ANOVA METHOD APPLIED TO PEMFC AGEING FORECASTING USING AN ECHO STATE NETWORK

S. Morando[1,2], S. Jemei[1,2], R. Gouriveau[2,3], N. Zerhouni[2,3], D. Hissel[1,2]

[1] *FEMTO-ST (UMR CNRS 6174) Energy department*

[2] *FC-LAB (FR CNRS 3539), Techn'Hom, 90010 Belfort Cedex, France*

[3] *FEMTO-ST (UMR CNRS 6174) AS2M department*

Abstract - According to the International Energy Agency, an increase of the requests of energy of 40% could arise in the next decades, mainly due to the emergence of developing countries. The problem with the nowaday energy system is the use of fossil energy, which is limited and attempt to disappear in the near future. Thus an energy transition has to begin in order to replace the fossil fuels and anticipate their disappearance. Consequently, in recent years, the promotion and development of renewable energy have been realized. One of this renewable energy, the energy vector hydrogen, appears to be a promising solution, mainly due to interesting performance of Fuel Cells (FC) systems and hydrogen abundance on Earth (it is still important to underline that the hydrogen does not exist in natural form). However, this research area is still subject to scientific and technological bottlenecks. One of these major bottlenecks preventing the industrialization of FC systems is it limited useful lifetime. It is therefore important to develop reliable tools for the diagnosis and prognosis of FC system in order to optimize its efficiency. The aim of this article is to present the results of a sensibility analysis applied to a prognosis tools called Echo State Network.

Keywords – Analysis of Variance, Echo State Network, PEMFC, Ageing forecasting

1. INTRODUCTİON

The Prognostic and Health Management (PHM) is a discipline involved in the process of industrial maintenance. The objective in PHM is to estimate the Remaining Useful Life (RUL) of a system by making prediction on this future behavior. Using this RUL it will be possible to know the time and the relevant part of the system where a fault will be occurred in order to to perform a preventive maintenance to avoid this degradation. This theme becomes important in industrial research as mentioned in recent papers on the Condition Based Maintenance (CBM) [1-4].

Three main prognosis approaches can be distinguished [3,5]: model-based [6], data-based [7,8] and hybrid method (which is a combination of model-based and data-based). Data-based methods such as Artificial Neural Network (ANN), aims to estimate the ageing behaviour of the process without the need of all knowledge about the system physical phenomenon. Neural networks are in fact parsimonious universal approximators. Unfortunately, this kind of approaches requires a huge number of experimental data and consequently the deployment of such algorithms can be demanding in terms of calculation time, mainly due to the "trial and error" learning method, which is a real problem for industrial applications, thus it requires the creation of real-time algorithms. Among the various methods of this area, the tool chosen here is called Echo State Network (ESN). An ESN consists in the use of a dynamical neurons reservoir where the training step consists in performing a linear regression. The computation time of this algorithm is thus shorter while maintaining an efficient model. Created in 2001, ESN propose a better human brain paradigm than traditional ANN, and are based on a reservoir of neurons randomly connected each other. Nevertheless, the parameter definition is one of the main bottlenecks of this kind of tools consequently the aim is to clearly define an automatic parameter design for this tool applied to the prognostic of FC system. The first step of this heuristic creation is to apply a sensibility analysis to the ESN parameters.

This article will be divided in three main parts. The first part consists in a brief presentation of the tool used, and then the ANOVA (ANalysis Of VAriance) method will be presented in a second part. As all the methods will be presented at this point, the last part concerns the realized experiments and the results.

2. ECHO STATES NETWORK AS A NEW RNN PARADİGM

2.1. BACKGROUND OF ESN

The Reservoir Computing is a specific architecture Recurrent Neural Network (RNN) using a reservoir of neurons enabling a better paradigm of the human brain. Created in the early 2000s with the work of Jaeger [9,10] on Echo State Network, the generic term of Reservoir Computing appeared only in 2006 while new forms of Reservoir neural network have already been developed: the ESN, the backpropagation-decorrelation (BPDC) [11] and Liquid-State machine [12]. The interesting performance of this new generation of neural networks have attracted considerable interest from the other researchers interested in making improvements to the structure of this powerful tool. Consequently some ESN extensions can now be found in the literature such as les decoupled ESN [13], Evolino [14], leaky integrator [15] and more recently the ESQN (for Echo State Queuing Network) [16]. Other researchers studies about the ESN suggests some improvements on the reservoir, and more precisely on reservoir neurons such as replace them with infinite impulse response filters neurons [17], or adapt the wavelet neurons ESN [18] the interest would reside here in the combination of the advantages of ESN and wavelet transform [19,20] .

In term of applications, ESN can be found in various domains such as:

- Medical domain, for example with the prediction of the patients dialysis needs [21], where the ESN are compared to support vector machines (SVM [22]) and naive Bayesian classification.

-Economic sciences, especially in order to predict the evolution of the stock price in short term [23].

-In the physical domain, in the theme of energy with short-term prediction of an electrical load [24]. ESN are also used in optics, with the work of the Mr. Larger team [25]. Nevertheless, it is in signals recognition than the more ESN applications can be found as classifier in speech recognition [26] or in grammatical structure [27].

-Obviously in PHM with the comparison of two different structures of ESN applied to the Data Challenge PHM 2008 [28].

2.2. *How does ESN works?*

According to [16], an ESN can be divided into three main identities: the first is the input layer, where neurons receive information from their environment; the second is the reservoir of neurons. This reservoir, with recurrence inside, makes it possible a kind of "expansion" and memory effect on input into a larger space. From this perspective, the idea of Reservoir is similar to the expansion function used in the Kernel methods, such as in Support Vector Machine [16,29]. This projection of the input via the reservoir improves the linear separation of the data. The last part is the output weight matrix (calculation detailed in 2.3).

It is important to clarify some notations. N_{res} is the number of neurons in the reservoir, K the number of input(s) and L the number of output(s). These three parameters are useful for the matrix description:

-W_{inp} is the weight matrix between input layer and the reservoir (N_{res} rows, K columns),

-W_{res} is the reservoir (N_{res} rows, N_{res} columns),

-W_{out} is the weight matrix between the reservoir and the output layer (L rows and (N_{res} +K) columns),

-W_{feed} is the output feedback weight matrix (N_{res} rows L columns).

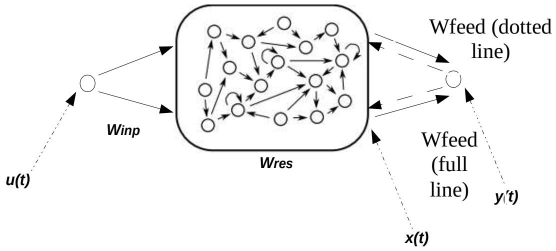

Fig.1. Basic Structure of ESN (Wfeed is in full line and Wout in dotted line).

2.3. *Learning algorithm*

The learning algorithm of an ESN is divided into two main steps.

Firstly, the ESN have to be trained with the « training data set », and it begins with the calculation of the reservoir update $\tilde{x}(n)$ as following in equation (1) [30]:

$$\tilde{x}(n) = f(W_{inp}.u(n) + W_{res}.x(n-1)) \quad (1)$$

With $\tilde{x}(n)$ the reservoir update, $u(n)$ the ESN input and $x(n-1)$ the reservoir output previous value.

This reservoir update calculation makes it possible to compute its output. This calculation involves the echo state property α in equation (2):

$$x(n) = (1-\alpha).x(n-1) + \alpha.\tilde{x}(n) \quad (2)$$

Then the ESN output is calculated using the results previously obtained using equation (2):

$$y(n) = W_{out}.x(n) + W_{feed}.y(n-1)) \quad (3)$$

However W_{feed} matrix is optional so the calculation of the ESN output given by equation (3) can be simplified by equation (4):

$$y(n) = f(W_{out}.x(n)) \quad (4)$$

Secondly, the learning algorithm is applied in order to reduce the Mean Square Error (MSE) between the target values (here denoted y_{target}) and the ESN output (denoted $y_{predicted}$).

$$MSE = \frac{1}{Nd}.\sum_1^{Nd}(y_{target}(n) - y_{predicted}(n))^2 \quad (5)$$

Which is equivalent to:

$$MSE = \frac{1}{Nd}.\sum_1^{Nd}(y_{target}(n) - f(W_{out}.x(n)))^2$$

(6)

With Nd the number of studied samples, that means the samples total number minus the number of dismissed samples due to the initial condition of the different matrix. The goal is now to find the best W_{out} weights matrix corresponding to the lowest MSE possible result, achieved by a linear regression, formula in equation (7):

$$W = (A^T.A - \lambda.I)^{-1}.A^T.B \quad (7)$$

In this case, W represents W_{out}, A represents the output of the reservoir and B the target output signal.

3. ANOVA METHOD APPLİED TO ESN İN THE CASE OF PEMFC AGEİNG FORECASTİNG

3.1. ANOVA METHOD

The use of ESN involves a question of the setting tool. The question at issue now is how, according to the signal under study, set the ESN to have the best possible results directly. The first step to resolve this problem is to perform a parametric sensitivity analysis, making it possible to quantify the impact of the change in model parameters on the model output. The major interest in this case is to prioritize the different parameters of the ESN (the studied parameters will be briefly described in the experiment section).

Created by G. Taguchi in the 50s [31-33], ANOVA is a collection of statistical models and procedures for simultaneous comparisons between several medium to determine meaningful relationship exists between variables. This method has been repeatedly applied to the field of fuel cells [34, 35].The principle is to perform a series of tests of an experimental design. For example, consider a system with two parameters A and B, and having one output Y. The experimental plan is to compare the different values of the output Y for each combination of A and B possible. In this example , A and B have two levels , low and high , giving four possibilities as shown in the following table showing the experimental plan :

I: Example of an Experimental Plan

Level of Parameter A	Level of Parameter B	Output value
Low	Low	Y1
Low	High	Y2
High	Low	Y3
High	High	Y4

From the results obtained in the previous table, the global average is calculated (in equation (8)):

$$\bar{Y} = \frac{Y1+Y2+Y3+Y4}{4} \quad (8)$$

Then each parameter level variance is calculated as following:

$$Var(Y_a) = \frac{Y1+Y2}{2} - \bar{Y} \quad (9)$$

$$Var(Y_A) = \frac{Y3+Y4}{2} - \bar{Y} \quad (10)$$

$$Var(Y_b) = \frac{Y1+Y3}{2} - \bar{Y} \quad (11)$$

$$Var(Y_B) = \frac{Y2+Y4}{2} - \bar{Y} \quad (12)$$

Respectively, $Var(Y_a)$ represents the parameter A variance at the low level and $Var(Y_A)$ the parameter A mean at the high level. It is the same principle for $Var(Y_b)$and $Var(Y_B)$ for the parameter B. Then the Total Square Summation (TSS) and the Difference Square Summation (DSS) are computed (i is the parameter studied A or B):

$$TSS = \sum Y_i^2 - \frac{\sum Y^2}{experience\ total\ number} \quad (13)$$

$$DSS_i = \frac{experience\ total\ number}{level\ number} \cdot \sum \bar{Y_\iota} - \bar{Y}^2 \quad (14)$$

Then each primary influence is calculated as follows:

$$Influence_i = \frac{DSS_i}{TSS} \quad (15)$$

3.2. PARAMETERS STUDİED

- The number of reservoir neurons N_{res} is one of the most important parameters. To find a better linear combination of the signals in order to find the best target, it is better to define a large reservoir (it is not uncommon to find a reservoir containing more than 10^4 neurons [36]).

- The reservoir connectivity c, which represents the percentage of non-zero weights in the reservoir. It can take values between 0 and 1.

- The spectral radius mathematically corresponds to the maximum value of this matrix eigenvalues. In ESN, the spectral radius is used to scale the non-zero elements of W_{res}.

- The Echo State Property α also called leaking rate. It is an important notion of ESN and it corresponds to the previous reservoir output importance, the echo, as shown in equation 2. Its value has to be determined in the range [0;1], and the more important this value is, the less important becomes the reservoir echo.

3.3. EXPERİMENTAL PLAN

ANOVA is a statistical tool, thus it was necessary to have enough. The ANOVA study was based on the following experimental design:

- Four parameters studied :

- The number of neurons in the reservoir
- The spectral radius of the reservoir matrix
- The value of the Echo State Property
- Connectivity of neurons in the reservoir

- Three levels with parameters:
  - Low (50 to the number of neurons, 0.2 for the other parameters)
  - Middle (250 for the number of neurons, 0.5 for the other parameters)
  - a high level (500 for the number of neurons, 0.8 for the other parameters)

Each experience is the result of a single simulation for a given set of parameters. Thus $N_{exp} = n_{niveaux}{}^{n_{paramètres}} = 3^4 = 81$ experiments have been realized. Each experiment corresponds to 100 simulation mean results therefore 8100 simulations were performed. The data come from the 1100 hours long duration test on fuel cells performed under the COCONPAC project. Only the stack is considered here (8 cells of 100 cm² active areas) and only the voltage is predicted. The nominal current density of the cells is 0.70 A/cm², the nominal temperature is 80°C. At time t= 0;48;185;348;515;658;823;991h, some characterizations have been realized (Polarization and Electrochemical Impedance Spectroscopy). Due to those characterizations, the signal must be pre-processed before. Here, the aberrant points are firstly deleted and then a moving average filtering have been applied (Figure 3).

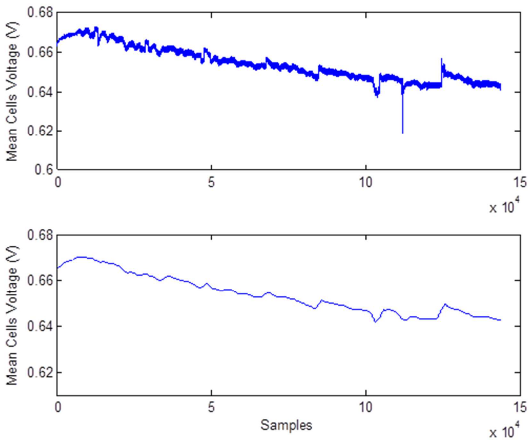

Fig.2.Data without aberrant points above and filtered data below

Each simulation consist in the forecast of the next 20 hours mean cells voltage as explained in Figure 2 (the sampling period is 30 seconds thus a 20 hours prediction represents a 2500 samples prediction). Only the stack with constant load is considerate here and the stack is supposed to be in healthy mod (only the natural ageing of the system is taken into account).

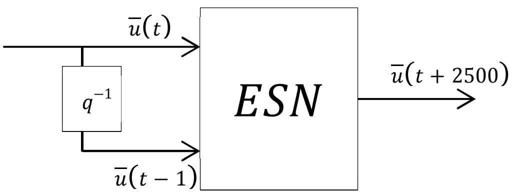

Fig.2.Forecasting with an ESN.

Then the results are collected and the Mean Average Percentage Error is calculated (equation 16):

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{\hat{y}_t}\right| \qquad (16)$$

The following Figure is an example taken randomly among all the simulations. The real mean cells voltage is in blue (also called the target) and the ESN output is in green (also called the estimation).

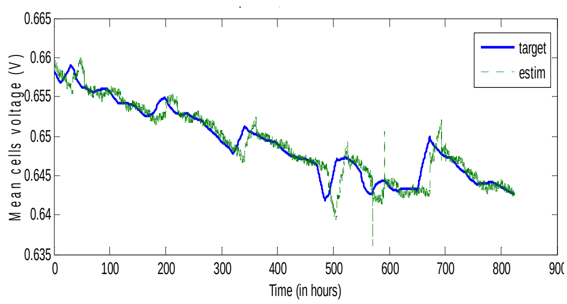

Fig.3.Example of forecasting result with a=0.5, SR=0.5, c=0.5 and N=50.

3.4. *Results*

The results of the ANOVA method are given in Table 2. The classification of the most influent parameters in the case of mean cell voltage prediction can be established. The most influential parameter for this signal is the spectral radius of the reservoir matrix with 68% influence on the result. This is justified by the "flooding" of input information during its diffusion through the reservoir in case of a too high spectral radius value. Spectral radius defines the weights scale, that is to say, the importance given to information that passes between two neurons, meaning that this parameter cannot be dissociated from the number of neurons in the reservoir. That is why the interaction of Neurons Numbers and Spectral Radius is so important. Reservoirs with too many neurons, which themselves have too important links between them will completely flood information and the outputs results will be inaccurate.

The Echo State Property has a low influence value. This can be explained by the use of an ESN structure with output feedback for the experiments. Moreover, this value is never equal to 0 or 1, which are the two extreme values of these parameters so there is always an echo of previous reservoir output in this case.

II: ANOVA study results (interactions between parameters included)

Parameters	Influence (in %)
Echo state Property	0,04
Number of neurons in W_{res}	9,90
Spectral Radius	68,13
Reservoir Conectivity	0,14
Interaction α / N	0,09
Interaction α / SR	0,08
Interaction α / c	0,11
Interaction N / SR	19,83
Interaction N / c	0,05
Interaction SR / c	0,29
TOTAL	98.66
Residual	1.34

Conclusion

It is important to underline that this article and these simulations do not improve the ESN theory and fuel cells systems; it is just an analysis of variance applied to a new concept of neural network in order to forecast the ageing of a Fuel Cell System.

The application of the ANOVA method is satisfactory because it remains only 1.34 % residue (maybe due to the random condition of W_{res}). Moreover, the influence of ESN parameters on a mean cells voltage forecasting result is now known in detail.

This application of the ANOVA has prioritized parameters of ESN and gives information about what parameters it will focus on the implementation of a heuristic for the automatic configuration of an ESN based on the processed signal. Some parameters must have a focus on them: the number of neurons inside the reservoir and their scaling value, the spectral radius of the reservoir weight matrix. At the opposite, some parameters can be unheeded such as connectivity, but only for this case. The aim now is to clearly define the link between these two most important parameters and the signal studied.

Acknowledgements

The used of COCONPAC (project funded by the CNRS) experiment data were greatly appreciated. This study has been realized thanks to the region of Franche-Comté funds.

References

[1] Ciarapica FE, Giacchetta G. Managing the condition based maintenance of a combined-cycle power plant:an approach using soft computing techniques.J of Loss Prevention in the Process Industries, vol. 19, 2006, pp. 316-325.

[2] Heng A,Zhang S,Tan A,Matwew J. Rotating machinery prognostic: State of the art, challenges and opportunities. Mechanical Systems and Signal Processing, vol. 23, pp. 724-739, 2009.

[3] A. K. S. Jardine, D. Lin, et D. Banjevic, « A review on machinery diagnostics and prognostics implementing condition-based maintenance », Mechanical Systems and Signal Processing, vol. 20, no 7, 2006, pp. 1483 -1510.

[4] M. Lebold et M. Thurston, « Open standards for Condition-Based Maintenance and Prognostic Systems », presented at the 5th Annual Maintenance and Reliability Conference, 2001.

[5] G. Vachtsevanos, F.-L. Lewis, M. Roemer, and A. Hess, Intelligent Fault Diagnosis and Prognosis for Engineering Systems. New Jersey, Hoboken: Wiley & Sons, 2006

[6] C. S. Byington, M. Watson, M. J. Roemer, T. R. Galic, et J. J. McGroarty, « Prognostic enhancements to gas turbine diagnostic systems », in 2003 IEEE Aerospace Conference, 2003. Proceedings, vol. 7, 2003, pp. 3247-3255.

[7] M. Karpenko et N. Sepehri, « A neural network based fault detection and identification scheme for pneumatic process control valves », in 2001 IEEE International Conference on Systems, Man, and Cybernetics, vol. 1, 2001, pp. 93-98.

[8] Wang W, "An adaptative predictor for dynamic system forecasting", Mechanical Systems and Signal Processing, vol.21, 2007, pp.809-823.

[9] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks", Technical Report GMD 148, German National Research Center for Information Technology, 2001.

[10] Jaeger, H. Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the echo state network approach. *Technical report*. GMD—German National Research Institute for Computer Science, 2002.

[11] J. J. Steil, "Backpropagation-decorrelation, Online recurrent learning with O(N) complexity", Proceedings of the International Joint Conference on Neural Networks, vol. 2, 2004, pp. 843–848.

[12] W. Maass, T. Natschlager and H. Markram, "Real-time computing without stable states : A new frame-work for neural computation based on

perturbations", Neural Compututations, vol.14, 2002, pp. 2531-2560.

[13] B. Zhang et Y. Wang, « Echo state networks with decoupled reservoir states », in IEEE Workshop on Machine Learning for Signal Processing, MLSP 2008, 2008, pp. 444‑449.

[14] J. Schmidhuber, D. Wierstra, M. Gagliolo, and F. Gomez, "Training recurrent networks by evolino," Neural Comput., vol. 19, 2007, pp. 757–779.

[15] H. Jaeger, M. Lukosevicius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky-integrator neurons," Neural Network, vol. 20, 2007, pp. 335–352.

[16] S. Basterrech et G. Rubino, « Echo State Queueing Network: A new reservoir computing learning tool », in 2013 IEEE Consumer Communications and Networking Conference (CCNC), 2013, pp. 118‑123.

[17] G. Holzmann et H. Hauser, « Echo state networks with filter neurons and a delay&sum readout », Neural Networks, vol. 23, 2010, pp. 244-256.

[18] A. Deihimi, O. Orang, et H. Showkati, « Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction », Energy, vol. 57, 2013,pp. 382‑401.

[19] D. Jonathan, B. Michael et F. Sébastien. *Les Ondelettes.* Université Libre de Bruxelles. Printemps des Sciences. Bruxelles. Belgique.

[20] I. Daubechies. *Ten Lectures of Wavelets*. Rutgers University and AT & T bell Laboratories, Sept.1992.

[21] Ongenae, F., et al., Time series classification for the prediction of dialysis in critically ill patients using echo state networks. Eng. Appl. Artif. Intel. ,2012.

[22] Cortes, Corinna; and Vapnik, Vladimir N.; "Support-Vector Networks", Machine Learning, vol.20, 1995

[23] X. Lin, Z. Yang, et Y. Song, « Intelligent stock trading system based on improved technical analysis and Echo State Network », *Expert Systems with Applications*, vol. 38, 2011, pp. 11347-11354.

[24] M. J. A. Rabin, M. S. Hossain, M. S. Ahsan, M. A. S. Mollah, et M. T. Rahman, « Sensitivity learning oriented nonmonotonic multi reservoir echo state network for short-term load forecasting », in *2013 International Conference on Informatics, Electronics Vision (ICIEV)*, 2013, pp. 1-6.

[25] L. Larger et J. M. Dudley, « Nonlinear dynamics: Optoelectronic chaos », *Nature*, vol. 465, 2010, pp. 41-42.

[26] M. D. Skowronski et J. G. Harris, « Automatic speech recognition using a predictive echo state network classifier », Neural Netw, vol. 202007, p. 414‑423.

[27] M. H. Tong, A. D. Bickett, E. M. Christiansen, et G. W. Cottrell, « Learning grammatical structure with Echo State Networks », Neural Netw, vol. 20, 2007, p. 424‑432.

[28] Yu Peng, Hong Wang, Jianmin Wang, Datong Liu, Xiyuan Peng "A Modified Echo State Network based Remaining Useful Life Estimation Approach", Proceedings of PHM 2012, Denver CO, July 2012

[29] B. Schrauwen, D. Verstraeten, et J. V. Campenhout, « An overview of reservoir computing: theory, applications and implementations », in *Proceedings of the 15th European Symposium on Artificial Neural Networks*, 2007, pp. 471–482.

[30] Lukoševičius, M., "A practical guide to applying echo state networks", Neural Networks: Tricks of the Trade, vol. 7700, 2012, pp. 659-686.

[31] Fowlkes WY, Creveling CM. Engineering methods for robust product design: using Taguchi methods in technology and product development. Reading. MA: Addison-Wesley; 1995.

[32] Lochner RH, Matar JE. Designing for quality: an introduction to the best of Taguchi and western methods of statistical experimental design. Milwaukee: WI: ASQC Quality; 1990.

[33] Ross PJ. Taguchi techniques for quality engineering. 2nd ed. McGraw-Hill; 1995.

[34] K.-Y. Chang, « The optimal design for PEMFC modeling based on Taguchi method and genetic algorithm neural networks », *International Journal of Hydrogen Energy*, vol. 36, 2011, pp. 13683-13694.

[35] B. Wahdame, D. Candusso, et J.-M. Kauffmann, « Study of gas pressure and flow rate influences on a 500 W PEM fuel cell, thanks to the experimental design methodology », *Journal of Power Sources*, vol. 156, 2006, pp. 92-99.

[36] F. Triefenbach, A. Jalalvand, B. Schrauwen, et J.-P. Martens, « Phoneme recognition with large hierarchical reservoirs », in Advances in Neural Information Processing Systems, vol. 23, 2010, pp. 2307-2316.