2018 IEEE 9th Power, Instrumentation and Measurement Meeting (EPIM)

Genetic algorithm applied to the specialization of neural networks for the forecast of wind and solar generation

D. Vallejo, E. Cornalino, and R. Chaer, *Senior-Member, IEEE*

***Abstract*-- This work presents the training of a Neural Network (NN) using Genetic Algorithm (GA) for medium-term wind and solar power generation forecasts. Wind speed, solar radiation, and other meteorological variables forecasts are provided by a meteorological service and are used as the NN's inputs. The NN's output is the generated power forecast. A comparison between the previously used model and the new one is shown and future improvements are discussed.**

***Index Terms*-- neural networks, wind and solar power mid-term generation forecast, genetic algorithm**

I. Introduction

Wind and solar power generation forecast is becoming more important in systems with high penetration for these technologies. In the case of Uruguay, the installed wind power capacity exceeds the average load and generates about 40 % of the countries' energy (YTD 2018-07-19) [1]. Installed solar power capacity is much lower and its contribution to the energy supply is 3.4 %. However, in the future it is expected that its participation will increase significantly.

Wind power forecast is an important issue in countries with high penetration of wind power, such as Uruguay, Denmark, Germany, Spain, and the USA and each one is working in developing reliable systems that are constantly evaluated [2].

Deep neural networks have been used for probabilistic short-term wind power forecasts in Germany with success using discrete target classes (bins) [3]. In Ireland, a back propagation (BP) training method with the improvement of an Ant Lion Optimizer has been implemented [4]. These papers and most works found in current publications are focused in short-term forecasts, in this paper mid-term (168 hours) forecasts are calculated. Activation functions different from the classic sigmoid have successfully been used, such as Wavelet Neural Networks (WNN) [5].

Under these conditions, the system's optimal management greatly depends on the ability to anticipate the power generated by these sources. Regarding the prediction of wind and solar generation, on the one hand there are the difficulties associated with the prediction of the resources, wind velocity and solar radiation. On the other hand, it is necessary to improve the prediction of the generation of the different farms based on the conditions that influence it.

Attending to the resolution of this last aspect from the Administration of the Electric Market (ADME), it is sought to develop a simple forecasting model, of easy operation and maintenance that allows adjusting to each power plant, achieving to predict the expected generation with the greatest accuracy.

The prediction of the generation of a wind or solar farm from the information of the available wind velocity or solar radiation is not easily reducible to the application of the operating curve of the respective power plant. There are multiple variables that influence the operation of the plants that, for example, in the case of wind generation, for the same wind speed module measured at one point, the power generated by the plant can be drastically different. To illustrate this, the case of the power curve tests is carried out on a single wind turbine not affected by wakes of other wind turbines or by complexities of the environment. This limits the operating conditions to ranges of speed and turbulence that exclude any type of anomalous or extreme events. Even in these conditions, the tests are usually given with an uncertainty of around 5 %. The generation of a plant composed of dozens of wind turbines distributed in parks of thousands of square meters is much more complex to determine and requires significant simplifications if it is wanted to address its physical modeling. For this reason and given the speed of calculation required to obtain real-time forecasts, it is necessary to approach the problem with mathematical models that allow multiple variables to be considered without falling back into the detailed modeling of each plant.

So far, the wind forecast is using a simplified model based on only 3 parameters and the solar forecast is even more simplified. Both are described in Section II.

In this paper a perceptron-like architecture based model is proposed. It applies to both wind and solar energy and, as expected, in the first one wind speed is more relevant, and in the second, solar radiation.

II. Previous model

Until this study, wind and solar power generation models were calculated in ADME using the models presented in this section. The function that relates wind velocity (v) with generated power per wind farm (P) is the following:

$$P(v)=\begin{cases} 0 & if \quad v \le v_{min} \\ \dfrac{P_0}{1+e^{-\alpha(v-v_m)}} & if \quad v_{min} < v \le v_{max} \\ \dfrac{v_{th}-v}{v_{th}-v_{max}} P_0 & if \quad v_{max} < v < v_{th} \\ 0 & if \quad v \ge v_{th} \end{cases}$$

The parameters (α, P_0 and v_m) were calibrated to minimize the Mean Squared Error (MSE) between the model and the SCADA measurement. For v_{th}, v_{max} and v_{min} the values 20 m/s, 25 m/s and 0.2 m/s were used to represent the cut-in, cut-out and cut-out transitions.

For solar power generation the following model was used:

$$P_1(E)=max\{0 \ , \ (a_0 E^2+a_1 E+a_2) P_0\}$$
$$P=min\{P_{1,} P_0\}$$

Which relates solar radiation (E) with output power per solar farm (P), where P_0 is the effective installed capacity of the farm. The same coefficients were used for every farm and they were calibrated to minimize the MSE of all the farms against the SCADA power measurement. The values obtained for a_0, a_1, and a_2 are 4.47×10^{-7}, 9.76×10^{-4}, and -7.32×10^{-3} respectively.

A meteorological forecast is provided by a meteorological service twice a day with a 168 hour horizon for wind speed and solar radiation among other quantities at the wind or solar farm's geographic coordinates.

III. Proposed Model

In this work, a model based in a perceptron-like architecture [6] using more meteorological quantities than the previous model, such as wind direction, air density, and temperature, is proposed. The inputs are weighed by multiplicative factors, then added among themselves and a bias value, and finally passed through a saturation sigmoid function to produce the output, which is per unit power.

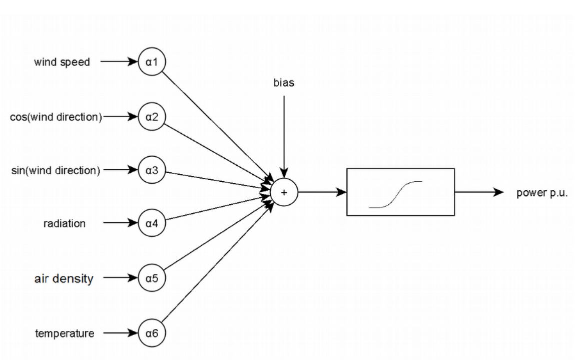

Figure 1: Architecture of a perceptron neuron.

This architecture is shown in Figure 1. and the perceptron output is given by:

$$P(t)=\frac{1}{1+e^{-\left[\left(\sum_{k=1}^{k=N}\alpha_k x_k(t)\right)+\beta\right]}}$$

Where the α_k's are the weighing factors, the x_k's are the input variables, β is the bias value, N is the number of input variables and P is the output of the perceptron. This architecure can be nested in what is called a Multilayer perceptron (MLP). Multiple perceptrons (neurons) can be used in the same layer and multiple layers can be used to form a NN. In this work, three different combinations are used and the final output is not passed through the activation function.

To generate the data used for this training, time series for past forecast were constructed by concatenating the first twelve hours of each forecast (since there is a new forecast every twelve hours). For the output signal, the SCADA power measurement is considered, along with the set-point and individual generators availability signals. With these three signals, a fourth one is generated which is per unit of power per available generator and it is not valid where there is an operative restriction (i.e. the setpoint is below the installed capacity), this signal is what the output of the network has to match.

In this first version, the input variables considered are wind speed, wind direction (in cosine and sine components), radiation, air pressure and temperature. All of them in their predicted values for the moment in which the power forecast is wanted. There are a lot of other variables that may be relevant in determining the generated power whose contribution will be evaluated comparatively to include them, if it is beneficial, to obtain new versions of this model. Some of these variables to evaluate can be, for example:

- previous power measurement
- turbulence intensity
- solar radiation received integrated in a previous time interval to be determined or another indirect indicator of the expected atmospheric state [7]

IV. Training Method

The network was trained via a GA where the weighing factors and bias values were initially randomly selected for each individual (set of values) to create an initial population (set of individuals). Each individual is evaluated by calculating the mean squared error of the output against the SCADA measurement and the lower that value is, the better that individual is considered. To generate new individuals, twp of the best individuals of the previous generation are mixed and mutation is allowed for the new one.

After a significant amount of iterations of this algorithm, the best individual is chosen and its values are used as the weighing factors and bias for the network.

The implementation of this training was made in the OddFace platform [8]. OddFace stands for distributed optimization of functions of high evaluation cost in Spanish. This tool, which was developed in ADME, made it possible to make many evaluations in a short period of time. OddFace is a platform that allows the distributed search of the optimum of an optimization problem by means of a set of Broker Agents, which share the information of the explored areas of the Domain of the Problem, by means of a central database. This is shown schematically in Figure 2. Each instance of the program (i.e. broker agent, solver, worker, shown as T_i in Figure 2) evaluates the objective function for a given individual. These workers run independently from each other in a different core of a server or even in a different server and communicate the result of the evaluation to a database (SDB in Figure 2). The workers also consult the database to have the information of individuals that have already been evaluated, individuals that are pending evaluation and the best individuals to use as "parents" to generate a new one with the GA.

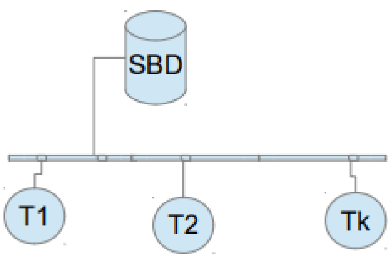

Figure 2: Database server and workers. Extracted from [8].

The exploration of domain D is then carried out in a distributed manner, by several agents, using the information of all the evaluations permanently shared among the agents. Each agent, using an algorithm to estimate a new point, generates a proposal for a new evaluation point and generates an evaluation with a set of realizations. This gives an "estimate" of the objective function at that point. If the point had already been calculated, the new information is integrated improving the existing estimate (if the function is not deterministic).
This particular problem does not have a high cost of evaluation (in processing time) and there is only one realization for each signal (deterministic) but many evaluations are needed, so the OddFace tool is suitable to solve it.

V. Results

To evaluate the performance of the resulting network, the Mean Squared Error (MSE) of the new model is compared against the previous one. The MSE is calculated as follows:

$$MSE=\frac{1}{N}\sum_{k=1}^{k=N}\left(P_{r[k]}-P_{f[k]}\right)^2$$

Where P_r stands for the value from the SCADA and P_f the forecasted value. The signal P_r is considered not valid where there is an operation restriction or the SCADA signal is originally invalid. Those terms are not considered for the error and are not considered for the training either.

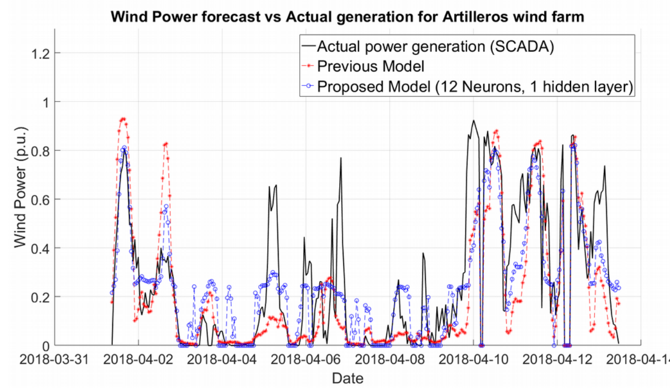

Figure 3: Wind Power forecast and real generation obtained from SCADA.

To make this comparison, wind and a solar farms were chosen over a one year time window and both method were applied. For the training series the time window was from 2017-07-19 to 2018-03-31 and for the evaluation series the interval was from 2018-04-01 to 2018-07-19. The sample period is one hour, so 6144 and 2616 samples were used to have a roughly 70/30 training to evaluation period ratio.
Three different topologies were used to fit the data; just one neuron, two neurons in a hidden layer, and twelve neurons in a hidden layer (six for solar farms). These different topologies were used to evaluate how much does the model improve by adding complexity to the network. The twelve neurons for the wind farms and the six for the solar ones were intended to specialize in different directions of the wind and different parts of the day respectively.

Table 1 shows the results of the comparison of wind farms and Table 2 of solar ones.

	Previous method	One neuron	Two neurons in hidden layer	Twelve neurons in hidden layer
Melowind	0.0635	0.0842	0.0699	0.0778
Carapé 1	0.0605	0.0755	0.0872	0.0839
Artilleros	0.0380	0.0403	0.0413	0.0395

Table 1. Mean squared error for different wind farms.

	Previous method	One neuron	Two neurons in hidden layer	Six neurons in hidden layer
La Jacinta	0.0140	0.0206	0.0215	0.0191
Raditon	0.0139	0.0245	0.0182	0.0231
Alto Cielo	0.0198	0.0386	0.0368	0.0540

Table 2. Mean squared error for different solar farms.

Figure 3 shows the comparison between the previous model and the NN with twelve neurons in a single hidden layer for Artilleros wind farm.

VI. CONCLUSIONS AND FUTURE WORK

This is a first step from which it is intended to continue advancing with the objective of generating a power forecast model with maximum certainty adaptable to all existing and future plants integrated into the system.

As it was shown in the comparatives results, there was not an improvement in any of the farms for the selected evaluation interval.

There are a lot of other variables that may be relevant in determining the generated power, that have not been explored in this instance and, as mentioned above, it will be evaluated to include them, if it is beneficial, to improve this version of the model.

Likewise, a more complex network could be used for the wind model to get better results, either including more layers or information about the previous values of the input signals.

An activation function different from the sigmoid could be used, such as a wavelet function as used in [5].

The downside of the GA is that there is no certainty that the best individual obtained will not be bested if the algorithm continues running for more time. For this study, approximately one hundred and fifty thousand evaluations were carried for each problem. However, it has been observed, particularly in the cases with more neurons, that new best individuals keep being generated. This means that, given enough evaluations, a model that surpasses the previous method could be obtained. In the moment of the end of this study, the GA was still generating better individuals so these results can still improve with more evaluations.

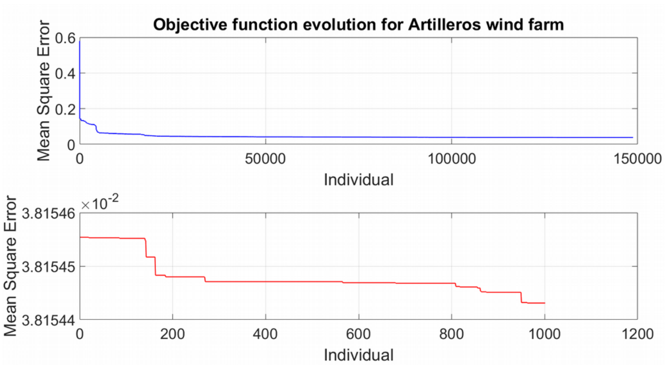

Figure 4: Evolution of the minimum attained for the objective function of Artilleros wind farm up to certain individual

In Figure 4 it can be observed that the minimum attained for the objective function up to certain individual decreased drastically in the first thousands of individuals and then at a lower rate but it is still decreasing in the last hundreds of individuals.

Finally, in future work, the uncertainty must be incorporated into the model both in the input variables and in the output obtained as power forecast. This is essential for the proper evaluation of risks in decision making in the electrical dispatch due to the uncertainty of multiple random variables involved.

VII. REFERENCES

[1] 2018 [Online] Available: http://www.adme.com.uy/
[2] J. Dobschinski, *Wind Power Forecasting Error Distributions*, 2012.
[3] M. Felder, *Probabilistic short term wind power forecasts using deep neural networks with discrete target classes*, 2018.
[4] F. Jiang, *Short-term Wind Power Forecasting Based on BP Neural Network with Improved Ant Lion Optimizer*.
[5] W. Yicong, *Short-term Wind Power Forecasting by Genetic Algorithm Of Wavelet Neural Network*, 2014.
[6] V. Kecman, *Learning and Soft Computing,* The MIT Press, 2001.
[7] S. Wharton, *Athmospheric stability affects wind turbine power collection,* Enviromental Research Letters, vol. 7, n. 1.
[8] R. Chaer, E Coppes, F Barreto, C Tutté, F Maciel, M Forets, E Cornalino, M Gurín, M Álvarez, F Palacios, D Cohn. *Memoria Final Proyecto A#II-FSE2009-18* https://simsee.org/simsee/biblioteca/anii_fse_2009_18/memoria_fse_2009_18_MejorasSimSEE.pdf