management. Diagnosis can find out the faults during the PEMFC operation and give suggestions about fault removal, as it can identify the different fault types and isolate the fault location. At the same time, prognosis means to give a prediction about the health state development in the long term; thus the whole

system can be arranged in advance. Also, the long-term degradation can be identified and the corresponding measures can be taken to reduce the degradation and improve the lifespan.

Several diagnosis and prognosis methods have been proposed and applied to maintain the PEMFC health, and the lifetime of PEMFC has increased during the last years. However, the targeted lifetime is not yet achieved, especially for real applications under complex conditions, such as vibrations, extremely high or low temperatures, dust pollution, etc. Therefore, it is not an easy task to maintain normal operations and improve the lifetime. More practical diagnostic and prognostic strategies are needed. Because PEMFC faults can occur and evolve very quickly without warning under complex operating conditions, real-time diagnostic capability is a critical requirement for system preservation. Therefore, the fault can be detected quickly enough, so that to react immediately to reduce the damage. Also, the cost of the diagnosis and prognosis should be reduced, so that to widely apply it in commercial applications.

To face those problems, the diagnosis and prognosis of the PEMFC system are researched in this thesis. The objective is to find more practical and quick diagnosis and prognosis methods to contribute to the development and extension of PEMFC applications.

More precisely, three different diagnostic algorithms and one prognostic method are proposed. The first diagnostic method is based on the voltage fluctuation and autoregressive model, and the faults related to different stoichiometry, cooling circuit temperature, and relative humidity levels can be classified accurately. The second and third diagnostic methods are all based on electrochemical impedance spectroscopy (EIS), and they are all on-line validated on a industrial level digital signal processor (DSP) system. One of them is achieved by an equivalent circuit model (ECM) and adaptive neuro-fuzzy inference system (ANFIS), while the other is proposed according to the quick detective EIS features. Both methods can recognize the fault related to flooding, drying-out, and mass transport in the PEMFC. A prognostic method is also proposed based on the frequency domain Kalman filter (FDKF) which compared to its equivalent in the time domain offers the simplicity of implementation and a much better computational performance.

All the proposed methods have been validated using experimental results from a test bench dedicated to PEMFC diagnostics. This bench offers different operating and test conditions with access to measurements of the main variables of the system. The fuel cell under consideration is a 12-cell PEMFC operated under different operating conditions of stoichiometry, cooling circuit temperatures and relative humidity levels.

The structure of the thesis is as follows:

- In Chapter 1, we provide the background knowledge of the PEMFC principles and reviewed the main diagnostic technologies. Firstly, the PEMFC principles, common PEMFC faults, and the concept of PEMFC diagnosis and prognosis are presented in detail. Then the literature research about diagnosis methodologies is given, and the diagnostic methods are divided into two groups, i.e. model-based and data-based methods. Further, according to the former research works, the challenges and solutions of real-time PEMFC diagnosis are analysed, and the main objectives and methods of the thesis are introduced.

- Chapter 2 is about the diagnosis based on high-frequency voltage fluctuation data and an autoregressive model (AR model). A novel diagnosis approach to classifying different fault operating conditions of the PEMFC system is proposed based on the stack voltage fluctuation patterns, which are identified by the autoregressive model (AR model). Rather than using the voltages of each cell in a stack for diagnosis, the proposed method focuses on the stack voltage fluctuation over time. As only the voltage of the stack rather than each cell needs to be collected, this method is less complex and more practical. For the first time, the AR model is applied to extract fluctuation features from the PEMFC stack voltage, and then the fault conditions are classified by several widely applied classifiers. The effectiveness of the proposed diagnosis method is demonstrated experimentally under 9 single-fault conditions and 8 multi-fault conditions. Those faults are related to the cathode stoichiometry, anode stoichiometry, cooling circuit temperature, and the relative humidity level. The diagnostic accuracy is 99% for single-fault conditions and 93% for multi-fault conditions, and it is more time-saving compared with the diagnosis based on the singularity analysis method. Moreover, the effects of sampling frequency and sample window length on the diagnosis effectiveness are first time studied and discussed.

- Chapter 3 focuses on two diagnosis methods based on EIS. The first of them is based on the equivalent circuit model (ECM) and adaptive neural fuzzy inference system (ANFIS), and another diagnostic method is based on two easily available EIS features. Both of them are validated on an industrial-level embedded system in real time.

  Firstly, a new diagnostic method based on ECM and ANFIS is proposed. A new parameter identification method that combines genetic algorithm (GA) and Levenberg–Marquardt (LM) algorithm is proposed to identify the fractional-order ECM, in which the anode impedance, cathode impedance, and mass transfer are all considered. This new method allows better exploitation of the EIS diagrams, and the internal relationships between the fault conditions and the ECM parameters are thoroughly analysed according to it. Then, based on these relationships, a new diagnostic algorithm based on k-means clustering and ANFIS is designed to precisely identify several faults that can occur in the PEMFC, such as membrane flooding, drying, and mass transfer fault. Finally, the effectiveness of this method is demonstrated experimentally through the exploitation of EIS data under different faults and operating conditions of the PEMFC. Also, the implementation of real-time PEMFC diagnosis is achieved on an industrial-level digital signal processor (DSP) based on the proposed diagnostic algorithm. The framework for linking on-line EIS measurement, diagnosis, and control of PEMFC is proposed, and the implementation of real-time diagnosis is highlighted. In the DSP, the ECM parameters are identified by a computationally efficient optimization method that is suitable for industrial processors, namely, Powell's algorithm. In the second phase, the adaptively tunable ANFIS is applied to fault diagnosis. It is experimentally demonstrated in real time, proving that the diagnosis method is accurate and practical.

  Secondly, a quick diagnosis method based on EIS is proposed and validated. As the traditional EIS diagnosis with wide frequency range measurement is time-consuming, a new diagnostic method is proposed based on two easily obtain EIS features, i.e. the zero-phase impedance and the turning phase. The zero-phase impedance is the average of EIS points with a small phase angle in the high-frequency zone, and it can represent the water condition in the membrane. The turning phase is the impedance phase that connects two semicircles in

the EIS curve, and it can represent the gas flow condition in the gas diffusion layer (GDL). It is experimentally demonstrated that the proposed features can represent the health state of the PEMFC. The feature space can be divided into 7 fault zones by K-nearest neighbours (KNN) classification and all the cases are assigned to the right classes. Both two features can be extracted from EIS only by several measurements, and the diagnostic algorithm is validated in real time on an industrial-level DSP system. The total diagnostic time is only 2.506 seconds, proving that the diagnosis method is quick and practical.

- Chapter 4 is devoted to developing a prognostic method based on the frequency domain Kalman filter (FDKF). As the degradation seriously affects the durability and cost of the PEMFC, a novel model-driven method based on the FDKF and voltage degradation model is proposed to predict the degradation of PEMFC in the frequency domain. The advantage of the proposed FDKF method is that it can process the data in groups; thus the computation time can be greatly reduced with high accuracy. Two degradation experiments under constant and dynamic currents have been used to demonstrate its prognosis performances under different conditions and different training times. Compared with the traditional time domain extended Kalman filter method and literature, it has been demonstrated that the proposed one has higher accuracy and requires much less calculation time.

- Finally, a general conclusion is given according to the studies of this thesis, and the possible improvement and perspectives for future work are given.

The results of this thesis work have been published in several international peer-reviewed journals and in a national conference.

- Y. Ao, S. Laghrouche, D. Depernet, and K. Chen, “Lifetime prediction for proton exchange membrane fuel cell under real driving cycles based on platinum particle dissolve model,” International Journal of Hydrogen Energy, 2020.

- Y. Ao, K. Chen, S. Laghrouche, and D. Depernet, “Proton exchange membrane fuel cell degradation model based on catalyst transformation theory,” Fuel Cells, vol. 21, no. 3, pp. 254-268, 2021.

- Y. Ao, S. Laghrouche, D. Depernet, and K. Chen, “Proton Exchange Membrane Fuel Cell Prognosis Based on Frequency-Domain Kalman Filter,” IEEE Transactions on Transportation Electrification, vol. 7, no. 4, pp. 2332-2343, 2021.
- Y. Ao, S. Laghrouche, and D. Depernet, “Diagnosis of proton exchange membrane fuel cell system based on adaptive neural fuzzy inference system and electrochemical impedance spectroscopy,” Energy Conversion and Management, vol. 256, pp. 115391, 2022.
- Y. Ao, S. Laghrouche, D. Depernet, and D. Candusso, “Diagnosis of proton exchange membrane fuel cell system based on autoregressive model,” fédération nationale hydrogène (FRH2) du CNRS, Aussois France, 2022/06/01.

Chapter 1

Review of PEMFC diagnosis and prognosis methods

1.1 Introduction

During the last 20 years, the world energy system has greatly changed and a lot of traditional fossil fuels have been replaced by renewable energy. Hydrogen is readily accessible, storable, carbon-free, and clean, so it is getting more and more attention [1]. To make use of hydrogen, fuel cell technology is one of the most important solutions. The fuel cell is a kind of technology that can convert chemical energy (often from hydrogen and oxygen) into electricity by electrochemical reactions. When hydrogen is applied as the fuel, the only by-products are only water and heat, so it is very clean [5].

There are various types of fuel cells. The most important difference between those technologies is the type of electrolyte used to separate the anode and cathode. Namely, the mainstream fuel cells can be grouped as alkaline fuel cells (AFC), polymer electrolyte membrane or proton exchange membrane fuel cells (PEMFC), phosphoric acid fuel cells (PAFC), molten carbonate fuel cells (MCFC), and solid oxide fuel cells (SOFC) [6] [5]. The comparison of different kinds of fuel cells are given in table 1.1.

As can be seen from the table, both PAFC, MCFC, and SOFC should operate under relatively high temperatures, so they need a large subsystem to maintain the

TABLE 1.1: Comparison of different kinds of fuel cells

Types	Moving ions	Operating temperature	efficiency	Main characters	Applications
AFC	OH^-	65-220 °C	45-60%	intolerant to CO_2; high efficiency; low cost for catalyst; high cost for gas purification	has been used in the space program
PEMFC	H^+	60-80 °C	45-50%	high power density; can operate under low temperature; high catalyst cost	has been applied to mobile vehicles or stationary power station
PAFC	H^+	205 °C	40%	low power density; high gas pollution resistance;can be supplied by various fuels	more suitable for stationary and CHP applications
MCFC	CO_3^{2-}	650 °C	45-55%	low catalyst cost; can supply various fuels; carbonate ions should be supplemented; pollution of carbon dioxide	suitable for large-scale stationary power plant and CHP application
SOFC	O^{2-}	600-1000 °C	45-60%	low catalyst cost; various fuel supply; high efficiency; long start-up time; low degradation; High temperature compatibility issues of the components	stationary and CHP applications

temperature. As a result, they all have big sizes and it will be uneconomic to have a small application based on them. At the same time, as the AFC is intolerant to CO_2, it is not easy to be applied in harsh conditions. All in all, PEMFC is now the most popular fuel cell technology among them, and it has already been applied to both mobile and stationary applications.

1.2 Basis of PEMFC diagnosis and prognosis

1.2.1 PEMFC structure and mechanism

PEMFC structure is like a sandwich. The components and basic mechanism of PEMFC can be shown in figure 1.1 [7].

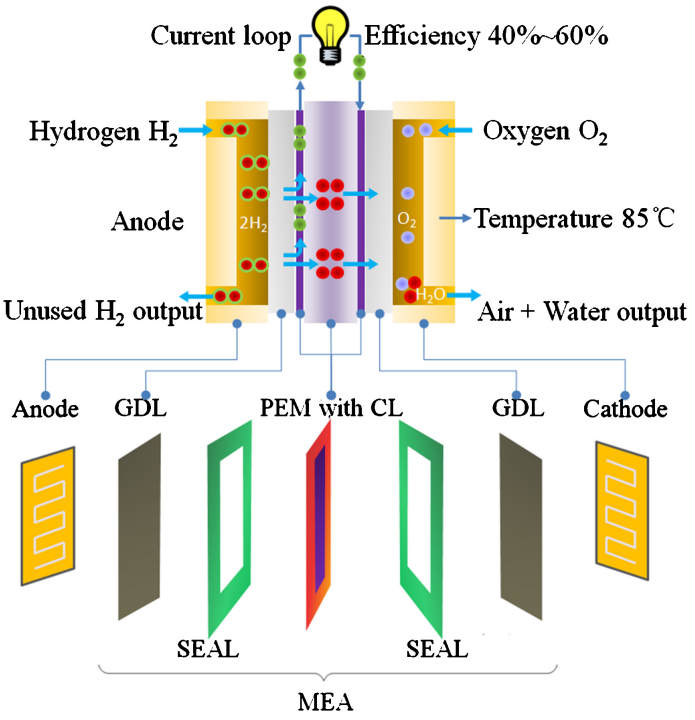


FIGURE 1.1: Basic structure and mechanism of PEMFC [7]

First, reactants enter the fuel cells through the channels of the bipolar plates (BP). The bipolar plates have three functions. First, they contain flow channels, which help organize the gas path. One of the most important objectives of gas flow channel design is to make the gas concentration more even over the whole reaction surface. Thus, the current distribution over the surface will be even, and the efficiency and performance of PEMFC can be improved. Also, the BPs are also the electricity collectors and heat conductors, therefore they should be good conductor for both electricity and heat. Finally, the BPs should be solid enough, so that to support and protect all the stack components as well as bear the pressure for the sealing [5].

Secondly, the hydrogen and air spread through the gas diffusion layers (GDLs) to reach the reaction positions. Usually this layer is made of porous material, so that to distribute the gas to the reaction site and take out the products (water). They must also conduct electricity and heat produced.

Then there are reactions in the catalyst layers (CLs). The basic reactions in PEMFC can be shown as:

$$
\begin{aligned}
&H_2 = 2H^+ + 2e^-,\ Anode \\
&\frac{1}{2}O_2 + 2H^+ + 2e^- = H_2O,\ Cathode \\
&H_2 + \frac{1}{2}O_2 = H_2O - \Delta H + electricity,\ Overall\ reaction
\end{aligned}
\tag{1.1}
$$

The reactions should be catalysed by the catalyst layer. The hydrogen loses electrons at the anode, and the electrons and protons reach the cathode by the outer circuit and membrane respective. Therefore, the load on the outer circuit can be driven by the current. CLs are typically constructed by adhering catalyst particles to the porous support. The most widely applied catalyst in PEMFC is made from expensive metal, such as Pt, while the support is usually carbon-based porous material. Therefore, one of the most important tasks for catalyst research is to improve the catalyst efficiency, but also to reduce the utilization of expensive metals, so that to reduce the cost [5].

The electrolyte in PEMFC is the membrane between the two catalyst layers. It is specially designed to only let protons pass the membrane from the anode to the cathode. Therefore, the electron can move in one direction, so that to provide electricity. It is usually made of perfluorosulfonic acid (PFSA). The membrane, catalyst layers, and GDLs are usually attached together by seal products, and they are called membrane electrode assembly (MEA). MEA is the core of a fuel cell, and it is usually very thin [5].

For each cell, the provided voltage can be calculated by the reaction energy. The plot of voltage under different static currents is called polarization curve [8]. A representative polarization curve of fuel cells can be shown in figure 1.2 [5].

Generally, the output voltage can be represented by the theoretical potential and different voltage losses, i.e. activation losses, ohmic losses, and concentration losses [9].

The theoretical potential can be obtained according to the reaction energy of reaction 1.1, which can be calculated by equation 1.2.

$$
E_r = \frac{-\Delta G}{nF} \tag{1.2}
$$

Where

$$
\Delta G = \Delta H - T\Delta S \tag{1.3}
$$

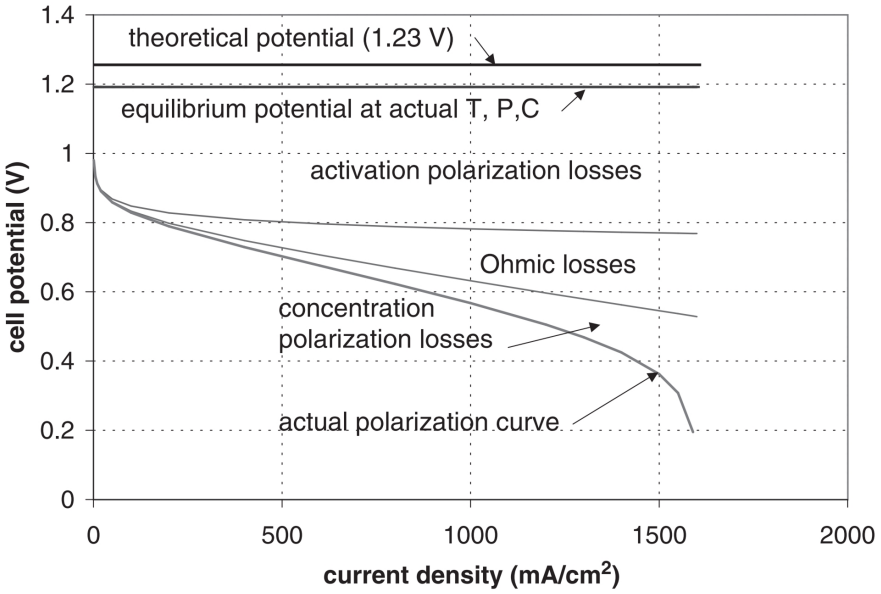


FIGURE 1.2: A typical polarization curve and voltage losses in PEMFC [5]

Where the ΔH is the change of enthalpy of the reaction; E_r is the equilibrium potential; ΔG is the Gibbs free energy; n is the number of electron transport of the reaction; F is the Faraday constant, which is 96485; T is the reaction temperature in Kelvin; ΔS is the change of entropy of the reaction.

According to the Gibbs free energy under normal conditions (25°C, atmospheric pressure, the produced water is liquid), the theoretical potential is 1.23 V. However, the equilibrium potential is slightly lower under actual temperature and pressure [9].

The voltage losses are divided into three parts. The activation losses represent the loss caused by the electrode kinetics. There are activation losses at both anode and cathode, but the losses at the anode are usually much smaller than that of the cathode because the reaction in the anode is much easier. Therefore, the activation losses at the anode are usually omitted and the activation losses of a fuel cell can be calculated by equation 1.4 [5].

$$\Delta V_{act} = \frac{RT}{\alpha F} \ln(\frac{i}{i_0}) \tag{1.4}$$

Where ΔV_{act} is the activation losses; R is the ideal gas constant; α is the transfer coefficient; i is the current density; i_0 is the exchange current density.

The Ohmic loss is the voltage loss caused by the resistance of the flow of ions in the electrolyte and the flow of electrons through the conductive components. The losses can be calculated according to Ohm's law as equation 1.5.

$$\Delta V_{Ohm} = iR_i \tag{1.5}$$

Where the ΔV_{Ohm} is the Ohmic losses; R_i is the internal electrical resistance of the fuel cell.

The concentration losses are caused by the quick consummation of reactants at the electrodes. When the current is close to the limit current, the concentration gradients are established rapidly, thus causing voltage losses. It can be represented by the equation 1.6.

$$\Delta V_{conc} = \frac{RT}{nF} \ln(\frac{i_L}{i_L - i}) \tag{1.6}$$

Where the ΔV_{conc} is the concentration losses; i_L is the limit current density. This loss is only noteworthy when the current density is high.

Therefore, the actual voltage output of the fuel cell can be calculated by the theoretical potential and the losses as equation 1.7. More details about the output voltage and voltage losses can be found in the references [9].

$$V = E_r - \Delta V_{act} - \Delta V_{Ohm} - \Delta V_{conc} \tag{1.7}$$

The PEMFC stack is made up of several cells and the BPs are shared between two inner cells, so as to have a bigger power capacity by superposing the voltage of all the cells.In a stack, there are more design issues than in a single cell, because the distribution of gases along the cells as well as the heat must be better managed [10].

1.2.2 PEMFC faults

As both fluid flow, chemical reactions, heat transfer, and electrical processes are involved in PEMFC, it is a complex task to manage all the processes. If the operating

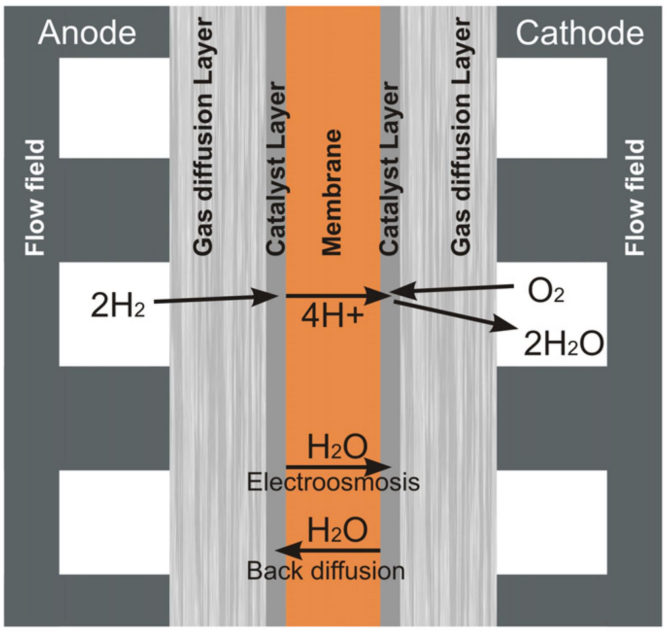


FIGURE 1.3: A diagram of the water transport in PEMFC [12]

parameters are not set correctly or the components degrade, the fault may happen and harm PEMFC durability and reliability. Fault means that the system is not operating under the ideal conditions or state. The faults of PEMFC can happen in all aspects, including gas leakage, heat management fault, water management fault, mechanical force damage, electrochemical reaction fault, etc [11].

1.2.2.1 Water faults

One of the most important problems in PEMFC is the water management fault. Different from other types of fuel cells, the electrolyte in PEMFC is the membrane, and it is solid-state. Therefore, to let the proton pass through the membrane, the membrane should be moist. The water transport in PEMFC can be shown in figure 1.3 [12]. During the operation, the water is produced in the cathode, and it should be taken out by the gas flow. At the same time, the water can pass the membrane from the cathode to the anode by diffusion. Proton will take water molecules from the anode to the cathode by drag effect, so that to pass the membrane with the help of the water molecules.

If too much water is produced or the water produced is not removed in time by the airflow, the accumulated water blocks the flow path. In addition, because the catalyst is covered with liquid water, the reaction zone is reduced. Also, as the catalyst is covered

by liquid water, the reaction area will reduce. Therefore, the performance of PEMFC will be greatly reduced. On the contrary, when the humidity in the PEMFC is insufficient, the heat conductibility of the membrane will greatly decrease; thus the performance will be reduced. Even worse, the membrane has the risk of breaking when it is extremely dry; thus the structure may be damaged permanently. Therefore, flooding and drying-out are the most important water defects in the PEMFC [13].

1.2.2.2 Temperature faults

Except for water management, temperature management is also important for the normal operation of PEMFC. All the components need to work at their optimal temperature to prevent degradation and guarantee overall performance. For example, the water saturate pressure is decided by local temperature, and the relative humidity will greatly change if the temperature is unstable or uneven. If the temperature is too low, the liquid water will appear in the GDL and cause a flooding fault, which is very harmful to the gas transfer and will cause the voltage to decrease and fluctuate [14]. On the contrary, if the temperature is too high, the membrane and catalyst layer will dehydrate; thus the drying out fault may appear. What's more, there may exist some hot spots in extreme conditions [15], which can cause damage to the component structure. Those damages are irreparable, and even the security of the system cannot be guaranteed.

To be aware of the temperature situation in the PEMFC, temperature mapping tools can be applied to obtain the temperature distribution inside the PEMFC. A typical temperature mapping result is shown in figure 1.4 [16]. The thermal unevenness, location of extreme hot points and the dynamic change of temperature field during the operation researchers have been studied [16].

1.2.2.3 Other faults

Several other defects can occur when operating conditions are not good or when there is degradation. In particular the following:

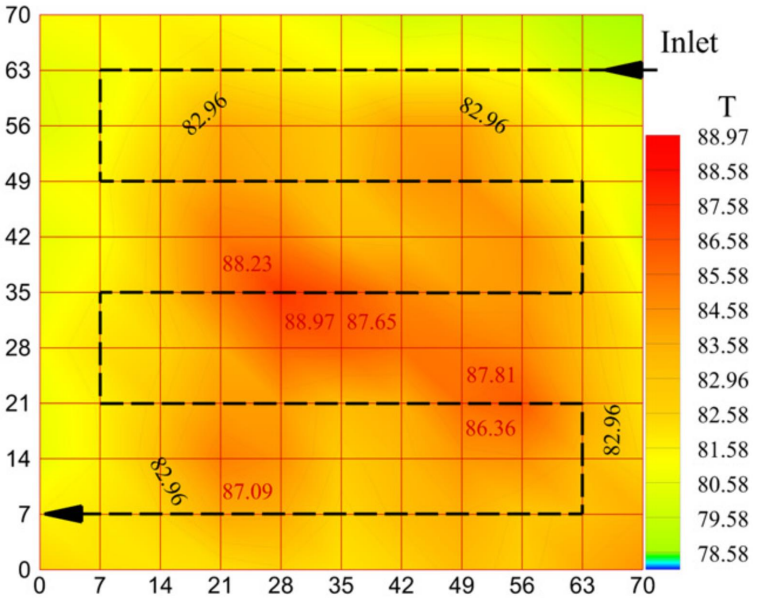


FIGURE 1.4: A typical temperature mapping result [16]

1. Poisoning. As the membrane and catalyst may react with some impurity species in the gas supply, the membrane and catalyst will be poisoned and damaged, sometimes even under very slight concentrations. For example, the SO_2 can cause dysfunction of catalyst layer [17]. Therefore, it is also a major fault source in PEMFC.

2. Mechanical damage. Normally the stack is well installed and protected, however, as the real operating conditions are complex, it is possible that the inner structure is mechanically damaged. Also, as the MEA is very thin, especially the membrane, the unbalanced gas pressure in the anode and cathode may also cause damage to the components.

Also, there are other faults that are related to the degradation of the components. The importance of different components can be listed as table 1.2 [11]. It can be seen that the health state of the membrane is the most important, followed by the electrode. It is therefore particularly important to reduce defects in these components.

TABLE 1.2: Criteria for components importance ranking [11]

Criteria	Bipolar plates	GDL	Electrodes	Membrane	Sealing gaskets
Does the component has a role in producing the output energy?	Yes(+)	Yes(+)	Yes(+)	Yes(+)	No(-)
Does a failure leading to a loss of power exist?	Yes(+)	Yes(+)	Yes(+)	Yes(+)	Yes(+)
What is the importance of this power loss?	Weak(+)	Weak(+)	Strong(++)	Strong(++)	Weak(+)
Is there a failure that prevents the components from partially or fully performing their functions?	No(-)	No(-)	No(-)	Yes(+)	Yes(+)
Does a failure leading to stack death exist?	No(-)	No(-)	No(-)	Yes(+)	No(-)
Does the degradation vary with the current profile required?	Yes(+)	Yes(+)	Yes(+)	Yes(+)	Unknown
Total	2	2	3	7	1
Component ranking	3	4	2	1	5

1.2.3 PEMFC diagnosis and prognosis

1.2.3.1 PEMFC application situation

As is addressed above, the different fuel cells have their advantage and disadvantage, and the main application of fuel cells nowadays are based on PEMFC. They have been applied to different situations, including portable, auto-mobile, and stationary applications [2]. However, it is the auto-mobile field that has the most passion, and a bright future is expected.

For commercialization, one of the most important factors is the cost, and the cost is affected by the production quantity in turn. An estimation of the relationship between the production and cost for fuel cell system is given in figure 1.5 [18].

According to the economic, society, and technology analysis, several major communities have proposed their targets for the production quantity of light-duty vehicles. The detailed target of the US department of energy (DoE), the EU fuel cells and hydrogen joint undertaking (FCHJU), and the strategy council road maps from China and Japan can be summarized in table 1.3.

Also, to compete with other technologies, the cost for each power unit is quite important. According to the estimation of technology development and production quantity, the estimation of cost per power unit is given in table 1.4. It is clear that the different

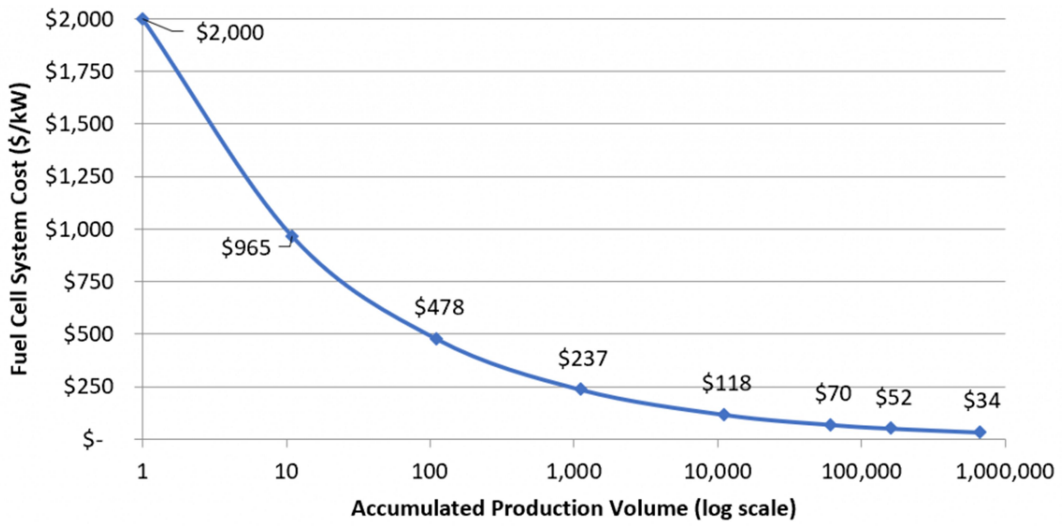


FIGURE 1.5: Estimation of fuel cell system cost relative to production volume [18]

TABLE 1.3: Assumptions for fuel cell system production for light-duty automotive

Target-setting entity	2020	2025	2030	2040
U.S. DOE	500000 systems per year	500000 systems per year	500000 systems per year	not given
EU FCHJU	100000 systems per year	100000 systems per year	100000 systems per year	not given
China	1000 systems per year	10000 systems per year	100000 systems per year	not given
Japan	40000 FCVs	200000 FCVs	800000 FCVs	3-6 millions FCVs

entities give a significantly different target for the short-term estimation, while they all give an optimist estimation of the long-term development. Maybe it is because the current situation is quite different for each entity, or also because they are estimated by different methods. No matter how much the cost is, it is believed that the fuel cell system can be an affordable and competitive technology after 20 years [18].

The average durability target can be shown in table 1.5. The target of heavy-duty vehicles is much longer than the light-duty because the light-duty vehicles usually face more complex working conditions. However, for both light-duty and heavy-duty vehicles, the lifetime should almost double until 2035 compared to 2020; therefore it is quite a challenging task.

TABLE 1.4: Assumptions for fuel cell system production for light-duty automotive

Target-setting entity	2020	2025	2030	2040
U.S. DOE	40$/kW	30$/kW	not given	25$/kW
EU FCHJU	66$/kW	55$/kW	44$/kW	not given
China	225$/kW	120$/kW	30$/kW	not given
Japan	72$/kW	45$/kW	36$/kW	18$/kW

TABLE 1.5: Durability target for light-duty vehicles and heavy-duty vehicles

Vehicles	2020	2025	2035
light duty	5000 hours	6000 hours	8000 hours
heavy duty	15000 hours	22000 hours	30000 hours

To have a lower cost and achieve successful commercialization, the main challenges and obstacles that should be dealt with are listed in figure 1.6 [18]. As can be seen from the figure, durability is very important and it can be improved in different aspects, including the component level, cell level, and stack level. To increase the stack lifetime, the diagnosis is an important part of the control system, as emphasized by the red box. The on-board diagnosis is just put forward, and it should be a mature technology before 2040.

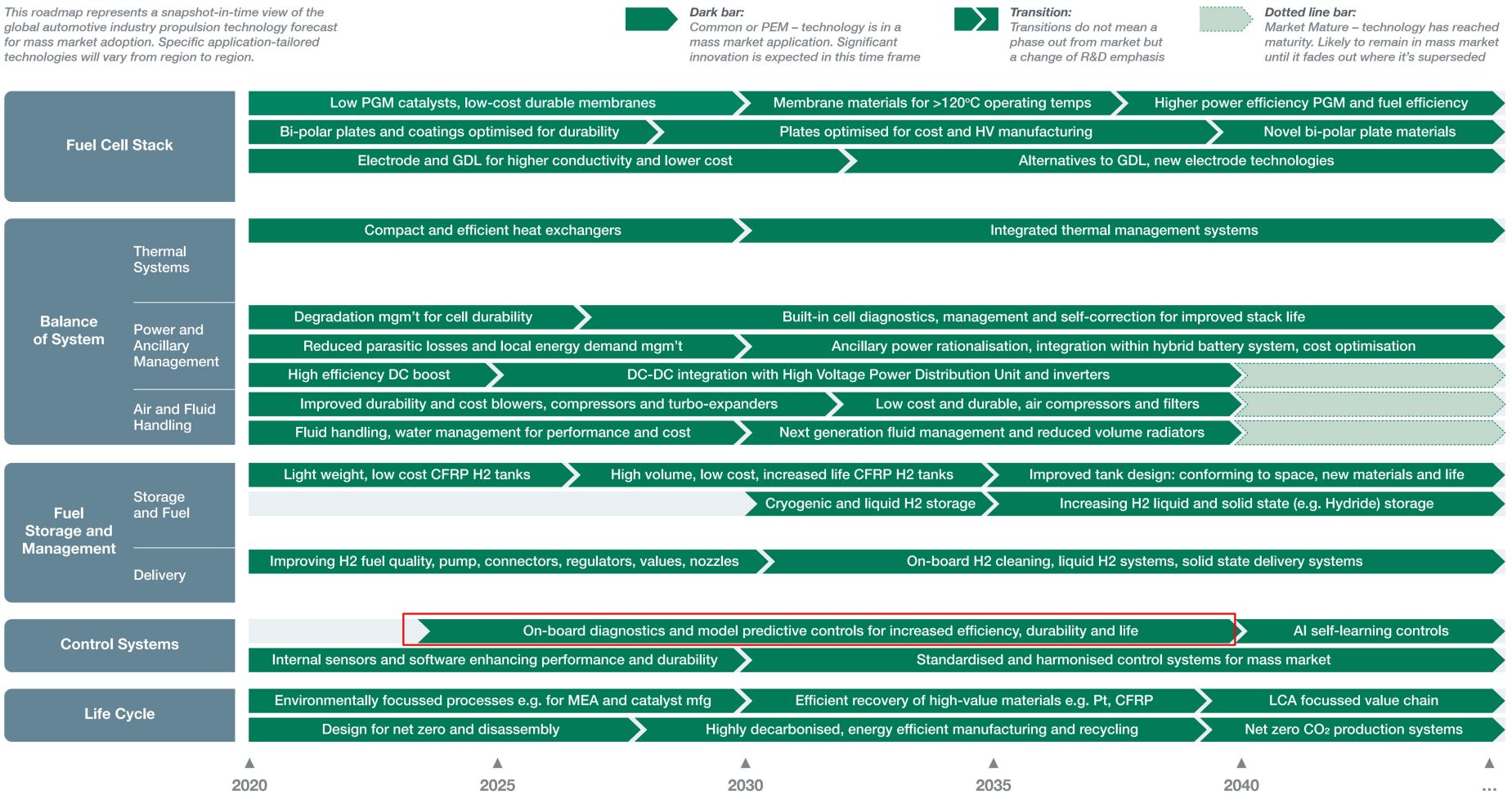


FIGURE 1.6: The main challenges and tasks for fuel cell development [18]

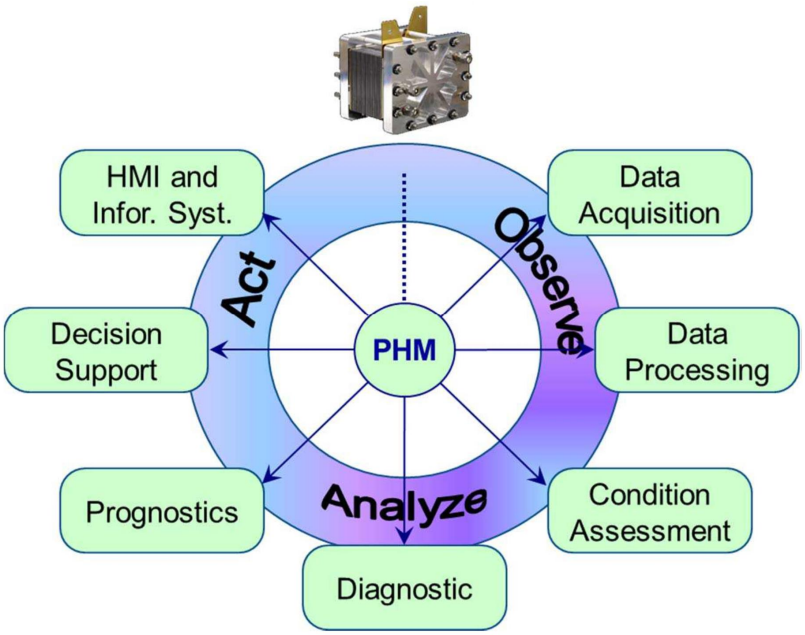


FIGURE 1.7: The position of diagnosis and prognosis in PEMFC health management [4]

1.2.3.2 Position of diagnosis and prognosis in health management

As is addressed above, the diagnosis and prognosis are important for the PEMFC system. The diagnosis and prognosis take the core places in the whole PEMFC health management field [4], and it was concluded as figure 1.7. The whole process can be divided into observation, analysis, and action, and both diagnosis and prognosis belong to analysis. They are based on the data from sensors and provide support for decisions.

According to the reference [19], diagnosis is the identification of the nature and cause of a certain phenomenon. Diagnosis is used in many different disciplines, to determine ”cause and effect”. In systems engineering, it is typically used to determine the causes, mitigation, and solutions of symptoms.

The diagnosis can be carried out at different levels, to improve the performance and detect the potential degradation. However, the diagnosis of this thesis should be distinguished from the diagnosis used in component-level research. For example, in the development of a catalyst layer with good performance, the catalyst may be tested under different unfavourable conditions, and to find what causes the performance to decrease and why [17]. Then according to the diagnosis result, new technology may be applied to build the catalyst. However, for the diagnosis of the PEMFC system, the main

task of the diagnostic tools is to maintain the operating performance, i.e. monitoring the health state of the whole system and giving notice to the controller if it is off the pre-set routine.

Further, the prognosis is another important concept in health state management. Different from diagnosis, the prognosis is applied to predict the long-term development of the PEMFC performance, and give suggestions to prolong the lifetime or arrange the power system [20].

1.3 Review of experimental tools for PEMFC system diagnosis

Different faults can be detected by diagnosis, so that to provide information for maintaining the normal operation of PEMFC and help increase the reliability and durability of PEMFC. There are two phases for the diagnosis of PEMFC. The first phase is the data acquisition by experiment tools, and the second phase is the algorithm design to analyse the data and obtain the diagnostic decisions [3]. The two phases react with each other. Usually, the mechanisms of the experiment can decide and inspire the data processing method and diagnostic algorithm design; thus some particular methods can be applied to particular experiment data [21]. At the same time, the methods to apply the data in practice should be analysed and considered at the beginning, to design better experiments to improve the data's practicality [22].

The main PEMFC diagnosis tools and methodologies are summarized in figure 1.8. The existing diagnostic tools and methodologies are reviewed in the following sections, respectively.

According to the experiment mechanisms and the types of obtained data, the diagnostic tools can be divided into two categories, i.e. electrochemical tools and physical/chemical tools [21]. The electrochemical tools are based on electrochemical reactions, and the electrical response such as potential and current can be measured as diagnostic data. According to the response, the inner processes can be inferred, so that to indicate the inner health state. Different kinds of stimulation can be used for diagnosis, and they can give information on different aspects. The widely used electrochemical tools include

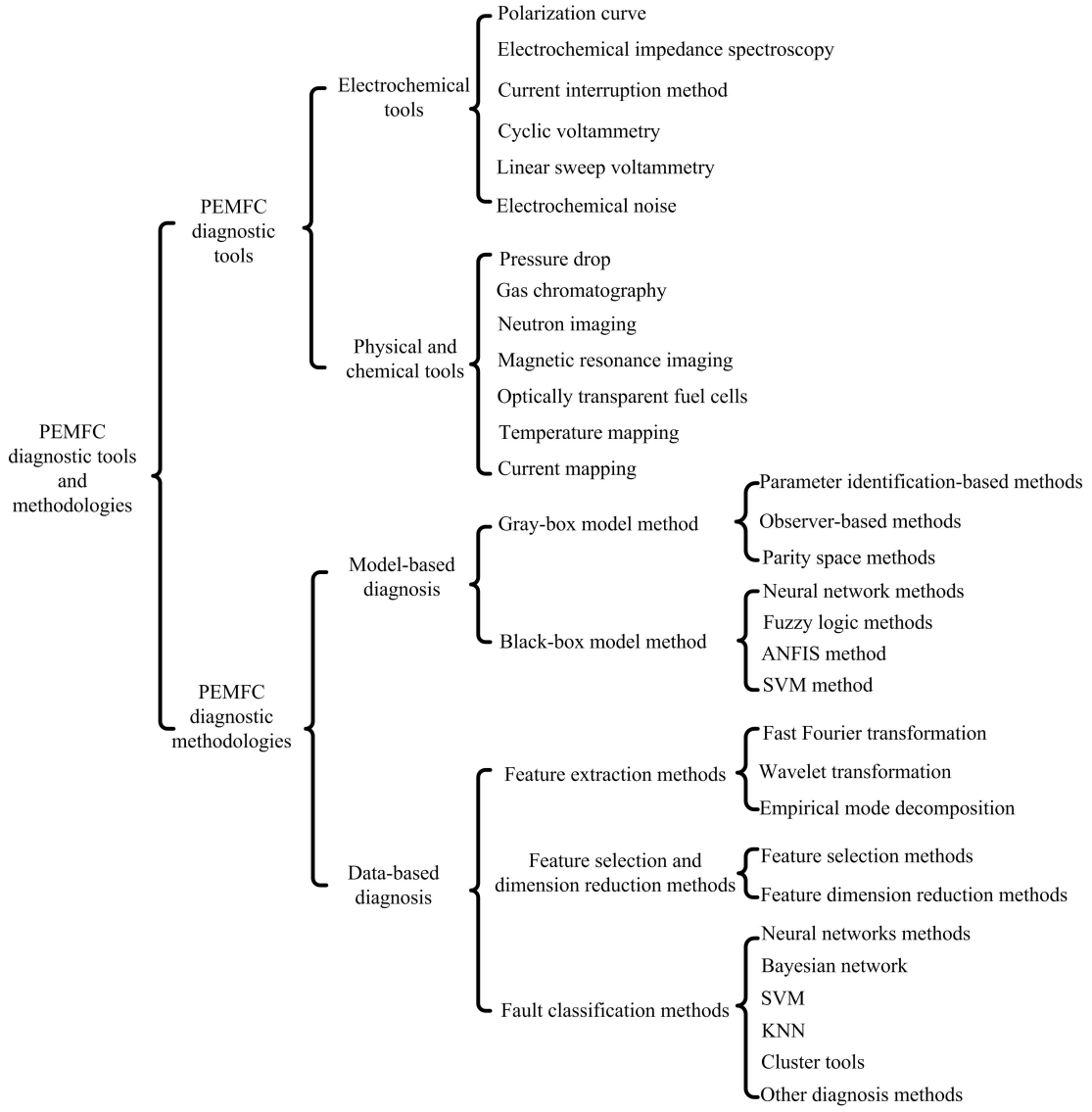


FIGURE 1.8: Synthesis of tools and methods for the diagnosis of PEMFCs

polarization curve, current interruption (CI), electrochemical impedance spectroscopy (EIS), cyclic voltammetry (CV), and linear sweep voltammetry (LSV), etc. [22].

Except for the electrochemical tools, there are also other experimental tools based on the physical/chemical processes in PEMFC [22]. Physical and chemical processes also play an important role during the operation, such as gas transportation, water balance, temperature maintenance, etc. The fault mechanisms of these processes should be understood by physical and chemical parameters, and the main physical/chemical diagnostic tools include gas chromatography, neutron imaging, magnetic resonance imaging, optically transparent fuel cells, temperature mapping, current mapping, etc.

However, the physical and chemical information used in the above methods can only

be obtained under specific ex situ conditions and they are not suitable for on-line diagnosis of the PEMFC system. Therefore, only electrochemical tools are presented and discussed in this thesis.

1.3.1 Polarization curve

For an electrochemical system like PEMFC, the plot of the output voltage to the output current is called the polarization curve. The polarization curve can be measured by setting several static currents (or voltage), and it can be used to represent the static performance of a PEMFC [8]. The typical polarization curve is given in figure 1.2.

As the polarization curve can give valuable information about the voltage of fuel cells under different currents, it is usually applied as an important indicator of the health state of the fuel cell [23] [24]. D. Bezmalinovic et al.[25] researched the degradation of PEMFC by the measurement of CV, EIS, and polarization curve. The analysis showed that the polarization curve can give similar conclusions about the degradation compared to CV and EIS. M. Mohsin et al.[26] used the polarization curve as a characterization tool and studied the performance of PEMFC after several voltage cycling tests under different relative humidity.

However, the voltage losses usually cover each other, and it is difficult to distinguish the contributions of different processes, so the polarization curves can only give qualitative conclusions. N. Fouquet et al.[27] carried out experiments and measured the polarization curve under both drying out and flooding fault conditions. Even though the polarization curves under both fault conditions are different from the normal condition, the polarization curves under drying out condition and flooding condition are very similar to each other, so it is impossible to distinguish the root causes of the fault.

Therefore, the polarization curve is an important health state indicator, but it is not suitable to distinguish the root fault causes only based on it. Also, as the measurement of the polarization curve needs to set the current at different levels, which will interfere with normal operation, it is impossible to detect it during the operation.

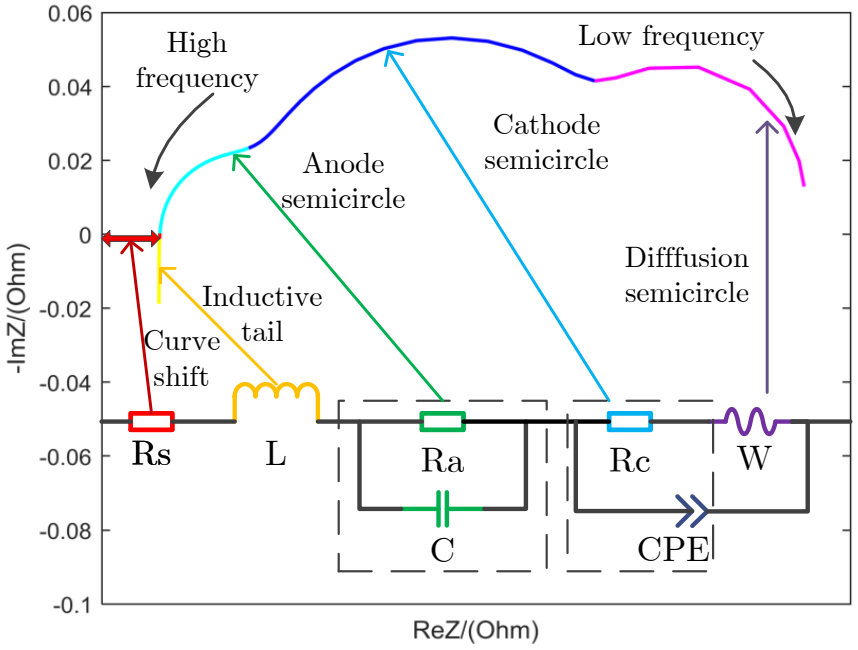


FIGURE 1.9: A typical Nyquist plot of EIS

1.3.2 Electrochemical impedance spectroscopy

EIS is another widely applied tool in the PEMFC diagnosis [28]. To carry out EIS, a small sine signal is superposed to the current (voltage) of PEMFC, and the voltage (current) response can be recorded. The magnitude and phase difference between the voltage and current can be obtained, and it is called electrochemical impedance. By adjusting the frequency of the sine signal, impedance spectroscopy can be obtained.

The typical representation methods of EIS result are Nyquist plot and Bode plot, which are shown in figure 1.9 and 1.10, respectively.

The Bode plot gives the modulus and phases under different frequencies, respectively. The Bode plot can give the frequency information, but it separates the impedance into two parts, which makes it less intuitive. Compared to the Bode plot, the Nyquist plot is more frequently applied. The Nyquist plot uses the real part of the impedance as the x-axis, while the imaginary part is the y-axis. Even though the exact frequency

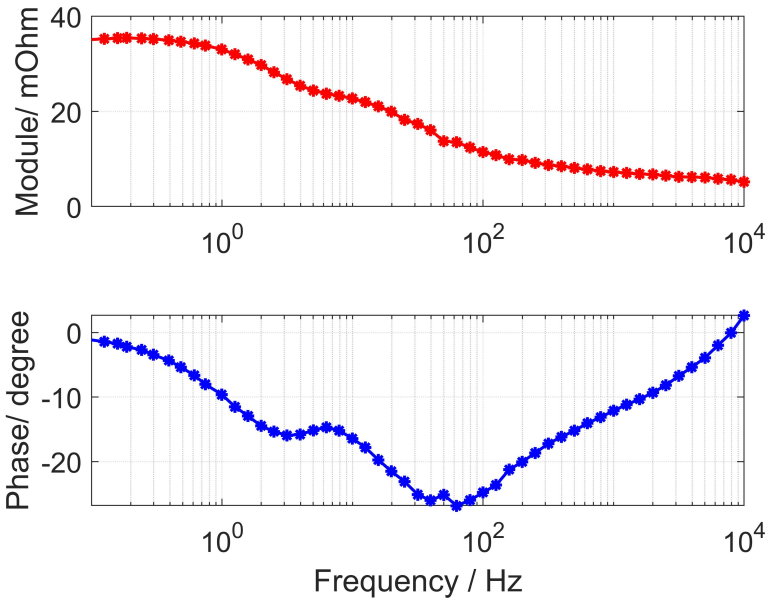


FIGURE 1.10: A typical Bode plot of EIS

information is not provided in the Nyquist plot, the transfer of the impedance from high frequency to low frequency is more intuitive than the Bode plot [29].

The impedance in different frequency ranges can represent different processes in the PEMFC. As can be seen in figure 1.9, normally there are three arcs that represent the impedance of the anode, cathode, and mass transfer process, respectively.

In the high-frequency zone, there is a small arc that represents the impedance of the anode, and it even disappears in some PEMFC. The anode arc is small because the oxidization of hydrogen is very easy to proceed with compared to the cathode. At the same time, the storage of electrons on the double electrode layers can be regarded as an electrical capacity, so the high-frequency arc can be represented by a parallel connection of resistance and a capacitor [30]. When the frequency decreases to the middle area, the dominant impedance can be represented by another arc, and it represents the cathode process. As the reduction of oxygen is much more difficult compared to the anode, the impedance arc is much bigger than the anode arc.

At very low frequency, the mass diffusion process in the gas diffusion layer (GDL) is dominant compared to anode and cathode impedances. The GDL is usually composed of porous material, and there are both resistance and storage effects for the reactants to reach reaction positions, so it will affect the reaction intensity. It is noteworthy that the impedance at low frequency is quite different for different PEMFCs, and it is supposed

to be affected by the structures of PEMFCs, such as the different shapes of the flow channel or structure of GDL [31].

There is always a curve shift of the real part at the high-frequency area, and it is observed almost in all PEMFC. This offset can represent the electrical resistance in the PEMFC, such as the resistance for ions to pass the electrolyte and for electrons to pass electrodes.

An important method to analyse the EIS data is to build equivalent circuit models (ECM) [32]. As shown in figure 1.9, the main idea of ECM is to match the EIS behaviour by electrical elements, such as electrical resistance, electrical capacitor, inductors, etc. The complex PEMFC system can be represented by an electrical circuit, so the inner state of the PEMFC can be represented by the elements of the circuit.

However, the ideal electrical capacitor cannot match the experiment results in most cases, so the constant phase element (CPE) was proposed and widely applied. The difference between CPE and an ideal electrical capacitor is that the CPE can be regarded as a capacity whose interface is rough, and the impedance of a CPE can be expressed as the parallel connection of numerous capacitors [33]. The CPE character of EIS is caused by the heterogeneity in the PEMFC, and the impedance of a CPE can be given as equation 3.1.

$$Z_{CPE} = \frac{1}{Qs^{\alpha}} \tag{1.8}$$

Where the Z_{CPE} is the impedance of CPE; Q is the nominal capacitance of CPE; s is the complex variable; α is the coefficient. When α is 1, the impedance of CPE is equivalent to a capacitor.

Another element is usually introduced to represent the mass diffusion impedance, i.e. the Warburg element. This kind of element is deduced from the semi-infinite diffusion phenomenon, and its impedance can be shown as equation 1.9.

$$Z_w = A\frac{\tanh(\sqrt{s\tau})}{\sqrt{s\tau}} \tag{1.9}$$

EIS is an important characterization tool for performance verification, such as the contamination effects of SO_2 [17], the effects of the micro-porous layer [34], etc. E. Balogun et al.[35] studied the PEMFC performance and durability with two different binders, and the EIS, polarization curve, and other characterization methods were applied to show the performance of different technologies. R. Pan et al.[36] used an ECM to fit the EIS data during a degradation experiment. The fitted ECM parameters show a good correlation with the degradation time; thus they are applied as indicators to predict the future degradation.

In conclusion, the EIS is a very important tool for PEMFC diagnosis. Compared to the polarization curve, it can distinguish the impedances of different components and different processes, so that to give more information about fault detection and fault isolation. What's more, as the measurement of EIS is based on adding small AC waves on a big DC, it has the potential to be applied in on-line diagnosis [37] [38]. Even though the accurate detection of EIS relies on special cumbersome equipment, nowadays more and more researchers tried to integrate the EIS measurement with existing control devices such as DC-DC converter [39] [40] [41] [42]. Therefore, EIS is quite a promising tool for both laboratory and practical utilization.

1.3.3 Current interruption

Another useful alternating current (AC) tool is the CI method. The basic idea of CI is to suddenly cut off the circuit, and record the change of voltage with the time immediately [43]. A typical result of CI can be shown in figure 1.11 [44].

The voltage response of CI can be divided into two parts: the immediate voltage increase, and the slow rise part. As has been analysed above, the inner resistance of PEMFC can be divided into a pure resistance part and another part with storage effects, so the immediate voltage increase can represent the voltage loss caused by pure resistance. For the slow rise part, it is usually taken as the activation loss. L. Carina et al.[45] applied the CI to the research of porous electrodes, and the over-voltage change in the whole process was analysed . M.A. Rubio et al.[46] carefully deduced the mathematical equations of the voltage response during CI, and the ECM was applied. The ECM parameters identification method was also proposed, and it was applied to the diagnosis of cathode flooding.

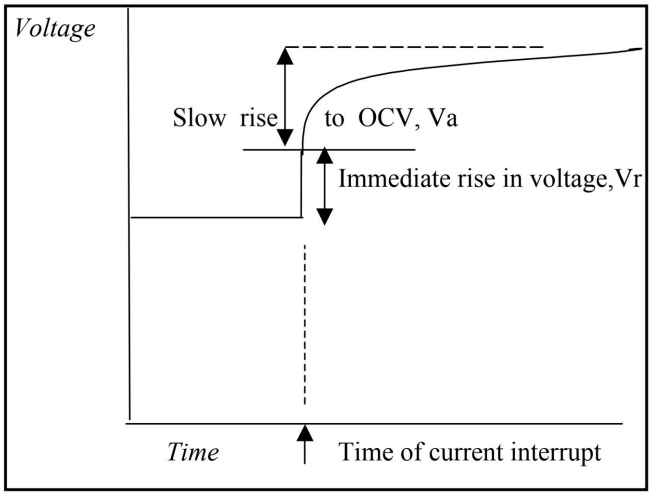


FIGURE 1.11: A typical result of current interruption for PEMFC [44]

In summary, CI is also a useful tool to know the inner electrical state of PEMFC. It is easier to perform than the EIS and also saves time, but it reveals less information than the EIS. There are also some variants based on the traditional CI, such as suddenly changing the current magnitude but not necessarily to zero; thus it may be applied to more conditions. Except for the electrical condition, no other conditions need to be modified during the diagnosis, so it is possible to have a wider application.

1.3.4 Cyclic voltammetry

Compared with other methods, CV is featured by its ability to provide information about the catalyst [47] [48]. CV has been widely used as an important tool to characterize the electrochemical active surface area (ECSA) [49]. The basic idea of CV is to impose a zigzag wave and record the current response, and the typical voltage imposed on the system as shown in figure 1.12(a) [50]. During the experiment, the anode is usually filled with hydrogen to provide a reference electrode, while the inert gas is supplied to the cathode and the corresponding CV result as shown in figure 1.12(b) [50]. The voltage increase process corresponds to the oxidation process of the catalyst, while the voltage decrease process corresponds to the reduction process. The relationship between the current and voltage is non-linear, and there is a hysteresis phenomenon [51].

As the CV is sensitive to the catalyst state and especially to the ECSA, it is widely applied as a characterization tool for the health state and performance assessment of

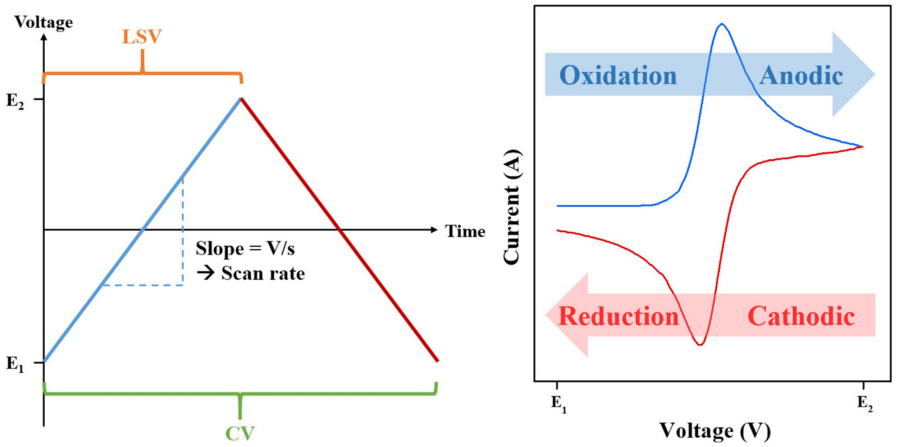


FIGURE 1.12: (1) A typical voltage applied for CV measurement and (2) corresponding CV result [50]

catalyst [52]. C. Roth et al.[53] synthesized a new Pt-Ru fuel cell catalyst, and it was compared with a commercial catalyst by CV. F.A. Bruijn et al.[54] studied the influence of carbon dioxide on PEMFC anode by experiment, and CV was applied to analyse the poison effect of CO_2 with different concentrations.

However, CV is not suitable for in situ applications because the experimental conditions are too strict. A special environment and gases are required during diagnosis, and the method is only suitable for simple electrochemical devices, but not for complex PEMFC systems. However, the attractive point of this method is that special information about the catalyst surface can be provided, which is unavailable for other methods. Therefore, this method can be used in the laboratory to research the degradation mechanisms or ex situ diagnosis applications.

1.3.5 Linear sweep voltammetry

LSV is similar to CV, but only the voltage increase part is applied. The anode is filled with hydrogen, while the cathode is filled with inert gases [55]. As the voltage increases with time linearly, the current response will change according to the inner state and processes. As is widely known, the membrane is not perfectly impassable for hydrogen; thus there is crossover current and it causes voltage losses. The detected current is composed of crossover current and short-circuit current, so it is a useful tool

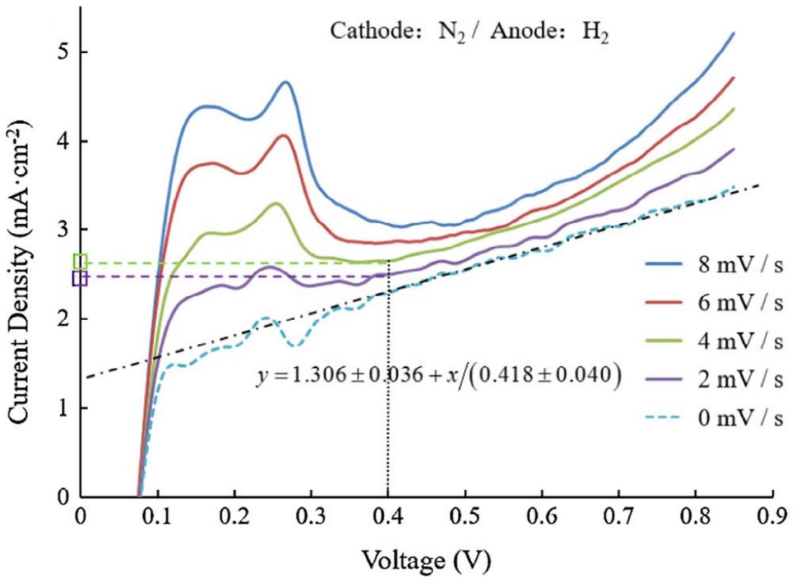


FIGURE 1.13: A typical LSV result [56]

to characterize the crossover current of the membrane. A typical LSV result is shown in figure 1.13 [56].

As LSV can give information about the crossover current and indicate the health state of the membrane, it is widely applied to characterization and diagnosis. J. Kang et al.[57] studied the degradation of a PEMFC after wet and dry gas cycles. The CV and LSV were applied and the crossover current increased with the degradation process.

Similar to CV, both the gas and other conditions cannot be satisfied when the PEMFC is operating, so the LSV is not suitable for on-line applications because of the special experimental requirement. Therefore, the method is only useful for laboratory experiments and ex situ diagnosis.

1.3.6 Electrochemical noise

During the operation of PEMFC, usually the current is controlled, so the voltage output is an important indicator of the performance assessment because it will directly decide the output power. However, it is found that not only the voltage magnitude is important, but the voltage fluctuations, i.e. the electrochemical noise (EN), is also a useful indicator of the inner health state, so it may be used as a diagnosis tool.

EN method has been applied as a diagnostic tool for a long time, but mainly in the field of electrochemical corrosion and coating process. EN is determined by electrochemical processes, but it is also affected by related processes such as mass transfer, gas evolution transfers, electrode corrosion, water transport and temperature distribution [58]. Therefore, it is still quite an open but also difficult area to analyse the basic mechanisms behind the phenomenon. However, even though there is no model to deal with the mechanisms, some experiments have been carried out to discuss the potential to use this phenomenon as a diagnosis tool.

B. Legros et al.[59] evaluated the possibility to apply EN as a diagnosis tool. They analysed the EN in the frequency domain by fast Fourier transformation. Power spectrum density (PSD) was calculated and it was sensitive to the change of operation condition. A typical PSD result is shown in figure 1.14 [60]. M.A. Rubio et al.[61] analysed the EN data by wavelet analysis, and the magnitudes under different frequencies were compared and discussed. R. Malizia et al.[60] created the term ”internal intermittence”, i.e. the statistical features in the time domain including standard deviation, skewness, and flatness in small-time windows. They found that the internal intermittence was sensitive to water management, and they dramatically increase under both flooding and drying out conditions. E.A. Astafew et al.[62] [63] published several research papers about using EN as a diagnosis tool in PEMFC. They analysed the spectral power density and frequency dependency relationship and connected the equivalent circuit with the noise, which is very inspiring. They also proposed the hypothesis of the EN formation mechanisms according to the experiment data, and some other statistic features were also researched [64][65].

EN is a natural part of the output voltage, and normally the voltage output of PEMFC applications can be easily recorded without extra costs. Therefore, a great advantage of EN is the on-line property, because the diagnosis can be done during operation without any intervention. So far, the EN has been analysed for different PEMFC systems by some frequency domain transformation tools, but no conclusion has been widely confirmed. Essentially, this is caused by the lack of understanding of EN generating mechanisms, and the critical experiment parameters may have changed between different experiments. For example, the EN caused by unstable gas supply and the EN caused by water management fault cannot be distinguished yet. Therefore,

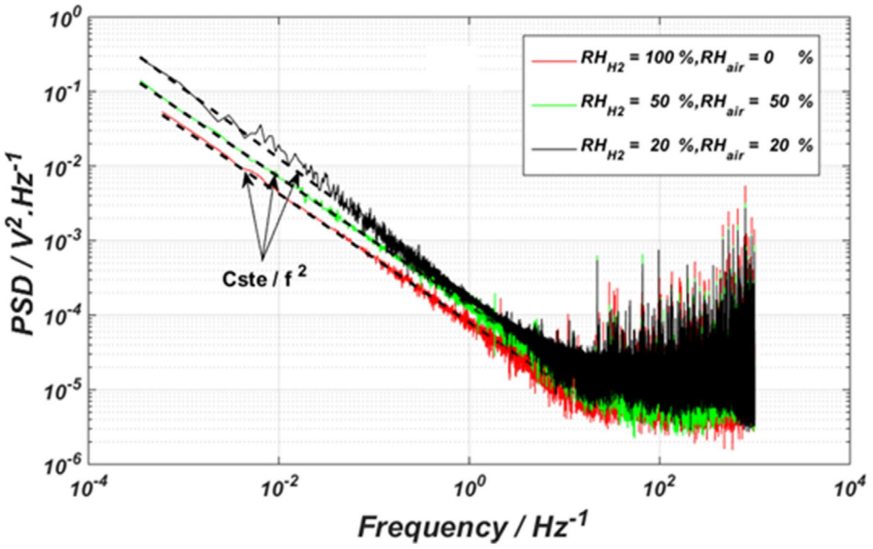


FIGURE 1.14: A typical power spectrum density analysis of electrochemical noise result [60]

it is quite an interesting field to find out the basic mechanisms. Meanwhile, as the implementation is very easy, it is also significant to develop some other algorithms.

1.4 Diagnosis algorithms review

As the diagnostic data are collected, the diagnostic algorithms can be applied to give diagnostic results. According to whether using a model, the diagnostic methodologies can be divided into two categories: model-based methods and data-based methods [66]. Model-based methods are based on the models that simulate the processes inside PEMFC; thus the model output can be compared with real output to generate residuals, which can be employed to decide the health state of the PEMFC system. On the contrary, data-based methods are completely based on data processing or classification algorithms, and no model is needed. These methods are widely applied as it is still difficult to build precise physical models for the complex PEMFC system, and also because of the flourishing development of artificial intelligent methods (AI) [67].

1.4.1 Model-based Diagnosis

For the model-based methods, the basic idea is to calculate the supposed output based on a model and compare the results with the real output, and the fault can be declared

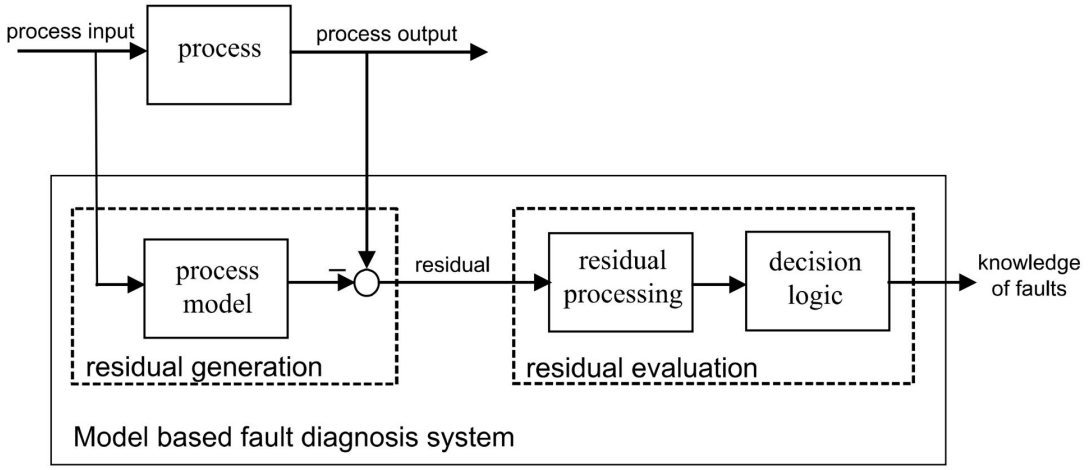


FIGURE 1.15: Typical structure of model-based Diagnosis methods [68]

if the result residuals are beyond the threshold. Therefore, the model-based diagnosis is also called residual-based diagnosis. The process to decide the correspondence between residuals and faults is called residual evaluation. A typical diagram of model-based methods can be shown in figure 1.15 [68].

According to how much the model is related to physical mechanisms, they can be divided into three types, i.e. white-box model, grey-box model, and black-box model. The white-box model is totally deduced by physical knowledge, and all the relationships and coefficients can be expressed explicitly. In these models, the physical mechanisms can be fully described by ordinary or partial differential equations or other mathematical forms [66]. For example, the Nernst-Planck equation, Butler–Volmer equation, and Fick's laws are widely used to describe the electron transfer process. However, as the PEMFC is a very complex system that is related to mass diffusion, heat transfer, electrochemical reactions, and chemical reactions, it is still impossible to build an accurate and universal white-box model for the PEMFC system.

To apply the observed physical and chemical rules in the PEMFC model, a more practical method is to build grey-box models [66]. Unlike white-box models, the relationship in a grey-box model is partly based on physical mechanisms, and partly based on experimental induction. Therefore, a lot of empirical formulas are applied in grey models, and the coefficients are usually fitted according to the experiment data.

As the behaviours of PEMFC systems are quite complex and non-linear, a lot of unknown variables and processes can affect the performance of PEMFC. Therefore,

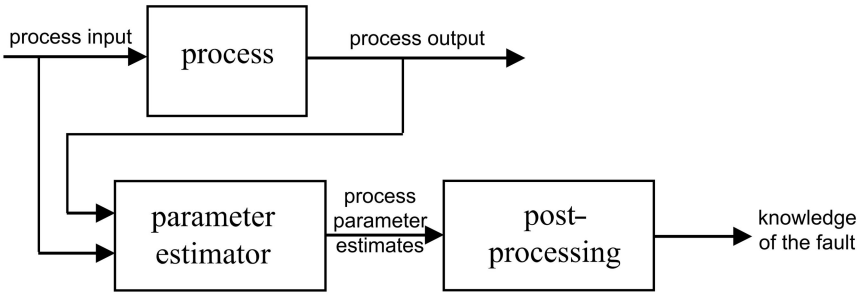


FIGURE 1.16: Typical Diagnosis process based on parameters identification [68]

models based on analytical equations are often difficult to design. Thus, black box models are proposed and applied to modelling. In black-box models, no physical relationships and mechanisms are needed, and the model can be obtained by training based on the experiment data; as a result they are widely used nowadays [66].

1.4.1.1 Diagnosis based on grey-box models

As has been addressed above, accurate and universal white-box models are hard to build; thus most diagnosis research is based on grey-box models. Those methods can be divided into three categories, i.e. parameter identification-based methods, observer-based methods, and parity space methods.

(1) Parameter identification-based methods

The diagnosis of a PEMFC system can be achieved by identifying the model parameters. A typical diagnosis process based on parameters identification is shown in figure 1.16 [68]. The model parameters can be identified according to the system input and output, so the obtained model parameters under the real condition can be compared with the parameters under normal conditions. Several parameter identification methods were applied, including the least square method, genetic algorithm (GA), etc.

One of the most widely applied diagnosis models is the polarization curve model. S. Wu et al.[24] applied the polarization curve data to analyse the ohmic and activation loss gradient, and they were employed as residuals to distinguish the fault conditions of normal conditions.

The system time constant was also applied to the diagnosis of PEMFC [69]. As obtaining EIS is too complex in measurement time and equipment, the voltage

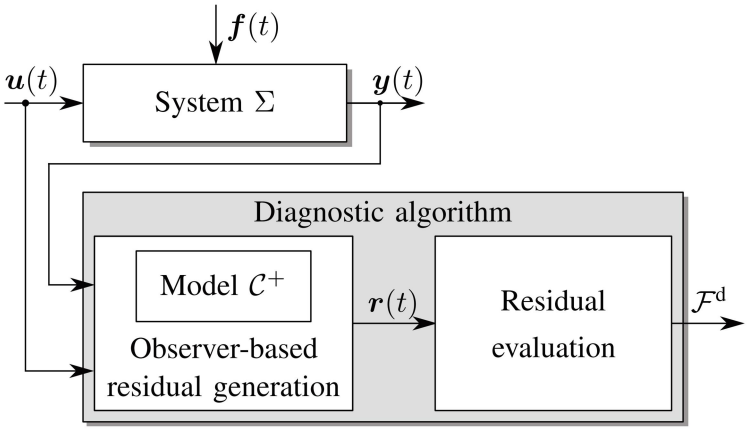


FIGURE 1.17: Typical structure of observer-based diagnosis [70]

response to current steps was applied as a diagnosis tool, i.e. CI data. The evolution of the time constant during the flooding process can be identified, and the time constant can be an indicator of flooding fault.

(2) Observer-based methods

Another important model-based diagnosis method is the observer-based method. Different from parameter identification methods, the main idea of observer-based methods is to design all kinds of observers to obtain the inner state variables or to reconstruct the undetected inputs, thus generating residuals and indicating the health state. A typical diagram of observer-based diagnosis method is shown in figure 1.17 [70].

(3) Parity space methods

Parity space method is another method to design the residual generator. The system model can be linearised in a discrete subspace, and the residual can be built by parity relations. Therefore, it is represented by linear algebraic equations, which is more convenient for computation. The diagram of parity space diagnosis method can be shown in figure 1.18 [71].

Q. Yang et al.[72] focused on the residual generation by parity space method for the PEMFC air supply subsystem. The state equation model was reduced to 3 orders, and the non-linear analytical redundancy relations were applied to design the residual generator. The simulation showed that all three faults can be detected by the residuals. The parity space method was also extended to the whole PEMFC system, thus the

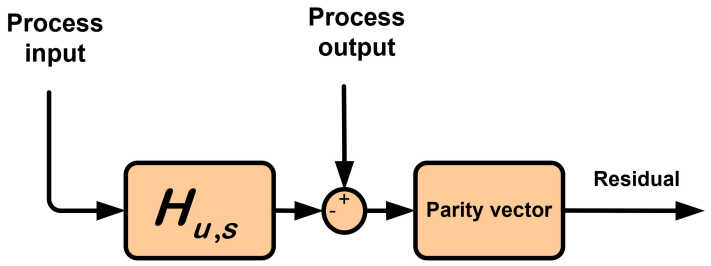


FIGURE 1.18: Parity space Diagnosis method [71]

system faults were detected, such as flooding and drying out [73]. Z. Li et al.[74] also connected the parity space strategy with data-driven methods, and the sophisticated modelling process can be omitted. Their method was also proved by experiment result.

The research works about PEMFC diagnosis based on grey-box models can be summarized in table 1.6. All the methods focused on the design of the residual generator, and the residual evaluation and signature rules are also important. Sensitivity analysis of the residuals to different faults is a useful tool for the fault signature, and advanced classification tools can also be applied to decide the fault.

TABLE 1.6: PEMFC diagnosis based on grey-box models

Methods	Models	Characteristics	Advantages
Parameter identification	polarization curve model [24];	Gradients of different phases from polarization curves were applied; results were compared with wavelet packet transform;	It can be applied to on-line diagnosis;
	equivalent circuit model for mass transportation [75];	the flow rate of oxygen, hydrogen and water in the PEMFC was analogized to the equivalent circuit;	the saturated conditions were considered by including liquid phase in the diffusion layers;
	equivalent electrical circuit model for current interruption [69];	voltage response of CI was used to fit the circuit, and parameters were compared with normal condition to generate residuals;	the time-constant spectrum and the dynamic flooding process for 10 minutes was demonstrated;
Observer-based	PEMFC operating model [76];	four residuals were generated and relative sensitivity was applied to isolate the faulty conditions;	faulty conditions can be separated by sensitivity even though all residuals are affected in the same direction;
	PEMFC linear parameter varying (LPV) model [77] [78];	the fault conditions can be isolated by relative sensitivity, and adaptive threshold was applied;	both the operating condition fault and sensor fault can be identified;
	First-order differential-algebraic PEMFC model [79];	residuals were created by high-index sub-models;	sub-models can be chosen by the structural analysis and causal computation;
	PEMFC model simplified by linear canonical variety analysis [80];	Both the inverse model method and Kalman filter method were applied to generate residuals;	non-linear PEMFC model can be approximated by linear state-space model with small order;
	PEMFC model simplified by consistency relations design [81];	The sub-models can be chosen by consistency relations design, and a residual can be designed for each fault;	each fault can be indicated by one individual residual, no need of residual evaluation;
	PEMFC air supply model [82];	a modified super-twisting (ST) sliding mode observer was designed to generate residuals;	the fault signal can be reconstructed and the state can be estimated during load change;
	PEMFC thermal model [83];	four residuals were generated according to the thermal model;	the fault of pumps, thermocouples, and heat exchangers can be isolated;

Continued

Continued

Methods	Models	Characteristics	Advantages
Parity space method	air supply subsystem model [72] and PEMFC model [73] ; parity vectors identified by orthogonal projection approach [74];	the PEMFC system model was linearised, residuals were generated based on parity space method; the residuals were generated based on the parity space, and SVM was used for residual evaluation;	both sensor, actuator and PEMFC water management fault can be detected and isolated; no need of sophisticated modelling, the parity space can be identified directly by data processing

1.4.1.2 Diagnosis based on black-box models

PEMFC is a complex system including a lot of processes that cannot be described by accurate mathematical models. Therefore, a more convenient method is to build the black-box models. Only the inputs and outputs data are needed to build the black-box model, so the complex physical mechanisms can be ignored, which is very practical.

As the residuals are created by the model results, the method to build a black-box model is critical for the diagnosis. Normally the models can be built based on artificial intelligent tools, and they can be divided into several types, i.e. neural network methods (NN), fuzzy logic methods, and adaptive neuro-fuzzy inference system (ANFIS) [66].

(1) Neural network models

The most widely used black-box models are all kinds of neural network methods. An artificial neural network (ANN) is a very good tool to build models for non-linear systems because it can represent different kinds of non-linear behaviours [84]. The nods in an ANN are called neurons, and their organizational structures can represent different topologies. According to the number of layers, they can be categorized as single-layer and multi-layer networks. According to the data transmission method, they can be divided into feed-forward, back propagation, and recurrent neural networks. They all have different characters and can be applied to particular cases. A typical structure of ANN can be shown in figure 1.19.

S. Jemei et al.[85] created an ANN model for an embedded fuel cell, and both static and dynamic models were considered. N. Y. Steiner et al.[86] moved one step forward towards the diagnosis based on the ANN model. Not only were the outputs of the PEMFC system predicted, but they were also compared to the experimental outputs to generate residuals for flooding and drying fault diagnosis. Rather than building a model to calculate the voltage output, S. Laribi et al.[87] [88] applied the ANN to model the EIS data. J. Park et al.[89] focused on the thermal management system, and an ANN model was built to simulate the relationship between the signals and faults. The current distribution was also predicted by the ANN model according to the given current density, temperature, pressure, stoichiometric ratio, and relative humidity [90].

(2) Fuzzy logic models

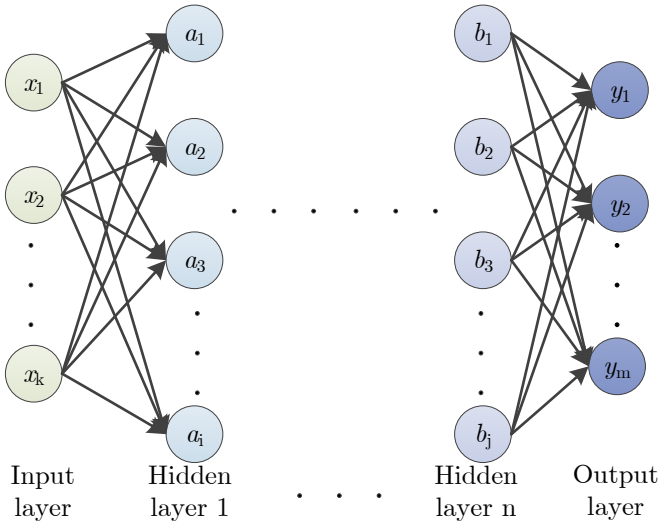


FIGURE 1.19: A typical structure of ANN

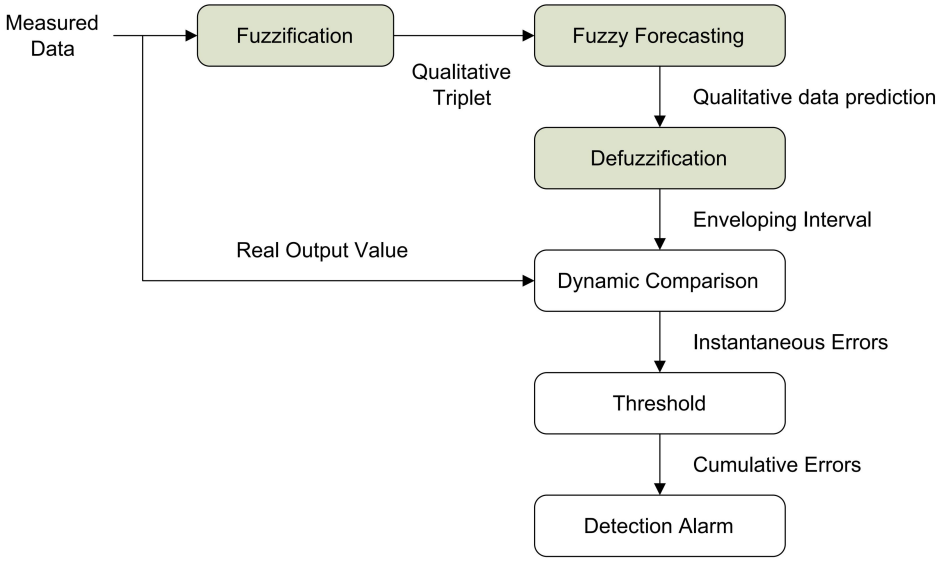


FIGURE 1.20: The typical structure of fuzzy logic Diagnosis [91]

Another widely used method is fuzzy logic models. The typical diagram of fuzzy logic diagnosis can be found in figure 1.20 [91]. Those methods are developed by imitating the human decision process, i.e. replacing the continuous input with the qualitative description, and finding the relationship based on the fuzzified inputs and output. Therefore, the fuzzy logic methods can be represented by some if-then relationships and their logical combination.

A. Escobet et al.[91] applied the fuzzy logic model to calculate the output, and the pattern recognition technique was applied to evaluate the residuals. B. Davies et al.[92]

developed the decision rules according to expert knowledge and experiment findings. M. Pei et al.[93] used the EIS experimental data as input to the fuzzy logic method, and they extracted the features of EIS data by fitting the equivalent circuit, rather than just choosing some points. G. Rubio et al.[94] applied the features of the current interruption experiment to decide the fault cases.

(3) ANFIS models

ANFIS method is another tool to build models for diagnosis purposes. It is the connection of neural networks and fuzzy logic; thus it combines their advantages. Different from the fuzzy logic method, the membership functions and decision rules are decided by training, rather than defined manually according to expert knowledge. Therefore, no particular rules are needed, which reduces the intervention of the researchers. What's more, compared with ANN methods, the data needed to train the model is also reduced because of the use of fuzzy logic [95].

K. Hustesen et al.[96] applied the ANFIS model to simulate the behaviours of PEMFC. The fuel cell temperature, current density, and carbon monoxide concentration of anode supply gas were set as inputs, and the voltage was the aim output of the model. K. Mammar et al.[97] also applied the ANFIS to PEMFC modelling, but they focused on the relationship between EIS data and humidity. The water situation was researched under different conditions, and it was compared with the output observation. T. Wilberforce et al.[98] also used the ANFIS method to predict the PEMFC performance, and the polarization curve can be predicted accurately.

The existing research based on the black-box model can be summarized in table 1.7. The applied methods, inputs, output, and characters of the study are presented.

TABLE 1.7: PEMFC diagnosis methods based on black-box models

Models	methods	inputs	outputs	Characteristics
Neural network	feed-forward NN and recurrent NN [85];	current, hydrogen and oxygen flow rates, temperature, air humidity level;	voltage;	the polarization curves can be predicted, and it corresponded well with experiments;
	Elman neural network [86];	current, air inlet flow rate, stack temperature, dew point temperature;	voltage output and pressure drop;	predicted outputs were compared with the real outputs to generate residuals for diagnosis of flooding and drying out;
	ANN [87] [88];	operating time and relative humidity;	ECM parameters;	the ECM parameters can be predicted and applied to diagnosis;
	ANN [89];	current, voltage, stack inlet and outlet temperature, radiator outlet temperature, pump outlet pressure and the fan control signal;	fault conditions;	6 ANN sub-models were built to reveal different faults;
	ANN [90];	current density, temperature, pressure, stoichiometric ratio, and relative humidity;	current density on 25 parts of the current collector;	the current distribution can be predicted by ANN, and it corresponded well with the experiment;
Fuzzy logic	Fuzzy inductive reasoning model and pattern recognition [91];	stack current and compressor voltage;	oxygen excess ratio, stack voltage, compressor current, compressor speed;	the predicted outputs were compared with the experiment data, and the method was effective even with 20 dB noise;
	Fuzzy logic and expert knowledge [92];	voltage, cycle number, anode stoichiometry, and humidity change, stack temperature;	dehydration certainty	expert knowledge was applied to decide the fuzzy logic rules;
	Fuzzy logic model and ECM [93];	ECM parameters identified according to EIS data;	water content	the fuzzy logic rules were directly decided by faulty mechanisms;
	Fuzzy decision tree [94];	voltage slope change, pressure change, and voltage oscillation;	water content	the current interruption experiment result and other features were employed as inputs;

Continued

Continued

Models	methods	inputs	outputs	Characteristics
ANFIS	ANFIS model [96];	fuel cell temperature, current density, CO concentration in anode supply gas;	voltage;	the voltage of high-temperature PEMFC was verified under different temperatures, current density and CO concentration;
	ANFIS model based on EIS [97];	EIS curve;	relative humidity;	water activity can be estimated based on EIS and ANFIS;
	Multiple regression analysis (MRA) and ANFIS model [98];	oxygen and hydrogen flow rate, oxygen pressure, hydrogen pressure;	current and voltage;	the performance of PEMFC was predicted based on the parameter selection by MRA;

1.4.2 Data-based Diagnosis algorithms

Different from the model-based methods, the data-based diagnosis methods need no models for the PEMFC system, and the diagnosis decisions can be directly obtained based on the diagnostic data [67]. Therefore, the inputs of data-based methods are features, and the output is the fault category. Classification or cluster methods are applied to decide the fault categories by features, including different kinds of NN, fuzzy logic, support vector machine (SVM), Bayesian network (BN), etc.

The choice and processing of features are quite important for data-based diagnosis. The features can be any useful information about the operation processes and health state of PEMFC [99]. The polarization curve data, the EIS data, current interruption data, and other operating parameters like the pressure, the temperature, the current, and the voltage are all useful. To obtain the most useful information with fewer calculation burdens, the feature selection and feature reduction methods were widely used. The purpose of feature selection methods is to select the most representative data in a group of data, and the selection can be based on the analysis of information amount. The feature reduction methods are used for reducing the dimension of the inputs, and the most important features can be obtained by the transformation of coordinates or other methods. Principal component analysis (PCA), discriminant function analysis (DFA), kernel principal component analysis (KPCA), and kernel discriminant function analysis (KDFA) are the most used feature reduction tools [100].

If the features are not explicit, some feature extraction tools are needed to obtain the most important information from the original data. The most widely applied feature extraction methods include the Fourier transform (FT), wavelet transform (WT), etc. Therefore, feature extraction, feature reduction, and fault classification are three important processes for PEMFC diagnosis. The typical data-based diagnosis methods and tools can be summarized in figure 1.21 [101].

The objectives are different for the feature extraction tools, feature reduction tools, and fault classification algorithms [67]. Generally, a fault classification algorithm is indispensable, because it is used to build the relationship between the feature data and the fault types. Secondly, the feature reduction tools are usually applied to reduce the computational burden, when there are too many samples to deal with and the feature

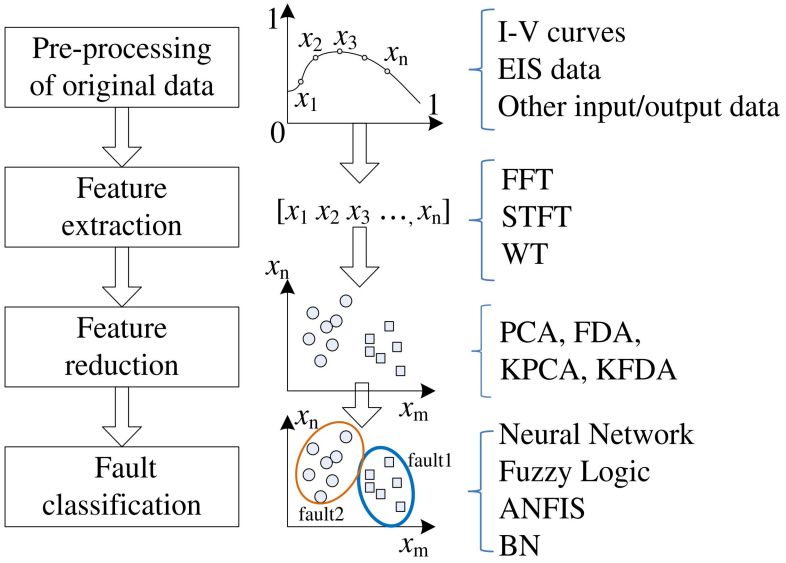


FIGURE 1.21: Typical processes of data-based Diagnosis methods [101]

dimension is too big. On the contrary, the feature extraction tools are only used in some particular problems where the data show some characters in the frequency domain or other aspects. Therefore, these methods are summarized below in order of their necessity.

1.4.2.1 Feature extraction methods

Actually, feature extraction from the original data set is quite a particular task for different kinds of data and different problems. Therefore, a lot of different methods have been applied based on the physical characteristics of the measurement process [102]. For time-series data such as voltage and pressure, some frequency domain information or other information can be extracted from the original data by domain conversion tools, including Fourier transform, wavelet transform methods, and empirical mode decomposition.

(1) Fourier transform

Fourier transform of a time-series signal can transform the signal from the time domain to the frequency domain, and some frequency domain information will appear. Therefore, the Fourier transform is widely used [14]. In real applications, the signals are always discrete, thus discrete Fourier transform (DFT) is applied, and FFT is a simple implementation method of DFT.

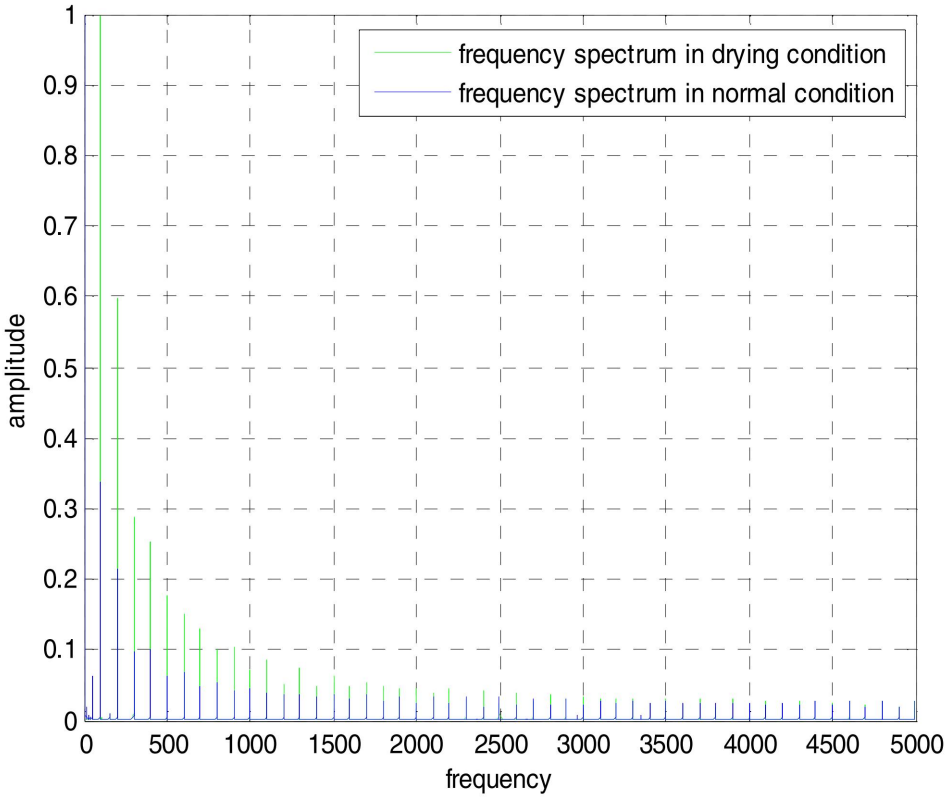


FIGURE 1.22: A typical FFT result of PEMFC voltage [103]

The signal frequency-domain behaviour can be different for faulty conditions and normal conditions. Usually, the gas pressure and voltage output data can be transformed to the frequency domain, because the fluctuation mode contains some health state characters of the fuel cell stack. A typical diagram of the voltage frequency domain spectrum can be shown in figure 1.22 [103]. It can be seen that the amplitude is totally different for different conditions.

J. Chen et al.[102] observed the oscillation of the voltage at the high air stoichiometry, and the FFT was performed for the pressure drop. M.A. Rubio et al.[61] carried out the FFT and wavelet transformation for the voltage data, and the result was compared with the EIS data. A.H. Detti et al.[103] and N.J. Steffy et al.[104] applied the FFT to the voltage data under drying out and flooding conditions, and the total harmonic distortion of the frequency spectrum was applied to the diagnosis. R. Maizia et al.[60] applied the FFT to analyse the electrochemical noise, and the power spectral density under different frequencies was obtained and compared under different situations.

(2) Wavelet transform

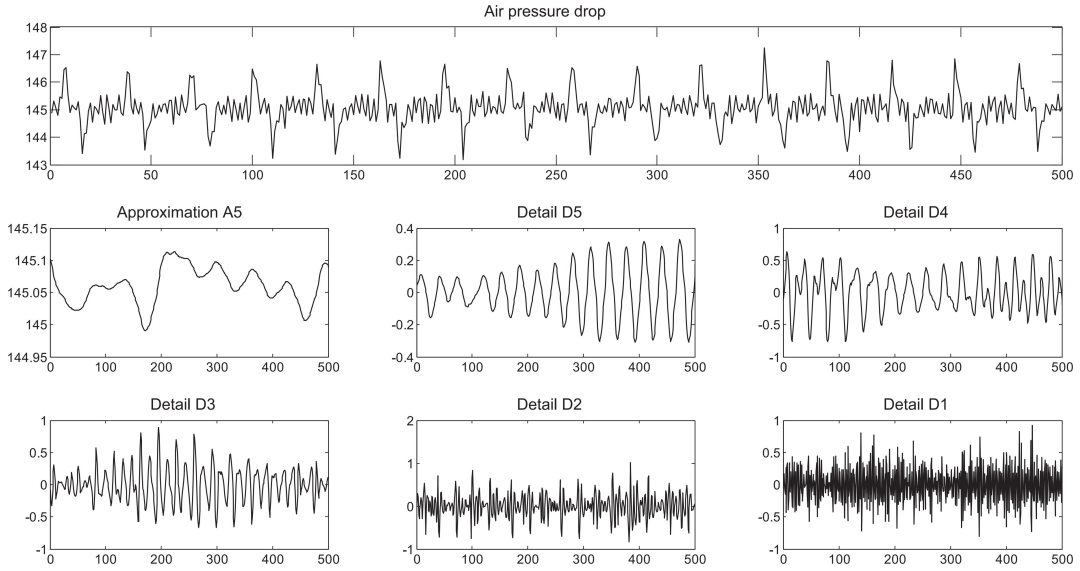


FIGURE 1.23: A typical wavelet transform result of PEMFC pressure drop [105]

A great limitation of FFT is the loss of time information, i.e. the frequency spectrum can represent the signal frequency behaviour during the whole process, but the transient behaviours cannot be represented. However, the transient state change is quite normal in the PEMFC system, and FFT cannot be applied to dynamic conditions. To show the transient behaviour, a novel method is proposed by using slide windows to carry out FFT, i.e. STFT, but it is difficult to decide the window length. If the window is too small, the signal cannot represent some low-frequency behaviour, and the frequency resolution is poor. On the contrary, the time resolution will be insufficient when the window is too wide. Therefore, it is not very convenient to deal with the dynamic signal with FFT [103].

Compared to the Fourier transform, the wavelet transform can give information in the frequency and time domain. The basic idea of this method is to extract the signals from wavelets of different lengths, which allows representing the behaviour of low and high frequencies with an appropriate resolution. A diagram of the wavelet transform process can be shown in figure 1.23 [105].

For the diagnosis of PEMFC, a lot of research dealt with the time series operating data by wavelet transformation. N.Y. Steiner et al.[106] proposed a non-intrusive diagnosis algorithm by wavelet transform of voltage data. Further, D. Benouioua et al.[107] [108] applied the wavelet transform to a multi-fractal formalism, and the singularity of the multi-fractal spectra was employed as a signature of the faulty operating conditions.

J. Kim et al.[109] analysed the discrimination ability based on wavelet decomposition, and the statistical quantities were applied to represent the state of health of the PEMFC stack. M. Ibrahim et al.[110] compared the discrete wavelet transformation (DWT) with continuous wavelet transformation, and they found that the DWT was advantaged in the localization and execution time.

The electrochemical noise was also analysed by the wavelet transformation, and it was compared with the FFT method and proved more effective [61]. E. Pahon et al.[111] [105] not only applied the wavelet transformation to deal with voltage data but also pressure drop data. The WT was also used to diagnose the water content fault according to the energy intensity of reconstructed vibrating voltage [112].

(3) Empirical mode decomposition

Some other methods were also applied to extract features from time-series data, such as empirical mode decomposition (EMD) [113]. EMD can be used to decompose the original signal into several signals based on the mode function of the original signal, so no extra basis functions are needed. But it is not widely used yet.

1.4.2.2 Feature selection and dimension reduction methods

One of the objectives of the PEMFC diagnosis algorithm design is high computational efficiency so that it can be used for on-line applications. However, as various kinds of data can help with the diagnosis, the feature dimension is usually too big and the diagnosis time is too long. To obtain enough information with fewer features, several feature selection methods and feature dimension reduction methods have been proposed and applied in the PEMFC diagnosis [100]. Although the goals of the feature selection method and the feature dimension reduction methods are both to reduce feature dimensions, the internal logic is not the same. For feature selection methods, the objective is to find and drop the useless features according to some analysis and statistical criteria, while the other important features will not change. On the contrary, for feature dimension reduction methods, the new features are created according to the original features data, and it extracts useful information from all the features.

The feature selection methods and feature dimension reduction methods are applied for different purposes. Feature selection is used when the dimension is redundant and

certain features are useless, while features dimension reduction methods are employed when all the features carry some information for the diagnosis.

(1) Feature selection methods

There are three feature selection methods used in the PEMFC diagnosis, i.e. the variance analysis method, coefficient correlation method, and information gain-based feature selection method [100]. The variance analysis method is applied to characterize the degree of dispersion within a feature, as a large degree of dispersion means that the difference between samples is big enough to distinguish them. Usually, a threshold is set and the features with variance smaller than the threshold will be abandoned. The basic idea of the coefficient of correlation method is to find the correlation between the input features and the output fault categories. The bigger the correlation coefficient is, the feature is more useful to the classification. Another feature selection method is the information gain-based method, also known as the information entropy-based method. This kind of method is based on the information entropy theory, and the amount of information can be defined and quantitatively calculated. Therefore, the aim of feature selection is to find the features of the biggest information gain.

Z. Zheng et al.[114] [115] applied both the variance analysis method and coefficient correlation method to select the most useful features from presupposed features. L. Huang et al.[116] proposed a diagnosis algorithm based on the decision tree, and the detailed calculation method of information gain was introduced. With the rank of information gain for all features, the stack voltage, inlet stack air pressure, inlet hydrogen temperature, and total power were chosen as features to carry out the diagnosis.

(2) Feature dimension reduction methods

Although the feature dimensions can be directly cut down by feature selection methods, it is usually risky because some useful information for the diagnosis may be thrown. To decrease the feature dimensions while making full use of the information of all the features, feature dimension reduction methods are proposed. The widely applied dimension reduction methods include PCA, KPCA, FDA, and KFDA.

1. PCA: It is the most used dimension reduction method in both the PEMFC diagnosis field and other fields. The basic idea is to map the features into a new orthogonal

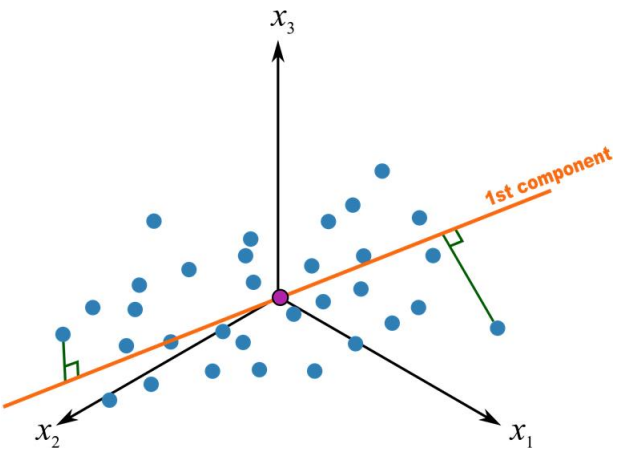


FIGURE 1.24: The principle of PCA method

feature space, where the new orthogonal axis is chosen according to the biggest variance direction [100]. A typical diagram of the principle of the PCA method is shown in figure 2.8. According to this method, the new features are ranked according to variance. Therefore, the new features in the last places can be dropped because they carry almost no information about the classification, and the dimensions can be reduced. It is a kind of unsupervised dimension reduction method because no classification labels are needed. There are two methods to decompose the matrix and obtain the new features, i.e. eigenvalue decomposition and singular value decomposition (SVD).

2. KPCA: The KPCA is quite similar to PCA, and the only difference is that a non-linear kernel function is applied to transform the data into a higher dimension linear space. In this space, the PCA method can be carried out to find the principal directions, and the useless features can be dropped.

3. FDA: FDA is a supervised dimension reduction method, so the labels are also applied to reduce the dimension of the data set. It is also known as linear discriminant analysis (LDA) [117]. The basic idea of the FDA is to map the features in new directions, and the objective is to ensure the variance within a class is the smallest while the variance between the classes is the biggest. The variance within a class can be represented by a within-class divergence matrix, while the variance between classes can be represented by an interclass divergence matrix. The optimization objective is to make the generalized Rayleigh quotient of the matrix the smallest. A diagram of the principle of the FDA is shown in figure 1.25. As the FDA is a supervised method, the dimension after reduction has to be smaller than the class number.

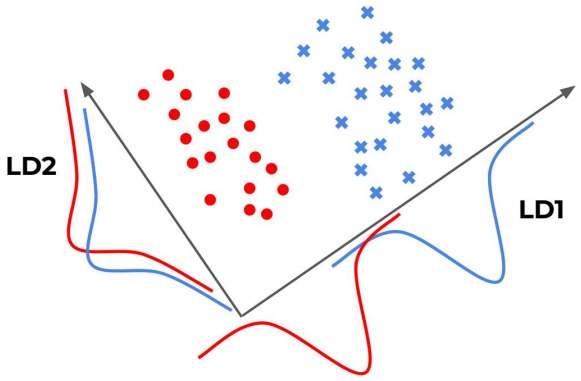


FIGURE 1.25: The principle of FDA

4. KFDA: Similar to the relationship between the PCA and KPCA method, the KFDA is also the combination of kernel function and FDA. By adding a kernel function, the non-linear data can be mapped to linear space, and the FDA can be applied. However, the calculation burden is much heavier than the normal FDA.

As all those methods have their own characteristics, they are suitable for different situations. They are usually tried and compared in the PEMFC diagnosis research.

1.4.2.3 Fault classification methods

The basic idea of fault classification methods is to find the relationship between the fault types and the features. To build this kind of relationship, some tools are widely used, including neural networks (NN), Bayesian networks, support vector machine (SVM), and k-nearest neighbour (KNN). Some other methods are also summarized.

However, those classification methods only dealt with the supervised classification problem, in which the fault samples have already been labelled by the fault types. To deal with non-supervised classification problems, cluster tools are widely applied. As the non-supervised diagnosis problem is not very common in the research of PEMFC, only the k-means clustering method and fuzzy cluster method are explained and summarized.

(1) Neural networks methods

The most widely used data-based diagnosis methods for PEMFC are all kinds of NN. As has been addressed above, the classification method is a kind of model where the

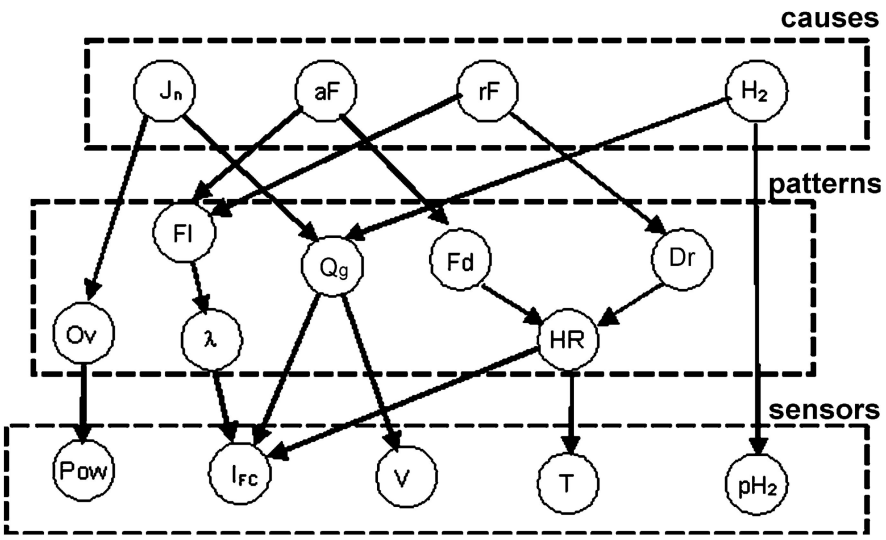


FIGURE 1.26: Typical structure of Bayesian network for PEMFC Diagnosis [124]

inputs are the features and the output is the fault type. NN is very suitable for modelling this kind of system, especially when the relationship between the inputs and output is complex.

J. Kim et al.[118] applied the hamming neural network to the output voltage pattern recognition, and the voltage pattern can represent the health state of PEMFC. S. Morando et al.[119] [120] proposed an ANN with reservoir computing optimization structure, and the air stoichiometry fault, cooling water fault, CO poisoning, and natural degradation was considered. C. Jeppesen et al.[121] also focused on the diagnosis of flooding and drying out in a PEMFC, but the EIS data was employed. M. Shao et al.[122] applied 4 back propagation neural networks (BPNN) sub-models to detect the health state of the subsystems, and the diagnosis of the whole system was achieved. X. Gu et al.[123] used a long short-term memory (LSTM) network model for PEMFC diagnosis, and the flooding fault was detected.

(2) Bayesian network

Another widely used diagnostic tool is the Bayesian network. It is developed based on Bayesian theory, which is a statistical tool to analyse the probability between events. A Bayesian network is a directed acyclic graph, where the nodes represent the random variables, and the directed lines between the nodes represent the conditional probability. The structure of a typical Bayesian network can be shown in figure 1.26 [124].

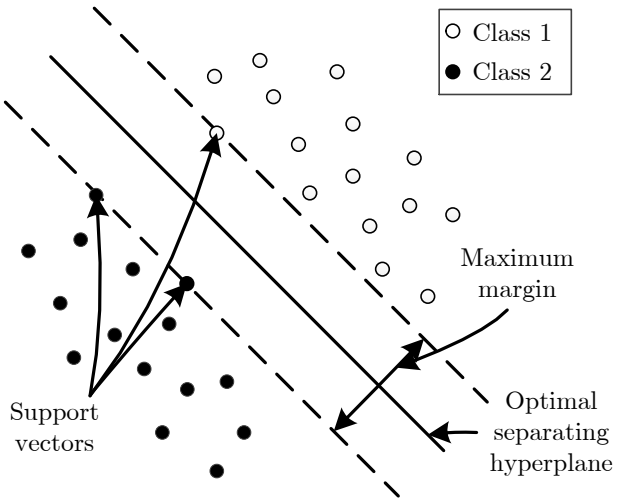


FIGURE 1.27: The basic mechanism of SVM

L. Riascos et al.[125] [124] constructed a Bayesian network structure by Bayesian-Score and Markov Chain Monte-Carlo algorithm, and they were combined to represent the final probability. S. Wasterlain et al.[126] used the EIS data to decide the water condition in a PEMFC. The real and imaginary parts of 6 points detected under particular frequencies were set as inputs to a naive Bayes classifier. L. Mao et al.[127] also applied the Bayesian network to the PEMFC fault diagnosis. The failure modes and the fault of the sensor were analysed, and the details of the Bayesian network were given.

(3) SVM classification

A widely used classification method is SVM. In SVM, the hinge loss is employed to calculate the empirical risk and the structural risk, so it is a robust classifier [128]. Normally, SVM belongs to the generalized linear classifier, but the kernel method can also be added to the SVM method to deal with non-linear classification. The basic idea of SVM is to find the hyperplane with the maximum margin. The mechanism of SVM is shown in figure 1.27.

As the SVM method is very accurate and the dimensions of features will not greatly affect the computational efficiency, it is widely applied to classify different fault types in PEMFC. Z. Li et al.[129] [130] employed the voltages of individual cells in a stack as features to indicate the fault types, and the SVM was used as a classifier. Also, they

developed the spherical-shaped multiple-class support vector machine (SSM-SVM), which can also detect the potential novel failure mode [131] [132].They also tried to diagnose based on magnetic measurement. The magnetic field can be measured by non-intrusive tools, and used for the current density distribution estimation [133]. Further, they developed a highly compacted embedded system-in-package hardware, to monitor the cell voltages and carry out the proposed diagnosis method [134]. J. Yu et al.[128] developed the traditional SVM method by employing the cuckoo search method for the optimization of the SVM model, rather than grid research in a traditional method. L. Mao et al.[135] focused on the feature selection for diagnosis, and the SVM was applied to compare the effect of different combinations of features. At the same time, I. Lim et al.[136] developed a diagnosis method for the thermal management system under various current density.

(4) KNN classification

Another widely applied classification method in PEMFC diagnosis is the KNN algorithm. It is one of the most simple algorithms for classification, and the basic idea is to represent the category of a sample by the majority categories of several nearest neighbours Although the algorithm is simple, the computation burden is relatively heavy because the distance between every two samples should be calculated.

R. Onanena et al.[137] carried out a diagnosis according to the EIS data and KNN. The confusion matrices were obtained, and both two classification methods can obtain high accuracy. D. Benouioua et al.[107] applied the KNN method to the fault classification based on voltage singularity measurement. The classification tools included both SVM and KNN. Z. Li et al.[129] also used the individual cell voltages as features to compare 4 different feature reduction methods and 3 classification methods.

(5) Clustering tools

The clustering algorithms are put forward to deal with the unsupervised classification problems, i.e. find similar samples and distinguish them from other groups when no labels for the samples. The most widely applied clustering tools in PEMFC diagnosis include the k-means clustering method and fuzzy cluster methods.

K-means clustering algorithm is a very simple algorithm for unsupervised learning. The basic idea is to find the centres that can minimize the total distance from cluster

members to the centres [138]. Firstly, the centres are given randomly, then the distances from sample points to the centres are calculated and compared, and the centres will be updated according to the samples until the minimum is reached.

J. Liu et al.[138] focused on the diagnosis of a PEMFC used for the tramway. Six faults were researched, and the discrete hidden Markov model fault diagnosis strategy based on the K-means method was proposed for the first time. W. Pan et al.[139] used the k-means clustering method to compare the effects of various features for the PEMFC diagnosis. Z. Liu et al.[140] proposed a novel method to extract features from voltage data. The FDA was used as the feature reduction tool and the k-means clustering method was applied for classification.

Another important clustering algorithm is the fuzzy clustering method. The fuzzy clustering method gives a degree of membership for each sample, rather than the binary partition rule in the k-means clustering algorithm. Therefore, the possibility of a sample belonging to a cluster is not 1 or 0, but a possibility between 1 to 0. With this fuzzy process, the fuzzy clustering algorithm is more suitable for problems where the boundaries between the clusters are not so clear.

M. Buchholz et al.[141] used the polarization curve as diagnosis data, and the fuzzy clustering algorithm was applied to divide the samples with different faults. D. Hissel et al.[142] proposed a method to diagnose the operation state of PEMFC for transportation applications. With the fuzzy clustering algorithm, three clusters can be identified. Z. Zheng et al.[143] applied the EIS data under different currents and different air stoichiometry to the diagnosis of a purge fuel cell. The fuzzy c-means clustering method was applied to classification, and the conditions were divided into oxygen starvation/flooding conditions and normal status. They also applied the same method to the experiments of another fuel cell [114]. Further, they developed another online diagnosis method called double-fuzzy methodology. The double-fuzzy method was composed of a fuzzy clustering algorithm and a fuzzy logic algorithm [115]. R. Petrone et al.[144] also applied the fuzzy clustering to another fuel cell, and 6 clusters with different water content states were obtained. F. Han et al.[145] developed another diagnosis method by connecting the fuzzy clustering, artificial bee colony algorithm, and SVM. The fuzzy clustering algorithm was employed to optimize the original dataset, while the SVM was used to diagnose the faults.

(6) Other diagnosis methods

Except for the widely used diagnosis methods listed above, several other methods were also tried in the PEMFC diagnosis field. The Gaussian mixture model is based on the Gaussian distribution function, and it can divide the samples according to the possible distribution. It was applied to the diagnosis classification and compared with the KNN and SVM methods [129].

J. Liu et al.[146] applied the decision-making tree classifier to diagnose membrane drying out and hydrogen leakage failure based on the operating parameters. The decision-making tree classifier is very similar to the human decision process, and a lot of if-else relationships are put forward to obtain the final decision; thus it is relatively simple. Also, a relevance vector machine (RVM) was applied to the diagnosis in reference [13]. RVM is quite a new algorithm which is put forward in the year 2000. It is similar to SVM but it is trained according to automatic relevance determination, and the RVM is more computationally efficient than SVM. R. Lin et al.[147] applied the random forest (RF) algorithm to PEMFC diagnosis. It was based on the decision-making tree classifier, and several trees were included in the RF to make decisions. As the decisions of different trees may be different, the final decision was voted by all decisions.

N. Zhou et al.[148] introduced the extreme gradient boosting (XGBoost) algorithm to the fault classification field. The classification accuracy and precision of the XGBoost algorithm were better than convolutional neural network (CNN), LSTM, CNN-SVM, CNN-LSTM, and the calculation time was much shorter. H. Dong et al.[149] also applied the XGBoost algorithm, and it had higher diagnostic accuracy than SVM and KNN. An extreme learning machine (ELM) was also applied to the diagnosis of PEMFC. J. Liu et al.[150] applied both the kernel ELM algorithm based on cell voltages and the on-line sequential ELM algorithm based on operating non-electrical parameters.

The existing research about data-based diagnosis methods is summarized as table 1.8. A lot of research only used the operating parameters as features and no feature extraction or feature reduction methods were applied. On the contrary, for some diagnosis tools such as polarization curve, EIS, and current interruption, usually special feature extraction tools were applied. The most popular classifiers include ANN, SVM, and KNN, as they are relatively simple. Also, the different classification methods

and feature reduction methods were compared in a lot of research. However, there is no diagnosis method that is absolutely good for all problems or for all features. Therefore, different classifiers and different feature reduction methods are usually tried and compared to find the best solution.

TABLE 1.8: PEMFC diagnosis researches by data-based methods

Methods	features	feature extraction	feature selection/ reduction	diagnosed faults	remarks
ANN	voltage response for current interruption [118];	physical analysis;	none;	degradation degree;	voltage loss can be an indicator of degradation degree;
	stack current and voltage data [119] [120];	FFT and STFT;	none;	fault of air stoichiometry, cooling water fault, CO poisoning, and natural degradation;	ANN with reservoir computing (RC) structure was used, and the algorithm can be applied on-line;
	two angles under high frequency and low frequency of EIS data[121];	physical analysis;	none;	low and high cathode stoichiometry, high CO concentration in anode gas, high anode methanol vapour concentrations, and low anode stoichiometry;	the other fault conditions can be detected accurately except for high methanol vapour concentration fault;
	voltage, current, temperature, flow rate, the pressure of air and hydrogen [122];	none;	none;	stack cooling system fault, increase of fuel crossover fault, air delivery system fault, and hydrogen delivery system fault;	4 NN sub-model was applied to each fault diagnosis;
	operating parameters of the whole system [123];	none;	none;	flooding fault;	LSTM network was applied to deal with time-series data, and the fault can be pre-diagnosed;
Bayesian network	power, current, voltage, temperature, pressure [125] [124];	none;	none;	fault in the air fan, refrigeration system, growth of current crossover, and low H_2 pressure;	diagnosis decision can be made at the early stage to avoid permanent damage;

Continued

Continued

Methods	features	feature extraction	feature selection/ reduction	diagnosed faults	remarks
	6 points of EIS data [126];	none;	none;	water management faults ;	classification accuracy is 91% ;
	8 operating parameters [127];	none;	none;	flooding, incorrect bipolar plate torque excess heat, contamination from impurity, Pt loss, migration, and agglomeration;	the probability of fault causes can be ranked;
SVM	voltages of individual cells in a stack [129] [130] [131];	none;	PCA, FDA, KPCA, and KFDA;	low pressure and high-pressure fault, flooding and drying out fault, low air stoichiometry fault;	the diagnosis results by different features reducing methods and classifiers were compared, and the FDA with SVM method was the best;
	16 operating parameters [128];	none;	none;	leakage of hydrogen, low pressure of pumps, temperature, voltage over-range, low pressure of input air;	the cuckoo search algorithm was added to a traditional SVM method, and it was proved better;
	individual cell voltage in a stack [132];	Shapelet transform;	none;	low and high-pressure fault, drying out fault, low air stoichiometry fault;	sphere-shaped multi-class (SSM) SVM was applied to detect new faults;
	magnetic fields data [133];	none;	FDA;	low and high air stoichiometry fault, drying out and flooding fault, low and high-pressure fault, low hydrogen stoichiometry fault;	the diagnosis based on magnitude field was better than that based on individual cell voltage;

Continued

Continued

Methods	features	feature extraction	feature selection/ reduction	diagnosed faults	remarks
	statistical features of stack voltage [135];	wavelet transformation and wavelet packet transformation;	KPCA;	drying out and flooding fault;	the diagnosis by different features were compared;
	temperature, pressure, and fan control signals [136];	none;	none;	pump degradation, radiator fouling, tube clogging, fan disabled, pump disabled faults;	thermal management faults under different current densities were detected;
KNN	real and imaginary parts of 6 points of EIS [137];	physical analysis;	correlation-based feature selection;	flooding and drying out;	the FDA was applied as a classifier and compared with KNN;
	voltage singularity [107];	wavelet transformation;	mutual information difference (MID) and mutual information quotient (MIQ);	cathode stoichiometry, anode stoichiometry, pressure, temperature, CO impurity faults;	SVM was compared with KNN, and the combination of MIQ and KNN was the best;
	voltage frequency domain features [103];	Fourier transform (FFT) and total harmonic distortion (THD);	none;	drying out and flooding faults;	K-means method and KNN were compared;
K-means clustering	operating parameters [138];	none;	none;	inlet temperature, low pump pressure, outlet temperature, voltage over-range, low air pressure and hydrogen leakage;	the KNN was used to identify the unqualified points;

Continued

Continued

Methods	features	feature extraction	feature selection/ reduction	diagnosed faults	remarks
	statistical features of voltage [139];	statistical analysis;	none;	flooding and drying out;	the dynamic fault formation process can be indicated;
	2D features of voltage [140];	image features extraction tools;	FDA;	flooding and drying out;	the diagnosis based on 2D features is better than 1D features;
fuzzy clustering	polarization curve [141];	physical analysis;	none;	high and low pressure, high and low stoichiometry, high and low relative humidity, temperature fault;	4 faults can be isolated and other 3 faults can be detected;
	two features of EIS data [142];	none	none;	flooding and drying out fault;	FC stack behaviour was under evolution;
	features of EIS data [143] [114] [115] [144];	physical analysis	variance and correlative coefficient analysis;	flooding and drying out fault;	fuzzy clustering and fuzzy logic methods were combined to obtain the diagnosis rule;
decision-making tree	20 operating parameters [146];	none;	none;	membrane drying and hydrogen leakage;	the accuracy was 98.5% ;
RVM	individual cell voltages [13];	none;	FDA;	flooding;	the algorithm was on-line adaptive;
RF	20 operating parameters [147];	none;	PCA;	faulty and normal cases;	RF algorithm was compared with Adaboost, KNN, and ANN methods;

Continued

Continued

Methods	features	feature extraction	feature selection/ reduction	diagnosed faults	remarks
XGBoost	operating parameters [149];	none;	PCA;	membrane drying, hydrogen leakage and unknown fault;	it was compared with SVM and KNN;
	operating parameters of vehicle [148];	none;	none;	3 fault levels;	it was compared with CNN, LSTM, CNN-SVM, CNN-LSTM;
ELM	cell voltages and operating parameters [150];	none;	none;	4 degrees of high air stoichiometry fault;	two different ELM model were applied simultaneously and combined, and it was compared with SVM and BPNN;

Also, there exist some research work where both model-based algorithms and data-based algorithms are combined together. For example, L. Mao et al.[151] proposed a sensor selection method based on an ANFIS model to calculate the voltage output with different combinations of sensors. With this model, the SVM was carried out to compare the diagnostic accuracy of selected sensors and all sensors. As model-based and data-based methods have different advantages, it is worth trying to connect them to have better diagnosis performance.

1.5 Challenges and solutions for PEMFC health management

1.5.1 Challenges of PEMFC diagnosis and prognosis

According to the analysis above, the diagnostic electrochemical tools include polarization curve, EIS, CI, CV, LSV, and EN. Among them, CV and LSV can give information about the catalyst and membrane flow current respectively, but they are only suitable for laboratory applications because the particular conditions required for the gases cannot be met in practice. Polarization curve, EIS, CI, and EN can be applied to in situ diagnosis. The polarization curve and EIS are the most widely used now. CI method is another easier tool for the dynamic feature assessment than EIS, but it is not widely applied yet [152]. The EN is also applied and is getting more and more attention.

For the diagnosis methodologies, both model-based and data-based diagnostic methods have been applied. However, according to the review above, several challenges remaining to be solved in PEMFC health management can be concluded as follows.

1. As it is difficult to describe the PEMFC in a universal mathematical form, a major trend of PEMFC diagnosis is to build black-box models or directly obtain the fault types of classification algorithms, i.e. by data-based methods. However, even though the relationship between the signal and fault can be related to data-based diagnostic methods, the inner physical mechanisms are ambiguous, so the generality is questionable.

2. For the data used in the diagnosis, a great restriction is that they should be able to be obtained when the PEMFC is operating or can be easily obtained ex situ. However, former research did not emphasize the real-time ability of the diagnostic data.

3. More feature extraction methods should be discovered, and it is especially important to connect the feature extraction methods with the data measurement system. To make sure that the feature really represents the fault characters, they should be related to the basic physical mechanisms from the beginning.

4. For the classification methods of data-driven diagnosis, according to the 'no free lunch rule', there is no method that is suitable for all diagnostic problems with all kinds of data. Therefore, finding the proper algorithm for different data is another challenge.

5. As the application of PEMFC diagnosis is critical, one important direction is the on-line diagnosis implementation. Even though a lot of researchers declare that their algorithm can be applied on-line, the methods are not tried on industry-level systems yet.

1.5.2 Study objectives and methods

According to the analysis above, the diagnosis technology of PEMFC is still an open area. The main challenges include the generality of the algorithm, real time obtaining data, interpretability of features, the matchup between problem and methods, industrial practicability, etc.

Therefore, to contribute to the PEMFC health management and focus on the practicability challenge, several diagnostic and prognostic algorithms are proposed and validated in this thesis.

First of all, the real time obtaining ability of diagnostic data is considered. Two types of data are used in the study, i.e. the stack voltage data and EIS data. As was analysed above, both of them are electrical signals, so they can be detected in situ without interrupting the normal operation of the PEMFC power system. It is quite important as the signals basically decide the practicability of the method [3].

Secondly, to overcome the problem of generality for data-based diagnostic methods, hybrid methods are applied to build the relationship between the fault and signals. In the first phase, the models are built based on physical mechanisms to extract the diagnostic features. Therefore, the features are interpretable and they can be generalized to other cases. Also, the data-driven methods are also applied in this thesis. As artificial intelligence methods are quickly developing, advanced classification methods are used to distinguish the different fault conditions. By combining the feature extraction based on the model and the classification algorithms, this thesis uses the hybrid method to deal with the generality and feature interpretability problems.

Thirdly, practicability is emphasized in this thesis. As the real-time application is quite important, a DSP system is applied to validate the algorithms based on EIS. Therefore, it is possible to apply the proposed algorithms in a wide range of industrial applications.

Further, not only the diagnosis should be carried out quickly, but also a quick prognosis is important to save decision time and resolve sources [20]. Therefore, a quick and accurate prognosis method is proposed based on FDKF, and the method is validated under both static and dynamic conditions.

1.6 Conclusions

In this chapter, the basic knowledge about PEMFC diagnosis and prognosis are introduced, so that to give a background of the thesis. First of all, the PEMFC structure, PEMFC faults, application situation, and the concept of diagnosis and prognosis are explained. Also, the existing experimental diagnostic tools and algorithms are clearly reviewed. Only several electrical tools have the potential to be applied to the PEMFC system in real time. The diagnosis algorithms can be divided into model-based and data-based diagnosis, and both of them have their advantages.

Based on the literature research and analysis, the challenges for real-time diagnosis are proposed. Also, the solutions to face those challenges are explained. The main challenges include the generality and industrial practicability of the diagnosis and prognosis methods, and this thesis focused on the implementation and practical application of the methods. Therefore, the whole background and the problems to be solved are given in this chapter.

In the rest of the thesis, the proposed novel diagnosis and prognosis methods are presented one by one, including three diagnosis methods and one prognosis research. In the next chapter, a diagnostic method based on the stack voltage fluctuation and the autoregressive model is presented, in which the electrical noise fluctuation mode is described by autoregressive model coefficients.

Chapter 2

Diagnosis based on voltage fluctuation

2.1 Introduction

To achieve a real-time diagnosis, more and more researchers try to apply the variables that can be easily obtained during operation, including the voltage, gas pressure, etc. [153] [151]. The output voltage was applied as a diagnostic tool in several studies.

Z. Li et al.[131] applied the voltages of each cell as features, and the spatial inhomogeneity in a PEMFC cell can be represented. SVM was applied to classify the different fault conditions in their search. More and more new classification tools have also been applied to decide the faults at the cell voltages, and the result has been compared with other widely used classifiers [154]. However, the features are also important for diagnosis. Therefore, in addition to proposing new classification methods, finding efficient feature extraction methods is another important task for PEMFC diagnosis [155].

In these studies [118] [150], the applied features are the voltages of each cell, i.e. $\{V_{cell_1}, V_{cell_2}, ...V_{cell_n}\}$. From these features, the spatial distribution characteristic of voltage can be revealed because the positions of the cells will greatly affect the voltage output of an individual cell. However, all the cell voltages should be monitored in those

methods, which is both costly and inconvenient when there are hundreds of cells in industrial applications.

On the contrary, except for the voltage spatial distribution, another interesting and important information is the stack voltage fluctuations over time, which is also called electrochemical noise [63]. To apply the voltage fluctuation as a diagnosis tool, only the stack voltage need to be considered, and the time dependency is emphasized. Therefore, the data applied in the diagnosis is the stack voltage at different moments, i.e. $\{V_{t_1}, V_{t_2}, ...V_{t_k}\}$. For these time-series data, the voltage data sampled under high frequency can be separated into voltage profiles by windows that contain certain data, so the diagnosis can be carried out for each profile within a certain period.

The most common fluctuation observed in electrochemical devices is the ”1/f noise”, because the noise intensity is inversely proportional to frequency. The PEMFC voltage fluctuation can be caused by mixed complex mechanisms, such as the water content fluctuation, resistance impurity, electron transmission pulse, etc [156]. It is still an open area to find out the dominant mechanisms of voltage fluctuations in PEMFC system, and several researchers have discussed it [59].

Even though the mechanism of the fluctuation is not totally clear yet, it is sure that diagnostic information exists in the fluctuations, and the interrelation between operating fault and fluctuation pattern was proved by both time domain and frequency domain methods. The voltage fluctuation has been analysed by different methods, such as wavelet transformation [157], statistical features [58], power spectral density [59], and high-order moments [158], etc.

To achieve quantitative diagnosis based on voltage time series data, D. Benouioua et al.[107] [159] employed the singularity analysis of the voltage time-series data as features to distinguish the operating conditions. The singularity spectrum is the multi-fractal spectrum calculated by wavelet transform-based multi fractal formalism, and the KNN was applied to identify the fault conditions using features extracted from the spectra. Also, the high-frequency and low-frequency parts of the voltage singularity spectrum were compared with EIS, which explored the possibility to link EIS and singularity spectra [157]. However, it needs about 4.5 minutes to obtain a voltage profile. The mathematical complexity of the method is relatively high, and the accuracy of the fault classification can still be improved. Rather than implementing

singularity analysis, we propose to look for other effective features that can be more easily identified, and in a shorter time. According to the reference [160] [161], as the autoregressive model (AR model) can represent the recurrence interval of fluctuations, it is a useful tool to describe the fluctuation patterns.

In this chapter, a novel data-driven method is proposed to achieve a quick and accurate diagnosis based on the stack voltage fluctuation. In the proposed method, the voltage fluctuation pattern is extracted by the AR model, then the AR model coefficients can be applied as features to classify the different fault conditions. The main advantage of this method is that only stack voltage data are needed, and the voltage profile can be obtained every 1 second, so it is quicker and more practical. The diagnostic method is experimentally demonstrated under extensive fault operating conditions. Nine single fault conditions and 8 multi-fault conditions are researched, and those conditions relate to the fault of the cathode stoichiometry (FSC), anode stoichiometry (FSA), cooling circuit temperature (T), and the relative humidity level (RH). The diagnosis is then carried out by several classifiers (ANN, ELM, KNN, and SVM) under different hyper-parameters, and both the accuracy and computational times are compared. For the first time, the quantitative effect of voltage sampling frequency and sample window length on the diagnosis accuracy is studied. It proves that a higher sampling frequency or longer data profile is beneficial to diagnosis accuracy.

This chapter is organized as follows: in Section 2.2, the diagnosis method based on the AR model and classification is addressed. In Section 2.3, the experiment conditions and related data are explained. Then the diagnostic result is obtained and analysed in Section 2.4, and it is also compared with other research works. Some discussions about the effects of sampling frequency and window length are given in Section 2.5. Finally, the main conclusions are summarized in Section 2.6.

2.2 Diagnosis based on AR model

2.2.1 PEMFC diagnosis process based on AR model

The overall PEMFC diagnosis process based on output voltage data and AR model can be concluded in figure 2.1. The diagnosis process can be divided into the offline period

(training) and online period (testing), and it can be described as follows.

1. Collect experiment data from different PEMFC conditions for training. In this research, the voltage data are measured under 3000 Hz frequency on a PEMFC stack, and both single-fault and multi-fault conditions are researched.
2. For the collected voltage data, intercept voltage profiles by windows. There are 140 profiles (70% of total profiles) for each condition and 3000 data in each profile.
3. The coefficients of the AR model can be calculated for each profile and the features of the diagnosis can be obtained. The exact feature extraction method is explained in algorithm 2.1.
4. The classifiers can be trained based on the features. Different classifiers (ANN, ELM, KNN, SVM) are compared, and the hyper-parameters of the classifiers are also researched.
5. For the on-line period, firstly the detected stack voltage data can be intercepted as profiles by the window, then the features can be obtained and applied to the trained classifiers. By comparing the diagnosis results with the real conditions, the diagnosis accuracy of different methods can be evaluated and compared.

2.2.2 Feature extraction by AR model

2.2.2.1 AR model principle

The relationship between data fluctuation pattern and AR model coefficients has been analysed and confirmed [160] [161]. The AR model can describe the development of time series data [162], and it was applied to system identification, future trend forecasting [163], system control [164] [165], etc. It was also applied to modelling and predicting PEMFC voltage [163], but never to PEMFC diagnosis.

In an AR model, the output data at time instant n can be represented by the linear combination of p previous data, which is shown as equation 2.1. The biggest advantage of the AR model is that no external inputs are needed, thus the model can be built only

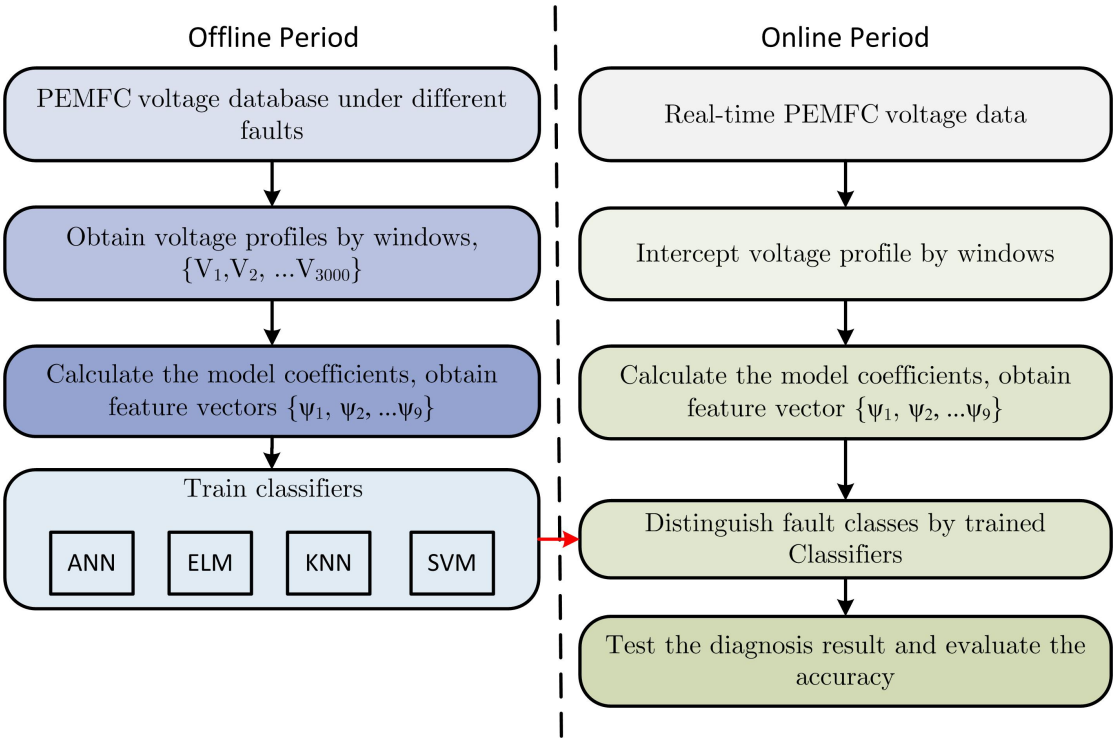


FIGURE 2.1: The diagnosis processes based on voltage data and AR model

based on the time-series data itself.

$$y_n = \sum_{i=1}^{p} y_{n-i} \times \psi_i + e_n \tag{2.1}$$

Where y is the time-series data, i.e. PEMFC voltage in this study; n is the time index; p is the order of the AR model, which should be artificially specified; ψ_i is the coefficient for i_{th} lag data. e_n is a white noise whose mean is 0.

Coefficients means the degree that the current output is decided by former data. As the different fluctuating patterns can lead to an AR model with different coefficients, the coefficients can be directly applied as features for pattern identification, i.e. diagnosis in our case.

2.2.2.2 Determination of model order

To have an accurate AR model, the model order should correspond to the characteristic of the data, therefore the determination of AR model order is one of the most important tasks. Some features may not be caught if the model order is too small, while there is

a risk of over-fitting if the model order is too high. To handle this problem, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) have been proposed to search for a balance between the accuracy and model complexity [166]. The AIC and BIC can be calculated by equations 2.2 and 2.3, respectively.

$$AIC = 2k - 2\ln(L) \tag{2.2}$$

$$BIC = k\ln(m) - 2\ln(L) \tag{2.3}$$

Where k is the number of model parameters, i.e. the order of the AR model in this research; L is the likelihood function, which can reflect the accuracy of the fitting; m is the number of the sample data.

The difference between AIC and BIC is that the BIC also considers the effect of a data number, and it tends to give a model with fewer parameters. In both criteria, the complexity term and the accuracy term have opposite signs, so we can decide the order by choosing the model with the smallest AIC or BIC. Both AIC and BIC are calculated in this research, and they will be analysed in Section 2.4.

2.2.2.3 Calculation of model coefficients

When the order of the model is determined, the coefficients of every lag in the AR model can be obtained by solving the Yule-Walker equation, as shown by equation 2.4.

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ . \\ . \\ . \\ \gamma_p \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_{-1} & \cdots & \gamma_{1-p} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{2-p} \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ . \\ . \\ . \\ \psi_p \end{bmatrix} \tag{2.4}$$

$$\gamma_k = E((y_i - \mu)(y_{i-k} - \mu)) \tag{2.5}$$

Where the γ_k is the expectation of the auto-covariance for k order lags; μ is the expectation of the series data. There are a lot of methods to solve the Yule-Walker equation, such as the least square method, covariance method, and Burg method [167]. In this research, the Burg method is applied, as it needs no assumption about the out-of-range data.

2.2.2.4 Feature vector

By the identification of the AR model, the coefficients of every lag in equation 2.1 can be set as the feature vector of the voltage profile. The feature vector of the i_{th} profile can be noted as F_i with dimension p, which can be given as equation 2.6.

$$F_i = [\psi_1 \ \psi_2 \ \psi_3 \ ..... \ \psi_p] \tag{2.6}$$

The total process to extract features from voltage fluctuation data can be summarized as algorithm 2.1. Therefore, the AR model coefficients can be calculated and set as diagnostic features.

Algorithm 2.1: feature extraction by AR model

```
Load time-series voltage data V;
intercept voltage profiles by windows with certain length;
for AR order =1 to 100 do
   for profile index=1 to total profile number do
      calculate AR model coefficients according to equation 2.4;
      calculate AIC and BIC according to the equation 2.2 and 2.3;
   end for
   compute average AIC and BIC;
end for
select AR order by comparing AIC and BIC of different AR orders;
take the AR model coefficients of selected order as features in equation 2.6;
```

2.2.3 Classification methods

With the characteristics identified, the objective of the diagnosis is to find the correspondence between the characteristics and the operating conditions. A lot of classification methods have been used in the diagnosis of PEMFC. Four widely used methods are applied and compared in this research work: KNN, ANN, ELM and SVM.

2.2.3.1 K-nearest neighbors method

KNN is one of the simplest methods for classification [168]. The principle of KNN is to find the closest points of the aim point and then classify it to the class that appears most times around it. The most important parameter that affects the classification accuracy is the number of points that are considered effective for the classification. If the number of points is too small, there is a risk of over-fitting, because the noise can greatly affect the result. But if the number of points is too big, there is also the risk of under-fitting, because some unimportant points will affect the classification. Therefore, the effect of different numbers of points are studied in this research, and they are compared with other methods. The detailed process of KNN is given as algorithm 2.2.

Algorithm 2.2: KNN classification

```
Load train samples, train sample classes, test samples, hyper-parameter k;
for test sample index=1 to total test sample number do
   for train sample index =1 to total train sample number do
      calculate distance between the train sample and test sample;
   end for
   rank the training samples by distance from smallest to largest;
   take k nearest samples, rank the number of each class;
   return the class that appears most;
end for
```

2.2.3.2 Artificial neural network

The second method applied in this research is the ANN method [84]. In the ANN, there is an input layer, an output layer, and several hidden layers. In each layer, there are several nodes, which are called neurons. The input of a neuron is the linear combination of all the inputs, as shown by equation 2.7. And the input is transferred by a certain function, such as the sigmoid function shown by equation 2.8. The parameters of the network are adjusted according to the training data, so that the output of this network is as close as possible to the real output.

$$z_i = w_{1,i} \cdot x_1 + w_{2,i} \cdot x_2 + ...w_{k,i} \cdot x_k \tag{2.7}$$

$$a_i = g(z_i) = \frac{1}{1 + e^{-z_i}} \tag{2.8}$$

The ANN method can approximate any continuous system in theory, as long as the number of layers is high enough. However, there is a risk of over-fitting if the structure is too complex, and the calculation will also be enormous. Therefore, the number of hidden layers is one of the most important hyper-parameters for the ANN structure. The ANN with different layer numbers is analysed and compared later. The process of ANN is given as algorithm 2.3.

Algorithm 2.3: ANN classification

```
Load train samples, train sample classes, test samples;
load hyper-parameter hidden layer number;
build ANN structure;
initialization ANN: randomly assign parameter values;
set precision and maximum learning times;
for learning time index =1 to maximum learning times do
  for train sample index =1 in total train sample number do
    choose one training sample and correspond class;
    calculate the input and output of every neuron;
    calculate the partial derivations of output layer neurons according to output error;
    update the connection weights of hidden layer neurons;
    update the connection weights of input layer neurons;
  end for
  calculate the sum of output errors;
  if error < set precision or learning times > maximum learning times
    break;
end for
for test sample index= 1 to total test sample number do
  calculate output according to trained model;
end for
```

2.2.3.3 Extreme learning machine

Another widely used classification method is the ELM method. ELM is a particular training method that can be applied to a single hidden layer neural network [150]. Compared with the traditional ANN training method, ELM can give accurate results with fewer calculations. In traditional training methods, the gradient descent methods are widely applied. However, a big disadvantage of this kind of method is the low efficiency. On the contrary, the ELM method can set the parameters in a random way,

so it can give results with fewer calculations. As there is only one hidden layer in the ELM method, the number of neurons in the hidden layer is a critical parameter that affects the classification accuracy, and the effects of neuron numbers are researched and compared. The process of ELM is given as algorithm 2.4.

Algorithm 2.4: ELM classification

```
Load train samples, train sample classes, test samples;
load hyper-parameter hidden neuron number;
build ELM structure;
initialize ELM hidden layer weights randomly;
calculate hide layer response matrix;
calculate output layer weight;
for test sample index= 1 to total test sample number do
   calculate output according to trained model;
end for
```

2.2.3.4 Support vector machine

The SVM is also a widely used classifier as it can give accurate results quickly [134]. The SVM method can give the maximum-margin hyperplane of the samples, and classify the new samples based on it. The aim of SVM is to find the optimal hyperplane with the maximum distance from the nearest samples. To solve linearly inseparable problems by SVM, the kernel functions can be applied to transfer them into linear separable cases. More details about the SVM method can be found in the reference [134]. As polynomial functions are widely used as kernel functions, the polynomial functions with different orders are applied and compared in this research. The process of SVM is given as algorithm 2.5.

Algorithm 2.5: SVM classification

```
Load train samples, train sample classes, test sample;
load hyper-parameter kernel function order;
calculate correlation value of the feature vectors according to kernel function;
calculate covariance matrix space;
calculate characteristic coefficients;
scale covariance matrix by characteristic coefficients;
calculate model parameters;
for test sample index= 1 to total test sample number do
   calculate output according to trained model;
end for
```

2.3 Experiment data for PEMFC diagnosis

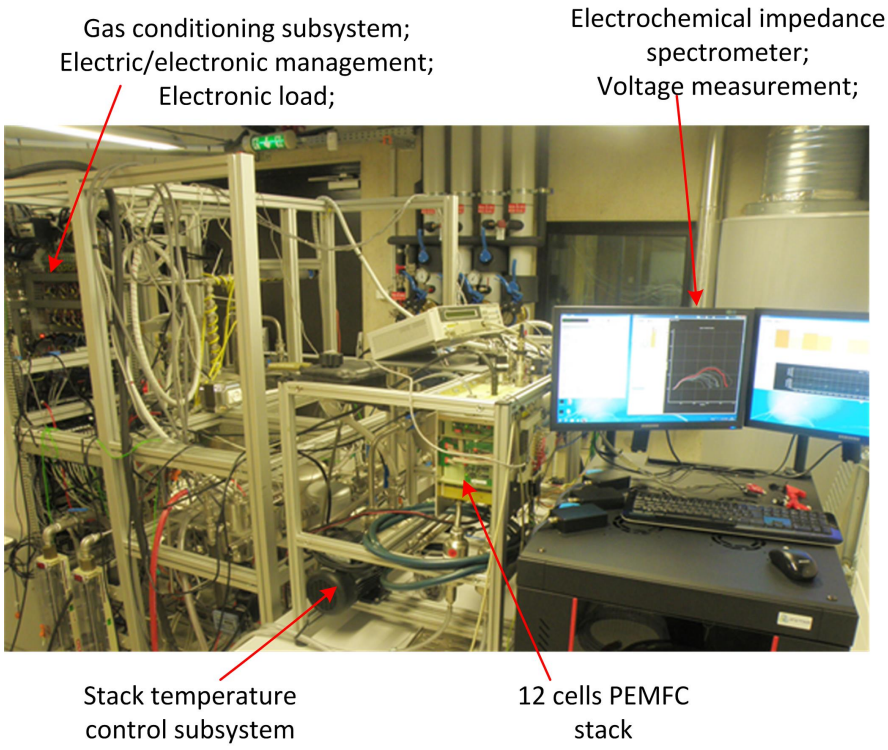
2.3.1 Experiments

To develop diagnosis strategies for a new PEMFC application, the experiments have been carried out using a test bench developed in FCLAB (Belfort, France). The experiments are a part of the "Decentralized energy production" project, directed by EFFICACITY, the French R&D Institute for urban energy transition. The experimental test bench is shown in figure 2.2(a). The monitoring and controls of the test bench parameters are done through National Instruments materials and dedicated software developed with Labview$^{\mathrm{TM}}$, and the control interface can be given in figure 2.2(b). More details about the test bench can be found in the reference [159].

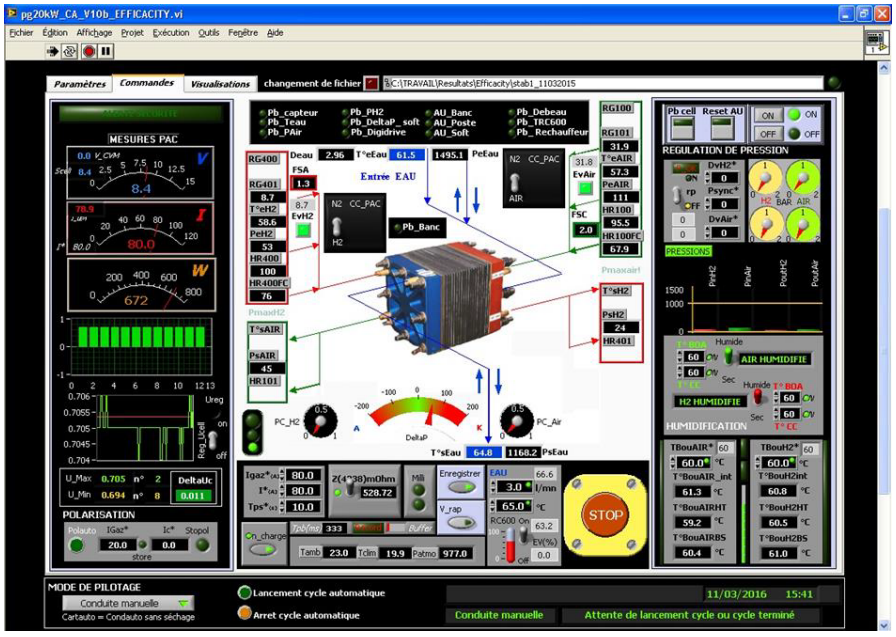
The investigated stack is with 12 cells, as shown in figure 2.3. The characteristics and nominal operating parameters of the stack are given in table 2.1. Air is supplied at the cathode, while a mixed fuel (75$\%$ of H_2 and 25$\%$ of CO_2) is supplied to the anode. The mixed fuel is set at this ratio to mimic a reformat because the PEMFC is designed with the ability to operate with a gas reforming system.

TABLE 2.1: Parameters of the investigated PEMFC stack and reference operating conditions

Parameter	Value
Number of cells	12
Electrode active surface	196 cm^2
Gas distributor plates	graphite
Fuel used during experiment	75% H_2+25% CO_2
Coolant flow (deionized water)	3 l/min
Anode stoichiometry (H_2 and CO_2 mix)	1.3
Cathode stoichiometry (air)	2
Anode inlet pressure	111 kPa
Air inlet pressure	106 kPa
Max. anode - cathode pressure gap	20 kPa
Temperature of the cooling circuit	70 °C
Anode relative humidity	50%
Cathode relative humidity	50%
Nominal current	80 A



(a)



(b)

FIGURE 2.2: (a) The test bench in FCLAB and (b) the human-machine interface of the bench

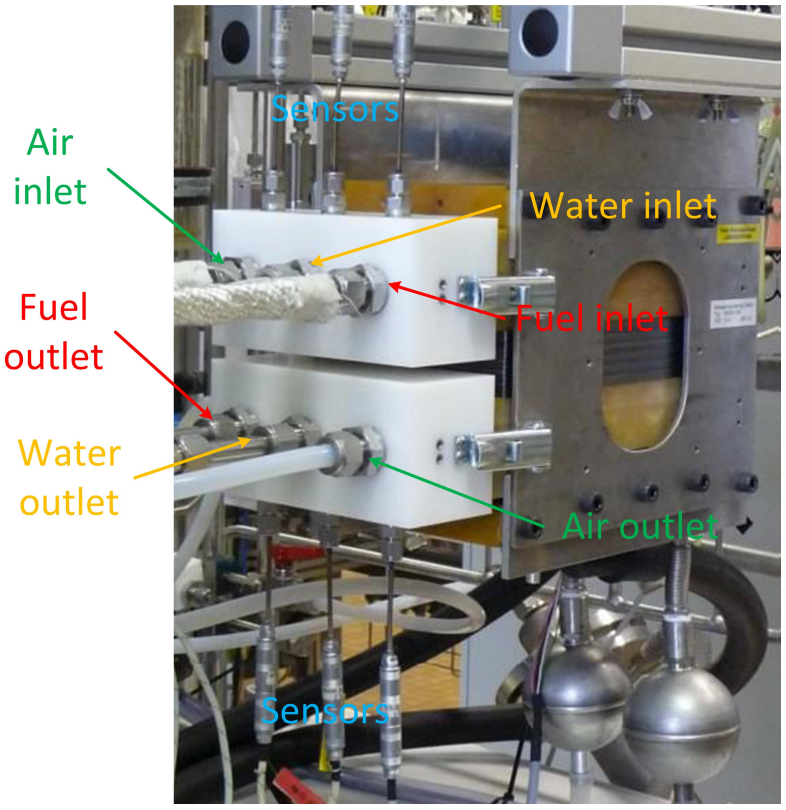


FIGURE 2.3: The investigated PEMFC stack

2.3.2 Available data

During the experiments, several operating parameters can be monitored and recorded. The available data from the experiments can be shown in figure 2.4. First of all, the system operating parameters are all monitored, including the gas pressure, gas flow rate, temperature of different components and materials, etc. Secondly, characterization tools are applied in the experiments, so that to indicate the health state of the PEMFC system. Both EIS and polarization curve are measured for each operating condition. In particular, the voltage data is recorded with high frequency, to study the fluctuation features of the stack voltage.

2.3.3 Single-fault conditions

To study the fault conditions of the PEMFC system in the experiments, different fault conditions have been reproduced by the adjustment of four operation parameters, i.e. cathode stoichiometry (SC), anode stoichiometry (SA), cooling circuit temperature (T),

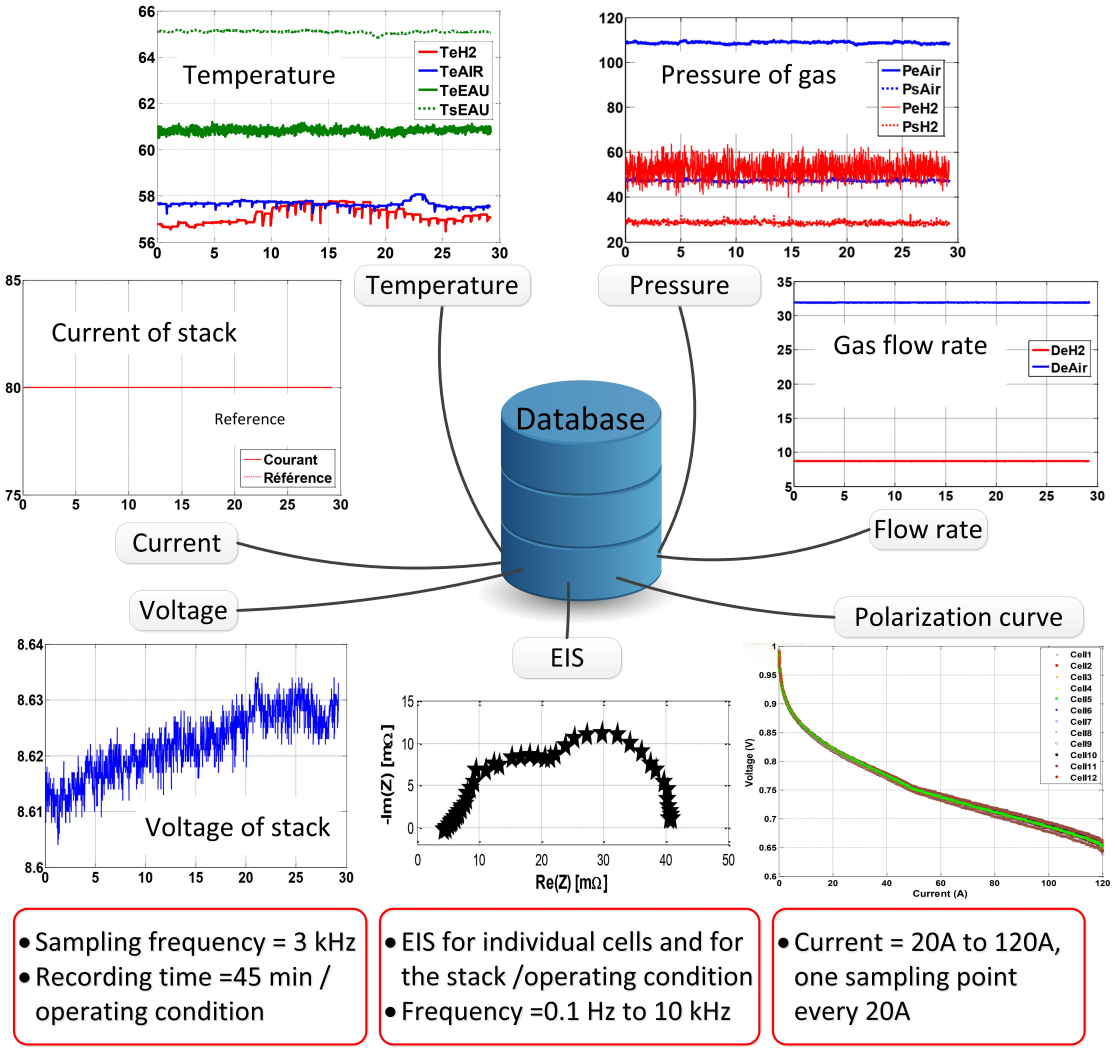


FIGURE 2.4: Available data of the experiments

and the relative humidity level (RH) (by controlling the temperature of the humidifier). Both single-fault conditions and multi-fault conditions have been tested.

Eight experiments under single fault conditions were carried out by setting the 4 parameters to higher or lower values than those corresponding to the reference conditions. The detailed operation parameters under different conditions are shown in table 2.2.

During the experiment, the current, the voltage of the stack and cells, the pressure of the gas inlet and outlet, the flow rates of gas and cooling water, the relative humidity rate, and temperature at different positions are measured. Because voltage fluctuations can reflect the state of the PEMFC, high-frequency voltage sensing can provide more information that is not available through low-frequency sensing.

TABLE 2.2: The operation parameters applied to single-fault conditions

Parameters	Ref	DFSCH	DFSCL	DFSAH	DFSAL	DTH	DTL	DRHH	DRHL
SC	2	2.6	1.6	2	2	2	2	2	2
SA	1.3	1.3	1.3	1.5	1.2	1.3	1.3	1.3	1.3
T (°C)	70	70	70	70	70	72	65	70	70
RH(%)	50	50	50	50	50	50	50	54	46

Ref: Reference/normal condition;
DFSCH: cathode flow fault, higher than normal;
DFSCL: cathode flow fault, lower than normal;
DFSAH: anode flow fault, higher than normal;
DFSAL: anode flow fault, lower than normal;
DTH: stack temperature fault, higher than normal;
DTL: stack temperature fault, lower than normal;
DRHH: relative humidity fault, higher than normal;
DRHL: relative humidity fault, lower than normal.

In this research, sliding windows are applied to obtain voltage profiles, and the length of the window is set to 3000 points so that one voltage profile can be obtained every second. Examples of the voltage profiles under different conditions are shown in figure 2.5. The different faults are related to the inner health state of PEMFC, such as the water content in the cell and/or different mass transport conditions. The voltage data are quite fluctuant, and the voltage levels are quite similar for different fault conditions. Therefore, it is practically impossible to assess the faults only by the voltage level, while analysing the fluctuation patterns of the time-series data is a promising method.

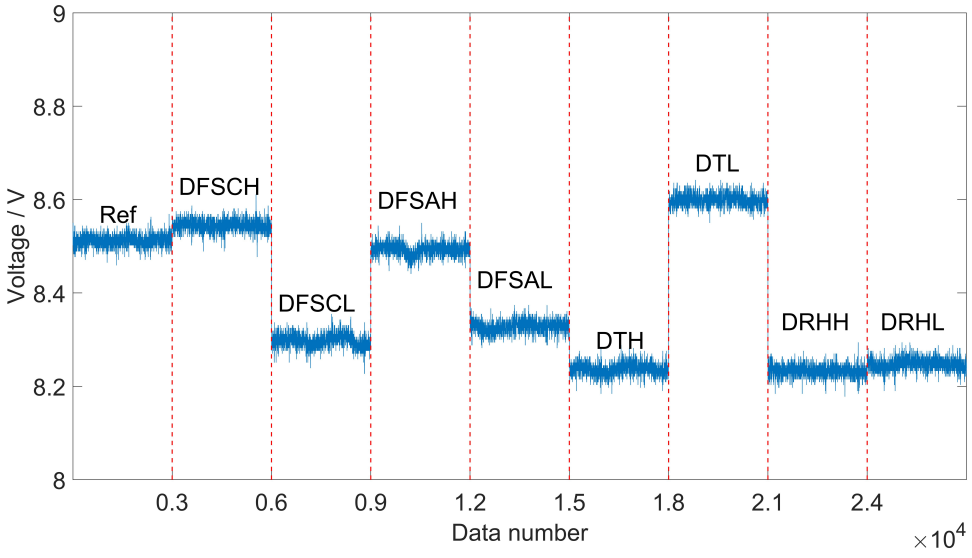


FIGURE 2.5: Examples of voltage samples under reference and different fault operating conditions

2.3.4 Multi-fault conditions

In addition to the single-fault conditions, some multi-fault conditions are also studied. The multi-fault conditions are obtained by changing the same operation parameters as for single-fault conditions, i.e. the SC, SA, T, and RH. However, two different faults are included in each multi-fault condition. 8 multi-fault conditions are considered and the detailed operation parameters are shown in table 2.3.

TABLE 2.3: The operating parameters under multi-fault conditions

Parameters	Ref	DFSC	DFSA	DT	DRH	DT+DFSC	DT+DFSA	DT+DRH
SC	2	$\frac{2.6}{1.6}$	2	2	2	$\frac{2.6}{1.6}$	2	2
SA	1.3	1.3	$\frac{1.5}{1.2}$	1.3	1.3	1.3	$\frac{1.5}{1.2}$	1.3
T (°C)	70	70	70	$\frac{72}{65}$	70	65	65	65
RH(%)	50	50	50	50	$\frac{54}{46}$	50	50	54

Ref: reference/normal condition;
DFSC: cathode flow fault, higher or lower than normal;
DFSA: anode flow fault, higher or lower than normal;
DT: stack temperature fault, higher or lower than normal;
DRH: relative humidity fault, higher or lower than normal;
DT+DFSC: cathode flow fault with lower temperature fault;
DT+DFSA: anode flow fault with lower temperature fault;
DT+DRH: relative humidity fault with lower temperature fault;

2.4 Results and analysis

In this section, the results obtained by the proposed diagnosis methods are presented, and different classifiers are compared and analysed. Both single-fault conditions and multi-fault conditions are researched, and 200 voltage profiles are applied for each condition. 70% of the data are applied as training data and 30% data are used to test the accuracy of the diagnosis.

2.4.1 Model order and coefficients distribution

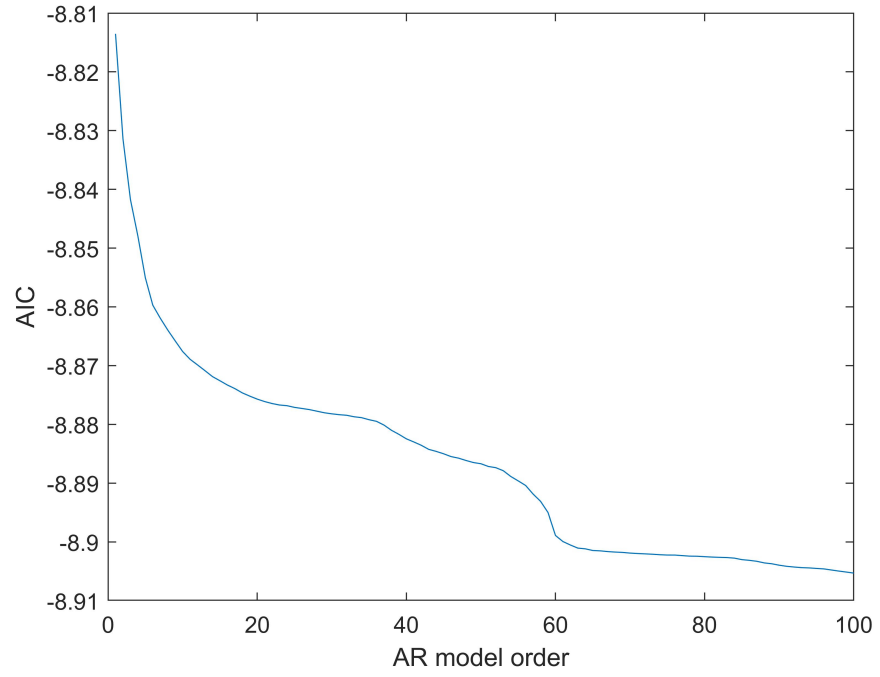
The AR model order can be decided by the AIC and BIC. The AIC and BIC under different orders are plotted in figure 2.6(a) and 2.6(b), respectively. The AIC does not reach the minimum until 100 orders, and the BIC has the smallest value when the order is 9. As a small model order will accelerate the diagnosis process, the AR model order can be chosen as 9 according to BIC in this study.

The coefficients of the model can be calculated by the Burg method. To see the distribution of the coefficients under different conditions, the box plot for the first-order and second-order coefficients under different conditions are shown in figure 2.7. Here, the condition numbers 1 through 9 mean the 9 single fault conditions given in Table 2.2. The box plot gives the median, quartiles, and upper and lower bounds. There are also some points that are out of bounds, which means they are abnormal. The distributions of the coefficients have different ranges under different conditions; as a result they are related to the fault conditions. Each coefficient can provide a part of information about the operating condition of the PEMFC; thus the conditions can be classified by the combination of the coefficients.

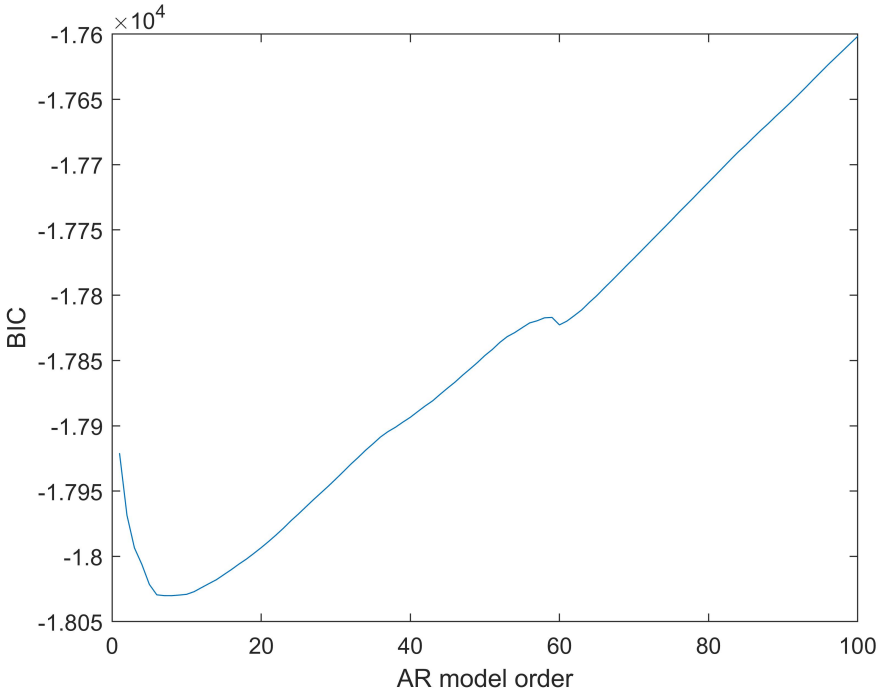
To show the spatial distribution of features visually, principal component analysis (PCA) is applied to find the most representative features, and the 2 dimensions and 3 dimensions representations of the main features of samples can be given in figure 2.8(a) and 2.8(b), respectively. The samples under different conditions are represented by different colours. There are some overlaps between different conditions when only 2 features are applied, but most of them can be separated when the third feature is applied. Therefore, the AR model coefficients are effective features that can represent different operating conditions.

2.4.2 Diagnostic accuracy for single-fault conditions

70% of the samples are set as train data, and both the SVM, KNN, ANN, and ELM methods with different hyper-parameters are researched and compared. As the hyper-parameters can decide the structure of the classification methods, they should be artificially set rather than tuned automatically. 30% data is set as a test subset, i.e. 60 voltage profiles are tested for each operating condition.



(a)



(b)

FIGURE 2.6: (a) The AIC and (b) the BIC for different AR model orders

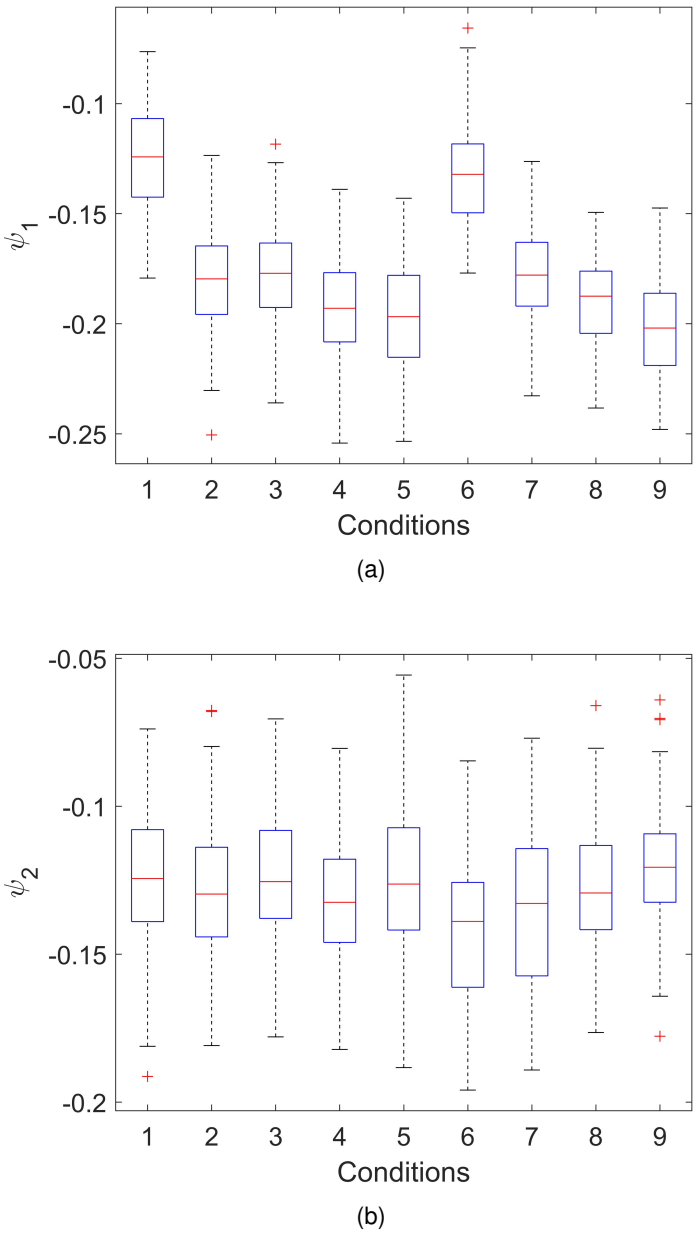
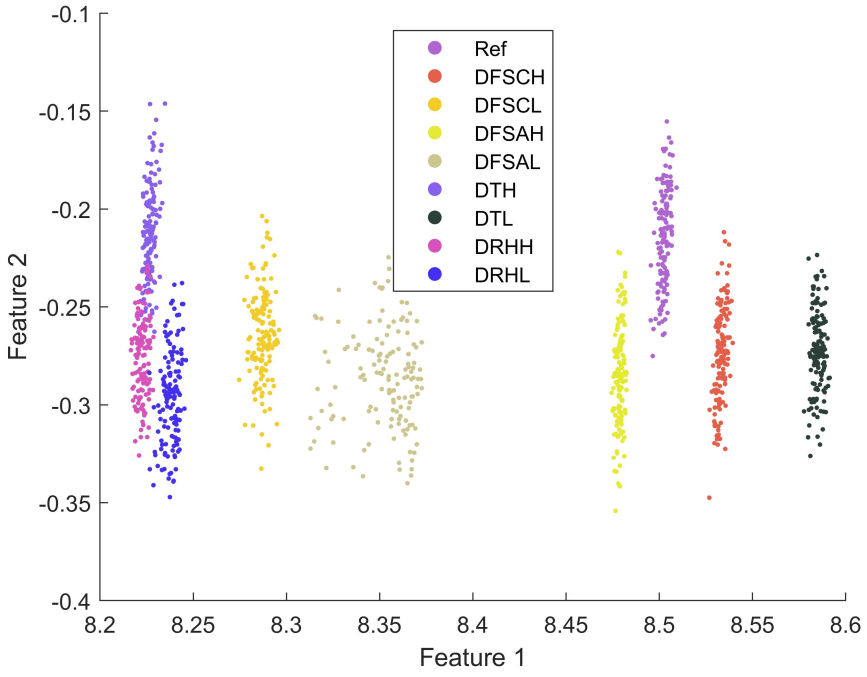
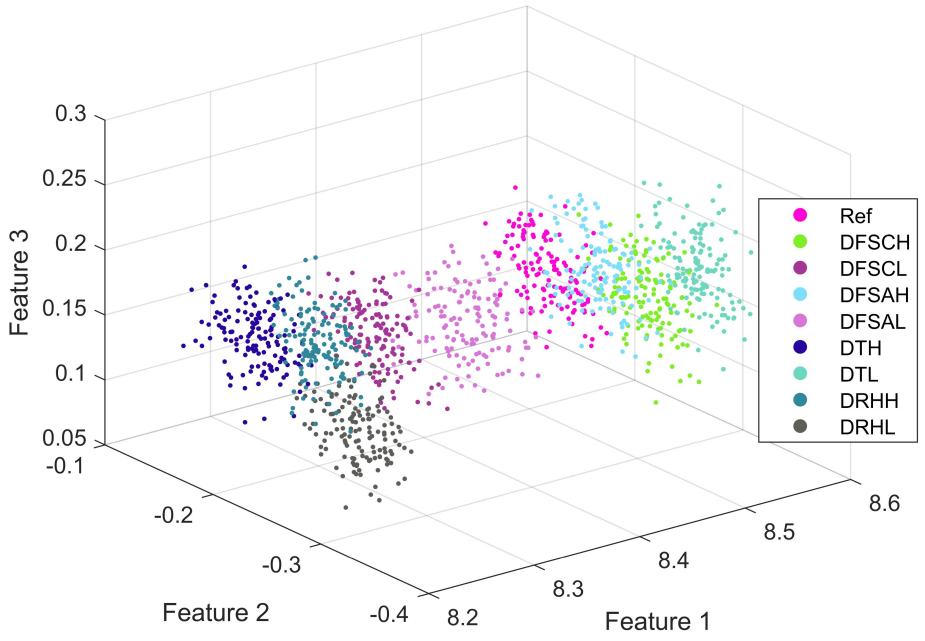


FIGURE 2.7: The box plot of (a) first-order coefficient and (b) second-order coefficient of AR model under different fault operating conditions



(a)



(b)

FIGURE 2.8: (a) The cluster based on 2 main features and (b) cluster based on 3 main features

The classification results by SVM is given as the confusion matrix in table 2.4. There are only 1 misclassified sample for the DTH condition, 2 misclassified samples for the DHH condition, and no misclassified samples for other conditions.

TABLE 2.4: The classification result based on AR model and SVM for single-fault conditions

Real conditions \ Classified conditions	C1	C2	C3	C4	C5	C6	C7	C8	C9
Ref	60	0	0	0	0	0	0	0	0
DFSCH	0	60	0	0	0	0	0	0	0
DFSCL	0	0	60	0	0	0	0	0	0
DFSAH	0	0	0	60	0	0	0	0	0
DFSAL	0	0	0	0	60	0	0	0	0
DTH	0	0	0	0	0	59	0	1	0
DTL	0	0	0	0	0	0	60	0	0
DHH	0	0	0	0	0	1	0	58	1
DHL	0	0	0	0	0	0	0	0	60

To quantitatively evaluate the diagnosis accuracy, the recall rate, precision rate and F1 accuracy are introduced. The recall rate is the ratio of the detected samples and all the existing samples of the real condition, which can be calculated by equation 2.9. The precision rate is the ratio of the right detected samples and all detected samples for a certain condition, and it can be calculated by equation 2.10. To balance the recall rate and precision rate, the F1 accuracy is the harmonic average of recall rate and precision rate, which can be calculated by equation 2.11.

$$Recall = \frac{TP}{TP + FN} \tag{2.9}$$

$$Precision = \frac{TP}{TP + FP} \tag{2.10}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{2.11}$$

Where the TP is the right detected samples, which is the number on the diagonal line of the confusion matrix; the FN is the undetected samples for a real condition, i.e. the sum of the numbers of each row except the diagonal line number; FP is the wrongly detected samples, i.e. the sum of each column except the diagonal number. The recall rate, precision rate, and F1 accuracy of each condition are given in table 2.5. The

mean F1 accuracy is 99.44%, which proves that the proposed diagnosis method is very accurate.

TABLE 2.5: The classification recall rate, precision rate and F1 accuracy for single-fault conditions

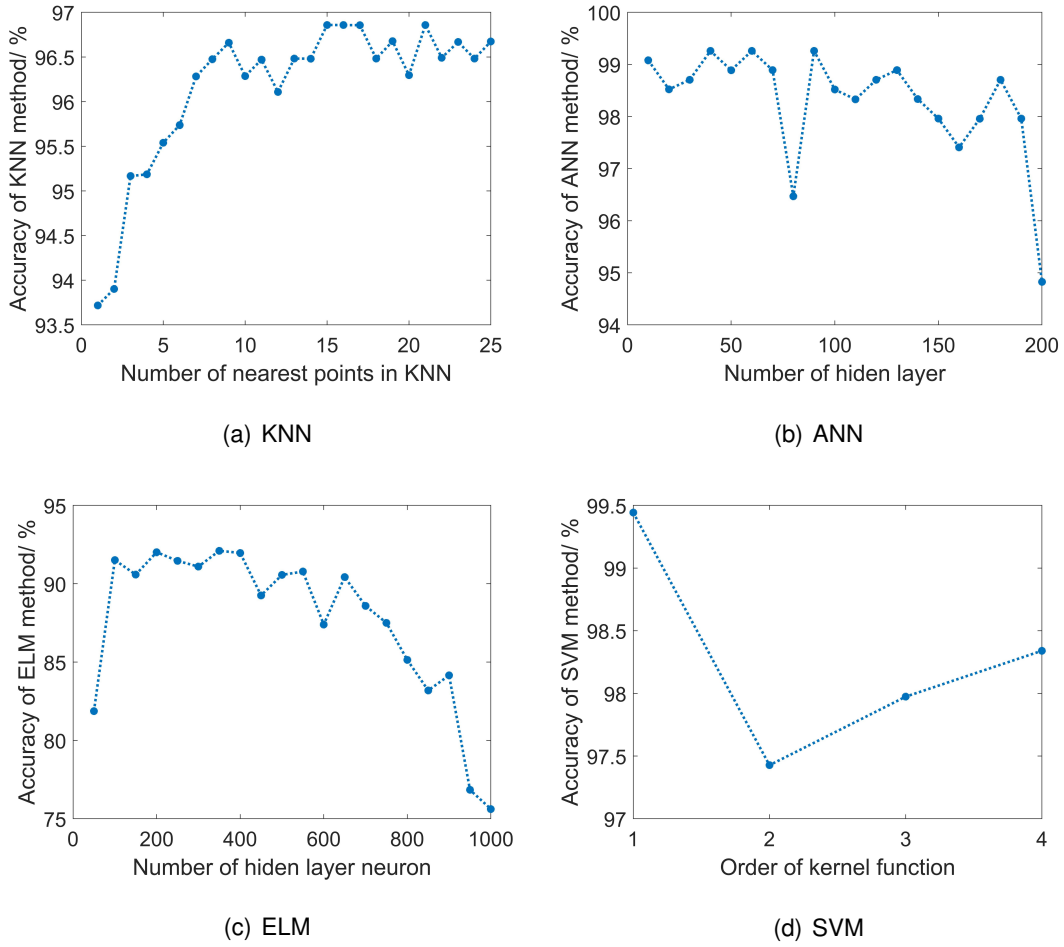
Conditions	Recall rate / %	precision rate / %	F1 / %
Ref	100	100	100
DFSCH	100	100	100
DFSCL	100	100	100
DFSAH	100	100	100
DFSAL	100	100	100
DTH	98.33	98.33	98.33
DTL	100	100	100
DHH	96.67	98.31	97.48
DHL	100	98.36	99.17
Average	99.44	99.44	99.44

The diagnosis is carried out by different classifiers in this research, i.e. KNN, ANN, ELM as explained in Section 3.2. The key hyper-parameters of those methods are researched, so that to find the parameters that can obtain the best classification result. The effect of the number of nearest points for KNN accuracy is plotted in figure 2.9(a); the effect of hidden layer number for ANN is shown in figure 2.9(b); the effect of neuron numbers for ELM is plotted in figure 2.9(c); and the effect of polynomial function order for SVM is shown in figure 2.9(d). For all four methods, the best F1 accuracy of different methods are given in table 3.1.

The SVM method can give the best result than other methods, and the accuracy of the ANN method is almost the same as SVM. In this study, the ELM method is worse than SVM and ANN, it may be because there is only one hidden layer in ELM and it is not sufficient to describe the actual relationship. Actually, the main advantage of ELM is the calculation speed compared with other methods, according to its design [150]. The good accuracy of SVM may come from its particular classification logic. However, all four classification methods can give accuracy of more than 92%, which also proves that the features extracted by the AR model from voltage fluctuation data can represent the operating conditions, and it is very effective for the PEMFC diagnosis.

Also the computational burden is another important index for PEMFC diagnosis because it will decide whether the algorithm can be applied in a real application. In this research, the different algorithms all run on a computer with a processor AMD A8-4500M 1.90 GHz, and the computational time is also listed in table 3.1. It can be seen

that the SVM is not the most time-saving, while the ELM method is the fastest. However, the computational time of all four methods are of the same magnitude, and they are very short compared to the measurement time which is in second magnitude. Therefore, the computational time is of little importance compared to diagnosis accuracy, and the SVM and ANN methods are the most suitable for the PEMFC diagnosis algorithm.



(a) KNN (b) ANN (c) ELM (d) SVM

FIGURE 2.9: The effect of hyper-parameters for the diagnosis accuracy of different methods: (a) KNN (b) ANN (c) ELM (d) SVM.

TABLE 2.6: The F1 accuracy and computational time of different classification methods

method	accuracy / %	computational time (s)
KNN	96.86	0.0425
ANN	99.26	0.0893
ELM	92.1	0.0327
SVM	99.44	0.0756
literature [159]	95.5	

The same problem was also studied by another method in literature [159], in which the singularity of the voltage fluctuation data was applied as features. The singularity

is an important feature for time-series data calculated by wavelet analysis, and the fault operating conditions are classified by the KNN method. In the research, the diagnosis accuracy is 95.5%, thus the accuracy of the proposed method is higher than the literature, meaning that the AR model coefficients are very effective features.

2.4.3 Diagnostic accuracy for multi-fault conditions

The same method is also applied to multi-fault conditions. The accuracy and computational time of the 4 methods are listed in table 2.7. Similar to single fault conditions, the ANN and SVM can give the most accurate result. However, the ANN method is a little more accurate than SVM in this case, and the performance of the KNN and ELM methods is poor. As has been addressed above, the accuracy of the classifier will be affected by the features and problems, so the ANN and SVM methods are more suitable in the diagnosis based on the AR model. The computational time has little difference for the 4 methods, so the accuracy is the emphasis for the choice of the classifier.

TABLE 2.7: The best F1 accuracy and correspond computational time by different classification methods for multi-fault condition

method	accuracy / %	computational time (s)
KNN	85.75	0.0748
ANN	93.18	0.0385
ELM	80.8	0.0590
SVM	91.77	0.0756
literature [159]	90	

The confusion matrix of the classification by ANN for multi-fault conditions is shown as table 2.8, and the recall rate and precision and F1 accuracy are shown in table 2.9. For multi-fault conditions, the different fault operating conditions may result in similar health problems in PEMFC and affect each other, and the situation is more complex and the voltage fluctuation is less regular. Therefore, the accuracy of the multi-fault conditions is a little lower than single-fault conditions. However, the accuracy of the proposed method is 93.18%, and it is superior compared to the diagnosis accuracy of literature [159], which is 90%. Therefore, the proposed method can provide better diagnostic accuracy for both single-fault conditions and multi-fault conditions.

TABLE 2.8: The classification result based on AR model and ANN for multi-fault conditions

Real conditions \ Classified conditions	C1	C2	C3	C4	C5	C6	C7	C8
Ref	96	0	5	0	0	0	13	6
DFSC	0	113	1	3	1	0	0	2
DFSA	8	3	109	0	0	0	0	0
DT	0	0	0	117	3	0	0	0
DH	0	0	0	6	114	0	0	0
DT+DFSC	0	1	0	0	0	119	0	0
DT+DFSA	3	3	0	0	0	1	112	1
DT+DH	5	0	0	0	0	0	0	115

TABLE 2.9: The classification recall rate, precision rate and F1 accuracy for multi-fault conditions

Conditions	Recall rate / %	precision rate / %	F1 / %
Ref	85.71	80.00	82.76
DFSC	94.17	94.17	94.17
DFSA	94.78	90.83	92.77
DTH	92.86	97.50	95.12
DH	96.61	95.00	95.80
DT+DFSC	99.17	99.17	99.17
DT+DFSA	89.60	93.33	91.43
DT+DH	92.74	95.83	94.26
Average	93.20	93.23	93.18

2.5 Discussions

As has been addressed above, the voltage fluctuation data can reveal important features of the PEMFC system, but the quantitative research about the effect of sampling frequency and window length on the diagnosis accuracy has not been studied yet. Therefore, the diagnosis based on data of different sampling frequencies and different window lengths is compared and analysed here.

As the voltage data was detected at 3000 Hz, we can obtain the down-sampled data with lower frequencies, and the data window can also be set as different lengths. By applying the voltage data with different frequencies and different window lengths in the proposed diagnosis algorithm, the diagnosis accuracy can be obtained in figure 2.10. According to the figure, the longer is the window length and the higher is the sampling frequency, the higher is the diagnosis accuracy. For example, to obtain a diagnosis accuracy higher than 98.69%, 0.8 s data is needed under the sampling frequency of 3000 Hz, while 2 s is needed under the frequency of 2000 Hz. Therefore,

with a higher sampling rate, the signal contains more information, which significantly reduces the diagnostic time. It also proves that high-frequency sampling is meaningful for PEMFC diagnosis. What's more, with the current measurement technique, the sampling frequency can be set to 3000 Hz without extra costs, so it is feasible to make use of fluctuation patterns by measurement with relatively high frequency.

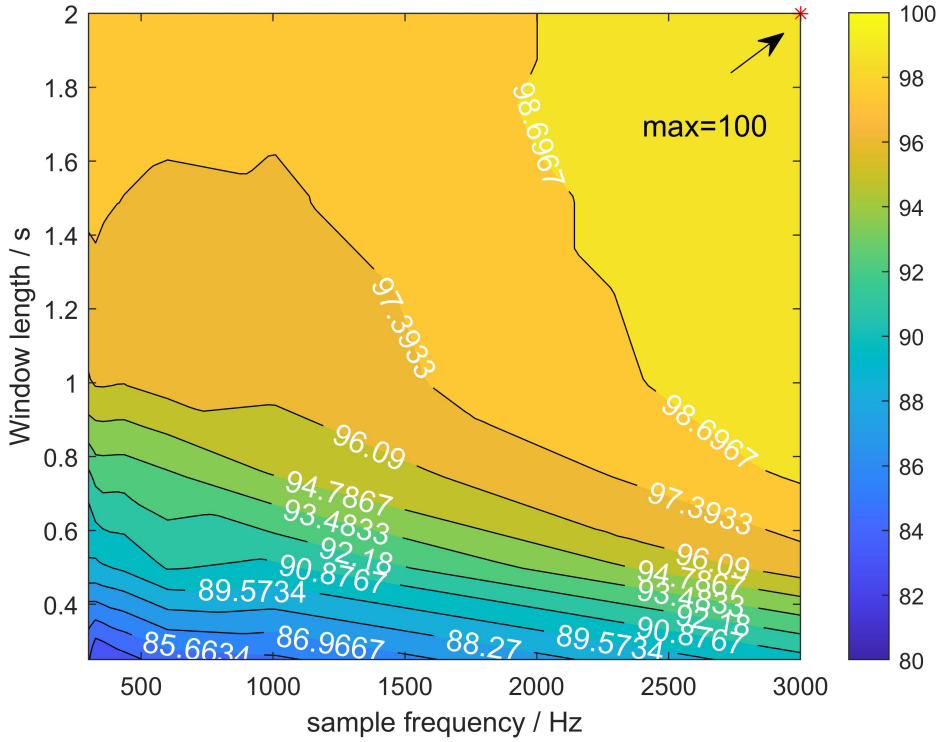


FIGURE 2.10: The diagnosis accuracy under different sampling frequencies and different window lengths

Several studies have focused on the explanation of the source and characteristics of the voltage fluctuation, by considering both electrochemical reaction aspects (effects of electric double layers, reduction–oxidation mechanisms in the electrodes, reactant diffusion in the GDLs), and more global thermal/fluid issues in the PEMFC [63] [158] [156]. However, it is still complex to describe precisely the physical phenomena involved in PEMFCs, especially to propose models capable of faithfully reproducing the evolution of the cell voltage morphology under different operating conditions and over a wide range of frequencies. The models based on current knowledge of PEMFC are therefore still relatively unsuitable for diagnostic applications. Nevertheless, the phenomenological approaches could also be useful in the development of data-driven

diagnostic strategies, for example, providing signatures and relevant descriptors for fault patterns.

2.6 Conclusions

In this chapter, a novel PEMFC diagnosis method is proposed based on the voltage fluctuation data, where the features are extracted from the AR model. The AR model coefficients are first time applied as features to classify the PEMFC operating conditions. The distribution of coefficients is related to the different fault conditions; as a result the AR model can well describe the voltage time-series data and provide information for voltage fluctuation patterns. With the proposed diagnosis method, nine single faults operating conditions and 8 multi-fault conditions of the PEMFC system can be classified accurately, and the data acquirement time is only 1 second, which is very important for early fault detection. The accuracy of the proposed method is 99.44% for the single-fault condition and 93.18% for the multi-fault conditions. Compared to other methods described in the literature, the diagnostic approach proposed in this article has two main advantages, i.e. the high accuracy of the classification results and the reduced time required to generate and exploit the fault patterns. Four different classification methods were applied and compared in this research work, which showed that the SVM and ANN methods are more suitable than the KNN and ELM methods. A higher detection frequency and a wider sample window can increase diagnostic accuracy. Therefore, increasing the sampling frequency can help to obtain enough data for accurate diagnosis within a shorter measure time, which is beneficial for timely diagnosis and treatment.

Except for the voltage fluctuation pattern being a useful diagnosis feature, the EIS technology is another experimental tool with great potential for PEMFC diagnosis. Therefore, in the next chapter diagnostic methods based on EIS are presented.

Chapter 3

Diagnosis based on electrochemical impedance spectroscopy

3.1 Introduction

Different kinds of data can be used for PEMFC diagnosis, such as polarization curve data, operating parameters, EIS data, and other physical parameters [21]. The polarization curve is a useful diagnostic tool because only electrical data need to be collected. However, the faults of different causes cannot be distinguished by the polarization curve because the voltage losses cover each other [169]. On the contrary, EIS is an important diagnostic tool that can reflect the inner conditions of various components [29].

As has been addressed in chapter 1, EIS is a diagnostic tool that has been widely applied to electrochemical devices such as fuel cells and Li-ion batteries [170]. The basic idea of EIS is to detect the electrical response under small fluctuation excitation with a certain frequency, and the frequency domain behaviour can reflect the inner state of PEMFC. M. Zhiani et al.[171] applied the EIS to characterize a PEMFC under different thermal and pressure stresses. The EIS at different operation times was analysed and it corresponded well with theoretical supposition. R. Pan et al.[36] found that the electrical resistance of EIS was almost proportional to the voltage during the

long-term operation, thus it could also be applied to the prognosis and predict the degradation. However, it is still an open question to make use of it in diagnosis.

Therefore, two diagnostic methods are proposed based on the EIS technology in this chapter. In the first part, the diagnosis method based on an equivalent circuit model (ECM) and adaptive neural fuzzy inference system (ANFIS) is addressed in section 3.2. The second method is presented in section 3.3, and it is based on two quick detective EIS features.

3.2 Diagnosis by equivalent circuit model and adaptive neuro-fuzzy inference system

3.2.1 Introduction

ECM is an important tool to analyse EIS data [172]. In ECM, the PEMFC components can be represented by electrical elements; hence the PEMFC behaviours can be analysed like circuits. Different kinds of ECM have been proposed for different applications, as well as for PEMFC. The most used electrical elements for PEMFC include resistance, capacitor, inductance, etc. However, the classical electrical elements cannot describe the total characters of the PEMFC because there are some particular processes that affect the PEMFC electrical characters. Some other non-linear elements are used, including the CPE and Warburg element, etc. Those elements are also called non-integer order elements or fractional-order elements because they introduce the fractional-order differential operation into the circuit. Fractional-order ECM has been applied to electrochemical devices a long time ago [173]. Q. Zhu et al.[174] applied the fractional-order ECM to the lithium-ion battery, and the state of charge can be estimated. M. Iftikhar et al.[175] built a fractional-order ECM of PEMFC and proved that the fractional-order model can give more accurate results than the integer-order model. S. Mohammad et al.[176] applied the process modelling method to analyse the full set of processes in the PEMFC, and a very comprehensive ECM was provided.

The ECM parameters can give important information about the inner state of PEMFC, so they can be applied to diagnosis. N. Fouquet et al.[27] used Randle's model with

CPE to describe the EIS data, and 5 ECM parameters were applied to indicate the water condition in PEMFC. However, this model was not accurate enough because the anode impedance was omitted. Also, only membrane water content faults were researched qualitatively in this research, and no quantitative diagnosis method was given. Few studies focused on the quantitative diagnosis based on EIS. S. Wasterlain et al.[177] tried to achieve diagnosis by Bayesian networks, and the position coordinates of six points of EIS data were applied as features to distinguish different faults. However, as the real part and imaginary part of EIS points are directly applied as features, the physical mechanism and behaviour pattern behind the EIS curve were not analysed and extracted. Also, the diagnostic accuracy was only 91% in their research. Therefore, there is still a gap to real applications, and more accurate diagnostic methods are necessary.

In this section, a new diagnostic methodology based on an ANFIS is proposed to fill the gap between laboratory research and application. A fractional-order ECM is applied to describe the PEMFC EIS behaviours, and the impedances of both membrane, anode, cathode, and gas diffusion layers are included. A new parameter identification method that combines genetic algorithm (GA) and Levenberg–Marquardt (LM) algorithm is proposed to identify the ECM. 10 ECM parameters can be obtained by the proposed identification method with both a wide range and good accuracy. The proposed method is validated by the experiments under different fault conditions, thus the relationships between the ECM parameters and fault causes are thoroughly analysed and understood, which can provide reference to the fault diagnosis. Then a novel diagnosis algorithm based on ANFIS is proposed according to the relationship between ECM parameters and fault causes. For the first time, the ECM parameters are applied to quantitative diagnosis. This method can detect the conditions with the same fault by clustering, and then classify the new samples into the right fault groups by ANFIS. The proposed diagnosis method is validated by the experiments, and all the fault causes can be distinguished accurately, including the membrane flooding, drying out, and mass transfer fault. Also, the proposed method is compared to another EIS diagnostic method and proved to be superior.

Also, a framework is proposed to connect the EIS on-line measurement, real-time diagnosis, and fault-tolerant control, thus trying to achieve a closed-loop strategy to improve the PEMFC durability and reliability. Based on the framework, the proposed

diagnosis strategy is further developed on an industrial-level digital signal processor (DSP). A computationally efficient ECM parameter identification method, i.e. Powell's algorithm, is applied to extract diagnostic features from EIS data, which is suitable for the computational-resource-limited system as DSP. Then the fault classes can be directly predicted by ANFIS on DSP. The ANFIS can be adaptively tuned on-line in the DSP, so it is suitable for long-term application. The diagnosis method is experimentally demonstrated on DSP in real time, and the full diagnosis process can be achieved within 250 seconds, thus it is practical for real applications.

In part 3.2.2 the proposed PEMFC ECM is addressed. In part 3.2.3, the EIS result is analysed according to the ECM and verified by the experiment. Then the diagnosis algorithm is explained and validated in part 3.2.4. The implementation method on DSP is presented in section 3.2.5, then the implementation result is addressed in section 3.2.6. Finally, the main conclusions of this section are summarized in section 3.2.7.

3.2.2 Diagnostic methods based on equivalent circuit model

3.2.2.1 Fractional-order equivalent circuit model

The PEMFC EIS can be analysed by ECM. Different ECM can be applied to different devices. Randle's model [178] [179] is one of the most basic models that include one capacitor, as shown in figure 3.1(a). Here the capacitor C_{dl} is used to describe the storage of electrons in the double layer of the electrode; the resistance R_c is used for the resistance of the oxygen reduction reaction, and the parallel connection of C_{dl} and R_c represent the charge transfer process; R_m is the resistance of the proton to pass the membrane and the electrons to flow through the wire. The biggest advantage of Randle's model is the simplicity because there are only 3 elements, thus it is widely applied. However, even though the impedance of the anode is usually smaller than the cathode, neglecting them will introduce noteworthy errors, so Randle's model is not suitable for our PEMFC system.

Compared to Randle's model, the 2-capacitor model is more accurate for PEMFC because both the impedance of the anode and cathode are included [29] [180], which is shown in figure 3.1(b). In this model, the anode is also represented by the parallel connection of a capacitor and resistance, so the electron storage and reaction

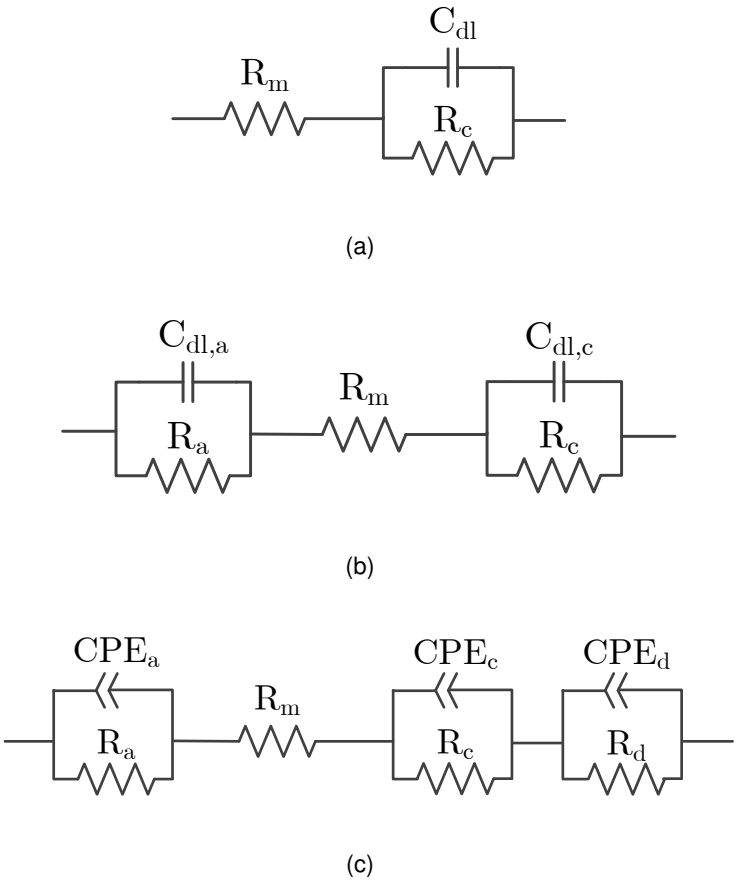


FIGURE 3.1: The equivalent circuit models of PEMFC: (a) Randle's model (b) model with 2 capacitors (c) model with 3 constant phase elements

resistance of the anode are also considered. However, research shows that there is also additional mass diffusion resistance under certain conditions [181], which cannot be represented by this model. The mass diffusion impedance represents the resistance when the gas passes the GDL, especially on the cathode side.

A more accurate ECM [30] applied in this research is shown in figure 3.1(c). In this model, the PEMFC impedance can be represented by four parts, i.e. membrane resistance, anode impedance, cathode impedance, and mass diffusion impedance. The anode impedance, cathode impedance, and mass diffusion impedance are all represented by the parallel connection of resistance and CPE. The impedance of a CPE can be given as equation 3.1, and it is the same as a capacitor when the order is 1. The CPE is caused by the unevenness of the plates or the heterogeneity of the dielectric, and it can be represented by the connection of a lot of RC circuits. A more detailed deduction of the impedance of CPE can be found in the reference [182]. The R_d in mass diffusion impedance is caused by the flow resistance when the air and

hydrogen diffuse through the GDL, and the CPE$_d$ represents the storage of gas in the porous materials.

$$Z_{CPE}(s) = \frac{1}{Cs^{\alpha}} \tag{3.1}$$

Where the Z_{CPE} is the impedance of CPE; s is the complex variable in Laplace space; C is the effective capacitance of the CPE; α is the order of the CPE, which is between 0 and 1.

By connecting all the impedance, the total impedance of the circuit can be represented as:

$$Z_{total}(s) = \frac{R_a}{1 + R_a C_a s^{\alpha_a}} + R_m + \frac{R_c}{1 + R_c C_c s^{\alpha_c}} + \frac{R_d}{1 + R_d C_d s^{\alpha_d}} \tag{3.2}$$

Where the Z_{total} is the total impedance of the circuit; the subscript a, c and d represent the parameters of the anode impedance, cathode impedance, and mass diffusion impedance, respectively.

3.2.2.2 Parameters identification of ECM model

The aim of parameter identification is to find the values of the unknown parameters of the ECM so that the identified model is the closest to the experimental EIS data. Essentially this is an optimization process, so we can use the GA to search for the rough global optimal parameters, and then use the LM algorithm to get the precise solution.

The mean squared error (MSE) is applied as the fitting criterion, which is shown as equation 4.13. In the equation, $Re(Z)_i$ and $Im(Z)_i$ is the real part and imaginary part of the experiment data, respectively, and $\hat{Re}(Z)_i$ and $\hat{Im}(Z)_i$ are the real and imaginary parts of the impedance calculated by the model according to the equation 3.2, respectively. i is the sequence number of measurements, and the EIS frequency applied in our experiment is from 0.1 Hz to 10000 Hz. N is the number of measured points, and fifty-one points can be obtained for each EIS curve. The LM method is a stable method for non-linear optimization problems, as it combines the advantages and

avoids the disadvantages of the Gauss-Newton algorithm and gradient descent method [183].

$$L = min(\frac{\sum[(Re(Z)_i - \hat{Re}(Z)_i)^2 + (Im(Z)_i - \hat{Im}(Z)_i)^2]}{N}) \tag{3.3}$$

The LM algorithm is based on Gauss-Newton method, the optimization function is defined as $f(x)$, where x is the vector of tunable parameters, the quadratic Taylor expansion can be given as:

$$f(x_k + s) \approx q^{(k)}(s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s \tag{3.4}$$

Where $s = x - x_k$, to make $q^{(k)}(s)$ the minimum, get:

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) = x_k - G_k^{-1} g_k \tag{3.5}$$

Where the G_k is the Hessian matrix, and $g_k = \nabla f(x_k)$. However, the calculation of the Hessian matrix is complex, also the positive definiteness of the matrix cannot be ensured, so the Hessian matrix can be approximated as:

$$G \approx J^T J \tag{3.6}$$

Where the J is the Jacobian matrix, and,

$$x_{k+1} = x_k - (J^T J)^{-1} g_k \tag{3.7}$$

This is the basic idea of the Gauss-Newton algorithm. However, the reversibility of matrix $J^T J$ is not sure. To increase the stability and convergence of the algorithm, the LM method uses $J^T J + \mu I$ to replace the Hessian matrix, where I is the identity matrix and μ is a scaler, so:

$$x_{k+1} = x_k - (J^T J + \mu I)^{-1} g_k \tag{3.8}$$

The value of μ can be modified to adapt to the algorithm and decide the search rate. When μ is big, the LM method is close to the gradient descent method, and it is close to the Gauss-Newton method when the μ is small. Therefore, it is a combination of the

two methods so that to obtain the best compromise between the search rate and the algorithm stability. The LM algorithm can be given as algorithm 3.1.

Algorithm 3.1: LM algorithm

give initial point x_0, u_0;
for $k = 1$ to k_{max} **do**
  calculate g_k, J and G_k;
  if $|g_k| < \epsilon$
    end;
  if $G_k + \mu_k I$ is non-positive definite
    $\mu_k = 4\mu_k$, until $G_k + \mu_k I$ is positive definite;
  solve equation $(G_k + \mu_k I)s_k = -g_k$, obtain s_k;
  calculate $f(x_k + s_k)$, $q^{(k)}(s_k)$, and $r_k = \frac{\Delta f_k}{\Delta q^{(k)}}$;
  if $r_k < 0.25$
    $\mu_{k+1} = 4\mu_k$;
  else if $r_k > 0.75$
    $\mu_{k+1} = \frac{\mu_k}{2}$;
  else $\mu_{k+1} = \mu_k$;
  if $r_k \leq 0$
    $x_{k+1} = x_k$;
  else $x_{k+1} = x_k + s_k$;
end for
end

However, this method may obtain a local minimum solution around the initial value, therefore the initial value is very important for the LM method. To overcome the initial value setting problem, GA optimal method can be used to search the global minimum firstly, and it needs no initial values. GA is a parallel stochastic search optimization method, which simulates natural genetics and biological evolution. However, as it searches for the global optimal solution in a large interval, the precision is relatively low. Therefore the GA and LM methods can be combined together to identify the model parameters in this research so that to take the advantages of both of them.

The parameter identification method proposed in this research is shown in figure 3.2. As we can see, when the optimization problem is given, a lot of populations are created to represent different possible solutions in the GA method. Then the bad solutions are abandoned according to the evaluation environment. If the solution is not yet good enough, i.e. the fitness value smaller than the criterion F_s, GA operators will work. There reproduction, crossover, and mutation are carried out, then those solutions will be checked again until the solution is qualified. By this GA method, the rough globally optimal solution is obtained, which will be set as the initial solution for the LM algorithm,

thus the LM algorithm can search for more precise results. The LM method will run over and over again until the difference of the MSE between the two steps is smaller than the setting precision. By the combination of the GA and LM methods, the globally optimal solution can be found with high precision.

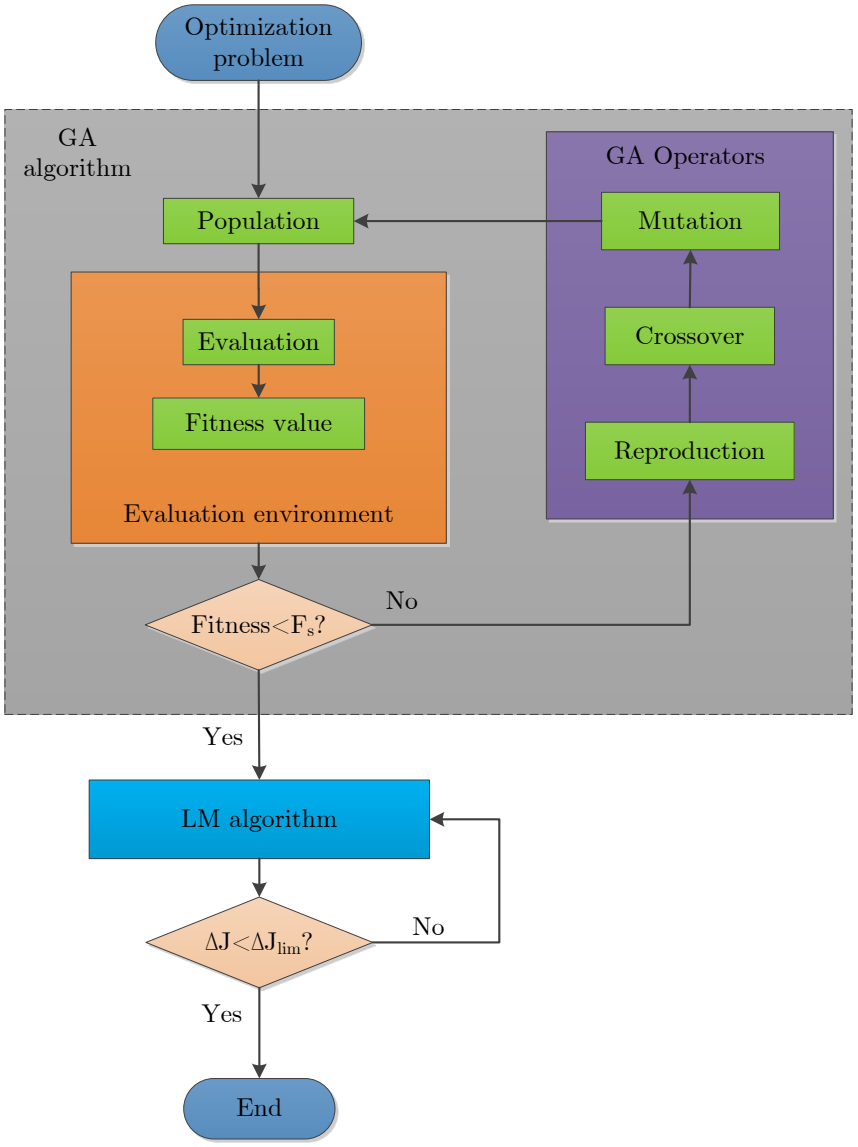


FIGURE 3.2: The parameter identification process by a combination of GA and LM method

To evaluate the effect of the proposed parameter identification method, the proposed method is compared with two other widely applied optimization methods, i.e. the particle swarm optimization (PSO) and simulated annealing algorithm (SAA). PSO is a kind of evolutionary computation, which is inspired by the predation behavior in flocks of

birds. The basic idea is to find the optimal solution through cooperation and information sharing among individuals in the group [184]. SAA is a stochastic optimization algorithm based on the Monte-Carlo iterative solution strategy, which is based on the similarity between the annealing process of solid matter in physics and general combinatorial optimization problems [185]. Using the EIS data from experiment No. 1, the identified results and the squared errors by PSO, SAA, and GA-LM algorithm are compared in table 3.1. It can be seen that the proposed GA-LM algorithm has the smallest error compared to PSO and SAA methods. Also, the parameters show great variation for different optimization methods, which indicates that the optimization problem is highly non-linear. Therefore, an accurate parameter identification method is necessary.

TABLE 3.1: The comparison of identified ECM parameters and errors of different methods

Methods	R_a	C_a	α_a	R_m	R_c	C_c	α_c	R_d	C_d	α_d	error
PSO	0.0107	1.8600	0.7594	0.0058	0.0111	4.7683	0.7202	0.0144	7.5882	1.0000	4.228×10^{-5}
SAA	0.0049	4.6886	0.8013	0.0052	0.0095	6.1178	0.5961	0.0240	5.7614	0.6628	1.384×10^{-4}
GA-LM	0.0038	0.9095	0.7223	0.0053	0.0129	0.5766	1.0000	0.0196	4.9494	1.0000	2.664×10^{-5}

3.2.3 EIS results and ECM validation

3.2.3.1 EIS experiment conditions

The EIS data was collected from the same experiments as explained in section 2.3. The detailed operating conditions for all the EIS detection experiments are given as table 3.2.

3.2.3.2 EIS under different currents

During the operation of PEMFC, the current usually changes at different levels to fit the load change. The PEMFC behaviours are quite different under different current levels, so it is meaningful to study the effects of the current level. In our experiments, the current is set as 20A, 40 A, 60 A, 80 A, 100 A, 120 A, respectively, and the EIS under different currents are shown in figure 3.3. The proposed parameters identification method is applied, and identified parameters are given in table 3.3.

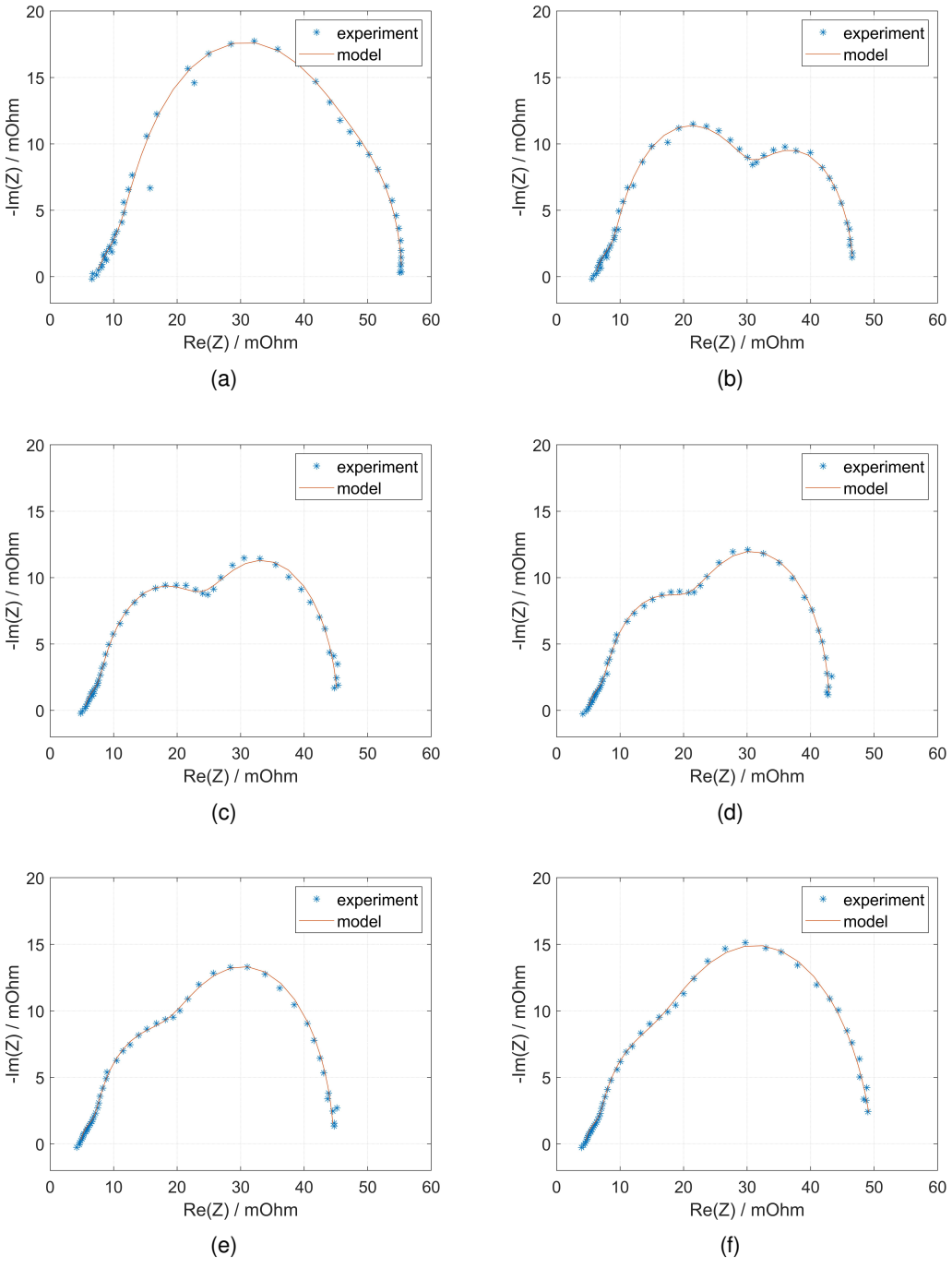


FIGURE 3.3: The EIS of experiments and model under different currents: (a) 20 A (b) 40 A (c) 60 A (d) 80 A (e) 100 A (f) 120 A

TABLE 3.2: Operating parameters of each experiment

No.	T(°C)	SC	SA	RH(%)	current(A)	No.	T(°C)	SC	SA	RH(%)	current(A)
1	65	2	1.3	50	80	27	65	1.6	1.5	50	80
2	65	2	1.3	50	80	28	65	2.6	1.5	50	80
3	65	2	1.3	50	120	29	65	2.6	1.3	50	80
4	65	2	1.3	50	100	30	70	2.6	1.3	50	80
5	65	2	1.3	50	80	31	70	2	1.3	50	80
6	65	2	1.3	50	60	32	70	2	1.3	50	80
7	65	2	1.3	50	40	33	70	2	1.3	50	80
8	65	2	1.3	50	20	34	70	2	1.3	50	120
9	65	2	1.3	50	80	35	70	2	1.3	50	120
10	65	2	1.3	50	80	36	70	2	1.3	50	100
11	65	2	1.3	50	80	37	70	2	1.3	50	80
12	65	2	1.3	50	80	38	70	2	1.3	50	60
13	65	2	1.3	50	80	39	70	2	1.3	50	40
14	65	2	1.3	50	80	40	70	2.3	1.3	50	120
15	65	2	1.3	50	80	41	70	2.3	1.3	50	120
16	65	2	1.3	50	80	42	70	2.6	1.3	50	120
17	65	2	1.3	50	80	43	70	2.6	1.3	50	80
18	65	2	1.5	50	80	44	70	1.6	1.3	50	80
19	70	2	1.3	50	80	45	70	2	1.5	50	80
20	70	2	1.3	50	80	46	70	2	1.2	50	80
21	62	2	1.3	50	80	47	70	2	1.3	50	80
22	62	2	1.3	50	80	48	70	1.6	1.5	50	80
23	65	2	1.3	46	80	49	70	2.6	1.5	50	80
24	65	1.6	1.3	46	80	50	70	2	1.3	50	80
25	65	2.6	1.3	46	80	51	72	2	1.3	50	80
26	65	2.6	1.3	46	80	52	70	2	1.3	54	80

TABLE 3.3: The identified ECM parameters under different currents

Parameters	R_a	C_a	α_a	R_m	R_c	C_c	α_c	R_d	C_d	α_d
20 A	0.0057	1.4449	0.6282	0.0072	0.0340	0.4712	0.9573	0.0085	10.719	1.0000
40 A	0.0033	0.4989	0.7775	0.0062	0.0214	0.4824	0.9672	0.0158	8.9344	1.0000
60 A	0.0027	0.5888	0.7682	0.0054	0.0164	0.5919	0.9537	0.0208	6.0515	0.9772
80 A	0.0030	0.6433	0.7634	0.0049	0.0135	0.6198	0.9765	0.0216	4.4173	1.0000
100 A	0.0026	0.4247	0.8232	0.0048	0.0115	0.7325	0.9692	0.0258	3.5383	0.9640
120 A	0.0025	0.4146	0.8325	0.0046	0.0077	0.7785	1.0000	0.0350	3.1134	0.8713

As can be seen from the figures, the impedance calculated by the ECM model is very close to the experiments. When the frequency is high, i.e. the left part where the imaginary part is close to 0, there is a small arc that represents the impedance of the anode. The arc is very small, and it means that the impedance is small in the anode. This is consistent with other studies because the hydrogen oxidation reaction in an anode catalyst is much easier than the reaction in the cathode. Then, when the frequency is middling, there is an obvious arc that represents the impedance of cathode charge transfer. Thirdly, when the frequency is small, i.e. the right part of the EIS plot, there is a third arc that represents the impedance of mass diffusion.

Comparing the EIS under different currents in figure 3.3, it can be found that the cathode

arc is smaller and smaller when the current increases, while the mass diffusion arc is higher and higher. As the current increases, more oxygen is consumed and the charge transfer resistance will decrease. However, more water is produced in the stack, so the diffusion of gas to catalyst is blocked, and the arc of the anode does not change obviously.

This phenomenon can also be seen by the table 3.3. As the current increases, the R_a, R_m, R_c and C_d decrease, while the C_c and R_d increase. As has been addressed above, more water is produced with the increase of the current, and the humidity in the whole stack will increase. With the increase of humidity, the resistance of the membrane decrease because the water content is helpful for the proton to traverse the membrane. At the same time, the resistance in the anode and cathode will decrease because the catalytic reactions also benefit from the humidity. However, as the humidity increases, the resistance for reactant gas to enter the reaction point increases because the passages in porous material are blocked by the water. As for the capacitance, the storage of gas in the GDL will decrease as more porous places are taken by water, so the capacitance C_d that represents the gas storage ability decreases. What's more, the increase of cathode capacitance C_c may be related to the dielectric coefficient of the catalyst double layer, which can be caused by the increase of water in the double layer. For the other parameters, i.e. C_a, α_a, α_c, and α_d, there are no significant changes for them, which means that they are not sensitive to the stack current.

The internal state of the PEMFC will be affected by the water generated when the current density is high, and the high level of water content may cause noise in the EIS measurement. The generated water will cause high water content in the membrane, catalyst layer, and gas diffusion layer, thus affecting the PEMFC behaviours as well as EIS measurement noise. It is especially crucial that the measured EIS curve should be relatively smooth, as the experiments in this research work, so that proving that the inner condition is stable under the current stimulation under different frequencies. To guarantee that the noise is relatively small and negligible, one of the most important things is to keep the operating conditions within the design limitations. Otherwise, the behavior pattern of PEMFC may be quite unstable; thus the noise may take dominating place and fail the EIS measurement.

3.2.3.3 EIS under different fault conditions

As has been explained in Section 2.3, the EIS under different operating conditions are recorded and they can be analysed by the proposed model. The experiments with the fault of SC, SA, temperature, and RH are researched.

(1) EIS under different SC

The EIS under higher and lower SC fault conditions (FSCH and FSCL, respectively) are shown in figure 3.4. It can be seen that the results of the model match the experiments well, and the regularity is obvious. As the SC increases, the total impedance decreases. The cathode arc and diffusion arc change in different magnitudes, i.e. the cathode arc increases slower than the mass diffusion arc when the SC decreases.

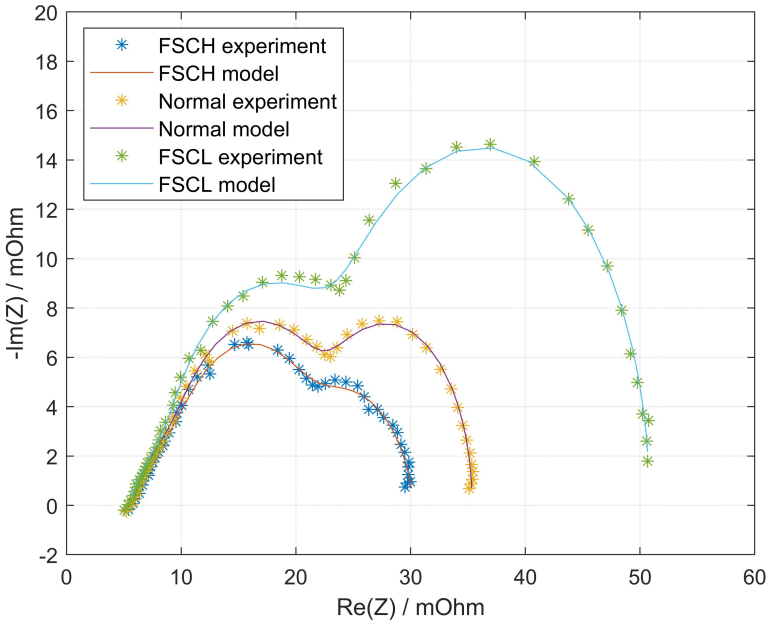


FIGURE 3.4: The EIS of experiment and model under SC fault conditions

The same conclusion can also be drawn according to the identified parameters as shown in table 3.4. As the SC decreases, the resistance of cathode charge transfer increases because the oxygen concentration at the catalyst is limited. The mass diffusion impedance is the most affected in those cases. The mass diffusion resistance increases when the SC decreases, while the capacitance decreases. This is caused by the shortage of oxygen in the cathode. As the oxygen is insufficient, the mass diffusion process is the main restriction, so the diffusion resistance increases. Also, the capacitance decreases because the air storage in the cathode is insufficient.

TABLE 3.4: The identified ECM parameters under different fault conditions

Parameters	R_a	C_a	α_a	R_m	R_c	C_c	α_c	R_d	C_d	α_d
Normal	0.0049	0.8076	0.7242	0.0058	0.0119	0.4638	1.0000	0.0128	6.4416	1.0000
FSCH	0.0052	1.0031	0.6921	0.0059	0.0098	0.5210	1.0000	0.0095	11.412	0.8468
FSCL	0.0045	1.1765	0.6897	0.0055	0.0136	0.5266	1.0000	0.0273	4.6193	1.0000
FSAH	0.0038	0.5252	0.7688	0.0058	0.0122	0.4325	1.0000	0.0125	5.9386	1.0000
FSAL	0.0017	0.3565	0.8647	0.0056	0.0082	0.5526	1.0000	0.0394	13.8461	0.5000
FTH	0.0048	0.5956	0.7450	0.0063	0.0104	0.4134	1.0000	0.0169	9.0078	0.8005
FTL	0.0029	1.2408	0.7053	0.0050	0.0132	0.6097	0.9771	0.0184	5.0479	1.0000
FRHH	0.0023	0.3316	0.8299	0.0053	0.0149	0.6334	0.9120	0.0175	5.1134	0.9946
FRHL	0.0030	1.0328	0.7172	0.0051	0.0141	0.7087	0.9365	0.0196	4.8441	1.0000

(2) EIS under different SA

The EIS result under higher and lower SA conditions (FSAH and FSAL, respectively) can be seen in figure 3.5. Different from the SC fault conditions, the EIS of SA fault conditions have little difference for the anode and cathode impedance, but they are quite different for mass diffusion impedance. The mass diffusion arc is very big for the low SA condition. According to the table 3.4, both the mass diffusion resistance and the capacitance increase a lot when the SA decreases. This can be explained by the shortage of hydrogen at the anode and the surplus of oxygen in the cathode. As the supply of hydrogen is insufficient, the mass diffusion resistance increases. At the same time, the oxygen in the cathode is not all consumed, so the air storage increases, and the capacitance increases.

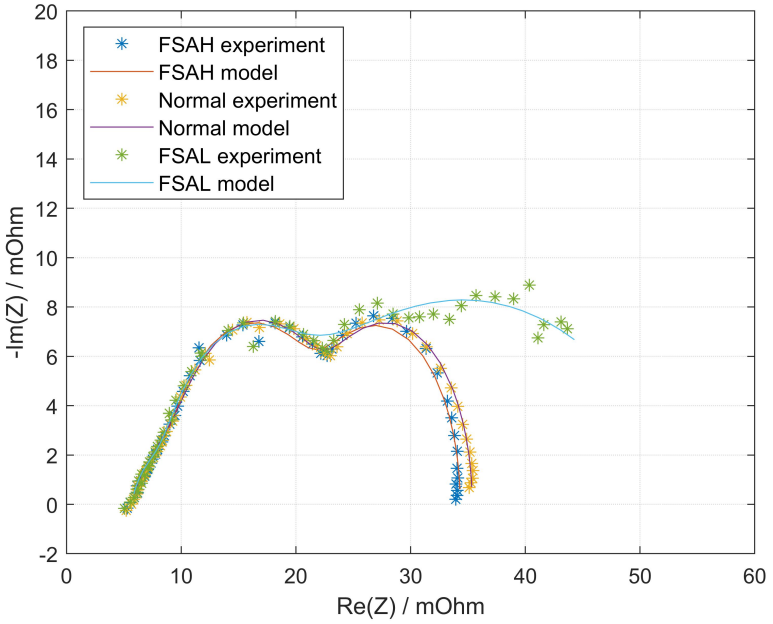


FIGURE 3.5: The EIS of experiments and model under SA fault conditions

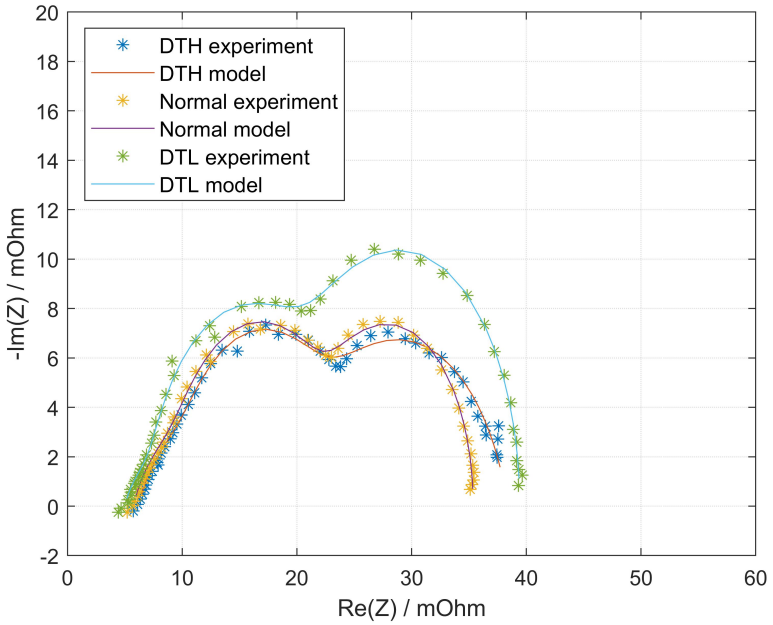


FIGURE 3.6: The EIS of experiments and model under temperature fault conditions

(3) EIS under different temperatures

The effects of stack temperature are studied, too. The EIS under higher and lower temperature fault conditions (FTH and FTL, respectively) is shown in figure 3.6. It can be seen that the cathode arc increases when the temperature decreases and the mass diffusion arc at low frequency is big in both lower and higher temperature cases. When temperature changes, the most important feature is the resistance of the membrane, i.e. the resistance at high frequency when the imaginary part is 0. There is a great change of membrane resistance which is not observed in other cases. According to the table 3.4, the membrane resistance decreases with the decrease of temperature, and this should be related to the water content and other physical properties of the membrane because high temperature means low humidity in the membrane. As the stack temperature affects almost all the processes in the stack, the impedances caused by oxidation-reduction reactions, water transport between the cathode and anode, and the mass transfer in the GDL all change under different operating temperatures. The gas diffusion resistance under both higher and lower temperature cases are bigger than the normal case, which shows that the temperature under the normal case is favourable for the gas to pass the GDL.

(4) EIS under different RH

To directly research the effect of RH, the RH of the input gas is adjusted. The EIS under higher and lower RH fault conditions (FRHH and FRHL, respectively) is given in figure 3.7. It is clear that the impedance under normal condition is the smallest compared to the high and low humidity cases. Both the cathode arc and mass diffusion arc increase when the humidity deviates from the normal value, and the change of mass diffusion impedance is more obvious than the cathode impedance. Also, the impedance under low RH case is bigger than under high RH case. According to the parameters shown in table 3.4, both resistance of the cathode and mass diffusion increase under higher and lower RH fault conditions, which indicates that both the reactions at the cathode and the mass diffusion in GDL rely on the appropriate humidity level.

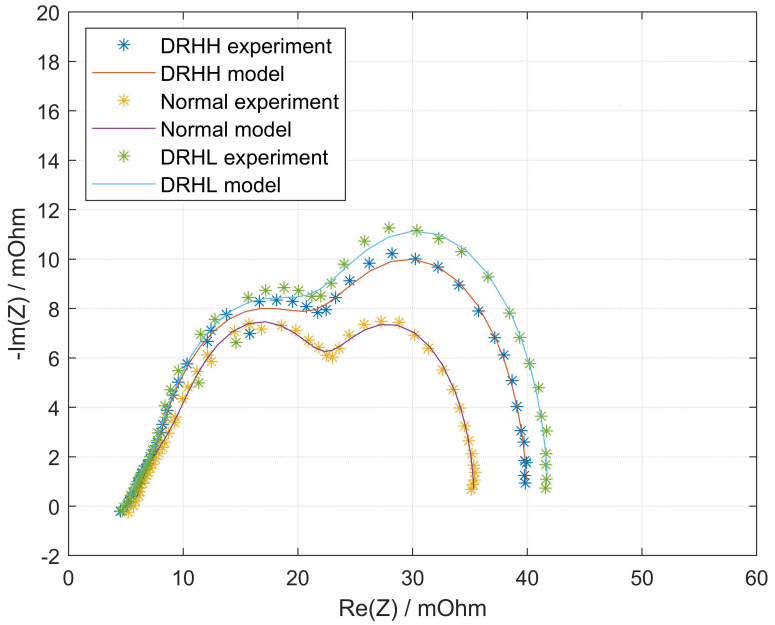


FIGURE 3.7: The EIS of experiment and model under relative humidity fault conditions

The different operating parameters can also affect each other. As addressed in the reference [10], the thermal effect in PEMFC becomes more critical at higher current density and/or lower gas diffusion layer thermal conductivity, and the thermal management and stack cooling are a significant engineering challenge. Joule heating and water generation are two main effects that affect the inner state of the PEMFC when current changes, and it will affect the water content in the membrane, gas diffusion layer, and the reaction resistance in catalyst layers. Actually, almost all parameters affect the PEMFC condition by affecting the water conditions in the different parts of PEMFC. Of course, the temperature in different parts is also important, but the temperature

change can be regarded as a kind of water content fault to some extent, because the temperature will affect the water saturation and the relative humidity. Therefore, monitoring the water content in different components is one of the most important tasks for PEMFC diagnosis. The operating parameters affect each other in a complex way. For example, the change of stoichiometry will not only change the mass concentration in the reaction position, the rest gas will take out water from the gas diffusion layer, and finally affect the water content in different components. It is possible that the change of different operating parameters will lead to similar fault symptoms. However, this situation exists in almost all diagnosis problems, because it is impossible to collect the information of all aspects. To carry out a diagnosis with good resolution and fineness, as much as possible useful features of the fault should be collected, so that they can be compared in different aspects. However, there is a compromise between the diagnosis accuracy and diagnosis cost.

3.2.4 Diagnosis result based on ANFIS

3.2.4.1 Classification by k-means clustering

According to the analysis above, it is obvious that the proposed ECM can represent the EIS behavior, and the parameters can represent the different fault causes. The most sensitive ECM parameters include the R_m, R_c, C_c, R_d and C_d. According to the parameters of all 52 samples, the k-means clustering algorithm is applied here to classify the experiments into different fault groups.

K-means clustering is widely applied to unsupervised classification problems. The basic idea of k-means is to find the smallest sum of the distance between the cluster members and the cluster centres where several distance representation methods can be chosen [138]. In this research, the squared Euclidean distance is applied. The number of groups is set according to the silhouette values, which can indicate the similarity degree of a sample to the group.

To maintain the interpretability, no feature extraction methods are applied to change the original dataset here. As there are five possible parameters to represent the fault conditions, the features applied to the classification are selected by the method of exhaustion. The different combinations of features are applied to the k-means

clustering method, and it is found that the R_m and C_d can be applied to classify the experimental samples into four groups with good precision. The classification result is given in figure 3.8.

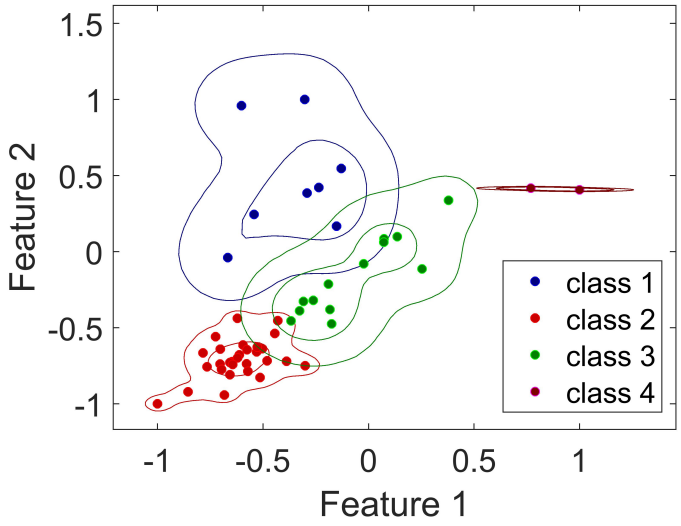


FIGURE 3.8: Classification of experiments by k-means clustering algorithm

The first feature is R_m, i.e. the membrane resistance of PEMFC, and it can represent the total water content condition of the membrane. The second feature is the mass transfer capacitance C_d, and it can represent the gas diffusion condition in the GDL. Even though the water content in the membrane and the water content in GDL will affect each other, they represent different fault causes, and the mitigation measures are different. Combining the classification result with the experimental conditions given in table 3.2, we can conclude that class 1 is the mass transfer fault conditions because the membrane resistance is small but the capacitance is big; class 2 is the slight flooding fault conditions because the membrane resistance and capacitance are both small; class 3 is normal conditions because the membrane resistance and capacitance are both in middle; class 4 is slightly drying out conditions as both membrane resistance and capacitance are big.

3.2.4.2 Diagnosis by ANFIS

According to the analysis above, the experiments can be classified into 4 groups, and the task of diagnosis is to find out the right fault class of the operating state. The most

used diagnostic methods include model-based and data-based methods, and the data-based method is applied in this research because no PEMFC model is needed; thus it is more convenient to carry out.

Many data-based methods can be applied to PEMFC diagnosis, such as various neural networks. However, the neural networks based methods usually need a lot of data to tune the massive parameters, so it is not suitable for our case. To find the diagnostic regularity from a limited database, fuzzy logic is usually applied. Fuzzy logic is a kind of method to make decisions according to the qualitative description. Different from the boolean operation, the fuzzy logic method uses the membership degree to represent the state rather than yes or no; thus it is similar to the human decision process. However, the membership function and the decision rules of fuzzy logic should be provided by researchers according to the expert knowledge of the system, which is quite tricky because the inner process of PEMFC is not well understood yet.

To take advantage of both neural network methods and fuzzy logic algorithm, the ANFIS is proposed [97]. The fuzzy logic is applied to make decisions in ANFIS, but the parameters of the fuzzy system are tuned automatically by the adaptive neural network method, and no presupposed rules are needed. Therefore, it is quite suitable for PEMFC diagnosis here. A typical ANFIS system can be represented in figure 3.9.

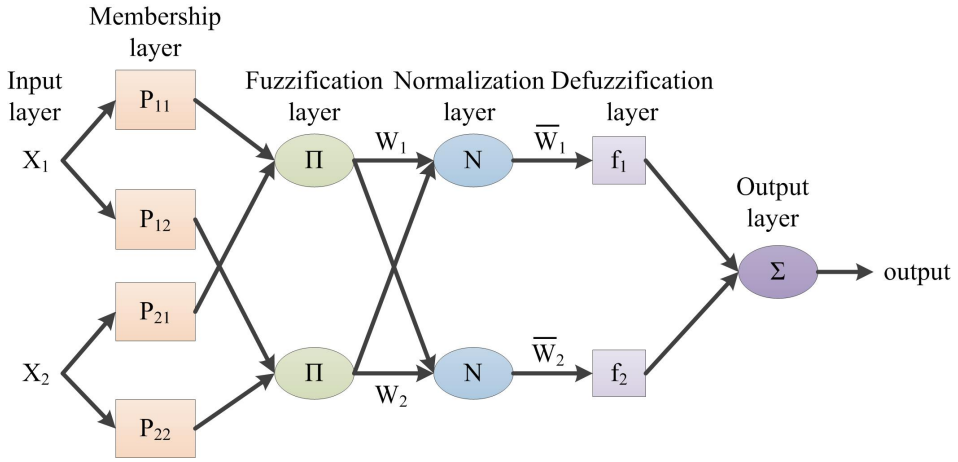


FIGURE 3.9: The typical structure of ANFIS

There are 6 layers in ANFIS. Firstly, the input layer gives the input features. Then the membership function is used to give the degree that the input feature belongs to this class. The widely used membership functions include the triangle function, Gaussian

function, and trapezoidal membership function. For example, the Gaussian function can give the membership as equation 3.9:

$$P(x) = e^{-(\frac{x-a}{\sigma})^2} \tag{3.9}$$

Where P is the membership; x is the feature; a and σ are the parameters that decide the position and shape of the membership function.

The third layer is the fuzzification layer, there fuzzification rules are applied to decide the degree of truth of a possible case. For example, the rule can be ”if feature 1 is big and feature 2 is big, then the output is big ”. To show the rule mathematically, it gives the degree of truth for the output to be ”big” as:

$$W_1 = P_{11} \times P_{21} \tag{3.10}$$

Where W is the degree of truth of a possible case.

Then the normalization layer is applied to make the degrees of truth comparable with each other, calculated as:

$$\bar{W_1} = \frac{W_1}{W_1 + W_2 ... + W_n} \tag{3.11}$$

Where $\bar{W_1}$ is the normalized degree of truth.

Fifthly, the output of this case is calculated by defuzzification function. It is usually decided by the input features, and a widely used function is the linear function, such as:

$$f(x_1, x_2 ...) = b_0 + b_1 \times x_1 + b_2 \times x_2 ... + b_n \times x_n \tag{3.12}$$

Where f is the output of a case; $b_0, b_1 ... b_n$ are the parameters. Finally, the case outputs are summed to give the final output as equation 3.13.

$$Output = \bar{W_1} \times f_1 + \bar{W_2} \times f_2 ... + \bar{W_n} \times f_n \tag{3.13}$$

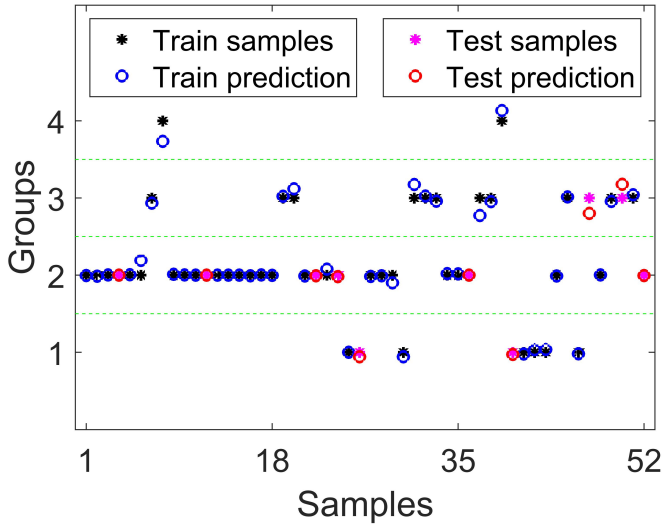


FIGURE 3.10: The diagnostic results by ANFIS

Therefore, the membrane resistance R_m and the GDL capacitance C_d can be applied as inputs to the ANFIS, and the output is the groups that the experiment samples belong to. The database is divided into the train part and test part randomly, and 20% of the data is set as test samples. The diagnostic result can be shown in figure 3.10.

The four fault groups are labelled as 1,2,3,4, respectively, and the experiment result and the predictions of ANFIS are represented as asterisks and circles, respectively. Also, the training samples and test samples are partitioned by different colours. As the ANFIS is a kind of continuous method, the predictions of ANFIS are also continuous. Therefore, there are some small deviations between the predicted output and the real group label. However, with the threshold given by the green lines between every two groups, all the samples can be assigned to the right groups, and both train and test predictions correspond well with the real conditions. Therefore, the proposed diagnosis method based on EIS and ANFIS is very accurate and effective.

A former research [177] also achieved PEMFC diagnosis based on EIS data. Six points under certain frequencies in the EIS data were chosen, and the position coordination of those points were applied as diagnostic features. The Bayesian network was applied to classify the faults according to the chosen features, and the flooding and drying-out faults were researched. The diagnostic accuracy is 91% in their research, so obviously the proposed method is superior to their method. The main advantage of the proposed method is that the ECM is applied to extract features from all the EIS data, rather than

only depending on several special points. In this way, the noises and stochastic errors of EIS measurement under different frequencies can be counteracted, and more useful information can be revealed compared to their method.

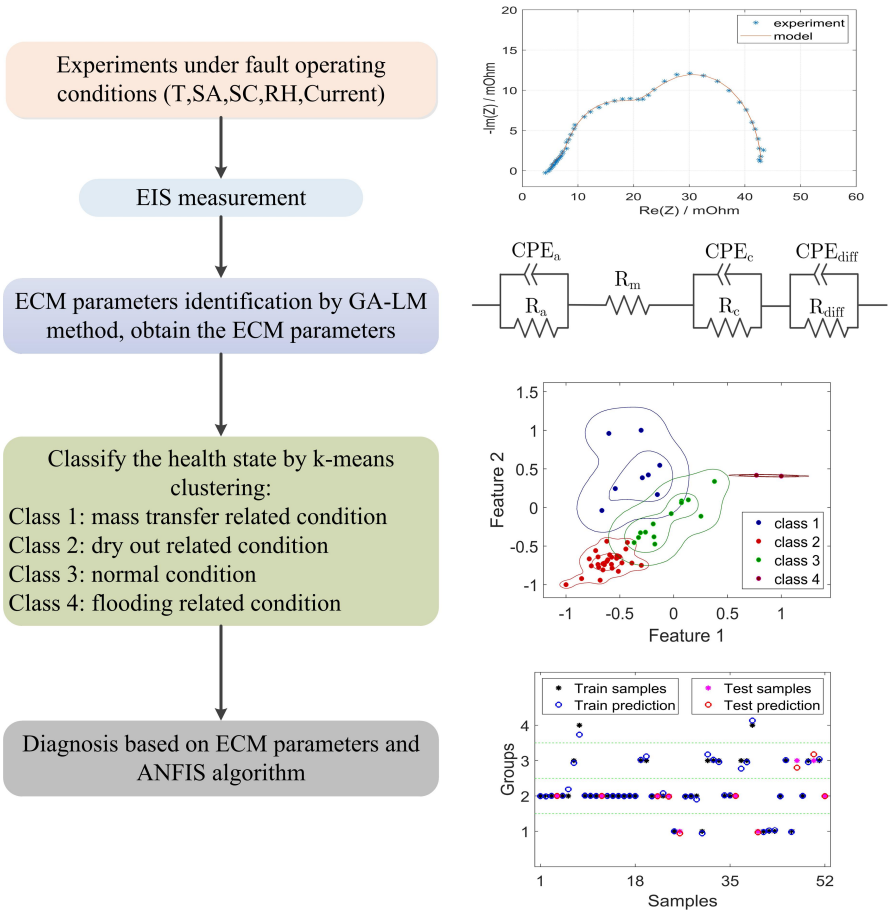


FIGURE 3.11: The diagnostic methodology by using EIS, ECM, k-means clustering, and ANFIS method

To summarize the diagnosis methodology, the overall processes are given in figure 3.11. It is clear that the method is full-fledged with all the processes of EIS measurement, ECM parameter identification, and diagnosis. As more and more researchers have put forward the methods to measure the EIS data on-line [38], the proposed diagnostic method is possible to be applied to on-line diagnosis by connecting with on-line EIS measurement.

Even though only discrete experimental conditions are studied in this research, it is also applicable to other conditions which are not too far from the studied cases. The biggest character of the proposed method is that the diagnostic features are extracted from EIS data by ECM, and the ECM parameters have physical meanings. Therefore, compared with those features without physical meaning or with unclear meanings, the change

of proposed ECM features is continuous with the change of operating parameters. Therefore, it can be interpolated and appropriately extrapolated. However, there is a limit for extrapolation, and it is not researched in this research. For the out-of-range cases, the most important thing is that the operating parameters should not exceed the design limitation of the PEMFC; otherwise the severe fault will totally change the behavior pattern of PEMFC. In this case, it is possible that the proposed method will lose efficacy. Even when the operating parameters are close to the design limitations, the proposed method should be tested first before being applied.

The applicable conditions of the method should be noticed. The proposed method is useful under stable conditions, as the information of all components can be detected by EIS and extracted by ECM. Also, the proposed method is robust to a certain level of noise, because the parameters are extracted using all measurement points rather than using a few points as in the reference [177], so the measurement noise and stochastic error can be counteracted by this method. As the noise of EIS measurement will disturb the health state information extraction, the diagnostic method may fail under severe noise. In terms of noise problems during EIS measurement, it is especially crucial that the measured EIS curve should be relatively smooth, proving that the inner state is stable under the current stimulation of different frequencies. To guarantee the noise is relatively small and negligible, one of the most important things is to keep the operating conditions within the design limitations. Otherwise, the behavior pattern of PEMFC may be quite unstable; thus the noise may take the dominating place and fail the EIS measurement.

3.2.5 Implementation methods on DSP system

3.2.5.1 Real-time implementation task for diagnosis

The proposed method is valid under laboratory conditions on a test bench. To apply it in real applications, an on-line EIS measurement-diagnosis-control framework can be proposed based on the research above. The procedures and objectives of the framework can be shown in figure 3.12. There are three stages. The first stage is on-line EIS measurement by integrating it with the power converter controller, which

has been achieved in previous work [38]. Therefore, the EIS can be measured on-board without extra cumbersome and expensive impedance measurement devices. The diagnosis works focus on stage 2, i.e. on-line diagnosis strategy implementation in industrial-level processors based on EIS data. With this implementation, the diagnostic decisions can be obtained on board quickly enough. Further, stage 3 should be achieved to control the PEMFC actuators based on the diagnostic decisions. As a result, the PEMFC can be adjusted in time when there are faults, and it can be operated with better performance. To connect the on-line EIS measurement and practical diagnosis-based control strategy, the proposed diagnosis strategy is further developed on an industrial-level DSP.

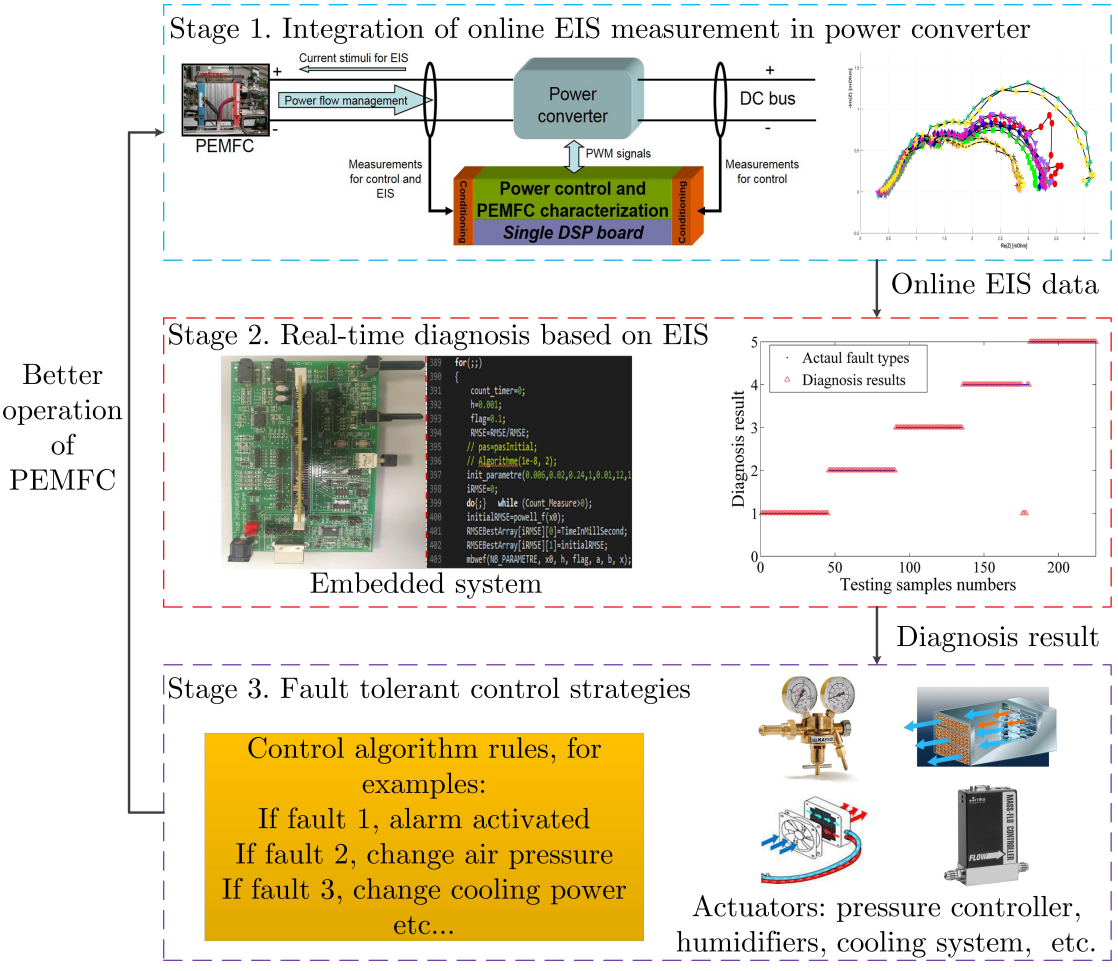


FIGURE 3.12: The framework to achieve on-line EIS measurement, diagnosis, and control

3.2.5.2 Overall processes for implementation

To achieve real-time diagnosis on DSP system, the main processes can be shown in figure 3.13. It can be described as follows:

1. As the emphasis is to implement the diagnosis algorithms, the EIS data is collected in advance and saved in the computer to facilitate it. To simulate the EIS measurement in a real situation [38], the PC can act as an emulator to provide EIS data in real time for diagnosis. The EIS data is transported to the DSP system in real time with a specific transport protocol using the DSP serial communication interface (SCI).

2. The diagnosis processes are all carried out in a DSP. For the first phase, the ECM parameters can be identified according to the EIS data. As the computational resource and memory are limited in the DSP system, Powell's algorithm is applied as it is simpler and more practical.

3. For the second phase, the diagnosis is implemented based on ANFIS by taking ECM parameters as features, and there are two periods, i.e. the adaptive training period and the test period. Both of them are achieved on-line in DSP, and the predicted results are compared with the real classes to evaluate the accuracy.

3.2.5.3 EIS Data measurement and transmission

To carry out EIS measurement, a small magnitude sinusoidal current is superimposed on the PEMFC system, and the voltage response is recorded to calculate the impedance. The impedance can be obtained as equation 3.14 [38].

$$Z_{FC}(f) = \frac{V_z(f)}{I_z(f)} \tag{3.14}$$

Where the $Z_{FC}(f)$ is the PEMFC impedance at frequency f; $V_Z(f)$ is the AC components of voltage response to the frequency f; $I_Z(f)$ is the AC components of the current at frequency f. However, as it is difficult to measure the high frequency current and voltage with high precision, most of the EIS are measured by a cumbersome

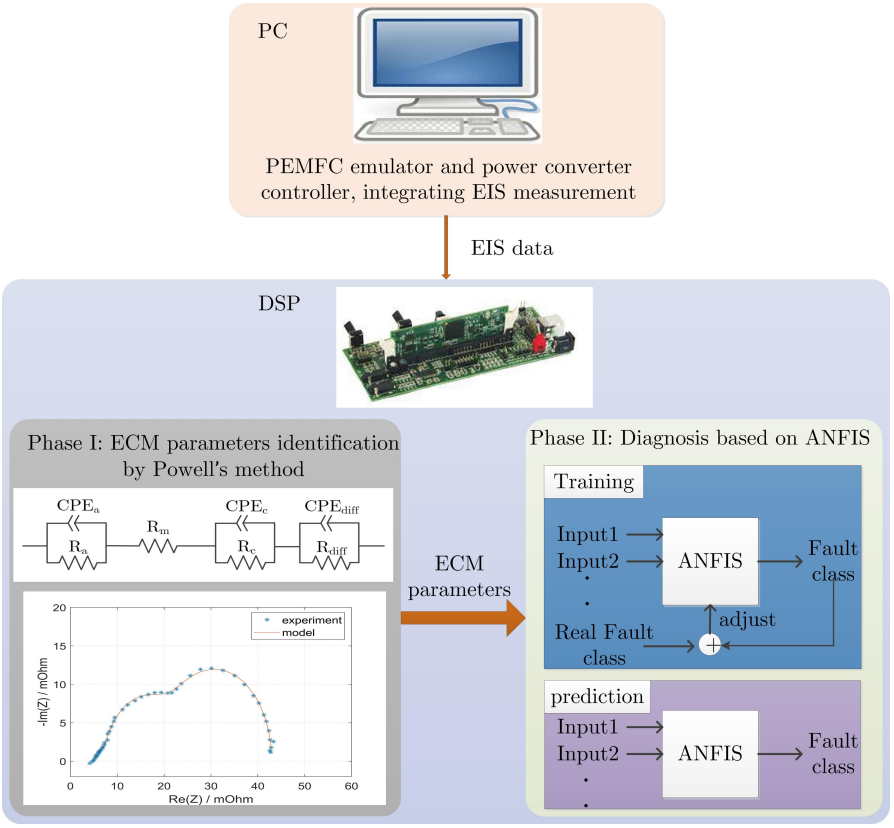


FIGURE 3.13: The overall diagnostic processes on DSP system

equipment, i.e. impedance spectrometer. As the equipment is too heavy and too big, and also because it is too expensive, it is impossible to measure EIS on board for vehicular applications.

However, the on-line measurement of EIS has been researched in former work [38]. The superimposed current signal and voltage detection can be achieved by a power converter so that no other extra equipment is needed. As shown in figure 3.12, the EIS measurement signal can be performed in parallel with the PEMFC power control, thus multi-targets can be achieved at the same time on a single DSP board. As we focus on the implementation of diagnosis in this research, the measurement process is simulated by a PC and not repeated. More details about the on-line detection of EIS data by the power converter can be found in former research.

The impedance can be represented by a complex number for each frequency, i.e. a real part and an imaginary part. In this research, 51 points of different frequencies are measured for each condition, and the frequency ranges from 9996 Hz to 0.0978 Hz. It should be noticed that our experiment data is not measured by the same system as former research [38], and the frequency range is wider.

The DSP system is the TI (Texas Instruments) Peripheral Explorer kit with TMS320F28335 FPU (Floating-Point Unit), and the structure can be shown in figure 3.14. The maximum clock frequency of the DSP is 150 MHz, and the memory is 256k for 16 bits flash and 32k for 16 bits single-access RAM. The SCI pins are applied to receive EIS data, combined with a specific interrupt service routine.

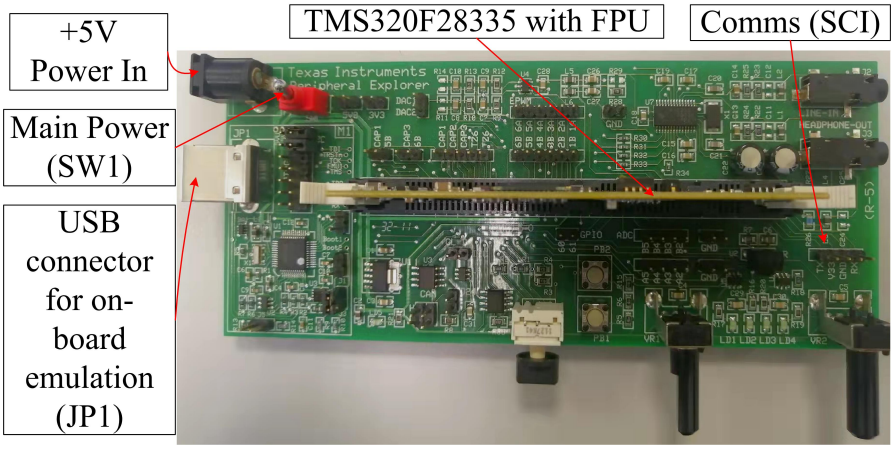


FIGURE 3.14: Structure of applied DSP board

The impedance data is transported from the computer to the DSP system by RS232 protocol in this research, and the transmission is carried out once for each group of impedance data (51 points), including the frequencies, real parts, and imaginary parts. The transmission time is very short compared to the sampling time, and the reception of data in the DSP system is carried out by continuous scanning during a certain waiting time, thus avoiding data leaks. The method to acquire data is given as algorithm 3.2.

Algorithm 3.2: EIS data transmission

Computer side (Sending):	DSP side (Receiving):
for each operation condition **do**	**for** each operation condition **do**
collect impedance of different frequencies;	wait and count time;
if all frequencies have been detected; **do**	**if** preset waiting time is up; **do**
send data to serial port;	scan serial port to receive data;
end for	**end for**

3.2.5.4 Fractional-order equivalent circuit model for DSP

The ECM applied in the DSP is also the 3-arc model as shown in figure 3.1(c). The calculation of the complex impedance on the DSP board should be separated into real parts and imaginary parts. Taking the cathode impedance as an example, the real part

and imaginary part can be calculated by equation 3.15 and 3.16, respectively. The details of the derivation can be found in Appendix A.

$$Z_{c-Re}(f) = \frac{R_c + R_c^2 C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2}}{(1 + R_c C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2})^2 + (R_c C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2})^2} \tag{3.15}$$

$$Z_{c-Im}(f) = \frac{-R_c^2 C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2}}{(1 + R_c C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2})^2 + (R_c C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2})^2} \tag{3.16}$$

Therefore, the total impedance of PEMFC can be separately given as real part and imaginary part as:

$$\begin{aligned} Z_{FC-Re}(f) =& Rm + \frac{R_a + R_a^2 C_a (2\pi f)^{\alpha_a} \cos\frac{\pi\alpha_a}{2}}{(1 + R_a C_a (2\pi f)^{\alpha_a} \cos\frac{\pi\alpha_a}{2})^2 + (R_a C_a (2\pi f)^{\alpha_a} \sin\frac{\pi\alpha_a}{2})^2} \\ &+ \frac{R_c + R_c^2 C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2}}{(1 + R_c C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2})^2 + (R_c C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2})^2} \\ &+ \frac{R_d + R_d^2 C_d (2\pi f)^{\alpha_d} \cos\frac{\pi\alpha_d}{2}}{(1 + R_d C_d (2\pi f)^{\alpha_d} \cos\frac{\pi\alpha_d}{2})^2 + (R_d C_d (2\pi f)^{\alpha_d} \sin\frac{\pi\alpha_d}{2})^2} \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} Z_{FC-Im}(f) =& - \frac{R_a^2 C_a (2\pi f)^{\alpha_a} \sin\frac{\pi\alpha_a}{2}}{(1 + R_a C_a (2\pi f)^{\alpha_a} \cos\frac{\pi\alpha_a}{2})^2 + (R_a C_a (2\pi f)^{\alpha_a} \sin\frac{\pi\alpha_a}{2})^2} \\ &- \frac{R_c^2 C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2}}{(1 + R_c C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2})^2 + (R_c C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2})^2} \\ &- \frac{R_d^2 C_d (2\pi f)^{\alpha_d} \sin\frac{\pi\alpha_d}{2}}{(1 + R_d C_d (2\pi f)^{\alpha_d} \cos\frac{\pi\alpha_d}{2})^2 + (R_d C_d (2\pi f)^{\alpha_d} \sin\frac{\pi\alpha_d}{2})^2} \end{aligned} \tag{3.18}$$

3.2.5.5 ECM parameters identification on DSP board

A lot of optimization methods are based on the idea of gradient descent method, such as the Levenberg–Marquardt algorithm in former research. However, as shown in the equation 3.17 and 3.18, the objective function is very complicated in our research, and the gradient calculation will be too computationally intensive and time-consuming, especially for a DSP system with limited computation resources. To overcome this problem, Powell's algorithm has been implemented in the DSP system [186]. However, the initial value is also important for Powell's method as for all other local search methods, so the GA is applied first to obtain the rough global solution. Since the EIS is

similar under different conditions for the same PEMFC, the initial solution only needs to be provided once, so the GA solution can be found off-line before diagnosis.

Powell's method is a kind of direct search method, where the conjugate direction is applied to accelerate the search speed. First of all, for an optimization problem with n parameters, n linearly independent search directions can be given as $\{\mathbf{d}_0, \mathbf{d}_1 ..., \mathbf{d}_{n-1}\}$. Give an initial solution as $\mathbf{x}_0$ and learn rate λ, take $\mathbf{y}_0 = \mathbf{x}_0$, then search for the solution of minimum objective function for each direction. i.e.

$$f(\mathbf{y}_{j-1} + \lambda \mathbf{d}_{j-1}) = min(f(\mathbf{y}_{j-1} + \lambda \mathbf{d}_{j-1})), \quad for \quad j = 1, 2, ..., n \tag{3.19}$$

$$\mathbf{y}_j = \mathbf{y}_{j-1} + \lambda \mathbf{d}_{j-1} \tag{3.20}$$

For each one-direction search, the golden-section search method is applied. The basic idea is to cut the searching region into three segments according to the golden ratio, and abandon one of the segments by comparing the objective function value of the cutting points. The searching region is shortened until the permissible error is reached; thus obtaining the solution with the smallest objective function.

Secondly, the acceleration search can be achieved by changing the search direction. Taking

$$\mathbf{d}_n = \mathbf{y}_n - \mathbf{y}_0 \tag{3.21}$$

If $\mathbf{d}_n < Err$, where Err is the permissible error of the solution, then this is a feasible solution and the algorithm ends. Otherwise, find the new learning rate by looking for the best solution in the direction $\mathbf{d}_n$ from $\mathbf{y}_n$, so that

$$f(\mathbf{y}_n + \lambda \mathbf{d}_n) = min(f(\mathbf{y}_n + \lambda \mathbf{d}_n)), \quad for \quad \lambda = \lambda, 2\lambda, 3\lambda ... \tag{3.22}$$

Then the new start point can be set as

$$\mathbf{x}_{k+1} = \mathbf{y}_n + \lambda \mathbf{d}_n \tag{3.23}$$

Where the k is the iteration index. And the new search directions can be obtained by abandoning the first direction and adding the new direction, i.e.

$$\mathbf{d}_j = \mathbf{d}_{j+1}, \quad for \quad j = 0, 1, ..., n-1 \tag{3.24}$$

With the new start point and new search directions, the solution moves toward the minimum without calculating the gradient. However, the search directions should keep linearly independent. The linear independence may be broken because of degradation or ill condition during the search; thus the solution cannot converge as it only searches in lower dimension space.

To solve the defect of the original method, a correcting algorithm is added to check the linear independence and correct it when the linear independence is not satisfied. The direction which has the biggest contribution to the objective function can be given as:

$$\begin{aligned} \Delta_m &= max\{f(\mathbf{x}_k^{(j)}) - f(\mathbf{x}_k^{(j+1)})\}, \quad for \quad j = 0, 1, ..., n-1 \\ &= f(\mathbf{x}_k^{(m)}) - f(\mathbf{x}_k^{(m+1)}) \end{aligned} \tag{3.25}$$

And the linear independence can be ensured if

$$\begin{aligned} f_3 < f_1, \quad and \\ (f_1 - 2f_2 + f_3)(f_1 - f_2 - \Delta_m)^2 < 0.5\Delta_m(f_1 - f_3)^2 \end{aligned} \tag{3.26}$$

Where

$$\begin{aligned} f_1 &= f(\mathbf{x}_k^{(0)}) \\ f_2 &= f(\mathbf{x}_k^{(n)}) \\ f_3 &= f(\mathbf{x}_k^{(n)} + \lambda \mathbf{d}_n) \end{aligned} \tag{3.27}$$

If equation 3.26 is true, then the direction $\mathbf{d}_m$ can be abandoned and add $\mathbf{d}_n$ to the research direction, i.e.

$$\mathbf{d}_j = \mathbf{d}_{j+1}, \quad for \quad j = m, m+1, ..., n-1 \tag{3.28}$$

Otherwise, if the equation 3.26 is not held, the search direction will stay the same as the last iteration until the permissible error is reached. The total Powell's algorithm can be summarized in figure 3.15.

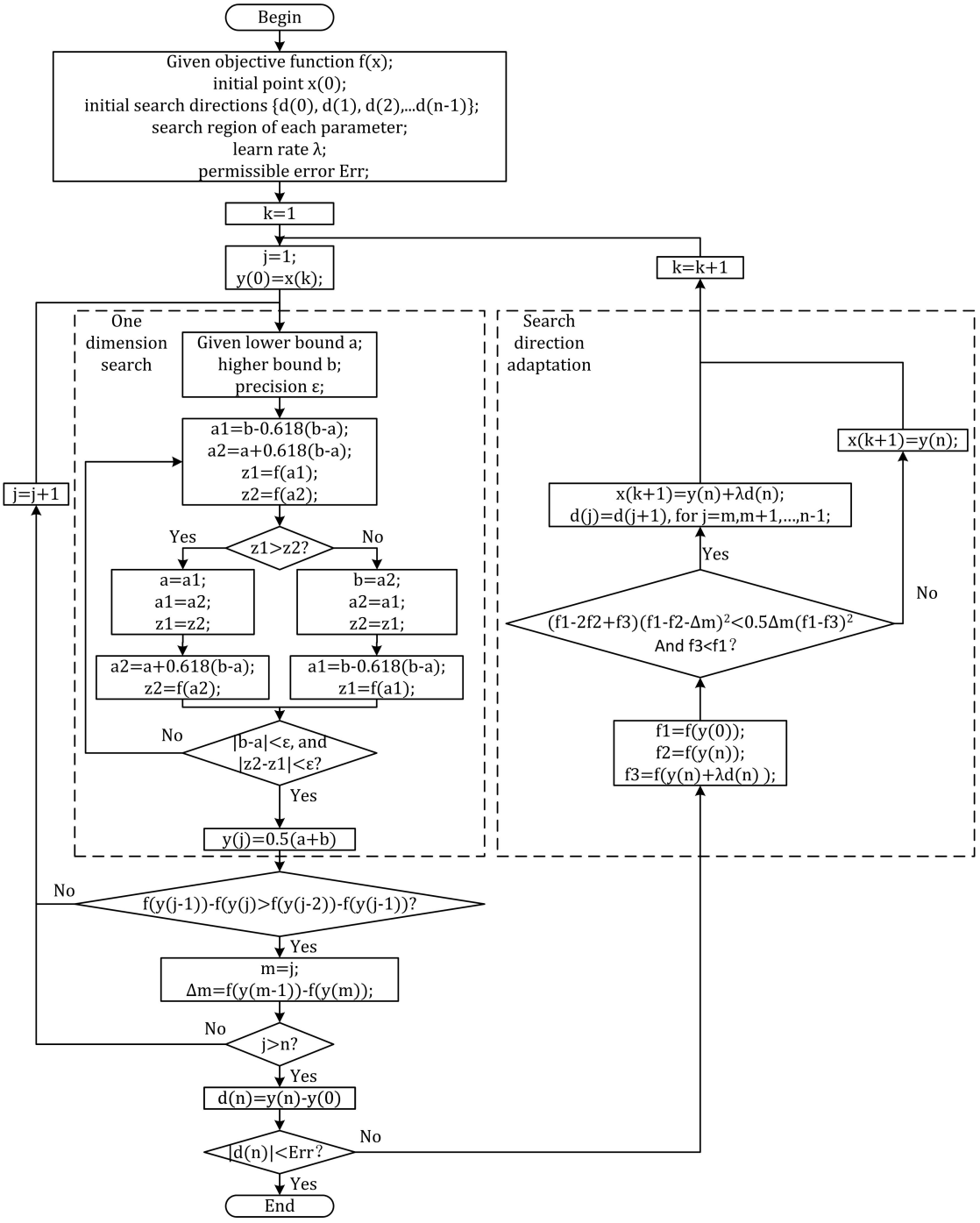


FIGURE 3.15: Powell method for parameter identification

3.2.5.6 Diagnosis by ANFIS on DSP

As proved above, the identified ECM parameters can be applied as diagnostic features in ANFIS. ANFIS needs to be adaptively updated on the DSP during operation, so the tuning process needs to be explained in more detail.

Tuning is necessary to find the proper parameters of ANFIS, and the back-propagation algorithm is applied in this research. The basic idea is to calculate the error of the result, then propagate the error from the output layer to the input layer by the chain rule. The parameters can be adjusted according to the derivative to minimize the output error.

For example, the error of output is:

$$e = (\hat{y} - y)^2 \tag{3.29}$$

Then the derivative is:

$$\delta\hat{y} = \frac{\partial e}{\partial \hat{y}} = 2(\hat{y} - y) \tag{3.30}$$

And the error can be passed to the previous layer as:

$$\delta f_n = \frac{\partial e}{\partial \hat{f_n}} = \frac{\partial e}{\partial \hat{y}} \frac{\partial \hat{y}}{f_n} = \delta\hat{y}\bar{W}_n \tag{3.31}$$

and

$$\delta\bar{W}_n = \frac{\partial e}{\partial \bar{W_n}} = \frac{\partial e}{\partial \hat{y}} \frac{\partial \hat{y}}{\bar{W}_n} = \delta\hat{y} f_n \tag{3.32}$$

Once again, the partial derivative of parameters $b_0, b_1, b_2, ...b_n$ can be calculated as:

$$\delta b_0 = \frac{\partial e}{\partial \bar{b_0}} = \frac{\partial e}{\partial f_n} \frac{\partial f_n}{b_0} = \delta f_n \times 1 \tag{3.33}$$

$$\delta b_n = \frac{\partial e}{\partial \bar{b_n}} = \frac{\partial e}{\partial f_n} \frac{\partial f_n}{b_n} = \delta f_n x_n \tag{3.34}$$

The partial derivative of other parameters can also be calculated by the chain rule. And the parameters can be tuned by:

$$p_{k+1} = p_k + \lambda \delta p \tag{3.35}$$

Where p represents the tunable parameters, i.e. the a, b, c in equation 3.9 and the $b_0, b_1, ..., b_n$ in equation 3.12. k is the iteration index, and λ is the learn rate.

Therefore, in each tuning iteration there is one forward-propagating process to calculate the variables and output, and one back-propagation process to calculate the partial derivative of each parameter. The tuning process continues until the output error reaches the set precision or when the set time is up, as given in algorithm 3.3.

Algorithm 3.3: ANFIS training algorithm

give ANFIS structure;
do
  calculate each variable and ANFIS output by forward propagation;
  calculate the partial derivatives of each variable by back propagation;
  update each tunable parameter according to the partial derivative;
Until output error $<$ setting precision **or** training time $>$ setting time
End

3.2.6 Implementation results on DSP

3.2.6.1 Parameter identification result

The ECM parameters are identified according to the proposed GA-Powell's method. The MSE can be applied as the accuracy judgement for the parameter identification process, and the convergence process for experiment No. 1 can be shown in figure 3.16. It can be seen that the MSE reaches 3.232×10^{-5} at 18.79 seconds, and it is 3.147×10^{-5} at 69.59 seconds. Therefore, the identification precision improves with time, but the magnitude of improvement is smaller and smaller after a certain time. Actually, the ideal solution can be found within 19 seconds as the identification result has no big difference after 19 seconds. There are plateau as a better solution is not found during the period.

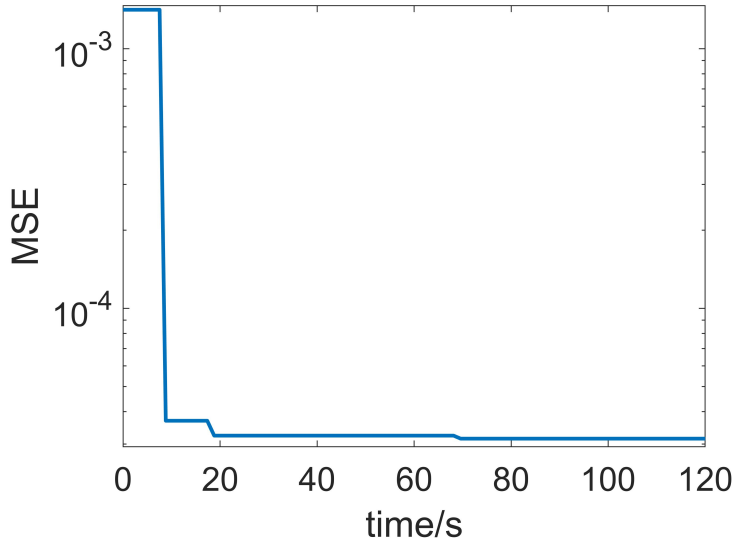


FIGURE 3.16: The convergence process of parameter identification by Powell's method

As shown in table 3.5, the MSE of the proposed method is compared with other methods above [187], including particle swarm optimization (PSO), simulated annealing algorithm (SAA), and the combination of GA and Levenberg–Marquardt algorithm (GA-LM). According to the MSE, the proposed method is more precise than PSO and SAA methods, but the MSE of the proposed method is slightly bigger than the GA-LM method. However, it is acceptable as the difference between them is minor. On the contrary, the LM method is based on the gradient, and the computational burden is too heavy for our DSP system. The LM method is also tried in the DSP system, but no acceptable results can be given within a limited time. Therefore, the proposed GA-Powell's method is accurate and suitable for parameter identification on the DSP.

TABLE 3.5: The comparison of errors between different parameter identification methods

Methods	error
PSO	4.228×10^{-5}
SAA	1.384×10^{-4}
GA-LM	2.664×10^{-5}
GA-Powell	3.147×10^{-5}

3.2.6.2 Diagnosis results on DSP

As was proved in our former research, the identified membrane resistance R_m and the GDL capacitance C_d can be applied as features to the ANFIS, and the output is the fault classes that the experiment samples belong to. The experimental samples can be divided into two parts: training part and the test part. The first 50% of the data is set as training samples, while the last 50% of the samples are set as tests. For training part, the ECM parameters and the fault classes are all provided to the ANFIS, so that the ANFIS can be tuned. In the test part, only the ECM parameters are inputs and the fault classes are predicted.

The ANFIS is adaptively tuned, i.e. after each prediction, the real fault result is used as a training reference to adjust the ANFIS parameters. Therefore, the ANFIS model can be on-line tuned by more samples and adapt to the inner state change in the long term. The ANFIS convergence process can be shown in figure 3.17. With the new training sample, the MSE of the ANFIS prediction result can drop to 8×10^{-5} with 15 seconds of tuning. Therefore, the ANFIS can converge to an acceptable level within a short time.

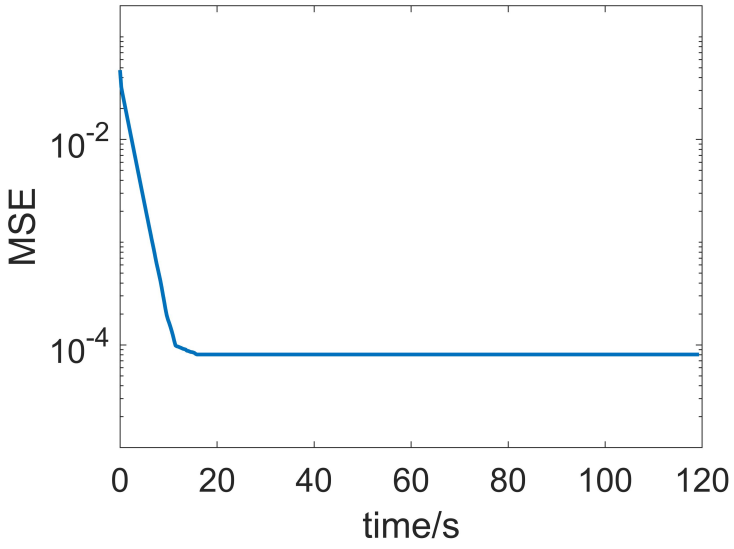


FIGURE 3.17: The MSE convergence during ANFIS tuning process on DSP

The fault classes of testing experiment samples are predicted by the ANFIS, and the comparison between the predicted fault class and the real fault class can be shown in figure 3.18. The experiment results are represented by asterisks, while the model predictions are represented by circles. It can be seen that all the predictions are close

to the experiment result. With the threshold given by the green lines between every two groups, all the samples can be assigned to the right classes.

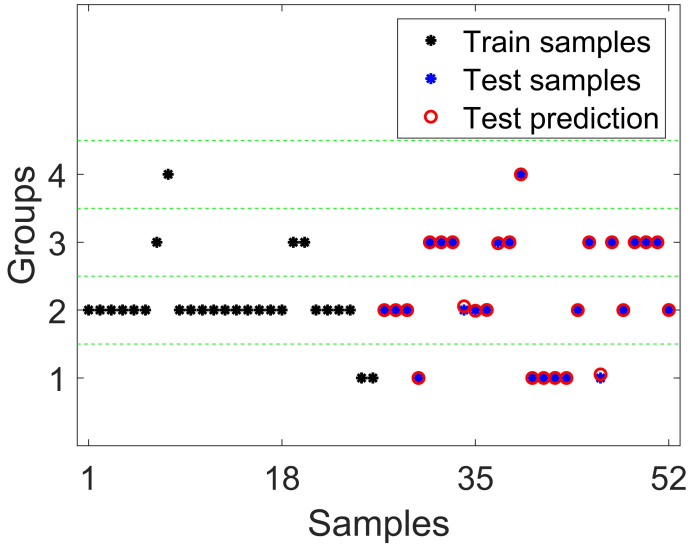


FIGURE 3.18: The ANFIS diagnosis results by DSP board

Compared to research above [187], a great improvement is that the ANFIS can be adaptively tuned in DSP; thus it can adapt to the EIS data when the inner state of PEMFC has changed. For example, if the PEMFC degrades after a long-term operation, the range of ECM parameters may transform even when operating parameters are the same. Therefore, the original ANFIS cannot predict the right fault classes in this situation, while the ANFIS on DSP can still work if the ANFIS parameters can be tuned according to the recent fault data.

3.2.6.3 Real-time ability

The most important advantage of the proposed method is the real-time ability. As all the diagnosis processes are achieved in real time on a DSP board, the proposed method can be applied to industrial applications. The EIS obtain process is simplified by the pre-saved experimental data on the computer, but the time for real EIS measurement is respected in the research. According to the previous research [38], the total measurement time for each EIS curve can be calculated according to equation 3.36.

$$T = \sum_{i=1}^{N_p} \frac{N_c}{f_i} \tag{3.36}$$

Where the T is the total time cost for each curve; N_p is the number of measurement points for each EIS curve, and it is 51 in our research; N_c is the number of sine periods for each frequency, and it is set 5 according to reference [38]; f_i is the frequency of the i_{th} point. Therefore, taking the parameters into the equation, the measurement time for each curve is 250 seconds.

The time cost for each process is shown in figure 3.19. All the diagnostic algorithms are carried out on the DSP board. After the EIS measurement, the data transportation to the DSP board only needs 0.01 seconds, which is very short and negligible. During the EIS measurement, the diagnosis on the DSP board works simultaneously. The diagnosis time can be cut into 3 segments during the EIS measurement process. One hundred and twenty seconds is reserved for Powell's method to identify the ECM parameters, and another 120 seconds is reserved to tune the ANFIS parameters and give the fault type prediction. There is also some waiting time to guarantee that the DSP is ready to receive the data when the measurement is done and the data is sent, which is flexible because some other minor programs and computations in DSP also take some time. All the clocks are controlled by the DSP interruption program.

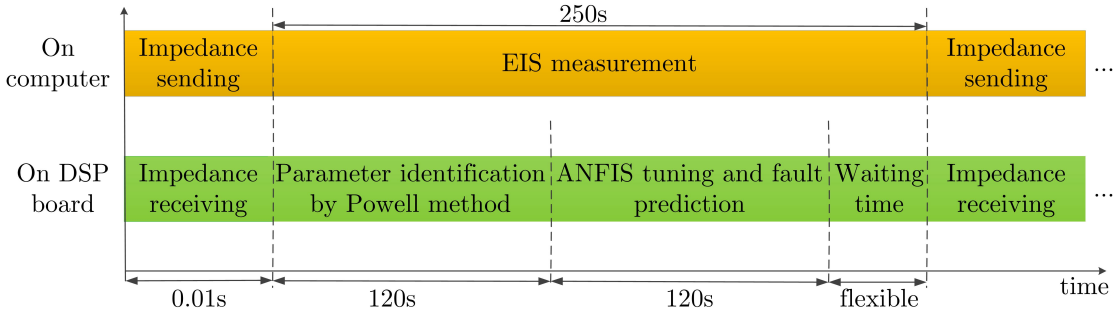


FIGURE 3.19: The time cost for every process in real-time diagnosis

As was discussed in reference [188], the EIS data collection time is much shorter compared to the polarization curve method, so the proposed diagnosis methodology can be a more practical tool than polarization-based diagnosis. Indeed, the diagnostic time based on EIS is relatively longer than the diagnosis based on single-point time signals such as voltage or pressure magnitude signals. However, as EIS is based on physical mechanisms and is more interpretable, there is more useful inner health state

information in EIS data. Also, as the development of drying out and flooding faults usually takes several minutes, the diagnosis delay by our method is acceptable.

The main limit is the measurement, not the diagnostic algorithm. More and more new EIS measurement technology are also proposed, such as using the pseudo-random binary sequence (PRBS) as a perturbation signal to detect EIS information [189] rather than the traditional sinusoidal signal. With the development of measurement technology, if the measurement noise is small enough, it is also possible that the number of sine periods in equation 3.40 can be reduced. In addition, the time measurement can be significantly reduced if less frequency points are measured for each EIS curve. For example, only 18 seconds is required when 28 frequency points are measured in reference [38]. If the frequencies are well selected, the ECM features can still be extracted by our method when the measurement points are reduced. Within the proposed framework, the diagnosis performance can be improved with both the development of EIS detection technology and diagnostic algorithm.

3.2.7 Conclusions

In this section, a fractional-order ECM model is applied to analyse the EIS result of the PEMFC system, and a novel diagnostic method based on ANFIS is put forward. Further, the proposed diagnostic method is implemented and validated in real time on a DSP system. The applied ECM considers the impedance of the anode, cathode, membrane resistance, and mass diffusion, and it corresponds well with experiment EIS. The EIS behaviours can be well understood in the physical aspect. Also, the proposed parameter identification method is very effective. As the current increases, the impedances of the anode, cathode, and mass diffusion resistance are quite different. At the same time, the properties of different parts will change under fault conditions, and they can be used as indicators to alarm the fault development. The fault caused by membrane flooding, drying-out, and mass diffusion problem can be revealed by EIS. Based on the inner relationship between the fault causes and ECM parameters, the proposed diagnosis method based on ANFIS is very accurate to identify the fault causes of different experimental conditions. Also, compared to the former research [177], this method is more accurate. By using the ECM and Powell's algorithm, the EIS features can be extracted in real time on the DSP system conveniently. Powell's

algorithm is very effective to identify the ECM parameters, and it is very suitable for systems with limited computing resources. The proposed diagnosis method can be successfully implemented in the DSP system in real time. The whole process can be achieved within 250 seconds. It is a meaningful step from laboratory research toward real applications for EIS-based diagnosis technology.

3.3 Real-time diagnosis based on quick detective EIS features

3.3.1 Introduction

In former research, as the EIS measurement should be carried out under a wide frequency range, it takes too much time for obtaining a full spectroscope curve. For example, a typical EIS curve covers the frequency range from 0.1 mHz to 100 kHz [190]. If 5 periods are measured for each frequency [38], the measurement process will take more than 10 minutes for a full curve [190], which limits its application. During the measurement, the PEMFC inner state may have changed, and the uncertainty increases compared to other quick detection parameters. Also, even though the inner state can be kept static during each measurement, the measurement is too time-consuming to obtain a lot of data for quantitative diagnosis. Only one research [177] achieved a quantitative diagnosis for PEMFC based on EIS. Six points were chosen equidistantly from the full EIS curve, and the coordinates of the points were applied as features. However, the diagnostic significance of the chosen points is not clear, and the diagnosis accuracy is only 91%. Our former research also applied the equivalent circuit model (ECM) to extract features from the whole EIS curve, and the diagnostic accuracy is high. However, it still takes a considerable time for the diagnosis, which limits its application. Therefore, to apply EIS technology to real diagnostic applications, an accurate but also quick quantitative diagnosis methodology is necessary.

To achieve a quick diagnosis based on EIS, two novel features are proposed, which can be quickly measured and applied to a quantitative diagnosis. The features include the zero-phase impedance and the turning phase. The zero-phase impedance comes from high-frequency resistance (HFR), which represents the magnitude of the impedance at

high frequency when the impedance phase is 0, i.e. the left cross point of the EIS curve and real axis in the Nyquist plot [31]. A lot of research works have proved that it is an effective indicator of the inner membrane humidity; consequently it can be applied to the diagnosis of flooding and drying out. However, it is difficult to detect the exact HFR whose phase is zero, because the corresponding frequency is not certain in different conditions. Therefore the concept of zero-phase impedance is raised to calculate the HFR by an average of impedances in the range of small phase zone.

Another feature is the turning phase. In former research, the maximum absolute phase has been applied to indicate the PEMFC health state [115] [142]. However, the frequency corresponding to the maximum absolute phase is not constant under different conditions, so it is not easy to find it quickly. As an alternative, the turning phase can be detected under a certain frequency. Therefore, there is no need to measure and compare a lot of EIS points to obtain the maximum phase.

As a result, the most important health state information can be indicated by only measuring the impedance under a few frequencies. Therefore, the measuring time can be greatly reduced and on-line diagnosis is possible. In this section, a quick quantitative diagnosis is achieved and validated based on the two EIS features proposed above. The novel EIS features are explained and proved effective. The dependence of zero-phase impedance and turning phase on the PEMFC health state is proved. The method to detect them is also given, so that the EIS features can be quickly obtained and applied to diagnosis. The whole methodology is experimentally demonstrated. The effects of different operating conditions are considered in extensive experiments. A novel PEMFC diagnosis method is proposed, which is based on the proposed quick detective EIS features and KNN. The feature space can be separated into different fault zones, and the diagnosis is very accurate. The proposed method is also proved to be superior to the former method. Further, the procedure to generalize this method to other PEMFCs is also given, and the diagnostic algorithm is validated in real time on a DSP system. The diagnostic result is accurate and the diagnostic time is short, thus proving that the proposed method is practical.

The methods applied in this research are explained in section 3.3.2, including features analysis, diagnostic methods, and the implementation method on the DSP system.

Then the diagnostic results are given in section 3.3.3. Finally, the conclusions about the proposed method are given in section 3.3.4.

3.3.2 Methodology

The overall diagnostic process is explained in this section. The methodology includes two main parts. The first part deals with EIS features extraction, which is composed of the zero-phase impedance and the turning phase. The relationship between the features and operating conditions is also analysed and validated. The second part deals with the KNN diagnosis method based on the proposed features, where the details of the method are explained. Also, the overall processes and the DSP implementation method are presented.

3.3.2.1 Zero-phase impedance

The concept of zero-phase impedance is firstly raised in reference [31], and it is an expansion and actualization of the widely used high-frequency resistance (HFR), i.e. the cross point of the EIS curve and the real axis in Nyquist plot [87]. At this point, the impedance has a zero phase, and the PEMFC acts like a pure resistance, thus the magnitude of the resistance is called HFR. It is wildly recognized that the HFR is a good indicator of membrane water content because it can represent membrane resistance. The HFR is high when the membrane is dry and it is lower under flooding fault, so it can be applied to the diagnosis of drying out and flooding fault [191]. However, it is not an easy task to accurately detect the HFR whose phase is exactly zero, because the frequency that corresponds to the HFR always changes under different operating conditions. Therefore, the zero-phase impedance is proposed to solve the measurement problem, which is more practical.

The basic idea is to take the average of the impedances in the zero-phase zone ($\pm 3^{\circ}$) as zero-phase impedance. The zero-phase zone can be shown in figure 3.20. The scatter points represent the impedance measurements under different currents, i.e. experiments No. 3 to No. 8 in table 3.2. The model result is based on an ECM in our former research [187], which is employed to fit the measurement points and smooth

the curve. The frequency increases from the right side to the left side. At the high-frequency zone, the curve crosses the real axis, and the impedance magnitude of the cross point is called HFR. However, the measurement points are scattered, and they still have imaginary parts, so none of them correspond exactly to the real axis. Therefore, it can be seen that a zero-phase zone is marked, where the impedance phase is within $-3°$ and $+3°$. In this zone, the impedance measurement points are close to the real axis, and their average can be applied to replace HFR, i.e. zero-phase impedance.

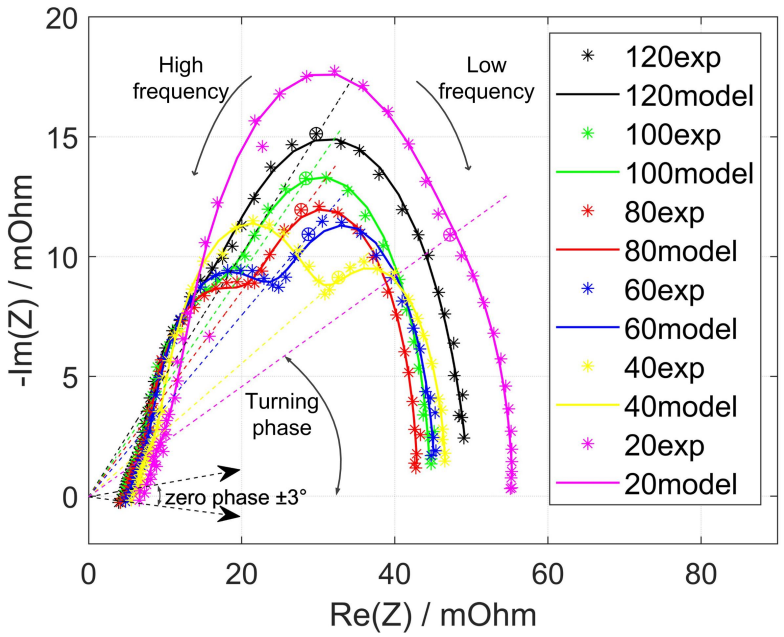


FIGURE 3.20: The zero-phase impedance and turning phase in Nyquist plot under different currents

The impedance magnitude and phase can be better shown in the Bode plot, as their development with frequency can be clearly given respectively. The magnitude of zero-phase samples can be shown in figure 3.21. In general, the impedances of zero-phase samples increase when the current decrease, which is related to the water content in the PEMFC. As the current decreases, the produced water also decreases, and the membrane is drier. As the proton conductivity of the membrane decreases, the PEMFC resistance increases.

The zero-phase impedance can be obtained by the average of the samples in the zero-phase zone, as equation 3.37.

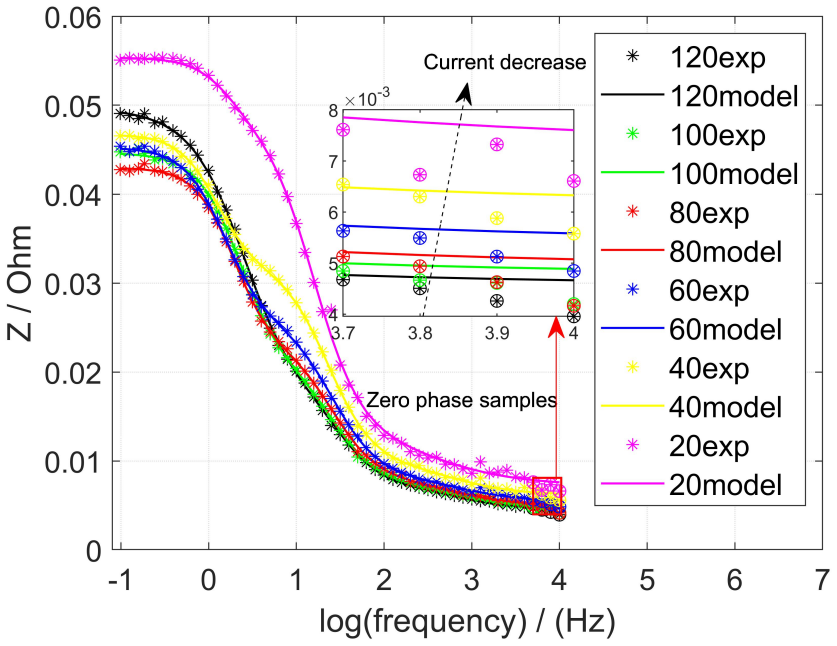


FIGURE 3.21: The zero-phase impedance samples in Bode plot under different currents

$$Z_{zp} = \frac{1}{n}\sum_{i=1}^{n} Z_i \tag{3.37}$$

Where the Z_{zp} is the zero-phase impedance; n is the sample number in the zero-phase zone; i is the index of the sample; Z_i is the impedance of the sample. In this way, the zero-phase impedance can be obtained quickly, also the random error of measurement can be reduced by the average.

According to equation 3.37, the calculated zero-phase impedance under different currents can be shown in figure 3.22. It is clear that the zero-phase impedance increases when the current decrease, and it changes monotonously. The zero-phase impedance changes rapidly with the drying process compared to the flooding process, showing that the zero-phase impedance is more sensitive to the drying process. Therefore, the experiments prove that the zero-phase impedance can be applied as an indicator of the membrane water situation.

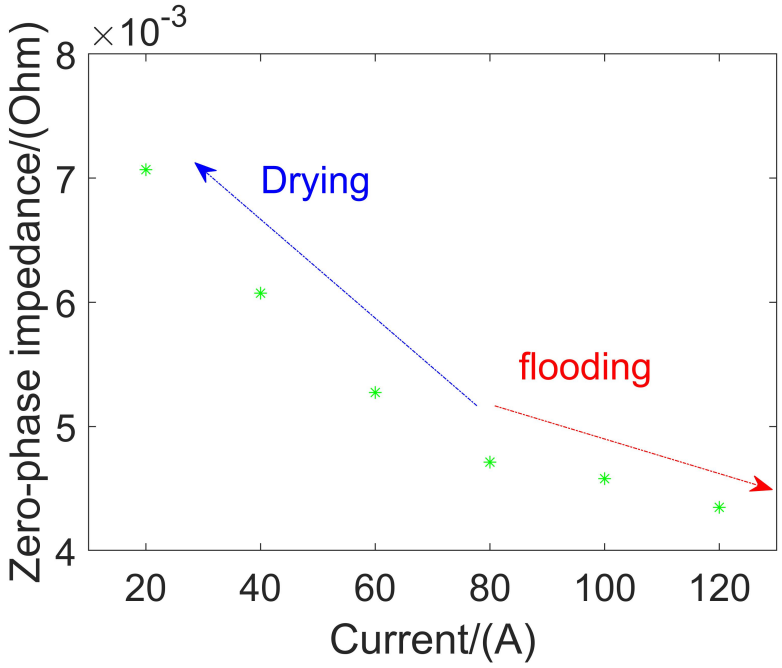


FIGURE 3.22: The zero-phase impedance under different current conditions

3.3.2.2 Turning phase

The second EIS feature is the turning phase. This impedance phase is an analogy and development of the maximum absolute phase proposed by reference [142]. The maximum absolute phase is an important factor that can represent the angle of the biggest semicircle, thus it can indicate the health state of the process that the semicircle represents. In most cases, this maximum phase appears at the peak of the cathode impedance semicircle, which represents the reaction resistance in the cathode. However, a similar problem as HFR also exists for the maximum absolute phase, i.e. the frequency that corresponds to the maximum phase changes under different conditions. Therefore, the whole EIS curve should be measured to find the maximum phase. To overcome this problem, the turning phase is proposed.

As shown in figure 3.20, the turning phase is the impedance phase that the curve turns from one main semicircle to another semicircle, and it can be better demonstrated by Bode plot in figure 3.23. This is a pivotal moment where the impedance phase enters a plateau in most cases at this frequency. A simple method of marking this frequency is to focus on the normal condition. The second peak phase point is exactly the turning phase, so the phase under this turning frequency can be identified as the turning phase for all cases. Compared with the maximum phase, the turning phase is

generally the second peak. It can be measured with a certain frequency; therefore it can be a convenient indicator. It locates at a lower frequency where the frequency is 2.496 Hz, which connects the cathode semicircle and mass diffusion semicircle.

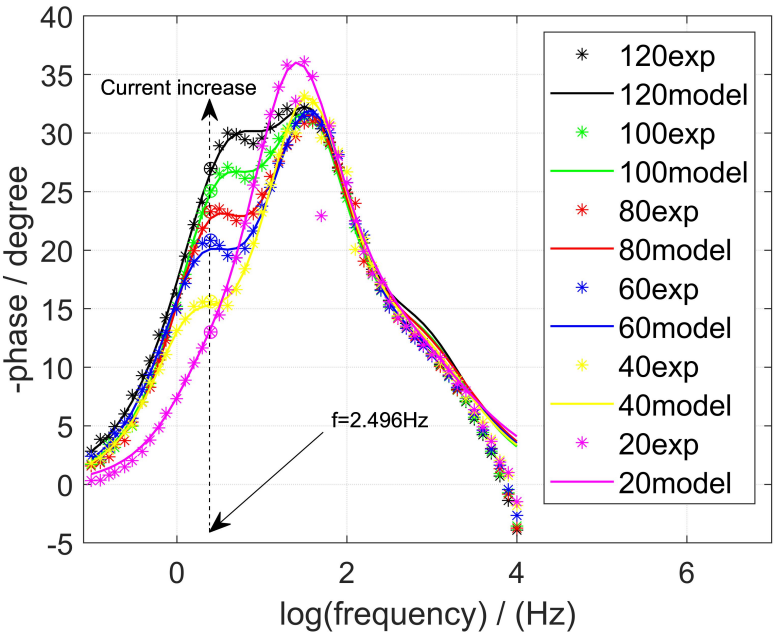


FIGURE 3.23: The turning phase in Bode plot under different currents conditions

Similar to the zero-phase impedance, the turning phase decreases when the PEMFC is under lower current, and it increases when the PEMFC is under high currents. The turning phase under different currents can also be clearly revealed by the figure 3.24. The relationship between the turning phase and the current is also monotonic, so it can be applied in diagnosis. On the contrary, the relationship between the maximum phase and current is disordered, which may be caused by the interaction between the processes of different components.

It is explainable because more water will be generated when the current increases and the water affect not only the membrane resistance. The produced water has been flushed out by the gas, and it affects the mass transfer in the GDL, so the impedance phase of the turning point will increase. Therefore, this feature is another useful indicator for PEMFC diagnosis, and it is complementary to the zero-phase impedance. Also, the turning phase is almost proportional to current magnitude, and it is more sensitive to the decrease of the current.

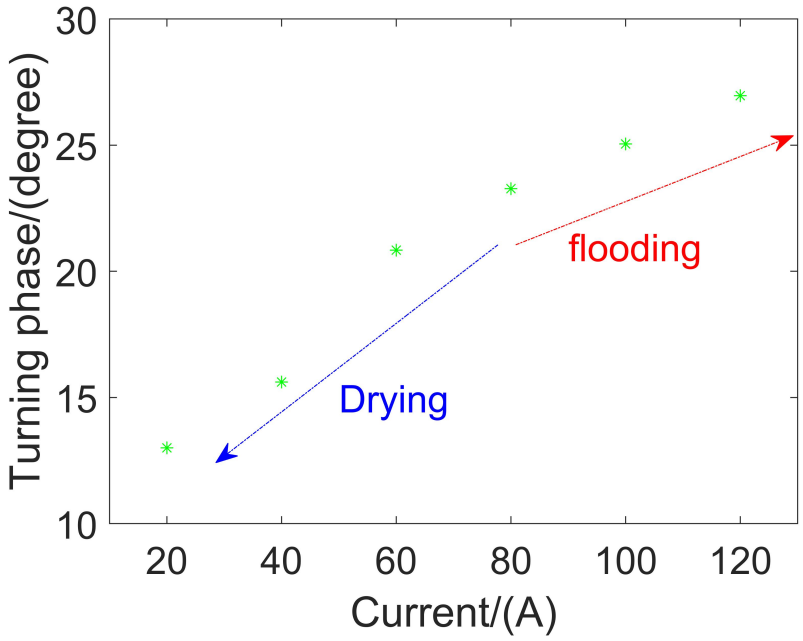


FIGURE 3.24: The turning phase under different currents

To fully understand the physical meaning of the proposed features, the zero-phase impedance and the turning phase of all the experiments are plotted in figure 3.25 and 3.26, respectively. Considering both the two features with the reference of experiment conditions in table 3.2, the features can be well explained and understood. For example, there are two extremely high points for the zero-phase impedance, i.e. experiments No. 8 and No. 39. It can be found that those two experiments are both carried out under extremely low current, that is, 20 A for No. 8 and 40 A for No. 39. Under the extremely low current, the generated water is very little; hence it is extremely dry in the PEMFC and the membrane resistance is enormous. Also, the drying-out fault of No. 39 is more severe than No. 8 while the current of No. 39 is higher, which is counter-intuitive. However, this is because they are tested under different temperatures. Number 8 is tested under 65 °C, while No. 39 is tested under 70 °C, thus the drying-out is intensified by the higher temperature. The other experiments can also be compared and explained in the same way, proving that it is a reasonable diagnostic feature.

It is the same case for the turning phase. Actually, the transformation of the turning phase is more complicated than zero-phase impedance, as the behavior of low-frequency semicircle is not yet fully understood and no consensus was reached. However, it is well recognized that it should be related to the mass transfer process in the GDL [142]. The turning phase can represent the magnitude of a low-frequency

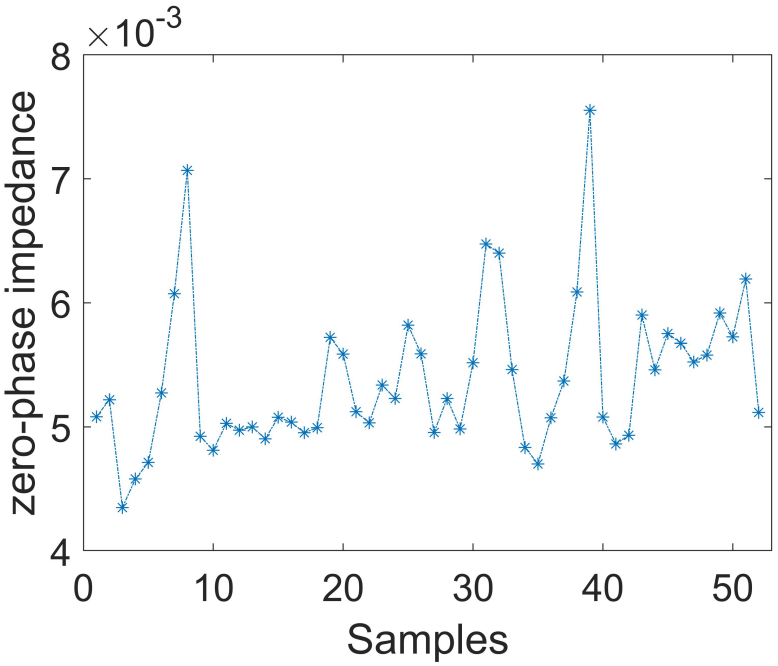


FIGURE 3.25: The zero-phase impedance of all experiment samples

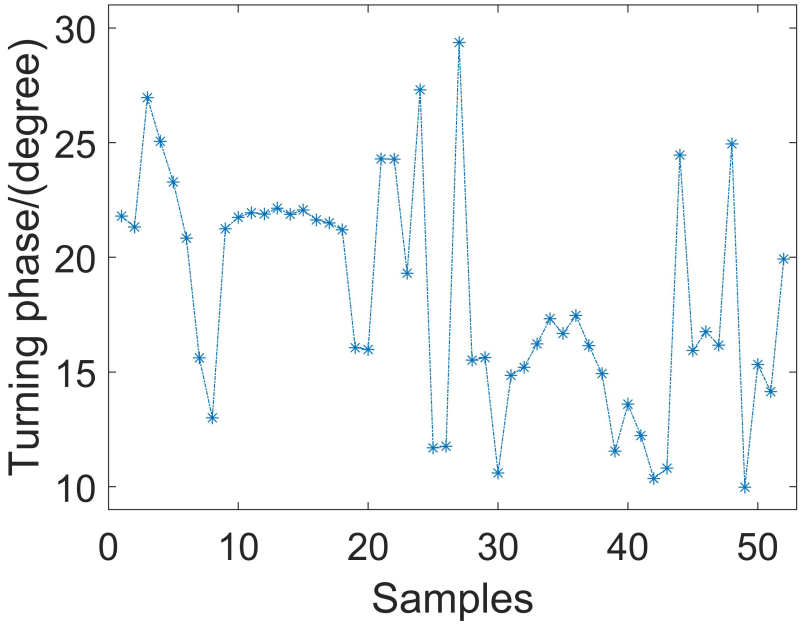


FIGURE 3.26: The turning phase of all experiment samples

semicircle, and it corresponds well with the gas supply conditions according to the experiments. For example, the three experiment samples with extremely high turning phases are experiments No. 3, No. 24, and No. 27. Number 3 is caused by the extremely high current, and No. 24 and No. 27 are caused by the low cathode stoichiometry. Although they are not under the same operating conditions, they can both be explained by the obstructed mass supply. In case No. 3, the water blocked the mass transfer passage, while in the other two cases the air is not sufficiently supplied in the inlet. On the contrary, all the extremely low turning phase appears when the cathode stoichiometry is much higher than the normal value, such as No. 30, No. 42, and No. 49. Therefore, we can draw the conclusion that the turning phase is a useful diagnostic feature that reveals the mass transfer goodness degree.

3.3.2.3 Fault labelling by k-means clustering

As the inner health state is affected by all the operating parameters, the fault classes cannot be directly given according to the operating parameters. The experiment samples can be labelled as different fault causes groups according to the two proposed EIS features by k-means clustering.

The k-means clustering is a widely applied unsupervised classification method to divide the samples into different groups according to the similarity between them. The algorithm of k-means clustering can be given in algorithm 3.4. First of all, k initial centres can be set. Then for each sample, the distance between the sample and different centres can be calculated. The city-block distance is applied in this research, and it gives the same weights to all features as equation 3.38.

$$d_{ij} = \sum_{l=1}^{n} |x_{il} - x_{jl}| \tag{3.38}$$

Where the d_{ij} is the distance between the point i and point j; the x_{il} and x_{jl} is the l_{th} coordinate of point i and point j, respectively; and n is the dimensions of the features, which is 2 in our case.

According to the distance between the sample and centres, the sample can be assigned to the group of the closest centre As all samples have been assigned to different groups,

Algorithm 3.4 : k-means clustering algorithm

give number of cluster k, samples, initial centres;
do
  calculate distance to centres for each sample;
  grouping each sample based on minimum distance;
  move centres to the centroids;
Until centres do not move
End

the group centres can be moved according to the positions of group members, i.e. the centroid calculated by equation 3.39.

$$x_{gl} = \frac{1}{G}\sum_{i=1}^{G} x_{il} \tag{3.39}$$

Where the x_{gl} is the l_{th} coordinate of centre of group g, G is the total number of group g, x_{il} is the l_{th} coordinate of i_{th} sample in group g.

As the group centres move, the distances between the samples and the centre also change. Repeat the distance calculation, group assignment, and centre adaptation over and over again until the centres do not move any more, then the centres and group members can be determined.

3.3.2.4 Fault classification by K-nearest neighbours method

KNN is one of the simplest methods for classification [168]. The principle of KNN is to find the closest points of the aim point and then classify it to the class that appears most times around it. The detailed process of KNN is given as algorithm 2.2.

3.3.2.5 Implementation processes on DSP system

The standard procedures for the real application are combed so that the proposed method can be generalized for any new PEMFC. Also, to validate the method, the industrial-level DSP system is applied to achieve the diagnosis in real time. The DSP system is the same as the one used in ECM-based diagnosis, as can be shown in figure 3.14. All feature detection, feature calculation, and adaptive KNN train, KNN prediction

are carried out on the DSP. The diagnostic processes are arranged by setting time, and the clock of the DSP is applied.

The first task is the obtaining of the two features in the real application. To obtain the zero-phase impedance, the main idea is to change the frequency to detect the samples within the zero-phase zone ($\pm 3^\circ$) according to feedback. The frequency should be increased if the phase angle is smaller than -3° and it can be decreased if the angle phase is bigger than $+3^\circ$ [31]. To measure the zero-phase impedance, two parameters should be decided, i.e. the basic frequency for the first measurement point and the frequency interval for each adjustment step. The basic frequency f_b can be decided according to the approximation of the frequency corresponding to HFR. This frequency is different for different PEMFCs, so it should be set according to the experiment. However, as the PEMFCs of the same batch are similar, the basic frequency needs only to be initialized once. Another parameter is the frequency interval Δf when each time the frequency is adjusted. If it is too small, too many samples will be measured and that wastes time. If it is too big, the samples of the zero-phase zone will be too few, and the random error will increase. A good way to decide the interval is to calculate it according to the sample number that we want, and it is recommended to be between 3 to 10. The total process for measurement of zero-phase impedance can be given in figure 3.27.

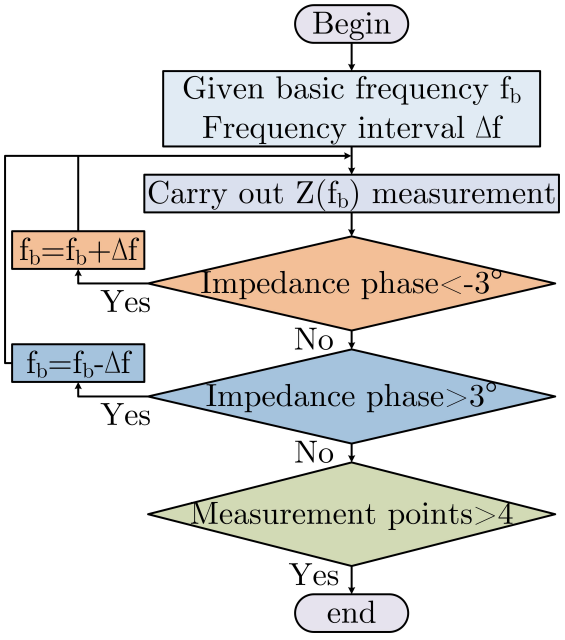


FIGURE 3.27: The zero-phase impedance measurement process

For the turning phase, it is also easy to be obtained. As has been addressed above, the turning phase is the second peak in the phase plot under nominal conditions. Therefore, the frequency that corresponds to the second peak phase can be found by experiment around this frequency, and the turning phase under the turning frequency f_t can be obtained immediately.

After the obtaining of two features, the diagnosis can be achieved with KNN. The experiments can be carried out on one PEMFC and applied to a batch. The total diagnosis procedure can be given in figure 3.28. The off-line period must be performed prior to the application, and the on-line portion can be performed while the PEMFC is in operation.

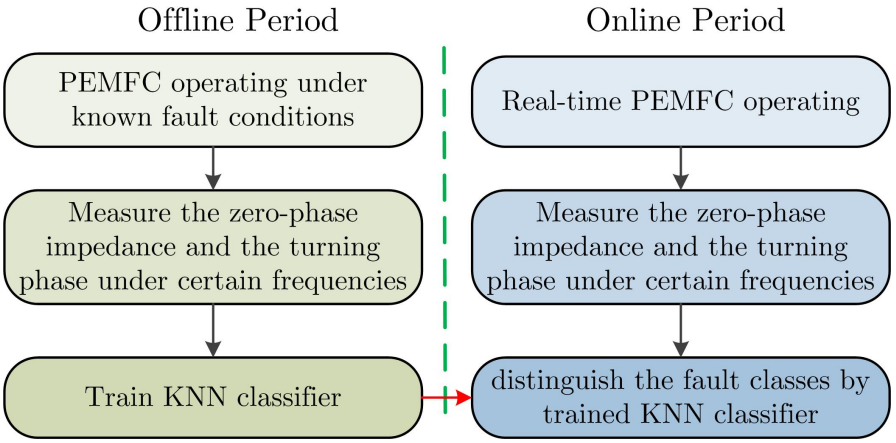


FIGURE 3.28: The overall diagnosis process

The total diagnosis time can be calculated for each diagnosis process. For the measurement time of zero-phase impedance and turning phase, it can be calculated according to the equation 3.40.

$$t = \sum_{i=1}^{n} \frac{N_p}{f_i} \tag{3.40}$$

Where t is the total time for measurement; i is the index of EIS measure points; n is the number of total measurement points, i.e. the number of zero-phase samples plus one turning phase sample; N_p is the number of sinusoidal periods for each frequency stimulation; f_i is the frequency of the samples. In our case, n is 5 because 4 zero phase samples and 1 turning phase sample are detected; the N_p can be set as 5 as

reference [38]; f_i is 9996 Hz, 7939 Hz,6306 Hz,5010 Hz, and 2.496 Hz for $i = 1$ to 5, respectively. Therefore, the total measurement time is 2.006 seconds.

3.3.3 Diagnostic results on DSP system

3.3.3.1 Fault labelling result by k-means clustering

According to the analysis above, the proposed two features can represent the different fault causes. The k-means clustering algorithm is applied to label the experiments as different fault groups, and the clustering result is given in figure 3.29.

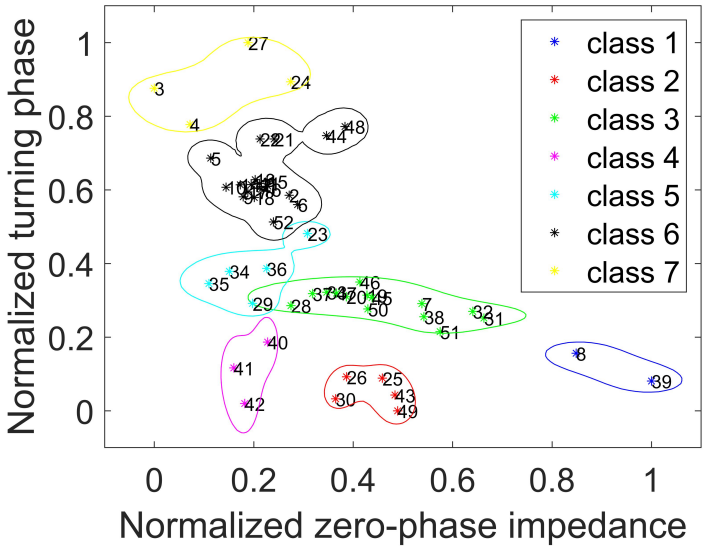


FIGURE 3.29: The fault groups labelled by k-means clustering

The first feature is the normalized zero-phase impedance, and the second phase is the normalized turning phase. Seven classes can be clustered according to the two feature coordination. The clustering result can be analysed by referring to the operating conditions shown in table 3.2. Class 1 is the fault group with big zero-phase impedance and small turning phase. Two samples belong to this class, and those two samples are of the extremely low current condition. As the current is low, the water content in the membrane is very little, so class 1 is the drying out fault condition. On the contrary, class 2 and class 3 are of middle zero-phase impedance, so the water condition in the membrane is good in those experiments. However, class 2 has a smaller turning phase compared to class 3, which is caused by the higher gas input stoichiometry of class 2. Therefore, class 2 is under oversupply condition while class 3 is the normal

condition. Class 4, 5, 6, and 7 are all with small zero-phase impedance, so all of those classes are kind of flooding. However, as different operating parameters are applied for the experiments, the flooding degree and cause reason are quite different, as a result they can be divided into different classes according to the turning phase. Class 4 is the samples with small zero-phase impedance and small turning phase, and they have high water content but also high gas supply flow. Therefore, class 4 is the flooding&oversupply case. Class 5 is only with flooding fault as the gas supply is normal. Class 6 is the case with both flooding and slight gas short supply fault, while class 7 is an extreme flooding and short-supply fault case. It can be noticed that some samples are not short-supplied by the inlet gas stoichiometry, but the turning phase is still high, and this is caused by the extremely high water content in the PEMFC. As the water will not only affect the membrane conductivity, the high water content in the GDL will also increase the air supply resistance, thus affecting the gas supply condition.

It can also be seen that some samples are close but they belong to different classes. It is because the health state is continuous in the PEMFC. As the parameters of the operating condition are close between experiment samples, the inner state is similar too. However, similar cases can be clustered according to the distance and the boundary between the classes are clear.

Also, the clustering results according to the proposed features are also compared with the result of ECM above. The results correspond well with former research, and the difference is that the experimental cases can be gathered in 7 groups rather than 4 groups as in former research. The mass transfer fault conditions of former research are divided into two groups with different water content, while the flooding fault cases of former research can be divided into 3 different groups with different causes. Therefore, the proposed features can give more details about the PEMFC inner health condition, and they can be applied to better distinguish the fault causes and also degrees.

3.3.3.2 Real-time diagnosis result on DSP system

As the experiments are labelled, the KNN can be applied to diagnosis. The database of proposed two features can be divided into the training part and the testing part. The 80% of the data is randomly set as training samples, while the rest 20% of the samples

are set as tests. Both the features and labels of the training part are applied to the KNN classifier, then the classes of test part are predicted by only the test features.

The diagnostic results by the DSP system can be shown in figure 3.30. The experiment samples are represented by crosses, while the model predictions are represented by circles. The green lines are the boundary between two groups. It can be seen that all the predictions correspond well with the experiment result, and the diagnosis result is very accurate. The KNN classification process is very quick, the time to diagnose one case is 0.5 seconds on the DSP system. Therefore, altogether the time cost for one full diagnosis is 2.506 seconds. It is very short compared with the fault development time, so it is quick and practical.

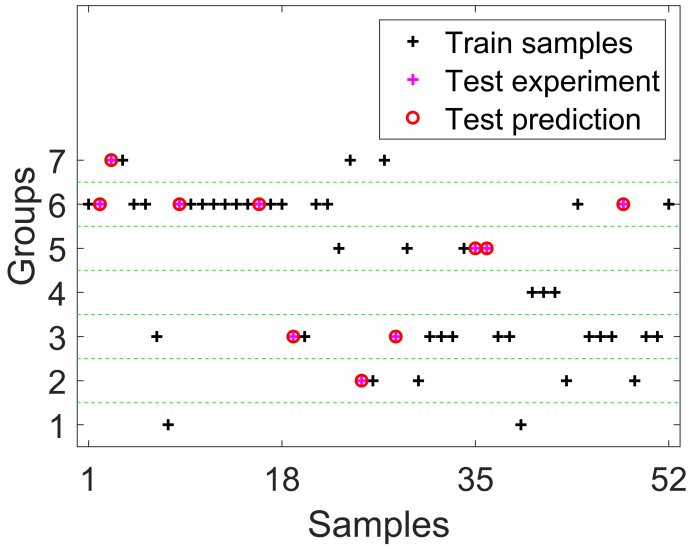


FIGURE 3.30: The diagnosis results by KNN in DSP system

Another former research also achieved quantitative diagnosis based on EIS [177]. 6 EIS points are chosen to replace the total EIS, and the diagnosis accuracy was only 91%. Therefore, the proposed method is more accurate than the former method.

3.3.4 Conclusions

In this research, a novel diagnosis method is achieved based on two EIS features that can be quickly detected. The proposed two features are effective health state indicators. The experiments prove that the zero-phase impedance is a good indicator of membrane water content situation, while the turning phase is a good indicator of

the mass diffusion condition in the GDL. Both gas supply problems and water blockage problems can lead to an increase in the turning phase; therefore it can reflect the health state in GDL. The KNN classification method can assign all the samples to the right fault causes classes. The procedures to generalize the proposed diagnostic method to any PEMFC are discussed and the detailed steps are provided. What's more, the algorithm is validated in real time on a DSP system, and the total diagnostic time is only 2.506 seconds. The proposed diagnosis method based on the proposed features and KNN is accurate and practical.

3.4 Conclusions

The EIS technology has great potential to be applied to on-board diagnosis, as it is rich in diagnostic information and it can be carried out with a minimal disruption of normal operation. In this chapter, two diagnostic methods based on EIS data are proposed. The first one is based on ECM and ANFIS, while the second method is based on two quick detective EIS features and KNN. Both of them are quite accurate, and they are all validated in real time by implementing them on the DSP system. Therefore, the proposed methods are practical, and they have the potential to be applied in real applications. In addition to diagnosis to find malfunctions, long-term health management is also important for PEMFCs. The following chapter proposes a prognostic method based on long-term voltage data and the Kalman filter in the frequency domain that offers a computational advantage over its time-domain counterpart.

Chapter 4

Prognosis by frequency domain Kalman filter

4.1 Introduction

During the operation, the performance of PEMFC declines irreversibly, which is called degradation [192]. It includes the degradation of different components such as bipolar, catalyst, membrane, and electrode, thus it is a complex multi-physics, multi-units, and interactive process [193]. A lot of research works have been carried out by experiments or theoretical analysis, and some degradation mechanisms have been proposed. However, it is still a tricky task to fully understand and model it.

To handle the problem of a short lifetime, prognostics and health management (PHM) are proposed to predict the degradation process and get the estimation of the remaining useful life (RUL) of PEMFC [4]. The existing prognosis methods can be divided into three categories: model-driven methods, data-driven methods, and hybrid methods.

The model-driven methods learn and predict the PEMFC degradation trend based on specific PEMFC degradation models. The degradation models can be divided into three categories, i.e. physical models, empirical models, and semi-empirical models [193] [194]. M. Bressel et al.[195] [196] predicted the degradation and RUL of PEMFC based on the extended Kalman filter (EKF) and a new empirical model, and the uncertainty of the prediction was quantified. At the same time, a

particle filter and a voltage degradation model were applied by M. Jouin et al.[197], and the degradation and RUL were predicted. Also, a model that considered the characterization disturbances and voltage recovery was researched by them [20]. Meanwhile, K. Chen et al.[198] researched the voltage degradation by unscented Kalman filter, and three different models were tried. The result was validated by three PEMFCs in different fuel cell electric vehicle (FCEV) with real load mission. Model-driven methods can give explicit expressions of degradation. However, as the degradation is highly non-linear, it is not easy to find a suitable model that can be widely used for different fuel cells [192].

The data-driven methods use the historic operation data to learn the inner relationship between them and then predict the degradation. Those methods need no specific models, and they are realized by artificial intelligence [199] [200]. In Ref. [201], the echo state network (ESN) was used to forecast the aging process, and the most influential parameters of ESN were analysed by analysis of variance method. M. Rui et al.[99] studied 8 experiments on 3 different PEMFCs by grid long short-term memory recurrent neural network, and the result corresponded well with experiments. At the same time, in Ref. [202], the degradation of the real mission FCEV was predicted by a combined method based on the wavelet analysis, ELM, and GA. Generally, the data-driven methods need more data, and there are more parameters to be tuned than model-driven methods. Also, those parameters have no physical meaning; therefore it is difficult to identify the causality based on the data-driven methods [193].

The hybrid methods are the combination of model-driven methods and data-driven methods. In Ref. [203], a hybrid prognostic method was proposed, which was based on SVM and regularized particle filter. Meanwhile, Z. Daming et al.[204] combined a non-linear autoregressive neural network (NANN) and particle filter (PF), as NANN was good at local fluctuation prediction while PF could give long-term degradation trends. The hybrid methods can combine the advantages of both model-driven and data-driven methods, but as more than two methods are connected, the complexity of the hybrid methods is usually higher than a single model-driven or data-driven method [205].

By the analysis above, it can be seen that the methods have their advantages and disadvantages. Usually, the model-driven methods can give the long-term trend with less computation, and the explicit relationship can help with decision-making. However,

the existing model-driven methods are still time-consuming when there is a lot of data to process, thus any reduction in the computation complexity is advantageous [206]. Therefore, a novel method based on FDKF is proposed in this chapter, which can achieve accurate prediction and is faster than traditional methods.

In this chapter, we contribute to PEMFC prognosis by introducing a new model-driven method based on the frequency domain Kalman filter (FDKF), which can achieve voltage degradation prediction quickly and accurately. FDKF is initially used on acoustic echo cancellation (AEC) problem [207] [208], and it is adapted to the PEMFC prognosis field for the first time. The prognosis process is achieved in the frequency domain. Four different voltage degradation models are researched by the proposed method, and the results are compared with each other. Different training times are also researched, which proves that the method is robust for both short-term and long-term predictions. The method is validated by two groups of experimental data, which were obtained under constant and dynamic current conditions respectively. The advantage of the proposed FDKF method is that it can process data in groups, so the computation time can be greatly reduced. The linear model, quadratic model, logarithmic model, and exponential model are employed and compared under the framework of the FDKF method. The degradation of PEMFC under different conditions can be predicted by the proposed method, which proves that the method is robust. The proposed FDKF method is compared with the EKF method. It was demonstrated that the proposed method is more time-saving and more accurate than the EKF method.

This chapter is organized as follows: in Section 4.2, the model-driven prognosis method based on FDKF is addressed. The framework of the method is given, as well as a detailed explanation of FDKF. Four voltage degradation models are also chosen here. In Section 4.3, the experiments and the data used in this research are explained in detail. The result is obtained and analysed in Section 4.4. Two case studies are applied to verify the method, and different models and training times are researched. The result of EKF is also compared with the FDKF method here. At the same time, the prediction horizon (PH) is calculated and proved more accurate than the literature. Finally, the main conclusions are summarized in Section 4.5.

To prevent confusion, rules of notations are as follows. The variables are taken in italic letters. The lower-case means that it is in the time domain, while upper-case means

frequency domain. Meanwhile, a bold variable means vector and matrix, while the normal letter means that it is a scalar.

4.2 Prognosis method based on FDKF

In this part, the model-driven prognosis method based on FDKF is presented. First of all, the framework of the model-driven method is discussed in Section 4.2.1. Then, the supposed voltage degradation models are given in Section 4.2.2. Thirdly, the FDKF is explained in detail in Section 4.2.3. Finally, the method to apply FDKF in the prognosis problem is addressed in Section 4.2.4.

4.2.1 Model-driven prognosis method based on FDKF

Model-driven prognosis methods are based on models that describe the degradation phenomenon in PEMFC. The overall framework of the model-driven prognosis method based on FDKF is shown in figure 4.1. Firstly, the operation data of two experiments are used. During the experiment, the temperature, pressure, current, and voltage data are all recorded. Then the characterization of the data can decide which kinds of data will be used in the prognosis. In the PEMFC prognosis problem, stack voltage is usually easy to obtain and it can represent the degradation of PEMFC, so it is taken as an indicator of the state of health of PEMFC. Thirdly, as has been discussed in the introduction, model-driven methods rely on degradation models, so four empirical models are chosen in this research. Then the experimental data can be used to get the estimation of the state variables and output voltage by the proposed FDKF method. Furthermore, the degradation can be predicted by the FDKF method based on the state variables during the training period, thus the prognosis can be achieved.

4.2.2 PEMFC voltage degradation models

The degradation of PEMFC has been researched by experiments as well as theoretical analysis. However, as the degradation is a non-linear and multi-physics process, it is difficult to get an exact model that perfectly corresponds to reality. In most research,

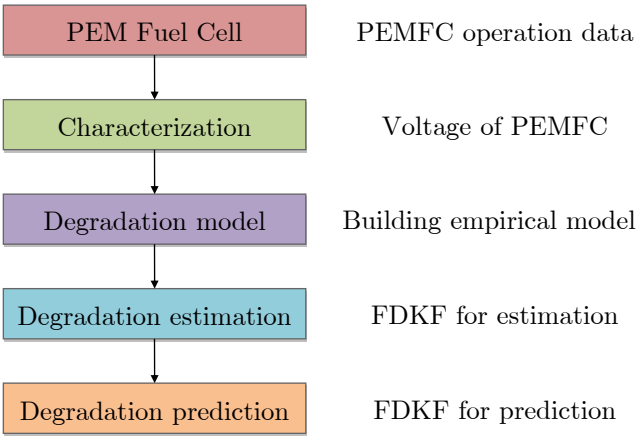


FIGURE 4.1: Model-driven prognosis method based on FDKF for PEMFC

it is given as a function of time with some undetermined parameters. Based on the Ref. [198] [209] [210], four different voltage degradation models are applied in this research. Namely, linear model, quadratic model, logarithmic model and exponential model, which are shown as equations 4.1, 4.2, 4.3 and 4.4, respectively.

$$Linear\ model : x_{i+1} = x_i - \alpha \tag{4.1}$$

$$Quadratic\ model : x_{i+1} = x_i - \alpha - \beta \cdot i \tag{4.2}$$

$$Logarithmic\ model : x_{i+1} = x_i - \alpha - \beta \cdot \ln(\frac{i+1}{i}) \tag{4.3}$$

$$Exponential\ model : x_{i+1} = x_i - \alpha - \beta \cdot (e^{\gamma \cdot i} - e^{\gamma \cdot i - \gamma}) \tag{4.4}$$

Where x is the PEMFC voltage; i is the time index; α, β, γ are undetermined parameters. The α is related to the voltage degradation rate under constant operating condition, and β, γ are the parameters related to voltage degradation under dynamic operating conditions, such as load current change. The degradation of PEMFC is in a near-linear trend under constant load and constant operating condition [210]. When the load changes, the PEMFC will have some transient processes, such as gas diffusion and water accumulation. During the transient process, the degradation is usually accelerated due to the worse environment for the components such as a catalyst, and it is usually in a logarithmic or exponential modes [200].

The linear model is widely used to describe the voltage degradation under normal

condition in a lot of references [198]. However, when the operating condition changes, this model is not enough to describe the degradation, therefore, other terms are introduced to amend it. So the quadratic term, logarithmic term, and exponential term are added in other models, respectively, so that to capture the behaviours under transient operating conditions [209].

4.2.3 Frequency domain Kalman filter

Originally, FDKF is widely used in AEC problem [211] [212]. In this part, we adapt the fuel cell prognosis problem into the form that can be solved by FDKF, which is shown in figure 4.2.

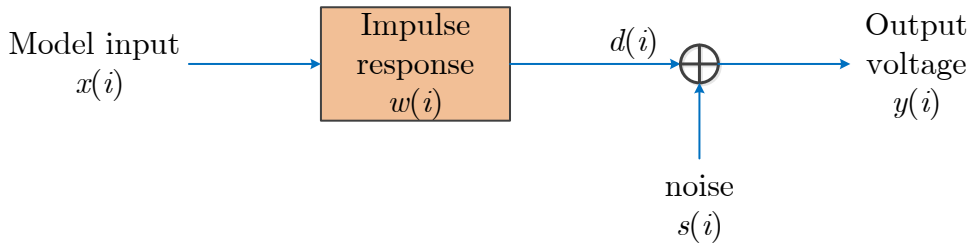


FIGURE 4.2: Fuel cell prognosis problem in time domain

Here i is the time index. $x(i)$ is the voltage signal from model. $w(i)$ is the impulse response path, which is unknown parameter decided by the fuel cell and the operating condition. The convolution of $x(i)$ and $w(i)$ forms $d(i)$. $s(i)$ is a white noise, and it forms the output signal $y(i)$ together with $d(i)$.

In this problem, the most important task is to find the path $w(i)$, so that to get rid of the noise and obtain $d(i)$. The relationship between $x(i)$, $w(i)$ and $d(i)$ can be shown as:

$$d(i) = x(i) * w(i) \tag{4.5}$$

Here the $*$ is the convolution operator. So for $y(i)$, we have:

$$\begin{aligned} y(i) &= d(i) + s(i) \\ &= x(i) * w(i) + s(i) \end{aligned} \tag{4.6}$$

To solve this problem, we can change the equations from the time domain to the frequency domain, so discrete Fourier transfer (DFT) operation is applied to a series of data. To achieve DFT, a window of data at length M is taken, and the window moves forward at a speed of R data each step. In this research we take $M = 256$ and $R = 128$. For step k, take a vector to represent the latest M data of $x(i)$:

$$\boldsymbol{x}(k) = [x((k-1)R+1), x((k-1)R+2), ..., x((k-1)R+M)]^H \tag{4.7}$$

Where the superscript H means the Hermitian transposition. So we get frequency domain input $\boldsymbol{X}(k)$ as:

$$\boldsymbol{X}(k) = diag\{\boldsymbol{F}\boldsymbol{x}(k)\} \tag{4.8}$$

Where $diag\{\}$ creates a new matrix that takes the vector as the main diagonal items of it. $\boldsymbol{F}$ is the Fourier matrix which can achieve the effect of Fourier transfer of a vector, so that to change the vector from the time domain to the frequency domain.

$$\boldsymbol{F}\boldsymbol{x}(k) = [X(1,k), X(2,k), ..., X(M,k)]^H \tag{4.9}$$

For $w(k)$, only the first $M - R$ order response is used, which is obtained from the last step, and we suppose that it can cover all the span of the impulse response. For the part where we ignore, add 0 to make it a M data vector.

$$\boldsymbol{w}(k) = [w(1,k), w(2,k), ..., w(M-R,k), 0, ..., 0]^T; \tag{4.10}$$

Also, in the frequency domain, we can get the complex frequency domain impulse response path as:

$$\begin{aligned} \boldsymbol{W}(k) &= \boldsymbol{F}\boldsymbol{w}(k) \\ &= [W(1,k), W(2,k), ..., W(M,k)]^T \end{aligned} \tag{4.11}$$

For $y(i)$ and $s(i)$, take the latest R data as a vector, it can be represented as:

$$\boldsymbol{y}(k) = [y((k-2)R+M+1), y((k-2)R+M+2), ..., y((k-1)R+M)]^T \tag{4.12}$$

$$\boldsymbol{s}(k) = [s((k-2)R+M+1), s((k-2)R+M+2), ..., s((k-1)R+M)]^T \tag{4.13}$$

Take $\boldsymbol{Q}^H = [\boldsymbol{0}_{R\times(M-R)}\ \boldsymbol{I}_{R\times R}]$, which is a R row M column matrix. It can take only the last R items from a vector of length M. According to the nature of convolution, the convolution of the time domain signal can be transferred to multiplication in the frequency domain, so we have:

$$\boldsymbol{d}(k) = \boldsymbol{Q}^H \boldsymbol{F}^{-1} \boldsymbol{X}(k) \boldsymbol{W}(k) \tag{4.14}$$

The equation 4.6 can be written as:

$$\begin{aligned} \boldsymbol{y}(k) &= \boldsymbol{d}(k) + \boldsymbol{s}(k) \\ &= \boldsymbol{Q}^H \boldsymbol{F}^{-1} \boldsymbol{X}(k) \boldsymbol{W}(k) + \boldsymbol{s}(k) \end{aligned} \tag{4.15}$$

Change it into the frequency domain and we can get:

$$\boldsymbol{Y}(k) = \boldsymbol{F}\boldsymbol{Q}\boldsymbol{Q}^H \boldsymbol{F}^{-1} \boldsymbol{X}(k) \boldsymbol{W}(k) + \boldsymbol{F}\boldsymbol{Q}\boldsymbol{s}(k) \tag{4.16}$$

Here the $\boldsymbol{Q}$ is used to add 0 in the first $M-R$ data of a vector, to make a vector change from length R to length M. Then, by taking:

$$\boldsymbol{C}(k) = \boldsymbol{F}\boldsymbol{Q}\boldsymbol{Q}^H \boldsymbol{F}^{-1} \boldsymbol{X}(k) \tag{4.17}$$

$$\boldsymbol{S}(k) = \boldsymbol{F}\boldsymbol{Q}\boldsymbol{s}(k) \tag{4.18}$$

We can get:

$$\boldsymbol{Y}(k) = \boldsymbol{C}(k)\boldsymbol{W}(k) + \boldsymbol{S}(k) \tag{4.19}$$

We can see that this equation gives the relationship between the frequency domain input signal $\boldsymbol{X}(k)$, frequency domain impulse response $\boldsymbol{W}(k)$, and the frequency domain output signal $\boldsymbol{Y}(k)$.

By taking $\boldsymbol{W}(k)$ as state variable, $\boldsymbol{C}(k)$ as measurement matrix, $\boldsymbol{Y}(k)$ as output and $\boldsymbol{S}(k)$ as noise, supposing the frequency domain response $\boldsymbol{W}(k)$ only change slowly between two steps, we can get the system equations as:

$$\begin{cases} \boldsymbol{W}(k+1) = A\boldsymbol{W}(k) + \Delta\boldsymbol{W}(k) \\ \boldsymbol{Y}(k) = \boldsymbol{C}(k)\boldsymbol{W}(k) + \boldsymbol{S}(k) \end{cases} \tag{4.20}$$

Where A is a constant close to 1. Take $\Delta\boldsymbol{W}(k)$ as white noise, then, in the frequency domain, the system can be represented as:

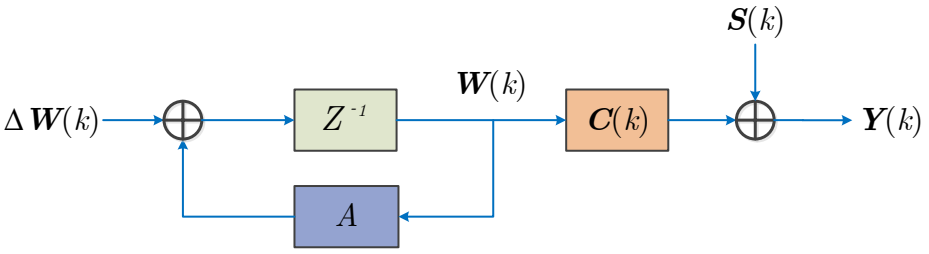


FIGURE 4.3: System of fuel cell prognosis problem in frequency domain

For the problems like figure 4.3, we can use Kalman filter to obtain $\boldsymbol{W}(k)$. Suppose that $\boldsymbol{S}(k)$ and $\Delta\boldsymbol{W}(k)$ are uncorrelated Gaussian noise, with the covariance matrix as $\boldsymbol{\Phi}_{SS}$ and $\boldsymbol{\Phi}_{\Delta\Delta}$, respectively. We can get the solution in Kalman filter form:

$$\boldsymbol{W}^{+}(k) = A\boldsymbol{W}(k-1) \tag{4.21}$$

$$\boldsymbol{P}^{+}(k) = A\boldsymbol{P}(k-1)A' + \boldsymbol{\Phi}_{\Delta\Delta} \tag{4.22}$$

$$\boldsymbol{K}(k) = \boldsymbol{P}^{+}(k)\boldsymbol{C}^{H}(k)[\boldsymbol{C}(k)\boldsymbol{P}^{+}(k)\boldsymbol{C}^{H}(k) + \boldsymbol{\Phi}_{SS}]^{-1} \tag{4.23}$$

$$\boldsymbol{W}(k) = \boldsymbol{W}^{+}(k) + \boldsymbol{K}(k)[\boldsymbol{Y}(k) - \boldsymbol{C}(k)\boldsymbol{W}(k)] \tag{4.24}$$

$$\boldsymbol{P}(k) = (\boldsymbol{I} - \boldsymbol{K}(k)\boldsymbol{C}(k))\boldsymbol{P}^{+}(k) \tag{4.25}$$

Where the $\boldsymbol{P}(k)$ is the estimation error covariance, $\boldsymbol{W}^{+}(k)$ and $\boldsymbol{P}^{+}(k)$ means the one step ahead estimation of $\boldsymbol{W}(k)$ and $\boldsymbol{P}(k)$. $\boldsymbol{K}(k)$ is the Kalman gain in the frequency domain. By this method, we can update $\boldsymbol{W}(k)$ step by step, then use it to predict the voltage output. The stability and convergence of the proposed FDKF method is given in the Appendix B.

In this research, the parameters A, $\boldsymbol{\Phi}_{SS}$, and $\boldsymbol{\Phi}_{\Delta\Delta}$ are tuning parameters, and they are chosen according to the criterion that the estimated voltage is the closest to the experimental data. So $A = 0.9999$, $\boldsymbol{\Phi}_{SS} = 10^{-2}\boldsymbol{I}_{M\times M}$, and $\boldsymbol{\Phi}_{\Delta\Delta} = 10^{-4}\boldsymbol{I}_{M\times M}$ are applied.

4.2.4 Prognosis by FDKF

In this part, we summarize the method to solve the prognosis problem with FDKF, and the diagram of the method is shown in figure 4.4. The prognosis process can be divided into two parts, i.e. the learning period and the prediction period. In the learning period, by the models given in Section 3.2, the $x(i)$ can be obtained. Group $x(i)$ in a block of M data, and at each step the DFT of it can give the frequency domain input $\boldsymbol{X}(k)$. By frequency domain impulse response $\boldsymbol{W}(k)$, the estimated frequency domain output $\hat{\boldsymbol{Y}}(k)$ can be obtained. At the same time, the real output data $y(i)$ is also transferred into the frequency domain, which is $\boldsymbol{Y}(k)$. Then the frequency domain error $\boldsymbol{E}(k)$ can be obtained by the subtraction of real frequency domain output $\boldsymbol{Y}(k)$ and the estimated output $\hat{\boldsymbol{Y}}(i)$. It will be used to update $\boldsymbol{W}(k)$, according to the equations 4.21 to 4.25. Finally, the Inverse Discrete Fourier Transform (IDFT) can change the estimated frequency domain output $\hat{\boldsymbol{Y}}(k)$ into the time domain. By this method, the impulse response updates and the output voltage can be estimated.

When the learning period is over, the process moves to the prediction period. During the prediction period, the $\boldsymbol{W}(k)$ of the learning period is used in the prediction. By the same process as the learning period, the $x(i)$ can be transferred to the frequency domain, and the predicted $\hat{\boldsymbol{Y}}(k)$ can be obtained by $\boldsymbol{X}(k)$ and $\boldsymbol{W}(k)$ according to equation 4.20. At last, the predicted voltage $\hat{y}(i)$ can be calculated by IDFT, thus the prediction of voltage degradation is achieved.

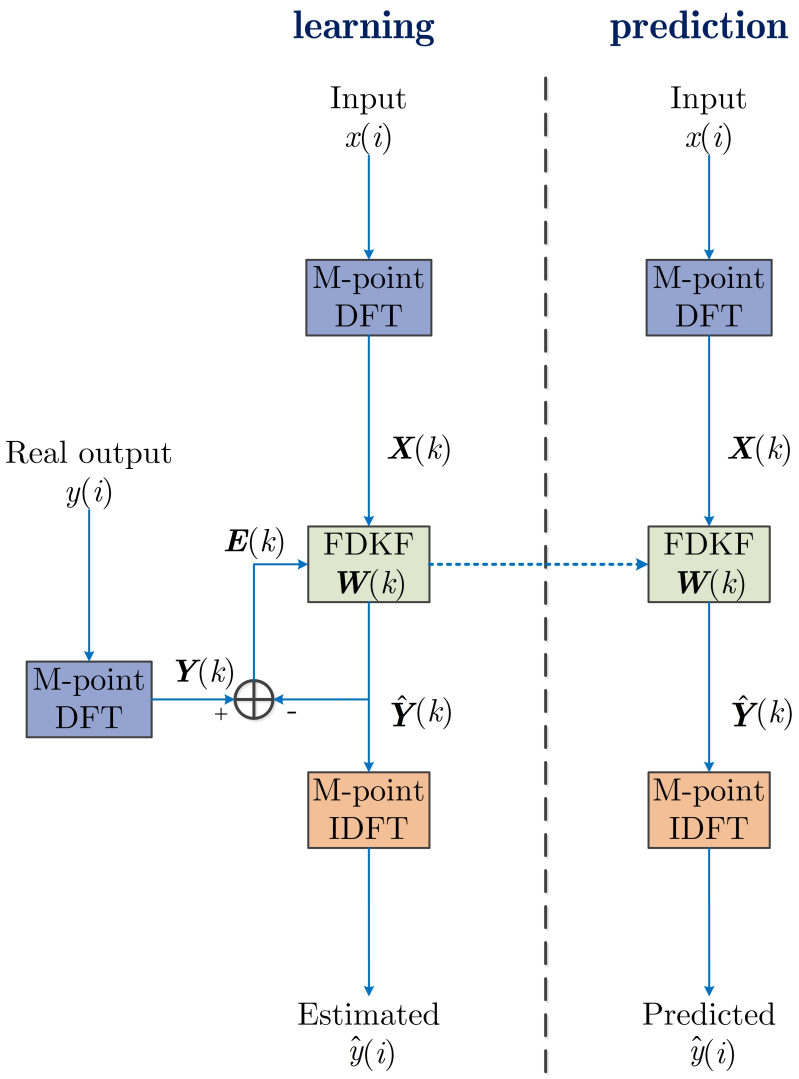


FIGURE 4.4: PEMFC prognosis by FDKF

4.3 Experiments and data

4.3.1 Experiments

Two experiments were carried out in the Federation for Fuel Cell Research (FCLAB) for prognosis, where two PEMFC stacks degraded under different operating conditions [213]. The experiment platform is shown in figure 4.5. In the first experiment, a PEMFC stack with 5 fuel cells operated under a constant current of 70A, and it lasted for more than 1100 h. It is called PEMFC1 in short. The main operating parameters of the experiment are shown in table 4.1. The second experiment was carried out on another 5-cells stack, but under the dynamic current, i.e., 70 A with 7 A high-frequency ripples, and it is called PEMFC2. The main operating parameters of

PEMFC2 are shown in table 4.2. The information of the stacks was recorded during the experiments, including the temperature, pressure, current, and voltage, etc. There were characterization experiments about every week during both of the degradation experiments, including the polarization curve and EIS measurement. The voltage has an obvious recovery after the characterization experiments, which enhances the non-linear property of voltage. More details about the experiments can be found in the reference [213].

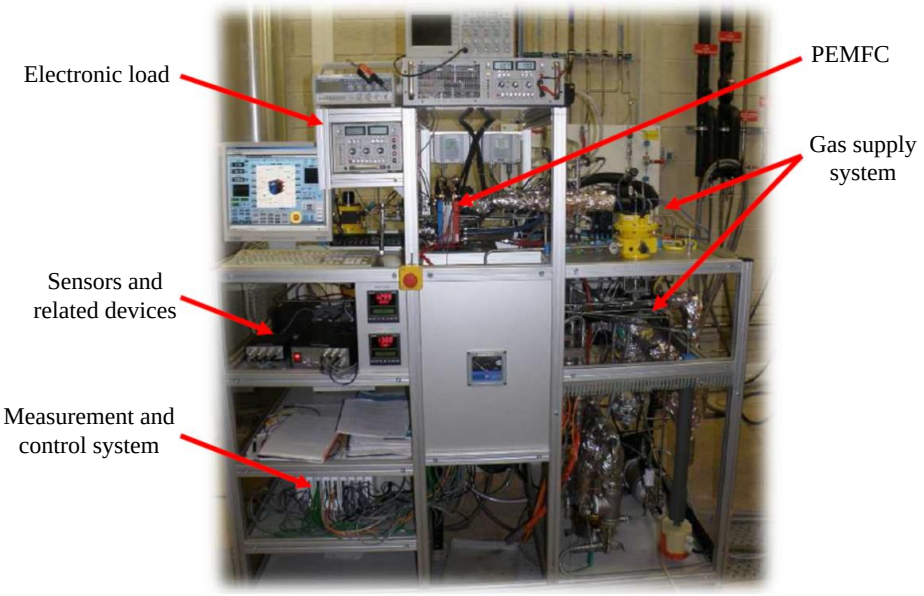


FIGURE 4.5: PEMFC degradation experiment platform in FCLAB [213]

TABLE 4.1: Experiment parameters of PEMFC1

Parameter	Value
Number of cells	5
Active area	$100cm^2$
Stack rated current	70A
Temperature	54 °C
Hydrogen pressure	1.3 bar
Relative humidity	50%

4.3.2 Data processing

The voltage data has to be processed before it is applied to the prognosis because there are random errors and violent fluctuations during the experiment. Gaussian kernel smooth method [214] is applied to process the voltage data in this research. It gives

TABLE 4.2: Experiment parameters of PEMFC2

Parameter	Value
Number of cells	5
Active area	100cm^2
Stack rated current	70A with 7A vibration
Temperature	54 °C
Hydrogen pressure	1.3 bar
Relative humidity	51.8%

a weight to every data according to Gaussian distribution and then gets the weighted average as the new value. However, it should be noticed that the future data should not be used to smooth the data in prognosis, so that to prevent data leakage. Therefore, only the left half of the Gaussian distribution is applied, while the right half is set as 0. For a time series data $d(t_i)$, where i is the index of the data, t_i is the corresponding time, the processed data $p(t_i)$ can be obtained as:

$$p(t_i) = \frac{\sum_{t_j=t_i-\Delta t}^{t_i} K(t_j) \cdot d(t_j)}{\sum_{t_j=t_i-\Delta t}^{t_i} K(t_j)} \tag{4.26}$$

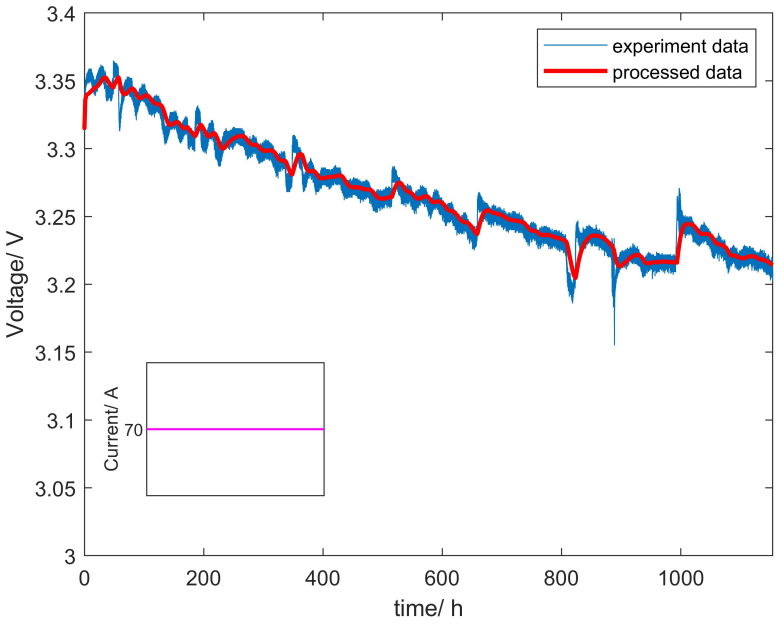
Where Δt is the smoothing interval, which is set 10 h in this case. The $K(t_j)$ is the Gaussian kernel function as:

$$K(t_j) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(t_j-t_i)^2}{2h^2}} \tag{4.27}$$

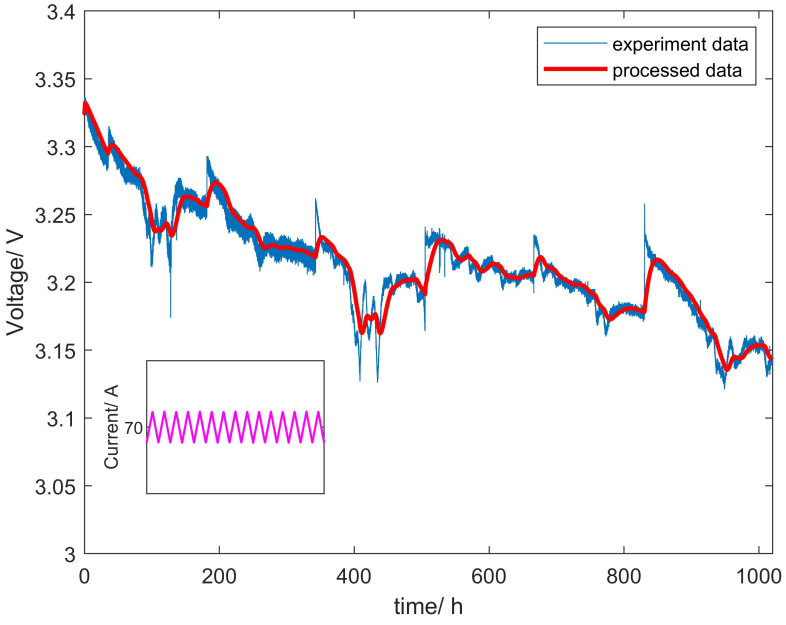
Where the h is the bandwidth, which is set as 500. The raw data and processed data of voltages are shown in figure 4.6(a) and 4.6(b) for PEMFC1 and PEMFC2, respectively. The corresponding operation currents of PEMFC1 and PEMFC2 are also shown in the figures. It can be seen that the processed data is smoother, but it can also represent the trend of original data. Afterwards, only the processed data is used.

4.4 Result and validation

In this section, the prediction of PEMFC voltage is obtained according to the proposed method. Two case studies are carried out, where the experimental data of PEMFC1 is



(a)



(b)

FIGURE 4.6: Voltage raw data and processed data for (a) PEMFC1 and (b) PEMFC2

used in Section 4.4.1 while PEMFC2 is used in Section 4.4.2. At the same time, the EKF method is compared with the proposed FDKF method in Section 4.4.3.

In each case study, four different models are studied respectively, and predictions are carried out under different training length so that to achieve both short-term and long-term prediction. The models are trained under 50%, 60%, 70%, 80%, 90% of the total time respectively, and the rest part of the experimental data is used as verification. Mean relative error (MRE) and root mean square error (RMSE) are taken as the judgements of its performance, which can be calculated by equation 4.28 and 4.29, respectively.

$$MRE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{V}(i) - V(i)}{V(i)}\right| \tag{4.28}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{V}(i) - V(i))^2}{N}} \tag{4.29}$$

N is the number of data; $\hat{V}(i)$ is the predicted voltage; $V(i)$ is the real voltage output. The MRE is the mean relative errors, while RMSE is more sensitive to the big errors. By those two methods, the performance of a prediction can be measured.

4.4.1 Case study 1: voltage degradation prediction for PEMFC1

In this part, based on the PEMFC1 data, four voltage degradation models are researched under different training time. The predicted voltage and relative error are analysed. They are given in the following order: linear model, quadratic model, logarithmic model, and exponential model.

4.4.1.1 Linear model for PEMFC1

The training and prediction voltage and corresponding relative error under linear model for PEMFC1 are shown in figure 4.7(a) and 4.7(b), respectively. It can be seen that the training result follows the experimental data well, and the predicted voltage is close to the real output voltage. As there are characterization experiments during the

degradation experiment at some chosen points, the voltage sharply fluctuates and it is quite non-linear. The predicted voltage is affected by the training length, and the result is more accurate with longer training time.

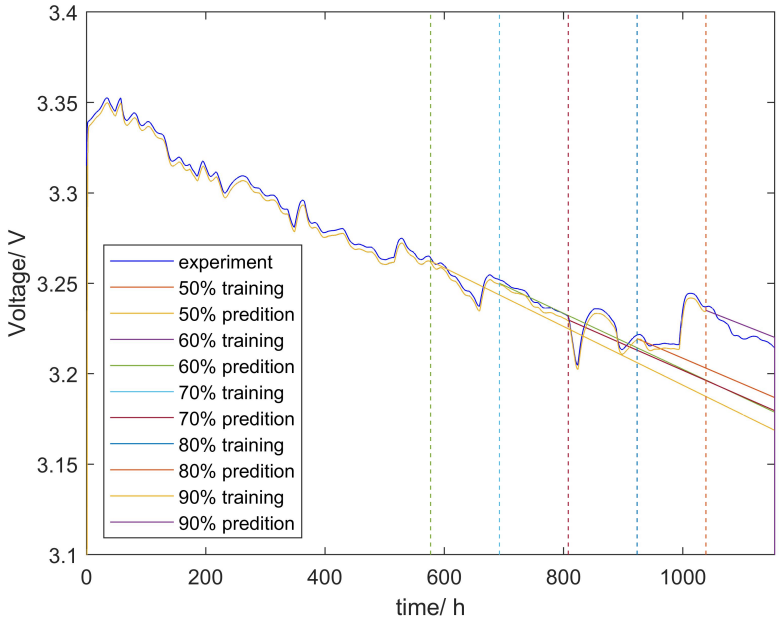
The relative error of the training period is smaller than 0.4%, so it follows the training data well. For the prediction period, the error arises when the experimental voltage abruptly changes, but the prediction error is always smaller than 1.8%, and the averaged RMSE is 0.0327 V. The experiment voltage suddenly rises at about 1000 hours, which is not recognized by the long-term prediction, so the error is relatively big. However, the predicted voltage under 90% training time corresponds well with the experiment. In summary, as some non-linear behaviours of PEMFC in the future are not considered, the long-term prediction is less accurate than short-term prediction.

4.4.1.2 Quadratic model for PEMFC1

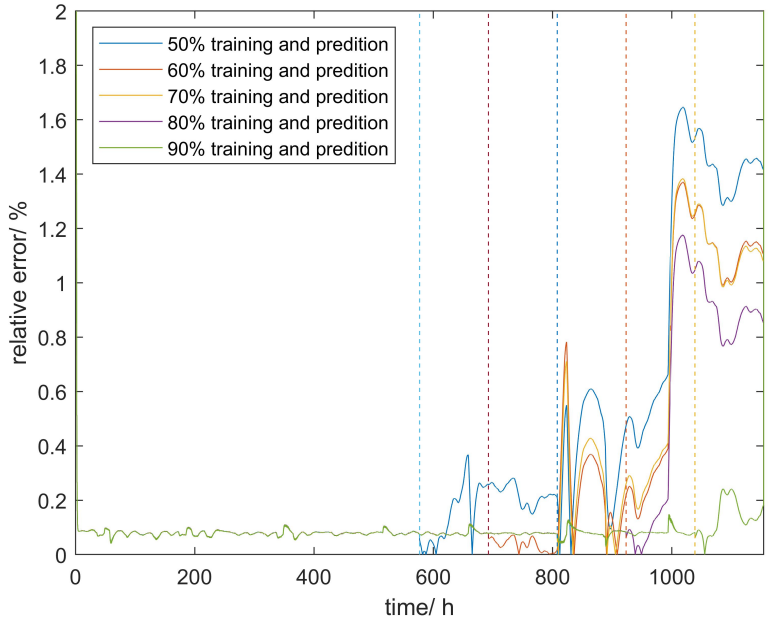
The quadratic model is the combination of a linear term and a quadratic term. Different training time is researched in this case, and the predicted voltage and relative error are shown in figure 4.8(a) and 4.8(b), respectively. As shown in figure 4.8(a), the predicted voltage can follow the actual voltage degradation trend for PEMFC1. As shown in figure 4.8(b), the relative error of this model is always smaller than 1.4%, while the maximum error of the linear model is 1.8%. The reason for the high accuracy of the quadratic model is that the quadratic term effectively considers the voltage recovery phenomenon of PEMFC1 caused by the characteristic experiment.

4.4.1.3 Logarithmic model for PEMFC1

The logarithmic model is the combination of a linear term and a logarithmic term. The training and prediction result and corresponding relative error under logarithmic model for PEMFC1 are shown in figure 4.9(a) and 4.9(b), respectively. It can be seen that the result of the logarithmic model is quite similar to the linear model, and it is more accurate than the linear model, which is shown by the average MRE in table 4.3. The reason for the high accuracy of the logarithmic model is that the logarithmic term effectively considers the voltage recovery phenomenon of PEMFC1 caused by the characteristic experiment. It is notable that the result is quite accurate before the sudden jump of

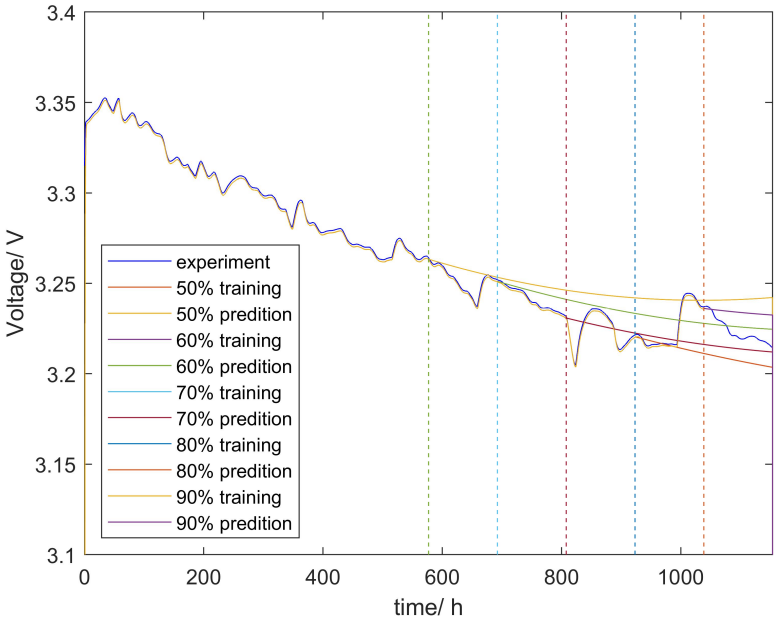


(a)

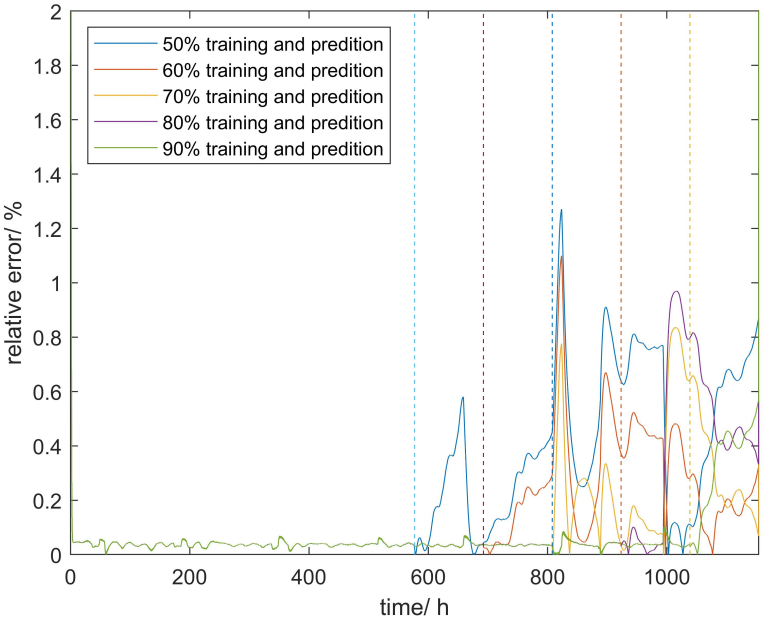


(b)

FIGURE 4.7: (a) The voltage prediction result based on FDKF under linear model and (b) the relative error under linear model for PEMFC1

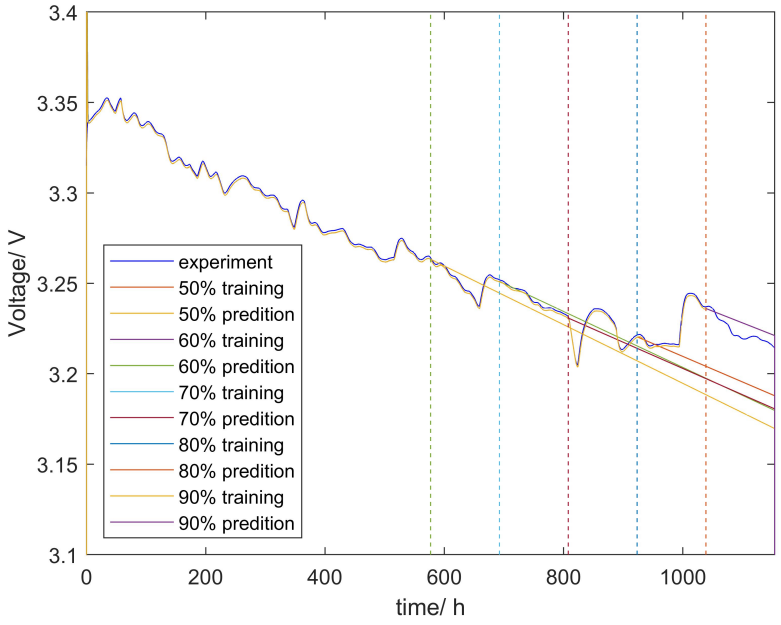


(a)

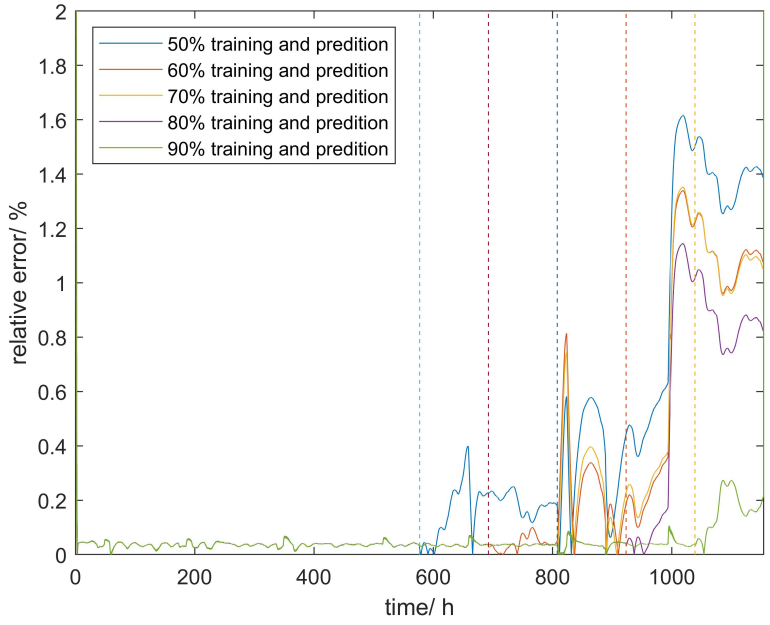


(b)

FIGURE 4.8: (a) The voltage prediction result based on FDKF under quadratic model and (b) the relative error under quadratic model for PEMFC1

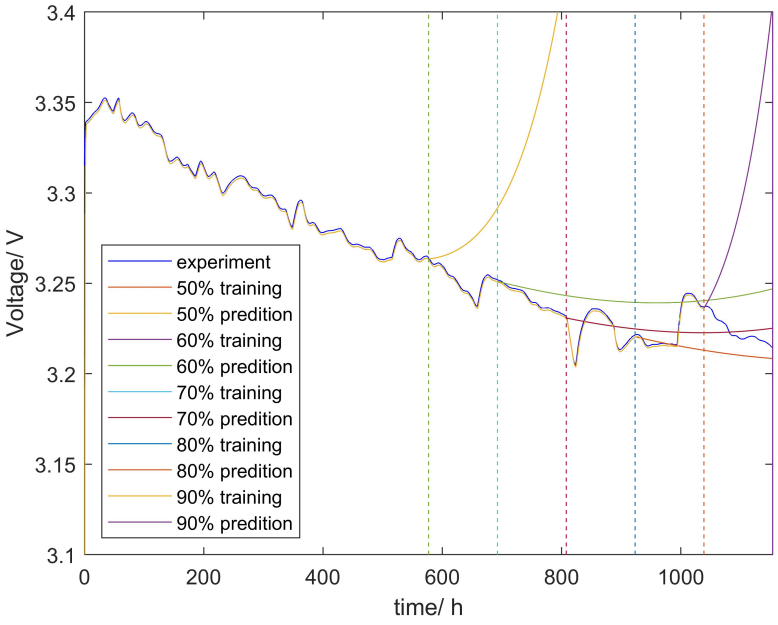


(a)

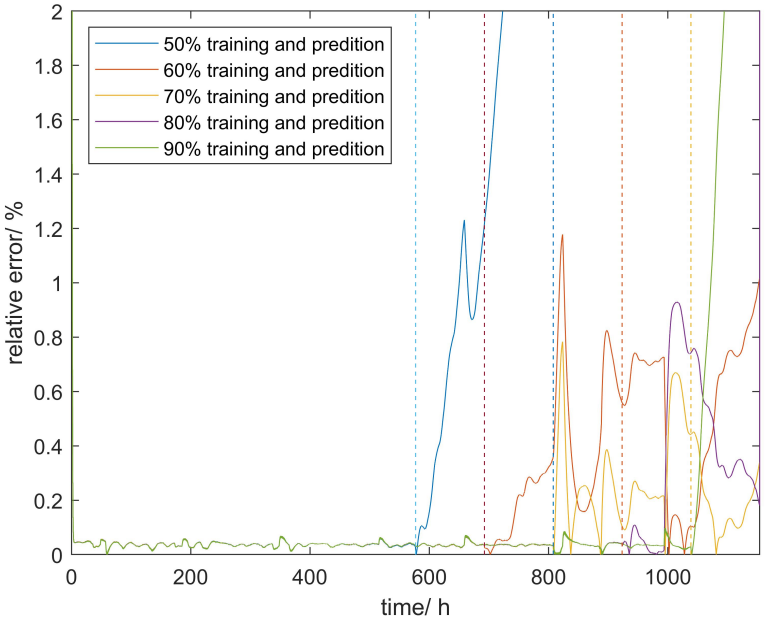


(b)

FIGURE 4.9: (a) The voltage prediction result based on FDKF under logarithmic model and (b) the relative error under logarithmic model for PEMFC1



(a)



(b)

FIGURE 4.10: (a) The voltage prediction result based on FDKF under exponential model and (b) the relative error under exponential model for PEMFC1

voltage at 1000h, where the relative error is less than 1%. Therefore, for some research that declares the end-of-life (EOL) of the stack at 800h [197], the prediction can be more accurate because the voltage degradation after EOL isn't concerned.

4.4.1.4 Exponential model for PEMFC1

For the exponential model, the voltage result and corresponding relative error for PEMFC1 are shown in figure 4.10(a) and 4.10(b), respectively. It can be seen that the result under 70% and 80% training time is accurate compared with experiment voltage, while the relative error under other training time is huge. The predicted voltage goes up dramatically because of the existence of an exponential term, which is far from reality. This proves that the exponential model is not robust for different training time, and it is not proper to be used in the method.

4.4.1.5 Comparison of different models

To compare those four models, the MRE and RMSE of predictions under different models and different training time are listed in table 4.3 and table 4.4, respectively. It can be seen that the average MRE of the quadratic model is the smallest, which is 0.4419%, and the average MRE of the linear model and logarithmic model are very close to that of the quadratic model, which are 0.6131% and 0.5999%, respectively. All those three models can give an acceptable result. They are also applied to PEMFC2 in the next section, to verify if they are effective under the dynamic current condition. The RMSE of the predictions give the same conclusions as MRE. As we can see, the RMSE of the exponential model is much higher than the linear model, quadratic model, and logarithmic model, and the exponential model is unacceptable.

TABLE 4.3: MRE (%) of prediction under different model and different training time for PEMFC1

model	50%	60%	70%	80%	90%	average
linear model	0.6825	0.5302	1.0096	0.6698	0.1736	0.6131
quadratic model	0.5449	0.3423	0.5276	0.4087	0.3861	0.4419
logarithmic model	0.6593	0.5142	0.9808	0.6438	0.2014	0.5999
exponential model	4.6888	0.5811	0.3844	0.3523	1.7269	1.5466

TABLE 4.4: RMSE (V) of prediction under different model and different training time for PEMFC1

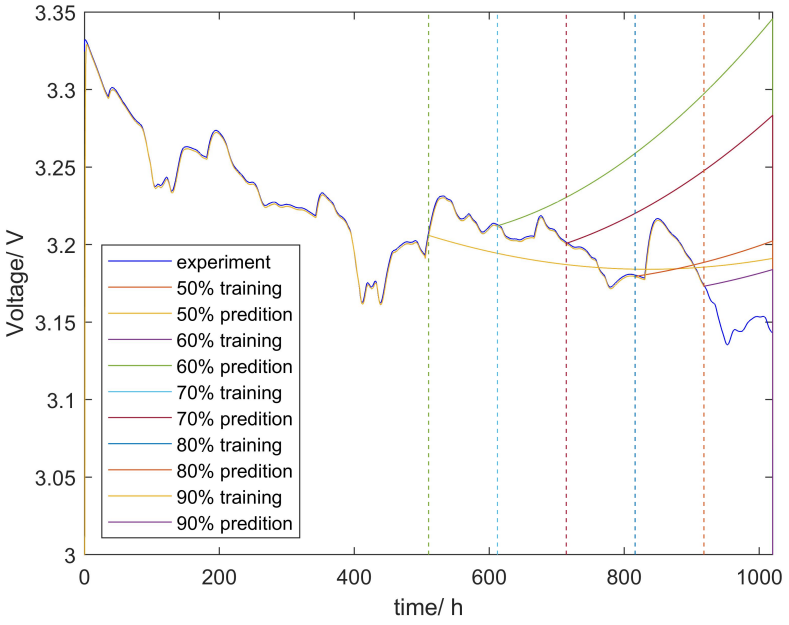
model	50%	60%	70%	80%	90%	average
linear model	0.0477	0.0357	0.0466	0.0273	0.0060	0.0327
quadratic model	0.0306	0.0115	0.0202	0.0146	0.0135	0.0181
logarithmic model	0.0467	0.0347	0.0456	0.0263	0.0070	0.0320
exponential model	0.4127	0.0321	0.0124	0.0119	0.0662	0.1071

4.4.2 Case study 2: voltage degradation prediction for PEMFC2

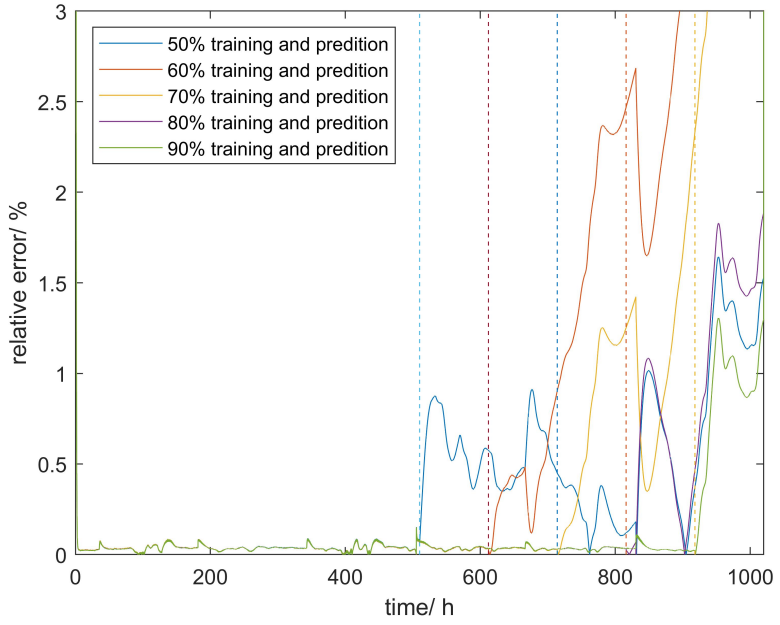
The experiment PEMFC2 is carried out under a dynamic current of 70A with 7A high-frequency ripples, whose operation conditions and voltage output can be found in Section II. The degradation pattern of PEMFC2 is different from PEMFC1, as PEMFC1 is obtained under constant current. To verify the FDKF prognosis method under the dynamic current case, it is also applied to PEMFC2. The FDKF method is also trained under 50%, 60%, 70%, 80%, 90% of the total time respectively, and the rest part of the experimental data is used as verification. In this section, the quadratic model, linear model, and logarithmic model are calculated, as the exponential model has already been proved unsuitable to be used in this method by the research above.

4.4.2.1 Quadratic model for PEMFC2

As the quadratic model has the smallest relative error in the PEMFC1 case, it is applied to PEMFC2 firstly. The training and prediction result and corresponding relative error under quadratic model for PEMFC2 are shown in figure 4.11(a) and 4.11(b), respectively. The result is very different under different training time. The model results under 60 % and 70 % training time corresponds badly with the voltage of PEMFC2, which could be caused by the non-linear feature of PEMFC2. Compared with the result of PEMFC1 based on the quadratic model, the quadratic term cannot accurately model the voltage recovery phenomenon of PEMFC2. Therefore, this model is unsuitable to be used in this case, especially for long-term prediction. As the result of the linear model and logarithmic model are acceptable in PEMFC1, both of them are calculated for PEMFC2 by the proposed FDKF method.



(a)



(b)

FIGURE 4.11: (a) The voltage prediction result based on FDKF under quadratic model and (b) the relative error under quadratic model for PEMFC2

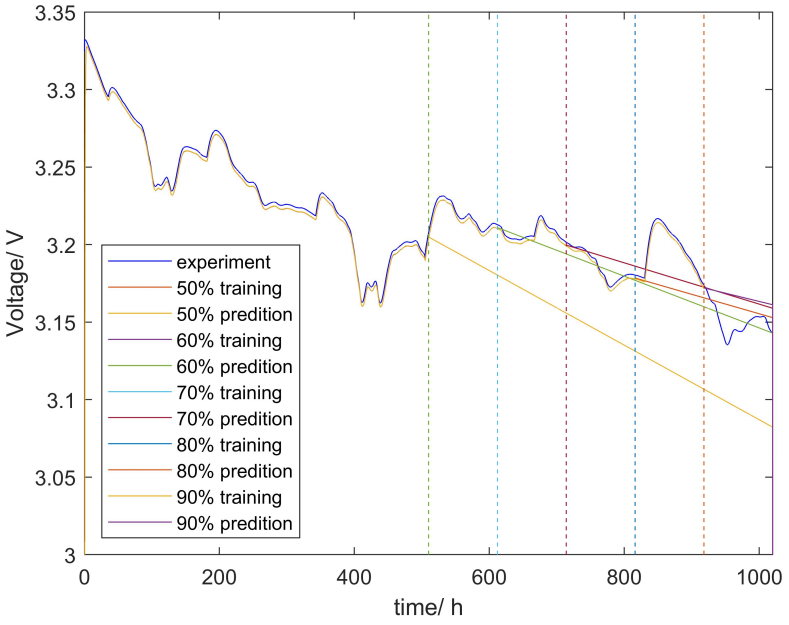
4.4.2.2 Linear model for PEMFC2

The voltage result and corresponding relative error under linear model for PEMFC2 is shown in figure 4.12(a) and 4.12(b), respectively. It can be seen that although the relative error arises when the voltage fluctuates, it is always smaller than 2.5%, thus the result of the linear model is more accurate than the quadratic model. As can be seen in table 4.5, the average MRE of the linear model is 0.5300%, which is much smaller than the quadratic model. The MRE of PEMFC2 is also smaller than the result of the linear model for PEMFC1. Thus, the linear model can be used in the FDKF method for both cases.

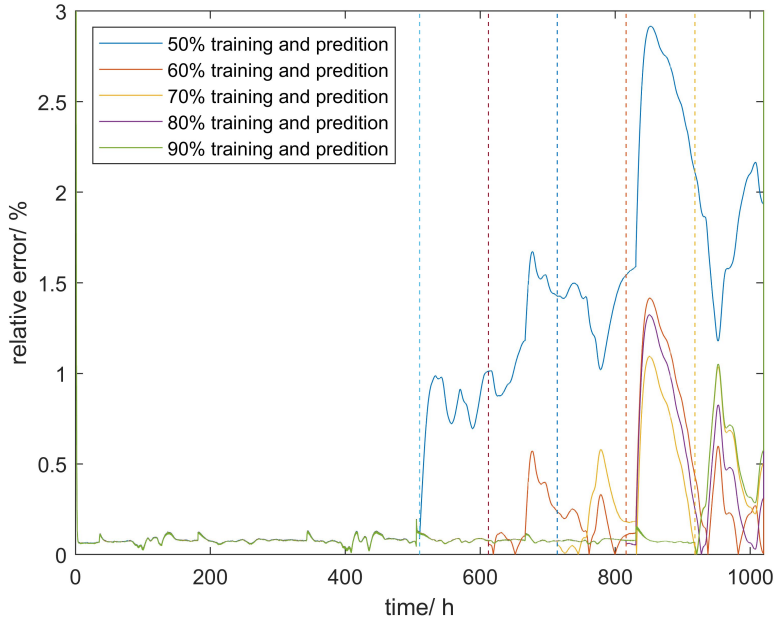
4.4.2.3 Logarithmic model for PEMFC2

For the logarithmic model, the training and prediction result and corresponding relative error for PEMFC2 are presented in figure 4.13(a) and 4.13(b), respectively. Again, the result is similar to that of the linear model. The predicted voltage is close to the actual degraded voltage, and the maximum relative error is less than 2.5%. The average MRE of this model is 0.5291%, which is slightly smaller than 0.5300% of the linear model, and it is also smaller than the MRE of the logarithmic model for PEMFC1. Compared with the result of PEMFC1 based on the logarithmic model, the logarithmic term can also accurately model the voltage recovery phenomenon of PEMFC2. Recall that the logarithmic model is also a little more accurate than the linear model for PEMFC1, we can conclude that the logarithmic model is the most suitable model when considering both PEMFC1 and PEMFC2. This agrees with the conclusion of the particle filter method [197].

The MRE and RMSE of different models under different training time for PEMFC2 are shown in table 4.5 and 4.6, respectively. It can be seen that although the MRE of the quadratic model is smaller than the linear model and logarithmic model under 50% training time, the error is high with longer training time. Therefore, the quadratic model is unsuitable for this case. The errors of the linear and logarithmic model are both small enough under all training time, which proves that they can achieve both short-term and long-term prediction, and they are robust for both two cases. The RMSE of the predictions shows the same tendency as MRE, i.e., the RMSE of linear and logarithmic

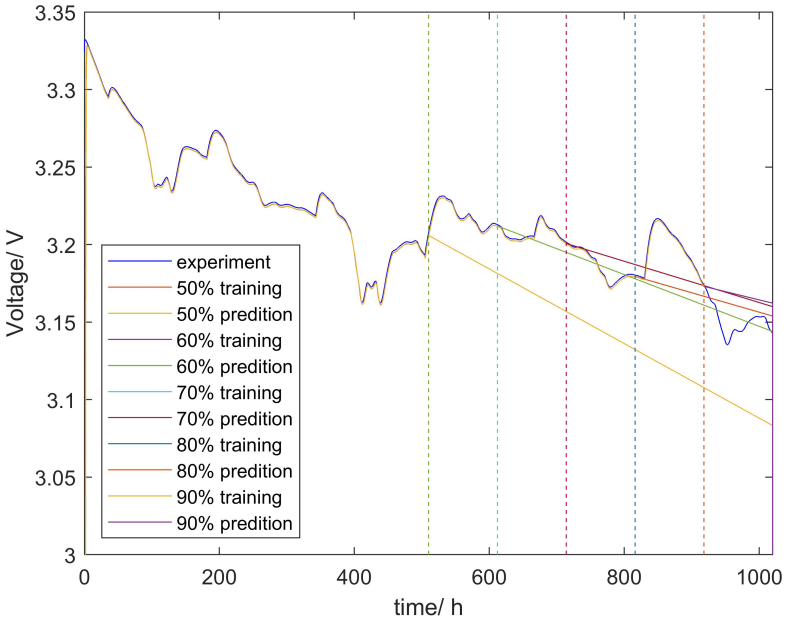


(a)

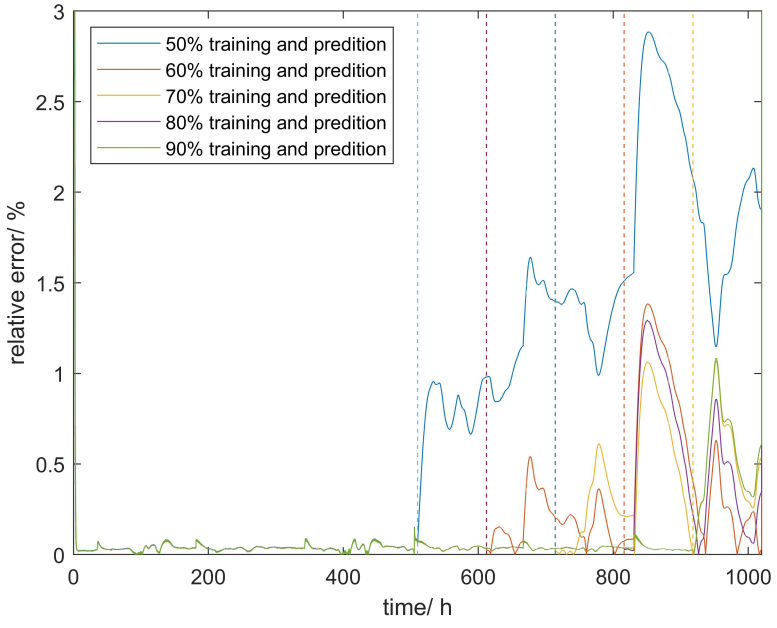


(b)

FIGURE 4.12: (a) The voltage prediction result based on FDKF under linear model and (b) the relative error under linear model for PEMFC2



(a)



(b)

FIGURE 4.13: (a) The voltage prediction result based on FDKF under logarithmic model and (b) the relative error under logarithmic model for PEMFC2

models are much smaller than the quadratic model, thus both linear and logarithmic model are accurate. Both MRE and RMSE of the linear model and logarithmic models are very close, so both MRE and RMSE can be used to evaluate the performance of prediction.

TABLE 4.5: MRE (%) of prediction under different model and different training time for PEMFC2

model	50%	60%	70%	80%	90%	average
quadratic model	0.7013	2.4230	1.7366	0.9723	0.7074	1.3081
linear model	0.8867	0.3784	0.4506	0.5611	0.3729	0.5300
logarithmic model	0.8545	0.3664	0.4595	0.5610	0.4041	0.5291

TABLE 4.6: RMSE (V) of prediction under different model and different training time for PEMFC2

model	50%	60%	70%	80%	90%	average
quadratic model	0.0653	0.1668	0.1110	0.0442	0.0237	0.0822
linear model	0.0335	0.0086	0.0185	0.0131	0.0138	0.0175
logarithmic model	0.0325	0.0090	0.0195	0.0139	0.0147	0.0179

4.4.3 Comparison with extended Kalman filter

Kalman filter can also be applied to degradation prediction in the time domain. However, a normal Kalman filter cannot solve non-linear problems. As an adaptation of the Kalman filter, EKF takes the Jacobi matrix of the nonlinear equation, so that it can be applied to a slightly non-linear problem. The EKF method has been applied to the fuel cell prognosis and proved advanced [195] [215] [174], thus it is applied to the PEMFC prognosis problem here, and it is carried out with three models for both PEMFC1 and PEMFC2 experiments. The computation time as well as the relative error of prediction are obtained and compared with the FDKF method.

The total number of data used in this research is 1120000 and 1040000 for PEMFC1 and PEMFC2, respectively. Therefore, the sampling frequency is 0.278 Hz. For both FDKF method and EKF method, the calculations are all carried out under the same test conditions, i.e. a personal computer with processor Intel(R)Xeon(R)W-2123CPU@3.60GHz; memory RAM 32Go; operation system Window 10 education and Matlab R2018a. For the linear model, quadratic model, and logarithmic model, the

voltage degradation is predicted for both PEMFC1 and PEMFC2 by FDKF and EKF, respectively. The computation time is shown in table 4.7. It can be seen that the FDKF method uses much less time than EKF for both experiments under all three models. The FDKF method can give results in several seconds, while the EKF method needs thousands of seconds. As has been addressed, the FDKF method is more efficient because it can process the data in groups, and it moves forward with R data at each step, which is the advantage of FDKF.

TABLE 4.7: Computation time (s) under different degradation models for EKF and FDKF

Method	linear model	quadratic model	logarithmic model	average
FDKF for PEMFC1	5.269	4.879	13.60	7.915
EKF for PEMFC1	1506	2456	2452	2138
FDKF for PEMFC2	4.319	4.422	11.20	6.647
EKF for PEMFC2	1507	1932	1895	1628

Compared with the EKF method, the FDKF method also has better performance in terms of accuracy. The MRE and RMSE under 50%, 60%, 70%, 80%, and 90% training time are averaged to represent the performance of the method, and they are shown in table 4.8 and 4.9, respectively. As we can see, the relative error of the EKF method under the quadratic model is very big, while that of the FDKF method is acceptable. Thus FDKF method can be used under more models. Furthermore, the error of the FDKF method is smaller than the EKF method under all three models, and the RMSE can give the same conclusion, which proves that the proposed method is advantageous. This is due to the different characteristics of FDKF and EKF method. The FDKF method deals with the data in the frequency domain, and it finds the basic structure of the data in the frequency domain so that to find the invariance of the data. For EKF, it uses the Taylor expansion to replace the non-linear function, then only the first-order term is reserved and the terms of higher order are neglected. In this way, a linear system can be obtained and the Kalman filter can be applied to the linearised system. The performance of EKF is worse than FDKF in our case because the real system is highly non-linear, and the linearisation process of the EKF method will introduce more error because of the neglected high-order terms. Therefore, the prediction error of the EKF method is bigger.

TABLE 4.8: Averaged MRE (%) under different degradation models for EKF and FDKF

Method	linear model	quadratic model	logarithmic model	average
FDKF for PEMFC1	0.6131	0.4419	0.5999	0.5516
EKF for PEMFC1	0.7718	4.2740	0.7717	1.9392
FDKF for PEMFC2	0.5300	1.3081	0.5291	0.7891
EKF for PEMFC2	1.0775	6.1435	1.0755	2.7655

TABLE 4.9: Averaged RMSE (V) under different degradation models for EKF and FDKF

Method	linear model	quadratic model	logarithmic model	average
FDKF for PEMFC1	0.0327	0.0181	0.0320	0.0276
EKF for PEMFC1	0.0345	0.1637	0.0345	0.0776
FDKF for PEMFC2	0.0175	0.0822	0.0179	0.0392
EKF for PEMFC2	0.0412	0.2344	0.0412	0.1056

4.4.4 Prediction horizon

According to the references [216], the prediction horizon (PH) is an important index for prognosis, thus the PH is calculated and compared with the references here. The PH is a measurement of the accuracy of the remaining useful life (RUL). At different prediction time, the real RUL and predicted RUL can be plotted to see the performance of the prediction. Usually, an acceptable error is set, which is α percent of the end of life (EOL). After a certain prediction time, the predicted RULs will all be within the acceptable error range; consequently, we can indicate how long the acceptable prediction can be obtained and guaranteed. In this work α is set as 10% according to most research. The EOL can be declared when the performance of the PEMFC is smaller than a critical line. In this work the critical line is set as 95% of the initial voltage, and the EOL is 936 hours. According to the proposed method, the RUL prediction and PH of PEMFC2 can be given in figure 4.14. The predicted RUL is within the acceptable region after 529 hours, and the PH can be obtained as 407 hours.

In the reference [217], the particle filter was applied and the voltage recovery phenomenon was considered in the PEMFC prognosis. In their research, the PH without consideration of recovery is 310 hours, and the PH calculated by the model with recovery is 380 hours. Therefore, the PH of the proposed method is longer, which means that the proposed method can give accurate prediction earlier, thus it is more accurate.

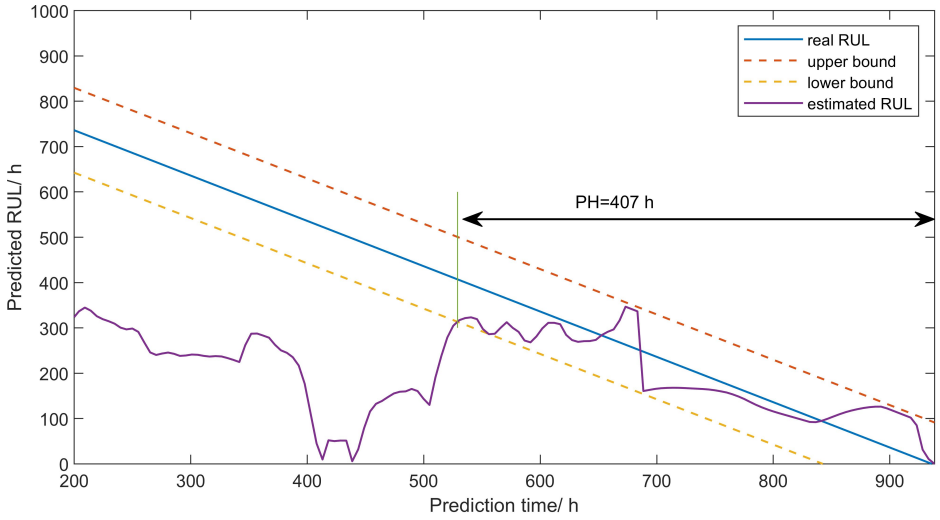


FIGURE 4.14: The prediction horizon for PEMFC2 by logarithmic model

4.5 Conclusions

In this chapter, we proposed a novel model-driven method for the PEMFC prognosis problem, which is based on FDKF and voltage degradation model. The advantage of this method is that it can process data in groups; therefore it is more efficient. The proposed method is verified by two case studies, and the voltage degradation is predicted both in the short-term and long-term under 4 degradation models. The result of the FDKF method is also compared with EKF and literature. With the proposed FDKF method, the quadratic model is only effective for PEMFC1, while both the linear and logarithmic models give acceptable results for both PEMFC1 and PEMFC2. In our research, the relative error of voltage degradation prediction is always smaller than 1.8% for PEMFC1, and it is smaller than 3% for PEMFC2. It proves that the proposed method is accurate and robust for both constant and dynamic current cases. The performance of the proposed FDKF method outweighs the EKF method both in terms of efficiency and accuracy. The proposed method can give results in several seconds, which is much more efficient than the EKF method. More importantly, the result of the FDKF method is also more accurate than the EKF method. The PH is also calculated and compared with other research, which proves that the proposed method is more accurate.

General Conclusion

In this thesis, we focus on PEMFC health management, more specifically, the diagnosis methods to distinguish operating faults and the prognosis algorithms to predict the long-term performance. As PEMFC is one of the most important technologies that can reduce the carbon dependence of energy systems, the diagnosis and prognosis methods are critical for increasing PEMFC durability and reliability. Three diagnostic methods and one prognostic method are proposed. The first diagnosis is based on the voltage fluctuation data. Two other diagnostic methods are based on EIS data, and they are validated in real time on an embedded system. The prognosis algorithm is based on voltage data and FDKF.

The main contributions and conclusions of the thesis can be listed as follows.

1. First of all, after reading massive relevant research works, we provide a basic introduction to the PEMFC diagnosis and prognosis. The basic knowledge is presented, including the PEMFC structure and mechanism, PEMFC faults, PEMFC application situation, and the position of diagnosis and prognosis in whole PEMFC health management. Further, the existing research works about experimental diagnostic tools and diagnostic algorithms are reviewed. Based on the analysis of former studies, the main challenges of PEMFC diagnosis and prognosis are summarized. Also, to face the challenges, the main objectives and possible solutions are presented in this thesis.

2. A diagnosis method is proposed based on the voltage fluctuation data and AR model. The voltage fluctuation patterns can be extracted by AR model, and the AR model coefficients can be applied directly as diagnostic features. The experiment set up in this research is introduced, and the available variables

measured under various operating conditions data are explained in detail. Based on the AR model coefficients, several fault classification methods are applied to distinguish the fault class, including the faults relevant to different temperature, relative humidity, and stoichiometry. The diagnostic result is very accurate for both single-fault and multi-fault conditions, and it is superior to the singularity analysis method of former research based on the same data.

3. Two diagnosis methods based on EIS data are proposed and validated in real time on an industrial level embedded system. As it is wildly acknowledged, EIS is a good diagnostic technology as it can reflect the health state of different processes in the PEMFC. The focus of the proposed diagnosis methods is to extract important diagnostic features from the original EIS data. The first feature extraction method is based on ECM. A novel fractional-order ECM is proposed, and the resistance of the membrane, anode, cathode, and gas diffusion are all considered in this model. The parameter fitting process based on GA-LM is proposed and validated, and the EIS features can be extracted accurately. Further, the diagnosis based on ANFIS is proposed to distinguish different inner faults, including flooding, drying-out, and mass transport fault. As ANFIS is a classification method with little complicity, it is suitable for the diagnosis of a computational resource-limited system. Also, the diagnosis based on ECM and ANFIS is validated by implementation on an industrial-level DSP system in real time. The detailed implementation processes and materials are explained, and the real-time ability of the method is clearly presented. Therefore, it is a practical method that can be applied in real applications.

4. Another diagnosis method based on two quick detective EIS features is also proposed, i.e. zero-phase impedance that reflects membrane resistance, and turning phase that reflects gas transport condition. The two features are explained and analysed with the experimental data under different operating conditions, proving that they can reflect the health state in the PEMFC. Also, a diagnostic method based on KNN is proposed by using the two features, and the fault of different degree flooding and gas blockage can be distinguished. Since the proposed features require only a few frequencies of EIS detection, the entire diagnostic method offers an important computational advantage. This has been proved through its implementation on a real-time DSP system.

5. To predict the long-term development of PEMFC performance, a prognostic method is proposed. The method is based on voltage data and FDKF, which is a novel prediction algorithm. The main advantage of this method is that it can proceed to the data by groups in the frequency domain, so it is quicker and more accurate, and it can save the computation resources. Two experiment cases have been researched, including static case and dynamic case. The result of different prediction times are also compared and discussed. Also, the results are compared with another prognosis method, proving that the proposed prognosis method is superior.

All in all, this thesis is a contribution to real-time diagnosis, fault classification, and prognosis of PEMFC. The proposed methods have the potential to be applied in real applications. However, there are still some limitations and more research should be devoted to the PEMFC diagnosis and prognosis. Here are some interesting research directions and perspectives.

1. First of all, even though the voltage fluctuation pattern can be extracted by autoregressive model in our work, the physical mechanism behind the pattern is not clear yet. More research should be carried out about the electrochemical noise of the PEMFC system. As a lot of existing studies have proved, the basic mechanism of EN method is related to the electrochemical reactions in the system; thus it can reflect on the characteristics of the system. Not only the phenomenon research but also mechanism research should be developed. Also, more feature extraction methods should be proposed to obtain the basic diagnostic information, and more practical diagnosis methods can be developed based on it.

2. For diagnosis based on EIS, the proposed algorithms can be implemented on industrial level embedded system in real time. However, the accurate real-time detection of EIS data with light and cheap equipment is still a challenge, which also limits its application. Therefore, designing appropriate circuits and carrying out more experiments to achieve convenient integrated EIS on-board detection technology is important and interesting.

3. For prognosis, the voltage is applied to represent the PEMFC health state in this research and most of other research. However, except for the stack variables such as stack voltage or output power, more inner state parameters may be detected as indicators, so that the health state development can be predicted by physical mechanisms of the inner process. As the PEMFC system is a complex system, the output parameters may be affected by some unknown process, and the degradation of components should be researched from physical aspects so that to give more convincing predictions.

4. Further, only diagnosis and prognosis are researched in this thesis. Developing diagnosis and prognosis methods that can be combined with a control algorithm is also important and interesting. As was presented above, the diagnosis and prognosis are a part of health management strategies, and they can provide information basis for the following decision process. Therefore, combining the diagnosis and prognosis with a decision-making algorithm, i.e. the control algorithm, is a promising direction. By combining these two steps together, there is no need for human beings to get involved in the whole diagnosis and control process, which will be a great step toward automatic self-learning control technology.

List of Figures

List of Tables

Appendix A. Derivation of real and imaginary parts of EIS

To calculate the real part and imaginary part of cathode impedance under different frequencies, take $s = i\omega$ into equation 3.2, we have:

$$Z_c(w) = \frac{R_c}{1 + R_c C_c (i\omega)^{\alpha_c}} \tag{A.1}$$

According to Euler's formula:

$$e^{ix} = \cos x + i \sin x \tag{A.2}$$

Take $x = \frac{\pi}{2}$, we have:

$$e^{i\frac{\pi}{2}} = i \tag{A.3}$$

Take equation A.3 and $\omega = 2\pi f$ into equation A.1:

$$Z_c(f) = \frac{R_c}{1 + R_c C_c (2\pi f e^{i\frac{\pi}{2}})^{\alpha_c}} \tag{A.4}$$

Reorganize the equation, we have:

$$Z_c(f) = \frac{R_c}{1 + R_c C_c (2\pi f)^{\alpha_c} e^{i\frac{\pi\alpha_c}{2}}} \tag{A.5}$$

Again use Euler's formula A.2, and take $x = \frac{\pi\alpha_c}{2}$, we have:

$$Z_c(f) = \frac{R_c}{1 + R_c C_c (2\pi f)^{\alpha_c} (\cos\frac{\pi\alpha_c}{2} + i\sin\frac{\pi\alpha_c}{2})} \tag{A.6}$$

To get rid of the i in the denominator, times both the denominator and nominator by $1 + R_c C_c (2\pi f)^{\alpha_c} (\cos\frac{\pi\alpha_c}{2} - i\sin\frac{\pi\alpha_c}{2})$, we can obtain:

$$Z_c(f) = \frac{R_c + R_c^2 C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2} - iR_c^2 C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2}}{(1 + R_c C_c (2\pi f)^{\alpha_c} \cos\frac{\pi\alpha_c}{2})^2 + (R_c C_c (2\pi f)^{\alpha_c} \sin\frac{\pi\alpha_c}{2})^2} \tag{A.7}$$

Therefore, the real part and imaginary part can be given as equation 3.15 and 3.16, respectively.

Appendix B. Proof of stability and convergence of FDKF method

The FDKF can be considered as a variable step-size frequency domain adaptive filter [211] [207], thus the solution can also be given as:

$$\boldsymbol{W}(k+1) = A[\boldsymbol{W}(k) + \boldsymbol{Q}\boldsymbol{\mu}(k)\boldsymbol{\Lambda}^{-1}(k)\boldsymbol{C}^{H}(k)\boldsymbol{E}(k)] \tag{B.1}$$

Where

$$\boldsymbol{E}(k) = \boldsymbol{Y}(k) - \boldsymbol{C}(k)\boldsymbol{W}(k) \tag{B.2}$$

$$\boldsymbol{\Lambda}(k) = \boldsymbol{C}^{H}(k)\boldsymbol{C}(k) \tag{B.3}$$

$$\begin{aligned}\boldsymbol{\mu}(k) &= \boldsymbol{C}(k)\boldsymbol{P}(k)\boldsymbol{C}^{H}(k)[\boldsymbol{C}(k)\boldsymbol{P}(k)\boldsymbol{C}^{H}(k) + \boldsymbol{\Phi}_{SS}]^{-1} \\ &= diag\{[\mu_0(k), \mu_1(k), ..., \mu_M(k)]^{T}\}\end{aligned} \tag{B.4}$$

Where the $\boldsymbol{\mu}(k)$ is called the variable step-size matrix. Take equation 4.24 and 4.25 into equation B.1, we can get:

$$\boldsymbol{P}(k+1) = A^{2}[\boldsymbol{I} - \frac{1}{2}\boldsymbol{\mu}(k)]\boldsymbol{P}(k) + \boldsymbol{\Phi}_{\Delta\Delta}(k) \tag{B.5}$$

As all the matrices are diagonal, every diagonal item in $\boldsymbol{P}(k)$ can be given as:

$$P_i(k+1) = A^2 \left[1 - \frac{1}{2} \frac{P_i(k)}{P_i(k) + \Phi_{SS,i}(k)/|C_i(k)|^2}\right] \times P_i(k) + \Phi_{\Delta\Delta,i}(k) \tag{B.6}$$

It can be seen that the coefficient $A^2 \left[1 - \frac{1}{2} \frac{P_i(k)}{P_i(k)+\Phi_{SS,i}(k)/|C_i(k)|^2}\right] < 1$ is always satisfied, thus the $P_i(k)$ can achieve an equilibrium after iterations. To find the steady state $P_i(\infty)$, we suppose that the far-end signal and noise signal are stationary, so that the $\Phi_{SS,i}(k)/|C_i(k)|^2$ can be replaced as η_i. And the $\Phi_{\Delta\Delta,i}(k)$ can be replaced by $(1 - A^2)\Phi_{WW,i}(k)$. Therefore, the steady state is:

$$P_i(\infty) = A^2 \left(1 - \frac{1}{2} \frac{P_i(\infty)}{P_i(\infty) + \eta_i}\right) P_i(\infty) + (1 - A^2)\Phi_{WW,i} \tag{B.7}$$

Then this equation can be solved and the positive solution is:

$$P_i(\infty) = \frac{1 - A^2}{2 - A^2} \left(\Phi_{WW,i} - \eta_i + \sqrt{\Phi_{WW,i}^2 + \eta_i^2 + \frac{2\eta_i \Phi_{WW,i}}{1 - A^2}}\right) \tag{B.8}$$

The solution can be normalized by dividing both sides with $\Phi_{WW,i}$, thus the normalized prediction distance $\bar{P}_i(\infty) = P_i(\infty)/\Phi_{WW,i}$, and take $\delta_i = \Phi_{WW,i}/\eta_i$, and the solution can be given as:

$$\bar{P}_i(\infty) = \frac{1 - A^2}{2 - A^2} \left(1 - \frac{1}{\delta_i} + \sqrt{1 + \frac{1}{\delta_i^2} + \frac{2}{(1 - A^2)\delta_i}}\right) \tag{B.9}$$

Take the solution into equation B.4, the steady state step size can be given as:

$$\mu_i(\infty) = \frac{\delta_i \bar{P}_i(\infty)}{\delta_i \bar{P}_i(\infty) + 1} \tag{B.10}$$

Therefore for all $0 < A \leq 1$, the FDKF is stable and can converge to a finite constant, and the converge rate can be decided by A and δ_i.

Bibliography

[1] B. Shaffer. Global energy trends: Demands for scientific innovation. *MRS Energy and Sustainability*, 6, 2019.

[2] S. Iida and K. Sakata. Hydrogen technologies and developments in japan. *Clean Energy*, 3(2):105–113, 2019.

[3] D. Hissel and M. C. Pera. Diagnostic and health management of fuel cell systems: Issues and solutions. *Annual Reviews in Control*, 42:201–211, 2016.

[4] M. Jouin, R. Gouriveau, D. Hissel, M. C. Péra, and N. Zerhouni. Prognostics and health management of pemfc - state of the art and remaining challenges. *International Journal of Hydrogen Energy*, 38(35):15307–15317, 2013.

[5] F. Barbir. *PEM fuel cells: theory and practice*. Academic press, 2012. ISBN 0123877105.

[6] S. P. S. Badwal, S. S. Giddey, C. Munnings, A. I. Bhatt, and A. F. Hollenkamp. Emerging electrochemical energy conversion and storage technologies. *Frontiers in Chemistry*, 2, 2014.

[7] J. Chen, H. Liu, Y. Huang, and Z. Yin. High-rate roll-to-roll stack and lamination of multilayer structured membrane electrode assembly. *Journal of Manufacturing Processes*, 23:175–182, 2016.

[8] C. Nitsche, S. Schroedl, and W. Weiss. Onboard diagnostics concept for fuel cell vehicles using adaptive modelling. In *IEEE Intelligent Vehicles Symposium, 2004*, pages 127–132.

[9] J. Zhang, J. Wu, and H. Zhang. *PEM fuel cell testing and diagnosis*. Newnes, 2013. ISBN 0444536892.

[10] H. Ju, H. Meng, and C.Y. Wang. A single-phase, non-isothermal model for pem fuel cells. *International Journal of Heat and Mass Transfer*, 48(7):1303–1315, 2005.

[11] M. Jouin, R. Gouriveau, D. Hissel, M.C. Pera, and N. Zerhouni. Degradations analysis and aging modeling for health assessment and prognostics of pemfc. *Reliability Engineering and System Safety*, 148(APR.):78–95, 2016.

[12] H. Markötter, I. Manke, R. Kuhn, T. Arlt, N. Kardjilov, M. P. Hentschel, A. Kupsch, A. Lange, C. Hartnig, J. Scholta, and J. Banhart. Neutron tomographic investigations of water distributions in polymer electrolyte membrane fuel cell stacks. *Journal of Power Sources*, 219:120–125, 2012.

[13] S. Zhou and J. S. Dhupia. Online adaptive water management fault diagnosis of pemfc based on orthogonal linear discriminant analysis and relevance vector machine. *International Journal of Hydrogen Energy*, 45(11):7005–7014, 2020.

[14] A. Benmouna, M. Becherif, D. Depernet, F. Gustin, H. S. Ramadan, and S. Fukuhara. Fault diagnosis methods for proton exchange membrane fuel cell system. *International Journal of Hydrogen Energy*, 42(2):1534–1543, 2017.

[15] R. Shimoi, M. Masuda, K. Fushinobu, Y. Kozawa, and K. Okazaki. Visualization of the membrane temperature field of a polymer electrolyte fuel cell. *Journal of Energy Resources Technology*, 126(4):258–261, 2004.

[16] H. Lin, T.F. Cao, L. Chen, Y.L. He, and W.Q. Tao. In situ measurement of temperature distribution within a single polymer electrolyte membrane fuel cell. *International Journal of Hydrogen Energy*, 37(16):11871–11886, 2012.

[17] Y. Zhai, K. Bethune, G. Bender, and R. Rocheleau. Analysis of the so2contamination effect on the oxygen reduction reaction in pemfcs by electrochemical impedance spectroscopy. *Journal of The Electrochemical Society*, 159(5):B524–B530, 2012.

[18] A. p. c. UK. Fuel cell roadmap 2020, 2020.

[19] Z. Li. *Data-driven fault diagnosis for PEMFC systems*. Thesis, 2014. Thèse de doctorat Automatique Aix-Marseille 2014 2014AIXM4335.

[20] M. Jouin, R. Gouriveau, D. Hissel, M. Pera, and N. Zerhouni. Prognostics of proton exchange membrane fuel cell stacks in a particle filtering framework including characterization disturbances and voltage recovery, 2014.

[21] J. Wu, X. Z. Yuan, H. Wang, M. Blanco, J. J. Martin, and J. Zhang. Diagnostic tools in pem fuel cell research: Part i electrochemical techniques. *International Journal of Hydrogen Energy*, 33(6):1735–1746, 2008.

[22] J. Wu, X. Zi Yuan, H. Wang, M. Blanco, J. J. Martin, and J. Zhang. Diagnostic tools in pem fuel cell research: Part ii: Physical/chemical methods. *International Journal of Hydrogen Energy*, 33(6):1747–1757, 2008.

[23] J. A. Salva, A. Iranzo, F. Rosa, E. Tapia, E. Lopez, and F. Isorna. Optimization of a pem fuel cell operating conditions: Obtaining the maximum performance polarization curve. *International Journal of Hydrogen Energy*, 41(43):19713–19723, 2016.

[24] S. Wu, X. Peng, L. Mao, A. Liu, L. Dong, and T. Wang. On-line fault diagnosis of proton exchange membrane membrane fuel cell using polarization curve. In *2020 International Conference on Sensing, Measurement and Data Analytics in the era of Artificial Intelligence (ICSMD)*, pages 7–10.

[25] D. Bezmalinovic, B. Simic, and F. Barbir. Characterization of pem fuel cell degradation by polarization change curves. *Journal of Power Sources*, 294:82–87, 2015.

[26] M. Mohsin, R. Raza, M. Mohsin ul Mulk, A. Yousaf, and V. Hacker. Electrochemical characterization of polymer electrolyte membrane fuel cells and polarization curve analysis. *International Journal of Hydrogen Energy*, 45(45): 24093–24107, 2020.

[27] N. Fouquet, C. Doulet, and C. Nouillant. Model based pem fuel cell state-of-health monitoring via ac impedance measurements. *Journal of Power Sources*, 159(2):p.905–913, 2006.

[28] X. Yuan, J. C. Sun, M. Blanco, H. Wang, J. Zhang, and D. P. Wilkinson. Ac impedance diagnosis of a 500w pem fuel cell stack: Part i: Stack impedance. *Journal of Power Sources*, 161(2):920–928, 2006.

[29] X. Yuan, H. Wang, J. C. Sun, and J. Zhang. Ac impedance technique in pem fuel cell diagnosis—a review. *International Journal of Hydrogen Energy*, 32(17): 4365–4380, 2007.

[30] T. Kim, S. Lee, and H. Park. A study of water transport as a function of the micro-porous layer arrangement in pemfcs. *International Journal of Hydrogen Energy*, 35(16):8631–8643, 2010.

[31] P. Ren, P. Pei, Y. Li, Z. Wu, D. Chen, S. Huang, and X. Jia. Diagnosis of water failures in proton exchange membrane fuel cell with zero-phase ohmic resistance and fixed-low-frequency impedance. *Applied Energy*, 239:785–792, 2019.

[32] S. Laribi, K. Mammar, M. Hamouda, and Y. Sahli. Impedance model for diagnosis of water management in fuel cells using artificial neural networks methodology. *International Journal of Hydrogen Energy*, 41(38):17093–17101, 2016.

[33] P. Zoltowski. On the electrical capacitance of interfaces exhibiting constant phase element behaviour. *Journal of Electroanalytical Chemistry*, 443(1):149–154, 1998.

[34] J. Shan, R. Lin, X. Chen, and X. Diao. Eis and local resolved current density distribution analysis on effects of mpl on pemfc performance at varied humidification. *International Journal of Heat and Mass Transfer*, 127:1076–1083, 2018.

[35] E. O. Balogun, N. Hussain, J. Chamier, and P. Barendse. Performance and durability studies of perfluorosulfonic acid ionomers as binders in pemfc catalyst layers using electrochemical impedance spectroscopy. *International Journal of Hydrogen Energy*, 44(60):32219–32230, 2019.

[36] R. Pan, D. Yang, Y. Wang, and Z. Chen. Health degradation assessment of proton exchange membrane fuel cell based on an analytical equivalent circuit model. *Energy*, 207:118185, 2020.

[37] G. Dotelli, R. Ferrero, P. G. Stampino, S. Latorrata, and S. Toscani. Pem fuel cell drying and flooding diagnosis with signals injected by a power converter. *IEEE Transactions on Instrumentation and Measurement*, 64(8):2064–2071, 2015.

[38] D. Depernet, A. Narjiss, F. Gustin, D. Hissel, and M.C. Péra. Integration of electrochemical impedance spectroscopy functionality in proton exchange membrane fuel cell power converter. *International Journal of Hydrogen Energy*, 41(11):5378–5388, 2016.

[39] G. Dotelli, R. Ferrero, P. G. Stampino, S. Latorrata, and S. Toscani. Diagnosis of pem fuel cell drying and flooding based on power converter ripple. *IEEE Transactions on Instrumentation and Measurement*, 63(10):2341–2348, 2014.

[40] H. Homayouni, J. DeVaal, F. Golnaraghi, and J. Wang. Voltage reduction technique for use with electrochemical impedance spectroscopy in high-voltage fuel cell and battery systems. *IEEE Transactions on Transportation Electrification*, 4(2):418–431, 2018.

[41] N. Katayama and S. Kogoshi. Real-time electrochemical impedance diagnosis for fuel cells using a dc–dc converter. *IEEE Transactions on Energy Conversion*, 30(2):707–713, 2015.

[42] H. Wang, A. Gaillard, and D. Hissel. A review of dc/dc converter-based electrochemical impedance spectroscopy for fuel cell electric vehicles. *Renewable Energy*, 141:124–138, 2019.

[43] W. J. Wruck, R. M. Machado, and T. W. Chapman. Current interruption—instrumentation and applications. *Journal of The Electrochemical Society*, 134(3):539–546, 1987.

[44] T. M. Koehler, D. B. Jarrell, and L. J. Bond. High temperture ceramic fuel cell measurement and diagnostics for application to solid oxide fuel cell systems. Report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2001.

[45] C. Lagergren, G. Lindbergh, and D. Simonsson. Investigation of porous electrodes by current interruption: Application to molten carbonate fuel cell cathodes. *Journal of The Electrochemical Society*, 142(3):787–797, 1995.

[46] M. A. Rubio, A. Urquia, and S. Dormido. Diagnosis of pem fuel cells through current interruption. *Journal of Power Sources*, 171(2):670–677, 2007.

[47] P. T. Kissinger and W. R. Heineman. Cyclic voltammetry. *Journal of Chemical Education*, 60(9):702, 1983.

[48] X. Sun, R. Li, D. Villers, J. P. Dodelet, and S. Désilets. Composite electrodes made of pt nanoparticles deposited on carbon nanotubes grown on fuel cell backings. *Chemical Physics Letters*, 379(1):99–104, 2003.

[49] R. N. Carter, S. S. Kocha, F. Wagner, M. Fay, and H. A. Gasteiger. Artifacts in measuring electrode catalyst area of fuel cells through cyclic voltammetry. *ECS Transactions*, 11(1):403–410, 2019.

[50] T. Kim, W. Choi, H.C. Shin, J.Y. Choi, J. M. Kim, M.S. Park, and W.S. Yoon. Applications of voltammetry in lithium ion battery research. *J. Electrochem. Sci. Technol*, 11(1):14–25, 2020.

[51] Y.C. Park, K. Kakinuma, M. Uchida, H. Uchida, and M. Watanabe. Deleterious effects of interim cyclic voltammetry on pt/carbon black catalyst degradation during start-up/shutdown cycling evaluation. *Electrochimica Acta*, 123:84–92, 2014.

[52] C. Rice, Y. Tong, E. Oldfield, and A. Wieckowski. Cyclic voltammetry and 195pt nuclear magnetic resonance characterization of graphite-supported commercial fuel cell grade platinum electrocatalysts. *Electrochimica Acta*, 43(19):2825–2830, 1998.

[53] C. Roth, N. Martz, F. Hahn, J. M. Léger, C. Lamy, and H. Fuess. Characterization of differently synthesized pt-ru fuel cell catalysts by cyclic voltammetry, ftir spectroscopy, and in single cells. *Journal of The Electrochemical Society*, 149 (11):E433, 2002.

[54] F. A. de Bruijn, D. C. Papageorgopoulos, E. F. Sitters, and G. J. M. Janssen. The influence of carbon dioxide on pem fuel cell anodes. *Journal of Power Sources*, 110(1):117–124, 2002.

[55] H. Wu, D. Wexler, and H. Liu. Durability investigation of graphene-supported pt nanocatalysts for pem fuel cells. *Journal of Solid State Electrochemistry*, 15(5): 1057–1062, 2011.

[56] P. Pei, Z. Wu, Y. Li, X. Jia, D. Chen, and S. Huang. Improved methods to measure hydrogen crossover current in proton exchange membrane fuel cell. *Applied Energy*, 215:338–347, 2018.

[57] J. Kang and J. Kim. Membrane electrode assembly degradation by dry/wet gas on a pem fuel cell. *International Journal of Hydrogen Energy*, 35(23):13125–13130, 2010.

[58] R. Maizia, A. Dib, A. Thomas, and S. Martemianov. Statistical short-time analysis of electrochemical noise generated within a proton exchange membrane fuel cell. *Journal of Solid State Electrochemistry*, 22(6):1649–1660, 2018.

[59] B. Legros, P. X. Thivel, Y. Bultel, and R. P. Nogueira. First results on pemfc diagnosis by electrochemical noise. *Electrochemistry Communications*, 13(12): 1514–1516, 2011.

[60] R. Maizia, A. Dib, A. Thomas, and S. Martemianov. Proton exchange membrane fuel cell diagnosis by spectral characterization of the electrochemical noise. *Journal of Power Sources*, 342:553–561, 2017.

[61] M. A. Rubio, K. Bethune, A. Urquia, and J. St-Pierre. Proton exchange membrane fuel cell failure mode early diagnosis with wavelet analysis of electrochemical noise. *International Journal of Hydrogen Energy*, 41(33):14991–15001, 2016.

[62] E. Astafev. Frequency characteristics of hydrogen-air fuel cell electrochemical noise. *Fuel Cells*, 18(6):755–762, 2018. https://doi.org/10.1002/fuce.201800102.

[63] E. Astaf'ev. Electrochemical noise measurement of polymer membrane fuel cell under load. *Russian Journal of Electrochemistry*, 54(6):554–560, 2018.

[64] E. Astaf'ev. Comparison of approaches in electrochemical noise analysis using an air–hydrogen fuel cell. *Russian Journal of Electrochemistry*, 56:156–162, 2020.

[65] E. A. Astafev, A. E. Ukshe, E. V. Gerasimova, Y. A. Dobrovolsky, and R. A. Manzhos. Electrochemical noise of a hydrogen-air polymer electrolyte fuel cell operating at different loads. *Journal of Solid State Electrochemistry*, 22(6):1839–1849, 2018.

[66] R. Petrone, Z. Zheng, D. Hissel, M. C. Péra, C. Pianese, M. Sorrentino, M. Becherif, and N. Yousfi-Steiner. A review on model-based diagnosis methodologies for pemfcs. *International Journal of Hydrogen Energy*, 38(17): 7077–7091, 2013.

[67] Z. Zheng, R. Petrone, M. C. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, and M. Sorrentino. A review on non-model based diagnosis methodologies for pem fuel cell stacks and systems. *International Journal of Hydrogen Energy*, 38(21):8914–8926, 2013.

[68] S. X. Ding. *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science and Business Media, 2008. ISBN 354076304X.

[69] T. Génevé, J. Régnier, and C. Turpin. Fuel cell flooding diagnosis based on time-constant spectrum analysis. *International Journal of Hydrogen Energy*, 41(1): 516–523, 2016.

[70] J. Lunze, S. Pröll, and F. Jarmolowitz. From structural analysis to observer-based residual generation for fault detection. In *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 491–498. ISBN 2162-1209.

[71] A. Q. Khan. *Observer-based fault detection in nonlinear systems*. Thesis, 2010.

[72] Q. Yang, A. Aitouche, and B. O. Bouamama. Residuals generation based on nonlinear analytical redundancy applied to air supply sub-system of fuel cell. In *7th Workshop on Advanced Control and Diagnosis*, pages 19–20, .

[73] Q. Yang, A. Aitouche, and B. O. Bouamama. Fault detection and isolation of pem fuel cell system by analytical redundancy. In *18th Mediterranean Conference on Control and Automation, MED'10*, pages 1371–1376, .

[74] Z. Li, R. Outbib, D. Hissel, and S. Giurgea. Diagnosis of pemfc by using data-driven parity space strategy. In *2014 European Control Conference (ECC)*, pages 1268–1273.

[75] A. Hernandez, D. Hissel, and R. Outbib. Modeling and fault diagnosis of a polymer electrolyte fuel cell using electrical equivalent analysis. *IEEE Transactions on Energy Conversion*, 25(1):148–160, 2010.

[76] T. Escobet, D. Feroldi, S. de Lira, V. Puig, J. Quevedo, J. Riera, and M. Serra. Model-based fault diagnosis in pem fuel cell systems. *Journal of Power Sources*, 192(1):216–223, 2009.

[77] S. De Lira, V. Puig, and J. Quevedo. Lpv model-based fault diagnosis using relative fault sensitivity signature approach in a pem fuel cell. *IFAC Proceedings Volumes*, 42(8):528–533, 2009.

[78] S. d. Lira, V. Puig, and J. Quevedo. Robust lpv model-based sensor fault diagnosis and estimation for a pem fuel cell system. In *2010 Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 819–824. ISBN 2162-1209.

[79] P. Polverino, E. Frisk, D. Jung, M. Krysander, and C. Pianese. Model-based diagnosis through structural analysis and causal computation for automotive polymer electrolyte membrane fuel cell systems. *Journal of Power Sources*, 357: 26–40, 2017.

[80] M. Buchholz, M. Eswein, and V. Krebs. Modelling pem fuel cell stacks for fdi using linear subspace identification. In *2008 IEEE International Conference on Control Applications*, pages 341–346, . ISBN 1085-1992.

[81] A. Rosich, R. Sarrate, and F. Nejjari. On-line model-based fault detection and isolation for pem fuel cell stack systems. *Applied Mathematical Modelling*, 38 (11):2744–2757, 2014.

[82] J. Liu, W. Luo, X. Yang, and L. Wu. Robust model-based fault diagnosis for pem fuel cell air-feed system. *IEEE Transactions on Industrial Electronics*, 63(5): 3261–3270, 2016.

[83] H. Oh, W.Y. Lee, J. Won, M. Kim, Y.Y. Choi, and S.B. Han. Residual-based fault diagnosis for thermal management systems of proton exchange membrane fuel cells. *Applied Energy*, 277:115568, 2020.

[84] R. P. Lippmann. Pattern classification using neural networks. *IEEE Communications Magazine*, 27(11):47–50, 1989.

[85] S. Jemei, D. Hissel, M. Pera, and J. M. Kauffmann. A new modeling approach of embedded fuel-cell power generators based on artificial neural network. *IEEE Transactions on Industrial Electronics*, 55(1):437–447, 2008.

[86] N. Yousfi Steiner, D. Hissel, P. Moçotéguy, and D. Candusso. Diagnosis of polymer electrolyte fuel cells failure modes (flooding and drying out) by neural networks modeling. *International Journal of Hydrogen Energy*, 36(4):3067–3075, 2011.

[87] S. Laribi, K. Mammar, Y. Sahli, and K. Koussa. Analysis and diagnosis of pem fuel cell failure modes (flooding and drying) across the physical parameters of electrochemical impedance model: Using neural networks method. *Sustainable Energy Technologies and Assessments*, 34:35–42, 2019.

[88] F. Z. Arama, K. Mammar, S. Laribi, A. Necaibia, and T. Ghaitaoui. Implementation of sensor based on neural networks technique to predict the pem fuel cell hydration state. *Journal of Energy Storage*, 27:101051, 2020.

[89] J. Y. Park, I. S. Lim, E. J. Choi, and M. S. Kim. Fault diagnosis of thermal management system in a polymer electrolyte membrane fuel cell. *Energy*, 214: 119062, 2021.

[90] J. Y. Park, Y. H. Lee, I. S. Lim, Y. S. Kim, and M. S. Kim. Prediction of local current distribution in polymer electrolyte membrane fuel cell with artificial neural network. *International Journal of Hydrogen Energy*, 46(39):20678–20692, 2021.

[91] A. Escobet, À. Nebot, and F. Mugica. Pem fuel cell fault diagnosis via a hybrid methodology based on fuzzy and pattern recognition techniques. *Engineering Applications of Artificial Intelligence*, 36:40–53, 2014.

[92] B. Davies, L. Jackson, and S. Dunnett. Expert diagnosis of polymer electrolyte fuel cells. *International Journal of Hydrogen Energy*, 42(16):11724–11734, 2017.

[93] M. Pei, C. Zhang, M. Hu, L. Jackson, and L. Mao. A fuzzy logic-based method for proton exchange membrane fuel cell fault diagnosis. In *2020 International Conference on Sensing, Measurement and Data Analytics in the era of Artificial Intelligence (ICSMD)*, pages 1–6.

[94] G. A. Rubio and W. E. Agila. A fuzzy model to manage water in polymer electrolyte membrane fuel cells. *Processes*, 9(6), 2021.

[95] S. Frizzo Stefenon, R. Zanetti Freire, L. dos Santos Coelho, L. H. Meyer, R. Bartnik Grebogi, W. Gouvêa Buratto, and A. Nied. Electrical insulator fault forecasting based on a wavelet neuro-fuzzy system. *Energies*, 13(2), 2020.

[96] K. K. Justesen, S. J. Andreasen, and S. L. Sahlin. Modeling of a htpem fuel cell using adaptive neuro-fuzzy inference systems. *International Journal of Hydrogen Energy*, 40(46):16814–16819, 2015.

[97] K. Mammar and S. Laribi. Application of adaptive neuro-fuzzy inference system techniques to predict water activity in proton exchange membrane fuel cell. *Journal of Electrochemical Energy Conversion and Storage*, 15(4), 2018.

[98] T. Wilberforce and A. G. Olabi. Performance prediction of proton exchange membrane fuel cells (pemfc) using adaptive neuro inference system (anfis). *Sustainability*, 12(12), 2020.

[99] 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. *Applied Energy*, 231:102–115, 2018.

[100] X. Zhao, L. Xu, J. Li, C. Fang, and M. Ouyang. Faults diagnosis for pem fuel cell system based on multi-sensor signals and principle component analysis method. *International Journal of Hydrogen Energy*, 42(29):18524–18531, 2017.

[101] Z. Zheng, R. Petrone, M. C. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, and M. Sorrentino. A review on non-model based diagnosis methodologies for pem fuel cell stacks and systems. *International Journal of Hydrogen Energy*, 38(21):8914–8926, 2013.

[102] J. Chen and B. Zhou. Diagnosis of pem fuel cell stack dynamic behaviors. *Journal of Power Sources*, 177(1):83–95, 2008.

[103] A. H. Detti, S. Jemei, S. Morando, and N. Y. Steiner. Classification based method using fast fourier transform (fft) and total harmonic distortion (thd) dedicated to proton exchange membrane fuel cell (pemfc) diagnosis. In *2017 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6, .

[104] N. J. Steffy, S. V. Selvaganesh, M. Kumar L, and A. K. Sahu. Online monitoring of fuel starvation and water management in an operating polymer electrolyte membrane fuel cell by a novel diagnostic tool based on total harmonic distortion analysis. *Journal of Power Sources*, 404:81–88, 2018.

[105] E. Pahon, N. Yousfi Steiner, S. Jemei, D. Hissel, and P. Moçoteguy. A signal-based method for fast pemfc diagnosis. *Applied Energy*, 165:748–758, 2016.

[106] N. Y. Steiner, D. Hissel, P. Moçotéguy, and D. Candusso. Non intrusive diagnosis of polymer electrolyte fuel cells by wavelet packet transform. *International Journal of Hydrogen Energy*, 36(1):740–746, 2011.

[107] D. Benouioua, D. Candusso, F. Harel, and L. Oukhellou. Pemfc stack voltage singularity measurement and fault classification. *International Journal of Hydrogen Energy*, 39(36):21631–21637, 2014.

[108] D. Benouioua, D. Candusso, F. Harel, and L. Oukhellou. Fuel cell diagnosis method based on multifractal analysis of stack voltage signal. *International Journal of Hydrogen Energy*, 39(5):2236–2245, 2014.

[109] J. Kim and Y. Tak. Implementation of discrete wavelet transform-based discrimination and state-of-health diagnosis for a polymer electrolyte membrane fuel cell. *International Journal of Hydrogen Energy*, 39(20):10664–10682, 2014.

[110] M. Ibrahim, U. Antoni, N. Y. Steiner, S. Jemei, C. Kokonendji, B. Ludwig, P. Moçotéguy, and D. Hissel. Signal-based diagnostics by wavelet transform for proton exchange membrane fuel cell. *Energy Procedia*, 74:1508–1516, 2015.

[111] E. Pahon, D. Hissel, S. Jemei, and N. Yousfi-Steiner. Relative wavelet energy as a diagnosis tool for pem fuel cells. In *2016 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6.

[112] T. Ma, W. Lin, Y. Yang, K. Wang, and W. Jia. Water content diagnosis for proton exchange membrane fuel cell based on wavelet transformation. *International Journal of Hydrogen Energy*, 45(39):20339–20350, 2020.

[113] C. Damour, M. Benne, B. Grondin-Perez, M. Bessafi, D. Hissel, and J.P. Chabriat. Polymer electrolyte membrane fuel cell fault diagnosis based on empirical mode decomposition. *Journal of Power Sources*, 299:596–603, 2015.

[114] Z. Zheng, M.C. Péra, D. Hissel, and M. Becherif. Diagnosis of proton exchange membrane fuel cell (pemfc) stack based on fuzzy clustering. In *5 th International Conference FDFC*, pages 16–18, .

[115] Z. Zheng, M.C. Péra, D. Hissel, M. Becherif, K.S. Agbli, and Y. Li. A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks. *Journal of Power Sources*, 271:570–581, 2014.

[116] L. Huang, Q. Zeng, and R. Zhang. Fuel cell engine fault diagnosis expert system based on decision tree. In *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 282–286.

[117] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien. Linear discriminant analysis: A detailed tutorial. *AI Communications*, 30:169–190, 2017.

[118] J. Kim, I. Lee, Y. Tak, and B. H. 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):4280–4289, 2012.

[119] S. Morando, M. C. Pera, N. Y. Steiner, S. Jemei, D. Hissel, and L. Larger. Fuel cells fault diagnosis under dynamic load profile using reservoir computing. In *2016 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6, .

[120] S. Morando, M. C. Pera, N. Y. Steiner, S. Jemei, D. Hissel, and L. Larger. Reservoir computing optimisation for pem fuel cell fault diagnostic. In *2017 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–7, .

[121] C. Jeppesen, S. S. Araya, S. L. Sahlin, S. Thomas, S. J. Andreasen, and S. K. Kær. Fault detection and isolation of high temperature proton exchange membrane fuel cell stack under the influence of degradation. *Journal of Power Sources*, 359:37–47, 2017.

[122] M. Shao, X.J. Zhu, H.F. Cao, and H.F. Shen. An artificial neural network ensemble method for fault diagnosis of proton exchange membrane fuel cell system. *Energy*, 67:268–275, 2014.

[123] X. Gu, Z. Hou, and J. Cai. Data-based flooding fault diagnosis of proton exchange membrane fuel cell systems using lstm networks. *Energy and AI*, 4:100056, 2021.

[124] L. A. M. Riascos, M. G. Simoes, and P. E. Miyagi. On-line fault diagnostic system for proton exchange membrane fuel cells. *Journal of Power Sources*, 175(1): 419–429, 2008.

[125] L. A. M. Riascos, M. G. Simoes, and P. E. Miyagi. A bayesian network fault diagnostic system for proton exchange membrane fuel cells. *Journal of Power Sources*, 165(1):267–278, 2007.

[126] S. Wasterlain, D. Candusso, F. Harel, X. François, and D. Hissel. Diagnosis of a fuel cell stack using electrochemical impedance spectroscopy and bayesian networks. In *2010 IEEE Vehicle Power and Propulsion Conference*, pages 1–6, . ISBN 1938-8756.

[127] L. Mao, L. Jackson, and B. Davies. Fault diagnosis of a polymer electrolyte membrane fuel cell using bayesian network. 2017.

[128] J. Yu, Q. Li, W. Chen, J. Liu, J. Li, and C. Xu. Fault classification of pem fuel cell systems based on cs-svm method. In *2019 IEEE Sustainable Power and Energy Conference (iSPEC)*, pages 2265–2269.

[129] Z. Li, R. Outbib, D. Hissel, and S. Giurgea. Data-driven diagnosis of pem fuel cell: A comparative study. *Control Engineering Practice*, 28:1–12, 2014.

[130] Z. Li, S. Giurgea, R. Outbib, and D. Hissel. Online diagnosis of pemfc by combining support vector machine and fluidic model. *Fuel Cells*, 14(3):448–456, 2014. https://doi.org/10.1002/fuce.201300197.

[131] Z. Li, R. Outbib, S. Giurgea, and D. Hissel. Diagnosis for pemfc systems: A data-driven approach with the capabilities of online adaptation and novel fault detection. *IEEE Transactions on Industrial Electronics*, 62(8):5164–5174, 2015.

[132] Z. Li, R. Outbib, S. Giurgea, and D. Hissel. Fault diagnosis for pemfc systems in consideration of dynamic behaviors and spatial inhomogeneity. *IEEE Transactions on Energy Conversion*, 34(1):3–11, 2019.

[133] Z. Li, C. Cadet, and R. Outbib. Diagnosis for pemfc based on magnetic measurements and data-driven approach. *IEEE Transactions on Energy Conversion*, 34(2):964–972, 2019.

[134] Z. Li, R. Outbib, S. Giurgea, D. Hissel, S. Jemei, A. Giraud, and S. Rosini. Online implementation of svm based fault diagnosis strategy for pemfc systems. *Applied Energy*, 164:284–293, 2016.

[135] L. Mao, Z. Liu, D. Low, W. Pan, Q. He, L. Jackson, and Q. Wu. An evaluation method for feature selection in proton exchange membrane fuel cell fault diagnosis. *IEEE Transactions on Industrial Electronics*, pages 1–1, 2021.

[136] I. S. Lim, J. Y. Park, E. J. Choi, and M. S. Kim. Efficient fault diagnosis method of pemfc thermal management system for various current densities. *International Journal of Hydrogen Energy*, 46(2):2543–2554, 2021.

[137] R. Onanena, L. Oukhellou, E. Côme, D. Candusso, D. Hissel, and P. Aknin. Fault-diagnosis of pem fuel cells using electrochemical spectroscopy impedance. *IFAC Proceedings Volumes*, 45(21):651–656, 2012.

[138] J. Liu, Q. Li, W. Chen, and T. Cao. A discrete hidden markov model fault diagnosis strategy based on k-means clustering dedicated to pem fuel cell systems of tramways. *International Journal of Hydrogen Energy*, 43(27):12428–12441, 2018.

[139] W. Pan, Y. Y. A. Abuker, and L. Mao. Investigation of feature effectiveness in polymer electrolyte membrane fuel cell fault diagnosis. In *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, pages 1–5.

[140] Z. Liu, M. Pei, Q. He, Q. Wu, L. Jackson, and L. Mao. A novel method for polymer electrolyte membrane fuel cell fault diagnosis using 2d data. *Journal of Power Sources*, 482:228894, 2021.

[141] M. Buchholz, G. Pecheur, J. Niemeyer, and V. Krebs. Fault detection and isolation for pem fuel cell stacks using fuzzy clusters. In *2007 European Control Conference (ECC)*, pages 971–977, .

[142] D. Hissel, D. Candusso, and F. Harel. Fuzzy-clustering durability diagnosis of polymer electrolyte fuel cells dedicated to transportation applications. *IEEE Transactions on Vehicular Technology*, 56(5):2414–2420, 2007.

[143] Z. Zheng, R. Petrone, M. C. Pera, D. Hissel, M. Becherif, and C. Pianese. Diagnosis of a commercial pem fuel cell stack via incomplete spectra and fuzzy clustering. In *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, pages 1595–1600, . ISBN 1553-572X.

[144] R. Petrone, E. Pahon, F. Harel, S. Jemei, D. Chamagne, D. Hissel, and M. C. Pera. Data-driven multi-fault approach for h2/o2 pem fuel cell diagnosis. In *2017 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–5.

[145] F. Han, Y. Tian, Q. Zou, and X. Zhang. Research on the fault diagnosis of a polymer electrolyte membrane fuel cell system. *Energies*, 13(10), 2020.

[146] J. Liu, Q. Li, W. Chen, Y. Yan, and L. Jiang. Fault diagnosis of pemfc systems based on decision-making tree classifier. In *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pages 1–5.

[147] R.H. Lin, Z.X. Pei, Z.Z. Ye, C.C. Guo, and B.D. Wu. Hydrogen fuel cell diagnostics using random forest and enhanced feature selection. *International Journal of Hydrogen Energy*, 45(17):10523–10535, 2020.

[148] N. Zhou, Q. Shao, J. Zhou, and H. Changjie. Fault classification of proton exchange membrane fuel cells for vehicles based on xgboost. In *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, pages 1054–1058.

[149] H. Dang, R. Ma, D. Zhao, R. Xie, H. Li, and Y. Liu. A novel diagnosis method of proton exchange membrane fuel cells based on the pca and xgboost algorithm. In *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, pages 3951–3956. ISBN 2577-1647.

[150] J. Liu, Q. Li, W. Chen, Y. Yan, and X. Wang. A fast fault diagnosis method of the pemfc system based on extreme learning machine and dempster–shafer evidence theory. *IEEE Transactions on Transportation Electrification*, 5(1):271–284, 2019.

[151] L. Mao, L. Jackson, and B. Davies. Effectiveness of a novel sensor selection algorithm in pem fuel cell on-line diagnosis. *IEEE Transactions on Industrial Electronics*, 65(9):7301–7310, 2018.

[152] H. Wang, X.Z. Yuan, and H. Li. *PEM fuel cell diagnostic tools*. CRC press, 2019. ISBN 1439839204.

[153] Z. Li, R. Outbib, S. Giurgea, D. Hissel, A. Giraud, and P. Couderc. Fault diagnosis for fuel cell systems: A data-driven approach using high-precise voltage sensors. *Renewable Energy*, 135:1435–1444, 2019.

[154] R. Ma, H. Dang, R. Xie, L. Xu, and D. Zhao. Online fault diagnosis for open-cathode pemfc systems based on output voltage measurements and data-driven method. *IEEE Transactions on Transportation Electrification*, pages 1–1, 2021.

[155] L. Mao, Z. Liu, D. Low, W. Pan, Q. He, L. Jackson, and Q. Wu. Evaluation method for feature selection in proton exchange membrane fuel cell fault diagnosis. *IEEE Transactions on Industrial Electronics*, 69(5):5277–5286, 2022.

[156] E. Astafev, A. Ukshe, R. Manzhos, Y. A. Dobrovolsky, S. Lakeev, and S. Timashev. Flicker noise spectroscopy in the analysis of electrochemical noise of hydrogen-air pem fuel cell during its degradation. *Int. J. Electrochem. Sci*, 12(3):1742, 2017.

[157] D. Benouioua, D. Candusso, F. Harel, X. François, and P. Picard. Characterization of low and high frequency phenomena in a pem fuel cell using singularity analysis of stack voltage. *Journal of Energy Storage*, 28:101298, 2020.

[158] S. Martemianov, F. Maillard, A. Thomas, P. Lagonotte, and L. Madier. Noise diagnosis of commercial li-ion batteries using high-order moments. *Russian Journal of Electrochemistry*, 52(12):1122–1130, 2016.

[159] D. Benouioua, D. Candusso, F. Harel, P. Picard, and X. François. On the issue of the pemfc operating fault identification: Generic analysis tool based on voltage pointwise singularity strengths. *International Journal of Hydrogen Energy*, 43 (25):11606–11613, 2018.

[160] B. Kaulakys. Autoregressive model of 1/f noise. *Physics Letters A*, 257(1):37–42, 1999.

[161] V. V. Morariu, L. Buimaga-Iarinca, C. VamoŞ, and Ş. M. ŞOltuz. Detrended fluctuation analysis of autoregressive processes. *Fluctuation and Noise Letters*, 07(03):L249–L255, 2007.

[162] Z. Qibin and Z. Liqing. Ecg feature extraction and classification using wavelet transform and support vector machines. In *2005 International Conference on Neural Networks and Brain*, volume 2, pages 1089–1092.

[163] A. H. Detti, N. Y. Steiner, L. Bouillaut, A. B. Same, and S. Jemei. Fuel cell performance prediction using an autoregressive moving-average arma model. In *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–5, . ISBN 1938-8756.

[164] Y.P. Yang, F.C. Wang, H.P. Chang, Y.W. Ma, and B.J. Weng. Low power proton exchange membrane fuel cell system identification and adaptive control. *Journal of Power Sources*, 164(2):761–771, 2007.

[165] D. C. Baillie and J. Mathew. A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings. *Mechanical Systems and Signal Processing*, 10(1):1–17, 1996.

[166] Y. Sakamoto, M. Ishiguro, G. K. Dordrecht, and D. Reidel. *Akaike Information Criterion Statistics*. 1986.

[167] E. J. Hannan and B. G. Quinn. The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2): 190–195, 1979. https://doi.org/10.1111/j.2517-6161.1979.tb01072.x.

[168] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[169] C. Brunetto, G. Tina, G. Squadrito, and A. Moschetto. Pemfc diagnostics and modelling by electrochemical impedance spectroscopy. In *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No.04CH37521)*, volume 3, pages 1045–1050 Vol.3.

[170] E. Lechartier, R. Gouriveau, M. Pera, D. Hissel, and N. Zerhouni. Static and dynamic modeling of a pemfc for prognostics purpose. In *Vehicle Power and Propulsion Conference*.

[171] M. Zhiani, S. Majidi, V. B. Silva, and H. Gharibi. Comparison of the performance and eis (electrochemical impedance spectroscopy) response of an activated pemfc (proton exchange membrane fuel cell) under low and high thermal and pressure stresses. *Energy*, 97:560–567, 2016.

[172] Q. Wang, Z. Hu, L. Xu, J. Li, Q. Gan, X. Du, and M. Ouyang. A comparative study of equivalent circuit model and distribution of relaxation times for fuel cell impedance diagnosis. *International Journal of Energy Research*, n/a(n/a), 2021. https://doi.org/10.1002/er.6825.

[173] T. Poinot and J. C. Trigeassou. A method for modelling and simulation of fractional systems. *Signal Processing*, 83(11):2319–2333, 2003.

[174] Q. Zhu, M. Xu, W. Liu, and M. Zheng. A state of charge estimation method for lithium-ion batteries based on fractional order adaptive extended kalman filter. *Energy*, 187.

[175] M. U. Iftikhar, D. Riu, F. Druart, S. Rosini, Y. Bultel, and N. Retière. Dynamic modeling of proton exchange membrane fuel cell using non-integer derivatives. *Journal of Power Sources*, 160(2):1170–1182, 2006.

[176] S. M. Rezaei Niya and M. Hoorfar. Process modeling of electrodes in proton exchange membrane fuel cells. *Journal of Electroanalytical Chemistry*, 747:112–122, 2015.

[177] S. Wasterlain, D. Candusso, F. Harel, X. François, and D. Hissel. Diagnosis of a fuel cell stack using electrochemical impedance spectroscopy and bayesian networks. In *2010 IEEE Vehicle Power and Propulsion Conference*, pages 1–6, . ISBN 1938-8756.

[178] A. Nasser Eddine, B. Huard, J.D. Gabano, and T. Poinot. Initialization of a fractional order identification algorithm applied for lithium-ion battery modeling in time domain. *Communications in Nonlinear Science and Numerical Simulation*, 59:375–386, 2018.

[179] A. Nasser-Eddine, B. Huard, J.D. Gabano, and T. Poinot. A two steps method for electrochemical impedance modeling using fractional order system in time and frequency domains. *Control Engineering Practice*, 86:96–104, 2019.

[180] J. H. Lee, J. H. Lee, W. Choi, K. W. Park, H. Y. Sun, and J. H. Oh. Development of a method to estimate the lifespan of proton exchange membrane fuel cell using electrochemical impedance spectroscopy. *Journal of Power Sources*, 195(18): 6001–6007, 2010.

[181] D. Malevich, E. Halliop, B. A. Peppley, J. G. Pharoah, and K. Karan. Investigation of charge-transfer and mass-transport resistances in pemfcs with microporous layer using electrochemical impedance spectroscopy. *Journal of The Electrochemical Society*, 156(2):B216, 2009.

[182] G. J. Brug, A. L. G. van den Eeden, M. Sluyters-Rehbach, and J. H. Sluyters. The analysis of electrode impedances complicated by the presence of a constant phase element. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 176(1):275–295, 1984.

[183] A. Ranganathan. The levenberg-marquardt algorithm. *Tutoral on LM algorithm*, 11(1):101–110, 2004.

[184] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, volume 4, pages 1942–1948. IEEE. ISBN 0780327683.

[185] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[186] M. J. D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(2): 155–162, 1964.

[187] Y. Ao, S. Laghrouche, and D. Depernet. Diagnosis of proton exchange membrane fuel cell system based on adaptive neural fuzzy inference system and electrochemical impedance spectroscopy. *Energy Conversion and Management*, 256:115391, 2022.

[188] C. Cadet, S. Jemeï, F. Druart, and D. Hissel. Diagnostic tools for pemfcs: from conception to implementation. *International Journal of Hydrogen Energy*, 39(20): 10613–10626, 2014.

[189] P. Manganiello, G. Petrone, M. Giannattasio, E. Monmasson, and G. Spagnuolo. Fpga implementation of the eis technique for the on-line diagnosis of fuel-cell systems. In *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, pages 981–986. ISBN 2163-5145.

[190] S. M. Rezaei Niya and M. Hoorfar. Study of proton exchange membrane fuel cells using electrochemical impedance spectroscopy technique – a review. *Journal of Power Sources*, 240:281–293, 2013.

[191] S. M. Rezaei Niya, R. K. Phillips, and M. Hoorfar. Process modeling of the impedance characteristics of proton exchange membrane fuel cells. *Electrochimica Acta*, 191:594–605, 2016.

[192] T. Sutharssan, D. Montalvao, Y. K. Chen, W.C. Wang, C. Pisac, and H. Elemara. A review on prognostics and health monitoring of proton exchange membrane fuel cell. *Renewable and Sustainable Energy Reviews*, 75:440–450, 2017.

[193] R. H. Lin, X. N. Xi, P. N. Wang, B. D. Wu, and S. M. Tian. Review on hydrogen fuel cell condition monitoring and prediction methods. *International Journal of Hydrogen Energy*, 44, 2018.

[194] J. Chen, D. Zhou, C. Lyu, and C. Lu. A novel health indicator for pemfc state of health estimation and remaining useful life prediction. *International Journal of Hydrogen Energy*, 42(31):20230–20238, 2017.

[195] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama. Extended kalman filter for prognostic of proton exchange membrane fuel cell. *Applied Energy*, 164 (feb.15):220–227, 2016.

[196] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama. Fuel cell remaining useful life prediction and uncertainty quantification under an automotive profile. In *Conference of the IEEE Industrial Electronics Society*.

[197] M. Jouin, R. Gouriveau, D. Hissel, M.C. Péra, and N. Zerhouni. Prognostics of pem fuel cell in a particle filtering framework. *International Journal of Hydrogen Energy*, 39(1):481–494, 2014.

[198] K. Chen, S. Laghrouche, and A. Djerdir. Fuel cell health prognosis using unscented kalman filter: Postal fuel cell electric vehicles case study. *International Journal of Hydrogen Energy*, 44(3):1930–1939, 2019.

[199] Y. Wu, E. Breaz, F. Gao, and A. Miraoui. A modified relevance vector machine for pem fuel cell stack aging prediction. *IEEE Transactions on Industry Applications*, 52(3):2573–2581, 2016.

[200] Y. Wu, E. Breaz, F. Gao, D. Paire, and A. Miraoui. Nonlinear performance degradation prediction of proton exchange membrane fuel cells using relevance vector machine. *IEEE Transactions on Energy Conversion*, pages 1–1, 2016.

[201] S. Morando, S. Jemei, D. Hissel, R. Gouriveau, and N. Zerhouni. Proton exchange membrane fuel cell ageing forecasting algorithm based on echo state network. *International Journal of Hydrogen Energy*, 42(2):1472–1480, 2016.

[202] 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 Conversion and Management*, 195:810–818, 2019.

[203] Y. Cheng, N. Zerhouni, and C. Lu. A hybrid remaining useful life prognostic method for proton exchange membrane fuel cell. *International Journal of Hydrogen Energy*, 43(27):12314–12327, 2018.

[204] 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*, 138(nov.1):1175–1186, 2017.

[205] D. Zhang. Contribution to prognostics of proton exchange membrane fuel cells : approaches based on degradation information at multiple levels. 2018.

[206] Ibrahim, Mona, Steiner, N. Yousfi, Jemei, Samir, Hissel, and Daniel. Wavelet-based approach for online fuel cell remaining useful lifetime prediction. *IEEE Transactions on Industrial Electronics*, 2016.

[207] F. Yang, G. Enzner, and J. Yang. Frequency-domain adaptive kalman filter with fast recovery of abrupt echo-path changes. *IEEE Signal Processing Letters*, 24 (12):1778–1782, 2017.

[208] G. Bernardi, T. V. Waterschoot, J. Wouters, and M. Moonen. Adaptive feedback cancellation using a partitioned-block frequency-domain kalman filter approach with pem-based signal prewhitening. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 25(9):1480–1494, 2017.

[209] M. S. Jha, G. Dauphin-Tanguy, and B. Ould-Bouamama. Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. *Mechanical Systems and Signal Processing*, pages 301–329, 2016.

[210] R. Lin, B. Li, Y. P. Hou, and J. M. Ma. Investigation of dynamic driving cycle effect on performance degradation and micro-structure change of pem fuel cell. *International Journal of Hydrogen Energy*, 34(5):2369–2376, 2009.

[211] G. Enzner and P. Vary. Frequency-domain adaptive kalman filter for acoustic echo control in hands-free telephones. *Signal Processing*, 86(6):p. 1140–1156, 2006.

[212] W. Fan, K. Chen, J. Lu, and J. Tao. Efficient improvement of frequency-domain kalman filter. 2018.

[213] G. R, H. M, H. D, J. S, J. M, L. E, S. Morando, E. Pahon, M.C. Pera, and N. Zerhouni. Ieee phm 2014 data challenge: outline, experiments, scoring of results, winners. 2014.

[214] J. K. Kimotho, T. Meyer, and W. Sextro. Pem fuel cell prognostics using particle filter with model parameter adaptation. In *2014 International Conference on Prognostics and Health Management*, pages 1–6.

[215] H. Cherragui, M. Bressel, M. Hilairet, and S. Giurgea. Fuel cells remaining useful life real-time estimation using an extended kalman filter in a hardware in the loop platform. In *2017 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6.

[216] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel. Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and Health Management*, 1:2153–2648, 2010.

[217] D. Zhang, C. Cadet, C. Bérenguer, and N. Yousfi-Steiner. Some improvements of particle filtering based prognosis for pem fuel cells. 2016.