A fuzzy logic and a neural network based wind power forecasting model

Georges Kariniotakis, Eric Nogaret, Georges Stavrakakis

▶ **To cite this version:**

Georges Kariniotakis, Eric Nogaret, Georges Stavrakakis. A fuzzy logic and a neural network based wind power forecasting model. 1996 European Union Wind Energy Conference, EUWEC'96, EWEA, May 1996, Göteborg, Sweden. pp.596-599. hal-04268994

HAL Id: hal-04268994

https://minesparis-psl.hal.science/hal-04268994

Submitted on 3 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1996 European Union Wind Energy Conference,
20-24 May 1996, Göteborg, Sweden.

A FUZZY LOGIC AND A NEURAL NETWORK BASED WIND POWER FORECASTING MODEL.

G. Kariniotakis, E. Nogaret

Ecole des Mines de Paris,
Centre d'Energétique,
06904 Sophia-Antipolis, France.

G. Stavrakakis

Technical University of Crete,
Dept. of Electronic & Computer Eng.,
73100 Chania, Crete, Greece.

ABSTRACT - In this paper, models based on fuzzy logic and recurrent high-order neural networks are developed for forecasting the power output of a wind park. These models appear several advantages over conventional models ; e.g. they permit to consider easily explanatory variables like wind speed, direction etc. Adaptive learning laws are derived for the estimation of the various parameters. The architecture of the models is optimised automatically using a non-linear optimisation technique that aims to substitute the usually applied trial-and-error method. The developed models have been tested using various time-series of wind power where they outperform persistence as well as conventional models. Finally, the paper presents issues on the on-line implementation of a forecasting model into an advanced control system for the optimal operation and management of a real autonomous wind-diesel power system.

Keywords: *Forecasting methods, Wind-Diesel Systems, Neural networks, Fuzzy Logic.*

1. INTRODUCTION

The integration of wind energy conversion systems (WECS) in power systems and especially in isolated ones, like the systems of islands or the ones of rural areas, may result in important environmental and economic benefits. However, the stochastic nature of the wind, imposes considerable difficulties on the optimal management of these systems.

The consideration of efficient wind power forecasts by the power system control and management functions may result in an improvement of the quality of service and a reduction of operating costs [2].

In the literature, various models, have been developed for wind speed or power prediction. The time scale of these models ranges from some seconds, when the objective is wind turbines control, to some minutes, or even hours, when the objective is economic dispatch and power system planning. In smaller time scales, the time-series methodology is generally applied, according to which, past values of the process are used to predict its future evolution. In higher time-scales, of some hours or more, models incorporating meteorological information are usually used. These models produce forecasts a limited number of times per day and necessitate more parameters and considerable computer time compared to time-series models.

In this paper, the time-series approach is followed in order to develop artificial intelligence based models for forecasting the power output of a wind park. Namely, models based on recurrent neural networks and fuzzy logic are built. The paper is mainly concerned with forecasts useful for the short-term scheduling of an isolated power system that is, with forecasts of the WECS power output profile for the next 2 or 3 hours with a time step in the order of 10 minutes

The time-series models found in the literature do not provide a significant improvement over the method of *persistence*. As a consequence, the latest one, being extremely simple (*the wind in all future time-steps will be equal to the wind now*), is almost exclusively used in the applications.

In [4], ARMA models for wind speed prediction achieve an improvement with respect to the persistence in the order of 5-12 % on 2-second data (horizon up to 20 seconds), and 9-14 % on 1-minute data (horizons up to 10 minutes) relative to the root mean square (RMS) criterion of errors. In previous work of the same author based on Kalman filter techniques, the RMS improvement on 1-minute data, is in the order of 4 to 10 %. The performance of the method becomes very small or negative when 10-minute averaged data are used. In [5], multi-layer feed-forward neural networks and radial basis functions are applied for wind power prediction. The improvement with respect to persistence is in the order of 11 % for one-step predictions of 1-minute data and 8 % for 10-min data. In [6], various models like ARMA and bilinear ones are developed for wind power. They achieve an improvement of 7 to 12 % with respect to the persistence, for a forecast horizon of 2 hours and time-steps of 30 min. Finally, statistically evaluated results on multi-step ahead forecasting of time-series with a time-step in the order of 10 min. are not met in the literature.

2. THE FORECASTING METHODOLOGY

Two main approaches may be followed in order to generate wind power forecasts :

(i) to develop an explicit prediction model for wind power, in which it will be possible to consider wind speed, wind direction etc. as explanatory variables.

(ii) to develop a prediction model for wind speed and a second model for the transformation of wind speed to power. As transformation model can be used the manufacturer's characteristic curve of the wind turbine power output as a function of wind speed. If however, the point of measurement of wind speed is different than the hub height of a wind turbine - as is the case in the case-study of the paper, then the transformation model should account for the spatial variations of the wind speed.

In the case-study of Paragr. 3, the consideration of the characteristic curve as a transformation model gave poor results, while an advanced transformation model did not give any advantages over the first approach. Similar results are also reported in [6].

2.1 The "Naive" Predictors

When an advanced forecasting model is not available, forecasts may be obtained with a minimal effort and data manipulation and can be based solely on the most recent information available. Such forecasts are referred to as *Naive* forecasts. One such method (*Persistence* or *Naive_1* method), is to use the most recent datum available $P(t)$, as forecast $\hat{P}(t+k/t)$ for each one of the future time-steps.

The forecasting errors over a set of data, are used to calculate various performance criteria like the RMS both for Naive and advanced methods. The benefit gained by the use of an advanced method is typically measured as a *percentage improvement* on a certain criterion (e.g. RMS) [1-6].

2.2 The Artificial Neural Network Model

In the last years, increasing interest has been shown for the use of recurrent neural networks for time-series forecasting where they are found to outperform the classical feed-forward neural networks (FNN). Recurrent networks dispose dynamic elements in the form of feedback connections. This distinguishes them from FNN, where the output of one neuron is connected only to neurons in the next layer. Here the class of *recurrent higher-order neural networks* (RHONN) is considered.

A RHONN has only one layer with n neurons, each one giving a forecast $\hat{y}_i$ for a different time-step of the forecast horizon. The input vector $\boldsymbol{u}$ consists of the n neurons outputs (recurrence) plus the m external inputs (e.g. past values of wind power, speed, direction etc.) :

$$\boldsymbol{u} = \left[u_1,\ldots u_n, u_{n+1},\ldots u_{m+n}\right]^T \triangleq \left[S(\hat{y}_1),\ldots S(\hat{y}_n), S(x_1),\ldots S(x_m)\right]^T$$

As remarked from the definition of $\boldsymbol{u}$ each input is passed though the function $S(.)$ which is a sigmoidal function :

$$S(x) = \alpha \frac{1}{1+e^{-\beta x}} - \gamma \qquad (1)$$

where α, β are positive real numbers and γ is a real number.

A RHONN is linear in the weights and allows higher-order interactions between neurons. E g. in a recurrent *second-order* NN, the total input to the neuron is not only a linear combination of the components u_j, but also of their products $u_j u_k$. Moreover, one can pursue along this line and include higher-order interactions represented by triplets $u_j u_k u_l$, quadruplets, etc. Each first or higher order term can be represented by z_k, which is defined as: $z_k \triangleq \prod_{j \in I_k} u_j^{d_j(k)}$, where $\{I_1, I_2, \ldots, I_L\}$ is a collection of L not-ordered subsets of $\{1,2,\ldots,m+n\}$, and $d_j(k)$ are non-negative integers.

Following the above definitions the global form of the model is given by the following difference equation :

$$\hat{y}_i(t+1) = a_i \hat{y}_i(t) + b_i \sum_{k=1}^{L} w_{ik}(t) z_k(t) \qquad (2)$$

where $\hat{y}_i$ is the state of the i-th neuron (i=1,…,n), a_i, b_i are constants, w_{ik} is the synaptic weight connecting the k-th input to the i-th neuron.

The stochastic gradient method is used to derive the following learning rules for the weights estimation :

$$w_{ik}(t+1) = w_{ik}(t) + \eta_i z_k(t) e_i(t) \qquad (3)$$

where $e_i(t) \triangleq \hat{y}_i(t) - y_i(t)$ denotes the prediction error, $y_i(t)$ is the measured value of the process and η_i is a small positive parameter denoting the *learning rate*.

Since the RHONN model is linear in the weights the learning process leads to globally optimal values of them.

2.3 The Fuzzy Logic based model.

In recent years the construction and use of fuzzy models has been accepted as a useful method when confronted with dynamic systems whose underlying dynamics are unknown, as is the case with wind, or too complex for analysis by conventional mathematical methods.

Being in the form of IF-THEN rules, a fuzzy model can be used as an explanatory model for the process itself whose underlying dynamics are unknown. In addition constraints on the operation of a wind park, like cut-in/out speed, may be easily considered. This is an advantage compared to neural networks models, which are generally understood as black-box models. A fuzzy model describes locally a dynamic process since at each rule a different model of the process is associated and a rule is activated only for certain regions of values of the input variables. This aspect was indeed a motivation in our application. Several characteristics of the wind, like non-stationarity and seasonallities, are favourable factors to apply local modelling.

In the literature several different types of fuzzy models are found. Here the Sugeno and Tagaki model [7] is used. It consists of a set of r fuzzy rules of the type :

$$\begin{array}{lll} R^1: & \text{IF } \left(x_1 \text{ is } A_1^1, \ldots, \text{ and } x_m \text{ is } A_m^1\right) & \text{THEN } \hat{y}^1 = g^1(x_1,\ldots,x_m) \\ \vdots & & \vdots \\ R^r & \text{IF } \left(x_1 \text{ is } A_1^r, \ldots, \text{ and } x_m \text{ is } A_m^r\right) & \text{THEN } \hat{y}^r = g^r(x_1,\ldots,x_m) \end{array}$$

where :

$x_1,\ldots,x_m$ are input variables of the system.

A_j^i is a fuzzy set [7] characterised by the membership function $\mu_i^j(x_j)$ with i=1,…,r and j=1,…,m.

$\hat{y}^i$ is the output of the i-th rule. The function g^i is a linear or non-linear function that implies the value of $\hat{y}^i$ when $x_1,\ldots,x_m$ satisfy the premise. Here, a linear function of the form $g^i(\boldsymbol{x}) = p_0^i + p_1^i x_1 + \ldots + p_m^i x_m$ is considered.

Since the variables $x_1,\ldots,x_m$ are past values (and eventually errors or exogenous variables) of the process, it is obvious that the consequent part of each rule is an ARMAX model.

The notion of fuzzy set is fundamental in fuzzy logic. If x_i represents "*wind speed*", then a number of fuzzy sets,

representing vague terms like "*small*", "*medium*", or "*large*" may be associated to this variable, as shown in Fig. 1.

Here, fuzzy sets are represented by Gaussian functions which are defined by :

$$\mu_{A_j^i}(x_j) = \exp\left(-\left(\frac{x_j - a_j^i}{b_j^i}\right)^2\right) \quad (4)$$

where $b_j^i > 0$ is the center value and $a_j^i \in \Re$ the dispersion of the Gaussian membership function.

When an input $\boldsymbol{x} = [x_1, x_2, \ldots, x_m]^T$ is presented to a rule *i*, the values of the membership functions are calculated. The product of these values gives the activation degree - or *weight* w^i - of the rule. When, for a given input, more than one rules

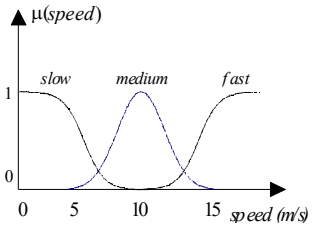

Figure 1 : Representation of fuzzy speeds. "Speed" is a linguistic variable with three terms "slow", "medium", and "fast" represented as fuzzy sets with the membership functions shown in the Figure.

are activated, the global output $\hat{y}$ of the model is the weighted average of the outputs $\hat{y}^i$ of the rules.
The fuzzy model may be written analytically as following :

$$\hat{y} = \frac{\sum_{i=1}^{r}\left(p_0^i + \sum_{j=1}^{m} p_j^i x_j\right)\cdot\prod_{j=1}^{m}\mu_{A_j^i}(x_j)}{\sum_{i=1}^{r}\prod_{j=1}^{m}\mu_{A_j^i}(x_j)} \hat{=} \sum_{i=1}^{r} w^i \hat{y}^i \quad (5)$$

The following learning rules are derived according to the stochastic gradient scheme. These rules are used to tune all the parameters of the model :

$$p_j^i(t+1) = p_j^i(t) - \eta_p \cdot w^i(t) \cdot \chi_j \cdot e(t) \quad \chi_j = \begin{cases} 1 & \text{if } j = 0 \\ x_j & \text{if } j > 0 \end{cases}$$

$$a_j^i(t+1) = a_j^i(t) - \eta_a \cdot w^i \cdot \left(\hat{y}^i - \hat{y}(t)\right) \cdot \frac{2\left(x_j - a_j^i(t)\right)}{b_j^i(t)^2} \cdot e(t)$$

$$b_j^i(t+1) = b_j^i(t) - \eta_b \cdot w^i \cdot \left(\hat{y}^i - \hat{y}(t)\right) \cdot \frac{2\left(x_j - a_j^i\right)^2}{b_j^i(t)^3} \cdot e(t) \quad (6)$$

2.4 The learning procedure

The learning laws are applied at the presentation of each data to the model. The presentation of the whole data set (the one reserved for learning) is characterised as *learning epoch*. Usually, a certain number of *K* epochs is necessary for the convergence of the learning process. The value of the learning rate is adjusted at each epoch in order to avoid residual fluctuations or instability for both neural and fuzzy types of models. Here, an adaptive *learning rate scheme* (LRS) is followed according to which the learning rate is reduced by a factor ρ_D if the sum of square prediction errors (*SSE*) during the *k*-th epoch has increased relatively to that of *k*-1. In the contrary case , the learning rate increases by a factor ρ_I :

$$\eta(k+1) = \begin{cases} \rho_I \cdot \eta(k) & \text{if } SSE(k) < SSE(k-1) \\ \rho_D \cdot \eta(k) & \text{if } SSE(k) \geq SSE(k-1) \end{cases} \quad (7)$$

where $\rho_D \leq \rho_I$ and $\rho_D < 1$. The initial learning rate value is η_0. If $\rho_D = \rho_I < 1$, then a LRS of "*search-then-converge*" type is obtained [1].

2.5 Modelling the wind

Consider now that there is a dynamic process, whose input/output relation is given by a general difference equation of the form :

$$\boldsymbol{y}(t+1) = f\left(\boldsymbol{y}(t), \boldsymbol{x}(t)\right) \quad (8)$$

where $\boldsymbol{y} \in \Re^n$ is the scalar process output, $\boldsymbol{x} \in \Re^m$ is the scalar input of the process and $f(.)$ is an unknown function. Models (2) and (5) are proven to be universal approximators of systems described by (8) if the later satisfy several mild conditions [1].

Considering the problem of the wind power prediction, it is assumed that the dynamic process $\boldsymbol{y}$, is the future wind power profile. The objective is then, to find a neural network or fuzzy model $f_M(.)$ that can approximate $f(.)$.

2.6 Optimisation of the forecasting model architecture

Usually in practise, the *trial-and-error* method is applied to test various alternative model architectures and choose the one with the optimal generalisation capability. As *generalisation,* is defined the capability of a model to predict data other than those on which it has been trained.

The architecture parameters that affect generalisation are the types of data used as input to the model (speed, direction...), the number of past values of each kind of data, and the parameters of the learning process - η_0, ρ_D, ρ_I, *K*. Concerning fuzzy models, the number of fuzzy sets associated to each input variable must also be estimated. The above parameters are optimised according to the cross-validation scheme and by using an algorithm based on the non-linear Simplex method of Box. The method has been adapted to optimise both discrete (e.g. number of past data) and continuous variables (e.g. learning rate) [1].

3. EXPERIMENTAL RESULTS & DISCUSSION

The case-study of the wind-diesel system of the Greek island of Lemnos, where 2 wind parks of a total capacity of 1.14 MW are installed [8 WTs of 55 kW, 7 WTs of 100 kW], is considered. The short-time scheduling of the above power system is performed by an advanced control system described in [2], and necessitates forecasts of the average wind turbines power of each wind park for an *horizon of 2 hours* with a *time-step of 10 minutes* [2].

A set of real data coming from the data acquisition system of the island - see Fig. 2 - were used to estimate the parameters of the models, while a second independent set of data - *test set* (on the right of the dashed line of Fig. 2) - was used to evaluate the forecasting models.

Three basic ANN configurations have been tested : (i) a network with 12 outputs, one for each time-step, denoted as ANN-a. (ii) A network with one output, which is used iteratively to give forecasts for all the 12 time-steps, denoted as ANN-b. It's architecture has been optimised by considering the forecast errors of all the time-steps. (iii) A network similar to the previous one, denoted as ANN-c, which has been optimised by considering only the errors of the first time-step. High-order terms up to two were considered in all cases, while the neurons are fully interconnected via recurrent links. The configuration of the fuzzy (FL) model is similar to ANN-b. The number of past values for each type of model where found to be :

	ANN-a	*ANN-b*	*ANN-c*	*F-L*
Wind power	2	8	6	6
Wind speed	1	3	5	0

The improvement on the RMS criterion gained by the four models over Persistence is given in Fig. 3. From the use of the advanced models there is a clear gain concerning

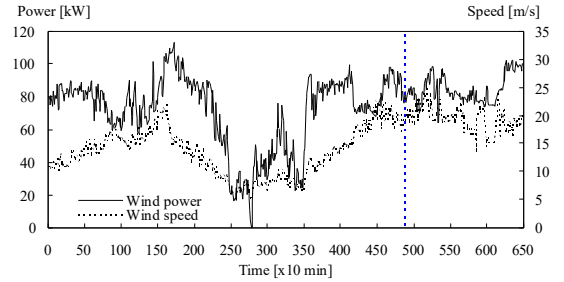

Fig. 2 : Average wind turbines power and wind speed data from the 2nd wind park of Lemnos (7x100kW).

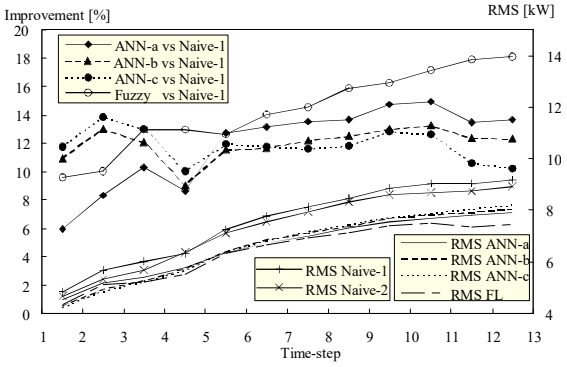

Fig. 3: RMS error of various models and improvement of ANN and Fuzzy logic (FL) models over simple methods.

multi-step forecasts. ANN-a and ANN-b have a better performance than ANN-c on forecasts of longer time-steps. This is because in the architecture optimisation of these networks their performance in the whole horizon has been considered. Concerning one step-ahead forecasts, a better performance is achieved by ANN-c, which has been exclusively built for such forecasts. Although better results can be expected if a different model is built for each time-step, this approach is undesirable from the on-line implementation point of view.

The fuzzy model was found to outperform any other type of model tested. The amelioration over Persistence on the RMS criterion was between 9% and 18% according to the time step - see Figure 3.

Finally, it is remarked that the length of the considered time-series is certainly small for a complete statistical evaluation of the developed models. However, this case-study, has been selected for presentation here due to the on-line implementation of the developed models as described in the next Paragraph. In addition, from tests with longer time-series from other sites, it was concluded that the above results can be considered as representative for the level of performance that can be obtained by the proposed models.

3.1 Implementation in the existing Control System

An advanced control system (CS) for the optimal operation and management of wind-diesel power systems has been developed in the frame of the EU project JOU2-CT92-0053 [2, 3]. The control system has been installed and is under evaluation in the wind-diesel power system of the island of Lemnos.

The CS aims to assist the power system operators by proposing optimal scenarios for the power system operation, so that maximum fuel saving is achieved, without deterioration of the quality of service to the consumers. The scenarios are generated by an economic dispatch module by taking into account load and wind power forecasts. Wind power forecasts are provided by a wind power forecasting (WPF) module. Various Naive predictors, as well as advanced models, have been integrated in the WPF module. The ANN-a RHONN model was chosen for implementation due to its globally better performance compared to ANN-b and ANN-c (the fuzzy model has been developed later than this implementation).

The utilisation of advanced forecasts by the control system, results (i) in a decrease of the total number of start/stops of the diesel units, which may arise up to 30 %; (ii) in a decrease of the loss of load events; (iii) in an increase of fuel savings; (iv) and finally in a higher utilisation of the available wind energy [2, 3].

4. CONCLUSIONS

In the present paper, artificial intelligence techniques, and namely adaptive neural networks and fuzzy logic based models, have been developed for time-series forecasting of wind power. Special attention has been given to the optimisation of the architecture of the forecasting models. Both types of models outperform Persistence, while the obtained results are superior to similar results of the known alternatives. However, the fuzzy model was found to have the best performance. This was mainly due to the aspect of local modelling characterising the fuzzy model. Furthermore, the fuzzy model being in the form of rules provides some insight to the unknown dynamics of the procedure as opposed to the ANN model which acts as a black box. The validation of the models using other wind time-series, as well as benchmark data, leaded to similar positive conclusions.

REFERENCES

[1] G. Kariniotakis, G. Stavrakakis, E. Nogaret, "Wind power forecasting using advanced neural network models", *paper accepted for publication at the IEEE Transactions on Energy Conversion*, IEEE PES 1996 Summer Meeting, 28/7-1/8 1996, Denver, Colorado, USA.

[2] E. Nogaret et al, "An advanced control system for wind-diesel power systems", *Procd. of the 1996 EUWEC,* 20-24 May, 1996, Götemborg, Sweden.

[3] ARMINES, NTUA, INESC, RAL, PPC, *Development and implementation of an advanced control system for the optimal operation and management of medium-sized power systems with a large penetration from renewable power sources*, Final report of EU-DG XII JOULE II project JOU2-CT92-0053. Edited by the Office for Official Publications of the European Communities, Luxembourg 1996.

[4] E.A. Bossanyi, "Stochastic wind prediction for wind turbine system control", *Procd. of the 7th BWEA Wind Energy Conference*, 27-29 March 1985, pp. 219-226.

[5] H.G. Beyer et al, "Short-term prediction of wind-speed and power output of a wind turbine with neural networks", *Procd. of the 5th European Wind Energy Conference*, EWEC'94, Thessaloniki, Greece, 1994, pp.349-352.

[6] ELSAM, *Wind power prediction tool in central dispatch centres*, Progress Report 3 of the EU project JOU2-CT92-0083, October 1993.

[7] T. Takagi, M. Sugeno, "Fuzzy identification and its applications to modeling and control", IEEE Transactions on Systems, Mans and Cybernetics, Vol. SMC-15, No. 1, January/February 1985.