COMPARATIVE ANALYSIS OF METEOROLOGICAL DATABASES AND TRANSPOSITION MODELS APPLIED TO PHOTOVOLTAIC SYSTEMS

MICHELLE K. DA SILVA[1], DANTE I. NARVAEZ[1], KAREN B. DE MELO[1],
TATIANE S. COSTA[1], THAIS G. SIQUEIRA[2] , MARCELO G. VILLALVA[1]

[1]*FEEC-School of Electrical and Computer Engineering – UNICAMP – Campinas-SP – Brazil*
[2]*ICT-Instituto de Ciência e Tecnologia – UNIFAL – Poços de Caldas-MG*

E-MAILS: {michelle.kitayama,dante.innar}@gmail.com, {karenbm,tatianec}@dsce.fee.unicamp.br, thais.siqueira@unifal-mg.edu.br, marcelo@fee.unicamp.br

Abstract— Estimation of the solar irradiation incidence on the surface of photovoltaic panels is a determinant factor for the analysis of risks and the expected energy generated by photovoltaic systems. One of the biggest challenges in this field is to obtain the incident radiation on tilted surfaces since the radiation data available in meteorological databases considers only the radiation on the horizontal plane. For the conversion of the radiation data from the horizontal to the tilted plane, there are several mathematical models available in the literature, such as Perez's model and Hay's model. The analysis described in this study seeks to investigate the impact of these transposition models as well as the accuracy of the NASA SSE and Meteonorm 7.1 irradiation databases, regarding the power generation of a photovoltaic system located at the Tanquinho Photovoltaic Power Plant in Brazil.

Keywords— Transposition Models, Photovoltaic Systems, Irradiation on Tilted Surfaces, Perez's Model, Hay's Model.

Resumo— A estimativa da irradiação solar incidente na superfície dos painéis fotovoltaicos é um fator determinante para a análise de riscos e expectativa de produção de energia em sistemas fotovoltaicos. Um dos grandes desafios nessa área é obter a irradiação incidente no plano inclinado, haja visto que os dados de irradiação disponibilizados através de bancos de dados meteorológicos, consideram apenas a irradiação no plano horizontal. Para a conversão desses dados para o plano inclinado existem na literatura diversos modelos matemáticos tais como o modelo de Perez e o modelo de Hay. A análise descrita neste estudo busca investigar o impacto desses modelos de transposição e a precisão dos bancos de dados de irradiação da NASA SSE e Meteonorm 7.1, na geração de energia de uma planta fotovoltaica localizada na Usina Fotovoltaica de Tanquinho no Brasil.

Palavras-chave— Modelos de Transposição, Sistemas Fotovoltaicos, Irradiação no Plano Inclinado, Modelo de Perez, Modelo de Hay.

1 Introduction

The search for information related to the use of solar radiation as an alternative and sustainable energy source has been growing significantly in recent years. However, for an appropriate and optimized application of this system, it is necessary to have a detailed knowledge of the climatic characteristics and the incident solar radiation on the tilted surface of photovoltaic panels (Scolar, Martins, & Escobedo, 2003). From this, the initial stage for developing a photovoltaic system project is to evaluate the intensity of solar resources at the site of interest. Through this analysis, it is possible to obtain information on the potential of energy generation of the region and better estimate financial returns.

Several meteorological data platforms provide monthly and hourly average data of global radiation for the horizontal plane. Among these platforms, two of them stand out, the NASA SSE and Meteonorm 7.1 databases (Gueymard & Myers, 2008). The data contained in those databases are obtained through measurements made by solarimetric stations, estimated values through interpolations or values obtained through terrestrial mathematical models and based on satellite images (Oliveira, 2017). However, those data are generally only available for the horizontal plane. Therefore, it is necessary to apply mathematical models to estimate the radiation incidence on a tilted surface, considering that photovoltaic systems can be installed with a fixed slope or using solar tracker (Gueymard, 2009)

Gueymard (Gueymard, 1987) stated that the only apparent difference between the empirical transposition models is related to the way that they consider and analyse the diffuse radiation. In order to evaluate this difference, this article aims to investigate the performance of two transposition models: Hay's model (Hay, 1979) and Perez's *et al*. (Perez *et al.*, 1990). The accuracy of irradiation data from the NASA SSE and Meteonorm 7.1 datasets is also investigated. Both analyses are performed through simulation using the software PVSyst®. The energy generated by the solar plant simulated will be estimated by varying both the meteorological databases and the transposition models. The results obtained are compared with energy measured at Tanquinho Solar Photovoltaic Plant in Brazil. Also, the radiation data from the available sources were compared with data measured by a solarimetric station installed at the plant.

2 Meteorological and Solarimetric Databases

There are currently many sources of solar radiation data, commercial and public. Thevenard and Pelland (Thevenard & Pelland, 2013) evaluated several solar radiation databases and stated that they compile information from other sources, such as data measured in the field by solarimetric stations, data obtained through modelling and interpolations from satellites images. Each data sources creates uncertainties added to the system.

In the analysis described in the present work, two databases are considered: the NASA SSE and Meteonorm 7.1 databases. The mean global radiation values of each database are compared to each other and to the values measured by a solarimetric station installed at the Tanquinho Solar Photovoltaic Plant.

2.1 Meteonorm 7.1 Database

The version 7.1 of the Meteonorm database offers meteorological data measured from 1991 to 2010. This data is obtained through several solarimetric stations throughout many different locations. Although the data measured by solarimetric stations are the most accurate, there are cases in which this source is not available near the project site. In these cases, the database uses interpolations of satellite information to obtain irradiance data. Those interpolations generate disadvantages such as insertion of uncertainties, impossibility to obtain some meteorological parameters, difficulty in detecting multiple layers of clouds and lack of precision in areas with snow. Combined with the mentioned disadvantages those interpolations are modelled as a function of latitude and albedo in a way that the higher these values, the greater the uncertainties (Oliveira, 2017).

The uncertainties for the global radiation values are mainly based on three factors:

- Uncertainties generated by interpolations of field measurements and satellite images.
- Uncertainties generated due to the use of data transposition models and radiation decomposition.
- Uncertainties generated by measurements made locally, for example measurement errors and climatic variations.

2.2 NASA-SSE Database

This database was developed by the Prediction of Worldwide Energy Resource created by the National Aeronautics and Space Administration (NASA) with the objective of providing public information related to solar radiation and global meteorological data, especially for locations where no data is available, in order to encourage the development of renewable energy projects (Oliveira, 2017).

The data available in the NASA SSE database takes into account areas of 100 km per 100 km to perform the calculation of irradiance. These data were collected over a period of 22 years, from 1983 to 2005. This database provides the monthly averages data of global, direct and diffuse irradiance as well as meteorological information, those values are estimated from satellite images (Kumar & Waila, 2017)

Table 1 shows the mean values of global irradiance in the horizontal plane (W/m²) of each database under analysis and the measured irradiance data collected by the solarimetric station at the site.

2.3. Measured Data

The measured data used was collected via a solarimetric station installed at the Tanquinho Photovoltaic Power Plant (22°46’37.2688” S, 47°0’25.6536” W). The data were measured periodically in an interval of one minute, from the 1st of December 2014 to the 30th of November 2015.

The average values of global irradiation were performed similarly to the analysis done by Gomes & Escobedo (Gomes & Escobedo, 2007) and Riveros-Rosas (Riveros-Rosas et al., 2015). Where the statistical tool of arithmetic mean was used.

The sum of irradiance data measured in an one minute interval was divided by the number of minutes in one hour to obtain the hourly average global irradiance (I_g^h), as expressed in the equation (1)

$$I_g^h = \frac{1}{N_{min}} \sum_{i=0}^{N_{min}} I_g^{min,i} \tag{1}$$

Where N_{min} is the number of minutes in an hour and I_g^{min} is the irradiance data measured each minute.

The daily average of global irradiance (I_g^d) was obtained by the arithmetic mean of the hourly global irradiance, as shown by the equation (2)

$$I_g^d = \frac{1}{N_h} \sum_{i=0}^{N_h} I_g^{h,i} \tag{2}$$

Where N_h is the number of hours in a day.

Finally, the monthly average of global irradiance (I_g^m) was calculated by the arithmetic mean of the daily average of global irradiance in each month, as represented by the equation (3)

$$I_g^m = \frac{1}{N_d} \sum_{i=0}^{N_d} I_g^{d,i} \tag{1}$$

Where N_d is the number of days in each month.

3 Transposition Models to Predict the Diffuse Radiation on Tilted Surfaces

Similarly to the meteorological database, the transposition of the solar irradiance incident on a tilted plane from irradiance values on the horizontal plane generates uncertainties. The complex and changeable nature of the diffuse irradiation requires several assumptions in order to estimate its value from a particular portion of the hemisphere within the field of view of an inclined surface (Hay & McKay, 1985).

In order to find a meticulous model of data transposition, several authors have evaluated different models, both isotropic, that considers that the diffuse irradiation is equally spread in the same intensity and in all directions of the celestial dome, and anisotropic models which considers the anisotropic nature of the diffuse radiation (Gueymard, 1987), (Ineichen, 2011), (Reindl *et al.*,1990), (David *et al.*, Boland, 2013). The validation of the models was performed by comparing the data estimated with the data measured in loco by means of statistical analyses involving average error indicators. The Hay and Perez models in general present positive results in the representation of the diffuse radiation in the tilted plane.

3.1 Hay's Model

In 1979, Hay (Hay, 1979) proposed a model to estimate the diffuse component of the irradiance on tilted surfaces from irradiance data on the horizontal plane. In his model, the transposition factor presents a contribution of two different elements. One of them considers the isotropic background and the other a circumsolar component that varies inversely proportional to the zenith angle (θ_z) and the incident angle of the direct irradiation (θ_i). Hay has determined an index that is applied to each component considered in the model. This index is calculated by BHI/I_0, where BHI is the direct irradiance incident on the horizontal plane and I_0 is the extraterrestrial irradiance. This factor aims to translate the interference of the atmosphere into the diffuse irradiance so that for days of clear skies this interference tends to be smaller and the index tends to get closer to 1. Thus, the circunsolar component receives a higher weight, as can be observed in the equation (4)

$$DTI = DHI\left[\left(\frac{BHI}{I_0}\right)\frac{\cos\theta_i}{\cos\theta_z} + \left(1 - \frac{BHI}{I_0}\right)\frac{1+\cos\beta}{2}\right] \quad (4)$$

Where DHI is the diffuse irradiation on the horizontal plane, DTI is the diffuse irradiation on tilted surfaces and β is the tilt angle of the photovoltaic panels related to the ground.

3.2 Perez's Model

Although the Hay model is considered robust and used by many simulation software for photovoltaic systems, it does not consider a component that contributes reasonably to the amount of diffuse radiation on an inclined plane, called the horizon. Therefore, the influence of this factor was included in Perez's model. In addition, the models developed until then neglected some possible cases, such as the intensification of the circumsolar component in atmospheres that have high diffuse radiation fraction allied to an intense brightness (Baptista, 2016).

In order to overcome the limitations of the available methods at the time, Perez proposed a model applicable to a wider spectrum of celestial configurations, where the circumsolar and horizon components that overlap the isotropic dome are weighted by two coefficients F_1 and F_2, respectivelly. These coefficients vary according to the clarity and brightness index as well as empirical coefficients (Perez et al., 1986).

In 1987, the author proposed a new and simplified version of his model, where the coefficients F_1 and F_2 were redefined. The component F_2 can then assume negative values, which represents the exchange of brightness of the horizon for a zenithal brightness. The horizon brightness is common in days of clear sky and the zenith component is verified on cloudy days (Perez et al., 1987).

The model underwent through new changes in 1990, in order to make the clarity index independent of the zenith angle (Perez *et al.*, 1990). The result of the final version of the model expresses the diffuse irradiation on the tilted surface through the equation (5)

$$DTI = DHI\left[(1-F_1)\left(\frac{1+\cos\beta}{2}\right) + F_1\frac{a}{b} + F_2\sin\beta\right] \quad (5)$$

Where:
$a = max(0, \cos\theta_i)$
$b = max(\cos 85°, \cos\theta_z)$
F_1 is the circumsolar anisotropy coefficient.
F_2 is the horizon/zenith anisotropy coefficient.

4 Calculation of the errors of the models

The transposition models were simulated using the software PVSyst® and the generated energy results were validated by means of the systematic mean error (MBE), root mean square error (RMSE) and absolute error (MAE). These same figures of statistical merit were used for the comparative analysis of the meteorological databases.

The systematic mean error (MBE) indicates the average error of a long series of data. Thus, it acts to dilute a large difference between measured and simulated data. A negative MBE result indicates that the simulated values were underestimated, while a positive result indicates that the values were overestimated. A disadvantage of this method is that underestimated values can be offset by overstated values, thus making some major errors.

MBE is calculated by the sum of the deviations divided by the number of elements of the data set considered, in this case, 15 days, as shown by the equation (6).

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(E_{t,i} - E_{pv,I}) \quad (6)$$

Where E_t is the value obtained through simulation, when related to the energy generated, or global irradiance on the horizontal plane from databases, when related to the analyses of the meteorological data. The E_{pv} represents the energy generated and measured at the photovoltaic solar plant or the global irradiance on the horizontal plane measured at the photovoltaic solar plant by a solarimetric station, and N is the number of data elements in the analysis.

The RMSE is preferable when large variation between the data under analysis are undesirable, that means, even if the sequence has data with significantly low difference between the simulated and measured values, a data that is estimated far from the measured data will have a noticeable impact on the value of RMSE. This method has a sensitivity to work with the analysis of the behaviours of small plots of data, but it impairs the overall performance of a data series that presents the majority of the data close to the real ones and some few data discrepancies. Moreover, another disadvantage of the method is the impossibility of identifying whether the errors are overestimated or underestimated due to the absence of the negative signal. It can be calculated by the equation (7).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(E_{t,i} - E_{PV,I})^2} \qquad (7)$$

In order to evaluate the statistical deviation, the absolute error MAE was also analysed. This error considers the sum of the absolute deviation values by the number of samples as shown in the equation (8).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|E_{t,i} - E_{pv,i}\right| \qquad (8)$$

5 Photovoltaic Power Plant

The photovoltaic system under analysis is located at the Tanquinho Solar Photovoltaic Plant. The system considered is composed by 72 DA142-C1 modules in total from Dupont Apollo. Those modules are organized into 24 strings with 3 modules each. Totalizing 10.22kWp of installed photovoltaic power.

The strings are connected to a Fronious International inverter, model IG Plus 120V-3, operating voltage 230-500V and nominal power 10 kW. A Mavowatt 30 power meter, from Dranetz, is connected to the inverter's output. This equipment measured the energy generated by the strings in 30 seconds interval, from the 7th of June until the 21st of June 2016.

6 Results

Figure 1 shows that although the monthly irradiation data of the bases under analysis present values considerably different from the measured values, there is a trend followed throughout the months.

As expected, there is a disparity between the data from the different sources of information analysed. Table 1 shows that the standard deviations are considerably higher for the months of February and March, reaching 24.38% and 21.19%, respectively. These data disparities are consequences of several factors, such as the absence of data in historical series and measurement errors combined with the fact that the data were collected in different years. For Lorenzo (Lorenzo, 2003), these disparities are not linked initially to the accuracy of measuring equipment. The main factor that affects the accuracy of the data is of its stochastic origin and consequence of the random climatic variations.

It is also noticeable that these databases present data quite close to each other, even using different systems and data acquisition periods. Table 2 shows the standard deviation between Meteonorm 7.1 and NASA SSE data.

Simulations were performed with each meteorological database and meteorological data measured at the site using the software PVSyst®. For each meteorological database, the models of Perez and Hay were applied in the PVSyst® simulation in order to compare the impact of each one on the generation of energy in a photovoltaic system.

It is important to note that for this analysis, monthly averages of global irradiance were used in the horizontal plane. These averages insert an uncertainty of 4-6% when compared to the hourly means. The values of RMSE and MAE shown in Table 3 are obtained through equations (6), (7) and (8). The RMSE and MAE presented better results for simulations using the Perez's model in all databases for the considered location when compared to the Hay's model.

The results found using the Hay's model, despite presenting larger errors than the Perez's model, do not present significant statistical differences as can be seen in the Table 3.

For the study site, the Meteonorm database presented the worst results, reaching RMSE in the order of 13% for both transposition methods and MAE in the order of 10%. Such disparity can be seen from Figure 2 and Figure 3. Figure 3 shows the energy generated through simulations using the NASA SSE database and applying the Perez's and Hay's models. It is observed that the values obtained by simulation are closer to the real values than the results shown in the Figure 2, when the Meteonorm database was used.

Among the meteorological databases, simulations with real data presented considerably better results for both transposition models, as expected, although the period of data collection of irradiation did not coincide with the period of data collection of energy generated. Figure 4 shows the values of energy generated per day applying the different transposition models and using the meteorological data collected at the site. It can be seen that the simulated results are very close to the real values measured.

Table 1: Mean global irradiance values on the horizontal plane (W/m²) and standard deviation.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Set	Oct	Nov	Dec
Meteonorm	234.8	236.8	214.9	194.4	166.8	159.2	168	195.4	211.9	225.9	240.1	246.5
NASA/SSE	231.2	223.8	215.1	202.5	167.5	161.7	173.4	202.2	207.5	236.3	250.4	238.3
Measured data	257.5	271	251.7	192.1	171.8	154.6	150.3	178.5	208.3	205.6	250.8	219.9
Standard deviation	14.26	24.38	21.19	5.46	2.71	3.60	12.08	12.20	2.34	15.61	6.07	13.62

Table 2: Standard deviation between the Meteonorm 7.1 and NASA SSE databases.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Standard deviation	2.55	9.19	0.14	5.73	0.49	1.77	3.82	4.81	3.11	7.35	7.28	5.80

Table 3: RMSE and MAE of the energy generated obtained using Perez's model, Hay's model and using meteorological data from NASA SSE, Meteonorm and Measured Meteorological data on the site.

	NASA-SSE		Meteonorm 7.1		Measured Data	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Perez's Model	10.56	7.45	13.41	9.97	2.82	2.49
Hay's Model	10.59	7.52	13.54	10.11	3.27	2.87

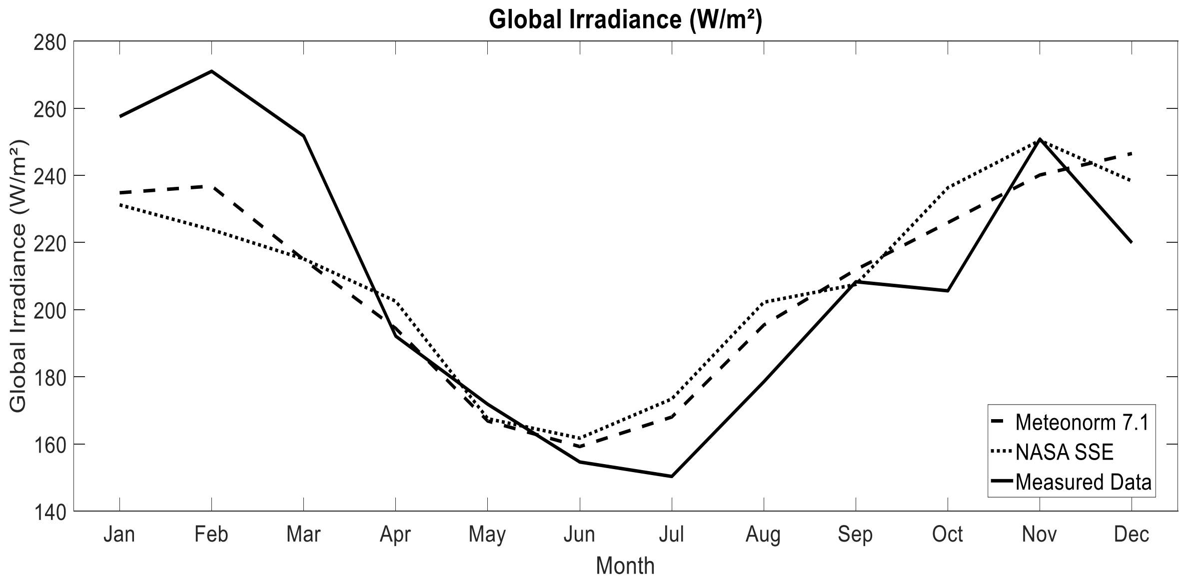

Figure 1: Mean global irradiation on the horizontal plane (W/m²)

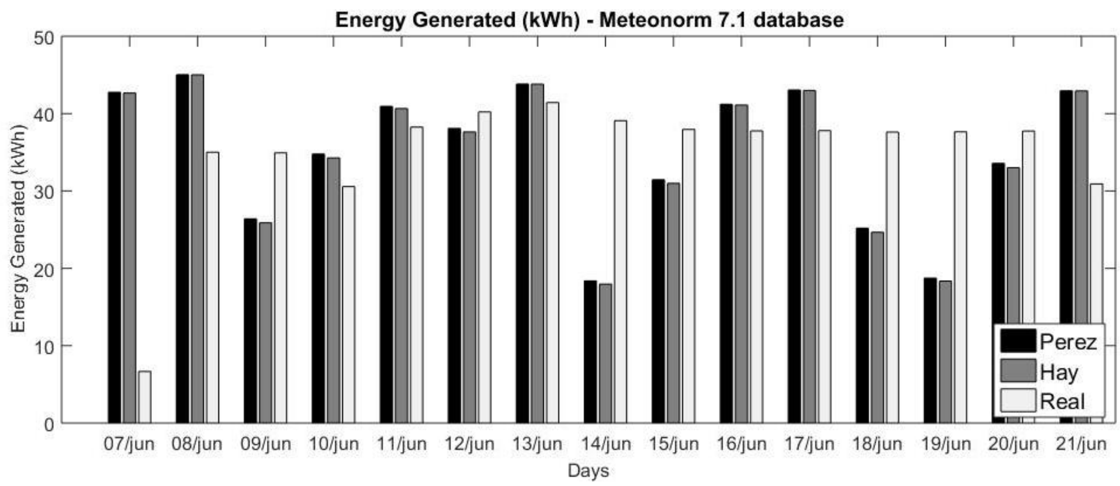

Figure 2: Energy generated (kWh) applying Perez model, Hay model and real data using Meteonorm 7.1 meteorological database

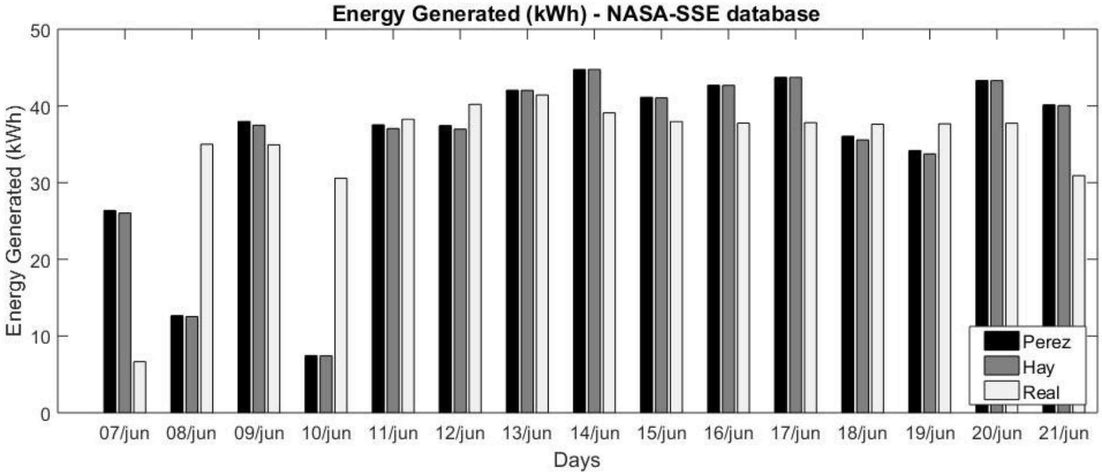

Figure 3: Energy generated (kWh) applying Perez model, Hay model and real data using NASA SSE meteorological database

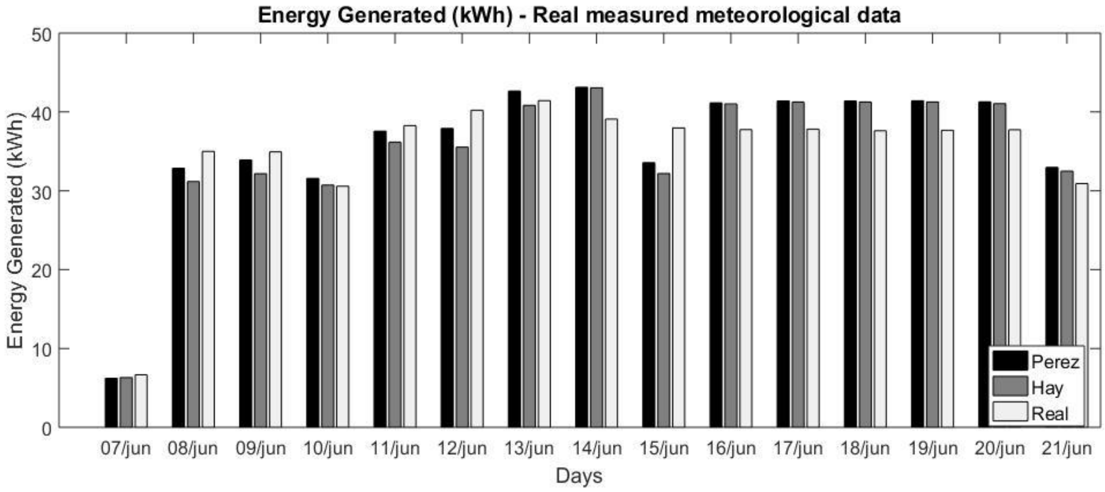

Figure 4: Energy generated (kWh) applying the model of Perez and Hay model using real meteorological data measured on the site.

7 Conclusion

Initially, the NASA SSE meteorological database showed a closer proximity to the actual data measured by the local solarimetric station, when compared to Meteonorm 7.1. However, the NASA SSE database provides updated data until 2005 and since then the land has been experiencing rapid increases in temperature and climate change, which could suggest lack of precision. In addition, the data available on this database is obtained through interpolation of satellite images and monthly averages, which could generate an increase of uncertainties. A more in-depth study related to the accuracy of this data is required. The analysis could include a larger amount of measured data in the locality, to prove the proximity of the irradiation values of the database.

Several authors have performed evaluations of different transposition methods, comparing the results with data collected in the field. However, the great majority of the studies were carried out considering only the Northern hemisphere and only the clear sky condition.

All transposition models have faults. In the present study, it can be seen that the Hay's model underestimates the diffuse irradiation, since it does not consider the horizon in the transposition calculations. The results obtained through the Perez's method were more accurate, which is expected, as it considers more variables in its calculations.

It is important to emphasize that although the Perez's and Hay's models are widely used due to the minor errors compared to other models, the accuracy of diffuse irradiation estimation for each model will vary according to the project site. The choice of the model depends on the characteristics and values of the solar radiation at the desired location, as well as the amount of solar radiation reflected in the soil (albedo). Due to the limited amount of analysis considering the Southern hemisphere, a more in-depth study is needed to evaluate the models that best suit the Brazilian conditions.

Acknowledgement

This work was supported by CNPq, CAPES, FAPESP (2016/08645-9), CPFL (ANEEL-PA3032-Campus Sustentável) and BYD Energy Brazil (MCTIC-PADIS).

Bibliographic References

Baptista, J. P. M. (2016). Distribuição Angular da Radiação Solar : Aplicação do Modelo de Perez.

David, Mathieu; Lauret, Philippe; Boland, J. (2013). Evaluating tilted plane models for solar radiation using.pdf. *Renewable Energy*, *51*, 124–131.

Gomes, E. N., & Escobedo, J. F. (2007). Estimativa Da Radiação Direta Na Incidência Horária , Diária E Mensal Em Função Do Índice De Claridade (Kt). In *I CBENS - I Congresso Brasileiro de Energia Solar*.

Gueymard, C. A. (1987). An Anisotropic Solar Irradiance Model for Tilted Surfaces and Its Comparison with Selected Engineering Algorithms. *Solar Energy*, *38*(5), 367–386.

Gueymard, C. A. (2009). Direct and Indirect Uncertainties in the Prediction of Tilted Irradiance for Solar Engineering Applications. *Solar Energy*, 83, 432-444

Gueymard, C. A., & Myers, D. R. (2008). Solar Radiation Measurement: Progress in Radiometry for Immproved Modeling.

Hay, J. E. (1979). CALCULATION OF MONTHLY MEAN SOLAR RADIATION.pdf. *Solar Energy*, *23*, 301–307.

Hay, J. E., & McKay, D. C. (1985). Estimating Solar Irradiance on Inclined Surfaces : A Review and Assessment of Methodologies Estimating Solar Irradiance on Inclined Surfaces : A Review and Assessment of Methodologies. *International Journal of Solar Energy*, *3*(May 2013), 203–240.

Ineichen, P. (2011). Global Irradiance on Tilted and Oriented Planes: Model Validations. *University of Geneva*, (February).

Kumar, P., & Waila, V. C. (2017). Parameters Considered In Efficient Designing Of Solar PV System, (3), 923–926.

Oliveira, L. G. M. (2017). *Avaliação De Fatores Que Influenciam Na Estimativa Da Geração E Operação De Sistemas Fotovoltaicos Conectados À Rede Elétrica*.

Perez, R., Ineichen, P., Seals, R., Michalsky, J., & Stewart, R. (1990). Modeling daylight availability and irradiance components from direct and global irradiance. *Solar Energy*, *44*(5), 271–289. https://doi.org/10.1016/0038-092X(90)90055-H

Perez, R., Seals, R., Ineichen, P., Stewart, R., & Menicucci, D. (1987). A New Simplified Version of the Perez Diffuse Irradiance Model for tilted surfaces. *Solar Energy*, *39*(3), 221–231. https://doi.org/10.1016/S0038-092X(87)80031-2

Perez, R., Stewart, R., Arbogast, C., Seals, R., & Scott, J. (1986). An Anisotropic Hourly Diffuse Radiation Model for Sloping Surfaces: Description, Performance Validation, site dependency evaluation. *Solar Energy*, *36*(6), 481–497.

Reindl, D. T., Beckman, W. A., & Duffie, J. A. (1990). Evaluation of Hourly Tilted Surface Radiation Models. *Solar Energy*, *45*(1), 9–17.

Riveros-Rosas, D., Arancibia-Bulnes, C. A., Bonifaz, R., Medina, M. A., Peón, R., & Valdes, M. (2015). Analysis of a solarimetric database for Mexico and comparison with the CSR model. *Renewable Energy*, *75*, 21–29.

Scolar, J., Martins, D., & Escobedo, J. F. (2003). Estimativa da Irradiação Total Sobre uma Superfície Inclinada a Partir da Irradiação Global na Horizontal. *Revista Brasileira de Geofisica*, *21*(3), 249–258.

Thevenard, D., & Pelland, S. (2013). Estimating the Uncertainty in Long-Term Photovoltaic Yield Predictions. *Solar Energy*, *91*, 432–445. https://doi.org/10.1016/j.solener.2011.05.006