Failure Prognosis Based on Relevant Measurements Identification and Data-Driven Trend-Modeling: Application to a Fuel Cell System

Mohand Djeziri, Oussama Djedidi, Samir Benmoussa, Marc Bendahan, Jean-Luc Seguin

processes

MDPI

Article

Failure Prognosis Based on Relevant Measurements Identification and Data-Driven Trend-Modeling: Application to a Fuel Cell System

Mohand Djeziri [1,*], Oussama Djedidi [1], Samir Benmoussa [2], Marc Bendahan [3] and Jean-Luc Seguin [3]

[1] CNRS (Centre National de la Recherche Scientifique), LIS (Laboratoire d'Informatique et Systèmes), Aix-Marseille University, Université de Toulon, 13007 Marseille, France; oussama.djedidi@lis-lab.fr
[2] Laboratoire d'Automatique et de Signaux de Annaba (LASA), Université Badji Mokhtar, Annaba 23000, Algeria; samir.benmoussa@univ-annaba.dz
[3] CNRS, IM2NP (Institut Matériaux Microélectronique et Nanosciences de Provence), Aix Marseille Univ, Université de Toulon, 13007 Marseille, France; marc.bendahan@im2np.fr (M.B.); j-l.seguin@univ-amu.fr (J.-L.S.)
* Correspondence: mohand.djeziri@lis-lab.fr

Abstract: Fuel cells are key elements in the transition to clean energy thanks to their neutral carbon footprint, as well as their great capacity for the generation of electrical energy by oxidizing hydrogen. However, these cells operate under straining conditions of temperature and humidity that favor degradation processes. Furthermore, the presence of hydrogen—a highly flammable gas—renders the assessment of their degradations and failures crucial to the safety of their use. This paper deals with the combination of physical knowledge and data analysis for the identification of health indices (HIs) that carry information on the degradation process of fuel cells. Then, a failure prognosis method is achieved through the trend modeling of the identified HI using a data-driven and updatable state model. Finally, the remaining useful life is predicted through the calculation of the times of crossing of the predicted HI and the failure threshold. The trend model is updated when the estimation error between the predicted and measured values of the HI surpasses a predefined threshold to guarantee the adaptation of the prediction to changes in the operating conditions of the system. The effectiveness of the proposed approach is demonstrated by evaluating the obtained experimental results with prognosis performance analysis techniques.

Keywords: remaining useful life; health index identification; discrete state model; trend modeling; fuel-cell systems

Citation: Djeziri, M.; Djedidi, O.; Benmoussa, S.; Bendahan, M.; Seguin, J.-L. Failure Prognosis Based on Relevant Measurements Identification and Data-Driven Trend-Modeling: Application to a Fuel Cell System. *Processes* **2021**, *9*, 328. https://doi.org/10.3390/pr9020328

Academic Editor: Mohd Azlan Hussain
Received: 26 January 2021
Accepted: 9 February 2021
Published: 11 February 2021

1. Introduction

In the specialized literature, we can distinguish two main families of approaches for fault diagnosis and failure prognosis: approaches using physical models and data-driven ones. These large families can then be decomposed according to the tools used in each work, as proposed by Lin et al. [1] and Djeziri et al. [2]. The use of physical or data-driven models is a function of the compromise that has to be found between the advantages and the drawbacks of each of them.

Model-based methods are well suited when the physical knowledge of the systems is rich enough to make modeling assumptions and to build physical models that best represent the real dynamics of a system. Renewable energy systems are among the systems of which physical knowledge is rich; it is even at the origin of the development of these systems. When the physical model is well constructed, it takes into account all the possible operating points of a system, which makes methods based on physical models more generic. However, the step of identifying the numerical values of the model parameters is done using data-driven methods, which can lead to the identification of inadequate numerical values with the physical meanings of the parameters.

Data-driven methods are easier to implement compared to model-based ones, and are perfectible thanks to the ability to update the models online by learning whenever new data become available. On the other hand, the performance of these methods depends directly on the quality and quantity of the available data. Often, data regarding degraded system operations are not available, and the cause-and-effect relationship between degradation and the variables measured is not formally demonstrated. So-called attribute selection techniques, such as ReliefF, allow automatic selection of relevant variables from a set of measured variables [3–5]. Two main approaches can be found in the literature: filter-based methods and wrapping-based ones. Filter-based methods evaluate datasets using their information content, such as inter-class distance and statistical dependence. Wrapping-based methods use a classifier to evaluate subsets according to their predictive accuracy using statistical sampling or cross-validation. Their limitation lies in the fact that they are random search algorithms for an optimal set of attributes, so it is possible to obtain a different set of attributes for each launch of the selection algorithm. These methods can be an alternative when the number of measured variables is large and when the physical relationship between the measured variables is not sufficiently known to make the selection through physical reasoning.

The hybrid method proposed in this paper aims to capitalize on all the available information (physical knowledge and data history) to build an algorithm for predicting the failure of dynamic systems. The physical model is used to formally identify the measured variables that carry useful information about the degradation process, and it can also be used for optimal sensor placement in fault diagnosis and failure prognosis contexts. The failure prediction is then carried out by a data-driven model and updated online. Identifying relevant features using a physical approach significantly reduces the uncertainty that can be generated by a large number of input variables used for learning in the trend model. Model updating is only calculated when needed according to the estimation error of the health index, and the update is done with the recursive least squares algorithm. The proposed method is applicable on a wide range of systems, and is easy to implement online. The physical model, which was built with the bond graph (BG) methodology [6] and experimentally validated in several works [7–11], is firstly used for structural analysis in order to identify the relevant measures that carry the information about the degradation process. The degradation due to wear is manifested by the gradual deviations of the characteristics of one or more system components from their nominal values. In the BG methodology, these components correspond to well-identified BG elements, and thanks to the notion of analogy, it is possible to represent all the hardware components and the physical phenomena in a system with BG elements.

In addition, the causal and structural properties of the BG make it possible to browse a BG as a directed graph, and thus to visualize the causal paths between the measurements and the BG parameters whose physical meanings are clearly known. Once the relevant measurements are identified, failure prognosis is achieved using a data-driven method. Thus, the use of the physical model will be limited to structural and causal analyses that require only a symbolic model, which is more generic and corresponds to a wide range of fuel cells. The complex steps of parameter identification and model implementation are not necessary in the approach proposed in this paper. After the identification of the measures carrying the degradation process, these raw signals are preprocessed to extract the useful information; then, the usability of the processed signals as health indicators for failure prognosis is analyzed. The processing of the raw signals is performed using an automatic smoothing process based on the Savitzky–Golay filter [12], where the sliding window value that conserves only the appearance of the signal by removing noise and outliers is calculated automatically. The usability analysis is carried out according to the scores of the monotonicity metric [13]. The evolution trend of the health indices (HIs) is modeled by a data-driven state model and is updated periodically to adjust the Remaining Useful Lifetime (RUL) prediction to changes in condition monitoring. The RUL is calculated through the temporal projection of the HI estimates up to the total failure threshold.

The rest of the paper is organized as follows: Related works are given in Section 2. Section 3 presents the proposed approach. Structural analysis for HIs identification, measurement smoothing, and the analysis of the usability of the smoothed HI for failure prognosis are developed in Section 4. The trend modeling method and RUL prediction are described in Section 5. The application to the fuel cell system and the obtained results are shown and analyzed in Section 6, and a conclusion is given in the last section.

2. Related Works

Fuel-cell systems are increasingly used in the fields of transport and energy production. This evolution motivated the scientific community to develop physical models that are widely used for simulation [14], control [7], and condition monitoring [9].

The fuel cell is a multi-physics system with interactions of hydro-pneumatic, electrical, chemical, and thermal phenomena. The BG formalism, as a unified tool for all domains of physics, is naturally used for the modeling of this system [7,8,11,15,16].

In the field of fault diagnosis, Yang et al. [17] used BG for the identification of detectable faults in a fuel cell; then, the principle of analytical redundancy was used for the generation of HIs. In the context of failure prognosis, the bond graph methodology is often associated with data-driven trend models, as in [18], where the authors used a bond graph model in linear fractional transformation (LFT) form, initially proposed in [19], for the generation of HIs, as well as a particle filter for failure prognosis.

Although the proposed BG models differ in granularity, they are causally and structurally equivalent. The difference in granularity essentially comes from modeling assumptions and from the purpose for which the model is built. For example, Saisset et al. [7] created a dynamic model of a proton exchange membrane fuel cell (PEM-FC) for analyzing the interactions between a fuel cell and DC/AC converters without considering the hydraulic losses in the gas channels and the hydration changes of the membrane, and the model validation was limited to the electrical performance. The BG model proposed in [16,20] was built for control purposes on the basis of the model of Saisset et al. [7]; by including a more refined description of the cooling process, neglecting heat transfer between individual layers of the cell to simplify the thermal dynamics, disregarding the pressure losses within channels and constant membrane hydration, and using a controller for the solenoid valve to control hydrogen flow rate and the air compressor as transfer functions, the coefficients were tuned by using an iterative approach. This improvement and adaptation of BG models to the context of use is possible thanks to their modular and hierarchical appearance. The system is seen as a set of interconnected causal subsystems, where the model of each subsystem can be modified or enhanced independently of other subsystems, as long as its causal link with other subsystems is correctly defined. This BG property has been effectively highlighted in the model of Vasilyev et al. [11], where the BGs model of the PEM-FC was built in an object-oriented modeling environment. The model proposed in Vasyliev et al. [11] was structurally equivalent to the models presented in [16,20]; however, it included the main physical phenomena occurring in the fuel-cell system by considering the thermal and water management, the chemical composition of the inlet gases, and the electrolyte ionic resistance. This object-oriented presentation furnishes an opportunity to study the system as a whole while keeping the possibility of analyzing individual cells located in different parts of the stack.

In the context of fault diagnosis and failure prognosis, Lin et al. [1] proposed a classification of existing methods into three categories: model-based methods [21], filter-based methods [22], and data-driven methods [23–25]. Wu et al. [26] proposed a method that simultaneously performs fault diagnosis and failure prognosis for the solid oxide fuel cell (SOFC), where a least squares support vector machine (LS-SVM) classifier was used to identify the fault type, and hidden semi-Mark models (HSMMs) were employed to estimate the RUL, assuming that the fuel cell passes through several intermediate degradation states before reaching total failure. The use of a new neural network paradigm, based on the Echo State Network, which is part of a group of reservoir computing methods, is explored in [27]

for the aging forecasting of PEM-FC. The signals used for failure prognosis are filtered by the combination of the wavelet filter and Hurst coefficient. An algorithm of optimal sensor selection for predicting PEM-FC performance was proposed in [28]. The method is based on the largest gap and exhaustive brute-force searching techniques; then, the optimal sensor set is studied to predict fuel-cell performance using test data from a PEM-FC system. The Elman neural network state prediction model combined with multivariable dynamic response analysis of an SOFC system was used in [29] to predict the future voltage of a solid oxide fuel cell under various operating modes, including system start–stop, long-term operation, and hot standby. Unlike the majority of work employing stack/cell voltage as a direct link for RUL predictions, the method proposed in [30] is based on the stack's ohmic area-specific resistance (ASR), and consists of three steps: Firstly, the current ASR value is estimated using an unscented Kalman filter based on a lumped stack model. Then, the temporal drift in the ASR is recursively estimated using a linear Kalman filter. Finally, based on the identified drift model, a Monte Carlo simulation is performed to predict the RUL. D. Hissel and M. C. Perra [31] dealt with the state of the art and the motivations regarding diagnostic and state-of-health estimation methodologies applied to fuel cells, and presented a selection of recent developments and experimentations. In their conclusion, the authors of [31] presented the RUL prediction of fuel cells as a research theme that is still open, which will contribute to the development of the design and large-scale operation of fuel cells.

3. RUL Prediction Schema

The presented approach is given in three steps, as illustrated in Figure 1. The first two steps are offline analysis steps. The first one concerns the use of a generic BG model of the system for causal and structural analysis in order to identify the causal path between measures and all BG elements whose variation represents a degradation. The presence of the causal path formally demonstrates that the measurement is sensitive to the fault considered. The second step of the offline analysis is to demonstrate that the variables selected in the first step satisfy the monotonicity condition, which is necessary for the usability of the selected measurement as an HI for failure prognosis. Then, the initial values of the state model are identified. The last step is the online monitoring, where the state model is implemented and updated online, and then an RUL is predicted by a time projection of the HI estimation until a predefined failure threshold.

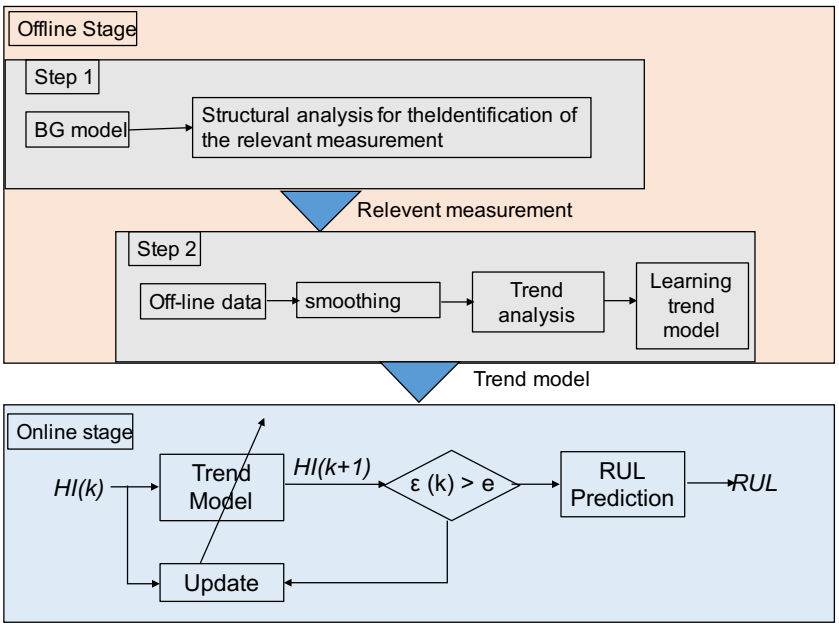

Figure 1. Overview of the RUL estimation process.

4. Health Indices Identification

4.1. Structural Analysis

The notion of causality is one of the major advantages of the BG methodology. It allows one to highlight, in the block diagram sense, the relationships of: cause–effect, input–output, and known–unknown variables. The causality rule in a BG model is defined, by convention, as illustrated in Figure 2: The causal stroke is placed near (or far from) the element for which the effort (or flow) is an input (known).

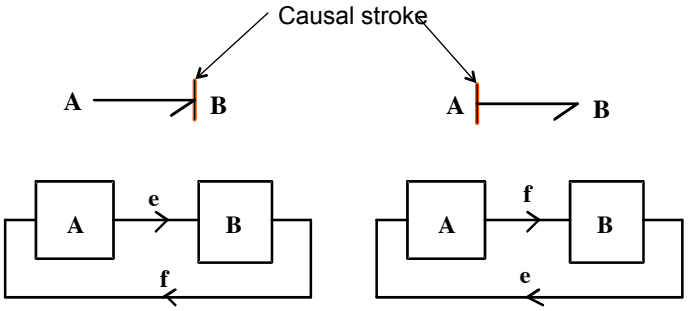

Figure 2. Causality assignment convention.

A causal path in a junction structure (0, 1, T F, or GY) is defined as an alternation of bonds and elements (R, C, or I) called nodes, such that all nodes have a complete and correct causality, and two bonds of the causal path have opposite causal orientations in the same node. Depending on the causality, the variable crossed is effort or flow. To change this variable, it is necessary to pass through a junction element, GY, or through a passive element (I, C, or R).

Thanks to the causal and structural properties, an analysis of the observability of the considered faults can be done directly on the BG model of the system by following the causal paths relaying the faults to the sensors. To illustrate these properties, let us consider the example of the mechanical system given in Figure 3a and its BG model, which is given in Figure 3b. The causal path relaying the input force $Se : F$ and the measured output $De : Fc$ is illustrated in the model in Figure 3c and is given as follows:

$$Se : F \rightarrow e1 \rightarrow e2 \rightarrow I : M \rightarrow f2 \rightarrow f4 \rightarrow f5 \rightarrow C : 1/K \rightarrow e5 \rightarrow e6 \rightarrow De : Fc \quad (1)$$

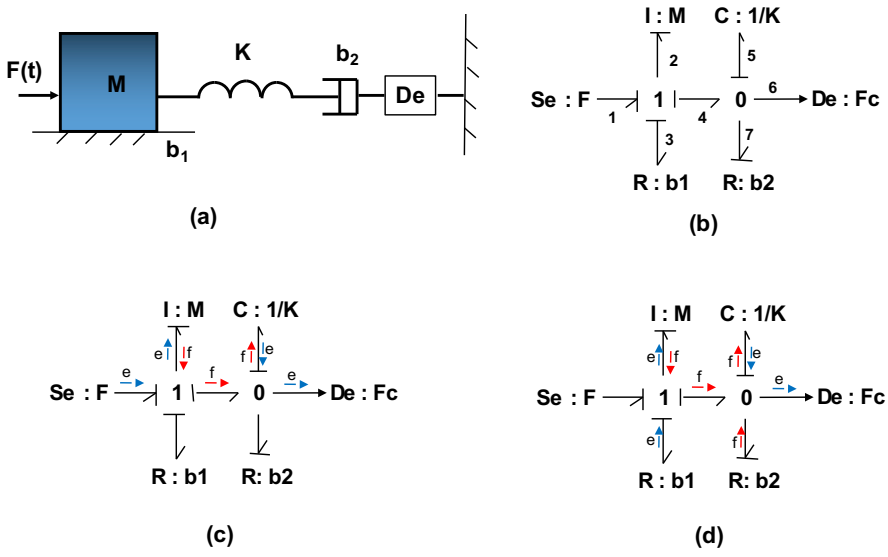

Figure 3. Illustration of an example of the causal paths following from an input to an output: (**a**) example of the mechanical system, (**b**) its corresponding BG model, (**c**,**d**) causal paths.

The same reasoning is considered with an external fault or a component fault; the causal path makes it possible to visualize the fault path in the system up to the measured output. A component fault results in a deviation of the corresponding BG element from the nominal value identified during normal operation. In the example of the mechanical system in Figure 3, a degradation of the damper will result in a deviation of the BG element $R : b2$ from its nominal value. Since there is an indirect causal path between this R element and the measure $De : Fc$ (illustrated in Figure 3d and given in Equation (2), this measure will be sensitive to this degradation, and will consequently carry the information on the degradation process of the damper.

$$R : b2 \rightarrow f7 \rightarrow f5 \rightarrow e5 \rightarrow e6 \rightarrow De : Fc \tag{2}$$

4.2. Preprocessing

The sensor measurements are extremely noisy raw signals, which do not satisfy the properties required for the use of these raw measurements for failure prognosis. The most common method used to avoid the noise is smoothing. The smoothing methods are multiple, and every method has its advantages and limitations. In this work, an automatic smoothing process based on the Savitzky–Golay (SG) filter is considered. Unlike conventional filters, which require noise characterization, the adopted filter makes it possible to obtain the smoothed signal and its derivatives through a simple and fast calculation.

The Savityzky–Golay filter formalism is given by Equation (3) for equally spaced input data x_i, y_j, with $j = 1, \ldots, n$, and n is the data amount.

$$g_i = \sum_{i=-m}^{m} c_i y_{k+i}, \tag{3}$$

where $i = -m, \ldots, \lambda, \ldots, m$ and λ is the center point index. A table of Savitzky–Golay coefficients is obtained using $J = 2m + 1$ as a window length and the polynomial degree k [32,33]. The performance of Savitzky–Golay filters depends on the choice of the polynomial order and the window length. In this work, the sliding window J is computed automatically by keeping only the appearance of the signal and cancelling noise and the outlier's components. The polynomial order is fixed to 2 to avoid any intense wiggling in the smoothed signal. The proposed calculation of the J value is based on the similarity between the kernel estimation and Savitzky–Golay filter.

$$\tilde{f}(y) = \frac{1}{n} \sum_{i}^{n} w(y - y_i; h), \tag{4}$$

where $\tilde{f}$ is the kernel estimator, n is the size of the dataset y, and w is the kernel function whose variance is controlled by the parameter h. h is called the smoothing function or the bandwidth. In order to smooth $\tilde{f}$, the optimum formula to calculate h is given by:

$$h = \left(\frac{4}{3n} \right)^{\frac{1}{5}} \sigma, \tag{5}$$

where σ is the standard deviation of the signal. In order to avoid over-smoothing when dealing with non-normal data, σ is adjusted as follows.

$$\tilde{\sigma} = \frac{median\{abs(y_i - \tilde{\mu})\}}{0.6745}, \tag{6}$$

where $\tilde{\mu}$ denotes the median of the sample.

4.3. Usability Analysis of the Smoothed Signal for Failure Prognosis

The degradation process is a progressive and irreversible phenomenon, so the HIs that carry information on the degradation process must have a high level of monotonicity for better prognostic performance.

The monotonicity analysis of HIs generated by data-driven methods has been the subject of several research works [13,34]. The monotonicity metric presented below is the one proposed in [13,34], as its score is easily interpretable (between 0 and 1), where 1 indicates the most satisfactory and 0 the least satisfactory level of the specific HI property:

$$Monoton = \frac{1}{N}\left|\sum_{i}^{N} M_i\right|, \tag{7}$$

where M_i is the monotonicity of a single run-to-failure trajectory, given by:

$$M_i = \frac{n_i^+}{n_i - 1} - \frac{n_i^-}{n_i - 1}, i = 1, \ldots, N, \tag{8}$$

where n_i is the total number of observations in the i^{th} run-to-failure trajectory and n_i^+ (n_i^-) is the number of observations characterized by a positive (negative) first derivative.

5. Trend Modeling and RUL Prediction

The state model was introduced to build a matrix representation from physical models in the form of differential equations, as this representation is most suitable for performance analysis and structural analysis. In this case, the parameters of the model have a physical meaning. The data-driven state model proposed in this paper is a state model whose parameters do not have a clear physical meaning, but are identified from the data [35]. The order of the model is determined according to the profile of the health index available for learning, or by calculating the Hankel singular values for models of different orders. States with the smallest Hankel singular values are the most appropriate [36].

The model is then implemented and updated online for HI estimation and RUL prediction. The prediction of the RUL is performed at each update time by the temporal projection of the model estimate until the output (HI estimated) exceeds the total failure threshold.

The formal description of the data-driven state model is given as follows:

$$\begin{cases} x(n+1) = Ax(n) + Ke(n) \\ HI(n) = Cx(n) + e(n) \end{cases}, \tag{9}$$

where $x(n)$ is the state vector and contains the slot size, $HI(n)$ is the health index, $e(n)$ is the noise, and A, C, K are to be identified.

The Kalman filter and its variants (particle filter, strong-tracking filter, etc.) are widely used for online updating of the parameters of a prediction model [37]. In this case, the model parameters are updated with each sample. Then, a second instance of the model is necessary to generate the temporal projection of the output until it reaches the threshold, which is expensive in computation time and causes a permanent update of the model, making the decision-making difficult. So, in this work, we opted for a different approach where the update is only calculated when needed, according the HI estimation error, and the update is done with the recursive least squares algorithm [38]. This approach frees up resources for the calculation of the time projection of the health index. In addition, the projection is only carried out when the model is updated. Otherwise, the RUL is equal to the value predicted during the last update, which gives more visibility to maintenance operators for decision-making. This approach also allows the choice of samples used for updating for better accuracy.

The RUL can be defined as a function of the operating state of the system, which is given by the health index ($HI(t)$) as follows:

$$RUL(t|HI_t) = T - t|T > t,\ HI(t), \tag{10}$$

where T represents a random variable of the time to failure, t is the current time, and $HI(t)$ is the past condition profile up to the current time. The pseudocode used for RUL prediction and model updating is given in Algorithm 1.

Algorithm 1 Pseudocode for RUL prediction and model updating

Begin
Define endOfLifeThreshold ▷ The limit defining the end of life
Define timeBeforeEstimation ▷ Time spent gathering data
Define timeBetweenUpdates ▷ Time to wait before updating the model
Define samplingTime

while $t <$ timeBeforeEstimation **do** ▷ Gather initial set of data.
    HI.append(read($Utot$)) ▷ Append the data set with the current $Utot$ Value
$i = 0$
while HI(t) > endOfLifeThreshold **do** ▷ Gather initial set of data.
    if $i = 0$ **then**
        predictionModel = identifyStateSpaceModel(HI, orders)
        newModel = predictionModel ▷ Initialization
    else if $i > 0$ **then**
        while $t <$ timeBeforeEstimation **do** ▷ Gather current data.
            HI.append(read($Utot$))
        newModel = identifyStateSpaceModel(HI, orders) ▷ Estimate new model
            ▷ Get the performance of the new model (nM) and the prediction model (pM)
    $SSE_{nM} = \sum_{k=1}^{n}(HI(k) - y_{nM}(k))^2$
    $SSE_{pM} = \sum_{k=1}^{n}(HI(k) - y_{pM}(k))^2$
    if $SSE_{nM} <= SSE_{pM}$ **then**
        predictionModel = newModel
    $\hat{y}$ = predict(predictionModel, HI, numberOfSteps)
    timeOfRUL[i] = t
    $T = \underset{t}{\text{argmin}} f(t) = \{t \mid f(t) = \min_t |HI - \hat{y}|\}$
    RUL$[i] = T - t|T > t,\ HI(t)$
End

6. Application to a Fuel Cell

The considered system is a PEM-FC with a power of up to 1 kW (electrical power), proposed by FCLab Research Federation in the data challenge of the Prognostics and health management (PHM) conference [39]. The system is composed of five-cell stacks with an active area of 100 cm^2 for each cell. The PEM-FC is realized here with commercial membranes, diffusion layers, and machined flow distribution plates. The nominal current

density of the cells is 0.70 A/cm^2. Their maximal current density is 1 A/cm^2. The available measurements are given as follows:

- Stack temperature, gas flows, and air and hydrogen hygrometry rates;
- Inlet and outlet flows (of hydrogen, air, and cooling water), inlet and outlet pressures (of hydrogen and air), and temperatures (of incoming and outlet hydrogen, air, and cooling water).

6.1. Structural Analysis

BG models of fuel cell have been established in several works [7,9,11] and validated experimentally. In this paper, we use the model of the one-cell fuel cell (FC) presented in Figure 4 for structural analysis.

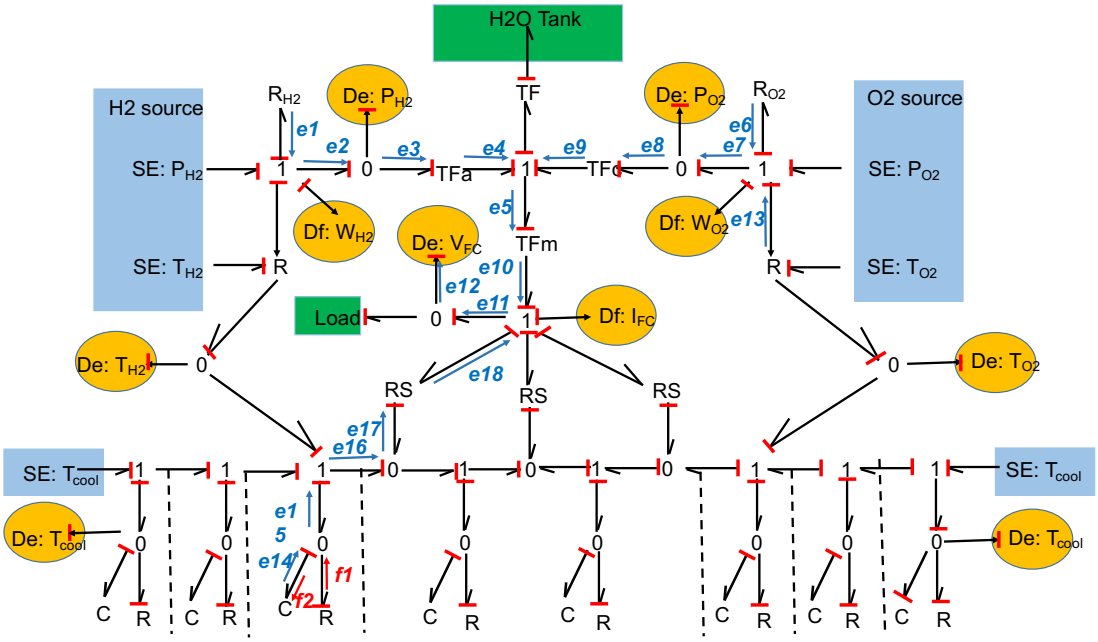

Figure 4. Causal BG model of the considered FC.

The power variables used are: torque and angular velocity (τ, ω), pressure and mass flow (P, m), temperature and flow of enthalpy (T, H), chemical potential and molar flux (μ, n), and voltage and current (U, i). These variables are respectively associated with the mechanical, hydraulic, thermal, chemical, and electrical domains. As regards the chemical reaction of the cell, two types of power variables are used: the chemical potential and the molar flux for the transformation phenomena, as well as the chemical affinity A (J/mol) and the speed of the reaction J.

The transformers TFa and TFc modulated by the molar flow represent chemical transformations that are, respectively, relative to the flow of dihydrogen through the anode and the flow of oxygen through the cathode. The transformer TFe represents the transformation of Gibbs free energy (variation of enthalpy) into electrical energy assuming that Gibbs free energy can be converted into electrical energy. The theoretical potential E of the fuel cell is related to Gibbs free energy ΔG, called free enthalpy, as follows:

$$E = -\frac{\Delta G}{nf}, \tag{11}$$

where n is the number of electrons exchanged during the reaction and F is the Faraday constant (96,485 C/mol). In the proposed BG model, there are 10 sensors that make it possible, respectively, to measure the mass flow, the hydrogen and oxygen temperatures, the hydrogen and oxygen pressures, the current at the terminals of the stack ,and the

voltage. The elements $RH2$ and $RO2$ represent the valves that control the hydrogen and oxygen input flow rates. These R elements are, in general, modeled by an information bond representing the valve control signal. The R element represents friction through the diffusion layers. The thermal losses are modeled by the RS element, which corresponds to an active resistance and represents the losses in the activation layer, the losses in the diffusion layers, and the losses due to the concentration.

Potential degradations of the fuel cell are caused by:

- Fouling: Can be modeled by a deflection of the hydraulic elements $RH2$ and $RO2$;
- Degradation of the cooling system: Can be modeled by the deviation of thermal elements R and the C elements;
- Membrane wear: Can be modeled by a deviation of the TF modules, anode side TFa, cathode side TFc, or in the membrane TFm.

After analysis of the causal paths between the BG elements and the measured outputs, the signature matrix of the Table 1 is built. It illustrates the causal links between each BG element and each measurement, and shows that the measurement causally related to all the BG elements is that of the voltage VFC; the corresponding causal paths are illustrated in Figure 4. It is also noted that these causal paths are direct causal paths, which means a direct transfer of the degradation phenomenon towards the measurement.

Table 1. Fault signature matrix for the fuel cell.

Physical Phenomenon	BG Elt	PH2	PO2	WH2	WO2	VFC	IFC	TH2	TO2	TCool
Fouling on the O2 side	RO2		1			1				
Fouling on the H2 side	RH2	1				1				
Degradation of the cooling system	R					1				1
Degradation of the cooling system	C					1				1
Stack degradation on the anode side	TFa			1		1				
Stack degradation on the cathode side	TFc				1	1				
Degradation of the membrane	TFm			1	1	1				

To verify that the information carried in the measurement of the voltage is presented in a usable way for trend modeling and failure prognosis, the raw signal is processed using an SG filter, and the scores of the monotonicity metric are carried and analyzed.

6.2. Preprocessing

The profile of the available PEM-FC voltage in the presence of degradation is given in Figure 5. This figure shows that although the profile of the voltage has a tendency, the signal is highly noisy and presents abrupt variations and some outliers.

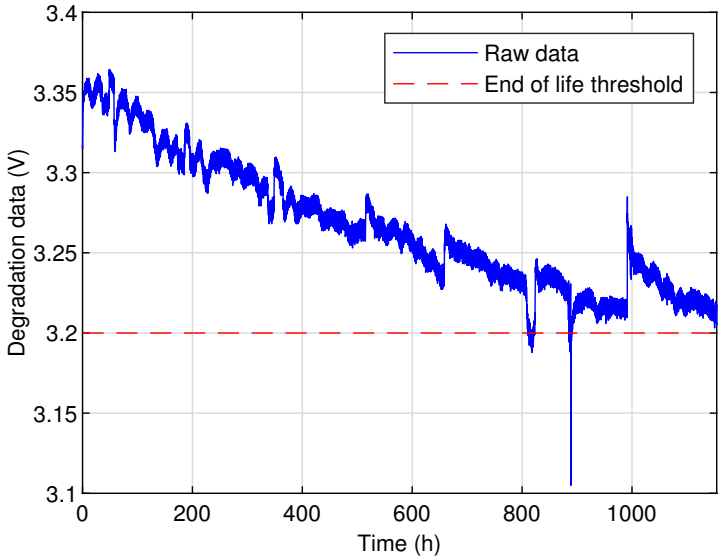

Figure 5. Profile of the available measurement of PEM-FC voltage.

The result of the application of the SG filter on the raw signal is given in Figure 6. This result shows that the initial profile of the degradation signal is preserved, but the noises and outliers are eliminated.

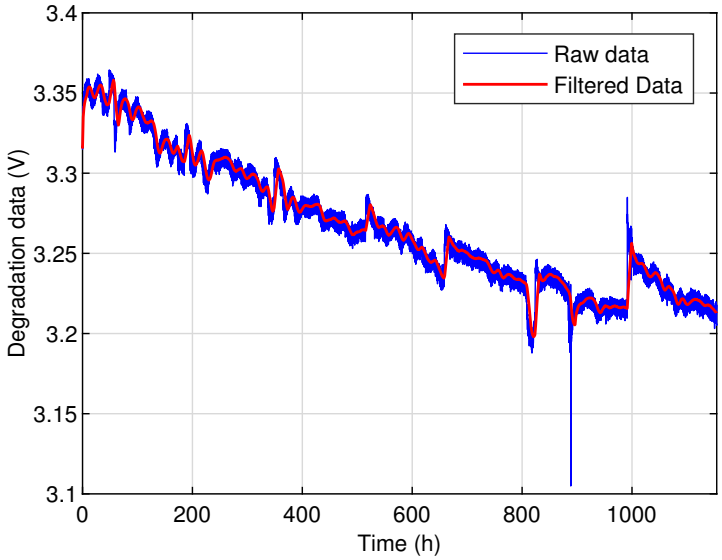

Figure 6. Smoothing result of the raw signal using SG filter.

6.3. Usability Analysis of the Smoothed Signal for RUL Prediction

The monotonicity metric is applied to the smoothed voltage measurement, and the results are given in Figure 7. The monotonicity score is around 0.82, thus demonstrating that the stack voltage is usable for failure prognosis. This result also shows that, compared to the other variables measured on the PEM-FC, the stack voltage is the most appropriate for failure prognosis, as it has the highest monotonocity score.

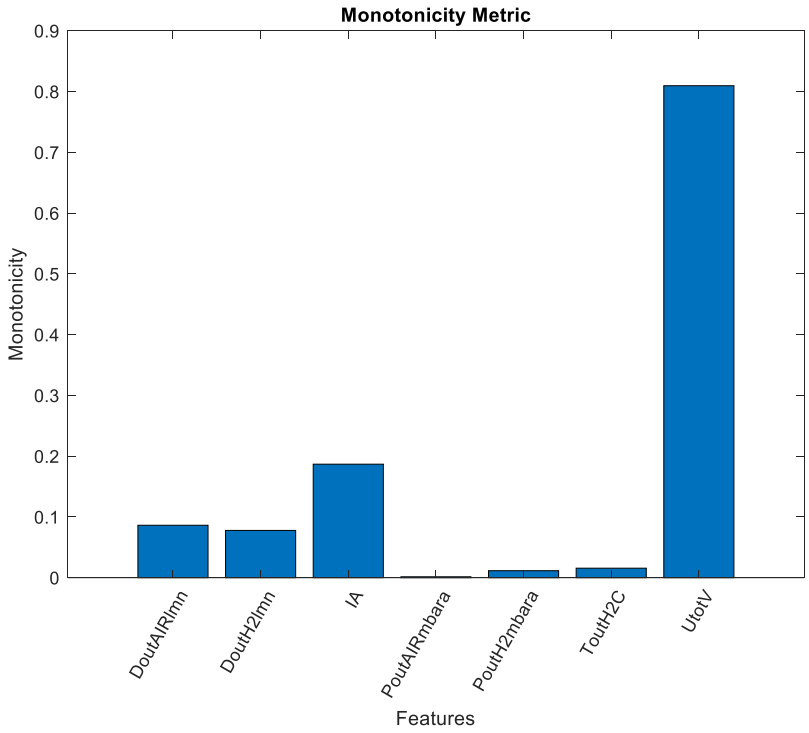

Figure 7. Score of Monotonicity analysis of the smoothed signal.

6.4. HI Estimation-Updating and RUL Prediction

The result of the estimation and updating of the HI is given in Figure 8. This result shows that after each update, a degradation trajectory is generated by the data-driven state model as a function of the current state of the system (value of the HI at the present time). This update makes it possible to take accelerations or decelerations of the degradation process into account. It can be noticed that each update gives an end-of-life (EOL) value. This property is used to calculate a sliding RUL that takes all the dynamics observed on the health index into account.

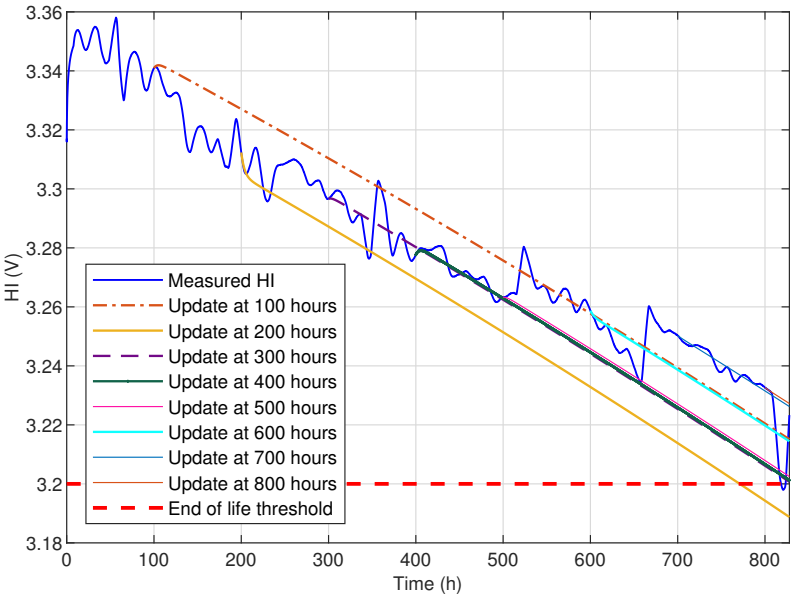

Figure 8. Result of the HI estimation, updating and time-projection until the failure threshold.

The evolution of the RUL estimation over time is given in Figure 9. It shows that the prediction of the RUL follows the changes in the degradation process of the system,

which is manifested by acceleration periods of the HI's progress. Updating the data-driven state model allows one to obtain an RUL estimation error of less than 100 h over a forecast horizon of 800 h, as shown in Figure 9.

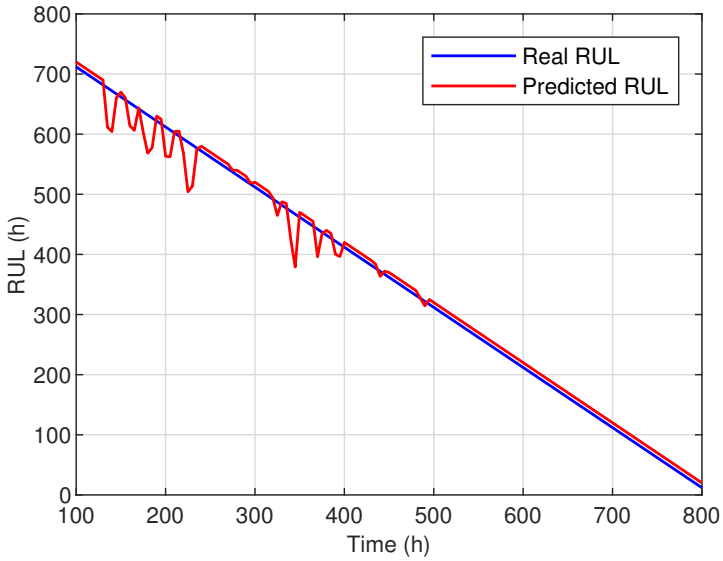

Figure 9. Result of the online estimation and updating of the health index.

The RUL estimation error is given in Figure 10. It shows that the estimation error of the RUL remains acceptable, within a confidence interval of less than 60 h. The prediction error is then reduced considerably when one approaches the real RUL at time t = 500 h, where the prediction error represents less than 20 h, i.e., less than a day of operation.

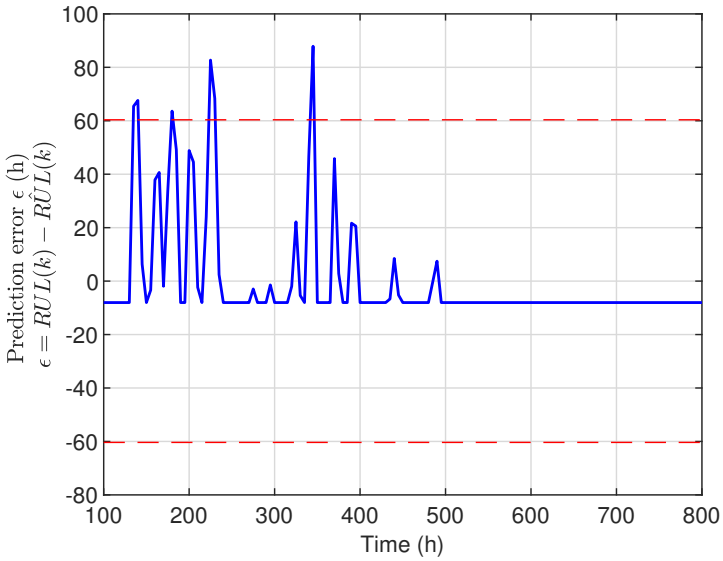

Figure 10. RUL prediction prror.

To better quantify the performance of the failure prognosis method proposed in this work, the relative error and the mean absolute percentage error (MAPE) are calculated and given in Figure 11. This figure shows a relative error of less than 20% and an MAPE of less than 5%. These results show that the performance of the proposed method is acceptable.

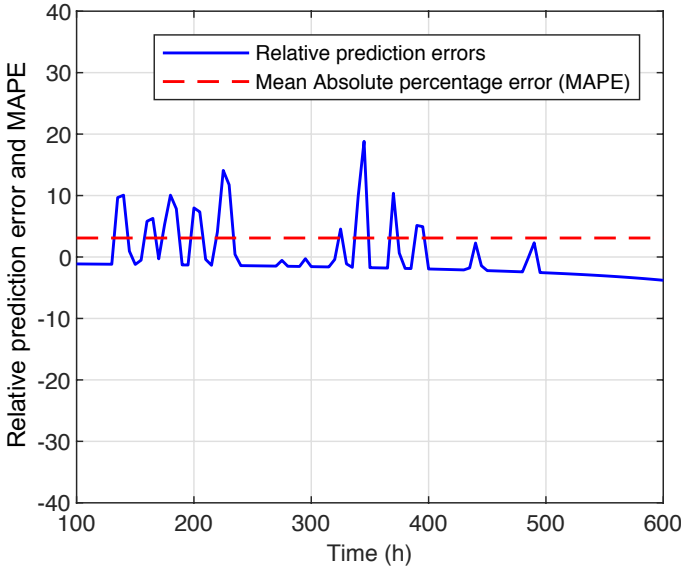

Figure 11. Relative prediction error and MAPE.

7. Conclusions

In this work, a generic method of failure prognosis of dynamic systems is proposed and applied to a PEM-FC. The first step concerns the system modeling for structural analysis and the identification of the measurement carrying the degradation process. The measured raw signal is then smoothed using a Savitsky–Golay filter, and is then analyzed using the monotonicity metric to check the usability of the measurement identified as the HI for failure prognosis. The trend of the identified HI is modeled using a data-driven state model whose order is identified by calculating the Hankel singular values for models of different orders, and parameters are identified and updated online using the least squares method. The RUL is then predicted and updated online through the time projection of the HI estimation until the failure threshold. The application of the proposed method to the PEM-FC shows its relevance and its ease of deployment. The results obtained show its effectiveness and the robustness of the RUL prediction to the changes in condition monitoring.

Author Contributions: Conceptualization, M.D., O.D., and S.B.; Formal analysis, M.D., O.D., and S.B.; Methodology, M.D.; Formal analysis, M.D., O.D., and S.B.; Investigation, M.D., O.D., and S.B.; Software, O.D. and M.D.; Supervision, M.D.; Writing—original draft, M.D.; Writing—review and editing, M.D., O.D., S.B., M.B., and J.-L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, R.H.; Xi, X.N.; Wang, P.N.; Wu, B.D.; Tian, S.M. Review on hydrogen fuel cell condition monitoring and prediction methods. *Int. J. Hydrogen Energy* **2019**, *44*, 5488–5498. doi:10.1016/j.ijhydene.2018.09.085.
2. Djeziri, M.; Benmoussa, S.; Zio, E. Review on Health Indices Extraction and Trend Modeling for Remaining Useful Life Estimation. In *Artificial Intelligence Techniques for a Scalable Energy Transition: Advanced Methods, Digital Technologies, Decision Support Tools, and Applications*; Sayed-Mouchaweh, M., Ed.; Springer International Publishing: Cham, Switzerland, 2020; pp. 183–223, doi:10.1007/978-3-030-42726-9_8.

3. Kira, K.; Rendell, L.A. The feature selection problem: Traditional methods and a new algorithm. In Proceedings of the 10th National Conference on Artificial Intelligence, San Jose, CA, USA, 12–16 July 1992; pp. 129–134.
4. Somol, P.; Pudil, P.; Kittler, J. Fast branch & bound algorithms for optimal feature selection. *IEEE Pattern Anal. Mach. Intell.* **2004**, *26*, 900–912.
5. Chen, X.W. An improved branch and bound algorithm for feature selection. *Pattern Recognit. Lett.* **2003**, *24*, 1925–1933.
6. Karnopp, D.; Rosenberg, R. *System Dynamics: A Unified Approach*; A Wiley-Interscience Publication; Wiley: Hoboken, NJ, USA, 1975.
7. Saisset, R.; Fontes, G.; Turpin, C.; Astier, S. Bond Graph model of a PEM fuel cell. *J. Power Sources* **2006**, *156*, 100–107, doi:10.1016/j.jpowsour.2005.08.040.
8. Hung, Y.; Lin, P.; Wu, C.; Hong, C. Real-time dynamic modeling of hydrogen PEMFCs. *J. Frankl. Inst.* **2008**, *345*, 182–203, doi:10.1016/j.jfranklin.2007.08.004.
9. Chatti, N.; Ould-Bouamama, B.; Gehin, A.; Merzouki, R. Signed Bond Graph for health monitoring of PEM fuel cell. In Proceedings of the 5th International Conference on Fundamentals and Development of Fuel Cells, Karlsruhe, Germany, 16–18 April 2013.
10. Benchouia, N.; Hadjadj, A.; Khochemane, L.; Bouziane, M. Bond graph modeling approach development for fuel cell PEMFC systems. *Int. J. Hydrogen Energy* **2014**, *39*, 15224–15231, doi:10.1016/j.ijhydene.2014.05.034.
11. Vasilyev, A.; Andrews, J.; Jackson, L.; Dunnett, S.; Davies, B. Component-based modelling of PEM fuel cells with bond graphs. *Int. J. Hydrogen Energy* **2017**, *42*, 29406–29421, doi:10.1016/j.ijhydene.2017.09.004.
12. Diaf, Y.; Benmoussa, S.; Djeziri, M. Time series model improving with automatic Savitzky—Golay filter for Remaining Useful Life estimation. In Proceedings of the 12th International Conference on Integrated Modeling and Analysis in Applied Control and Automation, Lisbon, Portugal, 18–20 September 2019; pp. 32–38.
13. P.Baraldi.; Bonfanti, G.; Zio, E. Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics. *Mech. Syst. Signal Process.* **2018**, *102*, 382–400.
14. Pravin, P.S.; Gudi, R.D.; Bhartiya, S. *Dynamic Modeling of an Integrated Reformer Membrane Fuel Cell System*; IFAC PapersOnline; IFAC: New York, NY, USA, 2017; pp. 10790–10795.
15. Peraza, C.; Diaz, J.D.; Arteaga-Bravo, F.J.; Villanueva, C.; Gonzalez-Longatt, F. Modeling and simulation of PEM Fuel cell with bond graph and 20sim. In Proceedings of the American Control Conference, Seattle, WA, USA, 11–13 June 2008; pp. 5104–5108.
16. Mzoughi, D.; Allagui, H.; Bouaicha, A.; Mami, A. Modeling and testing of a 1.2-kW Nexa fuel cell using bond graph methodology. *IEEJ Trans. Electr. Electron. Eng.* **2015**, *10*, 527–538.
17. Yang, Q.; Aitouche, A.; Ould-Bouamama, B. Structural Diagnosability of Fuel Cell Stack System Based on Bond Graph Tool. *IFAC Proc. Vol.* **2009**, *42*, 728–733, doi:10.3182/20090630-4-ES-2003.00121.
18. Jha, M.; Dauphin-Tanguy, G.; Bouamama, B.O. Particle Filter Based Hybrid Prognostics for Health Monitoring of Uncertain Systems in Bond Graph Framework. *Mech. Syst. Signal Process.* **2016**, *75*, 301–329.
19. Djeziri, M.; Bouamama, B.O.; Dauphin-Tanguy, G.; Merzouki, R. LFT Bond Graph Model-Based Robust Fault Detection and Isolation. In *Bond Graph Modelling of Engineering Systems*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 105–133.
20. Mzoughi, D.; Allagui, H.; Khili, N.; Mami, A. Bond graph modeling and control of a single PEM cell. In Proceedings of the I6th International Renewable Energy Congress (IREC), Sousse, Tunisia, 24–26 March 2015.
21. Petrone, R.; Zheng, Z.; Hissel, D.; Pera, M.C.; Pianese, C.; Sorrentino, M.; Becherif, M.; Yousfi-Steiner, N. A review on model-based diagnosis methodologies for PEMFCs. *Int. J. Hydrogen Energy* **2013**, *38*, 7077–7091.
22. Kandidayeni, M.; Macias, A.; Amamou, A.; Boulon, L.; Kelouwani, S.; Chaoui, H. Overview and benchmark analysis of fuel cell parameters estimation for energy management purposes. *J. Power Sources* **2018**, *380*, 92–104.
23. Javed, K.; Gouriveau, R.; Zerhouni, N.; Hissel, D. Data-driven prognostics of proton exchange membrane fuel cell stack with constraint based summation-wavelet extreme learning machine. In Proceedings of the 6th International Conference on Fundamentals and Development of Fuel Cells, FDFC, Toulouse, France, 3–5 February 2015; pp. 1–8.
24. Liu, H.; Chen, J.; Hou, M.; Shao, Z.; Su, H. Data-based short-term prognostics for proton exchange membrane fuel cells. *Int. J. Hydrogen Energy* **2017**, *42*, 20791–20808.
25. Zhu, L.; Chen, J. Prognostics of PEM fuel cells based on Gaussian process state space models. *Energy* **2018**, *149*, 63–73.
26. Wu, X.; Ye, Q. Fault diagnosis and prognostic of solid oxide fuel cells. *J. Power Sources* **2016**, *321*, 47–56.
27. Morando, S.; Jemei, S.; Hissel, D.; Gouriveau, R.; Zerhouni, N. Proton exchange membrane fuel cell ageing forecasting algorithm based on Echo State Network. *Int. J. Hydrogen Energy* **2017**, *42*, 1472–1480.
28. Mao, L.; Jackson, L. Selection of optimal sensors for predicting performance of polymer electrolyte membrane fuel cell. *J. Power Sources* **2016**, *328*, 151–160.
29. Wua, X.L.; Xua, Y.; Xuea, T.; Zhaoa, D.; Jianga, J.; Denga, Z.; Fub, X.; Lia, X. Health state prediction and analysis of SOFC system based on the data- T driven entire stage experiment. *Appl. Energy* **2019**, *248*, 126–140.
30. Dolenc, P.B.; Boškoski, M.S.A.P.J. State of health estimation and remaining useful life prediction of solid oxide fuel cell stack. *Energy Convers. Manag.* **2017**, *148*, 993–1002.
31. Hissel, D.; Pera, M. Diagnostic & health management of fuel cell systems: Issues and solutions. *Annu. Rev. Control* **2016**, *428*, 201–211.

32. Savitzky, A.; Golay, M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639.
33. Steinier, J.; Termonia, Y.; Deltour, J. Smoothing and differentiation of data by simplified least square procedure. *Anal. Chem.* **1972**, *44*, 1906–1909.
34. Coble, J. Merging Data Sources to Predict Remaining Useful Life—An Automated Method to Identify Prognostic Parameters. Ph.D. Thesis, University of Tennessee, Knoxville, TN, USA, 2010.
35. Ljung, L. *System Identification—Theory for the User*; PTR Prentice Hall: Upper Saddle River, NJ, USA, 1999.
36. Yin, H.; Zhu, Z.; Ding, F. Model order determination using the Hankel matrix of impulse responses. *Int. J. Hydrogen Energy* **2017**, *24*, 797–802.
37. Djeziri, M.; Benmoussa, S.; Mouchaweh, M.; Lughofer, E. Fault diagnosis and prognosis based on physical knowledge and reliability data: Application to MOS Field-Effect Transistor. *Microelectron. Reliab.* **2020**, *110*, 113682.
38. Zhong, K.; Ma, J.; Han, M. Online prediction of noisy time series: Dynamic adaptive sparse kernel recursive least squares from sparse and adaptive tracking perspective. *Eng. Appl. Artif. Intell.* **2020**, *91*, 103547, doi:10.1016/j.engappai.2020.103547.
39. Gouriveau, R.; Hilairet, M.; Hissel, D.; Jemea, S.; Jouin, M.; Lechartier, E.; Morando, S.; Pahon, E.; Péra, M.C.; Zerhouni, N. IEEE 2014 PHM data challenge. In Proceedings of the International Conference on Prognostics and Health Management (PHM14), Cheney, WA, USA, 22–25 June 2014.