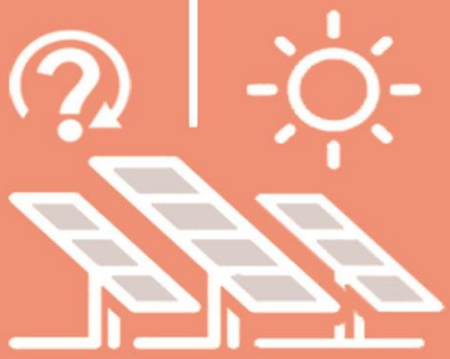
CSTEP-WS-2018-04

Performance Comparison of Solar Photovoltaic Models for Plant Siting in India

Kelly Klima
Harshid Sridhar
Vaishalee Dash
Mridula Dixit Bharadwaj
Parveen Kumar
Jay Apt



SERIIUS
CARNEGIE MELLON UNIVERSITY PITTSBURGH PENNSYLVANIA 1900
CSTEP® CENTER FOR STUDY OF SCIENCE, TECHNOLOGY & POLICY

Performance Comparison of Solar Photovoltaic Models for Plant Siting in India

Dr Kelly Klima[1]

Harshid Sridhar[2]

Vaishalee Dash[2]

Dr Mridula Dixit Bharadwaj[2]

Dr Parveen Kumar[2]

Prof. Jay Apt[1, 3]

December, 2018

[1]*Department of Engineering and Public Policy, Carnegie Mellon University*
[2] *Center for Study of Science, Technology and Policy*
[3] *Tepper School of Business, Carnegie Mellon University*

Executive Summary

India plans to add 100 GW of solar electric power generation by 2022 to an existing system (with an installed capacity of close to 330 GW, as on January 2018, from all sources). Selection of sites for such a large infrastructure investment is definitely an important decision. Here, we examine the performance of four models of annual solar photovoltaic production that are appropriate for site selection against actual generation data from utility-scale plants in the Indian state of Gujarat. We find that a simple Bird clear sky index predicts the annual PV plant production with an error of 14±5%, while the satellite data (at 10 km × 10 km resolution) in National Renewable Energy Laboratory's National Solar Radiation Database (NSRDB) has a prediction error of 9±6%. Data from solar monitors at India's Solar Radiation Resource Assessment (SRRA) stations, which are 46–95 km away from the power plants, has a prediction error of 7±3%. However, a power plant model using SRRA weather data to predict output correction for solar cell temperature performs best with an error of 6±6%. The inter-annual variability of the annual mean irradiation for 2000–2014 in Gujarat is ±3.6% for direct normal irradiance (DNI). Thus, we find that for site selection based on annual PV production, the inter-annual variability of irradiance is larger than the differences between models; in the absence of coincident data, satellite data (NSRDB) could be a preferred choice.

Introduction

Many countries throughout the world are investing in large-scale solar photovoltaic (PV) plants. In the United States, Hawaii (with its high cost of oil generation) is already at socket parity (solar generation can produce electricity at a cost equivalent to the retail electricity rate) [1]. Moreover, the U.S. Department of Energy's SunShot Initiative has the goal of reducing the unsubsidised cost of utility-scale solar energy to 6 U.S. cents per kWh by 2020 and 3 U.S. cents/kWh in 2030 [2]. In India, the Ministry of New and Renewable Energy's National Solar Mission aims to achieve 100 GW of solar power capacity by 2022 [3]. To choose sites effectively for such significant solar targets, planners require accurate predictions of annual solar output, aided by representative weather data (such as modelled irradiation or historical weather data).

The United States has several models capable of informing PV siting, such as the Sandia National Laboratory's PV Performance Modelling Collaborative (PVPMC) [4], the National Renewable Energy Laboratory's (NREL) System Advisor Model (SAM) [5] and NREL's PV Watts [6]. India does not yet have such a model attuned to conditions in the country. Rather, research in India focuses on determining the solar potential and creating comprehensive databases of historical weather [7] that includes one-minute resolution data. For example, Kumar and Umanand demonstrate a theoretical model for calculating global insolation on a horizontal surface in India that is within 20% of actual insolation with no information on cloud coverage [8]. Further, Karakoti et al. examine non-linear solar radiation models for predicting the monthly average daily diffuse radiation for 12 locations in India and find that the cubic equation can be used for the prediction [9].

Given these studies, many researchers have started using irradiance and other geographic properties to suggest site locations [10,11,12,13]. Some researchers have also attempted to forecast solar power generation in India. For example, Ashraf and Chandra use artificial neural network based models for forecasting electricity generation of grid-connected solar PV with "reasonable" errors of up to 15% or more [14]. Studies for actual locations with coincident weather data suggest the actual generation can be predicted within 1.4% [15]; however, there are hardly any studies without coincident irradiance data.

Since plant siting is the goal of this study (as opposed to forecasting daily or hourly generation), we examine the representative model performance, integrated over a year, for four locations of PV power plants in the Indian state of Gujarat. We analyse the annual performance of four solar PV generation models used to inform the siting of PV plants against the actual generation data. These models are as follows:

- A simple irradiance model (the Bird clear sky index)
- A model based on gridded hourly solar and meteorological data from the National Solar Radiation Database
- A model using observed weather data from weather stations near the four plants for which we have output data

- A power plant model which uses that weather data to incorporate PV cell performance as a function of observed temperature and observed insolation as well as details such as the number and type of modules

We organise the research note as follows. Section 2 describes the methods, including the datasets (of actual generation data and four models of plant performance) and the metrics used to compare the datasets. Section 3 presents the results, and Section 4 contains the discussion and conclusions.

Methodology

We analyse PV plant generation data and four models for four representative power plants in Gujarat, India, for one year (February to December, 2014). This section describes the datasets, models and comparison metrics in detail.

Observed Generation Data

The observed generation data consists of real-time output measurements from solar photovoltaic plants in Gujarat, India. As described in a previous publication [16], data is available from 50 PV plants (whose capacities range between 5 and 330 MW) and is provided approximately every minute beginning February 17, 2014. Since, the meteorological data (described in Section 2.2.3) ends on December 31, 2014, we limit our time period from February 17, 2014 to December 31, 2014. Out of these 50 plants, we select four that are closest to the weather monitoring stations in the Indian SRRA network. None of these four plants have capacity increase over this time period.

In order to have uniform comparison of generation data across models, we implement two data-cleaning steps. Some of the plants have negative generation at night (typically because of electricity needs of support equipment such as air conditioners and power electronics); generation at these times is adjusted from negative to zero. Moreover, if at some point from 10 a.m. to 5 p.m. local time, generation is found to be unchanging for more than 15 minutes, we assume there is a malfunction in the sensor or telemeter and remove the entire day from our analysis.

The location and other details of the plants are collected from project reports submitted to the Clean Development Mechanism (CDM) under United Nations Framework Convention on Climate Change (UNFCCC) [17,18,19,20] and verified via Google Maps satellite data (see Appendix A, Table A.1).

Existing Models Applied for the Study (Models 1–4)

Here, we describe three previously published meteorological models (in order of increasing complexity) that a decision-maker might use to quickly estimate the actual generation data for site screening. Each meteorological dataset provides information on (a) the direct radiation normal to the beam at the Earth's surface (DNI, W/m^2), (b) the diffuse radiation incident upon a horizontal surface (DHI, W/m^2) and (c) the global radiation incident upon a horizontal surface (GHI, W/m^2). For each dataset, we convert the data into a model of net effective solar radiation incident on the tilted panel (G_T, W/m^2) and then to a representative generation (P_{plant}, MW) as described in Sections 2.3. and 2.4.

For the three models, we assume that the solar panels are in a fixed-tilt configuration, in which each panel is tilted at an angle θ_T from the horizontal (we verified from photographs of the four sites that none are tracking arrays). We also assume that the panels face south. In such a case, it is usual to mount the solar panels at an angle equal to the latitude of the site. Although this is common practice, we note that some utility-scale sites are constructed at a different angle. Indeed, the one plant for which we were able to find the module tilt (Plant B) was tilted at 15°, while its latitude is 23°. For the remaining three plants, information was not available on the module tilt. This is of little consequence, because the difference in the annual output of a Gujarat plant tilted at 23° and another at 15° is only 1.1%, as per the NREL PV Watts model [21].

Model 1: Bird Clear Sky Index

Model 1 is the radiation data derived from the Bird clear sky index. The model assumes a cloudless sky and, thus, a smooth radiation profile. We use Bird clear sky index [22] with inputs of the latitude, longitude and altitude (Table A.1); an assumed foreground albedo of 0.2; the time zone of India (UTC+5.5[1]) and the default settings in Table A.2. Most of the assumed values are model defaults within the typical value range. The barometric pressure value is also a model default value and is a slight adjustment from sea level values to account for slight altitude elevations. We convert this data into a model of net effective solar radiation incident on the tilted panel (G_T, W/m^2) and then to a representative generation (P_{plant}, MW) as described in Sections 2.3. and 2.4.

Model 2: NSRDB Gridded Weather Data

Model 2 is gridded actual weather data from the National Solar Radiation Database (NSRDB v.2.0.0) [23]. The NSRDB comprises solar and meteorological data interpolated for 0.1° latitude by 0.1° intervals (approximately 10 km x 10 km at Gujarat's latitude). Data is provided hourly from 1998 to 2014; we use the subset corresponding to February 17, 2014–December 31, 2014. We use the NSRDB user interface [24] to extract the DNI, DHI and GHI at the nearest grid point corresponding to each plant in Table A.1. We then convert this data into a model of net effective solar radiation incident on the tilted panel (G_T, W/m^2) and then to a representative generation (P_{plant}, MW) as described in Sections 2.3. and 2.4.

[1] UTC refers to Coordinated Universal Time

Model 3: Solar Radiation Resource Assessment (SRRA) Observed Weather Data

Model 3 comprises observed weather data compiled by the National Institute of Wind Energy (NIWE), India and shared by the National Institute of Solar Energy (NISE) [25]. Details of the SRRA stations closest to Plant A (Gandhinagar station), Plant B (Bhogat station), Plant C (Bhogat station) and Plant D (Gandhinagar station) are described in Table A.1. We use the data from February 17, 2014 to December 31, 2014 at 1-minute intervals (with 0.44%–5.30% of data missing, as described in Table A.1). The datasets include 1-minute average values of air temperature (T_{amb}, in °C); wind speed at 6-m height (WS, in m/s) and solar radiation components, namely global horizontal irradiance incident upon a horizontal surface (GHI, in W/m^2) and direct normal irradiance at the Earth's surface (DNI, in W/m^2). We assume the data has been provided corresponding to UTC + 05:30. We then convert this data into a model of net effective solar radiation incident on the tilted panel (G_T, W/m^2) and then to a representative generation (P_{plant}, MW) as described in Sections 2.3. and 2.4.

Model 4: Power Plant Model

Finally, we describe a fourth model that is not only based on metrological data, but also incorporates variations in PV cell performance based on panel temperature. The performance of solar PV cells is sensitive not only to the incident solar radiation, but also to the cell temperature. We calculate the cell temperature of the solar panel and then determine the power generated by the solar PV plant, accounting for the efficiency of the power electronics as well. This model is then used to generate results for each of the four plants given in Table A.1 for 1-minute intervals over January 1, 2014 to December 31, 2014. The equations used in the model are described in Appendix B.3.

Conversion of Models 1–3 into Net Effective Solar Radiation Incident on a Tilted Panel

This section provides basic equations for converting models 1, 2 and 3 into their respective G_T, the net effective solar radiation incident on a tilted panel (W/m^2). We first describe the solar angle calculation. Then, we provide the calculation of the net effective solar radiation incident on a tilted panel.

Solar Angle Calculation

The values for the solar zenith angle (θ_z, the complement of the solar elevation) vary in complexity between models. In this analysis, we calculate the appropriate solar angles at minute-wise resolution using Duffie and Beckman [26] for all datasets.

This algorithm takes as an input the apparent solar time, which directly tracks the motion of the sun and thus calculates the position of the sun every minute at every point in the year. However, the weather data has been provided in either the local time or in the universal time clock (in India, UTC

+ 05:30), which tracks a theoretical "mean" sun with the noon 24 hours apart. Therefore, we use the equation of time to convert the local time to the apparent solar time at the weather station.

Also note that a visual inspection of Figures 1–4 suggests there is a 25–30 minute lag between Dataset 1 and the concurrent Datasets 3 and 4. The time discrepancy (which varies slightly between the plants A–D) is consistently observed for each plant throughout the simulation time period. Since the time discrepancy is consistent, it is not a function of the apparent solar time / equation of time. It might be possible that SRRA reports data in a different time zone than India's (UTC +5:30), but we could not confirm this. Since, we could not determine the reason for a non-uniform time differential, we provide the results with the time discrepancy as given.

We use the Duffie and Beckman solar calculator to calculate the solar angles needed for estimating radiation on a tilted panel for all the models. However, the data in Models 1 and 2 use National Oceanic and Atmospheric Administration's (NOAA's) solar position calculator [27]. Further, we check whether our results are dependent on the type of solar calculator used. We find that the correlation of the Duffie and Beckman solar calculator and the NOAA solar position calculator is greater than or equal to 0.9999 for all plants. We thus note that using a different solar position calculator does not affect the results significantly.

Calculation of Net Effective Solar Radiation Incident on a Tilted Panel

The net effective radiation on a tilted panel is calculated by multiplying the module tilt factors with the individual components of radiation, namely DNI, DHI and GHI. Details of the calculation process are listed in Appendix B.1.

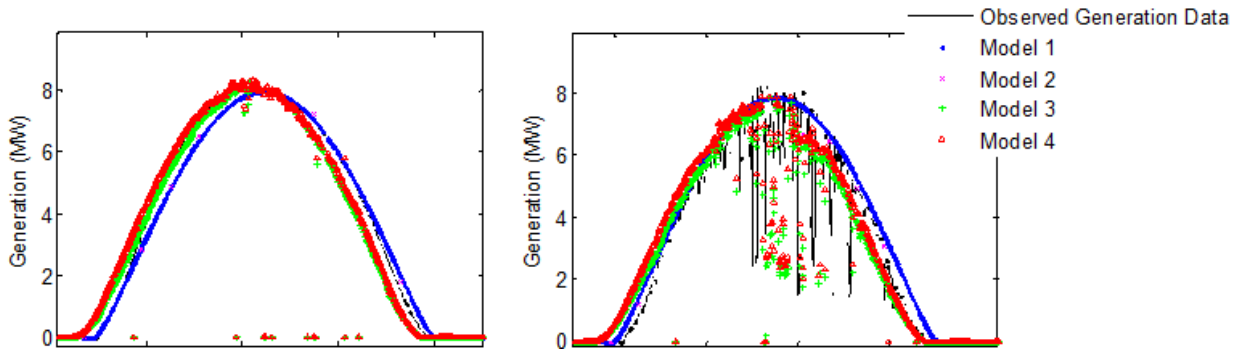


Figure 1. Scaled datasets for Plant A. Left: March 19. Right: Sept 22

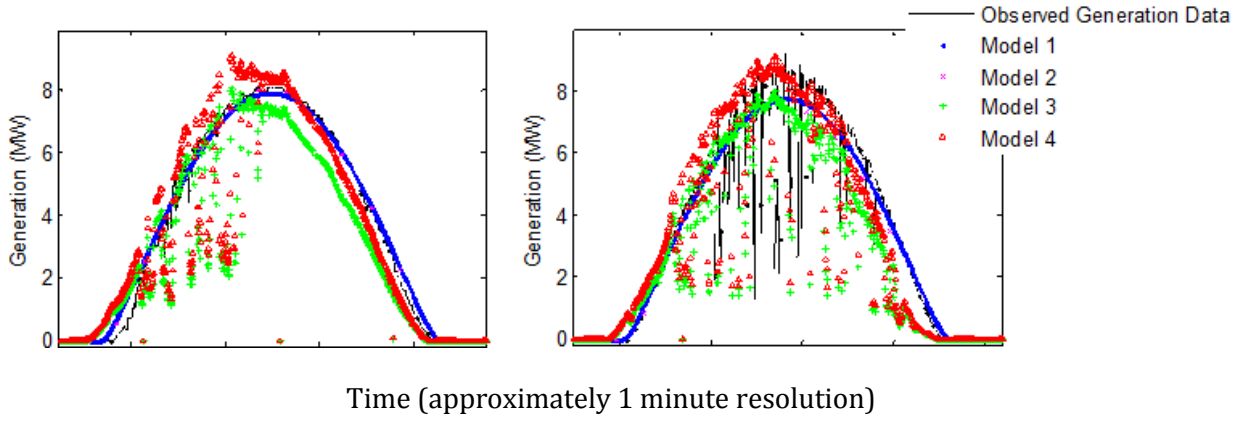


Figure 2. Scaled datasets for Plant B. Left: March 25. Right: Sept 23

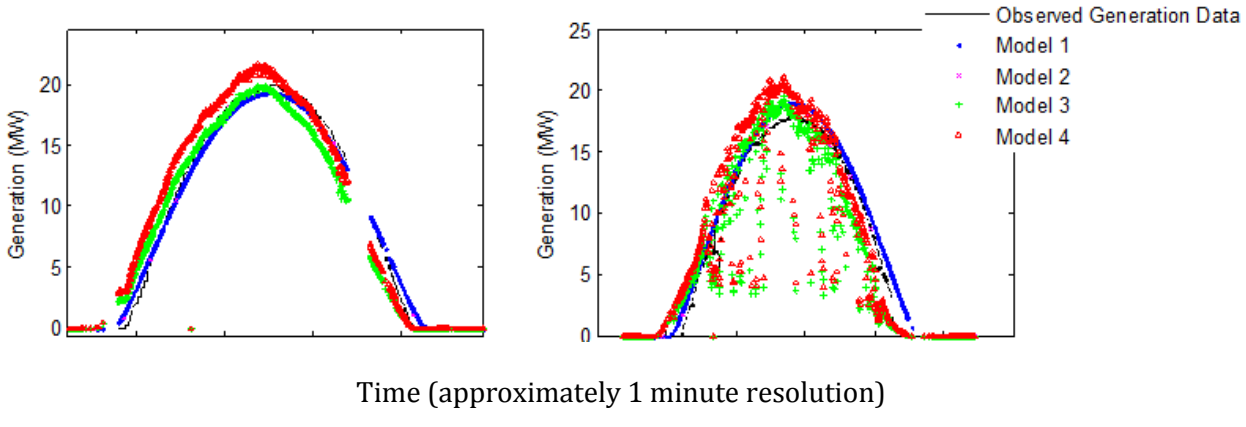


Figure 3. Scaled datasets for Plant C. Left: March 20. Right: Sept 23

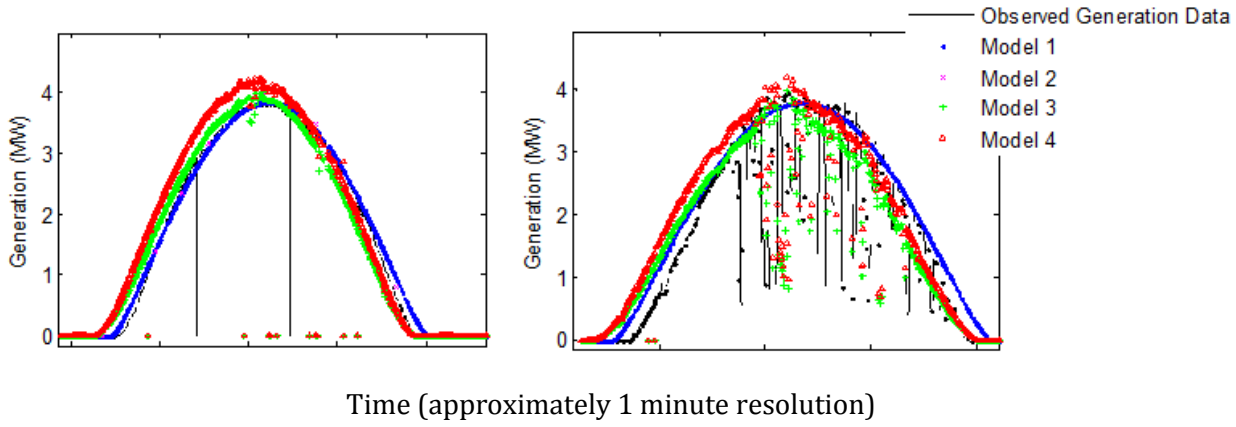


Figure 4. Scaled datasets for Plant D. Left: March 19. Right: Sept 21

Conversion of Models 1–3 into Representative Generation

Next, we need to convert the predictions of Models 1–3 from units of W/m^2 to MW. We observe that none of the four PV plants produce their nameplate capacity at any time during the year, with the exception of very brief periods when cloud focusing occurred. Unlike some utility-scale plants in the United States, the output of the Gujarat plants we examine is not limited by the rating of the power electronics. That is, on a cloud-free day, the output follows a smooth curve without being capped near the maximum output. The equations for converting the radiation data into representative generation in MW are listed in Appendix B.2.

While any day could be chosen for this scaling, recall that we are interested in simulating the annual performance at a site. Furthermore, because the power plant model does not account for the ageing of infrastructure or lack of maintenance (e.g., dust on the cells), the scaling should consider two days at different times of the year. Hence, we examine the two days in the middle of the year: the solar equinoxes (the time or date at which the sun crosses the celestial equator, when the day and the night are of an equal length) of March 20, 2014 and September 22, 2014. Using data from a relatively cloudless day on or near the two equinoxes for each power plant, we inspect and calculate an average η_{plant}. Figures 1–4 show the resulting days (8 total), including Model 4 (as described in Section 2.2.6). Table A.4 shows the resulting scalar values. Note that there is no temporal shifting of the datasets. Such a scaling approach may not be the best to compare models; however, it provides a crude boundary condition assessment.

Model Performance Metrics

We seek to determine the level of model sophistication required to screen potential PV sites based on annual generation. An ideal case would be to have temporally and spatially coincident solar resource data for better PV plant siting; however, such data is not always available and sometimes does not exist. Model assessment by correlation to the actual plant output would be a natural choice, and the closest available stations in the NSRDB gridded weather data are less than 10 km from the plant. However, this data is at hourly resolution only and data of a finer resolution is required for the power plant model (Model 4). These data inputs (the observed weather data) are 36–95 km from the nearest plant. Thus, there will be a significant lag in cloud signals between the datasets, which necessitates the knowledge of the cloud speed and direction to use even lagged correlation coefficients at a high time resolution.

Instead, we use metrics that examine performance assuming meteorological data integrated over the entire year. Here, we present two metrics that compare performance of the four models against the actual generation data.

Temporal Matching of Datasets

After the data cleaning, we conduct two separate time step matches, resulting in a dataset with a resolution of approximately (a) 2 minutes and (b) 1 hour. Note that because of data cleansing of the actual generation data, these datasets exist for only a subset of the days in 2014.

Since the actual generation data is available at uneven time stamps, we need to temporally match the data. For each point at minute resolution (Models 1, 3 and 4), we identify a point within the actual generation data that is within 30 seconds from the chosen point. For the times without a matching data point, we drop the point in all four datasets. This results in an approximately 2-minute resolution combined dataset for the actual generation data and Models 1, 3 and 4. For Plants A–D, this results in matched datasets of 137,658; 98,039; 102,971 and 92,661 points, respectively.

In addition, we want to compare these data and models with the NSRDB gridded weather data (model 2), which has hourly resolution. Therefore, we conduct a second, separate matching. For each hourly point in Model 2, we identify a point within the combined dataset for the actual generation data and Models 1, 3 and 4 that is within 30 seconds. For the times without a matching data point, we drop the point for all five datasets. This results in an approximately 60-minute resolution combined dataset for the actual generation data and Models 1, 2, 3 and 4. For Plants A–D, this results in matched datasets of 1,823; 1,174; 1,270 and 1,269 points, respectively.

Metrics with Scaled Datasets

Given that the first three models do not predict the power output in MW, we compare the model performance integrated over the full year using metrics with scaled datasets.

We use a metric with scaled datasets meant to highlight the performance of the models over the year. This is appropriate when using representative weather data and can be considered a good metric for siting plants.

Ratio of summations= *annual summation (actual generation data) / summation (model data)*.

Thus, a ratio less than 1 indicates that the model dataset overestimates the actual generation. Similarly, a value more than 1 shows the model dataset underestimating actual generation. This metric reveals how the power plant model (Model 4) performs overall in comparison with the actual generation data and also shows how well the assumed scaling for Models 1–3 performs.

Results and Discussion

The ratio of the observed annual production to the modelled annual production—which we term as the ratio of summations (Tables 1 and 2), a metric designed specifically for a siting decision—shows that the plant model (Model 4) performs slightly better than the other models for three of the four plants. On an average, Model 4 performs the best for all plants, with a mean error of 6±6% at an hourly resolution and 5±5% at an approximately 2-minute resolution.

Table 1. Ratio of summations for Models 1–4 to the actual generation data at ~2-minute resolution[2]

Model	Ratio of summations for each plant				Average difference from "perfect" of 1.0	
	A	B	C	D	Mean	Standard Deviation
1	0.755	0.848	0.798	0.777	21%	4%
2	-	-	-	-	-	-
3	0.963	1.114	1.120	0.975	7%	5%
4	1.011	0.957	0.990	0.881	5%	5%

Table 2. Ratio of summations for Models 1–4 to the actual generation data at hourly resolution[3]

Model	Ratio of Summations for each plant				Average difference from "perfect" of 1.0	
	A	B	C	D	Mean	Standard Deviation
1	0.814	0.915	0.898	0.815	14%	5%
2	0.853	0.965	0.966	0.870	9%	6%
3	0.948	1.083	1.103	0.957	7%	3%
4	0.999	0.933	0.981	0.867	6%	6%

The simple Bird clear sky index predicts the annual PV plant production of the 4 plants with an error of 14±5%, while the methods that incorporate local weather data predict the annual production with an error of 9±6%, 7±3% and 6±6% respectively (where uncertainties are the formal standard errors for the four plant locations). To put these differences between the models into perspective, we note that the inter-annual variability of the annual mean irradiation for 2000–2014 in Gujarat is ±3.6% for direct normal irradiance (DNI), ±3.1% for diffuse horizontal irradiance (DHI) and ±1.4% for global horizontal irradiance (GHI) [28]. Thus, although the models are statistically significantly different (p<0.01), the inter-annual variability of irradiance is larger than the differences between models. Further, while one-minute data is essential for many aspects of renewable integration studies, here they are shown to not provide a significant improvement over hourly data for site screening.

[2] A value nearer to one indicates a better performance. Note: Since Model 2 is of a different temporal resolution, this metric is not meaningful in comparison with other datasets, and hence has been removed.
[3] A value nearer to one indicates a better performance of the model.

Of course, Models 1–4 are but one part of the process of siting a PV plant. Additional constraints in urban areas might require other environmental characteristics leading to suboptimal siting [29, 30, 31], such as forcing siting on inclined terrain [32] or near landfills [33].

Conclusion

In this paper, we examine the performance of four models of annual solar photovoltaic production that are appropriate for site selection against actual generation data from utility-scale plants in the Indian state of Gujarat. Unlike other papers, we examine the difference between actual generation data and a number of different predictive models, some of which use nearby (but not coincident) weather data. We find that a simple Bird clear sky index predicts the annual PV plant production with an error of 14±5%, while the satellite data in National Renewable Energy Laboratory's 0.1° by 0.1° National Solar Radiation Database (NSRDB) has a prediction error is 9±6%. Stations belonging to India's SRRA are approximately 46–95 km away from the power plants; thus, the data has a prediction error of 7±3%. However, a power plant model using SRRA data to predict power output performs best with an error of 6±6% as it accounts for both radiation and temperature effects on the module performance. The inter-annual variability of the annual mean irradiation for 2000–2014 in Gujarat is ±3.6% for direct normal irradiance (DNI). Thus, we find that for site selection based on annual PV production, the inter-annual variability of irradiance is larger than the differences between models. Thus, in the absence of coincident data, satellite data (NSRDB) could be a preferred choice.

Acknowledgement

We thank the National Institute of Solar Energy, India for providing the solar radiation data for our study.

Funding

This research is based upon work supported by the Solar Energy Research Institute for India and the United States (SERIIUS), funded jointly by the U.S. Department of Energy subcontract DE AC36-08G028308 (Office of Science, Office of Basic Energy Sciences, and Energy Efficiency and Renewable Energy, Solar Energy Technology Program, with support from the Office of International Affairs) and the Government of India subcontract IUSSTF/JCERDC-SERIIUS/2012 dated 22nd Nov. 2012. Jay Apt received partial support from the Carnegie Mellon Climate and Energy Decision Making Center (CEDM), formed through a cooperative agreement between the National Science Foundation and CMU (SES-0949710).

Nomenclature

a_{CT}	empirically determined coefficient establishing the upper limit for module temperature at low wind speeds and at high effective radiation on the panel
b_{CT}	empirically determined coefficient establishing the rate at which module temperature drops as wind speed increases
DHI	diffuse radiation incident upon a horizontal surface (W/m^2)
DNI	direct normal irradiance upon a horizontal surface (W/m^2)
GHI	global horizontal irradiance incident upon a horizontal surface (W/m^2)
G_{ref}	reference solar radiation (1000 W/m^2 at Standard Reporting Condition)
G_T	net effective solar radiation incident on a tilted panel (W/m^2)
K_T	temperature coefficient of power for a given module (%/°C)
N_{panels}	number of panels in a plant
P_{module}	power for a given module (W)
P_{plant}	representative generation (W)
P_{ref}	specified power rating of the module (in W) at standard test conditions (Wp)
T_{amb}	1-minute average values of air temperature (°C)
T_{cell}	cell temperature inside the module (°C)
T_m	back-surface module temperature (°C)
T_{ref}	reference temperature (25°C at Standard Reporting Condition)
ΔT	empirically determined T_{cell} - T_m at G_{ref} (°C)

WS	wind speed at 10-m height (m/s)
η_{PCU}	efficiency of the inverter/power conditioning units
θ_T	module tilt angle (°)
θ_z	solar zenith angle (°)
γ_A	solar azimuth angle (°)
γ_T	surface azimuth angle (°)
η_{plant}	empirical scaling to each plant
NC	nameplate capacity (MWp)
R_d	module tilt factor for the diffused component of radiation
R_g	module tilt factor for the global radiation component
R_b	module tilt factor for the direct or beam radiation component

References

1 Hagerman, S.; Jaramillo, P.; Morgan, M.G. (2016). Is rooftop solar PV at socket parity without subsidies? Energy Policy. 89, 84–94.

2 USDOE Office of Energy Efficiency & Renewable Energy. Sunshot Initiative Goals. https://energy.gov/eere/sunshot/sunshot-initiative-goals Last accessed January 16, 2017.

3 Government of India Ministry of New and Renewable Energy. Scheme/ Documents. http://www.mnre.gov.in/solar-mission/jnnsm/introduction-2/ Last accessed January 16, 2017.

4 PV Performance Modeling Collaborative. PVPMC Home. https://pvpmc.sandia.gov/ Last accessed January 16, 2017.

5 Gilman, P. (2015). SAM Photovoltaic Model Technical Reference. National Renewable Energy Laboratory. 59 pp.; NREL/TP-6A20-64102

6 http://pvwatts.nrel.gov/

7 Hummon, M.; Cochran, J.; Weekley, A.; Lopez, A.; Zhan, J.; Stoltenberg, B.; Parsons, B.; Batra, P.; Mehta, B.; Patel, D. Variability of Photovoltaic Power in the State of Gujarat Using High Resolution Solar Data. Technical Report NREL/TP-7A40-60991 . March 2014.

8 Kumar, Ravinder; Umanand, L. (2005). Estimation of global radiation using clearness index model for sizing photovoltaic system. Renewable Energy. 30(15), 2221-2233.

9 Karakoti, Indira; Pande, Bimal; Pandey, Kavita. (2011). Evaluation of different diffuse radiation models for Indian stations and predicting the best fit model. Renewable and Sustainable Energy Reviews. 15(5), 2378-2384

10 Ghazanfar Khan, Shikha Rathi (2014). Optimal Site Selection for Solar PV Power Plant in an Indian State Using Geographical Information System. International Journal of Emerging Engineering Research and Technology . 2(7),260-266.

11 Amit Jain, Rajeev Mehta & Susheel K. Mittal (2011). Modeling Impact of Solar Radiation on Site Selection for Solar PV Power Plants In India. International Journal of Green Energy. 8(4), 486-498.

12 Ravindra M.Moharil; Prakash S.Kulkarni. (2009). A case study of solar photovoltaic power system at Sagardeep Island, India.Renewable and Sustainable Energy Reviews. 13(3), 673-681.

13 Debyani Ghosh; P.R.Shukla; Amit Garg; P.Venkata Ramana (2002). Renewable energy technologies for the Indian power sector: mitigation potential and operational strategies Renewable and Sustainable Energy Reviews. 6(6), 481-512.

14 Imtiaz Ashraf, A. Chandra (2004). Artificial neural network based models for forecasting electricity generation of grid connected solar PV power plant. International Journal of Global Energy Issues (IJGEI), 21(1/2).

15 Sharma, V.; Chandel, S.S. (2013). Performance analysis of a 190 kWp grid interactive solar photovoltaic power plant in India. Energy. 55(15), 476-485.

16 Klima, K.; Apt, J. (2015). Geographic smoothing of solar PV: Results from Gujarat. *Environmental Research Letters*. 10(10).

17 United Nations Framework Convention on Climate Change (UNFCCC) - Clean Development Mechanism (CDM), "Project Design Document for small scale CDM project activities - 10 MW solar photovoltaic power plant in Rajkot, Gujarat (India)," New Delhi, 2012.

18 United Nations Framework Convention on Climate Change (UNFCCC) - Clean Development Mechanism (CDM), "Clean Development Mechanism small scale program activity design document form -5 MW solar PV power project, Sabarkantha District, Gujarat, India," Ahmedabad, 2013.

19 United Nations Framework Convention on Climate Change (UNFCCC) - Clean Development Mechanism (CDM), "Clean Development Mechanism project design document form - 10 MW solar PV power project by Azure Power," New Delhi, 2012.

20 United Nations Framework Convention on Climate Change (UNFCCC) - Clean Development Mechanism (CDM), "Project Design Document for CDM project activities - 25 MW Mithapur solar power project," Mumbai, 2012.

21 National Renewable Energy Laboratory. PVWatts Calculator. http://pvwatts.nrel.gov Last accessed January 16, 2017.

22 Bird, R. E., and R. L. Hulstrom, Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces, Technical Report No. SERI/TR-642-761, Golden, CO: Solar Energy Research Institute, 1981

23 National Renewable Energy Laboratory. National Solar Radiation Database (NSRDB). https://nsrdb.nrel.gov/nsrdb-viewer Last accessed January 16, 2017.

24 National Renewable Energy Laboratory. NSRDB Data Viewer. https://maps.nrel.gov/nsrdb-viewer/ Last accessed January 16, 2017.

25 Solar Radiation Resource Assessment database 2014-15, maintained by National Institute of Wind Energy (NIWE), India. Weblink:http://niwe.res.in/department_srra.php

26 Duffie, J.A.; Beckman, W.A. (2013). *Solar Engineering of Thermal Processes: Fourth Edition*. ISBN: 978-0-470-87366-3

27 USDOC National Oceanic and Atmospheric Administration. Earth System Research Lab. Solar Position Calculator. http://www.esrl.noaa.gov/gmd/grad/solcalc/azel.html Last accessed January 16, 2017.

28 Sengupta, M.; Weekley, A.; Habte, A.; Lopez, A.; Molling, C.; Heidinger, A. (2015). Validation of the National Solar Radiation Database (NSRDB) (2005–2012). NREL/CP-5D00-64981.

29 EERE chapter 8

30 Doris, E.; Lopez, A.; Beckley, D. (2013). Geospatial Analysis of Renewable Energy Technical Potential on Tribal Lands. DOE/IE-0013 .

31 Macknick, J.; Quinby, T.; Caulfield, E.; Gerritsen, M.; Diffendorfer, J.; Haines, S. (2014). Geospatial Optimization of Siting Large-Scale Solar Projects. Technical Report. NREL/TP-6A50-61375 .

32 Charabi, Y.; Rhouma, M.B.H.; Gastli, A. (2014). Siting of PV power plants on inclined terrains. *International Journal of Sustainable Energy*. 35(9) DOI: 10.1080/14786451.2014.952298

33 Kiatreungwattana, K.; Mosey. G.; Jones-Johnson, S.; Dufficy, C.; Bourg, J.; Conroy, A.; Keenan, M.; Michaud, W.; Brown, K. (2013). Best Practices for Siting Solar Photovoltaics on Municipal Solid Waste Landfills. Technical Report. NREL/TP-7A30-52615 .

Appendix A

Table A.1. Power plant characteristics used in this analysis

Designation	Plant A	Plant B	Plant C	Plant D
Project Name	10 MW Solar PV Power Project by Azure Power	10 MW Solar PV plant in Rajkot	25 MW Mithapur Solar PV Power Project	5 MW Solar PV power project
Latitude and Longitude	23.4431 °N, 73.2008 °E	21.7364 ºN, 70.12 ºE	22.4074 °N, 68.9911 °E	23.3041 ºN, 73.3085 ºE
Altitude	133m	48m	3.5m	116m
Plant Details	Khadoda and Shinol Village, Modasa & Dhansura Tehsil, Sabarkantha District, Gujarat, India	Meravadar Village, Upleta Tehsil, Rajkot District, Gujarat, India	Mithapur Village, Dwarka Taluk, Jamnagar District	Sherdi Shakhandi Village, Dhansura Town, Sabarkantha District, Gujarat, India
Solar Radiation Resource Assessment (SRRA) Station Location	PDPU campus, Gandhinagar	132 kV Enercon Station, Bhogat,	132 kV Enercon Station, Bhogat,	PDPU campus, Gandhinagar
SRRA Station Number	1834	1799	1799	1834
SRRA Station Latitude and Longitude	23.1549 ºN, 72.6669 ºE	22.0287 ºN, 69.2611 ºE	22.0287 ºN, 69.2611 ºE	23.1549 ºN, 72.6669 ºE
Percent Missing Data in SRRA	5.30%	0.44%	0.44%	5.30%
Distance Between Plant and SRRA Weather Station (Google maps)	36 km	95 km	50 km	68 km
Installed Capacity	10.21 MW_{DC}	10.00 MW_{DC}	25.01 MW_{DC}	4.99 MW_{DC}
Promoter	Azure Power (Haryana) Pvt. Ltd.	Green Infra Solar Energy Ltd.	Tata Power Renewable Energy Ltd.	Aatash Power Private Ltd.

Data Start Date/Time (DD-MM-YYYY HH:MM UTC)	08-02-2014 10:54	01-02-2014 00:00	01-02-2014 00:00	08-02-2014 10:54
Data End Date/Time (DD-MM-YYYY HH:MM UTC)	31-12-2014 23:59	31-12-2014 23:59	31-12-2014 23:59	31-12-2014 23:59
Inverter Manufacturer and Rating	Unknown - 630 kW x 16	SMA SC 630 CP – 630 kW x 16	ABB - 500 kW x 34, SMA 800 kW x 10	ABB - 500 kW x 10
Assumed Inverter Efficiency	95%	98%	96%	96%
Module Manufacturer	Suntech – Multi crystalline	First Solar – Thin Film CdTe	Tata BP Solar - Multi crystalline, Suntech - Multi crystalline, Canadian Solar - Multi crystalline	REC peak energy black series – Multi crystalline
Model	STP280-24Vd	FS 380	TBP 4230X/3230X, STP 230-20/WD, CS6P-235P, CS6P-240P	REC245PE BLK
Temperature Coefficient of Maximum Power (%/ºC)	-0.47%	-0.25%	-0.442%, -0.44%, -0.43%, -0.43%	-0.4%
Module Tilt	Assumed equal to latitude	15°	Assumed equal to latitude	Assumed equal to latitude

Table A,2: Assumed values in the Bird clear sky index model

Metric	Typical Value	Assumed Value
Barometric pressure	1013 mb at sea level	1006 mb
Ozone thickness of atmosphere	0.05–0.4 cm	0.35 cm
Water vapour thickness of atmosphere	0.01–6.5 cm	4 cm
Aerosol optical depth at 500 nm (unit less)	0.02–0.5	0.35
Aerosol optical depth at 380 nm (unit less)	0.1–0.5	0.35
Forward scattering of incoming radiation (unitless)	0.85	0.85

Table A.3: Rg and R_d values (a) as used in the main paper and (b) using the common assumption of 2D or 3D isotropy

Dataset	R values (unitless)			
	Plant A	Plant B	Plant C	Plant D
R_g, As used in paper, assuming 2D isotropy	0.0413	0.0170	0.0377	0.0408
assuming 3D isotropy	0.0756	0.0686	0.0726	0.0783
R_d, As used in paper, assuming 2D isotropy	0.9587	0.9830	0.9623	0.9592
assuming 3D isotropy	0.9244	0.9314	0.9274	0.9217

Table A.4. Scaling factors used to fit Datasets 2–4 to the actual generation data

Dataset	Scaling Factor			
	Plant A	Plant B	Plant C	Plant D
η_{plant}	90%	80%	78%	78%

Appendix B

Calculation of net effective solar radiation incident on a tilted panel

By definition, G_T is

$$G_T = DNI * R_b + DHI * R_d + \rho * GHI * R_g \quad \textbf{(1)}$$

where *DNI* is the direct radiation normal to the beam at the Earth's surface (W/m²); *DHI* is the diffuse radiation incident upon a horizontal surface (W/m²); *GHI* is the global radiation incident upon a horizontal surface (W/m²); ρ is the foreground's albedo (unitless) and R_b, R_d, and R_g are module tilt factors for each component of the radiation (unitless).

Further, by definition, *GHI*, *DNI* and *DHI* are related by the solar zenith angle as:

$$GHI = DNI * \cos\theta_Z + DHI \quad \textbf{(2)}$$

and, therefore, R_b is defined geometrically as:

$$R_b = \cos\theta = \cos\theta_Z \cos\theta_T + \sin\theta_Z \sin\theta_T \cos(\gamma_A - \gamma_T) \quad \textbf{(3)}$$

where θ is the angle of incidence of the sun rays on the tilted plane, γ_A is the solar azimuth angle (measured in degrees clockwise from north) and γ_T is the surface azimuth angle (zero due south, negative towards east and positive towards west). Therefore, the only assumptions made are for ρ, R_d and R_g. For all datasets in the paper, we assume ρ = 0.2, a common assumption for land.

The values for R_d and R_g vary in complexity between models. In the main paper, we assume 2D isotropy holds [1], and thus:

$$R_d = \frac{1+\cos\theta_T}{2} \quad \textbf{(4)}$$

$$R_g = \frac{1-\cos\theta_T}{2} \quad \textbf{(5)}$$

For comparison, R_d and R_g are listed in Table A.3. Values are similar between plants with the assumed module tilt equal to the latitude; this is because the latitudes are very similar to one another for the plants considered in our study.

We note that more recent literature has derived 3D isotropy values [2]; those have been listed in Table A.3 for reference. The 3D isotropic values are calculated to be within 5% of the highly complex 3D non-isotropic values for a tilt (θ_T) of 22° [4]. However, these values have not yet been widely adopted in power plant modelling, and hence we use the 2D values in the main paper.

Conversion of solar radiation into generation for Models 1–3

For each power plant with a nameplate capacity of NC (units of MW), we identify a scaling factor to convert G_T to P_{plant} (f, units of MWm²/W). Note that because Models 1–3 are all representations of G_T, this scaling factor is a function only of the plant:

$$P_{plant}(Plant, Model) = G_T(Plant, Dataset) * f(Plant) \quad \textbf{(6)}$$

To identify the appropriate scaling factor, we first assume that the plant would reach its maximum capacity at least once per year, at the time of the year with the maximum G_T for the Bird sky index (the summer solstice). For each plant, this value reaches an annual maximum of approximately 1000 W/m², and so we would empirically find:

$$\mathrm{f}(Plant, Model) \approx 0.001NC \quad \textbf{(7)}$$

where 0.001 has units of m²/W. We then apply an empirical scaling to each plant of η_{plant} (unit less), or:

$$\mathrm{f}(Plant, Model) \approx 0.001NC * \eta_{plant} \quad \textbf{(8)}$$

Calculation of solar power generation in Model 4

The details of equations used in calculation of solar power as per the power plant model (Model 4) is shown below.

Calculation of cell temperature of the solar panel

The temperature of the solar cell is dependent on the meteorological data and cell configuration. For the meteorological data, we use Model 3 (Section 2.2.3). For the cell configuration, we assume a flat-plate module type of glass/cell/polymer sheet and an open rack module mount configuration. Given these assumptions, the back-surface module temperature (T_m, in °C) is estimated following King et. al. (2004) [3] as:

$$T_m = G_T * exp(a_{CT} + b_{CT} * WS) + T_{amb} \quad \textbf{(9)}$$

where a_{CT} is the empirically determined coefficient (unitless) establishing the upper limit for module temperature at low wind speeds and at a high effective radiation on the panel ([35], -3.56), b_{CT} is the empirically determined coefficient establishing the rate at which module temperature drops as wind speed increases ([35], -0.075 s/m) and the resulting exponential is in units of °Cm²/W. Note that while the wind data was given at a hub height of 6 m, here *WS* indicates the 10-m height velocity. We checked the effect of this variation and found that this roughly increases the wind velocity by 5%, which has a negligible effect on T_m. Given this, the cell temperature inside the module (T_{cell}, in °C) is:

$$T_{cell} = T_m + \Delta T \left(\frac{G_T}{G_{ref}}\right) \quad \textbf{(10)}$$

where G_{ref} is the reference solar radiation (1,000 W/m² at Standard Reporting Condition) and ΔT is the empirically determined T_{cell} - T_m at G_{ref}, or from [35], 3°C.

Calculation of the power generated by a solar PV plant

Given the cell temperature, the power for a given module (P_{module}, in W) is estimated following Whitaker et al. (1991) [4] as:

$$P_{module} = P_{ref} \left(\frac{G_T}{G_{ref}}\right) \left[1 + K_T\left(T_{cell} - T_{ref}\right)\right] \quad \textbf{(11)}$$

where (P_{ref}) is the specified power rating of the module (in W) at standard test conditions of G_{ref} (1000W/m²) and T_{ref} (25°C) and K_T is the temperature coefficient of maximum power for the given module (%/ºC) (see Table A.1). This power is then scaled by the total number of panels in the plant (N_{panels}) to estimate the power output of the entire plant (P_{plant}) as:

$$P_{plant} = N_{panels} * P_{module} * \eta_{PCU}/100 \quad \textbf{(12)}$$

where η_{PCU} is the efficiency of the inverter / power conditioning units (unitless). Note that in Table A.1, we assume the inverter efficiency to be slightly less than the rated efficiency to account for the performance fluctuation of the inverter.

Additional power plant model assumptions

A few characteristics of PV systems are not modelled. These include dust/soiling on the module power output, partial shading due to cloud cover, module degradation and additional reductions in power because of conduction losses in lines and transformers. When such effects are taken into account, we expect generation to decrease [5], [6] and [7].

1 Liu, B.Y.H.; Jordan, R.C. (1963). The long-term average performance of flat-plate solar energy collectors. *Sol Energy*. 7, 53.

2 Badescu, V. 3D isotropic approximation for solar diffuse irradiance on tilted surfaces. Renewable Energy. Volume 26, Issue 2, June 2002, Pages 221–233

3 King, D.L.; Boyson, W.E.; Kratochvil, J.A. (2004). Photovoltaic array performance model SAND2004-3535 http://prod.sandia.gov/techlib/access-control.cgi/2004/043535.pdf Last accessed January 16, 2017.

4 Whitaker, C.M.; Townsend, T.U.; Wenger, H.J.; Iliceto, A.; Chimento, G.; Paletta, F. (1991). Effects of irradiance and other factors on PV temperature coefficients. *Photovolt. Spec. Conf. 1991., Conf. Rec. Twenty Second IEEE.* 1, 608–613.

5 Charabi, Y.; Gastli, A. 2013. Integration of temperature and dust effects in siting large PV power plant in hot arid area. *Renewable Energy*. 57, 635-644.

6 Annathurai, V.; Gan, C.K.; Baharin, K.A.; Ghani, M.R.A. (2016). Shading Analysis for the Siting of Solar PV Power Plant. *Journal of Engineering and Applied Sciences.* 11(8), ISSN 1819-6608

7 Jordan, D.C.; Kurtz, S.R. (2012). *Photovoltaic Degradation Rates — An Analytical Review.*

NREL/JA-5200-51664