CHEMICAL ENGINEERING TRANSACTIONS

VOL. 33, 2013

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at: www.aidic.it/cet

Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791

DOI: 10.3303/CET1333155

A Genetic Algorithm and Neural Network Technique for Predicting Wind Power under Uncertainty

Ronay Ak[a], Yan-Fu Li[a], Valeria Vitelli[a], Enrico Zio*[a,b]

[a]Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy - Électricité de France, École Centrale Paris, Chatenay-Malabry, France, and Supelec (École Supérieure d'Électricité), Plateau de Moulon, Gif-sur-Yvette, France.
[b]Energy Department, Politecnico di Milano, Via Ponzio 34/3 Milan, 20133 Italy
enrico.zio@ecp.fr

Wind speed uncertainty, and the variability in the physical and operating characteristics of turbines have a significant impact on power system operations such as regulation, load following, balancing, unit commitment and scheduling.
In this study, we consider historical values of wind power for predicting future values taking into account both the variability in the input and the uncertainty in the model structure. Uncertainty in the hourly wind power input is presented as intervals of within-hour variability.
A Neural Network (NN) is trained on the interval-valued inputs to provide prediction intervals (PIs) in output. A multi-objective genetic algorithm (namely, non-dominated sorting genetic algorithm–II (NSGA-II)) is used to train the NN. A multi-objective framework is adopted to find PIs which are optimal for accuracy (coverage probability) and efficacy (width).

1. Introduction

The power output of a wind turbine mainly depends on the local wind speed, and the physical and operating characteristics of the turbine. Wind speed changes according to weather conditions, in time scales ranging from minutes to hours, days and years (Kavasseri and Seetharaman, 2009); this aleatory behavior induces a corresponding variability in the power output. Uncertainty and variability of wind power have significant effects on wind integrated power system operations, such as regulation, load following, balancing, unit commitment and scheduling (Lew et al., 2011; Lei et al., 2009).
Several works can be found in the literature which focus on providing a forecasting tool in order to predict wind speed and power. The approaches proposed therein can be classified as physical, i.e. making use of numerical weather prediction (NWP) models (Giebel et al., 2006), statistical, i.e. data-driven methods comprising also artificial intelligence methods like neural networks (NN) and fuzzy logic (Kusiak et al., 2009), or a combination of both (Sideratos and Hatziargyriou, 2007). Most research has focused on point prediction of wind power with crisp input data. However, in order to reflect the variability of the phenomenon, it is important to provide uncertainty evaluation of its prediction considering variability in the input, which can be due to measurement errors and imprecise, incomplete and uncertain information.
Interval-valued representation (Moore et al., 2009), which means considering an interval enclosing real observations instead of real quantities themselves, can be used to reflect the variability (e.g. bounding wind speeds in a given area, minimum and maximum of daily temperature, etc.) and uncertainty (e.g. strongly skewed wind speed distributions, imprecise reliability of the components, etc.) in the observed crisp, single-valued measurements. Muñoz San Roque et al. (2007) propose an interval multi-layer perceptron (iMLP) model capable of handling input and output interval data, whereas the weights and biases defining the network are single-valued. They test the model by application to the forecasting of daily electricity prices intervals. Similarly, Garcia-Ascanio and Maté (2010) make a comparison between iMLP and the vector autoregressive (VAR) model for multivariate time series, adapted to interval time series (ITS), in order to forecast the monthly electric power demand per hour in Spain from 2006 to 2007.

Please cite this article as: Ak R., Li Y., Vitelli V., Zio E., 2013, A genetic algorithm and neural network technique for predicting wind power under uncertainty, Chemical Engineering Transactions, 33, 925-930 DOI: 10.3303/CET1333155

This work aims at doing short-term (1-hour ahead) wind power prediction taking into account both the variability in the input and the uncertainty in the model structure (Pandya et al., 2013). With the purpose of exploring the effects of using interval-input wind power data on the prediction accuracy and robustness, uncertainty in the input (hourly wind power) is represented by an interval which captures the within-hour variability. Unlike the existing papers on wind speed and power prediction, which use single-valued hourly wind power input (Kusiak et al., 2009) obtained as a within-hour average, we give an interval representation to the hourly inputs by using two approaches (see Section 4), which quantify in two different ways the within-hour variability.

To tackle the prediction problem, a data-driven learning approach is used, i.e. a neural network (NN) trained on the basis of experimental data (historical wind power data with a time-step of 5 min). The network uses interval-valued data but its weights and biases are crisp (i.e. single-valued). For the training of the NN, we implement a multi-objective genetic algorithm (namely, non-dominated sorting genetic algorithm–II (NSGA-II)) to find the optimal parameters (weights and biases) of the NN. The network maps interval-valued data into an interval output, providing the prediction intervals (PIs) of the wind power. The PIs are optimized both in terms of accuracy (coverage probability) and dimension (width). The prediction interval coverage probability (PICP) represents the probability that the set of estimated PI values will contain a certain percentage of the true output values. Prediction interval width (PIW) simply measures the extension of the interval as the difference of the estimated upper bound and lower bound values.

The paper is organized as follows. Section 2 briefly introduces the basic concepts of interval-valued NNs for PIs estimation. In Section 3, the use of NSGA-II for training an interval-valued NN to estimate PIs is described. Experimental results on the real case study of short-term wind power prediction are given in Section 4. Finally, Section 5 concludes the paper with an analysis of the results obtained and some ideas for future improvements.

2. NNs and PIs

Neural networks (NNs) are a promising statistical data-driven method capable of learning complex nonlinear relationships among variables, from observed data. A NN is an interconnected assembly of individual processing units (neurons). Information is passed between these units along interconnections. An incoming connection has two values associated with it, an input value and a weight (Kalogirou, 2001). The neurons are connected by weights and convert input data $D = \{(x_n, y_n), n = 1, 2, \ldots, n_p\}$ into output values by using a sigmoid transfer (activation) function. Figure 1 shows the scheme of a multiple-input neuron with the associated information processing through this neuron.

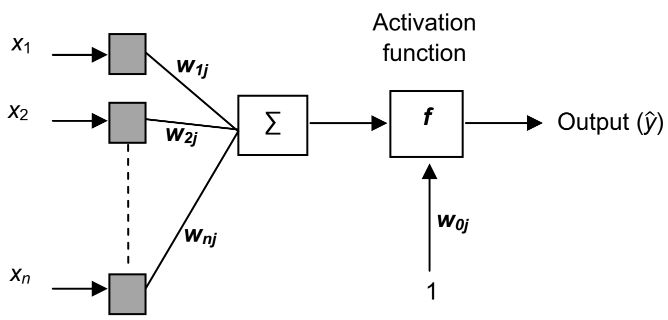


Figure 1: Multiple input neuron (Nazzal et al., 2008)

A PI is defined by upper and lower bounds that include a future unknown value with a predetermined probability, called confidence level ($1 - \alpha$). The formal definition of a PI with crisp data can be presented as follows:

$$Pr\left(\hat{y}^-(x) < y(x) < \hat{y}^+(x)\right) = 1 - \alpha \tag{1}$$

where $\hat{y}^-(x)$ and $\hat{y}^+(x)$ are the estimated lower and upper bounds corresponding to input x; the confidence level ($1 - \alpha$) refers to the expected probability that the true value of $y(x)$ lies within the prediction interval $[\hat{y}^-(x), \hat{y}^+(x)]$.

When interval-valued data are used as input, each input pattern x_i is represented as an interval $x_i = [x_i^-, x_i^+]$ where $x_i^- \leq x_i^+$ are the lower and upper bounds (real values) of the input interval, respectively. With the same formulation, it is natural to describe each estimated output value $\hat{y}_i$, corresponding to the i-th sample x_i, as $\hat{y}_i = [\hat{y}_i^-, \hat{y}_i^+]$, where $\hat{y}_i^- \leq \hat{y}_i^+$ are, respectively, the estimated lower and upper bounds of the prediction interval in output.
The mathematical formulation of the PICP and PIW measures has been given in Khosravi et al. (2011). In this work we have modified these two measures to adapt them to interval-valued input and output data:

$$PICP = \frac{1}{n_p}\sum_{i=1}^{n_p} c_i \tag{2}$$

where n_p is the number of samples in the training or testing sets, and

$$c_i = \begin{cases} 1 & y_i \subseteq [\hat{y}_i^-, \hat{y}_i^+] \\ \frac{diam(y_i \cap \hat{y}_i)}{diam(y_i)} & y_i \not\subset [\hat{y}_i^-, \hat{y}_i^+] \wedge y_i \cap \hat{y}_i \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $y_i = [y_i^-, y_i^+]$ where $y_i^- \leq y_i^+$ are the lower and upper bounds (real values) of the output interval, respectively, and *diam*() indicates the width of an interval.
Concerning PIW, we consider the following quantity:

$$\text{NMPIW} = \frac{1}{n_p}\frac{\sum_{i=1}^{n_p}(\hat{y}_i^+ - \hat{y}_i^-)}{t_{max} - t_{min}} \tag{4}$$

where NMPIW is the Normalized Mean PIW, and t_{min} and t_{max} represent the true minimum and maximum values of the targets (i.e., the bounds of the range in which the true values fall).

3. NSGA-II optimization of a NN for PIs estimation

NSGA-II generates a Pareto optimal solution set, rather than a single solution, by comparing different solutions via an elitist approach, i.e., a fast non-dominated sorting and crowding-distance estimation procedure (Konak et al., 2006). The practical implementation of NSGA-II on our specific problem involves two phases: initialization and evolution. These can be summarized as follows (Ak et al., 2013):

Initialization phase:
Step 1: Split the input data set into training (D_{train}) and testing (D_{test}) subsets.
Step 2: Fix the maximum number of generations and the number of chromosomes (individuals) Nc in each population. Each chromosome codes a solution by G real-valued genes, where G is the total number of parameters (weights and biases) in the NN: thus, each chromosome represents a NN. Set the generation number n =1. Initialize the first population P_n of size Nc, by randomly generating Nc chromosomes (corresponding to NNs).
Step 3: For each input vector x in the training set, compute the lower and upper bound outputs of the Nc NNs.
Step 4: Evaluate the two objectives PICP and NMPIW for the Nc NNs; then, one pair of values 1 — PICP (for minimization) and NMPIW is associated to each of the Nc chromosomes in the population P_n.
Step 5: Rank the chromosomes (vectors of G values) in the population P_n by running the fast non-dominated sorting algorithm (Konak et al., 2006) with respect to the pairs of objective values, and identify the ranked non-dominated fronts $F_1, F_2, \ldots, F_k$ where F_1 is the best front, F_2 is the second best front and F_k is the least good front.
Step 6: Apply to P_n a binary tournament selection based on the crowding distance (Konak et al., 2006), for generating an intermediate population S_n of size Nc.
Step 7: Apply the crossover and mutation operators to S_n, to create the offspring population Q_n of size Nc.
Step 8: Apply Step 3 onto Q_n and obtain the lower and upper bound outputs.
Step 9: Evaluate the two objectives in correspondence of the solutions in Q_n, as in Step 4.

Evolution phase:
Step 10: If the maximum number of generations is reached, stop and return P_n. Select the first Pareto front F_1 *as* the optimal solution set. Otherwise, go to Step 11.
Step 11: Combine P_n and Q_n to obtain a union population $R_n = P_n \cup Q_n$.
Step 12: Apply Steps 3-5 onto R_n and obtain a sorted union population.
Step 13: Select the *Nc* best solutions from the sorted union to create the next parent population P_{n+1}.
Step 14: Apply Steps 6-9 onto P_{n+1} to obtain Q_{n+1}. Set $n = n + 1$; and go to Step 10.
Finally, the best front in terms of ranking of non-dominance and diversity of the individual solutions is chosen, and testing of the trained NN with optimal weight values is performed using the data of the testing set.

4. Experiments and results

In this Section, results of the application of the proposed method to short-term wind power forecasting with interval-input data are detailed. The considered 5-min wind power data have been measured for Canunda, a region of South Australia. The actual situation in Canunda wind farm is characterized by the presence of 23 turbines capable of generating approximately 46 megawatts (MW) of electricity, enough to provide the power needs of around 30,000 homes (GDF SUEZ, 2010).
The wind power data set, covering the period from January 18, 2012 till March 13, 2012, has been downloaded from the website AEMO (2012). As 5-min data have been collected, there are 12 wind power values for each hour. The raw data set includes 6,000 samples among which the first 80 % (the first 4,800 samples) is used for training and the rest for testing. The real wind power changes from 0 MW to 43.35 MW with an unstable behavior. Figure 2 shows the behavior of 5-min wind power values only in the first 24 hours, for the sake of clarity: one can appreciate the within-hour variability in each individual hour.
In order to represent hourly wind power as an interval, this 5-min data have been converted to interval-input data with two approaches, named “min-max” and “mean”: the former obtains hourly intervals by taking the minimum and the maximum values of the wind power per hour; in the latter approach, instead, the within-hour mean ($\overline{x}_i$) and the standard deviation (s_i) of 12 5-min wind power data have been computed, and then one-standard deviation intervals have been obtained as $[\overline{x}_i - s_i, \overline{x}_i + s_i]$.

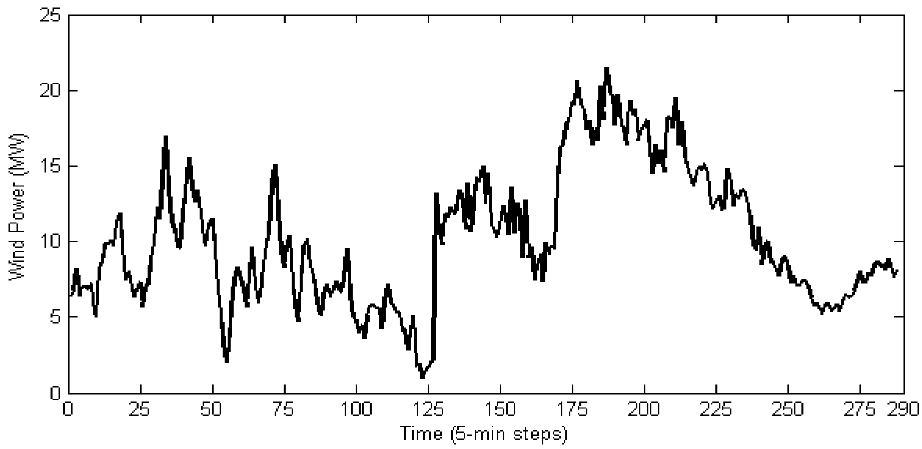
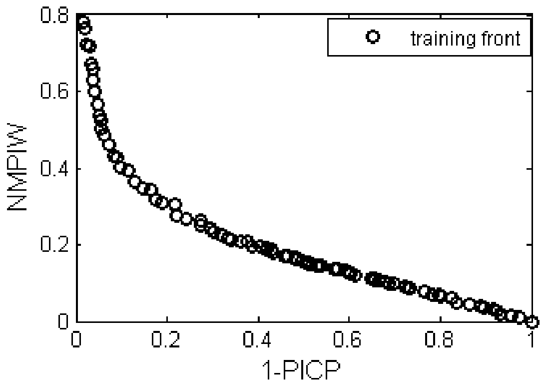


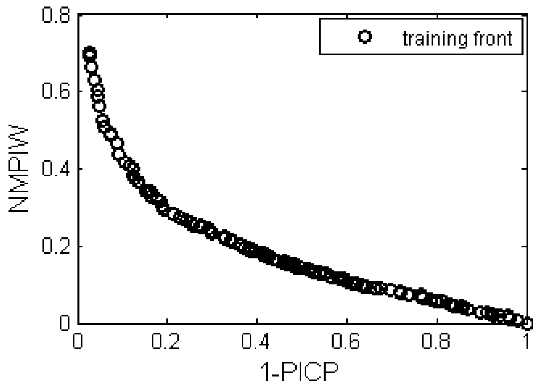
Figure 2: The 5-min wind power data set used in this study: first 24 h

The architecture of the NN consists of one input, one hidden and one output layers. The number of input neurons is 1, since the historical wind power value W_{t-1} is used as input variable for predicting W_t in output; the number of hidden neurons is set to 10 after a trial-and-error process; the number of output neurons is 1 which results in interval-valued estimations. As activation functions, the hyperbolic tangent function is used in the hidden layer and the logarithmic sigmoid function is used at the output layer. All data have been normalized within the range [0.1, 0.9]. In NSGA-II, the population size (*Nc*) is set to 50 and the number of generations (MaxGen) to 500; this latter is used as termination condition.
To account for the inherent randomness of NSGA-II, five different runs have been performed and an overall best non-dominated Pareto front has been obtained from the five individual fronts. Figure 3 illustrates the first (best) Pareto front found after training the NN on interval data constructed by the min-

max approach (a) and mean approach (b). Given the overall best Pareto set of optimal solutions (i.e. optimal NN weights), one has to select one (i.e. one trained NN) for use. For both interval construction approaches, the solution has been chosen as the one with smallest NMPIW among those with $PICP \geq 0.9$: 90 % CP and interval width of 0.494 for min-max, and 91 % CP with 0.439 interval width for the mean approach. The results on the test dataset give a coverage probability of 82 % and an interval width of 0.357 for the min-max approach, and 80 % CP with 0.380 interval width for the mean approach. Figure 4 shows the prediction intervals estimated on the test data set by the trained NN corresponding to the Pareto solution, for the min-max approach.



(a)



(b)

Figure 3: The overall best Pareto front obtained by training the NN for 1h-ahead wind power prediction: (a) min-max approach (b) mean approach

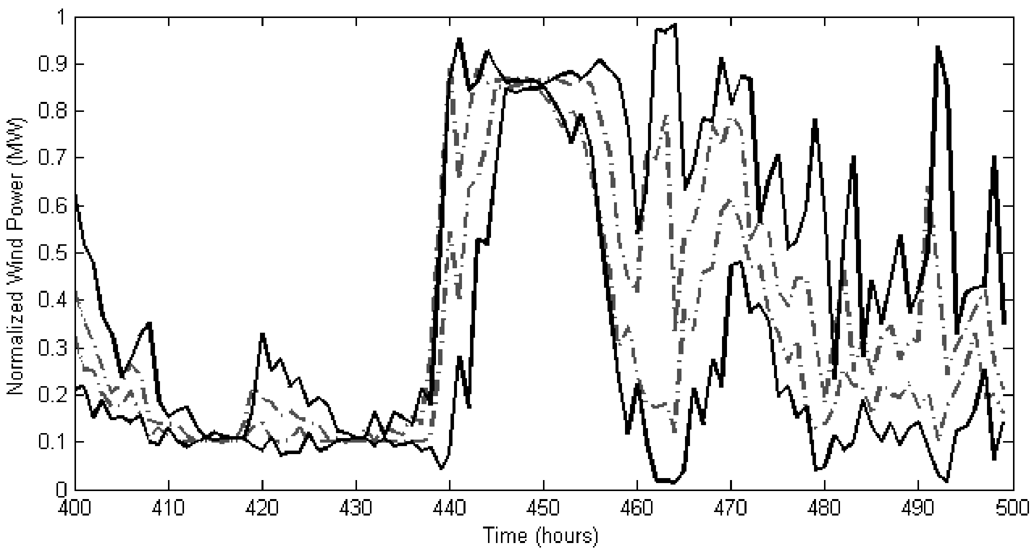


Figure 4: Estimated PIs for 1h ahead wind power prediction on the test data set (solid lines), and interval-valued wind power data (constructed by the min-max approach) included in the test data set (dashed line)

From the results illustrated in Figure 4, one might say that PIs obtained via the min-max approach are capable of capturing the peak points (highest and lowest) of the target output. The drop of coverage probability from 90 % in the training to 82 % in the testing dataset, which results in tighter interval widths, can be due to the particular nature of the data at hand showing a remarkable variability in the test data with respect to the training data. Another reason could be the insufficient number of observations in the training and testing samples. These claims should be confirmed by using larger experimental sets of data.

5. Conclusion

Variations in the generated power due to variability in wind turbines power generation can lead to serious problems, especially in the day-ahead commitments of generation resources to meet the electric demand.
In the work presented in this paper, we have represented by intervals the wind power variability in a given time horizon, 1 h in our case. The goal is to contribute to the understanding, representation and analysis of the uncertainty associated to wind power generation prediction. The original contributions of the work are to handle the prediction problem with uncertain inputs in a multi-objective framework. In fact, rather than optimizing parameters and subsequently obtaining the outputs, we directly map the interval inputs into PIs (interval outputs), which are optimal both in terms of coverage and width. Moreover, we explore two different approaches to represent input variability, and to quantify potential uncertainties in the outputs. The results obtained show that NNs are promising for handling interval input data accounting for uncertainties.
As for future research, the use of different approaches for better interval representation of wind power data will be explored to further increase the accuracy of the predictions. The extension of the approach for other engineering applications will also be pursued.

References

AEMO (Australian Energy Market Operator), 2012, Wind generation data, <www.aemo.com.au/Electricity/Data/Market-Management-System-MMS/Generation-and-Load> accessed 14.11.2012.

Ak R., Li Y., Vitelli V., Zio E., Droguett E. L., Jacinto C. M. C., 2013, NSGA-II-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment, Expert Systems with Applications, 40 (4), 1205-1212.

GDF SUEZ Australian Energy, Canunda wind farm brochure, 2010, <www.gdfsuezau.com/uploads/2010/01/CanundaBrochure4ppv2.pdf> accessed 14.11.2012.

Giebel G., Badger J., Marti I., Louka P., Kallos G., Palomares A. M., Lac C., Descombes G., 2006, Short-term forecasting using advanced physical modeling—the results of the ANEMOS project, European Wind Energy Conference (EWEC), Athens, Greece.

Kalogirou S. A., 2001, Artificial neural networks in renewable energy systems applications: a review, Renewable and Sustainable Energy Reviews, 5, 373-401.

Kavasseri R. G., Seetharaman K., 2009, Day-ahead wind speed forecasting using f-ARIMA models, Renewable Energy, 34, 1388-1393.

Khosravi A., Nahavandi S., Creighton D., Atiya A. F., 2011, Lower upper bound estimation method for construction of neural network-based prediction intervals, IEEE Transactions on Neural Networks, 22, 337-346.

Konak A., Coit D. W., Smith A. E., 2006, Multi-objective optimization using genetic algorithms: a tutorial, Reliability Engineering and System Safety, 91, 992-1007.

Kusiak A., Zheng H., Song Z., 2009, Short-term prediction of wind farm power: a data mining approach, IEEE Transactions on Energy Conversion. 24, 125-136.

Lei M., Shiyan L., Chuanwen J., Hongling L., Yan Z., 2009, A review on the forecasting of wind speed and generated power, Renewable and Sustainable Energy Reviews, 13, 915-920.

Lew D., Milligan M., Jordan G., Piwko R., 2011, The value of wind power forecasting, 91st American Meteorological Society Annual Meeting, the Second Conference on Weather, Climate, and the New Energy Economy Washington, USA, January 26, 2011, <http://www.nrel.gov/docs/fy11osti/50814.pdf>, accessed 14.11.2012.

Moore R.E., Kearfott R.B., Cloud M.J., Eds., 2009, Introduction to Interval Analysis. Society for Industrial Mathematics, Philadelphia, USA

Muñoz San Roque A., Maté C., Arroyo J., Sarabia A., 2007, iMLP: applying multi- layer perceptrons to interval-valued data, Neural Processing Letters, 25, 157–169.

Nazzal J. M., El-Emary I. M., Najim S. A., 2008, Multilayer perceptron neural network (MLPs) for analyzing the properties of Jordan oil shale, World Applied Sciences Journal, 5, 546-552.

Pandya N., Gabas N., Marsden E., 2013, Uncertainty analysis of Phast's atmospheric dispersion model for two industrial use cases, Chemical Engineering Transactions, 31, 97-102, DOI: 10.3303/CET1331017.

Sideratos G., Hatziargyriou N. D., 2007, An advanced statistical method for wind power forecasting, IEEE Transactions on Power Systems, 22(1), 258-265.