

Dependence among complex random variables as a fuel cell condition indicator

Biljana Mileva Boshkoska^{a,b,*}, Pavle Boškosi^b, Andrej Debenjak^{b,c}, Đani Juričić^b

^aFaculty of Information Studies, Ulica talcev 3, SI-8000 Novo mesto, Slovenia

^bJožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

^cJožef Stefan International Postgraduate School, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

Abstract

As various faults alter the PEM fuel cell impedance characteristic over a broad frequency range, the electrochemical impedance spectroscopy is frequently employed for the purpose of condition monitoring. The proposed methodology treats the impedance components among different frequencies as dependent complex random variables. The information about fuel cell condition is incorporated into the dependence structure of these complex random variables. This dependence is described through the corresponding joint cumulative density function by employing copula functions. The benefits of such an approach are threefold: (i) the estimation of the joint cumulative density function requires only several measurements of a fuel cell in a fault-free condition, (ii) the procedure is computationally efficient, and (iii) the output of the copula function is directly used as an overall unit-free condition indicator. The approach was evaluated on a kW-range PEM fuel cell stack subjected to water management faults of various severities. The results show that the condition indicator corresponds with the severity of the induced faults.

Keywords: PEM Fuel Cell, Condition Monitoring, Flooding, Drying, Dependence Structure, Copula Functions

1. Introduction

A proton exchange membrane (PEM) fuel cell is an electrochemical device that converts the chemical energy of hydrogen directly into electrical energy [1]. When both electricity and heat are utilised in a cogeneration system, the overall efficiency of such a system can reach up to 90% [2]. However, the large-scale market penetration of PEM fuel cell technology is still impeded due to the durability and reliability issues of the technology [3–5]. Studies conducted over the last two decades have pointed out several kinds of faults influencing the optimal exploitation of PEM fuel cell systems [6, 7]. These faults include corrosion of the electrodes and degradation of membranes [8–11], catalyst and membrane poisoning [12–15], and water management faults (i.e., flooding of gas channels and membrane drying) [16–18]. Detection of these faults can be achieved by employing effectual diagnostics techniques [19, 20]. Such effective diagnostic information can be beneficial in various applications [21–23].

The majority of faults occurring in a PEM fuel cell affect its impedance characteristic. Consequently, the faults can be detected with a multitude of approaches based on electrochemical impedance spectroscopy (EIS) [24–28]. EIS is a technique of analysing electrochemical device through measurements of its electrical impedance characteristic. Typically, small-amplitude electrical current perturbation signals in the form of sine waves are applied to the fuel cell. By measuring the voltage response of the fuel cell, the impedance characteristic can be easily estimated [29]. In the field of PEM fuel cells, the EIS technique was already proven to be effective in detecting fuel cell flooding, membrane drying, and anode poisoning [30–35].

Despite the effectiveness of the EIS technique, current implementations have three major drawbacks:

1. The application of sine waves as perturbation signals covering a wide frequency region is time-consuming, hence inappropriate for online condition monitoring (CM).
2. Determining appropriate threshold values requires a priori characterisation of a fuel cell, which in many cases is infeasible.
3. Performing EIS characterisation over a large num-

*Corresponding author.

Email address: biljana.mileva@unm.fis.si (Biljana Mileva Boshkoska)

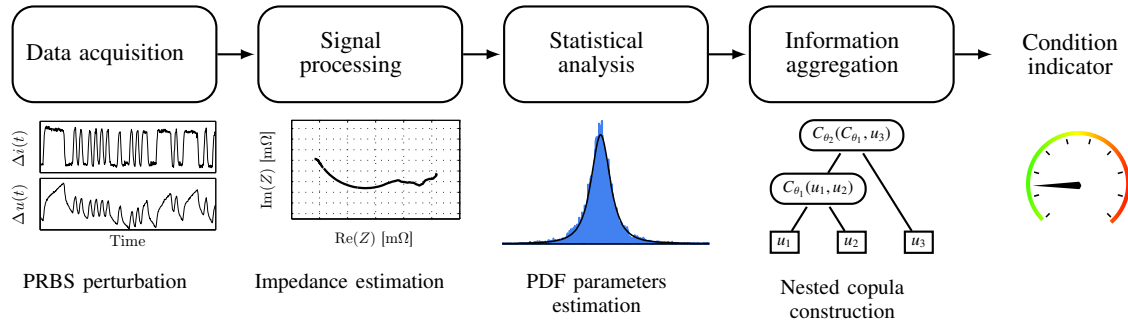


Figure 1: Schematic representation of the complete condition monitoring process

ber of frequencies results in a feature set with large cardinality. Therefore, specifying a single overall value describing the fuel cell condition is of significant practical merit.

The first issue can be resolved by employing a pseudo-random binary sequence (PRBS) perturbation signal and continuous wavelet transform (CWT) [36], which significantly decreases the time for impedance estimation. In an attempt to address the second issue, Bošković and Debenjak [37] provided a solution in which the individual impedance components are treated as complex random variables. Consequently, the probability density functions (PDFs) of the complex impedance components at a particular frequency exist and were mathematically derived [37]. The corresponding threshold values were determined by specifying the probability of false alarm (PFA). This study deals with the third issue by treating the impedance values at individual frequencies as dependent complex random variables. Using copula functions, their joint cumulative distribution function (CDF) is estimated and used to calculate the fuel cell condition indicator (CI). The complete condition monitoring process, from the data acquisition to the calculation of the CI, is shown in Figure 1.

Since the fuel cell condition affects the impedance components over a broad frequency range, in the statistical context, this implies that the complex impedance values at various frequencies should be regarded as dependent random variables. An appropriate description of the dependence among random variables is their multidimensional joint CDF. Specifying the joint CDF is a difficult task especially for a large number of dimensions [38], such as the issue at hand. Copula functions offer a solution to this problem by estimating the joint CDF using the marginal distributions. In the context of EIS, the marginal distributions are the CDFs of impedance at specific frequencies of interest. The multi-

variate joint CDF, specified in such a way, is capable of describing various types of dependences, including non-linear ones. So far, copulas were successfully applied in the petroleum industry [39], finance [40, 41], hydrology [42], biology [43], change detection in images [44], quality control [45], and machine learning [46]. To the authors' knowledge, this is the first application of this concept for condition monitoring of fuel cell systems.

The application of this concept has three practical merits. First, the initial estimation of the fault-free multivariate joint CDF is time efficient since it relies only on several measurements acquired while a fuel cell is in a fault-free condition. Second, the evaluation process is computationally efficient, which makes it suitable for implementation as an online CM system. Finally, as the copula function itself is a CDF, its output is a probability of observing a particular impedance characteristic. Therefore, the output is regarded as a unit-free CI describing the overall fuel cell condition.

The paper is organised as follows: Section 2 presents the way features are generated and discusses the statistical properties of the complex impedance values. The concept of copula functions is presented in Section 3. The complete CI computation process is presented in Section 4. Finally, Section 5 presents the evaluation results of the proposed technique, employed on an 8.5 kW fuel cell stack.

2. Feature set

The straightforward approach for impedance estimation is to excite the fuel cell using sine current excitation signal $i(t)$ with frequency f_0 and to measure the corresponding voltage signal $u(t)$. The impedance $Z(f_0)$ is then calculated as a ratio between the Fourier transform coefficients of the two signals [29]:

$$Z(f_0) = \frac{\mathcal{F}[u(t)]}{\mathcal{F}[i(t)]} \Big|_{f_0} \quad (1)$$

In order to cover a broad frequency range, the sine signals have to be separately generated and applied for each frequency of interest, which is time inefficient. An attempt to overcome this issue is by using compound multisine signals [47]. However, even in such a case, the impedance values are calculated only at discrete frequencies that are present in the generated multisine signal.

2.1. Impedance estimation with PRBS and CWT

In order to jointly cover a broad frequency range and perform the required measurements in a reasonably short time period, a broadband excitation is required. In this study, the PRBS signal is used as proposed by Debenjak et al. [36]. The PRBS can be regarded as sufficiently close to stationary random noise, and in a predefined frequency band, it exhibits power spectral density similar to the one of the white noise. Furthermore, the amplitudes of its frequency components have Gaussian distribution [48].

Using such an excitation signal, the Fourier analysis is unable to describe transient phenomena, located in a narrow time interval. Therefore, it is necessary to perform time-frequency analysis for the estimation of the impedance. The possible candidates are the short-time Fourier transform or wavelet transform. The former offers only fixed time-frequency resolution, whereas the latter one allows flexibility in the time-frequency resolution by interchangeably achieving good time resolution for the high-frequency events, and good frequency resolution for the low-frequency events. Therefore, in this study, the CWT approach based on the Morlet mother wavelet was employed for the impedance estimation. With properly tuned CWT parameters, this approach provides reliable impedance results along the entire frequency band. Additionally, it yields statistical information about the confidence interval of the impedance measurement.

The process of impedance estimation starts by superimposing the PRBS on the DC current $i_{dc}(t)$ supplied by the fuel cell. As a result, the fuel cell responds with the corresponding voltage $u(t)$ signal, which is directly related to the cell impedance. Performing the CWT with the Morlet mother wavelet on both $i(t)$ and $u(t)$ signals results in two sets of complex wavelet coefficients:

$$\begin{aligned} Wi(\tau, s) &= \langle i(t), \psi_{\tau, s}(t) \rangle = \int_{-\infty}^{\infty} i(t) \psi_{\tau, s}^*(t) dt, \\ Wu(\tau, s) &= \langle u(t), \psi_{\tau, s}(t) \rangle = \int_{-\infty}^{\infty} u(t) \psi_{\tau, s}^*(t) dt, \end{aligned} \quad (2)$$

where $\psi_{\tau, s}(t)$ is the translated and scaled version of the

Morlet wavelet [49]:

$$\psi_{\tau, s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (3)$$

and

$$\psi(t) = \pi^{-1/4} \left(e^{-j\omega_0 t} - e^{-\omega_0^2/2} \right) e^{-t^2/2}. \quad (4)$$

The relation between scale s and frequency f reads as follows:

$$\frac{1}{f} = \frac{4\pi s}{\omega_0 + \sqrt{2 + \omega_0^2}} \quad (5)$$

and, since the translation τ directly corresponds to time t , the impedance $z(t, f)$ is calculated as a ratio of the corresponding complex wavelet coefficients (2) [36]:

$$z(t, f) = \frac{Wu(t, f)}{Wi(t, f)}. \quad (6)$$

This segment of the condition monitoring process is presented by the first two modules in Figure 1.

2.2. Probability distribution function of impedance at individual frequency

By treating the fuel cell as a piecewise linear system, the response to such an excitation preserves the statistical properties of the PRBS excitation signal. Since the CWT is a linear transformation, the signal's properties are preserved in the calculated wavelet coefficients. Therefore, the wavelet coefficients (2) should be regarded as random variables as well.

In order to perform a statistical analysis of the impedance components, the PDF of the impedance components at a particular frequency $z(t, f)|_{f=f_0}$ has to be determined. This is a two-step process. First, the PDFs of the wavelet coefficients (2) have to be derived. Second, the PDF of their ratio (i.e., $z(t, f)|_{f=f_0}$) has to be derived in a closed form, bearing in mind that wavelet coefficients $Wi(t, f)$ and $Wu(t, f)$ are dependent complex random variables.

The wavelet coefficients $Wi(t, f)$ and $Wu(t, f)$ can be rewritten as

$$\begin{aligned} Wi(t, f) &= \Re\{Wi(t, f)\} + j\Im\{Wi(t, f)\}, \\ Wu(t, f) &= \Re\{Wu(t, f)\} + j\Im\{Wu(t, f)\}. \end{aligned} \quad (7)$$

Since the voltage $u(t)$ is the fuel cell response to the current excitation $i(t)$, and the excitation has the shape of PRBS, these coefficients are dependent zero-mean Gaussian circular complex random variables [37, 50]:

$$\begin{aligned} \Re\{Wi(t, f)\}, \Im\{Wi(t, f)\} &\sim \mathcal{N}\left(0, \frac{\sigma_i^2}{2}\right) \\ \Re\{Wu(t, f)\}, \Im\{Wu(t, f)\} &\sim \mathcal{N}\left(0, \frac{\sigma_u^2}{2}\right). \end{aligned} \quad (8)$$

Under these conditions, the PDF of the complex impedance (6) at a particular frequency f_0 reads as follows [37]:

$$p_Z(z) = \frac{1 - |\rho|^2}{\pi \sigma_u^2 \sigma_i^2} \left(\frac{|z|^2}{\sigma_u^2} + \frac{1}{\sigma_i^2} - 2 \frac{\rho_r z_r - \rho_i z_i}{\sigma_u \sigma_i} \right)^{-2}, \quad (9)$$

where z_r and z_i are real and imaginary components and $\rho = \rho_r + j\rho_i$ is a complex correlation coefficient, such that $|\rho| \leq 1$. For the purpose of condition monitoring, the module $|z|$ of the complex impedance is of particular interest. Under the same assumption for the stochastic nature of the impedance z , the PDF of its module reads as follows [37]:

$$p_{|z|}(|z|) = \frac{2\sigma_i^2 \sigma_u^2 (1 - \rho^2) |z| (\sigma_i^2 |z| + \sigma_u^2)}{[(\sigma_i^2 |z|^2 + \sigma_u^2)^2 - 4\rho^2 \sigma_u^2 \sigma_i^2 |z|^2]^{3/2}}, \quad |z| \geq 0, \quad (10)$$

where $|\rho| \leq 1$ is the correlation coefficient. The CDF of $|z|$ reads as follows [51]:

$$P_{|z|}(z) = \frac{1}{2} - \frac{\sigma_u^2 - z^2 \sigma_i^2}{2 \sqrt{4z^2 (1 - \rho^2) \sigma_i^2 \sigma_u^2 + (\sigma_u^2 - z^2 \sigma_i^2)^2}}, \quad (11)$$

where $z \geq 0$.

The parameters σ_u , σ_i and the correlation coefficient ρ can be estimated through the calculated wavelet coefficients as [52, 53]:

$$\begin{aligned} E\{Wu(t, f_0)Wu(t, f_0)^*\} &= \frac{\sigma_u^2}{2} \\ E\{Wi(t, f_0)Wi(t, f_0)^*\} &= \frac{\sigma_i^2}{2} \\ E\{Wi(t, f_0)Wu(t, f_0)^*\} &= \frac{\sigma_u \sigma_i}{2} \rho. \end{aligned} \quad (12)$$

2.3. Dependence structure among the impedance components

The deterioration of the fuel cell condition alters the impedance characteristic over a wide frequency range. Since the impedance at individual frequency is a random variable, the impedance values should be regarded as dependent random variables. Therefore, in order to properly determine the presence and the root cause of the fault, the overall change in the impedance characteristics has to be characterised. A proper way of describing the joint behaviour of random variables is through the corresponding joint CDF. In this study, the joint CDF is estimated by using copula functions.

3. Basics of copula functions

The joint PDF of two continuous random variables X_1 and X_2 is equal to:

$$p_{X_1, X_2}(x_1, x_2) = p_{X_2|X_1}(x_2|x_1)p_{X_1}(x_1) = p_{X_1|X_2}(x_1|x_2)p_{X_2}(x_2), \quad (13)$$

where $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$ are the marginal PDFs, and $p_{X_1|X_2}(x_1|x_2)$ and $p_{X_2|X_1}(x_2|x_1)$ are the corresponding conditional PDFs. The marginal distributions either are known or can be easily estimated from the available data. On the other hand, when x_1 and x_2 are dependent, the conditional PDFs are generally unknown and are usually difficult to estimate. Consequently, the derivation of the joint PDF is not trivial. An elegant way to connect the marginal PDFs of the random variables with their joint PDF is given through copulas.

3.1. Copula functions

For two random variables X_1 and X_2 , which are defined on the same probability space, the joint CDF $P_{X_1, X_2}(x_1, x_2)$ defines the probability of events in terms of the simultaneous occurrences of X_1 and X_2 . An explicit estimation of the joint CDF is a difficult task due to possibly complicated forms and the dimensionality issues [38]. Copulas simplify this process by expressing the joint CDF of dependent random variables as a function of their marginal distributions.

A bivariate copula function $C : [0, 1]^2 \rightarrow [0, 1]$ is an aggregation function that satisfies [54]:

1. $C(x, 0) = C(0, x) = 0$ and $C(x, 1) = C(1, x) = x$ for all $x \in [0, 1]$ (boundary conditions);
2. $C(x_1, y_1) - C(x_1, y_2) - C(x_2, y_1) + C(x_2, y_2) \geq 0$, for all $x_1, y_1, x_2, y_2 \in [0, 1]$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$ (2-increasing property).

Let the two random variables X_1 and X_2 have marginal CDFs $P_{X_1}(x_1)$ and $P_{X_2}(x_2)$, respectively, and joint CDF $P_{X_1, X_2}(x_1, x_2)$. According to Sklar's theorem [54, 55], there exists a multivariate CDF $C_\theta(u, v)$, where θ is a parameter that has to be estimated, and u and v are uniformly distributed random variables on the unit interval $[0, 1]$. Therefore, the joint CDF $P_{X_1, X_2}(x_1, x_2)$ reads as follows:

$$P_{X_1, X_2}(x_1, x_2) = C_\theta(P_{X_1}(x_1), P_{X_2}(x_2)) = C_\theta(u, v), \quad (14)$$

where u and v are obtained by using the inverse probability integral transform [56]:

$$\begin{aligned} u &= P_{X_1}(x_1), \quad u \sim \mathcal{U}(0, 1), \\ v &= P_{X_2}(x_2), \quad v \sim \mathcal{U}(0, 1). \end{aligned} \quad (15)$$

From (14), it follows that copula $C_\theta(u, v)$ is a function that couples the marginal CDFs $P_{X_1}(x_1)$ and $P_{X_2}(x_2)$ of the random variable $X = (X_1, X_2)$ with its joint distribution $P_{X_1, X_2}(x_1, x_2)$.

From the large collection of bivariate copulas, the paper addresses the Archimedean family, where copulas are constructed by using the following relation [57]:

$$C_\theta(u, v) = \varphi_\theta^{[-1]}(\varphi_\theta(u) + \varphi_\theta(v)) \quad (16)$$

where $u = P_{X_1}(x_1)$, $v = P_{X_2}(x_2)$, and $\varphi_\theta(\cdot)$ is called a generator function and $\varphi_\theta^{[-1]}(\cdot)$ is

$$\varphi_\theta^{[-1]}(t) = \begin{cases} \varphi_\theta^{-1}(t), & \text{if } 0 \leq t \leq \varphi_\theta(0); \\ 0, & \text{if } \varphi_\theta(0) \leq t \leq \infty. \end{cases} \quad (17)$$

The generator function $\varphi_\theta(t) : [0, 1] \rightarrow [0, \infty]$ must be continuous and strictly decreasing. Table 1 shows different types of Archimedean copulas constructed with the different generators $\varphi_\theta(t)$ [54]. Each of the families in Table 1 describes a different dependence structure. The selection is performed intuitively, as shown in Figure 2, Clayton and Gumbel copulas exhibit best results in the tails of the distribution, while Frank copula performs best under the bell of the distribution.

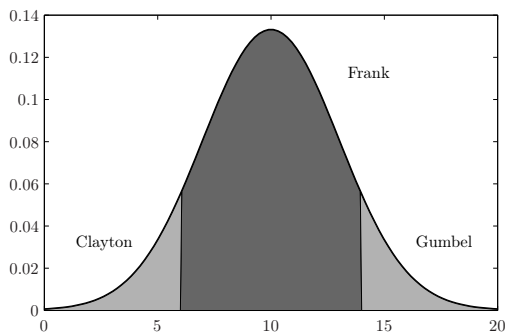


Figure 2: Regions of best performance for different Archimedean copulas

3.2. From bivariate to multivariate copulas

The copula, as defined by (14), couples two random variables into their joint CDF, hence the name bi-variate copula. However, majority of problems are multidimensional, and thus, the bivariate copulas should be extended to accommodate multidimensional cases. There are two approaches to extending to higher dimensional copulas [40, 58]:

1. The first approach extends the bivariate copula to multivariate copula using only one dependence parameter θ . Such copulas are known as exchange-

able Archimedean copulas [59]. The main drawback is that copulas for higher dimensions are tedious to derive [60].

2. The second approach uses bivariate copulas to form a hierarchical structure with at most $n - 1$ dependence parameters θ_i for n random variables. This construction of multivariate copulas is called fully nested Archimedean copula (FNAC) [59, 61].

As shown in Figure 3, FNAC is a treelike structure that is obtained using an iterative procedure that starts with coupling two random variables. For this particular FNAC structure, u_1 and u_2 are coupled into copula $C_{\theta_1}(u_1, u_2)$ with parameter θ_1 . In all subsequent iterations, the obtained copula is coupled with a new random variable; for example, copula C_{θ_1} is coupled with u_3 into $C_{\theta_2}(C_{\theta_1}, u_3)$ with parameter θ_2 and so on. The final output of the topmost copula reads

$$C_{\theta_4}(u_1, u_2, u_3, u_4, u_5) = C_{\theta_4}(u_5, C_{\theta_3}(u_4, C_{\theta_2}(u_3, C_{\theta_1}(u_1, u_2)))) \quad (18)$$

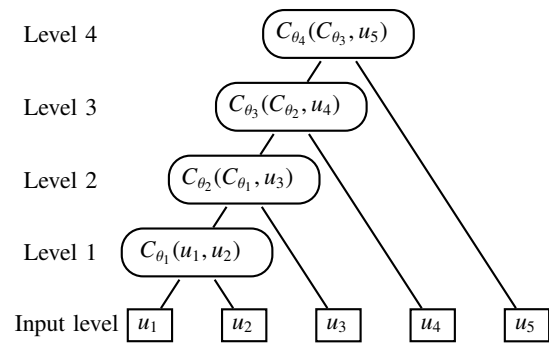


Figure 3: Fully nested Archimedean copula

Generally, a FNAC structure with n input variables has $n - 1$ parameters θ_i . The final function (18) represents a valid copula only if the following condition is fulfilled [62]:

$$\theta_1 > \theta_2 > \dots > \theta_{n-1} \quad (19)$$

where θ_1 is the parameter of the most nested copula, θ_2 is the parameter of the second most nested copula, and so on. The estimations of the values of θ_i , $i = 1, \dots, n - 1$ are obtained using the maximum likelihood algorithm.

Sometimes, it is not possible to build a FNAC structure that satisfies condition (19) for the desired order of variables in the FNAC. In such a case, a valid FNAC could be found in the set of all FNACs obtained by permuting the order of the variables entering the FNAC [63].

	$C_\theta(u, v)$	$\varphi_\theta(t)$	Solve $\left(\frac{\partial C_\theta(u, v)}{\partial u} = q, v\right)$
Clayton	$\left[\max(u^{-\theta} + v^{-\theta} - 1, 0)\right]^{-1/\theta}$	$\frac{1}{\theta} (t^{-\theta} - 1)$	$(1 - u^{-\theta} + (qu^{1+\theta})^{-\frac{\theta}{1+\theta}})^{-\frac{1}{\theta}}$
Frank	$-\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right)$	$-\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$\frac{1}{\theta} \log \frac{-e^\theta(1-q+qe^{\theta u})}{-e^\theta+qe^\theta-qe^{\theta u}}$
Gumbel-Hougaard	$\exp\left(-\left[(\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right)$	$\ln \frac{1 - \theta(1-t)}{t}$	only numerical solution

Table 1: Archimedean copulas

4. Condition indicator formation

The process of computing the CI by using the impedance joint CDF is composed of several steps. First, from the available frequency range, a low cardinality frequency set is selected, in a way that the fuel cell characteristic is preserved. Second, the impedance values (6) are calculated at the selected frequencies by using CWT (2). Afterwards, the parameters of the PDFs (9)–(10) are estimated by employing relation (12). It has to be noted that for each frequency f_i , the PDF of the wavelet coefficients (2) has different parameters. As a result, n different sets of parameters are calculated, one for each frequency f_i . In order to use the copula concept, the selected impedance values are transformed into a uniform random variable by using the CDF (11). Finally, the CI is computed as the output of the estimated multivariate copula.

4.1. Selection of the appropriate frequencies

Impedance values (6) are calculated on a broad frequency range. The results are represented by a Nyquist plot, as shown in Figure 4. Usually, the number of selected frequencies is quite large, which leads to a high cardinality feature set. The first step is to reduce the number of features while at the same time preserving the necessary diagnostic information.

In the context of PEM fuel cells, the shape of the Nyquist plot suggests that the characteristic is predominantly capacitive, and it embodies two depressed semicircles, as shown in Figure 4:

- The low-frequency one, $f \leq 5$ Hz, which is attributed to diffusion processes
- The high-frequency semicircle, $f > 5$ Hz, which is attributed to kinetic processes

Therefore, it is reasonable to choose a similar number of impedance values from both frequency regions sampled at logarithmically equal intervals. As a result, the final feature set comprises $N_f = 10$ impedance values that preserve the information about the impedance

shape. This feature set with reduced cardinality is subsequently used for the estimation of the fuel cell condition. A possible future improvement of the feature selection process can be achieved by employing, for instance, correlation-based feature selection methods, sequential forward(backward) feature selection, or minimum redundancy maximum relevance feature selection [64, 65].

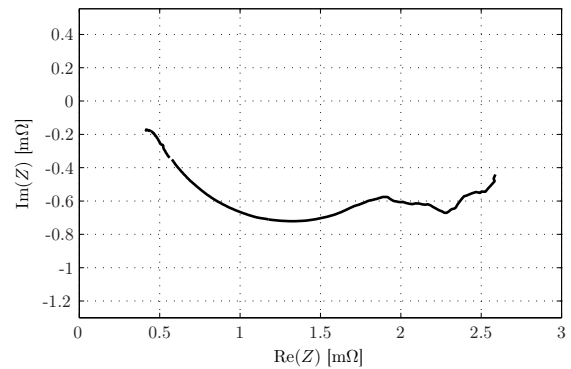


Figure 4: Nyquist plot of the PEM fuel cell impedance.

4.2. Estimating the copula parameters

First, the impedance is calculated by using the measured current $i(t)$ and voltage $u(t)$ signals. For each measurement, a complex matrix $Z(t, f)$ is obtained as a result of the CWT (6), which represents the instantaneous amplitude and phase of the impedance at each time moment over the selected frequency range. The complex matrix $Z(t, f)$ has dimensions $m \times N_f$, where $m = \text{sampling frequency} \times \text{duration}$ is the number of time moments at which $u(t)$ and $i(t)$ were measured, in the case at hand, $m = 40000$, $N_f = 10$.

By using the calculated impedance values at each of the selected frequencies f_i , $i = 1, \dots, N_f$, the corresponding parameters (12) of the CDFs are estimated. Afterwards, using the corresponding CDF, each value of the matrix $Z(t, f)$ is transformed into its uniform feature. As a result, a matrix \mathbf{U} with dimension $m \times N_f$

is obtained. Finally, the copula θ_i parameters (19) are estimated by employing the matrix \mathbf{U} .

4.3. Copula output as a fuel cell condition indicator

Since the copula function is a CDF, its output is a probability of observing a particular combination of uniformly distributed values $\mathbf{U} = [u_1, \dots, u_{N_f}]$. Through the experiment, for each measurement sequence q , a FNAC structure is built using uniform matrix \mathbf{U}_q . The process is shown in Figure 5.

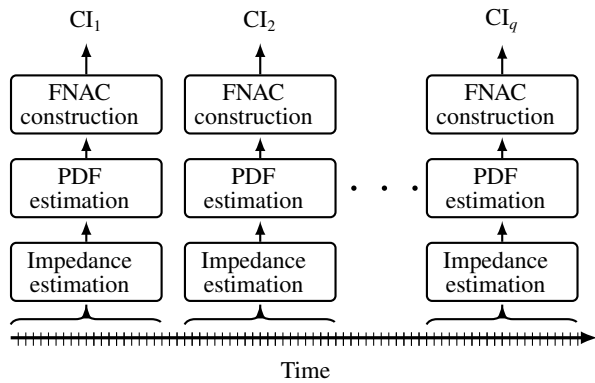


Figure 5: Measurement sequences in time

The calculated values from the initial fault-free measurements $q = 1$ are used as a baseline for CI. Since a deterioration in the condition will alter the impedance statistical characteristics, for the initially calculated copula, such a state will be regarded as a random event with a low likelihood. The low likelihood events (i.e., faults) lie in the tails of the PDF. As the output of CDFs is the probability $P[U \leq u]$, the output of the copula will be close to one for events that lie in the tail that is in the $(+, +, \dots, +)$ hyperoctant.¹ On the other hand, the output of the copula will be close to zero for events that lie in the tail that is in the $(-, -, \dots, -)$ hyperoctant. For a two-dimensional PDF, the $(+, +)$ quadrant is marked with 1 and the $(-, -)$ quadrant is marked with 2 in Figure 6. The CI baseline will be located somewhere in between.

5. Evaluation results

For the purpose of evaluation of the proposed approach, a set of $q \in [1, 105]$ impedance characteristics were estimated, each based on 40-second measurements. Figure 7 shows an example of the measured

¹The notation $(+, +, \dots, +)$ specifies the area determined by positive semiaxes in all dimensions. A two-dimensional special case is shown in Figure 6.

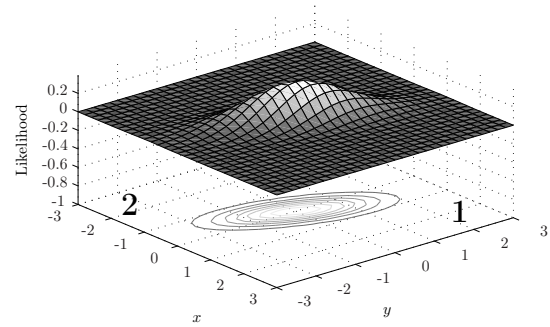


Figure 6: Description of the regions for high and low copula outputs

current $i(t)$ and voltage $u(t)$ signals, which were further used for feature extraction.

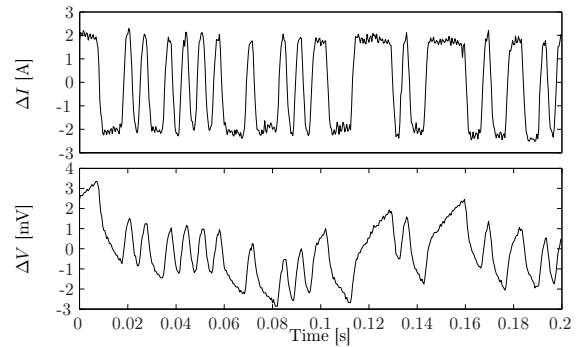


Figure 7: Example of current and voltage signals during the experiment

5.1. Experimental procedure

The experiment was performed on a commercially available PEM fuel cell system HyPM HD 8 produced by the Hydrogenics Corporation. The stack consists of 80 PEM fuel cells, each with surface area of 200 cm², providing 8.5 kW of electric power in total. The fuel cell system operates on pure hydrogen and ambient air. The impedance was measured on individual cells of the stack, where the PRBS perturbation signal was applied in galvanostatic mode.

During the experiment, the fuel cell system was kept at constant operating and environmental conditions at a temperature of 50°C, stoichiometry 2.5. The temperature of the airflow was kept constant at 50°C, and the relative humidity was controlled in order to introduce drying or flooding conditions. At the anode side, the fuel cell was fed with pure and dry hydrogen at a constant temperature of 20°C. The DC current operating point I_{dc} was set to 70 A, resulting in the stack voltage of 55 V, whereas the amplitude of the superposed PRBS waveform was set to 2 A.

The experiment went through five stages. In the first stage of approximately 11 minutes, 18 measurement sets were acquired in the fault-free state at which the relative humidity of the inlet air was kept at 9%. Then air fed to the fuel cell was slightly dried down to a relative humidity of 5% for about 10 minutes, between the 12th and the 22nd minute, in order to induce membrane drying, and again humidified to 9%. Afterwards, between the 34th and the 58th minute, saturated air with a relative humidity of 100% was fed into the fuel cell system, inducing the flooding of fuel cells. Finally, the air humidity was decreased down to the initial 9%. Using such a cycle, the response of the fuel cell under different fault scenarios with various fault severities was evaluated.

5.2. Time evolution of condition indicator

Following the procedure described in Section 4, the CI was computed for each of the $q \in [1, 105]$ measurements. The time evolution of the CI is presented in Figure 8, where the experiment time is represented on x -axis.

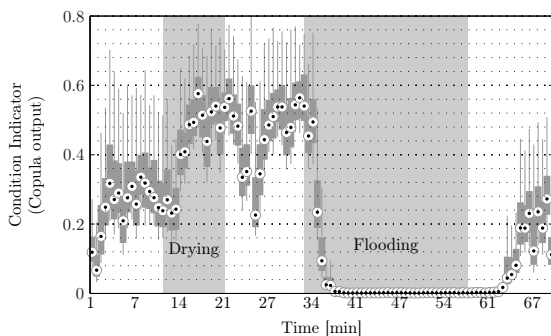


Figure 8: Evolution of the health index (copula output)

At the beginning of the experiment, the copula value defines the CI baseline of a fault-free condition. Dry inlet air of 5% relative humidity causes fuel cells to dry out, hence the increase of the CI. This indicates that some of the impedance components changed, and consequently, the observed measurements belong to a low likelihood tail. On the other hand, the saturated inlet air with a relative humidity of 100% causes the flooding of fuel cells. As a result, the CI reaches almost zero. At the end of the experiment, the humidity of the inlet air was set back to nominal values, resulting in fuel cell recovery, causing the CI to return to the baseline values similar to the ones from the beginning of the experiment.

It has to be noted that the experiment was performed on an industrial-grade fuel cell system. As such, the

built-in controls managed various parameters such as air and fuel flow (stoichiometry) as well as temperature. Therefore, these parameters were not strictly held within sharp boundaries with laboratory precision. As a result, some fluctuations can be observed in the CI values throughout the experiment (Figure 8). Regardless of these fluctuations, the fault regions are distinctly visible, which confirms the link between CI and the actual fuel cell condition.

The results show that the proposed CI (i.e., copula output) can also be directly related to the fault severity. The departure of the copula values for the drying fault from the baseline level is smaller compared with the copula values during the flooding fault, which corresponds to the induced fault severity.

6. Conclusion

This paper presents an approach for fuel cell condition monitoring by employing PRBS excitation signals, CWT for impedance estimation and copula functions for information fusion. The calculated large cardinality feature set comprises dependent complex random variables, which represent fuel cell impedance at individual frequencies. The fuel cell condition is hidden in the dependence structure of these complex random variables. The application of copula functions allows the proper characterisation of the dependence by estimating the multidimensional joint CDF. Therefore, the CI, defined as the output of the copula structure, is a probability of observing particular impedance values and is directly related to the fuel cell condition as well as the severity of the present fault.

The proposed approach has several practical merits. First, for the purpose of fault detection, the proposed CI can be implemented without any prior characterisation of the fuel cell under various faults. The CI baseline can be estimated from data acquired solely from measurements under a fault-free operation. Second, the approach is computationally efficient. Finally, the estimation of PDF parameters and the transformation of the acquired data into uniform random variables exist in closed form. This makes the proposed approach a suitable solution for an embedded condition monitoring system.

The applicability of the proposed approach is not limited either to PEM fuel cells or to water management faults. It can be applied in an unmodified form wherever the EIS approach is applicable.

Acknowledgments

The work is supported by the Creative Core FISNM-3330-13-500033 “Simulations” project funded by the European Union, the European Regional Development Fund. The authors also acknowledge the financial support of the Slovenian Research Agency through the Research Programme P2-0001.

- [1] F. Barbir, *PEM Fuel Cells: Theory and Practice*, Elsevier, 2005.
- [2] V. Jovan, M. Perne, J. Petrovčič, *Energ. Convers. Manage.* 51 (2010) 2467–2472.
- [3] J. Wu, X. Z. Yuan, J. J. Martin, H. Wang, J. Zhang, J. Shen, S. Wu, W. Merida, *J. Power Sources* 184 (2008) 104–119.
- [4] X.-Z. Yuan, H. Li, S. Zhang, J. Martin, H. Wang, *J. Power Sources* 196 (2011) 9107–9116.
- [5] M. Fowler, R. F. Mann, J. C. Amphlett, B. A. Peppley, P. R. Roberge, *Reliability issues and voltage degradation*, John Wiley & Sons, Ltd, 2005.
- [6] F. A. de Bruijn, V. A. T. Dam, G. J. M. Janssen, *Fuel Cells* 8 (2008) 3–22.
- [7] W. Schmittinger, A. Vahidi, *J. Power Sources* 180 (2008) 1–14.
- [8] H. R. Colón-Mercado, B. N. Popov, *J. Power Sources* 155 (2006) 253–263.
- [9] R. Borup, J. Meyers, B. Pivovar, Y. S. Kim, R. Mukundan, N. Garland, D. Myers, M. Wilson, F. Garzon, D. Wood, P. Zelenay, K. More, K. Stroh, T. Zawodzinski, J. Boncella, J. E. McGrath, M. Inaba, K. Miyatake, M. Hori, K. Ota, Z. Ogumi, S. Miyata, A. Nishikata, Z. Siroma, Y. Uchimoto, K. Yasuda, K.-i. Kimijima, N. Iwashita, *Chem. Rev.* 107 (2007) 3904–3951.
- [10] P. T. Yu, Z. Liu, R. Makharia, *J. Electrochem. Soc.* 160 (2013) F645–F650.
- [11] J. D. Fairweather, D. Spornjak, A. Z. Weber, D. Harvey, S. Wessel, D. S. Hussey, D. L. Jacobson, K. Artyushkova, R. Mukundan, R. L. Borup, *J. Electrochem. Soc.* 160 (2013) F980–F993.
- [12] C. Farrell, C. Gardner, M. Ternan, *J. Power Sources* 171 (2007) 282–293.
- [13] X. Cheng, Z. Shi, N. Glass, L. Zhang, J. Zhang, D. Song, Z.-S. Liu, H. Wang, J. Shen, *J. Power Sources* 165 (2007) 739–756.
- [14] W.-M. Yan, H.-S. Chu, M.-X. Lu, F.-B. Weng, G.-B. Jung, C.-Y. Lee, *J. Power Sources* 188 (2009) 141–147.
- [15] G. Postole, A. Auroux, *Int. J. Hydrog. Energy* 36 (2011) 6817–6825.
- [16] H. Li, Y. Tang, Z. Wang, Z. Shi, S. Wu, D. Song, J. Zhang, K. Fatih, J. Zhang, H. Wang, Z. Liu, R. Abouatallah, A. Mazza, *J. Power Sources* 178 (2008) 103–117.
- [17] N. Yousfi-Steiner, P. Moçotéguy, D. Candusso, D. Hissel, A. Hernandez, A. Aslanides, *J. Power Sources* 183 (2008) 260–274.
- [18] T. Ous, C. Arcoumanis, *J. Power Sources* 240 (2013) 558–582.
- [19] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, *Int. J. Hydrog. Energy* 38 (2013) 15307–15317.
- [20] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, *Int. J. Hydrog. Energy* 39 (2014) 481–494.
- [21] B. Pregelj, D. Vrečko, V. Jovan, *J. Power Sources* 196 (2011) 9419–9428.
- [22] A. J. Verhage, J. F. Coolegem, M. J. Mulder, M. H. Yildirim, F. A. de Bruijn, *Int. J. Hydrog. Energy* 38 (2013) 4714–4724.
- [23] B. Pregelj, D. Vrečko, J. Petrovčič, V. Jovan, G. Dolanc, *Applied Energy* 137 (2015) 64–76.
- [24] X.-Z. Yuan, H. Wang, J. C. Sun, J. Zhang, *Int. J. Hydrog. Energy* 32 (2007) 4365–4380.
- [25] J. Wu, X. Z. Yuan, H. Wang, M. Blanco, J. J. Martin, J. Zhang, *Int. J. Hydrog. Energy* 33 (2008) 1735–1746.
- [26] R. Petrone, Z. Zheng, D. Hissel, M. Péra, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner, *Int. J. Hydrog. Energy* 38 (2013) 7077–7091.
- [27] Z. Zheng, R. Petrone, M. Péra, D. Hissel, M. Becherif, C. Pianese, N. Y. Steiner, M. Sorrentino, *Int. J. Hydrog. Energy* 38 (2013) 8914–8926.
- [28] S. M. Rezaei Niya, M. Hoorfar, *J. Power Sources* 240 (2013) 281–293.
- [29] X.-Z. Yuan, C. Sons, H. Wang, J. Zhang, *Electrochemical Impedance Spectroscopy in PEM Fuel Cells, Fundamentals and Applications*, Springer, London, 2010.
- [30] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, B. Ould-Bouamama, *J. Power Sources* 159 (2005) 905–913.
- [31] W. Merida, D. Harrington, J. L. Canut, G. McLean, *J. Power Sources* 161 (2006) 264–274.
- [32] J.-M. Le Canut, R. M. Abouatallah, D. A. Harrington, *J. Electrochem. Soc.* 153 (2006) A857–A864.
- [33] A. Debenjak, M. Gašperin, B. Pregelj, M. Atanasijević-Kunc, J. Petrovčič, V. Jovan, *J. Mech. Eng.* 59 (2013) 56–64.
- [34] S. Cruz-Manzo, R. Chen, *J. Electrochem. Soc.* 160 (2013) F1109–F1115.
- [35] M. Gašperin, P. Boškosi, A. Debenjak, J. Petrovčič, *Fuel Cells* 14 (2014) 457–465.
- [36] A. Debenjak, P. Boškosi, B. Musizza, J. Petrovčič, Đ. Juričić, *J. Power Sources* 254 (2014) 112–118.
- [37] P. Boškosi, A. Debenjak, *J. Power Sources* 268 (2014) 692–699.
- [38] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, 1986.
- [39] M. Al-Harthy, S. Begg, R. B. Bratvold, *J. Petrol. Sci. Eng.* 57 (2007) 195–208.
- [40] M. Fischer, C. Kock, S. Schluter, F. Weigert, *Quant. Financ.* 9 (2009) 839–854.
- [41] E. Bouyé, V. Durrleman, A. Nikeghbaliand, G. Riboulet, T. Roncalli, 2000, *Copulas for finance - a reading guide and some applications*.
- [42] C. Genest, A. C. Favre, *J. Hydrol. Eng.* (2007).
- [43] J. M. Kim, Y. S. Jung, E. A. Sungur, K. H. Han, C. Park, I. Sohn, *BMC Bioinformatics* (2008).
- [44] G. Mercier, G. Moser, S. B. Serpico, *IEEE T. Geosci. Remote.* 46 (2008).
- [45] B. Mileva Boshkoska, M. Bohanec, P. Boškosi, Đ. Juričić, *J. Intell. Manuf.* (2013) 1–13.
- [46] S. Jaimungal, E. K. H. Ng, in: *ECML PKDD 2009*, pp. 628–643.
- [47] C. Brunetto, A. Moschetto, G. Tina, *Electric Power System Research* 79 (2008) 17–26.
- [48] E. O. Doebelin, *System Modeling and Response*, John Wiley & Sons Inc, 1980.
- [49] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, 1992.
- [50] P. O. Amblard, M. Gaeta, J. L. Lacoume, *Signal Proc.* 53 (1996) 1–13.
- [51] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*, Springer Science, New York, 2006.
- [52] E. Ollila, *IEEE Signal Processing Letters* 15 (2008) 841–844.
- [53] A. Khurshid, Z. Al-Hemyari, S. Kamal, *Pak. J. Stat. and Oper. Res.* 8 (2012) 645–654.
- [54] R. B. Nelsen, *An Introduction to Copulas*, 2nd ed., Springer, New York, 2006.
- [55] A. Sklar, *Distributions with Fixed Marginals and Related Topics - Random Variables, Distribution functions, and Copulas - A Personal Look Backward and Forward*, volume 28, Institute of Mathematical Statistics, Hayward, CA., 1996.

- [56] L. M. Leemis, S. K. Park, *Discrete-Event Simulation: A First Course*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2005.
- [57] H. Joe, *Multivariate Models and Dependence Concepts*, Chapman and Hall, 1997.
- [58] C. Savu, M. Trade, in: *International Conference on High Frequency Finance*, Konstanz, Germany.
- [59] D. Berg, K. Aas, *Eur. J. Financ.* 15 (2009) 639–659.
- [60] P. Trivedi, Z. David, *Copula Modeling: An Introduction for Practitioners*, World Scientific Publishing, 2006.
- [61] M. Hofert, in: *Copula Theory and its Applications*, Proceedings of the Workshop held in Warsaw 25–26 September, 2009, *Lecture Notes in Statistics*, Springer, 2010, pp. 147–160.
- [62] S. T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*, Elsevier/North Holland, 2003.
- [63] B. Mileva Boshkoska, M. Bohanec, *Int. J. Decis. Support Syst. Technol.* 4 (2012) 42–58.
- [64] R. Onanena, L. Oukhellou, D. Candusso, F. Harel, D. Hissel, P. Aknin, *Int. J. Hydrog. Energy* 36 (2011) 1730–1739.
- [65] D. Benouioua, D. Candusso, F. Harel, L. Oukhellou, *Int. J. Hydrog. Energy* 39 (2014) 21631–21637.