Future Sustainability
Open Access Journal
ISSN 2995-0473

May 2024| Volume 02 | Issue 02 | Pages 01-07

Journal homepage: https://fupubco.com/fusus

https://doi.org/10.55670/fpll.fusus.2.2.1

Future Publishing LLC

Article

Validation of satellite-derived solar irradiance datasets: a case study in Saudi Arabia

A. F. Almarshoud
Department of EE, College of Engineering, Qassim University, Saudi Arabia

ARTICLE INFO

Article history:
Received 01 December 2023
Received in revised form
31 December 2023
Accepted 07 January 2024

Keywords:
GHI prediction, Solar irradiance prediction,
Solar irradiance estimation,
Satellite-derived irradiance

*Corresponding author
Email address:
dr_almarshoud@qec.edu.sa
dr_almarshoud@QU.edu.sa

DOI: 10.55670/fpll.fusus.2.2.1

ABSTRACT

A robust dataset of Surface Solar Irradiance is essential for secure competitive financing for solar energy projects. Rating agencies and lenders alike require verification of the solar-resource dataset for utilizing each solar energy project, as this can be translated directly into expected electrical energy and revenues. The accuracy of the dataset and the variability of solar radiation, as recorded by historical solar data, play a significant role in estimating the future performance of the project and its budget. The historical observed solar irradiance datasets by local stations are the best and most reliable for a specific site, but they are not always available for long and continuous periods in any location, especially in arid areas. So, the importance of historical solar radiation datasets derived from satellite-based models arises here. This paper validates the historical modeled datasets of the three most famous satellite-based commercial prediction models (SolarGIS, SUNY, and Solcast) against the observed dataset by six ground stations in Saudi Arabia under different climatic zones. The validation method has been implemented using the standard error metrics: Maximum Absolute Error (MAE) and relative Maximum Bias Error (rMBE). The validation process showed that, in the case of GHI, the discrepancy between observed and predicted values is narrow, while in the case of DNI, the discrepancy is wide. Also, the predicted GHI values are more accurate than predicted DNI values, and -in general- the values predicted by the SUNY model are less accurate than those predicted by SolarGIS and Solcast models for both GHI and DNI. The resultant of this validation process could be accepted not for the six locations under study only but, also for deserts and arid areas across Saudi Arabia and might be extended to similar arid areas around the world.

1. Introduction

The financing of large solar projects requires detailed diligence and allocation of technical and commercial risks; one of the principal risks is the knowledge of solar resources. The intermittent nature of the solar resource made its assessment essential for determining the performance of solar-power projects and securing financing for them in the long term. A robust dataset Surface Solar Irradiance (SSI) is essential for secure competitive financing for solar energy projects. Generally, financing communities consider solar resources to be stable on an annual basis when compared to other renewable energy resources. Therefore, rating agencies and lenders alike require verification of the solar-resource dataset for utilizing each solar energy project, as this can be translated directly into expected electrical energy and revenues. The accuracy of the dataset and the variability of the solar radiation, as recorded by historical solar data, play significant roles in estimating the future performance of the project and its budget. Information concluded from historical solar resource datasets may be used to make energy policy decisions, design solar energy systems for specific locations, choose optimum energy conversion technology, and operate installed solar energy projects. Historical solar resource datasets may be the result of local measurement stations, satellite-based estimation methods, or numerical weather prediction methods. There is no doubt that the observed solar irradiance datasets by local stations are the best and most reliable for a specific site, but they are not always available for long and continuous periods in any location, especially arid areas. So, the importance of historical solar radiation datasets derived from satellite-based models arises here. In the last decade, the King Abdullah City for Atomic and Renewable Energy [1], as the lead Saudi Arabia governmental agency for renewable energy, has developed the Renewable Resource Monitoring and Mapping (RRMM) Solar Measurement Network, which currently involves more than 50 metrological stations distributed over the area of Saudi Arabia. These metrological stations record the solar irradiance data in addition to other weather data; the details of the RRMM network are summarized in [2]. The RRMM network can be accessed via the Saudi Arabia Renewable Resource Atlas website [1]. The RRMM currently has a historical dataset that reaches up to 7 years for most locations with high resolution. But these metrological stations are concentrated in the main cities and towns over the wide area of Saudi Arabia that reaches up to 2M km^2; this means that the solar energy projects that are planned to be constructed in desert and arid areas are still suffering from a lack of reliable data of solar irradiance. So, the historical solar radiation datasets derived from satellite-based models are the only way to assess the performance of solar energy projects in arid areas. In this work, the researchers will validate three historically modeled datasets against ground measurements under different climatic zones in Saudi Arabia; if the validation result becomes acceptable, then we can depend on the modeled irradiance datasets for deserts and arid areas over Saudi Arabia; this is the aim of this study.

2. Evolving of satellite-based solar irradiance estimation

Estimating Surface Solar Irradiance (SSI) from satellites began in the 1960s with considerable errors in predicted data, but after 2000, satellite-based estimation of SSI had become increasingly mature, and many sensors have been employed. In addition to advances in sensors, many more sophisticated algorithms have been developed that take into account detailed radiative transfer processes. By these algorithms, multi-channel satellite observations are often combined to quantitatively determine the states of the surface and atmosphere. Satellite data from several sources are also combined to compensate for gaps [3-8]. The algorithms to estimate SSI from satellite data may be classified into two categories: Methods based on radiative transfer processes, which involve the acquisition of atmospheric spectral properties, and statistical methods that depend on the top of atmosphere reflectance observed by satellites, which is usually proportional to the cloud transmission such as Heliosat method [9-11] and Perez' method [12]. Statistical methods are more frequently used to estimate SSI on longer timescales such as daily, monthly, or yearly. In general, statistical methods have good accuracy due to their tuning property. To date, mature high-resolution SSI datasets with global coverage are still rare except for specific regions or are available on the basis of one specific satellite like the Geostationary Operational Environmental Satellite [13- 16]. The accuracy of SSI substantially improves as the timescale increases, the accuracy of hourly SSI is always slightly better than of instantaneous SSI, and the monthly SSI is more accurate than daily SSI when compared with ground observations. Also, the accuracy of SSI in clear-sky cases is better than cloudy-sky cases. Moreover, there are a few commercial high-resolution SSI datasets with near-global coverage. For example, the SolarGIS/SGIS (https://solargis.com), the SUNY/SolarAnywhere (https://data.solaranywhere.com), and Solcast (http://solcast.com). These commercial datasets are considered semi-empirical algorithms that typically include two operational models: the Clear-sky and Cloud-sky models. The early embodiment of semi-empirical models is referred to as the contribution of Cano et al. [9], which evolved over the years into the Heliosat model series [17-20]. These commercial datasets follow the same principle but differ in the source of data and in the fine details of operation models, so a difference in the accuracy of output data is expected. for more details about these commercial datasets [8, 21].

3. Validation methodology

For the purpose of assessing the risk of solar-resource to support the project due diligence and financing, the most critical consideration is the total annual solar energy available in the location, which is typically characterized as total insolation (kWh/m^2) or as daily average (kWh/m^2/day) [21]. The type of solar irradiance to be estimated depends on the technology used for energy production. For concentrating solar power systems (CSP) or concentrating photovoltaic (CPV), the direct normal irradiance (DNI) must be estimated, while for the non-concentrating systems (PV), primarily global horizontal irradiance (GHI) must be estimated [21]. So, in this study, the historical dataset of (GHI & DNI) will only be validated on a monthly basis as daily-average irradiation (kWh/m^2/day). For this purpose, the historical solar datasets of three commercial models (SolarGIS, SUNY, and Solcast) have been selected to be validated with historical datasets of six ground observation stations. These stations cover several distinct climatic environments ranging from arid (Buraydah) in the middle to humid (Jeddah) on the west coast and (Dhahran) on the east coast to cold (Tabuk) in the northern, to warm (Najran) in the southern to (Taif) in top of mountains. The validation period spans about five years, from June 1, 2013, to August 1, 2018. Table 1 illustrates the characteristics of the observation stations, and Figure 1 shows their locations.

Figure 1. Locations of selected observation stations

Table 1. Characteristics of selected observation locations

Location	Latitude	Longitude	Elevation	Avg. Temp. °C	Annual GHI kWh/m²	Annual DNI kWh/m²
Tabuk	28.38° N	36.48° E	781 m	23	2285	2617
Jeddah	21.49° N	39.24° E	76 m	30.7	2117	1774
Taif	21.43° N	40.49° E	1518 m	23.7	2310	2288
Najran	17.63° N	44.54° E	1187 m	26.7	2449	2263
Buraydah	26.34° N	43.76° E	688 m	26.5	2219	2055
Dhahran	26.30° N	50.14° E	75 m	27.9	2037	1847

4. Validation metrics

Various metrics have been proposed in the literature to quantify the accuracy of solar irradiance forecasts. Overall bias and dispersion are the criteria that have been used to gauge the accuracy of solar irradiance models. The metrics recommended to quantify these criteria are the mean bias error (MBE) and its relative (rMBE) for quantifying the overall bias, and the root mean square error (RMSE) and the mean absolute error (MAE) for quantifying the dispersion [22, 23]. Many researchers prefer the mean absolute error (MAE) over the RMSE as a measure of dispersion because it is less sensitive to distant outliers and less subject to interpretation when expressed as a percentage [21, 24]. So, in this study, the two error metrics (MAE & rMBE) have been selected for the validation process and defined as [25]:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|\hat{y}_t - y_t| \qquad (1)$$

$$MBE = \frac{1}{N}\sum_{t=1}^{N}\hat{y}_t - y_t \qquad (2)$$

Where $\hat{y}_t$ and y_t are predictions and observations at time step t, respectively, and N is the total number of samples. then,

$$rMBE = 100\frac{MBE}{\bar{y}} \qquad (3)$$

Where MBE is normalized by the mean of observations:

$$\bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_t \qquad (4)$$

5. Results and Discussion

The Maximum Absolute Error (MAE) for the GHI and DNI has been calculated using formula (1), and the results are represented in Figures 2 & 3 for all regions under study; it is clear from these two figures that the SUNY model has the highest MAE value in most regions under study, while the SGIS and Solcast are the less. Figures from (4) – (15) show the variation of observed values of GHI & DNI and predicted ones of the three estimation models (SIGS, SUNY, and Solcast) in addition to the percentage of their relative Main Bias Error rMBE for the six regions under study using formula (3). In Tabuk City, Figure 4 shows no significant difference between observed and predicted values of GHI for the three estimation models; this conclusion is clear from the small values of rMBE in most months. Also, it is clear from the small values of MAE shown in Figure 2. While in the case of DNI, as shown in Figure 5, there are some differences between observed and predicted values in a few months, especially in the case of the SUNY model, where its rMBE may reach up to 16%. The values of the Solcast model satisfy the best matching with observed values where its rMBE does not exceed 8%; this result is confirmed by the value of MAE, as shown in Figure 3.

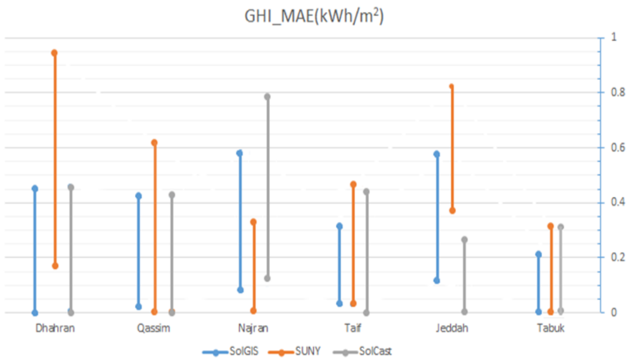

Figure 2. Maximum absolute error of GHI for locations under study

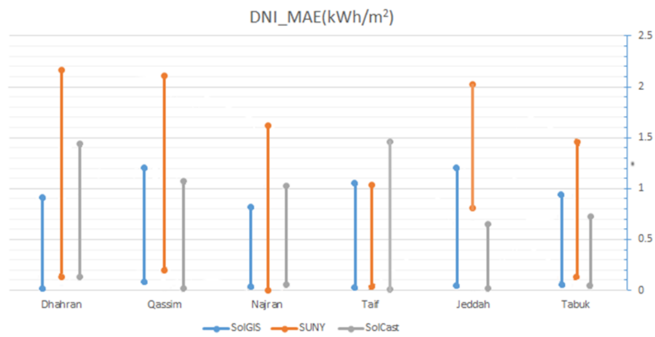

Figure 3. Maximum absolute error of DNI for locations under study

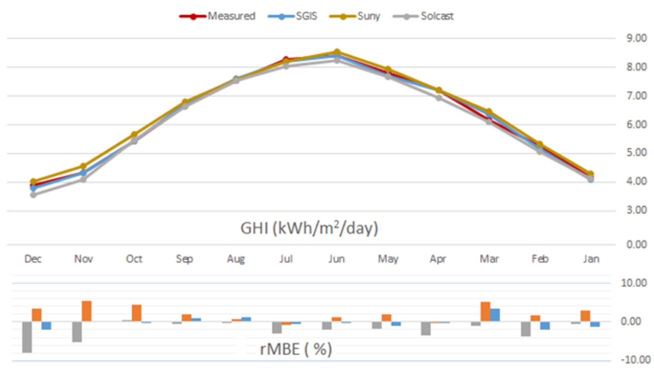

Figure 4. The relative maximum bias error of GHI for Tabuk City

In Jeddah city, Figure 6 shows no significant difference between observed and predicted values of GHI for the Solcast model, but for SUNY and SGIS models, the rMBE reached up to 14% and 9%, respectively. This big discrepancy between observed and predicted values in the case of SUNY and SGIS models is confirmed by their MAE shown in Figure 2. In the case of DNI, as shown in Figure 7, SGIS and Solcast models are in the best matching with observed values for most of the year, but SUNY model values are very far, with rMBE reaching up to 45%; this result is confirmed by the value of MAE as shown in Figure 3 (Taif City: Figure 8 and Figure 9).

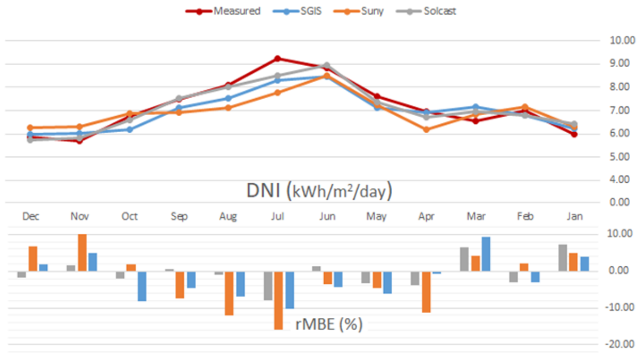

Figure 5. The relative maximum bias error of DNI for Tabuk City

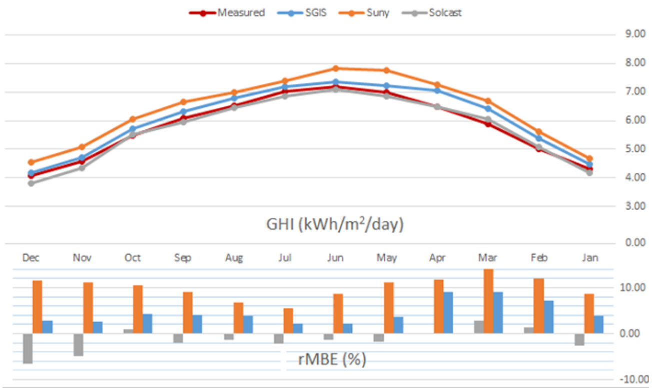

Figure 6. The relative maximum bias error of GHI for Jeddah City

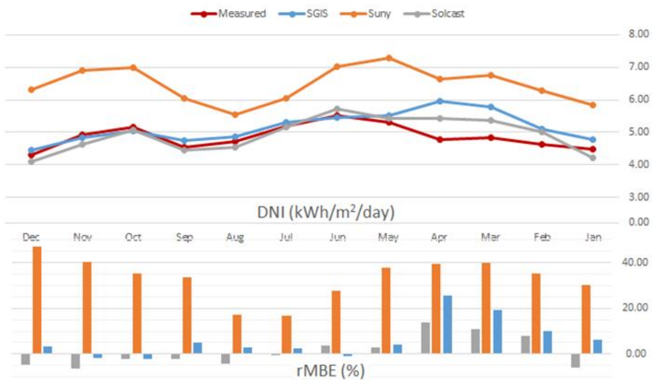

Figure 7. The relative maximum bias error of DNI for Jeddah City

In Najran city, Figure 10 shows no significant difference between observed and predicted values of GHI for all models most of the year in general; the best matching is done by the SUNY model, while the maximum rMBE is done by the Solcast model. The MAE values in Figure 2 show the superiority of the SUNY model in this case. In the case of DNI, as shown in Figure 11, SGIS and Solcast are close to observed values most of the year, with rMBE reaching up to 16%, but the SUNY model did the best matching for half of the year and in another half did the worst with rMBE reaches to 38%, this result is confirmed by its MAE value which reaches up to 1.6 kWh as shown in Figure 3.

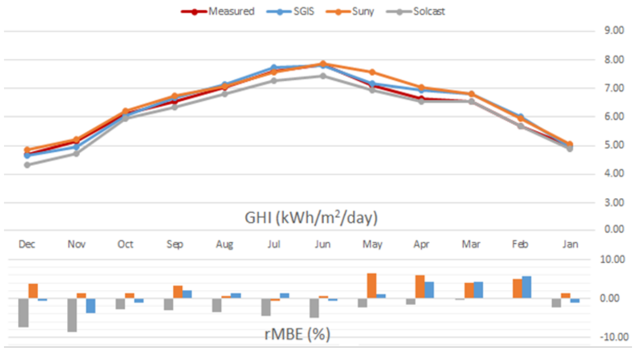

Figure 8. The relative maximum bias error of GHI for Taif City

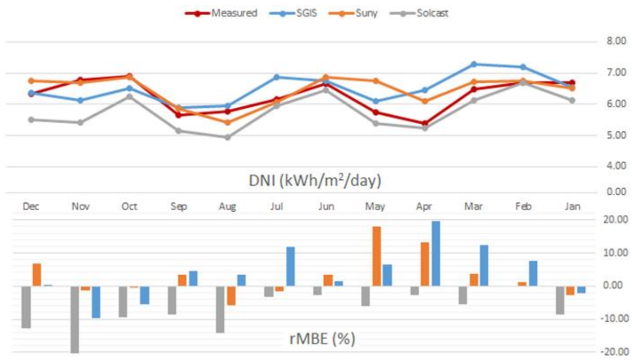

Figure 9. The relative maximum bias error of DNI for Taif City

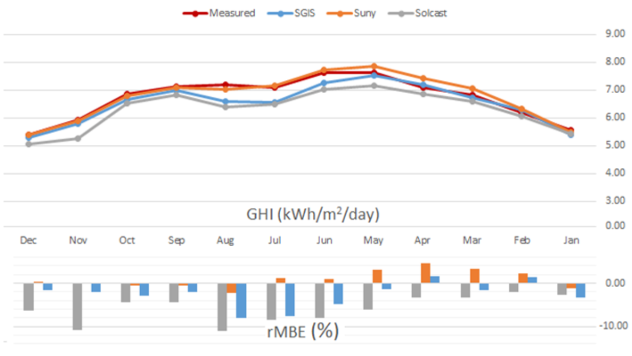

Figure 10. The relative maximum bias error of GHI for Najran City

In Buraydah City, Figure 12 shows no significant difference between observed and predicted values of GHI all the year for SGIS and Solcast models and most of the year for SUNY models where the rMBE is less than 10% for all models. The MAE values in Figure 2 confirm this conclusion. In the case of DNI, as shown in Figure 13, the rMBE is less than 15% for all models in the second half of the year, while in the first half, it reaches high values: 45% for SUNY, 28% for SGIS and 25% for Solcast. Figure 13 shows the superiority of SGIS and Solcast models over the SUNY model in this case.

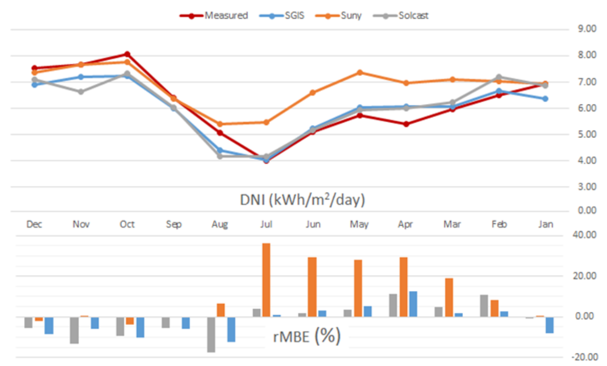

Figure 11. The relative Maximum Bias Error of DNI for Najran City

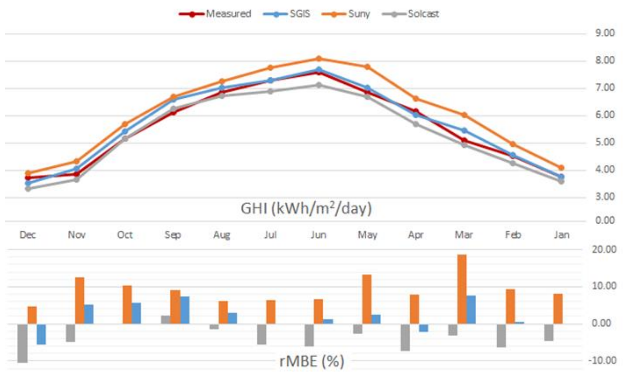

Figure 14. The relative maximum bias error of GHI for Dhahran City

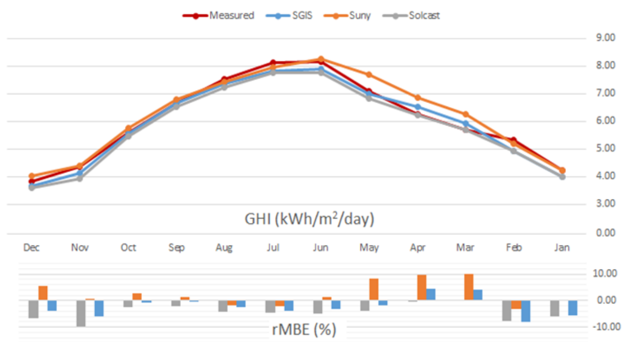

Figure 12. The relative maximum bias error of GHI for Buraydah City

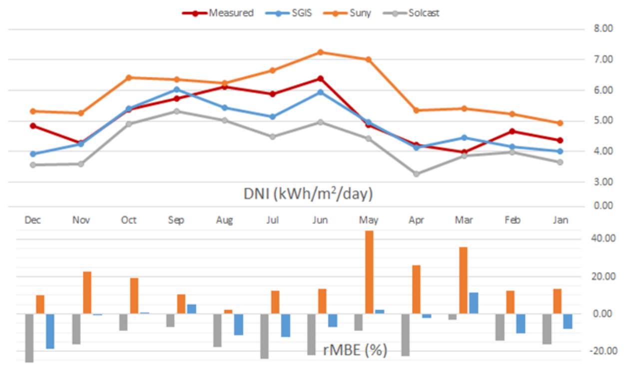

Figure 15. The relative maximum bias error of DNI for Dhahran City

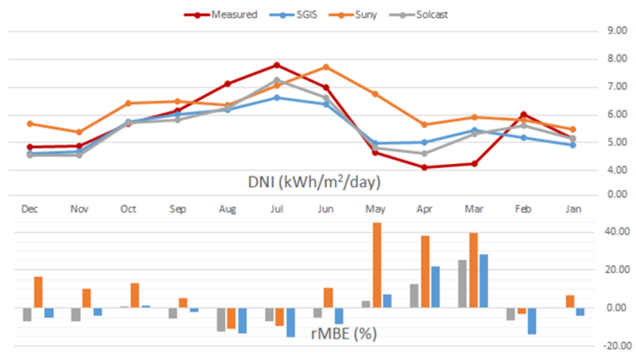

Figure 13. The relative maximum bias error of DNI for Buraydah City

In Dhahran City, Figure 14 shows the best matching was done by SGIS with rMBE less than 8%, then Solcast with rMBE less than 11%, while SUNY did the worst matching with rMBE which reaches up to 18%, the MAE in Figure 2 confirms the superiority of SGIS and Solcast in this case. Regarding DNI, Figure 15 shows a high discrepancy for all models; the rMBE reaches up to 45% for SUNY, 28% for Solcast, and 18% for SGIS; this result is confirmed by their MAE values in Figure 3, which reaches up to 2.2 kWh for SUNY model.

6. Conclusions

The historical predicted solar irradiance datasets for the last five years of six locations across Saudi Arabia have been validated against the observed dataset by local stations for the same period. The validation process has been implemented using the standard error metrics: Maximum Absolute Error (MAE) and relative Maximum Bias Error (rMBE). The predicted data sets have been collected from the most famous commercial solar irradiance datasets: SolarGis, SUNY, and Solcast. The validation process showed that, In the case of GHI, the discrepancy between observed and predicted values is narrow, its rMBE is in the range of 10%, and its MAE is less than 0.6 kWh, especially for SolarGIS and Solcast models, while in the case of DNI, the discrepancy between observed and predicted values is wide, especially in case of SUNY model which its rMBE reaches up to 45%, and its MAE reaches up to 2.2 kWh. So, the GHI-predicted values are more accurate than the DNI-predicted values, and the values predicted by the SUNY model are less accurate than those predicted by SolarGIS and Solcast models for both GHI and DNI. It is clear from these results that GHI values, especially those predicted by SolarGIS and Solcast, can be accepted not only for locations under study but also for deserts and arid areas across Saudi Arabia. Is this conclusion valid for the whole globe? We think it is valid for deserts that are less cloudy between the Cancer and Capricorn lines.

Acknowledgment
Researchers would like to thank all the entities who provided us with data used in this research: the King Abdullah City for Atomic and Renewable Energy K.A. CARE (http:// rratlas.energy.gov.sa), SolarGIS (https://solargis.com), SUNY (https://data.solaranywhere.com), and Solcast (http://solcast.com).

Ethical issue
The author is aware of and comply with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The author adheres to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement
The manuscript contains all the data. However, more data will be available upon request from the author.

Conflict of interest
The author declares no potential conflict of interest.

References
[1] K.A.CARE. Renewable Resource Atlas, King Abdullah City for Atomic and Renewable Energy (K.A.CARE), Saudi Arabia, ⟨http:// rratlas.energy.gov.sa⟩.
[2] Zell Erica, Gasim Sami, Wilcox Stephen, Katamoura Suzan, Stoffel Thomas, Shibli Husain, Engel-Cox Jill, Subie Madi Al. 2015. Assessment of solar radiation resources in Saudi Arabia. Sol Energy 119:422–38. http://dx.doi.org/ 10.1016/j.solener.2015.06.031.
[3] Huang, G.H., Li, X., Ma, M.G., Li, H.Y., Huang, C.L., 2016. High resolution surface radiation products for studies of regional energy, hydrologic and ecological processes over Heihe river basin, northwest China. Agric. For. Meteorol 230:67–78.
[4] Huang, G.H., Ma, M.G., Liang, S.L., Liu, S.M., Li, X., 2011. A LUT-based approach to estimate surface solar irradiance by combining MODIS and MTSAT data. J. Geophys. Res.-Atmos. 116.
[5] Kim, H.Y., Liang, S.L., 2010. Development of a hybrid method for estimating land surface shortwave net radiation from MODIS data. Remote Sens. Environ 114:2393–2402.
[6] Tang, W.J., Qin, J., Yang, K., Liu, S.M., Lu, N., Niu, X.L., 2016. Retrieving high-resolution surface solar radiation with cloud parameters derived by combining MODIS and MTSAT data. Atmos. Chem. Phys 16:2543–2557.
[7] Zhang, Y., He, T., Liang, S.L., Wang, D.D., 2018. Estimation of all-sky instantaneous surface incident shortwave radiation from Moderate Resolution Imaging Spectroradiometer data using optimization method. Remote Sens. Environ 209:468–479.
[8] Huang, G., Li, Z., Li, X., Liang, S., Yang, K., Wang, D., and Zhang, Y. 2019. Estimating surface solar irradiance from satellites: Past, present, and future perspectives, Remote Sens. Environ 233: 111371. https://doi.org/10.1016/j.rse.2019.111371
[9] Cano, D., Monget, J.M., Albuisson, M., Guillard, H., Regas, N., Wald, L., 1986. A method for the determination of the global solar-radiation from meteorological satellite data. Sol. Energy 37:31–39.
[10] Qu, Z., Oumbe, A., Blanc, P., Lefevre, M., Wald, L. 2012. A new method for assessing surface solar irradiance: Heliosat-4. In: Geophysical Research Abstracts.
[11] Rigollier, C., Lefevre, M., Wald, L. 2004. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Sol. Energy 77:159–169.
[12] Perez, R., Schlemmer, J., Hemker, K., Kivalov, S., Kankiewicz, A., Gueymard, C. 2015. Satellite-to-irradiance modeling - a new version of the SUNY model. In: IEEE 42nd Photovoltaic Specialist Conference.
[13] Bisht, G., Venturini, V., Islam, S., Jiang, L. 2005. Estimation of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days. Remote Sens. Environ. 97:52–67.
[14] Hollmann, R., Mueller, R.W., Gratzki, A., 2006. CM-SAF surface radiation budget: first results with AVHRR data. Atmospheric Remote Sensing: Earth's Surface, Troposphere, Stratosphere and Mesosphere - Ii 37:2166–2171.
[15] Deneke, H.M., Feijt, A.J., Roebeling, R.A., 2008. Estimating surface solar irradiance from METEOSAT SEVIRI-derived cloud properties. Remote Sens. Environ 112:3131–3141.
[16] Forman, B.A., Margulis, S.A., 2009. High-resolution satellite-based cloud-coupled esti- mates of total downwelling surface radiation for hydrologic modelling applications. Hydrol. Earth Syst. Sci. 13:969–986.
[17] Beyer, H.G., Costanzo, C., Heinemann, D., 1996. Modifications of the Heliosat procedure for irradiance estimates from satellite images. Sol. Energy 56:207–212.
[18] Schillings, C., Mannstein, H. and Meyer, R., 2004. Operational method for deriving high resolution direct normal irradiance from satellite data. Solar Energy 76:475-484.
[19] Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. Sol. Energy 73: 307–317.
[20] Cebecauer, T. and Suri, M., 2010. Accuracy improvements of satellite-derived solar resource based on GEMS re-nnalysis aerosols. Proceedings of: SolarPACES 2010 Conf., Perpignan, France.
[21] Jan Kleissl. 2013. Solar Energy Forecasting and Resource Assessment. Elsevier Inc, 1st edition.
[22] Espinar, B., Ramı ́ rez, L., Drews, A., Beyer, H.G., Zarzalejo, L., Polo, J., Martı ́ n, L. 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. Solar Energy 83 (1):118–125.
[23] Meyer, R., Gueymard, C., Ineichen, P. 2011. Proceedings of SolarPACES Conference. Stan-dardizing and benchmarking of modeled DNI data products. Granada.
[24] Hoff, T.E., Perez, R. 2012. Predicting Short-Term Variability of High-Penetration PV. Proc. World

Renewable Energy Forum (ASES Annual Conference), May, Denver, CO.
[25] Jamie M. Bright, 2019. Solcast: Validation of a satellite-derived solar irradiance dataset. Solar Energy, Vol (189), 435-449. https://doi.org/10.1016/j.solener.2019.07.086