Energy conversion and management 195 (2019) 76–85 doi: 10.1016/j.enconman.2019.05.004

Fault detection of fuel cell systems based on statistical assessment of impedance data

Martin Stepančič[a,b,*], Đani Juričić[a], Pavle Boškoski[a]

[a] Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[b] Jožef Stefan International Postgraduate School, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

Abstract

Accurate online health assessment of fuel cell systems is a key for the timely mitigation and maintenance actions to be taken in order to maximise reliability of operation and useful life span of the cells. The majority of approaches rely on occasional probing of the system with small-amplitude signals around an operating point. The responses are then used to create either a parametric or a non-parametric model of the linearised system dynamics. However, during the probing session, the measurements might be corrupted with random noise and disturbances. Consequently, the evaluated parameters, being points on the impedance curve, parameters of the equivalent circuit models or the distribution of relaxation times, contain some uncertainty. That fact is largely ignored in the state of the art techniques, meaning that only mean value estimates are taken into account in the further analysis. In this paper we use a non-parametric two-sample Kolmogorov-Smirnov test to detect a change in the internal condition by evaluating changes at each frequency point on the Nyquist curve. Moreover, we show that in some cases it is even possible to isolate the fault origin from the pattern of detected changes. The applicability of the approach is demonstrated on the detection of water management faults of an industrial proton exchange membrane fuel cell system.

Keywords: Kolmogorov-Smirnov test, electrochemical impedance spectroscopy, distribution of relaxation times, wavelet transform, hypothesis testing, fault detection.

1. Introduction

During their operation, fuel cell systems might encounter a variety of fault and degradation modes. Timely detection and isolation of the root cause is therefore essential in order to take timely mitigation and maintenance actions. Hence, one can help increase system reliability and extend its useful life. The majority of health assessment approaches rely on occasional system probing with small-amplitude signals around an operating point. The responses are then used to describe the local linearised dynamics in terms of either a parametric or a non-parametric model. Due to random noises and disturbances that affect

*Corresponding author.
Email address: martin.stepancic@ijs.si (Martin Stepančič)

the measurements, the evaluated parameters that are values of the impedance curve, parameters of the equivalent circuit models or distribution of relaxation times, tend to contain some uncertainty.

A change in the internal health condition, due to either a degradation mechanism or a fault, can affect a range of features like the shape of the impedance curve, or one or more parameters of the associated equivalent circuit model (ECM). Finding the relationships between changes in the pattern of the available features and changes in health condition is of great importance for the effective on-line condition monitoring.

In the majority of the works published so far, the researchers have mainly paid attention to the relative change in the features at the end and at the start of the experiment, i.e. before and after the appearance of a fault, while little attention has been dedicated to early detect of incipient change in the features *during the online operation*. Finding out the root cause for the fault turns to be even more challenging. In this paper we propose a simple and effective approach to the incipient fault detection and (possibly) fault isolation, which relies on revealing changes at each point of the impedance curve. It builds on the statistical hypothesis testing and takes into account the stochastic nature of the impedance data.

Fault detection is performed in two main steps. The first one is *feature extraction*, i.e. evaluation of some characteristic quantities out of the available measurement data. Ideally, a feature should be sensitive to at least one or more fault modes while, at the same time, remaining insensitive to random noises and disturbances. The most often exploited features are parameters of the equivalent circuit model [1, 2], distribution of relaxation times [3–5], geometrical properties of the impedance curve [6–8], and parameters of the physical model [9–13].

In the second step, it should be determined whether a feature has changed or not. A conventional and rather straightforward way to do this is to check whether a feature exceeds a prescribed threshold [14]. The idea, however, has two weaknesses. First, due to noise and disturbances in the measurements, the features may frequently cross the threshold values, hence causing intermittent alarms. That effect is referred to as *diagnostic instability* [15] and can quickly cause a loss of confidence in the diagnostic system by the end user. One remedy is to apply longer measurement sessions, which might help to filter out the effects of noise. A better way is to consider the stochastic nature of the measurements and make use of methods for statistical decision making. Second, selecting optimal threshold values requires substantial background knowledge. More precisely, one requires information on the sensitivity of a feature with respect to the fault. If the sensitivity is high, higher thresholds could be used, which is good for reducing problems with false alarms due to noise. In the opposite case, a feature may change slightly even if a considerable fault is present. In that case, the thresholds should be put lower. In practice, however, the sensitivities of the symptoms to the faults are only rarely available.

This paper is based on a different rationale. Short, more frequent, and statistically independent measurement sessions allow for the use of the entirely data driven framework for detecting statistically significant changes in the evaluated features. There are few published results that treat this aspect in the field of electrochemical systems [16–19]. For the approach to work, it is important to define the reference behaviour of the system. Then we can choose either the null hypothesis, which says that the current data do not differ from the reference data, or the alternative, i.e. that a change occurred.

Great care should be devoted to the assumptions under which stochastic analysis is performed. The test statistic from [16] is based on impedance evaluated for each wavelet scale s separately. Under strict assumptions of linearity, the wavelet coefficients are results of a random process and in ideal conditions obey the Rayleigh distribution [17, 20]. Even a minute departure from the assumed distribution will make the hypothesis test inconsistent. A possible remedy is not to wrestle with the assumptions and make use of non-parametric statistical tests, which rely on the *empirical* distributions. The nonparametric approach requires two samples of the test statistic, e.g. one from the fault-free operating condition and the other from the current measurements. The continuous wavelet transform (CWT) presented in [16] is used. Since CWT is an ergodicity-preserving transformation, the set of wavelet coefficients evaluated at different translations in time could be treated as if they were obtained from multiple successive experiments. Hence, it becomes possible to employ the nonparametric hypothesis testing to detect changes in the electrochemical impedance spectroscopy (EIS) curves.

In this paper, we use the nonparametric Kolmogorov-Smirnov (KS) test. A remarkable property of the KS test statistic is that it has a known distribution irrespective of the distribution of wavelet coefficients. In the context of EIS, the wavelet coefficients are treated as elements of a statistical sample. The statistical test evaluates whether there is a significant change between the current and the reference sample. The KS test is performed at each wavelet scale (which corresponds to particular frequency). Consequently, the KS test, in essence, evaluates whether there is a significant change in the impedance at each frequency.

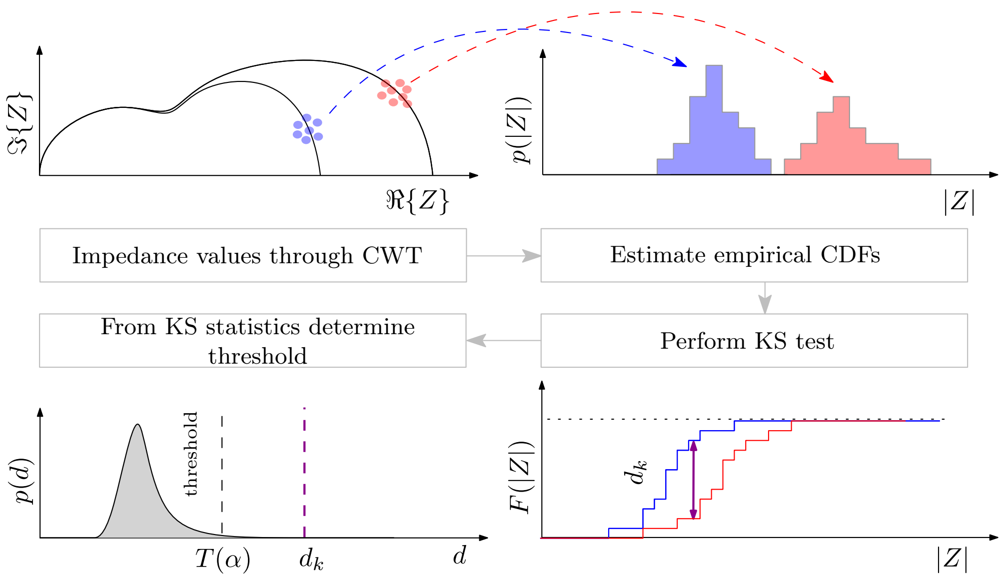


Figure 1: Schematic representation of the proposed approach

The proposed approach is graphically shown in Figure 1. Repeated evaluation of impedance at a given frequency will result in an ensemble of impedance values scattered like a cluster around the true value (shown in the leftmost plot). To cope with the randomness in impedance realisations one can simply extend the measurement session and perform averaging over the results. However, in order to keep the measurement sessions reasonably short, we simply evaluate the empirical distributions of the CWT coefficients for each frequency. With those empirical distributions, it is possible to perform a two-sample KS test in order to make a judgement about the changes in the impedance points. The threshold governing the decision is defined by the tolerated false alarm rate α as well as the sample size through the so-called power of the test.

In what follows, we will first explain the rationale for the approach in Section 2. The aim is to clarify the relationship between the internal fuel cell condition and external behaviour captured within the EIS curve. Cell condition is reflected in the parameters of the ECM and its eigenmodes, which are explicit in the distribution of relaxation times (DRT). We will stress that a change in a particular component of the ECMs affects a particular part of the impedance curve. In Section 3, a brief description of the CWT and its application for performing impedance spectroscopy is provided. Section 4 presents the KS based hypothesis testing with guidelines for selecting the sample size and the significance level. Finally, the experimental results obtained on a proton exchange membrane (PEM) fuel cell system are presented in Section 5.

2. The rationale for the approach: a simulation study

Changes on the Nyquist curve are related to the changes in the internal parameters of the system. A fault or a degradation mode will affect a particular frequency range of the impedance curve. The ECM captures the most relevant internal processes in the cell. Its parameters bear clear physical interpretation and can be associated with certain fault modes. The models include a particular component called constant phase element Q whose impedance is[1] [23]:

$$Z_Q(j\omega) = \frac{1}{(j\omega)^\alpha Q}, \tag{1}$$

where $\alpha_i \in \mathbb{R}^+$ is the order of the pole. For the special case $\alpha = 1$, the constant phase element reduces to a capacitor.

A sufficiently accurate model of fuel cell linearised dynamics [4, 5, 24] can be represented by a series of parallel connected resistors R and constant-phase elements Q (cf. Figure 2) as:

$$Z(j\omega) = R_0 + \sum_{i=1}^{k} \frac{R_i}{1 + (j\omega)^{\alpha_i} R_i Q_i}, \tag{2}$$

where R_0 is the series resistance, R_i and Q_i are the parameters of each pole and $\alpha_i \in \mathbb{R}^+$ is the order of the i^{th} pole.

[1] The impedance of Q element is not uniquely defined, for instance there are examples where $Z_Q(j\omega) = \frac{1}{(j\omega Q)^\alpha}$ [21]. Depending on the definition, the units of Q also vary. This paper follows the notation as stated in [22].

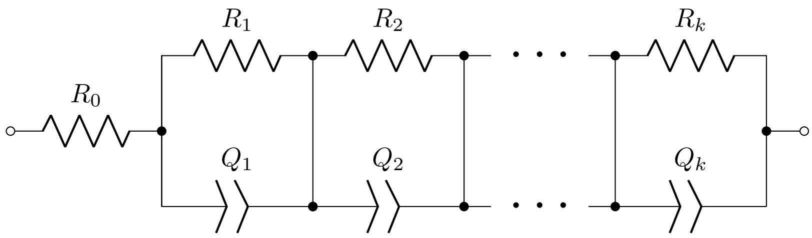


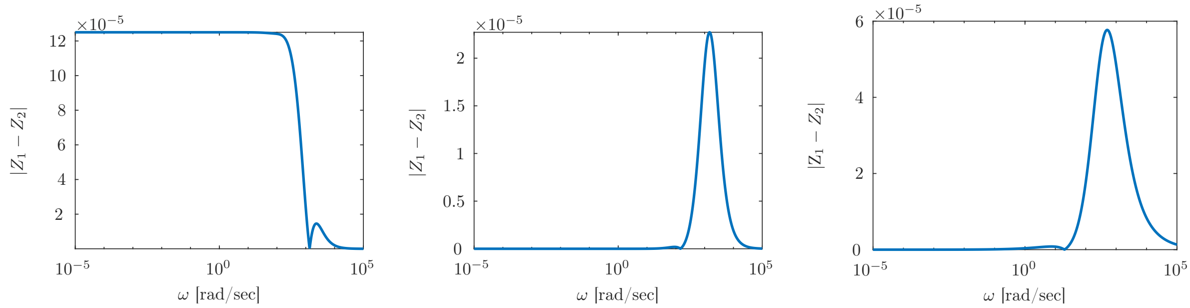
Figure 2: Schematic representation of the ECM (2)

Variation in a parameter of the transfer function (2) implicates a specific effect on the shape of the Nyquist curve. For example, a change in the resistance R_i affects the Nyquist curve in the interval $\omega < \omega_i$. The influence of the α_i parameter cannot be seen in the low-frequency region $\omega \ll \omega_i$ because at low frequencies $1 + (j\omega)^{\alpha_i} R_i Q_i \approx 1$. On the other hand, for higher frequencies, the second term vanishes, which means that some influence of α_i is visible only around ω_i. A similar situation is when Q_i undergoes a change, the implications of which can be seen in the vicinity of ω_i.

The influence of ECM parameters on the shape of the Nyquist curve can be most easily demonstrated with a simple numerical example. Let $Z(j\omega)$ be the second-order system with the following transfer function:

$$Z(j\omega) = R_0 + \frac{R_1}{1 + (j\omega)^{\alpha_1} R_1 Q_1} + \frac{R_2}{1 + (j\omega)^{\alpha_2} R_2 Q_2}, \tag{3}$$

with $R_0 = 0.1\Omega$, $R_1 = 25\text{m}\Omega$, $Q_1 = 2.5\text{F}s^{\alpha_1 - 1}$, $R_2 = 50\text{m}\Omega$, $Q_2 = 0.2\text{F}s^{\alpha_2 - 1}$, $\alpha_1 = 0.6$ and $\alpha_2 = 0.8$. Figure 3 shows the absolute change of the impedance modulus after a 5% increase of the parameters R_2, Q_2 and α_2. Taking into consideration that the resonance frequency of the 2^{nd} pole is $\omega_2 = 100$ rad/sec, the observed changes are in line with the above analysis.



(a) Variation by changing R_2 (b) Variation by changing Q_2 (c) Variation by changing α_2

Figure 3: Module of differential impedance $|\Delta Z| = |Z(\theta + \Delta\theta)| - |Z(\theta)|$ where $\theta \in (R_2, Q_2, \alpha_2)$

Remark 1

In practice, changes in the Nyquist curve are usually inferred by visual inspection and characterised in a qualitative way. Therefore, there is a clear need for a systematic way of detecting and quantifying such changes preferably in a timely and computationally efficient way.

Remark 2

Generally, published results treat only large changes in the Nyquist curves and try to associate them with the internal fault mechanism. Incipient changes, which are natural in the early stage of the fault, are typically neglected. An additional question arises whether it is possible to infer about the origin of the fault mechanism just from the change pattern.

We will show that the proposed approach provides a solution to the issue of detection of faults. Furthermore, it has also isolation capabilities, which is a valuable feature on top of the detection.

3. Impedance evaluation by means of wavelet signal processing

A sufficiently rich data set is essential for performing non-parametric statistical hypothesis testing. In the context of impedance analysis, that means acquiring sets of successive independent measurements over a short period of time that will be used for estimation of the impedance spectrum. An efficient way of acquiring such a data set is by performing time-frequency analysis in terms of CWT.

3.1. Continuous wavelet transform

The wavelet transform enables flexible selection of the desired time-frequency resolution thanks to the concepts of scaling. The main building blocks are particular waveforms with compact support called wavelets. To perform CWT, the wavelet function $\psi(t)$ is translated and scaled by using two additional parameters, u and s, respectively:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right). \tag{4}$$

The scale parameter s determines the frequency localisation of the mother wavelet. The translation parameter u defines the time location where the CWT is performed. Finally, using the wavelet function (4), CWT of a square integrable function $f(t) \in \mathbf{L}^2(\mathbb{R})$ is [25]

$$Wf(s,u) = \int_{-\infty}^{\infty} f(t)\psi^*_{u,s}(t)\,dt, \tag{5}$$

where $\psi^*_{u,s}(t)$ denotes the complex conjugate of (4).

Since the EIS analysis requires information about the amplitude and phase of the excitation and response signals, only complex wavelet functions can be considered. Out of the many, the Morlet and the Log-Normal wavelet functions offer superior time-frequency resolutions [26]. Furthermore, for both wavelets, the Fourier transforms exist in closed form. That enables computationally efficient evaluation of (5) in the frequency domain [20, 26]. Therefore, for EIS purposes, the CWT should be performed using either of those two mother wavelet functions. The subsequent analysis is performed using the Morlet wavelet with an additional parameter being the central frequency ω_0.

The Morlet wavelet is expressed as follows [26]:

$$\psi(t) = \pi^{-\frac{1}{4}} \left(e^{-j\omega_0 t} - e^{-\frac{\omega_0}{2}} \right) e^{-\frac{t^2}{2}}. \tag{6}$$

The wavelet can be visualised as a complex exponential carrier with frequency ω_0 multiplied by a Gaussian window. The Morlet wavelet's scale parameter s and the actual frequency is linked through the following relation:

$$\frac{1}{f} = \frac{4\pi s}{\omega_0 + \sqrt{2 + \omega_0^2}}. \tag{7}$$

Further details regarding the properties of the Morlet wavelet and the application of CWT for EIS analysis can be found in [20, 27].

The CWT (5) is defined for continuous signals. However, in reality, we are usually dealing with digitally sampled signals of finite length $f[k]$ with $0 \leq k \leq N_t$. The CWT results into $N \leq N_t$ complex wavelet coefficients at each scale s_i, i.e.

$$Wf(u, s_i) = [Wf(u_1, s_i), Wf(u_2, s_i), \ldots, Wf(u_N, s_i)].$$

The number of wavelet coefficients N is always smaller than the number of available samples N_t due to the so-called cone of influence. This is the region of the analysed signal around the translation point u that is within the support of the wavelet at a particular scale. When dealing with finite length signals, the border effects at $k = 0$ and $k = N_t - 1$ must be removed from the analysis, hence decreasing the number of valid wavelet coefficients [28].

In the remaining text the translation parameter u and scale parameter s are replaced with time t and frequency f, respectively. Since we are dealing with finite length signals and CWT is performed only within the window where the signal was observed, the translation parameter can be directly related to the time t of the observed signal. On the other hand, the relation between the scale parameter s and frequency f is given by (7).

3.2. Evaluation of impedance from the wavelet coefficients

The straightforward way of extracting impedance data from the signals is by calculating the ratio of the Fourier transform of the excitation electric current $i_{\text{cell}}(t)$ and the resulting voltage $u_{\text{cell}}(t)$ as

$$Z(j\omega) = \frac{U_{\text{cell}}(j\omega)}{I_{\text{cell}}(j\omega)} = \frac{\mathscr{F}\{u_{\text{cell}}(t)\}}{\mathscr{F}\{i_{\text{cell}}(t)\}}. \tag{8}$$

The resulting impedance at each frequency is just an estimate of its mean value over the observation time.

On the other hand, through CWT analysis both time and frequency information are preserved. The CWT analysis of $u_{\text{cell}}(t)$ and $i_{\text{cell}}(t)$, with the Morlet wavelet, is a set of complex wavelet coefficients:

$$\begin{aligned} Wi_{\text{cell}}(t, f) &= \Re\{Wi_{\text{cell}}(t, f)\} + j\Im\{Wi_{\text{cell}}(t, f)\}, \\ Wu_{\text{cell}}(t, f) &= \Re\{Wu_{\text{cell}}(t, f)\} + j\Im\{Wu_{\text{cell}}(t, f)\}. \end{aligned} \tag{9}$$

The impedance is then the ratio of the wavelet coefficients (9) as:

$$Z(t,f) = \frac{Wu_{\mathrm{cell}}(t,f)}{Wi_{\mathrm{cell}}(t,f)}. \tag{10}$$

Selection of the excitation signal plays the key role in proper estimation of the impedance over the required frequency region. Following the results of Boškoski et al. [20], the work below employs pseudo-random binary sequence (PRBS) for $i_{\mathrm{cell}}(t)$. The essential property of PRBS is that its power spectral density in a certain frequency band closely resembles that of the white noise. Therefore, by using just a single excitation, it is possible to calculate the impedance values over the complete frequency interval spanned by the PRBS excitation.

4. Statistical approach to the change detection

The impedance curve evaluated with the CWT is not deterministic but is considered as a realisation of a random process. Randomness arises from varying experimental conditions, outer disturbances, and random phenomena in the system as well as in sensory instrumentation. Moreover, the wavelet transform at a specific time-instant can be viewed as the estimator of the frequency characteristic on a limited time window. In turn, the modulus and phase of the impedance at the given frequency are random variables. If the same excitation were repeated many times and each time CWT were evaluated for a particular frequency we would get not one, but an ensemble of values.

Since no a priori distribution of the impedance values is assumed, we rely on the available data. In other words, the goal is to evaluate the empirical distributions of the impedance values for each frequency.

Assuming a set of impedance values is selected to represent the nominal state, the question is how can one infer about the change in their distribution. Although tests such as the median test, the Mann-Whitney test, or the parametric t test may be used, they turn out to be sensitive to the differences in means or medians of two distributions. Moreover, they may not be able to detect differences in other instances, such as differences in variances [29]. One of the advantages of the two-sided KS test is that both tests are consistent against all types of differences that may exist between the two distribution functions [29]. Therefore, we use KS thanks to its appealing properties when reasoning with empirical distributions.

The statistical hypothesis test requires a test statistic whose distribution is known under the null hypothesis. The null hypothesis is rejected if the test statistic of the current measurement lies in a highly improbable region of its distribution under the null hypothesis. The KS test compares a statistical sample with either known distribution (one-sample KS test) or another statistical sample with unknown distribution (two-sample KS test). In our case, we focus on the two-sample KS test, also known as the Smirnov test [29].

Let F_{observ} and F_{nominal} be the empirical cumulative distributions of two samples of the impedance modulus at a given frequency calculated by (10). Let the sample size of F_{observ} and F_{nominal} be equal to

$n \leq N$, where N is the number of valid wavelet coefficients. The former is generated from the sample of current data whereas the latter is generated from a reference sample. The one-sided two-sample KS test is based on the following null hypothesis $\mathcal{H}_0^+$ and its converse hypothesis $\mathcal{H}_0^-$:

$$\mathcal{H}_0^+ : \forall x \in \mathbb{R}, F_{\text{observ}}(x) \geq F_{\text{nominal}}(x), \qquad \mathcal{H}_0^- : \forall x \in \mathbb{R}, F_{\text{observ}}(x) \leq F_{\text{nominal}}(x). \tag{11}$$

Let the corresponding test statistics be D_n^+ and D_n^-, respectively, where the corresponding index n denotes the sample size:

$$D_n^+ = \sup_{-\infty<x<\infty} \left(F_{\text{observ}}(x) - F_{\text{nominal}}(x)\right), \qquad D_n^- = \sup_{-\infty<x<\infty} \left(F_{\text{nominal}}(x) - F_{\text{observ}}(x)\right). \tag{12}$$

An example of the test statistic is illustrated in Figure 4. The exact distribution of the test statistic D_n^+ (or D_n^-) is described as [29, 30]

$$\Pr(D_n^+ \leq d) = 1 - \binom{2n}{n + \lfloor dn \rfloor}\binom{2n}{n}^{-1}, \tag{13}$$

where $\lfloor dn \rfloor$ is the greatest integer less than or equal to dn. The logical conjunction of both hypotheses in (11) is equivalent to the null hypothesis of the two-sided two-sample KS test with the hypothesis $\mathcal{H}_0 : \forall x \in \mathbb{R}, F_{\text{observ}}(x) = F_{\text{nominal}}(x)$, which is used for testing the equality of two empirical distributions. The statistical hypothesis test of $\mathcal{H}_0$ is the two-sample two-sided KS test with the test statistic

$$D_n = \sup_{-\infty<x<\infty} |F_{\text{nominal}}(x) - F_{\text{observ}}(x)| \tag{14}$$

and probability distribution [30]

$$\Pr(D_n \leq d) = 1 - 2\sum_{i=1}^{\lfloor dn \rfloor} (-1)^{i+1} \binom{2n}{n + i\lfloor dn \rfloor}\binom{2n}{n}^{-1}. \tag{15}$$

The numerical computation of the distributions (13) and (15) is challenging and may induce a source of numerical error. There are various asymptotic and numerical methods addressing this issue that improve the numerical evaluation of these distributions. Proposed implementations can be found in [30–32].

4.1. Power of the KS test

The performance of the statistical test is determined by two design parameters:

1. the significance level α (or the Type I error) defined as $\Pr(\text{reject } \mathcal{H}_0 | \mathcal{H}_0 \text{ is true})$ and
2. the power of the test $\Pr(\text{reject } \mathcal{H}_0 | \mathcal{H}_1 \text{ is true})$, i.e. the probability of rejecting the null hypothesis $\mathcal{H}_0$ when the alternative hypothesis $\mathcal{H}_1$ is true.

There are guidelines for specifying the value of the significance level α. The assessment of the power of the test is generally difficult, since the alternative hypothesis $\mathcal{H}_1$ is usually unknown.

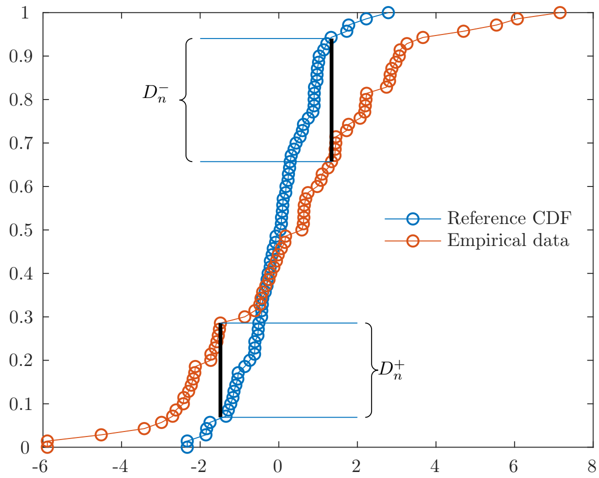


Figure 4: The comparison of two empirical distributions obtained from two samples of different population illustrates the value of the one-sided two-sample KS test statistic D_n^+ or D_n^-. Each empirical distribution is obtained from $n = 70$ realisations of two independent, normally distributed random variables with zero mean and unequal σ: the standard deviation in the blue and red colour corresponds to 0.2 and 2, respectively.

The power of the statistical test depends not only on α but also on the sample size n. The sample size must be chosen carefully; it should be just enough large to provide a sufficient amount of statistical power. If the chosen sample size is too large, the null hypothesis of the detection algorithm might be rejected due to some insignificant small effects (statistical artefacts). On the contrary, small sample size could make the detection too conservative, i.e. only larger deviations will be detected. Therefore, for the *a priori* power analysis, the goal is to determine the optimal sample size n in order to achieve sufficient power of the test for a particular significance level α.

The power of the test depends on the distribution of the alternative hypothesis $\mathcal{H}_1$, which is generally unknown. It answers the question: *"What is the significant change of the observed value for which an alarm should be triggered?"*. Typically, it is impossible to provide the analysis without knowing the actual change of the observed value. However, in the context of impedance analysis, we provide a specific statistical power analysis based on the assumption that the change in the wavelet coefficients is a constant $\mu \in \mathbb{R}$. For a predetermined value μ, where the alternative hypothesis $\mathcal{H}_1$ is known, the analysis of the statistical power is based on estimation of the probability of a missed alarm $\Pr(D_n \leq T|\mu = \mu_i), \mu_i \neq 0$, where $T(\alpha)$ is the threshold for triggering the alarm and depends on the significance level α. The probability of a missed alarm is estimated from multiple repetitions of the statistical test with a different samples from the same population. This is a frequentist approach, which requires an appropriate amount of data, the lack of which might be avoided up to an extent through "bootstrapping". In our case, we prepare two groups of M samples with each sample containing n wavelet coefficients at a specific wavelet scale. All wavelet coefficients from the first group are left unchanged while the wavelet coefficients from the second group are modified with

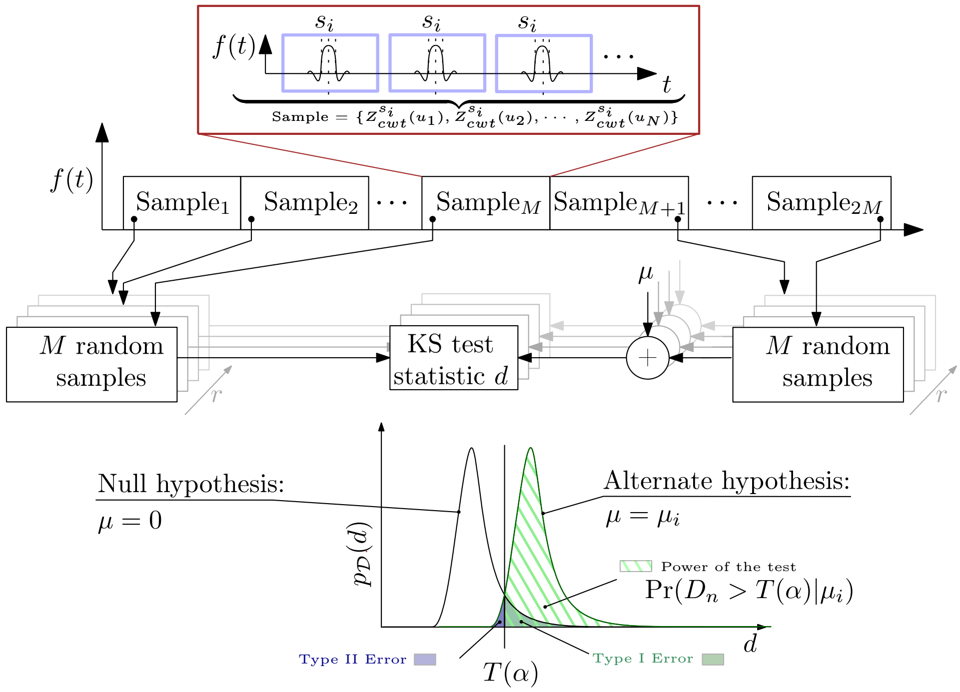


Figure 5: The scheme for empirical evaluation of the power of the KS test. The critical value $T(\alpha)$ of the test determines the boundaries of the acceptance region $[0, T(\alpha)]$ of the test where $\Pr(D_n > T(\alpha)|\mu = 0)$ equals to the significance level α.

the addition of the same value μ_i to the wavelet coefficients. Furthermore, each sample from the first group is compared element-wise to the second group through the KS test, such that we obtain M outcomes of the test. The power of the test is estimated with the number of outcomes rejecting the null hypothesis divided by the number of all outcomes M. The empirical evaluation scheme for the power of the test is shown in Figure 5. It enables power estimation for a predetermined choice of μ_i, α and n. The elements of each statistical sample consist of n randomly selected wavelet coefficients from a particular scale s_i at different translations u taking into consideration that the selected wavelets do not overlap. Such assessment of the statistical power with various choices of sample size n and significance level α provides more insight into the choice of n and α.

4.2. Application of the KS test to the impedance data

The application of the KS test on the wavelet coefficients is shown in Figure 6. Let the impedance time-frequency wavelet coefficients (10) be calculated over N_s frequencies (scales) with least N valid wavelet coefficients at each frequency f_i, $i \in [1, \ldots, N_s]$. For each frequency, a sample of size $n \ll N$ is randomly selected from coefficients from non-overlapping wavelets, resulting in a $n \times N_s$ matrix comprising randomly selected wavelet coefficients. This process is repeated for each measurement section. The two-sample KS test is then performed between the current and the reference measurements for each row (i.e. frequency f_i).

The following null hypothesis describes a two-tailed statistical test with sample size n and significance level α:

$$\mathcal{H}_0 : D_n^{f_i} = \sup_x |F_{\text{ref}}^{f_i}(x) - F^{f_i}(x)| > T(\alpha),\ T(\alpha) \simeq \sqrt{-\frac{1}{n}\ln(1-\alpha)}, \tag{16}$$

where α is the desired significance level with the corresponding approximate value of the threshold $T(\alpha)$. The critical value $T(\alpha)$ in the right-hand part of (16) is determined from the approximated inverse cumulative distribution of (15), following the procedure described by Press et al. [32]. Such an approximation is valid only for $\alpha \leq 0.3$. In essence, the null hypothesis (16) says whether the one-dimensional empirical distribution $F^{f_i}(x)$ of the wavelet coefficients at frequency f_i differs from the reference one $F_{\text{ref}}^{f_i}(x)$.

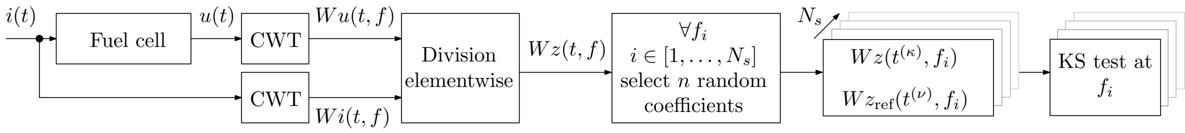


Figure 6: Evaluation scheme for performing the two-sample KS test using CWT coefficients for each frequency (scale) f_i. Time moments $t^{(\kappa)}$ and $t^{(\nu)}$ are sets of n randomly chosen time instances within the observation window and may generally differ among different frequencies.

5. Experimental results

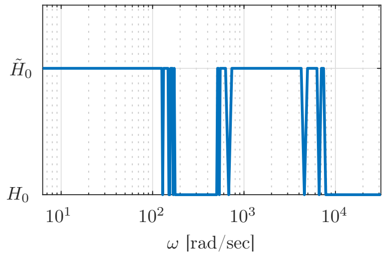
The proposed approach was first evaluated on the simulated two-pole fractional order system (3). The second case includes experiments with water management fault performed on industrial grade PEM fuel cell system. The analysis of both examples followed the procedure described in Section 4.2.

5.1. Simulated example

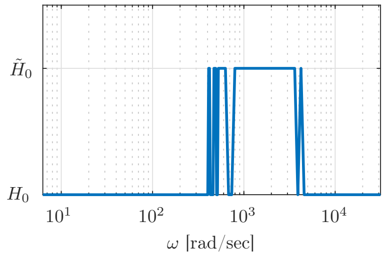
The simulation of the test system (3) was performed using the Gründwald-Letnikov scheme [33]. The excitation signal was noise-free PRBS. The noise was added to the system's response by specifying the desired signal-to-noise ratio. The reference empirical distributions of the system (3) were obtained using the initial parameter values. The changes were simulated by modifying the resistance R_2, constant phase element Q_2 and fractional order α_2 by 5%.

The results of the KS tests for each of the three changes are shown in Figure 7. The plots show the frequency regions where the null hypothesis (16) can be rejected. The results have to be analysed together with the plots shown in Figure 3. The frequency regions where the KS test rejects the null hypothesis coincide with the ones in which the change of the impedance is most pronounced.

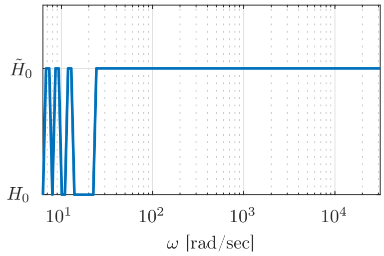
For all three cases the results of the KS test exhibit altering values before and after frequency regions where the change in the impedance becomes significant. These effects are the result of the added noise levels (in this case 1%) and the significance level for rejecting the null hypothesis, which in this numerical example



(a) Change of R_2



(b) Change of Q_2



(c) Change of α_2

Figure 7: Frequency regions where null hypothesis H_0 can be rejected based on the KS test on the numerical example for system (3)

was set to 10^{-3}. Those values were selected in order to emphasise the importance of systematic selection of the sample size n and significance level α when dealing with real-world noisy signals.

5.2. Experimental validation

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 cm^2 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. Figure 8 shows an example of the measured current $i(t)$ and voltage $u(t)$ signals, which were further used for feature extraction.

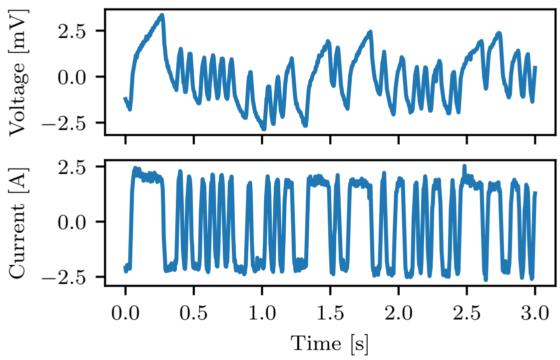


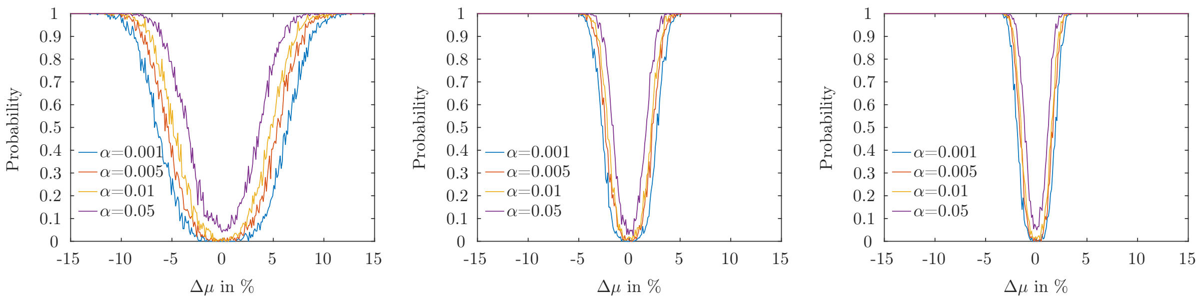
Figure 8: A sample of the current probing signal and the resulting cell voltage signal acquired during an experiment

During the experiment, the temperature of the airflow was kept constant at 50°C, stoichiometry at 2.5, and the relative humidity was controlled in order to examine the response over various conditions. On 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 80 A resulting in a stack voltage of 55 V. The PRBS amplitude was set to 4% of the I_{dc} value.

The experiment went through three phases in which the humidity of the inlet air was changed three times. In the first interval, the initial humidity of the inlet air at 9.6% spans the first eight measurement sessions. The second interval had lowered humidity of the inlet air and is between the 9^{th} and the 13^{th} measurement sessions. This is followed with the interval between the 14^{th} and the 17^{th} measurement sessions. The final interval, between the 9^{th} and the 26^{th} measurement, is the interval with increased humidity. The complete data set includes 28 measurements, each lasting for 40-seconds, and which were acquired within 120 minutes of operation. Each of the 28 measurements was analysed using the CWT approach as described in Section 4.2.

5.3. Empirical power of the test

As stated in Section 4.1, to perform the KS test one has to determine the sample size n and the significance level α. Following the aforementioned procedure, the empirical evaluation of the test's power is performed on impedance data at the frequency $f = 20$Hz. Three sample sizes are considered $n \in \{100, 500, 1000\}$ and variations of the impedance mean values μ_i of $\pm 50\%$. The results are presented in Figure 9.



(a) n=100 (b) n=500 (c) n=1000

Figure 9: Empirical power of the test for three different sample sizes n and four different significance levels α

Three observations must be made.

1. First, for smaller sample sizes $n = 100$, sufficiently high power of the test is achieved only for larger discrepancies of the mean values between the reference and the test distribution, i.e. for $\Delta\mu > 10\%$.

2. Second, by increasing the sample size n, the KS test becomes more sensitive, i.e. the same power of the test is achieved for $\Delta\mu > 5\%$.

3. Finally, the value of the parameter α also influences the power of the test, although its influence is significantly smaller than the sample size n.

Hence, the entire analysis of the experimental results was performed using sample size $n = 1000$ and $\alpha = 10^{-2}$.

5.4. Detection of water management faults

Humidity has a profound effect on the conductivity of PEM fuel cells [34–36]. As stated by Yuan et al. [37], the elevated water content directly affects kinetics thus altering the contact between the Pt catalyst and the ionomer. Conversely, a lack of water decreases the contact surface of the catalyst with the ionomer as well as its proton conductivity [38].

During the operation, there are two effects that influence the water concentration of the membrane. The first one is the so-called electro-osmotic drag under which protons H^+ travelling through the membrane drag water molecules towards the cathode side. The second one is back diffusion, under which the water that is on the cathode side diffuses towards the anode. As described by Ji and Wei [39], at low current densities, back diffusion will prevail, while at high current densities, electro-osmotic drag will prevail over back diffusion and thus the anode will tend to dry out even if the cathode is well hydrated. For the PEM fuel cell stack under test, the experiment was performed at low current density. As a result, the humidification of inlet air has significant influence on the stack performance.

The intervals where the inlet air humidity was changed are visible on the KS rejection map shown in Figure 10. The colormap shows the areas where the KS test rejects the null hypothesis (coloured red) and the regions where changes were not significant (coloured green).

The initial water production at the inlet humidity of 9.6% appears to be sufficiently high. As a result, the fuel cell condition departs from the initial condition. At the 5$^{\text{th}}$ measurement, this effect is visible only in the high frequency region above 100Hz. By decreasing the inlet humidity below 6.5%, the fuel cell stack slowly dries and its condition becomes similar to the initial one. This is clearly visible around the 15$^{\text{th}}$ measurement. The sudden significant increase in the inlet air humidity to 80% after the 18$^{\text{th}}$ measurement causes a fast change in the outcome of the KS test over the whole observed frequency interval.

Besides the detection of changes in inlet air humidity, the application of the KS test on the impedance data offers two additional benefits:

1. The KS is shown to be sensitive even to minute variations of the impedance characteristics.

2. The use of one-sided KS tests allows for some restricted fault isolation.

Sensitivity of the KS test results. The sensitivity of the KS test can be demonstrated by analysing EIS curves from three measurements. The first one is the comparison between the 1$^{\text{st}}$ measurement, which is the reference, and the 3$^{\text{rd}}$ measurement. From the KS rejection map shown in Figure 10, one can see that for the 3$^{\text{rd}}$ measurement, the KS test rejects the null hypothesis for the frequency bands ∼1Hz, ∼9Hz, 40-60Hz and 200-400Hz. The EIS curves for these two measurements are shown in Figure 11(a). The first two intervals

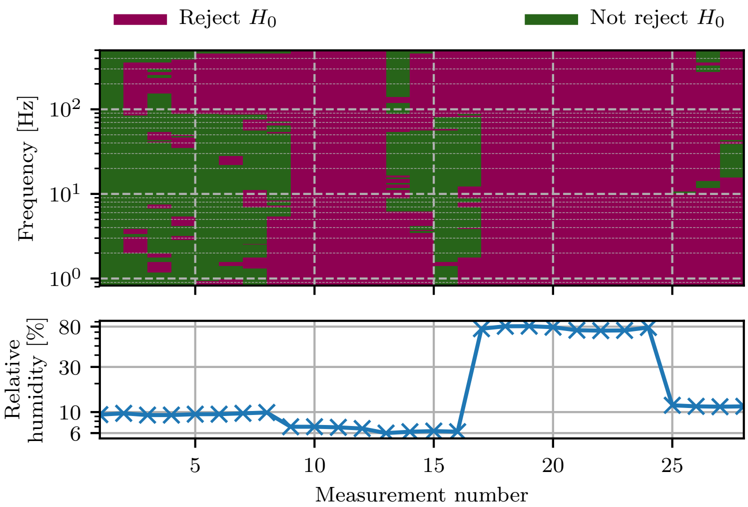


Figure 10: Two-sided KS test rejection map

are clearly visible as a change in the EIS curve. However, for the last two intervals, changes in the EIS curve are not so clear. Therefore, without the extra information provided by the KS test rejection map, these changes might have been left unnoticed.

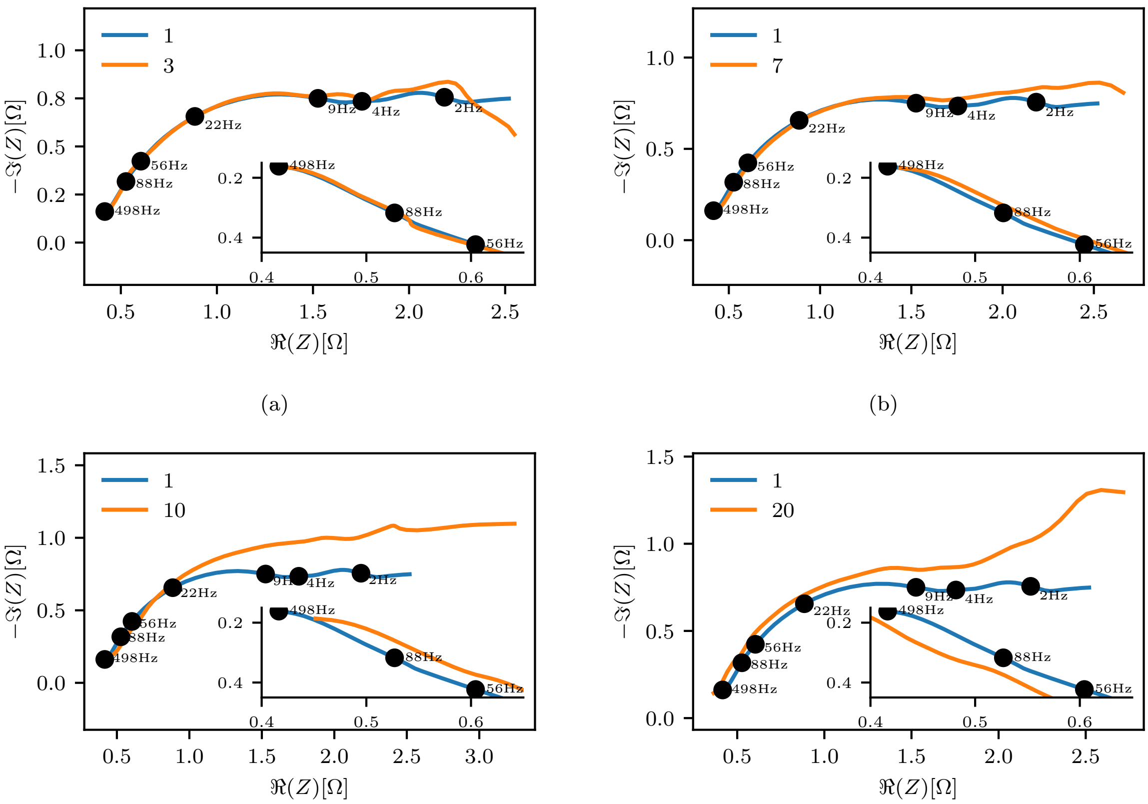
Similar observations can be made for the second case, i.e. between the 1$^{\text{st}}$ and the 7$^{\text{th}}$ measurement. The comparison of the EIS curves is shown in Figure 11(b). Here one can observe the opposite effect. The only frequency interval where the impedance can be regarded as unchanged is from 10-50Hz. In this interval, the changes in the EIS curves are not significant in order to reject the null hypothesis despite the small visible deviations between the two EIS curves. Outside this interval, the changes are significant and the null hypothesis can be rejected.

Finally, for the 10$^{\text{th}}$ and the 20$^{\text{th}}$ measurements, the KS map from Figure 10 shows significant changes on all frequency bands. This also can be confirmed by simple comparison of the EIS curves in Figure 11(c) and (d).

This analysis shows that the KS test offers a systematic way of detecting the frequency regions where the impedance characteristic is significantly changed. Such an approach completely overpowers the visual inspection and requires no prior expert knowledge of impedance analysis.

One-side (small and large) KS tests. The above analysis was performed by taking into consideration two-sided KS test. The one-sided KS test provides additional insight, whether the change in the impedance characteristics is either due to an increase or a decrease of particular impedance components. The rejection maps of both one-sided tests are shown in Figure 12(a) and (b).

One can observe two main effects. For the interval with lower humidity, the impedance values tend to



(a) (b)

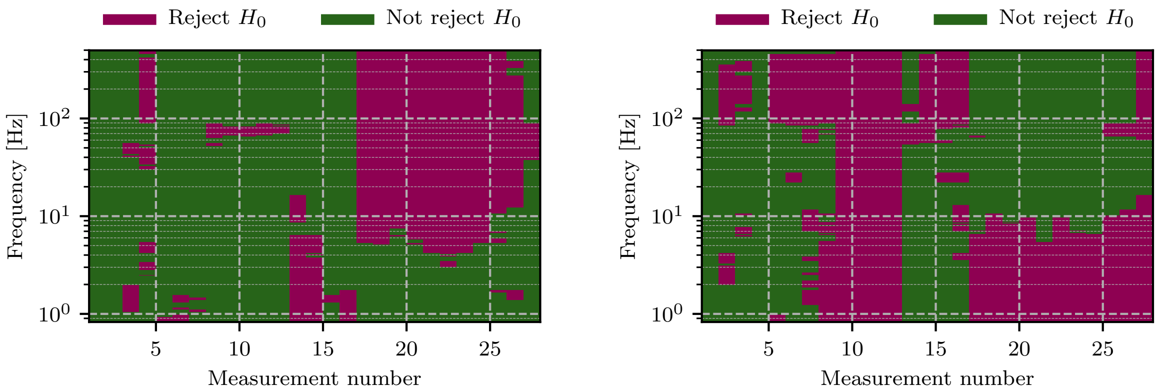
(c) (d)

Figure 11: Comparison of EIS curves for selected characteristic measurement

be higher throughout the frequency interval. On the other hand, for the interval with increased humidity, low frequency impedance components $f < 10$Hz exhibit higher values, whereas for $f > 10$Hz, they exhibit lower values. The information from the KS rejection maps can be confirmed by analysing the changes in the impedance amplitudes for the 10^{th} and the 20^{th} measurement, which are shown in Figure 13.

6. Conclusion

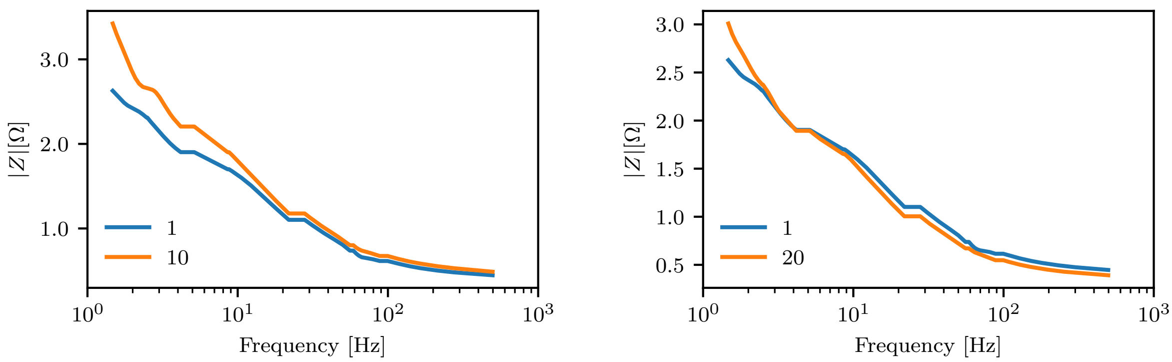
In this paper, we introduce a systematic procedure for monitoring the internal health condition of fuel cells by revealing changes on the Nyquist curve. The main idea is to apply the KS hypothesis to detect changes in the empirical distributions of the Nyquist modulus evaluated from repetitive system probing and processing the signals with complex wavelet transform. The frequency bands where the KS test rejects the null hypothesis indicate where the impedance curves are changed significantly. The KS statistic is useful for quantifying the severity of the change. Furthermore, the one-sided tests are shown to be able to perform even fault isolation in some cases.



(a) Small side KS test

(b) Large side KS test

Figure 12: One sided KS test rejection map



(a)

(b)

Figure 13: Comparison of the impedance amplitude

A nice property of the proposed KS hypothesis testing for EIS analysis is the easy tuning of the design parameters, i.e. the significance level α and the sample size n. One has just to define the desired power of the test and probability of false alarm (Type I error). For the evaluated PEM fuel cell, the significance level was set at $\alpha = 10^{-2}$ and sample size $n = 1000$. With such parameters, the results indicate that the KS test is capable of detecting changes in the impedance values bigger than 1% with power of the test more than 0.6.

Acknowledgements

The authors acknowledge the projects L2-7663, PR-2561 and research core funding No. P2-0001 were financially supported by the Slovenian Research Agency. The authors acknowledge financial support through the project MEMPHYS (grant agreement No 735533) within the framework of the Fuel Cells and Hydrogen 2

Joint Undertaking under the European Union's Horizon 2020 research and innovation programme, Hydrogen Europe and Hydrogen Europe research. The authors are grateful to the anonymous reviewers for their helpful comments on the initial version of the manuscript.

[1] P. Polverino, M. Sorrentino, C. Pianese, A model-based diagnostic technique to enhance faults isolability in Solid Oxide Fuel Cell systems, Applied Energy 204 (2017) 1198 – 1214, ISSN 0306-2619.

[2] S. M. R. Niya, R. K. Phillips, M. Hoorfar, Process modeling of the impedance characteristics of proton exchange membrane fuel cells, Electrochimica Acta 191 (2016) 594 – 605, ISSN 0013-4686.

[3] B. Liu, H. Muroyama, T. Matsui, K. Tomida, T. Kabata, K. Eguchi, Analysis of Impedance Spectra for Segmented-in-Series Tubular Solid Oxide Fuel Cells, Journal of The Electrochemical Society 157 (12) (2010) B1858–B1864.

[4] A. D. Franklin, H. J. De Bruin, The fourier analysis of impedance spectra for electroded solid electrolytes, physica status solidi (a) 75 (2) (1983) 647–656, doi:10.1002/pssa.2210750240.

[5] M. Heinzmann, A. Weber, E. Ivers-Tiffée, Advanced impedance study of polymer electrolyte membrane single cells by means of distribution of relaxation times, Journal of Power Sources 402 (2018) 24 – 33, ISSN 0378-7753.

[6] J. Deseure, Coupling RTD and EIS modelling to characterize operating non-uniformities on PEM cathodes, Journal of Power Sources 178 (1) (2008) 323 – 333, ISSN 0378-7753.

[7] D. Bezmalinovic, B. Simic, F. Barbir, Characterization of PEM fuel cell degradation by polarization change curves, Journal of Power Sources 294 (2015) 82 – 87, ISSN 0378-7753.

[8] S. M. R. Niya, M. Hoorfar, Study of proton exchange membrane fuel cells using electrochemical impedance spectroscopy technique – A review, Journal of Power Sources 240 (2013) 281 – 293, ISSN 0378-7753.

[9] G. A. Futter, P. Gazdzicki, K. A. Friedrich, A. Latz, T. Jahnke, Physical modeling of polymer-electrolyte membrane fuel cells: Understanding water management and impedance spectra, Journal of Power Sources 391 (2018) 148 – 161, ISSN 0378-7753.

[10] T. Jahnke, G. Futter, A. Latz, T. Malkow, G. Papakonstantinou, G. Tsotridis, P. Schott, M. Gérard, M. Quinaud, M. Quiroga, A. Franco, K. Malek, F. Calle-Vallejo, R. F. de Morais, T. Kerber, P. Sautet, D. Loffreda, S. Strahl, M. Serra, P. Polverino, C. Pianese, M. Mayur, W. Bessler, C. Kompis, Performance and degradation of Proton Exchange Membrane Fuel Cells: State of the art in modeling from atomistic to system scale, Journal of Power Sources 304 (2016) 207 – 233, ISSN 0378-7753.

[11] G. Zhang, K. Jiao, Multi-phase models for water and thermal management of proton exchange membrane fuel cell: A review, Journal of Power Sources 391 (2018) 120 – 133, ISSN 0378-7753.

[12] C. Bao, W. G. Bessler, Two-dimensional modeling of a polymer electrolyte membrane fuel cell with long flow channel. Part II. Physics-based electrochemical impedance analysis, Journal of Power Sources 278 (2015) 675 – 682, ISSN 0378-7753.

[13] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, Đ. Juričić, State of health estimation and remaining useful life prediction of solid oxide fuel cell stack, Energy Conversion and Management 148 (2017) 993 – 1002, ISSN 0196-8904.

[14] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess, B. Wu, Intelligent Fault Diagnosis and Prognosis for Engineering Systems, Wiley, 2006.

[15] A. Rakar, D. Juričić, P. Ballé, Transferable belief model in fault diagnosis, Engineering Applications of Artifcial Intelligence 12 (5) (1999) 555–567.

[16] P. Boškoski, A. Debenjak, B. M. Boshkoska, Rayleigh copula for describing impedance data–with application to condition monitoring of proton exchange membrane fuel cells, European Journal of Operational Research 266 (1) (2018) 269–277.

[17] P. Boškoski, A. Debenjak, Optimal selection of Proton Exchange Membrane fuel cell condition monitoring thresholds, Journal of Power Sources 268 (2014) 692–699.

[18] M. Unser, T. Blu, Fractional Splines and Wavelets, SIAM Review 42 (1) (2000) 43–67.

[19] P. Costamagna, A. D. Giorgi, G. Moser, S. B. Serpico, A. Trucco, Data-driven techniques for fault diagnosis in power

generation plants based on solid oxide fuel cells, Energy Conversion and Management 180 (2019) 281 – 291, ISSN 0196-8904.

[20] P. Boškoski, A. Debenjak, B. M. Boshkoska, Fast Electrochemical Impedance Spectroscopy, Springer, 2017.

[21] P. Zoltowski, On the electrical capacitance of interfaces exhibiting constant phase element behaviour, Journal of Electroanalytical Chemistry 443 (1) (1998) 149 – 154, ISSN 1572-6657.

[22] M. Sluyters-Rehbach, Impedances of electrochemical systems: Terminology, nomenclature and representation part I: Cells with metal electrodes and liquid solutions, Pure and Applied Chemistry 66 (9) (1994) 1831–1891.

[23] A. Lasia, Electrochemical Impedance Spectroscopy and its Applications, Springer-Verlag, New York, doi:10.1007/978-1-4614-8933-7, 2014.

[24] H. Schichlein, A. Müller, M. Voigts, A. Krügel, E. Ivers-Tiffée, Deconvolution of electrochemical impedance spectra for the identification of electrode reaction mechanisms in solid oxide fuel cells, Journal of Applied Electrochemistry 32 (8) (2002) 875–882, ISSN 1572-8838.

[25] S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way, Elsevier Academic Press, 3 edn., ISBN 9780080922027, 2008.

[26] D. Iatsenko, Nonlinear Mode Decomposition, Springer Theses, Springer International Publishing, 2015.

[27] D. Iatsenko, P. V. McClintock, A. Stefanovska, Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms, Digital Signal Processing 42 (2015) 1 – 26, ISSN 1051-2004.

[28] C. Torrence, G. P. Compo, A Practical Guide to Wavelet Analysis, Bulletin of the American Meteorological Society 79 (1998) 61–78.

[29] W. J. Conover, Practical nonparametric statistics, Wiley New York, 3rd ed., 1999.

[30] J. W. Pratt, J. D. Gibbons, Kolmogorov-smirnov two-sample tests, in: Concepts of Nonparametric Theory, Springer, 318–344, 1981.

[31] R. Simard, P. L'Ecuyer, et al., Computing the two-sided Kolmogorov-Smirnov distribution, Journal of Statistical Software 39 (11) (2011) 1–18.

[32] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes 3rd edition: The art of scientific computing, Cambridge university press, 2007.

[33] I. Podlubny, Fractional Differential Equations, vol. 198 of *Mathematics in Science and Engineering*, Elsevier, 1999.

[34] G. Hinds, M. Stevens, J. Wilkinson, M. de Podesta, S. Bell, Novel in situ measurements of relative humidity in a polymer electrolyte membrane fuel cell, Journal of Power Sources 186 (1) (2009) 52 – 57, ISSN 0378-7753.

[35] S. Jeon, J. Lee, G. M. Rios, H.-J. Kim, S.-Y. Lee, E. Cho, T.-H. Lim, J. H. Jang, Effect of ionomer content and relative humidity on polymer electrolyte membrane fuel cell (PEMFC) performance of membrane-electrode assemblies (MEAs) prepared by decal transfer method, International Journal of Hydrogen Energy 35 (18) (2010) 9678 – 9686, ISSN 0360-3199, hE (Hydrogen Systems and Materials For Sustainability).

[36] Q. Yan, H. Toghiani, H. Causey, Steady state and dynamic performance of proton exchange membrane fuel cells (PEMFCs) under various operating conditions and load changes, Journal of Power Sources 161 (1) (2006) 492 – 502, ISSN 0378-7753.

[37] X.-Z. Yuan, C. Sons, H. Wang, J. Zhang, Electrochemical Impedance Spectroscopy in PEM Fuel Cells, Fundamentals and Applications, Springer, London, 2010.

[38] J. Song, S. Cha, W. Lee, Optimal composition of polymer electrolyte fuel cell electrodes determined by the AC impedance method, Journal of Power Sources 94 (1) (2001) 78 – 84, ISSN 0378-7753.

[39] M. Ji, Z. Wei, A Review of Water Management in Polymer Electrolyte Membrane Fuel Cells, Energies 2 (4) (2009) 1057–1106.