Energy Reports 8 (2022) 6687–6706

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Review article

Fuel cell diagnosis methods for embedded automotive applications

J. Aubry[a,b,*], N. Yousfi Steiner[b], S. Morando[a], N. Zerhouni[b], D. Hissel[b]

[a] SYMBIO, Vénissieux, France
[b] Univ. Bourgogne Franche-Comté, FEMTO-ST, FCLAB, CNRS, Belfort, France

ARTICLE INFO

Article history:
Received 21 February 2022
Received in revised form 22 April 2022
Accepted 11 May 2022
Available online xxxx

Keywords:
PEM fuel cell
Embedded diagnostic
Fuel cell system degradation
Fuel cell system durability

ABSTRACT

Fuel cell durability being one of the technical bolts regarding the technology industrialization in the automotive sector, durability improvement methods are particularly relevant. Fault tolerant control process enables to increase fuel cell durability by detecting and correcting fuel cell faults in real time. Fuel cells are prone to faults because they are very sensitive to operating conditions. In vehicle application, fault risk is exacerbated as dynamic conditions are often encountered. Dynamic conditions make the fuel cell control harder because it impacts reactants supply, thermal management, water management… If not corrected, those faults degrade the fuel cell and reduce its remaining useful lifetime.

Fault tolerant control consists in diagnosing faults, then taking corrective actions to resolve those faults. This article treats the diagnosis part, which consists in detecting and identifying faults, in vehicle application. Vehicle applications engender several constraints as the reduced cost, the hydrogen usage and computation limitations or the safety regulations for algorithms implementation. Three steps are necessary for diagnosis: real time measurements, useful information extraction, and classification. In this article, a state-of-the-art of methods for each of these steps independently is presented. In the last section, useful explanations to convert offline diagnosis algorithm into an embedded diagnosis tool are provided.

Contents

* Corresponding author at: Univ. Bourgogne Franche-Comté, FEMTO-ST, FCLAB, CNRS, Belfort, France.
E-mail addresses: julie.aubry@symbio.one (J. Aubry), nadia.steiner@femto-st.fr (N.Y. Steiner), simon.morando@symbio.one (S. Morando), noureddine.zerhouni@ens2m.fr (N. Zerhouni), daniel.hissel@femto-st.fr (D. Hissel).

https://doi.org/10.1016/j.egyr.2022.05.036

Glossary

PEFMC	Proton Exchange Membrane Fuel Cell
FDI	Fault detection and isolation
FTC	Fault Tolerant Control
EIS	Electrochemical Impedance Spectroscopy
PCA	Principal Component Analysis
LDA	Linear discriminant Analysis
(K)FDA	(Kernel) Fisher Discriminant Analysis
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
SVM	Support Vector Machines
DB scan	Density Based clustering
KNN	K-Nearest-Neighbor
GMM	Gaussian Mixture model
DA	Discriminant Analysis
FFT	Fast Fourier Transform

1. Introduction

Global warming, air pollution and health issues incite the use of other transportation propulsion solutions rather than traditional thermal engines. Electric vehicles having no local greenhouse gases emission, they present a particular interest for vehicular applications. This trend intensified the last few years and the global electric car stock reached 5 million vehicles in 2019 (Mawonou et al., 2021). Hydrogen fuel cells are electrochemical devices converting the chemical energy contained in hydrogen into electrical energy. Fuel cell vehicles only require several minutes to fill the tank and their autonomy is higher than the one of battery vehicles. Proton Exchange Membrane Fuel Cells (PEMFCs) are a promising technology for vehicular application due to their high efficiency and low operating temperature (around 80 °C).

However, fuel cell durability and reliability are limiting factors to their commercialization. Fuel cells are prone to faults. Faults are anomalies leading to performance losses and degradation (Andújar et al., 2018; Zheng et al., 2013; Petrone et al., 2013; Zheng et al., 2021). The five main faults are fuel cell flooding, drying, fuel or oxidant starvation and catalyst poisoning. As presented in Section 3, those faults are not specific to automotive fuel cell systems, but their cause are related to the ancillaries, which are specific to the automotive application. To avoid early degradation, diagnosis tools are very useful as they enable to detect and identify faults in real time. From Fault Detection and Identification (FDI), it is possible to take corrective actions to correct the faults. Diagnosis with associated corrective actions is called Fault Tolerant Control (FTC).

A bench of different methods and approaches can be used to perform diagnosis. Andújar et al. (2018) presented the existing methods enabling to diagnose faults in the hydrogen line, air/cooling subsystem, electrical circuit and even in the stack itself on air-cooled polymer electrolyte fuel cells. Yang et al. (2021) reviewed recent studies dealing with Solid Oxide Fuel Cell diagnosis.

Diagnosis methods are usually separated into model based and non-model-based approaches depending if diagnosis requires a fuel cell model or not (for more information, please refer to Section 5). Zheng et al. (2013) introduced non-model based approaches while Petrone et al. (2013) reviewed the different model-based diagnosis methods used to detect faults in fuel cells. Yuan (2020) presented new developments dealing with model-based observers, also called internal state observers, applied to fuel cells.

Benmouna et al. (2017) presented both models based and non-model-based diagnosis methods with an illustrative example of anode flooding. Lin et al. (2019b) reviewed recent fuel cell diagnosis studies. The authors separated the different diagnosis techniques into model based, data based and filter-based approaches. Tang (2020) reviewed recent studies using EIS to diagnose fuel cell faults. Electrochemical Impedance Spectroscopy (EIS) is a fuel cell characterization method widely used for diagnosis purposes.

As mentioned above, multiple studies present the state of the art of offline fuel cell diagnosis. However, online diagnosis and more particularly embedded diagnosis for vehicle applications is much less presented in the literature. In automotive applications,

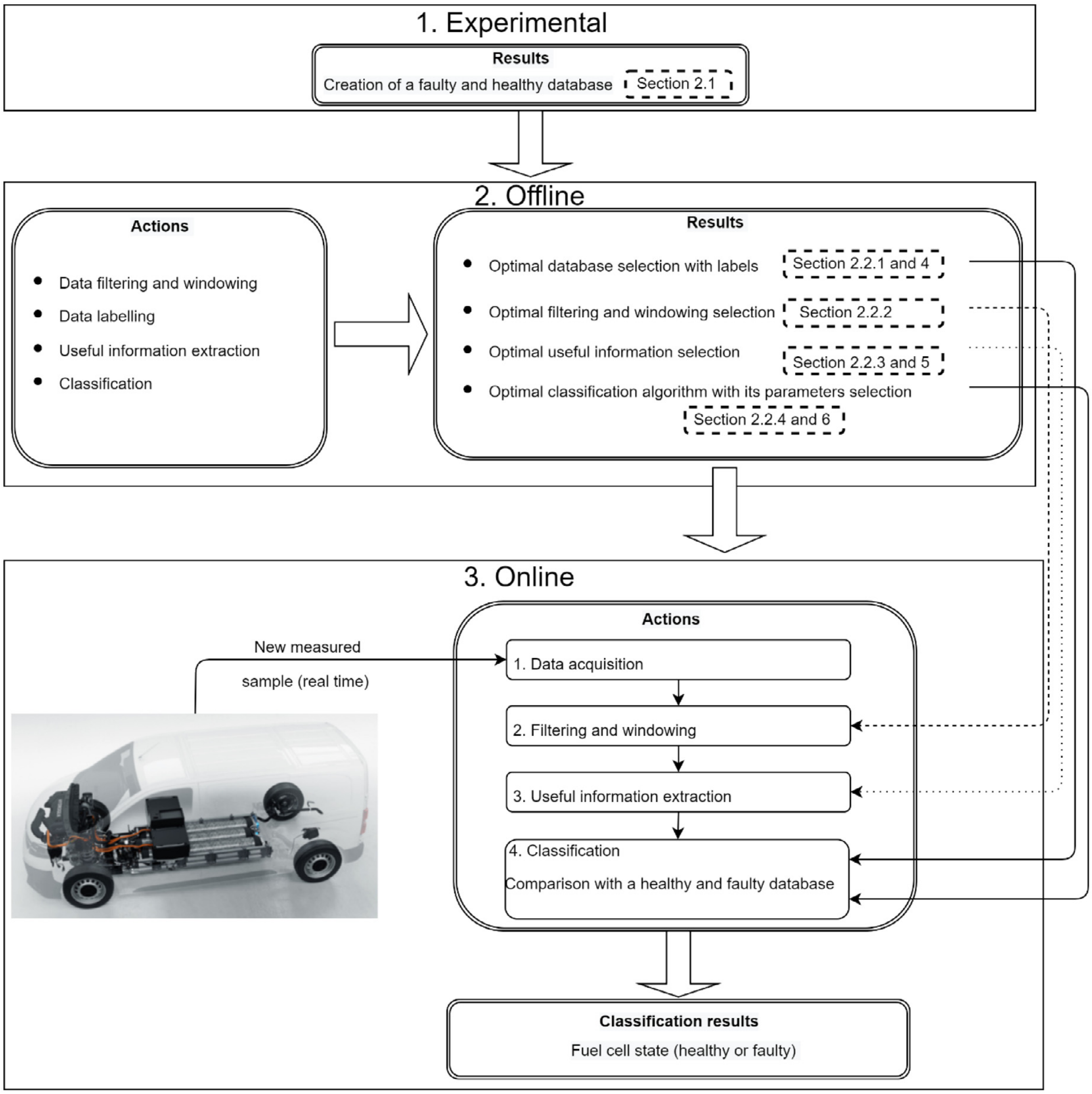


Fig. 1. General process of embedded fault diagnosis tool development: experimental data, online and offline process.

a lot of constraints and challenges are encountered for embedded diagnosis tool development process. Those constraints are addressed in Section 2.4.

This article objective is to clearly present the existing methods and approaches to develop an embedded fuel cell diagnosis tool. Diagnosis tool development mainly consists in 3 steps: the experimental step, the offline step, and the online step. The different approaches for each of these steps will be presented in this paper. Note that embedded diagnosis is necessarily online, but online diagnosis can be delocalized if there is a communication path with an external computer.

The second section of this article presents an overview of online embedded diagnosis tool development process, so the reader understands the global context of this article and the constraints of embedded application. The third section describes common fuel cell faults and their causes in fuel cell vehicles. Fuel cell diagnosis requires real time data acquisition (treated in the fourth section of this article), useful information extraction (presented in Section 5) and classification (explained in Section 6). Section 7 provides to the reader some tips to embed the algorithm in a vehicle. Finally, Section 8 concludes this article.

2. General presentation of fault detection and isolation (FDI) tool development

The "Fault detection and isolation" tool (FDI) allows the detection and identification of faults occurring in a system. This information is very useful when applied to fuel cell systems as corrective actions can be taken, leading to an increase of the fuel cell lifetime (Dijoux, 2019; Aubry et al., 2020).

The goal of this section is to explain the FDI tool conception process in detail, specifically for embedded application. As presented in Fig. 1, the development of an onboard FDI tool can be separated into three major steps. They are the experimental, offline and online steps. Those steps are presented in the firsts three sub-sections, which are not specific to automotive applications. The fourth section is dedicated to the explanation of the additional constraints linked to the embedded automotive application.

2.1. Experimental part

Experimental data is required to develop an FDI tool. The amount and type of data depends on the useful information extraction method. There are two main categories of useful information extraction methods: the ones that are based on a fuel cell model and the ones that are not.

The experimental data can be composed of the different measured fuel cell operating conditions (temperature, pressure, current… variations). This data is used for model-based diagnosis methods (explained in Section 5.2), which consists of comparing the outputs of a fuel cell model with the same outputs of real-time measurements. The amount of required experimental data is low if the model is based on physical equations (grey box model) and large if it is a black box model as data is required to train the model. For model-based diagnosis methods, experimental data is required to calibrate the fuel cell model.

The experimental data can also be composed of the different measured fuel cell operating conditions, when faults are deliberately generated on the fuel cell. This data is used for non-model-based methods (presented in Section 5.1), which consist of identifying fault patterns. All the faults that one wants to detect must be experimentally generated. This leads to a substantial amount of data.

The experimental data generation is one of the trickiest parts of diagnosis. Data should accurately represent the phenomena occurring in vehicular applications, even though data collection is performed on a test bench. Data quality is of primary importance for diagnosis as it uses machine learning tools. The quality, reliability and repeatability of the experiments are crucial (Cadet et al., 2014).

2.2. Offline part

The objective of the offline part is to select the database, the data processing method, the useful information extraction method, and the classifier.

To achieve this, multiple methods are tested offline, on a computer. The methods leading to the best chance of detecting and identifying a fault online are then selected.

2.2.1. Database formation and labeling

In diagnosis applications, the real time measurements are compared to a database during the classification. This database substantially influences the diagnosis results. The data may come from different sensors or characterization techniques. As the selection of the measurements and characterization method is very important, a focus is presented on this subject in Section 4.

Labeling the database is another important step in the diagnosis tool conception. Labeling consists in attributing a label or a class (typically healthy, flooded, dried….) to the experimental data. This step is especially required for non-model based useful information extraction methods. This is tricky because it requires the identification of exact fault, the moment when it is created, as well as its duration. Also, if two faults are generated at the same time, this could bias the database and the diagnosis results.

2.2.2. Filtering and windowing

Experimental data may require filtering to isolate the useful information. Filters are especially required when a fast Fourier transform is used to extract the useful information of a signal, to avoid edge artifacts. Multiple tapering windows may be used.

Moreover, the data acquisition frequency can be optimized to obtain good diagnosis result without requiring large storage space. Usually, many data samples are acquired at once, in blocks. This is called interrupt buffered acquisition. The number of samples in the blocks corresponds to the window length. This parameter also has an impact on diagnosis results.

Finally, data can be normalized or standardized. Normalization consists in rescaling the data in between 0 and 1. Standardization consists in rescaling the data to have a means of 0 and a standard deviation of 1. Data normalization or standardization are used to be able to compare various data sources, which are not initially on the same order of magnitude (for example stack temperature and cell voltage). This avoids giving one parameter more importance than another.

2.2.3. Useful information extraction and selection

Numerous methods exist to extract useful information. They may or may not require a fuel cell model. A synthesis of the useful information extraction methods is presented in Section 5.

For non-model-based methods, multiple tools exist to select the best features. The first type is dimensions reduction methods such as PCA (Principal Component Analysis), LDA (Linear discriminant Analysis), FDA (Fisher Discriminant Analysis), KFDA (Kernel Fisher Discriminant Analysis) (Li et al., 2014). In dimension reduction methods, only the most relevant information is kept. With this method, the non-relevant or redundant information can be removed which results in a better classification result since the noise is lower. Moreover, this minimizes the data storage space required in the vehicle.

For example, Li et al. (2018) and Zheng et al. (2021) used a PCA algorithm to directly select the features while removing the non-relevant information.

Another method to select the useful information is the GINI calculation. This statistical formula determines the impurity of a feature as presented on Fig. 2. The GINI impurity is a measure of the likelihood of an incorrect classification. Its value varies between 0 and 1, with 0 being the best feature, with no impurity and 1 being a totally impure feature. The idea is to compare the false positive and false negative samples when the feature is activated to the true positive and true negative samples. This can be used to grade the features as well as their threshold and select the best ones.

For model-based method, the residual between the model and the real time measurement is the useful information. Therefore, the useful information selection consists in selecting the model.

2.2.4. Classification algorithm selection

A classification algorithm determines the category to which a set of data belongs (typically faulty, fault type or healthy categories). As for useful information extraction methods, classifiers are numerous. Therefore, they are presented in a dedicated section (see Section 6).

The classifier selection consists in choosing the algorithm providing the best performances for the studied case, without requiring too much computational time (this parameter is limited by the automotive microcontroller performances).

Multiple classification algorithms have internal variables. Their optimization should also be studied to obtain the best classification probability.

2.3. Online part

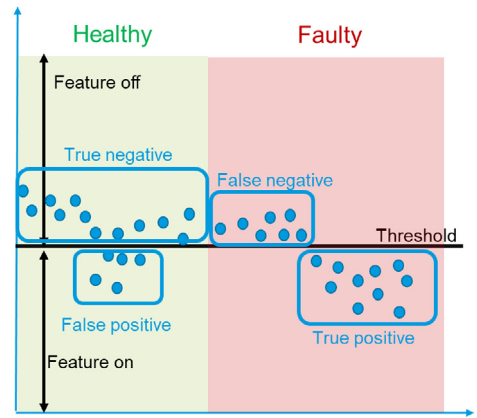
Online diagnosis detects and identifies a fault in real time, during fuel cell operation. The process is described in Fig. 1.

First, the real time data is acquired. The nature of the data corresponds to the selected data type during the offline phase

	True class: Faulty	True class: Healthy
Feature on	True positive = TP	False positive = FP
Feature off	False negative = FN	True negative = TN

$$\text{Impurity feature on} = 1 - \left(\frac{TP}{TP+FP}\right)^2 - \left(\frac{FP}{TP+FP}\right)^2$$

$$\text{Impurity feature off} = 1 - \left(\frac{FN}{FN+TN}\right)^2 - \left(\frac{TN}{FN+TN}\right)^2$$



$$GINI = \text{Impurity feature on} \times \frac{TP+FP}{TP+FP+FN+TN} + \text{Impurity feature off} \times \frac{TN+FN}{TP+FP+FN+TN}$$

Fig. 2. GINI feature impurity coefficient explanation.

(see Sections 2.2.1 and 4). Typically, the data type is cell or stack voltage, temperature, pressure… measurements. If characterization methods are used as EIS or polarization curves, the resulting parameters (as the impedance, the exchange current density…) can also be used as input data.

Then, these real time data samples are filtered with the parameters determined offline (see Section 2.2.2).

An additional computation is performed to extract the useful information from the real time data. The useful information extraction method has also been preliminarily determined offline (see Sections 2.2.3 and 5)

Finally, the real-time extracted useful information is compared to the selected embedded experimental database. This comparison is performed using a classifier, which type and internal parameters have been determined during the offline classification algorithm selection (see Sections 2.2.4 and 6).

In some model-based applications, a complex classifier is not required. The classification is performed by evaluating the value of the residual between the model and the real time measurement. Typically, if the difference between the real time voltage and the voltage determined by the model in healthy conditions is high, the fuel cell is considered as faulty (for more information, refer to Section 6). In this case, only a threshold value is required to identify whether the fuel cell is in healthy or faulty condition. However, different thresholds values, or multiple fuel cell models (healthy and faulty) must be embedded to be able to identify the fault.

The process to build a real time fault detection and identification algorithm suitable for fuel cells has been described in this section. All these methods must be adapted to the automotive constraints. The next subsection presents main constraints that are encountered in embedded automotive applications.

2.4. Embedded systems diagnosis constraints

Automotive fuel cell diagnosis is usually embedded in the vehicle to be able to detect and identify a fault in real time. This specification leads to several constraints. The main ones are the computation time, the storage space, the safety regulations, the cost and the volume/weight of the diagnosis system. Moreover, the diagnosis frequency should be high enough to detect the fault early enough and correct it before degradations occur.

2.4.1. Computational time limitations

Diagnosis requires solving many equations. In a vehicle, these equations are computed by a microcontroller. Microcontrollers are compressed microcomputers on a chip used in embedded systems. Microcontrollers have a limited capacity in terms of computation speed and internal storage capacity. As Fault Detection and Identification is developed for fault tolerant purposes, the frequency of the diagnosis must be high enough to allow a rapid fault correction (before irreversible degradation occurs).

The two steps requiring the highest computational time are the useful information extraction and the classification. Therefore, the computational time available on the automotive microcontroller should be considered for selecting the methods.

2.4.2. Memory space

In a vehicle, data storage capacity is limited. The comparative database to which real time samples are compared during classification must be stored in the vehicle microcontroller. This requires internal memory space. Thus, the database should be adapted to the available storage capacity. However, a minimum of data samples is required as the diagnosis should be adapted to the largest number of situations in a vehicle. In other words, the diagnosis should work properly at the beginning of fuel cell life, but also after several hundred or thousand hours of operation. Therefore, the embedded data quality and representativeness are crucial for diagnosis application.

2.4.3. Safety regulations

A vehicle controller is subject to very strict regulations to avoid safety issues as malfunctions can have dramatic consequences. Diagnosis can be implemented to a vehicle controller, but the controller is also required for the normal driving use. Thus, the additional diagnosis code layer must be highly checked to avoid any malfunctions that could lead to unwanted actions on the driving experience.

2.4.4. Cost and volume

In the automotive industry, competition is fierce, so the manufacturing cost is of primary importance. Thus, the cost of each additional sensor, additional storage space, microcontroller and all additional equipment must be considered. This directly impacts the choice of measurement and characterization type.

Also, the volume of the equipment should be as low as possible as the space is a critical parameter in automotive applications. For example, the cooling circuit might have to be bigger to properly cool down the electronics.

Diagnosis may also be used to limit the number of redundant sensors while ensuring the security of the system as presented in Ref. Behravan et al. (2019).

2.4.5. Impact on the driving experience

Several diagnosis methods can be based on electrochemical characterization methods as polarization curves, or Electrochemical Impedance Spectroscopy (EIS). However, polarization curves and EIS measurements require several minutes to be performed. During that time, the produced power is not flexible. This can be problematic in fuel cell automotive systems because the power demand is not known in advance and the produced power may be lower than the power demand. A battery or an ultracapacitor can provide the power required by the vehicle during the characterization in some cases but there should not have any impact on the driving experience.

2.5. Partial synthesis

The different parts in diagnosis tool development are an experimental part, an offline part, and an online part. The optimum algorithms and associated parameters are determined during the offline part, using the experimental data. The online part is the integration of the best diagnosis method in the vehicle. The embedded automotive application generates a lot of constraints presented in Section 2.4.

This section objective is to give the reader an overview of the global workflow of diagnosis tool development process. Several steps require deeper explanations because a lot of methods exist, and they are of primary importance for diagnosis results. Therefore, they are the core of separated sections. This is the case for measurement and characterization methods (Section 4), useful information extraction methods (Section 5) and classification algorithms (Section 6).

Diagnosis is required because faults occur in fuel cells. To understand the context of the study, the next section is dedicated to the presentation and explanation of the major fuel cell faults.

3. PEMFC faults and failures in an automotive system

Fuel cells are prone to faults, especially when the current profile is dynamic. This section presents the most frequent and degrading fuel cell faults. Note that those faults are present in all fuel cell applications. However, the ancillaries of the fuel cell system differ in automotive applications, so the causes of the failures too. This section provides the faults causes in automotive applications.

3.1. Fuel cell automotive system presentation

A fuel cell needs ancillaries to operate. Those ancillaries allow the reactants supply and their humidification, the stack and reactants temperature regulation and the reactants and cooling pressure regulation. Before enouncing the different failure modes, a presentation of the automotive fuel cell system is necessary to understand faults causes.

Note that a fuel cell vehicle powertrain is not only composed of the fuel cell and its ancillaries. Electric motors, batteries, DC/DC converters, an overall system cooling circuit, and multiple other components are also required. However, this article is focused on the diagnosis of the fuel cell only since it is the least reliable component compared to its price. Thus, only the components leading to possible and correctable fuel cell faults are presented.

3.1.1. Reactant's supply and pressure regulation

The air supplied to the fuel cell comes from the ambient. The air compressor extracts the ambient air and compresses it to generate a cathodic flow. Ambient air must be filtered to avoid CO, NOx, and SOx to be supplied to the fuel cell as these particles are very damaging for the fuel cell as presented in Section 3.2. A counter pressure valve is present at the cathode outlet to adapt the cathodic pressure.

Hydrogen is stored in a tank at several hundred bars. It should be depressurized to several bars before entering in the fuel cell. Thus, an expander is placed upstream of the fuel cell. Fuel cells operate with an over-stoichiometric factor. This means that a higher quantity of reactants than the amount needed to produce the required power is injected. This over-stoichiometry avoids reactants starvation, which is very damaging for the fuel cell (see Section 3.2). However, releasing the excess of hydrogen to the exhaust is not an option for cost and safety reasons. Thus, the non-consumed hydrogen is recirculated and reinjected to the fuel cell inlet. Hydrogen recirculation is usually carried out using an injector.

3.1.2. Reactant's humidification

Fuel cell membranes must be humidified to transport H^+ protons. Jung et al. (2007) described all possible humidification methods for fuel cell systems. In automotive systems, membrane humidifiers are usually used to humidify the inlet gases as they are not very complex and can be easily implemented (Firouzjaei et al., 2020). The working principle of membrane humidifiers is that hot and wet gas or water coming from the output of the fuel cell is supplied to one side of the membrane humidifier. Water and heat diffuse through the membrane humidifier and humidify the cold inlet reactants that diffuse on the other side of the membrane (Hwang et al., 2012). The inlet gas humidity is controlled by intermittently bypassing the humidifier.

3.1.3. Fuel cell cooling circuit

The cooling circuit controls the stack temperature. A coolant liquid flows into the fuel cell cooling channels. This liquid must resist to negative temperatures to avoid any freezing during winter. A pump circulates the coolant, and the coolant is cooled down using a radiator (air exchanger). The radiator fan speed is adjusted to adapt the coolant temperature. Three parameters are regulated to adjust the stack temperature: the coolant pump speed, the fan speed and the flowrate of coolant exchanged with the fan (the rest of the coolant is bypassed).

The cooling circuit also contains a heating element. This helps the stack to heat up during cold or even freeze start (see Fig. 3).

3.2. Fuel cell faults

Fuel cells are prone to faults. The main ones are flooding, drying, anodic and cathodic starvation and catalyst poisoning. Those faults have been widely presented in the literature. Therefore, the faults mechanisms will not be detailed, only an overview of the fault types, their causes and consequences are summarized in Table 1. For more information, please refer to the references of Table 1.

Hopefully, the reversible faults can be corrected. This is the reason why early detection of faults is very important. The objective of this paper is not to go into details in faults mitigation strategies. The reader can refer to Refs. Gallo et al. (2020), Dijoux et al. (2017) and Aubry et al. (2020) for more information. Gallo et al. (2020) proposed corrective actions for SOFC faulty conditions regarding the fault magnitude. A Failure Mode and Effect Analysis is presented in Aubry et al. (2020) in which corrective actions for flooding, drying, starvation and catalyst poisoning are suggested while in Dijoux et al. (2017), a Petri net is developed to schedule corrective actions for fuel cell faults.

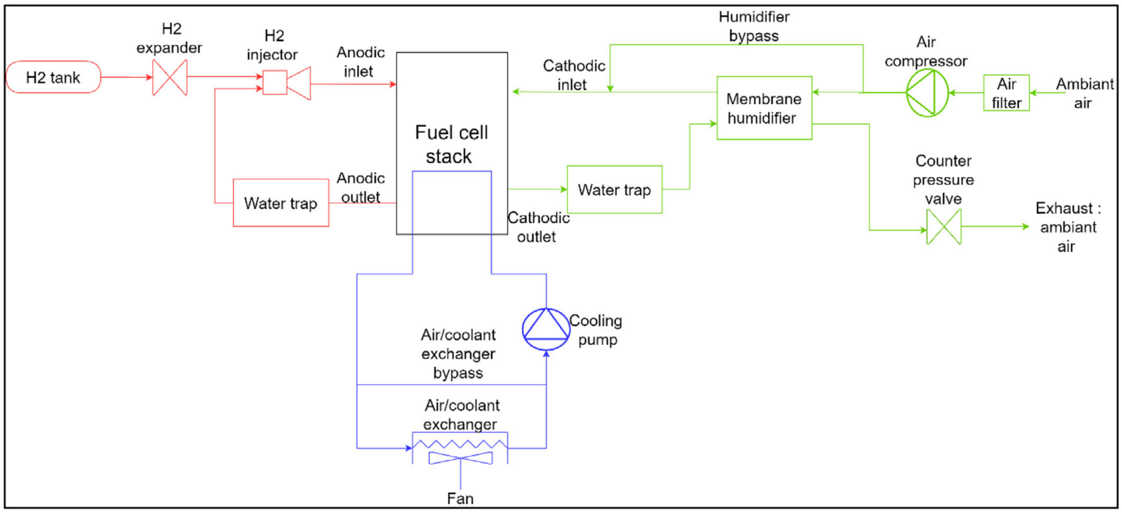


Fig. 3. Scheme of the anodic, cathodic and cooling loop of an automotive fuel cell system.

Table 1
Faults causes and consequences in an automotive system.

Fault	Brief description	Causes	Consequences	References
Flooding	Liquid water molecules block the active sites. This induces starvation phenomenon	Tear in the membrane humidifier[a] Cooling oversizing[b] No radiator bypass[b] Low stoichiometric factor[b] Pressure higher than the saturation pressure[b]	• Performance losses • Starvation	Yousfi-Steiner et al. (2008): *degradation mechanisms associated with water management issues* Owejan et al. (2009): *two phases flow modeling in fuel cells.* Lim et al. (2021): *failure modes in the cooling circuit system presentation* Polverino et al. (2017): *humidification or stack temperature control problems causing drying*
Drying	Lack of water hinders H^+ protons to cross the fuel cell membrane	Humidifier[a] Coolant pump[a] Radiator fan[a] Leak in the cooling circuit[a] Problem in humidification control[b] Problem in stack temperature control[b] High stoichiometric factor[b] Cooling under sizing[b] Problem in humidification control[b] Problem in stack temperature control[b]	• Increase of fuel cell membrane resistance • Increase of gas crossover • Thermal fuel cell membrane degradation • Chemical fuel cell membrane degradation	
Anodic starvation	Lack of hydrogen on the anodic active sites	H2 circuit leakage[a] Flooding[a] Obstruction in the gas channels[a] Obstruction in the active areas[a] Catalyst poisoning[a] Too fast load demand[b] Nitrogen purge badly regulated (purge intervals, duration…)[b] Injector control problem[b]	• Cell reversal phenomenon • platinum oxidation • Carbon corrosion • Catalyst degradation	Yousfi-Steiner et al. (2009): *causes and consequences of starvation on PEMFCs* Ren et al. (2020): *starvation faults in automotive conditions* Liang et al. (2009): *uneven current distribution problems assessment during anodic starvations* Liu et al. (2006): *anodic and cathodic comparison starvations on a 30* cm^2 *PEMFC* Gerard et al. (2010): *experimental study on cathodic starvation*
Cathodic starvation	Lack of oxygen on the cathodic active sites	Compressor failure[a] Leak on the air circuit[a] Flooding[a] Obstruction in the gas channels[a] Obstruction in the active areas[a] Catalyst poisoning[a] Too fast load demand[b] Compressor control problem[b]	• Hydrogen Pumping • Cell reversal phenomenon • Catalyst degradation • Thermal membrane degradation	
Catalyst poisoning	Poison species as CO or H_2S are adsorbed by platinum	Air filter deficiency Unpure hydrogen fueling Internal pollutants	• Loss of active area	Zamel and Li (2011): *contaminants in PEMFCs*

[a]Main component failure possible causes.
[b]Control/sizing mistakes.

4. Measurement and characterization methods

Fault Detection and Identification require real time measurements to evaluate the state of health of the fuel cell. These real time data are obtained by sensors measurements or characterization methods. In the first subsection, the different passive sensors used for fuel cell diagnosis purposes are presented. In the second subsection, a focus is made on an active characterization method: Electrochemical Impedance Spectroscopy (EIS).

4.1. Passive measurement

Sensors are of primary importance to detect and identify faults occurring in the fuel cell system. Different sensor types have been used for diagnosis purposes. It appears that the most common sensors are stack voltage, cathode inlet and outlet pressure, and stack temperature sensors.

Li et al. (2014) and Li et al. (2016) chose cell voltages to diagnose fuel cell faults. Fuel cell voltage depends on the performances and state of health of the fuel cell. Thus, it is an indicator for fault detection and identification. Stack voltage is usually not sufficient for diagnosis as the fault can be hidden by the measurement noise. This is particularly the case for vehicle applications as the required power is high. Therefore, fuel cell stacks are composed of several tens of cells in series. It is often observed that the voltages of the cells located at different positions are different (Li et al., 2016). Moreover, analyzing cell voltage enables to locate the fault and detect it before it spreads to the whole stack.

However, measuring single cell voltages in automotive applications is not always possible due to the required data storage space and the number of sensors. Moreover, the pins taking the measure must be very thin to avoid any short circuit, so their mechanical resistance might be insufficient because of vehicles vibrations. Therefore, measuring voltage of packets of several cells may be a good compromise.

Lim et al. (2021) used 5 temperature sensors and 5 pressure sensors to detect failures in the cooling circuit. The large number of sensors enables diagnosis precision and reliability. However, this may not be feasible in a real automotive system because the trend is to reduce the sensors for economic (and reliability) reasons.

Yousfi Steiner et al. (2011) used voltage measurement and the pressure difference between the outlet and the inlet of the cathodic channel to detect flooding and drying. The pressure difference is indeed a good indicator for flooding as water droplets movements generate pressure oscillations across the electrode.

4.2. EIS characterization

Electrochemical Impedance Spectroscopy (EIS) characterization consists in injecting a small AC current at different frequencies. The resulting voltage is analyzed, and the fuel cell impedance can be extracted. These excitations are repeated at multiple frequencies to obtain the impedance at those frequencies. Usually, the AC current amplitude is equal or less than 10% of the continuous current component.

The impedance is a state indicator: the membrane resistance, the gas diffusivity and other internal parameters can be obtained. Faults impacting those physical parameters, EIS measurements is a powerful tool to dynamically characterize a fuel cell and detect failure modes. Yang et al. (2021) reviewed the different existing EIS studies used to diagnose faults in SOFCs. The authors mainly presented papers using EIS to extract fuel cells model parameters. Zhang et al. (2021) presented the different existing PEMFC characterizations methods and their online application.

Toyota used EIS measurements in the Toyota Mirai to compute the membrane resistance (Maruo et al., 2017). This enables to evaluate the water quantity present in the stack during freeze stop and freeze start phases.

Fairweather et al. (2011) used EIS to model a Randles equivalent circuit. The authors used pseudo random binary sequences (PRBSs) as excitation signal. Onanena et al. (2012) detected flooding and drying based on EIS measurements. Bouaicha et al. (2017c) presented an embedded EIS for the Nexa PEM fuel cell. The authors used capacitors to generate the sinusoidal perturbation and the STM32F4 microcontroller.

The implementation of EIS in automotive applications is not easy. First, the AC component generation is not easily feasible in a classical electric vehicle powertrain system. The integration of an additional apparatus being able to inject an alternative current is usually too expensive for automotive applications. Moreover, the steady state is required for the measurement duration. The measurement duration depends on the injected AC current frequency.

However, various solutions have been presented to overcome those constraints.

For example, Narjiss et al. (2008) proposed to inject the AC current through the DC/AC/DC converter. This converter is an indispensable element in the classic fuel cell system to connect the fuel cell to the DC bus. Thus, there is no additional cost. Depernet et al. (2016) continued the work by proposing a control strategy to avoid that EIS measurements perturbed the DC bus voltage. In that sense, the EIS is suspended during load transient or when the battery or ultracapacitor state of charge is low. Therefore, the EIS procedure does not interfere the driving experience.

To reduce EIS duration time, Al Nazer et al. (2013a) proposed broadband excitations for electric vehicles. The concept is to inject multiple frequencies at once to obtain the impedance at multiple frequencies at the same time. In a first article (Al Nazer et al., 2013a), the authors proposed five different signal types: white noise, pseudo random binary sequences (PRBSs), swept sine, swept square; and square wave. In a second article (Al Nazer et al., 2013b), the authors selected pseudo random binary sequence and square waves. These signals have been chosen to facilitate the implementation in embedded applications as square waves are easier to inject than sine waves. The authors compared classical EIS (generated by sine waves) with square pattern signals. They concluded that square waves and pseudo random binary sequence signals led to a biased broadband identification, but with a very low dispersion, which is suitable for embedded EIS applications. In this article, the electronical circuit they used to generate the square waves is presented. It is composed on a transistor, a microcontroller, and a resistance.

Lu et al. (2019) injected a current pulse to generate the AC component at multiple frequencies at once. A continuous wavelet transform is applied to the voltage and current signals and the information is stored in the wavelet coefficients. The authors extracted the impedance from the wavelet coefficients distribution and the maximum likelihood estimation. This process is made at different frequencies to obtain the EIS spectrum.

Debenjak et al. (2015) presented their system to perform fuel cell EIS measurements for embedded applications. The conception steps of the device are presented in detail. The AC current is generated through the DC/DC converter. The authors applied a continuous wavelet transform using the complex Morlet mother wavelet. The voltage and current complex coefficients were extracted to estimate the impedance. The fault detection was based on a threshold value.

4.3. Partial synthesis

Inlet data is crucial for diagnosing faults in PEMFCs. Two main approaches can be found in the literature: passive measurements (voltage, pressure, temperature) and characterization methods (polarization curves, Electrochemical Impedance Spectroscopy (EIS)). The main advantage of passive measurements is that it does not require additional sensors. However, the data analysis is more complex because it is not a direct state indicator. Polarization curves are a static characterization method, but the characterization time is high and not suitable to vehicle applications. EIS measurements is presented in the second subsection. It is a powerful tool to dynamically characterize the fuel cell and the resulting EIS spectrum is an accurate state indicator. However, the AC component generation, and the steady state required during the characterization time make it hard to implement EIS in a vehicle. Solutions overcoming the problems are presented in Section 4.2.

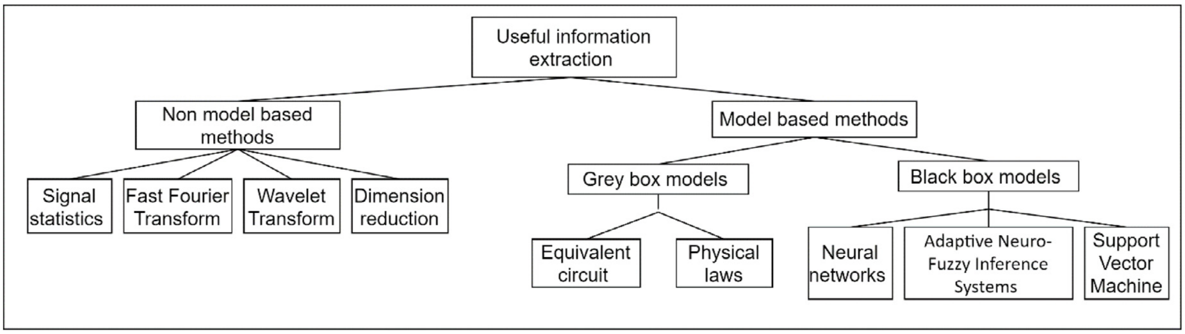


Fig. 4. Schematic summary of the different useful information extraction methods.

5. Useful information extraction methods

To perform diagnosis, the characteristics of the fuel cell response in faulty condition must be identified. This is the useful information identification step. There are two main approaches for useful information identification: the model based one and the non-model based one.

The model-based approach consists in comparing the real time measurements to a healthy fuel cell model. The non-model-based approach consists in processing the data to identify faulty fuel cell response characteristics.

Note that it is common to split diagnosis methods into model based and data-based methods in the literature. However, black box model methods are usually classified into data-based methods because they require a lot of data, although they use fuel cell model. In this article, useful information extraction methods are classified into model based and non-model-based methods to avoid any ambiguity (see Fig. 4).

5.1. Non model based

Non model based useful information extraction methods consists in analyzing signal patterns typically representative of the fault. Usually, these patterns are not directly observable and further data processing is often necessary. In this article the non-model-based data processing methods are separated into 3 groups: the statistical analysis, the time frequential analysis and dimensions reduction methods. Zheng et al. (2013) proposed a detailed review on the non-model-based methods used for fuel cell diagnosis. This review has been recently completed in Ref. Wang et al. (2021), where the authors present recent work using signal processing, pattern recognition, principal component analysis and Bayesian network methods.

5.1.1. Statistical analysis

Non model based methods goal is to identify features representing a fault. Statistics are widely used in signal processing. The main advantage is their ease of implementation as they do not require many computations and are easily interpretable.

Statistics are used in a lot of fields. Simple statistics as the mean, the standard deviation, the variance, etc. give information on the signal pattern. These basic features evolution is generally representative of the state of health of the fuel cell. For example, when a fault occurs, the mean voltage decreases. Flooding generates voltage oscillation as the water droplets blocks the gas channels before being expelled. Thus, the voltage oscillates, and the voltage variance increases.

Some statistics calculations describe the data distribution as the kurtosis and the skewness factors. They are more advanced statistics, yet their computation complexity is low and interpretation easy. The skewness factor is a measure of the asymmetry of the probability distribution. This enables to have an indication on the data distribution and on the trend of the signal. The skewness factor is complementary to the kurtosis factor as it also gives indication on the distribution function shape. The kurtosis factor represents the tailedness of the distribution function. If all the data points are centered around the same value, the kurtosis factor is less than 3. It is equal to 3 when there is a normal distribution and above 3 when there are extreme values.

5.1.2. Time frequential analysis

The frequential approach usually gives a lot of information in the signal processing field. In fuel cell diagnosis applications, two main methods are used for time frequency feature extraction: the windowed Fast Fourier Transform and the Wavelet Transform. For more time frequency features extraction methods information, please refer to Krishnan and Athavale (2018).

The Fast Fourier Transform (FFT) is a widely known algorithm that decomposes a temporal signal into a frequential signal. The most represented frequencies in the signal are highlighted and this can be very useful to detect a faulty condition. Note that the FFT is used with a sliding window as data are acquired over time. There is generally a compromise to make between time resolution (small time window) and frequency resolution (large time window). Also, the signal is usually transformed by a tapering function because edge artifacts could be observed on non-periodic signals (Broughton and Bryan, 2018).

The Wavelet transform is semblable to the FFT except that the tapering function is a wavelet function. Multiple types of wavelets can be used to transform the signal. The main advantages of the wavelet transform are their good time/ frequency resolution, and they recognize easier signal singularities than FFT. Two types of Wavelet transform exist: the continuous wavelet transform, and the discrete wavelet transform. In the discrete wavelet transform, the signal is separated into the high frequency part coefficients named approximations and the low frequency part coefficients named details.

From the FFT and the wavelet transform, it is possible to directly use the frequencies amplitude or wavelet coefficients or to perform a deeper analysis as for example the total harmonic distortion calculation. The total harmonic distortion is a measure of the signal frequencies spread which can be interesting to detect a fault.

5.1.3. Dimension's reduction

Some dimension reduction methods can be used to extract features. Indeed, the PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis) or FDA (Fisher Discriminant Analysis) consist of changing the observation space to reduce the number of observation dimensions. In other words, we change the axes orientation to be in a plane where we keep a maximum of information (for example, in PCA, the maximal information corresponds to variance maximization). Reducing the number of dimensions reduces the number of variables to represent information. Some studies use these methods after extracting features (statistics or frequencies for example). Other studies use the PCA on extracted raw data (voltage or pressure for example). The resulting PCA matrix has no physical meaning, it just represents data information in a different way. These methods can also be associated with a kernel function. The use of kernel functions enables to get further than linear transformation and to deal with more complex data patterns.

5.2. Model based

One approach to diagnose whether a fuel cell is in faulty condition or not is to build a model of the fuel cell describing its normal operating condition. This model will be then compared to a measurement of the operating fuel cell. The difference between the model and the measurement is called residual. If the residual is high, that may mean that a fault is occurring.

Different types of models exist. Physical models, based on the physical principles of the fuel cells are called "white box" models. However, those models are rare in the fuel cell field as many phenomena are not accurately understood and it is hard to determine the parameters values of the analytic equations. For fuel cells, physical equations are usually adapted by using empirical coefficients. Those coefficients are mainly obtained from experimental data. This type of model is called "grey box" models. Other models are not based on physical equations but are directly calibrated thanks to experimental data. This type of model, that can be seen as a mathematical model of the fuel cell, is called "black box model". A detailed review on the model-based methods used for fuel cell diagnosis can be found in Petrone et al. (2013). Recently, Wang et al. (2021) completed the recent literature regarding PEFMC diagnosis methods. The authors separated the model-based methods into analytical model-based methods and black box model methods. In their paper, the analytical model method groups parameter identification methods, observer-based methods and parity space methods while the black box model methods gather neural networks, fuzzy logic and machine learning methods. The review (Zhao et al., 2021) presents recent PEMFC models (physic based or not) used for real time control.

5.2.1. Grey box models

A common output of a fuel cell model is the stack voltage. Analytic voltage models can be based on the Butler–Volmer, Nernst or Fick's laws. The pressure is also an important parameter to model as pressure oscillations are representative of flooding for example. Some pressure models are presented in Section 5.3.

An interesting grey box model approach is the equivalent circuit models. The idea is to consider the fuel cell as an electric circuit composed of resistors, capacitors, and eventually inductances, constant phase elements or other electrical components. This type of model is easily understandable and the electrical model components represent physical fuel cell properties as the membrane resistance. The most common fuel cell equivalent model is the Randle's model. This circuit is composed of two resistors, a capacitor, and a Warburg element. One resistance represents the membrane resistance. The capacitor represents the charge accumulation phenomenon in the double layer area. The activation and the concentration losses are modeled by the other resistance and the Warburg element represents the mass transfer phenomenon.

5.2.2. Black box models

Black box models consist in finding mathematical equations representing the fuel cell by calibrating it using experimental data. Neural networks are very useful for this type of application.

As presented in Petrone et al. (2013), adaptive neuro-fuzzy inference systems (ANFIS) algorithm also uses neural networks. The fuel cell is modeled with fuzzy logic, but the membership function and the fuzzy rules are optimized thanks to a neural network.

In addition to classification applications (see Section 6.2), Support Vector Machines (SVM) can be used to model a fuel cell. Indeed, SVM can resolve regression problems. The principle of SVM regression is not to minimize the error in the least square error sense as classic linear regression methods do. In the SVM, a flexibility is introduced to have a certain error acceptance in the model.

5.3. Use of information extraction methods for fuel cell diagnosis

Statistical features have been used in Lin et al. (2019a). The authors extracted 9 statistical features (minimum, maximum, mean, median, the smaller and bigger quartile, standard deviation, skewness and kurtosis) of sensors data. Dimension reduction methods (principal component analysis, factor analysis and linear discriminant analysis) were applied to select the best features.

Benouioua et al. (2014a) applied a multi fractal analysis based on the wavelet transform modulus maxima to detect anodic starvation, cathodic starvation and sub pressure faults on PEMFCs.

Pahon et al. (2016) used the detail coefficients of a discrete wavelet transform to diagnose the high air stoichiometry fault on a PEMFC. Stack and cells voltages and the air pressure drop were used as input signals.

A PEMFC water content model has been performed by using the energy intensity of reconstructed vibrating voltage (EIV) based on wavelet transformation in Ma et al. (2020). The authors used the Sym20 wavelet basis function, and their model was applied to detect anodic flooding.

Mao et al. (2018) performed a Kernel principal component analysis to reduce the dimension number of sensor data to four. Then, a wavelet packet transform was applied to extract the features. Finally, the two most relevant features were selected using a Singular Value Decomposition technique.

Principal Component Analysis (PCA) was used by Zheng et al. (2021) to extract the most relevant features for air leakage and fuel starvation detection on an SOFC. PCA was applied on 74 variables and 10 have been selected as the correlation between the variable and the faults were high.

Mohammadi et al. (2015) applied an FFT to the voltage signal in order to detect fuel cell drying and flooding. The authors identified the 7 first harmonics as being the relevant features for drying and flooding detection.

Model based methods have also widely been used for fuel cell diagnosis purposes.

Gu et al. (2021) developed a PEMFC model to detect flooding using a Long Short Term Memory (LSTM) neural network. They implemented their flooding diagnosis tool to a microcontroller and tested it on an automotive fuel cell system. The algorithm was developed offline using Matlab and implemented online using Simulink.

Lu et al. (2019) modified a Randle model to better fit to the fuel cell. The Warburg was replaced by parallel resistance and capacitor and replaced the double layer capacitor by a Constant Phase Element (CPE). They used EIS generated with a current pulse injection and treated with a Morlet wavelet transform. The authors used a nonlinear least square method to identify the model components. Three model parameters are selected as features: the ohmic resistance, the polarization resistance and the resistance associated to mass transfer.

Laribi et al. (2016) also used the Randle model with a CPE to model the fuel cell under different hydration conditions. A feed forward neural network with one hidden layer has been used to identify the parameters of the Randel model in real time. The bias and weights are adapted using the gradient descent algorithm. The objective is to determine the hydration level in real time to adapt the control by adjusting the operating parameters.

Another equivalent circuit model has been presented in Bouaicha et al. (2017a) and Bouaicha et al. (2017b). An anode, membrane and cathode models are put in series to model a Nexa PEMFC. The elements of the fuel cell model are resistors, capacitors, and a voltage source. An additional model composed of a resistor and an inductance represents the impedance of the electric wires and other additional components.

A model based diagnosis method was employed in Ref. Steiner (2010). In their studies, the authors modeled a healthy PEMFC using an Elman Neural Network. The Elman Neural Network is a recurrent neural network with one hidden layer. The modeled variables were the cathodic pressure for flooding detection (Steiner, 2010) together with the stack voltage to detect both flooding and drying (Yousfi Steiner et al., 2011).

In their study, Lee et al. (2019) used five different residuals to detect faults in the stack, the air supply system, the water management system, the thermal management system and the fuel supply system. Those residuals have been extracted thanks to five models. The faulty thresholds have been defined as 3 times the standard deviation of the residuals obtained with nominal operating data.

Petrone et al. (2019) used a Fouquet model to extract internal PEMFC properties. To accelerate the impedance parameters identification, the authors bounded the parameters values by the intervals. Those intervals were found using a branch and bound method.

Arama et al. (2020) modeled the internal resistance at high frequency and the biasing resistance at low frequency to estimate the fuel cell state, flooded or dried. The authors used a neural network with 3 neurons input layer, two hidden layers composed of 20 and 10 neurons and two neurons in the output layer. The dynamic gradient descent algorithm was used to estimate the weights of the neural network.

With a similar approach, Laribi et al. (2019) used a feed forward neural network with two hidden layers of 20 and 10 neurons. This neural network enables to estimate the Randles model parameters, namely the membrane resistance, the double layer capacitance, the polarization resistance, the diffusion resistance, and the diffusion time constant. The values of those parameters indicated the health sate of the fuel cell and enabled to detect drying or flooding.

Pohjoranta (2015) modeled the voltage of a multi stack SOFC using six recurrent neural networks with 5 hidden neurons and 10 past inputs. They used the Apros and Matlab tools to extract online residuals and detect fuel cell degradation.

In Zheng et al. (2017), a reservoir computing method is used on frequential signals to diagnose PEMFCs faults. The studied faults were carbon monoxide poisoning, cathodic starvation, defective cooling and natural degradation.

Shao et al. (2014) developed a fuel cell model based on the Tafel equation and a heat transfer equation. This model was coupled to an ensemble ANN to diagnose faults in the PEFMC system.

A back propagation neural network is used in Wu and Zhou (2016) to model a healthy fuel cell. Voltage and cathodic pressure residuals are compared to a threshold to detect if the fuel cell is in faulty conditions or not. The inputs model are the hydrogen and air inlet pressure, the injected water flowrate in the humidifier, and the coolant flowrate. The authors proposed a control reconfiguration to adjust the operating parameters when a fault is detected.

Lebreton et al. (2015) also modeled voltage and cathodic pressure using a 3-layer feed forward neural network to detect flooding. The authors proposed a self-tuning PID to correct the fault in real time.

Barzegari et al. (2019) modeled the fuel cell voltage operating in a dead-end mode using ANN. Dead end mode cause water accumulation and flooding. In a first study, the authors used an ANN with one hidden layer. The oscillations due to the accumulation of water were not correctly modeled. Thus, the authors added one hidden layer and tested multiple architectures and activation functions. The best results were obtained with 14 neurons at each hidden layer.

It is worth noticing that the previous authors used different ANN structures (different number of hidden layers, different number of neurons per layer…). Defining the ANN structure is one of the trickiest part in artificial neural network model development. In fuel cell modeling application, the structure should be adapted to the stack type, the modeled fault type, the available computational capacity…

Vural et al. (2009) used an Adaptive Neuro-Fuzzy Inference System (ANFIS) to model a PEMFC voltage. The inputs on the model were the current density, fuel cell temperature, anode and cathode dew points and anodic and cathodic pressures. The authors implemented their model using Matlab.

As mentioned in Section 5.2.2, Support Vector Machines (SVM) can be used for classification or modeling purposes. In that sense, Zhong et al. (2006) modeled a Ballard PEMFC using SVM. The authors used a Radial Base Gaussian Function as a kernel function of the SVM. The inputs of the models were only current density and stack temperature.

One limit to model based methods to diagnose fuel cell faults is the evolution of the fuel cell response over ageing. To overcome those problems, some authors modeled fuel cell considering their degradation for prognosis purposes.

Jouin et al. (2013a) proposed a state of the art of the existing ageing PEMFC models. The authors clearly presented prognosis and health management process. Morando et al. (2016) modeled the fuel cell voltage by using an Echo State Network, which is a reservoir computing method. This prognosis method enables to predict the voltage. The voltage signal was filtered by the Symlet 5 wavelet, and the Hurst exponent was evaluated. Another prognosis approach is particle filters. This approach has been used to predict the remaining useful lifetime of fuel cells in Jouin et al. (2013b) and of hybrid electric vehicle prognosis applications in Yue et al. (2021).

Fuel cell diagnosis might become obsolete with time because fuel cell response changes with ageing/ degradation. Combining diagnosis and prognosis enables to diagnose faults not only at the beginning of fuel cell life, but also when degradation already occurred.

5.4. Partial synthesis

Useful information extraction is either done by analyzing and processing data (non-model-based methods) or to compare real time measurements to a fuel cell model. Sections 5.1 and 5.2 present the theory of common useful information extraction methods of non-model-based approaches and model-based approaches respectively. Section 5.3 is dedicated to the presentation of practical studies for useful information extraction in fuel cells. All those studies are summarized in Table 2. It is interesting to notice that a lot of methods have been applied offline but only few studies have been implemented to be used in embedded applications.

Table 2
Useful information extraction methods references.

Useful information extraction method	Offline useful information extraction method	Embedded useful information extraction method
Signal statistics	Lin et al. (2019a)	
FFT	Mohammadi et al. (2015) and Zheng et al. (2017)	
Wavelet transform	Benouioua et al. (2014a), Pahon et al. (2016), Ma et al. (2020) and Lu et al. (2019)	Debenjak et al. (2015)
Dimension reduction	Mao et al. (2018), Zheng et al. (2021) and Lin et al. (2019a)	
Equivalent circuit model	Lu et al. (2019), Laribi et al. (2016), Bouaicha et al. (2017a,b) and Petrone et al. (2019)	
Physical laws model	Lee et al. (2019) and Shao et al. (2014)	
Neural networks	Steiner (2010), Arama et al. (2020), Laribi et al. (2019), Shao et al. (2014), Wu and Zhou (2016), Lebreton et al. (2015) and Barzegari et al. (2019)	Gu et al. (2021) and Pohjoranta (2015)
Adaptive Neuro-Fuzzy Inference Systems	Vural et al. (2009)	
Support Vector Machine	Zhong et al. (2006)	

6. Classification algorithms

Once the relevant information has been extracted, the classifier allows to determine whether the real time extracted samples correspond to a fuel cell faulty or healthy state.

A very simple way to classify is to determine thresholds that separate healthy from faulty data. This type of classification is also called linear classification. References using this method are available in Section 6.3. However, most of the time, this threshold separation is not sufficient to accurately determine the samples class because the separation can be more complex than a simple line. In this sense, machine learning tools are used. These tools enable to recognize complex patterns, in numerous dimensions, and accurately classify the new samples.

Several classifiers are based on distance calculation as knn, kmeans algorithms. Others depend on statistics such as the Bayesian classifiers. It is also possible to model the pattern of data for different class, as support vector machines and neural networks do. Some classifiers rely on the principle of decisions trees such as random forest and decision tree algorithms. Finally, the Density Based clustering algorithm (DB scan algorithm) combines distance calculation and the density function of the points to group the points having similar properties. Numerous other classification algorithms exist. However, the ones enunciated above are the most common ones used for diagnosis methods.

The classification algorithms are grouped into 3 groups in this article: the supervised classification algorithms, the non-supervised classification algorithms and a third group that contains either non machine learning tools, or ensemble methods that can use supervised and non-supervised classifiers.

Supervised classification algorithms tools compare a new sample to an existent database with known labels. On the contrary, unsupervised classifiers group the existent database with the new sample without prior information of the classes. Note that for comprehension purposes, the algorithms are presented as binary classifiers in the next sections: there are only two classes. It is however possible to apply those methods to multiple classes.

In this section, the firsts two subsections explain the different supervised and unsupervised classifiers. The third subsection present classification methods using thresholds and ensemble methods. References of the use of classifiers for fuel cell diagnosis are given in the fourth subsection.

6.1. Supervised classification algorithms

Supervised classification algorithms gather the k-nearest-neighbor (KNN), the Support Vector Machine (SVM), the decision tree, the random forest, the naïve bayes, neural networks classifiers, etc. The term supervised means that the labels of the training dataset are known. In other words, when a new sample is classified by a supervised classifier, it is compared to a bench of other samples, which labels have been identified beforehand.

As presented on Fig. 6a, the k-nearest-neighbor (KNN) algorithm consists in:

- calculating the distance of a new incoming point with all the database points (of which labels are known).
- The attributed class to the new incoming point is the majority class of the k nearest points.

The two variables of the KNN algorithm are the choice of k and the type of distance (Euclidian, Manhattan, cosine…). More details on the KNN algorithm operation and application can be found in Li et al. (2014). The main advantage of the KNN algorithm is that it is easy to understand its principle and use. However, the distances between the new sample and all the database points must be computed at each iteration, which makes it time consuming.

Three mains naïve bayes classifiers exist: the gaussian naïve bayes classifier, the multinomial naïve bayes classifier and the Bernoulli naïve bayes classifier. Their names stand for the utilization of the Bayes theorem and the term naïve refers to the naïve assumption that the variables are independent. The classifiers determine the likelihood of being in a class or in another. The class that is the most probable is chosen. The Gaussian Naïve Bayes algorithm (presented in Fig. 6e) employs a normal distribution function. This gaussian distribution function enables to estimate a probability density function that is a statistical estimation of the similitude between the new sample and each class. The advantages of the gaussian naïve bayes classifier is that the probability of having predicted the good class is given. Moreover, only the mean and the standard deviation for each class and each feature must be stored. That drastically reduces the required storage space since for the other algorithms each data point feature must be stored. However, this algorithm is strongly dependent on the data and features quality since it has no additional parameters and is only based on statistics.

Support Vector Machine (SVM) is another powerful classification tool. A SVM classifier separates the data by creating a hyperplane. SVM uses the best separating hyperplane by maximizing the distance between the separating line and the extreme samples (this consists of maximizing the margin on Fig. 6b). These extreme samples are named the support vectors.

Another approach is the decision tree one. The principle is to build a tree composed of decision nodes and leaf nodes. Fig. 6c presents a very simple decision tree applied to fuel cell faults.

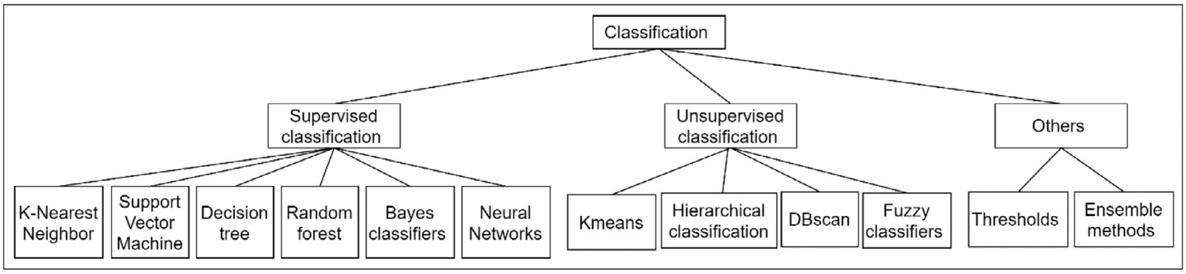


Fig. 5. Schematic summary of the different classification algorithms.

The main advantage of this method is that its representation is straightforward, and the results are easily interpretable. The algorithm principle is as follows:

- First, the best feature is determined. It can be made by the measure of entropy or the GINI coefficient calculation (see Section 2.2.3). The best feature is the one that is the most representative of the fault.
- Then, the decision rule splits the data into smaller subsets.
- The second-best feature is selected, and the classification is made once again. This is repeated for all the features.

A variant of the decision tree classifier is the random forest algorithm. As presented on Fig. 6d, the idea is to build a forest of multiple decision trees. The selected class is the mode of the prediction of all the decision trees in the forest. Random forest belongs to ensemble method which combines multiple classifiers. Using multiple decision trees usually improves the classification accuracy.

Neural networks can also be used for classification purposes (see Fig. 6f). The training of the neural network for classification problems consists in adjusting the weights and bias values to accurately predict the correct class for a dataset (whose labels are known). Once the weights and bias have been optimized (this is made offline), a new sample can be classified.

6.2. Clustering algorithms

Unsupervised classification algorithms, also named clustering algorithms consist of grouping a data set into a fixed number of groups. The biggest difference between supervised and non-supervised classifiers is that the labels are not known in advance for unsupervised classifiers. In other words, non-supervised classifiers consist in grouping the most similar samples of a dataset together. This type of classification is very interesting as preliminary labeling is not necessary. However, the biggest constraint of this type of classification is to select the right number of groups, or the threshold to consider a region as dense. Unsupervised classification algorithms gather the K-means algorithm, the hierarchical classifications, the DBscan (density-based spatial clustering of applications with noise) algorithm and fuzzy classifiers.

The K-means algorithm, presented on Fig. 7a, resembles to the KNN algorithm, but the KNN is supervised while the K-means is not.

- At the first iteration, K centers are randomly set. These are the centroids of each group.
- Then, the distance between the centers and each data point is computed.
- The closest data from each center are grouped in the same cluster.
- The barycenter from each cluster is computed and set as the new centers.
- Step 2, 3 and for are repeated until the stopping criterion is reached.

The K-means algorithm is fast and can easily be implemented on a large amount of data. In this algorithm, there are several parameters to set. The first one is K which corresponds to the number of clusters. The second one is the stopping criterion. It can be a fixed number of iterations or the barycenter stability (they do not move between two iterations). The third one is the chosen distance. As for the KNN algorithm, the Euclidian distance is the most common, but the Manhattan distance, the cosine distance… can also be used. The initialization can also be improved (as done in the K-means++ algorithm) to speed up the convergence and avoid local optimum. In that case, the initial centers are not randomly set but chosen to be coherent with data distribution.

Another approach consists in introducing fuzzy logic in K-means algorithm. This algorithm is named the C-means algorithm (see Fig. 7d). To introduce fuzzy logic, the distance between the point and the barycenter is considered. Thus, a membership degree to the cluster can be introduced. The smallest the distance between the point and the center is, the highest the membership degree is.

Hierarchical clustering is also a powerful technique to group similar data together. Its representation (see Fig. 7b) uses a dendrogram. The principle for the ascendant hierarchical classification is:

- Start with setting one cluster per data point.
- Group the 2 datapoints that are the most similar.
- Continue until obtaining only 1 cluster (all the data is grouped together)

Once this dendrogram has been computed, it is possible to choose the number of desired clusters. The biggest advantage of this clustering type is that once the dendrogram has been build, it is very easy and fast to compare the number of chosen clusters.

DBscan (Fig. 7c) means density-based spatial clustering of applications with noise. The main difference between this method and the other clustering methods is than the number of clusters does not have to be set in advance. The fixed number of clusters is replaced by a similitude criterion. The principle of this algorithm is that clusters are defined by high density regions separated by low density regions. In other words, when samples are similar, they are grouped in the same area and the density is high. The two parameters of this algorithm are the distance type (Euclidian, cosine…) and the minimum number of points to consider a region as dense. This parameter should be defined with care because this has a big impact on the classification results.

6.3. Other classifiers

A very simple way of classifying that is widely used is to fix a threshold value. This method is particularly used in

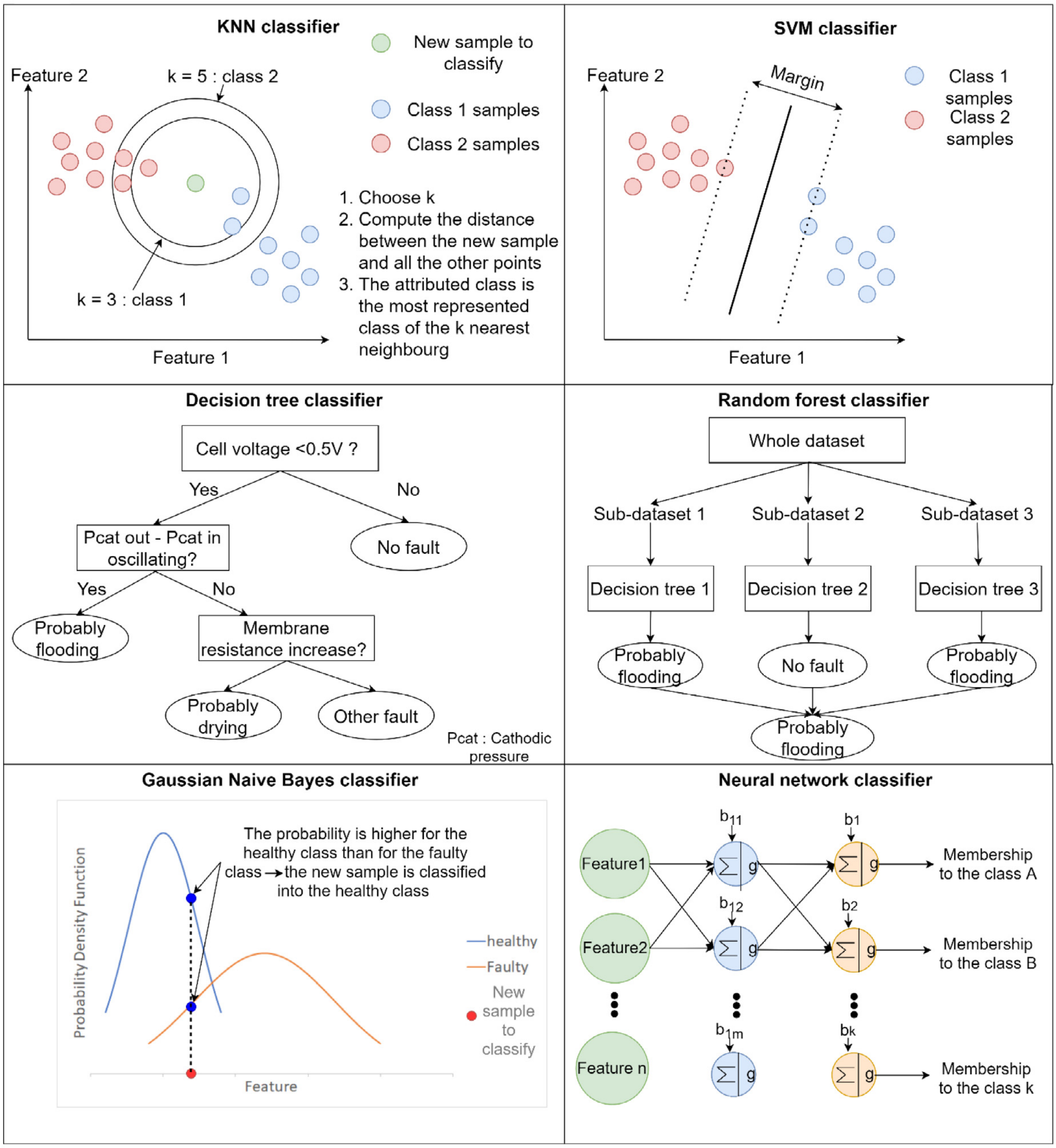


Fig. 6. Illustration of supervised classification algorithms working principle.

model-based methods (see Section 5.2). In model-based methods, the generated residuals increase when a fault occurs. The linear classification consists in comparing the residual to the threshold value: if the residual is above the threshold, it is classified into the faulty class. The main disadvantage of this method is the choice of the threshold value. It is often determined by trial-and-error methods.

Ensemble method is simply a combination of multiple classification methods. The idea is instead of choosing the best classification algorithm (which is often tricky), multiple classification algorithms are used, and the attributed class considers the results of multiple methods. This improves the classification robustness a lot. A common ensemble method is the random forest algorithm which uses multiple decision trees. The classification methods presented in the previous sections can also be combined. This is also an ensemble method.

6.4. Use of classifiers for fuel cell diagnosis

KNN algorithm has been widely used to perform fuel cell diagnosis. Onanena et al. (2012) tested a KNN algorithm and a multiclass LDA classifier to detect flooding and drying on a PEMFC. EIS experimental data have been treated and two set of features have been selected and compared. The first set represented the real or imaginary parts of the EIS spectrum at 4 different frequencies. The second set was composed of extracted physical parameters of the EIS measurements, namely the internal resistance, the polarization resistance, and the value of the maximal phase. On the two tested data set in this article and after a k fold cross validation, the KNN classifier revealed to have better performances than the multiclass LDA classifier.

Li et al. (2014) also used the KNN algorithm to detect flooding and drying faults on a 20 cell PEMFC. The considered signals were

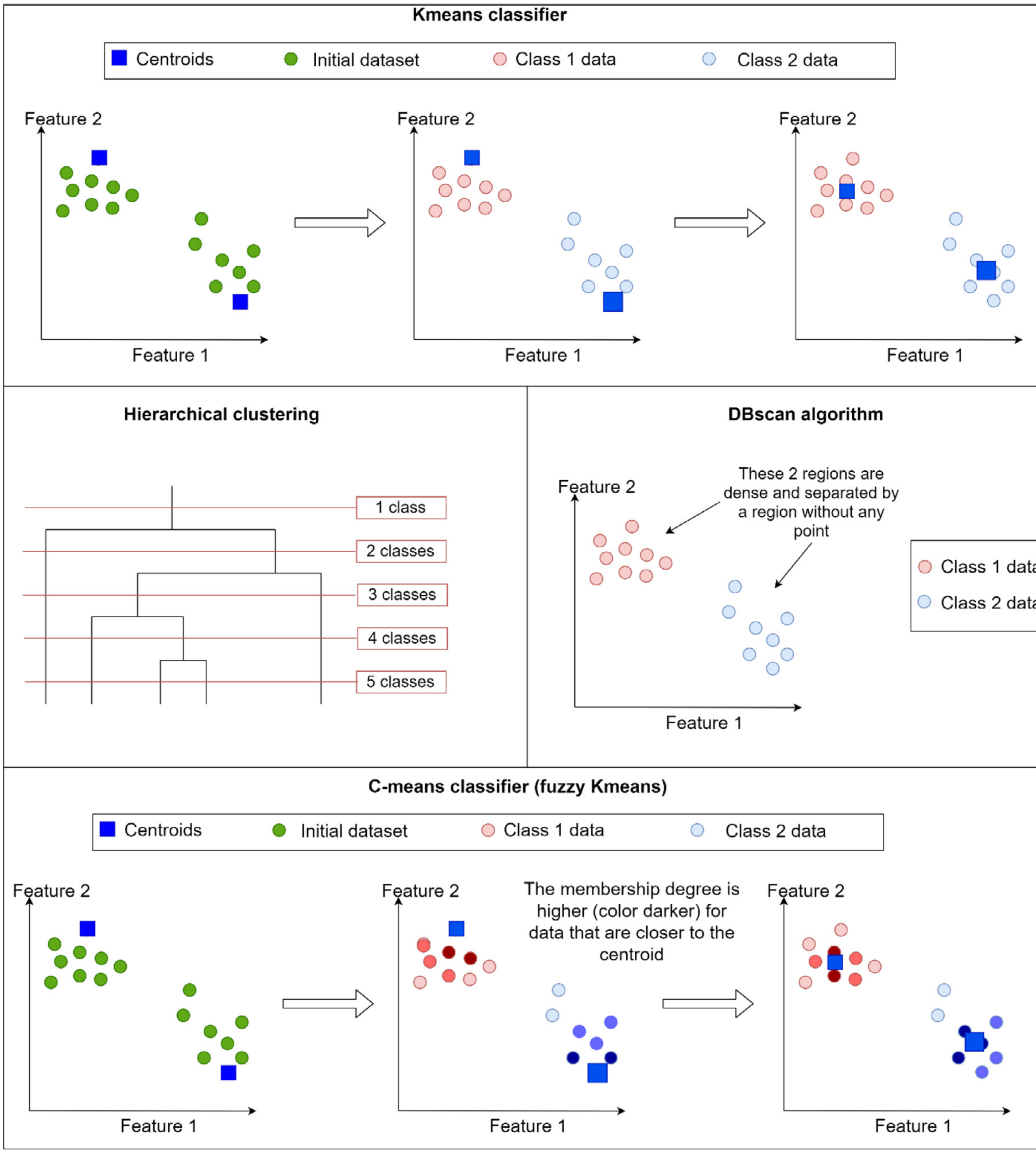


Fig. 7. Illustration of unsupervised classification algorithms working principle.

cell voltages and four set of features have been automatically extracted by using FDA, KFDA, PCA and KPCA. In their study, they also performed classification using a gaussian mixture model (GMM) and a Support Vector Machine (SVM) classifier. It resulted that with the dataset analyzed, all the classifiers had a good classification rate above 90% on their testing dataset.

Benouioua et al. (2014b) used a KNN and a SVM classifier to detect anodic starvation, cathodic starvation, and sub pressure faults. The features have been extracted with wavelet transform modulus maxima algorithm and selected with a Minimum Redundancy Maximum Relevance method. On the dataset tested by the authors, the KNN classifier appeared to have equal or better classification rate than the SVM classifier.

Lee et al. (2019) proposed a hierarchical fault diagnosis at the component level. They detected faults in the stack, the air supply system, the water management system, the thermal management system, and the fuel supply system. They used a model-based approach. They first detected the faulty state when residuals where greater than thresholds values. Those residuals were then classified with 5 different classifiers: a KNN, a SVM, an Artificial Neural Network, a Naïve Bayes classifier, and a Discriminant Analysis (DA) method. Faults were generated on a fuel cell test bench and all the classifiers were able to detect those faults.

Mao et al. (2018) used a K-means classifier to detect flooding. They used a wavelet packet transform to extract the features. They reduced the dimensions with both a KPCA (Kernel Principal Component Analysis) and a SVD (Singular Value Decomposition) algorithms to classify the relevant information only.

Lin et al. (2019a) developed a diagnosis algorithm to detect fuel cell faults. After testing seven classifiers: decision tree, random forest, Adaboost, KNN, ANN, and two SVM, the authors selected the random forest algorithm as it was the one that had the better results to detect faulty conditions on their data set.

Escobet (2014) used a Fuzzy Inductive Reasoning methodology to detect and identify faults in a fuel cell system. First, the detection process consists of comparing measured data which are fuzzified and defuzzified to thresholds. The fault identification step is made possible by modeling the different faults. The Fuzzy Inductive Reasoning uses a variant of the KNN rule.

Lim et al. (2021) diagnosed five component failures of the fuel cell cooling system. The detected failures were the pump performance degradation, the radiator fouling, the tube clogging, the fan disabled and the pump disabled states. They used 5 temperature and 5 pressure sensors to identify component failures with a SVM classifier.

Wu and Ye (2016) developed a least square support vector machine classifier to detect anode poisoning and cathode humidification on a solid oxide fuel cell. They added two hidden semi-Mark models to compute the remaining useful lifetime of the fuel cell.

Refs. Li et al. (2018) and Zheng et al. (2021) used a SVM classifier combined with a PCA to diagnose a solid oxide fuel cell. In Li et al. (2018), the authors used a multi label SVM with a Radial Basis Function kernel classifier to detect fuel and air leakage at different locations in the fuel cell system. The main advantage of a multi label classifier is that it requires much less data than a single label classifier because it detects the combination of multiple faults with single faults data. The authors used temperature, pressure, and gas flow rate data in different location in the system. Then, a PCA is applied to reduce the dimensions while keeping the relevant information. The data comes from a solid oxide fuel cell system model developed by the authors. Zheng et al. (2021) detected air leakage and fuel starvation in a SOFC. They also used a Radial Basis Function as a kernel for the SVM classifier on experimental data from a 1 kW SOFC. Before applying PCA, the sensor data that had the highest variance has been selected. The authors obtained better results with the SVM classifier than with the random forest algorithm and the artificial neural network used as a classifier.

In Ref. Costamagna et al. (2019b), the authors used a SOFC system model to train their classifiers. This solution avoids time consuming experiments to calibrate the machine learning algorithms. Fours faults were detected: stack degradation, air leakage, fuel leakage and reformer degradation. The authors dealt with the incapacity of supervised machine learning classifier to accurately detect unknown faults patterns. To overcome the problem, they first generated the training database using a SOFC system model. Then, they used a domain adaptation technique, with signals containing random errors of maximum 2% to update the dataset and be able to detect off design operating conditions.

Li et al. (2019) presented a spherical shaped multi-class support vector machine with a diagnosis rule to detect faults of PEMFCs. The authors used high precise cell voltage sensors as input data and a fisher discriminant analysis to extract relevant features of cell voltage data. An incrementation of the database to increase the robustness of fault classification with fuel cell ageing is also proposed. The approach has been implemented in a microcontroller for online applications.

Lu et al. (2019) used a binary tree SVM to detect drying, flooding and air starvation. They used equivalent circuit parameters as features. The authors developed and implemented a fast EIS measurement method. The diagnosis system has been tested online by sending the measured data through USB to a computer. The host computer uses C++ and Matlab to estimate EIS and perform diagnosis.

In Ref. Zhou and Dhupia (2020), the authors used a relevance vector machine to classify features obtained by an Orthogonal Linear Discriminant Analysis. The authors presented an adaptation of the database to be robust to fuel cell ageing. The features were single cell voltages projected in a 5-dimensional feature space. The authors incremented the database when the posterior probability of fault detection was above a threshold.

Park et al. (2021) used seven multi-layer perceptron neural networks having each a single output to diagnose thermal faults in a PEMFC. Each neutral network is associated with one fault or the normal state and classifies weather the fuel cell is in this faulty condition or not. The studied failure modes in this article are the stack performance degradation, sudden pump shut down, sudden radiator fan shut down, pump performance degradation, tube clogging and radiator fouling.

Li et al. (2015) used a Directed Acyclic Graph Support Vector Machine to detect short circuit, a cooling water circulation problem, high and low air stoichiometry, and CO poisoning. Directed Acyclic Graph Support Vector Machine consists of solving multiple binary SVM which enables to distinguish two classes. Fisher discriminant analysis is performed on voltage signals to extract features.

Fan et al. (2013) detected PEMFC faults using a naive bayes classifier. The classifier inputs were residuals computed with an analytical model and experimental data.

Kim et al. (2012) developed a Hamming Neural Network as a pattern recognition tool to detect faulty cells in a PEMFC fuel cell. The authors used the voltage measurement to detect anomalies in the fuel cell.

Polverino et al. (2015) presented a diagnosis method for SOFCs systems. The authors used healthy and faulty models to generate residuals. The fault detection and identification has been performed using thresholds values to classify the residuals.

In Ref. Steiner (2010), a healthy PEMFC using an Elman Neural Network has been modeled. In their studies, the authors classified the residuals into flooding or non-flooding class by defining a threshold (Steiner, 2010). In another paper the same approach was presented to also detect drying (Yousfi Steiner et al., 2011). The thresholds have been obtained after several error-trials.

Wu and Zhou (2016) also classified residuals by defining voltage and cathodic pressure residuals thresholds. The authors proposed a control adjustment when faults were detected.

An ensemble ANN method was used in Shao et al. (2014) to detect cooling circuit failure, the increase of fuel crossover and faults in the air or hydrogen supply system. Four back propagation ANN were composing the ensemble ANN. The classification accuracy of each of the sub-ANN was between 75 and 85% while the ensemble ANN appears to have a good classification rate of 92%.

The fault location is identified in Mohammadi et al. (2015) by classifying frequential voltage features thanks to a feed forward two layer ANN.

6.5. Partial synthesis

Classification takes the useful information extracted as an input and give the class (in our case faulty, healthy and type of faults) as an output. Classification algorithm is probably the most highlighted section in the literature. In this article, classification methods are grouped into supervised classification, unsupervised classification and other methods that gather thresholds classification and a combination of several classifiers (called ensemble methods). A scheme with the different classification approaches is presented on Fig. 5.

The first 3 subsections explain the principle of each classification algorithm. Section 6.4 presents the use of classification algorithms for fuel cell diagnosis in the literature. Table 3 summarizes the different papers found in the literature and group it by method. As for useful information extraction methods, it is interesting to notice that only few papers deal with the implementation of classification algorithms for embedded applications.

Table 3
Offline and online classification methods references.

Classification method	Offline classification references	Online classification references
KNN	Onanena et al. (2012), Li et al. (2014), Lee et al. (2019) and Lin et al. (2019a)	
SVM	Li et al. (2014), Lee et al. (2019), Lin et al. (2019a), Lim et al. (2021), Li et al. (2018), Zheng et al. (2021), Costamagna et al. (2019a,b) and Li et al. (2015)	Li et al. (2016, 2019) and Lu et al. (2019)
Decision tree	Lin et al. (2019a)	
Random forest	Lin et al. (2019a), Zheng et al. (2021) and Costamagna et al. (2019b)	
Bayes classifier	Lee et al. (2019)	
Neural networks	Lee et al. (2019), Lin et al. (2019a), Zheng et al. (2021), Park et al. (2021), Shao et al. (2014) and Mohammadi et al. (2015)	
Kmeans	Mao et al. (2018)	
Thresholds	Steiner (2010), Yousfi Steiner et al. (2011), Polverino et al. (2015) and Wu and Zhou (2016)	Lebreton et al. (2015)
Ensemble method	Lin et al. (2019a), Zheng et al. (2021) and Costamagna et al. (2019b)	

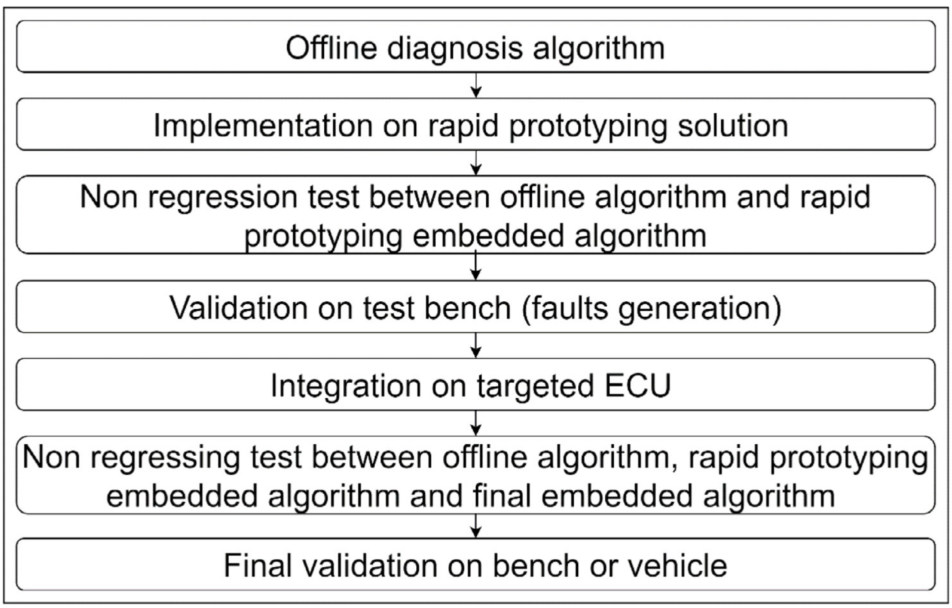


Fig. 8. From offline to embedded diagnostics.

7. From offline to embedded diagnostics

From the last sections, it appears that the offline diagnostics has largely been covered in the literature, yet few studies extend the diagnostic tool to a product that can be used in a vehicle. The offline step is mandatory to design the diagnosis algorithm, however, some additional work is needed to implement the diagnostic tool in a vehicle.

Fig. 8 enunciates the main steps to implement an offline diagnosis algorithm in a vehicle.

First, a rapid prototype is created. The objective is to test the algorithm in a microcontroller which has more flexibility than an automotive Electronic Control Unit (the automotive microcontroller). Rapid prototyping systems are flexible systems that help algorithm developers worldwide bring ideas to life in a real environment, with real sensors and actuators, and in real time. With rapid prototyping, the algorithm can be tested in real-time and rapidly corrected, because there are less constraints than on an automotive Electronic Control Unit.

Secondly, the algorithm is validated using the same database than the one used offline. The offline algorithm results should be very similar to the real time algorithm results (results may slightly differ, due to variable data precision). With this non-regression test, mistakes in the algorithm implementation can be detected.

Thirdly, the rapid prototype is connected to a test bench. Faults are deliberately generated on this test station, to test if the rapid prototype correctly detects the faults.

Once the rapid prototype has been validated, the algorithm can be implemented in the targeted automotive Electronic Control Unit (ECU). Note that the diagnosis algorithm can be implemented to an additional ECU specifically dedicated to diagnosis, or on the existing ECU (dedicated to control). The rapid prototype hardware must be selected according to the type of ECU, as the programming languages should be compatible. As mentioned in Section 2.4, the type of ECU should be chosen with care, with specifications that are adapted to the vehicle constraints. It should also apply with automotive safety regulations, to the diagnosis algorithm complexity and storage requirements, whilst remaining at low cost and volume.

Then, the algorithm is tested using the same database than the one used offline and for the non-regression test of the rapid prototype. Once again, the results of the ECU should be very similar to the rapid prototype ones and to the offline algorithm ones.

Finally, the ECU containing the diagnosis algorithm can be tested on the vehicle. Faults can be generated in the vehicle by imposing associated operating conditions. To that aim, one must have access to the control unit of the fuel cell. The embedded algorithm can then be validated in the vehicle.

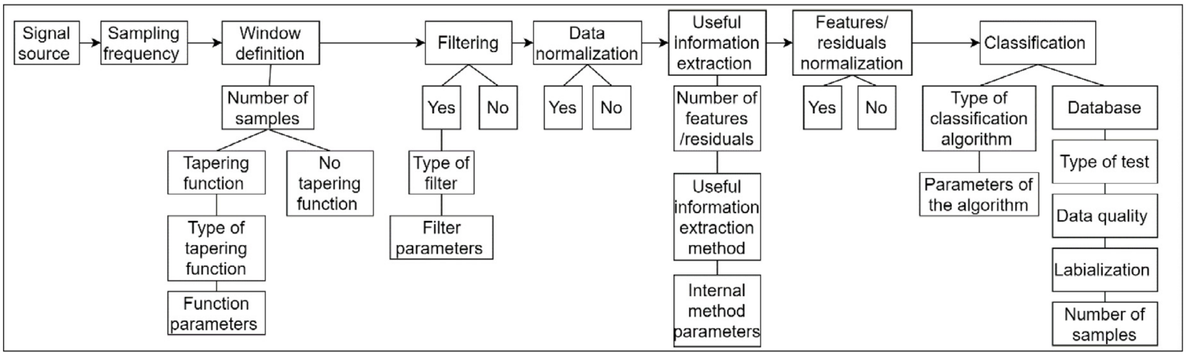


Fig. 9. Summary of the parameters impacting diagnosis.

8. Conclusion

Fuel cells are prone to faults during operation. Fault occurrence can be limited by a good fuel cell control but not avoided. Faults cause fuel cell degradation and decrease the remaining useful lifetime of the fuel cell system.

Fuel cell diagnosis enables real-time fuel cell state of health estimation. Fault tolerant control actions can be taken to correct faults and limit degradation. In other words, fuel cell diagnosis and fault tolerant control improve the fuel cell lifetime while using the best operating range.

In this article, the main methods to diagnose a fuel cell stack in a vehicle during operation have been presented. The objective of this paper is to give to the reader an overview of the existing methods to diagnose fuel cells and their constraints in vehicles.

As mentioned in Section 2, the database, the filter type, the useful information extraction method, and the classifier method must be determined considering the embedded constraints. The methods selection is a compromise between the diagnosis accuracy and reliability on one side and the cost, volume, and impact on the driving experience on the other side. Note that the computational time and the memory can be included in the cost constraint as it is possible to use more efficient microcontrollers, but the associated cost increases. The safety regulations are of course very important, however, the algorithms themselves are not questioned but the way of their implementation is. Therefore, it is not necessary to take the safety regulations into consideration for the algorithms' choice. Finally, the used sensors must also be noninvasive sensors because they must be implemented in a vehicle.

Regarding the measurements and characterization methods, two main approaches exist for vehicle applications: the one with passive sensors (voltage, pressure…) and the one with EIS characterization. The main advantage of using passive sensors is that there is no additional cost as those sensors are required for the system control. EIS measurements give more information than passive sensors on the fuel cell internal state. However, it is not easy to generate sinusoidal current signal in a vehicle and the fuel cell must be in a stabilized operating state during the measurement. In other words, EIS measurements give more information than passive sensors measurements, but their implementation is harder, and the associated cost is higher than the one for passive sensors.

Model-based and non-model-based methods can be used to extract useful information for diagnosis purposes. The chosen method depends on the available computation capacity in the vehicle microcontroller, on the storage capacity, and on the available experimental data available. The useful information extraction is the most important step of the diagnosis process. The features or model quality, and accordingly the database quality, drastically influence the classification results.

Classification methods are various. Obtaining a standardized optimal classifier is not possible, as its relevance strongly depends on the application. The main restrictive parameter in embedded application is the computational complexity of the algorithm because the Electronic Control Unit (ECU) must be capable of performing the calculations in a matter of seconds. Some algorithms as the gaussian naïve bayes classifier can also reduce the required storage space attributed to classification as only the mean and the standard deviation must be stored. Once again, it is important to highlight that the classification accuracy is mainly linked to the quality of its inputs: the useful information extraction method chosen and the database quality.

In addition to measurement or characterization method, useful information extraction method and classification method, a bench of other parameters influence diagnosis results, as presented on Fig. 9.

The signal source corresponds to the chosen measurement or characterization method presented in Section 4. Typically, this corresponds to cell or stack voltage measurements, parameters from EIS or polarization curve characterizations, temperature or pressure sensors…The sampling frequency of the measurement also impacts diagnosis results. When decreasing the sampling frequency, the data resolution decreases. As a result, some faults may become undetectable. When increasing the sampling frequency, the required storage capacity increases. Furthermore, with a higher data resolution, more perturbation signals (noise) visually appear in the data (see Shannon theorem).

Data acquisition windows also has an influence on diagnosis results. The number of samples per window as well as the tapering function type act as a filter. Varying those parameters (number of samples per window and tapering function type) is a good practice to check the robustness of diagnosis results.

Multiple normalization or standardization techniques exist. Normalization is required when the inputs for features extraction or classification do not have the same order of magnitude (for example temperature and voltage). It is common to normalize the data between 0 and 1 or standardize the data between −1 and 1.

Finally, useful information extraction methods have been explained in further details in Section 5 and classification methods in Section 6. The parameters enunciated on Fig. 9 are not specific to automotive application, and even not specific for fuel cell applications. The same workflow can be used for other diagnosis applications. However, in addition to the parameters represented on Fig. 9, one should also optimize the computational time, and the storage capacity for embedded applications. The final objective is to obtain the most accurate and universal diagnosis tool (valid for the highest number of situations in a vehicle).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing

interests: This work has been supported by Symbio, a Faurecia Michelin hydrogen company.

This work has also been supported by the EIPHI Graduate School (contract ANR-17-EURE-0002) and the Region Bourgogne Franche-Comté.

Acknowledgments

We would like to thank Luc Leydier (Symbio) for having taken the time to read through this paper and provide his comments. This work has been supported by Symbio, a Faurecia Michelin hydrogen company, France.

This work has also been supported by the EIPHI Graduate School (contract ANR-17-EURE-0002) and the Region Bourgogne Franche-Comté.

References

Al Nazer, R., Cattin, V., Granjon, P., Montaru, M., Ranieri, M., 2013a. Broadband identification of battery electrical impedance for HEVs. IEEE Trans. Veh. Technol. 62 (7), 2896–2905. http://dx.doi.org/10.1109/TVT.2013.2254140.

Al Nazer, R., Cattin, V., Granjon, P., Montaru, M., Ranieri, M., Heiries, V., 2013b. Classical EIS and square pattern signals comparison based on a well-known reference impedance. In: 2013 World Electric Vehicle Symposium and Exhibition (EVS27), Barcelona. pp. 1–7. http://dx.doi.org/10.1109/EVS.2013.6914874.

Andújar, J.M., Segura, F., Isorna, F., Calderón, A.J., 2018. Comprehensive diagnosis methodology for faults detection and identification, and performance improvement of air-cooled polymer electrolyte fuel cells. Renew. Sustain. Energy Rev. 88, 193–207. http://dx.doi.org/10.1016/j.rser.2018.02.038.

Arama, F.Z., Mammar, K., Laribi, S., Necaibia, A., Ghaitaoui, T., 2020. Implementation of sensor based on neural networks technique to predict the PEM fuel cell hydration state. J. Energy Storage 27, 101051. http://dx.doi.org/10.1016/j.est.2019.101051.

Aubry, J., Steiner, N.Y., Morando, S., Zerhouni, N., Hissel, D., 2020. Fault tolerant control of a proton exchange membrane fuel cell based on a modified failure mode and effect analysis. In: 2020 IEEE Vehicle Power and Propulsion Conference (VPPC), Gijon, Spain. pp. 1–5. http://dx.doi.org/10.1109/VPPC49601.2020.9330864.

Barzegari, M.M., Rahgoshay, S.M., Mohammadpour, L., Toghraie, D., 2019. Performance prediction and analysis of a dead-end PEMFC stack using data-driven dynamic model. Energy 188, 116049. http://dx.doi.org/10.1016/j.energy.2019.116049.

Behravan, A., Meckel, S., Obermaisser, R., 2019. Generic fault-diagnosis strategy based on diagnostic directed acy- clic graphs using domain ontology in automotive applications. p. 5.

Benmouna, A., Becherif, M., Depernet, D., Gustin, F., Ramadan, H.S., Fukuhara, S., 2017. Fault diagnosis methods for proton exchange membrane fuel cell system. Int. J. Hydrog. Energy 42 (2), 1534–1543. http://dx.doi.org/10.1016/j.ijhydene.2016.07.181.

Benouioua, D., Candusso, D., Harel, F., Oukhellou, L., 2014a. Fuel cell diagnosis method based on multifractal analysis of stack voltage signal. Int. J. Hydrog. Energy 39 (5), 2236–2245. http://dx.doi.org/10.1016/j.ijhydene.2013.11.066.

Benouioua, D., Candusso, D., Harel, F., Oukhellou, L., 2014b. PEMFC stack voltage singularity measurement and fault classification. Int. J. Hydrog. Energy 39 (36), 21631–21637. http://dx.doi.org/10.1016/j.ijhydene.2014.09.117.

Bouaicha, A., Allagui, H., Aglzim, E.-H., Rouane, A., Mami, A., 2017a. Validation of a methodology for determining the PEM fuel cell complex impedance modelling parameters. Int. J. Hydrog. Energy 42 (17), 12738–12748. http://dx.doi.org/10.1016/j.ijhydene.2017.01.114.

Bouaicha, A., Allagui, H., Mami, A., Aglzim, E.-H., Rouane, A., 2017b. Parameters identification of the complex impedance model of the PEM fuel cell using matlab/simulink. In: 2017 International Conference on Green Energy Conversion Systems (GECS), Hammamet, Tunisia. pp. 1–6. http://dx.doi.org/10.1109/GECS.2017.8066124.

Bouaicha, A., Trabelsi, A., Allagui, H., Mami, A., Aglzim, E.-H., Rouane, A., 2017c. Study and design of a controlled oscillator for an embedded system to measure the complex impedance of PEM fuel cell. In: 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), Monastir. pp. 173–178. http://dx.doi.org/10.1109/STA.2017.8314935.

Broughton, S.A., Bryan, K., 2018. Discrete Fourier Analysis and Wavelets: Applications to Signal and Image Processing, second ed. John Wiley, Hoboken, NJ.

Cadet, C., Jemeï, S., Druart, F., Hissel, D., 2014. Diagnostic tools for PEMFCs: from conception to implementation. Int. J. Hydrog. Energy 39 (20), 10613–10626. http://dx.doi.org/10.1016/j.ijhydene.2014.04.163.

Costamagna, P., De Giorgi, A., Moser, G., Pellaco, L., Trucco, A., 2019a. Data-driven fault diagnosis in SOFC-based power plants under off-design operating conditions. Int. J. Hydrog. Energy 44 (54), 29002–29006. http://dx.doi.org/10.1016/j.ijhydene.2019.09.128.

Costamagna, P., De Giorgi, A., Moser, G., Serpico, S.B., Trucco, A., 2019b. Data-driven techniques for fault diagnosis in power generation plants based on solid oxide fuel cells. Energy Convers. Manage. 180, 281–291. http://dx.doi.org/10.1016/j.enconman.2018.10.107.

Debenjak, A., Petrovcic, J., Boskoski, P., Musizza, B., Juricic, D., 2015. Fuel cell condition monitoring system based on interconnected DC–DC converter and voltage monitor. IEEE Trans. Ind. Electron. 62 (8), 5293–5305. http://dx.doi.org/10.1109/TIE.2015.2434792.

Depernet, D., Narjiss, A., Gustin, F., Hissel, D., Péra, M.-C., 2016. Integration of electrochemical impedance spectroscopy functionality in proton exchange membrane fuel cell power converter. Int. J. Hydrog. Energy 41 (11), 5378–5388. http://dx.doi.org/10.1016/j.ijhydene.2016.02.010.

Dijoux, É., 2019. Contrôle tolérant aux défauts appliqué aux systèmes pile à combustible à membrane échangeuse de protons (pemfc).

Dijoux, E., Benne, M., Yousfi Steiner, N., Grondin Perez, B., Pera, M.-C., 2017. Active fault tolerant control strategy applied to PEMFC systems. In: 2017 IEEE Vehicle Power and Propulsion Conference (VPPC), Belfort. pp. 1–5. http://dx.doi.org/10.1109/VPPC.2017.8330967.

Escobet, A., 2014. PEM fuel cell fault diagnosis via a hybrid methodology based on fuzzy and pattern recognition techniques. Eng. Appl. Artif. Intell. 14.

Fairweather, A.J., Foster, M.P., Stone, D.A., 2011. Battery parameter identification with pseudo random binary sequence excitation (PRBS). J. Power Sources 196 (22), 9398–9406. http://dx.doi.org/10.1016/j.jpowsour.2011.06.072.

Fan, L., Huang, X., Yi, L., 2013. Fault diagnosis for fuel cell based on Naive Bayesian classification. 11 (12), 7.

Firouzjaei, V.K., Rahgoshay, S.M., Khorshidian, M., 2020. Planar membrane humidifier for fuel cell application: Numerical and experimental case study. Int. J. Heat Mass Transfer 10.

Gallo, M., Costabile, C., Sorrentino, M., Polverino, P., Pianese, C., 2020. Development and application of a comprehensive model-based methodology for fault mitigation of fuel cell powered systems. Appl. Energy 279, 115698. http://dx.doi.org/10.1016/j.apenergy.2020.115698.

Gerard, M., Poirot-Crouvezier, J.-P., Hissel, D., Pera, M.-C., 2010. Oxygen starvation analysis during air feeding faults in PEMFC. Int. J. Hydrog. Energy 35 (22), 12295–12307. http://dx.doi.org/10.1016/j.ijhydene.2010.08.028.

Gu, X., Hou, Z., Cai, J., 2021. Data-based flooding fault diagnosis of proton exchange membrane fuel cell systems using LSTM networks. Energy AI 100056. http://dx.doi.org/10.1016/j.egyai.2021.100056.

Hwang, J.J., Chang, W.R., Kao, J.K., Wu, W., 2012. Experimental study on performance of a planar membrane humidifier for a proton exchange membrane fuel cell stack. J. Power Sources 215, 69–76. http://dx.doi.org/10.1016/j.jpowsour.2012.04.051.

Jouin, M., Gouriveau, R., Hissel, D., Péra, M.-C., Zerhouni, N., 2013a. Prognostics and health management of PEMFC – state of the art and remaining challenges. Int. J. Hydrog. Energy 38 (35), 15307–15317. http://dx.doi.org/10.1016/j.ijhydene.2013.09.051.

Jouin, M., Gouriveau, R., Hissel, D., Péra, M.-C., Zerhouni, N., 2013b. Prognostics of PEM fuel cell in a particle filtering framework. Int. J. Hydrog. Energy 14.

Jung, S.H., Kim, S.L., Kim, M.S., Park, Y., Lim, T.W., 2007. Experimental study of gas humidification with injectors for automotive PEM fuel cell systems. J. Power Sources 170 (2), 324–333. http://dx.doi.org/10.1016/j.jpowsour.2007.04.013.

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

Krishnan, S., Athavale, Y., 2018. Trends in biomedical signal feature extraction. Biomed. Signal Process. Control 43, 41–63. http://dx.doi.org/10.1016/j.bspc.2018.02.008.

Laribi, S., Mammar, K., Hamouda, M., Sahli, Y., 2016. Impedance model for diagnosis of water management in fuel cells using artificial neural networks methodology. Int. J. Hydrog. Energy 41 (38), 17093–17101. http://dx.doi.org/10.1016/j.ijhydene.2016.07.099.

Laribi, S., Mammar, K., Sahli, Y., Koussa, K., 2019. Analysis and diagnosis of PEM fuel cell failure modes (flooding & drying) across the physical parameters of electrochemical impedance model: Using neural networks method. Sustain. Energy Technol. Assess. 34, 35–42. http://dx.doi.org/10.1016/j.seta.2019.04.004.

Lebreton, C., et al., 2015. Fault tolerant control strategy applied to PEMFC water management. Int. J. Hydrog. Energy 40 (33), 10636–10646. http://dx.doi.org/10.1016/j.ijhydene.2015.06.115.

Lee, W.-Y., Oh, H., Kim, M., Choi, Y.-Y., Sohn, Y.-J., Kim, S.-G., 2019. Hierarchical fault diagnostic method for a polymer electrolyte fuel cell system. Int. J. Hydrog. Energy S0360319919339564. http://dx.doi.org/10.1016/j.ijhydene.2019.10.145.

Li, S., Cao, H., Yang, Y., 2018. Using multi-label pattern identification. J. Power Sources 14.

Li, Z., Outbib, R., Giurgea, S., Hissel, D., Giraud, A., Couderc, P., 2019. Fault diagnosis for fuel cell systems: A data-driven approach using high-precise voltage sensors. Renew. Energy 135, 1435–1444. http://dx.doi.org/10.1016/j.renene.2018.09.077.

Li, Z., Outbib, R., Giurgea, S., Hissel, D., Li, Y., 2015. Fault detection and isolation for polymer electrolyte membrane fuel cell systems by analyzing cell voltage generated space. Appl. Energy 148, 260–272. http://dx.doi.org/10.1016/j.apenergy.2015.03.076.

Li, Z., Outbib, R., Hissel, D., Giurgea, S., 2014. Data-driven diagnosis of PEM fuel cell: A comparative study. Control Eng. Pract. 28, 1–12. http://dx.doi.org/10.1016/j.conengprac.2014.02.019.

Li, Z., et al., 2016. Online implementation of SVM based fault diagnosis strategy for PEMFC systems. Appl. Energy 164, 284–293. http://dx.doi.org/10.1016/j.apenergy.2015.11.060.

Liang, D., Shen, Q., Hou, M., Shao, Z., Yi, B., 2009. Study of the cell reversal process of large area proton exchange membrane fuel cells under fuel starvation. J. Power Sources 194 (2), 847–853. http://dx.doi.org/10.1016/j.jpowsour.2009.06.059.

Lim, I.S., Park, J.Y., Choi, E.J., Kim, M.S., 2021. Efficient fault diagnosis method of PEMFC thermal management system for various current densities. Int. J. Hydrog. Energy 46 (2), 2543–2554. http://dx.doi.org/10.1016/j.ijhydene.2020.10.085.

Lin, R.-H., Pei, Z.-X., Ye, Z.-Z., Guo, C.-C., Wu, B.-D., 2019a. Hydrogen fuel cell diagnostics using random forest and enhanced feature selection. Int. J. Hydrog. Energy S0360319919339382. http://dx.doi.org/10.1016/j.ijhydene.2019.10.127.

Lin, R.-H., Xi, X.-N., Wang, P.-N., Wu, B.-D., Tian, S.-M., 2019b. Review on hydrogen fuel cell condition monitoring and prediction methods. Int. J. Hydrog. Energy 44 (11), 5488–5498. http://dx.doi.org/10.1016/j.ijhydene.2018.09.085.

Liu, Z., Yang, L., Mao, Z., Zhuge, W., Zhang, Y., Wang, L., 2006. Behavior of PEMFC in starvation. J. Power Sources 157 (1), 166–176. http://dx.doi.org/10.1016/j.jpowsour.2005.08.006.

Lu, H., Chen, J., Yan, C., Liu, H., 2019. On-line fault diagnosis for proton exchange membrane fuel cells based on a fast electrochemical impedance spectroscopy measurement. J. Power Sources 430, 233–243. http://dx.doi.org/10.1016/j.jpowsour.2019.05.028.

Ma, T., Lin, W., Yang, Y., Wang, K., Jia, W., 2020. Water content diagnosis for proton exchange membrane fuel cell based on wavelet transformation. Int. J. Hydrog. Energy 45 (39), 20339–20350. http://dx.doi.org/10.1016/j.ijhydene.2019.11.068.

Mao, L., Jackson, L., Davies, B., 2018. Investigation of PEMFC fault diagnosis with consideration of sensor reliability. Int. J. Hydrog. Energy 43 (35), 16941–16948. http://dx.doi.org/10.1016/j.ijhydene.2017.11.144.

Maruo, T., et al., 2017. Development of fuel cell system control for sub-zero ambient conditions. 2017-01-1189. http://dx.doi.org/10.4271/2017-01-1189.

Mawonou, K.S.R., Eddahech, A., Dumur, D., Beauvois, D., Godoy, E., 2021. State-of-health estimators coupled to a random forest approach for lithium-ion battery aging factor ranking. J. Power Sources 484, 229154. http://dx.doi.org/10.1016/j.jpowsour.2020.229154.

Mohammadi, A., Djerdir, A., Yousfi Steiner, N., Khaburi, D., 2015. Advanced diagnosis based on temperature and current density distributions in a single PEMFC. Int. J. Hydrog. Energy 40 (45), 15845–15855. http://dx.doi.org/10.1016/j.ijhydene.2015.04.157.

Morando, S., Jemei, S., Hissel, D., Gouriveau, R., Zerhouni, N., 2016. Proton exchange membrane fuel cell ageing forecasting algorithm based on echo state network. Int. J. Hydrog. Energy 9.

Narjiss, A., Depernet, D., Candusso, D., Gustin, F., Hissel, D., 2008. On-line diagnosis of a PEM fuel cell through the PWM converter. p. 9.

Onanena, R., Oukhellou, L., Côme, E., Candusso, D., Hissel, D., Aknin, P., 2012. Fault-diagnosis of PEM fuel cells using electrochemical spectroscopy impedance. IFAC Proc. 45 (21), 651–656. http://dx.doi.org/10.3182/20120902-4-FR-2032.00114.

Owejan, J.P., Gagliardo, J.J., Sergi, J.M., Kandlikar, S.G., Trabold, T.A., 2009. Water management studies in PEM fuel cells, part I: Fuel cell design and in situ water distributions. Int. J. Hydrog. Energy 34 (8), 3436–3444. http://dx.doi.org/10.1016/j.ijhydene.2008.12.100.

Pahon, E., Yousfi Steiner, N., Jemei, S., Hissel, D., Moçoteguy, P., 2016. A signal-based method for fast PEMFC diagnosis. Appl. Energy 165, 748–758. http://dx.doi.org/10.1016/j.apenergy.2015.12.084.

Park, J.Y., Lim, I.S., Choi, E.J., Kim, M.S., 2021. Fault diagnosis of thermal management system in a polymer electrolyte membrane fuel cell. Energy 214, 119062. http://dx.doi.org/10.1016/j.energy.2020.119062.

Petrone, G., Zamboni, W., Spagnuolo, G., 2019. An interval arithmetic-based method for parametric identification of a fuel cell equivalent circuit model. Appl. Energy 242, 1226–1236. http://dx.doi.org/10.1016/j.apenergy.2019.03.136.

Petrone, R., et al., 2013. A review on model-based diagnosis methodologies for PEMFCs. Int. J. Hydrog. Energy 38 (17), 7077–7091. http://dx.doi.org/10.1016/j.ijhydene.2013.03.106.

Pohjoranta, A., 2015. Validation of neural network-based fault diagnosis for multi-stack fuel cell systems: Stack voltage deviation detection. Energy Procedia 9.

Polverino, P., Frisk, E., Jung, D., Krysander, M., Pianese, C., 2017. Model-based diagnosis through structural analysis and causal computation for automotive polymer electrolyte membrane fuel cell systems. J. Power Sources 357, 26–40. http://dx.doi.org/10.1016/j.jpowsour.2017.04.089.

Polverino, P., Pianese, C., Sorrentino, M., Marra, D., 2015. Model-based development of a fault signature matrix to improve solid oxide fuel cell systems on-site diagnosis. J. Power Sources 280, 320–338. http://dx.doi.org/10.1016/j.jpowsour.2015.01.037.

Ren, P., Pei, P., Li, Y., Wu, Z., Chen, D., Huang, S., 2020. Degradation mechanisms of proton exchange membrane fuel cell under typical automotive operating conditions. Prog. Energy Combust. Sci. 80, 100859. http://dx.doi.org/10.1016/j.pecs.2020.100859.

Shao, M., Zhu, X.-J., Cao, H.-F., Shen, H.-F., 2014. An artificial neural network ensemble method for fault diagnosis of proton exchange membrane fuel cell system. Energy 67, 268–275. http://dx.doi.org/10.1016/j.energy.2014.01.079.

Steiner, N.Y., 2010. Model-based diagnosis for proton exchange membrane fuel cells. Math. Comput. Simulation 13.

Tang, Z., 2020. Recent progress in the use of electrochemical impedance spectroscopy for the measurement, monitoring, diagnosis and optimization of proton exchange membrane fuel cell performance. J. Power Sources 26.

Vural, Y., Ingham, D.B., Pourkashanian, M., 2009. Performance prediction of a proton exchange membrane fuel cell using the ANFIS model. Int. J. Hydrog. Energy 34 (22), 9181–9187. http://dx.doi.org/10.1016/j.ijhydene.2009.08.096.

Wang, J., et al., 2021. Recent advances and summarization of fault diagnosis techniques for proton exchange membrane fuel cell systems: A critical overview. J. Power Sources 500, 229932. http://dx.doi.org/10.1016/j.jpowsour.2021.229932.

Wu, X., Ye, Q., 2016. Fault diagnosis and prognostic of solid oxide fuel cells. J. Power Sources 321, 47–56. http://dx.doi.org/10.1016/j.jpowsour.2016.04.080.

Wu, X., Zhou, B., 2016. Fault tolerance control for proton exchange membrane fuel cell systems. J. Power Sources 324, 804–829. http://dx.doi.org/10.1016/j.jpowsour.2016.05.066.

Yang, B., et al., 2021. Solid oxide fuel cell systems fault diagnosis: Critical summarization, classification, and perspectives. J. Energy Storage 34, 102153. http://dx.doi.org/10.1016/j.est.2020.102153.

Yousfi Steiner, N., Hissel, D., Moçotéguy, Ph., Candusso, D., 2011. Diagnosis of polymer electrolyte fuel cells failure modes (flooding & drying out) by neural networks modeling. Int. J. Hydrog. Energy 36 (4), 3067–3075. http://dx.doi.org/10.1016/j.ijhydene.2010.10.077.

Yousfi-Steiner, N., Moçotéguy, Ph., Candusso, D., Hissel, D., 2009. A review on polymer electrolyte membrane fuel cell catalyst degradation and starvation issues: Causes, consequences and diagnostic for mitigation. J. Power Sources 194 (1), 130–145. http://dx.doi.org/10.1016/j.jpowsour.2009.03.060.

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

Yuan, H., 2020. Model-based observers for internal states estimation and control of proton exchange membrane fuel cell system: A review. J. Power Sources 17.

Yue, M., Al Masry, Z., Jemei, S., Zerhouni, N., 2021. An online prognostics-based health management strategy for fuel cell hybrid electric vehicles. Int. J. Hydrog. Energy 46 (24), 13206–13218. http://dx.doi.org/10.1016/j.ijhydene.2021.01.095.

Zamel, N., Li, X., 2011. Effect of contaminants on polymer electrolyte membrane fuel cells. Prog. Energy Combust. Sci. 37 (3), 292–329. http://dx.doi.org/10.1016/j.pecs.2010.06.003.

Zhang, X., Zhang, T., Chen, H., Cao, Y., 2021. A review of online electrochemical diagnostic methods of on-board proton exchange membrane fuel cells. Appl. Energy 286, 116481. http://dx.doi.org/10.1016/j.apenergy.2021.116481.

Zhao, J., Li, X., Shum, C., McPhee, J., 2021. A review of physics-based and data-driven models for real-time control of polymer electrolyte membrane fuel cells. Energy AI 6, 100114. http://dx.doi.org/10.1016/j.egyai.2021.100114.

Zheng, Z., et al., 2013. A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems. Int. J. Hydrog. Energy 38 (21), 8914–8926. http://dx.doi.org/10.1016/j.ijhydene.2013.04.007.

Zheng, Z., et al., 2017. Brain-inspired computational paradigm dedicated to fault diagnosis of PEM fuel cell stack. Int. J. Hydrog. Energy 42 (8), 5410–5425. http://dx.doi.org/10.1016/j.ijhydene.2016.11.043.

Zheng, Y., et al., 2021. Data-driven fault diagnosis method for the safe and stable operation of solid oxide fuel cells system. J. Power Sources 490, 229561. http://dx.doi.org/10.1016/j.jpowsour.2021.229561.

Zhong, Z.-D., Zhu, X.-J., Cao, G.-Y., 2006. Modeling a PEMFC by a support vector machine. J. Power Sources 160 (1), 293–298. http://dx.doi.org/10.1016/j.jpowsour.2006.01.040.

Zhou, S., Dhupia, J.S., 2020. Online adaptive water management fault diagnosis of PEMFC based on orthogonal linear discriminant analysis and relevance vector machine. Int. J. Hydrog. Energy 45 (11), 7005–7014. http://dx.doi.org/10.1016/j.ijhydene.2019.12.193.