HAL open science

Lifespan Prediction of Proton Exchange Membrane Fuel Cell System

Zhiguang Hua

To cite this version:

Zhiguang Hua. Lifespan Prediction of Proton Exchange Membrane Fuel Cell System. Other. Université Bourgogne Franche-Comté, 2021. English. NNT : 2021UBFCA013 . tel-03793342

HAL Id: tel-03793342

https://theses.hal.science/tel-03793342

Submitted on 1 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

PRÉPARÉE A L'UNIVERSITÉ DE TECHNOLOGIE DE BELFORT-MONTBÉLIARD

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

Doctorat de Génie Electrique

Par
Zhiguang Hua

Prédiction de la durée de vie du système de pile à combustible à membrane échangeuse de protons basée sur la réseau echo state

Thèse présentée et soutenue à Belfort, le 08/12/2021

Composition du jury:

M. Christophe Turpin	Directeur de Recherche au CNRS, Laboratoire LAPLACE, Président du jury
M. Loïc Boulon	Professeur à l'Université du Québec à Trois-Rivières, Rapporteur
M. Belkacem Ould Bouamama	Professeur à l'Université de Lille 1, Rapporteur
Mme. Nada Zamel	Scientifique senior à Institut Fraunhofer pour les systèmes d'énergie solaire ISE, Examinateur
Mme. Huizhi Wang	Maître de conférences à Imperial College London, Examinateur
M. Fei Gao	Professeur à l'Université de Bourgogne Franche-Comté, Directeur de thèse
Mme. Marie-Cécile Péra	Professeur à l'Université de Bourgogne Franche-Comté, Codirectrice de thèse
Mme. Zhixue Zheng	Maître de Conférences à l'Université de Lorraine & CentraleSupélec, Encadrante de thèse

PH.D. THESIS OF THE UNIVERSITY BOURGOGNE FRANCHE-COMTÉ

PREPARED AT THE UNIVERSITY OF TECHNOLOGY OF BELFORT-MONTBÉLIARD

Doctoral School n°37
Engineering Sciences and Microtechnologies

Doctor of Philosophy (Ph.D.) in Electrical Engineering

By
Zhiguang Hua

Lifetime prediction of proton exchange membrane fuel cell system based on echo state network

Thesis presented and defended in Belfort, le 08/12/2021

Composition of jury:

M. Christophe Turpin	Senior researcher of CNRS, LAPLACE Laboratory, Chairman of jury
M. Loïc Boulon	Professor of University of Quebec at Trois-Rivières, Reviewer
M. Belkacem Ould Bouamama	Professor of University of Lille 1, Reviewer
Mrs. Nada Zamel	Senior Scientist of Fraunhofer Institute for Solar Energy Systems ISE, Examiner
Mrs. Huizhi Wang	Senior Lecturer of Imperial College London, Examiner
M. Fei Gao	Professor of University of Burgundy Franche-Comté, Supervisor
Mrs. Marie-Cécile Péra	Professor of University of Burgundy Franche-Comté, Co-supervisor
Mrs. Zhixue Zheng	Associate Professor of University of Lorraine & CentraleSupélec, Co-supervisor

SPIM

Titre: Prédiction de la durée de vie du système de pile à combustible à membrane échangeuse de protons basée sur la réseau echo state

Mots-clés: Piles à combustible, Pronostic, Durée de vie utile restante, Data-driven, Réseau echo state

Résumé: La pile à combustible à membrane échangeuse de protons est considérée comme un dispositif de conversion de puissance prometteur dans différents domaines, néanmoins, la durée de vie limitée est l'un des principaux obstacles à leur déploiement industriel à grande échelle. Le pronostic basée sur les données vise à estimer la durée de vie utile restante sans avoir besoin d'une connaissance complète des phénomènes physiques se produisant dans le système. En tant que structure améliorée d'un réseau de neurones récurrent, l'echo state network a démontré de meilleures performances, en particulier en réduisant la complexité de calcul et en accélérant le taux de convergence. Des indicateurs couramment employés tels que la tension ou la puissance ne permettent pas toujours de donner une image fidèle de l'état de santé de la pile. Aussi, un nouvel indicateur de santé nommé taux de perte de puissance relatif, facile à calculer, est introduit dans cette thèse. Une structure d'algorithme l'echo state network multi-entrées et multi-sorties est développée et testée pour différents profils de mission de la pile. Sur la base de l'indicateur taux de perte de puissance relatif, une méthode combinée, à savoir la transformée en ondelettes discrète et l'echo state network est proposée pour traiter les caractéristiques multi-échelles de temps et améliorer la précision de la prédiction à long terme. Une solution associant l'approche par les données, un algorithme génétique, un echo state network et la transformée en ondelettes discrètes est proposée pour améliorer les performances de prédiction. Les performances des algorithmes proposés sont évaluées sur plusieurs cas d'études expérimentaux dans des conditions de fonctionnement en régime permanent, quasi-dynamique et dynamique.

Title: Lifetime prediction of proton exchange membrane fuel cell system based on echo state network

Keywords: Fuel cells, Prognostics, Remaining useful life, Data-driven, Echo state network

Abstract: The proton exchange membrane fuel cell is considered as a promising power conversion device in different areas, nevertheless, the limited lifetime is one of the key barriers to their large-scale industrial deployment. The data-driven prognostic method aims to estimate the remaining useful life without the need for complete knowledge about the system's physical phenomena. As an improved structure of the recurrent neural network, the echo state network has demonstrated better performances, especially in reducing the computational complexity and accelerating the convergence rate. To overcome the weakness of static health indicators commonly used (e.g., voltage and power), a convenient and practical health indicator named relative power-loss rate is proposed in this thesis. A multi-input and multi-output echo state network structure is developed and tested for the steady-state, quasi-dynamic, and full dynamic operating conditions of the fuel cell. Based on the relative power-loss rate indicator, a combined method, namely discrete wavelet transform and ensemble echo state network is proposed to deal with the multi-timescale features and improve the long-term prediction accuracy. A data-driven approach of discrete wavelet transform-echo state network-genetic algorithm is proposed to further improve the prediction performance. The performance of the proposed approaches is evaluated by different experimental datasets under steady-state, quasi-dynamic, and full dynamic operating conditions separately.

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

Université Bourgogne Franche-Comté
32, avenue de l'Observatoire
25000 Besançon

The value of a man should be seen in what he gives and not in what he is able to receive.

Albert Einstein (1879-1955)

Acknowledgement

The works of this thesis were done in FEMTO-ST Institute in the framework of FCLAB activities, Univ. Bourgogne Franche-Comté, UTBM, CNRS.

Firstly, I want to thank Prof. Christophe Turpin, Prof. Loïc Boulon, Prof. Belkacem Ould Bouamama, Dr. Nada Zamel, Dr. Huizhi Wang for their kindly help in preparation for my thesis defense. I am very grateful to my supervisors Prof. Fei Gao, Prof. Marie-Cécile Pera, and Dr. Zhixue Zheng. During my Master's period, I did some works about motor design and control. I knew nothing about fuel cells before I started this research topic. Nevertheless, the encouragements and supports from my supervisors help me to hold on explore the academic problems. I also remember the scene when we have the first group meeting, and they told me I can keep patient and study step by step. In the first year of my Ph.D., I was mainly busy reading a lot of literature. If I have some question, no matter how simple it is, they would explain to me carefully. Prof. Gao and Prof. Pera give me some insightful suggestions and Dr. Zheng usually discuss the details with me including the modeling methods, implementation processes, etc. All my improvement of research skills and the publication of papers cannot leave the patient guidance and encouragement from them. Besides, I would like to especially thank Dr. Elena Breaz for her suggestions on my future research career, Dr. Elodie Pahon for she provided the experimental datasets for my thesis, and Dr. Zhongliang Li from Aix-Marseille University for he gave me some insightful advice especially the prognostic methods.

Secondly, I would like to thank Prof. Manfeng Dou, Assoc. Prof. Dongdong Zhao, Assoc. Prof. Yuanlin Wang, Prof. Yigeng Huangfu, who are from Northwestern Polytechnical University, Xi'an, China. They give me a lot of help in pursuing my Master's Degree and give me the necessary encouragement to finish the Ph. D. Degree in France. Besides, I would like to thank all my dear colleagues in UTBM: Dr. Rui Ma, Dr. Chen Liu, Dr. Hanqing Wang, Dr. Hao Bai, Dr. Huan Luo, Dr. Yu Wu, Dr. Yang Zhou, Dr. Shengrong Zhuo, Mr. Qian Li, Mr. Tianhong Wang, Mr. Liangzhen Yin, Mr. Xinyang Hao, Mr. Xiaolei Ye, Mr. Rui Yang, Mr. Yuchen Liu, Mr. Ning Li, Mr. Zihan Peng, Mr. Lei Gao, for they have helped me a lot no matter in the academic researches and daily life.

Finally, I would like to convey my sincere gratitude to my family. My parents support me both financially and spiritually on the way to my dream. I hope they can keep fit during their works. My girlfriend Miss Jie SUN accompanies and supports me no matter what happens to me. I am so lucky to meet such a kind and pretty girl in my life. I hope she can do very well for her Ph.D. thesis and I believe that we will fulfill each other's dreams in the following years.

Publication list

International Journals

[1] **Z. Hua**, Z. Zheng, M.-C. Péra, and F. Gao, "Remaining useful life prediction of PEMFC systems based on the multi-input echo state network," *Appl. Energy*, vol. 265, p. 114791, May. 2020.

[2] **Z. Hua**, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, "Remaining useful life prediction of PEMFC systems under dynamic operating conditions," *Energy Convers. and Manage.*, vol. 231, p. 113825, Mar. 2021.

[3] **Z. Hua**, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, " Multi-timescale lifespan prediction for PEMFC systems under dynamic operating conditions," *IEEE Trans. Transp. Electrific.*, PP. 1-1, Aug. 2021.

[4] **Z. Hua**, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, " Lifespan prediction for proton exchange membrane fuel cells based on wavelet transform and echo state network," *IEEE Trans. Transp. Electrific.*, pp. 1-1, Oct. 2021.

International Conferences

[1] **Z. Hua**, Z. Zheng, F. Gao, and M.-C. Péra, "Challenges of the remaining useful life prediction for proton exchange membrane fuel cells," *in IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, pp. 6382-6387, Oct. 2019.

[2] **Z. Hua**, Z. Zheng, M.-C. Péra, and F. Gao, "Data-driven prognostics for PEMFC systems by different echo state network prediction structures," *in ITEC 2020 - IEEE Transportation Electrification Conference and Exposition*, pp. 495-500, Aug. 2020.

[3] **Z. Hua**, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, "Health indicators for PEMFC systems life prediction under both static and dynamic operating conditions," *in IECON 2020 - the 46th Annual Conference of the IEEE Industrial Electronics Society*, pp. 3963-3968, Oct. 2020.

[4] **Z. Hua**, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, "Lifetime prediction of PEMFC systems based on echo state network," *in FRH2 2021 - the French Research Network on Hydrogen Energy*, Jun. 2021.

Abbreviations

PEMFC	Proton exchange membrane fuel cells
FCEV	Fuel cell electric vehicle
μ-CHP	Micro-combined heat and power
ICE	Internal combustion engines
RUL	Remaining useful life
CBM	Condition-based maintenance
HI	Health indicator
AI	Artificial intelligence
EIS	Electrochemical impedance spectrometry
ECM	Equivalent circuit model
ECSA	Electro-chemical surface area
RSVAR	Regime switching vector autoregressive
UKF	Unscented Kalman filter
EKF	Extended Kalman filter
OCV	Open-circuit voltage
HIL	Hardware in the loop
PF	Particle filter
EMR	Energetic macroscopic representation
MPPT	Maximum power point tracking
PDF	Probability density function
APF	Adaptive particle filter
WT	Wavelet trans form
DWT	Discrete wavelet transform
NARNN	Non-linear autoregressive neural network
ARMA	Autoregressive and moving average
TDNN	Time delay neural network
PAM	Physical aging model
RPF	Regularized particle filter
LSSVM	Least square support vector machine
GP	Gaussian process
ANFIS	Adaptive neuro fuzzy inference systems
PSO	Particle swarm optimization
AUKF	Adaptive unscented Kalman filter
LSTM	Long short-term memory
ARIMA	Autoregressive integrated moving average
SW-ELM	Summation wavelet-extreme learning machine
ANN	Artificial neural networks

GMDH	Group method of data handling
SaDE	Self-adaptive differential evolutionary
GA	Genetic algorithm
GNNM	Grey neural network model
BPNN	Backpropagation neural network
MEA	Mind evolutionary algorithm
RVM	Relevance vector machine
SVM	Support vector machine
RNN	Recurrent neural network
G-LSTM	Grid long short-term memory
SAE-DNN	Sparse autoencoder-based deep neural network
ESN	Echo state network
BB-BC	Big bang and the big crunch
FSC	Cathode stoichiometric ratio
ANOVA	Analysis of variance
LPV	Linear parameter varying
SISO-ESN	Single-input and single-output ESN
MIMO-ESN	Multi-input and multi-output ESN
MAF	Moving average filtering
FT	Failure threshold
ANN	Artificial neural network
FFNN	Feed-forward neural network
BoL	Beginning of life
EoL	End of life
RPLR	Relative power-loss rate
SoH	State of health

Table of contents

General Introduction 1

Chapter 1. Comparative study on remaining useful life prediction for PEMFC system 11

1.1 PEMFC system and its degradation phenomena 11

1.1.1 Working principles of PEMFC system 11

1.1.2 Thermodynamics and degradation of the PEMFC stack 13

1.2 Challenges of RUL prediction and the datasets in this thesis 17

1.2.1 Prognostic and health management of PEMFC 17

1.2.2 Experimental datasets for RUL prediction 18

1.2.3 Challenges of remaining useful life prediction 24

1.3 Literature review of remaining useful life prediction 29

1.3.1 Model-based methods 29

1.3.2 Data-driven methods 38

1.4 Reasons about why choose the echo state network 43

1.4.1 The background of echo state network 43

1.4.2 Research fundamental of echo state network in PEMFC system 44

1.5 Objectives of this thesis 47

Chapter 2. Multi-input and multi-output echo state network of test-SS and test-QD 53

2.1 Brief introduction 53

2.2 Mathematical background of ESN 53

2.3 Implementation process of MIMO-ESN 59

2.4 Experimental results 62

2.4.1 MIMO-ESN under steady-state operating condition 62

2.4.2 MIMO-ESN under quasi-dynamic operating condition 66

2.5 Multi-input discussion 70

2.6 Chapter summary 73

Chapter 3. Double-input echo state network of three dynamic tests 75

3.1 Brief introduction 75

3.2 Dynamic health indicator extraction 75

3.3 Implementation process of DI-ESN 81
3.4 Experimental results of dynamic tests 83
3.5 Sensibility analysis of parameters 90
3.6 Chapter summary 93
Chapter 4. Multi-timescale predictions of ensemble echo state network 95
4.1 Brief introduction 95
4.2 Discrete wavelet transform 95
4.3 Implementation process of DWT-EESN 97
4.4 Experimental results under three dynamic tests 99
4.4.1 Long-term prediction of single ESN 99
4.4.2 Long-term prediction of DWT-EESN 101
4.5 Chapter summary 107
Chapter 5. Prediction efficiency improvement by genetic algorithm and discrete wavelet transform 109
5.1 Brief introduction 109
5.2 Genetic algorithm parameter optimization 110
5.3 Implementation process of DWT-ESN-GA 113
5.4 Experimental results under different mission profiles 117
5.4.1 Prediction results of three tests 117
5.5 Relative wavelet energy discussion 124
5.6 Chapter summary 125
Chapter 6. Conclusions and perspectives 127
6.1 Conclusions 127
6.2 Perspectives 129
List of figures 133
List of tables 139
References 141

General introduction

Energy and environment are the focal points attracting the eye of the international community. According to the survey of Economics and Technology Research Institute (ETRI) which belongs to China National Petroleum Corporation (CNPC), the world's population will rise to 9.77 billion in 2050, and energy demand will increase by 50 %. For the next 30 years, the demand for coal and oil would be declined steadily, and the share of natural gas and renewables consumption will gradually increase. Meanwhile, the global warming caused by carbon dioxide (CO_2) emissions is still tough, and a breakthrough in hydrogen energy technology including clean manufacturing, safe storage, and transportation would relieve the pressure [1]. The proportion of different types of energy consumption on the earth in the next 30 years is shown in Fig. I.

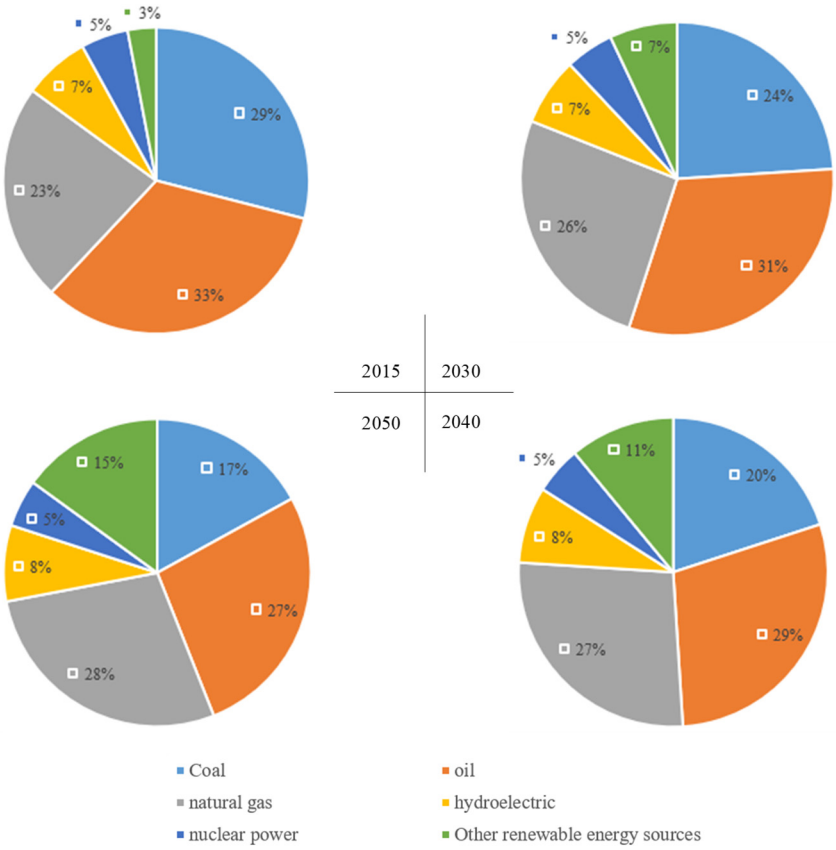


Fig. I Proportion of different types of energy consumption on the earth in the next 30 years [1].

Fuel cells (FCs) can reduce fossil fuel consumption and greenhouse gas emissions, and they are drawing more and more attention in the last few decades [2], [3]. The advantages of FCs are

- **High efficiency**. The FCs directly convert diverse fuels into electricity efficiently without internal combustion. In the internal combustion engines (ICE), the chemical energy in the fuels should be transferred into heat energy first. Then the heat energy would be converted into mechanical energy, and finally to produce

electrical energy. Thus, the transfer efficiency of FCs is higher than ICE. Nevertheless, the energy efficiency of PEMFC is much less than that of a battery pack.

- **High reliability**. The FCs can be regarded as the all-solid mechanical structure, and it means that there are no moving parts in FCs. Compared with the ICE, the FCs have the potential of high reliability and stability, and the FCs would be pretty quiet during their working because of no moving parts.
- **Environmental friendly.** In the FCs system, hydrogen (or other fuel) and oxygen (or fresh air) are the inputs, and the by-products are heat and water usually [4], [5].
- **Easy scaling between power and capacity.** The sizing in terms of power depends only on the fuel cell's size, and the capacity of FCs is determined by the size of fuel and oxidant tanks. Compared with batteries, the FCs have great potential applications for they can provide power for the demand system as small as a laptop computer (1-W range) and as large as a utility power station (megawatt range). Besides, the FCs can be quickly recharged by refueling which is very convenient for the users.

The classification of fuel cells and their properties are described in Tab. I. In all types of FCs, the proton exchange membrane fuel cells (PEMFC) system is the most popular one for its low operating temperature (50 °C - 80 °C) and solid electrolyte, which eliminates the risk of leakage. The PEMFC is welcomed in the area of fuel cell electric vehicles (FCEV), backup power, portable devices, etc. The total shipments and Megawatts (in MW) by application and by fuel cell type are shown in Fig. II.

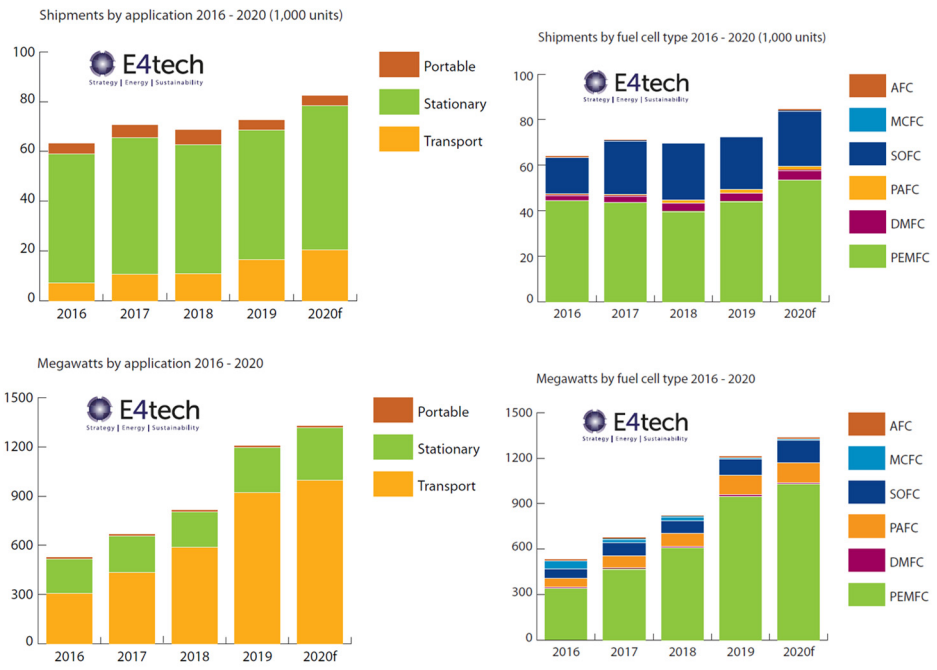


Fig. II The total shipments and Megawatts (in MW) by applications and by fuel cell types [6].

Tab. I The classification of fuel cells and their properties.

Kinds	PEMFC	AFC	PAFC	MCFC	. SOFC
Operating temperature	50 – 80 °C	65 – 200 °C	180 – 250 °C	600 – 700 °C	750 – 1000 °C
Moving ions	H^+	OH^-	H^+	CO_3^-	O^{2-}
Electrolyte	Solid (polymer membrane)	Liquid (KOH)	Liquid (H_3PO_4)	Liquid (carbonate ion salt)	Solid (oxide-based ceramic)
Fuel compatibility	H_2, CH_3OH	H_2	H_2	H_2, CH_4, other	H_2, CH_4, CO, other
Electrical efficiency	60 % (by H_2)	60 %	40 %	50 %	60 %
Advantages	Low operating temperature, High power-to-weight ratio	Non-noble metal catalysts, High efficiency	Insensitive to CO, Mature technology	Avoid using noble metal catalysts, No effects from CO	Avoid using noble metal catalysts, No effects from CO
Weaknesses	Sensitivity to CO, Need noble metal catalysts	Risk of electrolyte leakage, Sensitive to CO_2	Need noble metal catalysts, Cannot start at a low temperature	Cannot start at a low temperature, Risk of electrolyte leakage	Problem of sealing, Heat-resisting material
Application areas	Transportation, Portable and backup power	Space-flight, Submarines	Micro-combined heat and power, Distributed generation	Electric utility of military, Distributed generation	Auxiliary power units, Micro-combined heat, and power

Even though the PEMFC has some advantages over ICE and primary batteries, they also possess some bottlenecks in terms of cost and performance [7]-[9], as shown in Tab. II.

Tab. II Statistics about fuel cell light-duty vehicles (including cars) from FCH 2 JU [10].

Parameter	Unit	State of the art		FCH 2 JU target		
		2012	2017	2020	2024	2030
cost	€/kW	500	100	60	50	40
Areal power density	W/cm^2	-	1.0	1.5	1.8	2.0
PGM loading	g/kW	-	0.4	0.17	0.08	0.05
durability	h	2500	4000	5000	6000	7000
Cell volumetric power	kW/L	-	5.0	7.3	9.3	10.0

Cost: Based on the statistics of the United States Department of Energy (U.S. DOE) Fuel Cells Technologies Office (FCTO) on an automotive PEMFC system (80 kW), the noble metal catalysts and bipolar plates (e.g., graphite) are the largest cost components of the PEMFC stack [11]. Besides, mass production and application of fuel cell components can help reduce the costs of membrane and gas diffusion layers to some extent. The costs of different components in the PEMFC stack are shown in Fig. III.

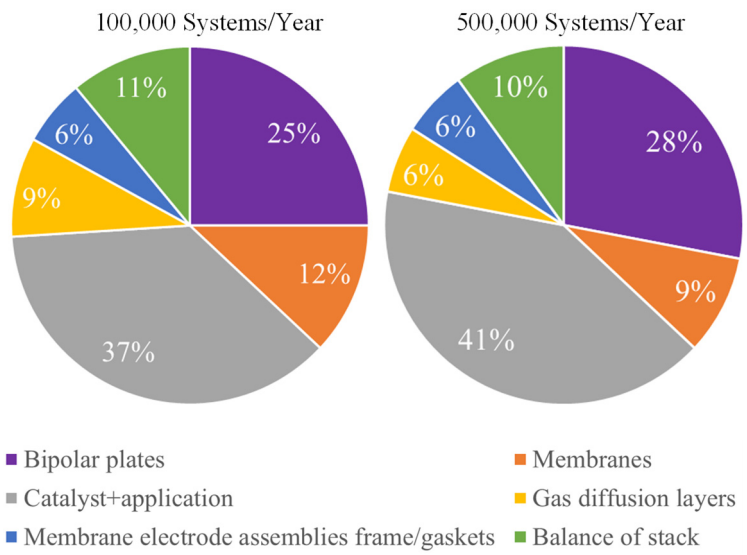


Fig. III Costs of different components in the PEMFC stack [11].

Increasing the activity of catalysts, reducing the content of Platinum group metals (PGM), and seeking inexpensive catalyst (PGM-free) materials are the research focuses for long-term applications of catalyst layers [12]. Graphite, metals, and polymer composites are commonly used materials for PEMFC's bipolar plates. Graphite is the most popular one, and it has been realized the large-scale commercial applications. Nevertheless, graphite fabrication is expensive, and it has the weakness of being fragile.

Some anti-corrosive metals (e.g., aluminum, nickel, and stainless steel) can be used as a substitute for graphite because of their cheap fabrication process and strong mechanical strength. The formation of oxides on the metal surface would increase the contact resistance, and this shortcoming can be partially solved by the utilize of anti-corrosive surface coatings. The chemical stability of surface coatings in the long term needs to be enhanced. The polymer composites have the advantages of corrosion resistance, easy fabrication, and low density, etc. Nevertheless, the electrical conductivity needs to be further improved by adding conducting material (e.g., graphite and graphene). To sum up, the cost reduction of the PEMFC stack depends on the development of materials. For the whole PEMFC system, the hydrogen's production, delivery, and storage, the auxiliary equipment, and the drivetrain maintenance are other important costs. According to recent statistics of the Fuel Cells and Hydrogen Joint Undertaking (FCH 2 JU) in Europe [10], the costs of fuel cell systems in some applications are shown in Tab. III.

Tab. III Costs (unit: €/kW) of fuel cell systems in different applications.

Applications	State of the art		FCH 2 JU target		
	2012	2017	2020	2024	2030
Fuel cell light duty vehicles①	500	100	60	50	40
Fuel cell electric buses②	3500	1500	900	750	600
Fuel cell forklifts (10 kW)③	4000	NA	2500	1250	450
Fuel cell electric aircraft④	3500	>20000	20000	6000	3000

Notes:

1. When evaluating costs, the hydrogen consumption, maintenance, overheads, and profits are excluded.

2. The cost is evaluated when manufactured at a volume of 100000 units/year for ①, a volume of 250 units/year (the year 2020), 500 units/year (the year 2024), and 900 units/year (the year 2030) for ②, a volume of 20000 units/year for ③.

3. The power level is 10 kW for ③ and 15 - 50 kW for ④ in the ram air turbine (RAT) - emergency power system.

According to recent statistics of the Fuel Cells and Hydrogen Joint Undertaking (FCH 2 JU) in Europe, the technical goals of the costs are 60 €/kW in 2020 and 40 €/kW in 2030 for fuel cell light-duty vehicles (including cars). The evaluation costs in 2012 and 2017 were 500 €/kW and 100 €/kW if manufactured at a volume of 100000 units/year [10]. In recent years, the PEMFC system becomes more and more competitive with its mass production.

Performance: In March 2019, the National Renewable Energy Laboratory (NREL) of U.S. DOE evaluated the performance of the on-road fuel cell electric vehicles (FCEV)

[13]. The evaluation results were based on 230 on-road vehicles from six manufacturers (Hyundai, Honda, Toyota, etc.) that had accumulated more than 7.4 million miles. FCEV evaluation project partners are shown in Fig. IV.

Fig. IV Fuel cell electric vehicle evaluation project partners [13].

The system's performance of peak energy efficiency, power density, specific power, volumetric capacity, gravimetric capacity, and durability is shown in Fig. V. The results were presented as a ratio of the current value to the U.S. DOE's 2020 target, and they showed that all the key performances of FCEV still have much room for improvement when compared to the target value. The status is indicated as a fraction of the targets.

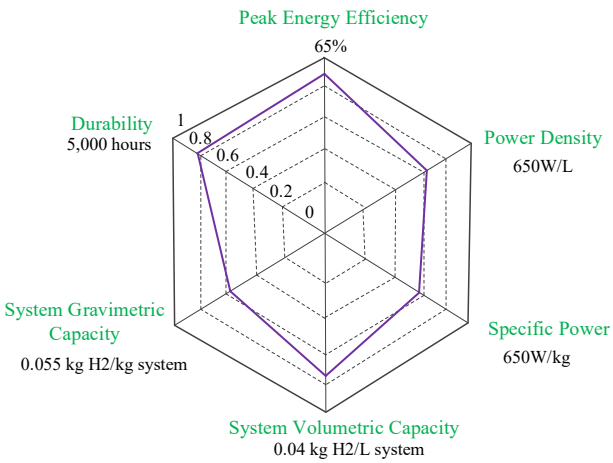


Fig. V Summary of key FCEV metrics versus U.S.DOE targets (fuel cell system level).

Exploring novel materials with better properties or enhancing the properties of available materials by integration strategies is the original power to improve the overall performance of fuel cells. Other directions of efforts to improve the current performance of the PEMFC system are as follows [12]:

- Developing hydrogen compression and solid-state hydrogen storage technologies (e.g., method of metal doping) to improve the hydrogen storage efficiency.

- Improving the catalyst activity to accelerate the separation of protons (H^+) and electrons (e^-) in hydrogen, and decreasing the catalyst tolerance to a poisonous substance (e.g., CO and sulfur).
- Exploring the more excellent proton exchange membrane electrolytes with higher transfer efficiency (for H^+ and e^-) and longer durability.
- Using the most advanced integration technologies to assemble state-of-the-art components of the membrane electrode assembly (MEA), and finally improving MEA's power density.
- Building accuracy models to understand the PEMFC system's working properties and degradation state, and validation of the models by in-situ and ex-situ experimental data.
- Applying advanced approaches to manage the sub-systems such as reactants (H_2 and O_2), temperature (cooling), and products (water and heat) management loops.
- Identifying the degradation mechanisms of the PEMFC system and exploring approaches to realize the degradation prediction.
- Utilizing high-performance hardware components such as sealing gaskets (usually rubber polymer) and bipolar plates.

According to recent statistics of the FCH 2 JU in Europe [10], the durability of fuel cell system in some applications are shown in Tab. IV.

Tab. IV Durability (unit: h) of fuel cell system in different applications.

Applications	State of the art		FCH 2 JU target		
	2012	2017	2020	2024	2030
Fuel cell light duty vehicles	2500	4000	5000	6000	7000
Fuel cell electric buses	10000	16000	20000	24000	28000
Fuel cell electric trains	N/A	12000	20000	25000	30000
Fuel cell electric aircraft	2000	5000	10000	15000	20000

Notes:

1. The 10 % power degradation is used to metric the durability in all durability tests. The load current of testing is around 55 % - 65% rated stack current.

2. The 10 % power drop is not a criterion of end of life (EoL), and different manufacturers may define the EoL in different applications.

3. The catastrophic stack failures are not addressed, and effects of transient operation, startup, and shutdown should be considered.

Fuel cell applications generally require adequate performance to be maintained over long periods. The FCH 2 JU targets of durability lifetime are 5000 h in 2020 and 7000 h in 2030 with less than 10 % power degradation in light-duty vehicle transportation applications. In the years 2012 and 2017, the lifetime was 2500 h and

4000 h separately [10]. Compared to the gasoline ICE, the cost and the durability of fuel cells have a large room to improve.

The durability prediction of PEMFC is always a research focus in both academia and industry. Developing some methods to predict the degradation state of the PEMFC system is helpful for the uses to take some actions in advance to extend its service life. And lots of prognostic methods have been proposed to evaluate the degradation condition in the past 10 years. Based on research status, this thesis focuses on exploring the lifespan prediction methods based on machine learning, and the method of echo state network (ESN) is selected. The research route of this thesis is summarized in Fig. VI.

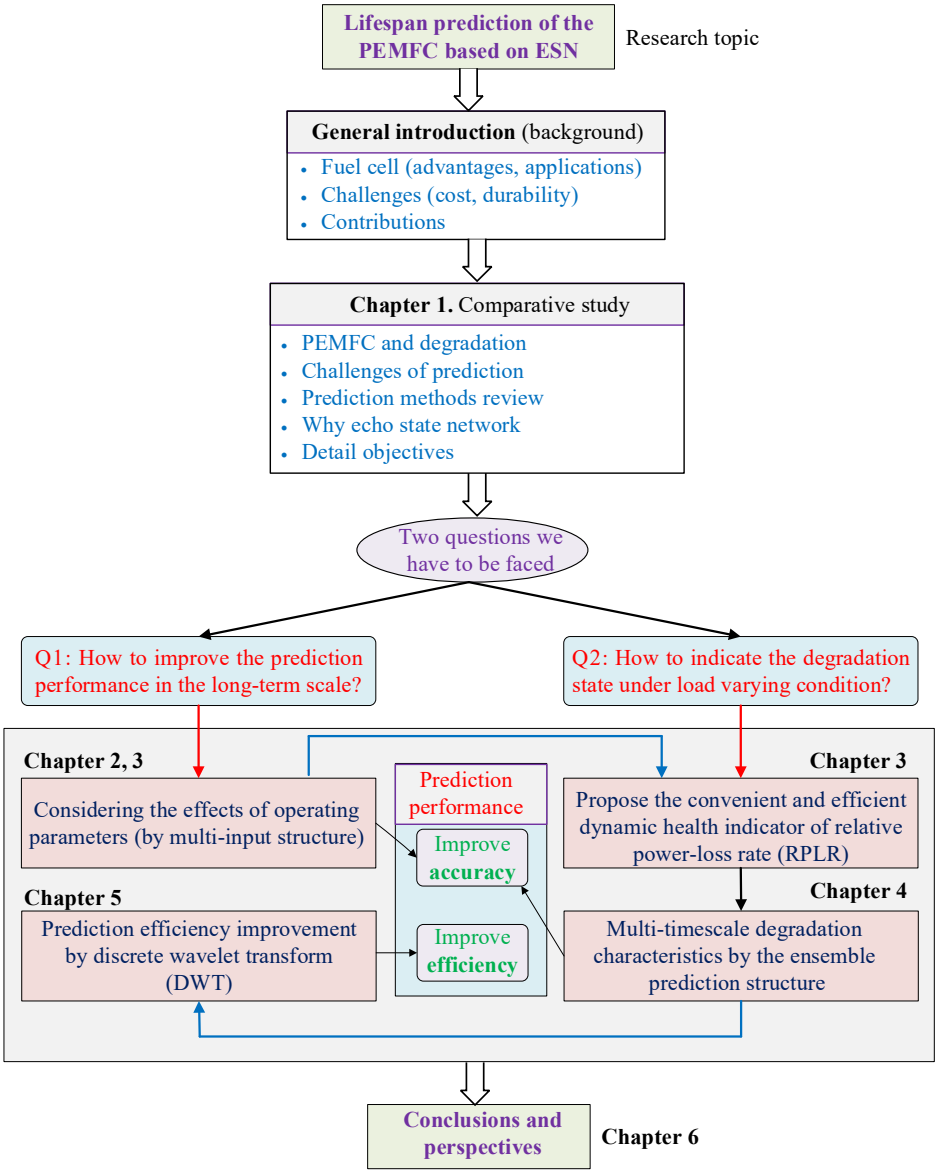


Fig. VI Research route of this thesis.

The contributions are briefly summarized as

- Improving the prediction accuracy by the multiple input structure which considering the effects of operating parameters (**Chapter 2 and Chapter 3**).

- Proposing a dynamic health indicator (HI) to indicate the degradation state under the dynamic operating conditions (**Chapter 3**).
- Dealing with the multi-timescale degradation characteristics by the ensemble prediction structure and improving the prediction accuracy at the same time (**Chapter 4**).
- Enhancing the prediction efficiency by the discrete wavelet transform's "compression and reconstruction" technique (**Chapter 5**).

Chapter 1. Comparative study on remaining useful life prediction for PEMFC system

In this chapter, the working principles of the PEMFC system and the degradation phenomena inside the PEMFC stack are first introduced. Then, a brief description of the prognostic and health management (PHM) and the process of prognostic are presented. After that, the experimental datasets for remaining useful life (RUL) prediction under the steady-state, quasi-dynamic, and full dynamic are also presented. The challenges of RUL prediction are given in 5 aspects. After a comparative literature review, the objectives of this thesis are detailed into 4 aspects.

1.1 PEMFC system and its degradation phenomena

1.1.1 Working principles of PEMFC system

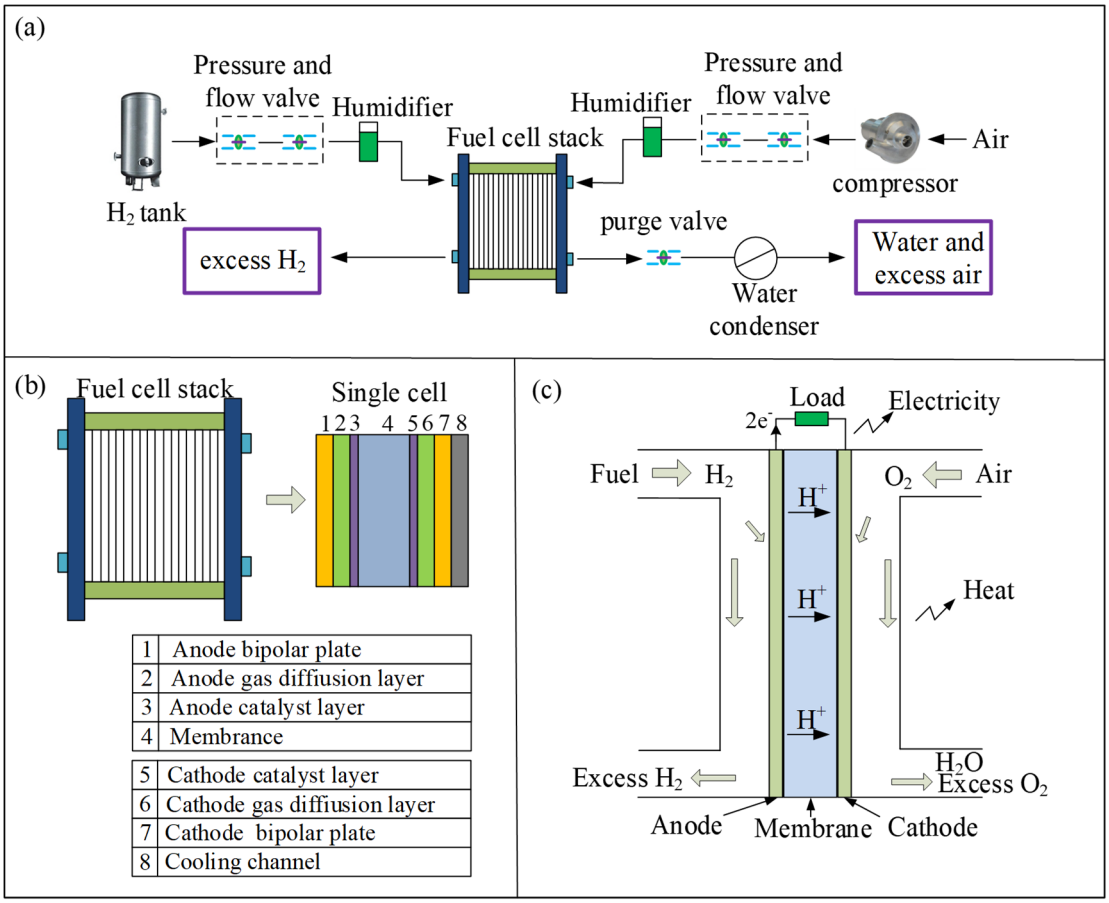


Fig. 1 - 1 The PEMFC system: (a) schematic of the system, (b) structure of fuel cell stack and single-cell, and (c) working principle of fuel cell [19].

Generally, a simple single fuel cell consists of 8 layers (Fig. 1 - 1(b)), and a number of cells are assembled in series as a fuel cell stack. In most cases, the cooling channel is assembled in the bipolar plate. Two reactant loops are added to the PEMFC system: hydrogen loop and air loop (Fig. 1 - 1(a)). Other ancillary loops like cooling and load management are also important for the PEMFC system. Most of the time, a hydrogen

tank supplies the fuel for the anode, and a compressor provides oxide to the cathode. The pressure and flow rates of the reactants can be adjusted by the valves, and the reactants should be humidified before being transmitted to the stack. Meanwhile, the protons are delivered through the membrane and the electrons are transferred through the external load to produce electricity (Fig. 1 - 1(c)). In brief, the whole function of PEMFC is converting chemical energy directly to electrical energy with water and heat as the by-products. Anode and cathode reaction equations are

$$\begin{cases} \mathrm{H_2} \rightarrow 2\mathrm{H^+} + 2e^- \text{ (Anode)} \\ \frac{1}{2}\mathrm{O_2} + 2\mathrm{H^+} + 2e^- \rightarrow \mathrm{H_2O} \text{ (Cathode)} \end{cases} \tag{1-1}$$

The overall reaction equation of FCs is

$$\mathrm{H_2} + \frac{1}{2}\mathrm{O_2} \rightarrow \mathrm{H_2O} + \text{electricity} + \text{heat} \tag{1-2}$$

The working principle of PEMFC can be divided into 3 steps:

Step 1: Continuous supply of hydrogen and air to the anode and the cathode. On both sides of bipolar plates, gas channels are grooved to increase the reaction areas. The demand for the reactants is increasing with the load. The fuel cell stack would be "starved" if the fuel and air are not supplied quickly and sufficiently enough. Oxygen excess ratio is defined as the ratio between oxygen entering the cathode and oxygen consumed due to the electrochemical reaction. To avoid irreversible damages to the membranes, the value of the oxygen excess ratio needs to be optimized with the mission profile [20].

Step 2: Electrochemical reaction of reactants. Hydrogen is decomposed into protons and electrons in the anode. The electrons are transferred through the external electric circuit, and the protons are transferred through the membrane. Then in the cathode, oxygen, protons, and electrons are combined to produce water. During the electrochemical reaction, suitable catalysts are used to increase the reaction speed and efficiency. The activity of the catalyst is strongly dependent on the reacting temperature. Thus, keeping a proper temperature is very important to ensure the stability of the reaction procedure [21].

Step 3: Purging of the by-products and use of electricity. For the PEMFC system, the by-products are water and heat. Heat can be used for recycling. Water vapors are expelled by the flow of the unreacted air and hydrogen. They should be removed in time. Otherwise, they may induce a "flooding" fault. The excessive water will block the reactant pathways, thus decreasing the reaction efficiency. The movement of electrons in the external circuit produces electricity. With the help of power electronic converters, the electricity can be used for lighting, grid connection, motor drive system,

etc. [22]

1.1.2 Thermodynamics and degradation of the PEMFC stack

The PEMFC system is a complex intersection of many research fields, e.g., electrochemistry, thermodynamics, and hydromechanics. The material characteristics, structure designing, assembly method, external environment, and operating conditions would affect its working performance and durability.

The current versus voltage (*i* - *U*) measurement, also known as the polarization curve measurement, is the most ubiquitous characterization technique to reveal the static performance of the PEMFC for it shows the output voltage under a given current. The typical *i* – *U* curve of a PEMFC is shown in Fig. 1 - 2. A large PEMFC would produce more energy than a small PEMFC in general, and standardizing the current in terms of the area can improve the comparability. Thus, the current in the *x*-axis is replaced by the current density (Ampere per square centimeter) [5].

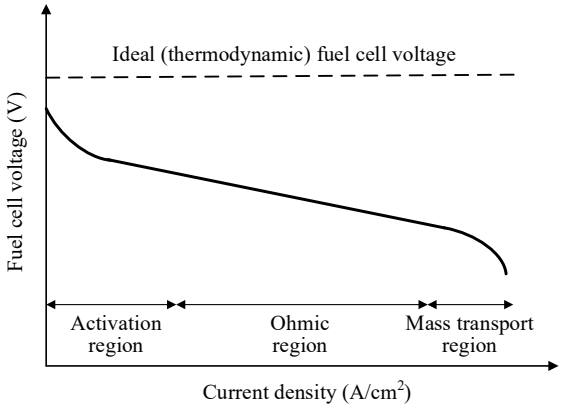


Fig. 1 - 2 The *i* - *U* curve of a PEMFC system.

In practice, the actual output voltage is less than the ideal thermodynamically value. Besides, the output voltage is also decreasing with the current's increasing. All these have limited the total delivered power (product of the current and voltage) of the PEMFC, and the power density curve of a PEMFC is shown in Fig. 1 - 3 [5].

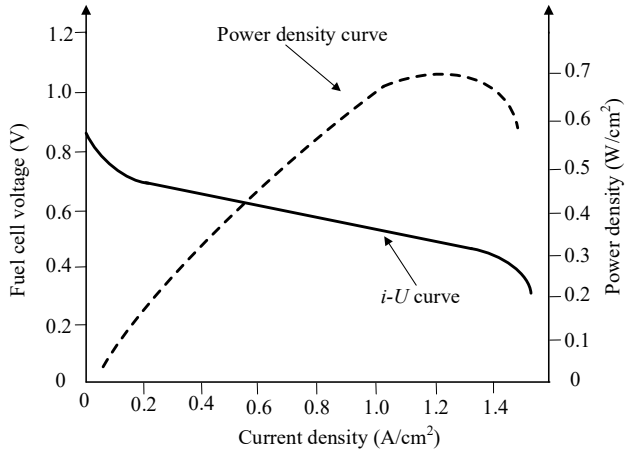


Fig. 1 - 3 The *i* - *U* and power density curves of a PEMFC system.

The real output voltage (U) of PEMFC can be expressed as

$$U = E_{thermo} - \eta_{act} - \eta_{ohmic} - \eta_{conc} \tag{1-3}$$

Where E_{thermo} is the thermodynamically predicted output voltage, η_{act} is the activation losses due to reaction kinetics, η_{ohmic} is the ohmic losses from ionic and electronic conduction, η_{conc} is the concentration losses due to mass transport.

At the constant temperature and constant pressure condition, the maximum output electrical power (W_{elec}) of a system can be expressed by the negative of the change in Gibbs free energy (g_{rxn})

$$W_{elec} = -\Delta g_{rxn} \tag{1-4}$$

The symbol Δ represents the state change between the final state and the initial state. The output electrical power can also be represented by the electrical potential difference E and charge Q

$$W_{elec} = E \cdot Q = E \cdot n \cdot F \tag{1-5}$$

Where n is the number of moles of transferred electrons, F is the Faraday's constant. Then

$$\Delta g_{rxn} = -E \cdot n \cdot F \tag{1-6}$$

The reversible voltage E^0_{thermo} of a hydrogen-oxygen fuel cell under the standard-state condition (25 °C, 101.325 kPa) is

$$E^0_{thermo} = -\frac{\Delta g^0_{rxn}}{n \cdot F} = -\frac{-237000\ \mathrm{J}/mol}{(2mol\ e^-\ /\ mol\ reactant) \times (96400\ \mathrm{C}/mol)} = 1.23\mathrm{V} \tag{1-7}$$

Where Δg^0_{rxn} is the change in Gibbs free energy under the standard-state condition, and the superscript "0" denotes the standard-state condition (25 °C, 101.325 kPa). Multiple single cells in series can increase the output voltage. In practice, the working conditions of PEMFC are far from the standard-state condition, e.g., 3×101.325 kPa – 5×101.325 kPa and 50 – 80 °C for FCEV application. In the standard-pressure (101.325 kPa), the reversible voltage at temperature T (E^T_{thermo}) can be expressed as

$$E^T_{thermo} = E^0_{thermo} + \frac{\Delta s}{n \cdot F}(T - T_0) \tag{1-8}$$

Where Δs is the entropy changes (a negative value for PEMFC), T_0 is the temperature under the standard-state condition. Thus, the reversible PEMFC voltages tend to decrease with increasing temperature, so a lower temperature is better to keep

the E_{thermo}^{T} at a higher value. However, the kinetics loss tends to increase with decreasing the temperature. The performance of PEMFC tends to increase with increasing the temperature on the whole even if the E_{thermo} would be decreased in practice. The reversible voltage which affected by pressure can be expressed as

$$(\frac{dE}{dp})_T = -\frac{\Delta v}{n \cdot F} = -\frac{\Delta n_g RT}{n \cdot F \cdot p} \tag{1-9}$$

Where p is the pressure, Δv is the volume change of reactants, Δn_g is the changes in the total number of moles during the reaction. To further analyze the reversible voltage, the Nernst equation of the PEMFC can be expressed as

$$E_{thermo} = E_{thermo}^{T} - \frac{RT}{2F} \ln \frac{a_{\mathrm{H_2o}}}{a_{\mathrm{H_2}} a_{\mathrm{O_2}}^{1/2}} \tag{1-10}$$

Where a_i is the activity of material i. The $a_{\mathrm{H_2O}}$, $a_{\mathrm{H_2}}$, and $a_{\mathrm{O_2}}$ are the activity of H_2O, H_2, and O_2 separately. Based on the Nernst equation, increasing the partial pressure of the reactants can improve the E_{thermo}^{T}. However, the effects of pressure increasing are slight for the pressure is calculated by the natural logarithm function. Thus, it is not economically advisable to raise the voltage by increasing the pressure based on the thermodynamics principle.

Based on the Butler-Volmer equation, the net current density j can be shown as

$$j = j_0 (\mathrm{e}^{\alpha nF\eta/RT} - \mathrm{e}^{-(1-\alpha)nF\eta/RT}) \tag{1-11}$$

Where η is voltage loss (i.e., activation losses), α is the transfer coefficient (0.2-0.5), j_0 is the current density under equilibrium conditions. For a large η_{act} (50-100 mV at 25 °C), the Butler-Volmer equation can be reduced to

$$j = j_0 \, \mathrm{e}^{\alpha nF\eta_{act}/RT} \tag{1-12}$$

Then the η_{act} can be re-expressed by the Tafel equation

$$\eta_{act} = -\frac{RT}{\alpha nF} \ln j_0 + \frac{RT}{\alpha nF} \ln j = a + b \lg j \tag{1-13}$$

Where a and b (Tafel slope) are the experience parameters. In the Ohmic region, the voltage loss is caused by the resistance of charge transport, and the η_{ohmic} obeys Ohm's law.

$$\eta_{ohmic} = i \cdot R_{ohmic} = i \cdot (R_{elec} + R_{ionic}) \tag{1-14}$$

The ionic contribution to R_{ohmic} tends to dominate for the electronic charge transport is more easy than ionic. Mass transport means the procedure of supplying

reactants and removing products, and the total mass transport losses η_{conc} can be expressed as

$$\eta_{conc} = \frac{RT}{nF}\ln\frac{j_L}{j_L - j} + \frac{RT}{nF}\ln\frac{j_L}{j_L - j} = \frac{RT}{nF}(1+\frac{1}{\alpha})\ln\frac{j_L}{j_L - j} \tag{1-15}$$

The general form can be shown as

$$\eta_{conc} = c\ln\frac{j_L}{j_L - j} \tag{1-16}$$

Where c is a constant value decided by experience. The actual output voltage of a PEMFC can be expressed by the thermodynamically predicted output voltage and the activation losses, ohmic losses, and concentration losses.

The degradation is related to the PEMFC's design and assembly process, component characteristics, material aging, etc. Once a PEMFC stack is assembled and in use, the degradation will occur during its whole life until the breakdown. The output voltage at a certain current (about half load) can be regarded as the outward manifestation of degradation, and exploring the degradation characteristics helps analyze the output performance and improve the service life of a PEMFC system. Generally speaking, the longer the PEMFC stack is operated in transient or varying load profiles conditions, the stronger is the degradation and the shorter is the lifetime. The degradation reasons of some components are cursorily summarized in Tab. 1 - 1.

Tab. 1 - 1 Degradation reasons of some components.

Main component	Main degradation causes
Membrane	Physical breach by mechanical stresses, decomposition of polymer material, punctures or pinholes by the freezing and melting of water, dehydration, direct combustion of crossover gases, chemical deterioration, temperature, relative humidity, pH value.
Catalyst layer	Contamination of impurities, sintering or migration of Pt particles on the carbon support, detachment and dissolution of Pt into the electrolyte, catalyst particle coarsening, corrosion of the catalyst carbon support, temperature, relative humidity, pH value.
Gas diffusion layer	Chemical attack (e.g., OH-), electrochemical oxidation, corrosion of the dual-layer carbon-based porous material, change of hydrophobicity, temperature, relative humidity, pH

Bipolar plate	Chemical corrosion and oxidation (for metals) of bipolar plate material, deformation of the coating material, deformation of the bipolar plate by the compressive forces, temperature, relative humidity.

1.2 Challenges of RUL prediction and the datasets in this thesis

1.2.1 Prognostic and health management of PEMFC

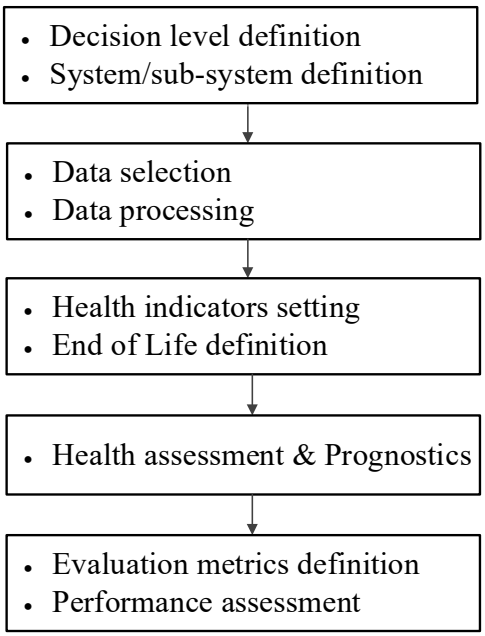


Fig. 1 - 4 The stages of the prognostic process [58].

PHM has the ability to estimate the future health state based on the past working profiles and the current operations, and then to extend the lifespan based on the condition-based maintenance (CBM) [23]. The key process of PHM is prognostic, and the International Organization for Standardization (ISO) defines the concept of prognostics: "prognostics is the estimation of time to failure and risk for one or more existing and future failure modes". The guideline of the prognostic stages is shown in Fig. 1 - 4.

For the PEMFC system, the sensors are used to acquire the data at a certain sampling rate (e.g., 1 Hz). After the decision of the level (stack or single-cell), the data processing layer performs the signal transformations and features extraction, reduction, and selection. After processing the data, the health indicators are settled and the end of life is defined. The system's current state of health (SoH) can be assessed by different prognostic methods (model-based and data-driven methods). In the prognostic process, the future condition of the system, subsystem, or component can be predicted by the

monitored system (propagation from causes to effects). In the PEMFC system, the main objective of prognostic is to realize the prediction of RUL. Finally, a lot of metrics and performance assessments [14], [15] can be used to evaluate the prediction results. Be aware, the prognostic does not address catastrophic stack failures.

1.2.2 Experimental datasets for RUL prediction

The durability data of FCs are scarce for it is not easy to test such a complex system for a long time. And it is even more difficult for the general researcher to acquire the durability data of the fuel cell equipped system such as the FCEV, stationary, and emergency backup power applications.

To promote the development of prognostic methods, the "IEEE PHM 2014 Data Challenge" (Data Challenge for short) was launched by the IEEE Reliability Society, FCLAB research federation, FEMTO-ST Institute, and the Laboratory of excellence ACTION [27]. Lots of research works are developed based on the open datasets of Data Challenge which tested under the steady-state operating condition (test-SS for short) and the quasi-dynamic operating condition (test-QD for short).

For the FCEV, the mission profiles are time-varying and they are determined by the driving behavior, road condition, energy demand, etc. The open data of test-SS and test-QD are tested under ideal circumstances, and they are defective to simulate the real operating conditions in practice. Thus, three dynamic tests under the micro-combined heat and power (μ-CHP) conditions were performed (382 h for test-D1, 1000 h for test-D2, and 405 h for test-D3) in the framework of the French National Research Agency (ANR) project, PROPICE "Prognostic and Health Management of PEM Fuel Cell Systems" [140]. The main purpose of this project is to develop prognostic methods to estimate the RUL of PEMFC systems.

1) Steady-state and quasi-dynamic datasets

The Data Challenge focused on estimating the RUL of a PEMFC. Both academic (from university) and professional teams (from industry) have joined in this Data Challenge, and the top-scoring participants are the winners of this competition.

The PEMFC test bench is adapted for a 1 kW electrical power fuel cell. The stack has 5 cells, and each cell has an active area of 100 cm^2. The nominal current density of the stack is 0.7 A/cm^2, and the maximal current density is 1.0 A/cm^2. The hydrogen loop and air loop are two reactant loops of the test bench. Two independent boilers (air and hydrogen boilers) are placed upstream of the stack to realize the reactant humidification. The air boiler is heated to get the desired relative humidity of the air, and the hydrogen boiler remains at room temperature. The flow rate of reactants is adjusted by the pressure and flow valves to avoid the FC stack "starvation". The temperature of the stack is controlled by a cooling water system. Moreover, the FC stack enables normal and accelerated aging tests under constant and dynamic operating conditions. The physical

parameters in the test bench can be measured and controlled in order to manage the PEMFC operating conditions as accurately as possible. Health monitoring data like the voltage, current, temperature, etc. are monitored by different sensors.

Two new fuel cell stacks under the steady-state and quasi-dynamic operating condition were implemented for more than 1000 h: the first test was operated under a static current operating condition (FC1, 1050 h), and the second test was operated under a quasi-dynamic current operating condition (FC2, 1020 h). The constant load current of 70 A is imposed on the aging test of FC1. In the aging test of FC2, a triangular ripple current of 7 A with 5 kHz is superimposed to the constant current of 70 A. The current waveforms used for the durability test are shown in Fig. 1 - 5.

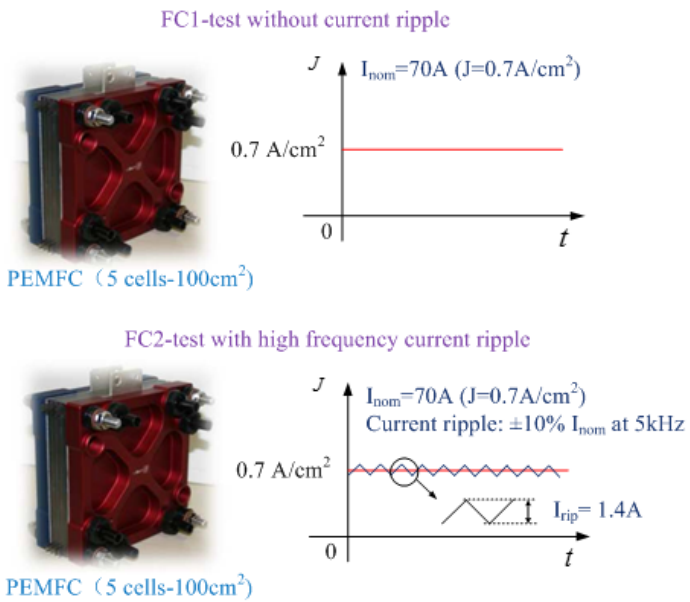


Fig. 1 - 5 Current waveforms used for the durability test of FC1 and FC2.

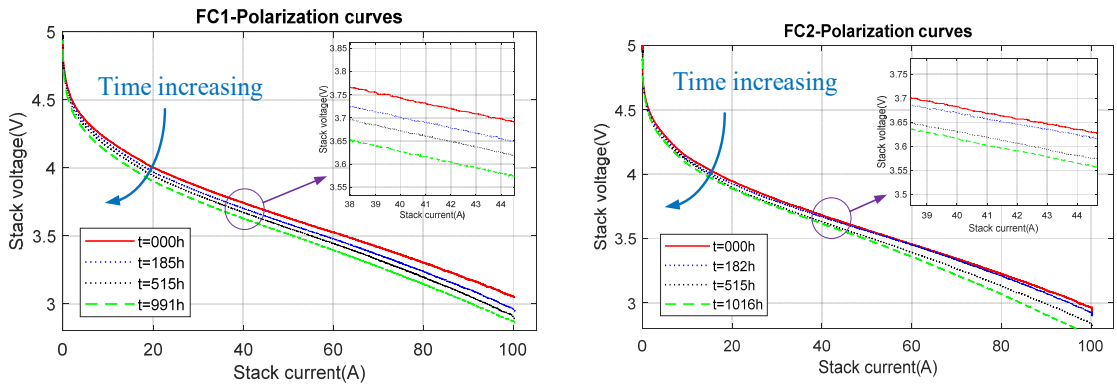
Some of the controllable and operating parameters are presented in Tab. 1 - 2.

Tab. 1 - 2 Physical and operating parameters of the test bench.

Parameter	Control range
Temperature	20 - 80 °C
Cooling flow	0 - 10 L/min
Gas temperature	20 - 80 °C
Gas humidification	0 - 100 % RH
Air flow	0 - 100 L/min
H_2 flow	0 - 30 L/min
Gas pressure	0 - 2 Bar
Fuel current	0 - 300 A
$T_{in}H_2$; $T_{out}H_2$	Inlet/Outlet temperatures of H_2 (°C)
T_{in}Air ; T_{out}Air	Inlet/Outlet temperatures of air (°C)
T_{in}Wat ; T_{out}Wat	Inlet/Outlet temperatures of water (°C)

$P_{in}H_2$; $P_{out}H_2$	Inlet/Outlet pressure of H_2 (mBar)
$P_{in}Air$; $P_{out}Air$	Inlet/Outlet pressure of air (mBar)

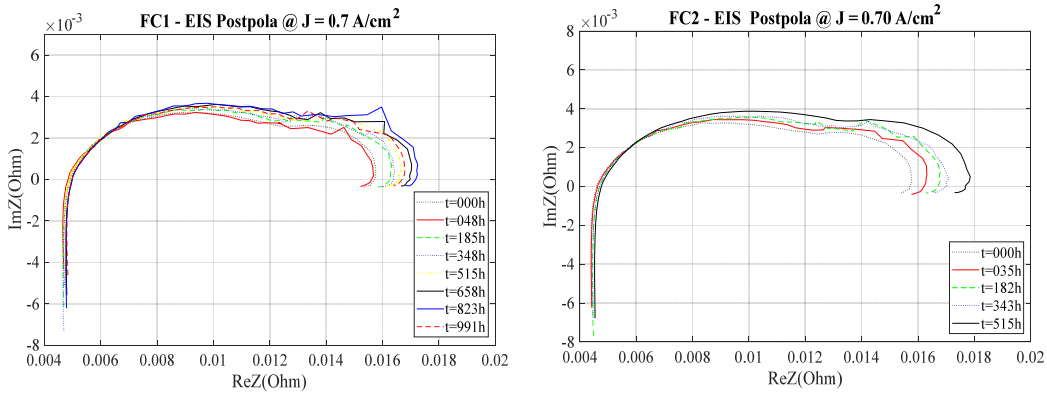
The polarization curves measurements and electrochemical impedance spectroscopy (EIS) tests are implemented in the test bench to analyze the static and quasi-dynamic properties of the PEMFC and are carried out once a week [25], [26]. Polarization curves of the single cells and stack were measured under a current ramp from 0 A/cm^2 to 1 A/cm^2 in 1000 s. The polarization curves measured at four different time steps are shown in Fig. 1 - 6.



(a) (b)

Fig. 1 - 6 Polarization curves: (a) FC1, (b) FC2.

In general, the longer the duration time is, the more serious the stack degradation is. More properly, the single-cell and stack voltage decrease as time grows, which depicts the degradation phenomena of the test bench. The EIS tests of the two tests are shown in Fig. 1 - 7.



(a) (b)

Fig. 1 - 7 EIS tests at different time: (a) FC1, (b) FC2.

In the PEMFC system, the stack voltage sensors are easy for installation and implementation, and the voltages are always used for control purposes. For convenience, the stack output voltage is also widely accepted as a HI of PEMFC systems. In Data

Challenge, the stack voltage is monitored with the acquisition interval of 30 s. The performance of prediction methods depends highly on the quantity and quality of the data. The raw data contain noises and large peaks (corresponding to the weekly characterizations), which would lead to a disturbance in the degradation prediction. The mission profiles of two long-term tests are shown in Fig. 1 - 8.

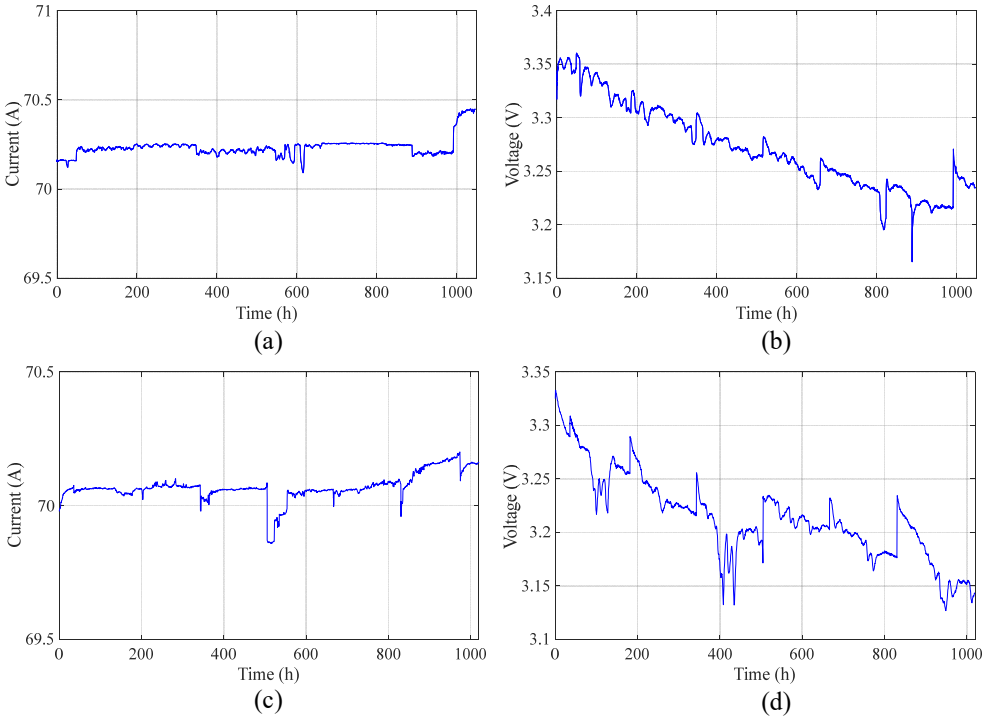


Fig. 1 - 8 The mission profiles of two long-term tests: (a) stack current of test-SS, (b) stack voltage of test-SS, (c) stack current of test-QD, (d) stack voltage of test-QD.

2) Full dynamic datasets

The experimental data come from the stationary PEMFC system which contains eight fuel cells and are obtained under dynamic mission profiles. The electrical power of the test bench is 1.0 kW, and the active area of a single cell is 220 cm^2. Some of the nominal specifications and working parameters are shown in Tab. 1 - 3.

Tab. 1 - 3 Nominal specifications and working parameters of the experimental platform.

Parameters	Control range
Dimensions	(220×160×186) mm
Temperature	80 ℃
Current density	0.5 A/cm^2
Anode/cathode stoichiometry	1.5/2
Anode/cathode inlet pressure	150/150 kPa
Cooling flow rate	2 L/min
Pressure drop	30 kPa

$T_{in}H_2$; $T_{out}H_2$	Inlet/Outlet temperatures of H_2 (°C)
T_{in}Air; T_{out}Air	Inlet/Outlet temperatures of air (°C)
T_{in}Wat; T_{out}Wat	Inlet/Outlet temperatures of water (°C)
$P_{in}H_2$; $P_{out}H_2$	Inlet/Outlet pressure of H_2 (mBar)
P_{in}Air; P_{out}Air	Inlet/Outlet pressure of air (mBar)

For all the durability tests, the polarization curves and the EIS under different current densities are also measured about once a week. The acquisition frequency of stack voltage and stack current is 1 Hz. Both the inlet and outlet reactant properties (flow rate, temperature, pressure, humidification, etc.) could be controlled during the tests, and various sensors are installed to monitor the operating conditions [45].

The mission profiles of three long-term tests are shown from Fig. 1 - 9 to Fig. 1 - 11.

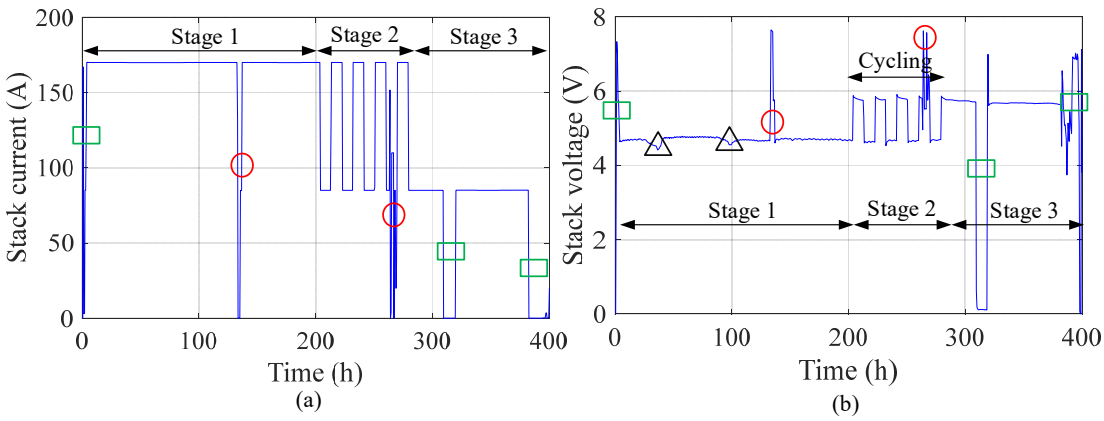


Fig. 1 - 9 The mission profiles of test-D1: (a) stack current of test-D1, (b) stack voltage of test-D1.

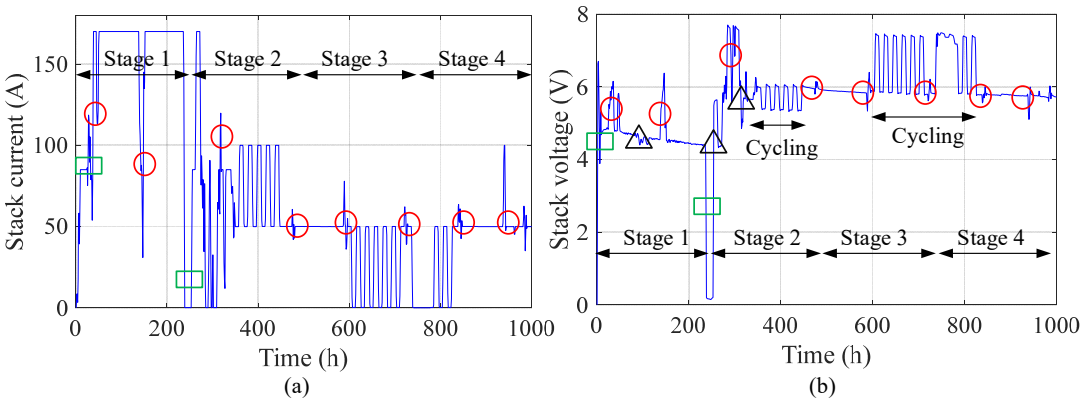


Fig. 1 - 10 The mission profiles of test-D2: (a) stack current of test-D2, (b) stack voltage of test-D2.

The transient effects of stack current on the stack voltage are marked in red circles. The effects of start and stop operations on stack voltage are marked in green rectangles. Besides, the operating parameters (e.g., $T_{in}H_2$ in this case) would also influence the

stack voltage. The effects of $T_{in}H_2$ on the stack voltage are marked in black triangles. The EIS tests and the polarization curves are shown from Fig. 1 - 12 to Fig. 1 - 14.

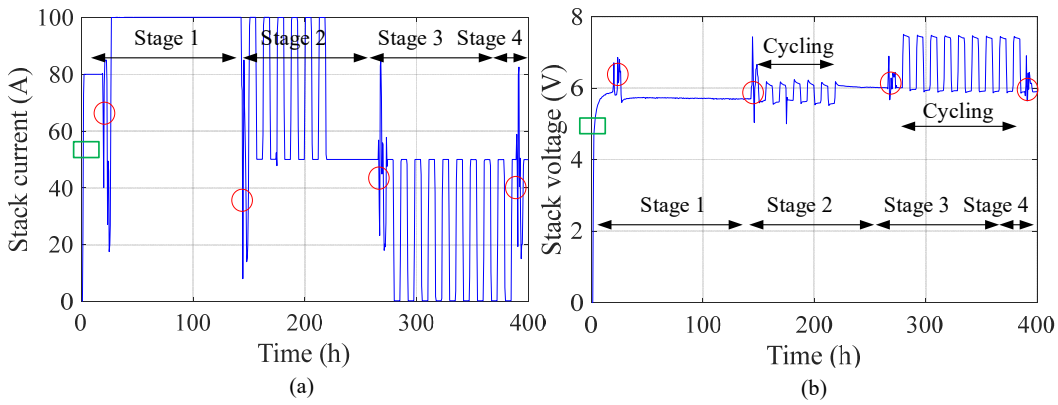


Fig. 1 - 11 The mission profiles of test-D3: (a) stack current of test-D3, (b) stack voltage of test-D3.

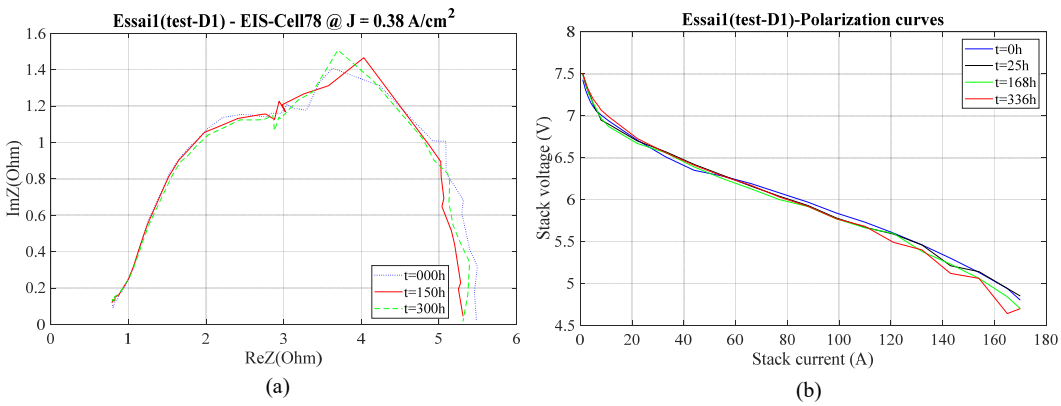


Fig. 1 - 12 The EIS tests and polarization curves of test-D1: (a) EIS tests of test-D1, (b) polarization curves of test-D1.

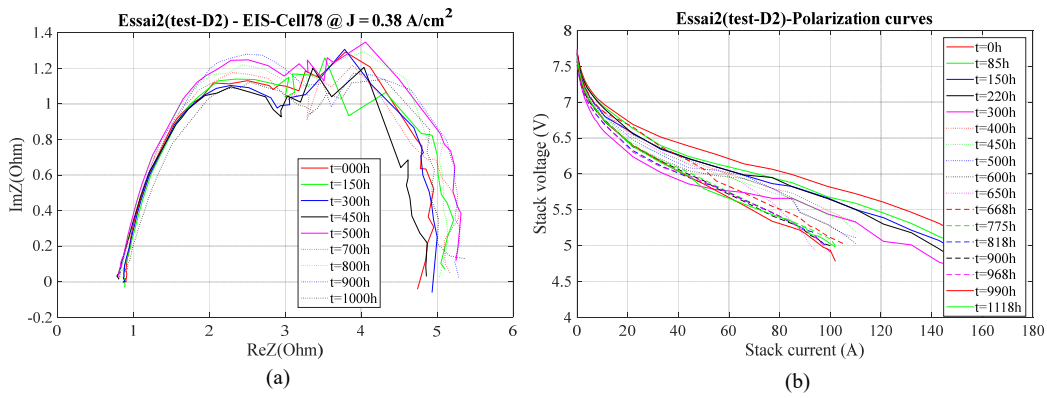


Fig. 1 - 13 The EIS tests and polarization curves of test-D2: (a) EIS tests of test-D2, (b) polarization curves of test-D2.

Test-D1: The 382 h duration data have three operating conditions and can be divided into 3 stages. The stack current and stack voltage are presented in Fig. 1 - 9 (a) and Fig. 1 - 9 (b).

(1) In stage 1 (0-200 h), the load current density is 0.77 A/cm^2 (maximum).

(2) In stage 2 (200 h-300 h), the load is cycling between 0.77 A/cm^2 and 0.38 A/cm^2.

(3) In stage 3 (300 h-382 h), the load current density is 0.38 A/cm^2.

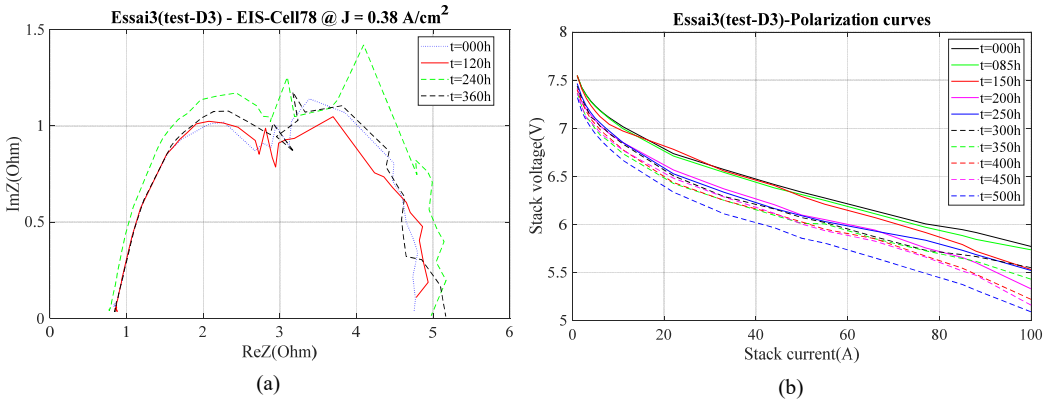


Fig. 1 - 14 The EIS tests and polarization curves of test-D3: (a) EIS tests of test-D3, (b) polarization curves of test-D3.

Test-D2: The 1000 h duration data are divided into 4 stages and each stage lasts for 250 h. The stack current and stack voltage are presented in Fig. 1 - 10 (a) and Fig. 1 - 10 (b).

(1) In stage 1(0-250 h), the load current density is maximum at 0.77 A/cm^2.

(2) In stage 2 (250 h-500 h), the mission profile changes dynamically between no load (0 A/cm^2), 0.77 A/cm^2, and 0.38 A/cm^2 from 250 h to 380 h. The load current density alternates between 0.45 A/cm^2 and 0.23 A/cm^2 from 380 h to 500 h.

(3) In stage 3 (500 h-750 h), the load current density is 0.23 A/cm^2 from 500 h to 600 h. The load current density alternates between 0 A/cm^2 and 0.23 A/cm^2 from 600 h to 750 h.

(4) In stage 4 (750 h-1000 h), the load current density alternates between 0 A/cm^2 and 0.23 A/cm^2 from 750 h to 850 h. The load current density is 0.23 A/cm^2 in the rest time of stage 4.

Test-D3: The 405 h duration data are also divided into 4 stages. The stack current and stack voltage are presented in Fig. 1 - 11 (a) and Fig. 1 - 11 (b).

(1) In stage 1 (0-125 h), the load current density is 0.36 A/cm^2 (0-25 h) and it changes to 0.45 A/cm^2 during the rest of the time.

(2) In stage 2 (125 h-250 h), the load current density changes between 0.45 A/cm^2 and 0.23 A/cm^2 from 125 h to 225 h. It stays at 0.23 A/cm^2 during the rest of the time.

(3) In stage 3 (250 h-375 h), the load current density alternates between 0 A/cm^2 and 0.23 A/cm^2.

(4) In stage 4 (375 h-405 h), the load current density is 0.23 A/cm^2.

1.2.3 Challenges of remaining useful life prediction

1) Decision level definition

Different physical levels (stack, cell, and component) of the PEMFC system have various degradation origins and impacts [16]-[18]. The degradation phenomena of different components (bipolar plates, catalyst layers, proton exchange membrane, etc.) happen on different timescales (from μs to h) and space scales (from nm to m). Carbon corrosion, platinum dissolution or reorganization, membrane degradation, and structure stress are caused by a variety of factors. The degradation rates of different components are also not the same, and it is difficult to fully understand the degradation properties of different components and to identify the internal parameters of the PEMFC system. The time scales in a PEMFC system are shown in Fig. 1 - 15.

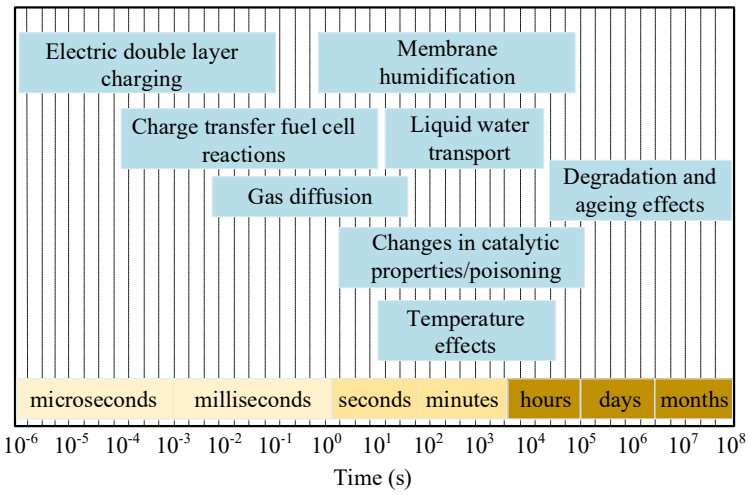


Fig. 1 - 15 Time scales in a PEMFC system [58].

The degradation observation at the component level (ex-situ) of the component is easier and more sensible than at the in-situ single-cell level. In the same way, the detection of the degradation of a single cell is more observable than at the stack level. Furthermore, the operating conditions which highly impact the degradations are more difficult to control at the stack level due to the inhomogeneities from one cell to another. Nevertheless, the data measurements should be implemented at the FCs stack level from a practical point of view. Until today, the deterioration mechanisms of the various fuel cell components are not all completely understood. Otherwise, all cells do not degrade in the same way, and cells near the edges tend to degrade faster. Fortunately, most of the degradation processes have an effect on the HIs which can be used to evaluate the SoH. There are different health indicators like the voltage, power, current, impedance at different frequencies, etc. The output voltage and output power are the most commonly used HIs, and they are easy to be detected by the sensors.

2) Data selection and data processing

Inevitably, the disturbances and fault signals would be monitored by the sensors along with the degradation-related features in practices, and the nondegenerate signal has a strong impact on the degradation prediction performance. Disturbances in the raw data should be pre-processed, and various filtering techniques have been used to select

the data which include the degradation information.

Taking the dataset of Data Challenge, for example, the output stack voltages are shown in Fig. 1 - 16. The stack voltages have the monotonic decline tendency under the steady-state and quasi-dynamic operating conditions. Nevertheless, the recoveries after the EIS measurements and the faults would affect the prediction accuracy. For example, in Fig. 1 - 16 (b), the faults at 100 h and 400 h would lead to an early prediction.

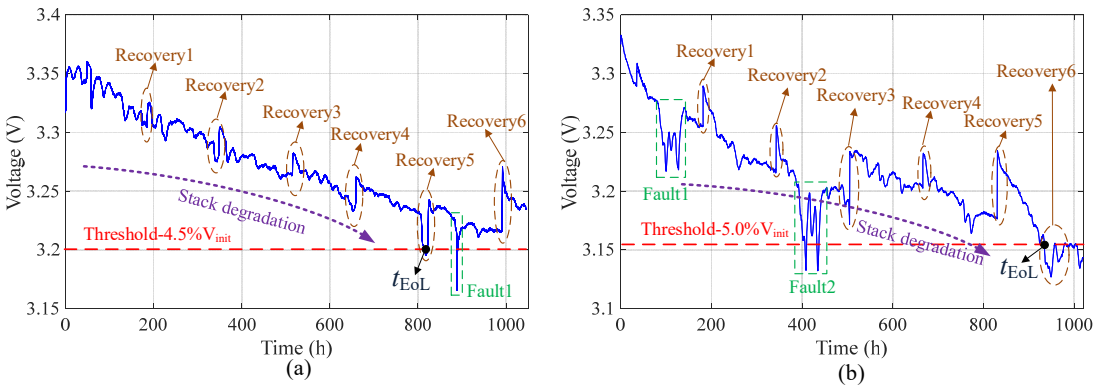


Fig. 1 - 16 The output stack voltages of Data Challenge: (a) FC1, (b) FC2.

Adequate and high-quality data can improve the precision of mathematical equations in the model-based methods, and they are also helpful for training the input and output relationships in the data-driven methods. A proper data pre-processing method is the key process of the prognostic.

Filtering is a frequently used technology in the literature, the **rloess** filter is used in [77] and [78], and locally weighted scatterplot smoothing method is utilized in [90], Gaussian-weighted moving average filter is used in [91], Gaussian Kernel-based smoother is employed in [59], and the moving average filtering (MAF) method is presented in [96], [106].

For each filtering method, the filtering coefficients have a significant effect on the filtering results, and they are difficult to be decided in practice. Unfortunately, trial-and-error is still the common method nowadays. For MAF, a too-large moving window may distort the original signal (eliminate the effects of characterization phases), and a too-small moving window may contain lots of noise during the degradation-related features extraction. With a span value of 0.9, the effects of characterization phases at the time interval of 168 h are also filtered out by **rloess** filter [77]. With a moving window of 31, the recoveries and faults are retained by MAF [106].

Besides, the filtering method cannot identify the degradation features that belong to which part (components) of the PEMFC system. So, developing advanced methods to extract the degradation-related features is an urgent problem to be solved. Besides, combining the diagnosis and prognostic has the potential to classify the degradation features, fault information, and disturbance signals.

3) End of life definition and health indicators

The definition of the end of life (EoL) is still an open question. Reaching an absolute consensus on the EoL threshold to measure the degradation state for different applications is impossible, and it is usually defined by the users themselves.

According to the definition of the U.S. DOE, the EoL can be regarded as when the PEMFC reaches a certain performance (voltage or power at a constant load current) decay (e.g., 10 % for the vehicle application, 20 % for μ-CHP, and 20 % for portable devices). Voltage or power is the external manifestation of the degradation at the component level [120]. The time when the system reaches the EoL was regarded as the time of failure threshold (FT) in the Data Challenge ({3.5, 4.0, 4.5, 5.0, 5.5} % loss of initial power), i.e., the old fuel cell stack should be replaced by a new one. Be aware, there is a difference between the FT and EoL threshold in practice. Failure time means the PEMFC is broken and doesn't produce power anymore. When the PEMFC stack reaches the FT, it must be fixed or replaced by a new one. The EoL time means the produced power of the PEMFC stack cannot fit anymore the application requirement, and it still has the ability to work. Considering the economic factors, the stack should be replaced when it reaches the EoL threshold in most conditions.

The durability performance of fuel cells should be equivalently durable and reliable to compete with ICE. That is the reason why 5000 h is set as the target for FCEV, i.e., the equivalent driving range is about 250000 km when the average driving speed is about 50 km/h. It is important to point out that the experimental environment of the U.S. DOE is under some strict conditions [121]. For the on-road testing of FCEV:

- The power loss is tested under about 55 % - 65 % rated stack current, and the transient operating, startup, and shutdown should also be included.
- Durability performance is evaluated based on the voltage degradation to 10 % lower than the beginning of life voltage, and 10 % voltage drop level is a U.S. DOE metric for assessing on-road fuel cell durability.
- 10 % voltage drop metric does not represent the real EoL of original equipment manufacturers (OEM). A higher or lower level of power degradation can be accepted in different applications.

At the constant load current, the voltage (or power) would have a nonreversible decreasing tendency as a whole regardless of the causes of degradation. This is the reason why voltage and power are usually regarded as the HIs in the literature [48], [97], [59], [77], [74], and they are classified as static HIs in this thesis.

4) Health assessment and prognostic

Over the last few years, various prognostic methods have been proposed. According to whether an analytical PEMFC model exists, these strategies can be categorized broadly into the **model-based** method and **data-driven** method. Through the analysis of the literature, the RUL prediction methods of the PEMFC system are

shown in Fig. 1 - 17.

The prognostic principle of the model-based method can be described as building the degradation models of the PEMFC system and then realizing the RUL prediction by the techniques of signal processing (e.g., wavelet transform), statistics (e.g., Markov chain), filtering methods (e.g., extended Kalman filter), and machine learning (ML) methods. By the way, some scholars called the approach based on both models and filtering (or ML) methods the hybrid method. In this thesis, the hybrid method is also classified to the model-based category since the models are necessary for this condition.

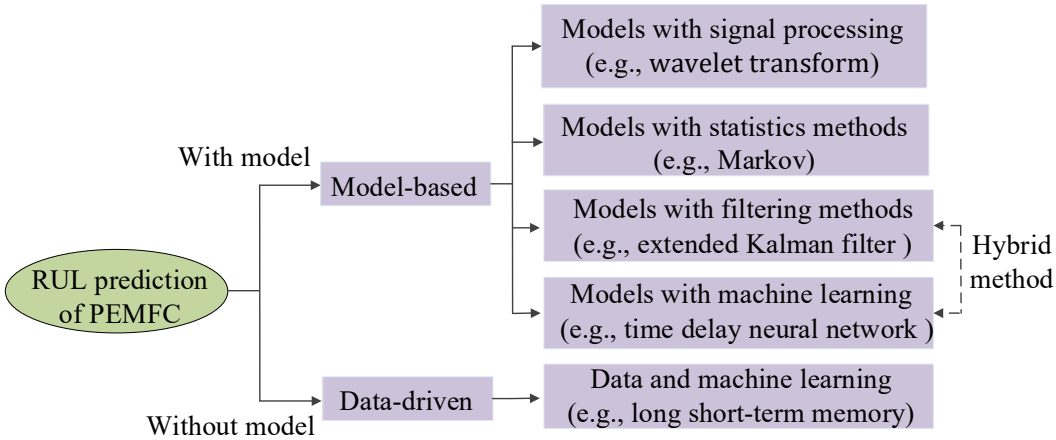


Fig. 1 - 17 The RUL prediction methods of the PEMFC system.

Most of the literature has realized the prognostic offline. Nevertheless, online prediction is more interesting in practice. According to the author's knowledge, [64] is the first paper to realize the online RUL prediction of the PEMFC system. Later, [65] and [66] have further improved the online prediction accuracy. Thus, more and more methods that can realize the lifespan's online prediction should be proposed and tested. Besides, the dynamic health indicators, RUL prediction accuracy, and efficiency are also need to be further improved in the long-term scale.

5) Performance evaluation criteria

To evaluate the prediction performance, one of the common criteria of percent error on RUL estimate ($\%Er_{FT}$) was defined by the Data Challenge in 2014. The definition of FT is important before using the criterion of $\%Er_{FT}$. Nevertheless, the FT is defined by the users in different conditions, and the failure thresholds (FTs) are 6 %, 10 %, and 15 % in [78]. In the Data Challenge, {3.5, 4.0, 4.5, 5.0, 5.5} % of initial power have been used as FTs.

$$\%\mathrm{Er}_{\mathrm{FT}} = \frac{t_{RUL}^{act} - t_{RUL}^{pre}}{t_{RUL}^{act}} \times 100 \tag{1-17}$$

Percentage error ($\%Er_{FT}$) is used to evaluate the error between the actual RUL (t_{RUL}^{act}) and the prediction one (t_{RUL}^{pre}). There are two cases for the prediction results: a) the estimation is smaller than the actual RUL, which means it is an early prediction, or

b) the estimation is greater, which means it is a late prediction.

The underestimates and overestimates are not considered in the same manner. In practice, good performance of estimations relates to early predictions of RUL (i.e., cases where $\%Er_{FT} > 0$), with a deduction to early removal, and more severe deductions for RUL estimations that exceed actual RUL (i.e., cases where $\%Er_{FT} \leq 0$). The score of the accuracy of the RUL estimate is defined as

$$A_{FT} = \begin{cases} exp^{-\ln(0.5)\cdot(\%Er_{FT}/5)} & if\ \ \%Er_{FT} \leq 0 \\ exp^{+\ln(0.5)\cdot(\%Er_{FT}/20)} & if\ \ \%Er_{FT} > 0 \end{cases} \tag{1-18}$$

The final score of all RUL predictions is defined as being the mean of all A_{FT}, and winners are the participants who get the highest scores

$$\text{Score}_{RUL} = \frac{1}{5}\sum\nolimits_{FT}(A_{FT}) \tag{1-19}$$

The root mean square error (RMSE) and mean average percentage error (MAPE) are two other commonly used criteria in the literature, and they can be used to quantitatively estimate the error between the real value and the forecasted value. RMSE is commonly used to quantify the difference between the real signal ($y_i^{target}(n)$) and its predicted value ($y_i^{predict}(n)$) during the RUL time (m data points). MAPE is also a quantification measurement between the real value ($y_i^{target}(n)$) and its predicted value ($y_i^{predict}(n)$).

$$\begin{cases} \text{RMSE} = \sqrt{\dfrac{1}{m}\sum\limits_{i=1}^{m}(\mathbf{y}_i^{predict}(n) - \mathbf{y}_i^{target}(n))^2} \\ \text{MAPE} = \dfrac{1}{m}\sum\limits_{i=1}^{m}\left|\dfrac{\mathbf{y}_i^{predict}(n) - \mathbf{y}_i^{target}(n)}{\mathbf{y}_i^{target}(n)}\right| \end{cases} \tag{1-20}$$

The criterion of prediction horizon (PH) can be used to evaluate the prediction ability, and it is defined as the time between start prediction and ending prediction. There do not need the new measured data during the PH. The longer the PH, the better the prediction ability.

In a word, using these criteria together is a more proper choice for the users to evaluate the prediction performance no matter for the short-term (PH ≤ 24 h), mid-term (24 h < PH ≤ 168 h), or long-term prediction (PH > 168 h). And finding other criteria to evaluate the prediction in the long term is one of the research focuses in the next step.

1.3 Literature review of remaining useful life prediction

1.3.1 Model-based methods

The degradation state of the PEMFC system can be modeled by physical or empirical equations, and the model's accuracy would have a direct impact on the prediction results. To improve the modeling accuracy, statistical techniques, filtering

methods, and ML methods have been used to assist the modeling process. This combination structure can also provide an opportunity to combine both the advantages of model-based (simple) and data-driven methods (accurate).

1) Models prepared for RUL prediction

Building a degradation model is fundamental to the model-based method. The polarization curve (i - U), as stated earlier, can be used to reflect the static performance of the PEMFC. Besides, the EIS can also be used to analyze the dynamic characteristics of the PEMFC for the impendence is tested under different frequencies. After building the models to fit the polarization curves and EIS curves, the degradation-related parameters can be extracted from the models. Then, the degradation state can be evaluated by predicting the tendency of these parameters.

A pattern recognition method based on the polarization curves and EIS is proposed by Onanena *et al*. to predict the lifetime of the PEMFC system [24]. An empirical model is used to extract the static features from polarization curves, and both expert knowledge and parametric models are used to extract the dynamic features from EIS. The parameters which are more correlated to the degradation are selected to help the users to estimate the operating time.

Based on the EIS measurements and linear regression technique, the impedance behavior estimation is realized in [25]and [26]. With the same models in [24], the mathematical equations are used to model the real part of EIS, and a high-order polynomial regression (5-order) is used to model the imaginary part of EIS in [25] by Vianna *et al*. Besides, the gradient descent approach is utilized to optimize the parameters of the mathematical equations, and the leave-one-out method is adopted to decide the order of polynomial regression. Based on the same datasets of [25], the equivalent circuit model (ECM) and voltage-resistance model are combined to realize the impedance prediction in [26] by Taejin *et al*. The ECM with 10 parameters is explored to express the physical characteristics of the PEMFC system, and then 4 parameters which can reflect the degradation phenomena are selected by the model reduction method. The effectiveness of these models is validated by 2 datasets of Data Challenge which were tested under the steady-state (FC1, 70 A) and quasi-dynamic (FC2, 63 A - 77 A) operating conditions [27]. Besides, the models in [25] and [26] are two winners of the Data Challenge from the industrial and academic participants, and final sores of them are 7.8932×10^{-6} and 5.3177×10^{-6} separately.

Similar to [24], a prognostic-oriented model is proposed by Lechartier *et al.* in [28]. The physical-based model is composed of two parts based on the measurements of polarization curves and EIS. The static part is built by the physical mechanism (Butler-Volmer law) of the PEMFC system, and the dynamic part is built by ECM. Finally, the proposed model is also validated on the datasets from Data Challenge.

Under the nominal and power cycle operating conditions, two numerical models are proposed by Robin *et al*. to describe the Pt surface degradation caused by the carbon support degradation [29]. A physical model which focuses on the electro-chemical surface area (ECSA) is proposed by Polverino *et al*., and this model has the potential for online RUL prediction [30]. The activation losses in this model are represented by Tafel's equation, and the exchange current density is a function of ECSA. After fully understanding the decreasing rate of ECSA, the degradation model about the cell voltage is formulated. Finally, the RUL can be estimated by comparing the current ECSA with its initial value. Besides, the effects of current density (0.05 A/cm^2-0.5 A/cm^2) and temperature (30 °C - 80 °C) on the voltage decay are also analyzed. Results show that the voltage decay has a high speed when the current density is at the lower (0.05 A/cm^2) and higher levels (0.5 A/cm^2). And the voltage decay speed is decreasing with the increase of temperature (maximum 80 °C).

Based on the vehicular operation records and the test results in the laboratory, Chen *et al*. built a model to realize the online lifespan prediction which considers the effects of the environmental and voltage degradation [31]. In practice, the economic lifespan can be regarded as the service time when the average cost of the PEMFC system reaches the minimum value. And the total consumption in the proposed model is made up of manufacturing cost, hydrogen consumption, and attachment consumption (air compressor, cooling water, electronic control equipment, etc.). At last, the average cost of the PEMFC system and the economic lifespan are analyzed by considering different hydrogen prices and stack manufacture costs.

After a deep literature review on the degradation at different levels (component, cell, and stack), a semi-empirical aging model is proposed by Jouin *et al*. for the PEMFC system's health assessment and prognostics [32]. This model can describe the power behavior accuracy in its whole lifetime, and 4 datasets under the steady-state, quasi-dynamic, and dynamic operating conditions are used to validate the proposed model. The coefficients of correlation are higher than 0.96 in these 4 datasets, and this model is suitable for the RUL predictions. Similar to [32], an empirical voltage degradation model is proposed by Zhang *et al*. in [33]. And both the aging terms and the mission profiles are used to build the voltage model.

For the signal processing techniques, the wavelet transform (WT) and two degradation models i.e., polynomial regression and autoregressive integrated moving average (ARIMA) models are used to realize the online RUL prediction [64]. The data in one week is compressed by discrete wavelet transform (DWT), and then the models are utilized to predict the coefficients of the next week. Finally, the output voltages of the next week are reconstructed by the inversed discrete wavelet transform (IDWT). Results of FC1 and FC2 show that the data in 168 h (one week) can be shortened into

21 h (about one day), and fewer data are used in the models to realize the long-term (168 h) RUL prediction.

For the statistics techniques, the model-based method of regime switching vector autoregressive (RSVAR) is proposed by Hochstein *et al*. to realize the RUL prediction [63]. The changing trend of the stack voltage in the next step is represented by the form of probability based on the first-order Markov model. The method is validated by the datasets of Data Challenge in the long term (PH = 500 h) and gets the highest score (0.3592) from the industrial participants.

2) Model-based methods with filtering approaches

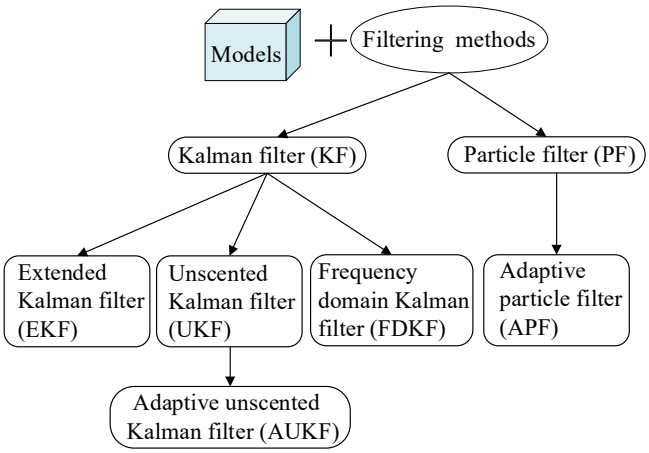


Fig. 1 - 18 The types of models with filtering methods.

In [34], a catalyst degradation model based on the physical deterioration mechanism is proposed by Zhang *et al*. This model is used to indicate the degradation rate of ECSA, and the area size can be indicated by the output voltage. Then the unscented Kalman filter (UKF) is utilized to predict the RUL of a PEMFC system. The prediction results are located in 95 % confidence intervals when the prediction horizon is about 300 h. In practice, it is inconvenient to measure the ECSA, and the measurements may disturb the behavior of the fuel cell stack. The same semi-empirical voltage model with an adaptive unscented Kalman filter (AUKF) algorithm is utilized to realize the RUL prediction in [35].

Chen *et al*. proposed a hybrid method that combined three empirical voltage models (linear, logarithmic, and exponential) and UKF to realize the degradation prediction [36]. Ten fuel cell electric vehicles are developed for the postal delivery mission under the Mobypost project. There are 40 cells at each PEMFC, and the rated power and current are 1 kW and 34 A separately. The effectiveness of the proposed method is validated by the datasets from the Mobypost vehicle. Results show that the logarithmic model has the best prediction performance in the short term (within 10 h).

Bressel *et al*. introduce the extended Kalman filter (EKF) to the long-term prognostic area of the PEMFC system [37] - [44]. In [37], a semi-empirical voltage

model is built based on the polarization curves of a single cell under the steady-state operating condition. The parameters of open-circuit voltage (OCV) E_0, exchange current i_0, overall resistance R, and limiting current i_L are calculated at each characterization time (168 h in this case) with the help of the Levenberg-Marquardt optimization method. Through analyzing these four parameters deviation, the R and i_L change a lot, and the E_0 and i_0 keep constant in the whole durability test. Thus, the R and i_L are regarded as the state variables, and then the EKF is used to estimate their variation tendency.

Afterward, the same model of [37] is further verified in [38], [39], and [40]. The dataset in [38] is tested under the steady-state operating condition with characterization disturbances and voltage recovery. Several simulations based on the EKF are performed to analyze the sensitivity of the parameters affected by the initial value of R and i_L. Nevertheless, the datasets in [38] and [37]are under steady-state operating conditions. After that, the method is verified by the dynamic datasets in [39] and [40] separately. The datasets in [39] are tested under an automotive profile. The prediction horizon is about 30 h (175 hours' durability test) when the results are located into $\pm$10 % bound of real RUL. And the EKF extraction process is implemented in [40] under the dynamic operating condition (μ-CHP) [45]. After the extraction of parameters by EKF, the model-based method of inverse first-order reliability method (IFORM) is used to extrapolate the degradation indicator until to the failure threshold. Besides, the RUL prediction results with a 90 % confidence interval are also presented, and the prediction horizon is about 200 h (1000 hours' durability test) when the results are located into $\pm$10 % bound of the real RUL. Later, the RUL real-time estimation is validated on the hardware in the loop (HIL) platform [41]. Recently, the frequency domain Kalman filter (FDKF) and the four voltage degradation models (linear, quadratic, logarithmic, and exponential) are combined by Ao *et al*. in [46]. This method is more efficient than the typical EKF because it can process the data in groups. Besides, the datasets of Data Challenge are used to verify its effectiveness in the long term.

Besides the semi-empirical voltage model of [37], a bond graph (BG) model of the PEMFC system is also proposed by Bressel *et al*. in [47] and [42]. Besides, the particle filter (PF) is used to replace EKF in [42]. Comparison results under the steady-state (FC1), quasi-dynamic (FC2), and dynamic (μ-CHP) operating conditions show that the BG-PF hybrid method could have higher accuracy than [37]. Similar to the model in [42], an aging tolerant control method is realized in [43] and [44]. This method aims at generating the reference current and considering the degradation state at the same time. Besides, the reference value of the gas can be calculated by the energetic macroscopic representation (EMR) model, and the maximum power which can be provided by the fuel cell stack is identified by the maximum power point tracking (MPPT) algorithm [43], [44].

The PF which is based on the Monte Carlo principle can predict RUL with a probability density function (PDF), and it can deal with the uncertainties at the same time. Jouin *et al*. have done lots of work about the application of PF in the prognostic area of the PEMFC system [48] - [54], and [55] - [58] reviewed the current methods of RUL predictions. Three empirical models (linear, log-linear, and exponential) are used to represent the degradation rate of the PEMFC system in [48], and then the PF is taken to assist the prediction procedure. Based on the datasets of Data Challenge, the log-linear model shows the best performance in the long-term perspective with a 90 % time interval. When the prediction horizon is 500 h, the prediction error is in the interval of [-25 h, +50 h].

Later, the prognostics which include characterization disturbances and voltage recovery are realized in [49], [50], and [51]. In [49] and [50], the aging model consists of three parts: 1) global power degradation is represented by a logarithmic and linear equation, 2) degradation's acceleration which is caused by the irreversible factors is represented by an exponential equation, and 3) the recovery is represented by an empirical equation. Then three paralleled PFs are used to deal with the models of each part separately, and the results are synchronized at the same time step. The prediction error is in the interval of [-28 h, +28 h] for a 950 hours' prediction horizon. To overcome the weakness of parameter initialization in [49] and further improve the prediction accuracy, an advanced empirical model is proposed in [51]. Compared to [49], the aging model in [51] consists of four parts, i.e., the degradation's acceleration is divided into two exponential equations. Four paralleled PFs are used to track the model, and the prediction error is within 5 % when the prediction horizon is 500 h. The datasets in [49] and [51] are the same, and different PFs in the paralleling structure could improve the prediction accuracy. Nevertheless, recoveries due to the characterization are in the regular time intervals. These empirical models are built under steady-state or quasi-dynamic operating conditions.

Based on the typical loss model, both the time and current are used to build a semi-empirical model in [52]. Then the PF is utilized to realize the long-term predictions based on the dataset of FC1, and the prediction horizon is about 500 h. Similar to [52], another two semi-empirical models are proposed in [53] and [54] to describe the power behavior based on the dataset of transportation application and μ-CHP separately. Nevertheless, the degradation predictions in [53] and [54] are realized in the short-term perspective.

The adaptive particle filter (APF) is used in [59] to predict the stack voltage in the long term based on the datasets of Data Challenge [27]. Similar to [48], Kimotho *et al*. tested another two empirical models (pure logarithmic and polynomial model) apart from the linear, log-linear, and exponential model. A self-healing factor is introduced to increase the adaptiveness, and its value is observed and updated after each

characterization. The prediction error is in the interval of [-25 h, +25 h] (under 5 %) when the prediction horizon is about 500 h. Based on the datasets of Data Challenge, this method gets the highest score (0.7760) from the academic participants. Nevertheless, the self-healing factor should be determined in advance, and the setting process is largely based on empirical knowledge.

The recovery phenomena caused by the measurements of polarization curves and EIS make the prognostic difficult [60], [61]. Because the degradation rates of each two characteristic measurements are not the same, and the sudden jumps phenomena would decrease the RUL prediction accuracy. The second-order polynomial model in [48] is used to describe the stack power, Zhang *et al*. also proposed other three functions to model the recovery state. Moreover, the parameters in the recovery model are extracted from the ECM of the EIS curves. Finally, with the help of PF, the RUL is predicted in the long-term (about 500 h) with the datasets of Data Challenge.

Zhou *et al*. proposed a multi-physical (electrical, fluidic, and thermal domain) aging model in [62]. And then the PF is used to predict the degradation in the long-term perspective, and the prediction horizon is 150 h and 250 h. Finally, 3 datasets (12 A/30 °C, 30 A/35 °C, and 44 A/40 °C) from the Ballard NEXA fuel cell stack are used to verify the robustness of this method.

The classification of model-based methods (models with filtering approaches) is shown in Tab. 1 - 4.

Tab. 1 - 4 Classification of model-based methods (models with filtering approaches).

Year	Reference	Methods	Health indicator	Operating condition	Prediction horizon
2012	[34]	Catalyst degradation model + UKF	Voltage	Dynamic	Long-term (≤ 300 h)
2017	[35]	Semi-empirical model + AUKF	Voltage	SS	Short-term (single-step)
2019	[36]	Three empirical models + UKF	Voltage	Dynamic	Short-term (≤ 10 h)
2015	[37]	Semi-empirical model + EKF	State of health	SS	Long-term (≤ 800 h)
2016	[38]	Semi-empirical model + EKF	State of health	SS	Long-term (≤ 1500 h)
2016	[39]	Semi-empirical model + EKF +IFORM	State of health	Dynamic	Long-term (≤ 120 h)
2016	[40]	Semi-empirical model + EKF +IFORM	State of health	Dynamic	Long-term (≤ 800 h)
2016	[42]	Semi-empirical model + PF	State of health	SS and dynamic	Long-term (≤ 900 h)
2021	[46]	Four empirical models	Voltage	SS and QD	Long-term

		+ FDKF			(≤ 500 h)
2014	[48]	Three empirical models + PF	Voltage	SS and QD	Long-term (≤ 850 h)
2014	[49], [50]	Three empirical models + paralleled PFs	Power	SS and QD	Long-term (≤ 950 h)
2016	[51]	Four empirical models + paralleled PFs	Power	SS	Long-term (≤ 500 h)
2015	[52]	Semi-empirical model + PF	Voltage	SS	Long-term (≤ 500 h)
2016	[53]	Semi-empirical model + PF	Power	Dynamic	Short-term (single-step)
2015	[54]	Semi-empirical model + PF	Power	Dynamic	Short-term (single-step)
2014	[59]	Five empirical models + APF	Voltage	SS and QD	Long-term (500 h)
2016	[61]	Four empirical models + PF	Power	SS and QD	Long-term (500 h)
2018	[60]	Four empirical models + PF	Power	SS and QD	Long-term (500 h)
2017	[62]	Multi-physical model + PF	Voltage	SS	Long-term (≤ 250 h)

3) Model-based methods with ML methods

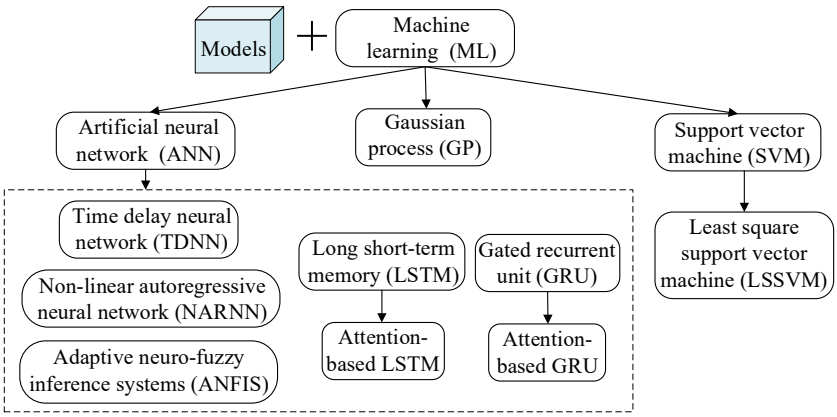


Fig. 1 - 19 The types of models with ML methods.

To improve the prediction accuracy of [64], the data-driven method of non-linear autoregressive neural network (NARNN) and an empirical fuel cell voltage model are combined to realize the degradation prediction of the PEMFC system [65]. The moving window method is used to improve the prediction accuracy, and the weights of these two methods are updated in each step (time interval is one week). Finally, this hybrid method is validated on the datasets of Ballard NEXA stack (36 A/40 ℃), proton motor 200 (PM200) fuel cell stack (64 A/58 ℃), and Data Challenge. Based on the same datasets of FC1 and FC2, the prediction accuracy of [65] is higher than [64] in terms of

RMSE. Afterward, Zhou *et al*. proposed another hybrid method that combined the autoregressive and moving average (ARMA) model and time delay neural network (TDNN) [66]. Before executing this hybrid method, a PF-based empirical model and physical aging model (PAM) are utilized to remove the non-stationary trend of the original stack voltage. The linear part is filtered by the ARMA model, and the nonlinear part is processed by the TDNN. This ARMA-TDNN method can handle the voltage signal in different patterns, and the improvements are validated by 3 datasets of Ballard NEXA stack (12 A/30 ℃, 30 A/35 ℃, and 44 A/40 ℃) and 2 datasets of Data Challenge.

To overcome the weaknesses of typical PF [66], the regularized particle filter (RPF) which is based on three state models (exponential, log-linear, and linear) is proposed by Cheng *et al*. in [67]. The hybrid approach is combined by the RPF and data-driven method of the least square support vector machine (LSSVM). The output voltage is first predicted by LSSVM, and then the predicted value is used in the RPF structure. Besides, the confidence interval (90 %) and PDF curve can be given by LSSVM-RPF, which are helpful for the users to make proper decisions. Based on the datasets of Data Challenge, the prediction performance of LSSVM-RPF is better than typical PF and RPF. Nevertheless, both in [66] and [67], the model-based method is realized in the long term, and the data-driven method is realized in the short term.

Apart from the PF, the Gaussian process (GP) can also present the degradation tend in a probability distribution form [68]. An empirical aging model is developed by Zhu *et al*., and the GP is used to predict the stack voltage under the steady-state operating condition (FC1). During the prediction, the GP model is updated every 50 h based on the new measured data. Thus, the prediction horizon of GP is 50 h.

Another hybrid prognostics method is proposed by Liu *et al*. in [69]. The semi-empirical voltage degradation model, AUKF, adaptive neuro-fuzzy inference systems (ANFIS), and particle swarm optimization (PSO) are combined to realize the long-term RUL prediction.

Yang *et al*. constructed a hybrid prognostic method in [70] and [71] by combining empirical models, PF, and long short-term memory (LSTM) network. The parameters of three models (linear, logarithmic, and exponential) are learned by PF, and the RUL is estimated in the long term by the LSTM network.

The LSTM network and autoregressive integrated moving average (LSTM-ARIMA) method are proposed in [72]. The LSTM-ARIMA structure can decrease the influences of overfitting and recovery phenomenon, and it is tested in the short term. The methods in [68] - [72] are verified by the offline datasets of Data Challenge.

Recently, Ma *et al*. proposed another hybrid prognostic method based on the EKF and LSTM network [73]. The internal parameters of limiting current, exchange current, and fuel cell resistance is extracted by a PAM. Then the EKF is adopted to estimate the three parameters based on the output voltage. Finally, the data-driven method of LSTM

is used to realize the voltage (or virtual steady-state voltage) prediction in the long term under different operating conditions. The classification of model-based methods (models with ML methods) is shown in Tab. 1 - 5.

Tab. 1 - 5 Classification of model-based methods (models with ML methods).

Year	Reference	Methods	Health indicator	Operating condition	Prediction horizon
2014	[63]	RSVAR + first-order Markov model	Voltage	SS and QD	Long-term (500 h)
2016	[64]	(Polynomial regression and ARIMA) + DWT	Voltage	SS and QD	Long-term (168 h)
2017	[65]	Empirical model + NARNN	Voltage	SS and QD	Long-term (168 h)
2018	[66]	(Empirical and physical model) + (ARMA and TDNN)	Voltage	SS	Short-term (single-step)
2018	[67]	Three empirical models + RPF + LSSVM	Voltage	SS	Short-term (single-step)
2018	[68]	Empirical model + GP	Voltage	SS	Mid-term (50 h)
2019	[69]	Semi-empirical model + AUKF +ANFIS	Voltage	SS and QD	Long-term (≤ 1000 h)
2017	[70],[71]	Three empirical models + PF + LSTM	Voltage	SS and QD	Long-term (400 h)
2019	[72]	ARIMA + LSTM	Voltage	SS and QD	Short-term (single-step)
2021	[73]	Physical model + EKF+ LSTM	Voltage + virtual voltage	SS, QD, and Dynamic	Long-term (≤ 450 h)

1.3.2 Data-driven methods

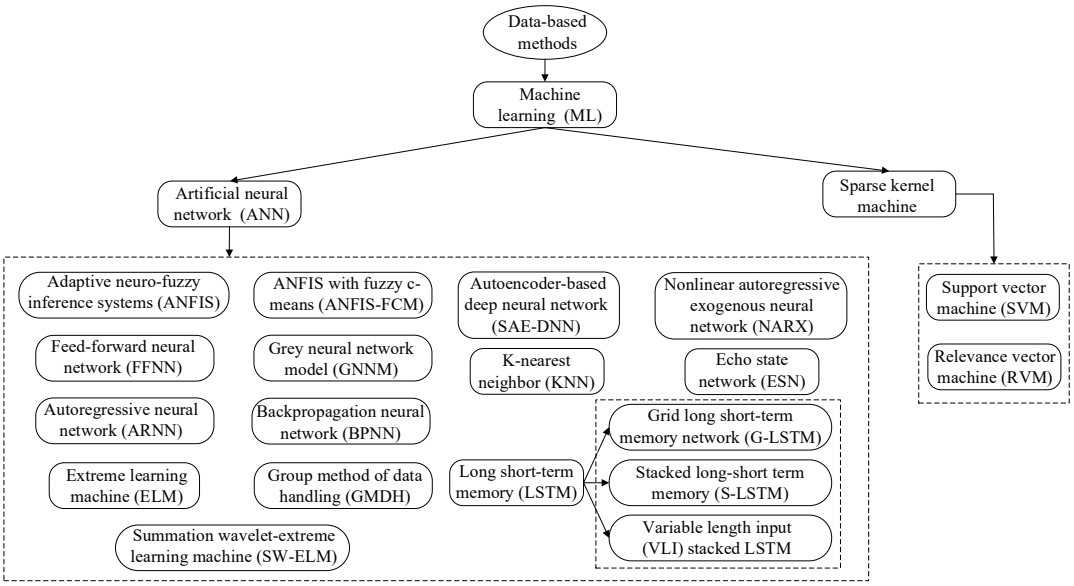


Fig. 1 - 20 The types of data-driven methods.

The data-driven methods do not need the precise mathematical models of the PEMFC system. Instead, the extrinsic attribute of the data can be learned and the RUL can be estimated based on the historical and current data. Thus, data-driven methods are usually used in cases where the degradation process is difficult to be modeled or the data are easy to be acquired. More generally, the data-driven method is like a "black box" which can be used to simulate the high-order nonlinear relations between the input and target data, and the new output data can be estimated by the trained "black box" and the new input data. During the working process of the PEMFC system, the electrical parameters (power, voltage, current, etc.) and the operating parameters (temperature, humidity, pressure, etc.) are measured by different kinds of sensors. In addition to the material degradation and decreasing of catalyst activity, these external parameters can also reflect the working properties of the PEMFC system. And understanding them sufficiently is helpful for the data-driven methods to predict the RUL efficiently. From the literature known so far, the data-driven methods are mainly based on the ML framework, and many kinds of ML approaches have been used in the RUL predictions of the PEMFC system.

The temporal variation of stack voltage is learned by ANFIS, and then the new stack voltages are predicted by the training structure. To improve the prediction accuracy, Silva *et al*. divided the voltages into the normal operation part and external perturbations part (e.g., transient process), and only the normal operation part is used to train the ANFIS structure [74]. This splitting approach can efficiently improve the prediction behaviors of two PEMFC stacks. The prediction horizon of these two long-term tests is about 500 h under the static (FC1) and quasi-dynamic (FC2). The long-term prediction is realized by the iterative approach and single-step ahead prediction process. The parameters are set by the trial-and-error process which relies heavily on the empirical data. A simple comparative study of ANFIS, feed-forward neural network (FFNN), and ARNN is presented in [75]. The ANFIS with fuzzy c-means (ANFIS-FCM) method is proposed by Liu *et al*. in [76] to realize the short-term prediction (single-step ahead) of stack voltage, and the parameters of ANFIS-FCM are optimized by PSO to improve the adjustment efficiency. Besides, the WT is also combined with ANFIS-FCM to obtain a better performance. The datasets in [74] and [76] come from Data Challenge (FC1 and FC2).

The summation wavelet-extreme learning machine (SW-ELM) is introduced to the PEMFC system by Javed *et al.*, and it is fully explored at a long-term prediction scale under steady-state operating conditions [77] - [79]. The SW-ELM can be regarded as the combination of artificial neural networks (ANN) and WT theory. Two activation functions are used for each hidden node of ANN, and two parameters' initialization (weights and bias) can decrease the effects of uncertain perturbations. Therefore, its ability to deal with nonlinearity and disturbance is enhanced. Long-term prediction

results with different training lengths have been researched in [77], and the prediction horizon is about 500 h. Nevertheless, the single prediction has some limitations in dealing with the uncertainties of degradation, disturbances of the operation conditions, parameter sensitivity, etc. To overcome the weakness of a single prediction, the SW-ELM with an ensemble of constraints is proposed in [78] and [79] to improve the prediction accuracy and robustness. One hundred predictions by SW-ELM with different learning parameters at each training length are implemented, and then the mean RUL and median RUL of all predictions are calculated and compared. Both in [78] and [79], the training length starts from 850 h (half lifespan of the stack), the prediction interval is 50 h, and the prediction horizon is from 50 h to 850 h. Besides, the data in [77] is also tested by the ensemble of constraints-based SW-ELM approach in [79]. In this dataset, training length starts from 600 h, and the prediction horizon is less than 550 h. Results show that the median RUL is more interesting because the prediction of them usually belongs to the underestimations, and the prognostic with ensemble structure can minimize the error probability than the individual prediction. Besides the steady-state operating condition, the ensemble SW-ELM approach is also tested by Javed *et al.* under the variable load [80]. The results of SW-ELM with or without incremental learning are compared under the durability tests for μ-CHP. In [80], the new measurement data are used for training when they are available, and the prediction is realized in the single-step ahead process in essence.

Based on the ensemble structure and WT principle [80], the approaches of group method of data handling (GMDH) network [81] and extreme learning machine (ELM) [82], [83] are used to deal with the multi-timescale features of PEMFC system in the short term. Firstly, the original signal is decomposed into several sub-waveforms by the WT. Secondly, the sub-waveforms are predicted by data-driven methods separately. Finally, an ensemble structure is utilized to integrate the multi-timescale features and improve the prediction accuracy. In [81], the original stack voltages are decomposed into 3 layers under the steady-state operating conditions (FC1) and 4 layers under and quasi-dynamic operating conditions (FC2). The stack voltages do not need to be normalized in the training process, and the short-term prediction is realized from a single-step ahead to 5-step ahead (about 5 h). Compared to [81], the GMDH is replaced by the enhanced ELM method in [82] to improve the prediction accuracy, and the self-adaptive differential evolutionary (SaDE) method is used to optimize the parameters of ELM. The prediction procedure of [82] is also similar to [78], and one hundred predictions are implemented to weaken the randomness of ELM. Besides, the influence of different training lengths is also analyzed with the same datasets of [81]. Later, the WT and ELM are used in [83] to realize the degradation prediction, and 3 decomposition layers are chosen for the dynamic datasets which are from MobyPost. Different from [82], the genetic algorithm (GA) is used to optimize the parameters of ELM, and stack

voltage prediction of single-step ahead is realized under the dynamic operating condition [83]. Nevertheless, the setting principle of the decomposition layer is unclear, and the physical meaning of the features in different layers needs to be further explored [81] - [83].

Also based on the datasets of FC1 and FC2, the grey neural network model (GNNM) method is proposed by Chen *et al.* to realize the single-step ahead output voltage prediction. The PSO is used to optimize the parameters of GNNM, and the moving window method is used to improve the prediction accuracy [84]. Later, the backpropagation neural network (BPNN) is applied to construct the relationship of inputs (current, humidity, temperature, etc.) and output (voltage), and it is tested by the datasets of FC1, FC2, and MobyPost [85]. Apart from the GA [83] and PSO [84], the mind evolutionary algorithm (MEA) is also used in [85] to optimize the initial weights and neuron numbers of BPNN. Besides, the BPNN with different activation functions are analyzed, and the voltage prediction results with traditional relevance vector machine (RVM) and support vector machine (SVM) are also compared. Nevertheless, in [81] - [85], the degradation prediction of the stack voltage is all realized in the short term (single-step ahead).

The modified RVM and self-adaptive RVM are put forward to predict the degradation of the PEMFC system by Wu *et al.* in [86] and [87] separately. Compared with the classical SVM, the RVM has higher selecting freedom in kernel functions, and fewer vectors are needed in RVM which can further decrease the computational complexity. Moreover, the RVM can realize the predictions with confidence intervals, i.e., it has the probabilistic prediction ability which is more interesting in practice than the single point prediction. Two experimental datasets (400 hours' durability testing) from the Ballard NEXA fuel cell stack under the steady-state operating conditions (30 A/ 35℃ and 36 A/40 ℃) are used to validate the Gaussian kernel-based RVM in [86]. The output voltage is used to indicate the degradation state. Prediction results in different training lengths are discussed, and the prediction horizon is 50 h, 100 h, and 160 h separately. Besides, the modified RVM has shown a better performance than SVM, and the confidence interval of 99.7 % is also given in the prediction results. Later, the self-adaptive RVM is proposed in [87] to improve the kernel width selection and prediction accuracy. Besides the 2 datasets in [86], the other 3 datasets are used to validate the prognostic performance, i.e., two tests (12 A/30 ℃ and 44 A/40 ℃) are from the Ballard NEXA fuel cell stack under the steady-state operating conditions, and the last test (30 – 97 A/58 ℃) is from PM200 fuel cell stack under the dynamic operating condition. The short-term predictions are realized based on the multi-step ahead prediction process, i.e., from single-step ahead (4 h) to 5-step ahead (20 h). The prediction performance is improved when compared with the classical SVM and original

RVM. Nevertheless, the effectiveness of RVM in the long-term timescale needs to be further researched especially under dynamic operating conditions.

LSTM network is an improved structure of the recurrent neural network (RNN) [88] - [90], and it can avoid the issues of gradient exploding and vanishing. The LSTM method is introduced to the degradation prediction of the PEMFC system by Ma. *et al*. [88]. Apart from the 4 datasets (12 A/30 ℃, 30 A/35 ℃, 36 A/40 ℃, and 44 A/40 ℃) from the Ballard NEXA fuel cell stack [87], another dataset with 1 hour time interval from PM200 fuel cell stack (64 A/58 ℃) is used to validate the effectiveness of LSTM network. Later, the grid long short-term memory network (G-LSTM) is proposed in [89] to further improve the prediction accuracy of output voltage, and this structure is tested in different operating conditions. Besides the 5 datasets from [88], the other 2 datasets from Data Challenge (FC1 and FC2) and 1 dynamic dataset (20 – 99 A/58 ℃) from the PM200 fuel cell stack (time interval is 4 h) are also used to test the G-LSTM network [89]. Moreover, the prediction accuracy is guaranteed by the updated sliding window size. Similar to the works in [88] and [89], the LSTM network is used in [90] to predict the output voltage of 2 datasets from Data Challenge. After dealing with the locally weighted scatterplot smoothing, the voltages are resampled at the time interval of 1 hour. Results show that the predictions by the LSTM network are more accurate than the BPNN both in FC1 and FC2. Later, Liu. *et al*. used another ML method of sparse autoencoder-based deep neural network (SAE-DNN) to predict the output power under the quasi-dynamic operating condition (FC2) [91]. The single-step ahead (1 hour) prediction is realized after the filtering and resampling of the stack power. Compared to K-nearest neighbor (KNN) and SVM regression algorithms, the SAE-DNN can reach a higher accuracy at different training lengths. Nevertheless, the output voltage/power predictions in [88] - [91] are realized by the single-step ahead prediction process. The prediction horizon in these four papers depends on the time intervals of the resampled data, i.e., the 1-hour or 4-hour ahead prediction is realized in these datasets. The classification of data-driven methods is shown in Tab. 1 - 6.

Tab. 1 - 6 Classification of data-driven methods.

Year	Reference	Methods	Health indicator	Operating condition	Prediction horizon
2014	[74]	ANFIS	Voltage	SS and QD	Long-term (500 h)
2019	[76]	ANFIS-FCM	Voltage	SS and QD	Short-term (single-step)
2015	[77]	SW-ELM	Voltage	SS	Long-term (≤ 500 h)
2015	[78]	Ensemble SW-ELM	Voltage	SS	Long-term (≤ 850 h)

2016	[79]	Ensemble SW-ELM	Voltage	SS	Long-term (≤ 850 h)
2016	[80]	SW-ELM	Power	Dynamic	Short-term (single-step)
2017	[81]	DWT-GMDH	Voltage	SS and QD	Short-term (≤ 5 h)
2019	[82]	DWT-ELM	Voltage	SS	Short-term (1 h)
2019	[83]	DWT-ELM	Voltage	Dynamic	Short-term (single-step)
2019	[84]	PSO-GNNM	Voltage	SS, QD and dynamic	Short-term (5 mins)
2020	[85]	MEA-BPNN	Voltage	SS	Short-term (single-step)
2016	[86]	Modified RVM	Voltage	SS	Long-term (≤ 160 h)
2016	[87]	Self-adaptive RVM	Voltage	SS and dynamic	Short-term (≤ 20 h)
2018	[88]	LSTM	Voltage	SS	Short-term (4 h)
2018	[89]	Grid-LSTM	Voltage	SS, QD and dynamic	Short-term (1h and 4 h)
2019	[90]	LSTM	Voltage	SS and QD	Short-term (1 h)
2019	[91]	SAE-DNN	Voltage	QD	Short-term (1 h)

1.4 Reasons about why choose the echo state network

1.4.1 The background of echo state network

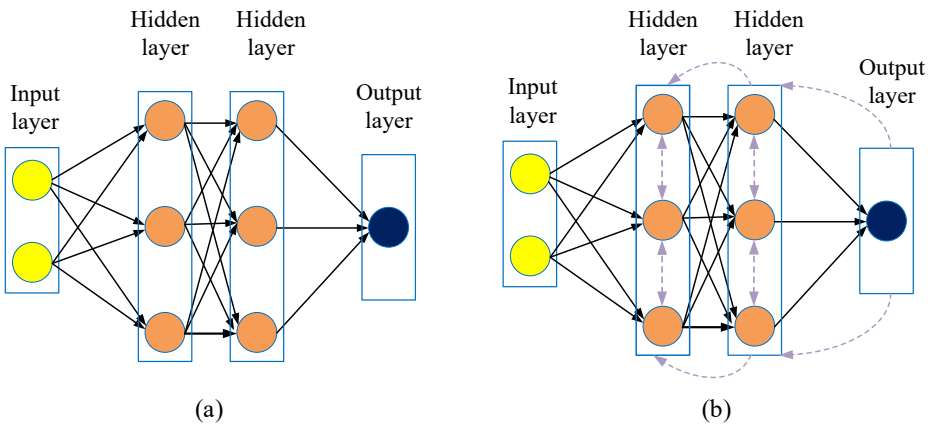


Fig. 1 - 21 Different network topologies of ANN: (a) FFNN, (b) RNN [94].

The ANN has been applied in different areas and it is used to construct and simulate the working principle of the brain. There are two typical structures of ANN:

feed-forward neural network (FFNN) whose information only flows forward and RNN which contains the feedback loops. The topologies of FFNN and RNN are shown in Fig. 1 - 21.

By imitating the behavior of biological neurons, the ANN can solve the problem of complex engineering. There do not exist recurrent connections in FFNN. Therefore, it lacks memory and cannot deal with temporal information. The RNN can deal with large dynamical memory and has high computational capabilities. It means that the current states of the neurons are not only affected by the inputs but also by the historical states of the network. The working principles of RNN are more similar to biological brains, the substrate of natural intelligence. Both the FFNN and RNN have been used in many fields.

Nevertheless, the training complexity is a common problem of ANN. The error back propagation (BP) is one of the efficient training rules to decrease the complexity in ANN especially for the FFNN. There are still some weaknesses when applying the BP to RNN such as bifurcations, slow convergence, expensive computing, and local optimum. These shortcomings still hinder the large-scale deployment of RNN in practical applications.

As a new design and training architecture, the echo state network (ESN) which was proposed by Prof. Jaeger *et al.*, has overcome the shortcomings of the traditional RNN [112]. This proposal is based on some evidence that the RNN has a satisfactory performance even without adaptation to all the weights in the network. Compared to the RNN, the hidden layer is replaced by a large randomly generated reservoir which reflects the dynamic topologies of the neurons. The weights (input, internal, and feedback) of ESN are global scaled to reach a desired dynamic state. And most of the weight matrices are randomly generated, and only the output weight matrix needs to be trained once. The readout of the reservoir is calculated via linear regression [113], [114]. As a result, the computational burden of ESN is much decreased. Echo state property (ESP) is an important indicator to represent the dynamic behavior of the reservoir. It should be considered carefully in designing the ESN, and the reservoir should wash out the neurons' initial states at a rate that is independent of the inputs [115] - [118]. On one hand, the neurons in the reservoir should be dynamic enough to decrease the computation complexity of output weights. On the other hand, too dynamic neurons may move the network to an unstable boundary. An optimized reservoir means that the neurons have rich enough dynamics, and the fading memory should also be preserved. Together with the liquid state machine (LSM) [109] and the back-propagation de-correlation (BPDC) [110], they are three paradigms of reservoir computing (RC) [95].

1.4.2 Research fundamental of echo state network in PEMFC system

The basic ESN structure has been introduced to the diagnostic and prognostic area

of the PEMFC system by Zheng *et al*. and Morando *et al*. respectively [92] - [95]. The fault diagnosis under dynamic load profile is realized by ESN in [92], the target fault types are stoichiometry value, pressure drop, temperature drops, and cooling circuit failure. Based on the work in [92], the parameters of ESN (leaking rate α, spectral radius ρ, and regularization parameter β) are optimized by the big bang and the big crunch (BB-BC) method in [93]. The fault types of carbon monoxide (CO) poisoning, low air flow rate, defective cooling, and natural degradation is discriminated efficiently in [94]. Besides, the key parameters (α, ρ, and β) of ESN are also investigated. Results have shown that the characteristics of ESN are not extremely critical concerning the parameters as long as they are kept within a certain range, and the suggestions in this task are: $\alpha > 0.3$, $4 < \rho < 10$, and $3 \times 10^{-5} \leq \beta \leq 5 \times 10^{-5}$. An excellent classification rate of 99.88 % for the offline training process and 92.43 % for new unknown dataset tests are obtained. Later, the faults diagnosis is directly detected in the model space (voltage-current model) rather than based on the original data space (voltage space) [95]. Based on the value of the output weight matrix ($\mathbf{W}_{out}$) of ESN, the influence of cathode stoichiometric ratio (FSC) is fully studied (the air side). Besides, another structure of ESN (time-delay reservoir) is used in [94] to realize the faults diagnosis.

At the same time, the ESN is introduced to the prognostic area of the PEMFC system by Morando *et al*. [96]. The ESN is used in [96] to predict the mean voltage of cells in the short term, and two prediction structures (direct approach and parallel approach) have been fully explored. The advantages of ESN and its parameter configurations are presented in [97], and the results of direct structure (in the short term) and iterative structure (in the long term) are compared. Different reservoir structures (random reservoir, delay line reservoir, cycle reservoir, etc.) have been compared in [98]. The double ESN structure is used after the wavelet filtering to improve the prediction accuracy in [99], and the parameter settings are the same with [97]. The influence of ESN parameters (α, ρ, number of reservoir neurons N, etc.) on the mean cell voltage prediction results is analyzed by analysis of variance (ANOVA) method, and it is helpful for the users to configure the parameters efficiently [100]. Based on the results of the ANOVA method in [100], the multi-reservoir ESN structure is developed in [101] to evaluate the mean voltage of cells in the long term, and it tries to avoid the parameters optimization process by setting the reservoirs with different spectral radius ρ. The voltage is used as the HI in the above articles [96] - [101], and the predictions are implemented in the single-step ahead or multi-step ahead under the steady-state and quasi-dynamic operating conditions.

To overcome the disturbance of load current, a novel HI of "virtual steady-state stack voltage" is proposed by Li *et al*. in [102]. A group of linear parameter varying (LPV) models is used to fit the sliding "voltage-current" segments, and then the

prognosis-oriented features are identified in the model space. This HI has the monotonic decline property even in the repeated load cycle test (0 A - 8 A), and it is more useful in practical use. Based on the dynamic HI in [102], an ensemble ESN structure is proposed in [103] to realize the long-term RUL prediction under the steady-state and dynamic operating conditions. This ensemble structure is similar to [80] - [83]. Nevertheless, the multi-step ahead prediction is realized in [103] by the iterative principle. This method is tested at different training lengths, and the acceptable prognostic can be achieved when the prediction horizon is within 350 h under both the steady-state and dynamic operating conditions.

Recently, the ESN and Markov chains are used to predict the degradation under a variable load profile (repeated load cycle) [104], and this method can obtain satisfying results without a great effort for data processing. Nevertheless, the HI is stack voltage, and the prediction is also realized by the single step ahead structure. Based on the double-input ESN structure, the measured ambient temperature and the predicted stack voltage are used as the double inputs to predict the RUL recently [108]. The stack voltage is predicted in the long-term in Vichard *et al.*'s work, and the influence of the temperature is analyzed based on the dataset from the Mobypost vehicle. In the above papers, the typical ESN prediction structure in the RUL prediction field can be expressed in Fig. 1 - 22. Where $\mathbf{W}_{out}$ is the output weight matrix in the training part, $\{(t_1, y_1),\cdots,(t_i, y_i)\}$ are the stack voltage data points in the training part. The y_1 and y_i are the stack voltages at time step t_1 and t_i (e.g., i = 2000). The application of ESN in degradation prediction is shown Tab. 1 - 7.

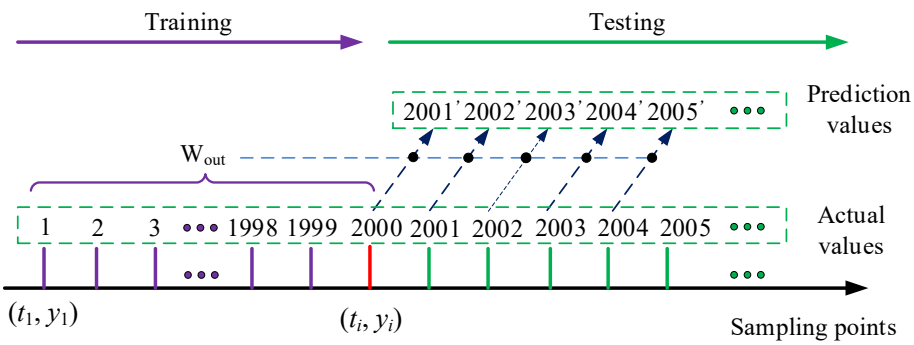


Fig. 1 - 22 The typical ESN structure in the work of Morando *et al*. and Mezzi *et al*.

Tab. 1 - 7 The application of ESN in degradation prediction.

Year	Reference	Methods	Health indicator	Operating condition	Prediction horizon
2013	[96]	ESN	Voltage	SS	Short-term (30 s)
2014	[97]	ESN	Voltage	SS	Long-term (1400 h)

2016	[99]	Double ESN	Voltage	SS	Long-term (1400 h)
2018	[101]	Multi-reservoir ESN	Voltage	SS	Long-term (≤ 900 h)
2019	[103]	Ensemble ESN	Virtual voltage	SS and Dynamic	Long-term (≤ 750 h)
2021	[104]	Ensemble ESN	Voltage	Dynamic	Short-term (single-step)
2020	[108]	Multi-input ESN	Voltage	SS	Long-term (4000 h)

1.5 Objectives of this thesis

The basic ESN structure has been used to realize the short-term RUL prediction or under the steady-state operating conditions in previous works. Nevertheless, more interesting tasks like the effects of the operating parameters, the proposition of dynamic HI, and the improvement of prediction performance, the RUL prediction under the long-term scale are should be further explored in the RUL prediction application, and this is also the motivity of this thesis. There are **4 objectives** in this thesis

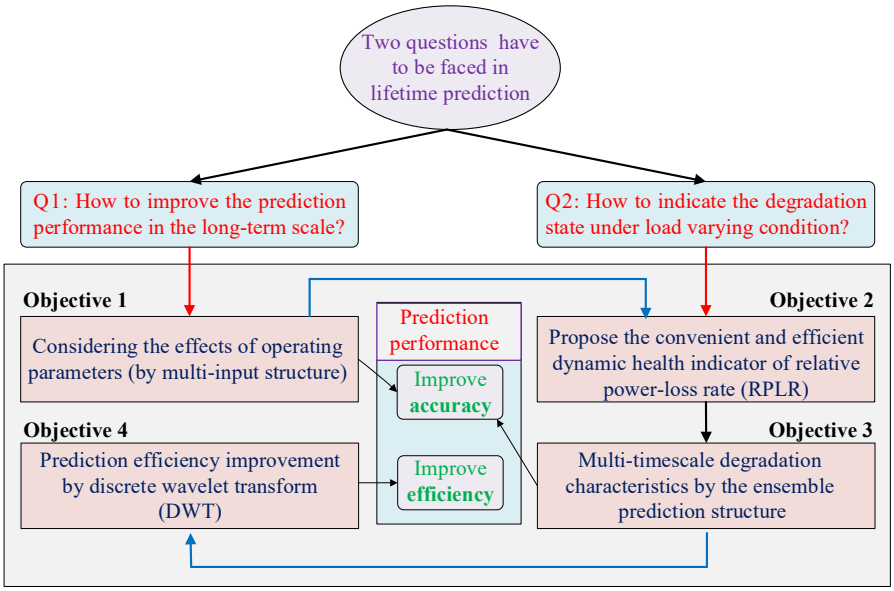


Fig. 1 - 23 The objectives in this thesis.

1) Analysis of the effects of operating parameters based on the improved ESN structure

In simple terms, the typical ESN can be regarded as the single-input and single-output ESN (SISO-ESN) structure in previous articles [96]-[104]. The historical voltages are regarded as the inputs and the outputs are the future voltages in the next step or the next several steps. The effects of four ESN training structures are discussed in [105]. Nevertheless, the current operating conditions, such as current, temperature,

and the pressures of the reactants (i.e., oxygen and hydrogen) can also contain important degradation information in practice. Especially, the stack current is a crucial operating parameter, since it is normally taken as the scheduling variable, and it can also reflect the operating conditions.

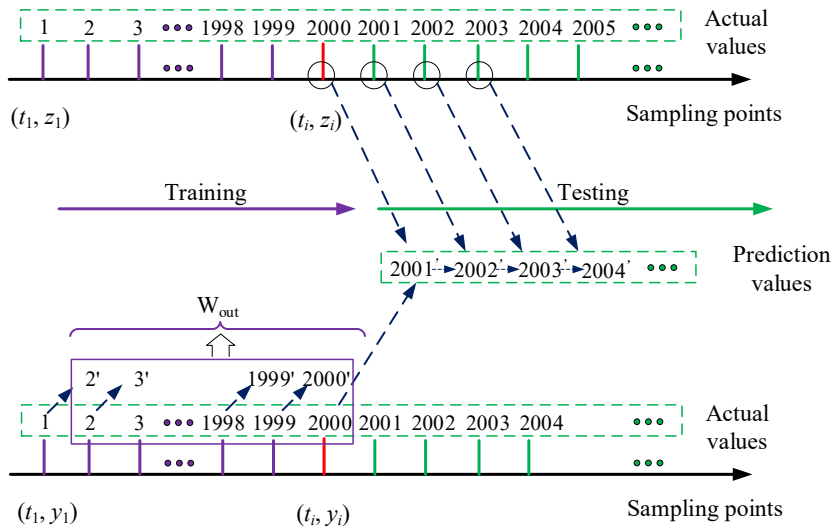


Fig. 1 - 24 The MIMO-ESN (2 inputs and 1 output) prediction structure.

Thus, the multi-input and multi-output ESN (MIMO-ESN) structures under the steady-state and quasi-dynamic should be developed [106], and the effects of temperature (H_2, air, and water), pressure (H_2 and air), and current would be fully researched. Later, this structure is validated under dynamic operating conditions (μ-CHP) to further explore the effects of variable load profile [107]. Compared with the SISO-ESN structure, the ESN with multiple inputs and multiple outputs is one research focus in this thesis. The MIMO-ESN prediction structure can be expressed in Fig. 1 - 24. Where $\mathbf{W}_{out}$ is the output weight matrix in the training part, $\{(t_1, y_1),\cdots,(t_i, y_i)\}$ and $\{(t_1, z_1),\cdots,(t_i, z_i)\}$ are the stack voltage and stack current data points in the training part. The y_1 and y_i are the stack voltages at time step t_1 and t_i, and the z_1 and z_i are the stack current at time step t_1 and t_i (e.g., $i = 2000$).

2) Propose the dynamic health indicator and realize the RUL prediction under the dynamic operating conditions

In practice, the static HIs can be much affected by the mission profiles, and their ability to indicate the degradation state will be limited especially under dynamic operating conditions. In the recent literature, the datasets from Data Challenge (static and quasi-dynamic) are usually used to test the prognostic methods. Nevertheless, the prognostic methods tested on the datasets from the dynamic operating conditions, especially from the FCEV and μ-CHP are more interesting for their practical use [103], [40]. Finding the more general dynamic HI is one of the objectives in this thesis.

The classification of existing HIs is shown in Tab. 1 - 8.

Tab. 1 - 8 The classification of HIs.

Working states	HIs	Advantages	Disadvantages
Static	Voltage, power	Easy to be measured online; Used for the control loop.	Disturbed by the load; Cannot be used for dynamic states.
	Resistance	Physical factor; Degradation related.	Based on polarization curve; Interrupted measurement.
	Hybrid	Accurate	Complex; Difficult to decide the weights.
Dynamic	Electrochemical catalytic surface area (ECSA)	Building models between ECSA and voltage; Cyclic voltammetry.	Offline; Interrupt the working of FC.
	EKF-based degradation factor (DF)	Simple (linearization hypothesis); Polarization curves.	Imprecise; large time interval (about one week)
	Virtual stack voltages (VSV)	A group of linear parameter varying (LPV) models, accurate, online.	Complex; More suitable for load cycling states.

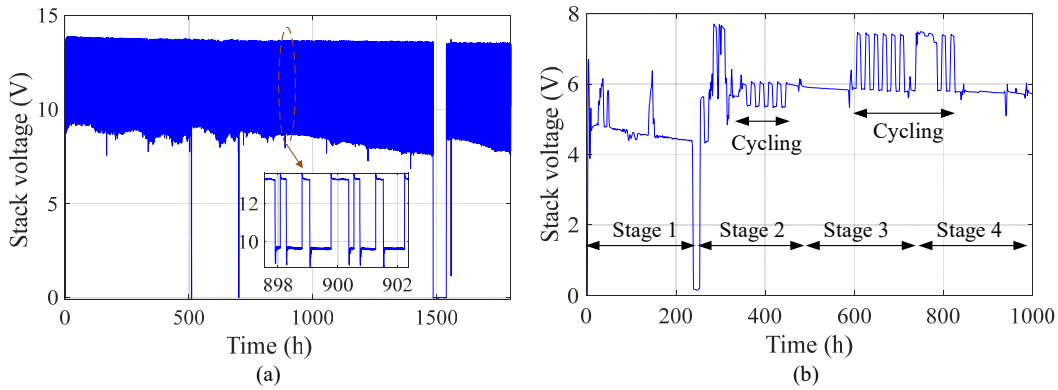


Fig. 1 - 25 The output stack voltages of FCEV and μ-CHP: (a) dataset in [103], (b) dataset in [40].

3) Improve the prediction accuracy by considering the multi-timescale degradation characteristics

The stochastic disturbances produced by the sensor noises during the measurements, unknown disturbances from the environment and operating parameters, load irregular varying caused by user habits, and the uncertainties of the prognostic methods themselves all would affect the prediction performance.

To improve the robustness and accuracy, an ensemble structure is proposed in [78], [80], [79], and [103]. Lots of prediction models with different parameter combinations are used in this structure, and the predictions are implemented at the same time. And then the mean or median value from the statistical data is chosen as the final RUL.

Compared with the single prediction, this ensemble structure could improve the robustness and accuracy from the probability point of view. Besides, the exploration of the methods which can provide the confidence intervals and PDF would also be welcomed in the next step.

Similar to the ensemble structure, another prognostic structure inspired by the "divide and rule" concept is proposed to deal with the multi-timescale features in [81], [82], and [83]. The data-driven methods of GMDH and ELM are used in this structure, the long-term predictions under the steady-state operating condition, and the short-term predictions under the dynamic operating condition are realized. Based on thesis ideas, the accuracy improvements considering the multi-timescale effects under the dynamic operating conditions are also the objective of this thesis.

4) Realize the mid-term and long-term prediction

The lifespan of the PEMFC system is in hundreds or thousands of hours, the short-term prediction (minutes to hours) cannot provide enough time for the users to take some actions in advance. So, the typical single-step ahead prediction ($(y_1, \cdots, y_t) \rightarrow \hat{y}_{t+1}$) does not have many practical meanings at all if the goal is to extend the lifespan even if the prediction accuracy looks satisfactory.

In the PEMFC system, RUL predictions in a mid-term (hours to tens of hours) or long-term horizon (hundreds to thousands of hours) are more useful for the users. The iteration of single-step ahead prediction is a common approach to realize the multi-step ahead prediction ($(y_1, \cdots, y_t) \rightarrow (\hat{y}_{t+1}, \cdots, \hat{y}_{t+H})$), where H is the time steps of the prediction horizon [80]. More specifically, prognostics in the mid-term and long-term horizon can be realized by the multi-step ahead prediction process. When PH ≤ 24 h, it is regarded as a short-term prediction. When 24 h < PH ≤ 168 h, it is regarded as the mid-term prediction. When PH > 168 h, it is regarded as the long-term prediction in this thesis. During the PH, it does not need the new measured data to train the prediction structure. For example, the mid-term and long-term predictions by ESN are shown in Fig. 1 - 26 and Fig. 1 - 27.

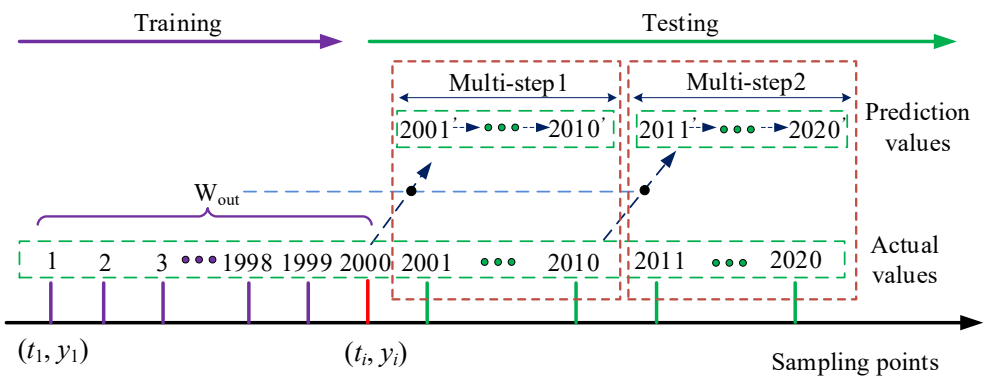


Fig. 1 - 26 Mid-term prediction structure by ESN.

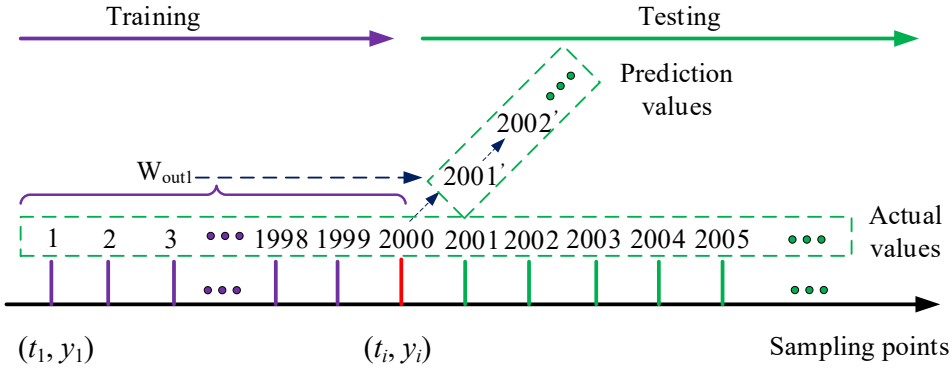


Fig. 1 - 27 Long-term prediction structure by ESN.

Chapter 2. Multi-input and multi-output echo state network of test-SS and test-QD

2.1 Brief introduction

To predict the degradation state of the PEMFC system, the single-input and single-output ESN (SISO-ESN) has been explored by Morando *et al*. [96], [97], [99]. In the SISO-ESN structure, the previous voltage is the single input and the predicted voltage is the single output. In practice, the ESN has the intrinsic property of dealing with multi-input and multi-output problems. Moreover, during the implementation of ESN, increasing the inputs improves the dimension of the output weight matrix and the ability to deal with non-linearity issues. The output weight matrix with a high dimension contains more system characteristics and can mimic more accurately the degradation phenomena in the PEMFC system. During the operation of PEMFC, different parameters such as stack current, the temperatures of hydrogen and air, the pressures of hydrogen and air can be easily obtained by the sensors. These parameters can also contain the degradation information of the PEMFC to some extent. Among all the operating parameters, the stack current is the most interesting one because it is normally taken as the scheduled variable. To improve the RUL prediction accuracy, different operating parameters are investigated together with the stack voltage as the inputs of ESN in this chapter. As the chosen HI of PEMFC, the stack voltage is regarded as the main output of the ESN. Other operating parameters can also be predicted at the same time. To the best of the author's knowledge, the use of ESN with multiple inputs and multiple outputs (MIMO-ESN) for the RUL prediction is originally explored and studied in this chapter. After the mathematical modeling and the parameter designing, the prediction performance of SISO-ESN and MIMO-ESN are verified and compared by the datasets of test-SS (FC1) and test-QD (FC2).

2.2 Mathematical background of ESN

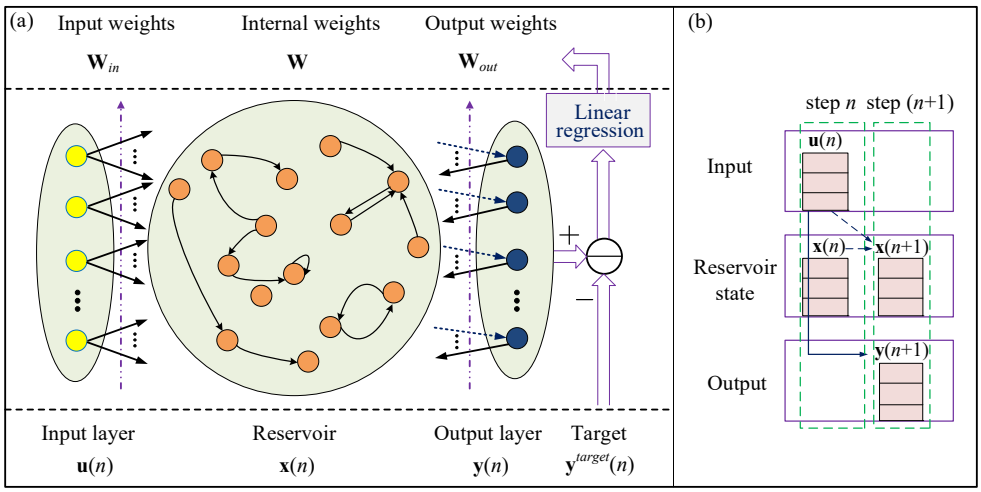


Fig. 2 - 1 The basic representation of echo state network: (a) ESN structure, (b) the updating state

in the reservoir at time step n.

The structure of a typical ESN is shown in Fig. 2 - 1 (a). The updating state of the neuron in the reservoir is described in Fig. 2 - 1 (b). The outputs of the reservoir at step (n+1) are affected by the inputs of step n, reservoir state of step n, and step (n+1). If the dimension of the input signal is K, the neuron number in the reservoir is N, and the dimension of the output signal is L, the external input signal $\mathbf{u} \in \mathbb{R}^{K\times 1}$, is a K-dimensional vector, the reservoir neuron activation signal $\mathbf{x} \in \mathbb{R}^{N\times N}$ is an N-dimensional square matrix, and the output signal $\mathbf{y} \in \mathbb{R}^{L\times 1}$ is an L-dimensional vector. At time step n, the forms of them are shown as

$$\begin{cases} \mathbf{u}(n) = (\mathbf{u}_1(n), \ldots, \mathbf{u}_K(n))^{\mathrm{T}}, \ n = 1, \ldots, B \\ \mathbf{x}(n) = (\mathbf{x}_1(n), \ldots, \mathbf{x}_N(n))^{\mathrm{T}}, \ n = 1, \ldots, B \\ \mathbf{y}(n) = (\mathbf{y}_1(n), \ldots, \mathbf{y}_{\mathrm{L}}(n))^{\mathrm{T}}, \ n = 1, \ldots, B \end{cases} \tag{2-1}$$

Where B is the total data points in the training part. With the external input signal $\mathbf{u}$, the time series prediction task aims to find a model to minimize the error between the target signal $\mathbf{y}^{target}$ and the output signal $\mathbf{y}^{predict}$. The error $E(\mathbf{y}^{target}, \mathbf{y}^{predict})$ which is represented by the Root Mean Square Error (RMSE) is given as

$$E(\mathbf{y}^{target}, \mathbf{y}^{predict}) = \frac{1}{L}\sum_{i}^{L}\sqrt{\frac{1}{B}\sum_{n=1}^{B}(y_i^{predict}(n) - y_i^{target}(n))^2} \tag{2-2}$$

Where $i \in [1, L]$ is the output dimension. $y_i^{target}(n)$ is the target output value of i-dimensional at step n, and $y_i^{predict}(n)$ is the predicted value of i-dimensional at step n. In the continuous-time domain, the dynamic model of ESN is given by

$$\dot{\mathbf{x}} = \frac{1}{c}(-\alpha\mathbf{x} + h(\mathbf{W}_{in}\mathbf{u} + \mathbf{W}\mathbf{x} + \mathbf{W}_{fb}\mathbf{y})) \tag{2-3}$$

Where $c > 0$ is the global scaling factor for the temporal dynamics, $0 < \alpha < 1$ is the leaking rate of reservoir neurons, and $h(\cdot)$ is a nonlinear transformation, and the sigmoid functions (tanh and Fermi) are commonly used. The $\mathbf{W}_{in} \in \mathbb{R}^{N\times K}$, $\mathbf{W} \in \mathbb{R}^{N\times N}$, and $\mathbf{W}_{fb} \in \mathbb{R}^{N\times L}$ represent the input weight matrix, internal weight matrix, and feedback weight matrix respectively. The value of these weight matrixes is randomly located from -0.5 to 0.5.

$$\begin{cases} \mathrm{Fermi}(x) = \dfrac{1}{1 + \exp(-x)} \\ \tanh(x) = 2\mathrm{Fermi}(2x) - 1 \end{cases} \tag{2-4}$$

The neurons' behaviors are related to the shape of the activation functions. During the selection of functions, a trade-off should be made between the dynamic richness and the computing complexity based on the premise of differentiable and boundedness.

According to the forward Euler discretization principle, the ESN model in the discrete-time domain can be approximated by

$$\mathbf{x}((t+1)\delta) = (1-\frac{\alpha\delta}{c})\mathbf{x}(t\delta) + \frac{\delta}{c}h(\mathbf{W}_{in}\mathbf{u}(t\delta) + \mathbf{W}\mathbf{x}(t\delta) + \mathbf{W}_{fb}\mathbf{y}(t\delta)) \tag{2-5}$$

Where δ is the discrete step size, which is determined by the sampling rate of the input signal. A proper value of δ could improve the discrete approximation ability and modeling stability. Using the notation $\mathbf{x}\big((t+1)\delta\big) = \mathbf{x}(n+1)$, the model of ESN can be transferred into the form

$$\mathbf{x}(n+1) = (1-\frac{\alpha\delta}{c})\mathbf{x}(n) + \frac{\delta}{c}h(\mathbf{W}_{in}\mathbf{u}(n) + \mathbf{W}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n)) \tag{2-6}$$

The output vector **y** is given by

$$\mathbf{y}((t+1)\delta) = g(\mathbf{W}_{out}[\mathbf{x}((t+1)\delta);\mathbf{u}(t\delta)]) \tag{2-7}$$

Where $g(\cdot)$ is nonlinear transformation (usually the identity or sigmoid function), $\mathbf{W}_{out} \in \mathbb{R}^{L\times(N+K)}$ is the output weight matrix. In the Euler discretization, the step size δ should be small enough to ensure the accuracy and stability of the approximation. The ESP is the essential feature of ESN which represents the reservoir working at a stable state, and the ESP can be designed as follows

$$\begin{cases} |\lambda|_{max}(\widehat{\mathbf{W}}) < 1 \ ; \ \textit{for} \text{ zero inputs} \\ |\lambda|_{max}(\widehat{\mathbf{W}}) \geq 1 \ ; \ \textit{for} \text{ non-zero inputs} \end{cases} \tag{2-8}$$

Where $|\lambda|_{max}(\widehat{\mathbf{W}})$ is the equivalent spectral radius (largest absolute eigenvalue) of $\widehat{\mathbf{W}}$ in the leaky integrator neuron, and the $\widehat{\mathbf{W}}$ can be represented as

$$\widehat{\mathbf{W}} = \frac{\delta}{c}\mathbf{W} + (1-\frac{\alpha\delta}{c})\mathbf{I} \tag{2-9}$$

Where **I** is the identity matrix. Besides, the neuron in the reservoir in a single update step should be less than its previous excitation. So, all the neurons in the reservoir should be under the rule of

$$1-\frac{\alpha\delta}{c} = 1-\alpha\xi \leq 0 \tag{2-10}$$

Where $\xi = \delta/c$ is the local scaling factor. This constraint is a necessary and insufficient condition for the ESP. In practice, the input signal usually pushes the neuron state from 0 to the smaller slop region of the hyperbolic tangent curve. So, the value of the spectral radius depends on the characteristics of the input signal and the expected output signal. Then the model of ESN can be rewritten as

$$\mathbf{x}(n+1) = (1-\alpha\xi)\mathbf{x}(n) + \xi h(\mathbf{W}_{in}\mathbf{u}(n) + \mathbf{W}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n)) \tag{2-11}$$

Assuming Γ represents an ESN with the weights of $\mathbf{W}_{in}$, $\mathbf{W}$, $\mathbf{W}_{fb}$, the leaking rate of α, and the scale gain of ξ, then the updated model of ENS can be expressed as

$$\begin{aligned}\frac{1}{\xi}\mathbf{x}(n+1) &= \frac{1}{\xi}(1-\alpha\xi)\mathbf{x}(n) + h(\mathbf{W}_{in}\mathbf{u}(n) + (\xi\mathbf{W})\frac{1}{\xi}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n)) \\ &= \frac{1}{\xi}(1-\tilde{\alpha})\mathbf{x}(n) + h(\mathbf{W}_{in}\mathbf{u}(n) + \tilde{\mathbf{W}}\frac{1}{\xi}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n))\end{aligned} \tag{2-12}$$

In the updated ESN of $\tilde{\Gamma}$, $\tilde{\alpha} = \xi\alpha$, $\tilde{\mathbf{W}} = \xi\mathbf{W}$. So $\tilde{\Gamma}$ has the same updated structure with Γ except for the scaling factor of $1/\xi$. The property of the updated internal weight matrix $\tilde{\mathbf{W}}$ can be scaled by its spectral radius, and the updated output vector $\tilde{\mathbf{y}}$ can be presented as

$$\tilde{\mathbf{y}}(n+1) = g(\tilde{\mathbf{W}}_{out}[\frac{1}{\xi}\mathbf{x}(n+1);\mathbf{u}(n)]) \tag{2-13}$$

When $\xi = \tilde{\xi} = 1$, the updated model and previous model are the alternative structures. Without loss of generality, the dynamic state in the reservoir and the outputs can be expressed as

$$\begin{cases}\mathbf{x}(n+1) = (1-\alpha)\mathbf{x}(n) + h(\mathbf{W}_{in}\mathbf{u}(n) + \mathbf{W}\mathbf{x}(n) + \mathbf{W}_{fb}\mathbf{y}(n)) \\ \mathbf{y}[n+1] = g(\mathbf{W}_{out}(\mathbf{x}[n+1];\mathbf{u}[n]))\end{cases} \tag{2-14}$$

Then the $\mathbf{W}_{out}$ can be calculated as

$$\mathbf{W}_{out} = \operatorname{argmin}\frac{1}{L}\sum_{i}^{L}\sqrt{\frac{1}{B}\sum_{n=1}^{B}(y_i^{predict}(n) - y_i^{target}(n))^2} \tag{2-15}$$

When the ridge regression is utilized, the $\mathbf{W}_{out}$ can be calculated by

$$\mathbf{W}_{out} = \mathbf{y}^{target}\mathbf{x}^{\mathrm{T}}(\mathbf{x}\mathbf{x}^{\mathrm{T}} + \beta\mathbf{I})^{-1} \tag{2-16}$$

Where β is the regularization parameter. The working principles of ESN can be divided into 4 steps and further summarized in **Algorithm 1**.

Step 1: Generate a reservoir and build up the ESN. The number of reservoir neurons N, spectral radius ρ of the internal matrix $\mathbf{W}$, and leaking rate α are determined. The input matrix $\mathbf{W}_{in}$ and internal matrix $\mathbf{W}$ are assigned respectively.

Step 2: Run the ESN and collect the reservoir activation states. The appropriate activation functions of $h(\cdot)$ and $g(\cdot)$ are selected, and the states of the neurons are initialized. The external input signal $\mathbf{u}$ is sent to the reservoir, and the neuron activation signal $\mathbf{x}$ and its updated state are calculated by Eq. (2-6).

Step 3: Calculate the output matrix. The training part aims to compute the output matrix $\mathbf{W}_{out}$ by Eq. (2-15) and Eq. (2-16). In the training dataset, the objective of the learning algorithm is minimizing the error between the target signal $\mathbf{y}^{target}$ the output signal of the ESN $\mathbf{y}^{predict}$.

Step 4: The new output vector $\hat{\boldsymbol{y}}$ is predicted. In the prediction dataset, the output weight matrix $\mathbf{W}_{out}$ and the new input vector $\hat{\boldsymbol{u}}$ are used to calculate the new outputs $\hat{\boldsymbol{y}}$ by Eq. (2-7).

Algorithm 1: Working principle of ESN

Input: $\mathbf{u}$, $\mathbf{y}^{target}$, B, K, N, L, α, ρ, β, $\mathbf{W}_{in}$, $\mathbf{W}$, $\mathbf{W}_{fb}$, $h(\cdot)$,$g(\cdot)$, $\hat{\boldsymbol{u}}$

Output: $\mathbf{W}_{out}$, $\hat{\boldsymbol{y}}$

Step1: Network initialization

$\mathbf{W}_{in} \in (-0.5, 0.5)$, $\mathbf{W} \in (-0.5, 0.5)$, $\mathbf{W}_{fb} = \mathbf{0}$,

$\alpha \in (0, 1)$, $\rho \in (0, 1.5)$, $N = 400$, $\mathbf{x}(0) = \mathbf{0}$

Step 2: Training based on historical input u

for n = 1 to B, **do**

update $\mathbf{x}$ by Eq. (2-6)

collect the state of $\mathbf{x}$ by Eq. (2-7)

collect $\mathbf{y}$ by Eq. (2-7)

end

Step 3: Calculate $\mathbf{W}_{out}$ by Eq. (2-15) and Eq. (2-16)

Step 4: Prediction based on new input $\hat{\boldsymbol{u}}$

for n = B+1 to B+m, **do**

calculate the new output $\hat{\boldsymbol{y}}$ by Eq. (2-7)

end

Performance evaluation ($\hat{\boldsymbol{y}}$ vs. $\hat{\boldsymbol{y}}^{target}$)

The parameters in ESN can be divided into three groups: assigned parameters, adjustable parameters, and calculated parameters.

The assigned parameters include the input weight matrix $\mathbf{W}_{in}$, recurrent weight matrix $\mathbf{W}$, input units K, and output units L. Gaussian distributions, symmetrical uniform, and normal distribution centered around zero are commonly used principles to generate the input weight matrix $\mathbf{W}_{in}$ and recurrent weight matrix $\mathbf{W}$. These two matrices are assumed to be fixed once they are generated randomly.

The adjustable parameters include the spectral radius ρ, the leaking rate α, the regularization parameter β, and the number of reservoir neurons N. The spectral radius is the maximal absolute eigenvalue of the matrix $\mathbf{W}$. It is an important parameter that controls the dynamic regime of the reservoir. To ensure the echo state property, the spectral radius is always less than 1 for zero inputs. Nevertheless, for nonzero inputs, the system usually has better performance when the spectral radius is bigger than 1 in practice. The leaking rate represents the update speed of the neurons in the reservoir.

The dynamic of the reservoir increases with the value of the leaking rate, and a large value of the leaking rate means that the output value of the reservoir at time step (t-1) has little impact on the reservoir state at time step (t). Generally speaking, when the reservoir includes more neurons, the performance would be better. It is easier to find a linear combination of the inputs to target outputs when the dimension of the reservoir increases. However, the computation time would increase at the same time. The selection of the number of reservoir neurons is to find a tradeoff between the prediction accuracy and the computation complexity. The regularization parameter is set by the accuracy of the target value and predicted value in the training part.

The calculated parameter is the output weight matrix $\mathbf{W}_{out}$, and it can be calculated by multi-linear regression (usually ridge regression). According to the above designing rules, the ESN for predicting the RUL of PEMFC is developed. The parameters of ESN can be determined by trial-and-error, grid-search, and some automatic optimization methods. The implementation framework of the basic ESN and some parameters are shown in Fig. 2 - 2 and Tab. 2 - 1.

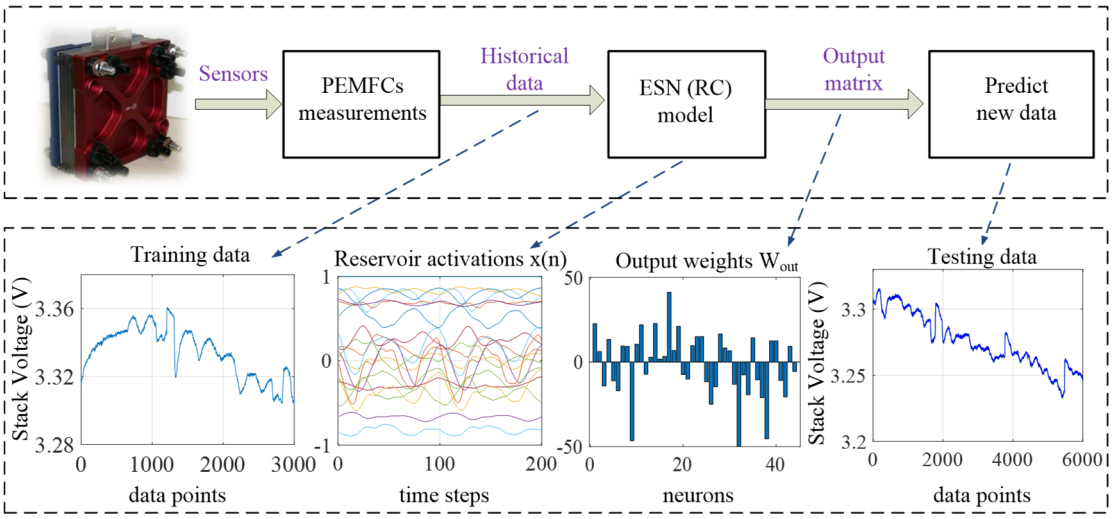


Fig. 2 - 2 Implementation framework of the ESN method.

Tab. 2 - 1 Key parameter of ESN.

Parameter	Values
reservoir neurons N	[50, 500]
leaking rate α	[0, 1]
spectral radius ρ	[0, 1.5]
regularization parameter β	[8×10^{-3}, 8×10^{-1}]
input weight matrix $\mathbf{W}_{in}$	[-0.5, 0.5]
recurrent weight matrix $\mathbf{W}$	[-0.5, 0.5]

2.3 Implementation process of MIMO-ESN

During the fuel cell running, many operating parameters are supervised to guide a more durable working lifespan, and some of the operating parameters have been shown in Fig. 2 - 3 and Fig. 2 - 4.

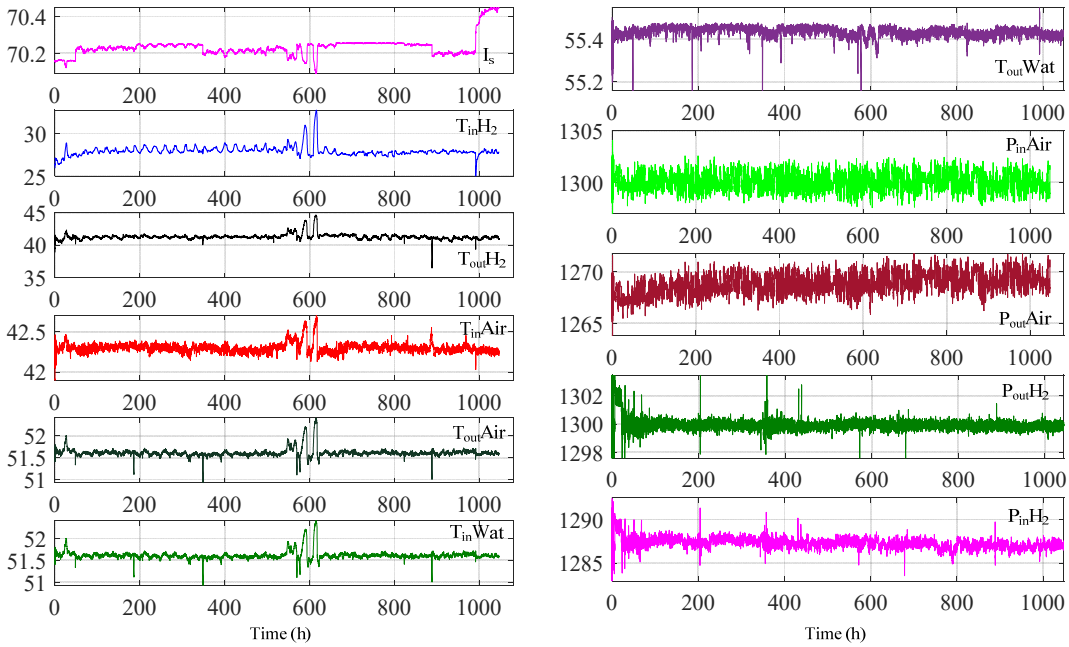


Fig. 2 - 3 Operating parameters of FC1.

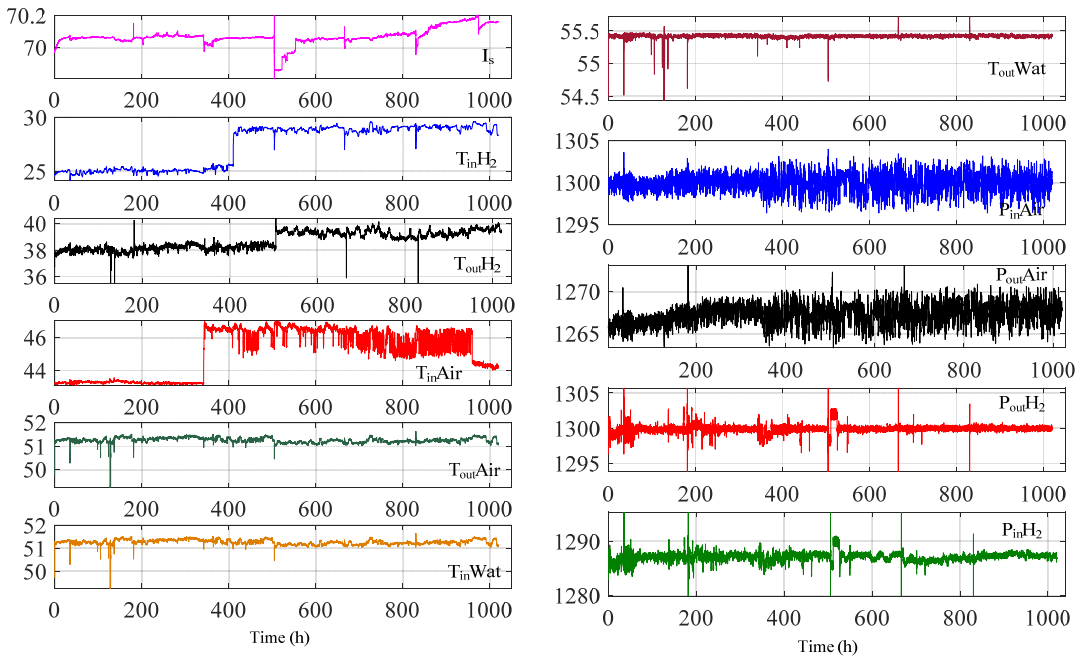


Fig. 2 - 4 Operating parameters of FC2.

The characteristics of all the parameters are shown in Tab. 2 - 2. Signal-to-Noise Ratio (SNR) is used to measure the quality of parameters, and it can be expressed as

$$SNR=10*\log_{10}(P_s / P_n) \tag{2-17}$$

Where P_s is the power of the signal and P_n is the power of the noise which is the difference between the raw data and the filtered data. Among all the parameters, the stack current (I_s) has the highest SNR both in FC1 and FC2.

In this task, MIMO-ESN is defined as

Single-input ESN: the input is the stack voltage (U_s).

Two-input ESN: the inputs are the combination of U_s and one of the parameters in Tab. 2 - 2.

Three-input ESN: the combination of U_s and two of the parameters in Tab. 2 - 2.

Tab. 2 - 2 The SNR of test-SS and test-QD.

Parameter	characteristics	SNR of test-SS (dB)	SNR of test-QD (dB)
I_s	Regulated (A)	**78.72**	**70.33**
$T_{in}H_2$	Measured (°C)	65.51	62.85
$T_{out}H_2$	Measured (°C)	45.58	43.06
$T_{in}Air$	Measured (°C)	55.25	52.32
$T_{out}Air$	Measured (°C)	61.42	57.18
$T_{in}Wat$	Regulated (°C)	55.39	52.65
$T_{out}Wat$	Regulated (°C)	59.42	56.12
$P_{in}Air$	Measured (mBar)	58.83	58.84
$P_{out}Air$	Regulated (mBar)	59.39	58.79
$P_{in}H_2$	Regulated (mBar)	50.93	51.57
$P_{out}H_2$	Regulated (mBar)	51.31	52.02

Therefore, there are C_{11}^1 =11 combinations in 2-input ESN, and C_{11}^2 = 55 combinations in 3-input ESN. The calculation flow chart of the MIMO-ESN method is shown in Fig. 2 - 5, where H denotes the number of the final step of prediction points (prediction horizon), p is the number of past discrete values used for prediction and q is the prediction step. The time interval between two points of the time series is a constant value (30 min).

For single-input ESN, an iterative one-step prediction method is utilized for the multi-step prediction. The sampling points of the stack voltage (U_s) can be expressed as $\{(t_1, y_1),\cdots,(t_i, y_i),\cdots,(t_H, y_H)\}$, where t_i represents the sampling time, y_i represents the stack voltage at the time t_i, and index H denotes the total number of data points. Firstly, the sampling points $\{(t_1, y_1),\cdots,(t_s, y_s)\}$ are trained to calculate the output weight matrix $\mathbf{W}_{out(s)}$. Once $\mathbf{W}_{out(s)}$ is obtained during the training process, it is fixed. Then based on $\mathbf{W}_{out(s)}$ and y_s, the voltage value of the next step is predicted as

$\hat{y}_{s+1}$. When the system reaches time t_{s+1}, $\hat{y}_{s+1}$ and $\mathbf{W}_{out(s)}$ are used to predict the $\hat{y}_{s+2}$. The same sequence is repeated for $\hat{y}_{s+3}$, $\hat{y}_{s+4}$, …, etc.

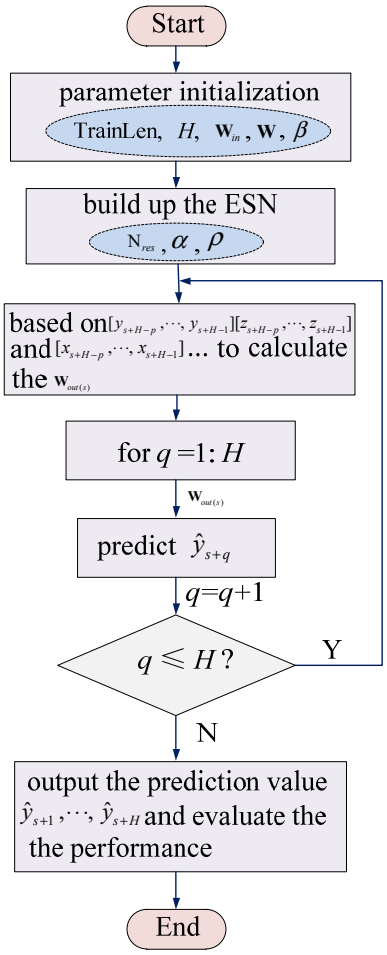


Fig. 2 - 5 Calculation flow chart of the MIMO-ESN prediction process.

For multi-input ESN, besides the one-step iterative prediction method for stack voltage, a one-step ahead prediction method is applied to the operating parameters. Taking the double inputs of stack voltage (U_s) and stack current (I_s) for example, the sampling data points can be expressed as $\{(t_1,y_1),\cdots,(t_i,y_i),\cdots,(t_H,y_H)\}$ and $\{(t_1,z_1),\cdots,(t_i,z_i),\cdots,(t_H,z_H)\}$ where z_i represents the stack current at the time t_i. Firstly, the sampling points $\{(t_1,y_1),\cdots,(t_s,y_s)\}$ and $\{(t_1,z_1),\cdots,(t_s,z_s)\}$ are trained to calculate the output weight matrix $\mathbf{W}_{out(s)}$. Then based on $\mathbf{W}_{out(s)}$, y_s and z_s, the voltage value of the next step is predicted as $\hat{y}_{s+1}$ and $\hat{z}_{s+1}$. After that, $\hat{y}_{s+1}$, z_{s+1} and $\mathbf{W}_{out(s)}$ are used to predict the $\hat{y}_{s+2}$ and $\hat{z}_{s+2}$. The same sequence is repeated for $\hat{y}_{s+3}$, $\hat{z}_{s+3}$, $\hat{y}_{s+4}$, $\hat{z}_{s+4}$,⋯, etc. Attention should be paid to the fact that the operating parameters such as stack current are usually scheduled variables and cannot be predicted in the same iterative way as the stack voltage. A major potential assumption herein is that the operating parameters can be schedulable or programmable such as in the homemade test benches and μ-CHP applications.

Based on the MIMO-ESN, the stack voltage is not only dependent on the historical

profile but also related to the operating conditions. Finally, the predicted voltages and the actual voltages are compared to evaluate the prediction performance. In a practical application, the output weight matrix W_{out} is firstly calculated offline, then the stack voltage at time step (t) and the W_{out} are used to predict the new value at the time step (t+1). The computing time during the prediction process is about 20 s (Matlab 2018a, 8G RAM, Core i5-2450 CPU @ 2.50 GHz) in this work and it is short enough for the several hundred hours' lifespan predictions. The block diagram of single-input and multi-input ESN is shown in Fig. 2 - 6. After the data measurement by sensors, the pre-processing means the filtering by moving average method.

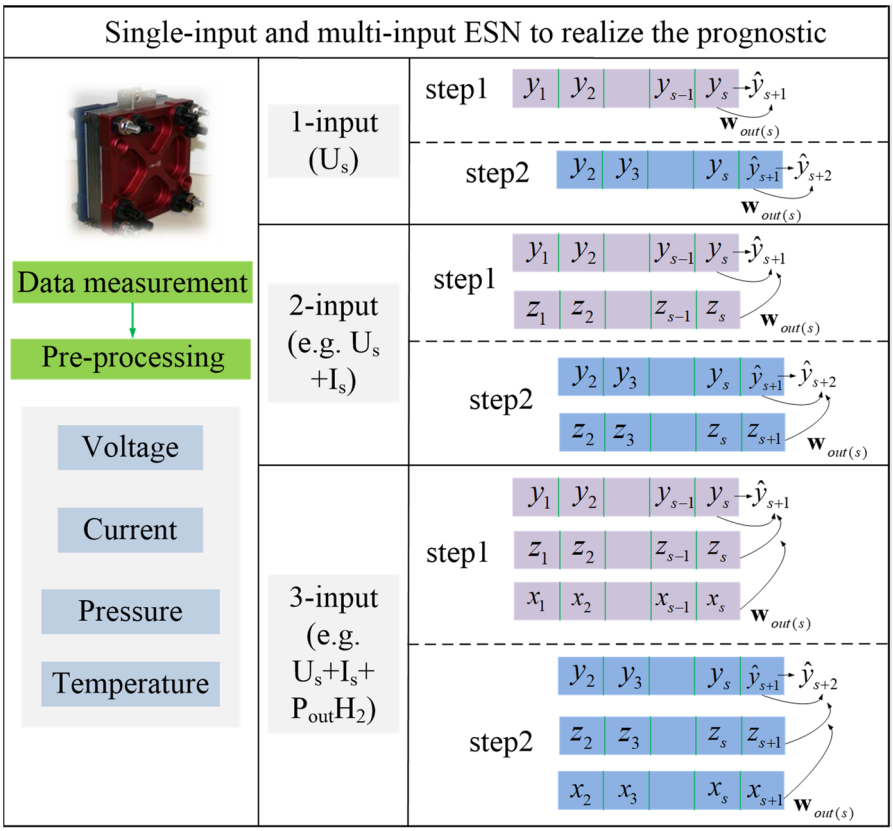


Fig. 2 - 6 Block diagram of single-input and multi-input ESN for one example with 2 steps in each case (where y_i represents U_s at time t_i, z_i represents the I_s at time t_i, x_i represents $P_{out}H_2$ at time t_i, $\hat{y}_{s+1}$ and $\hat{y}_{s+2}$ are next-step predicted stack voltage value, $\mathbf{W}_{out(s)}$ represents the output weight matrix).

2.4 Experimental results

2.4.1 MIMO-ESN under steady-state operating condition

In test-SS, the data that are used for the experiment come from a 1050 hours' duration test on the PEMFC stack. The data between 0 h and $t_{predict}$ are used for training, and the rest of the data are used for the prediction. The RUL time (t_{RUL}) can be considered as the time between the prediction time ($t_{predict}$) and the equivalent failure thresholds ($t_{failure}$). In fact, the failure time and the EoL time are not the same. Failure

time means the fuel cell is broken and doesn't produce power anymore. The EoL time means the produced power cannot fit anymore the application requirement, and it also has the ability to work. Considering the economic factors, the fresh stack should be used at the EoL time in most of the conditions. In the document of Data Challenge, the time at failure thresholds are regarded as the EoL time to calculate the RUL easily and validate the prediction methods simply. Thus, we also use this terminology in this work. In the Data Challenge, a certain voltage loss of initial voltage (V_{init}) has been regarded as the failure threshold. The actual values at different equivalent failure thresholds are 95.8 h (3.0 % V_{init}), 127.1 h (3.5 % V_{init}), 277.6 h (4.0 % V_{init}), 284.1 h (4.5 % V_{init}), and 354.5 h (5.0 % V_{init}) respectively. In the steady-state operation condition, the data from 0 h to 550 h are applied for training, and the data from 550 h to 1050 h are used for the prediction. With a moving window of 31 which is decided by trial and error, the recoveries and faults are retained by moving average filtering technique. The RUL prediction of FC1 based on single-input (U_s) ESN is shown in Fig. 2 - 7.

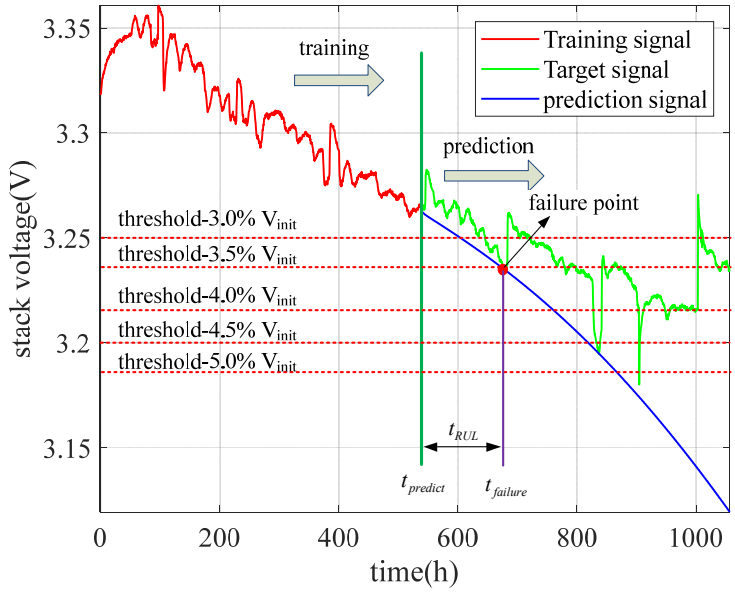


Fig. 2 - 7 The RUL prediction of FC1 based on single-input ESN.

With all the other parameters remaining unchanged, the number of inputs is increased to test the prediction performance. The RUL predictions of FC1 based on 2-input ESN are shown in Fig. 2 - 8 and the prediction results of 4 combinations at different equivalent failure thresholds are given in Tab. 2 - 3. The RMSE and MAPE of 2-input ESNs are lower than those of single-input ESN. Results represent that the prediction accuracy of 2-input ("U_s+T_{out}Air", "U_s+I_s", "U_s+T_{in}Wat", "U_s+$P_{out}H_2$") ESN is higher than that of single-input ESN. It also means that 2-input data are sufficient to mimic the degradation characteristics, and the prediction accuracy is improved by increasing another input to the ESN.

Furthermore, more parameters are regarded as the inputs of ESN to investigate its

prediction performance. The RUL predictions of FC1 based on 3-input ESN are shown in Fig. 2 - 9. The prediction results of 3-input ESN at different equivalent failure thresholds are given in Tab. 2 - 4. Results represent that some of the 3-input ESN ("U_s+$T_{in}H_2$+$P_{in}H_2$" and "U_s+$T_{out}H_2$+$T_{out}Air$") have a better prediction performance than 2-input ESN. Sometimes, they are worse than 2-input ESN ("U_s+$T_{in}Air$+$T_{out}Wat$" and "U_s+$P_{out}H_2$+$P_{in}H_2$").

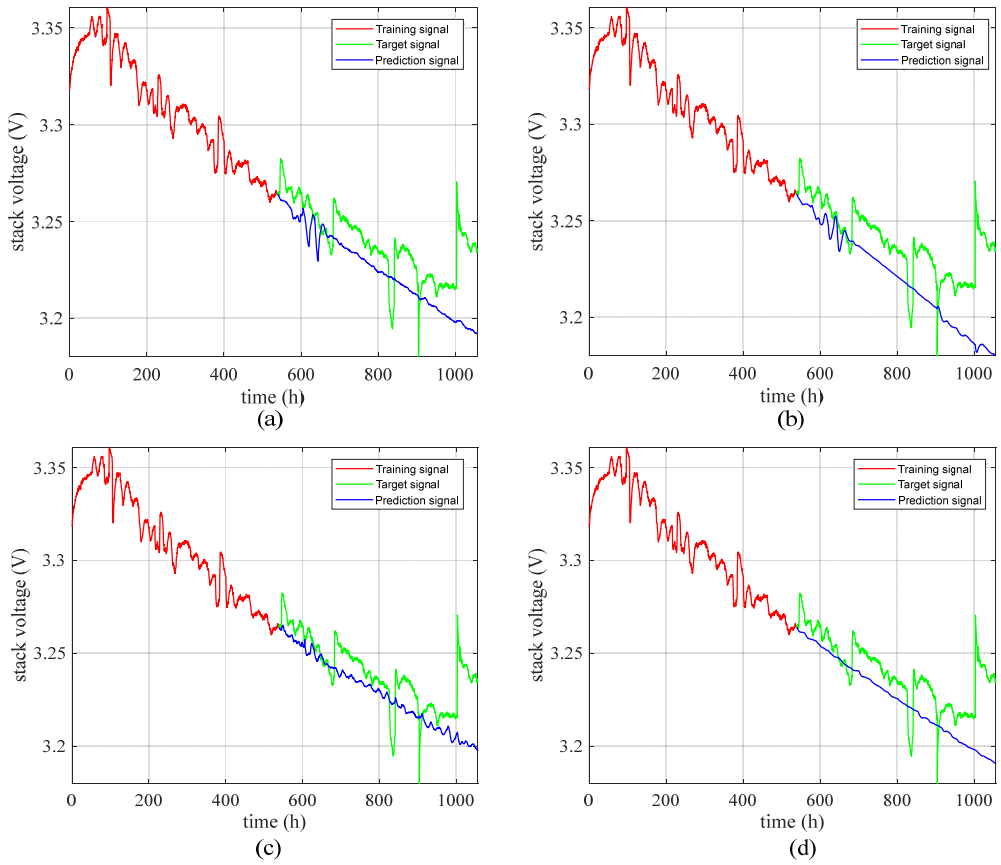


Fig. 2 - 8 The RUL prediction of FC1 based on 2-input ESN: (a) stack voltage (U_s) and outlet temperature of air ($T_{out}Air$). (b) stack voltage (U_s) and stack current (I_s). (c) stack voltage (U_s) and inlet temperature of cooling water ($T_{in}Wat$). (d) stack voltage (U_s) and outlet pressure of H_2 ($P_{out}H_2$).

Tab. 2 - 3 Prediction results of FC1 based on 2-input ESN.

Inputs	Prediction	$\%Er_{FT}$	RMSE	MAPE
	54.2	43.42	0.01188	0.00343
	136.8	-7.63	0.01077	0.00300
U_s	208.5	24.89	0.02039	0.00560
	272.3	4.15	0.02018	0.00551
	329.5	7.05	0.02710	0.00709
	45.5	52.51	0.01281	0.00341

U_s+ T_{out}Air	91.5	28.01	0.01146	0.00289
	320.0	-15.27	0.01190	0.00330
	441.2	-55.30	0.01211	0.00336
	500.0	-41.04	0.01339	0.00375
U_s+I_s	48.7	49.16	0.01143	0.00312
	175.5	-38.08	0.01048	0.00275
	281.8	-1.51	0.01295	0.00365
	366.5	-29.00	0.01297	0.00367
	456.0	-28.63	0.01532	0.00429
U_s+ T_{in}Wat	59.6	37.79	0.00902	0.00242
	209.3	-64.67	0.00808	0.00205
	344.6	-24.14	0.00931	0.00249
	500.0	-75.99	0.00972	0.00258
	500.0	-41.04	0.01106	0.00295
U_s+ $P_{out}H_2$	77.0	19.62	0.00910	0.00245
	194.7	-53.19	0.00815	0.00207
	310.0	-11.67	0.01017	0.00281
	431.6	-51.92	0.01045	0.00288
	500.0	-41.04	0.01236	0.00339

Tab. 2 - 4 Prediction results of FC1 based on 3-input ESN.

Inputs	Prediction	$\%Er_{FT}$	RMSE	MAPE
U_s +$T_{in}H_2$ +$P_{in}H_2$	71.6	25.26	0.00937	0.00239
	98.2	22.74	0.00930	0.00239
	358.5	-29.14	0.00850	0.00226
	500.0	-75.99	0.00930	0.00239
	500.0	-41.04	0.01051	0.00271
U_s +$T_{out}H_2$ +T_{out}Air	41.5	56.68	0.01047	0.00284
	186.8	-46.97	0.00956	0.00252
	347.5	-25.18	0.00959	0.00258
	500.0	-75.99	0.01001	0.00267
	500.0	-41.04	0.01075	0.00287
U_s +T_{out}Wat	73.2	23.59	0.00904	0.00236
	188.4	-48.23	0.00817	0.00207
	307.4	-10.73	0.01061	0.00293

$+T_{in}Air$	436.8	-53.75	0.01080	0.00298
	500.0	-41.04	0.01266	0.00349
U_s $+P_{out}H_2$ $+P_{in}H_2$	81.2	15.24	0.00789	0.00199
	226.1	-77.89	0.00722	0.00177
	370.3	-33.39	0.00819	0.00219
	500.0	-75.99	0.00898	0.00232
	500.0	-41.04	0.01027	0.00265

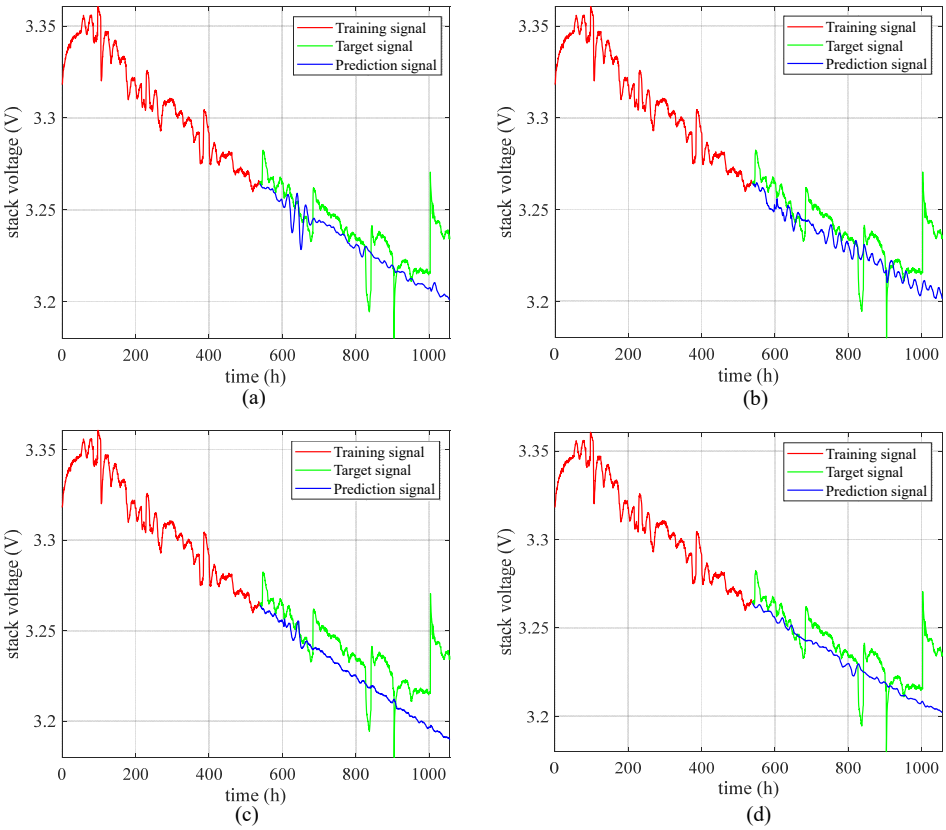


Fig. 2 - 9 The RUL prediction of FC1 based on 3-input ESN: (a) stack voltage (U_s), inlet temperature of H_2 ($T_{in}H_2$) and inlet pressure of H_2 ($P_{in}H_2$). (b) stack voltage (U_s), outlet temperature of H_2 ($T_{out}H_2$) and outlet temperature of air ($T_{out}Air$). (c) stack voltage (U_s), inlet temperature of air ($T_{in}Air$) and outlet temperature of cooling water ($T_{out}Wat$). (d) stack voltage (U_s), outlet pressure of H_2 ($P_{out}H_2$), and inlet pressure of H_2 ($P_{in}H_2$).

2.4.2 MIMO-ESN under quasi-dynamic operating condition

In test-QD, the data set come from a 1020 hours' duration test on the PEMFC. For FC2, the actual values at different failure thresholds are 21.4 h (3.5 % V_{init}), 194.2 h (4.0 % V_{init}), 209.7 h (4.5 % V_{init}), 384.3 h (5.0 % V_{init}), and 386.7 h (5.5 % V_{init}) respectively. The data from 0 h to 550 h are applied to training, and the data from 550 h to 1020 h are used for prediction. The RUL prediction of FC2 based on single-input

ESN is shown in Fig. 2 - 10. The RUL prediction of FC2 based on 2-input ESN is shown in Fig. 2 - 11 and the prediction results of 4 combinations at different equivalent failure thresholds are given in Tab. 2 - 5. The RMSE and MAPE of 2-input ESN are lower than those of the single-input one.

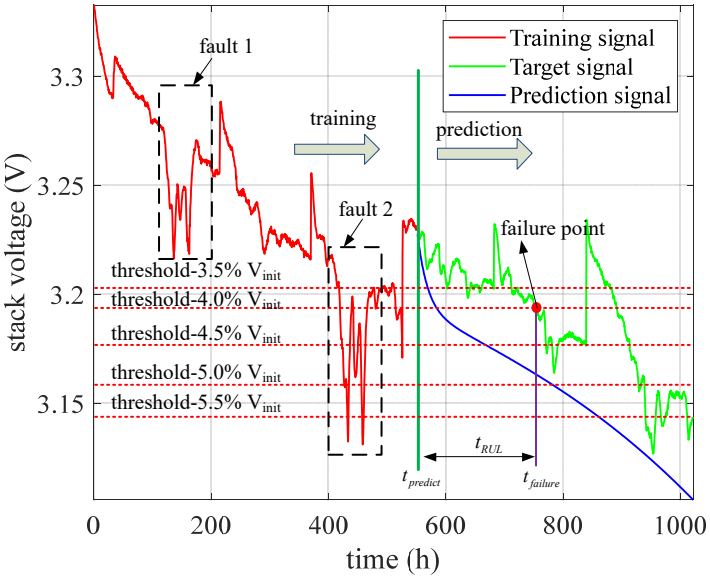


Fig. 2 - 10 The RUL prediction of FC2 based on single-input ESN.

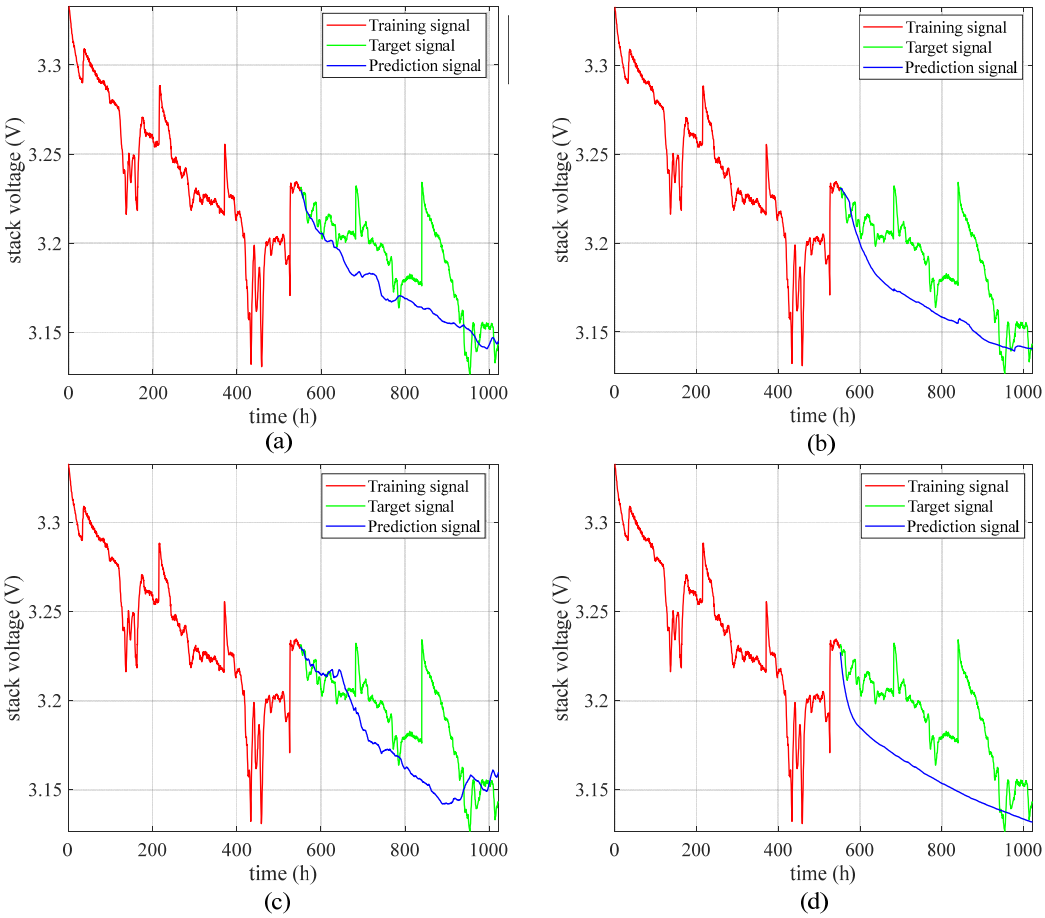


Fig. 2 - 11 The RUL prediction of FC2 based on 2-input ESN: (a) stack voltage (U_s) and outlet

temperature of air ($T_{out}Air$). (b) stack voltage (U_s) and stack current (I_s). (c) stack voltage (U_s) and inlet temperature of water ($T_{in}Wat$). (d) stack voltage (U_s) and outlet pressure of H_2 ($P_{out}H_2$).

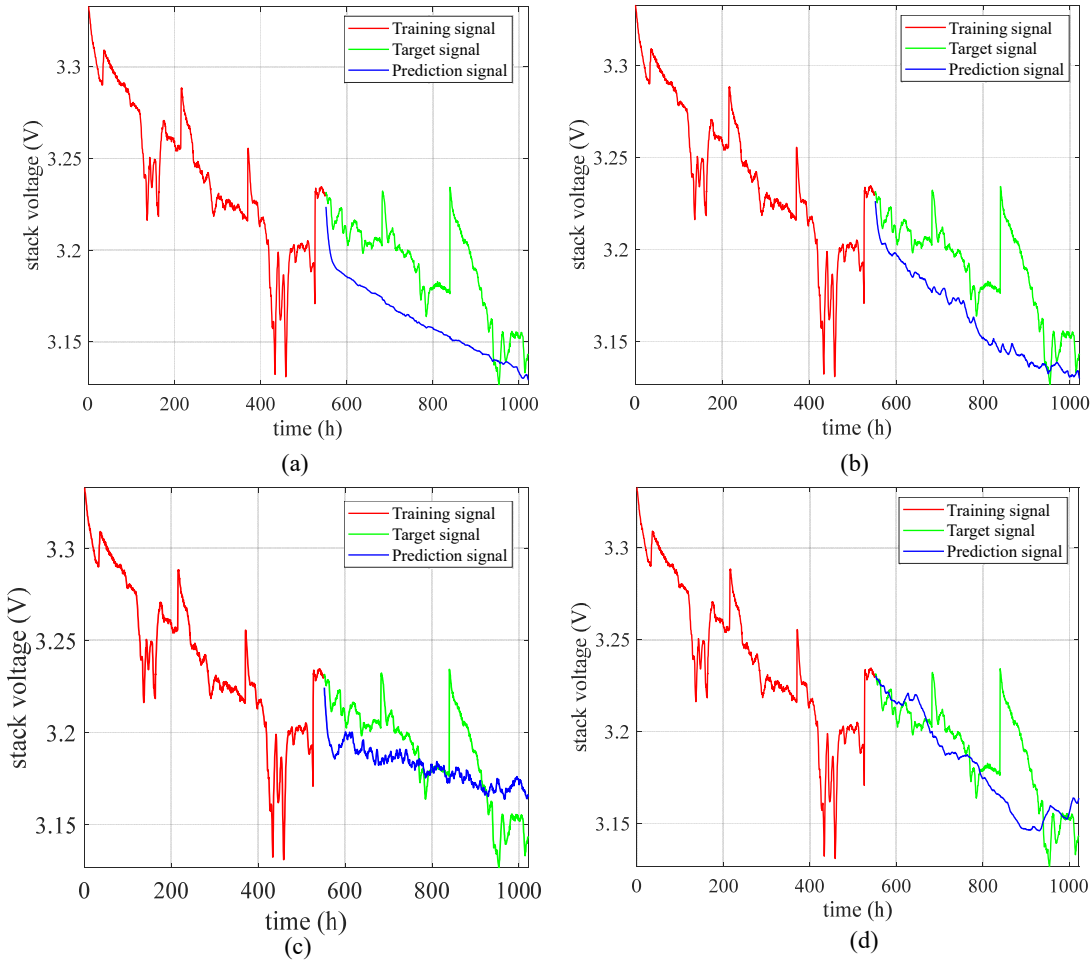


Fig. 2 - 12 The RUL prediction of FC2 based on 3-input ESN: (a) stack voltage (U_s), inlet temperature of H_2 ($T_{in}H_2$) and inlet pressure of H_2 ($P_{in}H_2$). (b) stack voltage (U_s), outlet temperature of H_2 ($T_{out}H_2$) and outlet temperature of air ($T_{out}Air$). (c) stack voltage (U_s), inlet temperature of air ($T_{in}Air$) and outlet pressure of air ($P_{out}Air$). (d) stack voltage (U_s), inlet temperature of water ($T_{in}Wat$), and inlet pressure of H_2 ($P_{in}H_2$).

Tab. 2 - 5 Prediction results of FC2 based on 2-input ESN.

Inputs	Prediction	$\%Er_{FT}$	RMSE	MAPE
	12.5	41.59	0.01839	0.00520
	36.0	81.46	0.02915	0.00868
U_s	122.5	41.58	0.02900	0.00866
	222.8	42.02	0.03811	0.01082
	314.2	18.75	0.03789	0.01078
	34.8	-62.62	0.00652	0.00154
	100.2	48.40	0.01864	0.00484

U_s+ T_{out}Air	189.1	9.82	0.01878	0.00495
	322.2	16.16	0.02532	0.00640
	428.5	-10.81	0.02495	0.00627
U_s+I_s	35.2	-64.49	0.00820	0.00205
	63.0	67.56	0.02692	0.00749
	117.7	43.87	0.02677	0.00751
	238.2	38.02	0.03271	0.00906
	383.6	0.80	0.03231	0.00892
U_s + T_{in}Wat	104.3	-387.38	0.00516	0.00126
	134.8	30.59	0.01639	0.00401
	173.1	17.45	0.01640	0.00408
	262.2	31.77	0.02959	0.00710
	335.8	13.16	0.02921	0.00700
U_s + $P_{out}H_2$	10.4	51.40	0.02166	0.00623
	27.9	85.63	0.03287	0.00989
	94.5	54.94	0.03268	0.00986
	202.5	47.31	0.03828	0.01116
	344.5	10.91	0.03783	0.01099

Under the quasi-dynamic condition, the 2-input ("U_s+T_{out}Air", "U_s+I_s", "U_s+T_{in}Wat", and "U_s+$P_{out}H_2$") ESN perform better than single-input ESN. They also represent that the prediction accuracy is improved by increasing the number of inputs. The prediction error of FC2 is bigger than that of FC1 when the training length (0-550 h) is the same for the prediction processes of FC2 have more perturbations than FC1. The results also mean that the ESN structure has more difficulties mimicking the quasi-dynamic operation when compared with the static operation.

The RUL predictions of FC2 based on 3-input ESN are shown in Fig. 2 - 12. The prediction results of 3-input ESN at different equivalent failure thresholds are given in Tab. 2 - 6. Results represent that some of the 3-input ESN ("U_s+T_{in}Wat+$P_{in}H_2$" and "U_s+P_{out}Air+T_{in}Air") have a better prediction performance than 2-input ESN. Sometimes, they are worse than 2-input ESN ("U_s+$T_{in}H_2$+$P_{in}H_2$" and "U_s+$T_{out}H_2$+T_{out}Air").

Tab. 2 - 6 Prediction results of FC2 based on 3-input ESN.

Inputs	Prediction	%Er_{FT}	RMSE	MAPE
U_s +$T_{in}H_2$	5.9	72.43	0.02594	0.00787
	18.2	90.63	0.03147	0.00959
	119.5	43.01	0.03122	0.00953

$+P_{in}H_2$	220.1	42.73	0.03608	0.01052
	376.5	2.64	0.03564	0.01036
	9.5	55.61	0.01677	0.00493
U_s	66.0	66.01	0.02405	0.00705
$+T_{out}H_2$	147.2	29.80	0.02404	0.00708
$+T_{out}Air$	232.1	39.60	0.03522	0.00974
	308.1	20.33	0.03487	0.00964
	6.0	71.96	0.02631	0.00799
U_s	15.5	92.02	0.02048	0.00587
$+T_{in}Air$	234.6	-11.87	0.01985	0.00561
$+P_{out}Air$	500.0	-30.11	0.02042	0.00520
	500.0	-29.30	0.02024	0.00516
	114.7	-435.98	0.00593	0.00141
U_s	148.1	23.74	0.01175	0.00300
$+T_{in}Wat$	246.5	-17.55	0.01140	0.00288
$+P_{in}H_2$	309.2	19.54	0.02466	0.00554
	500.0	-29.30	0.02435	0.00548

2.5 Multi-input discussion

In order to test all possible combinations, the prediction results based on all the 2-input ESN of FC1 and FC2 are shown in Tab. 2 - 7 and Tab. 2 - 8, respectively. The $RMSE_i$ (i =1, 2, …, 5) are the RMSE between the prediction time and the EoL time. In FC1, the $RMSE_5$ is the RMSE between 550.0 h to 904.5 h, and the $RMSE_5$ in FC2 is the RMSE between 550.0 h to 936.7 h. Therefore, the $RMSE_5$ is used to quantify the prediction accuracy. The $RMSE_5$ of single-input ESN in FC1 and FC2 are 0.02710 and 0.03789, respectively.

Results show that the $RMSE_5$ of randomly 2-input ESN is smaller than single-input ESN both in FC1 and FC2. The optimal combination in FC1 is the U_s with $T_{in}Wat$ (improved 59.20 %) and the optimal combination in FC2 is the U_s with $T_{out}Air$ (improved 34.15 %). All the other 3-input combinations are tested to verify the prediction accuracy. The results of FC1 and FC2 are shown in Tab. 2 - 9 and Tab. 2 - 10, respectively. The numbers (from 9 to 19) in Tab. 2 - 9 and Tab. 2 - 10 represent the inputs (from I_s to $P_{in}H_2$) in Tab. 2 - 7 and Tab. 2 - 8. Comparing the results of 2-input ESN and 3-input ESN in both FC1 and FC2, some of the 3-input ESN has a better performance than 2-input ESN, but some combinations have worse performance instead. Even some 3-input ESN combinations are worse than single-input ESN. This is most likely due to the interaction between different inputs.

In the results of 3-input ESN, “better” (in green) means the $RMSE_5$ is smaller than any 2-input combinations, “worse” (in purple) means the $RMSE_5$ is bigger than any 2-input combinations, and “middle” (in blue) means the $RMSE_5$ is in between of the 2 combinations of 2-input. For example, in Tab. 2 - 9, the “U_s+10+11” is “better” (in green), and the $RMSE_5$ of “U_s+10+11” (0.01216) is smaller than “U_s+10” (0.01294) or “U_s+11” (0.01380); the “U_s+9+12” is “middle” (in blue), and the $RMSE_5$ of “U_s+9+12” (0.01505) is bigger than “U_s+12” (0.01260) and smaller than “U_s+9” (0.01532); the “U_s+18+15” is “worse” (in purple), and the $RMSE_5$ of “U_s+15+18” (0.01463) is bigger than “U_s+18” (0.01236) and “U_s+15” (0.01440).

Note: There is a small mistake in our previous published paper (Z. Hua, *et al.*, "Remaining useful life prediction of PEMFC systems based on the multi-input echo state network," *Applied Energy.* [106]). The “Us+18+19” should be “better” (in green), and the $RMSE_5$ of “U_s+18+19” (0.01027) is smaller than “U_s+18” (0.01236) or “U_s+19” (0.01168). Thus, there are 40 “better” combinations and 11 “middle” combinations in FC1.

There are 40 “better” combinations and 11 “middle” combinations in FC1 and the “U_s+T_{out}Air(13)+T_{in}Wat(14)” has the best performance. There are 21 “better” combinations and 24 “middle” combinations in FC2 and the “U_s+T_{in}Air(12)+ P_{out}Air(17)” has the best performance. Besides, when comparing all the “better” combinations of 3-input to the 2-input ESN, there are 12 combinations in FC1 that have a smaller $RMSE_5$ than the 2-input ESN of “U_s+T_{in}Wat”. Nevertheless, there are 3 combinations (“U_s+9+12”, “U_s+12+17”, and “U_s+14+19”) in FC2 that could have a smaller $RMSE_5$ than the 2-input ESN of “U_s+T_{out}Air”.

The results show that 3-input ESN has the potential ability to improve the performance than 2-input ESN, especially under the static condition. But when considering the interaction of input variables, the selection of inputs is a rewarding area for further study. Considering the prediction accuracy and the computational complexity, the 2-input ESN has a top priority in both FC1 and FC2.

Tab. 2 - 7 The improvement results of FC1 based on 2-input ESN.

Inputs	$RMSE_5$	Improvement (%)
U_s	0.02710	--
U_s+T_{in}Wat(14)	**0.01106**	**59.20**
U_s+$P_{in}H_2$(19)	0.01168	56.90
U_s+$P_{out}H_2$(18)	0.01236	54.37
U_s+T_{in}Air(12)	0.01260	53.51
U_s+P_{in}Air(16)	0.01265	53.31
U_s+$T_{in}H_2$(10)	0.01294	52.24

U_s+P_{out}Air(17)	0.01336	50.71
U_s+T_{out}Air(13)	0.01339	50.59
U_s+$T_{out}H_2$(11)	0.01380	49.08
U_s+T_{out}Wat(15)	0.01440	46.84
U_s+I_s(9)	0.01532	43.47

Tab. 2 - 8 The improvement results of FC2 based on 2-input ESN.

Inputs	$RMSE_5$	Improvement (%)
U_s	0.03789	--
U_s+T_{out}Air(13)	**0.02495**	**34.15**
U_s+T_{in}Wat(14)	0.02921	22.90
U_s+T_{in}Air(12)	0.03023	20.21
U_s+$T_{out}H_2$(11)	0.03039	19.79
U_s+I_s(9)	0.03231	14.74
U_s+$T_{in}H_2$(10)	0.03242	14.43
U_s+P_{out}Air(17)	0.03292	13.13
U_s+T_{out}Wat(15)	0.03587	5.34
U_s+$P_{in}H_2$(19)	0.03589	5.30
U_s+PinAIR(16)	0.03715	1.96
U_s+$P_{out}H_2$(18)	0.03783	0.16

Tab. 2 - 9 The prediction results based on all 3-input combinations of FC1.

Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$
U_s+9+10	0.01179	U_s+10+12	0.01053	U_s+11+15	0.01231	U_s+12+19	0.01070	U_s+14+19	0.01020
U_s+9+11	0.01369	U_s+10+13	0.01158	U_s+11+16	0.01161	**U_s+13+14**	**0.00981**	U_s+15+16	0.01459
U_s+9+12	0.01505	U_s+10+14	0.01052	U_s+11+17	0.01227	U_s+13+15	0.01227	U_s+15+17	0.01461
U_s+9+13	0.01401	U_s+10+15	0.01155	U_s+11+18	0.01205	U_s+13+16	0.01221	U_s+15+18	0.01463
U_s+9+14	0.01119	U_s+10+16	0.01090	U_s+11+19	0.01120	U_s+13+17	0.01252	U_s+15+19	0.01374
U_s+9+15	0.01168	U_s+10+17	0.01149	U_s+12+13	0.01163	U_s+13+18	0.01242	U_s+16+17	0.01387
U_s+9+16	0.01187	U_s+10+18	0.01134	U_s+12+14	0.01120	U_s+13+19	0.01141	U_s+16+18	0.01166
U_s+9+17	0.01230	U_s+10+19	0.01051	U_s+12+15	0.01266	U_s+14+15	0.01142	U_s+16+19	0.01109
U_s+9+18	0.01193	U_s+11+12	0.01175	U_s+12+16	0.01100	U_s+14+16	0.01035	U_s+17+18	0.01210
U_s+9+19	0.01215	U_s+11+13	0.01075	U_s+12+17	0.01179	U_s+14+17	0.01068	U_s+17+19	0.01180
U_s+10+11	0.01216	U_s+11+14	0.01135	U_s+12+18	0.01125	U_s+14+18	0.01087	U_s+18+19	0.01027

Tab. 2 - 10 The prediction results based on all 3-input combinations of FC2.

Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$	Inputs	$RMSE_5$
U_s+9+10	0.03104	U_s+10+12	0.02763	U_s+11+15	0.03620	U_s+12+19	0.03097	U_s+14+19	0.02435
U_s+9+11	0.03055	U_s+10+13	0.03467	U_s+11+16	0.03334	U_s+13+14	0.02912	U_s+15+16	0.03511
U_s+9+12	0.02371	U_s+10+14	0.03659	U_s+11+17	0.03037	U_s+13+15	0.03288	U_s+15+17	0.03261
U_s+9+13	0.02727	U_s+10+15	0.03661	U_s+11+18	0.03497	U_s+13+16	0.02782	U_s+15+18	0.03782
U_s+9+14	0.03248	U_s+10+16	0.03377	U_s+11+19	0.03457	U_s+13+17	0.02709	U_s+15+19	0.03675
U_s+9+15	0.03865	U_s+10+17	0.03055	U_s+12+13	0.02817	U_s+13+18	0.02911	U_s+16+17	0.03146
U_s+9+16	0.03555	U_s+10+18	0.03560	U_s+12+14	0.03250	U_s+13+19	0.02765	U_s+16+18	0.03498
U_s+9+17	0.02954	U_s+10+19	0.03564	U_s+12+15	0.03202	U_s+14+15	0.03558	U_s+16+19	0.03295
U_s+9+18	0.03272	U_s+11+12	0.02523	U_s+12+16	0.02778	U_s+14+16	0.02946	U_s+17+18	0.02985
U_s+9+19	0.02982	U_s+11+13	0.03487	**U_s+12+17**	**0.02024**	U_s+14+17	0.02904	U_s+17+19	0.02831
U_s+10+11	0.02674	U_s+11+14	0.03625	U_s+12+18	0.03119	U_s+14+18	0.02825	U_s+18+19	0.03775

2.6 Chapter summary

In order to improve the prediction performance, a multi-input and multi-output ESN (MIMO-ESN) method is proposed in this chapter. Besides the commonly used stack voltage, the operating parameters, such as stack current, stack temperature, and the pressures of the reactants are also utilized as the inputs of ESN. The 1-input, 2-input, and 3-input ESN are designed and tested on the datasets which were acquired from the test bench for more than 1000 hours' duration. The feasibility and effectiveness of the proposed MIMO-ESN are verified under both static and quasi-dynamic operation conditions. Experimental results show that MIMO-ESN has the capability to improve the prediction accuracy, especially the 2-input ESN with an improvement of 59.20 % (static condition) and 34.15 % (quasi-dynamic condition) in terms of $RMSE_5$ respectively compared with the traditionally used SISO-ESN.

Chapter 3. Double-input echo state network of three dynamic tests

3.1 Brief introduction

An effective health indicator (HI) could help to define an explicit degradation state and improve the prediction accuracy. The HIs of voltage and power are usually used under steady-state and quasi-dynamic operating conditions due to their overall monotonic decreasing characteristics (even if some recovering episodes can occur), and their measurements of them are implemented easily in practice. Nevertheless, the static HIs are unable to be directly used under the dynamic operating conditions because they are sensitive to the mission profiles which leads to variations of the current which mainly impact the voltage and the power, without meaning aging. To overcome the weakness of static HIs, a convenient and practical HI named relative power-loss rate (RPLR) is proposed herein. According to the polarization curve at the beginning of life (BoL), the initial power under different mission profiles can be identified. Then the actual power is obtained by monitoring the current and voltage continuously. Finally, the RPLR is calculated based on the initial power and actual power.

Based on previous work, the double-input echo state network (DI-ESN) has a top priority in practical use. For the dynamic operating conditions, the historical RPLR is taken as the input of ESN and the predicted RPLR is regarded as the output in the single-input ESN (SI-ESN) structure. The parameters and RPLR can be used as the multiple inputs of ESN to improve the prediction accuracy. Nevertheless, the operating parameters are unavailable in most of the conditions. In the μ-CHP applications, the scheduled current is an extremely meaningful operating parameter because it could reflect the system's dynamic characteristics. Besides, the degradation speed is also related to the load current. Therefore, the DI-ESN structure is used to further explore the characteristics under the full-dynamic operating conditions (test-D1, test-D2, and test-D3), the scheduled stack current, and the RPLR are combined to predict the degradation tendency.

3.2 Dynamic health indicator extraction

In the steady-state and quasi-dynamic operating conditions, the deviations of voltage and power are considered to be influenced only by the aging degradation. Nevertheless, the voltage and power are also influenced by the mission profiles and even more specifically the current profile, and it is improper to take them as the HIs in the dynamic or time-varying operating conditions. Finding the dynamic HI is difficult and just a few papers have presented the handing methods. Two approaches are investigated. One way is to transform the dynamic condition into the static representation [102], [103]. Another way is to extract some monotonic parameters during the degradation [38], [40].

In [102], the current profile changes periodically between 0 A and 8 A, and the stack voltage also changes periodically between 9.5 V and 13.5 V. In the 1500 hours' durability test, the voltage is divided into about 1000 segments (1.5 h for one segment). A series of linear parameter varying (LPV) equivalent models are built to fit the fixed-size window voltage segments [102]. In each LPV model, the input is the real current and the output is the real voltage, and the output in the steady-state condition y_{inf} can be extracted by

$$y_{inf} = \frac{\sum_{j=0}^{n_b} b_j(p_{inf})u_{inf} + bias(p_{inf})}{\sum_{i=1}^{n_a} a_i(p_{inf}) + 1} \tag{3-1}$$

Where u_{inf}, p_{inf}, and y_{inf} are the input, scheduled variable, and output in a steady state. a_i and b_j are the coefficients which depend on the instantaneous value of $p(k)$. Then y_{inf} in each segment is connected to be the health indicator (virtual steady-state stack voltage) Later, an ensemble ESN structure is used in [103] to enhance the prognostic adaptability. Nevertheless, the extraction of HI is time-consuming, and the computing time is based on the number of segments. It is worth mentioning that this method is more suitable for the conditions in which the load varies regularly (repeating cycles), and the failure threshold is set as the same quantity as the last measurement [102].

In [38], an empirical model is used to fit each polarization curve (one measurement in each week)

$$V_{st} = n(E_0 - A \cdot T \cdot \ln(\frac{i}{i_0}) - R \cdot i - B \cdot T \cdot \ln(1 - \frac{i}{i_L})) \tag{3-2}$$

Where V_{st} is the stack voltage, n is the cell numbers of the stack, i is the current density, T is the temperature, A is the Tafel constant, and B is the concentration constant. The Levenberg-Marquardt optimization algorithm is utilized to obtain the parameters of open-circuit voltage E_0, exchange current density i_0, overall resistance R and limiting current density i_L. Compared with their initial value, the overall resistance R increase by more than 70 %, and the limiting current density i_L decreases by 60 %. Considering the R and i_L are linearly varying

$$\begin{cases} R(\mathrm{t}) = R_0(1 + \alpha(\mathrm{t})) \\ I_L(\mathrm{t}) = I_{L0}(1 - \alpha(\mathrm{t})) \\ \alpha(\mathrm{t}) = \beta t \end{cases} \tag{3-3}$$

Regarding the inputs are current load and temperature, the output is the voltage, the system states are α_k (an indicator of degradation) and β_k (speed of degradation). An extended Kalman filter (EKF) is adopted to evaluate the degradation condition. The EoL is the time when the indicator of EKF-based degradation factor (EKF-DF) α_k reaches to the maximum value (0.7). Then, the inverse first-order reliability method (IFORM) is

used in [40] to predict the RUL. Meanwhile, the confidence in the RUL prediction and the uncertainty of the estimation (caused by the covariance matrices) under the dynamic operating condition are also considered. Nevertheless, the HI comes from the periodic measurements of the polarization curves which are characterized every week in practice. Therefore, the extracted parameters are very limited to describing the complicated degradation phenomenon. Besides, the linearization hypothesis of overall resistance and limiting current density is inaccurate for the nonlinear PEMFC systems [119].

Finding other general and convenient HI under the dynamic operating condition is necessary for practical use. In the dynamic tests, the measurements of polarization curves at the BoL can be easily performed. Based on the polarization curves (voltage versus current) at the BoL, the BoL power curve (power versus current) can be obtained. Then a mathematical model is used to fit the BoL power curve to be able to compute easily the power for any value of the current. Then, the power supplied by the fresh stack (P_0) under different loads can then be computed by the model. In practice, during the lifetime of the stack (here the duration of the test), the current power (P) can be calculated from the stack current and stack voltage measurements. Based on the current power (P) and the BoL power (P_0) corresponding to the same current which is extracted from the models, the HI of RPLR is defined as

$$\mathrm{RPLR} = (P - P_0) / P_0 \tag{3-4}$$

So, the calculation of RPLR is conducted as follows:

Step1: measurements of BoL power (P_0) are carried out at different current values. A mathematical model of P_0 is built, in which parameters are identified thanks to the experimental values.

Step2: at time step t, current power (P_t) can be calculated by the stack current (I_t) and stack voltage (U_t).

Step3: calculate the BoL power (P_0-I_t) under the stack current (I_t) based on the model.

Step4: calculate the RPLR at time step t based on the current power (P_t) and BoL power (P_0-I_t). With the same method, the RPLR during its lifetime can be calculated. The calculation flow chart of the RPLR is shown in Fig. 3 - 1.

Compared with the LPV models under the dynamic operating conditions [102], the computational complexity of RPLR is lower. Compared with the linearization technique in [38], the time interval of each two points is no longer limited by the interval of polarization curves. Besides, the polarization curves are given in the whole life because they have been measured in advance [38]. To calculate the RPLR, only the polarization curve at the BoL needs to be performed, and the normal operation of the system would not be interrupted in a real system. This provides the implementation foundation for the

data-driven methods.

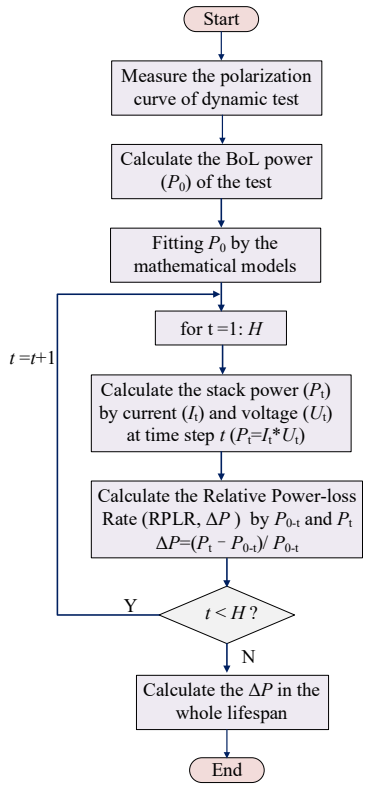


Fig. 3 - 1 The calculation flow chart of the RPLR.

To validate the effectiveness of the RPLR, three dynamic tests with different durations are implemented under various operating conditions. For the tuning of the parameters, the trust-region optimization method is used in the BoL power modeling process. The parameters and the accuracy of these three models are shown in Tab. 3 - 1. The polarization and the BoL power curves are shown in Fig. 3 - 2, and the typical current points of each experimental test are marked in red circles.

Tab. 3 - 1 Parameters of three mathematical models.

Test	Parameters			Accuracy	
D1	a_1	b_1	c_1	R-square	RMSE
	861.9	0.007438	0.003801	0.9999	2.588
D2	a_2	b_2	c_2	R-square	RMSE
	878.0	0.007241	0.004635	0.9999	2.717
D3	a_3	b_3	c_3	R-square	RMSE
	765.7	0.008437	0.004738	0.9999	2.206

In particular, the dynamic behavior of test-D2 is the harshest, and the degradation is severe after 300 h testing at maximal power. It is difficult to reach the maximal current value on the polarization curves, and then the limiting current value moves from 170 A to 100 A after 300 h testing due to the encountered degradation. The current and

voltage are resampled at a regular interval of half an hour. The assumption herein is that the degradation state in each time interval keeps the same. This is reasonable because the resampling time interval is short enough for the hundreds of hours' degradation.

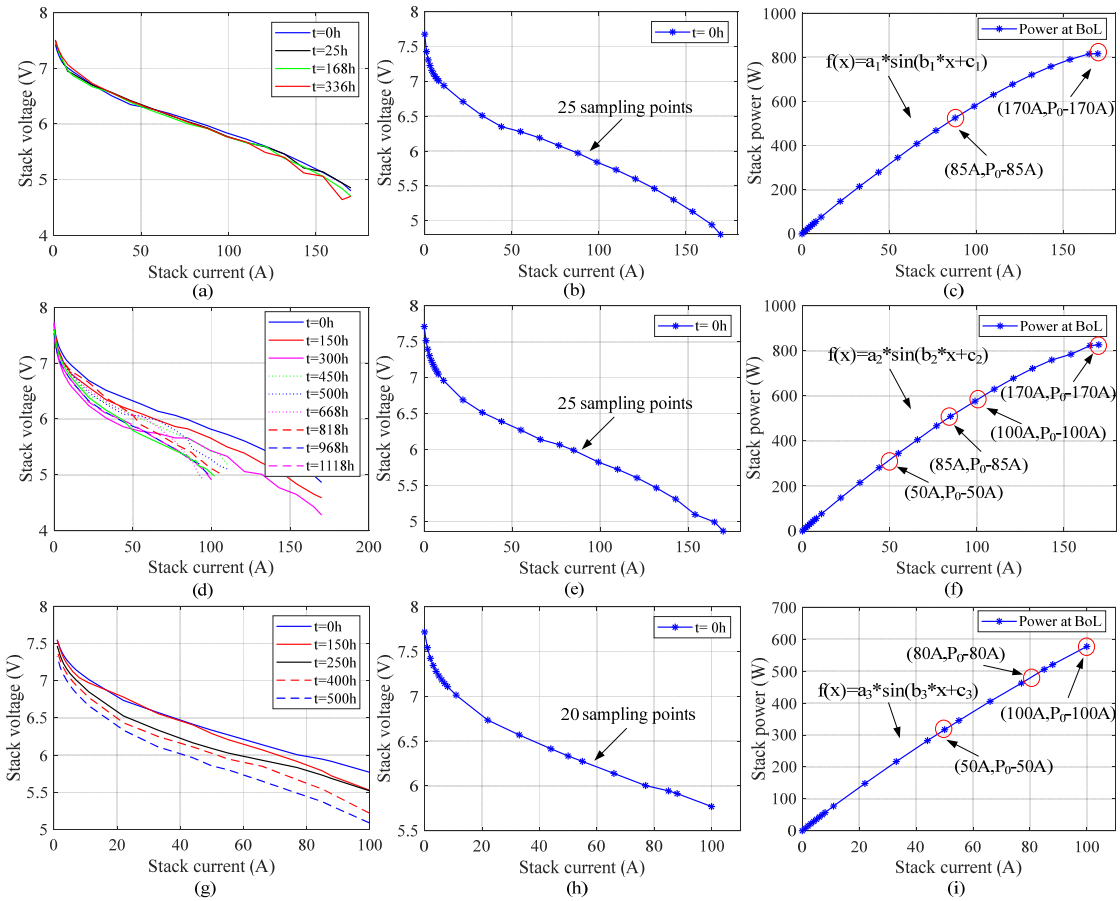


Fig. 3 - 2 The polarization curves and the BoL power curves of three long-term tests: (a) polarization curves during the whole life of test-D1, (b) polarization curve at the BoL of test-D1, (c) BoL power curve of test-D1, (d) polarization curves in the whole life of test-D2, (e) polarization curve at the BoL of test-D2, (f) BoL power curve of test-D2, (g) polarization curves in the whole life of test-D3, (h) polarization curve at the BoL of test-D3, (i) BoL power curve of test-D3.

The RPLR of the three tests is shown in Fig. 3 - 3. Results show that the RPLR in each test has a global monotone decreasing tendency and the current profile affects the degradation speed. A higher current level always leads to a higher degradation speed, e.g., in test-D2, $\triangle P_1$ during the first 300 h (170 A) is bigger than $\triangle P_2$ during the rest 700 h (50 A). An efficient HI is useful for the prognostic methods to predict the SoH and proposing an index to indicate the EoL time is also an important work for PEMFC systems. In practice, the U.S. DOE has pointed out that the definition of a proper index depends on the actual needs of users. It is worth mentioning that the meaningful definition of EoL is the time when the PEMFC cannot transfer the required power under safe conditions rather than the stack being out of use once it operates at a degradation

state. Due to the limitations of the experimental conditions, the duration of these three dynamic tests is all within 1000 h, they cannot meet the experimental environment of U.S. DOE. The last points in these three dynamic tests do not mean that they cannot work anymore or have reached their real EoL, and they just mean the time of end-of-test. Nevertheless, EoL could be regarded as one special case of end-of-test (power decay reaches certain defined levels). Without loss of generality, the end-of-test in this work is regarded as the time of EoL.

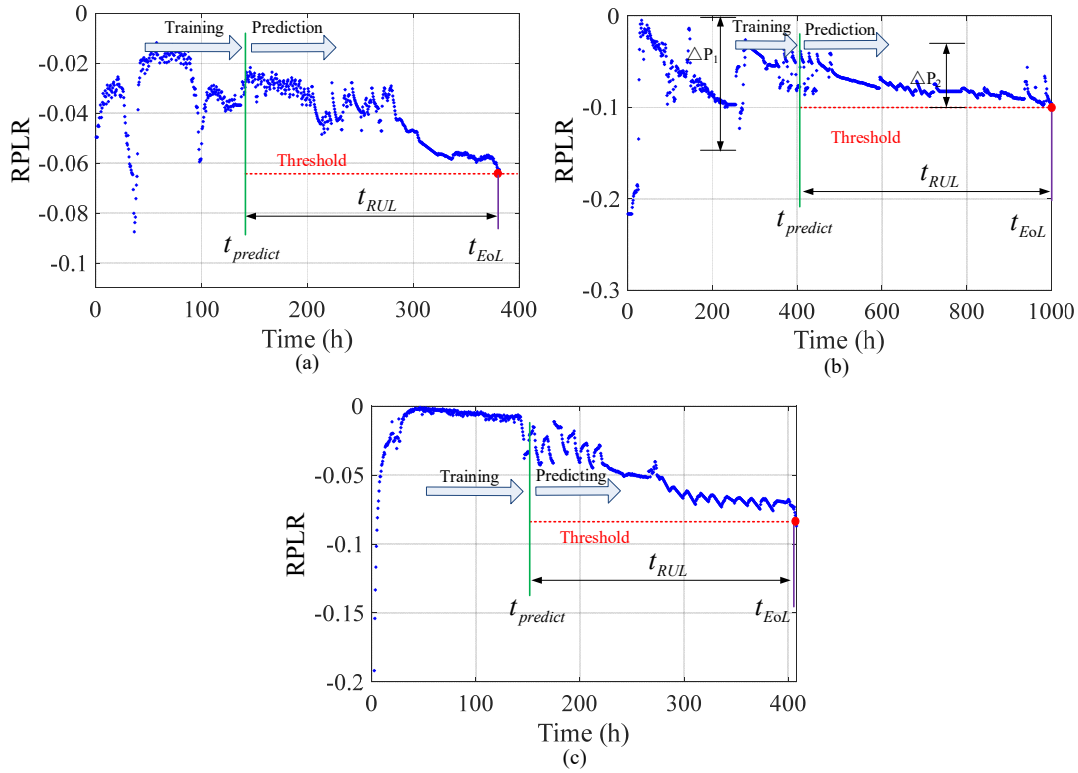


Fig. 3 - 3 The RPLR and RUL definition of 3 dynamic tests: (a) test-D1, (b) test-D2, (c) test-D3.

The first 50 h of data are abandoned due to the instability at startup. So, the training part is the data from 50 h to the time when the prediction starts ($t_{predict}$), and the prediction part is the data from $t_{predict}$ to the EoL time (t_{EoL}). The RUL (t_{RUL}) is usually represented by the time interval between $t_{predict}$ and t_{EoL}. In general, the PEMFC stack always has a satisfactory performance at the BoL and the $t_{predict}$ usually starts when the system has worked for a certain amount of hours. Besides, the prediction cannot be accurate at BoL because the historical data are inadequate for learning the degradation characteristics, especially for the data-driven methods. Based on the extracted HI of RPLR, the purpose of this work is to compare the prediction performance of DI-ESN and SI-ESN. So, in each test, the precision improvement of DI-ESN can be verified when these two structures (SI-ESN and DI-ESN) have the same EoL time. For comparison's purpose, the end-of-test in these three tests is regarded as the time of EoL. The end-of-test time for test-D1, test-D2, and test-D3 are 382 h, 1000 h, and 405 h

respectively.

3.3 Implementation process of DI-ESN

The operating parameters of test-D1, test-D2, and test-D3 are shown in Fig. 3 - 4, Fig. 3 - 5, and Fig. 3 - 6.

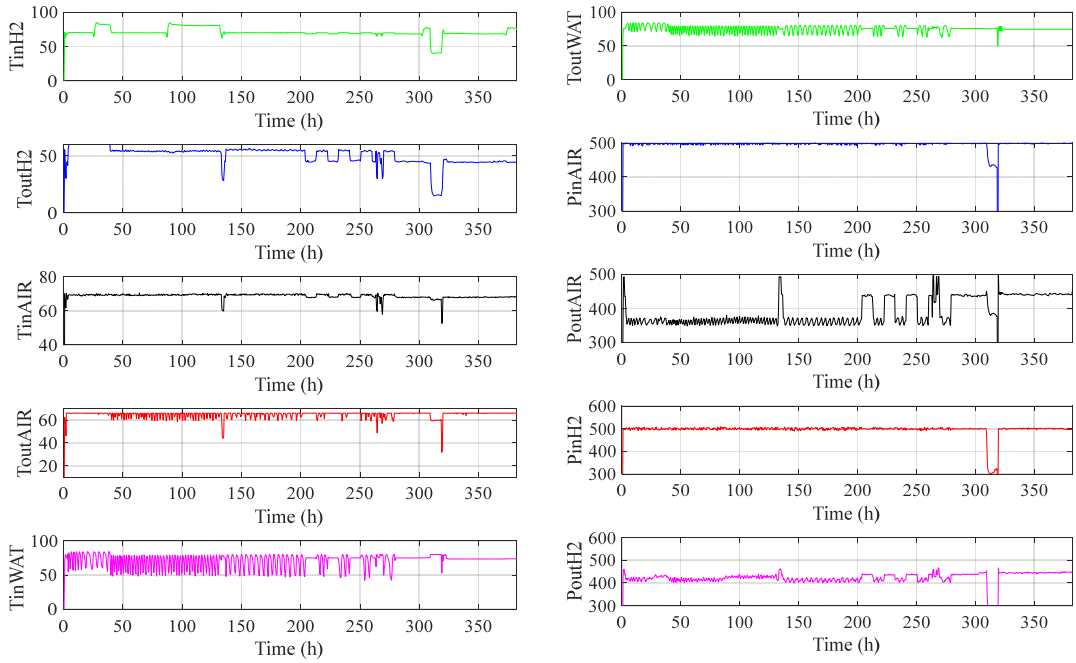


Fig. 3 - 4 Operating parameters of test-D1.

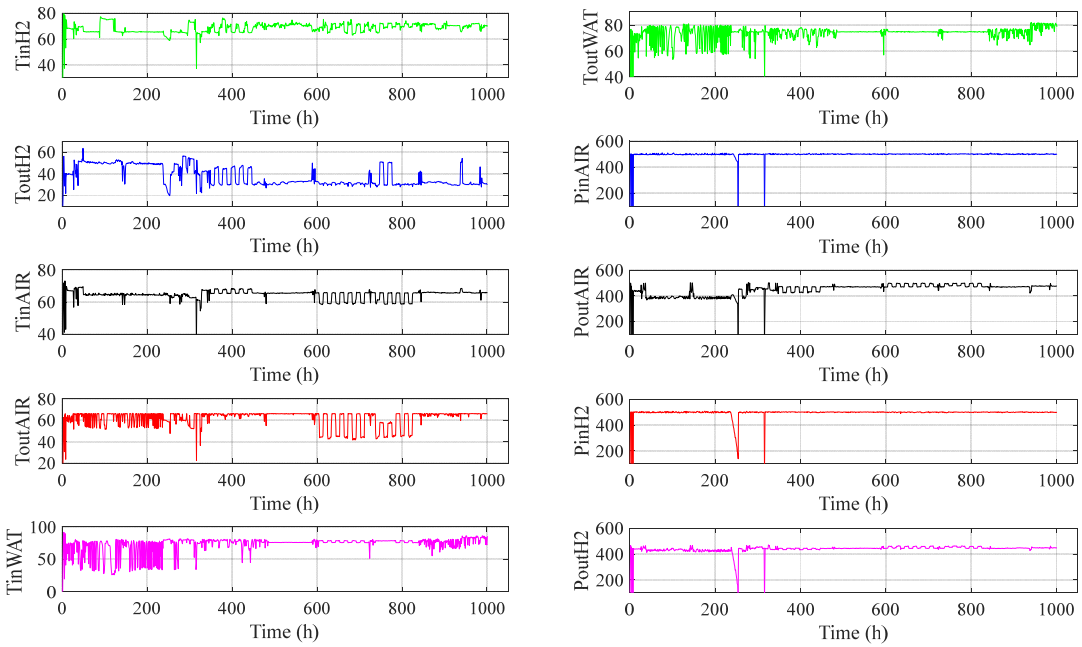


Fig. 3 - 5 Operating parameters of test-D2.

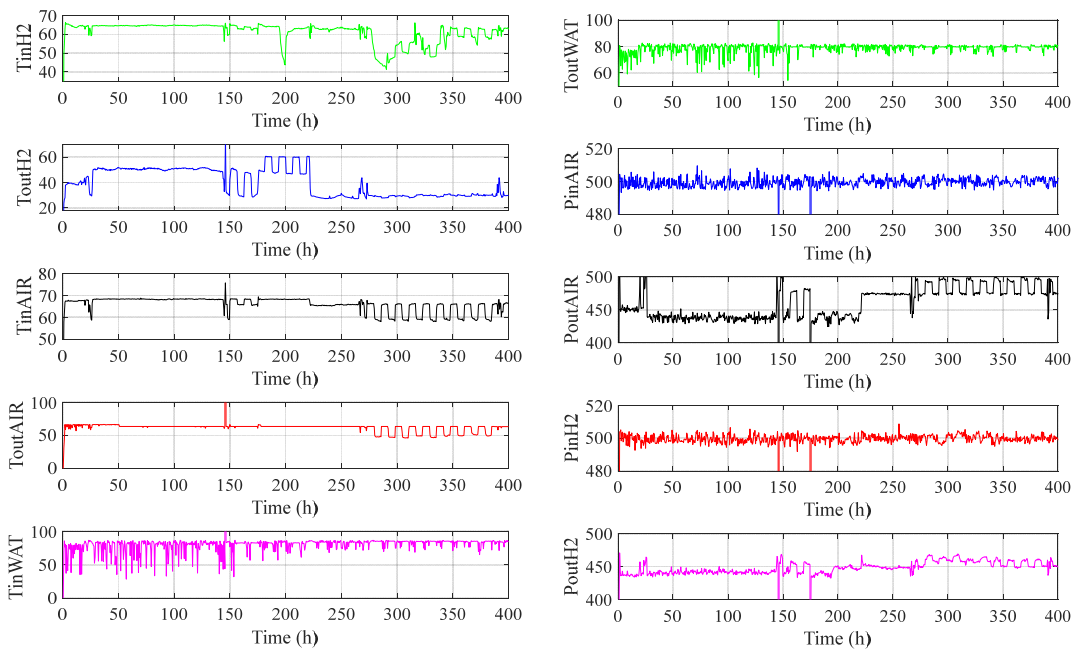


Fig. 3 - 6 Operating parameters of test-D3.

Once the RPLR is extracted, the SI-ESN and DI-ESN both with an iterative prediction structure are used to evaluate the RUL of the PEMFC systems [122]. It should be noted that the stack current is usually taken as a planning variable. This double-input long-term prediction pattern is more applicable for applications where the stack current is schedulable or programmable, such as the μ-CHP application where the load is related to the seasonal variations.

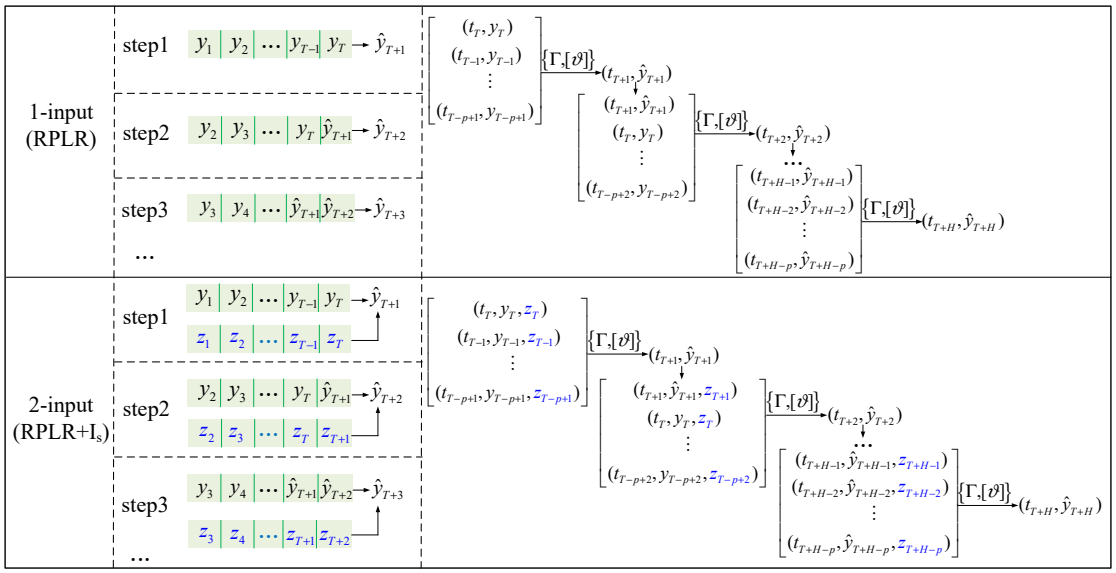


Fig. 3 - 7 The iterative process of SI-ESN and DI-ESN.

The iterative process of SI-ESN and DI-ESN is shown in Fig. 3 - 7. The influences of stack current are marked in blue. Where $\{\Gamma, \vartheta\}$ represents the prediction of one-step ahead ESN and $[\vartheta]$ represents the parameters in the ESN. And p is the number of

regressors, i.e., the quantity of previous data used for the prediction.

For SI-ESN (inputs and outputs are all voltages), the training part of RPLR can be expressed as $\{(t_1, y_1),\cdots,(t_i, y_i),\cdots,(t_T, y_T)\}$, where t_i is the sampling time, y_i is the RPLR at time t_i. The output matrix $\mathbf{W}_{out}$ of the reservoir is computed by the training part and its value stays the same after the training. With the one-step ahead prediction structure, the multi-step ahead prediction is realized and the RPLR in the prediction part can be expressed as $\{(t_{T+1}, \hat{y}_{T+1}),\cdots,(t_H, \hat{y}_H)\}$, where T is the time step start to predict and H is the number of predicted points.

For DI-ESN (inputs are voltages and currents, outputs are voltages), the historical RPLR $\{(t_1, y_1),\cdots,(t_i, y_i),\cdots,(t_T, y_T)\}$ and the corresponding stack current $\{(t_1, z_1),\cdots,(t_i, z_i),\cdots,(t_T, z_T)\}$ are regarded as the inputs, and the predicted RPLR $\{(t_{T+1}, \hat{y}_{T+1}),\cdots,(t_H, \hat{y}_H)\}$ is regarded as the output, where z_i represents the stack current at the time t_i. Also based on the one-step ahead prediction structure, the predicted RPLR, and the new scheduled stack current $\{(t_{T+1}, z_{T+1}),\cdots,(t_H, z_H)\}$ are used to be the inputs of DI-ESN in the prediction part.

All the simulations are implemented in the Matlab 2018a version, and the calculating time in the prediction part is less than 20 s. This execution time is rather short for the hundreds of hours of lifetime prediction. The target statement of RUL and its prediction process is presented in Fig. 3 - 8.

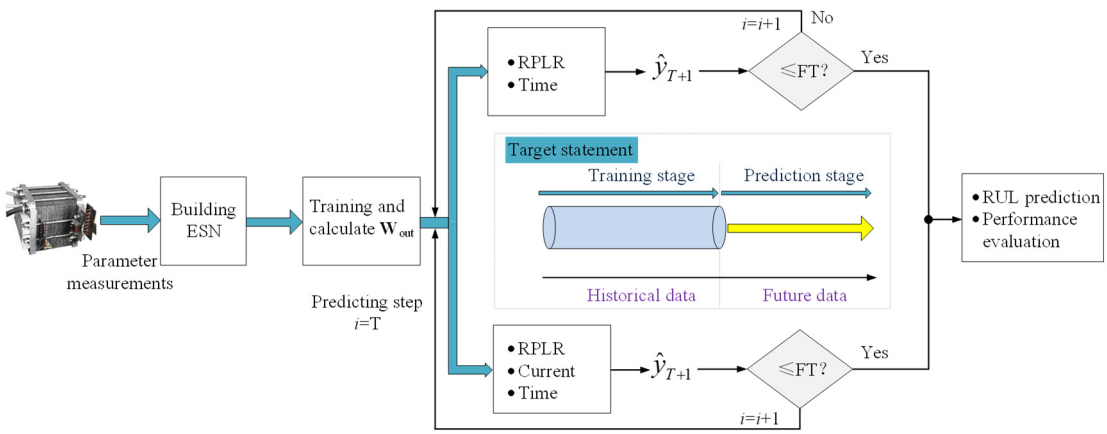


Fig. 3 - 8 Target statement of RUL and its prediction process.

3.4 Experimental results of dynamic tests

In dynamic tests, the results of SI-ESN are first presented. Afterward, the DI-ESN is implemented and compared with the SI-ESN. The satisfactory horizon (SH) is defined as the time when the prediction is located in an allowable error bound (= 5 %) around the true RUL to the EoL. The SH can be used to evaluate the prediction ability, and a longer SH means that the prediction accuracy meets the requirement at an early stage [123]. Besides, criteria of RMSE, MAPE, and $\%Er_{FT}$ are combined comprehensively to evaluate the prediction performances. In general, a small value of them means better

performance. In practice, underestimation is more meaningful and jeopardizes less the operation of the system because it could warn the users to adjust the operating conditions or replace the PEMFC in advance and then to prevent an incident. To analyze the influence of stack current (I_s) on the prediction effect, results of SI-ESN and DI-ESN with different training lengths are compared. All the predictions are implemented at about half of the whole data because premature prediction may have the problem of insufficient training data. The parameters (α, ρ, and β) of ESN in these three tests are determined based on the ANOVA sensibility analysis.

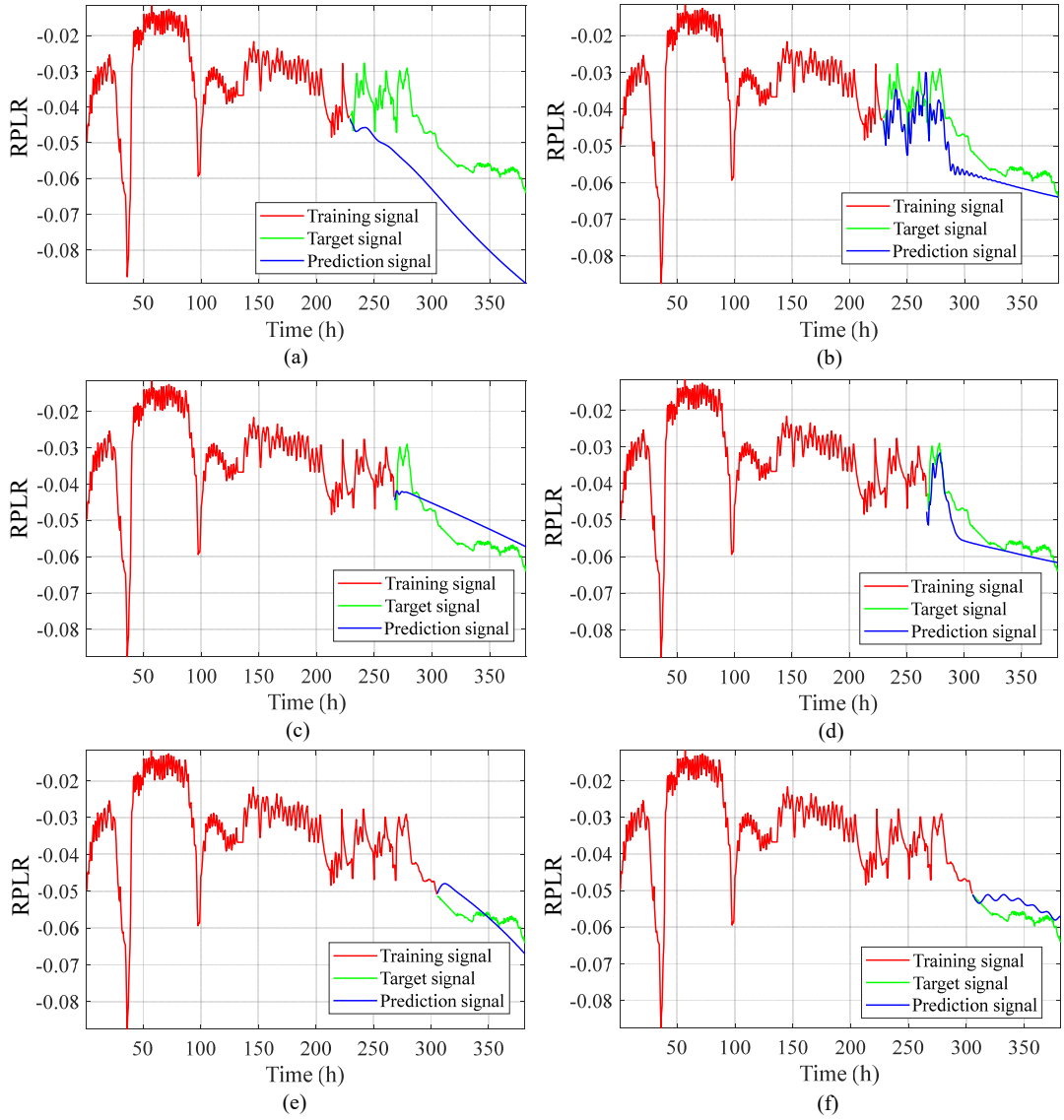


Fig. 3 - 9 The RUL prediction of test-D1: (a) SI-ESN with 60 % training, (b) DI-ESN with 60 % training, (c) SI-ESN with 70 % training, (d) DI-ESN with 70 % training, (e) SI-ESN with 80 % training, (f) DI-ESN with 80 % training.

In test-D1, the RUL predictions with 60 %, 70 %, and 80 % of the whole data (382 h) for training are presented in Fig. 3 - 9. They show that the DI-ESN (with stack current

I_s) has better performance than SI-ESN (without stack current I_s) especially during the load cycling period (200 h-300 h). On one hand, the DI-ESN could track the variation tendency of RPLR (ΔP) and on the other hand, the dynamic effectiveness of stack current is also superimposed to the RPLR prediction. These superposition effects are more pronounced in Fig. 3 - 9 (b) and Fig. 3 - 9 (d). In Tab. 3 - 2, the prediction results of test-D1 with different training lengths are given. With an increase of training length, the prediction error (%Er_{FT}, RMSE, and MAPE) has an overall downward trend both for the SI-ESN and the DI-ESN. The prediction results of DI-ESN are more accurate than those of SI-ESN as a whole. Nevertheless, two anomalies exist at 40 % and 50 % conditions from the view of RMSE and MAPE. Most probably, the prediction cannot be accurate if the training data are not enough. The two disturbances caused by the increase of $T_{in}H_2$ at 40 h and 100 h may also affect the prediction results. When considering the %Er_{FT}, the prediction RUL of DI-ESN is closer to the actual RUL both at 40 % and 50 % conditions. The RUL predictions of SI-ESN and DI-ESN in the whole lifespan of test-D1 are shown in Fig. 3 - 10. The SH of SI-ESN is 117 h and the SH of DI-ESN is 157 h.

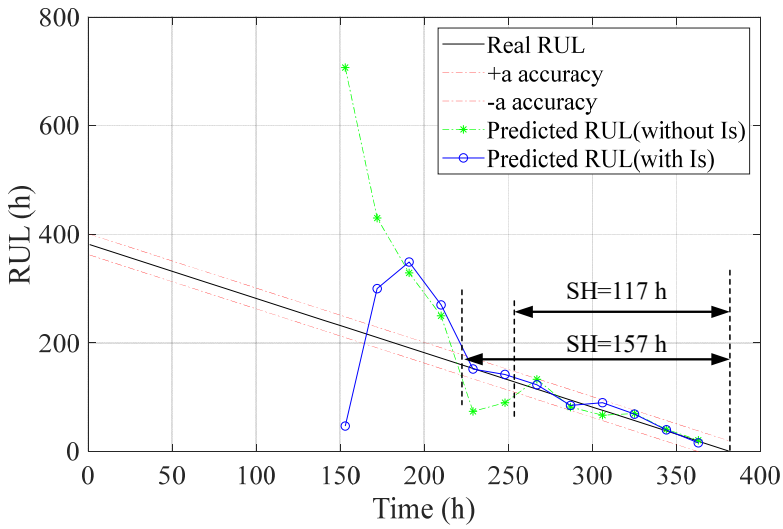


Fig. 3 - 10 The RUL results of SI-ESN and DI-ESN with 95 % probability bounds in test-D1.

Tab. 3 - 2 Prediction results comparison of test-D1.

ESN type	Training length (%)	Actual RUL (h)	Prediction RUL (h)	%Er_{FT} (%)	RMSE	MAPE
SI-ESN	40	229	707	-208.6	0.01869	0.57772
	50	191	329	-72.3	0.01067	0.24202
	60	153	74	51.7	0.01867	0.26439
	70	115	133	-15.7	0.00576	0.10435

	80	76	67	11.8	0.00429	0.06791
	90	38	41	-7.9	0.00360	0.05453
DI-ESN	40	229	47	79.4	0.02755	1.67360
	50	191	349	-82.7	0.04238	1.88170
	60	153	152	0.7	0.00650	0.11326
	70	115	123	-7.0	0.00411	0.06186
	80	76	94	-17.8	0.00373	0.06184
	90	38	40	-5.2	0.00231	0.03187

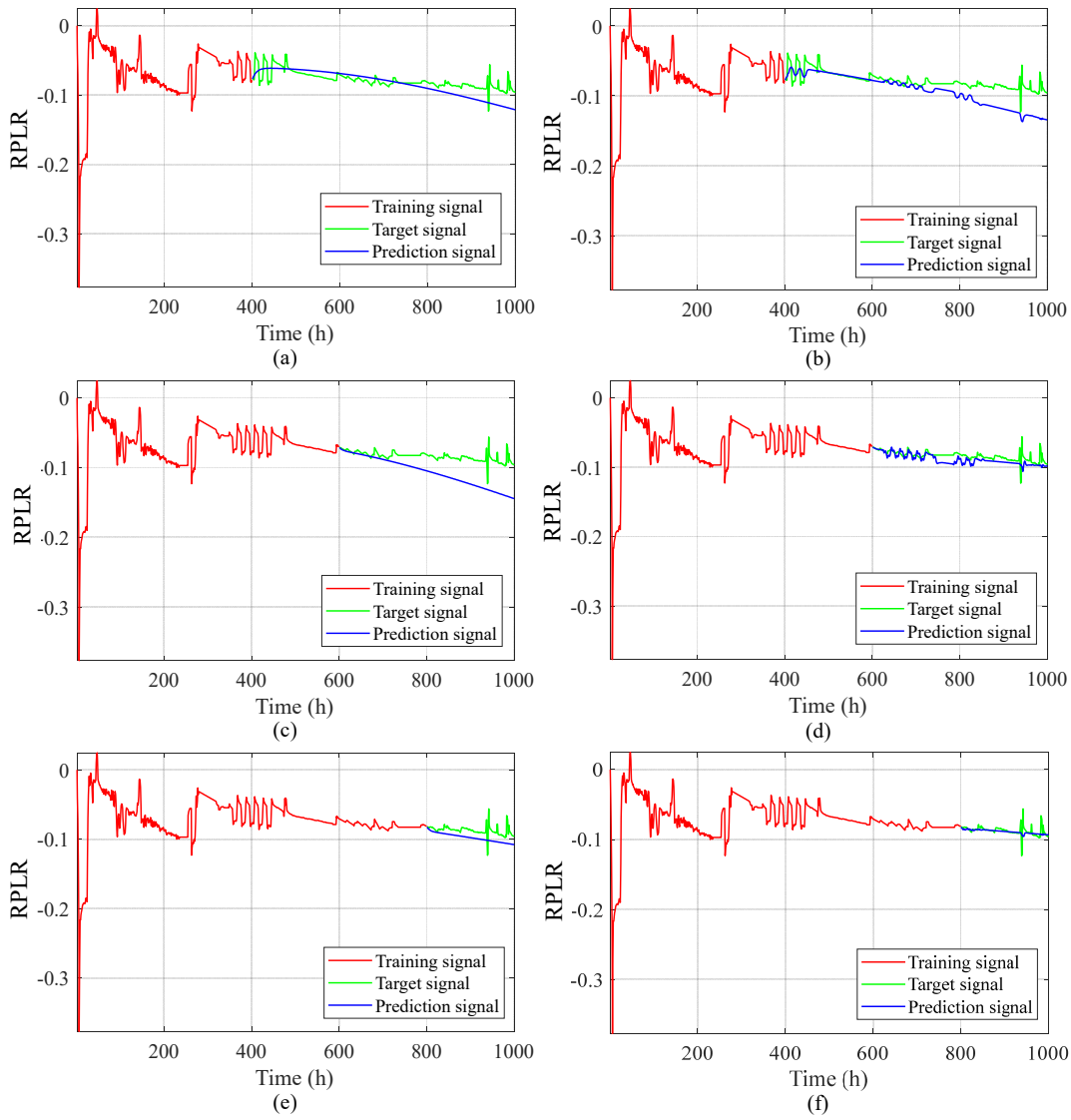


Fig. 3 - 11 The RUL prediction of test-D2: (a) SI-ESN with 40 % training, (b) DI-ESN with 40 % training, (c) SI-ESN with 60 % training, (d) DI-ESN with 60 % training, (e) SI-ESN with 80 % training, (f) DI-ESN with 80 % training.

Tab. 3 - 3 Prediction results comparison of test-D2.

ESN type	Training length (%)	Actual RUL (h)	Prediction RUL (h)	$\%Er_{FT}$ (%)	RMSE	MAPE
SI-ESN	40	600	468	22.0	0.01331	0.12065
	50	500	497	0.6	0.00879	0.09849
	60	400	174	56.5	0.02819	0.19113
	70	300	110	63.3	0.02422	0.18926
	80	200	122	39.0	0.01244	0.10711
	90	100	90	10.0	0.00885	0.05098
DI-ESN	40	600	390	35.0	0.02065	0.14497
	50	500	438	12.4	0.00788	0.07125
	60	400	342	14.5	0.00810	0.06666
	70	300	215	28.3	0.01197	0.10634
	80	200	138	9.0	0.00631	0.03575
	90	100	40	0.0	0.00440	0.02720

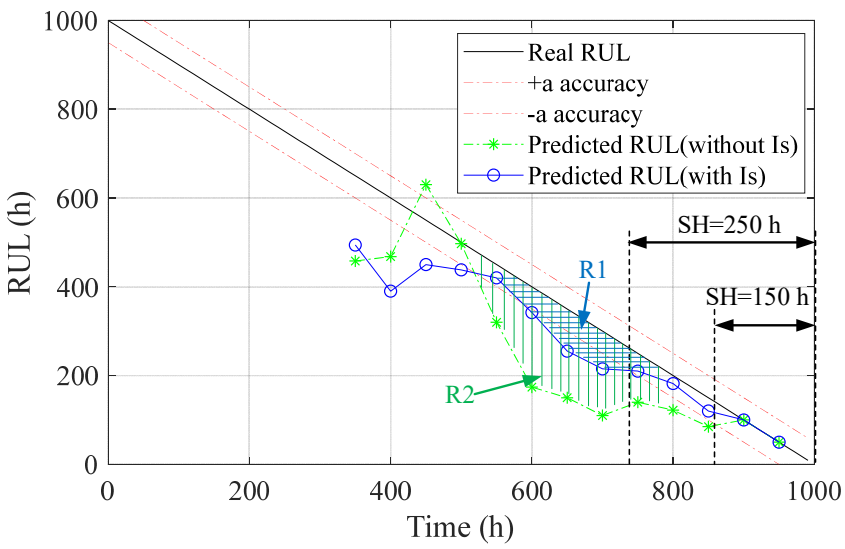


Fig. 3 - 12 The RUL results of SI-ESN and DI-ESN with 95 % probability bounds in test-D2.

In test-D2, the RUL predictions with 40 %, 60 %, and 80 % of the whole data (1000 h) for training are presented in Fig. 3 - 11. Similar to test-D1, the DI-ESN has better performance than SI-ESN and the superposition effects of stack current (400 h-500 h, and 600 h-800 h) are more pronounced in Fig. 3 - 11 (b) and Fig. 3 - 11 (d). The prediction results of test-D2 with different training lengths are shown in Tab. 3 - 3. The overall prediction performance of DI-ESN is better than SI-ESN. No matter for the SI-ESN or DI-ESN, the $\%Er_{FT}$, RMSE, and MAPE have a decreasing tendency with the

increase of training length. Similar to test-D1, two abnormal stages exist also at 40 % and 70 % conditions from the view of RMSE and MAPE. The limiting data and the disturbance caused by the $T_{in}H_2$ may affect the prediction result. This phenomenon is weakened when more data are added to the training part (from 60 %).

The RUL predictions of SI-ESN and DI-ESN in the whole lifespan of test-D2 are shown in Fig. 3 - 12. The SH of SI-ESN is 150 h and the SH of D-ESN is 250 h. The results of test-D2 also show that the system dynamic has a crucial effect on the prediction performance. In Fig. 3 - 12, an abnormal region exists both in SI-ESN (R1) and DI-ESN (R2) due to the repeat cycling condition from 500 h to 800 h (current repeating between 0 A and 50 A). This cycling condition would deteriorate the prediction results and increase the error. The abnormal region R1 of DI-ESN is smaller than the abnormal region R2 of SI-ESN. It means that the DI-ESN has stronger robustness under the dynamic load condition.

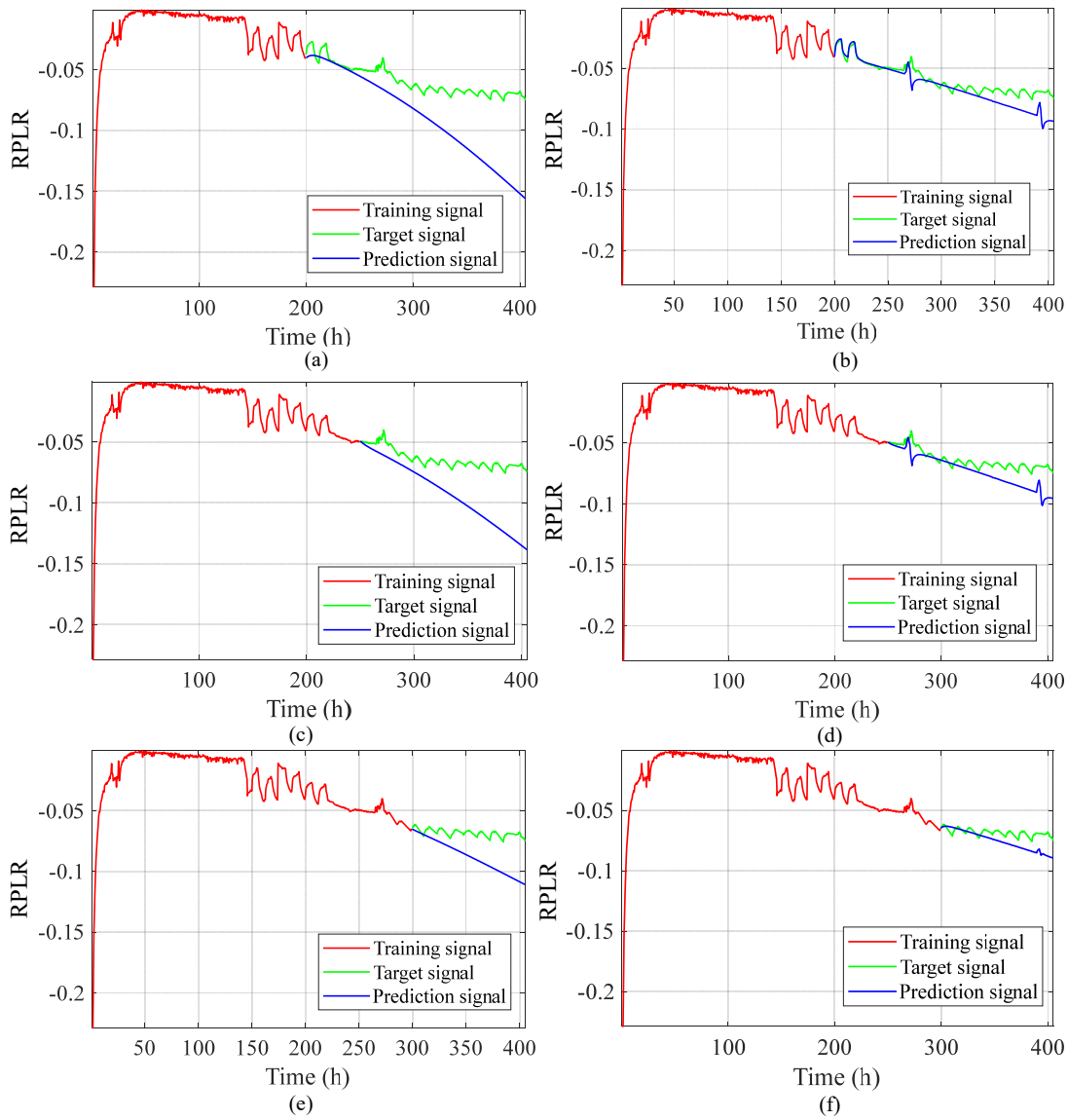


Fig. 3 - 13 The RUL prediction of test-D3: (a) SI-ESN with 50 % training, (b) DI-ESN with 50 % training, (c) SI-ESN with 65 % training, (d) DI-ESN with 65 % training, (e) SI-ESN with 75 % training, (f) DI-ESN with 75 % training.

Tab. 3 - 4 Prediction results comparison of test-D3.

ESN type	Training length (%)	Actual RUL (h)	Prediction RUL (h)	$\%Er_{FT}$ (%)	RMSE	MAPE
SI-ESN	30	283.5	116.5	58.9	0.13765	0.54349
	40	243	127	47.7	0.03882	0.27855
	50	202.5	99.5	50.9	0.03396	0.26779
	60	162	83	48.8	0.02254	0.20479
	70	121.5	75	38.7	0.01407	0.15470
	80	81	84	-3.7	0.00165	0.02045
DI-ESN	30	283.5	255	10.2	0.01100	0.26457
	40	243	183	24.7	0.00953	0.08614
	50	202.5	139	31.6	0.01199	0.11355
	60	162	114	29.6	0.00937	0.09378
	70	121.5	82	32.9	0.00976	0.11173
	80	81	83	-2.5	0.00182	0.02273

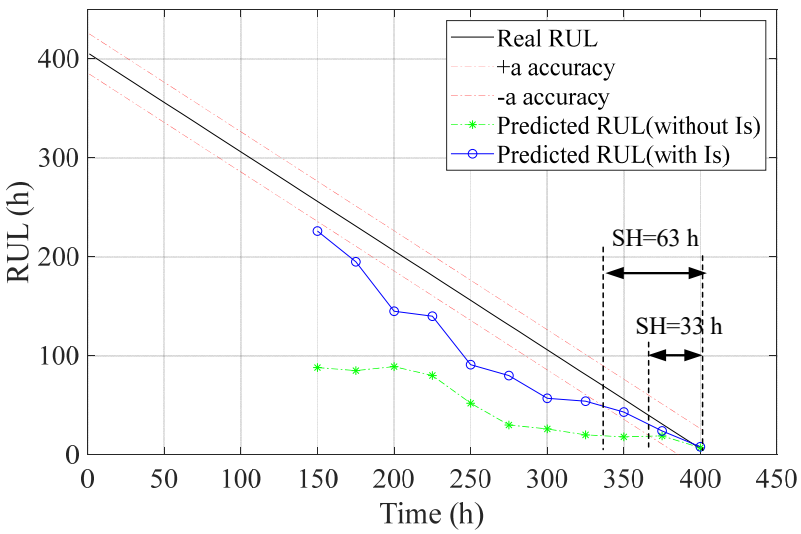


Fig. 3 - 14 The RUL results of SI-ESN and DI-ESN with 95% probability bounds in test-D3.

In test-D3, the RUL predictions with 50 %, 65 %, and 75 % of the whole data (405 h) for training are presented in Fig. 3 - 13. The superposition effects of stack current are more pronounced in Fig. 3 - 13 (b) and Fig. 3 - 13 (d). The prediction results of test-D3 with different training lengths are shown in Tab. 3 - 4. The RUL predictions of SI-ESN and DI-ESN in the whole lifespan of test-D3 are shown in Fig. 3 - 14. The results of

test-D3 are similar to those of test-D1 and test-D2. They also show that the DI-ESN structure performs better than SI-ESN. The SH of SI-ESN is 33 h and the SH of DI-ESN is 63 h. These three dynamic tests also show that the evaluation error is large at the BoL because of a lack of training data. This phenomenon is validated between 40 % - 60 % training length both in SI-ESN and DI-ESN structure. With the increasing training data, the magnitude of the error decreases as the prediction time gets closer to EoL.

3.5 Sensibility analysis of parameters

In industrial applications, parameter design is one of the bottlenecks of ESN's implementation and the classification of these parameters has been presented before. Determining these parameters reasonably is significant to improve the dynamic performance of the reservoir, and a detailed parameter manually setting guide is presented in [117].

In general, the number of input K and output L is determined by the requirement of the users. The input weight matrix $\mathbf{W}_{in}$, internal weight matrix $\mathbf{W}$ are generated randomly (e.g., Gaussian distributions) and regarded to be fixed once they are assigned. In fact, the dynamic characteristics of ESN have slight differences when using different generation methods, and these assigned weight matrixes are optimized in [124].

A larger value of reservoir neurons N is more conductive to find the linear combination between the input and output signals, and using a reservoir containing more than 10^4 neurons is a common phenomenon. The spectral radius ρ is the maximal absolute eigenvalue of the $\mathbf{W}$, and it is always less than 1 to ensure the ESP for the zero-input case. Enough evidence also proves that smaller than 1 is not the necessary condition of ESP, especially for the nonzero input. And the reservoir could also have better properties when extending the value of ρ. The leaking rate α indicates the dynamic performance of the reservoir, and the echo property becomes more important when the α becomes less important. In other words, a small value of α means that the value of the current step has a great influence on the next step. The parameters of N, ρ, and α are analyzed in [100].

Nevertheless, the inputs of ESN are the RPLR and scheduled stack current (I_s) in the dynamic PEMFC lifespan prediction task, and the ρ ($\rho \in (0,1)$) is better to be expended to a larger range ($\rho \in (0,1.5)$). The training part aims to calculate the output weight matrix $\mathbf{W}_{out}$, and the regularization parameter β has an important direct effect on the calculation result. Compared with N, analyzing β has a greater practical significance for the prediction results. So, the leaking rate α ($\alpha \in (0,1)$), spectral radius ρ ($\rho \in (0,1.5)$), and regularization parameter β ($\beta \in (8 \times 10^{-3}, 8 \times 10^{-1})$ are selected as the key parameters in the task of sensibility analysis.

The objectives of parameter analysis are the quantification of the impact of the parameters on the model output and studying the interactions among the parameters. ANOVA is a collection of statistical models and procedures to compare the effects of different variables [125]. The ANOVA-representation of integrable function $f(x)$ in the q-dimensional space (I^q) is shown as

$$\begin{aligned} f(x) &= f_0 + \sum_{s=1}^{n} \sum_{1 \le i_1 < \cdots < i_s \le q}^{n} f_{i_1 \cdots i_s}(x_{i_1}, \cdots, x_{i_s}) \\ &= f_0 + \sum_i f_i(x_i) + \sum_{1 \le i < j \le q} f_{ij}(x_i, x_j) + \cdots + f_{12 \cdots q}(x_1, \cdots, x_q) \end{aligned} \tag{3-5}$$

Where x are the input variables, $f(x)$ are the output variables, and $1 \le i < \cdots < j \le q$. The premise of Eq. (3-5) is shown as

$$\int_0^1 f_{i_1 \cdots i_s}(x_{i_1}, \cdots, x_{i_s}) dx_k = 0 \quad \text{for } k = i_1, \cdots, i_s \tag{3-6}$$

Indeed, the members in Eq. (3-6) are orthogonal and the form of them can be expressed as

$$\begin{cases} \int f(x) dx = f_0 \\ \int f(x) \prod_{k \ne i} dx_k = f_0 + f_i(x_i) \\ \int f(x) \prod_{k \ne i,j} dx_k = f_0 + f_i(x_i) + f_j(x_j) + f_{i,j}(x_i, x_j) \end{cases} \tag{3-7}$$

A detailed description of ANOVA is given in [126]. The RMSE is used to estimate the prediction performance. A smaller RMSE indicates that the prediction has a preferable performance. The process of parameters sensibility analysis is as follows

Step1: The data of test-D1 and the SI-ESN structure are chosen as the specimen of parameter analysis, and 60 % are used for training, and 40 % are used for prediction.

Step2: Setting the key parameters (α, ρ, β) by the trial-and-error method, then the ANOVA analysis is used to compute the effects and the contributions of these parameters on the prediction results.

Step3: Based on the results of ANOVA analysis, the parameters of different training lengths of test-D1 and the parameters for the other two tests (test-D2 and test-D3) with different training structures can also be determined.

The parameters are divided into three different levels in step2: leaking rate α (0.3 for low level, 0.6 for middle level, and 0.9 for high level), spectral radius ρ (0.5 for low level, 1.0 for middle level, and 1.5 for high level), and regularization parameter β (8×10^{-3} for low level, 8×10^{-2} for middle level, and 8×10^{-1} for high level). Each experience is the result of a single simulation of a specific combination of parameters.

There are thus 3^3=27 experiment results. In test-D1, the results of different combinations are shown in Tab. 3 - 5 and the influence of different parameters is shown in

Tab. 3 - 6.

Tab. 3 - 5 Prediction results of different combinations.

Parameters	RMSE	Parameters	RMSE	Parameters	RMSE
$\alpha_{low}+\rho_{low}+\beta_{low}$	0.00688	$\alpha_{mid}+\rho_{low}+\beta_{low}$	0.00899	$\alpha_{hig}+\rho_{low}+\beta_{low}$	0.01000
$\alpha_{low}+\rho_{low}+\beta_{mid}$	0.01001	$\alpha_{mid}+\rho_{low}+\beta_{mid}$	0.01098	$\alpha_{hig}+\rho_{low}+\beta_{mid}$	0.01053
$\alpha_{low}+\rho_{low}+\beta_{hig}$	0.00747	$\alpha_{mid}+\rho_{low}+\beta_{hig}$	0.00770	$\alpha_{hig}+\rho_{low}+\beta_{hig}$	0.00829
$\alpha_{low}+\rho_{mid}+\beta_{low}$	0.00518	$\alpha_{mid}+\rho_{mid}+\beta_{low}$	0.00525	$\alpha_{hig}+\rho_{mid}+\beta_{low}$	0.00559
$\alpha_{low}+\rho_{mid}+\beta_{mid}$	0.00691	$\alpha_{mid}+\rho_{mid}+\beta_{mid}$	0.00539	$\alpha_{hig}+\rho_{mid}+\beta_{mid}$	**0.00498**
$\alpha_{low}+\rho_{mid}+\beta_{hig}$	0.00616	$\alpha_{mid}+\rho_{mid}+\beta_{hig}$	0.00749	$\alpha_{hig}+\rho_{mid}+\beta_{hig}$	0.00623
$\alpha_{low}+\rho_{hig}+\beta_{low}$	0.01782	$\alpha_{mid}+\rho_{hig}+\beta_{low}$	0.04008	$\alpha_{hig}+\rho_{hig}+\beta_{low}$	0.06222
$\alpha_{low}+\rho_{hig}+\beta_{mid}$	0.01017	$\alpha_{mid}+\rho_{hig}+\beta_{mid}$	0.00950	$\alpha_{hig}+\rho_{hig}+\beta_{mid}$	0.00835
$\alpha_{low}+\rho_{hig}+\beta_{hig}$	0.01136	$\alpha_{mid}+\rho_{hig}+\beta_{hig}$	0.00509	$\alpha_{hig}+\rho_{hig}+\beta_{hig}$	0.00752

Tab. 3 - 6 Influence of different parameters.

Parameters	Influence (%)
Leaking rate α	2.59
Spectral radius ρ	22.28
Regularization ρ	15.54
Interaction $\alpha*\rho$	4.15
Interaction $\alpha*\beta$	7.51
Interaction $\rho*\beta$	35.75
Error	12.18
Total	100

Results represent that setting the leaking rate α at a high level, the spectral radius ρ at a middle level, and the regularization parameter at a middle level could lead to a minimum RMSE (**0.00498**). They also show that the spectral radius ρ is the most influential among these three key factors with a contribution equal to 22.28 % of the total variance. The effect of the regularization parameter β is the second most significant factor with a contribution equal to 15.54 %. Any two of the three parameters have different interaction values. The most influential interaction parameters for this task are the spectral radius ρ and regularization parameter β with a contribution of 35.75 %. Regarding the best combination in Tab. 3 - 5 as a benchmark could improve the adjusting efficiency of these three parameters in test-D2 and test-D3. The key parameters in test-D2 and test-D3 are shown in Tab. 3 - 7. Thanks to the ANOVA

analysis, the influences of different parameters and their interactions are evaluated. And the trial-and-error research space of the parameters is reduced at the same time.

Tab. 3 - 7 Parameter levels and the typical values in different tests.

	Leaking rate α	Spectral radius ρ	Regularization parameter β
Low	0 - 0.3	0 - 0.5	$< 8\times10^{-3}$
Middle	0.3 - 0.6	0.5 - 1.0	8×10^{-3} - 8×10^{-2}
High	0.6 - 0.9	1.0 - 1.5	8×10^{-2} - 8×10^{-1}
Test-D2	0.9	0.7	8×10^{-2}
Test-D3	0.9	0.6	8×10^{-2}

3.6 Chapter summary

Results show that the RPLR is an efficient dynamic HI, and the ESN with double inputs could imitate the degradation properties more sufficiently and improve the prediction precision under the dynamic operating conditions. In general, the value of $\%Er_{FT}$, RMSE, and MAPE of DI-ESN are smaller than those of SI-ESN. The satisfactory horizon (SH) of DI-ESN is longer than that of SI-ESN in each dynamic test, and thus more time is given to the users to anticipate maintenance actions before EoL. In test-D1, the SH of SI-ESN is 117 h and the SH of DI-ESN is 157 h. In test-D2, the SH of SI-ESN is 150 h and the SH of DI-ESN is 250 h. In test-D3, the SH of SI-ESN is 33 h and the SH of DI-ESN is 63 h. In general, the DI-ESN has a relatively longer SH which means a better prediction ability. Besides, the effects and the contributions of the various ESN parameters and their combinations (leaking rate α, spectral radius ρ, regularization parameter β) are analyzed based on the ANOVA and multiple comparisons.

Chapter 4. Multi-timescale predictions of ensemble echo state network

4.1 Brief introduction

Low prediction accuracy is one of the weaknesses for current RUL prediction methods on the long-term scale, especially when the PEMFC system works under variable load profiles. To overcome the drawbacks of static HIs (e.g., voltage, and power), a novel dynamic HI of RPLR has been presented before to indicate the degradation state under the dynamic operating conditions.

Some scholars have made great efforts to enhance forecasting performance [81] - [83]. The original voltage sequence of the PEMFC systems is decomposed into multiple sub-waveforms in [81], and the group method of data handling (GMDH) is proposed to realize the short-term (maximum 5-step ahead, i.e., 5 h) degradation prediction. Similar to [81], the voltage waveforms are decomposed in [83] and [82], and the extreme learning machine (ELM) is applied to build the degradation model of each sub-waveform (1-step ahead, i.e., about 1 hour). In this prediction structure, the measured voltage of the last step is used to predict the new value of the next step. Simulation results demonstrate that the improved structures of GMDH and ELM are more accurate than the conventional methods. Nevertheless, the degradation predictions in [81] and [82] are at the static and quasi-dynamic working status, and all the predictions in [81] - [83] are realized on a short-term scale.

The approach of discrete wavelet transform and ensemble echo state network (DWT-EESN) is proposed to improve the precision of prediction on the long-term scale in this chapter. Inspired by the concept of "divide and rule", the RPLR which contains different timescale features is decomposed by the DWT. Decomposing dynamic HI has more practical significance than decomposing static HI because different load profiles under the dynamic working conditions would result in multi-timescale features. Then several individual ESNs with different dynamic behaviors are used to handle the features in different timescales. The ESN with a higher dynamic reservoir is used to deal with the fast-dynamic part of the feature (short timescale). On the contrary, the ESN with a lower dynamic reservoir is used to deal with the slow-dynamic part of the features (long timescale). Finally, the individual long-term prediction results in different timescales are lumped together to constitute the final results. Three μ-CHP experiments with different dynamic profiles (i.e., test-D1, test-D2, and test-D3) are carried out for validation purposes.

4.2 Discrete wavelet transform

The Fourier analysis has been widely used in the time-frequency transformation domain, and it could break up the original signal into sine waves with different

frequencies. Nevertheless, the time information would be lost during the decomposition. As an improvement type of the classical Fourier analysis, the short-time Fourier transform (STFT) can analyze the original signal in a different time and frequency domains. In STFT, the original signal at the time domain is divided into the same length segments. In other words, the windows of the Fourier transform are all at the same length, and it also has some weaknesses to deal with the time-varying non-stationary signal. The wavelet transform (WT) can analyze the local features (i.e., RPLR in this task) through dilation and translation, and the window length is varying [127]. In WT, the long-time intervals are adopted for the low-frequency parts of the original signal, and the short-time intervals are utilized for the high-frequency parts. Therefore, the WT has more advantages in analyzing the nonstationary signals, and it has been used in the diagnostic area of the PEMFC system [128]. The continuous WT is represented as

$$C(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi^{*}(\frac{t-b}{a})dt \tag{4-1}$$

Where *f(t)* is the original signal, $\Psi(t)$ is the mother wavelet, $\Psi^{*}(t)$ is the complex conjugate of $\Psi(t)$, a $(a \in \mathbb{R}^{+})$ and b $(b \in \mathbb{R})$ are the wavelet scale and shift parameters. The Daubechies (db3) wavelet is used here because it can redistribute the signal energy and compress most of the energy into the approximation component. A smaller a means a compressed wavelet and a larger a means a dilated wavelet. The scale factor of 2^{p} and shift factor of $2^{p}q$ are usually used to discretize the wavelet parameters, and the discrete WT is defined as

$$D(p,q)=2^{-p/2}\int_{-\infty}^{+\infty} f(t)\psi^{*}(2^{-p}t-q)dt \tag{4-2}$$

Where p and q are integers ($p \in \mathbb{Z},\ q \in \mathbb{Z}$). In practice, the DWT of the original signal can be calculated by the Mallat method [129] as follows

$$\begin{cases} S_{2^{p}}f(t)=\sum_{q} h_{q}S_{2^{p-1}}f(t-2^{p-1}q) \\ W_{2^{p}}f(t)=\sum_{q} g_{q}S_{2^{p-1}}f(t-2^{p-1}q) \end{cases} \tag{4-3}$$

Where $S_{2^{p}}$ is the smoothing operator, and $S_{0}f(t)=f(t)$, $W_{2^{p}}f(t)$ is the DWT of original signal *f(t)*, h_{q} and g_{q} are the coefficients of the orthogonal filter bank

$$\begin{cases} H_{low}(w)=\sum_{q} h_{q}e^{-iqw} \\ G_{\text{high}}(w)=\sum_{q} g_{q}e^{-iqw} \end{cases} \tag{4-4}$$

Then, signal *f(t)* could be represented as

$$f(t)=\sum_{q} cA_{0}\psi_{p,q}(t)=\sum_{q} cA_{1}\psi_{p-1,q}(t)+\sum_{q} cD_{1}\psi_{p-1,q}(t)=A_{1}(t)+D_{1}(t) \tag{4-5}$$

Based on the multiresolution signal decomposition technique, the signal in each

decomposition layer j ($0 \leqslant j \leqslant J$) could be decomposed into the approximation component (A_{j+1}) and the detail component (D_{j+1}). The A_{j+1} represents the low frequency or large-scale content of A_j, and the D_{j+1} represents the high frequency or small-scale content of A_j. Then, the $f(t)$ can be decomposed into J layers

$$f(t) = A_1(t) + D_1(t) = A_2(t) + D_2(t) + D_1(t) = A_j(t) + \sum_{j=1}^{J} D_j(t) \tag{4-6}$$

Essentially, the lifespan prediction of the PEMFC system can be translated into dealing with the nonstationary time series. The decomposition process of DWT is useful to handle the dynamic HI which has multi-timescale characteristics.

4.3 Implementation process of DWT-EESN

The in-situ electrical variables of voltage and current are measured by the sensors. After the data acquisition, the moving average filtering (window size is 31) method is used to remove the peaks and noises. Then the electrical variables (voltage and current) are resampled at the same time interval (0.5 h). The degradation timescale of the PEMFC system is in hundreds or thousands of hours. Therefore, regarding the aging state during each time interval as a constant value is a reasonable assumption. After the dynamic HI extraction, the time series of RPLR can be divided into the training part and the prediction part. The implementation process of ESN is shown in Fig. 4 - 1.

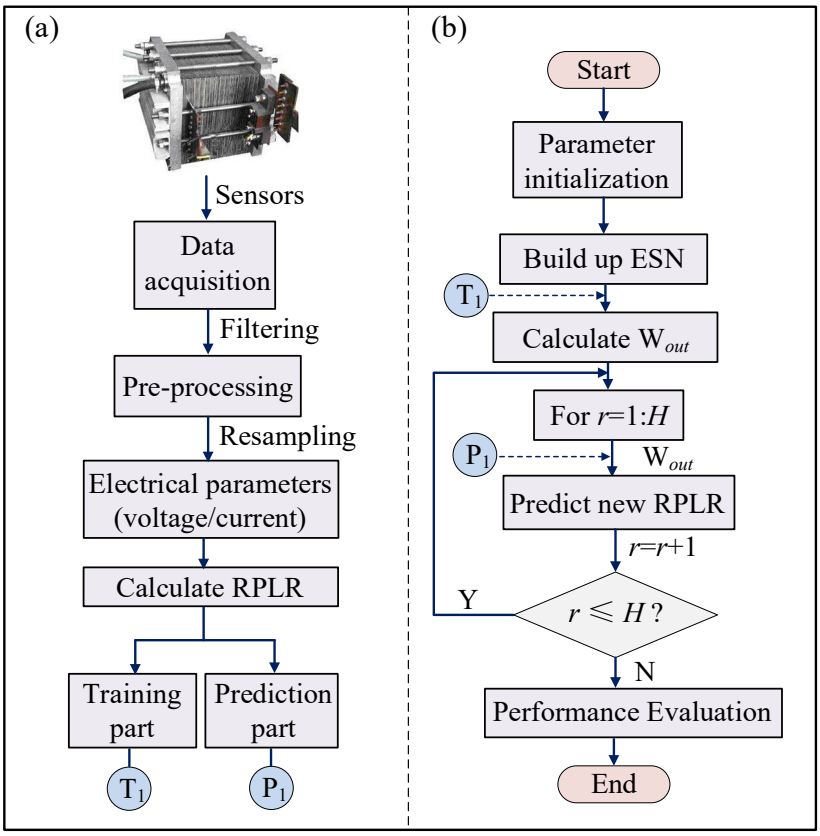


Fig. 4 - 1 Data processing and implementation process of the ESN: (a) data processing, (b) implementation process of ESN.

In practice, a single ESN is difficult to deal with the multi-timescale features at the same time. One effective thinking is decomposing the multi-timescale input signals at different levels and then composing the output features together. Inspired by this thinking, the DWT-EESN is proposed to deal with the multi-timescale features of the dynamic HI. The flowchart of DWT-EESN is described in Fig. 4 - 2.

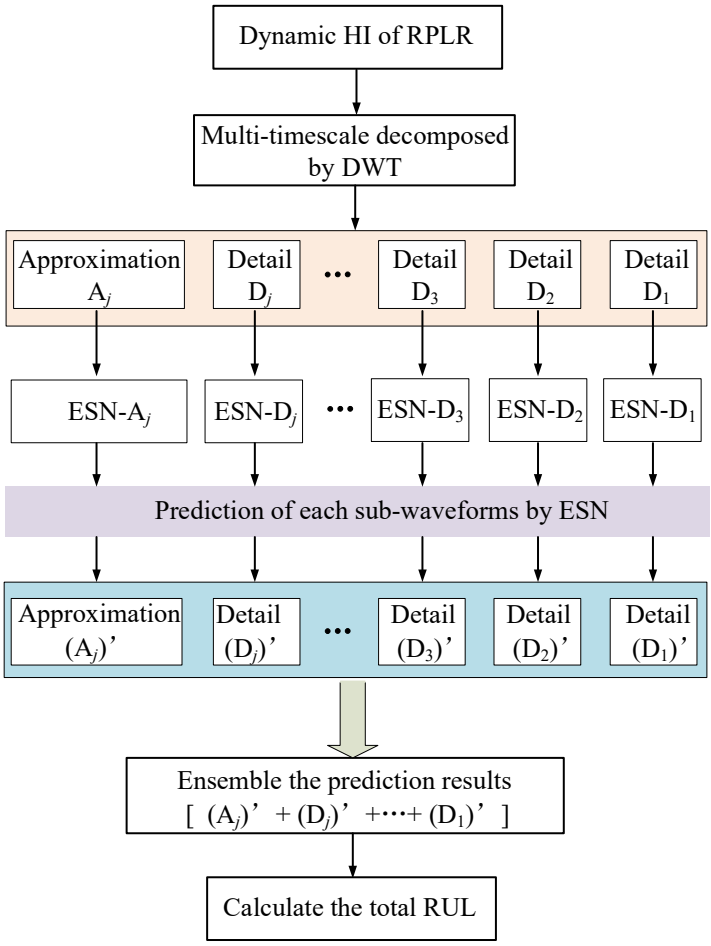


Fig. 4 - 2 The flowchart of DWT-EESN.

The DWT is used to decompose the HI of RPLR into sub-waveforms (A_j, D_j ,…, D_1) with different timescales. Afterward, several independent ESNs (ESN-A_j, ESN-D_j ,…, ESN-D_1) with different dynamic parameters are utilized to predict the sub-waveforms. In general, the dynamic state of the reservoir is increasing from A_j to D_1. The prediction results are represented as $(A_j)'$, $(D_j)'$,…, $(D_1)'$, and they are lumped together to calculate the total RUL. Before predicting the sub-waveforms by independent ESNs, the reconstructed signal (A_j, D_j ,…, D_1) are normalized to [0, 1].

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{4-7}$$

Where x' is the normalized data point, x is the original data point, $x_{\max}$ and $x_{\min}$ are the maximum and minimum data points separately in the whole dataset.

After the prediction, the results are inversely normalized to their original scales which can be represented by $(A_j)'$, $(D_j)'$,…, $(D_1)'$. At last, the individual results are summed at each time point to get the final results. The same normalization of sub-waveforms makes the dealing processes of different ESNs the same. The sum of inverse normalization results keeps the proportion of original sub-waveforms at different frequency domains.

In the current offline training process, the grid-search method is used to optimize two key parameters of ESN ($\alpha \in (0.1, 0.9)$, $\rho \in (0.1, 1.5)$). The parameter optimization process of grid-search is about 45 min, and the execution time of ESN by the best parameter combination is about 20 s. Besides, the proposed method of DWT-EESN has the potential to be used online which is also our next-step research focus. The parameter optimization can be implemented by offline optimization algorithms such as the grid research algorithm in the offline phase. Based on DWT, the RPLR of three dynamic tests can be decomposed into different sub-waveforms. Then the independent ESNs are used to predict the sub-waveforms in the long term separately. To assess the prediction performance of DWT-EESN, the criteria of RMSE and MAPE are used together in this task.

4.4 Experimental results under three dynamic tests

4.4.1 Long-term prediction of single ESN

To analyze the performance of DWT-EESN, single ESN's RUL predictions without the decomposition process are first performed as a benchmark. The inputs are the previous RPLR, and the outputs are the future RPLR. The prediction curve of dynamic tests with a single ESN is presented from Fig. 4 - 3 to Fig. 4 - 5, and the quantified results are shown in Tab. 4 - 1. All the predictions of RPLR are smooth curves, and the degradation tendency of the PEMFC system can also be obtained. It is worth noting that the operating parameters and environmental factors would have significant effects on the prediction performance. In the case of 90 % training length of test-D1, the abnormal points are probably caused by the outlet temperature of hydrogen ($T_{out}H_2$), and the inlet temperature of water (T_{in}Wat). In the case of 50 % training length of test-D2, the abnormal points are probably caused by the outlet temperature of hydrogen ($T_{out}H_2$), and the inlet temperature of air (T_{in}Air). In the case of 60 % training length of test-D3, the abnormal points are probably caused by the outlet temperature of hydrogen ($T_{out}H_2$), and the inlet temperature of air (T_{in}Air). Results show that most of the abnormal points are probably caused by the temperatures, and it means that temperature control plays an important role in the PEMFC system. Except for the influence of disturbance, in each test, the prediction error of RPLR has a declining trend with the increment of training length because the $\mathbf{W}_{out}$ of the reservoir would be more accurate with sufficient input samples.

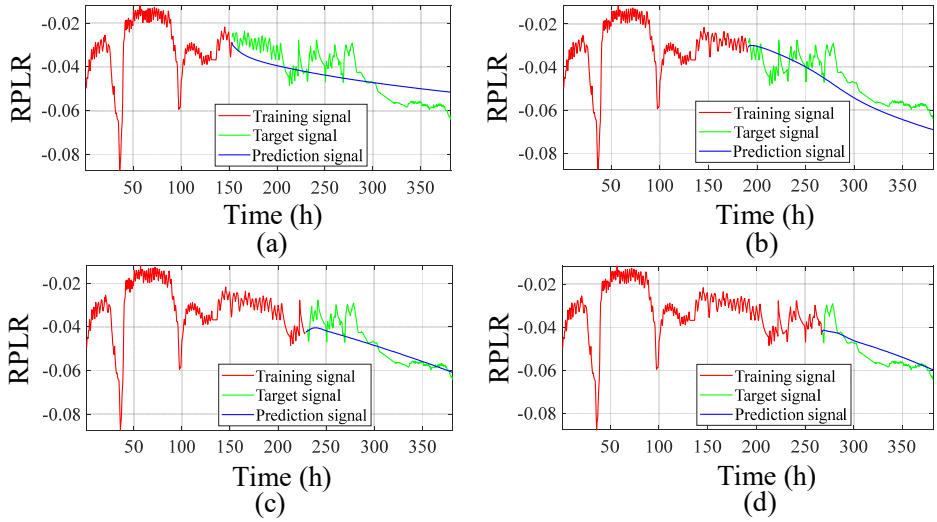


Fig. 4 - 3 The RPLR prediction with single ESN in test-D1: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

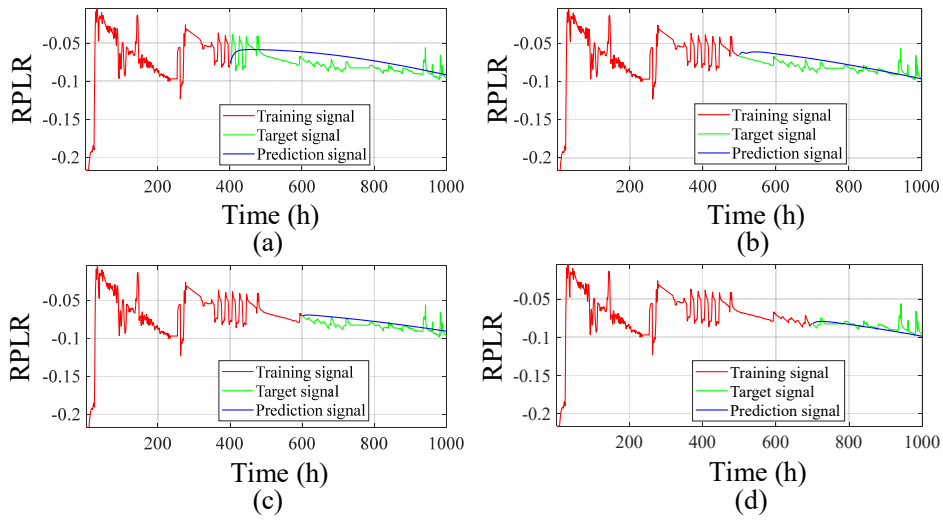


Fig. 4 - 4 The RPLR prediction with single ESN in test-D2: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

Tab. 4 - 1 Prediction results of single ESN.

Tests	Training length (%)	Leaking rate	Spectral radius	RMSE	MAPE
Test-D1	40	0.9	0.6	0.00727	0.14628
	50	0.5	1.2	0.00737	0.13677
	60	0.6	0.9	0.00529	0.08451
	70	0.1	1.0	0.00491	0.08269
	80	0.4	1.4	0.00359	0.05141
	90	0.9	0.3	0.00126	0.01666
Test-D2	40	0.9	0.7	0.01173	0.15735

	50	0.2	1.0	0.00911	0.10657
	60	0.7	0.3	0.00717	0.08006
	70	0.4	0.1	0.00606	0.04159
	80	0.3	1.2	0.00532	0.03775
	90	0.4	1.2	0.00720	0.05401
Test-D3	40	0.9	0.1	0.06024	0.44237
	50	0.2	0.9	0.00622	0.10488
	60	0.2	0.3	0.01623	0.14310
	70	0.8	1.2	0.00488	0.05449
	80	0.5	1.4	0.00375	0.04116
	90	0.5	1.4	0.00241	0.02896

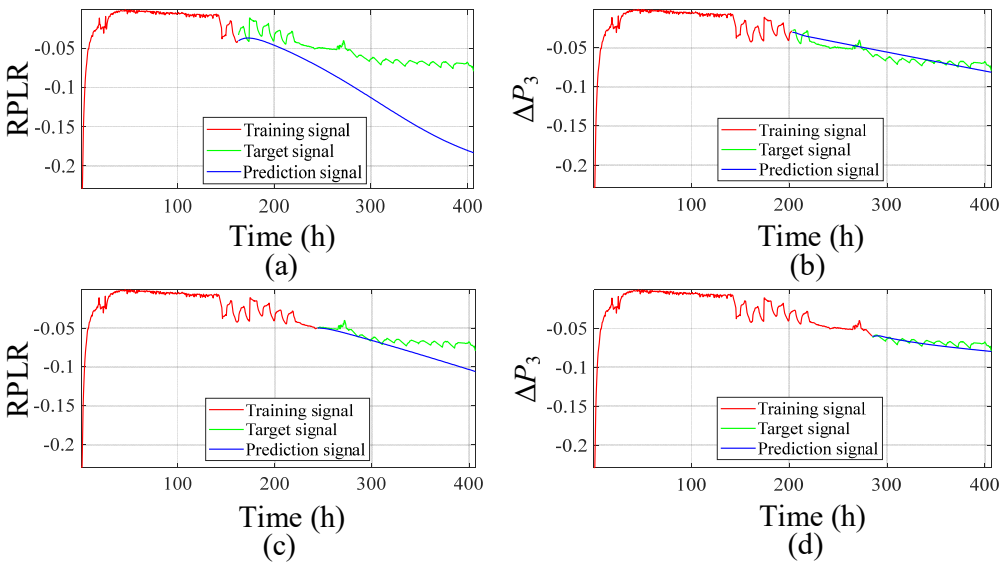


Fig. 4 - 5 The RPLR prediction with single ESN in test-D3: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

4.4.2 Long-term prediction of DWT-EESN

In this task, the initial RPLR can be decomposed into several sub-waveforms with different timescales. Furthermore, they are predicted on the long-term scale by different ESNs separately. The general design principle is that a more complex signal needs to be decomposed into more layers to obtain a higher prediction accuracy. On the contrary, more layers indicate a need for more ESNs to realize the prediction, and the computational complexity will also be increased.

Thus, the selection of decomposition layers is a process of finding the balance between prediction accuracy and computational complexity. In our current study, this selection was done empirically. Besides, the number of decomposition layers is highly related to the signal itself, and it would be varying in different applications. The lumped

decomposition layers have been decided by the trial-and-error method to get a better prediction performance than the single ESN, and the minimum decomposition layers of test-D1, test-D2, and test-D3 are five, seven, and six separately.

From the load profiles of these three dynamic tests, it can be observed that test-D2 is more complex than test-D3, and test-D3 is more complex than test-D1. The selected decomposition layers have also validated this general design principle. Besides, the parameter selection process in the DWT-EESN structure is the same as the single ESN to make sure of a fair comparison. Taking both three tests with 50 % training length, the long-term prediction results of sub-waveforms are presented from Fig. 4 - 6 to Fig. 4 - 8. Taking both three tests with 60 % training length, the long-term prediction results of sub-waveforms are presented from Fig. 4 - 9 to Fig. 4 - 11.

At a certain training length (e.g. 50 %), the prediction results of the sub-waveforms are lumped together to obtain the final RUL. The final RPLR prediction results by the DWT-EESN with a training length of 40 %, 50 %, 60 %, and 70 % are shown from Fig. 4 - 12 to Fig. 4 - 14 for the three dynamic tests.

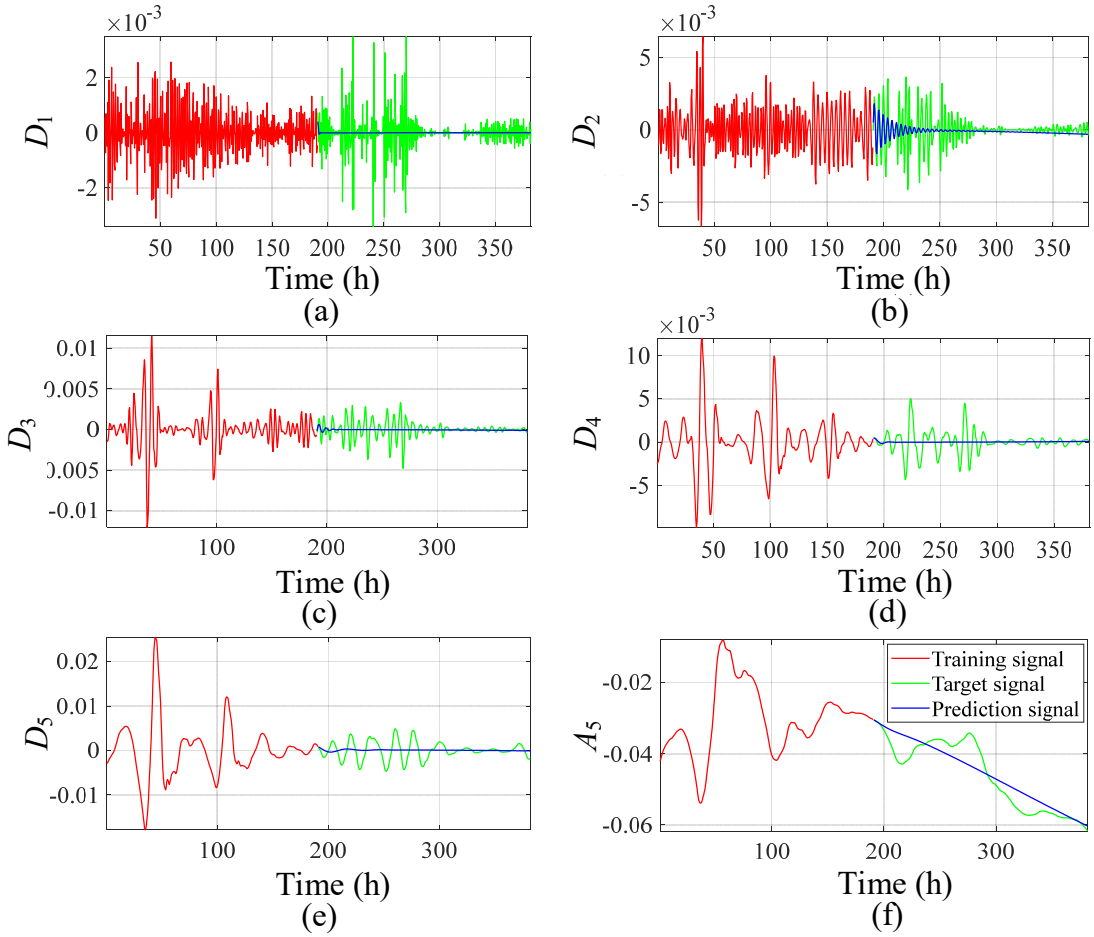


Fig. 4 - 6 Sub-waveforms prediction results of test-D1 in 50 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) A_5.

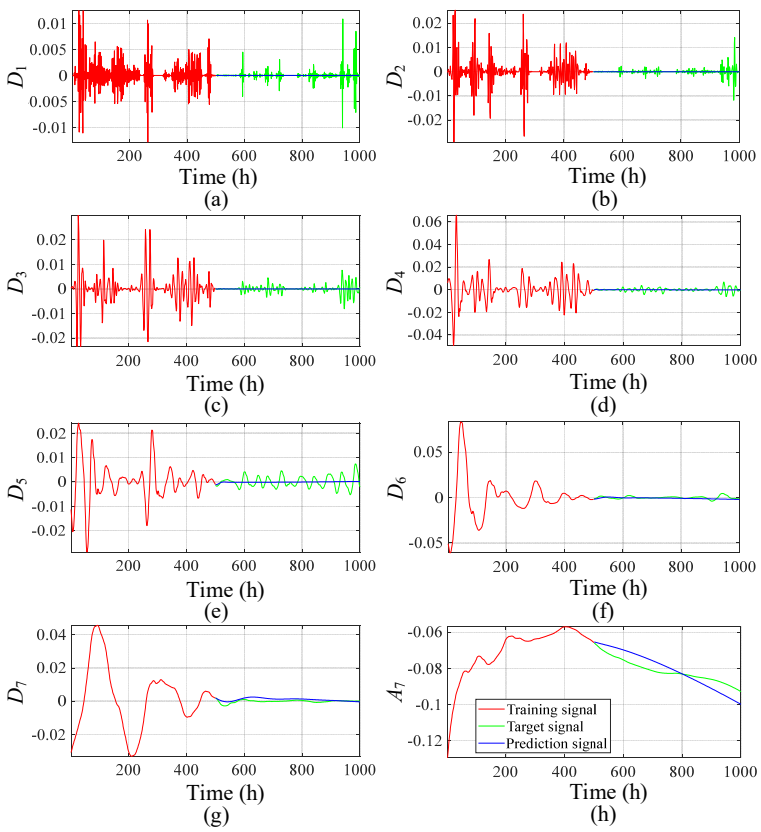


Fig. 4 - 7 Sub-waveforms prediction results of test-D2 in 50 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) D_6, (g) D_7, (h) A_7.

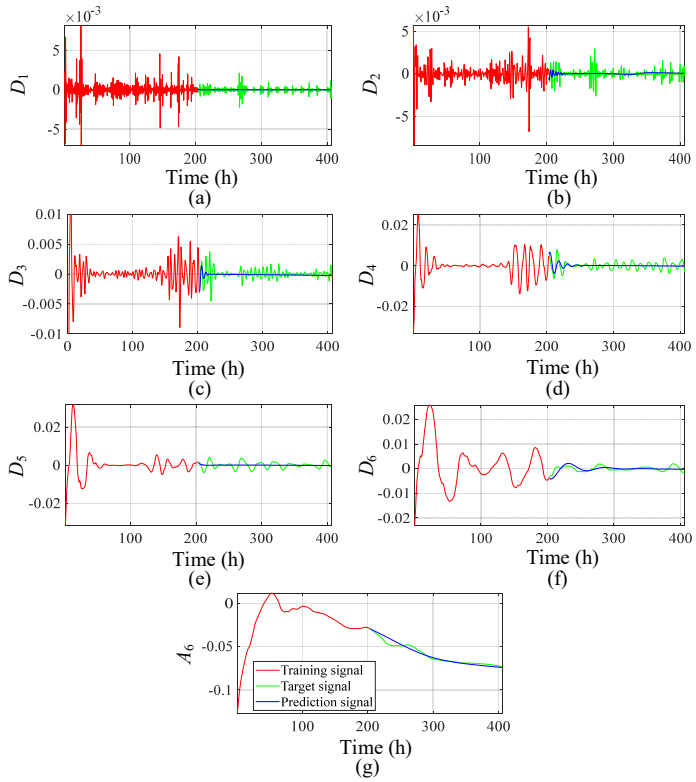


Fig. 4 - 8 Sub-waveforms prediction results of test-D3 in 50 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) D_6, (g) A_6.

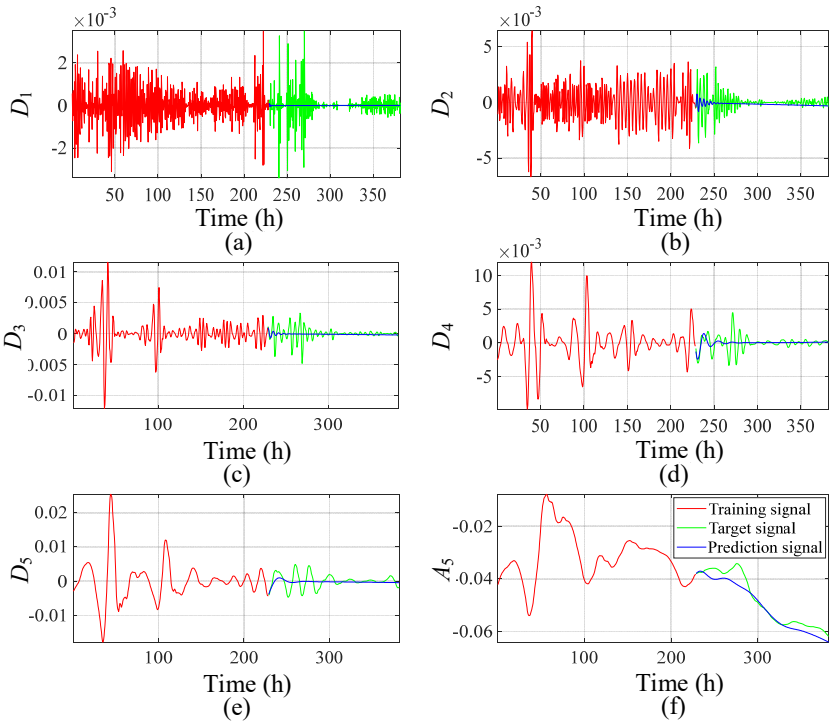


Fig. 4 - 9 Sub-waveforms prediction results of test-D1 in 60 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) A_5.

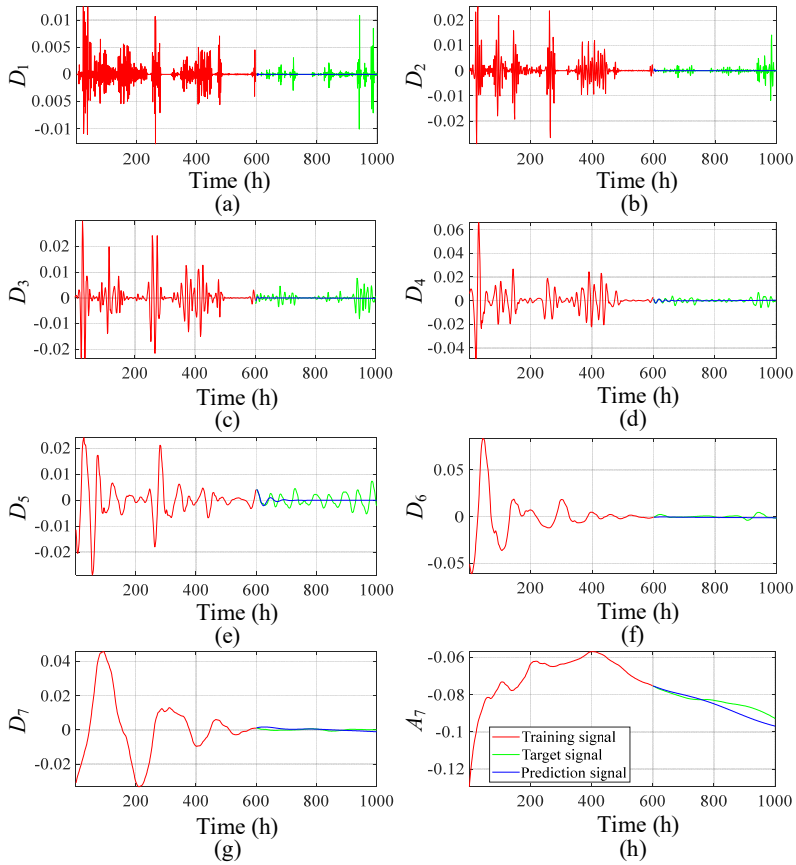


Fig. 4 - 10 Sub-waveforms prediction results of test-D2 in 60 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) D_6, (g) D_7, (h) A_7.

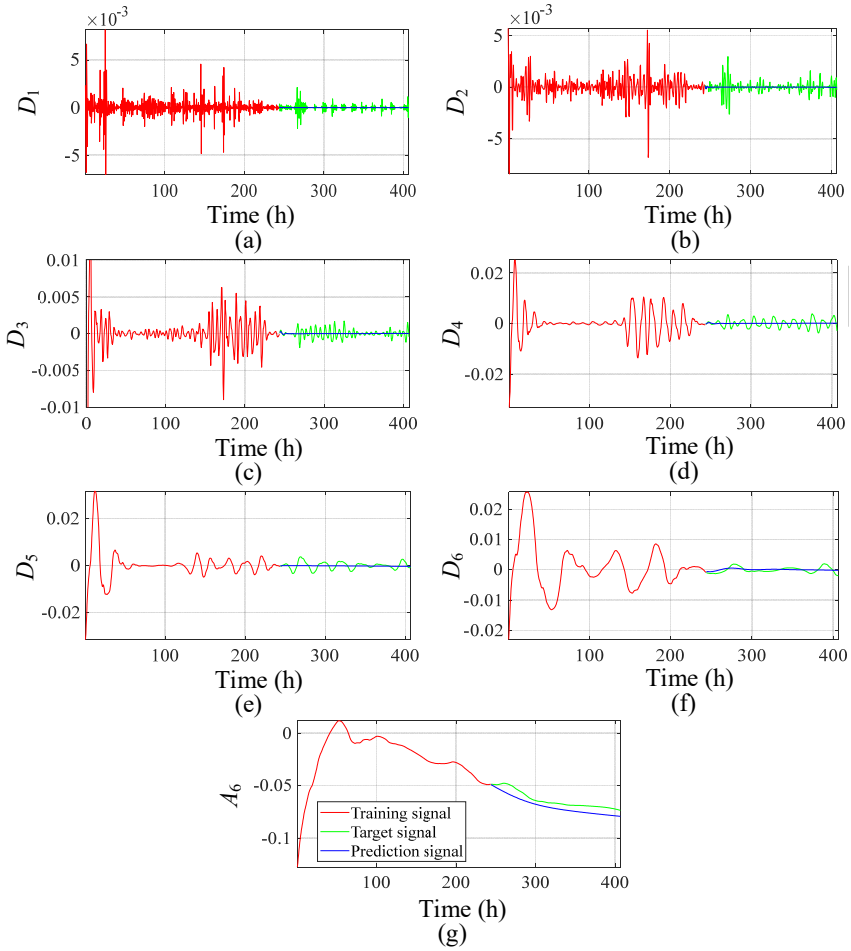


Fig. 4 - 11 Sub-waveforms prediction results of test-D3 in 60 % training length: (a) D_1, (b) D_2, (c) D_3, (d) D_4, (e) D_5, (f) D_6, (g) A_6.

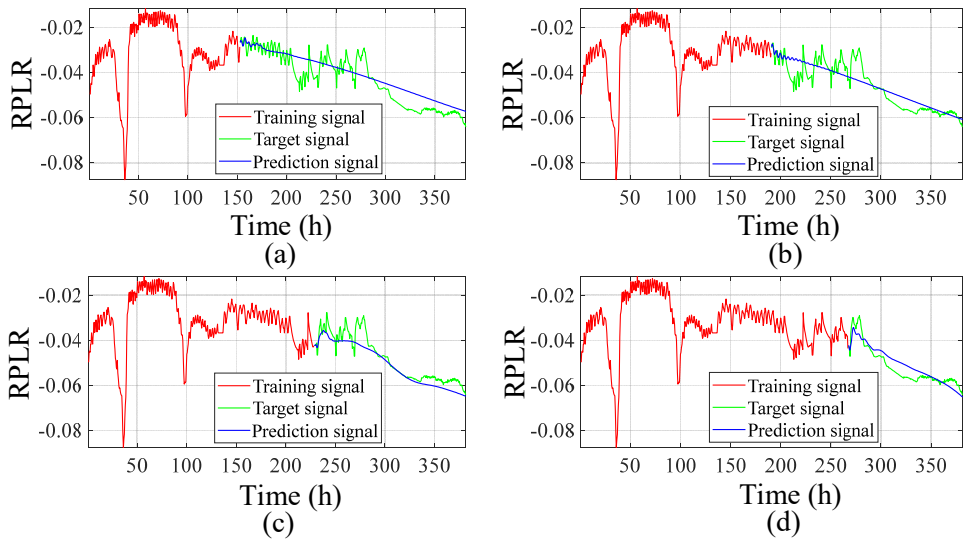


Fig. 4 - 12 The RPLR prediction with DWT-EESN in test-D1: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

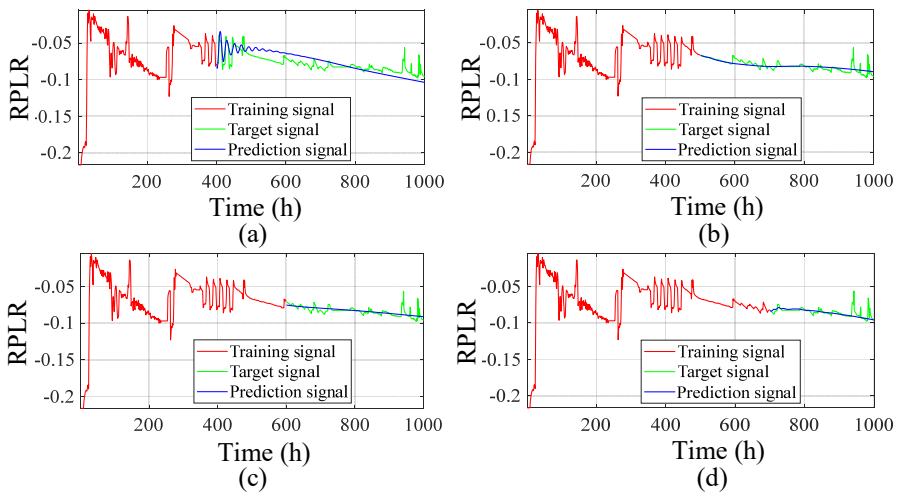


Fig. 4 - 13 The RPLR prediction with DWT-EESN in test-D2: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

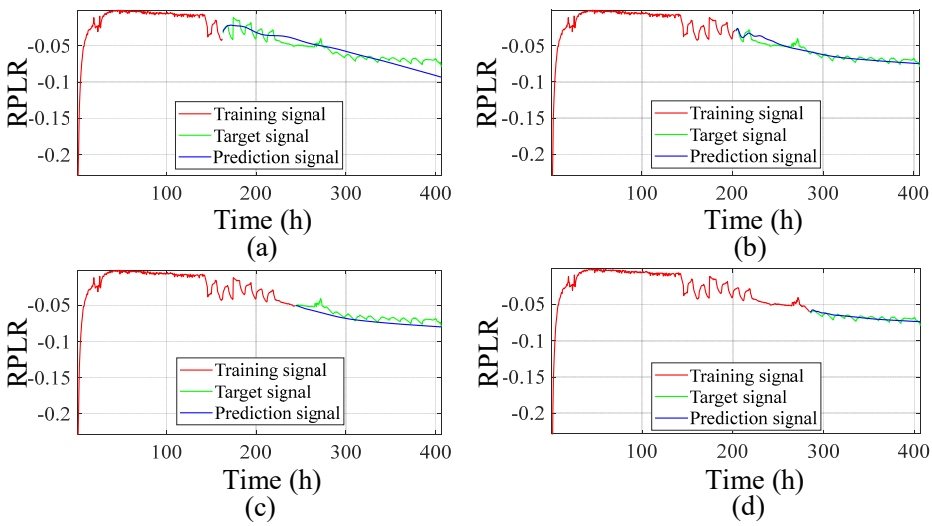


Fig. 4 - 14 The RPLR prediction with DWT-EESN in test-D3: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.

In the DWT-EESN method, the quantified prediction results of RPLR are shown in Tab. 4 - 2. The **notation of "↑"** represents performance improvement (in percentage) when compared to the prediction results of the single ESN. The DWT-EESN method could improve the prediction accuracy as a whole on a long-term scale. The maximum improvements of the three tests are marked out in bold separately in terms of RMSE and MAPE. In the DWT-EESN, the effects of the operating parameters are similar to the single ESN, i.e., 90 % training length of test-D1, and 60 % training length of test-D3. In the case of 50 % training length of test-D2, the effects of temperature are decreased. It shows that the DWT-EESN structure can also weaken the effects of the disturbance.

Tab. 4 - 2 Prediction results of DWT-EESN.

Tests	Training length (%)	RMSE	MAPE
Test-D1	40	0.00526 (↑27.64 %)	0.10597 (↑27.56 %)
	50	0.00494 (↑32.96 %)	0.08927 (↑34.73 %)
	60	0.00416 (↑21.34 %)	0.06438 (↑23.82 %)
	70	0.00368 (↑24.95 %)	0.06547 (↑20.82 %)
	80	**0.00163 (↑54.52 %)**	**0.02317 (↑54.94 %)**
	90	0.00126 (↑0.77 %)	0.01467 (↑11.93 %)
Test-D2	40	0.00977 (↑16.68 %)	0.11464 (↑27.14 %)
	50	0.00756 (↑17.04 %)	0.07260 (↑31.88 %)
	60	0.00555 (↑22.67 %)	0.07027 (↑12.23 %)
	70	**0.00464 (↑23.50 %)**	**0.03178 (↑23.59 %)**
	80	0.00459 (↑13.71 %)	0.03295 (↑12.72 %)
	90	0.00671 (↑6.84 %)	0.04318 (↑20.06 %)
Test-D3	**40**	**0.00891 (↑85.21 %)**	**0.14835 (↑66.46 %)**
	50	0.00415 (↑33.39 %)	0.06411 (↑38.87 %)
	60	0.00667 (↑58.91 %)	0.08575 (↑40.07 %)
	70	0.00265 (↑45.80 %)	0.03245 (↑40.45 %)
	80	0.00225 (↑39.96 %)	0.02681 (↑34.87 %)
	90	0.00224 (↑7.14 %)	0.02628 (↑9.26 %)

4.5 Chapter summary

Based on the RPLP indicator, a combined method, namely DWT-EESN is proposed to deal with the multi-timescale features and improve the long-term prediction performance. Compared with the short-term prediction (single-step or several-step ahead), the iterative prediction process is used in this study to realize the long-term RUL prediction. In this prediction process, the predicted RPLR of the last step is used to predict the new RPLR of the next step, and new measurements are not needed in the prediction phase. In test-D2, for example, the prediction horizon is 600 h, 500 h, and 400 h when the training lengths are 40 %, 50 %, and 60 % separately. Compared with a single ESN without the decomposition process, the DWT-EESN approach could improve the prediction accuracy under all these three dynamic tests. Properly speaking, in terms of RMSE, the maximum improvements of test-D1, test-D2, and test-D3 are 54.52 % (80 % training length), 23.50 % (70 % training length), and 85.21 % (40 % training length) separately. In terms of MAPE, the maximum improvements of test-D2, test-D1, and test-D3 are 54.94 % (80 % training length), 23.59 % (70 % training length),

and 66.46 % (40 % training length) separately.

Chapter 5. Prediction efficiency improvement by genetic algorithm and discrete wavelet transform

5.1 Brief introduction

To improve the prediction accuracy in the long term, four model-based methods of polynomial regression (PR), auto-regressive integrated moving average (ARIMA), PR with the discrete wavelet transform (DWT-PR), and ARIMA with the discrete wavelet transform (DWT-ARIMA) are proposed in [64]. Results on test-SS and test-QD show that the model's (PR and ARIMA) prediction accuracy can be improved when combined with DWT. To further improve the accuracy of [64], three methods of physical degradation model with the PF (PDM-PF), data-driven method of nonlinear autoregressive neural network (NARNN), and their hybrid structure (PDM-PF-NARNN) are proposed in [62]. Results on test-SS show that the method of PDM-PF-NARNN has the best performance in terms of root mean square error (RMSE).

Based on the previous research work, a data-driven approach of discrete wavelet transform-echo state network-genetic algorithm (DWT-ESN-GA) is proposed in this chapter to improve the RUL prediction performance. Firstly, the historical datasets are compressed by the discrete wavelet transform (DWT). Secondly, the approximation components of the original data are predicted in the compressed space by echo state network (ESN). Rather than predicting the degradation data themselves, their shortened coefficients are evaluated to decrease the prediction data points, i.e., from 2016 data points to 253 data points. Besides, a genetic algorithm (GA) is used to optimize the key parameters of ESN, and it can further increase the prediction accuracy. Finally, the inverse DWT is utilized to reconstruct the coming data based on the estimated approximation components. The performance of the proposed approach is evaluated by three different experimental tests under steady-state (test-SS), quasi-dynamic (test-QD), and full dynamic operating conditions (test-D2) separately. The durability tests under different operating conditions are about 1000 h, and six stages (each stage is 168 h) are divided in this chapter. The one-week-ahead degradation prediction is realized based on the prognostic method, i.e., the prediction horizon is 168 h.

The motivation of this study is to improve the RUL prediction efficiency on a long-term scale under both static and dynamic operating conditions. The major contributions of this chapter are summarized as follows:

- The health indicator of relative power-loss rate (RPLR) is proposed for the dynamic operating condition. Only the polarization curve at the beginning of life needs to be performed, and the normal operation of the system would not be interrupted in a real system.
- The degradation data are compressed by discrete wavelet transform (DWT) to shorten the prediction length. The degradation data can now be represented by their

approximation components based on the DWT in the training process.

- The genetic algorithm (GA)-optimized ESN is used to predict the approximation components of the degradation data. The predicted approximation components and inverse discrete wavelet transform (IDWT) are used to reconstruct the new degradation data in the prediction process.
- Based on the "decomposition-prediction-reconstruction" structure by the DWT-ESN-GA algorithm, the original data points during the prediction horizon can be compressed by a factor of 8, i.e., the degradation characteristics in 2016 data points (corresponding to 168 h) can be effectively represented by only 253 data points.

5.2 Genetic algorithm parameter optimization

There are several parameters of ESN, and the optimization of multiple parameters would contribute to its practical application. The parameters that need to be set in ESN are the number of neurons N in the dynamic reservoir, the leaking rate a, the spectral radius ρ of $\mathbf{W}$ (maximum eigenvalue), the regression coefficient β, reservoir connectivity, input scaling, and coefficients of different weight matrices.

The random and grid-search are two straightforward parameter optimization methods of ESN [117]. In general, the random method needs multiple attempts, and it can reach a roughly satisfactory performance with the help of expert experiences or the parameter's analysis of variance (ANOVA) [100]. The grid-search method is able to find the satisfactory parameter combination. Nevertheless, this exhaustive search technique is time-consuming. Some automatic global parameter optimization approaches have been used to improve optimization efficiency. The stochastic gradient descent is used for optimizing four parameters i.e., input scaling, ρ, a, and β [132]. The big bang-big crunch is utilized to optimize N, ρ, a, and β [104]. Nevertheless, the implementation processes of stochastic gradient descent and big bang-big crunch are computationally complex. Some other efficient optimization methods can also be used to optimize the parameters of ESN. In [124], the PSO is used to optimize the coefficients of weight matrices, and in [139], two leaking rates and two spectral radii are optimized by GA. The GA is inspired by the mechanisms of natural evolution which consist of selection, crossover, mutation, etc. And it has shown its satisfactory performance in optimizing the leaking rate and the spectral radius of ESN.

Based on the above work, four key parameters, i.e., N, a, ρ, and β are selected and optimized in this RUL prediction task. In general, a larger value of N could map a bigger reservoir space, and it would be easier to construct the linear recursive relation. It is necessary to choose a big N when handling a complex task. Nevertheless, computing is expensive when dealing with huge amounts of neurons even though the training structure of ESN has big advantages over other RNN methods. In this task, the neuron number N is selected in the range of [50, 500], and the data interval is 50, i.e.,

the number of neurons is increased by 50 at each try. After multiple trials, $N = 400$ is selected.

Except for N, the remaining 3 parameters (a, ρ, and β) are need to be identified. The grid-search method is used for comparison. Taking the grid steps of a, ρ, and β as 0.1, 0.01, 0.001 separately, there would be 13500 parameter combinations (execution time is about 3 days). Besides, the computational complexity grows exponentially with the number of parameters. As an efficient and widely used metaheuristic method for searching the optimal (or approximately optimal) solution, the GA is utilized in this task to optimize the $a \in (0,1)$, $\rho \in (0,1.5)$, and $\beta \in (8\times10^{-3}, 8\times10^{-1})$ of ESN. Besides, GA can make the most use of historical data to find the global optimum even if it has the stochastic searching property. The execution time of GA is about 20 min, and the timescale is reasonable for the 168 hours' degradation prediction. The objective of GA can be expressed as

$$y_{obj} = \arg\min f(\alpha,\rho,\gamma) = \operatorname{argmin} \frac{1}{L}\sum_{i}^{L}\sqrt{\frac{1}{B}\sum_{n=1}^{B}(y_i^{predict}(n) - y_i^{target}(n))^2} \tag{5-1}$$

Where L is the dimension of the output signal, B is the total data points in the training (or testing) part. The pseudo-code of GA is shown as

Algorithm 2: Working principle of GA

1: **Define the input and output variables**

input: population size L, elite number N_e, crossover probability P_c, mutation probability P_m, generation number G.

output: best fitness, best individual, parameter combinations.

2: **Parameter's initialization**

L=40, N_e=2, P_c=0.8, P_m=0.01, G=200.

3: **Generate the initial population randomly**

4: **Calculate the objective function and fitness of individuals**

for n=1 to G, **do**

select the parents based on the fitness of individuals

generate the population by elite

crossover, mutation

call **Algorithm 1** and evaluate the performance (based on RMSE)

update the population and rebuild the next generation

end

5: **Return the best fitness and its parameter combinations**

In RUL prediction of the PEMFC system, the degradation information has always more impacts on large time scales, and it has closer relationships with the low-frequency part of the original signal. Therefore, the approximation component is more interesting than the detail component at each decomposition level. In this task, the orthogonal wavelet db3 which belongs to the Daubechies wavelets family is used because it can compress the wavelet energy as much as possible to the approximation component. The waveforms of the Daubechies wavelets family are shown in Fig. 5 - 1 [141].

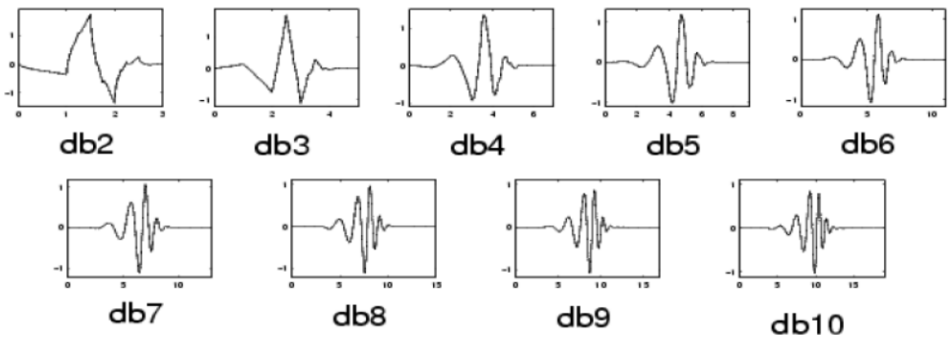


Fig. 5 - 1 The waveforms of Daubechies wavelets family [141].

Based on the GA-optimized ESN structure, the DWT is combined with ESN to improve the verification efficiency of RUL. For different experimental tests, the inputs of ESN are different in this work, i.e., power for steady-state, power for quasi-dynamic, and RPLR for dynamic operating conditions. In the considered datasets, the output power under the steady-state condition is from 222 W to 236 W. The output power under the quasi-dynamic condition is from 219 W to 233 W. And the RPLR under the dynamic condition is from -0.25 to 0. For the sake of fairness, all the input data of ESN are normalized to the same interval of [0, 1]. And then the output data of ESN are mapped to their original dimension by inverse normalization. The working principle of the DWT-ESN-GA approach is summarized in **Algorithm 3**.

Algorithm 3: Working principle of DWT-ESN-GA

1: **Define the input and output variables**

input: historical data D_w; **output**: future data $\widehat{D}_{w+1}$

2: **Data pre-processing**

filtering (moving-average), resampling (time-interval is 0.5 h),

normalization (from 0 to 1)

3: **RUL prediction**

for n=1 to N_w, **do**

DWT the data of previous weeks at level j (cA_{j+1} and cD_{j+1})

call **Algorithm 1 (Section 2.1)**, predict the future $c\hat{A}_{j+1}$ of the next week

call **Algorithm 2 (Section 5.2)**, optimize the parameters of ESN

IDWT the $c\hat{A}_{j+1}$ to the $\hat{D}_{w+1}$ of the coming week

end

4: **Return the results of the prediction side**

RMSE, MAPE, RUL, etc.

5.3 Implementation process of DWT-ESN-GA

After the data measurements (voltage and current) and data pre-processing (filtering and resampling), the health indicators can be calculated under different operating conditions, i.e., stack power for the steady-state and quasi-dynamic conditions, and RPLR for the full dynamic condition. The moving average filtering method is used, and the length of the moving window is empirically set to 31. The health indicators are shown in Fig. 5 - 2.

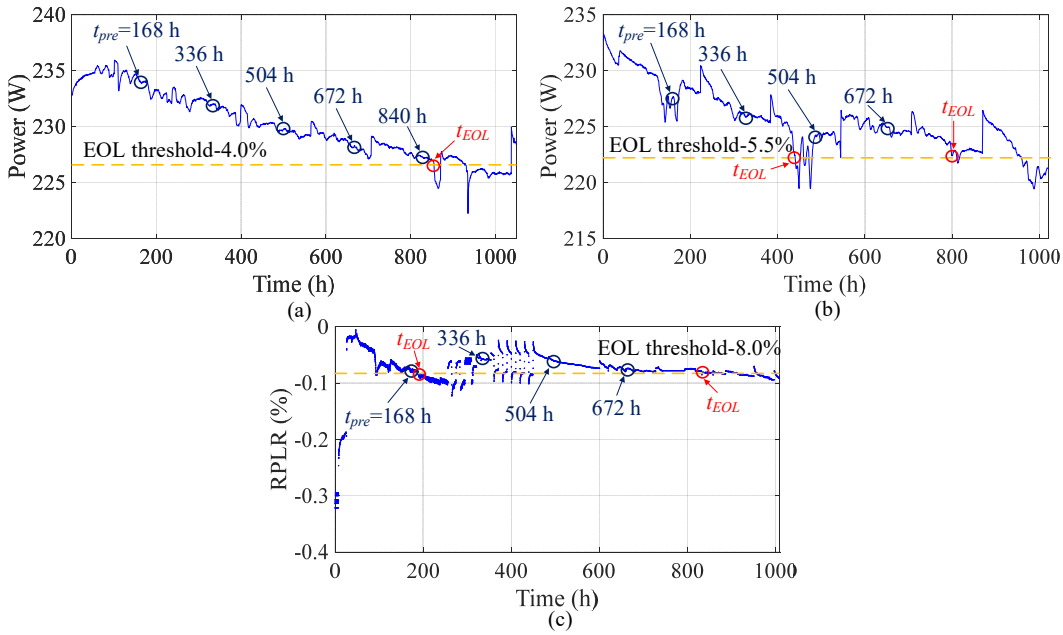


Fig. 5 - 2 Health indicators of three tests: (a) power of test-SS, (b) power of test-QD, (c) RPLR of test-D2.

The fuel cell stacks of test-SS and test-QD are the same, and the test-QD has more serious mission profiles. Therefore, 4.0 % and 5.5% of initial power are used as the EOL thresholds in test-SS and test-QD separately. The mission profile of test-D2 is the most complex in these three tests, and 8 % power loss is chosen as the EOL threshold to evaluate the prediction performance of RUL.

$$\begin{cases} t_{RUL}^{act} = t_{EoL} - t_{pre} \\ t_{RUL}^{pre} = t_{pEoL} - t_{pre} \end{cases} \tag{5-2}$$

The actual RUL (t_{RUL}^{act}) can be defined as the time between t_{pre} (the time when the prediction begins) and t_{EOL} (the time when the actual value reaches the EOL threshold). The predicted RUL (t_{RUL}^{pre}) can be defined as the time between t_{pre} and t_{pEOL} (the time when the prediction value reaches the EOL threshold).

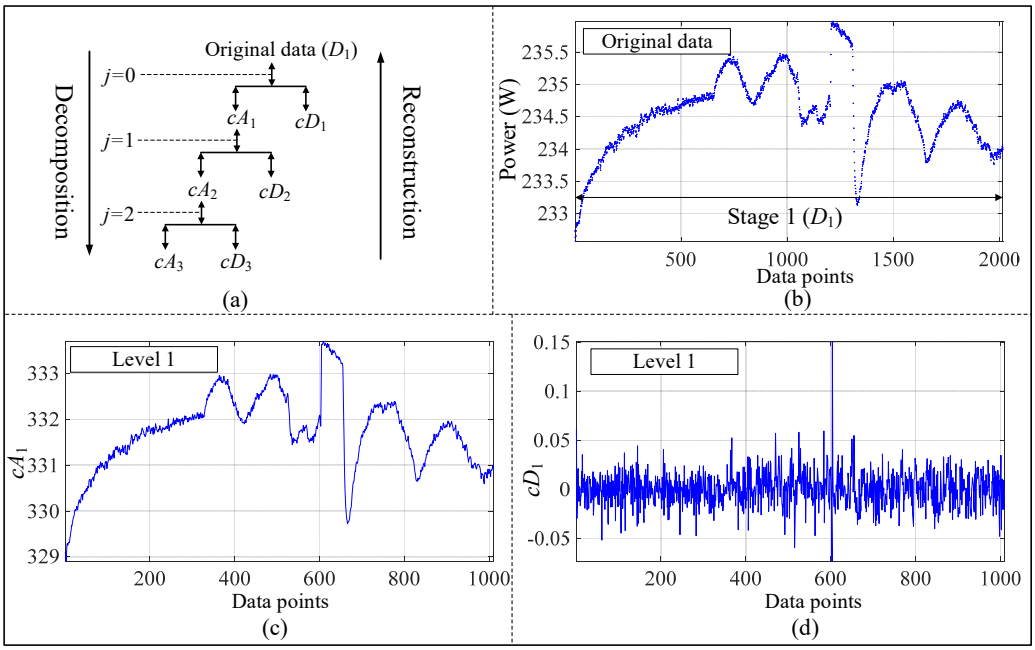


Fig. 5 - 3 The wavelet transform of test-SS: (a) decomposition and reconstruction process of the wavelet transform, (b) original power data in stage 1, (c) approximation component coefficient in level 1, (d) detail component coefficient in level 1.

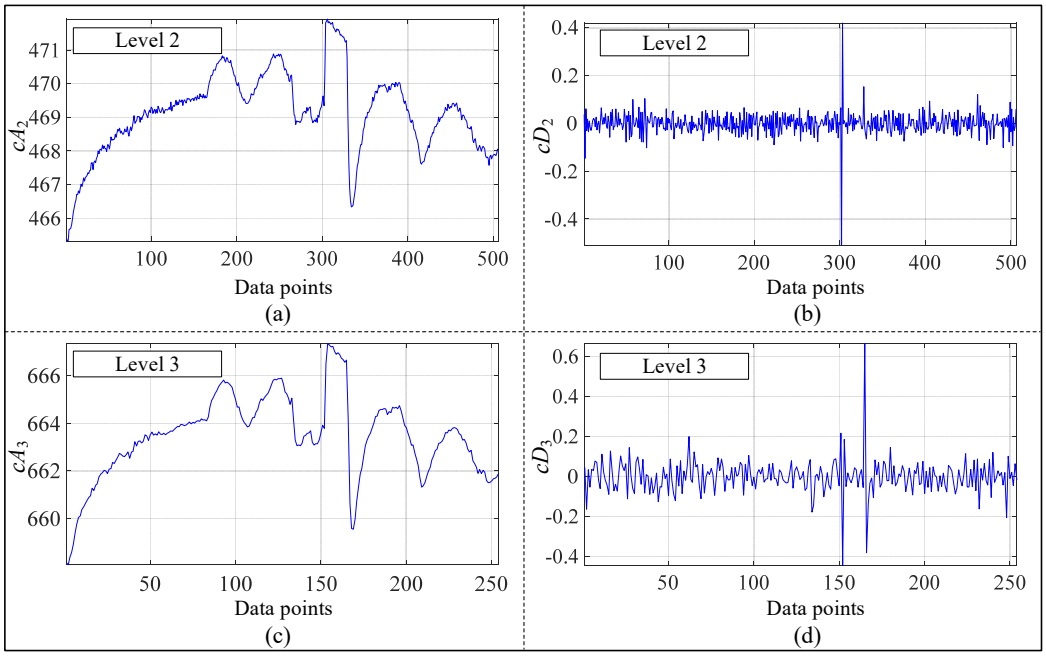


Fig. 5 - 4 The approximation component coefficients and detail component coefficients of level 2 and level 3 in stage 1: (a) cA_2, (b) cD_2, (c) cA_3, (d) cD_3.

Among the whole N stages, the historical data in the first (N-1) stages are used for

training, and the data of the Nth stage are used for testing. The decomposition and reconstruction process of wavelet transform at level 3 are shown in Fig. 5 - 3 (a). Taking the first stage's data of test-SS as an example (Fig. 5 - 3 (b)), the power in the first stage (D_1) is decomposed by the DWT method, and the approximation component coefficient (cA_1) and detail component coefficient (cD_1) in level 1 are shown in Fig. 5 - 3 (c) and (d). The approximation component coefficients (cA_2 and cA_3) and detail component coefficients (cD_2 and cD_3) in level 2 and level 3 are shown in Fig. 5 - 4 (a)-(d).

There are 2016 data points in each stage (e.g., D_1 or D_2), and the time interval of every two adjacent data points is 5 min. In Fig. 5 - 4 (c) and (d), the length of cA_3 and cD_3 are the same (253 points), and they are the half-length of cA_2 and cD_2 (506 points). The length of cA_2 and cD_2 is the half-length of cA_1 and cD_1 (1010 points). Results indicate that the length of cA_{j+1} and cD_{j+1} are the same, and they are the half-length of the cA_j and cD_j in the previous level (j=0, 1, 2). Based on this implementation process, the length of original power data can be shortened by the scale factor of 2^{j+1}.

The decomposition level j is decided by the trial-and-error principle. Based on the DWT process in Fig. 5 - 3 (a), the 2016 data points (168 h) can be compressed into 1010 data points at the first decomposition level (j = 0). And the 1010 data points can be compressed into 506 data points at the second decomposition level (j = 1). Then, the 506 data points can be compressed into 253 data points at the third decomposition level (j = 2). With the increase of j, the length of the original signal can be compressed shorter and shorter. At the same time, the original data would lose more and more high-frequency characteristics if only considering the approximation component in each level. The scale factor is selected as 2^3 (j = 2) in this task. On the one hand, the degradation's evaluation in the coming 2016 data points can be verified by only 253 data points. On the other hand, the property summation of the high-frequency part (cD_1 + cD_2+ cD_3) have slight effects (< 0.1 %) on the original signal in term of relative wavelet energy. The energy distributions of different decomposition layers in different tests are discussed in **Section 5.5**.

As the opposition process of DWT, the approximation component and detail component in the higher level (a higher j) can be combined to reconstruct the approximation component in the lower level (a lower j) by IDWT (Fig. 5 - 3 (a)). When j = 2, the cA_3 and cD_3 can be used to construct the cA_2. The cA_2 and cD_2 can be used to construct the cA_1 when j = 1, and the cA_1 and cD_1 can be used to construct the cA_0 when j = 0. In RUL prediction of the PEMFC system, the degradation information has always more impacts on large time scales, and it has closer relationships with the low-frequency part of the original signal. Therefore, the approximation component is more interesting than the detail component at each decomposition level. Only using the approximation component at j = 2 (cA_3) to reconstruct the cA_2, cA_1, and D_1 is reasonable because the low-frequency part in each level (cD_3 + cD_2 + cD_1) has almost no effects

on the RUL prediction. Compared with predicting the original data D_1 directly, the verification length can be reduced significantly when predicting the cA_3 with the help of DWT and IDWT.

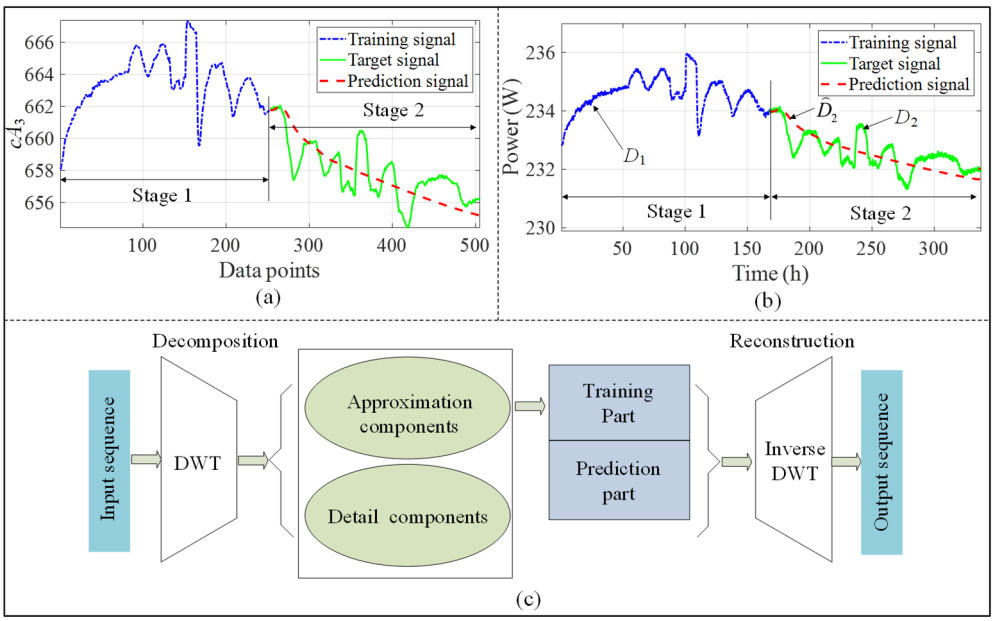


Fig. 5 - 5 Implementation structure of decomposition and reconstruction: (a) the cA_3 prediction in the first 2 stages, (b) the reconstructed power in the first 2 stages, (c) the implementation structure of DWT-ESN-GA.

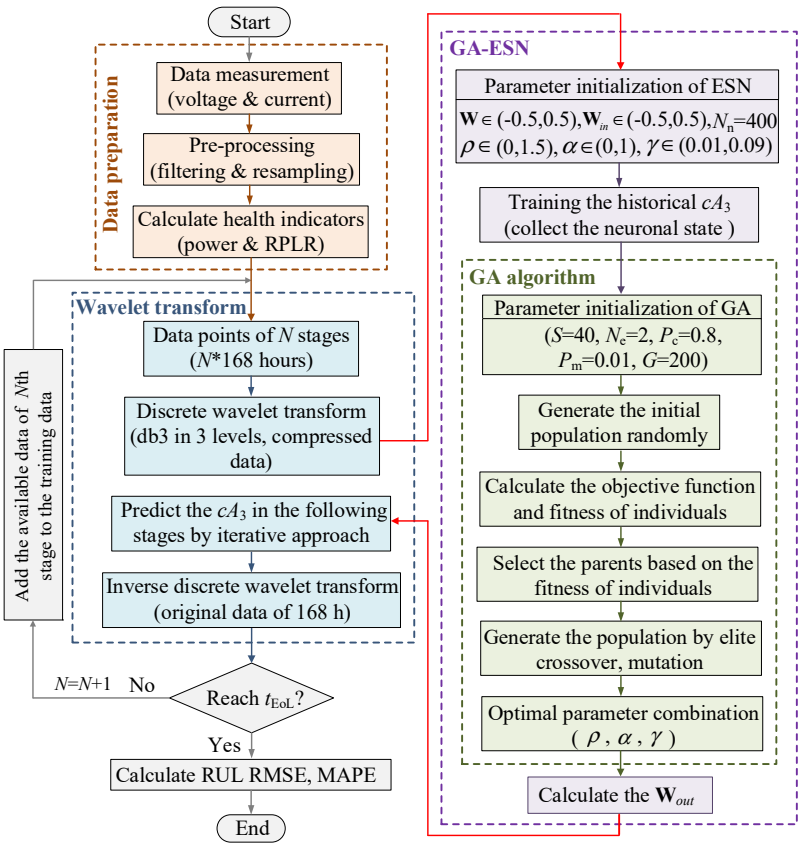


Fig. 5 - 6 The flowchart of DWT-ESN-GA.

The working principle of the proposed approach is shown in Fig. 5 - 5. Based on the data of the first 2 stages of test-SS, the GA-optimized ESN is used for training the cA_3 (253 points) in the first stage and predicting the $c\hat{A}_3$ (253 points) in the second stage ((Fig. 5 - 5 (a)). Finally, the power data in the second stage ($\widehat{D}_2$) is reconstructed by the predicted $c\hat{A}_3$ and IDWT method ((Fig. 5 - 5 (b)). The implementation structure of the proposed method is shown in Fig. 5 - 5 (c). The flowchart of the proposed DWT-ESN-GA is shown in Fig. 5 - 6.

5.4 Experimental results under different mission profiles

5.4.1 Prediction results of three tests

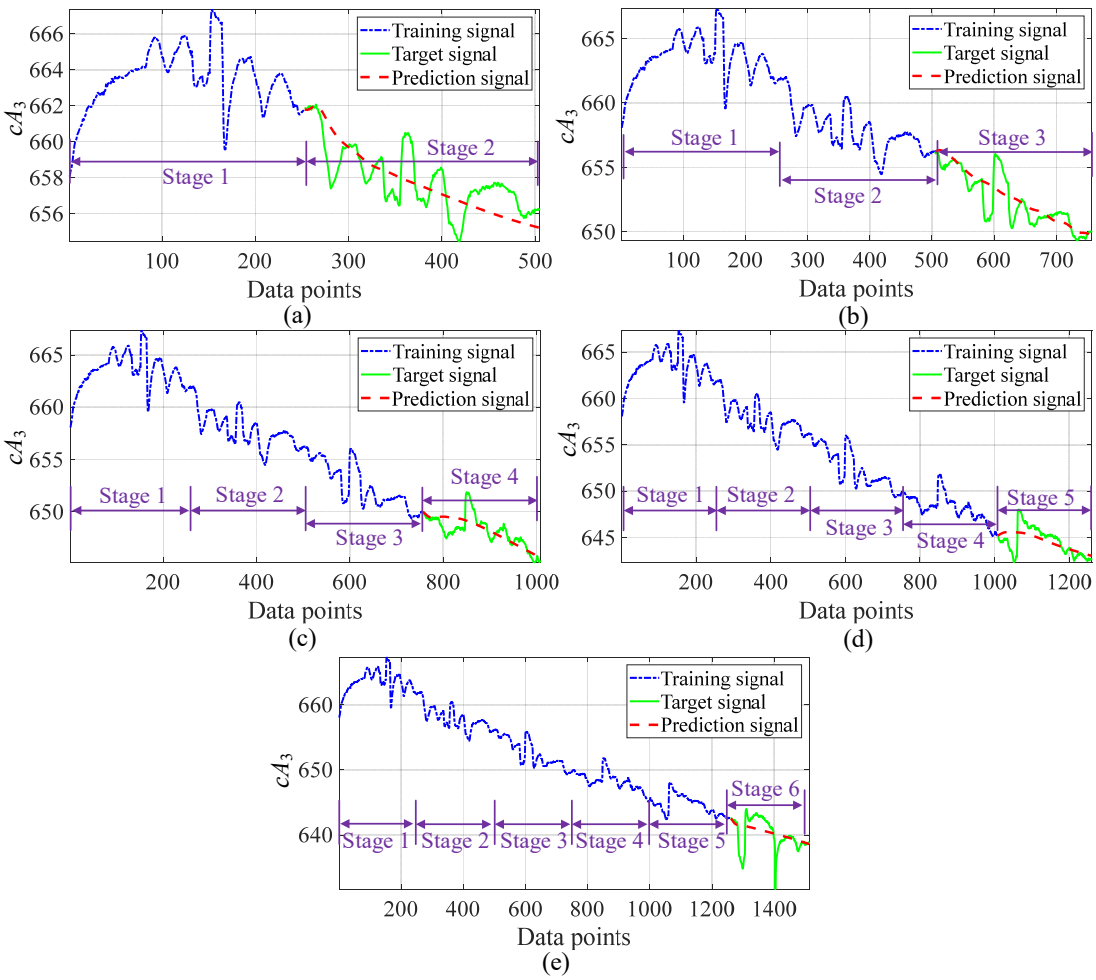


Fig. 5 - 7 The approximation component coefficients prediction of test-SS at level 3: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Based on the DWT, the approximation component coefficient predictions of test-SS, test-QD, and test-D2 at level 3 are shown from Fig. 5 - 7 to Fig. 5 - 9. There are 6 prediction stages in each operating condition, and the cA_3 prediction is performed at the end of each stage. The 1st prediction is performed at the end of stage 1, the historical

cA_3 of stage 1 (253 data points) are used for training and the ESN-GA method is used to predict $c\hat{A}_3$ of stage 2. The 2nd prediction is performed at the end of stage 2, the historical cA_3 of stage 1 and stage 2 (506 data points) are used for training and the ESN-GA method is used to predict $c\hat{A}_3$ of stage 3. By that analogy, the cA_3 prediction in all stages can be realized by a one-stage-ahead prediction process.

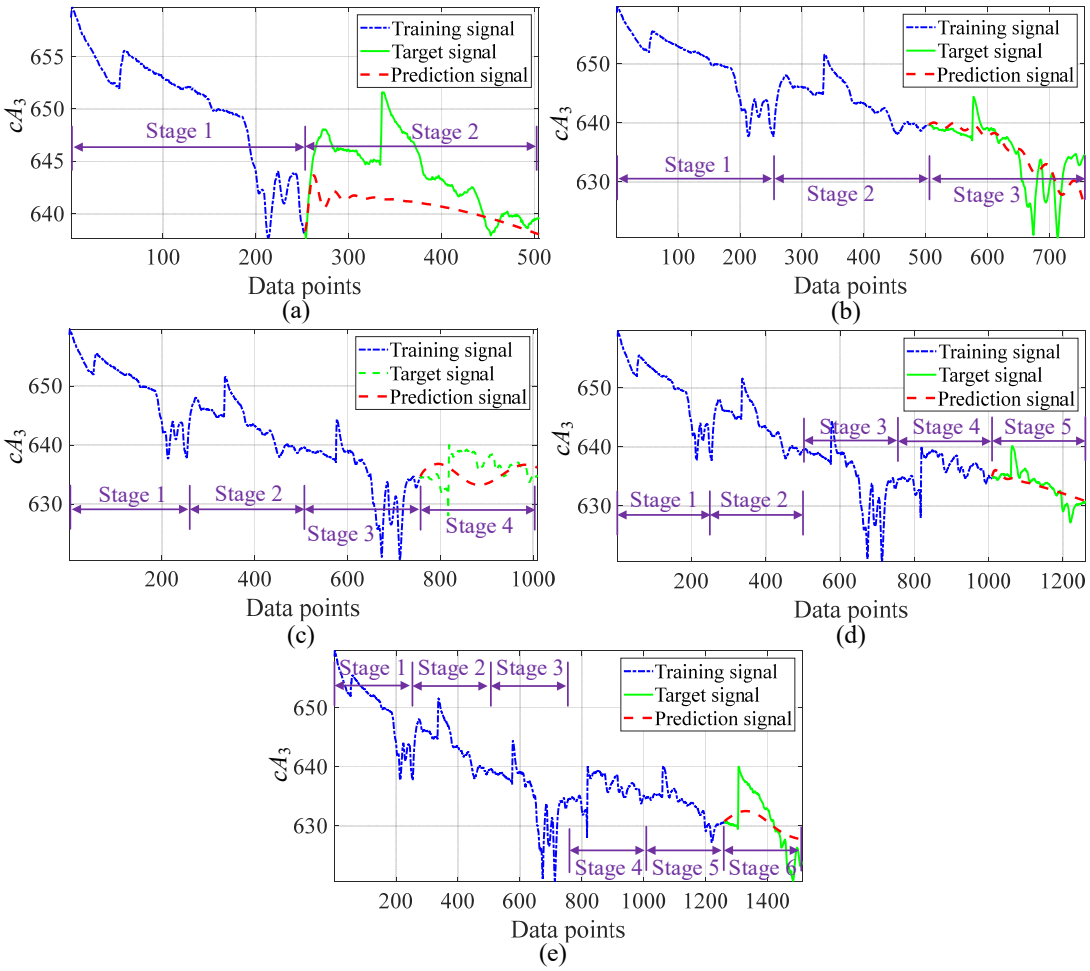


Fig. 5 - 8 The approximation component coefficients prediction of test-QD at level 3: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

The degradation predictions of test-SS, test-QD, and test-D2 are shown in Fig. 5 - 10, and the detailed results in each stage are shown from Fig. 5 - 11 to Fig. 5 - 13. Based on the predicted $c\hat{A}_3$ at level 3 and the IDWT method, the power prediction curve can be reconstructed to its original timescale. In each experimental test, the data of the first stage D_1 are used to predict the data of the second stage $\widehat{D}_2$. At the end of the second stage, the data of the first 2 stages (D_1 and D_2) are used together to predict the data of the third stage $\widehat{D}_3$. By that analogy, the RUL prediction can be realized in the whole durability testing by the one-week advanced horizon. The RMSE and MAPE of the three tests at different operation conditions are shown in Tab. 5 - 1 and Tab. 5 - 2.

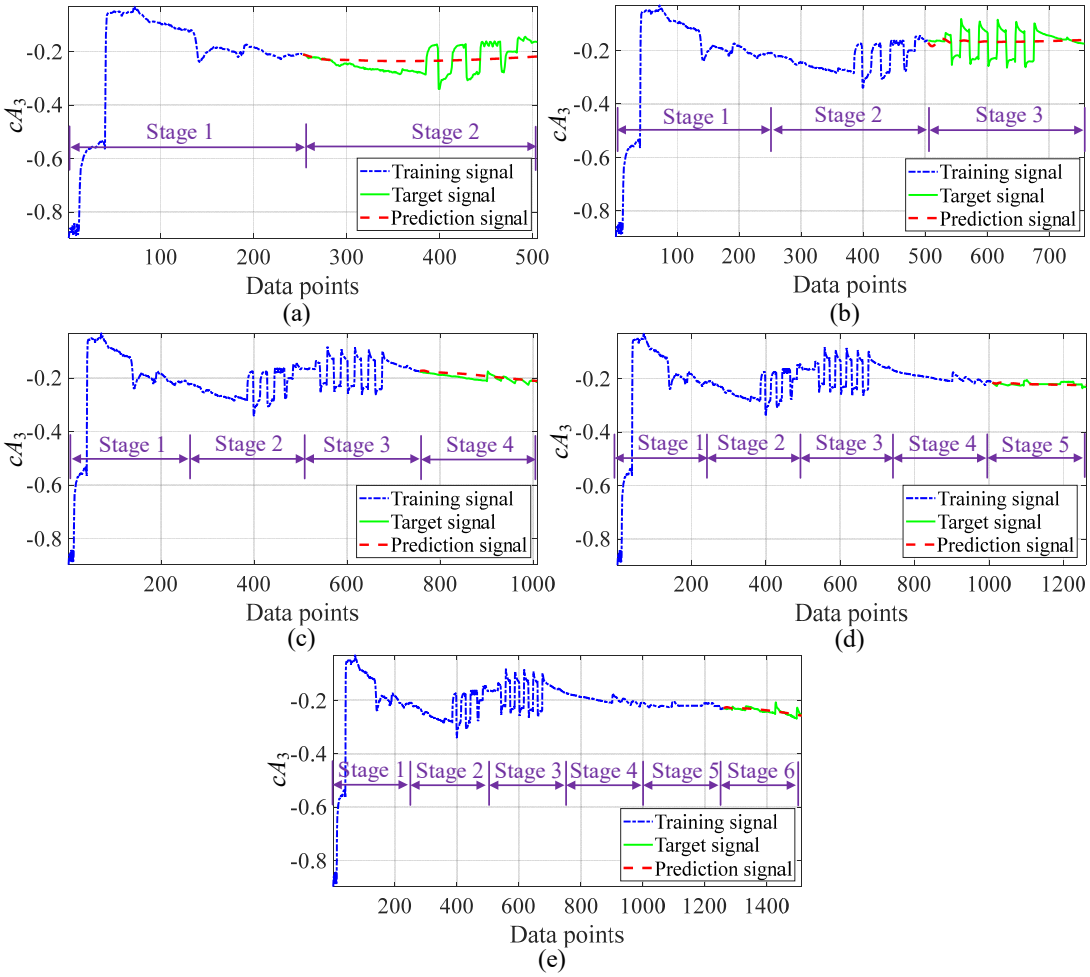


Fig. 5 - 9 The approximation component coefficients prediction of test-D2 at level 3: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

The power prediction of test-SS (Fig. 5 - 10 (a)) and test-QD (Fig. 5 - 10 (b)) are satisfactory because the prediction curve has a good reflection on the power's degradation tendency. The RMSE and MAPE of test-SS are lower than test-QD in every stage, and it means that the prediction accuracy will decrease with the increase of load dynamics. Besides, the prediction error would be severely affected by the disturbance, i.e., 850 h and 940 h of test-SS, 150 and 460 h of test-QD.

The proposed health indicator (RPLR) is defined as a percentage form in test-D2. Thus, the results cannot be compared with test-SS and test-QD directly. It is intuitively plausible that the prediction performance of test-D2 (Fig. 10 (c)) is worst because of its rich dynamic characteristics. It is difficult to acquire a decreasing tendency especially under the load cycling period (250 h - 450 h). This is the universality unsolved problem that should be explored further because simulating the suddenly dynamic with multi-step ahead is difficult for any prediction tools.

Comparing Fig. 5 - 11 to Fig. 5 - 13 with Fig. 5 - 7 to Fig. 5 - 9, the original data D_w (w=1, …, 5) are shortened by the DWT with a scale factor of 2^3, and then the approximation component of the third level ($c\hat{A}_3$) is predicted instead of a prediction of the original data.

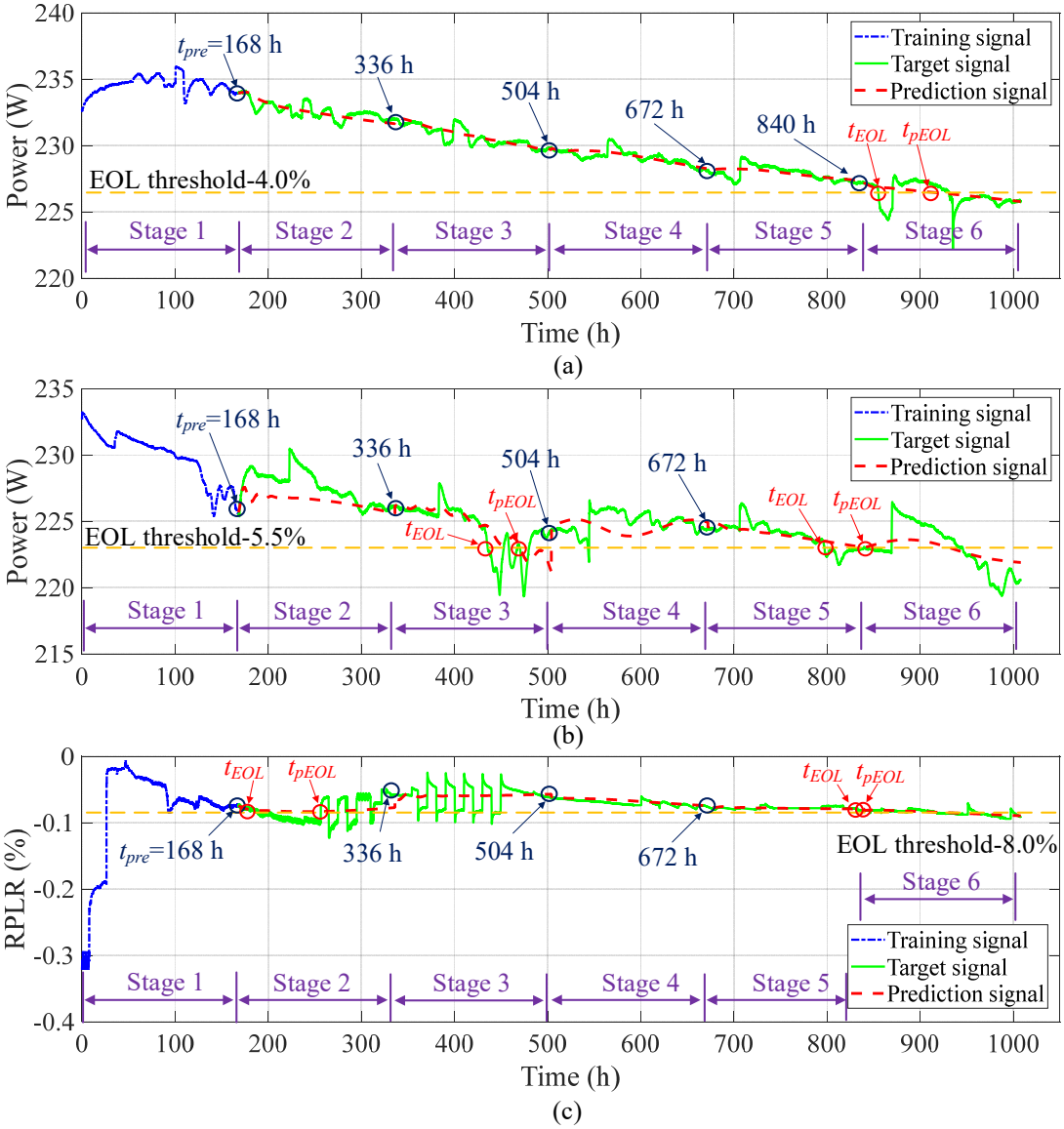


Fig. 5 - 10 The degradation predictions at different stages under different operating conditions: (a) test-SS, (b) test-QD, (c) test-D2.

Tab. 5 - 1 RMSE results of three tests in different stages.

Tests	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
Test-SS	0.3911	0.3928	0.3265	0.3960	0.7645
Test-QD	1.3311	1.2337	0.8240	0.6060	1.2510
Test-D2	0.0157	0.0164	0.0036	0.0019	0.0027

Tab. 5 - 2 MAPE results of three tests in different stages.

Tests	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
Test-SS	0.0013	0.0013	0.0011	0.0013	0.0023
Test-QD	0.0045	0.0042	0.0029	0.0020	0.0046
Test-D2	0.1716	0.2444	0.0438	0.0176	0.0226

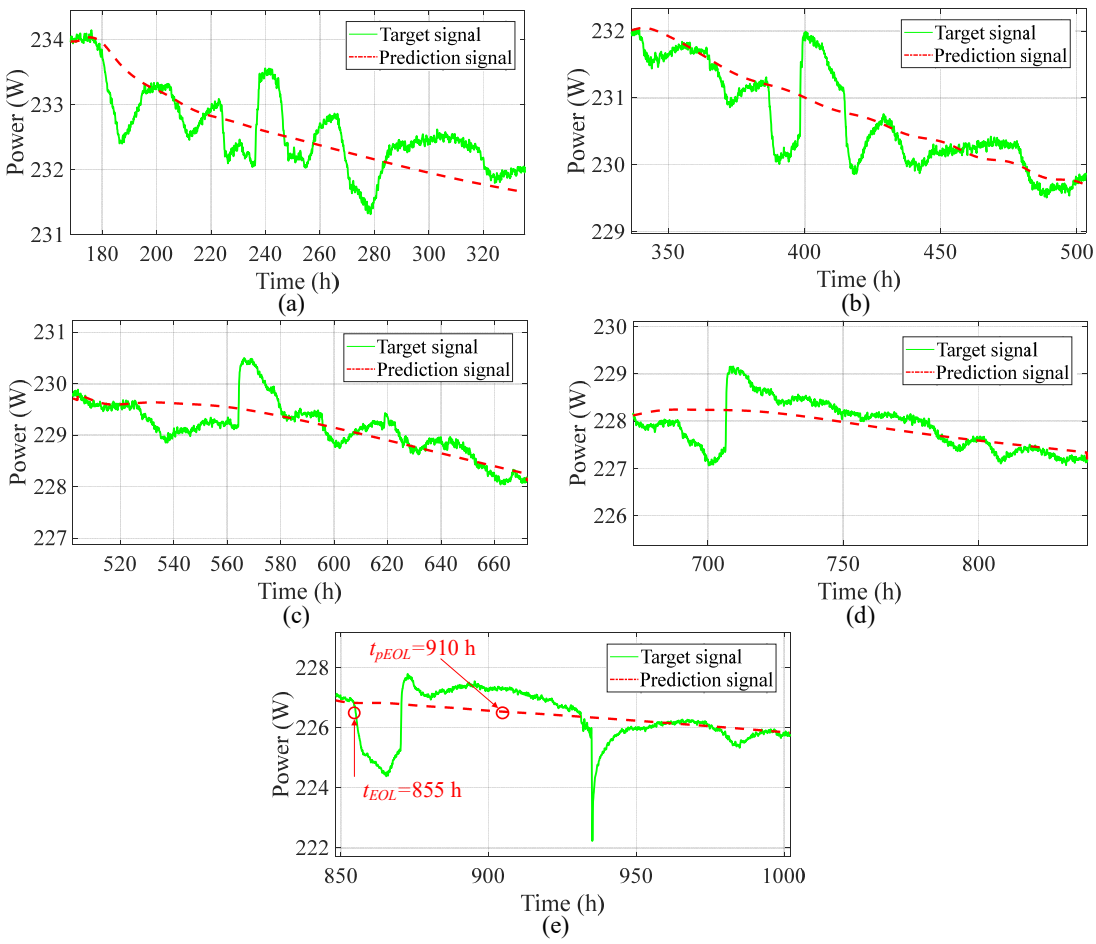


Fig. 5 - 11 The power prediction of test-SS: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Based on the defined EOL threshold, the RUL of different tests can also be analyzed.

For test-SS, the initial power is about 236 W, and the 4.0 % power loss is about 226.6 W. In stage 5 (Fig. 5 - 10(a)), the time when the prediction begins (t_{pre}) is at 840 h, the time when the actual value reaches EOL threshold (t_{EOL}) is at 855 h, and the time when the prediction value reaches EOL threshold (t_{pEOL}) is at 910 h. Thus, the t_{RUL}^{act} is 15 h, the t_{RUL}^{pre} is 70 h, and the error of RUL is 55 h.

For test-QD, the initial power is also 236 W, and the 5.5 % power loss is about

223 W. Without considering the fault disturbance (case 1), the t_{pre} is at 672 h, the t_{EOL} is at 795 h, and the t_{pEOL} is at 840 h. Thus, the t_{RUL}^{act} is 123 h, the t_{RUL}^{pre} is 168 h, and the error of RUL is 45 h. It is worth mentioning that the fault disturbance at 460 h would lead to the early prediction. In this condition (case 2), the t_{pre} is at 336 h, the t_{EOL} is at 435 h, and the t_{pEOL} is at 470 h. Thus, the t_{RUL}^{act} is 99 h, the t_{RUL}^{pre} is 134 h, and the error of RUL is 35 h (the error is 69 h in [64]).

For test-D2, the t_{pre} is at 672 h, the t_{EOL} is at 832 h, and the t_{pEOL} is at 840 h. Thus, the t_{RUL}^{act} is 160 h, the t_{RUL}^{pre} is 168 h, and the error of RUL is 8 h. Similar to test-QD, the maximum load current (170 A) in the first 2 stages (0 - 250 h) and the load cycling (50 A and 100 A) in stage 3 (350 h - 450 h) would also lead to an early prediction. In this case (case 2), the t_{pre} is at 168 h, the t_{EOL} is at 180 h, and the t_{pEOL} is at 310 h. Thus, the t_{RUL}^{act} is 12 h, the t_{RUL}^{pre} is 142 h, and the error of RUL is 130 h. The RUL results of different tests are shown in Tab. 5 - 3.

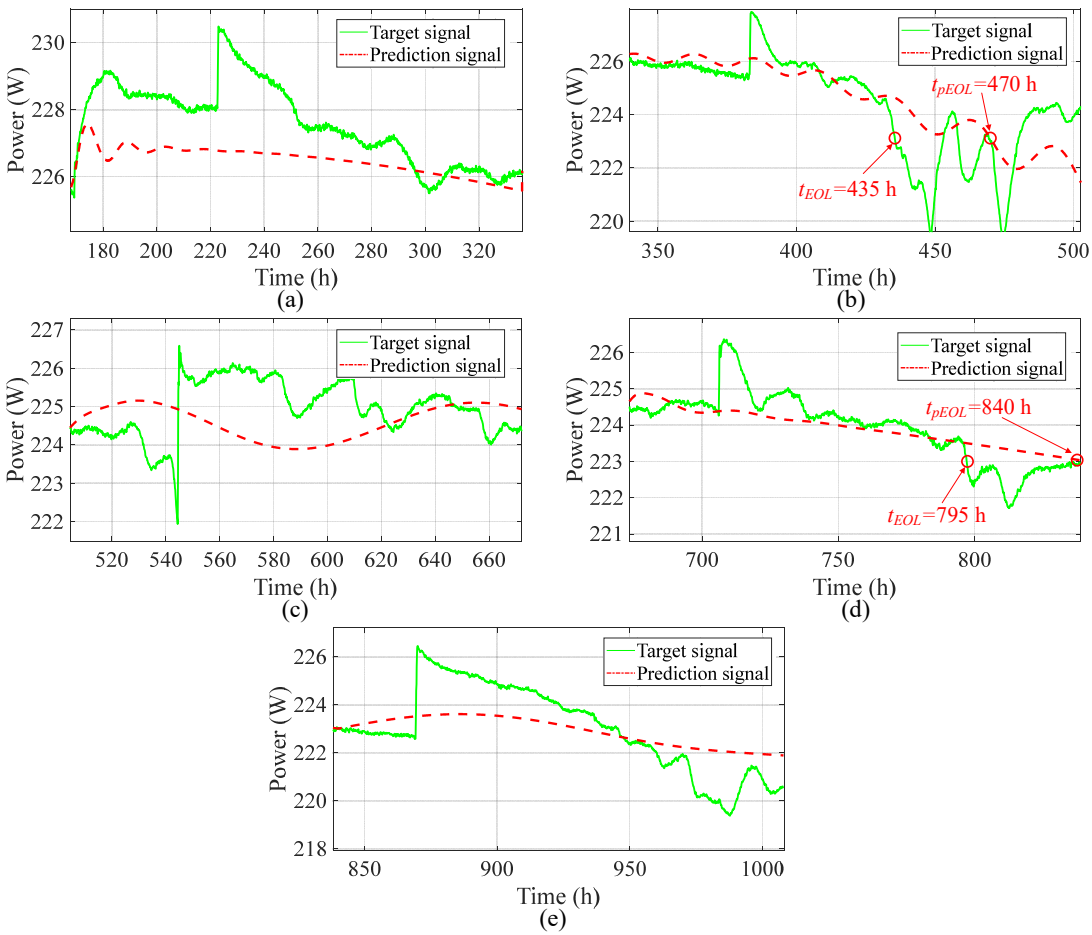


Fig. 5 - 12 The power prediction of test-QD: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

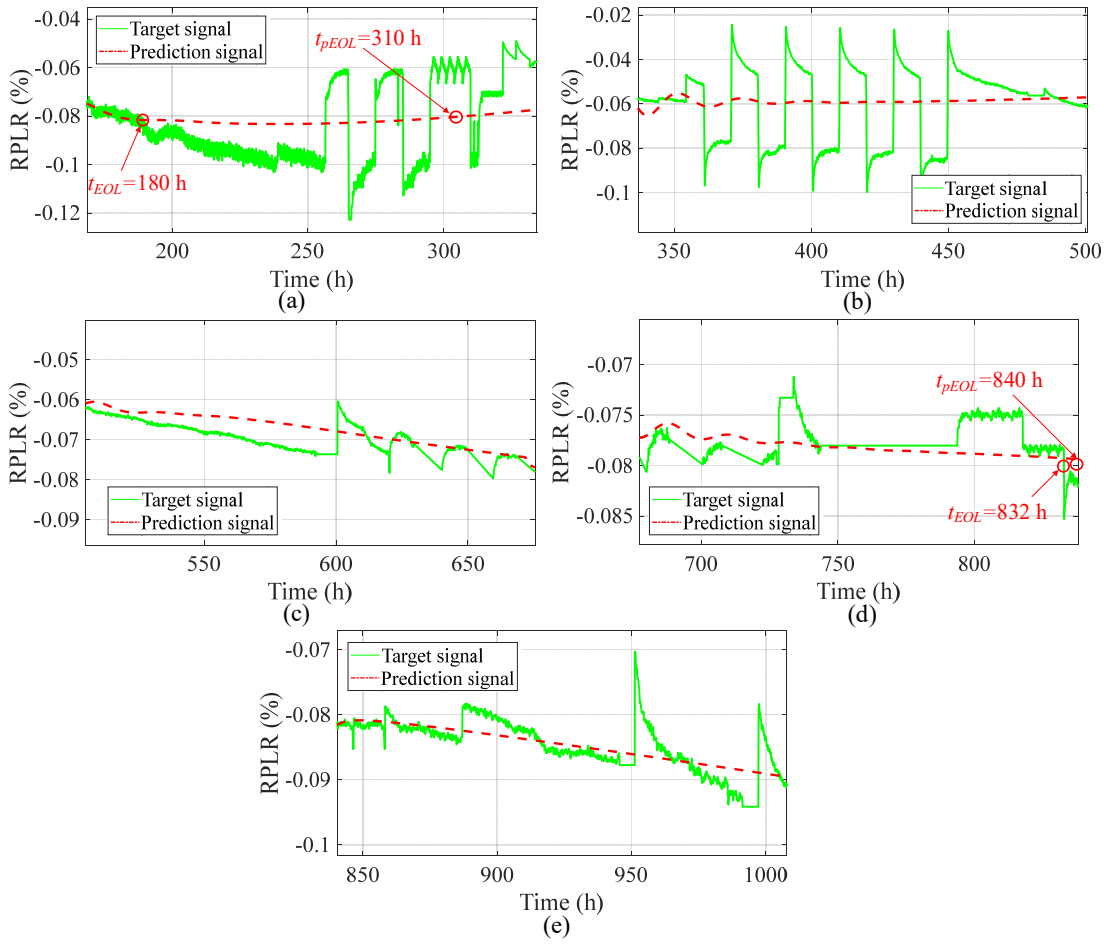


Fig. 5 - 13 The RPLR prediction of test-D2: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Tab. 5 - 3 The RUL of different tests

Tests	Cases	t_{pre}	t_{RUL}^{act}	t_{RUL}^{pre}	RUL error
Test-SS	--	840 h	15 h	70 h	55 h
Test-QD	Case 1	672 h	123 h	168 h	45 h
	Case 2	336 h	99 h	134 h	35 h
Test-D2	Case 1	672 h	160 h	168 h	8 h
	Case 2	168 h	12 h	142 h	130 h

Some model-based and data-driven methods have been performed under the steady-state and quasi-dynamic operating conditions, the RMSE comparison results of different methods are shown in Tab. 5 - 4 and Tab. 5 - 5. Both in the test-SS and test-QD, the proposed method of DWT-ESN-GA outperforms the four model-based methods of PR, ARIMA, DWT-PR, and DWT-ARIMA in all stages. In the steady-state

operating condition, the RMSE of DWT-ESN-GA is lower than PDM-PF at each stage. Compared with PDM-PF-NARNN and NARNN, the RMSE of DWT-ESN-GA are lower from stage 3 to stage 5, and they have a similar RMSE value in stage 6 which is mainly due to the recovery at 880 h. In a word, the proposed data-driven method has a satisfactory performance than the existing methods as a whole in terms of RMSE and MAPE.

Tab. 5 - 4 RMSE results comparison of test-SS.

Methods	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
PR	0.7	0.8	0.8	0.9	1.9
DWT-PR	0.5	0.8	0.7	0.7	1.6
ARIMA	0.6	0.8	0.6	1.1	1.1
DWT-ARIMA	0.5	0.4	0.6	0.7	1.1
PDM-PF	--	0.6184	0.6458	0.8026	0.8027
NARNN	--	0.5832	0.6377	0.8571	0.7523
PDM-PF-NARNN	--	0.4778	0.5145	0.6633	0.6265
DWT-ESN-GA	**0.3911**	**0.3928**	**0.3265**	**0.3960**	**0.7645**

Tab. 5 - 5 RMSE results comparison of test-QD.

Methods	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
PR	3.0	1.5	3.8	1.7	1.6
PR-DWT	1.4	1.6	1.2	1.3	1.3
ARIMA	1.5	1.5	1.5	1.8	2.1
ARIMA-DWT	1.3	1.6	1.2	1.8	1.9
DWT-ESN-GA	**1.3311**	**1.2337**	**0.8240**	**0.6060**	**1.2510**

5.5 Relative wavelet energy discussion

The DWT is used to shorten the length of the original data, and the prediction of power (or RPLR) is replaced by the prediction of its approximation components. Nevertheless, the high-frequency information would be lost during the data compression process. Relative wavelet energy (RWE) is used to quantitatively analyze the missing information in DWT. The approximation component and detail component wavelet energy at level j can be expressed as

$$\begin{cases} E_{j+1}^{cA} = \sum_{k} \left| C_{j+1}^{cA}(k) \right|^2 \\ E_{j+1}^{cD} = \sum_{k} \left| C_{j+1}^{cD}(k) \right|^2 \end{cases} \tag{5-3}$$

Where $C_{j+1}^{cD}(k)$ and $C_{j+1}^{cA}(k)$ are the reconstructed detail signal and approximation signal at time step k in level j. The total energy of the original signal is shown as

$$E_{sum} = E_{j+1}^{cA} + \sum_{j=0} E_{j+1}^{cD} \tag{5-4}$$

The RWE can be defined as the ratio between the dispersed energy and the total energy

$$\text{RWE} = \frac{E_{j+1}}{E_{tot}} * 100 \tag{5-5}$$

The RWE at different decomposition layers (j = 0, 1, 2) of these three tests is shown in Tab. 5 - 6.

Tab. 5 - 6 RWE at different decomposition layers.

	E_1^{cD}	E_2^{cD}	E_3^{cD}	E_3^{cA}
Test-SS	4.9×10^{-7}	9.0×10^{-7}	9.0×10^{-6}	≈100
Test-QD	4.8×10^{-7}	1.4×10^{-6}	4.3×10^{-6}	≈100
Test-D2	1.2×10^{-2}	2.2×10^{-2}	2.6×10^{-2}	≈100

The RWE of DC (E_{j+1}^{cD}) in all tests are small fractions of the total energy, and most of the energy is located in the approximation component. The RWE of the detail component increases with the decomposition level j. The missing information concentrate in the high-frequency part, and they are negligible during the DWT compression. Besides, the degradation timescale is in hundreds or thousands of hours. And the low-frequency part is more interesting in the lifespan prediction of the PEMFC system.

5.6 Chapter summary

The data-driven approach of DWT-ESN-GA is put forward to improve the lifespan validation efficiency of the proton exchange membrane fuel cells systems. Based on the technology of compression, prediction, and reconstruction, degradation characteristics in 2016 data points (corresponding to 168 h) can be effectively represented by only 253 data points. The health indicator of relative power-loss rate is proposed for the dynamic conditions to mitigate the impacts of mission profiles. The effectiveness of the proposed

method is validated under different operating conditions, and the improvements are significant when compared with other model-based or data-driven methods. In the steady-state operating condition, the maximum improvements of DWT-ESN-GA are 59.8 % (stage 6) and 64.0 % (stage 5) when compared with the model-based methods of PR and ARIMA in terms of root mean square error. The maximum improvements are 53.4 % (stage 4), 45.6 % (stage 4), 50.7 % (stage 5), and 40.3 % (stage 5) separately when compared with the model-based methods of DWT-PR, DWT-ARIMA, PDM-PF, and PDM-PF-NARNN. And the maximum improvements are 53.8 % (stage 5) when compared with the data-driven method of NARNN. In the quasi-dynamic operating condition, the maximum improvements of DWT-ESN-GA are 78.3 % (stage 4) and 66.3 % (stage 5) when compared with the model-based method of PR and ARIMA in terms of root mean square error. And the maximum improvements are 53.3 % (stage 5) and 66.3 % (stage 5) when compared with the model-based methods of DWT-PR and DWT-ARIMA. Besides, potential information loss in the compression process of discrete wavelet transform is fully discussed, and the discrete wavelet transform shorten principle is analyzed from the point of wavelet energy.

Chapter 6. Conclusions and perspectives

6.1 Conclusions

According to the problem of limited service life in the PEMFC system, this thesis is trying to realize the lifespan prediction under different operating conditions. Even though some researchers have done some meaningful work under the steady-state and quasi-dynamic operating conditions, just a few papers consider the lifespan prediction problems under the dynamic operating condition. On the one hand, the datasets under the dynamic operating condition are scarce because they are difficult to be measured. On the other hand, the traditional health indicator (e.g., voltage, power) cannot be used under the dynamic operating conditions anymore for the load current would affect them. Thus, proposing a new health indicator is fundamental to lifespan prediction under the dynamic operating condition. Besides, the effects of operating parameters in the RUL predictions have not been quantitatively analyzed to the authors' knowledge. During the RUL prediction, improving the prediction accuracy and prediction efficiency are helpful for the user to take some maintenance in advance.

To deal with the above issues, four aspects of work are developed. The first two aspects investigate the effects of operating parameters on the RUL prediction performance under the steady-state, quasi-dynamic, and full dynamic operating conditions separately. The third aspect is the works of how to deal with the multi-timescale features of the new health indicator and improve the prediction accuracy. The fourth aspect is the work of improving RUL prediction efficiency. The conclusions of this thesis are:

1) In the steady-state and quasi-dynamic operating conditions, a multi-input and multi-output ESN (MIMO-ESN) structure is developed to analyze the effects of operating parameters in the RUL predictions. Besides the commonly used stack voltage, the operating parameters, such as stack current, stack temperature, and the pressures of the reactants are also utilized as the inputs of ESN. Experimental results show that MIMO-ESN can improve the prediction accuracy, especially the 2-input ESN with an improvement of 59.20 % (static condition) and 34.15 % (quasi-dynamic condition) in terms of $RMSE_5$ respectively compared with the traditionally used single-input and single-output ESN (SISO-ESN).

2) Based on the polarization curve at the BoL and the variables' (voltage and current) continuous measurement, a dynamic health indicator named relative power-loss rate (RPLR) is proposed in this aspect. This dynamic health indicator can be used in practice due to its convenient extraction. Besides, the normal operation of the system would not be interrupted during the indicator's extraction process. Based on the results of the first aspect, the load current could reflect the system's dynamic, and combining

it with the RPLR could increase the dynamic property of ESN. Therefore, double-input ESN is utilized to enhance the prediction performance of lifespan under dynamic operating conditions. In general, the value of $\%Er_{FT}$, RMSE, and MAPE of double-input ESN are smaller than those of single-input ESN. For example, in test-D1, the RMSE at 70 % training length of single-input ESN is 0.00576, the RMSE at 70 % training length of double-input ESN is 0.00411. In test-D2, the RMSE at 70 % training length of single-input ESN is 0.02422, the RMSE at 70 % training length of double-input ESN is 0.01197. In test-D3, the RMSE at 70 % training length of single-input ESN is 0.01407, the RMSE at 70 % training length of double-input ESN is 0.00976. In test-D1, the satisfactory horizon of single-input ESN is 117 h and the satisfactory horizon of double-input ESN is 157 h. In test-D2, the satisfactory horizon of single-input ESN is 150 h and the satisfactory horizon of double-input ESN is 250 h. In test-D3, the satisfactory horizon of single-input is 33 h and the satisfactory horizon of double-input is 63 h. The satisfactory horizon of double-input ESN is longer than that of single-input ESN in each dynamic test, and thus more time is given to the user to anticipate maintenance actions before failure.

3) The discrete wavelet transform and ensemble echo state network (DWT-EESN) approach is proposed to deal with the multi-timescale features of RPLR and improve the long-term prediction performance. The degradation phenomena of the PEMFC are a complex process, and the dynamic health indicator of RPLR contains multi-timescale features. The discrete wavelet transform is used to decompose the RPLR into different timescales. Then different ESN with varying dynamic characteristics are used to deal with different features separately at the same time. Compared with the single ESN without decomposition process, the DWT-EESN approach could improve the prediction accuracy under all these three dynamic tests. Properly speaking, in terms of RMSE, the maximum improvements of test-D1, test-D2, and test-D3 are 23.50 % (70 % training length), 54.52 % (80 % training length), and 85.21 % (40 % training length) separately. In terms of MAPE, the maximum improvements of test-D1, test-D2, and test-D3 are 23.59 % (70 % training length), 54.94 % (80 % training length), and 66.46 % (40 % training length) separately.

4) The approach of discrete wavelet transform-echo state network-genetic algorithm (DWT-ESN-GA) is put forward to improve the lifespan validation efficiency of the PEMFC systems. In the previous work, the parameters are optimized by the trial and error method or grid-search method. These two methods are time-consuming especially for the multiple parameters of the ESN. To improve the optimization efficiency, the genetic algorithm is used to optimize the leaking rate a, the spectral radius ρ of the internal weight matrix, and the regression coefficient β. Results also show that the genetic algorithm is more efficient than the grid-search method. Based on the compression by DWT and the reconstruction by IDWT, the data's prediction in the

coming 2016 data points can be shortened into 253 data points. The validation effectiveness of the RUL prediction can be improved by this compression, prediction, and reconstruction process. Finally, the proposed method is validated under different operating conditions, and the improvements are significant when compared with other model-based or data-driven methods. For example, in test-SS, the maximum improvements are 53.4 % and 45.6 % separately when compared with the model-based methods of DWT-PR and DWT-ARIMA. In test-QD, the maximum improvements are 53.3 % and 66.3 % when compared with the model-based methods of DWT-PR and DWT-ARIMA.

6.2 Perspectives

Our works concentrate on analyzing the effects of operating parameters, proposing the new health indicators, dealing with the multi-timescale features, and improving the prediction efficiency during the RUL prediction. Even though we've been working on the above questions for about three years, there is still a long way to go on further. From my point of view, four aspects of work can be further developed on the RUL prediction of the PEMFC system.

1) The quantitative analysis of the uncertainties during the lifetime prediction. The filtering technique is usually used to pre-process the data (e.g., voltage), however, setting the filtering coefficient is a difficult task in practice. A too-large filtering window cannot filter the peaks efficiently and a too-small filtering window may lead to loss of some important information. It is a pity that there is no standard for the user to set the filtering coefficient because different users have different purposes. Thus, some other useless information would also exist in the voltage data after the filtering technique. Besides, the stochastic disturbances produced by the sensor noises during the measurements, unknown disturbances from the environment and operating parameters, load irregular varying caused by user habits, and the uncertainties of the prognostic methods themselves would also affect the lifespan prediction results. Quantitative analysis of the uncertainties and extracting the degradation-related information can make sure prediction accuracy. Different fault signals would also be measured along with the degradation data. Using the diagnostic technique to separate the fault signal in advance is necessary for the prognostic.

2) Finding more and more convenient and efficient dynamic health indicators. The degradation state of the PEMFC stack can be reflected by the health indicators, and an accurate health indicator is helpful for the users to take some preventive maintenance in advance and further extend the lifespan of the PEMFC stack. In general, the voltage and power can be regarded as health indicators under the steady-state operating condition. In this condition, the mission profile doesn’t change a lot, and the load current keeps at a constant value or changes in a small range. Thus, the voltage and

power would have a monotonic decreasing tendency which is mainly caused by the degradation. Nevertheless, the monotonic decreasing tendency of voltage or power would become less obvious because of the load current's large changing under the dynamic operating conditions. Thus, the traditional static health indicators cannot be used anymore. After reviewing the literature, different dynamic health indicators have been proposed, such as the "EKF-based degradation factor (EKF-DF)" and "virtual stack voltage (VSV)". Based on the detailed comparison, a dynamic health indicator named RPLR has been proposed in this thesis, and its effects have been validated on three dynamic datasets. The advantages and shortcomings of them have been analyzed in the body of the thesis. Nevertheless, exploring the health indicators is always the research focus for the RUL prediction especially under the dynamic operating condition. More and more other dynamic health indicators should be explored in the practical condition.

3) Exploration of the methods which can provide the confidence intervals (CI) and probability density function (PDF) would also be welcomed in the next step. Most of the current articles focus on the single-step or single curve prediction. The single-step prediction means the historical value in step t has already been known when predicting the value of step (t+1). This prediction process is realized in the short term, and the prediction horizon is the time interval of every two data points. Even if the error between the target value and predicted value seems small enough, the single-step prediction does not have much practice meanings for the RUL prediction because the degradation phenomena are in the hundreds of hours. Using all the historical data to realize only single-step prediction cannot reserve enough time for the users to take some maintenances in advance. The prediction horizon can be extended when combining the single-step prediction and the iterative process. The predicted value at step (t+1) is regarded as one of the inputs to predict the new value at step (t+2), and this process is the single curve prediction. Each prediction curve means one combination of results of the parameters no matter for the model-based or data-driven method. Thus, the parameter's effects should be considered during the prediction, and providing the prediction curves along with their confidence intervals and PDF can give more intuitive results for the decision-makers. Compared with the single curve prediction, the methods with the confidence intervals and PDF could improve the boundary from the probability point of view. Combining the ESN with particle filter or Gaussian process may be one of the solutions.

4) Exploring the prognostic method to realize the online prediction of RUL. Online prediction has more practical meanings than offline, and there are only a few articles that have done the related works as far as the author knows. The common idea of online prediction is building and training the mathematical models or "black box" by the historical data. And then the trained model or structures are used to realize the

prediction online. As mentioned before, the wavelet-based approach and some models (polynomial regression and ARIMA) are combined to realize the online prediction. For the data-driven method, the complex neuron network structure makes the online process more difficult to implement. The multiple parameters have also limited the online implementation process, and some online parameter optimization methods would be helpful. Thus, how to realize the online RUL prediction is worthy of study in the next step.

List of figures

Fig. I Proportion of different types of energy consumption on the earth in the next 30 years [1].
Fig. II The total shipments and Megawatts (in MW) by applications and by fuel cell types [6].
Fig. III Costs of different components in the PEMFC stack [11].
Fig. IV Fuel cell electric vehicle evaluation project partners [13].
Fig. V Summary of key FCEV metrics versus U.S.DOE targets (system level).
Fig. VI Research route of this thesis.
Fig. 1 - 1 The PEMFC system: (a) schematic of the system, (b) structure of fuel cell stack and single-cell, and (c) working principle of fuel cell [19].
Fig. 1 - 2 The *i* - *U* curve of a PEMFC system.
Fig. 1 - 3 The *i* - *U* and power density curves of a PEMFC system.
Fig. 1 - 4 The stages of the prognostic process [58].
Fig. 1 - 5 Current waveforms used for the durability test of FC1 and FC2.
Fig. 1 - 6 Polarization curves: (a) FC1, (b) FC2.
Fig. 1 - 7 EIS tests at different time: (a) FC1, (b) FC2.
Fig. 1 - 8 The mission profiles of two long-term tests: (a) stack current of test-SS, (b) stack voltage of test-SS, (c) stack current of test-QD, (d) stack voltage of test-QD.
Fig. 1 - 9 The mission profiles of test-D1: (a) stack current of test-D1, (b) stack voltage of test-D1.
Fig. 1 - 10 The mission profiles of test-D2: (a) stack current of test-D2, (b) stack voltage of test-D2.
Fig. 1 - 11 The mission profiles of test-D3: (a) stack current of test-D3, (b) stack voltage of test-D3.
Fig. 1 - 12 The EIS tests and polarization curves of test-D1: (a) EIS tests of test-D1, (b) polarization curves of test-D1.
Fig. 1 - 13 The EIS tests and polarization curves of test-D2: (a) EIS tests of test-D2, (b) polarization curves of test-D2.
Fig. 1 - 14 The EIS tests and polarization curves of test-D3: (a) EIS tests of test-D3, (b) polarization curves of test-D3.
Fig. 1 - 15 Time scales in a PEMFC system [58].
Fig. 1 - 16 The output stack voltages of Data Challenge: (a) FC1, (b) FC2.
Fig. 1 - 17 The RUL prediction methods of the PEMFC system.
Fig. 1 - 18 The types of models with filtering methods.
Fig. 1 - 19 The types of models with ML methods.
Fig. 1 - 20 The types of data-driven methods.

Fig. 1 - 21 Different network topologies of ANN: (a) FFNN, (b) RNN [94].
Fig. 1 - 22 The typical ESN structure in the work of Morando *et al*. and Mezzi *et al*.
Fig. 1 - 24 The MIMO-ESN (2 inputs and 1 output) prediction structure.
Fig. 1 - 25 The output stack voltages of FCEV and *μ-CHP*: (a) dataset in [103], (b) dataset in [40].
Fig. 1 - 26 Mid-term prediction structure by ESN.
Fig. 1 - 27 Long-term prediction structure by ESN.
Fig. 2 - 1 The basic representation of echo state network: (a) ESN structure, (b) the updating state in the reservoir at time step *n*.
Fig. 2 - 2 Implementation framework of the ESN method.
Fig. 2 - 3 Operating parameters of FC1.
Fig. 2 - 4 Operating parameters of FC2.
Fig. 2 - 5 Calculation flow chart of the MIMO-ESN prediction process.
Fig. 2 - 6 Block diagram of single-input and multi-input ESN for one example with 2 steps in each case (where *yi* represents Us at time *ti*, *zi* represents the Is at time *ti*, *xi* represents PoutH2 at time *ti*, $\hat{y}_{s+1}$ and $\hat{y}_{s+2}$ are next-step predicted stack voltage value, $\mathbf{Wout}_{(s)}$ represents the output weight matrix).
Fig. 2 - 7 The RUL prediction of FC1 based on single-input ESN.
Fig. 2 - 8 The RUL prediction of FC1 based on 2-input ESN: (a) stack voltage (Us) and outlet temperature of air (ToutAir). (b) stack voltage (Us) and stack current (Is). (c) stack voltage (Us) and inlet temperature of cooling water (TinWat). (d) stack voltage (Us) and outlet pressure of H2 (PoutH2).
Fig. 2 - 9 The RUL prediction of FC1 based on 3-input ESN: (a) stack voltage (Us), inlet temperature of H2 (TinH2) and inlet pressure of H2 (PinH2). (b) stack voltage (Us), outlet temperature of H2 (ToutH2) and outlet temperature of air (ToutAir). (c) stack voltage (Us), inlet temperature of air (TinAir) and outlet temperature of cooling water (ToutWat). (d) stack voltage (Us), outlet pressure of H2 (PoutH2), and inlet pressure of H2 (PinH2).
Fig. 2 - 10 The RUL prediction of FC2 based on single-input ESN.
Fig. 2 - 11 The RUL prediction of FC2 based on 2-input ESN: (a) stack voltage (Us) and outlet temperature of air (ToutAir). (b) stack voltage (Us) and stack current (Is). (c) stack voltage (Us) and inlet temperature of water (TinWat). (d) stack voltage (Us) and outlet pressure of H2 (PoutH2).
Fig. 2 - 12 The RUL prediction of FC2 based on 3-input ESN: (a) stack voltage (Us), inlet temperature of H2 (TinH2) and inlet pressure of H2 (PinH2). (b) stack voltage (Us), outlet temperature of H2 (ToutH2) and outlet temperature of air (ToutAir). (c) stack voltage (Us), inlet temperature of air (TinAir) and outlet pressure of air (PoutAir). (d) stack voltage (Us), inlet temperature of water (TinWat), and inlet pressure of H2 (PinH2).

Fig. 3 - 1 The calculation flow chart of the RPLR.
Fig. 3 - 2 The polarization curves and the BoL power curves of three long-term tests: (a) polarization curves during the whole life of test-D1, (b) polarization curve at the BoL of test-D1, (c) BoL power curve of test-D1, (d) polarization curves in the whole life of test-D2, (e) polarization curve at the BoL of test-D2, (f) BoL power curve of test-D2, (g) polarization curves in the whole life of test-D3, (h) polarization curve at the BoL of test-D3, (i) BoL power curve of test-D3.
Fig. 3 - 3 The RPLR and RUL definition of 3 dynamic tests: (a) test-D1, (b) test-D2, (c) test-D3.
Fig. 3 - 4 Operating parameters of test-D1.
Fig. 3 - 5 Operating parameters of test-D2.
Fig. 3 - 6 Operating parameters of test-D3.
Fig. 3 - 7 The iterative process of SI-ESN and DI-ESN.
Fig. 3 - 8 Target statement of RUL and its prediction process.
Fig. 3 - 9 The RUL prediction of test-D1: (a) SI-ESN with 60 % training, (b) DI-ESN with 60 % training, (c) SI-ESN with 70 % training, (d) DI-ESN with 70 % training, (e) SI-ESN with 80 % training, (f) DI-ESN with 80 % training.
Fig. 3 - 10 The RUL results of SI-ESN and DI-ESN with 95 % probability bounds in test-D1.
Fig. 3 - 11 The RUL prediction of test-D2: (a) SI-ESN with 40 % training, (b) DI-ESN with 40 % training, (c) SI-ESN with 60 % training, (d) DI-ESN with 60 % training, (e) SI-ESN with 80 % training, (f) DI-ESN with 80 % training.
Fig. 3 - 12 The RUL results of SI-ESN and DI-ESN with 95 % probability bounds in test-D2.
Fig. 3 - 13 The RUL prediction of test-D3: (a) SI-ESN with 50 % training, (b) DI-ESN with 50 % training, (c) SI-ESN with 65 % training, (d) DI-ESN with 65 % training, (e) SI-ESN with 75 % training, (f) DI-ESN with 75 % training.
Fig. 3 - 14 The RUL results of SI-ESN and DI-ESN with 95% probability bounds in test-D3.
Fig. 4 - 1 Data processing and implementation process of the ESN: (a) data processing, (b) implementation process of ESN.
Fig. 4 - 2 The flowchart of DWT-EESN.
Fig. 4 - 3 The RPLR prediction with single ESN in test-D1: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 4 - 4 The RPLR prediction with single ESN in test-D2: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 4 - 5 The RPLR prediction with single ESN in test-D3: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 4 - 6 Sub-waveforms prediction results of test-D1 in 50 % training length: (a) *D1,*

(b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *A5.*
Fig. 4 - 7 Sub-waveforms prediction results of test-D2 in 50 % training length: (a) *D1,* (b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *D6,* (g) *D7,* (h) *A7.*
Fig. 4 - 8 Sub-waveforms prediction results of test-D3 in 50 % training length: (a) *D1,* (b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *D6,* (g) *A6.*
Fig. 4 - 9 Sub-waveforms prediction results of test-D1 in 60 % training length: (a) *D1,* (b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *A5.*
Fig. 4 - 10 Sub-waveforms prediction results of test-D2 in 60 % training length: (a) *D1,* (b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *D6,* (g) *D7,* (h) *A7.*
Fig. 4 - 11 Sub-waveforms prediction results of test-D3 in 60 % training length: (a) *D1,* (b) *D2,* (c) *D3,* (d) *D4,* (e) *D5,* (f) *D6,* (g) *A6.*
Fig. 4 - 12 The RPLR prediction with DWT-EESN in test-D1: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 4 - 13 The RPLR prediction with DWT-EESN in test-D2: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 4 - 14 The RPLR prediction with DWT-EESN in test-D3: (a) in 40 % training length, (b) in 50 % training length, (c) in 60 % training length, (d) in 70 % training length.
Fig. 5 - 1 The waveforms of Daubechies wavelets family [141].
Fig. 5 - 2 Health indicators of three tests: (a) power of test-SS, (b) power of test-QD, (c) RPLR of test-D2.
Fig. 5 - 3 The wavelet transform of test-SS: (a) decomposition and reconstruction process of the wavelet transform, (b) original power data in stage 1, (c) approximation component coefficient in level 1, (d) detail component coefficient in level 1.
Fig. 5 - 4 The approximation component coefficients and detail component coefficients of level 2 and level 3 in stage 1: (a) *cA2,* (b) *cD2,* (c) *cA3,* (d) *cD3.*
Fig. 5 - 5 Implementation structure of decomposition and reconstruction: (a) the *cA3* prediction in the first 2 stages, (b) the reconstructed power in the first 2 stages, (c) the implementation structure of DWT-ESN-GA.
Fig. 5 - 6 The flowchart of DWT-ESN-GA.
Fig. 5 - 7 The approximation component coefficients prediction of test-SS at level 3: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.
Fig. 5 - 8 The approximation component coefficients prediction of test-QD at level 3: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.
Fig. 5 - 9 The approximation component coefficients prediction of test-D2 at level 3:

(a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Fig. 5 - 10 The degradation predictions at different stages under different operating conditions: (a) test-SS, (b) test-QD, (c) test-D2.

Fig. 5 - 11 The power prediction of test-SS: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Fig. 5 - 12 The power prediction of test-QD: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

Fig. 5 - 13 The RPLR prediction of test-D2: (a) prediction at stage 2, (b) prediction at stage 3, (c) prediction at stage 4, (d) prediction at stage 5, (e) prediction at stage 6.

List of tables

Tab. I The classification of fuel cells and their properties.
Tab. II Statistics about fuel cell light-duty vehicles (including cars) from FCH 2 JU [10].
Tab. III Costs (unit: €/kW) of fuel cell systems in different applications.
Tab. IV Durability (unit: h) of fuel cell system in different applications.
Tab. 1 - 1 Degradation reasons of some components.
Tab. 1 - 2 Physical and operating parameters of the test bench.
Tab. 1 - 3 Nominal specifications and working parameters of the experimental platform.
Tab. 1 - 4 Classification of model-based methods (models with filtering approaches).
Tab. 1 - 5 Classification of model-based methods (models with ML methods).
Tab. 1 - 6 Classification of data-driven methods.
Tab. 1 - 7 The application of ESN in degradation prediction.
Tab. 1 - 8 The classification of HIs.
Tab. 2 - 1 Key parameter of ESN.
Tab. 2 - 2 The SNR of test-SS and test-QD.
Tab. 2 - 3 Prediction results of FC1 based on 2-input ESN.
Tab. 2 - 4 Prediction results of FC1 based on 3-input ESN.
Tab. 2 - 5 Prediction results of FC2 based on 2-input ESN.
Tab. 2 - 6 Prediction results of FC2 based on 3-input ESN.
Tab. 2 - 7 The improvement results of FC1 based on 2-input ESN.
Tab. 2 - 8 The improvement results of FC2 based on 2-input ESN.
Tab. 2 - 9 The prediction results based on all 3-input combinations of FC1.
Tab. 2 - 10 The prediction results based on all 3-input combinations of FC2.
Tab. 3 - 1 Parameters of three mathematical models.
Tab. 3 - 2 Prediction results comparison of test-D1.
Tab. 3 - 3 Prediction results comparison of test-D2.
Tab. 3 - 4 Prediction results comparison of test-D3.
Tab. 3 - 5 Prediction results of different combinations.
Tab. 3 - 6 Influence of different parameters.
Tab. 3 - 7 Parameter levels and the typical values in different tests.
Tab. 4 - 1 Prediction results of single ESN.
Tab. 4 - 2 Prediction results of DWT-EESN.
Tab. 5 - 1 RMSE results of three tests in different stages.
Tab. 5 - 2 MAPE results of three tests in different stages.
Tab. 5 - 3 The RUL of different tests
Tab. 5 - 4 RMSE results comparison of test-SS.
Tab. 5 - 5 RMSE results comparison of test-QD.

Tab. 5 - 6 RWE at different decomposition layers.

References

[1] China national petroleum corporation (CNPC), "2050 World and China energy outlook," 2019. https://pdf.dfcfw.com/pdf/H3_AP201912151371851724_1.pdf?1576501894000.pdf

[2] J. M. Andújar and F. Segura, "Fuel cells: history and updating. a walk along two centuries," Renewable and Sustainable Energy Reviews, vol. 13, no.9, pp. 2309-2322, Dec. 2009.

[3] O. Z. Sharaf and M. F. Orhan, "An overview of fuel cell technology: fundamentals and applications," Renewable and Sustainable Energy Reviews, vol. 32, pp. 810-853, Apr. 2014.

[4] Y. Wang, K. S. Chen, J. Mishler, S. C. Cho, and X. C. Adroher, "A review of polymer electrolyte membrane fuel cells: technology, applications, and needs on fundamental research," Applied Energy, vol. 88, pp. 981-1007, Apr. 2011.

[5] R. O'hayre, S.-W. Cha, W. G. Colella, and F. B. Prinz, Fuel cell fundamentals: John Wiley & Sons, 2016.

[6] D. Hart, S. Jones, and J. Lewis, "The fuel cell industry review 2020," E4tech, Lausanne, Switzerland, Mar. 2021. https://fuelcellindustryreview.com/

[7] P. Pei and H. Chen, "Main factors affecting the lifetime of proton exchange membrane fuel cells in vehicle applications: A review," Applied Energy, vol. 125, pp. 60-75, Jul. 2014.

[8] R. Petrone, D. Hissel, M.-C. Péra, D. Chamagne, and R. Gouriveau, "Accelerated stress test procedures for PEM fuel cells under actual load constraints: State-of-art and proposals," International Journal of Hydrogen Energy, vol. 40, pp. 12489-12505, Sep. 2015.

[9] M. Mohammadi Taghiabadi and M. Zhiani, "Degradation analysis of dead-ended anode PEM fuel cell at the low and high thermal and pressure conditions," International Journal of Hydrogen Energy, vol. 44, no. 10, pp. 4985-4995, Feb. 2019.

[10] Europe fuel cells and hydrogen joint undertaking (FCH 2 JU), " State-of-the-art and future targets (KPIS), fuel cell light duty vehicles (including cars)," Multi - Annual Work Plan, 2014-2020. https://www.fch.europa.eu/soa-and-targets

[11] A. Wilson, G. Kleen, and D. Papageorgopoulos, "DOE hydrogen and fuel cells program record," U.S. DOE, Washington DC, United States, Tech. Rep., no. 17007, Sep. 2017.

[12] Hydrogen and Fuel Cell Technologies Office of U.S. DOE, "Comparison of fuel cell technologie", Washington DC, United States. Available online: https://www.energy.gov/eere/fuelcells/fuel-cells

[13] J. Kurtz, S. Sprik, G. Saur, and S. Onorato, "On-road fuel cell electric vehicles evaluation: overview," Nat. Renew. Energy Lab. (NREL), Golden, CO, United States, Tech. Rep., no. NREL/TP-5400-73009, Mar. 2019.

[14] 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, vol. 75, pp. 440-450, Aug. 2017.

[15] J. Zhao and X. Li, "A review of polymer electrolyte membrane fuel cell durability for vehicular applications: Degradation modes and experimental techniques," Energy Conversion and Management, vol. 199, p. 112022, Nov. 2019.

[16] I. Radev, K. Koutzarov, E. Lefterova, and G. Tsotridis, "Influence of failure modes on PEFC stack and single cell performance and durability," International Journal of Hydrogen Energy, vol. 38, no. 17, pp. 7133-7139, Jun. 2013.

[17] X. Zhang, Y. Rui, Z. Tong, X. Sichuan, S. Yong, and N. Huaisheng, "The characteristics of voltage degradation of a proton exchange membrane fuel cell under a road operating environment," International Journal of Hydrogen Energy, vol. 39, no. 17, pp. 9420-9429, Jun. 2014.

[18] P. Gazdzick, J. Mitzel, D. Garcia Sanchez, M. Schulze, and K. A. Friedrich, "Evaluation of reversible and irreversible degradation rates of polymer electrolyte membrane fuel cells tested in automotive conditions," Journal of Power Sources, vol. 327, pp. 86-95, Sep. 2016.

[19] Z. Hua, Z. Zheng, F. Gao, and M.-C. Péra, "Challenges of the remaining useful life prediction for proton exchange membrane fuel cells " in IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, pp. 6382-6387, Oct. 2019.

[20] D. Zhao, Z. Hua, M. Dou, and Y. Huangfu, "Control oriented modeling and analysis of centrifugal compressor working characteristic at variable altitude," Aerospace Science and Technology, vol. 72, pp. 174-182, Jan. 2018.

[21] V. Das, S. Padmanaban, K. Venkitusamy, R. Selvamuthukumaran, F. Blaabjerg, and P. Siano, "Recent advances and challenges of fuel cell based power system architectures and control – A review," Renewable and Sustainable Energy Reviews, vol. 73, pp. 10-18, Jun. 2017.

[22] Y. Zhou, A. Ravey, and M.-C. Péra, "A survey on driving prediction techniques for predictive energy management of plug-in hybrid electric vehicles," Journal of Power Sources, vol. 412, pp. 480-495, Feb. 2019.

[23] ISO13381-1. Condition monitoring and diagnostics of machines - prognostics - part1: general guidelines. International Standard, ISO; 2004.

[24] R. Onanena, L. Oukhellou, D. Candusso, F. Harel, D. Hissel, and P. Aknin, "Fuel cells static and dynamic characterizations as tools for the estimation of their ageing time," International Journal of Hydrogen Energy, vol. 36, no. 2, pp. 1730-1739, Jan. 2011.

[25] W. O. L. Vianna, I. P. d. Medeiros, B. S. Aflalo, L. R. Rodrigues, and J. P. P. Malère, "Proton Exchange Membrane Fuel Cells (PEMFC) impedance estimation using regression analysis," in 2014 International Conference on Prognostics and Health Management, 2014, pp. 1-8.

[26] K. Taejin, K. Hyunjae, H. Jongmoon, K. Keunsu, Y. Jungtaek, J. Joonha, et al., "A degenerated equivalent circuit model and hybrid prediction for state-of-health (SOH) of PEM fuel cell," in 2014 International Conference on Prognostics and Health Management, 2014, pp. 1-7.

[27] R. Gouriveau, M. Hilairet, D. Hissel, S. Jemei, M. Jouin, E. Lechartier, et al., "IEEE phm 2014 data challenge: outline, experiments, scoring of results, winners, " tech. rep., IEEE 2014 PHM Challenge, 2014.

[28] E. Lechartier, E. Laffly, M.-C. Péra, R. Gouriveau, D. Hissel, and N. Zerhouni, "Proton exchange membrane fuel cell behavioral model suitable for prognostics," International Journal of Hydrogen Energy, vol. 40, no. 26, pp. 8384-8397, Jul. 2015.

[29] C. Robin, M. Gerard, A. A. Franco, and P. Schott, "Multi-scale coupling between two dynamical models for PEMFC aging prediction," International Journal of Hydrogen Energy, vol. 38, no. 11, pp. 4675-4688, Apr. 2013.

[30] P. Polverino and C. Pianese, "Model-based prognostic algorithm for online RUL estimation of PEMFCs," in 2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol), 2016, pp. 599-604.

[31] H. Chen, P. Pei, and M. Song, "Lifetime prediction and the economic lifetime of Proton Exchange Membrane fuel cells," Applied Energy, vol. 142, pp. 154-163, Mar. 2015.

[32] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Degradations analysis and aging modeling for health assessment and prognostics of PEMFC," Reliability Engineering & System Safety, vol. 148, pp. 78-95, Apr. 2016.

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

[34] X. Zhang, P. Pisu, "An unscented Kalman filter based approach for the health-monitoring and prognostics of a electrolyte membrane fuel cell polymer," in 2012 Annual Conference of Prognostics and Health Management Society, pp. 1-9, Jan. 2012.

[35] H. Liu, J. Chen, C. Zhu, H. Su, and M. Hou, "Prognostics of proton exchange membrane fuel cells using a model-based method," IFAC-PapersOnLine, vol. 50, pp. 4757-4762, Jul. 2017.

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

[37] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Fuel cells remaining useful life estimation using an extended Kalman Filter," in IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, 2015, pp. 469-474.

[38] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Extended Kalman filter for prognostic of proton exchange membrane fuel cell," Applied Energy, vol. 164, pp. 220-227, Feb. 2016.

[39] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Fuel cell remaining useful life prediction and uncertainty quantification under an automotive profile," in IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, 2016, pp. 5477-5482.

[40] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Remaining useful life prediction and uncertainty quantification of proton exchange membrane fuel cell under variable load," IEEE Transactions on Industrial Electronics, vol. 63, pp. 2569-2577, Apr. 2016.

[41] 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), 2017, pp. 1-6.

[42] M. S. Jha, M. Bressel, B. Ould-Bouamama, and G. Dauphin-Tanguy, "Particle filter based hybrid prognostics of proton exchange membrane fuel cell in bond graph framework," Computers & Chemical Engineering, vol. 95, pp. 216-230, Nov. 2016.

[43] M. Bressel, D. Hissel, B. Bouamama, and M. Hilairet, Aging Tolerant Control of Proton Exchange Membrane Fuel Cell: a Model-Based Approach, 2017.

[44] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, "Model-based aging tolerant control with power loss prediction of Proton Exchange Membrane Fuel Cell," International Journal of Hydrogen Energy, Nov. 2018.

[45] E. Pahon, S. Morando, R. Petrone, M.-C. Péra, D. Hissel, N. Y. Steiner, et al., "Long-term tests duration reduction for PEMFC μ-CHP application," International Journal of Hydrogen Energy, vol. 42, pp. 1527-1533, Jan. 2017.

[46] 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, pp. 1-1, 2021.

[47] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, Dynamical modeling of Proton Exchange Membrane Fuel Cell and parameters identification, 2015.

[48] 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, vol. 39, pp. 481-494, Jan. 2014.

[49] M. Jouin, R. Gouriveau, D. Hissel, M. Péra, and N. Zerhouni, "Prognostics of proton exchange membrane fuel cell stack in a particle filtering framework including characterization disturbances and voltage recovery," in 2014 International Conference on Prognostics and Health Management, 2014, pp. 1-6.

[50] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Remaining useful life estimates of a PEM fuel cell stack by including characterization-induced disturbances in a particle filter model," in Conference Internationale Discussion on Hydrogen Energy and Applications, IDHEA'14., France, 2014, pp. 1-10.

[51] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Joint particle filters prognostics for proton exchange membrane fuel cell power prediction at constant current solicitation," IEEE Transactions on Reliability, vol. 65, pp. 336-349, Mar. 2016.

[52] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "PEMFC aging modeling for prognostics and health assessment," IFAC-PapersOnLine, vol. 48, pp. 790-795, Jan. 2015.

[53] M. Jouin, R. Gouriveau, D. Hissel, M. Cécile Péra, and N. Zerhouni, "Combined predictions for prognostics and predictive control of transportation PEMFC," IFAC-PapersOnLine, vol. 49, pp. 244-249, Jan. 2016.

[54] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Prognostics of PEM fuel cells under a combined heat and power profile," IFAC-PapersOnLine, vol. 48, no. 3, pp. 26-31, Jan. 2015.

[55] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "PHM of proton-exchange membrane fuel cells - a review," Chemical Engineering Transactions, vol. 33, pp. 1009-, Dec. 2013.

[56] 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, vol. 38, pp. 15307-15317, Nov. 2013.

[57] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Particle filter-based prognostics: Review, discussion and perspectives," Mechanical Systems and Signal Processing, vol. 72-73, pp. 2-31, May. 2016.

[58] M. Jouin, M. Bressel, S. Morando, R. Gouriveau, D. Hissel, M.-C. Péra, et al., "Estimating the end-of-life of PEM fuel cells: guidelines and metrics," Applied Energy, vol. 177, pp. 87-97, Sep. 2016.

[59] 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, pp. 1-6, 2014.

[60] D. Zhang, C. Cadet, N. Yousfi-Steiner, and C. Bérenguer, "Proton exchange membrane fuel cell remaining useful life prognostics considering degradation recovery phenomena," Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 232, no. 4, pp. 415 - 424, May. 2018.

[61] D. Zhang, C. Cadet, C. Bérenguer, and N. Yousfi-Steiner, "Some improvements of particle filtering based prognosis for PEM fuel cells," IFAC-PapersOnLine, vol. 49, no. 28, pp. 162-167, Jan. 2016.

[62] D. Zhou, Y. Wu, F. Gao, E. Breaz, A. Ravey, and A. Miraoui, "Degradation Prediction of PEM Fuel Cell Stack Based on Multiphysical Aging Model With Particle Filter Approach," IEEE Transactions on Industry Applications, vol. 53, no. 4, pp. 4041-4052, Jul./Aug. 2017.

[63] A. Hochstein, H. I. Ahn, Y. T. Leung, and M. Denesuk, "Switching vector autoregressive models with higher-order regime dynamics," in 2014 International Conference on Prognostics and Health Management, pp. 1-10, 2014.

[64] M. Ibrahim, N. Y. Steiner, S. Jemei, and D. Hissel, "Wavelet-based approach for online fuel cell remaining useful lifetime prediction," IEEE Transactions on Industrial Electronics, vol. 63, pp. 5057-5068, Aug. 2016.

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

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

[67] 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, vol. 43, no. 27, pp. 12314-12327, Jul. 2018.

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

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

[70] C. Yang, Z. Li, B. Liang, Q. Cui, W. Lu, and J. Yang, "A novel fusion strategy for failure prognostic of proton exchange membrane fuel cell stack," in 2017 - 36th Chinese Control Conference (CCC), 2017, pp. 7070-7076.

[71] C. Yang, Z. Li, B. Liang, W. Lu, X. Wang, and H. Liu, "A particle filter and long short term memory fusion algorithm for failure prognostic of proton exchange membrane fuel cells," in 2017 - 29th Chinese Control And Decision Conference (CCDC), 2017, pp. 5646-5651.

[72] R. Ma, Z. Li, E. Breaz, C. Liu, H. Bai, P. Briois, et al., "Data-fusion prognostics of proton exchange membrane fuel cell degradation," IEEE Transactions on Industry Applications, Apr. 2019.

[73] R. Ma, R. Xie, L. Xu, Y. Huangfu, and Y. Li, "A Hybrid Prognostic Method for PEMFC with Aging Parameter Prediction," IEEE Transactions on Transportation Electrification, pp. 1-1, 2021.

[74] R. E. Silva, R. Gouriveau, S. Jemeï, D. Hissel, L. Boulon, K. Agbossou, et al., "Proton exchange membrane fuel cell degradation prediction based on adaptive neuro-fuzzy inference systems," International Journal of Hydrogen Energy, vol. 39, pp. 11128-11144, Jul. 2014.

[75] L. Mao and L. M. Jackson, "Comparative study on prediction of fuel cell performance using machine learning approaches," 2016.

[76] H. Liu, J. Chen, D. Hissel, and H. Su, "Short-Term Prognostics of PEM Fuel Cells: A Comparative and Improvement Study," IEEE Transactions on Industrial Electronics, vol. 66, no. 8, pp. 6077-6086, Aug. 2019.

[77] K. Javed, R. Gouriveau, N. Zerhouni, D. Hissel, "Data-driven prognostics of proton exchange membrane fuel cell stack with constraint based summation-wavelet extreme learning machine," in the International Conference on Fundamentals & Development of Fuel Cells, pp. 1-8, 2015.

[78] K. Javed, R. Gouriveau, N. Zerhouni, and D. Hissel, "Improving accuracy of long-term prognostics of PEMFC stack to estimate remaining useful life," in 2015 IEEE International Conference on Industrial Technology (ICIT), 2015, pp. 1047-1052.

[79] K. Javed, R. Gouriveau, N. Zerhouni, and D. Hissel, "Prognostics of proton exchange membrane fuel cells stack using an ensemble of constraints based connectionist networks," Journal of Power Sources, vol. 324, pp. 745-757, Aug. 2016.

[80] K. Javed, R. Gouriveau, N. Zerhouni, and D. Hissel, "PEM fuel cell prognostics under variable load: a data-driven ensemble with new incremental learning," in 2016 International Conference on Control, Decision and Information Technologies (CoDIT), 2016, pp. 252-257.

[81] H. Liu, J. Chen, M. Hou, Z. Shao, and H. Su, "Data-based short-term prognostics for proton exchange membrane fuel cells," International Journal of Hydrogen Energy, vol. 42, pp. 20791-20808, Aug. 2017.

[82] X. Zhang, Z. Yu, and W. Chen, "Life prediction based on d-s ELM for PEMFC," Energies, vol. 12, 2019.

[83] K. Chen, S. Laghrouche, and A. Djerdir, "Degradation model of proton exchange membrane fuel cell based on a novel hybrid method," Applied Energy, vol. 252, p. 113439, Oct. 2019.

[84] 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, vol. 195, pp. 810-818, May. 2019.

[85] K. Chen, S. Laghrouche, and A. Djerdir, "Aging prognosis model of proton exchange membrane fuel cell in different operating conditions," International Journal of Hydrogen Energy, vol. 45, no. 20, pp. 11761-11772, Apr.2020.

[86] Y. Wu, E. Breaz, F. Gao, and A. Miraoui, "A modified relevance vectormachine for PEM fuel-cell stack aging prediction," IEEE Transactions on Industry Applications, vol. 52, no.3, pp. 2573-2581, May. 2016.

[87] 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, vol. 31, pp. 1570-1582, Dec. 2016.

[88] 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, vol. 231, pp. 102-115, Dec. 2018.

[89] R. Ma, E. Breaz, C. Liu, H. Bai, P. Briois, and F. Gao, "Data-driven prognostics for pem fuel cell degradation by long short-term memory network," in 2018 IEEE Transportation Electrification Conference and Expo (ITEC), pp. 102-107, 2018.

[90] J. Liu, Q. Li, W. Chen, Y. Yan, Y. Qiu, and T. Cao, "Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks," International Journal of Hydrogen Energy, vol. 44, no. 11, pp. 5470-5480, Feb. 2019.

[91] J. Liu, Q. Li, Y. Han, G. Zhang, X. Meng, J. Yu, et al., "PEMFC Residual Life Prediction Using Sparse Autoencoder-Based Deep Neural Network," IEEE Transactions on Transportation Electrification, vol. 5, no. 4, pp. 1279-1293, Dec. 2019.

[92] S. Morando, M.-C. Péra, 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), pp. 1-6, 2016.

[93] S. Morando, M.-C. Péra, 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), pp. 1-7, 2017.

[94] Z. Zheng, S. Morando, M.-C. Péra, D. Hissel, L. Larger, R. Martinenghi, A. B. Fuentes. "Brain-inspired computational paradigm dedicated to fault diagnosis of PEM fuel cell stack," International Journal of Hydrogen Energy, vol. 42, pp. 5410-5425, Feb. 2017.

[95] Z. Zheng, M.-C. Péra, D. Hissel, L. Larger, N. Y. Steiner, and S. Jemei, "Fault diagnosis of PEMFC systems in the model space using reservoir computing," in 2018 IEEE Vehicle Power and Propulsion Conference (VPPC), pp. 1-5, 2018.

[96] S. Morando, S. Jemei, R. Gouriveau, N. Zerhouni, and D. Hissel, "Fuel cells prognostics using echo state network," in IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, pp. 1632-1637, 2013.

[97] S. Morando, S. Jemei, R. Gouriveau, N. Zerhouni, and D. Hissel, "Fuel cells remaining useful lifetime forecasting using echo state network," in 2014 IEEE Vehicle Power and Propulsion Conference (VPPC), pp. 1-6, 2014.

[98] S. Morando, S. Jemei, D. Hissel, R. Gouriveau, and N. Zerhouni, "Predicting the remaining useful lifetime of a proton exchange membrane fuel cell using an echo state network," in International Discussion on Hydrogen Energy and Applications (IDHEA), France, 2014, pp. 1-9.

[99] 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, vol. 42, pp. 1472-1480, Jan. 2017.

[100] S. Morando, S. Jemei, D. Hissel, R. Gouriveau, and N. Zerhouni, "ANOVA method applied to proton exchange membrane fuel cell ageing forecasting using an echo state network," Mathematics and Computers in Simulation, vol. 131, pp. 283-294, Jan. 2017.

[101] R. Mezzi, S. Morando, N. Y. Steiner, M.-C. Péra, D. Hissel, and L. Larger, "Multi-reservoir echo state network for proton exchange membrane fuel cell remaining useful life prediction," in IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society, pp. 1872-1877, 2018.

[102] Z. Li, S. Jemei, R. Gouriveau, D. Hissel, and N. Zerhouni, "Remaining useful life estimation for PEMFC in dynamic operating conditions," in 2016 IEEE Vehicle Power and Propulsion Conference (VPPC), pp. 1-6, 2016.

[103] Z. Li, Z. Zheng, and R. Outbib, "Adaptive prognostic of fuel cells by implementing ensemble echo state networks in time varying model space, " IEEE Transactions on Industrial Electronics, vol. 67, pp. 379-389, Jan. 2020.

[104] R. Mezzi, N. Yousfi-Steiner, M.-C. Péra, D. Hissel, and L. Larger, "An echo state network for

fuel cell lifetime prediction under a dynamic micro-cogeneration load profile," Applied Energy, vol. 283, p. 116297, Feb. 2021.

[105] Z. Hua, Z. Zheng, M.-C. Péra, and F. Gao, "Data-driven prognostics for PEMFC systems by different echo state network prediction structures," in ITEC 2020 - IEEE Transportation Electrification Conference and Exposition, pp. 495-500, Jun. 2020.

[106] Z. Hua, Z. Zheng, M.-C. Péra, and F. Gao, "Remaining useful life prediction of PEMFC systems based on the multi-input echo state network," Applied Energy, vol. 265, p. 114791, May. 2020.

[107] Z. Hua, Z. Zheng, E. Pahon, M.-C. Péra, and F. Gao, "Remaining useful life prediction of PEMFC systems under dynamic operating conditions," Energy Conversion and Management, vol. 231, p. 113825, Mar. 2021.

[108] L. Vichard, F. Harel, A. Ravey, P. Venet, and D. Hissel, "Degradation prediction of PEM fuel cell based on artificial intelligence," International Journal of Hydrogen Energy, vol. 45, no. 29, pp. 14953-14963, May.2020.

[109] W. Maass, T. Natschläger, and H. Markram. "Real-time computing without stable states: a new framework for neural computation based on perturbations," Neural computation, vol. 14, pp. 2531-2560, 2002.

[110] J. J. Steil. "Backpropagation-Decorrelation: online recurrentlearning with O(N) complexity," in the IEEE International Joint Conference on Neural Networks, pp. 843-848, 2004.

[111] Z. Zheng, S. Morando, M.-C. Péra, D. Hissel, L. Larger, R. Martinenghi, A. B. Fuentes. "Brain-inspired computational paradigm dedicated to fault diagnosis of PEM fuel cell stack," International Journal of Hydrogen Energy, vol. 42, no. 8, pp. 5410-5425, Feb. 2017.

[112] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks-with an erratum note," Fraunhofer Institute for Autonomous Intelligent Systems, 2010.

[113] H. Jaeger, "Tutorial on training recurrent neural networks, covering BPTT, RURL, EKF and the 'Echo State Network' approach," Technical report GMD-German National Research Center for Information Technology, 2002.

[114] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, et al., "Recent advances in physical reservoir computing: A review," Neural Networks, vol. 115, pp. 100-123, Jul. 2019.

[115] H. Jaeger, M. Lukoševičius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky- integrator neurons," Neural Networks, vol. 20, pp. 335-352, Apr. 2007.

[116] H. Jaeger, "Long short-term memory in echo state networks: details of a simulation study," Technical Report No. 27, School of Engineering and Science, Jacobs University, Feb. 2012.

[117] M. Lukoševičius, "A practical guide to applying echo state networks," Neural Networks: Tricks of the Trade, Springer, 2012, pp. 659-686.

[118] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," Neural Networks, vol. 20, no. 3, pp. 391-403, Apr. 2007.

[119] B. Liu, H. Chen, T. Zhang, and P. Pei, "A vehicular proton exchange membrane fuel cell system co-simulation modeling method based on the stack internal distribution parameters monitoring," Energy Conversion and Management, vol. 197, p. 111898, Jan. 2019.

[120] J. Spendelow, J. Marcinkoski, and D. Papageorgopoulos, "Micro CHP fuel cell system targets," DOE Hydrogen and Fuel Cells Program Record, Mar. 2012.

[121] J. Kurtz, S. Sprik, G. Saur, and S. Onorato, "Fuel cell electric vehicle durability and fuel cell performance," Technical report NREL, Mar. 2019.

[122] R. Gouriveau and N. Zerhouni, "Connexionist-systems-based long term prediction approaches for prognostics," IEEE Transactions on Reliability, vol. 61, pp. 909-920, Dec. 2012.

[123] 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, vol. 1, pp. 2153-2648, Jan. 2010.

[124] N. Chouikhi, B. Ammar, N. Rokbani, and A. M. Alimi, "PSO-based analysis of echo state network parameters for time series forecasting," Applied Soft Computing, vol. 55, pp. 211-225, Jun. 2017.

[125] A. Azadeh, S. F. Ghaderi, and S. Sohrabkhani, "Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors," Energy Conversion and Management, vol. 49, pp. 2272-2278, Aug. 2008.

[126] I. M. Sobol, "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," Mathematics and Computers in Simulation, vol. 55, pp. 271-280, Feb. 2001.

[127] M. Ibrahim, S. Jemei, G. Wimmer, N. Y. Steiner, C. C. Kokonendji, and D. Hissel, "Selection of mother wavelet and decomposition level for energy management in electrical vehicles including a fuel cell," International Journal of Hydrogen Energy, vol. 40, no. 45, pp. 15823-15833, Dec. 2015.

[128] E. Pahon, N. Yousfi Steiner, S. Jemei, D. Hissel, and P. Moçoteguy, "A signal-based method for fast PEMFC diagnosis," Applied Energy, vol. 165, pp. 748-758, Mar. 2016.

[129] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, July 1989.

[130] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," Computer Science Review, vol. 3, no.3, pp. 127-149, Aug. 2009.

[131] H. Wang, C. Ni, and X. Yan, "Optimizing the echo state network based on mutual information for modeling fed-batch bioprocesses," Neurocomputing, vol. 225, pp. 111-118, Feb. 2017.

[132] L. A. Thiede and U. Parlitz, "Gradient based hyperparameter optimization in echo state networks," Neural Networks, vol. 115, pp. 23-29, Jul. 2019.

[133] H. Wang and X. Yan, "Optimizing the echo state network with a binary particle swarm optimization algorithm," Knowledge-Based Systems, vol. 86, pp. 182-193, Sep. 2015.

[134] T. Sergio and T. B. Ludermir, "PSO for reservoir computing optimization," in Artificial Neural Networks and Machine Learning – ICANN 2012, Berlin, Heidelberg, 2012, pp. 685-692.

[135] N. Chouikhi, R. Fdhila, B. Ammar, N. Rokbani, and A. M. Alimi, "Single- and multi-objective particle swarm optimization of reservoir structure in echo state network," in 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 440-447.

[136] M. J. A. Rabin, M. S. Hossain, M. S. Ahsan, M. A. S. Mollah, and M. T. Rahman, "Sensitivity learning oriented nonmonotonic multi reservoir echo state network for short-term load forecasting," in 2013 International Conference on Informatics, Electronics and Vision (ICIEV), 2013, pp. 1-6.

[137] S. Basterrech, E. Alba, and V. Snášel, "An experimental analysis of the Echo State Network initialization using the particle swarm optimization," in 2014 Sixth World Congress on Nature and Biologically Inspired Computing (NaBIC 2014), 2014, pp. 214-219.

[138] A. Ferreira and T. B. Ludermir, "Genetic algorithm for reservoir computing optimization," in 2009 International Joint Conference on Neural Networks, 2009, pp. 811-815.

[139] S. Zhong, X. Xie, L. Lin, and F. Wang, "Genetic algorithm optimized double-reservoir echo state network for multi-regime time series prediction," Neurocomputing, vol. 238, pp. 191-204, May. 2017.

[140] 2013-2016: ANR PROGELEC PROPICE: Prognostics et Health Management de systèmes Piles à Combustible de type PEMFC, available on http://www.propice.ens2m.fr/.

[141] I. Daubechies, "Ten lectures on wavelets," Rutgers University and AT&T Bell Laboratories, New Jersey, Second Printing, 1992.