Validation of SolarAnywhere Enhanced Resolution Irradiation in California

Mohammad Jamaly and Jan Kleissl

Dept. of Mechanical and Aerospace Eng., University of California, San Diego

June 26, 2012

Executive Summary

A 1 km high resolution solar resource dataset (SolarAnywhere, SAW) based on satellite data was developed by Clean Power Research under the CSI program. We compare the SAW estimates to solar irradiation measurements at 53 ground stations throughout California. At a single high quality ground site, SolarAnywhere was unbiased, but it overestimated ground measured irradiation by 3.7 +- 0.9% at the other sites. A larger and more consistent overestimate at all sites was observed during May – July and also in clear conditions, indicating that improvements to the clear sky model would remove at least some of the difference. Soiling may have affected the data quality of the ground sites and could explain some of the difference. Biases need to be further examined using high quality ground sites. SAW random errors were found to be small. Overall SAW is the most accurate publicly available solar resource dataset.

1. Introduction

Clean Power Research's commercially available SolarAnywhere (SAW) provides Global Horizontal Irradiation (GHI) and Direct Normal Irradiation (DNI) derived from Geostationary Operational Environmental Satellite (GOES) visible imagery at 1 km spatial and 30 minutes temporal resolution for California [1]. To obtain GHI, a cloud index is calculated from the reflectance in each pixel measured by the satellite. Instantaneous GHI for each pixel is then calculated by using the cloud index to decrease the irradiation calculated using a clear sky model that considers local and seasonal effects of turbidity [2].

Perez et al. [3] conducted a validation of an earlier version of the SAW algorithm against high quality ground measurements sites across the US (but outside California); the SAW dataset was found to have mean bias errors (MBE) between -5 and 15 W m-2 and root mean square errors (RMSE, based on hourly averages) ranging from 73-118 W m^{-2}. The objective of this study is to validate the accuracy of the enhanced resolution SolarAnywhere database in California against ground data.

2. Methodology

2.1. Datasets

SAW enhanced resolution provides satellite-derived irradiation with 30-min temporal and 1 km spatial resolutions. For comparison, ground measured irradiation data are analyzed; the California Irrigation Management Information System (CIMIS) with 124 active weather stations [4] and the NOAA Integrated Surface Irradiance Study (ISIS) network with one station in Hanford, CA [5].

CIMIS is operated by the Department of Water Resources (DWR) and has been operational since 1982 (CIMIS, 2009a). Each CIMIS station is equipped with a Li-Cor LI200S photodiode pyranometer, accurate under typical conditions to ±5% (CIMIS, 2009b). GHI is reported as an hourly average of 60 independent measurements within the hour [4].

The measured GHI of the ISIS in Hanford, CA and the 124 CIMIS stations are compared with the SAW GHI data of the corresponding pixel in which the stations are located. The analysis is conducted for yearly data in 2009 and 2010.

2.2. Interpolation between Datasets

SolarAnywhere provides 30-min irradiation centered at :00 and :30. CIMIS provides hourly averaged irradiation with an interval-ending timestamp at :00 and ISIS reports 3-min irradiation.

To compare SAW and CIMIS, the SAW :45 to :15 interval (reported at :00) is disaggregated into :45 to :00 and :00 and :15 and then aggregated with the :15 to :45 interval (reported at :30) to correspond to the hourly CIMIS interval. Consequently, the SAW clear sky index (*kt*) is calculated for each 30 min interval using 1 min clear sky irradiation based on the Ineichen model with Linke Turbidity from the Solar Radiation Data (SoDa) database [6]. Then, *kt* assigned to the :00 to :15, :15 to :45, and :45 to :00 intervals are multiplied by the clear sky irradiation for each 1 min interval to obtain hourly GHI. For SAW against ISIS comparison, the 3-min ISIS irradiation is averaged over the SAW time interval.

2.3. Comparison for Clear Sky Conditions

For the validation, clear sky conditions are considered separately to evaluate differences between modeled and measured GHI caused by atmospheric composition or aerosol optical depth and not by cloud cover or cloud optical depth. For clear skies the SAW irradiation is essentially calculated from the Ineichen clear sky model with climatological (monthly) turbidity from NREL's METSTAT database, which is a source of error given variable actual aerosol optical depth in the atmosphere. Clear conditions are assumed to exist if 0.85 < *kt* < 1.1 and *std*[*kt*(t-1h:t+1h)] < 0.03, where *std* is the standard deviation, and t is time. These expressions filter for large *kt* and low variability which is characteristic for clear conditions. This criterion had to be met by both CIMIS (or ISIS) and SAW so that only simultaneous and collocated clear data was considered. For reference, we compute clear conditions from the Ineichen model with the SoDA turbidity [7], [8].

2.4. Error Metrics

Mean Bias Error (MBE) describes persistent differences between two datasets. MBE of the SAW and CIMIS data is calculated as

$$\mathrm{MBE} = \frac{1}{N}\sum_{n=1}^{N}(GHI_{SAW} - GHI_{CIMIS}) \qquad \text{Eq. (1),}$$

where *N* is the number of samples. Also, the relative MBE is calculated as

$$\mathrm{rMBE} = \frac{MBE}{mean(GHI_{CIMIS})} \qquad \text{Eq. (2).}$$

The confidence Interval (CI) is calculated to determine the significance level of the difference between the datasets. Using yearly averaged MBE of all 124 CIMIS stations ($\overline{MBE_\iota}$), the CI of the difference between SAW and CIMIS datasets is calculated as

$$CI = \overline{MBE_\iota} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) \qquad \text{Eq. (3),}$$

where $\overline{MBE_\iota}$ is the average SAW bias error over all CIMIS stations, $Z_{\alpha/2}$ is the confidence level coefficient and equals to 1.96 for 95% confidence level (α=0.05), and σ is the standard deviation of MBE at all the stations.

To illustrate diurnal or seasonal patterns in the data, the GHI data are averaged over all days of each month for a given time-of-day (ToD) to yield $GHI_{CIMIS}(m, ToD), GHI_{ISIS}(m, ToD)$ and

$GHI_{SAW}(m, ToD)$ for m = 1,…, 12 and ToD = 1, …, 24 h. MBE is calculated for each ToD and each month separately to yield $MBE_{MT} = MBE(m, ToD)$ [9]. Relative MBE_{MT} is obtained by normalizing by the average of $GHI_{CIMIS}(m, ToD)$.

$$r\mathrm{MBE_{MT}}(\mathrm{m}, \mathrm{ToD}) = \frac{\mathrm{MBE_{MT}}(m, ToD)}{mean[GHI_{CIMIS}(m, ToD)]} \quad \text{Eq. (4).}$$

2.5. Data Quality Control

CIMIS provides an initial QC assessment based on procedures described by Meek and Hatfield (1994), issuing flags that allow the user to remove any data that appears faulty or erroneous [10]. These flags, detailed on the CIMIS website [4], restrict any data that contain obvious outliers or unphysical characteristics. CIMIS provides a further description of the QC method in the CIMIS technical manual [11]. Luoma and Kleissl (2012) reviewed data from each CIMIS station individually [12]; All flagged CIMIS data were excluded leaving 124 CIMIS stations with 70% or more available and high quality data. The same quality control is applied to the ISIS data and all flagged ISIS data (as described on NOAA website [5]) are not considered. Only data with solar zenith angle less than 75° are considered to avoid error in sensor cosine response and shading.

$r\mathrm{MBE}_{year}$ (for the whole data) and $r\mathrm{MBE}_{clear}$ (for the times with clear condition) between SAW modeled GHI and CIMIS and ISIS measured GHI is calculated. To exclude outliers, the CIMIS stations with $r\mathrm{MBE}_{year}$ or $r\mathrm{MBE}_{clear}$ out of the range of 0.25-0.75 quantiles are excluded. Therefore, 52 CIMIS stations, along with the ISIS station, are considered in this study (Fig. 1).

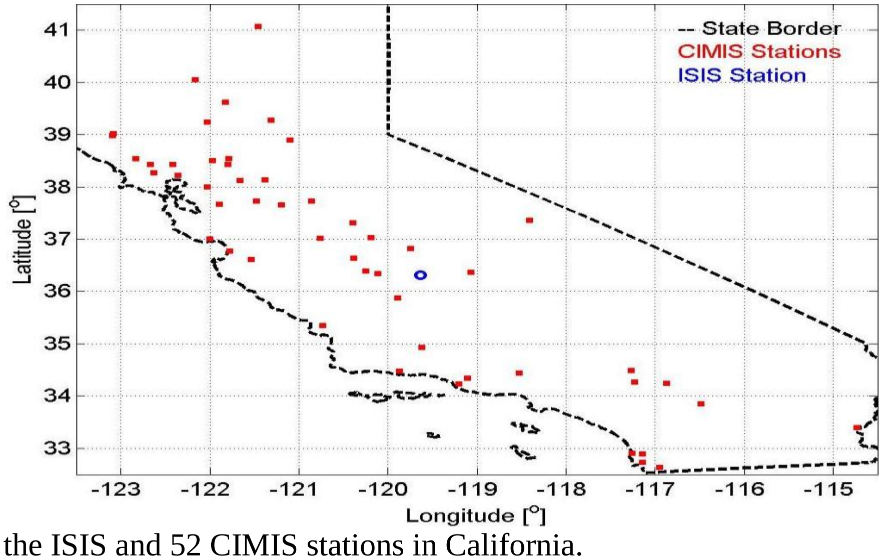


Fig. 1: Map of the ISIS and 52 CIMIS stations in California.

3. Results

3.1. Comparison of California-wide GHI averages across the year

The daily average of CIMIS, SAW, and clear sky GHI (based on the Ineichen model with Linke Turbidity from the SoDa database) averaged over 52 stations is computed for the year 2010. In Fig. 2, time series of daily averages through 2010 are shown. In Fig. 3, the daily averages are calculated only for clear sky conditions for both SAW and CIMIS data (based on the clear condition criteria defined in section 2.3). Generally, SolarAnywhere overestimates CIMIS irradiation data by 17 to 20 W m^{-2} or 3 to 5% throughout the year. The bias is somewhat larger for the clear periods at 21 to 31 W m^{-2}. During clear periods, SAW is larger than the clear sky model, while CIMIS measurements are smaller than the clear sky model.

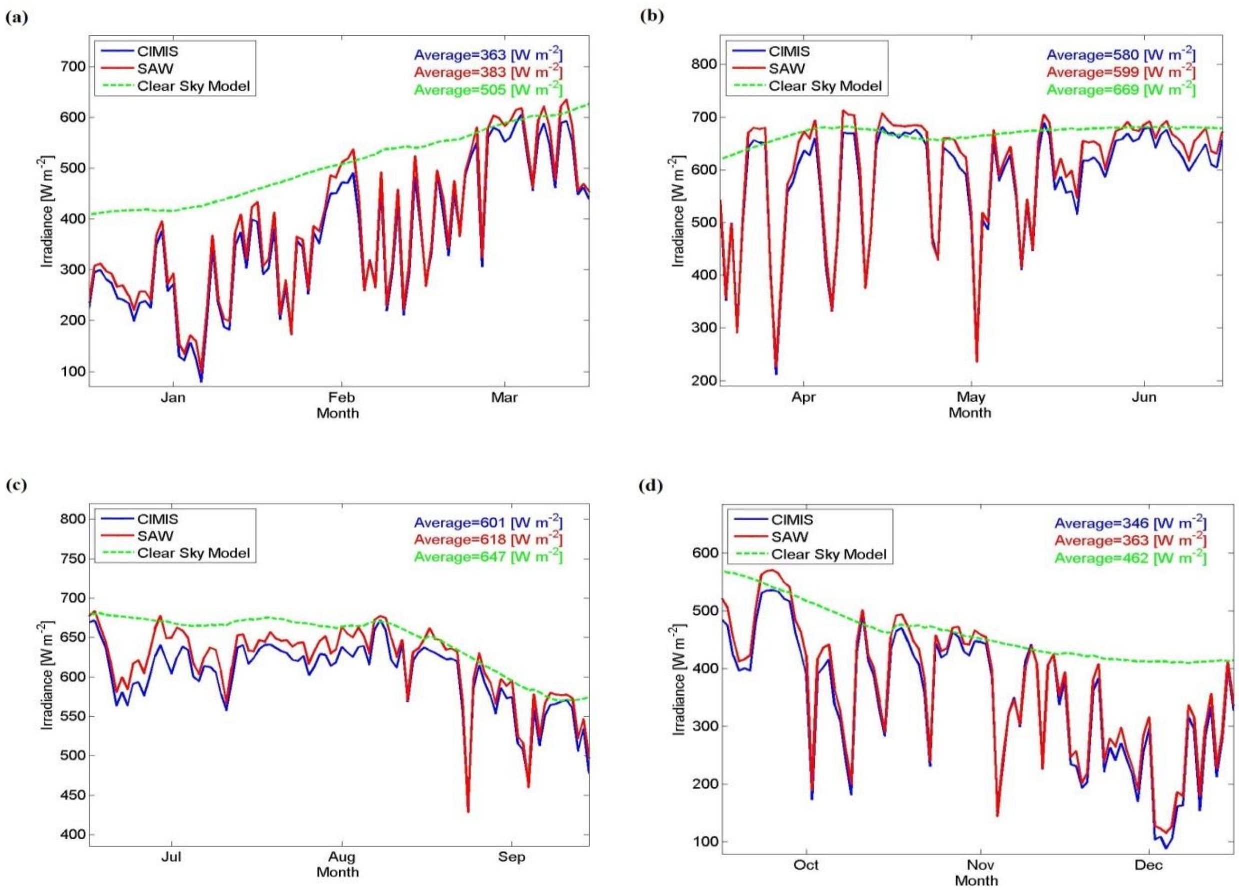


Fig. 2**, Daily average CIMIS versus SAW:** Mean daily (for SZA<75^{o}) irradiation averaged over 52 CIMIS stations and the collocated SAW pixels in 2010 broken out by season. The upper envelope of the CIMIS and SAW lines is expected to be similar to the clear sky model that is shown for reference. The day-to-day variability in the clear sky model is due to missing CIMIS data causing different CIMIS sites/ time steps to be chosen for each day.

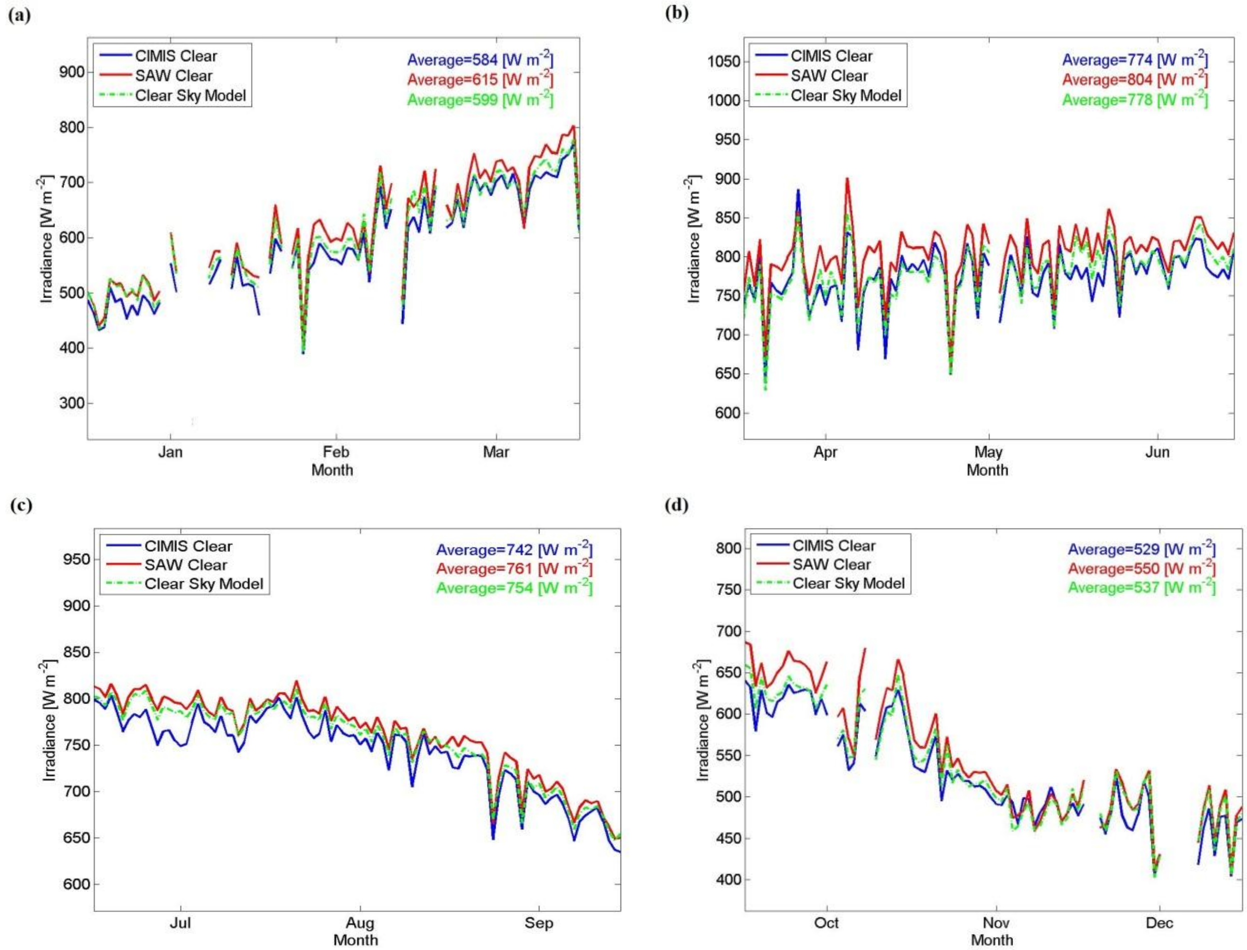


Fig. 3, **Daily average CIMIS versus SAW for clear conditions**: Mean SAW and CIMIS daily irradiation during clear sky condition (as defined in section 2.2) averaged over 52 stations in 2010. 34% of the daytime is clear. Days with less than 2 hours of clear sky data averaged across all stations are not shown. The day-to-day variability in the clear sky model is due to missing CIMIS data causing different CIMIS sites/ time steps to be chosen for each day.

The ISIS site in Hanford, CA is of special interest as data quality is expected to be higher and since SAW was calibrated and validated at this site before [13]. The daily average of ISIS GHI and SAW GHI for the whole year and the times with clear sky condition are shown in Figs. 4 & 5. The daily averages are calculated for the same available time steps of all the datasets. There is excellent agreement between SAW and ISIS with seasonal biases of less than 2%. The only exceptions are July and August when consistent over- and underestimates of 2 to 5% are observed. Similar results are observed in clear conditions (Fig. 5).

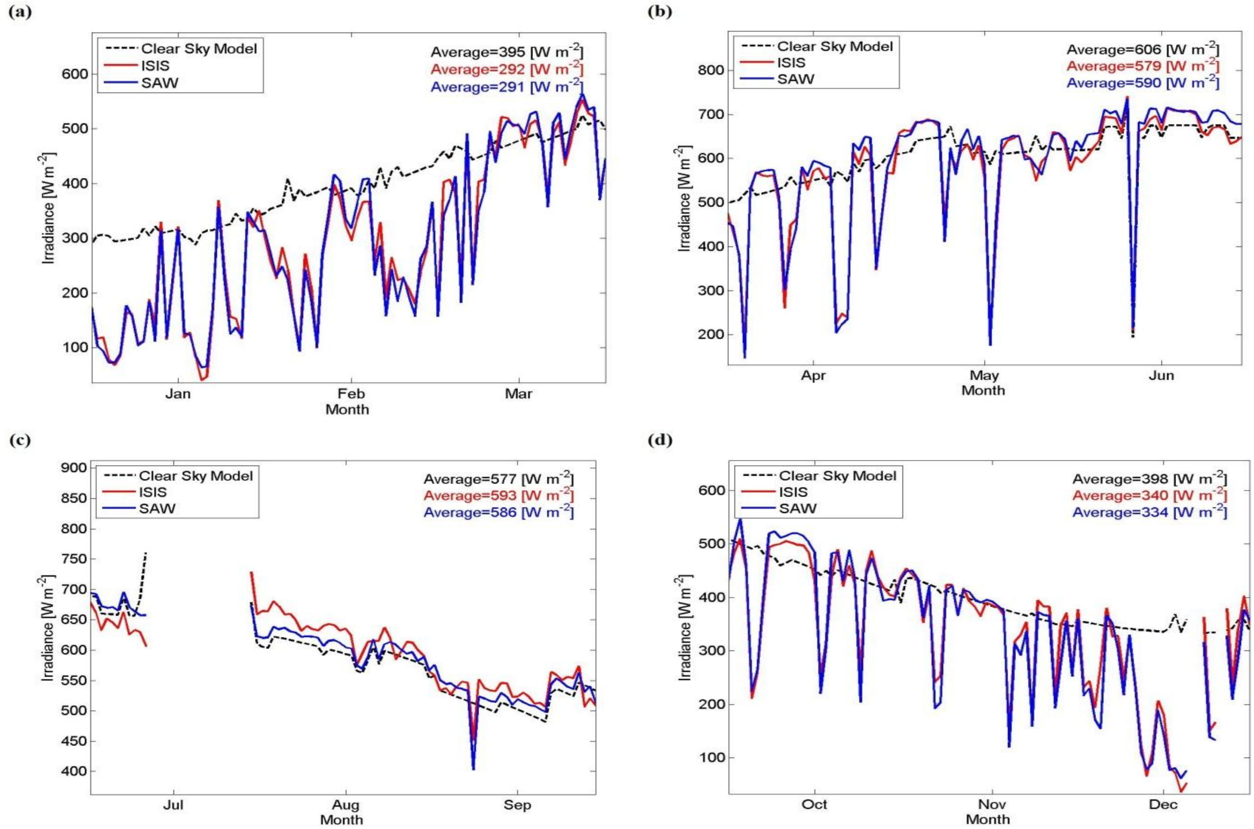


Fig. 4, **Daily average Hanford ISIS versus SAW**: Mean daily irradiation at the ISIS ground station in Hanford, CA for 2010. The day-to-day variability in the clear sky model is due to missing ISIS data causing different ISIS time steps to be chosen for each day.

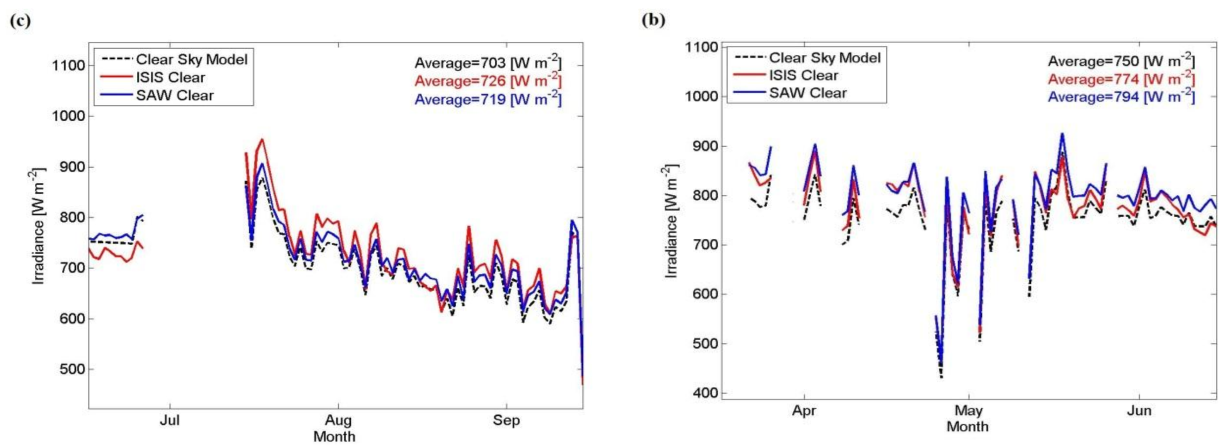
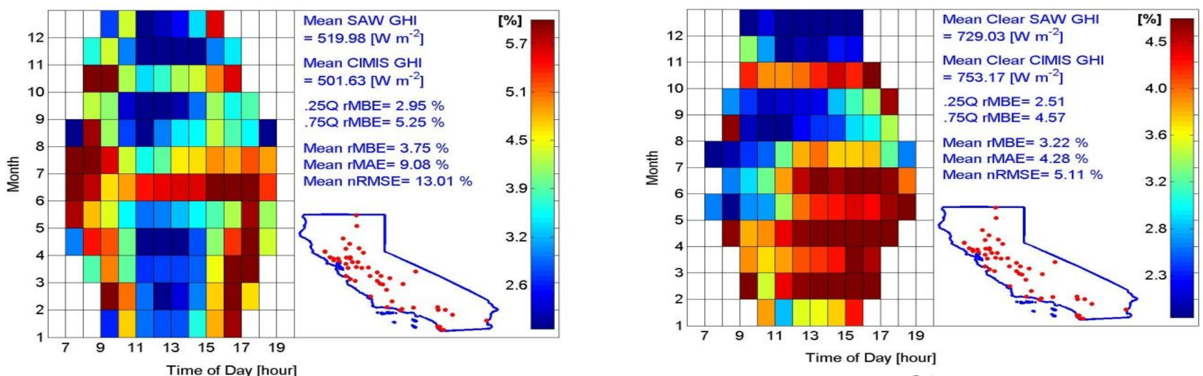


Fig. 5, **Daily average Hanford ISIS versus SAW for clear conditions**: Same as Fig. 4 but for clear sky condition in Hanford, CA. 35% of daytime data are considered clear. Winter months are not shown since few clear data points exist.

3.2. Climatologies of MBE by month and time-of-day

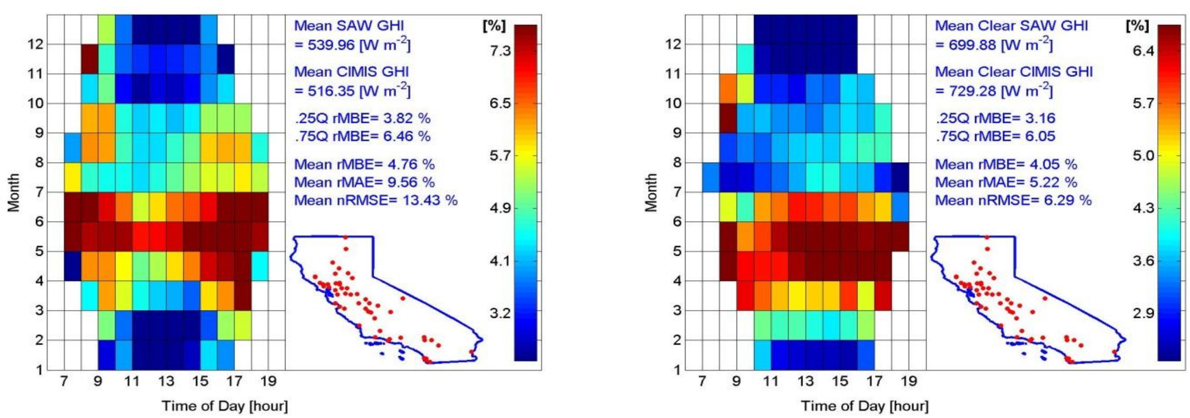
Averaged 2010 $rMBE_{MT}$ of all 52 CIMIS stations is shown in Fig. 6 (for both the whole dataset and the clear sky conditions). Overall bias error (SAW overestimates) of 18 W m^{-2} or 3.7% are observed, consistent with Fig. 2. MBE is 24 W m^{-2} or 3.2% in clear conditions. The biases are largest in June and to a lesser extent in May and July. For the rest of the year biases are less than 2% during midday, but larger in the (less important) morning and evenings. Clear sky biases are largest from February until June.

The same graphs for 2009 are shown in Fig. 7. Although MBEs are larger in 2009, the same trends are observed for both years. Since the 2009 and 2010 errors are also in agreement for other sub-regions and Hanford, CA, we will only show 2010 data from here onwards.



(a) (b)

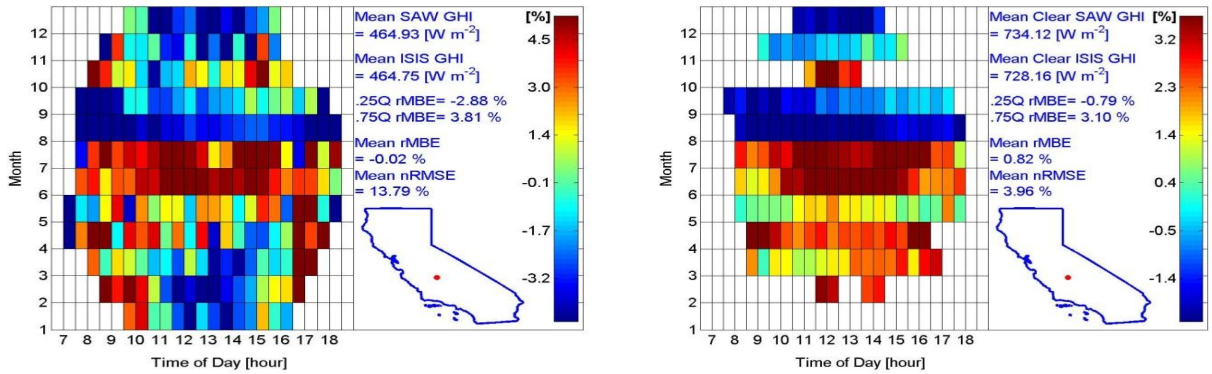
Fig. 6, **rMBE by month and time-of-day for CIMIS versus SAW for (a) all and (b) clear conditions:** $rMBE_{MT}$ of 2010 SAW and CIMIS data (averaged over 52 stations). (a) all data, (b) data in clear sky conditions. The caption indicates annual mean SAW and CIMIS GHI, 0.25 and 0.75 quantiles of $rMBE_{MT}$, annual mean rMBE, rMAE, and rRMSE. All relative errors are obtained from hourly data by dividing annual MBE, MAE, and RMSE by mean(GHI_{CIMIS}).



(a) (b)

Fig. 7: same as Fig. 6 but for 2009.

$rMBE_{MT}$ at Hanford, CA (Fig. 8) is qualitatively consistent with the overall trends. There is no bias on average over the year, but SAW overpredicts in June and July by 3 to 5%.



(a) (b)

Fig. 8: Same as Fig. 6 but for ISIS data at Hanford, CA. For clear conditions (b) only data with at least 5 points in a month - ToD bin are displayed.

Averaged $MBE_{MT,rel}$ of the costal stations and the inland stations are shown in Fig. 9 & 10 respectively. Since the trends for clear periods are generally similar as for the entire data, the clear data are not broken out separately from here-on-out. The coastal and inland regions were also examined subdivided into northern and southern sub-regions, but no significant north-south difference was observed (not shown). The rMBE is larger at coastal stations with 5% on average. Largest rMBE is still observed in May and June, but the difference to the other months is less pronounced than in the California-wide data (Fig. 6). The inland stations are more numerous and consequently the magnitude and trends are very similar to the overall data with an rMBE of 3.3%.

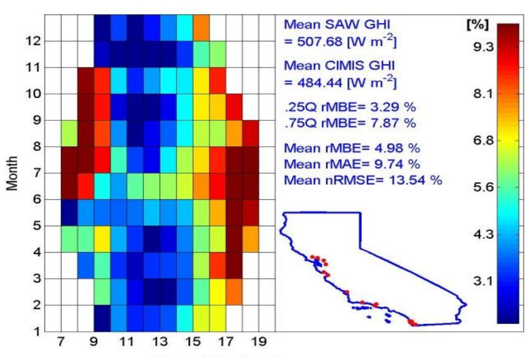


Fig. 9: Same as Fig. 6 but for coastal stations only (14 stations total).

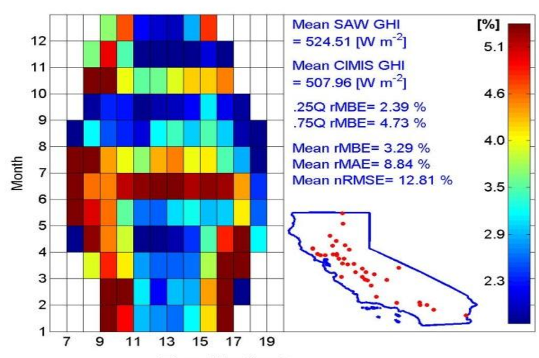


Fig. 10: Same as Fig. 6 but for inland stations only (38 stations total).

In Fig. 11, $rMBE_{MT}$ of 2 typical CIMIS stations, one coastal and one inland, are shown, which are similar to corresponding average $rMBE_{MT}$ in Figs. 9 & 10. Also, $MBE_{MT,rel}$ in Hanford, CA (Fig. 8) is similar to average $MBE_{MT,rel}$ for the inland CIMIS stations in Fig. 10.

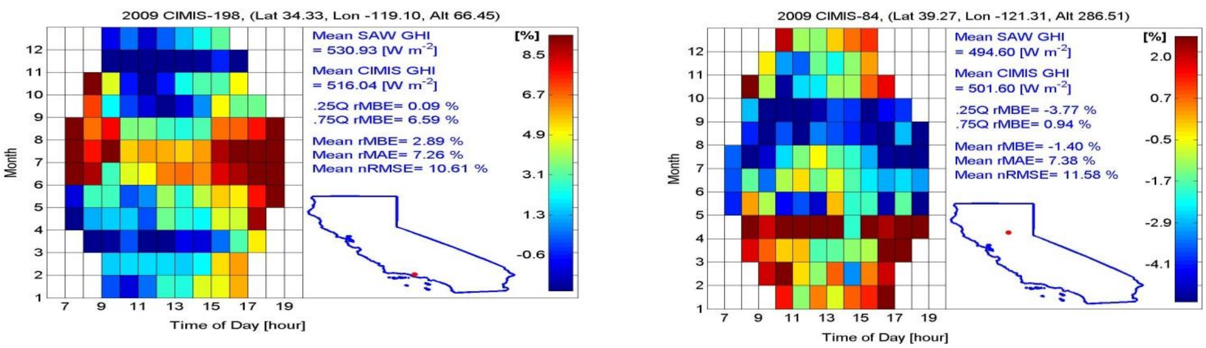
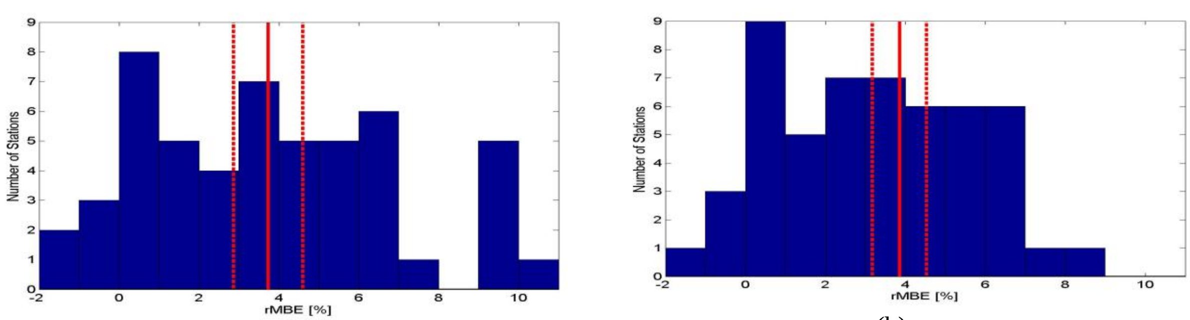


Fig. 11: Same as Fig. 6 but for CIMIS sites 198 and 84 only.

3.3. Overall MBEs and confidence intervals

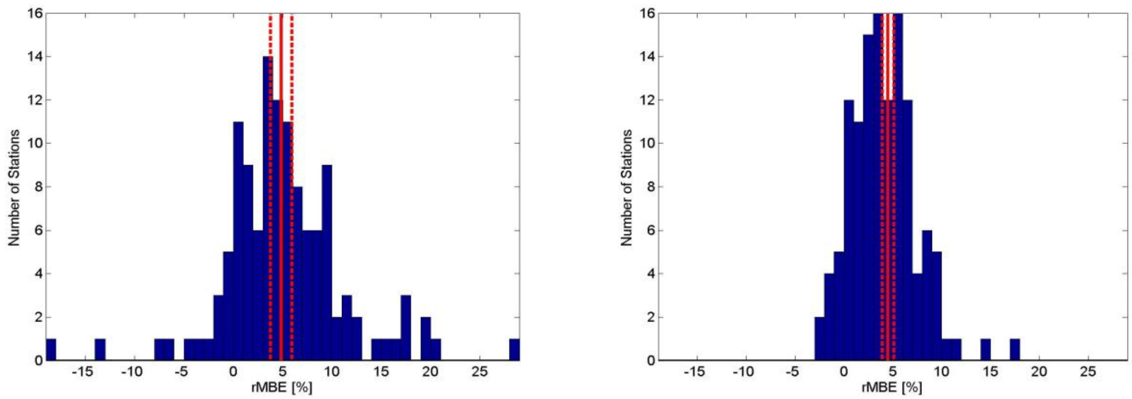
MBE_{year}, $rMBE_{year}$, MBE_{clear} and $rMBE_{clear}$ (averaged over the 52 high quality CIMIS stations) are 18.07 W m^{-2}, 3.74%, 24.38 W m^{-2}, and 3.86% with the corresponding confidence intervals of +-4.15 W m^{-2}, +-0.87%, 4.92 W m^{-2}, and 0.68%, respectively. Figure 12 shows histograms of $rMBE_{year}$ and $rMBE_{clear}$ for all the 52 CIMIS stations, which confirms that the $rMBE_{year}$ and $rMBE_{clear}$ for most of the CIMIS stations are close to the respective averaged values.

MBE_{year}, $rMBE_{year}$, MBE_{clear} and $rMBE_{clear}$ (averaged over all 124 quality-controlled CIMIS stations) are 23.68 W m^{-2}, 4.87%, 27.83 W m^{-2}, and 4.52% with the corresponding confidence intervals of +-5.25 W m^{-2}, +-1.08%, 4.44 W m^{-2}, and 0.59%, respectively. The errors and confidence intervals are slightly larger than the ones for 52 stations. Figure 13 shows histograms of $rMBE_{year}$ and $rMBE_{clear}$ for all the 124 CIMIS stations.



(a) (b)

Fig. 12: Histogram of (a) $rMBE_{year}$ and (b) $rMBE_{clear}$ for 52 high quality CIMIS stations. The red line shows the averaged respective rMBE and dashed red lines show the confidence interval for the 95% confidence level.



(a) (b)

Fig. 13: Same as Fig. 12 but for all 124 CIMIS stations.

4. Conclusions

Validations of the new state-of-the-art solar resource model for California (SolarAnywhere, SAW) were conducted using ground measurements. SAW is unbiased compared to the Hanford ISIS data (not surprising since the irradiation versus cloud index relationship was calibrated there). SAW overestimates the measured GHI at CIMIS stations by 18.07 +- 4.15 W m^{-2} or 3.7% +- 0.9% (95% confidence interval), on average. SAW is also biased large in clear conditions compared to the Ineichen / SoDa clear sky model and the CIMIS measurements.

Despite careful quality control by the authors, CIMIS stations have inferior sensors and are generally less well maintained than high quality solar resource sites such as ISIS. That may suggest that the differences between SAW & CIMIS are at least partially related to CIMIS measurement errors. Especially soiling of the ground sites due to infrequent cleaning likely explains some of the bias. However, persistent trends over the year likely indicate some underlying bias in SAW. Also for PV performance applications one could argue that soiling of PV panels will be even larger than for CIMIS sensors, so CIMIS measurements may be more reflective of performance of solar power plants. From our analysis the following recommendations emerge:

- The relative mean bias error ($rMBE_{MT}$, averaged over all 52 stations, Fig. 6) is largest in May through July. This is the most significant finding as it holds both for clear data and all data, 2009 and 2010, and for CIMIS and ISIS. The cause of this difference, however, is unclear.

- The SoDa turbidity climatology appears to be more accurate than the METSTAT turbidity database used in SAW, but the average observed CIMIS clear sky data lie in between (Fig. 3).
- $rMBE_{MT,}$ of the coastal stations is slightly larger than for inland stations. For the coastal stations the largest differences occurred in both morning (Jun.-Nov.) and evening (Mar.-Jul.). The morning differences could be related to marine layer clouds, while the evening differences show correlation to the clear sky model. However, overall the coastal differences are much smaller than those found for the National Solar Radiation Database [9] indicating an improvement in SAW compared to previous versions.

Random differences were not a focus of this study. For completeness we report that typical rMAEs (for hourly data and normalized by average irradiation) were 9% and typical RMSEs were 13%. Overall the SAW solar resource data are very accurate both in bias and random error.

5. References

1. Web-based Clean Power Research service database, SolarAnywhere. Available: https://www.solaranywhere.com/Public/About.aspx
2. Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F., "A new operational model for satellite-derived irradiances: Description and validation," Solar Energy, 73(5):307-317, 2002.
3. Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., and Hoff, T., "Validation of short and medium term operational solar radiation forecasts in the US," Solar Energy, 84(12):2161-2172, 2010.
4. CIMIS Sensor Specifications. Accessed on Mar 15 2011 at: http://wwwcimis.water.ca.gov/cimis/infoStnSensorSpec.jsp
5. Integrated Surface Irradiance Study (ISIS) Network. Accessed on Mar 15 2012 at: http://www.srrb.noaa.gov/isis
6. Solar Radiation Database (SoDa). Available: http://www.soda-is.com
7. R. Perez, P. Ineichen, K. Moore, M. Kmiecik, C. Chain, R. George, and F. Vignola, "A New Operational Satellite-to-Irradiance Model," Solar Energy, 73:307-317, 2002.
8. P. Ineichen, "Comparison of eight clear sky broadband models against 16 independent data banks," Solar Energy, 80:468–478, 2006.
9. A. Nottrott and J. Kleissl, "Validation of the NSRDB-SUNY global horizontal irradiance in California," Solar Energy, 84(10):1816-1827, 2010.
10. D.W. Meek and J.L. Hatfield, "Data quality checking for single station meteorological databases," Agricultural and Forest Meteorology 69 (1–2), 85–109, 1994.
11. S.O. Eching and D. Moellenberndt, "Technical Elements of CIMIS, the California Irrigation Management Information System: State of California, Resources Agency, Dept. of Water Resources, Division of Planning and Local Assistance," 66, 1998.
12. J. Luoma and J. Kleissl, Summary of Quality Control for 2010 CIMIS Data, Internal Report, 2012.
13. Tom Hoff, Clean Power Research, personal communication.