Energy Conversion and Management 64 (2012) 533–540

Contents lists available at SciVerse ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Analysis of forecast errors for irradiance on the horizontal plane

G.M. Tina [a,*], S. De Fiore [b], C. Ventura [a]

[a] Dipartimento di Ingegneria Elettrica, Elettronica e Informatica University of Catania, Viale Andrea Doria n. 6, 95125 Catania, Italy
[b] Softeco Sismat SpA Via De Marini 1, WTC Tower, 16149 Genoa, Italy

ARTICLE INFO

Article history:
Available online 19 August 2012

Keywords:
Forecast
Photovoltaic
nRMSE
Classification

ABSTRACT

A major challenge of the next years in global development will be the large scale introduction of renewable non-programmable (wind and solar) energy sources into existing energy supply structures. Due mainly to the variability of weather and shadow conditions, the total power production coming from photovoltaic plants in a specified future time period cannot be determined precisely, as it is a nondeterministic and stochastic process. This instability is caused by the dependence of PV generation on meteorological conditions: irradiance and temperature. If meteorological conditions can be forecasted with sufficient precision, it will be possible to estimate the energy a PV system will produce, making photovoltaic a more reliable electricity source. For this reason, in this paper, forecast solar radiation data, provided by two weather providers, have been analyzed. In order to verify their effectiveness, these forecast data have been compared with the measured ones and the errors have been calculated by means of the normalized Root Mean Square Error (nRMSE). Then an algorithm, that allows to classify a day as variable, cloudy, slightly cloudy or clear, has been implemented. Based on this classification, a maximum forecast error is determined. In this context, a neural network has been implemented, it allows to predict the nRMSE of a specific day knowing the percentages of the variable, cloudy, slightly cloudy or clear intervals (considered in that day) calculated on forecast data. Referring to Catania (Italy), experimental data are reported to demonstrate the potentiality of the adopted solutions.

1. Introduction

The sun is an extremely powerful energy source and solar radiation is by far the largest source of energy received by the earth. However, the large band variation in the wind speed and unpredictable solar radiation causes remarkable fluctuations of output power in offshore wind and photovoltaic system respectively. The sources like offshore wind and PV have inherently intermittent characteristics due to different climatic conditions. Both wind speed and solar radiations have unpredictable features which may reduce the capacity of the energy storage systems [1]. All plants, whether renewable or not, suffers from occasional breakdown, which also impacts on supply availability. The traditional response of network planners to the challenge of supplying a randomly varying system load is to design the system to acceptable levels of reliability [2]. Therefore, the increasing use of solar power as a source of electricity has led to increased interest in forecasting radiation over short time horizons [3]. A key point hereby is the prediction of renewable energy fluxes, typically for time scales from the sub-hourly range up to 2 days, depending on the given application. Forecasting of solar irradiance is in general significant for planning the operations of power plants which convert renewable energies into electricity. Examples are the storage management in stand-alone photovoltaic or wind energy systems, control systems in buildings, control of solar thermal power plants and the management of electricity grids with high penetration rates from renewable sources. Moreover, forecast information on the expected solar and wind power production is necessary for the management of electricity grids, for scheduling of conventional power plants and also for decision making on the energy market [4]. Predictability is, in fact, key to managing solar and wind power's variability and improved accuracy of solar and wind power prediction has a beneficial effect on the amount of balancing reserves needed, so the accurate forecasting of solar and wind power is important for their economic integration into the power system.

In particular, the possibility to predict the solar irradiance (up to 24 h or even more) can become, with reference to the grid connected photovoltaic plants, fundamental in making power dispatching plans and, with reference to stand-alone and hybrid systems, also a useful reference for improving the control algorithms of charge controllers [5]. Daily total solar radiation data, in fact, is considered such as the most important parameter in the meteorology, solar conversion, and renewable energy application, particularly for the sizing of PV systems, which represents an

* Corresponding author.
E-mail addresses: gtina@diees.unict.it (G.M. Tina), sebastiano.defiore@softeco.it (S. De Fiore).

0196-8904/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.enconman.2012.05.031

important part of the design in any PV system: stand-alone PV system, grid-connected PV system, PV-wind hybrid system, etc. [6]. Therefore, forecasting solar radiation is essential in all PV systems. Forecast meteorological data are, in fact, needed in such stand-alone systems, particularly that ones located in isolated sites, where data are not always available due to the non-availability of the meteorological stations. Moreover, over the last years the contribution of power production by PV systems to the electricity supply has become increasingly relevant [7]. With this increasing penetration of photovoltaic (PV) power systems into the utility network, the issue caused by the fluctuation and intermittence of PV power output draws more interest. This power output fluctuation, in fact, will impact the power system's stability. Therefore, forecast information is necessary in such grid-connected systems to improve the assessment of the output of a photovoltaic system beforehand, allowing a stable integration in the public electricity grid.

In this paper, first of all predicted solar radiation on a horizontal surface, given us by two weather forecast providers, have been compared with the measured ones, in order to determine their accuracy. Then, a method to classify each time step considered during a day as variable, cloudy, slightly cloudy or clear, has been developed. In this way, it is possible to understand if there is a correlation between the percentages of the intervals (based on the used time step) of a specific day that belong to each class and the error done on that day. Using a neural network, a correlation between these percentages and the error done in that day has been found.

2. Evaluation of the solar irradiation forecasts errors

To evaluate the errors of the forecasts of solar irradiation, different kinds of errors can be considered: the root mean absolute error, the mean absolute error, the mean relative variance and root mean square error. The Root Mean Square Error, RMSE, emphasizes more large deviation than small ones. The mean absolute error, MAE, provides a linear weighting between large and small deviations. The mean relative variance, MRV, provides a measure between the error of the forecast of solar irradiation in relation to the variation between the measured value and the average of all measured values. This measure quantifies the relative error and is proportional to the standard deviation.

In this paper, for the evaluation of the accuracy of solar radiation predictions, a non-dimensional form of the RMSE, the normalized Root Mean Square Error (nRMSE) has been used:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\cdot\sum_{n=1}^{N}\left(x_{forecast}(\Delta t) - x_{measured}(\Delta t)\right)^2} \tag{1}$$

$$\mathrm{nRMSE} = \frac{\mathrm{RMSE}}{\sqrt{\frac{1}{N}\cdot\sum_{n=1}^{N}\left(x_{measured}(\Delta t)\right)^2}} \tag{2}$$

where N is the number of available data, the variable $x_{forecast}(\Delta t)$ represents the forecast value of the global irradiance at the Δt time step, while $x_{measured}(\Delta t)$ represents the measured one. The RMSE and the nRMSE are, in fact, frequently used measures of the differences between values predicted by a model or an estimator and the values actually observed from the thing being modeled or estimated [8–12]. The nRMSE has been chosen because it allows to read in a more understandably way the errors, since it is a non-dimensional parameter.

3. Classification of daily solar radiation

The power output and economic viability of a PV plant depends heavily on the solar irradiance characteristics of its location. Therefore studies which allow to optimize direct normal irradiance forecasts are required in order to predict the potential power output of a plant. Such studies are based on daily irradiance records, often with a view to classifying days by cloud transition patterns and direct radiation variability, because the system response in concentrating power plants depends on the kind of day. For this reason, a new method based on Mathematical Morphology techniques for the classification of solar radiation curves has been presented in [13]. The main advantage of using the classification described in [13] is that it allows to keep the dynamics of the solar radiance curves in the analysis. While, in [14] they describe a new algorithm to characterize sky condition in intervals of 5 min using four categories of sun exposition: apparent sun with cloud reflection effects; apparent sun without cloud effects; sun partially concealed by clouds; and sun totally concealed by clouds. The algorithm can also be applied to estimate hourly and daily sky condition in terms of the traditional three categories: clear, partially cloudy and cloudy day.

Another example of classification procedure for solar irradiances is presented in [15] and discussed for five locations. This procedure uses fractal dimension analysis. It defines two thresholds of the fractal dimensions using first a heuristic method then a statistical one. This allows to determine three classes of days: clear sky day, partially clouded sky day and clouded sky day. While, in [16,17] they deal with the fractal modeling of daily solar irradiances measured with a sampling time of 10 min for 1 year.

Many authors use the clearness index as the starting point for their analysis. This index, k_t, is the mean of the ratio between measured global irradiance on the horizontal plane (G_{TH}) and extraterrestrial global horizontal irradiance (G_{OH}) at a specific site, day and time ($k_t = G_{TH}/G_{OH}$). A classification of daily distributions of the clearness index, k_t, by estimating a finite mixture of Dirichlet distributions without assuming any parametric hypothesis on these daily distributions, has been presented in [18]. Their results reveal four distinct distribution types each representing a different type of day, namely, "clear sky days", "intermittent clear sky days", "intermittent cloudy sky days" and "cloudy sky days".

In this paper, an algorithm used to classify days as variable, cloudy, slightly cloudy or clear, is presented. Practically, each day has been classified according to an algorithm based on the variation of an index of serenity, k. This index is calculated using the following formula:

$$k(t) = \frac{|x_{measured}(\Delta t) - x_{CSM}(\Delta t)|}{x_{CSM}(\Delta t)} \tag{3}$$

where Δt is the time step, $x_{measured}(\Delta t)$ and $x_{CSM}(\Delta t)$ indicate, respectively, the measured solar radiation value and the value of the solar radiation calculated using the ESRA clear sky model (CSM) [19]. This index is used to calculate the "moving average" (Eq. (4)) and the "moving function" (Eq. (5)), whose values are used in the algorithm to characterize a day:

$$MA = \frac{1}{Nm}\cdot\sum_{s=0}^{Nm-1} k(\Delta t - s) \tag{4}$$

$$MF = \sum_{s=0}^{Nm-1} |MA - k(\Delta t - s)| \tag{5}$$

where Nm is the number of the considered time steps used to calculate the moving average.

Studies presented in [20] describe why hourly average measurement data is insufficient to model PV system performance accurately. They show how, when averaging data each hour, short periods of high irradiance are combined with dull periods. However, as the power output of modules reacts quickly to changes of irradiance, while temperature changes are slow, modules will

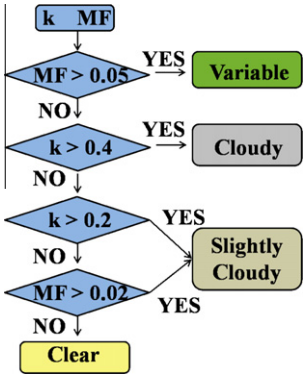

Fig. 1. Block diagram of the algorithm used to classified each time step considered during a day.

often give higher power than expected at lower module temperatures than calculated from hourly averages. Since the quick changes of irradiance during a day are essential to correctly classify a day, in this paper two different time steps, 10 min and 1 h, have been used and we have chosen $Nm = 3$.

Then the value of k and MF are used to characterize the weather condition of a day, using the algorithm described in the block diagram shown in Fig. 1. Applying this algorithm, each step considered during a day can be classified depending on the value of the index k and of MF, in order to evaluate the percentage of the time steps that during a day can be considered as variable, cloudy, slightly cloudy or clear. Our aim is to find a correlation between the values of these percentages and error done on the forecast data compared to the measured one (calculated using the nRMSE), in order to predict this error knowing the four values of percentages. Table 1 summarizes the different classes used in literature to classify days. As it is possible to note, the methods used in literature do not considered in the classification the variability of the weather during a day. Moreover, these methods associate only one class to each day, while in the method here used each day is characterized by four values that correspond to the percentages of the time steps considered during a day that belong to each class. That is very useful to correctly classify a day, especially if it is characterized by a changeable weather.

4. Clear sky model

To calculate the index of serenity (Eq. (3)), it is necessary to calculate the clear sky radiation on the horizontal surface, x_{CSM}.

The maximum radiation is defined as the radiation occurring on days with a clear or cloudless sky; it corresponds to the greatest possible value of global radiation per hour at the specified altitude [21]. Since there are a several numbers of variables to be calculated and they are unpredictable, an accurate evaluation of the solar radiation, even if it is a clear sky day, is next to never possible. Nonetheless, numerous methods have been developed to calculate the clear sky solar radiation. These methods allow to have a degree of accuracy that can be considered acceptable when this information must be used by engineers or architects.

In this application, a clear-sky model, which has been developed in the framework of the new digital European Solar Radiation Atlas (ESRA), has been used. It has been chosen because it seems to be the best one with respect to robustness and accuracy [19].

On a horizontal surface the total radiation is given by the sum of direct-beam (G_{BH}) and diffuse radiation (G_{DH}) [22].

$$G_{TH} = G_{BH} + G_{DH} \tag{6}$$

In accordance with Rigollier et al. [19] the direct irradiance on a horizontal surface (or beam horizontal irradiance) for clear sky, G_{BH}, is given by:

$$G_{BH} = I_{E0} \cdot sin(\gamma_s) \cdot exp(-0.8662 \cdot T_L(AM2) \cdot m \cdot \delta_R(m)) \tag{7}$$

where I_{E0} is the extraterrestrial irradiance, γ_s is the solar altitude angle, it is 0° at sunrise and sunset; $T_L(AM2)$ is the Linke turbidity factor for an air mass equal to 2; m is the relative optical air mass and finally $\delta_r(m)$ is the integral Rayleigh optical thickness.

The diffuse irradiance falling on a horizontal surface for clear sky (or diffuse horizontal irradiance), G_{DH}, is determined by:

$$G_{DH} = I_{E0} \cdot T_{rd}(T_L(AM2)) \cdot F_d(\gamma_s, T_L(AM2)) \tag{8}$$

In this equation, the diffuse radiation is expressed as the product of the diffuse transmission function at zenith (i.e. sun elevation is 90°), T_{rd}, and a diffuse angular function, F_d.

$$T_{rd}(T_L(AM2)) = -1.5843 \times 10^{-2} + 3.0543 \times 10^{-2} \cdot T_L(AM2) + 3.797 \times 10^{-4} \cdot T_L^2(AM2) \tag{9}$$

The diffuse angular function, F_d, depends on the solar elevation angle and is fitted with the help of second order sine polynomial functions:

$$F_d(\gamma_s, T_L(AM2)) = A_0 + A_1 \cdot sin(\gamma_s) + A_1 \cdot sin(\gamma_s)^{-2} \tag{10}$$

The coefficients A_0, A_1, and A_2, only depend on the Linke turbidity factor. They are unitless and are given by:

$$\begin{aligned} A_0 &= 0.26463 - 0.061581 \cdot T_L(AM2) + 0.0031408 \cdot T_L^2(AM2) \\ A_1 &= 2.0402 + 0.018945 \cdot T_L(AM2) - 0.011161 \cdot T_L^2(AM2) \\ A_2 &= -1.3025 + 0.039231 \cdot T_L(AM2) + 0.0085079 \cdot T_L^2(AM2) \end{aligned} \tag{11}$$

with a condition on A_0:

$$A_0 = \frac{0.0022}{T_{rd}}, \text{if} A_0 \cdot T_{rd} < 0.0022 \tag{12}$$

5. Neural network

Neural networks are being widely used in many fields of study. This could be attributed to the fact that these networks are attempts to model the capabilities of human brains [23]. Since the last decade, neural networks are being used as a theoretically sound alternative to traditional statistical models. Although neural networks (NNs) originated in mathematical neurobiology, the rather simplified practical models currently in use have moved steadily towards the field of statistics. Neural networks are being used in areas of prediction and classification, areas where regression models and other related statistical techniques have traditionally been used. These applications have been classified into various categories namely, accounting and finance, health

Table 1
Different classes used in literature to classify days.

Reference	Classification
[13]	Clear sky, intermittent clear sky, intermittent cloudy sky and cloudy sky
[14]	Clear, partially cloudy and cloudy
[15]	Clear sky, partially clouded sky and clouded sky
[16]	Clear sky, partly cloudy sky and completely cloudy sky
[17]	Clear sky, partially covered sky and overcast sky
[18]	Clear sky, Intermittent clear sky, Cloudy sky, Intermittent cloudy sky
This paper	Clear %, slightly cloudy %, cloudy %, variable %

and medicine, engineering and manufacturing, marketing and other important applications.

Artificial neural networks, originally developed to mimic basic biological neural systems, the human brain particularly, are composed of a number of interconnected simple processing elements called neurons or nodes [24]. Each neuron receives an input signal which is the total "information" from other neurons or external stimuli, processes it locally through an activation or transfer function and produces a transformed output signal to other neurons or external outputs. Each neuron is connected to other neurons by means of directed communication links, each with an associated weight. The weights represent information being used by the net to solve a problem; they are adapted during use to improve performance. In the typical single layer net, the input units are fully connected to output units but are not connected to other input units, and the output units are not connected to other output units. While, a multi-layer net is a net with one or more layers (or levels) of neurons (the so-called hidden units) between the input units and the output units. Typically, there is a layer of weights between two adjacent levels of units (input, hidden, or output). Multi-layer nets can solve more complicated problems than can single-layer nets, but training may be more difficult. However, in some cases, training may be more successful, because it is possible to solve a problem that a single-layer net cannot be trained to perform correctly at all.

5.1. Neural network to approximate the nRMSE function

The neural network here implemented has been used to approximate the nRMSE function. The values of the learning period, size of data used in the training, neurons on the hidden layer and learning rate were set on a trial and error basis. The neural network is composed by four inputs that represent the four percentages of a specific day and one output that is the correspondent nRMSE, while the number of neurons in the hidden-layer is set equal to ten and the number of epochs equal to 3000. The block diagram of the neural network used to approximate the nRMSE function is shown in Fig. 2. Moreover, a hyperbolic tangent sigmoid transfer function is used as neuron activation function of the hidden layer, a linear transfer function is used as the neuron activation function of the output layer, the learning rate is set equal to 0.001 and the learning algorithm used to implement the neural network is the Levenberg–Marquardt back-propagation. Using this algorithm the network trains function that updates weight and bias values according to Levenberg–Marquardt optimization [25]. In this work, first the percentages calculated using measured data have been used as input, then the nRMSE has been approximated having as inputs the four values of percentages calculated using the forecast data.

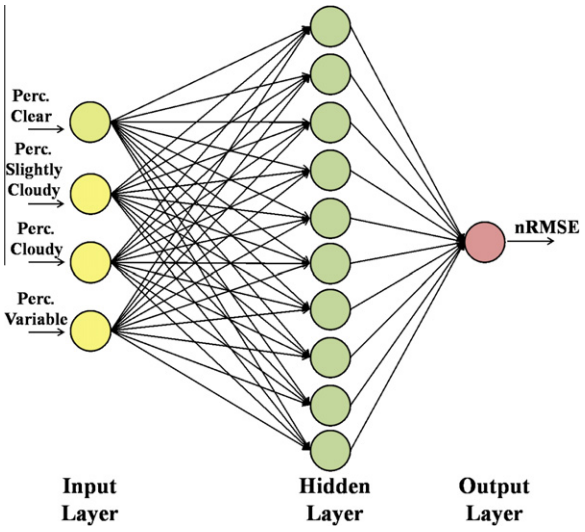

Fig. 2. Structure of the Neural Network used to approximate the nRMSE.

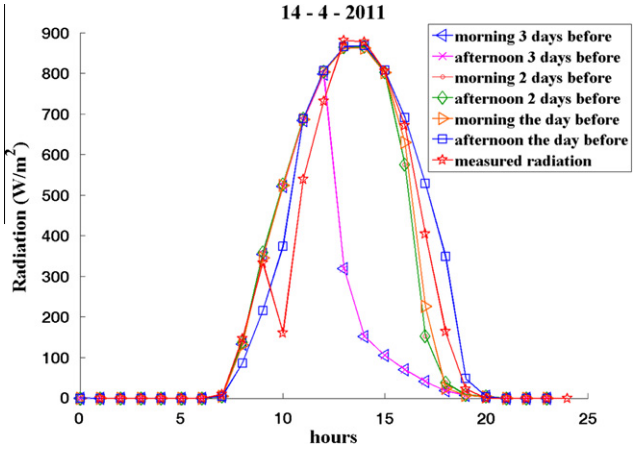

Fig. 3. Forecast data by the first provider; reported data represents the radiation calculated three, 2 days and the day before the 14th April 2011, twice a day.

6. Experimental setup

To test the accuracy of the hourly daily forecast of solar radiation, they have been compared with the ground measured ones. In particular, data relative to Catania (latitude 37°34′N, longitude 15°10′E, Italy) have been considered.

6.1. Forecast data

Meteorological data such as solar radiation, ambient temperature, relative humidity, wind speed, clearness index and sunshine duration, are accepted as dependable and widely variable sources. It is therefore required to be able to formulate forecasting and estimation models of these meteorological data. These data play a very important role in PV systems [26]. However, in many cases, meteorological forecast data can require the implementation of complex algorithms. For this reason, in this paper forecast weather data, given us by two weather providers, have been used.

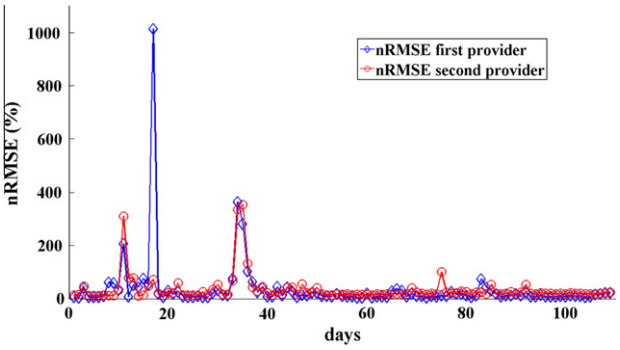

Fig. 4. Comparison between forecast data relative to the first provider and forecast data relative to the second provider; diamond blue line represents the nRMSE obtained comparing measured radiation with forecast data given us by the first provider here analyzed, while circle red line represents the nRMSE obtained comparing measured radiation with forecast data given us by the second provider here considered. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The first provider here considered uses the new 2 km WRF EMM (Eulerian Mass Model) model. It derives from the framework WRF with solver ARW and it is characterized by an extreme configurability and integrality with other models. This provider gives us the following forecast measures: solar radiation, temperature, wind speed and intensity, calculated in Catania (Italy); these data are provided as hourly data. The provided flow of data is integrated by four modeling runs per day, but the updating which are relevant for our target are only the first 72 h of forecast that are updated by the two runs of the high resolution model; outputs are provided twice a day: at 9:30 a.m. and at 9:30 p.m. As it is possible to note in Fig. 3, the forecast solar radiation calculated 3 days before is less accurate than the last recent.

While the second provider uses a system able to support a calculation power suitable for elaborations working from the global to the local scale. This provider gives us hourly data, outputs are provided once a day: at 2:00 a.m. 72 h before.

Fig. 4 shows the nRMSE calculated comparing measured radiation with forecast data given us by the two different providers here considered, that use two different methods to forecast the solar radiation. Fig. 4 shows that it is not possible to define which provider is more accurate, but it depends on the considered day. Maybe, a finer comparison could be done considered a longer interval of time. In our experiments we have used data relative to the first provided here considered since a greater amount of data are available.

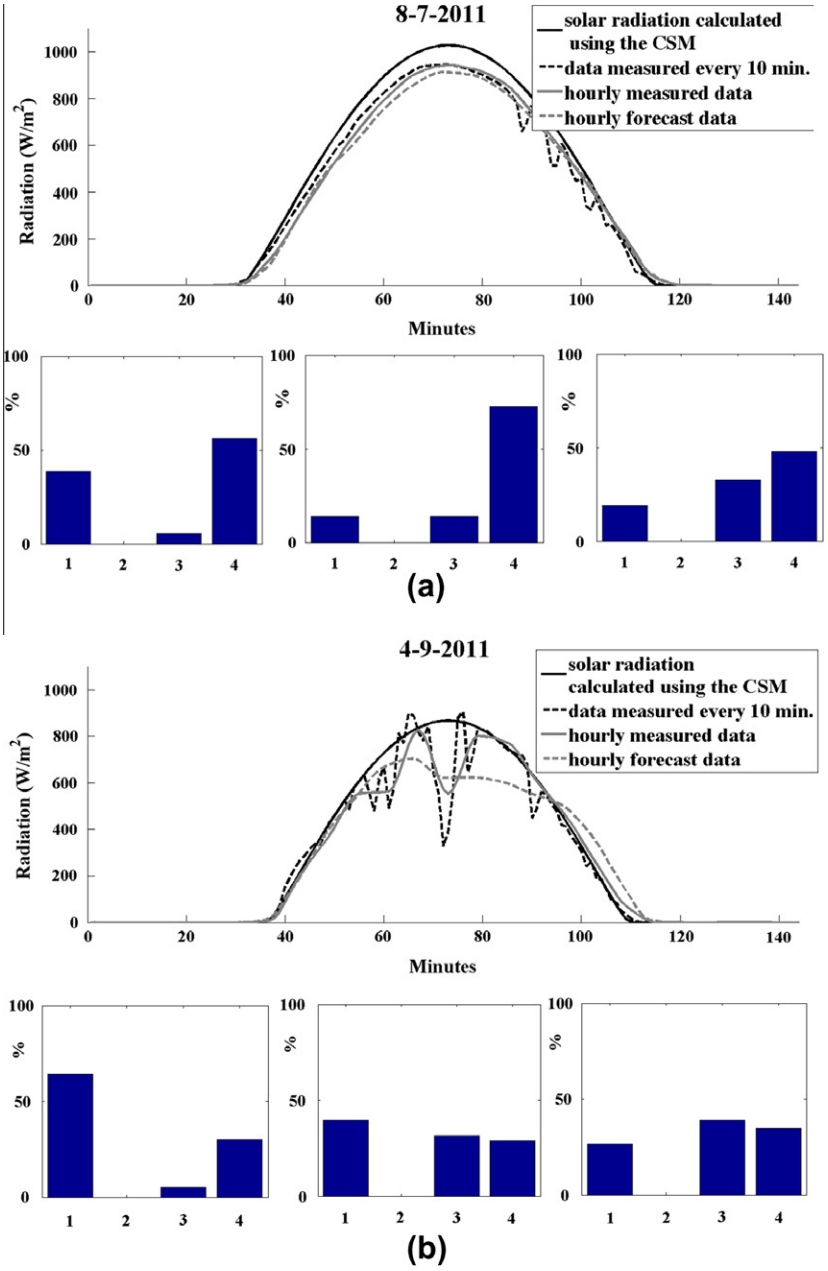

Fig. 5. Example of day classification; in figures: solar radiation calculated using the ESRA CSM [19] (black continuous line), solar radiation measured every 10 min (black dashed line), hourly measured solar radiation (gray continuous line) and hourly forecast solar radiation (gray dashed line); 1 indicates the variable percentage, 2 indicates the cloudy percentage, 3 indicates the slightly cloudy percentage, while 4 indicates the clear percentage; (a) data relative to a clear day: 8th July 2011; (b) data relative to a variable day: 4th September 2011.

Table 2
Values of percentages relative to the 8th July 2011.

Class	10 min	1 h	Forecast
1 – Variable %	38.5	13.7	19.2
2 – Cloudy %	0	0	0
3 – Slightly Cloudy %	5.5	13.7	32.9
4 – Clear %	56.2	72.6	47.9

Table 3
Values of percentages relative to the 4th September 2011.

Class	10 min	1 h	Forecast
1 – Variable %	64.4	39.7	26.4
2 – Cloudy %	0	0	0
3 – Slightly cloudy %	5.48	31.5	38.9
4 – Clear %	30.1	28.8	34.7

6.2. Ground measured data

For the investigation, we have used measured irradiance data provided by a weather station installed on the coverage of the DIEEI laboratories – University of Catania (Italy) buildings. Such weather station has been realized by the society WiSNAM s.n.c. in collaboration with the research group of the Electrical Systems for the Energy [27]. It answers to the demand of a suitable measurement and the primary energetic source (wind and solar energy) modeling. The weather station includes the following sensors: wind speed and direction, temperature, relative humidity and solar radiation.

7. Experimental results

In Italy the third feed-in incentive mechanism (D.M 6 August 2010) has introduced an increasing of the base tariff (+20%) connected with the programmability of the energy produced by the PV plant. The owner of the PV power plant (the size of the PV system ranges from 200 kW a 10 MW) has to respect an hourly scheduling (from 8:00 a.m. to 8:00 p.m.) with a maximum forecast error of 10%. Further he has to communicate this forecasted scheduling 1 day before the dispatching day. This profile must be followed for almost 300 days, that means the 82% of the days during a year.

In this paper an algorithm, that allows to classify a day as clear, slightly cloudy, cloudy or variable, has been developed. Fig. 5 shows two example of implementation of this algorithm, relative to a clear day (Fig. 5a) and a variable day (Fig. 5b). The forecast data are relative to the first provider before analyzed. Tables 2 and 3 show the four values of percentages relative to the 8th July 2011 and the 4th September 2011 respectively, considering data measured every 10 min, each hour or forecast data.

The algorithm has been applied using the solar radiation measured every 10 min, the hourly measured radiation and the hourly forecast radiation. Experiments show as the percentages calculated on hourly data are less accurate than that ones calculated on data measured every 10 min. That happens because, using hourly data, the dynamic of the radiation curve is lost; however, data with a short time step are rarely measured due to the costs involved with such measurements. A time step of 10 min seems to be a good tradeoff.

As it is possible to note, unlike methods presented in literature used to classify days, the method here developed allows to characterize each day using four values, correspond to the percentages of time steps that during a day can be classified as clear, slightly cloudy, cloudy or variable; moreover, one of the class considered in the method here used is "variable". These two properties can be very useful to identify a day characterized by a changeable weather.

The uncertainty of the forecast data have been evaluated using the nRMSE as main measure, taking into account the solar radia-

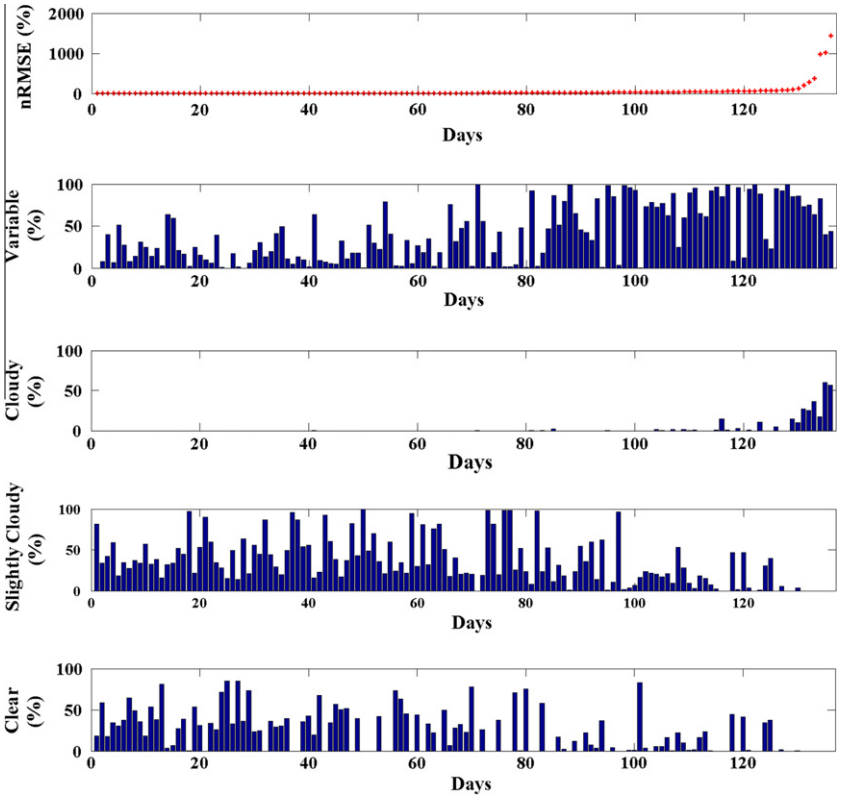

Fig. 6. The error between the measured and the forecast daily solar radiation, calculated using the nRMSE, is higher when both variable and cloudy percentages are high.

tion measurements in the period between January 2011 and August 2011, in which period both forecast (first provider here considered) and measured data were available for only 136 days. The solar radiation is partly absorbed and partly reflected by the atmosphere; this causes a drop of the energy that can be absorbed on the ground. This drop varies notably basing on the weather condition and in particular on the cloudiness. In Fig. 6 the percentages have been sorted based on the daily nRMSE. As it is possible to note, the nRMSE is higher when both variable and cloudy percentages are high. It demonstrates as the forecast uncertainty depends on the meteorological situation: situations with inhomogeneous clouds generally are more difficult to forecast than forecasts for clear sky days. In the nRMSE curve there are some very high values of nRMSE, those points correspond to days where the forecast data are completely different from the measured ones; that happens especially when the measured radiation corresponds to a variable day, while the provider forecasts a clear day. These results show quantitatively that this procedure can be successfully applied in sites characterized by mostly sunny or slightly cloudy days along the year, this situation is quite common in Mediterranean area. Anyway, considering data relative to Catania, the rules scheduled by the third feed-in incentive mechanism, are not respected, since only for the 44% of the days (61 days) during the 136 days analyzed there is an error less than 10%.

After having assigned four values of percentages to each day and having calculated the nRMSE, two neural networks have been implemented. The first one have been used to approximate the nRMSE using as input the four percentages calculated starting from measured data, while in the second one the inputs were the percentages calculated using forecast data. Fig. 7 shows the measured nRMSE and the two approximated nRMSE functions. In particular, Fig. 7a shows the comparison between the measured nRMSE and the nRMSE evaluated through the neural network using the percentages calculated starting from measured data, while Fig. 7b shows the measured nRMSE compared with the nRMSE evaluated through the neural network using the percentages calculated starting from forecast data. In Fig. 7b it is possible to note that some points of the nRMSE function evaluated using the percentages calculated starting from forecast data are negative, also if the nRMSE cannot assume negative values. That happens because, in this case, the nRMSE is not calculated analytically, but using the neural network.

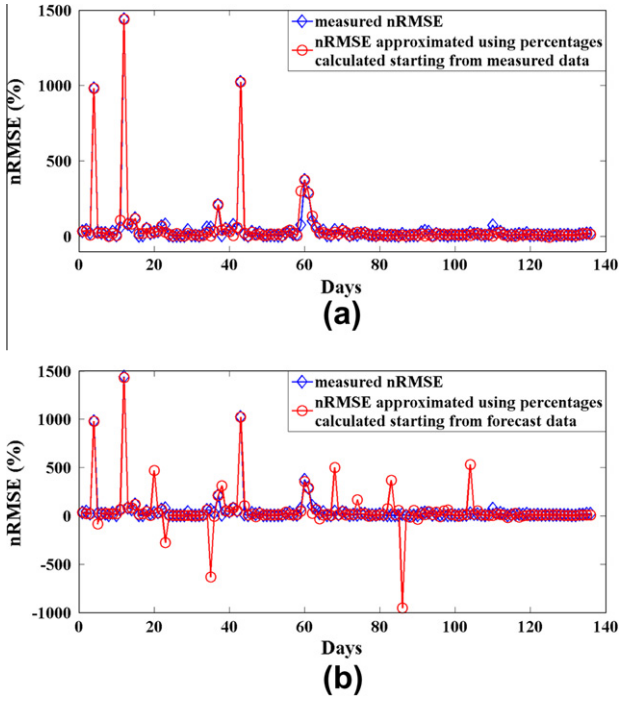

Fig. 7. Comparison between the measured nRMSE and the approximated ones; (a) Comparison between the measured nRMSE and the nRMSE evaluated through the neural network using the percentages calculated starting from measured data; (b) Comparison between the measured nRMSE and the nRMSE evaluated through the neural network using the percentages calculated starting from forecast data.

To evaluate the effectiveness of these approximations, the normalized Root Mean Square Error (Eq. (2)) has been used. For the first approximated function the global nRMSE was 4.9%, while for the second one it was 17%. As it is possible to note, in every case the approximated function that uses the percentages calculated starting from measured data is approximated better than the other one. Anyhow, the approximated nRMSE functions calculated using the neural networks here implemented can be considered a good approximation of the real one and therefore, it is possible to predict the value of the nRMSE of a specific day, the day before, using forecast data obtaining small errors.

8. Conclusion

The purpose of this work was threefold: first of all the accuracy of the forecast data, given us by two weather providers, have been evaluated using the normalized Root Mean Square Error (nRMSE) as mean measure, then an algorithm that allows to classify days through four values of percentages: variable, cloudy, slightly cloudy or clear, has been developed and finally a neural network that allows to approximate the nRMSE starting from the value of these percentages has been developed.

The main advantage of using the classification here proposed is that it can be very useful to identify a day characterized by a changeable weather, since four values are used to classify a day, correspond to the percentages of time steps that during a day can be classified as clear, slightly cloudy, cloudy or variable.

Analyzing the nRMSE and the correspondent four values of percentages, we have noted that the error done on the forecast data is lower if the slightly cloudy or clear percentages are higher that the variable or cloudy ones. It demonstrates as the forecast uncertainty depends on the meteorological situation: situations with inhomogeneous clouds generally are more difficult to forecast and show a lower accuracy than forecasts for clear sky days. Moreover, forecasting solar radiation using hourly data causes the lost of the dynamic of the solar radiation curve and therefore here a time step of 10 min has been used since it allows to keep the dynamics of the solar radiance curves in the analysis.

However, our target was to predict the nRMSE knowing the four values of percentages calculated using the forecast solar radiation, through the implemented neural network. To evaluate the effectiveness of the approximation, we have evaluated the nRMSE and it was equal to 4.9% when the input of the neural network were the four percentages calculated starting from measured data, while it was equal to 17% when the inputs were the percentages calculated using forecast data. These results can be considered satisfying and therefore it is possible to predict the value of the nRMSE of a specific day, the day before, using forecast data, with a good approximation.

References

[1] Ray PK, Mohanty SR, Kishor N. Proportional – integral controller based small-signal analysis of hybrid distributed generation systems. Energy Convers Manage 2011;52(4):1943–54.
[2] Mitchell K, Nagrial M, Rizk J. Simulation and optimization of renewable energy systems. Int J Electr Power Energy Syst 2005;27(3):177–88.
[3] Gordon. Predicting solar radiation at high resolutions: a comparison of time series forecasts. Solar Energy 2009;83(3):342–9.
[4] Lange M, Focken U. Physical approach to short-term wind power prediction. New York: Springer; 2005 [3-540-25662-8].

[5] Mellit A, Pavan AM. A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy. Solar Energy 2010;84(5):807–21.
[6] Mellit A. Sizing of photovoltaic systems: a review. Revue des Energies Renouvelables 2007;10(4):463–72.
[7] Lorenz E, Hurka J, Heinemann D, Beyer HG. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE J Spec Top Earth Observations Remote Sens 2009;2:2–10.
[8] Remund J, Perez R, Lorenz E. Comparison of Solar Radiation Forecasts for the USA. In: Proc 23rd European photovoltaic solar energy conference. Valencia, Spain; 2008.
[9] Torres JL, De Blas M, Garca A, de Francisco A. Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. Renew Energy 2010;35(6):1325–32.
[10] Lorenz E, Hurka J, Heinemann D, Beyer HG. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE J Special Top Earth Observations Remote Sens 2009;2:2–10.
[11] Torres JL, Garca A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Solar Energy 2005;79(1):65–77.
[12] Paoli C, Voyant C, Muselli M, Nivet ML. Solar radiation forecasting using ad-hoc time series preprocessing and neural networks. In: Proc of ICIC, vol. 1. 2009. p. 898–907.
[13] Romeo MG, Leon T, Mallor F, Santigosa LR. A morphological clustering method for daily solar radiation curves. Solar Energy 2011.
[14] Assuncao HF, Escobedo JF, Oliveira AP. A new algorithm to estimate sky condition based on 5 min-averaged values of clearness index and relative optical air mass. Theor Appl Climatol 2007;90(3–4):235–48.
[15] Harrouni S. Fractal classification of typical meteorological days from global solar irradiance: application to five sites of different climates. In: Badescue Viorel, editor. Modeling solar irradiance at earth surface. Berlin Heidelberg, Germany: Springer Verlag; 2008. p. 29–54.
[16] Maafi A, Harrouni S. Preliminary results of the fractal classification of daily solar irradiances. Solar Energy 2003;75:53–61.
[17] Harrouni S, Guessoum A, Maafi A. Classification of daily solar irradiation by fractional analysis of 10-min-means of solar irradiance. Theor Appl Climatol 2005;80:27–36.
[18] Soubdhan T, Emilion R, Calif R. Classification of daily solar radiation distributions using a mixture of dirichlet distributions. Solar Energy 2009;83(7):1056–63.
[19] Rigollie C, Bauer O, Wald L. On the clear sky model of the ESRA European Solar Radiation Atlas with respect to the heliosat method. Solar Energy 2000;68(1):33–48.
[20] Ransome S, Funtan P. Why hourly averaged measurement data is insufficient to model pv system performance accurately. In: 20th European PVSEC. Barcelona; 2005.
[21] Remund JP. Advanced parameters WP 5.2b: chain of algorithms: short- and long-wave radiation with associated temperature prediction resources. Report to the European commission; 2008. <http://www.soda-is.com/doc/d5-2-2_v3.pdf>.
[22] Markvart T, Castafier L. Practical handbook of photovoltaics: from fundamentals to applications. Elsevier Science Ltd.; 2003.
[23] Paliwal Mukta, Usha Kumar A. Neural networks and statistical techniques: a review of, applications. Expert Syst Appl 2009;36(1):2–17.
[24] Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: the state of the art. Int J Forecast 1998;14(1):35–62.
[25] More JJ. The Levenberg–Marquardt algorithm: implementation and theory. Lecture notes in mathematics, vol. 630. Berlin, Germany: Springer-Verlag; 1978. p. 105–16.
[26] Mellit A, Kalogirou S. Artificial intelligence techniques for photovoltaic applications: a review. Progr Energy Combust Sci 2008;34(5):574–632.
[27] Lorenz E, Hurka J, Heinemann D, Beyer HG. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE J Spec Top Earth Observations Remote Sens 2009;2:2–10.