Photovoltaic Power Production Estimation Based on Numerical Weather Predictions

F. E. Atencio Espejo, S. Grillo, L. Luini
Politecnico di Milano
p.za Leonardo da Vinci, 32
Dipartimento di Elettronica, Informazione e Bioingegneria
I-20133, Milano, Italy
frankeduardo.atencio@mail.polimi.it
samuele.grillo@polimi.it; lorenzo.luini@polimi.it

***Abstract*—A photovoltaic (PV) power generation system relying on meteorological forecast provided by the European Centre for Medium-Range Weather Forecast (ECMWF) (ERA5 database) is proposed. Three years of data collected from photovoltaic panels deployed in Milan, Italy, have been analyzed, both in clear-sky and cloudy conditions. The Ineichen-Perez model has been used as a reference for clear-sky conditions. The power measurements were compared with the power calculated using ECMWF, based on solar theory and technical characteristics of the PV plant in place. cumulative complementary distribution function (CCDF) and errors, have been calculated to determine the accuracy of the model. Results indicate a good agreement in terms of generated power statistics, showing that ERA5 data can be reliably used to design solar plants as well as to forecast their performance and energy production.**

***Index Terms*—forecasting performance, numerical weather prediction, photovoltaic forecasting.**

I. INTRODUCTION

In the latest years, both international and European policies have set the bases for the implementation and improvement of alternatives for energy production. As a consequence, the presence of renewable energy sources has been given a turn in the traditional energy sector. The low-emission generation, the decreasing costs, the increasing efficiency of the technology and the independence of fossil fuel are factors that make attractive the use of non-conventional renewable sources in order to solve the modern paradigm of the global energetic system. The total installed capacity in Italy, especially photovoltaic (PV) plants, has recently increased notably, thus producing a large impact on the distribution and transmission grid. Improvement in energy policies and incentives have encouraged the investments on renewable energy to fulfill the goals scenarios regarding climate change [1]. The connection of non-conventional renewable sources has raised problems in the system making necessary to improve the mechanisms of control and operation for these new power plants and the grid itself. PV plants production must be considered in the planning and operation of the system for the proper management of the grid. Predicting solar production reliably and accurately is not an easy task due to the intrinsic characteristics of the energy source related to the operation of solar technology. Forecasting models can be classified in two categories: indirect and direct forecasting models. The indirect forecasting model use various approaches as numerical weather prediction, image-based and statistical and hybrid artificial neural network as input to PV simulation softwares used in the industry. On the other hand, in the direct forecasting models, the PV power generation is forecast directly by means of historical data such as PV power output and associated meteorological measurements or predictions [2].

This contribution presents the implementation of a direct forecasting models of PV power generation based on historical data and on numerical weather prediction (NWP) products made available from the European Centre for Medium-Range Weather Forecast (ECMWF). Section II presents the theory on the solar radiation and describes the model used to calculate the solar radiation on the surface of the earth in clear sky conditions. Section III deals with the assessment of the production forecasts based on NWP data, which are compared against the real measurement collected in Milan. Section IV discussed the results, while Section V draws the conclusions.

II. GENERAL THEORY FOR SOLAR RADIATION ESTIMATION

The earth is subject to two main movements that affect the radiation coming from the sun. The rotational movement on its own central axis (24 hours to complete one cycle) and the movement of the planet around the sun along an elliptical orbit.

The solar constant, $I_0 = 1367.13\,\mathrm{W/m^2}$, is defined as the incident solar radiation on a unit area perpendicular to the beam direction outside the earth's atmosphere, at 1 astronomical unit ($1.495 \times 10^{11}\,\mathrm{m}$) of distance from the sun [3]. The solar declination δ is the angle between the equatorial plane and a line projected from the center of the sun to the center of the earth. It changes during the elliptical movement around the sun. During the equinox in spring and autumn the solar declination is $0°$, while on the summer and winter solstice, the solar declination reaches its maximum value ($23.45°$). The solar time is the time used in all the sun-angle relationships. It does not coincide necessarily with the civil time. Its difference is related to the distance between the observer's meridian and

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. DOI: 10.1109/PTC.2019.8810897

the reference longitude for the standard time. The equations of these variables are based on [4]. The hour angle, ω, is the angular displacement of the sun with respect to midday. It is equal to 0 at noon and is negative for morning hours and positive in the afternoon. It is obtained from the solar time considering that each hour is equal to 15° in order to make the one full day equivalent to 360°. The sunset hour, ω_s, varies according to the latitude φ. The farther from the equator, the greater the variation of the sunset hour along the year. The sunrise hour is the same as the sunset hour but with opposite sign. The zenith angle—denoted as θ_z—is the angle between a line perpendicular to the horizontal plane in the site and the sun's rays of incidence line to the surface. Its complementary angle is the solar elevation angle, α. The tilt angle, β, is the angle of inclination of the surface that receives the incident radiation. The surface azimuth, γ, is the angle of the projection on a horizontal plane of the normal to the surface from the local meridian. An orientation towards the equator represents an angle $\gamma = 0$. The solar azimuth, γ_s, refers to the angle, from the south, of the beam projection of the radiation on the horizontal plane. The air mass (AM) is the optical path length of a direct beam through the atmosphere, as a ratio of the vertical path length directly to the horizontal plane, proportional to the zenith angle.

Linke turbidity factor, T_{L}, describes the optical thickness of the atmosphere considering scattering and absorption due to water vapor and aerosol particles, as compared with a dry and clean atmosphere [3]. It summarizes the turbidity of the atmosphere and, thus, the attenuation of the direct beam solar radiation on a surface. A large value of the Linke turbidity factor represents a large attenuation of the extraterrestrial radiation due to the atmosphere [5].

A. Solar Radiation and Components

Global solar radiation $G(\beta, \gamma)$ is the total shortwave radiation received on a surface tilted at angle β and oriented towards γ. The direct component $B(\beta, \gamma)$ is the radiation received from the sun without any attenuation (scattering and/or absorption) through the atmosphere per unit area. For a tilted surface oriented to the equator, it is expressed as $B(\beta)$. The direct normal irradiance is the radiation reaching a surface always perpendicular to the rays coming from the sun at its actual position in the sky during the day, expressed as B_n. The diffuse radiation $D(\beta, \gamma)$ is the solar radiation per unit area reaching the surface from different paths because of the energy scattered by the atmosphere. The albedo radiation $R(\beta, \gamma)$ is the radiation reflected from the ground. If the albedo is unknown, a value of $\rho = 0.2$ is typically assumed for PV applications and used as a weight of global horizontal irradiance corrected to take into account the tilted angle [6].

B. Solar Models

Solar models are developed to describe the solar influence on the earth under specific conditions. The solar irradiance on the earth's surface has been broadly studied as a function of the sun elevation angle, water vapor in the atmosphere, aerosol concentration and altitude [7], [8]. The complexity of the models depends on the number of required inputs. One of the most important parameters is the position of the sun relative to the earth's surface. Therefore, analysis based on the solar theory is taken to describe the intensity of the sun radiation reaching the surface throughout the day. Clear sky models are used to estimate the irradiance level at a certain time instant and at a specific location assuming cloudless conditions. The application of these models is aimed to obtain results to be compared to measured data, in order to investigate the possible additional factors affecting the measured data. A large amount of studies can be found in the literature about clear-sky models. Some of them perform comparatives performance analysis to determine the global horizontal irradiance components based only on geometric calculation (zenith angle dependent) [9]–[12]. The monthly and daily insolation (energy during a period) is used to assess feasibility studies for solar energy systems, if the irradiation measurements are available. Such measurements are usually carried out be means of pyranometers that are sensitive to the wavelength of the solar radiation and to its angle of incidence [13]. If no measurements are available, the models for solar radiation can be used to estimate the radiation on the surface at a specific location. According to its acceptance in the literature [9], [10], [14], the model presented in [15], lately corrected by Ineichen and Perez [16], has been used in this study as a reference for the clear sky radiation in Milan, Italy.

III. METHODOLOGY

The main solar radiation data used in this work are extracted from the NWP products provided by the ECMWF, namely the ERA5 database, and the power generated by a set of PV modules installed on the rooftop of a building of Politecnico di Milano, Milan, Italy. The ERA5 database contains NWP data with temporal and spatial resolution equal to 1 hour and $0.28° \times 0.28°$, and offers a wide choice of products to be used according to the objective of the analysis. In the present work, we have used the shortwave radiation and the temperature to model the PV panels performance. Specifically, the products extracted from the database are listed in Table I.

The data measured by the PV panels have temporal resolution of 15 minutes. The output power is the average power during the acquisition period. Being the acquisition period smaller than the ERA5 forecast data (1 hour), the output power of the PV is also averaged for one hour. The Ineichen-Perez clear-sky model consider the variables explained previously in Section II and the site elevation h, in meters. Further explanation of the equations referred to this model can be found in [16].

A MATLAB code of the adjusted clear sky model by Ineichen and Perez was implemented using the open source tools developed by the PVPerformance Modelling Collaborative (PVPMC) group from Sandia National Laboratories [17]. The model enables to calculate the normal incident clear sky radiation on the surface for any day and time of the year. The

Table I
ERA5 PRODUCTS USED.

Variable	ERA5 code	range/value
Temperature at 2 m above surface	`t2m`	K
Surface solar radiation downwards, $G(0)$	`ssrd`	$\mathrm{J/m^2}$
Total sky direct solar radiation at surface, $B(0)$	`fdir`	$\mathrm{J/m^2}$
Surface clear sky solar radiation downwards, $G_{\mathrm{cs}}(0)$	`ssrdc`	$\mathrm{J/m^2}$
Total clear sky direct solar radiation at surface, $B_{\mathrm{cs}}(0)$	`cdir`	$\mathrm{J/m^2}$
Calculated total sky diffuse solar radiation at surface, $D(0)$ †	`ssrd - fdir`	$\mathrm{J/m^2}$
Calculated total clear sky diffuse solar radiation at surface, $D_{\mathrm{cs}}(0)$ †	`ssrdc - cdir`	$\mathrm{J/m^2}$

† Not an available product. Calculated from other products.

AM was calculated for every θ_z in every hour considered in the analysis. The Linke turbidity for the site was derived from the SoDa website services [18]. T_{L} is equal to 1 in the case of a perfectly dry and clean sky day. In turbid atmospheres, T_{L} can achieve values up to 6–7. Once global horizontal radiation, $G(0)$, and normal radiation, B_{ncI}, from the Ineichen-Perez model, are calculated, the diffuse component is

$$D(0) = G(0) - B_{\mathrm{ncI}} \cos\theta_z, \tag{1}$$

where the expression $B_{\mathrm{ncI}} \cos\theta_z$ represents the radiation on the horizontal plane. Power generation was calculated considering the tilt angle $\beta = 30°$ and the azimuth orientation to the south. The angle of incidence on the tilted surface θ is

$$\cos\theta = \cos\theta_z \cos\beta + \sin\theta_z \sin\beta \cos(\gamma_s - \gamma). \tag{2}$$

To obtain $B(\beta,\gamma)$ on the tilted surface, $B(0)$ and $B_{\mathrm{cs}}(0)$ coming from ERA5 database must be divided by the cosine of the solar azimuth, since these products are values of horizontal radiation and become values of normal radiation. When the incident radiation is falling on the back of the surface according to the tilt angle, $B(\beta,\gamma)$ is zero, otherwise it is obtained multiplying the normal radiation by the cosine of the incidence angle θ.

To obtain the tilted diffuse component $D(\beta,\gamma)$, the horizontal diffuse radiation $D(0)$ is multiplied by a factor that includes the tilted angle, considering the isotropic diffuse component approach.

The albedo irradiance is usually low or neglected, except for the cases of very bright surfaces (e.g., snow cover) that can strongly reflect the incident radiation. In this study the albedo irradiance has been supposed to be a portion of the global horizontal irradiance.

The global radiation $G(\beta,\gamma)$ on a tilted surface is the sum of the direct, diffuse and albedo radiation components

$$G(\beta,\gamma) = B(\beta,\gamma) + D(0)\frac{1+\cos\beta}{2} + \rho G(0)\frac{1-\cos\beta}{2}. \tag{3}$$

The power generation is calculated considering the parameters that characterize the PV Module Suntech STP250S-20/WD and the microinverter Aurora 300 W installed on the rooftoop of the Politecnico di Milano building. The data collected by the PV measurements from 2014 to 2016 were compared with the outcomes of the above-mentioned model relying on ERA5 data.

IV. RESULTS

Four different series of power generation were obtained for the same photovoltaic system in Politecnico di Milano. Figure 1 shows the power generated in the month of July in 2014 using as input the NWP products `ssrd`, `fdir` and `t2m` (see Table I) to calculate the PV generation. Each line represents the generation for a given day of the month. Observing the results, the bell shaped line (highest power production) is representative of clear sky conditions.

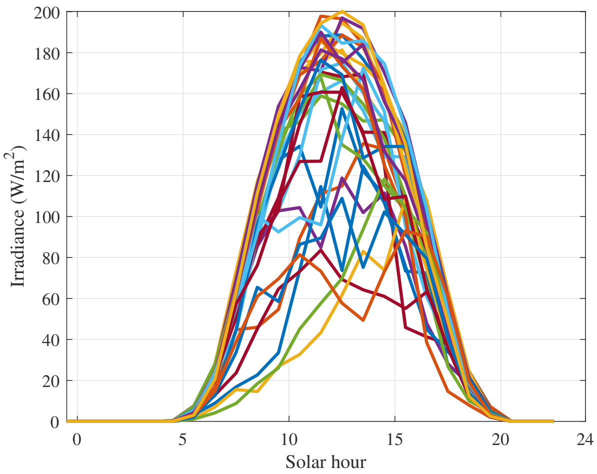


Figure 1. Daily PV production based on ERA5 forecasted in July 2014.

The clear-sky radiation products `ssrdc` and `cdir` in ERA5 data are computed with the same atmospherical conditions as temperature, Linke turbidity and pressure as the products `ssrd` and `fdir`, but regardless of the clouds in the atmosphere Figure 2 shows the comparison of the PV system power generation using the radiation of the Ineichen-Perez clear-sky model, ERA5 clear sky and non-clear sky data and the panel measurements for a specific dat (July, 31st 2014). The plot presents a very similar production for all datasets. On the other hand, Figure 3, which refers to a cloudy day (July, 30th of 2014), shows that only the measured power and the power predicted using ERA5 data are in good agreement. In general, the PV generation obtained from ERA5 data are consistent with the Ineichen-Perez model, in clear sky conditions, and with the local measurements collected in site, both in clear sky and cloudy conditions.

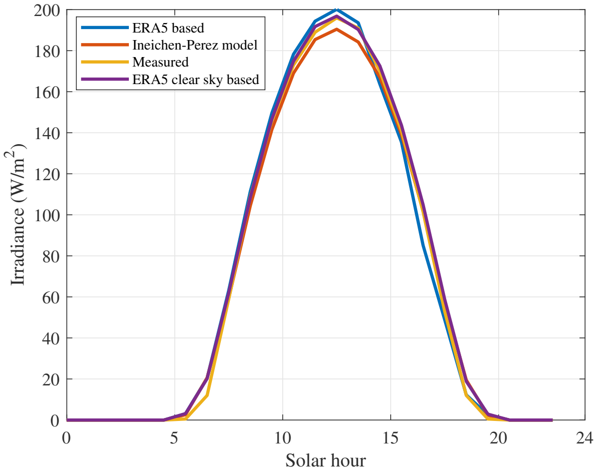


Figure 2. PV production for July 31st, 2014.

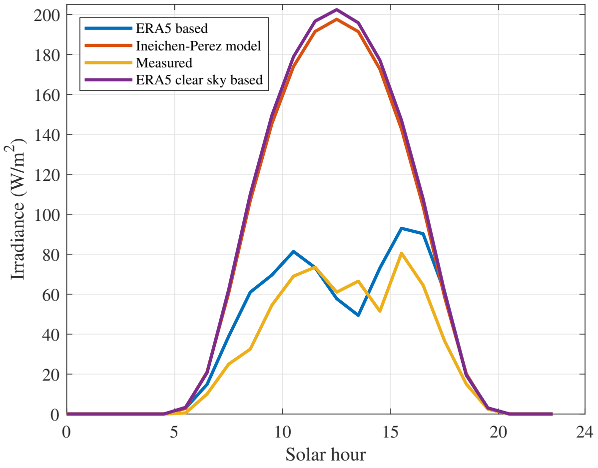


Figure 3. PV production for July 30th, 2014.

Table II reports the Pearson correlation coefficient obtained from the measurements collected using the PV panels and from the ERA5-derived power data: values higher than 0.9 are obtained for all years from 2014 to 2016, which corroborates the proposed approach based on the ERA5 database.

A CCDF was calculated for each year from 2014 to 2016 to assess the model performance also on a statistical basis. As shown in Figure 4 (next page), the clear sky model has a higher probability to produce the same amount of power if compared to the measured data and the power outputs obtained from the ERA5 NWP data. More importantly, the agreement between the data curve and the model curve is satisfactory, which suggests that the latter can be used as a reliable tool to design a PV power generation system on a global basis, i.e., depending on the local solar radiation characteristics, including the effects of clouds.

Table II
CORRELATION FACTOR FOR MEASURED POWER AND ESTIMATED ERA5-BASED POWER.

	2014	2015	2016
Correlation	0.904	0.912	0.917

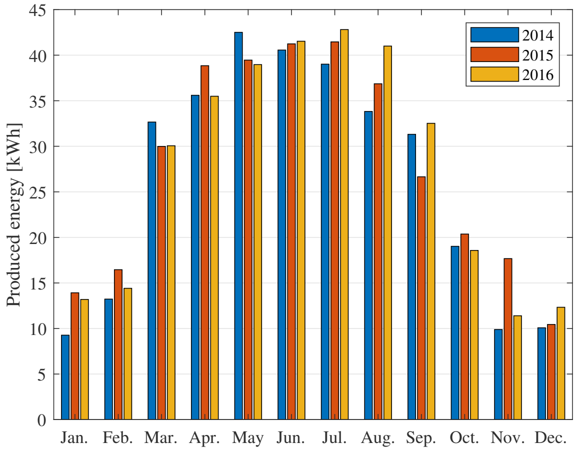


Figure 5. Monthly energy production.

A. *Errors*

Once data have been confronted a relation was obtained by a least squares method, fitting a polynomial curve with the form $y = ax + b$. Where, x is the output power derived from ERA5 database and y is the measured power output of the PV system. In Figure 6, a and b values are obtained using the 3-year dataset.

The performance of the ERA5-based prediction model was assessed by calculating the normalized mean absolute error (NMAE) and the normalized root mean square error (NRMSE)

$$\mathrm{NMAE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_{\mathrm{mes},i} - P_{\mathrm{est},i}|}{P_{\mathrm{n}}} \times 100 \tag{4}$$

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}|P_{\mathrm{mes},i} - P_{\mathrm{est},i}|^2}}{\max\left(P_{\mathrm{mes}}\right)} \times 100, \tag{5}$$

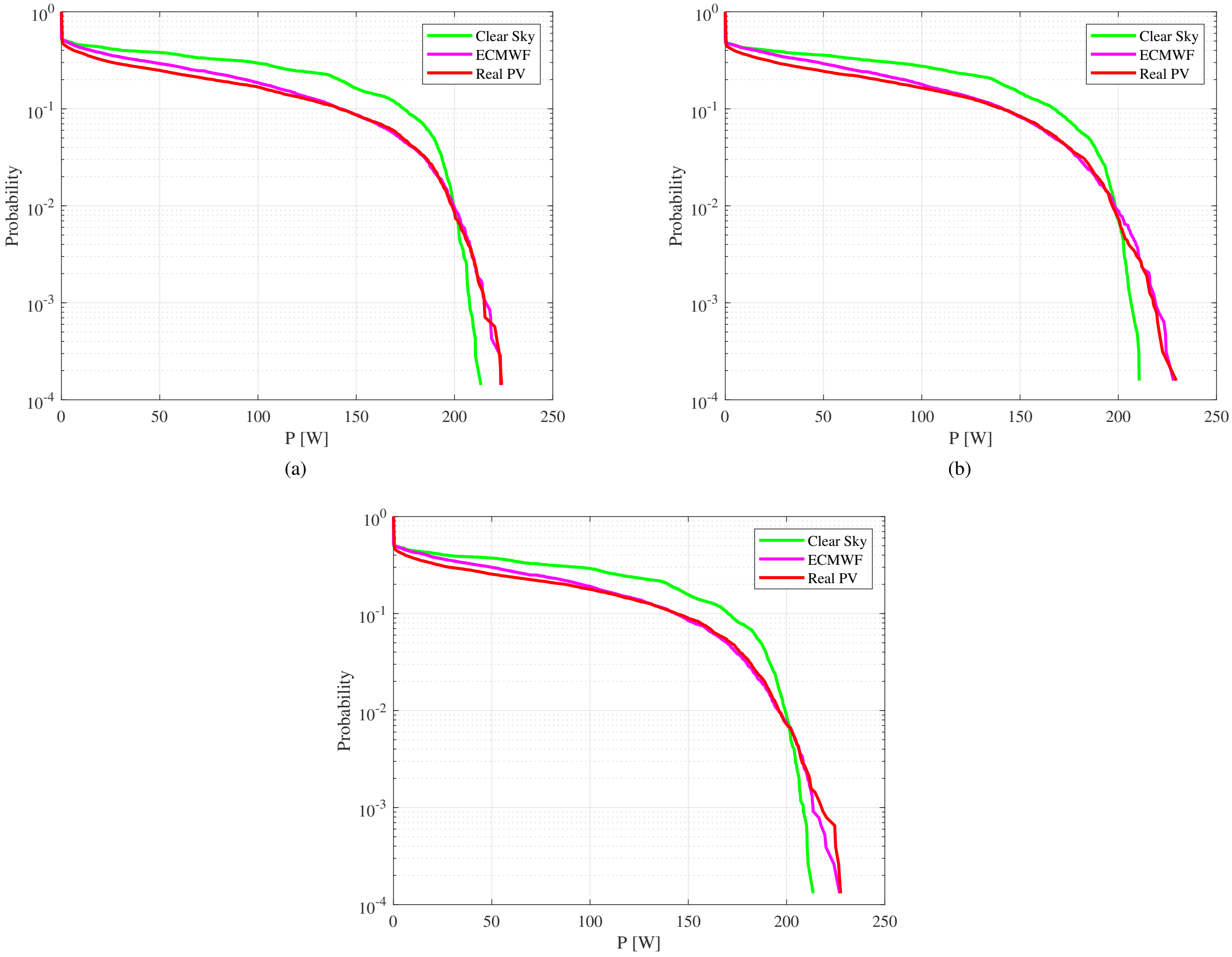
where $P_{\mathrm{mes},i}$ is the average power measured in the hour i, $P_{\mathrm{est},i}$ is the power predicted in the same hour i (ERA5-based model or Ineichen-Perez model), and P_{n} is the nominal capacity of the power plant. These error figures (ERA5 model vs. measurements) are presented in the Table III. Likewise, for the clear-sky case, we have compared the Ineichen-Perez model with the clear-sky prediction obtained from ERA5 data.

B. *Energy production*

Figure 5 reports, for each year, the monthly energy produced during a year based on the ERA5 predictions. Table IV (previous page) presents the equivalent hours in each month considering the ERA5 based predictions. In some months, no measurements are available, and accordingly, also ERA5 based outputs were omitted.

Table IV
MONTHLY PEAK-SUN EQUIVALENT HOURS.

Month	2014			2015			2016		
	ERA5	Measured	ERA5 adjusted	ERA5	Measured	ERA5 adjusted	ERA5	Measured	ERA5 adjusted
Jan.	51.3	37.0	51.3	80.3	62.5	73.4	74.9	64.5	74.9
Feb.	65.0	52.3	65.0	81.5	67.4	81.5	71.7	60.5	70.5
Mar.	146.6	133.4	146.6	134.8	119.7	134.8	134.4	49.9	64.2
Apr.	147.6	133.7	147.6	161.7	97.5	108.3	146.6	135.9	142.8
May	169.4	157.0	169.4	157.4	90.7	103.4	155.3	132.3	150.5
Jun.	159.7	148.5	159.7	162.6	128.8	140.7	163.9	94.4	101.7
Jul.	154.6	140.0	154.6	164.0	N.A.	N.A.	169.4	161.0	169.4
Aug.	137.8	46.1	47.3	150.2	N.A.	N.A.	167.4	85.0	89.8
Sep.	135.5	56.1	61.2	115.0	84.1	81.6	141.5	127.5	136.0
Oct.	92.6	25.1	32.4	98.7	86.2	98.7	88.2	81.9	88.2
Nov.	53.0	9.7	17.5	98.1	76.9	98.1	61.5	46.5	61.5
Dec.	61.2	42.3	61.2	62.0	54.1	62.0	76.0	63.2	76.0
Tot.	1374.4	981.2	1113.8	1466.4	867.9	982.6	1450.9	1102.5	1225.6



(a)

(b)

(c)

Figure 4. Yearly cumulative complementary distribution functions (CCDFs) of power generation: (a) 2014; (b) 2015; and (c) 2016.

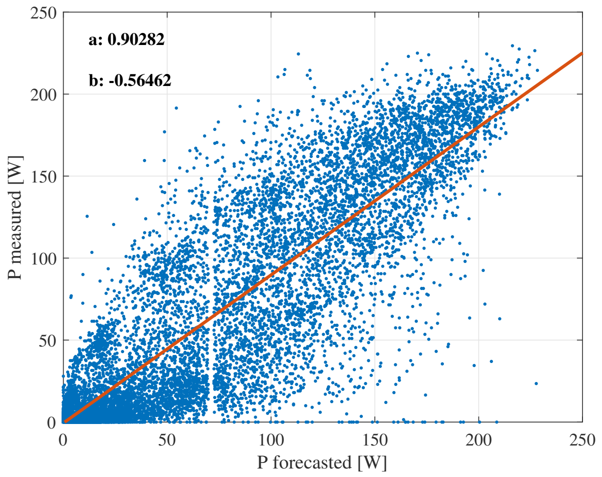


Figure 6. Regression line based on the years 2014–2016.

Table III
ERA5 DATABASE ERROR VALUES.

		2014	2015	2016
Real conditions	NMAE [%]	5.546	4.96	5.18
	NRMSE [%]	11.60	10.75	10.66
Clear-sky data	NMAE [%]	1.24	1.37	1.18
	NRMSE [%]	1.81	2.96	2.69

V. CONCLUSION

The use of NWP data, the ERA5 database recently made available by ECMWF, as input to a model to forecast PV power generation was assessed for a specific site (i.e., Milan, Italy). The predictions obtained by the model, which also include the effects of clouds, were compared against 3 years of power data collected by a set of PV solar panels installed at Politecnico di Milano premises. Results of the investigation indicate that the proposed model offers a satisfactory performance in forecasting the PV power generation, both in clear sky and cloudy conditions: indeed the correlation coefficient between the measurements and the model's outputs is always higher than 0.9, regardless of the considered year, with the NMAE and NRMSE values being around 5% and 11%, respectively.

Also results on a statistical basis confirm the good performance of the proposed model, with a good agreement between the generated power CCDF obtained from the measurements and from the outputs of the model. In addition, the normalized mean absolute error is approximately 13 W, which is nearly 5% of the nominal power.

Finally, the model performance was also tested only in clear sky conditions using as reference the outputs of the Ineichen-Perez model: also these results confirm the accuracy of the proposed model by showing even lower values of NMAE and NRMSE, approximately equal to 1.2% and 2.5%, respectively. In addition, there were obtained characteristic values for the studied location regarding equivalent sun-peak hours which can be used as a tool for deeper analysis of a PV system in the area.

Overall, although additional data are needed to perform further tests, results suggest that the proposed model based on ERA5 data can be reliably used to design solar plants as well as to forecast their performance in time.

REFERENCES

[1] "Shell energy scenarios to 2050," Shell International BV, Tech. Rep., 2011.

[2] M. Kudo, A. Takeuchi, Y. Nozaki, H. Endo, and S. Jiro, "Forecasting electric power generation in a photovoltaic power system for an energy network," *Electrical Engineering in Japan (English translation of Denki Gakkai Ronbunshi)*, vol. 167, no. 4, pp. 16–23, 2009.

[3] F. Kasten, "The linke turbidity factor based on improved values of the integral Rayleigh optical thickness," *Solar Energy*, vol. 56, no. 3, pp. 239–244, Mar. 1996.

[4] J. W. Spencer, "Fourier series representation of the sun," *Search*, vol. 2, no. 5, p. 172, 1971.

[5] J. Remund and D. Domeisen, "Aerosol optical depth and Linke turbidity climatology," IEA SHC Task 36, Bern, Tech. Rep., 2010.

[6] A. Luque and S. Hegedus, *Handbook of Photovoltaic Science and Engineering*, 2nd ed. John Wiley & Sons Inc, 2010.

[7] V. Badescu, "Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance," *Solar Energy*, vol. 61, no. 4, pp. 251–264, Oct. 1997.

[8] T. Stoffel, D. Renné, D. Myers, S. Wilcox, M. Sengupta, R. George, and C. Turchi, "Best Practices Handbook for the Collection and Use of Solar Resource Data," NREL, Tech. Rep., 2010. [Online]. Available: www.nrel.gov

[9] C. A. Gueymard, "Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models," *Solar Energy*, vol. 86, no. 8, pp. 2145–2169, Aug. 2012.

[10] J. S. Stein, C. W. Hansen, and M. J. Reno, "Global horizontal irradiance clear sky models : implementation and analysis." Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA (United States), Tech. Rep., Mar. 2012.

[11] P. Ineichen, "Comparison of eight clear sky broadband models against 16 independent data banks," *Solar Energy*, vol. 80, no. 4, pp. 468–478, Apr. 2006.

[12] ——, "Validation of models that estimate the clear sky global and beam solar irradiance," *Solar Energy*, vol. 132, pp. 332–344, Jul. 2016.

[13] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 3rd ed. John Wiley & Sons Inc, 2006.

[14] J. Polo, L. Zarzalejo, L. Martín, A. Navarro, and R. Marchante, "Estimation of daily Linke turbidity factor by using global irradiance measurements at solar noon," *Solar Energy*, vol. 83, no. 8, pp. 1177–1185, Aug. 2009.

[15] F. Kasten, "A Simple Parametrization of the Pyrheliometric Formula for Determining the Linke Turbidity Factor," *Meteorol. Rundsch.*, vol. 33, pp. 124–127, 1980.

[16] P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," *Solar Energy*, vol. 73, no. 3, pp. 151–157, Sep. 2002.

[17] PVPerformance Modeling Collaborative. Sandia National Laboratories. [Online]. Available: http://pvpmc.sandia.gov/

[18] Solar Radiation Data (SoDa). O.I.E. (Observation, Impacts, Energy). [Online]. Available: http://www.soda-pro.com/