Supervised learning of a regression model based on latent process. Application to the estimation of Fuel Cell lifetime

Raïssa Onanena[1], Faicel Chamroukhi[1], Latifa Oukhellou[1],
Denis Candusso[1,3], Patrice Aknin[1], Daniel Hissel[2,3]

[1]INRETS-LTN, 2 av de la butte verte, 93166 Noisy le Grand Cedex, France
[2]FEMTO-ST UMR CNRS 6174, Université de Franche-Comté, 90010 Belfort, France
[3]FCLAB, Rue Ernest Thierry-Mieg, 90010 Belfort Cedex, France

December 15 2009

Context: Predictive maintenance of the Fuel Cells (FCs)

- Fuel Cells (FCs) are widely used in many domains including transport
- They can offer high fuel economy
- Lower CO_2 emissions
- The stack is affected by the operating conditions (temperature, mechanical constraints on the membrane, electrode assemblies etc.)

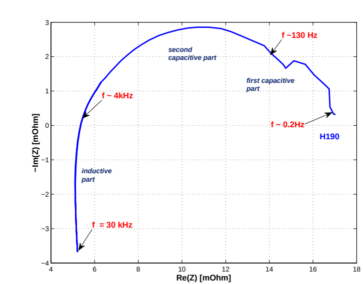
⇒ a predictive maintenance policy is needed

Aim

FC lifetime estimation using specific measurements acquired during the ageing study of the stacks

The Electrochemical impedance spectrum (EIS)

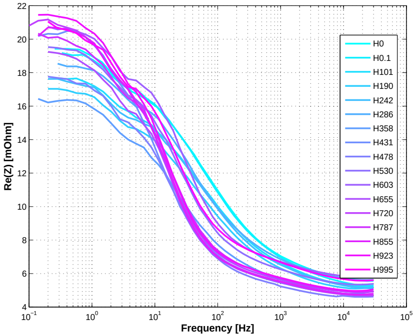
- Measurements of the Electrochemical Impedance Spectrum (EIS) are generally used for FC characterization



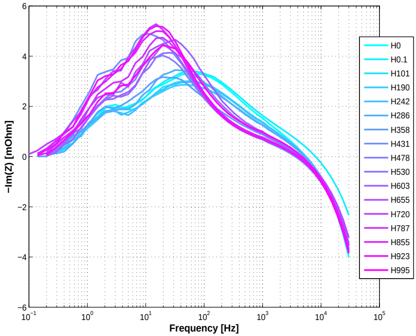
The impedance spectrum for the Fuel Cell consists of three regimes:

- a first capacitive arc ($f < 130$Hz) due to the diffusion phenomena
- a second capacitive arc ($130\text{Hz} <= f < 4\text{kHz}$) linked to the FC membrane charges
- a last inductive part arc which is present in high frequencies ($4\text{kHz} <= f$) due to the inductive behavior of connections

Evolution of the real and imaginary parts of the Electrochemical Impedance Spectrum (EIS) over time



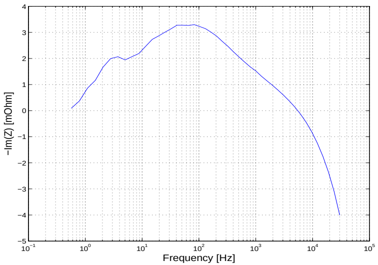
Real part



Imaginary part

Feature extraction from the imaginary part of the EIS

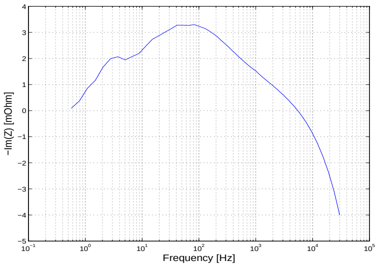
- The imaginary part of the spectrum is more informative and more complex than the real part
- Particularly, three regimes corresponding to the behaviour of the stack are perceptible:



- Smooth or abrupt changes between the different regimes

Feature extraction from the imaginary part of the EIS

- The imaginary part of the spectrum is more informative and more complex than the real part
- Particularly, three regimes corresponding to the behaviour of the stack are perceptible:



- Smooth or abrupt changes between the different regimes

⇒ The proposed solution: use an adapted regression model whose parameters will be used as the feature vector for each EIS

A regression model with a hidden logistic process

The data: $\{(x_1, f_1), \ldots, (x_n, f_n)\}$

- x_i: real dependent variable: - Imaginary part of the EIS
- f_i: independent variable representing the frequency

$$\forall i = 1, \ldots, n, \quad x_i = \boldsymbol{\beta}_{z_i}^T \boldsymbol{r}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),$$

- $z_i \in \{1, \ldots, K\}$ hidden variable: the class label of the component generating x_i
- $\boldsymbol{\beta}_{z_i} \in R^{p+1}$: regression coefficients of the sub-model z_i
- $\boldsymbol{r}_i = (1, f_i, \ldots, f_i^p)^T$: covariate vector in R^{p+1}

A regression model with a hidden logistic process

The data: $\{(x_1, f_1), \ldots, (x_n, f_n)\}$

- x_i: real dependent variable: - Imaginary part of the EIS
- f_i: independent variable representing the frequency

$$\forall i = 1, \ldots, n, \quad x_i = \boldsymbol{\beta}_{z_i}^T \boldsymbol{r}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),$$

- $z_i \in \{1, \ldots, K\}$ hidden variable: the class label of the component generating x_i
- $\boldsymbol{\beta}_{z_i} \in R^{p+1}$: regression coefficients of the sub-model z_i
- $\boldsymbol{r}_i = (1, f_i, \ldots, f_i^p)^T$: covariate vector in R^{p+1}

$\mathbf{z} = (z_1, \ldots, z_n)$ is a hidden logistic process

A regression model with a hidden logistic process

The data: $\{(x_1, f_1), \ldots, (x_n, f_n)\}$

- x_i: real dependent variable: - Imaginary part of the EIS
- f_i: independent variable representing the frequency

$$\forall i = 1, \ldots, n, \quad x_i = \boldsymbol{\beta}_{z_i}^T \boldsymbol{r}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),$$

- $z_i \in \{1, \ldots, K\}$ hidden variable: the class label of the component generating x_i
- $\boldsymbol{\beta}_{z_i} \in R^{p+1}$: regression coefficients of the sub-model z_i
- $\boldsymbol{r}_i = (1, f_i, \ldots, f_i^p)^T$: covariate vector in R^{p+1}

$\mathbf{z} = (z_1, \ldots, z_n)$ is a hidden logistic process

$z_i \sim \mathcal{M}\left(1, \pi_{i1}(\mathbf{w}), \ldots, \pi_{iK}(\mathbf{w})\right)$; where

$$\pi_{ik}(\mathbf{w}) = p(z_i = k; \mathbf{w}) = \frac{\exp\left(w_{k0} + w_{k1} f_i\right)}{\sum_{\ell=1}^{K} \exp\left(w_{\ell 0} + w_{\ell 1} f_i\right)},$$

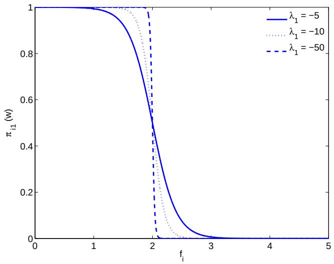
- $\mathbf{w} = (w_{10}, w_{11}, \ldots, w_{K0}, w_{K1}) \in R^{2K}$ the parameter vector for the K logistic functions

Flexibility of the logistic transformation

Variation of $\pi_{ik}(\mathbf{w})$ in relation to $\mathbf{w}$:

- Use the notation $\boldsymbol{w}_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$
- Example of two components:

$\alpha_1 = -2$

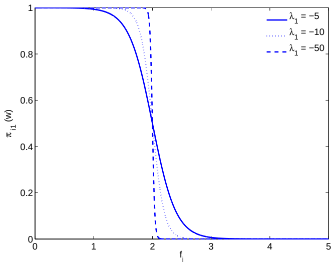


Flexibility of the logistic transformation

Variation of $\pi_{ik}(\mathbf{w})$ in relation to $\mathbf{w}$:

- Use the notation $\boldsymbol{w}_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$
- Example of two components:

$\alpha_1 = -2$



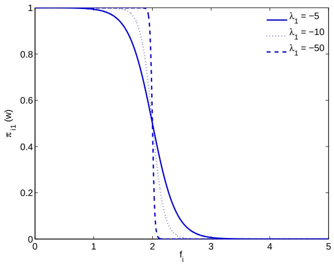
⇒ The parameter λ_k controls the quality of transitions (smooth/abrupt) between the regimes

Flexibility of the logistic transformation

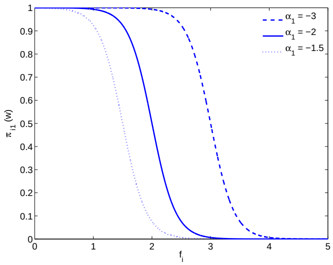
Variation of $\pi_{ik}(\mathbf{w})$ in relation to $\mathbf{w}$:

- Use the notation $\boldsymbol{w}_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$
- Example of two components:

$\alpha_1 = -2$



$\lambda_1 = -5$



$\Rightarrow$ The parameter λ_k controls the quality of transitions (smooth/abrupt) between the regimes

$\Rightarrow$ The parameter α_k is directly linked to the frequency at the transition point

Parameter estimation by maximum likelihood

- Derived mixture density

$$p(x_i;\boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_{ik}(\mathbf{w})\mathcal{N}(x_i;\boldsymbol{\beta}_k^T\boldsymbol{r}_i,\sigma_k^2)$$

- Model parameters

$$\boldsymbol{\theta} = \left(\mathbf{w},\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_K,\sigma_1^2,\ldots,\sigma_K^2\right)$$

- Log-likelihood of $\boldsymbol{\theta}$:

$$L(\boldsymbol{\theta};\mathbf{x}) = \sum_{i=1}^{n}\log\sum_{k=1}^{K} \pi_{ik}(\boldsymbol{w})\mathcal{N}(x_i;\boldsymbol{\beta}_k^T\boldsymbol{r}_i,\sigma_k^2).$$

- Maximization of $L(\boldsymbol{\theta};\mathbf{x})$ by a dedicated Expectation-Maximization (EM) algorithm [Dempster et al. 77].

Dedicated EM algorithm

Dedicated EM algorithm

Initialization: $\boldsymbol{\theta}^{(0)}$

Dedicated EM algorithm

Initialization: $\boldsymbol{\theta}^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) = E\Big[L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{z}) | \mathbf{x}, \boldsymbol{\theta}^{(q)}\Big]$$

$$= \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}\Big(x_i; \boldsymbol{\beta}_k^T \boldsymbol{r}_i, \sigma_k^2\Big)}_{Q_2(\boldsymbol{\beta}_k, \sigma_k^2 | k=1,\ldots,K)},$$

where $\tau_{ik}^{(q)} = p(z_i = k | x_i; \boldsymbol{\theta}^{(q)})$ is the posterior probability of the kth regime

Dedicated EM algorithm

Initialization: $\boldsymbol{\theta}^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) = E\Big[L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{z})|\mathbf{x}, \boldsymbol{\theta}^{(q)}\Big]$$

$$= \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}\big(x_i; \boldsymbol{\beta}_k^T \mathbf{r}_i, \sigma_k^2\big)}_{Q_2(\boldsymbol{\beta}_k, \sigma_k^2 | k=1,\ldots,K)},$$

where $\tau_{ik}^{(q)} = p(z_i = k | x_i; \boldsymbol{\theta}^{(q)})$ is the posterior probability of the kth regime

2. **Maximization step:** Compute $\boldsymbol{\theta}^{(q+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)})$

Dedicated EM algorithm

Initialization: $\boldsymbol{\theta}^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) = E\Big[L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{z}) | \mathbf{x}, \boldsymbol{\theta}^{(q)}\Big]$$

$$= \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}\big(x_i; \boldsymbol{\beta}_k^T \boldsymbol{r}_i, \sigma_k^2\big)}_{Q_2(\boldsymbol{\beta}_k, \sigma_k^2 | k=1,\ldots,K)},$$

where $\tau_{ik}^{(q)} = p(z_i = k | x_i; \boldsymbol{\theta}^{(q)})$ is the posterior probability of the kth regime

2. **Maximization step:** Compute $\boldsymbol{\theta}^{(q+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)})$

   1. Maximization of Q_2 w.r.t $\{\boldsymbol{\beta}_k, \sigma_k^2\}$ $(k = 1 \ldots, K)$: Analytic solutions

Dedicated EM algorithm

Initialization: $\boldsymbol{\theta}^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$
\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) &= E\Big[L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{z})|\mathbf{x}, \boldsymbol{\theta}^{(q)}\Big] \\
&= \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}\big(x_i; \boldsymbol{\beta}_k^T \mathbf{r}_i, \sigma_k^2\big)}_{Q_2(\boldsymbol{\beta}_k, \sigma_k^2 | k=1,\ldots,K)},
\end{aligned}
$$

where $\tau_{ik}^{(q)} = p(z_i = k | x_i; \boldsymbol{\theta}^{(q)})$ is the posterior probability of the kth regime

2. **Maximization step:** Compute $\boldsymbol{\theta}^{(q+1)} = \arg \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)})$

   1. Maximization of Q_2 w.r.t $\{\boldsymbol{\beta}_k, \sigma_k^2\}$ $(k = 1 \ldots, K)$: Analytic solutions
   2. Maximization of Q_1 w.r.t $\mathbf{w}$: a multiclass weighted logistic regression problem $\Rightarrow$ IRLS algorithm [Green 84, Jordan & jacobs 94]

Measurement approximation

As in standard regression, given the estimated parameters, x_i is approximated by its expectation:

$$\hat{x}_i = E(x_i;\hat{\boldsymbol{\theta}}) = \int_R x_i p(x_i;\hat{\boldsymbol{\theta}}) dx_i = \sum_{k=1}^{K} \pi_{ik}(\hat{\mathbf{w}})\hat{\boldsymbol{\beta}}_k^T \boldsymbol{r}_i$$

A sum of polynomials weighted by the logistic probabilities $\pi_{ik}(\hat{\mathbf{w}})$'s

$\Rightarrow$ Adapted for a smooth or abrupt transitions between the regression models.

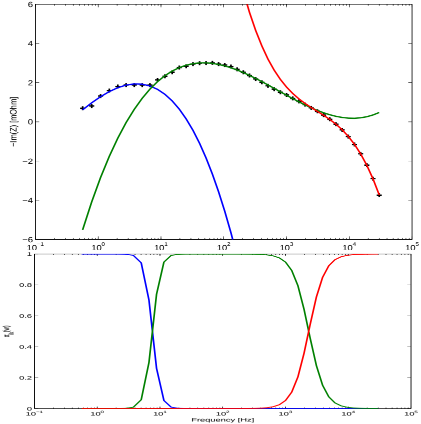
Segmentation

- The estimated class label $\hat{z}_i$ of x_i can be computed by the rule:

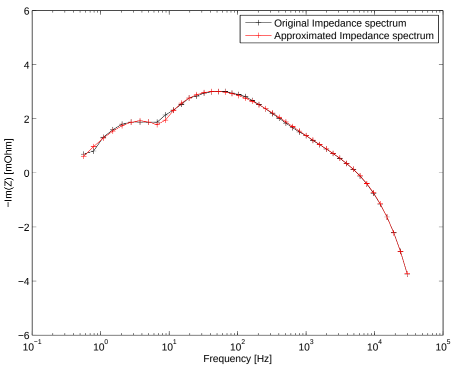
$$\hat{z}_i = \arg \max_{1 \leq k \leq K} \pi_{ik}(\hat{\mathbf{w}})$$

Case study:

- The impedance spectrums include 3 regimes which correspond to three behaviors of the stack → The number of regressive components is then set to $K = 3$
- The degree p of the polynomial regression is set to 3 which is adapted to the different regimes in the curves



The obtained approximation



linear regression model for the FC lifetime estimation

$LT_j = \alpha + \mathbf{b}^T \mathbf{y}_j + err_j$ for the EIS j where:

- LT_j : the duration time
- $\mathbf{y}_j = (\boldsymbol{a}_j, \boldsymbol{\theta}_j)^T$ features extracted from the real and the imaginary part
- $(\alpha, \mathbf{b})^T$ the vector of regression coefficients

linear regression model for the FC lifetime estimation

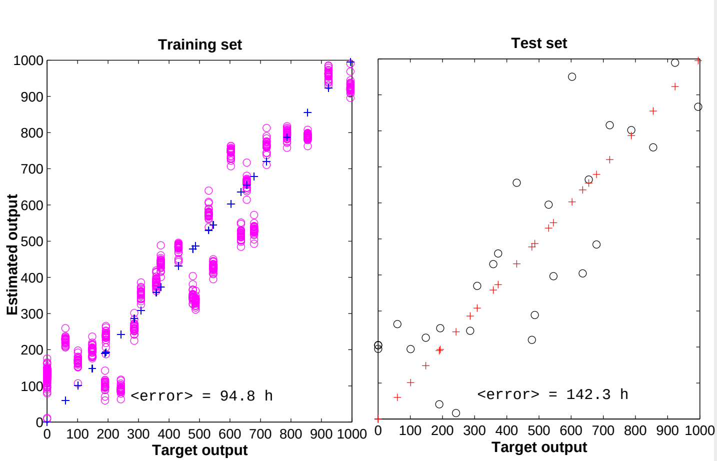
$LT_j = \alpha + \mathbf{b}^T \mathbf{y}_j + err_j$ for the EIS j where:

- LT_j : the duration time
- $\mathbf{y}_j = (\boldsymbol{a}_j, \boldsymbol{\theta}_j)^T$ features extracted from the real and the imaginary part
- $(\alpha, \mathbf{b})^T$ the vector of regression coefficients

Mean error (in hours) of duration time estimation using different input descriptors:

	Training set	Test set
Real part (dim=1) (a_2)	181.40	194.02
Imag. part (dim=3) ($\boldsymbol{\beta}_{23}, \boldsymbol{\beta}_{32}, \boldsymbol{\beta}_{34}$)	137.06	153.53
Real + Imag. parts (dim=7) ($\boldsymbol{\beta}_{21}, \boldsymbol{\beta}_{23}, \boldsymbol{\beta}_{24}, a_1, a_2, a_3, a_4$)	94.80	142.30

Duration time estimation obtained for the training set (left) and the test set (right):



Conclusion

- Supervised learning approach for Fuel Cell lifetime estimation from EIS measurements
- A probabilistic approach is used for feature extraction (from the imaginary part of the EIS)
  - The proposed model integrates a logistic process which makes possible to change smoothly within various possible regression models
  - Accurate modeling of the nonlinearities within the curves
  - Allows for automatically finding the three regimes corresponding to the behaviours of the stack

Thank you!