PhotoVoltaic Cost reduction, Reliability,
Operational performance, Prediction and Simulation

Title: COMPARATIVE STUDY OF NONPARAMETRIC AND PARAMETRIC PV MODELS TO FORECAST AC POWER OUTPUT OF PV PLANTS

Authors: Marcelo Pinho Almeida, Mikel Muñoz, Iñigo de la Parra, Oscar Perpiñán and Luis Narvarte

Paper in proceedings of conference: 31st European Photovoltaic Solar Energy Conference and Exhibition, Hamburg, 2015

COMPARATIVE STUDY OF NONPARAMETRIC AND PARAMETRIC PV MODELS TO FORECAST AC POWER OUTPUT OF PV PLANTS

Marcelo Pinho Almeida[a], Mikel Muñoz[b], Iñigo de la Parra[b], Oscar Perpiñán[c,d], Luis Narvarte[d]

[a] Instituto de Energia e Ambiente
Universidade de São Paulo
Avenida Professor Luciano Gualberto 1289, 05508-010 São Paulo, Brazil

[b] Departamento de Ingeniería Eléctrica y Electrónica
Universidad Pública de Navarra
Campus de Arrosadía, 31006, Pamplona, Spain

[c] Departamento de Ingeniería Eléctrica, Electrónica, Automática y Física Aplicada
Escuela Técnica Superior de Ingeniería y Diseño Industrial
Universidad Politécnica de Madrid
Ronda de Valencia 3, 28012 Madrid, Spain

[d] Instituto de Energía Solar
Universidad Politécnica de Madrid
Ciudad Universitaria s/n, Madrid, Spain

E-mail: marcelopa@iee.usp.br

ABSTRACT: In this paper, a comparison between two approaches to predict the AC power output of PV systems is carried out in terms of forecast performance. Each approach uses one of the two main types of PV modeling, parametric and nonparametric, and both use as inputs several forecasts of meteorological variables from a Numerical Weather Prediction model. Furthermore, actual AC power measurements of a PV plant are used to train the nonparametric model, to adjust the parameters of the different PV components models used in the parametric approach and to assess the quality of the forecasts. The approaches presented similar behavior, although the nonparametric approach, based on Quantile Regression Forests, showed smaller biased errors due to the machine learning tool used.

Keywords: PV output power forecast, Numerical Weather Prediction, Parametric PV model, Nonparametric PV model

1 INTRODUCTION

Forecast procedures to predict the AC power delivered to the grid by large ground mounted PV plants and smaller BIPV or BAPV systems are important for both plant owners and electric system operators in order to minimize technical risks and expenses related to the uncertainty of generation and to maximize profits.

A PV plant can be seen as a box with several inputs (irradiation, temperature and wind speed, for example) and one output, the AC power injected into the electrical grid. Two main types of approaches can be employed to estimate the AC power output given the required inputs:

- The parametric approach, which conceives the PV system as a white box where each subsystem can be modeled using a collection of parameters.
- The nonparametric approach, which conceives the PV system as a black box. It does not presume any knowledge of internal characteristics and processes of the PV system. Instead, it is a data-driven model that estimates the behavior of the system from a historical time series of inputs and outputs.

This paper presents the results of a comparative study of these two approaches, not to elect the "best" one to be used with PV power forecast, but to present their pros and cons. Predictions of both approaches have been compared with measured AC power from a PV plant as described in Section 4, and the results are presented in Section 5.

2 THE PARAMETRIC APPROACH

A parametric PV model relies on a set of sub-models to compute the AC power injected into the electrical grid, namely:

- Decomposition model that estimates diffuse and beam components from the global irradiance on the horizontal plane as the input.
- Transposition and shading models that estimate the effective irradiance on the generator plane from the diffuse and beam irradiances on the horizontal plane as the input.
- PV generator model that estimates DC power from the effective irradiance on the generator plane and the ambient temperature.
- Inverter model that estimates AC power from the DC power.
- Wiring and electrical equipment (transformers) models that estimate Joule and conversion losses in the way from the PV generator and inverter to the electrical grid.

This modeling requires precise and detailed information about the characteristics and behavior of each relevant component of the PV plant. This information is not always available so some simplifications and assumptions are needed, with the subsequent uncertainty in the output of these models. Consequently, the accuracy and precision of the estimations of a parametric model is driven by the performance of each sub-model and the accuracy of the measurements or estimations of each parameter, together with the accuracy of the irradiation and temperature (forecasts).

The main advantage of a parametric model is the possibility to compute the AC power output prior the construction of a PV plant, during the project and planning stages, by using, for example, the nameplate characteristics of the PV plant components. However, regardless if the calculations are made before or after the construction, the model always needs reliable parameters, i.e., it is crucial to know the internals of the PV plant as much as possible.

This scheme has been adopted in recent researches to forecast PV power. Lorenz et al. [1] derives solar irradiance forecasts based on forecasts of the global model of the European Centre for Medium-Range Forecasts (ECMWF) with a post-processing procedure, and PV power is derived with a set of physical modeling steps. Pelland et al. [2] uses photovoltaic simulation models with PV forecasts based on the spatially averaged solar irradiance forecasts derived from post-processing of a global numerical weather prediction model, namely Environment Canada's Global Environmental Multiscale (GEM) model.

The parametric approach analyzed in this paper uses as input variables predicted ambient temperature (T_a) and global horizontal irradiance (G_0). It has two steps:

- Step 1: Transform global horizontal irradiance into effective irradiance in the plane of the PV array (G_i) and then both ambient temperature and global horizontal irradiance into cell temperature (T_c).
- Step 2: Simulate the losses in each element of the PV installation.

Fig. 1 shows the diagram of a general configuration of a grid-connected PV system, which is composed by a PV generator, an inverter (MPPT + DC/AC converter), and a low voltage/medium voltage (LV/MV) transformer, that was considered in this study.

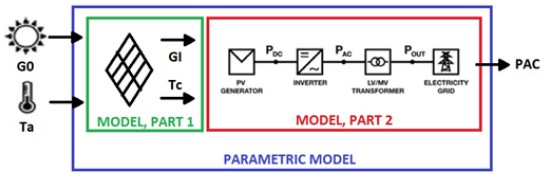


Figure 1: Diagram of a general configuration of a grid-connected PV system

The first part of Step 1 consists in estimating diffuse and beam components from the global irradiance on the horizontal plane. For this purpose, several hourly global-diffuse correlations have been proposed in the literature [3, 4]. In this paper, the Erbs correlation has been used. The decomposition model represents the first source of inaccuracy of the parametric model.

Other two calculation steps are then required: the translation of irradiance values from the horizontal surface to the plane of the PV modules and the reduction of generated power due to losses caused by shading, dirt, incidence and spectrum. The following sequence of calculations has been implemented:

- Position of the Sun, position of the PV generator surface, and incidence angle [5].
- Geometric shading on the PV generator.
- Irradiance on the PV generator plane [6, 7].
- Dirt and incidence losses [8].
- Shading losses [9].
- Spectral corrections [10].

Cell temperature is calculated on the basis of the ambient temperature and in-plane irradiance using the well-known Eq. 1.

$$Tc = Ta + \frac{NOCT - 20}{800} Gi \qquad (1)$$

Now considering Step 2, the PV generator performance has been modeled using the formulation proposed in [11] and [12].

The inverter is characterized by its nominal output power (P_I) and three experimental parameters (k_0, k_1 and k_2), which are used to calculate its efficiency, η_I, according to [13].

The power efficiency of the transformers, η_T, can be expressed as a function of the output power, P_{out}, using Eq. 2.

$$\eta_T = \frac{P_{out}}{P_{AC}} = \frac{P_{out}}{P_{out} + P_{Core} + P_{Cu}} \qquad (2)$$

where P_{Core} is the core losses and P_{Cu} is the copper losses, which can be calculated with Eq. 3.

$$P_{Cu} = P_{Cu,nom} \cdot \left(\frac{P_{out}}{P_T}\right)^2 \qquad (3)$$

where $P_{Cu,nom}$ is the copper losses when the transformer operates at its nominal output power, P_T. Power losses in DC and AC wiring are calculated using equations that are analogous to Eq. 3.

The parameters used in these mathematical models are mainly obtained from standard information, provided by manufacturers or promoters, which may be verified experimentally by on-site quality control testing procedures.

3 THE NONPARAMETRIC PV MODEL

Nonparametric PV models use only historical time series of meteorological variables and AC power measurements, so its accuracy depends mainly on the quality of the data. To illustrate how this feature could be useful, let's suppose that an electric system operator needs estimations of future generation of a PV plant, but he does not know anything about the plant, not even its nominal peak power. As system operators normally have access to the records of power output of generation plants, this data could be used to solve this problem. However, this characteristic also leads to its main disadvantage: the PV plant must exist and be operational for some time.

One interesting advantage of a nonparametric model is the potential to compensate systematic errors associated to the inputs. For example, if irradiance data has a systematic error, the model will learn to associate the incorrect irradiance with the correct AC power output value during the training process. When supplied with new data from the same source, the output will not be compromised if the same error persists.

The nonparametric approach has been implemented in several recent researches. Bacher et al. [14] forecasts hourly values of AC power of PV systems for horizons of up to 36 hours using adaptive linear time series models. Mandal et al. [15] forecasts one-hour-ahead power output of a PV system using a combination of wavelet transform and neural network techniques by incorporating the interactions of PV system with solar radiation and temperature data. Pedro and Coimbra [16] predicts 1 and 2 h-ahead solar power of a PV system comparing several forecast techniques without exogenous inputs such as Auto-Regressive Integrated Moving Average, k-Nearest-Neighbors, Artificial Neural Networks, and Neural Networks optimized by Genetic Algorithms. Zamo et al. [17] analyzes a mix of eight statistical methods to forecast PV power one day ahead in an hourly basis, and the Random Forests method presents the best results.

The nonparametric model analyzed in this paper was extensively detailed and validated in [18]. It forecasts AC power one day ahead with hourly resolution using Quantile Regression Forests (QRF) and gives statistical information about the quantiles of the hourly prediction. Besides, this study contributes with an analysis on how additional variability indexes, daily clearness index (*KTd*), training set length, training set selecting method and different configurations of predictors influence on the final results. Its methodology is as follows:

- Previous AC power measurements from a PV plant are collected.
- Forecasts of a set of Weather Research and Forecasting (WRF) variables (solar radiation, cloud cover, temperature, wind speed, etc.) from a Numerical Weather Prediction (NWP) model run by a meteorological institute are downloaded.
- Each WRF variable is processed to extract information about the value at the location of interest and its relation with the surrounding locations and previous forecasts. In addition, three calculated variables describing the Sun-Earth geometry are included in the predictor set: azimuth angle, altitude angle, and extra-terrestrial irradiance on the horizontal plane.
- The time series of processed WRF variables and AC power measurements is divided into two time series: train and test. The train time series comprises past values of both WRF variables and AC power, whereas the test time series contains only present WRF variables from the NWP model (forecasts).
- A machine learning tool (QRF) is trained with the train time series.
- Predictions of the median (quantile 0.5) and a confidence interval (quantiles 0.1 and 0.9) for the AC power are generated with the test time series.

Its code is freely available from the repository https://github.com/iesiee/PVF, which itself is a R package named *PVF* [19]. An online toolbox that implements this methodology is available at http://vps156.cesvima.upm.es:3838/predictPac.

4 COMPARISON PROCEDURES

Each approach has a specific performance evaluation, but they share the same inputs and desired outputs. The inputs are irradiance (*swflx*) and ambient temperature (*temp*) forecasts obtained from Meteogalicia, a meteorological institute of the Xunta de Galicia (Spain) that publishes regularly results from a regional mesoscale Numerical Weather Prediction (NWP) model, the Weather Research and Forecasting (WRF) [20].

The output is a database of real AC power measurements from a PV plant situated in southern Portugal, with a 5-s resolution measurement. It has an azimuthal one-axis tracker, with a receiving surface tilted 45°. The database was reduced to 1-h resolution due to the restrictions of the weather forecast data used. Table I summarizes the main characteristics of this PV plant.

Table I: PV plant characteristics

Peak Power (kWp)	Rated Power (kW)	Area (Ha)
45,600	38,500	250

A model performance is commonly evaluated by quantifying the discrepancy between forecasts and actual observations through the use of different statistics [21]. Because each performance statistic characterizes a certain aspect of the overall model performance, a complete evaluation needs the combined use of a collection of statistics tools. In this paper, the Mean Bias Error (*MBE*), the Root Mean Square Error (*RMSE*) and the Mean Absolute Error (*MAE*) will be used.

The performance of the nonparametric approach has been assessed using a leave-one-out cross-validation procedure:

- One day is extracted from the database to be the test set.
- The training set is constructed with 30 days extracted from the remaining days of the data set, according to the similarity between the empirical distribution function of the irradiance forecast for the day to be predicted and the day included in the database. These configurations were selected due to the good performance presented in [18].
- The QRF is trained with the training set and hourly AC power is predicted.
- The error between these predictions and AC power measurements for the test day is characterized with the performance statistics.

On the other hand, the performance of the parametric approach has been assessed with the following procedure:

- Hourly AC power for every day from the database is predicted. In order to do it, the inputs are entered in the sequence of mathematical models that represent the behavior of the PV system.
- The daily error between predictions and AC power measurements is characterized with the performance statistics.

5 RESULTS

The performance procedures were repeated for every day in the dataset, resulting in a massive collection of performance statistics. For ease of understanding, the results of each performance statistic have been aggregated with the quantiles 0.25, 0.5 and 0.75, hereafter denominated $QS_{.25}$; $QS_{.5}$ and $QS_{.75}$, respectively, to distinguish them from the quantiles of the predictions.

The results are grouped according to the *KTd* into three classes: cloudy days $0 \leq KTd < 0.532$, partially clouded days $0.532 \leq KTd < 0.678$ and clear days $0.678 \leq KTd \leq 1$. The ranges of *KTd* were selected so that each class comprises one third of the total number of days present in the database.

5.1 Statistical comparison

To make comparison between simulations easier, *MBE*, *RMSE* and *MAE* have been normalized in order to fall in a more restricted range of values. In statistic studies, it is common to normalize these statistics to the range, *max(**O**) – min(**O**)*, or the mean, *mean(**O**)*, of the observations *(**O**)*. For a statistical comparison, the first option was chosen to ensure most of the values fall in a range between 0 and 1. Therefore, the normalized statistics are *nMBE*, *nRMSE* and *nMAE*.

Tables II and III show the statistical performances of the parametric and the nonparametric approaches, respectively.

Table II: Quantiles $QS_{.25}$, $QS_{.5}$ and $QS_{.75}$ of the performance statistics for each *KTd* class using the parametric approach

Statistic	$0 \le KTd < 0.531$			$0.531 \le KTd < 0.687$			$0.687 \le KTd \le 1$		
	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$
nMBE	22.5%	**13.1%**	39.6%	-3.4%	**13.5%**	27.2%	0.5%	**7.0%**	12.1%
nRMSE	45.4%	**59.9%**	87.5%	26.8%	**34.4%**	44.0%	14.2%	**17.6%**	21.1%
nMAE	36.8%	**51.1%**	74.9%	21.3%	**27.5%**	35.6%	11.2%	**13.3%**	16.5%

Table III: Quantiles $QS_{.25}$, $QS_{.5}$ and $QS_{.75}$ of the performance statistics for each *KTd* class using the nonparametric approach

Statistic	$0 \le KTd < 0.531$			$0.531 \le KTd < 0.687$			$0.687 \le KTd \le 1$		
	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$	$QS_{.25}$	$\mathbf{QS_{.5}}$	$QS_{.75}$
nMBE	-17.5%	**3.8%**	18.7%	-7.9%	**0.2%**	5.6%	-4.1%	**0.2%**	2.3%
nRMSE	28.4%	**35.3%**	46.4%	8.8%	**18.3%**	26.6%	3.3%	**5.7%**	12.4%
nMAE	21.9%	**27.7%**	38.1%	6.1%	**13.4%**	21.1%	2.5%	**3.8%**	7.9%

The parametric approach shows a statistical performance somewhat worse, but this result is expected due to the uncertainties and errors mainly related to the quality of the input variables (weather forecasts), which can be partially suppressed by the QRF used in the nonparametric approach.

5.2 Daily energy production uncertainty comparison

Although the prediction of AC power output of PV plants using both parametric and nonparametric approaches have good statistical performance, some further analysis is necessary to assess the impacts on daily energy prediction. Two scenarios are accounted here: markets that penalize the daily energy error, for which the *MBE* is appropriate, and markets that penalize the hourly energy error, for which the *MAE* is preferred. In this context, these metrics are more useful if presented as an energy ratio, and thus they were normalized respect to the daily measured energy, resulting in *cvMBE* and *cvMAE*, respectively. Table IV presents the results for the quantile $QS_{.5}$, weighted with the energy generated by the PV plant under the corresponding *KTd* class.

Table IV: Weighted errors of energy forecast according to the *KTd* class

Statistic	Approach	$0 \le KTd < 0.531$	$0.531 \le KTd < 0.687$	$0.687 \le KTd \le 1$
cvMBE	Parametric	2.9%	4.7%	3.4%
	Nonparametric	1.2%	0.1%	0.1%
cvMAE	Parametric	9.3%	9.0%	6.1%
	Nonparametric	8.7%	6.5%	2.2%

Values of *cvMBE* for the nonparametric approach are smaller than those for the parametric approach, but this is expected due to the machine learning tool used. Nevertheless, the values obtained for the parametric approach are very good as well. Total daily energy is forecasted with a weighted *cvMBE* of less than 5% for both approaches and all *KTd* classes. Considering the nonparametric approach, the weighted *cvMBE* is less than 2% for cloudy days and it is only 0.1% for clear days.

In terms of hourly prediction, the performances of the two approaches are also good and even more alike, especially for cloudy or partially clouded days. Most of the difference between their performances is related to the bias the parametric method presents due to the errors of the forecasts used as inputs. The overall weighted *cvMAE* is less than 9.5% and it is around 2% for clear days using the nonparametric model.

6 CONCLUSION

A comparison between two approaches to forecast the AC power output of a PV system, one using a parametric PV model and the other a nonparametric PV model based on QRF, was made. Some points can be highlighted:

- Both approaches have state-of-the-art statistical performance. Besides, their performances in terms of daily and hourly energy prediction are very good.
- The two approaches have very similar performance, but the nonparametric is slightly better given the conditions of this study, especially for the biased metrics (*nMBE* and *cvMBE*) due to the machine learning tool used (Quantile Regression Forests).
- Daily energy production is forecasted with a weighted *cvMBE* of less than 5%. Considering the nonparametric model, this statistic is below 2% for cloudy days and it is only 0.1% for clear days.
- In terms of hourly prediction, most of the difference between approaches' performances is due to the bias the parametric method presents. The overall weighted *cvMAE* is less than 9.5% and it is around 2% for clear days using the nonparametric model.

When selecting one of the approaches, not only the accuracy must be considered, but also the application and the variables and parameters available.

ACKNOWLEDGEMENTS

This work has been partially financed by the Seventh Framework Programme of the European Commission with the Project Photovoltaic Cost Reduction, Reliability, Operational Performance, Prediction and Simulation (PVCROPS—Grant Agreement No. 308468).

REFERENCES

[1] E. Lorenz, T. Scheidsteger, and J. Hurka, Regional PV power prediction for improved grid integration. Progress in Photovoltaics: Research and Applications, (2010) 757–771.

[2] S. Pelland, G. Galanis, and G. Kallos, Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model. Progress in Photovoltaics: Research and Applications, 21(3) (2013) 284–296.

[3] J. F. Orgill, and K. G. T. Hollands, Correlation equation for hourly diffuse radiation on a horizontal surface. Solar Energy, 19(4) (1977) 357-359.

[4] D. G. Erbs, S. A. Klein, and J. A. Duffle, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Solar Energy, 28(4) (1982) 293-302.

[5] E. Lorenzo, Energy collected and delivered by PV modules. Photovoltaics Science and Engineering, (2011) 984–1042.

[6] J. E. Hay, and D. C. Mckay, Estimating Solar Irradiance on Inclined Surfaces : A Review and Assessment of Methodologies. International Journal of Solar Energy, 3(4-5) (1985) 203–240.

[7] R. Perez et al., A New simplified version of the Perez diffuse irradiance model for tilted surfaces. Solar Energy, 39(3) (1987) 221–231.

[8] N. Martin, and J. M. Ruiz, Calculation of the PV modules angular losses under field conditions by means of an analytical model. Solar Energy Materials and Solar Cells, 70 (2001) 25–38.

[9] J. M. Ruiz, A New Method for the Spectral Characterization of PV modules. Progress in Photovoltaics: Research and Applications, 310 (1999) 299–310.

[10] M. Fuentes et al., Application and validation of algebraic methods to predict the behavior of crystalline silicon PV modules in Mediterranean climates. Solar Energy, 81(11) (2007) 1396–1408.

[11] D. L. Evans, Simplified method for predicting photovoltaic array output. Solar Energy, 27(6) (1981) 555–560.

[12] C. R. Osterwald, Translation of device performance measurements to reference conditions. Solar Cells, 18(3–4) (1986) 269–279.

[13] M. Jantsch, H. Schmidt, and J. Schmid, Results on the concerted action on power conditioning and control. In 11th European photovoltaic Solar Energy Conference (1992) 1589–1592.

[14] P. Bacher, H. Madsen, H. A. Nielsen, Online short-term solar power forecasting. Solar Energy, 83 (2009) 1772–1783.

[15] P. Mandal, S. T. S. Madhirab, A. Ul haquec, J. Mengc, R. L. Pinedaa, Forecasting Power Output of Solar Photovoltaic System Using Wavelet Transform and Artificial Intelligence Techniques. Procedia Computer Science, 12 (2012) 332–337.

[16] H. T. C. Pedro, C. F. M. Coimbra, Assessment of forecasting techniques for solar power production with no exogenous inputs. Solar Energy, 86 (2012) 2017–2028.

[17] M. Zamo, O. Mestre, P. Arbogast, O. Pannekoucke, A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production. Solar Energy, 105 (2014) 792–803.

[18] M. P. Almeida, O. Perpiñán, L. Narvarte, PV power forecast using a nonparametric PV model. Solar Energy 115 (2015) 354 – 368.

[19] M. P. Almeida, O. Perpiñán, PVF. R package version 0.20 (2014). Available at: http://dx.doi.org/10.5281/zenodo.13348.

[20] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, J. G. Powers, A description of the advanced research WRF version 2. Tech. rep., National Center for Atmospheric Research (2005).

[21] C. Gueymard, A review of validation approaches and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. Renewable and Sustainable Energy Reviews, 39 (2014) 1024–1034.

COMPARATIVE STUDY OF NONPARAMETRIC AND PARAMETRIC PV MODELS TO FORECAST AC POWER OUTPUT OF PV PLANTS

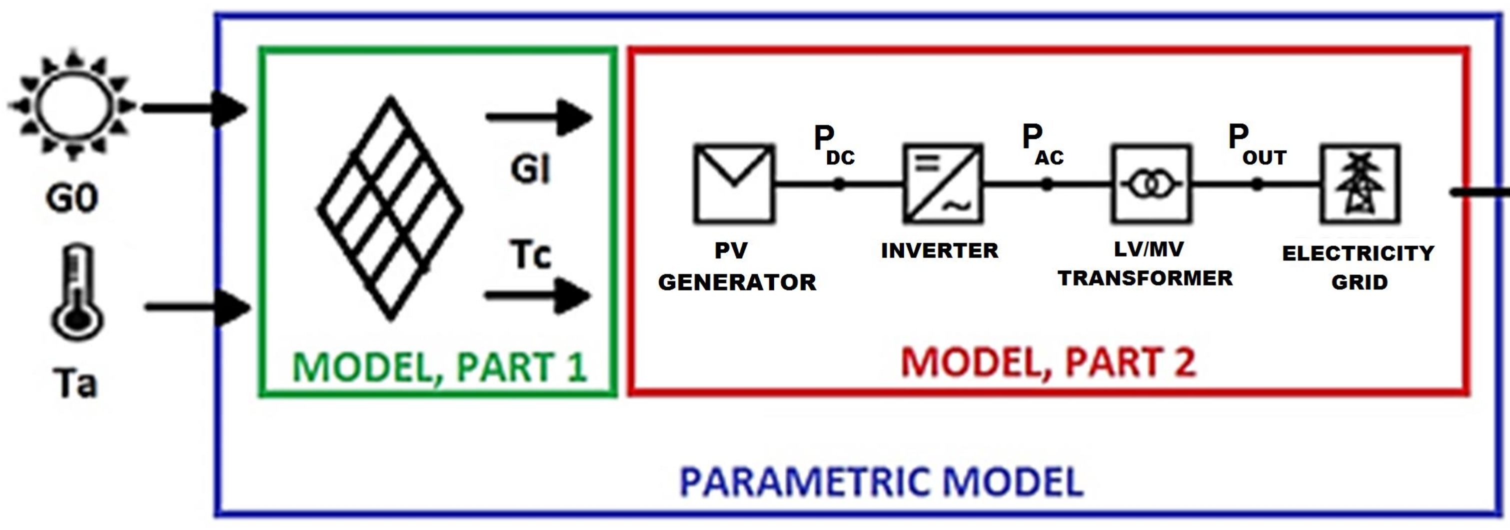
Marcelo Pinho Almeida[a], Mikel Muñoz[b], Iñigo de la Parra[b], Oscar Perpiñán[c,d], Luis Narvarte[d]

INTRODUCTION

The results of a comparative analysis of two PV models applied to AC power output forecast are presented, not to elect the "best" one, but to present their pros and cons. When selecting one of the approaches, not only the accuracy must be considered, but also the application and the variables and parameters available.

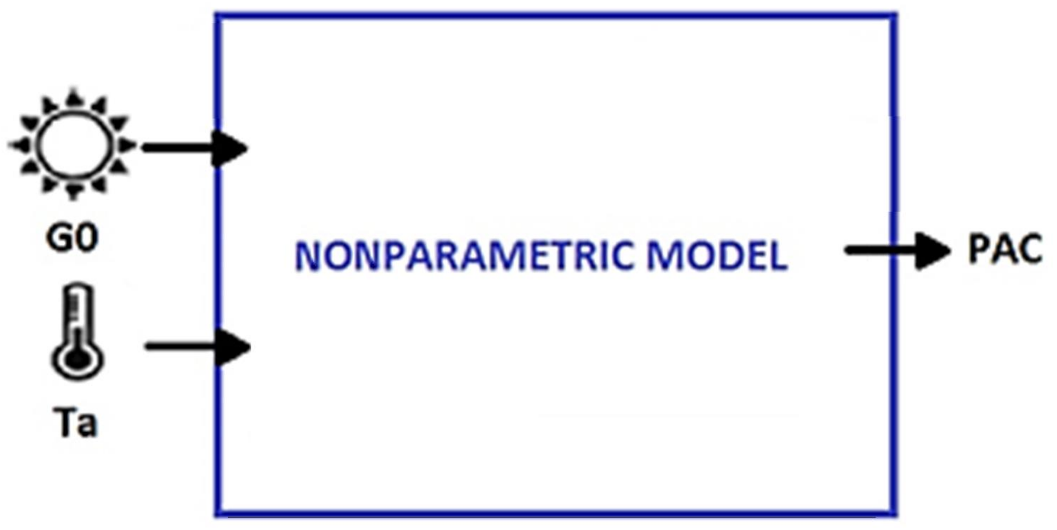
There are two approaches to model a PV system:

- The parametric:



A parametric modeling requires precise and detailed information about the characteristics and behavior of each relevant component of the PV plant. This information is not always available so some simplifications and assumptions are needed, with the subsequent uncertainty in the output of these models.

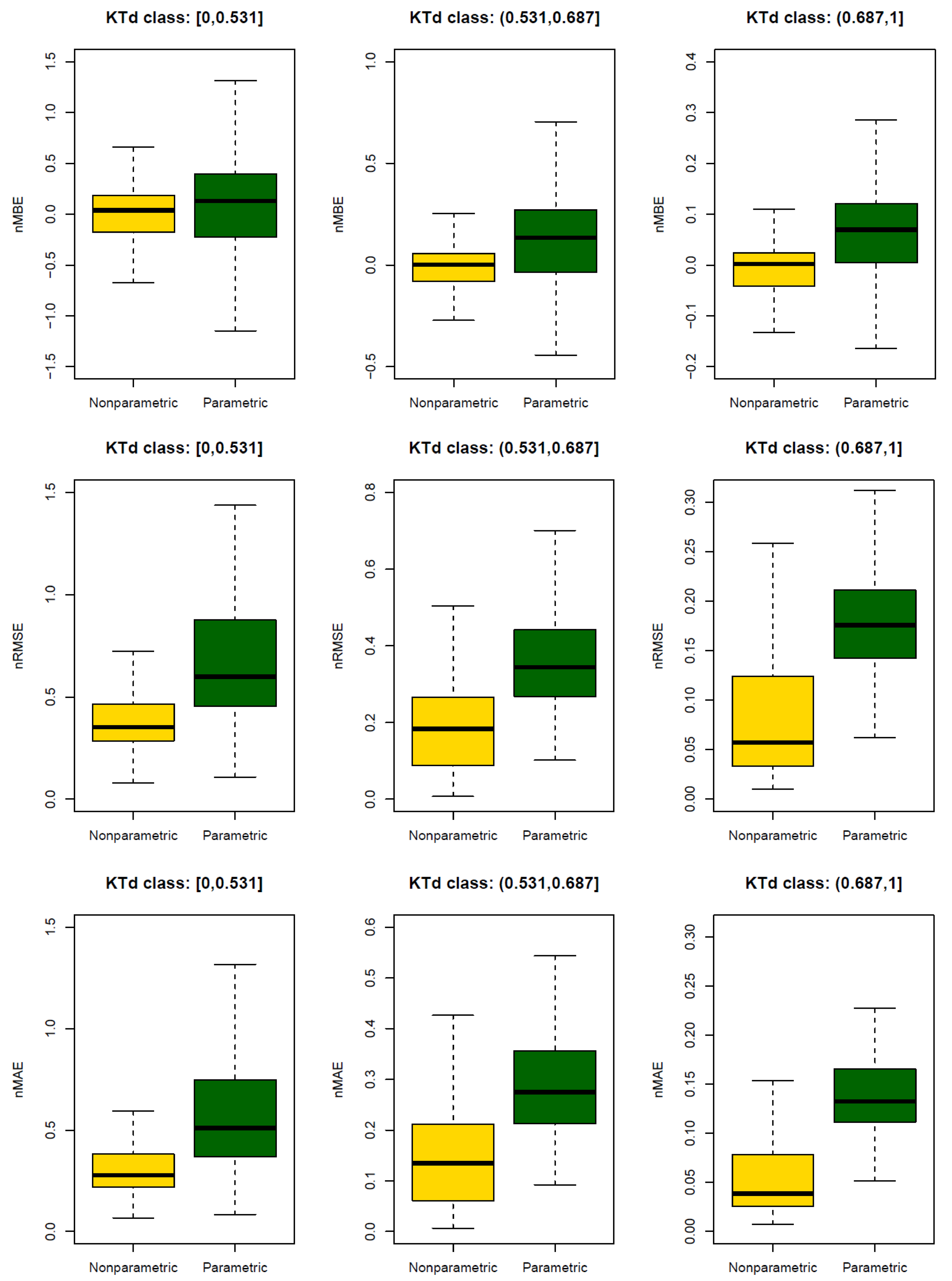
- The nonparametric:



Nonparametric PV models use only historical time series of meteorological variables and AC power measurements, so its accuracy depends mainly on the quality of the data. This characteristic leads to its main disadvantage: the PV plant must exist and be operational for some time. One interesting advantage of a nonparametric model is the potential to compensate systematic errors associated to the inputs.

STATISTICAL COMPARISON

The study was based on real AC power measurements from a 45.6 MWp PV plant situated in southern Portugal. It has an azimuthal one-axis tracker, with a receiving surface tilted 45°.



Errors of energy production forecast weighted by the energy generated by the PV plant under the corresponding *KTd* class were also calculated:

Statistic	Approach	*KTd* class		
		[0, 0.531]	(0.531, 0.687]	(0.687, 1]
cvMBE	Parametric	2.90%	4.70%	3.40%
	Nonparametric	1.20%	0.10%	0.10%
cvMAE	Parametric	9.30%	9.00%	6.10%
	Nonparametric	8.70%	6.50%	2.20%

CONCLUSION

- The PV models used in this study have state-of-the-art statistical performance in terms of daily and hourly energy prediction.
- The two approaches presented similar performance, but the nonparametric is slightly better given the conditions of the study.
- Due to the machine learning tool the nonparametric approach is based on (Quantile Regression Forests), it presented low values for biased metrics (*nMBE* and *cvMBE*).
- The lower statistical performance of the parametric approach is expected due to the uncertainties and errors mainly related to the quality of the input variables (weather forecasts), which can be partially suppressed by the machine learning tool used in the nonparametric approach.
- Daily energy production is forecasted with a weighted *cvMBE* of less than 5%. Considering only the nonparametric model, this statistic is below 2% for cloudy days and it is virtually none for clear days.

[a] Instituto de Energia e Ambiente
Universidade de São Paulo

[b] D. de Ingeniería Eléctrica y Electrónica
Universidad Pública de Navarra

[c] ETS de Ingeniería y Diseño Industrial
Universidad Politécnica de Madrid

[d] Instituto de Energía Solar
Universidad Politécnica de Madrid