Net demand short-term forecasting in a distribution substation with PV power generation

Eduardo Garcia-Garrido*, *Montserrat* Mendoza-Villena, *Pedro M*. Lara-Santillan, *Enrique* Zorzano-Alba, and *Alberto* Falces

Department of Electrical Engineering, University of La Rioja, Logroño, Spain.

Abstract. The integration of renewable energies, specifically solar energy, in electric distribution systems is increasingly common. For an optimal operation, it is very important to forecast the final net demand of the power distribution network, considering the variability of solar energy combined with the variability of the electric energy consumption habits of population. This paper presents the methodology followed to forecast the net demand in a power distribution substation. Two approaches are considered, the net demand direct prediction, and the indirect prediction with the forecasts of PV power generation and load demand. Artificial Neural Network (ANN) based models and autoregressive models with exogenous variables (ARX) are used to predict the net demand, directly and indirectly, for the 24 hours of the day-ahead. The methodology is applied to a medium voltage distribution substation and the direct and indirect forecasts are compared.

1 INTRODUCTION

The renewable energies are power sources required to reduce the greenhouse effect and to solve the problems of fossil fuels supplying. One of renewable energy sources (RES) more used is the solar photovoltaic (PV), due to it is an unlimited, clean, ecological, dispersed and free energy. The PV energy is the conversion of sunlight into electricity through solar panels. But the integration of PV power into the electric grids poses some difficulties. For an optimal operation of the power system it is fundamental to know the PV power generated in every moment, especially if the penetration of PV power is significant. This fact has propelled the development of short-term forecasting models for photovoltaic generation [1].

On one hand, the electricity market and the smart grid require accurate forecasting tools for carry out better demand side management tasks and more effective economic decisions. On the other hand, it is complex to predict the electric net demand due to different factors such as weather variables, social habits and seasonal characteristics [2].

According to the forecasting horizon, the forecasts can be classified in four categories: long-term (3 years-50 years), medium-term (2 weeks-3 years), short-term (hour, day or week) and very short-term (5 minutes-1 hour). The short-term load forecasting (STLF) is a decisive tool to ensure the balance between generation and demand [3] and to reduce the risk in decision-making power system planning and operational decisions [4].

Some researchers have published works describing short-term forecasting models applied to PV plants, load demand or net demand. They use different techniques for the forecasting model. Some of them include artificial neural network (ANN) based models or autoregressive (AR) based models [5,6].

The use of STLF has been identified as the proper tool to solve the problems derived from the integration of green (and very variable) energy sources into the electric power systems. In this context, the term of Net Demand (ND) is defined as electric load minus renewable generation. The ND prediction can be estimated of indirectly manner, that is, subtracting the renewable generation forecast to the electric load demand forecast; or can be forecasted directly. Bagheri *et al*. [7] carry out tests with the data of four weeks using five models: persistence, multilayer perceptron (MLP), radial basis function neural network (RBFNN), wavelet neural network (WNN) and three-phase Cascade Neural Network (CNN). The comparison of the forecasting results lead to the conclusion that the direct ND forecasting achieves lower errors than the indirect forecasting. They applied their methodology to practical power networks (regional or national power systems) like those of Alberta (Canada) or Ireland.

To this regard, van de Meer *et al*. [8] examined the difference between direct and indirect ND forecasts using static and dynamic Gaussian Processes (GP) on a set of residential consumers (300 customers) with a half-hourly basis. The authors conclude that there is no single best method (direct or indirect) that can be applied on any location and under any circumstance and they evaluate the performance of different models.

In this paper we define and apply a methodology to predict the ND, in a direct or indirect manner, using ANN based models and autoregressive with exogenous

* Corresponding author: eduardo.garcia@unirioja.es

variables (ARX) models for an electric distribution substation with significant PV power penetration.

This paper is structure as follows: in section 2 the methodology is described; section 3 presents the results obtained with the proposed methodology and finally, the conclusions are presented in section 4.

2 METHODOLOGY

2.1 Direct prediction of net demand

We define the ND at the connection point of the electric substation as the difference between the load demand (electricity consumption) and the generation of PV systems connected to this substation.

Three time series with mean hourly values can be identified: the load demand, the PV power generation and the net demand. As it was explained above, the availability of two of the time series is sufficient, as the third series can be calculated as sum or difference of the other two. Figure 1 plots an example with 14 days of the three time-series.

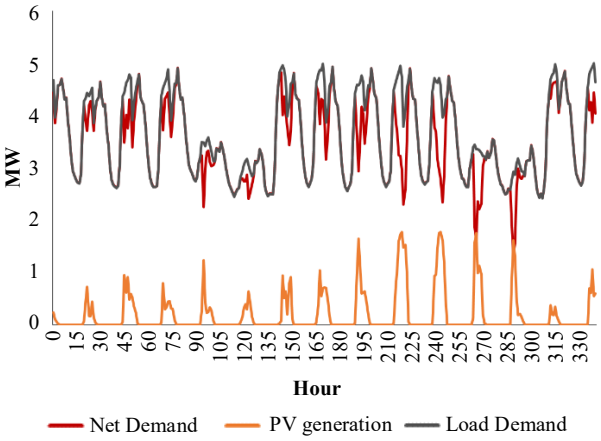

Fig. 1. Load demand, PV generation and Net demand. An example.

The three electric time series (ND, PV power and load demand) need to be complemented with other series corresponding to possible explanatory variables. The most used explanatory variables are weather variables, because there is a direct relation between them (for example between solar irradiance and PV power, or temperature and load demand). In order to develop an operational forecasting model for the next day, the explanatory variables must correspond to forecasts of the most important weather variables. These forecasts can be obtained from Numerical Weather Prediction (NWP) models, which are models that estimate the future values of the weather variables from initial conditions by solving the equations that govern the dynamic behaviour of the atmosphere. Usually NWP models work with grids of points, and the weather forecasts from the location of interest must be obtained by interpolation of the forecasts for the points of the analysis grid nearest to the location of the PV plant and consumers.

In our work, we predict the ND for the 24 hours of the next day using forecasting models based on two different techniques: ANN and ARX. The selected ANN model was the multilayer perceptron (MLP) with one hidden layer, since despite its simplicity, it is a model that behaves effectively in most of forecasting applications. In order to select a proper forecasting model, MLPs with different number of neurons in the hidden layer must be trained: the available data for the training process is randomly divided into two sets, the first one with 80% of the data used as the training data set, and the second with the resting 20% used as a validation data set. The different MLP models are trained using the back-propagation with momentum algorithm stopping the iterations (epochs) when the error with the validation data set begins to increase. The selected MLP model (with its specific number of neurons in the hidden layer) is that who achieves the lowest average error with the complete set (training and validation data sets).

The Root Mean Square Error, RMSE (1), and the Mean Absolute Percentage Error, MAPE (2) were used to assess the forecasting error (difference between real and forecast value).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{ND}_{\mathrm{forecast},i}-\mathrm{ND}_{\mathrm{actual},i}\right)^2} \quad (1)$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{\left|\mathrm{ND}_{\mathrm{forecast},i} - \mathrm{ND}_{\mathrm{actual},i}\right|}{\mathrm{ND}_{\mathrm{actual},i}} \quad (2)$$

where $ND_{forecast,i}$ is the ND forecast in the hour i, $ND_{actual,i}$ is the actual value of ND in the hour i, and n is the total number of cases (hours).

The direct forecasting model predicts future ND in the substation for the hour h of the next day (d+1). The input variables for the MLP and ARX models are: past values (ND in hour h of day d-1 and of day d-7), seven dummy variables to code the day d+1 (Monday, Tuesday-Wednesday-Thursday, Friday, Saturday, Sunday, regional or national festivity, and local festivity), and forecasted values for hour h of day d+1 of the selected weather variables (temperature, global horizontal irradiance, pressure, wind speed, wind direction, relative humidity, cloud fraction and rainfall) [9].

2.1 Indirect prediction of net demand

In this case, we predict the ND as the difference between the forecast of the PV power and forecast of the load demand. The forecasting models should be developed with the same technique as those used for the direct prediction in order to compare results. So, MLP models and ARX models were applied in both cases (direct and indirect prediction).

Therefore, two forecasting models must be developed for the indirect prediction of the net demand. The first forecasting model predicts the hourly PV power value for hour h of the day-ahead (d+1) and it is based only on forecasts of weather variables for hour h of day (d+1). The second forecasting model, which predicts hourly load demand for hour h of day (d+1), is based on past values

(load demand in hour *h* of day *d*-1 and day *d*-7), type of day (the seven dummy variables explained above) and forecasted values for hour *h* of day *d*+1 of the selected weather variables.

3 CASE STUDY

The proposed methodology was applied to the data corresponding to an electric substation 66/13.2 kV located in north of Spain, which feeds around three thousand electric consumers (domestic, commercial and industrial loads) and connects to the grid a PV plant with a capacity near 2 MW. The PV plant is composed of approximately 150 two-axis solar trackers. The available data correspond to the time series of hourly PV power generation and hourly ND for 30 months. These series were completed with forecasts of the set of selected weather variables obtained with an NWP model for each hour included in the time series. The weather forecasts corresponded to the values obtained from the NWP model at the first hours of the previous day (around 6:00 a.m.). The forecasts for the four nearest points of the grid of analysis used by the NMP model were interpolated to obtain the weather forecasts for the location of the electric substation (very near to the consumers).

The data was initially divided into two sets, the first one corresponding to the first 24 months that was used as training data set (for the MLP models) or adjustment data set (for the ARX models), and the second one with the data of the last six months that was used as testing data set.

Table 1 shows the number of neurons in the hidden layer for the selected MLP forecasting model. As it was mentioned in the section 2, we trained MLP models with neurons in the hidden layer in the range 3 to 40, selecting that model which achieved the lowest RMSE with all the training data set.

Table 1. Forecasting MLP models characteristics.

Model	Method	Forecasting	Number of neurons
MLP	Direct	Net Demand	25
	Indirect	PV generation	11
		Load Demand	21

Table 2 presents the variables used in the selected ARX model. As it was mentioned in section 2 the PV forecasting ARX model only uses forecasted values of weather variables, however the load demand forecasting model and the ND forecasting model use, in addition, past values and seven dummy variables. It should be noted that some of the selected input variables may have influence on the output of some of the models (x) but not in others (0).

Table 3 shows the results for the forecasting error obtained, using all the testing data set, with the different forecasting models and with the two strategies of prediction analysed (direct and indirect forecasting). As show in the Table 3 the forecasting model with the lowest RMSE and MAPE is a MLP. In addition, the results of the Table 3 show that the better strategy for the prediction of ND is the indirect method.

Table 2. Considered variables in ARX models

Variable	Forecasting		
	Direct Method	Indirect Method	
	Net Demand	Load Demand	PV generation
Temperature	x	x	x
Horizontal irradiance	x	x	x
Pressure	x	x	x
Wind speed	x	0	x
Wind direction	0	0	x
Relative humidity	x	x	x
Clouds fraction	x	x	x
Rainfall	x	x	x
Monday	x	x	Not considered
Tuesday-Wednesday-Thursday	x	x	Not considered
Friday	x	x	Not considered
Saturday	0	x	Not considered
Sunday	x	x	Not considered
Regional or national festivity	x	x	Not considered
Local festivity	x	x	Not considered
Actual in day-1	x	x	Not considered
Actual in day-7	x	x	Not considered

Table 3. RMSE and MAPE from ARX and MLP models.

Model	Method	RMSE (kW)	MAPE (%)
ARX	Direct	631.93	15.60
	Indirect	432.18	10.33
MLP	Direct	422.90	9.33
	Indirect	391.81	8.94

Figure 2 and Figure 3 show the hourly forecasts and actual values for net demand in a week of autumn 2010 (from Sunday to Saturday) using the direct and the indirect methods with the selected ARX model.

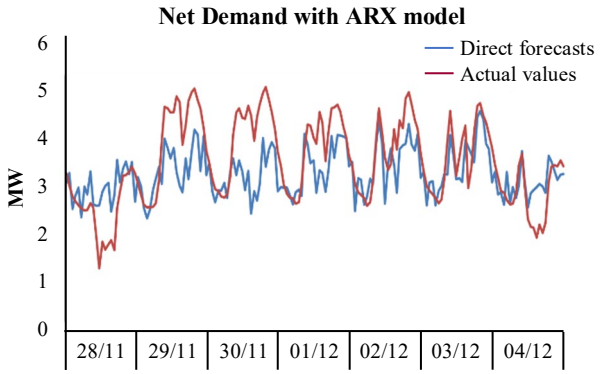

Fig. 2. Actual and direct forecasts for hourly net demand with ARX model for a week in autumn.

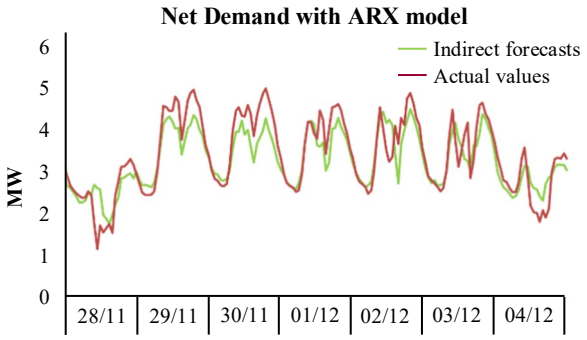

Fig. 3. Actual and indirect forecasts for hourly net demand with ARX model for a week in autumn.

Figure 4 and Figure 5 show the hourly forecasts and actual values for ND in the same week of autumn using the direct method and the indirect method and the selected MLP model.

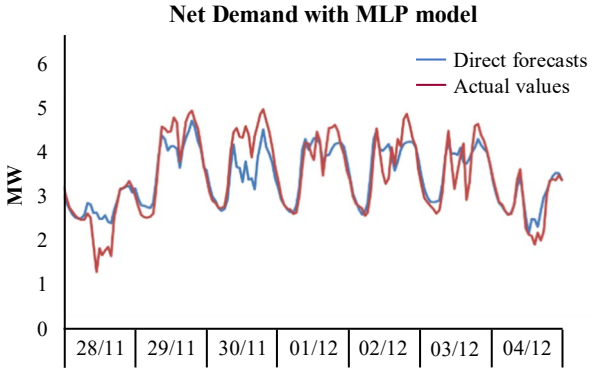

Fig. 4. Actual and direct forecasts for hourly net demand with MLP model for a week in autumn.

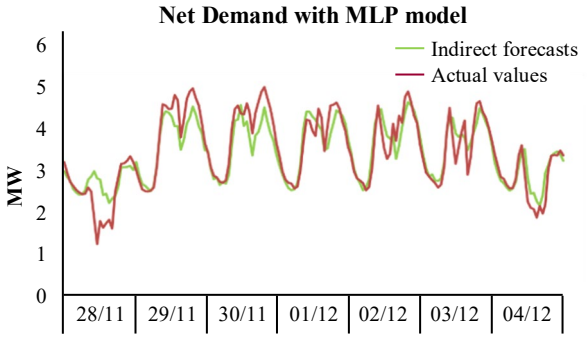

Fig. 5. Actual and indirect forecasts for hourly net demand with MLP model for a week in autumn.

Figures 1-5 show that indirect prediction with the MLP model for the ND (Figure 5) is the method that achieves the best results.

4 CONCLUSIONS

A forecasting methodology for the net demand in an electric distribution substation has been presented in this paper. The correct prediction of the ND for the day-ahead is important in order to operate the distribution network and to participate in the electricity market. Two different forecasting methods are explored: the first one, the direct prediction of the ND, and the second one, the indirect prediction using the forecasts of the PV power generated and of the power consumption. Both methods use, as explanatory variables, previous values of the predicted variable, weather forecasts and dummy variables, and both use the same techniques to build the forecasting models (ARX and MLP based models). This methodology is applied for an actual electric substation which feeds different types of electric consumers and connects a PV generation plant to the electric network. The results show that indirect prediction is more adequate for this task.

The authors would like to thank the "Ministerio de Economia y Competitividad" of the Spanish Government for supporting this research under the project ENE2016-78509-C3-3-P and the ERDF funds of the European Union.

References

1. H. Lan, C. Zhang, Y. Y. Hong, Y. He, S. Wen, Appl Energy **247**, 389 (2019)
2. K. Zor, O. Timur, A. Teke, *6th Int Youth Conf Energy, IYCE* (2017)
3. M. López, C. Sans, S. Valero, C. Senabre, Energies **12**, 7 (2019)
4. Y. He, Q. Xu, J. Wan, S. Yang, Energy **114**, 498 (2016)
5. P. Bacher, H. Madsen, H.Alborg, Sol Energy **83**, 1772, (2009)
6. H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, Y. Du, IEEE Access **7**, 78063 (2019)
7. M. Bagheri, K. Suieubek, O. Abedinia, M. S. Naderi, and M. S. Naderi, *Proc - 2018 IEEE Int Conf Environ Electr Eng* 8493990 (2018)
8. D. W. van der Meer, M. Shepero, A. Svensson, J. Widén, J. Munkhammar, Appl Energy **213**, 195 (2018)
9. L.A. Fernandez-Jimenez, A. Muñoz-Jimenez, A. Falces, M. Mendoza-Villena, E. Garcia-Garrido, P. M. Lara-Santillan, E. Zorzano-Alba, P. J. Zorzano-Santamaria, Renew Energy **44**, 311 (2012)

Authors' background

Your Name	Title*	Research Field	Personal website
Eduardo Garcia-Garrido	Senior lecturer	Power System planning Renewable Energy	https://investigacion.unirioja.es/investigadores/73/detalle
Montserrat Mendoza-Villena	Associate professor	Power System planning Renewable Energy	https://investigacion.unirioja.es/investigadores/126/detalle
Pedro M. Lara-Santillan	Associate professor	Power System planning Renewable Energy	https://investigacion.unirioja.es/investigadores/105/detalle
Enrique Zorzano-Alba	Associate professor	Power System planning Renewable Energy	https://investigacion.unirioja.es/investigadores/195/detalle
Alberto Falces	Lecturer	Power System planning Renewable Energy	https://investigacion.unirioja.es/investigadores/223/detalle

*This form helps us to understand your paper better, the form itself will not be published.

*Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor

* Corresponding author: eduardo.garcia@unirioja.es