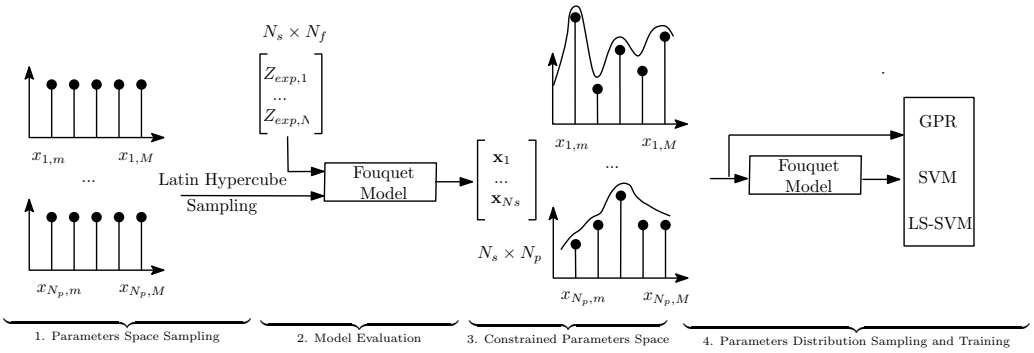
A fast fuel cell parametric identification approach based on machine learning inverse models

Graphical Abstract

A fast Fuel Cell Parametric Identification approach based on Machine Learning Inverse Models

Antonio Guarino, Riccardo Trinchero, Flavio Canavero, Giovanni Spagnuolo



Highlights

A fast Fuel Cell Parametric Identification approach based on Machine Learning Inverse Models

Antonio Guarino, Riccardo Trinchero, Flavio Canavero, Giovanni Spagnuolo

- Parameter identification of equivalent circuit model performed through Inverse Model;
- Parameter search space is first constrained in feasibility regions;
- Inverse Models built through GPR, SVM and LS-SVM machine learning regression models;
- The identification results are shown for both synthetic and experimental data;
- Performance and Resource Usage of the embedded implementation are explored.

A fast Fuel Cell Parametric Identification approach based on Machine Learning Inverse Models

Antonio Guarino[a], Riccardo Trinchero[b], Flavio Canavero[b], Giovanni Spagnuolo[a]

[a] *DIEM-University of Salerno, Via Giovanni Paolo II, 132, Fisciano (SA), 84084, Italy*
[b] *DET-Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino (TO), 10129, Italy*

Abstract

In this paper a computationally efficient optimization approach to the parametric identification of a fuel cell equivalent circuit model is presented. It is based on the inverse model and on machine learning regressions. During the training phase, the inverse model is built numerically by means of advanced regression approaches, i.e., the support vector machine regression, the least squares support vector machine regression and the Gaussian process regression. The training set is synthetically generated to the aim of exploring the parameter space and to characterize different stack operating conditions, including normal and faulty ones. The accuracy of the considered approaches is investigated by employing a test set including many experimental data, consisting of impedance spectra measured through the electrochemical impedance spectroscopy and referring to very different stack operating conditions. The results show that all the considered machine learning methods allow to identify the parameters of the fuel cell model with a low computational burden, so that they fit with the hardware resources of low cost embedded processors. This feature allows to envisage that the proposed approaches are good candidates for a model-based on-line diagnosis of fuel cell stacks.

Keywords: Parametric identification of Equivalent Circuit Model, Inverse Model, Machine Learning Regression

1. Introduction

Fuel Cells (FCs) and hydrogen-related technologies are a key point of the European Green Deal and of the Recovery Plan. In order to improve the

Life Cycle Assessment (LCA) of FCs' stacks, a significant lifetime has to be reached. Monitoring and Diagnostic Tools (MDTs) are helpful in evidencing a malfunctioning and the occurrence of operating conditions that stress the stack and reduce its Remaining Useful Life (RUL). Although many diagnostic approaches and methodologies are proposed in literature, very often they employ laboratory measurement systems and/or algorithms requiring significant hardware resources. Consequently, their on-line and on-site application becomes very challenging. At the same time, some efforts are on course to develop MDTs employing hardware systems and requiring computational efforts that are compatible with FC systems costs, diagnostic accuracy and time response. For instance, some H2020 funded projects, e.g. [1], have been aiming at pursuing the objective of having an Electrochemical Impedance Spectroscopy (EIS) -based MDT using ordinary power electronics and embedded processors to perform frequency domain measurements and detect a number of faults through on-line diagnostic algorithms. The diagnostic methods usually work in two spaces: one is the frequency domain, i.e., the stack impedance spectrum in the Bode or in the Nyquist planes, and the other is the multidimensional space of the parameters of a suitable non linear FC stack Equivalent Circuit Model (ECM). The latter requires a robust approach that takes the impedance spectrum as an input and gives at the output the set of ECM parameters. This task is quite challenging because of the number of parameters, of the wide search space and of the lack of a good guess solution. A further issue is the high non linearity between the function, which is usually considered as the objective of the optimization problem, e.g., the Root Mean Square Error (RMSE) between the ECM impedance and the experimental data, and the parameters thereof. Some approaches are proposed in literature, e.g. in [2], but they are based on specific assumptions about the signal-to-noise level affecting the data; consequently, they do not always ensure the convergence towards the absolute minimum RMSE value or they have a significant computational burden.

Machine learning (ML) has been using more and more in energy problems. An increasing number of ML-based approaches to monitoring, diagnosis and RUL prediction of batteries and FCs can be found in the recent literature. In [3, 4] multiple ML-based approaches have been employed to RUL estimation of a set of batteries by using charging, discharging, internal resistance and temperature data. The effectiveness of such approaches that operate on the battery early usage data is demonstrated. In [5], the battery state of health is estimated by employing mechanical and electrical features through

an Autoencoder and a Gaussian Process Regression (GPR) based estimation model. In [6, 7] an artificial neural network is used to estimate the remaining battery capacity. In [8], a number of variables, e.g., stack current and voltage, air supplier control signal, temperatures, air flow rates, is monitored and their values are compared with those ones computed through a suitable model. The patterns of the resulting differences are used to train supervised ML methods that are then able to classify different faults. In order to ensure a simultaneous faults detection, instead then addressing independent faults only, in [9] a deep learning network called stacked sparse autoencoder is proposed. The stack voltage is also the subject of the study presented in [10]: some relevant features are extracted and allow ML approaches to identify the fault.

FC degradation prediction in presence of a variable load is performed in [11] and [12] through a method based on a Genetic Algorithm (GA) and an Extreme Learning Machine (ELM). The long term dynamic behavior of the stack voltage is exploited in [13] to predict the performance degradation by using a Grid Long Short-Term Memory (G-LSTM) Recurrent Neural Network (RNN). A Support Vector Machine (SVM) classifier is adopted in [14] for detecting FC faults by using a 3D model. The validation of the multi-scale hybrid degradation index formulated in [15] and of the the approach aimed at predicting the FC RUL presented in [16] is done through a ML tool operating on experimental data. Moreover, a ML algorithm has been used in [17] to find the best trade off between the EIS measurement time and the fault detection accuracy through the impedance spectrum. The diagnosis is conducted in the frequency space only, without any parametric identification process. To this regard, the adoption of ML techniques for system parameters identification is quite rare in literature. Few examples refer to dynamical systems identification: e.g., in [18] the systems considered are described by differential-algebraic equations and the sum of the squared deviations of the values of the state vector's coordinates from their exact counterparts obtained through measurements at different time instants is considered as the objective function whose value is minimized through the adaptive random search method. In [19], transfer learning methods are used to address a challenge in state-space linear parameter-varying model identification and learning. Kernelized ML is adopted when the distributions of the training and testing sets are different.

This paper shows that the parametric identification of a Polymer Electrolyte Membrane (PEM) FC ECM can be effectively performed by using

the *inverse model* [20]. A mathematical model, i.e., a multivariate function, is achieved as first. Then, it exploits a set of experimental EIS spectra to identify the parameters of the widely recognised ECM presented in [21]. Since the parameters identification is carried out via a straightforward evaluation of the inverse model over the experimental values of the EIS spectra, the proposed approach ensures a good compromise between computational burden and accuracy, thus a performance that is better than the one the stochastic identification techniques presented in literature are able to guarantee [22][23][24].

The construction of the inverse model, which is the core of the proposed method, is not straightforward. Indeed, the inverse map is usually extremely non-linear and ill-posed, this meaning that a given EIS spectrum might be generated by more than one set of the ECM parameters. In order to mitigate the above issue, in this paper three advanced ML regression methods are considered. Specifically, the SVM regression [25], the Least Squares SVM regression (LS-SVM) [26] and the GPR [27] are adopted to construct the inverse model from synthetic data. The above techniques have been selected for two main reasons. First of all, they are non-parametric techniques in which the number of unknowns to be estimated during the model training is independent from the dimensionality of the input space, i.e., from the number of frequency points in the EIS spectrum. This allows to minimize the number of training samples required by the model training and thus the model complexity. Moreover, the regularizer used by the SVM and LS-SVM regressions, as well as the noise term in the GPR, can be used to mitigate the ill-posedness of the inverse model [28, 29]. The proposed approaches have been also implemented in a low cost embedded system to demonstrate their on-site and on-line potential; their performance is validated through a large test set consisting of experimental EIS spectra.

The experimental test set has been obtained during the HEALTHCODE H2020 project [1] on a PEM stack having 46 cells and an active area of 100 cm^2. The stack was tested at 0.4 A/cm^2 under different hydrogen and oxygen utilization and for different humidity levels. The considered dataset refers to cells number 1, 2, 23, 45, and 46 with measurements performed in the range from 50 mHz to 2 kHz for a total of 157 spectra. Further details on the testing procedure and on the study of the impact of water management and reactant utilization through EIS are disclosed in [30][31].

The paper is organized as follows. Section II is aimed at describing the proposed ML-based identification approach and Section III gives the details

of the application of the ML methods adopted. Section IV shows the procedure used to produce the synthetic data that have been used for the learning phase. Sections V and VI provide the analysis of the results based on synthetic and experimental data, respectively. Section VII shows the results of an approach validation performed through an embedded platform that is a candidate for the on-line application of the proposed approaches. The computational burden of the proposed approach in comparison with that one of an accurate stochastic approach proposed in literature is also provided. Conclusions end the paper.

2. Gray-Box System Identification via Forward Models

System identification aims at building a mathematical model on the basis of measured data [32]. Given the experimental FC impedance values $\mathbf{Z}_{exp} = [Z_{exp}(f_1), \ldots, Z_{exp}(f_{N_f})]^T \in \mathbb{C}^{N_f}$ in a given frequency range, the goal is to identify the optimal configuration $\mathbf{x}_*$ of the $N_p = 6$ free parameters $\mathbf{x} = [x_1, \ldots, x_{N_p}]^T$ of the Fouquet ECM $\bar{Z}_F(f; \mathbf{x})$ (see the Appendix and [21] for additional details), such that:

$$Z_{exp}(f) \approx \bar{Z}_F(f; \mathbf{x}_*), \tag{1}$$

Literature results allow to assume that the Fouquet model is able to fit, with a good accuracy, the experimental spectra.

The chosen ECM has a closed-form and it provides a gray-box model, i.e., a model built from a-priori knownledge, for the system identification. The Fouquet model is a forward map, or forward model, that for any configuration $\mathbf{x}$ of its parameters is able to provide the corresponding impedance value $\bar{Z}_F(f; \mathbf{x})$ at the desired frequency value f.

The identification problem (1) can be also formulated by means of an objective function whose value has to be optimized through a suitable algorithm. For instance, the configuration $\mathbf{x}_*$ can be estimated as the one that minimizes the average of the magnitudes of the squared error computed between the experimental impedance values collected in $\mathbf{Z}_{exp}$ and the corresponding ones predicted by the Fouquet model $\bar{Z}_F$ in a given frequency range:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}} \frac{1}{N_f} \sum_{i=1}^{N_f} |Z_{exp}(f_i) - \bar{Z}_F(f_i; \mathbf{x})|^2. \tag{2}$$

The properties of such an optimization problem, e.g. its convexity, depend on the structure of the forward model $\bar{Z}_F$. As an example, the problem at hand turns out to be convex when the model $\bar{Z}_F$ is so simple that it can be written as a linear combination of its input parameters $\mathbf{x}$. However, such minimization problem can be usually classified as a non-convex optimization with several local minima [33], which might require advanced optimization algorithms, as in [22][23].

Alternatively, provided that the candidate model $\bar{Z}_F(f;\mathbf{x})$ can be inexpensively evaluated for any set of the input parameters $\mathbf{x}$, the above optimization problem can be tackled via a grid search approach. It is a brute force optimization scheme, in which the N_p parameters can be interpreted as uniform distributed random variables in a given range (i.e., $x_i \sim \mathcal{U}(x_{i,min}, x_{i,max})$), then the model $\bar{Z}_F(f;\mathbf{x})$ is evaluated for a large number of realizations of the input random parameters drawn according to their probability distribution functions (PDFs). Thus, the optimal parameters set $\mathbf{x}_*$ can be estimated among all the considered realizations as the one that minimizes the cost function of the optimization problem (2). This approach is straightforward and does not require to calculate the gradient function of the cost function, but it turns out to be quite cumbersome and computational expensive. In fact, the accuracy and the reliability of optimal solution $\mathbf{x}_*$ predicted by such brute force approach heavily depends on the number of calls to the model $\bar{Z}_F$, and thus on the number of realizations of the random variables $\mathbf{x}$ considered during the optimization process. This limits the online use of such an approach through embedded systems.

3. Inverse Model for System Identification

The inverse model approach presented in this paper is an effective candidate for the FC ECM identification [20].

The inverse model denoted as $\mathcal{M}^{-1}$ goes in the opposite direction with respect to forward model $\bar{Z}_F(f;\mathbf{x})$ used in the conventional optimization approach (2); it is:

$$\mathbf{x}_* = \mathcal{M}^{-1}(\mathbf{Z}_{exp}), \tag{3}$$

where, the vector $\mathbf{x}_* \in \mathbb{R}_p^N$ collects the identified system parameters, $\mathbf{Z}$ is a vector collecting the experimental impedance values evaluated at the frequency $\mathbf{f} = [f_1, \ldots, f_{N_f}]^T$ and the inverse map $\mathcal{M}^{-1} : \mathbb{C}^{N_f} \to \mathbb{R}^d$.

Specifically, starting from the experimental impedance values $\mathbf{Z}_{exp}$, the inverse map $\mathcal{M}^{-1}$ is able to directly provide, without requiring to solve any optimization problem or to use an iterative algorithm, the optimal configuration of the ECM parameters $\mathbf{x}_*$. Indeed, by means of the inverse model, the identification task becomes as simple as a trivial function evaluation.

Unfortunately, for realistic cases in which the inverse map cannot be computed explicitly from the forward one, the construction of the inverse model is rather challenging. First of all, even if a mathematical relationship for the forward model is known, the inverse model might not be available in a closed-form. In such cases, the model must be constructed numerically with the help of regression or interpolation techniques, starting from a set of experimental and/or synthetic training data provided by the forward model. Moreover, inverse models are often ill-posed, since they can lead to a one-to-many map (i.e., the inverse model turns out to be a non-injective function), in the sense that a given set of value of the model output might be generated by more than one combinations of the system parameters.

Advanced ML regression techniques are promising candidates for building an accurate approximation of the inverse map in a high dimensional space. In this paper, the effectiveness of three ML regressions, i.e., SVM regression [25], LS-SVM regression [26] and the GPR [27], is investigated.

3.1. SVM and LS-SVM regression

Let us consider the problem of constructing the inverse map $\mathcal{M}^{-1}$, starting from a set of N_s training samples $\mathcal{D} = \{(\mathbf{x}_l, \mathbf{y}_l)\}_{l=1}^{N_s}$, in which the vector $\mathbf{x}_l \in \mathbb{R}^{N_p}$ collects the training realizations of the system parameters and $\mathbf{y}_l = [\mathrm{Re}\left\{Z_{1,l}\right\}, \ldots, \mathrm{Re}\left\{Z_{N_f,l}\right\}, \mathrm{Im}\left\{Z_{1,l}\right\}, \ldots, \mathrm{Im}\left\{Z_{N_f,l}\right\}]^T \in \mathbb{R}^{2N_f}$ is a real value vector collecting the corresponding real and imaginary parts of the FC impedance values, the latter being synthetic or experimental data. The primal space formulation of the inverse model $\tilde{\mathcal{M}}_i^{-1}$ for i-th component of the parameter vector $\mathbf{x}$ built via the SVM and LS-SVM regression, writes [25, 26]:

$$x_i \approx \mathcal{M}_i^{-1}(\mathbf{y}) = \langle \mathbf{w}_i, \boldsymbol{\phi}(\mathbf{y}) \rangle + b_i \tag{4}$$

where $\boldsymbol{\phi} : \mathbb{R}^{N_p} \to \mathbb{R}^D$ is a vector collecting D-basis functions [1], $\mathbf{w}_i \in \mathbb{R}^D$ is a vector collecting the regression unknowns and b_i is the bias term. The above formulation can be repeated to model all the components of the vector $\mathbf{x}$.

[1] For the sake of simplicity, it has been assumed that all the components of parameter vector x_i share the same basis functions.

For the LS-SVM and SVM regressions, the unknowns (i.e., $\mathbf{w}_i$ and b_i) are estimated by solving the following optimization problem:

$$\min_{\mathbf{w}_i, b_i} \frac{1}{2}\|\mathbf{w}_i\|_{L_2}^2 + \gamma_i \sum_{l=1}^{N_s} \ell(x_{i,l}, \langle \mathbf{w}_i, \boldsymbol{\phi}(\mathbf{y}_l)\rangle + b_i) \tag{5}$$

where $\ell(\cdot)$ is a cost function proving the error of the model on the training samples. This means that the coefficients $\mathbf{w}_i$ and the bias term b_i are estimated by minimizing the average of the cost function evaluated between the training data and the model predictions penalized by a L2 regularizer (i.e., $\|\mathbf{w}_i\|^2$). The latter term adds a penalty to the model accuracy on the training samples, thus mitigating the detrimental effects of the ill-posed problem and also reducing overfitting [34]. The effect of the regularizer in the optimization problem in (5) is tuned by the hyperparameter γ_i, usually estimated via cross-validation (CV) during the model training. A squared loss function is used as cost function for the LS-SVM regression[2], whereas the SVM regression uses the ε-insensitive loss function. Interested readers can refer to [25] for additional details.

It is important to notice that the primal space formulation of the LS-SVM and SVM regression provides a parametric model, in which the number of unknowns to be estimated (i.e., the cardinality of the vector $|\mathbf{w}| = D$) is equal to the number of basis functions considered in the liner expansion in (4). This means that the obtained model suffers the curse of dimensionality (i.e., a reduction of the model efficiency when either the model complexity and the number of parameters increases).

The dual formulation available for the LS-SVM and SVM regression allows mitigating the above issue and writes:

$$x_i = \sum_{l=1}^{N_s} \alpha_{i,l} k(\mathbf{y}, \mathbf{y}_l) + b_i, \tag{6}$$

where the coefficients $\alpha_{i,l}$ and the bias term b_i are estimated by the regression algorithm during the training process and $k(\cdot,\cdot) : \mathbb{R}^{2N_f \times 2N_f} \rightarrow \mathbb{R}$ is the kernel function defined as the inner product of between the basis function vectors, i.e., $k(\mathbf{y}, \mathbf{y}_l) = \langle \boldsymbol{\phi}(\mathbf{y}), \boldsymbol{\phi}(\mathbf{y}_l)\rangle$. Thanks to the kernel function $k(\cdot,\cdot)$, the dual

[2]For mathematical convenience, in the LS-SVM regression the squared loss function is multiplied by a constant term 1/2.

space formulation in (6) does not require an explicit definition of the basis functions $\boldsymbol{\phi}(\mathbf{y})$. This is the so-called "kernel trick" [25]. Several kernel functions with different mathematical properties and features are available [35]. The most common ones are:

- linear kernel: $k(\mathbf{y}, \mathbf{y}_l) = \mathbf{y}^T \mathbf{y}_l$;
- polynomial kernel of order q: $k(\mathbf{y}, \mathbf{y}_l) = (1 + \mathbf{y}^T \mathbf{y}_l)^q$;
- Gaussian radial basis function (RBF) kernel: $k(\mathbf{y}, \mathbf{y}_l) = \exp\left(-\|\mathbf{y} - \mathbf{y}_l\|^2 / 2\sigma^2\right)$, where σ is the kernel hyperparameter.

Hereafter, in this paper the Radial Basis Function (RBF) kernel will be adopted, since it has shown an excellent performance in high nonlinear regression problems [25, 28].

In the above dual formulation of the SVM and LS-SVM regression, the number of regression unknown $\alpha_{i,l}$ turns out to be independent from the dimensionality of the input space (i.e., the cardinality of $\mathbf{y}$). This is extremely useful to overcome the curse-of-dimensionality issue. Moreover, the regularizer used in the primal space formulation allows us to mitigate the ill-posed problem resulting from the inverse map.

It should also be noted that the unknowns of the dual form of the LS-SVM formulation can be suitably estimated as a solution of a linear problem by inverting a square matrix [26]. On the other hand, due to the ε-insensitive loss function, the convex optimization problem for the case of the SVM leads to a sparse solution, but it must be solved numerically [25].

3.2. GPR Regression

Differently from the LS-SVM and SVM regression, the GPR can be used to build a probabilistic version of the inverse surrogate model. Starting from the N_s training samples $\mathcal{D} = \{(\mathbf{x}_l, \mathbf{y}_l)\}_{l=1}^{N_s}$, the GPR allows to train a probabilistic inverse model able to estimate for any configuration of the input parameters a statistical interpretation of its prediction in terms of a Gaussian distribution, which provides additional information on its reliability [27]. Specifically, for any test sample $\mathbf{y}_*$, the output of a GPR model is:

$$x_{i,*} \sim p(x_{i,*} | \mathbf{y}_*, \mathcal{D}) = \mathcal{N}(\mu_{\mathbf{y}_*}, \sigma^2_{\mathbf{y}_*}) \tag{7}$$

in which $\mu_{\mathbf{y}_*}$ and $\sigma^2_{\mathbf{y}_*}$ are the so-called posterior mean and variance, computed as follows:

$$\mu_{\mathbf{y}_*} = \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{x} \tag{8}$$

$$\sigma^2_{\mathbf{y}_*} = k_{**} - \mathbf{k}_*(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k}_*^T \tag{9}$$

where $\mathbf{y} = [y_1, \ldots, y_{2N_f}]$, $\mathbf{K} \in \mathbb{R}^{N_s \times N_s}$ is the correlation matrix in which the covariance function k is evaluated on pairs of the training data (i.e., $K_{ij} = K(\mathbf{y}_i, \mathbf{y}_j)$), $\mathbf{k}_* = [k(\mathbf{y}_1, \mathbf{y}_*), \ldots, k(\mathbf{y}_L, \mathbf{y}_*)]^T$ is a column vector, $k(\mathbf{y}_*, \mathbf{y}_*)$ is a scalar, $\mathbf{I} \in \mathbb{R}^{N_s \times N_s}$ is the identity matrix and σ_n^2 is an hyperparameter representing the variance of a possible additive noise corrupting the training dataset. Similarly to the kernel function, the covariance function used in the GPR defines the correlation between the values of pairs of model output at different point in the parameters space. Hereafter in this paper, a squared exponential covariance function, also known as RBF kernel, is used for the GPR, i.e., $k(\mathbf{y}, \mathbf{y}') = \exp\left(-\|\mathbf{y} - \mathbf{y}'\|^2/2\sigma_l^2\right)$ [27], where σ_l is a hyperparameters estimated during the model training.

4. Learning framework for the inverse model

The LS-SVM, the SVM and the GPR regression models learn the inverse model through a suitable framework. To learn the inverse function of the Fouquet model through a data-driven regression technique, it is necessary to observe both the domain and the co-domain of the function $\mathbf{x} = \mathcal{M}^{-1}(\mathbf{y})$. However, given the very high dimensionality of the two spaces, one depending on the number of frequencies, the other on the number of parameters, an unconstrained generation of observations would be computationally intractable. It is worth to note that, since the forward model $\bar{\mathbf{Z}} = \bar{Z}_F(\mathbf{x}, \mathbf{f})$ is available, the sampling can be limited to the parameters space and the spectrum obtained through the evaluation of the direct model for each sampled parameters' set. However, even in this case, the training procedure would benefit from a bounded search space. From a computational perspective, bounding the parameters values between specific maxima and minima values drastically reduce the training time. Additionally, from a physical perspective, only certain combinations of parameters are feasible and lead to a meaningful impedance spectra, thus only certain regions of the parameters space are worth to be explored.

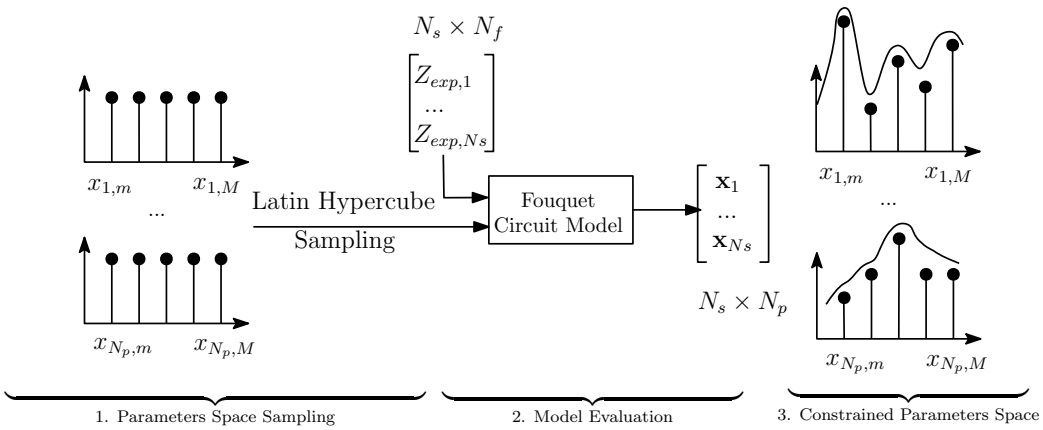


Figure 1: Generation of the training set. Each set of parameters is sampled through the LHS method and is processed with the Fouquet model in the Grid Search Procedure.

A synthetic dataset, which is obtained through a specific data-generating procedure shown in Fig. 1, has been employed for training purposes. It employs the grid search optimization briefly described in Sec. 2 together with the Fouquet ECM in which the candidate sets of parameters are sampled from the search space via a Latin Hypercube Sampling (LHS) scheme. In the figure, N_s refers to the number of training samples, $N_p = 6$ is the number of parameters involved in the Fouquet ECM, i.e. the dimension of the parameters vector $\mathbf{x}$, and $N_f = 48$ is the number of frequency values.

4.1. Parameters Space Sampling

Fig. 1 shows that, at the beginning of the procedure, the parameters space is sampled in the range between $\mathbf{x}_m$ and $\mathbf{x}_M$. These limits have been set after having analysed the experimental impedance spectra corresponding to the extreme fault conditions. Indeed, the experimental data have been obtained by inducing stack faults with an increasing level of severity. By moving the stack operating condition away from the nominal one, impedance modulus and phase change and lead to visible effects on the spectra, e.g. in terms of a magnification or reduction of the arcs in the Nyquist diagram. The effect depends on the specific fault, on its severity, and on the duration of the fault condition. For a specific operating condition imposed at stack level, the spectra of its cells that are connected in series may not exhibit the same alteration, because of inhomogeneities in the gas flows, constructive mismatching among cells and differences in cells degradation rate.

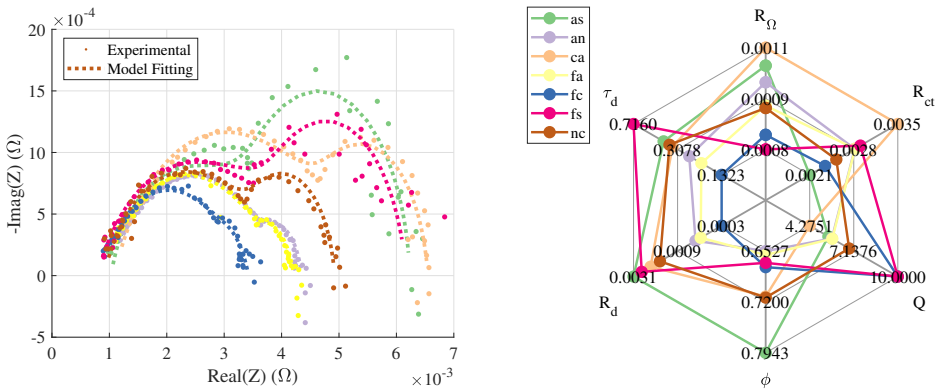


Figure 2: EIS Impedance curves (top panel) in nominal and faulty conditions identified offline and the corresponding parameters (bottom panel) defining the parameters space.

Thus, the boundaries for each parameter range have been fixed by picking, in a subset of curves referring to the same operating condition, the spectrum that is the farthest from the nominal one. The resulting spectra are considered as extreme ones and selected for a robust identification. The parametric identification has been performed according to the approach described in [36].

The fluctuation of the experimental measurements around the GA-based identified impedance curve is shown in Fig. 2. In the figure, the curves refer to the fitting while the dots are the EIS measurements of the farthest curves from the reference. Among the selected curves, the extreme values of each parameter are considered. Fig. 2 also shows the coverage of the parameters space obtained through the described procedure, together with the obtained identified parameters set for each specific operating condition. The colors are used to distinguish the different FC working conditions: Air Starvation (*as*), Anode and Cathode Drying (*an* and *ca*), Anode and Cathode Flooding (*fc* and *fa*) Fuel Starvation (*fs*) and Normal Conditions (*nc*).

4.2. Grid Search Optimization and the reduced parameters space

The parameters space defined in the previous subsection and depicted in Fig. 1 with the left-most stem plot, is sampled through a LHS procedure and evaluated during the grid search optimization to obtain the set of solutions $\mathbf{X}_{MC} = [\mathbf{x}_1, ..., \mathbf{x}_{N_s}]$ for the experimental impedance spectra $[\mathbf{Z}_{exp,1}, ..., \mathbf{Z}_{exp,N_s}]$. Thereby, the set of solutions $\mathbf{X}_{MC}$ is used to reduce the initial parameters space and build a prior discrete N_p-dimensional distribution of the parameters. This is represented by the rightmost stem plot in

Fig. 1. The obtained discrete distribution is smoothed with a Kernel Density Estimate (KDE) procedure so that it can be sampled to obtain an expanded training set of arbitrary dimension [37]. This procedure ensures that the three regression algorithms are focused on the plausible regions of the parameters space, which are those ones leading to impedance spectra that are representative of the FC operating conditions. Moreover, the availability of a continuous distribution from which an arbitrary number of spectra is drawn, allows to benchmark the performance of the regression algorithms by using different sizes of the training set.

5. Results on synthetic data

The proposed approaches are firstly validated by using the synthetic dataset. Both the training and test sets are sampled from the reduced parameters' space that is obtained as it has been described in Section 4.2. The reduced parameters' space is obtained by using the 30% and 70% of the experimental spectra, in order to create the training and the testing sets respectively. Thus, at this stage, the fractions of the experimental dataset reported above are exploited to estimate the parameters' distributions. The resulting distributions for the training and test sets are then sampled and processed through the Fouquet ECM to obtain the corresponding synthetic impedance spectra that are used for training and test purposes.

The use of spectra that do not come from experimental measurements but that have been generated by using the Fouquet ECM allows to analyse the performance of the regression algorithms. Indeed, to each synthetic spectrum, belonging either to the training set or to the test set, corresponds a set of parameters of the Fouquet ECM that guarantees the perfect fitting of the spectrum thereof. Consequently, modelling inaccuracies of the Fouquet ECM for the available experimental data do not affect the identification result and the performance of the regression approaches can be analysed in detail. The use of synthetic data also allows to study the impact of the number of training samples and of the noise affecting the data in the Nyquist plane, the latter being added through a simple numeric procedure applied to the test set. The noise has been added to the smooth impedance generated through the Fouquet ECM. The spectroscopy measurements are known to be affected by a stochastic noise, whose standard deviation is a function of the frequency and the magnitude of the impedance vector [38]. Therefore, to mimic the

noise of the available experimental spectra, the impedance used to build the synthetic test set has been altered according to (10):

$$Z(f) = \bar{Z}_F(f; \mathbf{x}_*) + \underbrace{(\beta_{re}(f) + j\beta_{im}(f))}_{n(f)}, \tag{10}$$

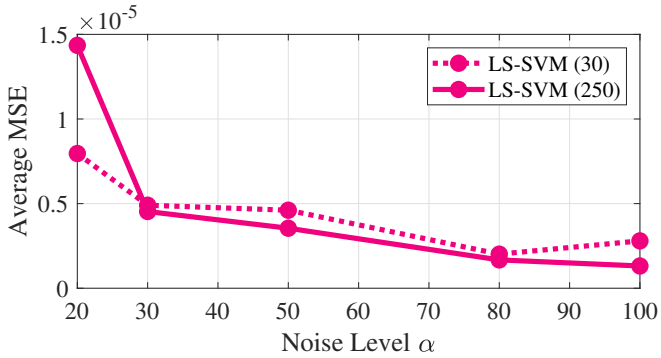
where $\beta_{re}(f), \beta_{im}(f) \sim \mathcal{N}(0, \sigma^2(f))$ are Gaussian distributed stochastic variables describing a non-stationary additive noise $n(f)$, in which the standard deviation is computed according to (11):

$$\sigma(f) = \frac{|\bar{Z}_F(f; \mathbf{x}_*)|}{\alpha} \tag{11}$$

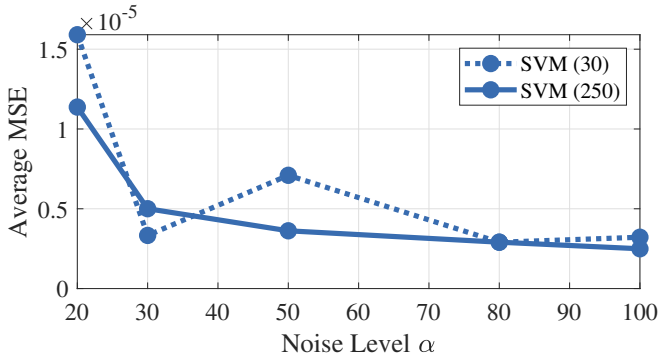
and α is used to change the noise level.

The amplitude of the noise term $n(f)$ is proportional to the impedance modulus, and affects its real and imaginary parts independently.

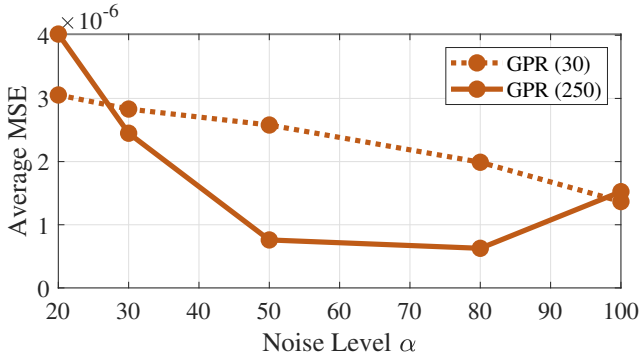
Fig. 3 shows the average performance, over three consecutive runs, of the three algorithms on the test set. The algorithms are trained to infer the vector of parameters $\mathbf{x}_*$ starting from the impedance Z. The evaluation of the Fouquet model returns the estimated impedance $\bar{Z}_F(f)(f; \mathbf{x}_*)$. The average MSE norm between the estimated spectra and the ones of the test set is used to compare the results. In Fig. 3(a) the LS-SVM performance underlines the impact of the noise level for two different sizes of the training set. For high levels of noise, thus α lower than 30, the average performance with 250 samples is worst than the one obtained with the reduced sample size of 30. When the noise level decreases, the impact of the increased dataset size becomes visible. At $\alpha = 100$ the performance index is halved. Similar results are also obtained for the GPR model (Fig. 3(c)) and very similar results between the two dataset dimensions have been reached when considering $\alpha = 100$. It follows that, with the proposed parameters space sampling, the impact of a reduced training set size is marginal and the size of the training set becomes important only for low levels of noise.



(a)



(b)



(c)

Figure 3: Influence of the number of training samples and of noise level α on the average MSE computed on the synthetic test set: (a) LS-SVM (b) SVM (c) GPR. Each point is the average of three training runs.

5.1. Hyperparameters optimization

For the SVM and the LS-SVM regressions, the optimization of the model hyperparameters is performed through a 10-fold CV [34] scheme and a Bayesian optimization [39]. Whilst, the GPR hyperparameters are optimized by maximizing the likelihood on the training set via a grid search approach. The above implementations automate the parameters tuning procedure, thus enabling the applicability of the methods to different ECMs with small changes, often outperforming the standard hand-tuning approaches. The procedure is executed on the hyperparameters domain defined as it follows. For the SVM and the LS-SVM these are defined in the range $C : (1 \times 10^{-6}, 5 \times 10^{4})$, $\gamma : (1 \times 10^{-6}, 5 \times 10^{4})$ and $\varepsilon : (1 \times 10^{-11}, 1 \times 10^{2})$ (only tuned for the SVM regression). For The GPR, the hyperparamers' ranges are $\sigma_l : (1 \times 10^{-5}, 1 \times 10^{5})$ and $\sigma_n^2 : (1 \times 10^{-5}, 1 \times 10^{5})$.

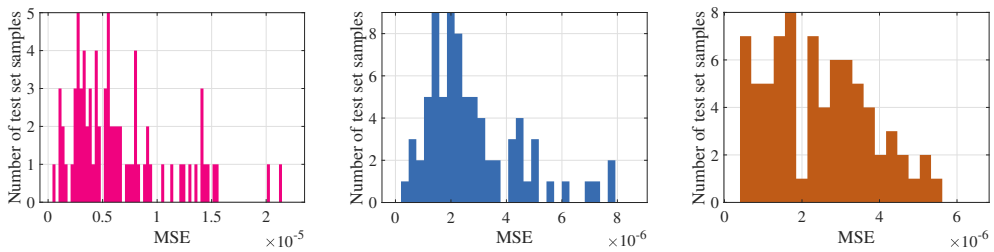
6. Experimental Results

The described procedure has been tested on an experimental dataset [1] described in the Introduction. The parameters space is built by using the procedure described in the previous section and the grid search optimization is employed to narrow the search space operating on 70% of the dataset, corresponding to 110 impedance spectra, while 30% of the set is left for the results evaluation. The histogram plots shown in Fig. 4 give a comprehensive view of the obtained results in terms of the MSE resulting by the comparison of the experimental data with the Fouquet impedance at the same frequency values. The MSE is obtained by using the parameters sets that are inferred by the three regression algorithms on all the spectra of the testing set. After, for each inferred parameters set, the Fouquet impedance is computed and the MSE is calculated with respect to the corresponding experimental data. Consequently, the MSEs histograms shown in Fig. 4 summarize the performance of the three approaches on all the spectra of the test set. Tab. 1 reports the identification results in terms of mean, median, and interquartile range (IQR) of the MSE and MAPE calculated over the spectra belonging to the test set. The mean is reported as an overall performance index and the median is provided to quantify the robustness of the result over the entire test set. The closer the mean and median values the better the identification performance. In this case, the error histogram approaches a normal distribution, indicating that there are no reconstructed spectra with an error that

Table 1: Mean, Median, and Interquartile Range (IQR) of the MSE and MAPE computed on the whole test set

	MSE ($\times 10^{-5}$)			**MAPE**		
	Mean	**Median**	**IQR**	**Mean**	**Median**	**IQR**
GPR	0.3179	0.2419	0.2017	0.9985	0.9858	0.0358
SVM	0.3403	0.2325	0.2109	1.0875	1.0844	0.0492
LS-SVM	0.7387	0.5537	0.5407	1.0863	1.0779	0.0438

differs significantly from the mean. The interquartile range is also provided to assess the variability of the obtained result.



(a) LS-SVM (b) SVM (c) GPR

Figure 4: Histogram of the MSE errors computed between the inverse model reconstructions of the impedance spectra and the experimental measurement of the testing set. Each bar corresponds to the number of spectra of the test set giving an MSE falling in the corresponding range.

The GPR regression, for a larger number of test set spectra, gives a smaller error, both in comparison with those ones ensured by the other methods. Moreover, the reduced spread of the values of the histogram plot indicates that the obtained performance is well generalized on the whole test set. The SVM shows an average performance similar to the GPR but with an increased spread of the error histogram. Contrarily, the LS-SVM shows a performance degradation and higher variability of the result. It is important to point out that the low performance of the LS-SVM regression could be due to the quite large range of variation the hyperparameters explored during the training phase. To quantify the results, different degrees of performance have been considered.

Experimental EIS spectra are affected by an evident noise and artifacts, thus those ones resulting with an error below 2×10^{-6} are considered as correctly identified. Fig. 4 shows that GPR and SVM guarantees that the

45% and the 41%, respectively, of the spectra in the test set are very well identified, because the error keeps below the defined threshold. The LS-SVM gives a 10% of the spectra that are identified with an error due to noise artifacts. The identification performance is reported in Fig. 5. In both the cases, the GPR and SVM spectra show optimal results with an MSE below 2×10^{-6}. The results shown in Fig. 5(a) and (b) are obtained through LS-SVM and are reconstructed with an error equal to 2.8×10^{-6} and 1.5×10^{-5}, respectively.

Also, the range $2 \times 10^{-6} < \text{MSE} < 2.5 \times 10^{-6}$ is giving valuable solutions. For the GPR, SVM and LS-SVM models, respectively 11% and 14% and 4% of the test set are reconstructed with this error. An example of the quality of the reconstruction obtained in this range is shown in Fig. 6.

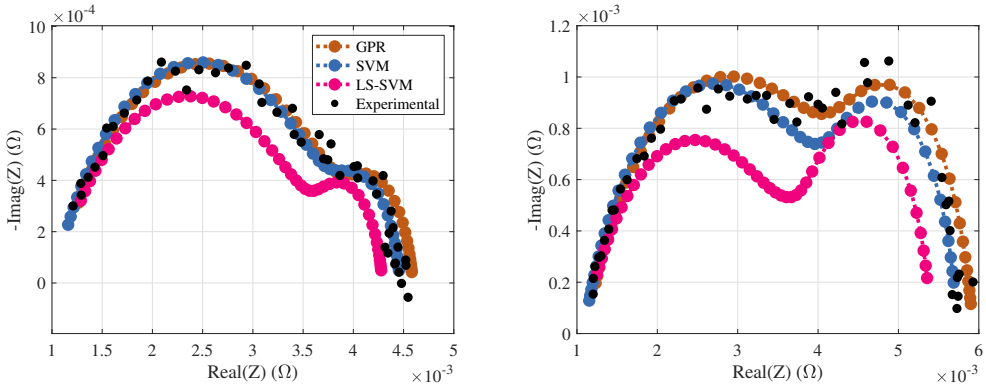


Figure 5: Best identification performance obtained by the GPR and SVM inverse models.

7. On Field Results

An important point of strength of the proposed inverse model is the short execution time observed during the phase of inference. Although the training is computationally expensive due to the dimensionality of the space to be explored and due to the automatic hyperparameters tuning, the trained model consists of a simple function, whose evaluation on the measured impedance spectra returns the set of parameters. For this reason, differently from other standard model identification procedures relying on iterative algorithms, the inference through the inverse model is a good candidate for the implementation on a low cost embedded system, since the impact on the resources

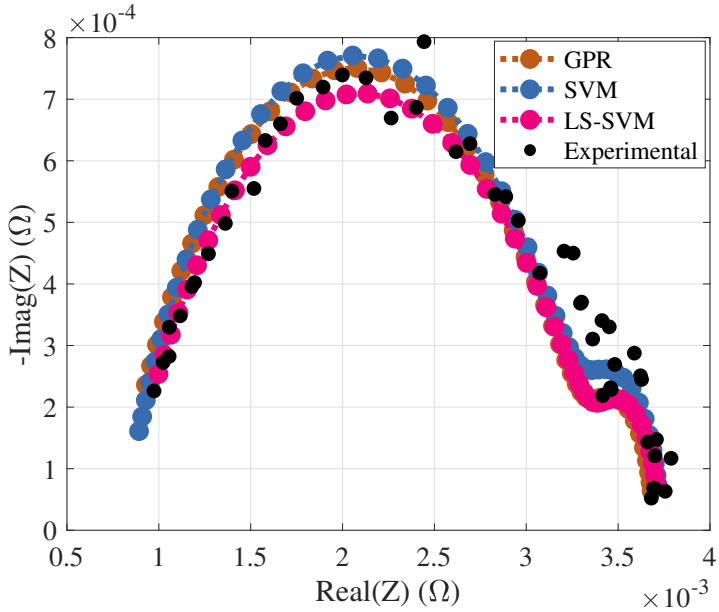


Figure 6: Middle quality identification obtained by the three inverse models.

usage is negligible and the execution time is deterministic. Both the latter aspects are in favor of the approach proposed in this paper with respect to any stochastic optimization procedure for parameters identification [22, 23, 40, 41, 42]. The three algorithms are implemented in Python and executed in a decentralized way. The training procedure, that is computationally expensive due to the hyperparameters optimization procedure, is executed on the cloud by employing the Google Colab platform. The remote machine is equipped with a Dual-Core Haswell Intel Xeon processor at 2.30 GHz and 12 GB of RAM. After the training procedure, the testing is executed on a Raspberry Pi 3 embedded system equipped with a Quad-Core 1.2 GHz Broadcom BCM2837 CPU and 1 GB of RAM and the execution times of the three algorithms have been considered as benchmark. In Fig. 7(a), the training time of the three models is reported as a function of the size of the training set. The training time is not only affected by the number of samples in the training set, but also by the number of hyperparameters to be tuned and the optimization scheme. This explains the scale difference between the three techniques. In fact, the training of the SVM regression turns out to be the most expensive one, since it requires the tuning of three

hyperparameters via the CV. On the other hand, the training is faster for the GPR because the optimization of the hyperparameter is carried out on the training samples directly, without requiring to compute the CV error. Fig. 7(b) reports the execution time for the inference of a set of parameters.

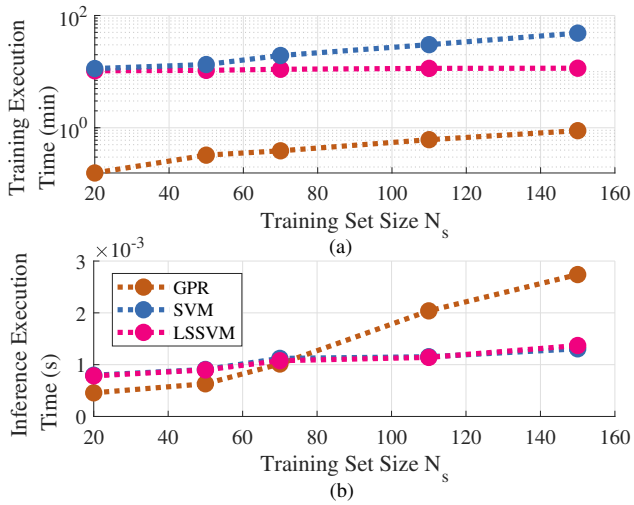


Figure 7: Training (a) and inference (b) execution times of the three regression algorithms.

The results clearly highlight the impact of the dimensionality of the training set on the inference time. Indeed, for the considered approaches, the model complexity increases with the number of training samples considered during the model training. The three algorithms infer the parameters set in few milliseconds. The chosen regressions trained with 110 samples perform the inference in 2 milliseconds for GPR and in slightly above 1 millisecond for the SVM and the LS-SVM models, respectively.

It is worth to have a comparison in terms of computation time also with approaches based on parameters identification from EIS data through the minimization of an objective function value [22][23][43][40]. The objective function is computed by fixing a set of Fouquet ECM parameters and a set of frequencies, then evaluating the Fouquet impedance and computing the MSE between the generated impedance and the experimental data. This computation has been performed in Python, like the other inverse models solutions, and run on the same Raspberry Pi 3 mentioned above. The average execution time is equal to 0.7 ms. This means that a competitive optimization

algorithm should be able to find the solution in less than 3 iterations to have an execution time that is comparable with that one of the GPR inference time, which is slightly above 2ms.

8. Conclusions

Machine learning regression models demonstrated to be an efficient choice to solve the inverse problem with minimal and deterministic execution time reported during inference. Leveraging the a priori knowledge on the ranges of the model parameters as well as the knowledge on the regions of the parameters space that are more likely to correspond to actual impedance measurements, the problem becomes computationally treatable with a training time that is lower than one minute for the GPR.

The three techniques, especially SVM and LS-SVM, have shown to be more sensitive to noise than to the dataset size. Moreover, for very high levels of noise, a big data set size has shown to be counterproductive. The LS-SVM showed weaker results in terms of reconstruction error due to its sensitivity to the hyperparameters' tuning and execution times similar to the SVM. Therefore it can be considered a valuable option when a small and noisy training set is available. Indeed, in the latter case the poor performance of the three techniques are comparable and the simplicity of the method is better exploited. The GPR and the SVM have shown very promising results both in simulated and experimental environment. GPR has been able to give an optimal reconstruction of the 45% of the spectra and a moderately good reconstruction of a further 11%. SVR gives a slightly lower accuracy that is compensated by the execution time required for the inference of the parameters, which is almost halved with respect to the execution time of the GPR.

Appendix A. Fouquet Equivalent Circuit Model

The Fouquet equivalent circuital model has already been adopted in [21][36][2] for the identification of PEMFC operating in faulty conditions, and it is defined as follows.

$$Z(f) = \bar{Z}_F(f;\mathbf{x}) = R_\Omega + \frac{\dot{Z}_{\mathrm{CPE}}(\omega)(R_{\mathrm{ct}} + \dot{Z}_{\mathrm{W}}(\omega))}{\dot{Z}_{\mathrm{CPE}}(\omega) + R_{\mathrm{ct}} + \dot{Z}_{\mathrm{W}}(\omega)}, \tag{A.1}$$

with

$$\dot{Z}_{\mathrm{CPE}}(f) = \frac{1}{Q(j\omega)^{\phi}}. \tag{A.2}$$

$$\dot{Z}_{\mathrm{W}}(f) = R_{\mathrm{d}} \frac{\tanh\sqrt{j\omega\tau_{\mathrm{d}}}}{\sqrt{j\omega\tau_{\mathrm{d}}}}, \tag{A.3}$$

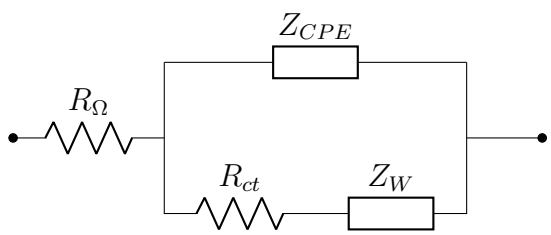


Figure A.8: The Fouquet equivalent circuit model.

Here $\omega = 2\pi f$ is the angular frequency. The equivalent circuit model makes use of a vector of $N_p = 6$ parameters $\mathbf{x} = [R_\Omega, R_{\mathrm{ct}}, Q, \phi, R_{\mathrm{d}}, \tau_{\mathrm{d}}]^T$.

R_Ω takes into account the losses due to the resistance of the electrolyte to the flow of protons, while R_{ct} represents the resistance at the electrode/electrolyte interface to the flow of charges. Q and ϕ are the parameter of the Constant Phase Element (CPE), used to model the frequency behaviour of the electrodes with rough or porous surfaces. The Warburg element is defined by R_{d}, that models the losses due to reactants' diffusion. The time constant of the diffusion process is specified with the parameter τ_{d}. The model is evaluated at the same frequencies where the EIS experiment was performed and $N_f = 48$ impedance points have been obtained.

References

[1] Real operation pem fuel cells health-state monitoring and diagnosis based on dc-dc converter embedded eis, Tech. rep., european H2020 project (2015).
URL `https://pemfc.health-code.eu/`

[2] G. Petrone, W. Zamboni, G.Spagnuolo, An interval arithmetic-based method for parametric identification of a fuel cell equivalent circuit model, Applied Energy 242 (2019) 1226 – 1236. doi:https://doi.org/10.1016/j.apenergy.2019.03.136.
URL `http://www.sciencedirect.com/science/article/pii/S0306261919305446`

[3] Z. Fei, F. Yang, K.-L. Tsui, L. Li, Z. Zhang, Early prediction of battery lifetime via a machine learning based framework, Energy 225 (2021) 120205. doi:https://doi.org/10.1016/j.energy.2021.120205.
URL https://www.sciencedirect.com/science/article/pii/S0360544221004540

[4] Z. Chen, N. Shi, Y. Ji, M. Niu, Y. Wang, Lithium-ion batteries remaining useful life prediction based on bls-rvm, Energy 234 (2021) 121269. doi:https://doi.org/10.1016/j.energy.2021.121269.
URL https://www.sciencedirect.com/science/article/pii/S0360544221015176

[5] S. Son, S. Jeong, E. Kwak, J. hyeong Kim, K.-Y. Oh, Integrated framework for soh estimation of lithium-ion batteries using multiphysics features, Energy 238 (2022) 121712. doi:https://doi.org/10.1016/j.energy.2021.121712.
URL https://www.sciencedirect.com/science/article/pii/S0360544221019605

[6] M. Zhang, T. Hu, L. Wu, G. Kang, Y. Guan, A method for capacity estimation of lithium-ion batteries based on adaptive time-shifting broad learning system, Energy 231 (2021) 120959. doi:https://doi.org/10.1016/j.energy.2021.120959.
URL https://www.sciencedirect.com/science/article/pii/S036054422101207X

[7] C. Qian, B. Xu, L. Chang, B. Sun, Q. Feng, D. Yang, Y. Ren, Z. Wang, Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries, Energy 227 (2021) 120333. doi:https://doi.org/10.1016/j.energy.2021.120333.
URL https://www.sciencedirect.com/science/article/pii/S036054422100582X

[8] W.-Y. Lee, H. Oh, M. Kim, Y.-Y. Choi, Y.-J. Sohn, S.-G. Kim, Hierarchical fault diagnostic method for a polymer electrolyte fuel cell system, International Journal of Hydrogen Energy 45 (47) (2020) 25733 – 25746, special Issue on HYPOTHESIS XIV. doi:https://doi.org/10.1016/j.ijhydene.2019.10.145.
URL http://www.sciencedirect.com/science/article/pii/S0360319919339564

[9] Z. Zhang, S. Li, Y. Xiao, Y. Yang, Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning, Applied Energy 233-234 (2019) 930 – 942. doi:https://doi.org/10.1016/j.apenergy.2018.10.113.
URL http://www.sciencedirect.com/science/article/pii/S0306261918316854

[10] D. Benouioua, D. Candusso, F. Harel, P. Picard, X. François, On the issue of the pemfc operating fault identification: Generic analysis tool based on voltage pointwise singularity strengths, International Journal of Hydrogen Energy 43 (25) (2018) 11606 – 11613, alternative Energies for Sustainability. doi:https://doi.org/10.1016/j.ijhydene.2017.09.177.
URL http://www.sciencedirect.com/science/article/pii/S0360319917338934

[11] K. Chen, S. Laghrouche, A. Djerdir, Degradation model of proton exchange membrane fuel cell based on a novel hybrid method, Applied Energy 252 (2019) 113439. doi:https://doi.org/10.1016/j.apenergy.2019.113439.
URL http://www.sciencedirect.com/science/article/pii/S0306261919311134

[12] K. Chen, S. Laghrouche, A. Djerdir, Proton exchange membrane fuel cell prognostics using genetic algorithm and extreme learning machine, Fuel Cells 20 (3) (2020) 263–271. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/fuce.201900085, doi:https://doi.org/10.1002/fuce.201900085.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/fuce.201900085

[13] R. Ma, T. Yang, E. Breaz, Z. Li, P. Briois, F. Gao, Data-driven proton exchange membrane fuel cell degradation predication through deep learning method, Applied Energy 231 (2018) 102 – 115. doi:https://doi.org/10.1016/j.apenergy.2018.09.111.
URL http://www.sciencedirect.com/science/article/pii/S0306261918314181

[14] K. V. S. Bharath, F. Blaabjerg, A. Haque, M. A. Khan, Model-based data driven approach for fault identification in proton exchange membrane fuel cell, Energies 13 (12) (2020) 3144. doi:10.3390/en13123144.
URL http://dx.doi.org/10.3390/en13123144

[15] H. Liu, J. Chen, D. Hissel, M. Hou, Z. Shao, A multi-scale hybrid degradation index for proton exchange membrane fuel cells, Journal of Power Sources 437 (2019) 226916. doi:https://doi.org/10.1016/j.jpowsour.2019.226916.
URL http://www.sciencedirect.com/science/article/pii/S0378775319309097

[16] H. Liu, J. Chen, D. Hissel, H. Su, Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method, Applied Energy 237 (2019) 910 – 919.

doi:https://doi.org/10.1016/j.apenergy.2019.01.023.
URL http://www.sciencedirect.com/science/article/pii/S0306261919300236

[17] B. Najafi, P. Bonomi, A. Casalegno, F. Rinaldi, A. Baricci, Rapid fault diagnosis of pem fuel cells through optimal electrochemical impedance spectroscopy tests, Energies 13 (14) (2020) 3643. doi:10.3390/en13143643.
URL http://dx.doi.org/10.3390/en13143643

[18] A. V. Panteleev, A. V. Lobanov, Mini-batch adaptive random search method for the parametric identification of dynamic systems, Automation and Remote Control 81 (11) (2020) 2026 – 2045. doi:https://doi.org/10.1134/S0005117920110065.
URL "https://link.springer.com/article/10.1134%2FS0005117920110065"

[19] Y. Bao, J. M. Velni, Data-driven linear parameter-varying model identification using transfer learning, IEEE Control Systems Letters 5 (5) (2021) 1579–1584. doi:10.1109/LCSYS.2020.3041407.

[20] H. Kabir, L. Zhang, M. Yu, P. H. Aaen, J. Wood, Q.-J. Zhang, Smart modeling of microwave devices, IEEE Microwave Magazine 11 (3) (2010) 105–118. doi:10.1109/MMM.2010.936079.

[21] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, B. Ould-Bouamama, Model based pem fuel cell state-of-health monitoring via ac impedance measurements, Journal of Power Sources 159 (2) (2006) 905–913. doi:https://doi.org/10.1016/j.jpowsour.2005.11.035.
URL https://www.sciencedirect.com/science/article/pii/S0378775305015946

[22] M. A. Kappel, R. Fabbri, R. P. Domingos, I. N. Bastos, Novel electrochemical impedance simulation design via stochastic algorithms for fitting equivalent circuits, Measurement 94 (2016) 344–354. doi:https://doi.org/10.1016/j.measurement.2016.08.008.
URL https://www.sciencedirect.com/science/article/pii/S0263224116304699

[23] M. A. Abud Kappel, F. C. Peixoto, G. M. Platt, R. P. Domingos, I. N. Bastos, A study of equivalent electrical circuit fitting to electrochemical impedance using a stochastic method, Applied Soft Computing 50 (2017) 183–193. doi:https://doi.org/10.1016/j.asoc.2016.11.030.
URL https://www.sciencedirect.com/science/article/pii/S1568494616305993

[24] Z. Sun, D. Cao, Y. Ling, F. Xiang, Z. Sun, F. Wu, Proton exchange membrane fuel cell model parameter identification based on dynamic differential evolution with collective guidance factor algorithm, Energy 216 (2021) 119056. doi:https://doi.org/10.1016/j.energy.2020.119056.
URL https://www.sciencedirect.com/science/article/pii/S0360544220321630

[25] V. N. Vapnik, The nature of statistical learning theory, Springer-Verlag New York, Inc., 1995.

[26] J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, 2002.

[27] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[28] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero, M. Swaminathan, Machine learning and uncertainty quantification for surrogate models of integrated devices with a large number of parameters, IEEE Access 7 (2019) 4056–4066.

[29] J.-M. Bourinet, Reliability analysis and optimal design under uncertainty - Focus on adaptive surrogate-based approaches, Habilitation à diriger des recherches, Université Clermont Auvergne (Jan. 2018).
URL https://tel.archives-ouvertes.fr/tel-01737299

[30] P. Moçotéguy, B. Ludwig, D. Beretta, T. Pedersen, Study of the impact of reactants utilization on the performance of pemfc commercial stacks by impedance spectroscopy, International Journal of Hydrogen Energy 46 (10) (2021) 7475–7488. doi:https://doi.org/10.1016/j.ijhydene.2020.11.197.
URL https://www.sciencedirect.com/science/article/pii/S0360319920344475

[31] P. Moçotéguy, B. Ludwig, D. Beretta, T. Pedersen, Study of the impact of water management on the performance of pemfc commercial stacks by impedance spectroscopy, International Journal of Hydrogen Energy 45 (33) (2020) 16724–16737. doi:https://doi.org/10.1016/j.ijhydene.2020.04.139.
URL https://www.sciencedirect.com/science/article/pii/S0360319920315421

[32] T. Söderström, P. Stoica, System identification, New York (N.Y.) : Prentice-Hall, 1989.

[33] S. Boyd, S. P. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.

[34] B. Ghojogh, M. Crowley, The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial (2019). arXiv:1905.12787.

[35] A. J. Smola, B. Schölkopf, A tutorial on support vector regression, Statistics and Computing 14 (3) (2004) 199–222. doi:10.1023/B:STCO.0000035301.49549.88.

[36] W. Zamboni, G. Petrone, G. Spagnuolo, D. Beretta, An evolutionary computation approach for the online/on-board identification of pem fuel cell impedance parameters with a diagnostic perspective, Energies 12 (22) (2019). doi:10.3390/en12224374.
URL https://www.mdpi.com/1996-1073/12/22/4374

[37] X. Yang, X. Cheng, Active learning method combining kriging model and multimodal-optimization-based importance sampling for the estimation of small failure probability, International Journal for Numerical Methods in Engineering 121 (21) (2020) 4843–4864.

[38] P. Agarwal, O. D. Crisalle, M. E. Orazem, L. H. Garcia-Rubio, Application of measurement models to impedance spectroscopy: II . determination of the stochastic contribution to the error structure, Journal of The Electrochemical Society 142 (12) (1995) 4149–4158. doi:10.1149/1.2048478.

[39] E. Brochu, V. M. Cora, N. de Freitas, A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning (2010). arXiv:1012.2599.

[40] E. A. Gouda, M. F. Kotb, A. A. El-Fergany, Jellyfish search algorithm for extracting unknown parameters of pem fuel cell models: Steady-state performance and analysis, Energy 221 (2021) 119836. doi:https://doi.org/10.1016/j.energy.2021.119836.
URL https://www.sciencedirect.com/science/article/pii/S0360544221000852

[41] W. H. Tanveer, H. Rezk, A. Nassef, M. A. Abdelkareem, B. Kolosz, K. Karuppasamy, J. Aslam, S. O. Gilani, Improving fuel cell

performance via optimal parameters identification through fuzzy logic based-modeling and optimization, Energy 204 (2020) 117976. doi:https://doi.org/10.1016/j.energy.2020.117976.
URL https://www.sciencedirect.com/science/article/pii/S0360544220310835

[42] Z. Yang, Q. Liu, L. Zhang, J. Dai, N. Razmjooy, Model parameter estimation of the pemfcs using improved barnacles mating optimization algorithm, Energy 212 (2020) 118738. doi:https://doi.org/10.1016/j.energy.2020.118738.
URL https://www.sciencedirect.com/science/article/pii/S0360544220318454

[43] E. A. Adinolfi, M. Gallo, P. Polverino, D. Beretta, S. S. Araya, C. Pianese, Ecm-based algorithm for on-board pemfcs diagnosis, in: W. Zamboni, G. Petrone (Eds.), ELECTRIMACS 2019, Springer International Publishing, Cham, 2020, pp. 103–115.