Simplified Model to Approach the Theoretical Clear Sky Solar PV Generation Curve through a Gaussian Approximation

J. Alpízar-Castillo*

Department of Electrical Engineering, Fidélitas University, COSTA RICA.

Abstract

Radiation forecast is the milestone of the solar energy industry, making possible the existence of the whole market. Solar radiation models allow scientists and engineers to predict the behaviour of a PV system to perform technical and economic analysis. Despite the existence of numerous models, most of them are highly complex or require massive amounts of data, limiting solar energy start-ups. As a result, a state-of-the-art review was performed and based on it this study proposed a simple model that allows emerging solar companies to create preliminary analysis for their clients. The proposed model calculates the instant power of the envelope curve of PV generation, based on the Gaussian bell equation, by using the daily specific energy and a deviation proportional to the sun hours of the geographical information as parameters. With accurate meteorological data, results showed an acceptable performance, with a more straightforward implementation, when compared against those reported in the literature.

Keywords: photovoltaic systems, PV generation curve, PV modelling, solar energy, solar forecasting

Nomenclature

Symbol	Variable	Units
E_{PV}	PV energy	kWh
E_{sd}	Daily specific energy	kWh/kW
η	Energy conversion safety factor	
P_n	Nominal PV power	kW
P_{PV}	Estimated PV power	kW
t_n	Noon time	h
t_{sr}	Sunrise time	h
t_{ss}	Sunset time	h
μ	Mean PV power time	h
σ	PV deviation	h

1. INTRODUCTION

To accurately determine the revenue of a PV system, the designer needs to compare the load profile against the solar generation curve, to estimate the energy that is directly consumed by the load and the one that needs to be sent back to the grid or stored. The load profile is usually measured. On the other hand, the estimation of the solar generation curve requires at least one of the following, due to its complexity: expensive software, high knowledge in databases, high knowledge in AI, or high computational resources, limiting the solar energy start-ups. Therefore, an easy implementation model is proposed.

---

*Corresponding author (**Tel:** +506 8950 9950)
Email address: Joel.AlpizarCastillo@gmail.com (J. Alpízar-Castillo)

Scholars have studied models for the radiation behaviour for about a century. For instance, in the 1920's [1] reported studies in forecasting weather and solar radiation, followed by the 1950's hourly values presented in [2]. Moreover, nowadays scientists have access to a wide variety of methods to forecast the solar radiation, that go from from statistical analysis to artificial intelligence (AI), even hybrid methods that combine two or more methods. Table 1 summarizes the most common methods in PV forecasting and their results, comparing the forecasts against measured data. The ranges indicate the best performance obtained; outliers were not considered. A deeper state-of-the-art review can be found in [3].

As observed in Table 1, the current models are complex and required specialized scientists, as mentioned in [4]. On the other hand, early simplified models based on exponential functions were presented in [7, 15, 16]. However, those studies were limited to radiation and not to generated power, and further analysis were not found in the literature. Hence, and because of the similar shape in the measured distribution of instant power during the day to a bell, this work presents a model based on the Gaussian bell equation to calculate the monthly average instant power produced by a PV system. The results showed an acceptable behaviour when compared against the state-of-the-art model's performance if accurate meteorological data is used as input.

Table 1: State-of-the-art summary in PV forecasting methods.

Classification	Method	Reference	Best accuracy range [%]
Meteorological models	Numerical weather prediction (NWP)	[4]	24.2 < error < 41.1
	Clear-sky solar irradiance–model: ASHRAE	[5]	5 < error < 30.8
	Clear-sky solar irradiance–model: Semi-sinusoidal model	[6]	9.13 < error < 35.1
	Clear-sky solar irradiance–model: Collares-Pereira and Rabl model	[7]	9.5 < error < 17.8
	Direct satellite measurement	[8, 9]	2 < error < 8
Statistical methods	Autoregressive integrated moving average (ARIMA)	[4, 10]	9.52 < error < 16.49
	Regression	[4]	30 < error < 41.2
AI	Artificial neural networks (ANN)	[4, 10–13]	4.75 < error < 16.77
	Adaptive neuro fuzzy–inference systems (ANFISs)	[10, 14]	6.82 < error < 12.32
	Least-squares support–vector machine (LS-SVM)	[10]	4.78 < error < 12.03
Hybrid methods		[4, 10]	3.43 < error < 6.57

2. MATHEMATICAL DEFINITION OF THE PROPOSED MODEL

As is well known, the equation for a Gaussian bell is given as

$$f(x) = a e^{-\frac{(x-b)^2}{2c^c}} \qquad (1)$$

where a refers to the maximum height of the bell, and the expression in the exponent defines its width. Also, Eq. (1) the probability density function when written as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (2)$$

As displayed in Fig. 1, the ideal PV generation curve is bell-shaped also, therefore, the hypothesis presented in this work is to define whether it can be accurately approached as a Gaussian bell, where the area under the curve is the energy produced by the system, i.e., to approach the power generated by the PV system as a bell-shaped function of time.

As the integral of the bell is also the energy produced by the PV system, it is obtained that

$$E_{PV} = \int_{-\infty}^{+\infty} a e^{-\frac{(x-b)^2}{2c^c}} dx = a|c|\sqrt{2\pi} \qquad (3)$$

Given Eqs. (2) and (3), the PV power can be defined as

$$P_{PV}(t) = \frac{E_{PV}}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^c}} \qquad (4)$$

where μ is the mean power time, σ is the deviation used to adjust the width of the bell, and E_{PV} is the output energy from the system, that can be estimated in terms of the specific energy production as

$$E_{PV} = \eta P_n E_{sd} \qquad (5)$$

where η is a conversion safety factor, P_n is the rated PV system power, and E_{sd} is the daily specific energy production per unit of power. The specific energy can be obtained through meteorological tools as SolarGIS (used in this study), Heliosat-2, EnMetSol, IrSolAv and others. Comparisons of those tools can be found in [17]. Note that Eq. (4) is bounded by the sun hours, thus, a valid expression for the whole day can be obtained if the sunrise time (t_{sr}) and the sunset time (t_{ss}) are considered, redefining Eq. (4) as

$$P_{PV}(t) = \begin{cases} 0 & t < t_{sr} \\ \frac{E_{PV}}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^c}} & t < t_{sr} < t < t_{ss} \\ 0 & t_{ss} \end{cases} \qquad (6)$$

3. COMPARISON WITH GROUND-MEASURED DATA

To obtain a monthly representative clear sky curve for the model's proof-of-concept, I analysed historical raw data from a 202 kW PV plant located in Alajuela, Costa Rica, to obtain a monthly representative clear sky curve, considered as the envelope curve of the daily data, as shown in Fig. 1.

As can be seen in Fig. 1, the distribution is not symmetric if an axis is defined at the noon time (t_n = 11:49), but slightly left skewed (μ = 12:40), a pattern that is consistent throughout the year, changing about half hour from the earliest to the latest in t_n, and about one hour in . Therefore, the mean power time is not the same of the noon time ($t_n \neq \mu$), but proportional, and must be defined according with the meteorological data of the geographical location and the month of the year.

For the present work, the relation between the noon time (taken from [18] and the mean power time (obtained from the data) is summarized in Table 2. The results show that the earlier the
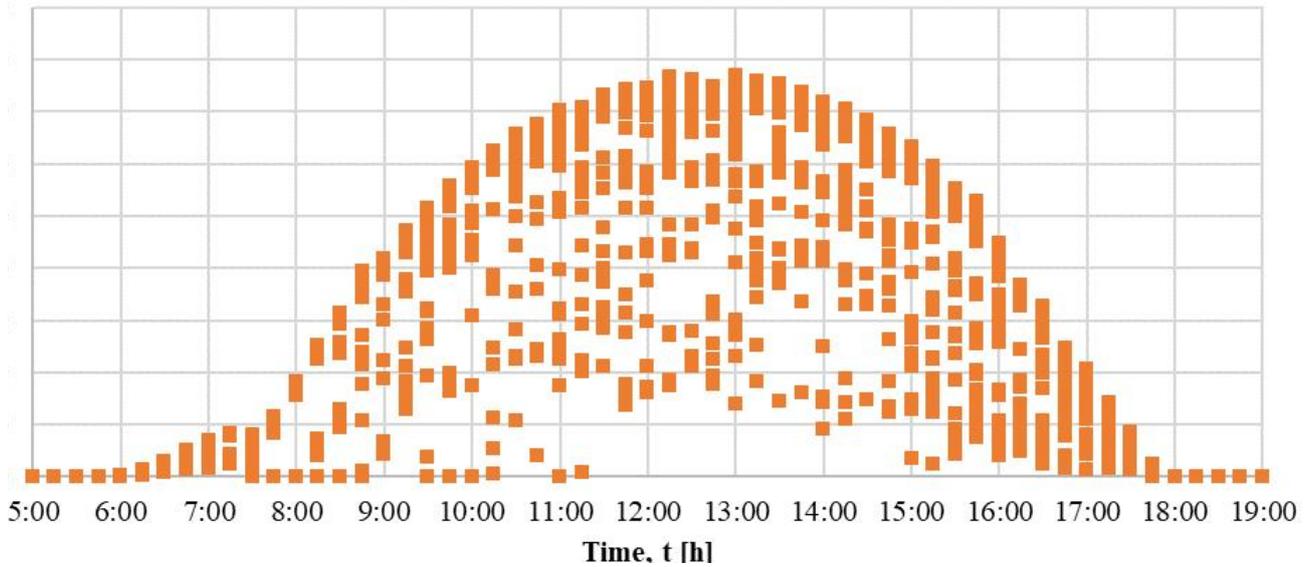
Figure 1: P_{PV} daily measures every 15 min during March, (t_{sr} = 5:46, t_{ss} = 17:52 and t_n = 11:49)

noon time, the closer the mean power time is to it.

Given the above, Eq. (6) must now include a new factor associated with the skewness of the data. A common solution to obtain a skewed bell is using the error function $erf(x)$. In this case, two error functions will be used to ensure the function crosses zero at the sunrise and at the sunset, obtained in Eq. (7).

$$
P_{PV}(t) = \begin{cases} 0 & t < t_{sr} \\ \frac{E_{PV}}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^c}} erf\left[-\frac{\alpha_1(t-t_{sr})}{\sigma\sqrt{2}}\right] erf\left[-\frac{\alpha_2(t-t_{ss})}{\sigma\sqrt{2}}\right] & t_{sr} < t < t_{ss} \\ 0 & t > t_{ss} \end{cases}
\tag{7}
$$

Table 2: Time parameters for the function per month.

Month	Noon time t_n	Mean power time μ
January	11:49	12:45
February	11:54	12:40
March	11:49	12:40
April	11:39	12:30
May	11:36	12:30
June	11:41	12:30
July	11:46	12:30
August	11:44	12:25
September	11:35	12:05
October	11:25	11:50
November	11:24	12:00
December	11:35	12:20

being α_1 and α_2 shape factors that are directly proportional to the slope of the bell near to the sunrise and the sunset, respectively. In this case is assumed that $\alpha_1 = -\alpha_2$, with $\alpha_1 < 0$.

4. RESULTS AND DISCUSSION

To analyze the performance of the model, is required to determine the deviations that minimize the calculated error between the observed load profile (measured every 15 minutes) and the calculated instant power. Several empiric expressions to calculate those deviations are proposed by [7] for similar methods, presented in [15, 16], but neither of them showed accurate approximations. A simple optimization was used to minimize the overall error in the instant power of the envelope curve and the results obtained using Eq. (7). The outcome deviations are shown in Table 3, and when compared against the noon times in Table 3, it is observed that the earlier the noon time, the smaller the deviation.

To visualize the model previously described, a normalized analysis of the average yearly curve will be performed, whose results are presented in Fig. 2. As can be seen, when using Eq. (6) to approach the envelope curve, there is a significant error near the sunrise and sunset, which is minimized using the error functions considered in Eq. (7). On the other hand, in both cases there are two error regions associated with the difference
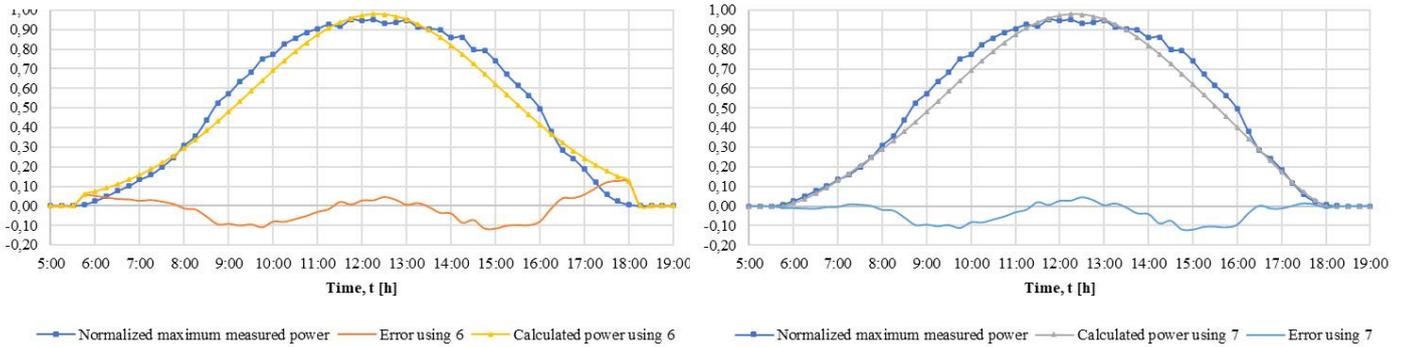
(a) Normalized results using Eq. 6 and the average yearly envelope curve.

(b) Normalized results using Eq. 7 and the average yearly envelope curve ($\alpha_2 = 3$).

Figure 2: Normalized results using Eqs. 6 and 7 and the average yearly envelope curve.

Table 3: Deviation of the function per month.

Month	Deviation , σ [h]
January	2,72
February	2,95
March	2,85
April	2,95
May	2,60
June	2,60
July	2,80
August	2,80
September	2,65
October	2,55
November	2,60
December	2,60

Table 4: Deviation of the function per month.

Month	Error [%]
January	-6.56
February	-8.66
March	-7.54
April	-8.32
May	-3.83
June	-3.83
July	-5.37
August	-5.14
September	-3.37
October	-2.54
November	-3.56
December	-4.04

in the width of the bell. The error can be minimized if the functions are deconvoluted into two new smaller bell functions. However, it would increase the complexity of the model, and the main goal is to keep the model as simple as possible. Furthermore, this model will be used as basis for a more realistic model, adding the effect of cloudiness, thus, it should be as simple as possible in order to simplify further works.

The errors in the obtained energy, associated with the implementation of Eq. (6) and (7), showed in Fig. 2, are -2.75% and -6.34%, respectively. Nonetheless, Eq. (7) presents a better fit, having mostly a negative error in the instant power and resulting in a more conservative approach, which is usually preferred, and therefore selected for the analysis. The process was repeated for each month, the obtained errors in energy are presented in Table 4 and the instant power errors are shown in Fig. 3.

The obtained results, shown in Table 4, are comparable to the reported results, presented in Table 1, showing high accuracy in the forecast, especially when its simplicity is considered. Likewise, the behaviour shown in Fig. 3 is also consistent with the literature, with a high density of instant errors below 10% and high error outliers spread among the day, as reported by [4, 9], since

the measurement data is not smooth, but the proposed model is shown in Fig.2a.

The obtained results, shown in Table 4, are comparable to the reported results, presented in Table 1, showing high accuracy in the forecast, especially when its simplicity is considered. Likewise, the behaviour shown in Fig. 3 is also consistent with the literature, with a high density of instant errors below 10% and high error outliers spread among the day, as reported by [4, 9], since the measurement data is not smooth, but the proposed model is shown in Fig. 2.

5. CONCLUSIONS AND LIMITATION

A novel straightforward model was proposed, allowing preliminary forecasts in the instant power that can be obtained from a PV system, with low computational cost when compared against the most common methods. The results showed a satisfactory performance if the normalized expected value is compared against the normalized envelope curve. However, the model is limited by the quality of the radiation data used as input, therefore, to obtain accurate results, accurate meteorological data is required. Further studies on the stochastic effect of external factors, like cloudiness, rain and wind would allow to obtain results closer to the real conditions.
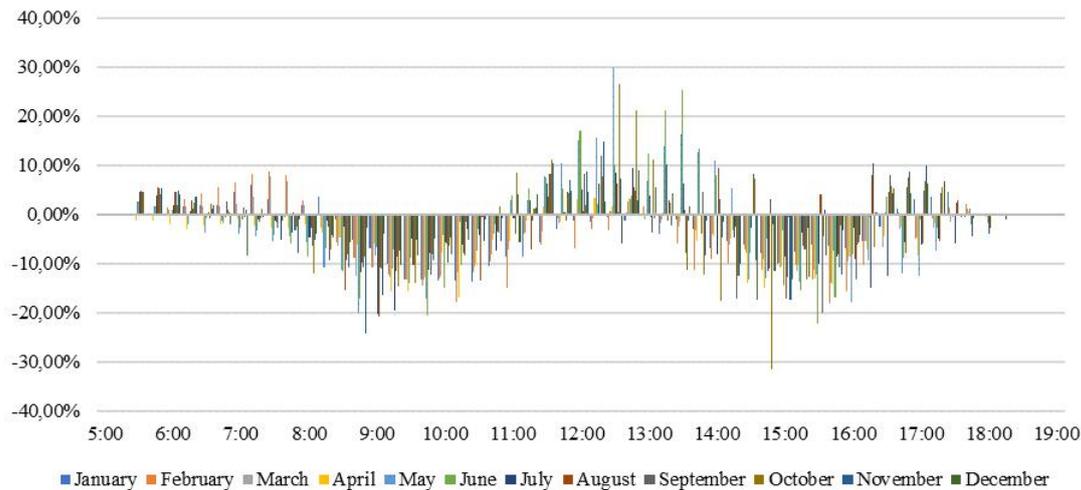
Figure 3: Instant P_{PV} error distribution per month.

Photovoltaic systems for self-use are an increasing market. Hence, solar radiation models are required for their technical and economic analysis. After evaluating the most common models, it was determined that most of them have complex requirements, mostly considerable amounts of data, high computational resources, and expertise. On the other hand, software are constantly developed to simplify the aforementioned analysis, nonetheless, their costs are high, and limit emerging companies. Given the above, the model presented in this work aims to be a tool for emerging companies in their early stages, while they can afford more complex and precise software.

References

[1] C. F. Marvin and H. H. Kimball, "Solar radiation and weather forecasting," *Journal of the Franklin Institute*, vol. 202, no. 3, pp. 273–306, 1926.

[2] A. Whillier, "The determination of hourly values of total solar radiation from daily summations," *Arch. Met. Geoph. Biokl. B.*, vol. 7, pp. 197–204, 1956.

[3] D. Yang, J. Kleissi, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining," *Solar Energy*, vol. 168, no. 1, pp. 60–101, 2018.

[4] G. Reikard, "Predicting solar radiation at high resolutions: a comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342–349, 2009.

[5] S. A. Al-Sanea, M. F. Zedan, and S. A. Al-Ajlan, "Adjustment factors for the ASHRAE clear-sky model based on solar radiation measurements in Riyadh," *Applied Energy*, vol. 79, no. 2, pp. 215–237, 2004.

[6] Y. El-Mghouchi, A. El-Bouardi, Z. Choullo, and T. Ajzoul, "New model to estimate and evaluate the solar radiation," *International Journal of Sustainable Built Environment*, vol. 3, no. 2, pp. 225–234, 2014.

[7] M. J. Jamil Ahmad and G. N. Tiwari, "Study of models for predicting the mean hourly global radiation from daily summations," *Open Environmental Sciences*, vol. 10, no. 2, pp. 6–14, 2008.

[8] V. García-Heller, R. Espinasa, and S. Paredes, "Forecast study of the supply curve of solar and wind technologies in Argentina, Brazil, Chile and Mexico," *Renewable Energy*, vol. 93, pp. 168–179, 2016.

[9] S. Pfenninger and I. Staffell, "Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, no. 1, pp. 1251–1265, 2016.

[10] Y. K. Wu, C. R. Chen, and H. A. Rahman, "A novel hybrid model for short-term forecasting in PV power generation," *International Journal of Photoenergy*, p. 569249, 2014.

[11] J. Cao and X. Lin, "Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1396–1406, 2008.

[12] J. Liu, X. Fang, X. Zhang, and C. Yang, "An improved photovoltaic power forecasting model with the assistance of aerosol index data," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 434–442, 2015.

[13] J. Liu, W. Fang, X. Zhang, and C. Yang, "An improved photovoltaic power forecasting model with the assistance of aerosol index data," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 434–442, 2015.

[14] M. Chaabene and M. Ben Ammar, "Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems," *Renewable Energy*, vol. 33, no. 7, pp. 1435–1443, 2008.

[15] P. C. Jain, "Comparison of techniques for the estimation of daily global irradiation and a new technique for the estimation of hourly global irradiation," International Centre for Theoretical Physics, Tech. Rep. IC/84/09, 1984.

[16] A. Baig, P. Akheter, and A. Mufti, "A novel approach to estimate the clear day global radiation," *Renewable Energy*, vol. 1, no. 1, pp. 119–123, 1991.

[17] P. Ineichen, "Five satellite products deriving beam and global irradiance validation on data from 23 ground stations," Université de Geneve, Tech. Rep. UNIGE: 23669, 2011.

[18] L. Eglitis. (2020, May) Amanecer y atardecer en Costa Rica. [Online]. Available: https://www.datosmundial.com/america/costa-rica/puesta-del-sol.php