Estimation of monthly Global Horizontal Irradiation pan India using spatial interpolation and comparing its deviations from standard dataset

Vijay Bhaskar
Senior associate, New Products, Amplus energy solutions private limited, India

Abstract

In the current scenario of increasing demand for solar photovoltaic (PV) systems, the need to predict their feasibility and performance is more than ever. Irradiance of a geographical location almost exclusively determines the generation possible via solar. Hence, accurate irradiance data is required to assess the value of solar PV systems. Emphasizing such need, this paper presents a method of estimating global horizontal irradiance (GHI) using the two dimensional (2-D) spatial interpolation technique. The proposed model is geo-agnostic and can estimate irradiance depending on the geographical range of the input data. This paper also compares the model predictions with a standard irradiation dataset in the industry. This comparison helps in getting insights regarding the spatio-temporal trends in recent times.

Keywords

Global horizontal irradiance, spatial interpolation, kriging, estimation, Data-driven

Introduction

PV systems are a predominant means of harnessing solar energy. They are cheaper than most means of renewable energy along with low periodic maintenance. They are also highly durable and easily scalable. Hence, there is a rapid demand for them worldwide. Since any decrease in PV generation can result in considerable financial penalties, it is important to predict the possible generation. Solar irradiance is a key factor used in yield assessment. Generation has an almost linear dependence on input irradiance. Thus, the estimation of irradiance is an important exercise for PV asset management. Solar irradiance is measured in various ways each with a different notion of incident radiation. S.M. Maleki et al [1] familiarizes with the requisite concepts. It dives deep into the formulation of such concepts and elucidates the mechanics of changing irradiance with time. D.Young et al [2] and D.Palmer et al [3] discuss the current methodologies used by the industrial as well as research communities not only to define the scope of the engineering challenges but also solving them. These provide a fairly lucid picture of the system and establish the context to understand the motivation behind this study and helps better appreciate the research work presented.
Spatial interpolation is a well known and effective technique in modelling distributions and parameters which are dependent on geography. Ryu, JS., et al [4] emphasizes the challenges of geostatistics and explains how interpolation techniques such as inverse distance weighted functions, kriging, etc., can be leveraged for accurate estimation. B. Bacchi et al [5] even presents the case of explaining the complexity of numerical weather prediction models. It circumvents the issue of weather prediction by employing spatial correlation techniques to explain rainfall trends. D. Perez-Astudillo et al [6] is a study which is similar to the one in this paper. It attempts to map GHI trends all over Qatar only using the weather station data. However, this paper goes a step further and performs interpolation pan India producing gap-filled estimates at every grid coordinates spaced 10 Kilometer (Km) apart i.e., 0.1° latitude/longitude difference.

This paper is organized as follows. Section 2 explains the methodology behind the estimation model. Section 3 discusses the obtained results and their accuracy. It also compares the deviations between the standard dataset and the model predictions. Section 4 summarizes the content and gives concluding remarks.

**Methodology**

Most of the PV projects are installed at an orientation(s) as per either design optimization or client-side requirements or both. Thus, irradiation sensor data are the measurements of Global Tilted Irradiance (GTI) in respective sites. In terms of plant monitoring, GTI makes more practical sense as it captures the incident irradiance on the tilted modules. But discerning inherent trends from such GTI data alone is complex as it majorly depends on the tilt and azimuth. GHI trends are also known to be more gradual and continually smoother than those of GTI. Hence, the idea is to transform GTI to GHI which will eliminate the model’s dependence on sensor orientation. This reduces the model complexity and simultaneously increases the accuracy of GHI estimates. Thus, we have used a transform domain approach to achieve the aim of predicting irradiance more accurately as summed up below. GTI data is firstly transformed into GHI data using the corresponding transposition factor (TF). The model is then built using the geographical parameters and input GHI values. The output essentially identifies spatial trends in GHI and uses them to produce GHI estimates pan India. This resulting GHI estimates can be transformed back to respective GTI using the inverse of TF.

**Training dataset**

As part of asset management, solar PV plants have irradiation sensors whose measures are sent to the servers on a real-time basis. This is incorporated into the analytics portal which is used for operations and monitoring. Thus, the data is organized for each plant with its geographical parameters (latitude and longitude) along with GTI measured by on-ground sensors. T-factors corresponding to each sensor orientation are also known which are used to obtain GHI values.
The above snippet is the sample training data showing ground sensor readings from the mentioned coordinates. The number of rows in the training set (number of sensors) varies depending on the month and year in question. This is because the source of data for this study is from the commissioned PV systems across India. The training data for August 2019 has more than 130 sensors across India.

Formulate a training dataset with the parameters \{\text{Latitude, Longitude, Monthly GHI, Month}\}.

- Use the 2-D spatial interpolation modules of Pykrige python library to perform gap-filling. Refer to the official Pykrige documentation [7] to understand the various modules of the library.

- Define a reasonable range and resolution of the geographical grid to perform interpolation so that inherent trends can be conspicuous.

- The optimization problem is defined as minimizing total absolute error of estimation (J) with respect to the kriging variogram parameters.

- Output dataset is of the same format as that of input one except that the former has GHI estimates for all co-ordinates within the grid.

- The feedback from the error minimization block is performed until we obtain the minima of J.
Flowchart of data flow

The results shown henceforth for August 2019 have been obtained using **2-D Ordinary kriging** in the geographical range corresponding to India (**+5.6 to +37.4** latitudes and **+67.8 to +97.6** longitudes) with a **0.1 grid resolution (10Km X 10Km)**.

The nomenclature below is defined to aid visual understanding in the results section and maintain uniformity.

- **Lat (+degrees)**: latitudes defining the geographical grid (rounded to 1 decimal place)
- **Lon (+degrees)**: longitudes defining the geographical grid (rounded to 1 decimal place)
- **Est.GHI (KWh/m²)**: estimated GHI for [Lat, Lon] coordinates from the model
- **Ref.GHI (KWh/m²)**: reference GHI (standard dataset) at respective coordinates
- **Est-Ref dev (%)**: percentage deviation of estimated GHI form reference GHI
- **Latitude (+degrees)**: latitudes of the training dataset
- **Longitude (+degrees)**: longitudes of the training dataset
- **Act.GHI (KWh/m²)**: measured GHI at respective coordinates
- **Est-Act dev (%)**: percentage deviation of estimated GHI form measured GHI
Results

Sample output - August 2019

| Lat  | Lon | Est.GHI | Month |
|------|-----|---------|-------|
| 23.6 | 68  | 128.96  | 8/2019|
| 23.5 | 68.1| 128.71  | 8/2019|
| 23.6 | 68.1| 128.94  | 8/2019|
| 23.7 | 68.1| 129.07  | 8/2019|
| 23.8 | 68.1| 129.25  | 8/2019|
| 23.9 | 68.1| 129.48  | 8/2019|
| 23.5 | 68.2| 128.64  | 8/2019|
| 23.6 | 68.2| 128.87  | 8/2019|
| 23.7 | 68.2| 129.07  | 8/2019|
| 23.8 | 68.2| 129.19  | 8/2019|
| 23.9 | 68.2| 129.35  | 8/2019|
| 24.  | 68.2| 129.51  | 8/2019|

Irradiation heat map - August 2019

Average daily GHI (KWh/m²)

Sample Training error - August 2019

| Latitude  | Longitude | Act.GHI | Lat | Lon | Est.GHI | Est-Act dev (%) |
|-----------|-----------|---------|-----|-----|---------|----------------|
| 31.59     | 74.962    | 149.48  | 31.6| 75  | 150.71  | 0.82           |
| 27.218    | 78.092    | 134.12  | 27.2| 78.1| 133.88  | -0.18          |
| 25.135    | 75.856    | 114.36  | 25.1| 75.9| 113.94  | -0.37          |
| 17.332    | 78.724    | 142.22  | 17.3| 78.7| 142.2   | -0.01          |
| 31.634    | 74.835    | 143.07  | 31.6| 74.8| 143.58  | 0.36           |
| 13.02     | 80.182    | 155.83  | 13  | 80.2| 150.11  | -3.67          |
| 12.874    | 80.138    | 151.85  | 12.9| 80.1| 153.75  | 1.25           |
| 12.818    | 77.53     | 123.99  | 12.8| 77.5| 122.85  | -0.92          |
| 11.046    | 76.899    | 130.4   | 11  | 76.9| 131.43  | 0.79           |
| 12.918    | 77.695    | 124.26  | 12.9| 77.7| 126.62  | 1.9            |
| 28.499    | 77.397    | 136.23  | 28.5| 77.4| 135.82  | -0.3           |
| 28.155    | 76.878    | 148.54  | 28.2| 76.9| 150.73  | 1.48           |

Error statistics:

Mean training error (μτ) = -0.032%
Standard deviation of error (στ) = 1.926%
Comparison between model estimates and standard dataset

Irradiation deviation (%) - August 2019

| lat | lon | Est.GHI | Ref.GHI | Est-Ref dev(%) |
|-----|-----|---------|---------|---------------|
| 31.6| 75  | 150.71  | 158.01  | -4.62         |
| 25.1| 75.9| 113.94  | 140.28  | -18.78        |
| 29  | 77.7| 141.71  | 147.67  | -4.03         |
| 16.3| 80.5| 143.24  | 141.49  | 1.23          |
| 30.2| 75.1| 170.21  | 164.55  | 3.44          |
| 31  | 75.8| 146.31  | 157.22  | -6.94         |
| 28.4| 77  | 149     | 148.99  | 0.01          |
| 28.5| 77.1| 144.81  | 148.49  | -2.48         |
| 20.3| 85.8| 111.87  | 137.73  | -18.77        |
| 19.1| 73.1| 96.14   | 127.62  | -24.67        |
| 28.9| 79.4| 154     | 152.15  | 1.22          |
| 26.9| 76.1| 136.73  | 147.03  | -7.01         |
| 18.5| 73.9| 134.68  | 133.21  | 1.11          |

Mean (Est-Ref dev) = μd = -8.76%
Standard deviation (Est-Ref dev) = σd = 9.67%

Discussion

- The black dots on the heat maps denote the locations of the irradiance sensors. Their considerable spread across India can be observed. This helps in learning majority spatial GHI trends if not all.
- The irradiation heat map for August 2019 is in line with actual weather data reported by Indian weather agencies. For instance, the trend of excessive rainfall reported in central India this august closely correlates that which is shown as considerably lower irradiation in the region.
- The model optimizes its variogram parameters by achieving the minima of total absolute error (%) which is evident in the mentioned statistics of training error.
- The standard dataset has been obtained from a reputed industrial vendor whose data points are a combination of long-term averaged values and satellite imagery. This is used as a reference to understand temporal GHI trends.
- As shown above, reference GHI considerably deviates from ground reality. In the sample results for August 2019, there is negative mean deviation and considerable standard deviation in the error distribution. This verifies the increasingly severe trend of monsoon pan India this August and the consequent irradiation shortfall.
- The major bottleneck in improving estimation accuracy is the quality and quantity of training data. Having a greater geographical spread of sensors and the ability to capture accurate measurements, the model’s capability to predict GHI trends can be better appreciated.
Conclusion

As part of our study, the increasing predominance of solar PV as a renewable source of energy is discussed. This has focused the attention on the need to have quality irradiation data. The above research has been as an endeavour to use a data-driven approach to solve the issue at hand. Hopefully, this work can showcase the power of using data-intensive techniques such as the one above to solve the many challenges in the energy industry especially those in solar. The model is built using irradiation sensor data pan India and used an effective spatial interpolation technique, kriging, to produce the gap-filled estimates. The statistical measures of estimate error are also mentioned which show impressive accuracy. Heat maps for respective months have also been produced for better visualization of GHI trends. An independent standard dataset is also compared with the estimates to better understand the temporal GHI trends with respect to long-term averaged values. The assessment of this work’s potential is for the industrial community to ascertain as this can have various use cases of immense business value.

Acknowledgement

This opportunity is to thank my managers Ankur Gupta, Vinayak Singh and Kranthi Tej for their overall guidance and nuanced domain expertise. Naresh Kumar and Oves Khan are part of our team’s full-stack developers and their help in proper and timely data acquisition is very much appreciated. Pooja has been a helping hand in enhancing my understanding at every stage of the project and for that, I thank her. The standard dataset is used by the design team for benchmarking PV performance and has been obtained from a reputed data services vendor.

References

1. S.M. Maleki, H.Hizam, C.Gomes. Estimation of Hourly, Daily and Monthly Global Solar Radiation on Inclined Surfaces: Models Re-Visited, Energies, Volume 10, issue 1, https://doi.org/10.3390/en10010134

2. D.Young, A.Nobre, R. Baker, T. Reindl. Large-area irradiance mapping, PV-Tech Power Volume 1

3. D.Palmer, I.Cole, T.Betts, R.Gottscalg. Interpolating and Estimating Horizontal Diffuse Solar Irradiation to Provide UK-Wide Coverage: Selection of the Best Performing Models, Energies, Volume 10, issue 2, https://doi.org/10.3390/en10020181

4. Ryu, JS., Kim, MS., Cha, KJ. et al. Kriging interpolation methods in geostatistics and DACE model, KSME International Journal (2002) 16: 619. https://doi.org/10.1007/BF03184811

5. B. Bacchi, N.T. Kottegoda. Identification and calibration of spatial correlation patterns of rainfall, Journal of hydrology, Volume 165, February 1995, https://doi.org/10.1016/0022-1694(94)02590-8

6. D. Perez-Astudillo, D. Bachour. Variability of measured Global Horizontal Irradiation throughout Qatar, Solar Energy, Volume 119, September 2015, https://doi.org/10.1016/j.solener.2015.06.045

7. Pykrige documentation, release 1.4.1, Jan 15, 2019, https://pykrige.readthedocs.io/en/latest/overview.html