Day-ahead probabilistic photovoltaic power forecasting models based on quantile regression neural networks

L. Alfredo Fernandez-Jimenez, Sonia Terreros-Olarte, Montserrat Mendoza-Villena, Eduardo Garcia-Garrido, Enrique Zorzano-Alba, Pedro M. Lara-Santillan, Pedro J. Zorzano-Santamaria, Alberto Falces
Department of Electrical Engineering
University of La Rioja
Logroño, Spain

***Abstract*—This paper presents the results obtained in the development of probabilistic short-term forecasting models of the power production in a photovoltaic power plant for the day-ahead. The probabilistic models are based on quantile regression neural networks. The structure of such neural networks is optimized with a genetic algorithm which selects the values for the main parameters of the neural network and the variables used as inputs. These input variables are selected among a set of variables which includes chronological, astronomical and forecasted weather variables related to the location of the power plant. The forecasts correspond to quantiles of the hourly power generation in the photovoltaic power plant for the daytime hours of the day-ahead. The forecasts are obtained in the first hours of the day, allowing their use for preparing bid offers for the day-ahead in electricity markets.**

Keywords— short-term forecasting, solar power forecasting, probabilistic forecasts

I. Introduction

The installation of power plants based on renewable energy sources has grown strongly in recent years. Governmental policies and social acceptance have stimulated this growth. By 2050, the production in power plants based on renewable energy is expected to meet all the demand [1]. Wind and solar photovoltaic energies are postulated as the more promising sources.

Photovoltaic (PV) systems are the most direct way to transform solar radiation into electric power. PV systems are widely used to produce electricity in areas isolated from electric power networks and can operate connected to the grids in zones where they are available. The global capacity worldwide in PV systems has reached at least 303 GW at the end of 2016 [2].

The growing capacity of PV plants and their increased penetration in the power systems, causes these generation facilities have to participate, as any other electric power producer, in the electricity markets. A PV electricity producer who participate in the electricity market follows the same market rules as any other producer, i.e., providing a schedule of its generation for the day-ahead market, and reducing his/her incomes if this schedule is not met. In this sense, it is crucial to dispose of accurate short-term forecasting tools able to provide estimations of the hourly energy production for a forecasting horizon from a few hours to nearly two days.

In the last decade many researchers have done an important effort developing short-term forecasting models for generating plants based on renewable energies, mainly wind farms and photovoltaic plants. Thus, in the international literature can be found in a wide variety of forecasting models for the electric power production in a PV plant [3][4]. Basically, two main approaches can be found in the international literature: indirect and direct forecasting models. Indirect models try to predict the power production in the PV plant by means of a prior forecast of the solar irradiation and a posteriori transformation to electric power, and direct models predict straightly the power production using statistical or machine learning techniques. The use of forecasts of weather variables (solar radiation, temperature, cloud cover, etc.) is widespread in both approaches. These forecasts are obtained from numerical weather prediction (NWP) models, that is, tools able to provide forecasts of the weather variables for the forecasting horizon.

The development of a new family of short-term forecasting models related to the power generation in renewable energy based power plants have captured the attention of researchers in the last few years. Forecasting models able to provide not only the spot or point forecast, i.e., the expected value for the variable of interest, but also additional information about its uncertainty have revealed as more useful to decision-makers related to electricity markets. Probabilistic models overcome this limitation providing information about the uncertainty associated with the predicted value.

Probabilistic models can provide as output values interval forecasts (prediction intervals) or density forecasts. In the energy field, most of the published models are focused on prediction intervals, since they can help to the agents of an energy market to trade with low risks. Quantile regression is one of the most popular techniques used to obtain prediction intervals in the form of conditional quantiles of the response variable.

This paper presents the methodology followed to develop a probabilistic solar power forecasting model for the day ahead. The model was based on the quantile regression neural network

978-1-5386-2085-4/17 $31.00 © 2017 IEEE
DOI 10.1109/EECS.2017.60

(QRNN) described by Taylor [5] and Chen [6], and implemented in the package “qrnn” [7] for the statistical program R [8]. The main parameters that define the structure of the neural network and the input variables, chosen from a set of available input variables, were selected using a genetic algorithm. In the training stage of the neural networks a 5 folds cross-validation scheme was used. The fitness function in the optimization process was the negative value of the average mean square error (MSE) obtained with the 5 cross-validation sets. The structure of the optimized model was used to build QRNN models for a set of quantiles and, finally, applied to a testing data set, no used in the training and optimization process, and its computational results analyzed.

The structure of the paper is as follows: section II outlines the theory of the quantile regression neural networks; section III describes the available input variables; section IV describes the optimization process; section V presents the computational results obtained with the models; finally, section VI presents the conclusions.

II. Quantile Regression Neural Networks

In statistics, quantiles are points that divide the range of the probability distribution of a random variable into contiguous intervals. The term quantile is usually used as synonymous of percentile. Quantile regression tries to estimate quantiles of the conditional distribution of a response variable as functions of observed covariates. The τ-quantile (with $0<\tau<1$) of a random variable corresponds to the value of the variable, q_τ, which accumulates a probability τ that the real value of the variable is equal or less than q_τ.

Linear quantile regression estimates quantiles of a random variable as a linear function of the covariates. Koenker and Bassett [9] defined the θth sample quantile as any solution to the minimization problem expressed in (1), where $\{x_t: t=1,\ldots,T\}$ is a sequence of K-vectors (explanatory variables), $\{y_t: t=1,\ldots,T\}$ is random sample on the regression process $u_t = y_t - x_t \boldsymbol{\beta}$ having distribution function $\boldsymbol{F}$, and $\boldsymbol{\beta}$ is a vector of parameters dependent on θ.

$$\min_{\boldsymbol{\beta}}\left[\sum_{t|y_t\geq x_t\boldsymbol{\beta}}\theta|y_t-x_t\boldsymbol{\beta}|+\sum_{t|y_t<x_t\boldsymbol{\beta}}(1-\theta)|y_t-x_t\boldsymbol{\beta}|\right] \tag{1}$$

The ability of artificial neural networks to approximate any non-linear function make them as candidates to build non-linear quantile regression models. One of the most popular neural networks used for forecasting purposes is the single hidden layer feedforward network, i.e. a set of n inputs, connected to each of m neurons in a single hidden layer, which are connected to an output. Instead of using the linear quantile function expressed in (1), a QRNN model estimates the θth quantile minimizing (2), where $\boldsymbol{X}_t$ represents the input variables, $\boldsymbol{W}$ the weights between the input layer and the hidden layer, $\boldsymbol{V}$ the weights between the hidden layer and the output layer, and $f(\boldsymbol{X}_t, \boldsymbol{V}, \boldsymbol{W})$ the output of the neural network. The other two parameters, λ_1 and λ_2, are regularization parameters which penalize the complexity of the network and avoid overfitting [5]. The optimal values of λ_1 and λ_2 and the number of neurons in the hidden layer can be established by cross validation.

$$\begin{aligned}\min_{\boldsymbol{V},\boldsymbol{W}}\Bigg(&\sum_{t|y_t\geq f(\boldsymbol{X}_t,\boldsymbol{V},\boldsymbol{W})}\theta\,|y_t-f(\boldsymbol{X}_t,\boldsymbol{V},\boldsymbol{W})|\\ &+\sum_{t|y_t<f(\boldsymbol{X}_t,\boldsymbol{V},\boldsymbol{W})}(1-\theta)\,|y_t-f(\boldsymbol{X}_t,\boldsymbol{V},\boldsymbol{W})|\\ &+\lambda_1\sum_{j,i}w_{ji}^2+\lambda_2\sum_j v_j^2\Bigg)\end{aligned} \tag{2}$$

The package “qrnn” [7] implements the QRNN with weight decay regularization, optional left censoring, and ensemble averaging via bootstrap aggregation o bagging, in order to avoid overfitting training data. The weight decay regularization penalizes large weights by adding a quadratic penalty term to the error function, although only the parameter λ_1 in (2) is considered.

III. Available Input Variables

The developed probabilistic forecasting models aim to forecast the hourly power production in a PV plant for the daytime hours of the day-ahead. The forecasts are carried out in the first hours of day D and correspond to the hourly power generation in the PV plant for all the daytime hours of the following day, D+1. The first step in the followed methodology corresponds to the obtaining of the explanatory variables of the variable to be forecasted.

The proposed forecasting model can use three types of input variables: chronological, astronomical and forecasted weather variables. The chronological variables are used to stand for the future instant corresponding to the forecasting horizon. The astronomical variables represent the position of the Sun with respect to the PV panels of the plant. The forecasted weather variables are the expected value of weather variables for the forecasting horizon and for the location of the PV plant. Table I shows the description of the chronological and astronomical variables.

TABLE I. CHRONOLOGICAL AND ASTRONOMICAL INPUT VARIABLES

Name	Description
SD	Sine of year fraction for *D*+1
CD	Cosine of the year fraction for *D*+1
SH	Sine of the day fraction for *h*
CH	Cosine of the day fraction for *h*
SDE	Sine of solar declination angle for *D*+1
CDE	Cosine of solar declination angle for *D*+1
SW	Sine of solar hour angle for *D*+1, *h*
CW	Cosine of solar hour angle for *D*+1, *h*
SA	Sine of solar elevation angle *D*+1, *h*
CA	Cosine of solar elevation angle *D*+1, *h*
SY	Sine of azimut angle *D*+1, *h*
CY	Cosine of azimut angle *D*+1, *h*
CI	Cosine of angle of incidence for *D*+1, *h*
KD	Correction factor for distance Sun-Earth

The fraction of the year (in radians, needed for variables SD and CD) is calculated by (3), where $O(D+1)$ represents the ordinal date of the day D+1. The fraction of the day (needed for variables SH y CH) is calculated by (4), where h represents the GMT hour, h, corresponding to the forecasting horizon. The solar declination, solar hour, solar elevation and azimuth correspond to variables related to the position of the Sun with respect to the location of the PV plant, and the angle of incidence represents the angle between the solar beam radiation and the surface of the PV panels [10]. The correction factor for the distance Sun-Earth represents the variation of such distance with respect to the mean distance and can be calculated by (5).

$$fy = 2\pi O(D+1)/365.25 \quad (3)$$

$$fh = 2\pi h/24 \quad (4)$$

$$\mathrm{KD} = 1 + 0.03344\, cos(fy - 0.04887) \quad (5)$$

The use of weather forecasts has proven its effectiveness in the accuracy of solar power forecasting models for the medium and long-term [11]. The forecasts of these variables are obtained by means of specific forecasting models, the abovementioned NWP models. NWP models are usually implemented as computer programs that simulate the atmospheric dynamics solving kinematic physical equations and estimating a future state from the current conditions. The results correspond to the estimated values (forecasts) for a set of weather variables for the nodes of a 3D grid for different instants in the future. NWP models can be classified according their spatial resolution as global, regional or mesoscale. Global NWP models have a global coverage (worldwide) and provide forecasts with a large temporal horizon (several days), although their temporal and spatial resolution is poor (the interval between two consecutive forecasts can be several hours and the nodes on the earth surface are very distant). Mesoscale NWP models offer proper temporal and spatial resolutions for their use for forecasting purposes in renewable energy applications. Mesoscale NWP models usually utilize the results of a global model as the initial conditions of the atmosphere.

In our study, we have used the forecasts of weather variables obtained with the mesoscale NWP model WRF-NMM [12], which provide us the forecasts for a set of weather variables for the nearest grid node from the location of the solar PV plant. The WRF-NMM is one of the most popular mesoscale NWP model in the scientific community. For the generation of forecasts for the weather variables, we used a two domains scheme. The use of nested domains allowed us to obtain forecasts with better temporal and spatial resolution in the zone where is located the PV plant. Fig. 1 shows the domains used, where the smaller domain, with a spatial resolution of 27 x 48 points on the earth surface is focused on the geographic location of the PV plant, which is located in the south-east of the Iberian Peninsula.

The initial values for the WRF-NMM model were obtained from the results of the GFS 1x1º global NWP model, corresponding to the cycle of the 00:00 GMT hour. The values for a set of weather variables and corresponding to the grid point of domain 2 nearest to the position of the PV plant (less than 1 km), were extracted from the output files from the WRF model. The results were the forecasts of the weather variables, shown in Table II, at half an hour intervals and for the 24 hours of the day

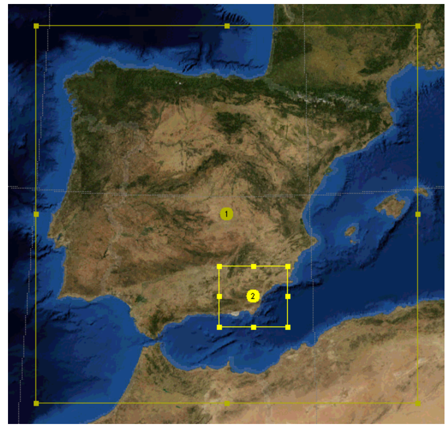

Fig. 4. Domains used for the WRF-NMM model.

ahead. These variable include solar radiation variables (the downward shortwave radiation at the earth surface, RSWIN, and the same value under the condition of clear-sky, RSWINC), cloud coverture at three height levels (low, medium and high), temperature (TH10), humidity (Q10), wind speed (the WRF model provided two orthogonal components at 10 m height, U10 and V10, from which the wind speed, WS10, and direction, represented by SINW and COSW, were calculated) and pressure near the earth surface (PSHLTR). Other variables, as snow cover, were not considered taking into account the position of the PV plant, where the probability of snow is negligible.

The forecasts were available before dawn of the day D, and the forecasts correspond to the variables for day D+1. These variables, obtained at intervals of half an hour, were transformed to average hourly values.

IV. Optimization of the QRNN

The QRNN model includes a set of parameters whose values can be optimized in order to provide better point forecasts for the hourly power generation in the PV plant. These parameters are the number of neurons in the hidden layer, the weight decay regularization factor (or penalty), and the use/not use of bagging in the training process. In previous works with the same data, we discovered that the best results, for the PV plant used in the study, were obtained without using bootstrap aggregation, so, we reduced the parameters to be optimized to the first two.

TABLE II. FORECASTED WEATHER VARIABLES

Name	Description
TH10	10-m potential temperature (K)
Q10	10-m specific humidity
CFRACL	Low cloud fraction (0 to 1)
CFRACM	Middle cloud fraction (0 to 1)
CFRACH	High cloud fraction (0 to 1)

Name	Description
RSWIN	Downward shortwave at Surface (W/m2)
RSWINC	Clear-sky equivalent of RSWIN (W/m2)
WS10	Wind speed at 10 m (m/s)
SINW	Sine of wind direction at 10 m
COSW	Cosine of wind direction at 10 m
PSHLTR	2-m pressure (Pa)

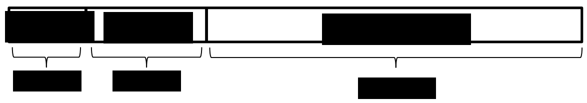

Fig. 5. Structure of the chromosome used in the optimization of the QRNN model.

In order to optimize the values of these parameters, we carried out an optimization process ruled by a genetic algorithm. The selection of the input variables, among those available ones, was also included in the optimization process. We used the functions included in the R package "GA" [13]. The adopted structure for the chromosome was binary digits with Gray encoding for the two parameters, and binary representation for the available input variables. The structure chosen for the chromosome is shown in Fig. 2.

The structure of the chromosome was the following: an array with 38 bits, the first five bits to represent the Gray code of the number of neurons in the hidden layer (form 00000 interpreted as 1 neuron, to 10000 interpreted as 32 neurons), the following 8 bits to represent the Gray code of the weight decay regularization factor or penalty (from 00000000 interpreted as 0, to 10000000 interpreted as 0.1), and the last 25 bits to code binary the input variables used by the model (one per available input variable, with a meaning of included variable if the corresponding bit is 1, and not included if it is 0).

The fitness function calculated the average MSE obtained with the 5 cross-validation data sets. The error corresponds to the difference between the output of the QRNN for the 0.5 quantile and the actual value.

V. COMPUTATIONAL RESULTS

We obtained the input variables described in the previous section for a PV power plant located in the south-east of Spain. The plant is composed of fixed panels with a total capacity around 7 MW. We have disposed of the data with the hourly power production for 35 months. The total number of available input variables was 25 (four chronological variables, ten astronomical variables and eleven forecasted weather variables). Only the data corresponding to daytime hours were considered in the database, which was divided into two sets: the training data set and the testing data set. The training data set included the data corresponding to the first 24 months, and it was used to train the models. The testing data set included the data corresponding to the last 11 months and it was used to evaluate the final models.

An optimization with a total of 100 generations with 80 individuals per generation was carried out using the training data set. Elitism was applied to the best four individuals, the crossover rate was 0.8 and the mutation rate 0.1; the fitness function was the negative value of average MSE with the five sets used as validation data in the 5-folds cross-validation process. The number of iterations to train the QRNN was set in 1500, and each network was trained 3 times, taking as result the best of the three trainings. The QRNN models used left censoring point at zero, just because the hourly power generation in the PV plant couldn't be negative. Fig. 3 plots the evolution

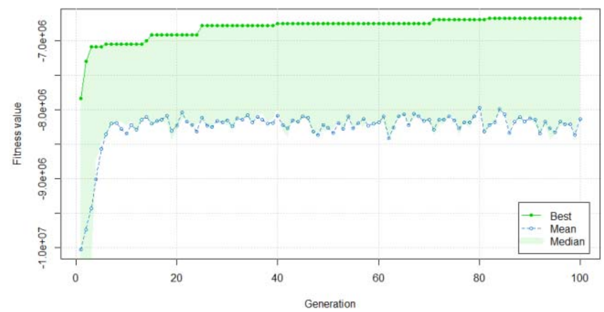

Fig. 6. Evolution of the fitness value in the optimization process.

of the fitness value of the best individual and of the mean value for all the individuals in the generation during the optimization process.

After the optimization process, the resulting best individual used as input variables SW, CW, CA, SY, CY, CI, KD, Q10, CFRACL, CFRACM, CFRACH, RSWIN, RSWINC and SINW. The number of neurons was 6 and the value for the weight decay regularization factor (penalty) was 0.3412.

Once the structure of the QRNN was selected with the optimization process, all the QRNN models corresponding to quantiles 0.05 to 0.95 at intervals of 0.05 were built. All the QRNN models had the same number of neurons in the hidden layer (6), the same value for the weight decay regularization factor (0.3412) and used the same inputs that the best individual obtained in the optimization process. The number of iterations was 1500 and the training repeated three times.

The 19 QRNN models (from quantiles 0.05 to 0.95) enabled us to provided probabilistic forecasts for the hourly power generation in the PV plant for the day-ahead. For example, Fig. 4 plots the results obtained for seven days in the testing set (not used to train the QRNN models). Notice that in the figure only the daytime hours are represented, and the values correspond to the actual value and to the quantiles 0.05 (Q0.05), 0.25 (Q0.25), 0.5 (Q0.5), 0.75 (Q0.75) and 0.95 (Q0.95).

In order to evaluate the performance of the developed QRNN models, we carried out two test, the first related to the point forecast provided by the QRNN for the 0.5 quantile, and the second one related to the reliability of the probabilistic forecasts obtained with the 19 QRNN models.

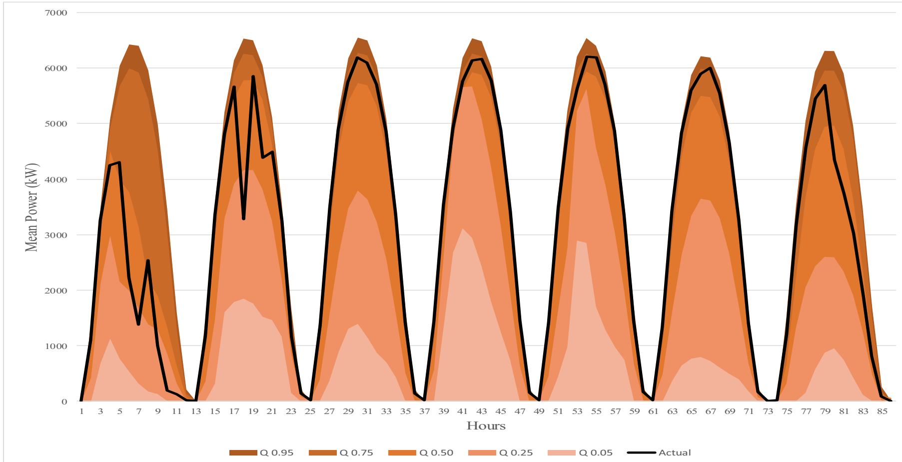

Fig. 4. Probabilistic power generation forecasts for one week in the testing set.

In the first test, we compared the RMSE obtained for the testing data set with the QRNN for the 0.5 quantile, with those obtained with other two point forecasting models. The first one was the persistence model, which provided as forecast the hourly power generation value for the same hour of the day that the corresponding to the forecasting horizon, but in the last day, i.e., the day D-1. The second model was a single hidden layer multilayer perceptron (MLP) neural network, also trained with 5-folds cross-validation, but using all the available variables as inputs. MLP models with different number of neurons in the hidden layer were trained with the backpropagation algorithm. The MLP model with the best results contained 18 neurons in the hidden layer and achieved a RMSE (average value with the five validation data sets) of 952.58 kW. Table III shows the results obtained with the QRNN model for the quantile 0.5 (Q0.5), for the persistence model (PS) and with the MLP model (MLP18) with the training and the testing data sets. As it is shown, Q0.5 model achieved the best result for the testing data set, and outperformed the persistence model results for both data sets.

The second test was carried out to calculate a new indicator related to the reliability of the probabilistic forecasts and designed to give a quantitative measure of such reliability. If the probabilistic forecasts were reliable, the number of occurrences of the hourly power generation variable should be almost the same for all the intervals limited for two consecutive quantiles from the 19 considered. We defined as reliability index the expressed by (6), where where $n_{obs,i}$ and $n_{tar,i}$ are the number of occurrences (in per unit) and the ideal number of occurrences (in per unit) for the interval i (between two consecutive quantiles), and NQ is the total number of intervals.

$$RIN = \left(1 - \sum_{i=1}^{NQ} \left| n_{obs,i} - n_{tar,i} \right| \right) \tag{6}$$

The probabilistic forecasts will be as good as next to 1 is the *RIN* value. In our case, the total number of intervals was 20, the ideal number of occurrences, in per unit, was 0.05, and the obtained *RIN* value for the testing data set was 0.88. Fig. 5 shows the occurrences, for the testing data set, of hourly power generation values in each interval, and the ideal number of occurrences (the 0.05% of the number of hours in the testing data set).

TABLE III. RMSE FOR POINT FORECAST

Model	RMSE training (kW)	RMSE testing (kW)
Q0.5	1074.35	969.47
PS	1357.28	1205.49
MLP18	952.58	1012.80

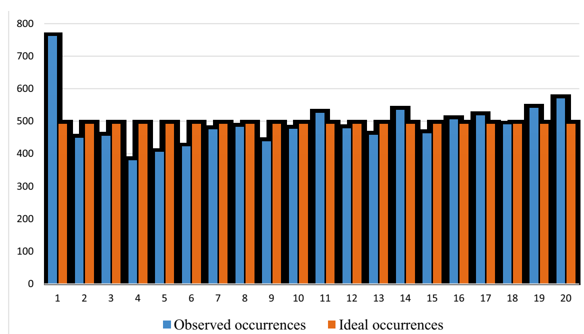

Fig. 5. Number of occurrences in each interval.

VI. Conclusions

The methodology followed to develop a probabilistic photovoltaic power forecasting model for the day ahead has been presented. The model is based on quantile regression neural networks which parameters have been optimized using a genetic algorithm. The genetic algorithm selected the optimal vales for the number of neurons in the hidden layer and the weight decay regularization factor, and also selects the set of input variables among those available. The input variables available include chronological, astronomical and forecasted weather variables.

The methodology has been applied to a real life PV plant with an installed power around 7 MW. The data of three years were used to develop the model: the data of the first two years was used to optimize and train the models and the data of the third year was used to test the resulting models. The results obtained using as point forecasts the corresponding to the QRNN for quantile 0.5 improved those one obtained with a persistence reference model and with the best MLP model using all the available inputs.

Further research is undergoing in order to improve the results analyzing the effect of alternative fitness functions in the optimization process, such as the reliability index or the tilted absolute value function (also known as pinball loss function), analyzing the effect of the inclusion of new forecasted weather variables, and developing QRNN models with different structure (number of neurons in hidden layer, penalty and used input variables) for each quantile.

ACKNOWLEDGMENT

The authors would like to thank the ‘‘Ministerio de Economia y Competitividad’’ of the Spanish Government for supporting this research under the projects ENE2013-48517-C2-2-R and ENE2016-78509-C3-3-P, and the ERDF funds of the European Union.

REFERENCES

[1] J. Hossain, “Wind Energy 2050,” World Wind Energy Association, 2015.

[2] International Energy Agency, “2016 Snapshot of global photovoltaic markets,” 2017.

[3] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, “Review of photovoltaic power forecasting,” *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.

[4] M. Q. Raza, M. Nadarajah, and C. Ekanayake, “On recent advances in PV output power forecast,” *Sol. Energy*, vol. 136, pp. 125–144, Oct. 2016.

[5] J. W. Taylor, “A quantile regression neural network approach to estimating the conditional density of multiperiod returns,” J. Forecast., vol. 19, no. 4, pp. 299–311, 2000.

[6] C. Chen, “A Finite Smoothing Algorithm for Quantile Regression,” J. Comput. Graph. Stat., vol. 16, no. 1, pp. 136–164, 2007.

[7] A. J. Cannon, “Quantile regression neural networks: Implementation in R and application to precipitation downscaling,” Comput. Geosci., vol. 37, no. 9, pp. 1277–1284, 2011.

[8] R Core Team, “R: A Language and Environment for Statistical Computing,” 2017. [Online]. Available: https://www.r-project.org/.

[9] R. Koenker and G. Bassett, “Regression quantiles,” Econometrica, vol. 46, no. 1, pp. 33–50, 1978.

[10] [G. N. Tiwari, A. Tiwari, and Shyam, Handbook of Solar Energy. Theory, Analysis and Applications. Springer, 2016.

[11] F. Barbieri, S. Rajakaruna, and A. Ghosh, “Very short-term photovoltaic power forecasting with cloud modeling: A review,” Renew. Sustain. Energy Rev., vol. 75, pp. 242–263, Aug. 2017.

[12] Z. Janjic and et al., “NMM Version 3 Modeling System User’s Guide,” 2014.

[13] L. Scrucca, “GA: A Package for Genetic Algorithms in R,” J. Stat. Softw., vol. 53, no. 4, pp. 1–37, 2013.