VAISALA
DIGITAL

White Paper
www.vaisala.com

Vaisala Global Solar Dataset 2019 Release

Methodology and Validation

Rev. 1, May 2020

Introduction

Solar energy production is directly correlated to the amount of radiation received at a project location. Like all weather-driven renewable resources, solar radiation can vary rapidly over time and space, and understanding this variability is crucial in determining the financial viability of a solar energy project.

The three components of irradiance most critical for determining solar installation production values are global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DIF). In this paper we are focused on validating GHI, or the total amount of radiation received by a horizontal surface, which is the primary resource in photovoltaic (PV) installations.

Most financing options for solar projects require information on expected yearly irradiance values as projects typically have to service debt one to four times per year. However, annual averages do not provide enough information to determine accurate annual irradiance and power production values.

Depending on the characteristics of a site, studies have shown that on average, annual irradiance means can differ from the long-term mean by 5% for GHI and by as much as 20% for DNI.[1] Thus, a long-term record of solar irradiance estimates is needed to calculate a realistic variance of production values.

The existing network of surface observation stations is too sparse to quantify solar resources at most potential sites. Also, a vast majority of stations only provide a limited short-term record of the resource (months to a few years), are rarely located near proposed sites, and are often plagued with measurement errors. Calculating site-specific solar irradiance values using geostationary satellite data is an accepted alternative.[2] Within the global atmospheric sciences community, satellite-derived values have proven to be more accurate than nearby surface observations for locations that are more than 25 km away from a ground station.[3]

Through its acquisition of 3TIER, Vaisala is the first organization, either public or private, to map the entire world's renewable resource potential at resolutions of 5 km or higher, providing a global blueprint for wind, solar, and hydro project development. Vaisala was the first to create a high-resolution, global solar dataset using a consistent satellite processing methodology to help clients determine solar variability at any site worldwide, from the prospecting stage through assessment and bankability.

In this paper, we will provide an outline of standard practices that should be followed to ensure accurate solar assessment. We will also describe the methodology Vaisala used to create its continually updated global solar dataset and provide results from an extensive validation study. Validation statistics by region are shown in the Appendix.

Solar Development Roadmap

Developing a solar project requires a large upfront investment. A standard development roadmap conserves time and money and ensures that the most promising projects are constructed. Each stage of development asks different questions about the solar resource and each stage requires varying degrees of information and financial investment.

Prospecting and Planning

The first step in building any solar energy project is identifying the regions most suitable for development. The price of energy, access to transmission, and environmental siting issues should all be taken into consideration, but the most essential variable is the availability of the solar resource — the "fuel" of the project. At this early stage, average annual and monthly solar irradiance values can be used to assess the overall feasibility of a particular site and to select the appropriate solar technology to be installed. Getting time series or typical meteorological year (TMY) data is an even better method, particularly when it is from the same data source you plan to use for financing. Having the same data source throughout the development process helps avoid a number of unpleasant surprises further down the development roadmap. Vaisala's online Solar Prospecting and Time Series Tools allow developers to quickly target the best locations for further investigation and identify red flags early in the process.

Design and Due Diligence

Once a promising site is identified, a more in-depth analysis is required to better quantify the long-term availability of the solar resource, to design technical aspects of the project, and to secure the upfront capital for construction. A common source of solar data used for this purpose is TMY data. A TMY dataset provides a 1-year, hourly record of typical solar irradiance and meteorological values for a specific location in a simple file format. Although not designed to show extremes, TMY datasets are based on a long time period and show seasonal variability and typical climatic conditions at a site. They are often used as an input to estimate average annual energy production.

While TMY data provide a good estimate of the average solar irradiance at a site, they are not a good indicator of conditions over the next year, or even the next 5 years. The U.S. National Renewable Energy Laboratory User Manual for TMY3 data explicitly states, "TMY should not be used to predict weather for a particular period of time, nor are they an appropriate basis for evaluating real-time energy production or efficiencies for building design applications or a solar conversion system."[4] Hourly time series covering a period of several years provide a much more complete record for calculating accurate estimates of solar resource variability.

Year-to-year variability has a significant impact on annual energy production. Many financial and rating institutions, as well as internal certification organizations, require 1-year P90 values to assess the economic feasibility of a project.[5] A 1-year P90 energy value indicates the production value that the annual energy output will exceed 90% of the time. A 1-year P90 value (as opposed to a 10-year P90 value) is typically mandatory because most solar projects have a lending structure that requires them to service debt one to four times a year, not one to four times every 10 years. If power production decreases significantly in a given year due to solar variability, debt on the project may not be able to be paid and the project could default on its loan. This is precisely what financiers are trying to avoid. The only way to determine 1-year P90 values acceptable to funding institutions is with long-term continuous data at the proposed site.

If collected properly, surface observations can provide very accurate measurements of solar radiation at high temporal resolution, but few developers want to wait the 10 years required to develop an accurate 1-year P90 GHI value or even the 5 years necessary for a P50 GHI value. Satellite-derived irradiance values can accurately provide a long-term, hourly time series of data without the expense and wait. However, satellite data cannot always capture the microscale features that affect a site. Therefore, a combination of short-term ground measurements and long-term satellite-derived irradiance values is ideal for assessing variability and project risk.

One method of combining short-term ground measurements with longer-term satellite data is a technique known as model output statistics (MOS). Vaisala pioneered the use of on-site observational data to validate and bias correct satellite-derived irradiance data. Our proprietary MOS technique uses an hourly multi-linear regression equation to remove bias and adjust the variance of the satellite model output to better match the observational data. The MOS equation for each observation station is trained over the observational period of record. The MOS equation is then applied to all time steps of the modeled dataset, so that corrections can be made for periods during which observational data are unavailable.

The value of performing MOS correction is that it captures the unique characteristics of a site through on-site observations and places them into the long-term historical perspective provided by the 3TIER Services modeled data. After validating the technique at many sites globally, Vaisala has determined that the resource model uncertainty can be reduced by 50% using this methodology.

These comprehensive solar resource assessments are used in a Solar Due Diligence Assessment to simulate the hour-by-hour electrical production of a specific, but yet-to-be-built solar generating station. A gold standard due diligence assessment includes a site adapted solar resource study and a net energy assessment. Production estimates are highly complex and involve dozens of specific assumptions and considerable exercise of professional judgment, which Vaisala's specialized and experienced personnel have amassed through assessing more than 46 GW of proposed solar projects globally, including preparing energy estimates for 6 GW.

Operations and Optimization

With more solar energy coming into the grid every day, effectively managing its integration is becoming increasingly important. Once a project is operational, forecasting plays a vital role in estimating hour- and day-ahead solar production and variability. This information is critical for estimating production, scheduling energy, managing a mixed energy portfolio, avoiding imbalance charges, and detecting reduced production days.

Some rudimentary numerical weather prediction (NWP) modeling systems have been introduced for this purpose. However, Vaisala has found that basic NWP models poorly estimate cloud cover, the single variable that most directly impacts solar energy production, and for this reason, has introduced advanced forecasting technologies incorporating machine learning to blend NWP models with observations to allow operators to more accurately schedule solar energy.

Recent solar irradiance observations from satellite-derived datasets or observations from on-site solar measurement stations can also be used to model the energy that a project should have produced based on actual weather conditions. Comparing modeled production with actual production helps identify underperforming projects and explain to what extent solar variation is impacting production. This periodic, ongoing reconciliation helps pinpoint maintenance and equipment issues, particularly for those with a geographically dispersed portfolio of projects.

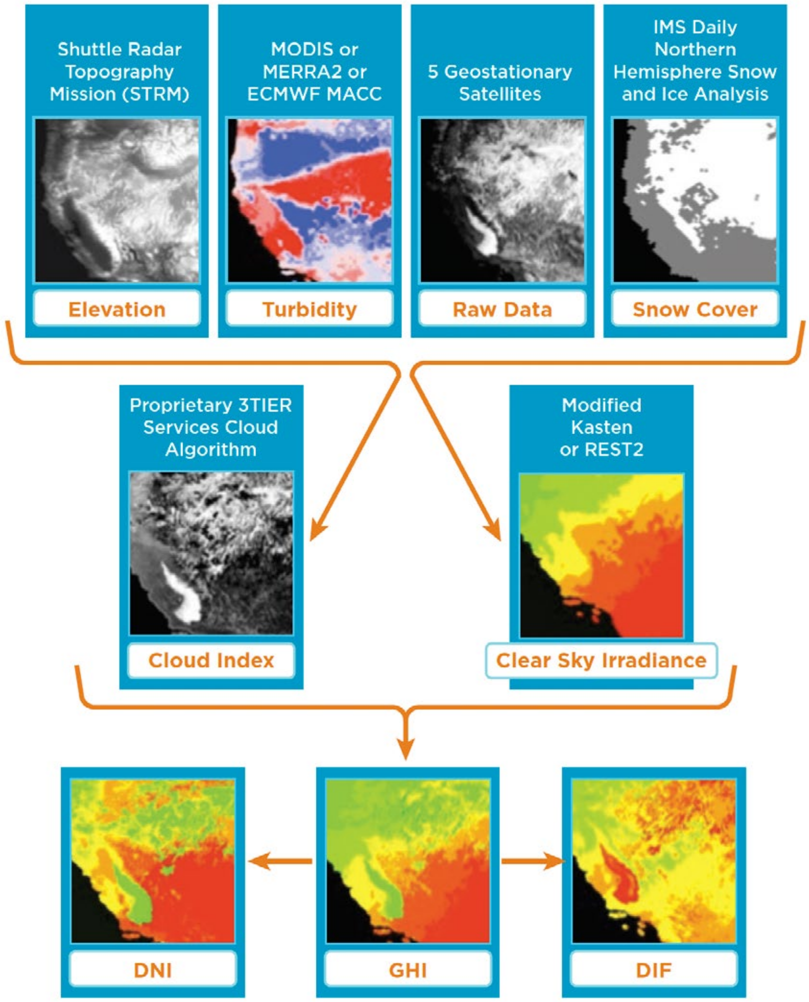
Vaisala's Solar Irradiance Modeling Methodology

Vaisala continues to maintain and improve upon its global, long-term, high resolution solar dataset, which was created using satellite observations from around the world. As discussed earlier in this document, satellite-derived data have proven to be the most accurate method of estimating surface solar irradiance beyond 25 km of a ground station. However, either technology requires special consideration. For example, if there is a dramatic elevation difference between a ground station and a project location, data from the ground station may not be representative of conditions at the project site. Satellite data accuracy can also be influenced by local terrain, such as in locations along coastlines or near dry lake beds.

Vaisala's main source of satellite observations is weather satellites in a geostationary orbit. These satellites have the same orbital period as the Earth's rotation and are thus stationary relative to a point on the earth. As a result, their instruments can make multiple observations of the same area with identical viewing geometry each hour. Vaisala's methodology uses visible satellite imagery to calculate the level of cloudiness at the Earth's surface. The resulting time series of cloudiness (or cloud index) is then combined with other information to model the amount of solar radiation at the Earth's surface. The outcome is an 20+ year dataset that provides hourly and sub-hourly estimates of surface irradiance (GHI, DNI, and DIF) for all of the Earth's land mass at a spatial resolution of approximately 3 km (2 arc minutes).

Vaisala's global solar dataset is based on two decades of half-hourly,

Figure 1. Vaisala's solar modeling methodology



high-resolution visible satellite imagery via the broadband visible wavelength channel. These data have been processed using a combination of peer-reviewed, industry-standard techniques and processing algorithms developed inhouse, including a cloud-index algorithm that produces consistent results when used with the large number of satellites that must be combined to construct a global dataset. With our methodology we currently produce five estimates of irradiance using different algorithms and inputs to provide our clients a full understanding of resource variability.

Despite the resolution of the dataset, some factors need to be taken into consideration by the user. Vaisala's global solar datasets do not directly account for local shades and shadows and, as a result, local conditions must be considered when interpreting the irradiance values. Also, in some areas with highly reflective terrain, such as salt flats and areas with permanent snow, the satellite algorithms have difficulty distinguishing clouds from the terrain. The cloudiness estimates in these areas are higher than they should be. As a result, the amount of GHI and DNI is underestimated and the DIF is overestimated. Known areas affected by this problem include highly reflective areas such as Lake Gairdner National Park in South Australia.

Satellite-based time series of reflected sunlight are used to determine a cloud index time series for every land surface worldwide. A satellite-based daily snow cover dataset is used to aid in distinguishing snow from clouds. In addition, the global horizontal clear sky radiation (GHC), or the amount of radiation in the absence of clouds, is modeled based on the surface elevation of each location, the local time, and the measure of turbidity in the atmosphere.

Vaisala employs two clear sky models. The first clear sky model used is a modified Kasten clear sky model[2] (hereafter referred to as Modified Kasten). The second is the REST2 9.0 model, a parameterized version of Gueymard's SMARTS

radiative transfer model.[6] Once GHC is determined using either the Modified Kasten methodology or the REST2 model, GHI is calculated by combining the cloud index values with the GHC values. In the Modified Kasten method, DNI is calculated from GHI using Perez's DIRINT model outlined in the 2002 paper. In the REST2 model, a modulation function is used to calculate DNI from the clear sky DNI value and the cloud index. For the calculated irradiance components, a calibration function is applied for each satellite region, based on a set of high-quality surface observations. For both models, diffuse is then calculated from GHI, DNI, and solar zenith angle.

Atmospheric turbidity describes the transparency of the atmosphere to solar radiation, and is primarily affected by aerosols and water vapor. Unfortunately, direct observations of turbidity are made at only a few locations. Vaisala ingests several sources of aerosol inputs and uses them in our various models including MODIS Atmosphere Daily Global Product, the ECMWF-MACC (European Centre for Medium-Range Weather Forecasts - Monitoring Atmospheric Composition and Climate) II reanalysis dataset, and MERRA2 (Modern-Era Retrospective analysis for Research and Applications, Version 2) reanalysis dataset. For the Modified Kasten method, turbidity is described by the Linke turbidity coefficient based upon the calculations outlined in Ineichen and Perez, 2002. We combine the data with another turbidity dataset that includes both surface and satellite observations to provide a turbidity measure that spans the period of our satellite dataset and is complete for all land surfaces. In the REST2 models, turbidity is estimated using aerosol optical depth (AOD) and Angstrom exponent, water vapor, and surface pressure taken from either the ECMWF-MACC dataset or the MERRA2 dataset. After testing, default values were chosen for other model input parameters: aerosol single-scattering albedo and asymmetry parameter, ozone concentration, and surface albedo.

Vaisala combines the above inputs to create five different versions of our global solar dataset. In each version the satellite imagery, snow data, topography, and albedo sources are the same. In all versions, the Vaisala proprietary cloud index calculation methodology is also used. The model variations come from different combinations of the clear sky models and turbidity inputs, as shown in Table 1.

In 2019, we released an updated version of the dataset. The Modified Kasten models (1.0-1.2) now ingest data from the latest MODIS aerosol products released by NASA, i.e., Collection 6.1 instead of Collection 5.1. In addition, the latest MODIS values have been applied during the years 2017 and 2018 in place of a static climatology. Reference parameter values used to describe aerosol characteristics have been updated in the REST2 models (2.0 and 2.1). Both sets of changes are intended to improve the representation of aerosols, which strongly affect the transmission of solar radiation through the clear atmosphere.

Table 1. Inputs to each of Vaisala's dataset models

Vaisala Model	Clear Sky Model	Turbidity Input Parameters	Turbidity Input Resolution Temporal / Spatial	Aerosol Source
1.0	Modified Kasten	AOD550 nm	Monthly/1°x1°	MODIS (Terra)
		Precipitable Water (cm)	Monthly/1°x1°	MODIS (Terra)
1.1	Modified Kasten	AOD550 nm	Monthly/1°x1°	MODIS (Terra, Aqua)
		Precipitable Water (cm)	Monthly/1°x1°	MODIS (Terra, Aqua)
1.2	Modified Kasten	AOD550 nm	Daily/1°x1°	MODIS (Terra, Aqua)
		Precipitable Water (cm)	Monthly/1°x1°	MODIS (Terra, Aqua)
2.0	REST2	AOD 380 nm - 1020 nm	3 hourly/0.7°	ECMWF-MACC
		Precipitable Water (cm)	3 hourly/0.7°	ECMWF-MACC
		Surface Pressure (Pa)	3 hourly/0.7°	ECMWF-MACC
2.1	REST2	AOD550 nm	Hourly/0.5° lat x 0.625° lon	MERRA2
		Precipitable Water (cm)	Hourly/0.5° lat x 0.625° lon	MERRA2
		Surface Pressure (Pa)	Hourly/0.5° lat x 0.625° lon	MERRA2
		Angstrom Exponent	Hourly/0.5° lat x 0.625° lon	MERRA2

Notes on the datasets

Vaisala 1.0

The Vaisala 1.0 dataset is the original dataset created by Vaisala (previously known as 3TIER) in 2009. It uses the Modified Kasten clear sky model and monthly average aerosol optical depth (AOD) from the MODIS Dark Target AOD retrieval algorithm.

Vaisala 1.1

The Vaisala 1.1 dataset, released in 2012, is the second dataset based on the Modified Kasten clear sky model. The main change from the Vaisala 1.0 dataset was to incorporate AOD from both Dark Target and Deep Blue MODIS retrieval algorithms.

Vaisala 1.2

Developed in 2014, the Vaisala 1.2 dataset is the third dataset variation using the Modified Kasten approach. The main change over the Vaisala 1.1 dataset was increasing the temporal resolution of MODIS AOD data from monthly averages to daily averages.

Vaisala 2.0

The Vaisala 2.0 dataset is the first dataset Vaisala created using the new REST2 clear sky model developed in 2016 and uses ECMWF-MACC data for the aerosol and water vapor inputs.

Vaisala 2.1

Also developed in 2016, the Vaisala 2.1 dataset is the second dataset Vaisala created using the new REST2 clear sky model. The main difference from the Vaisala 2.0 dataset is the use of MERRA2 for aerosol and water vapor inputs.

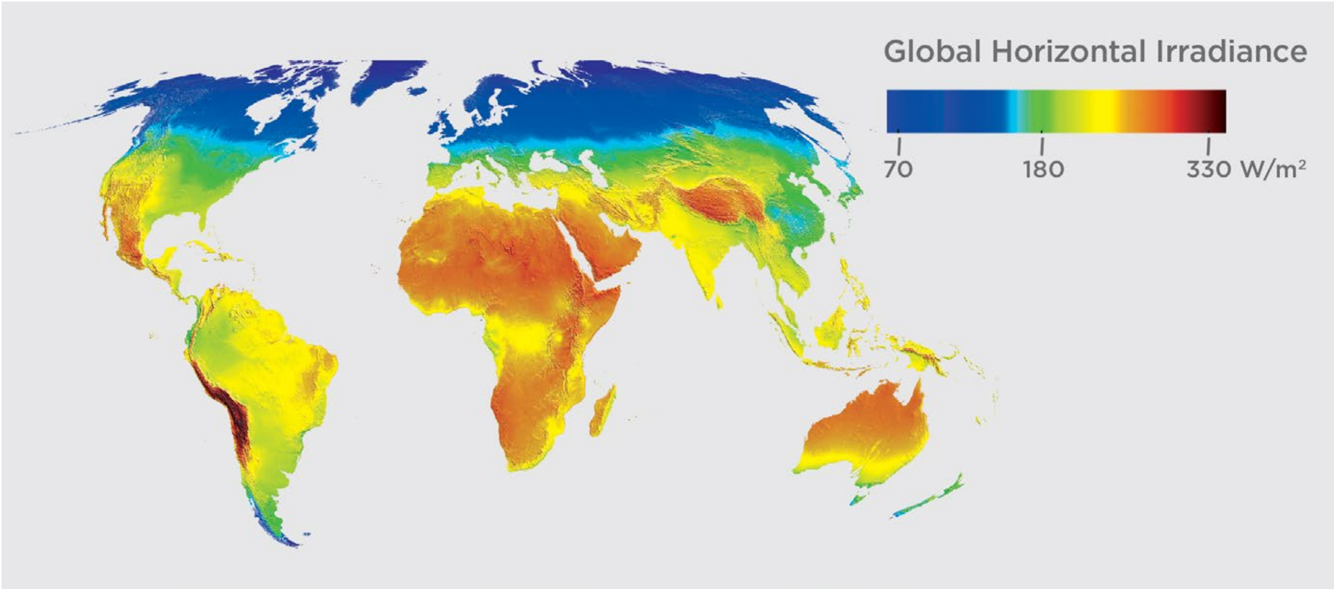
2019 dataset updates

Vaisala 1.0, 1.1, and 1.2

- Incorporate next generation MODIS aerosol product (Collection 6.1 replaces 5.1).
- Replace temporary aerosol optical depth climatology with up-to-date time varying values in most recent years (2017-2018).

Vaisala 2.0 and 2.1

- Refine reference parameter values used to describe aerosol characteristics.



Conclusion

Vaisala actively maintains all five versions of the dataset in order to give our clients a better understanding of local resource variability. In different regions, one version may perform more accurately than another due to local factors, such as pollution or dust, which are better represented by a particular aerosol optical depth product, or the location may have some seasonal irradiance variations that are captured with higher precision by one clear sky model compared to another. If all the models show very similar results, there can be high confidence in the irradiance values. However, sites that show a spread of irradiance values are good candidates for including ground station data in the assessment.

To give project developers higher confidence in our irradiance values and assessment results, Vaisala provides multiple datasets that use trusted underlying processing methodologies that allow clients to compare the results and find the one that best fits local conditions.

How should one choose the best dataset to use? The first step is to review the regional validation results contained in this paper and identify which model performs best in your region of interest. Please contact Vaisala for further details at the validation locations, so you can review results at sites closest to your project location. Secondly, if you have a ground station in the project area, compare the different data options for the concurrent period of time and evaluate which ones most closely match your ground data. Lastly, if you have no ground data to refer to, we don't recommend using the highest or the lowest time series record, but rather a version that is in the middle of the results and has good validation statistics in the region. The final consideration would be the technology employed. If a tracking photovoltaic or concentrating solar plant are under consideration, all other statistics being equal, we would suggest one of the REST2 based models because of the greater accuracy of the DNI irradiance component.

Validation of the Vaisala Global Solar Irradiance Dataset

An extensive validation of Vaisala's solar irradiance dataset was performed using observations from nearly 200 surface stations across the globe. In the study, Vaisala used stations from the World Climate Research Program and the Baseline Surface Radiation Network, national programs such as the Indian Meteorological Department and the Australian Bureau of Meteorology, the National Solar Radiation Database, and several other observational datasets. The various instruments used to measure GHI have different uncertainty estimates on an annual basis. The best equipment has uncertainty of less than 1% at a 95% confidence level, but most equipment deployed for solar project measurements is in the 1.5–2% range and some of the second class equipment deployed has closer to 4–6% uncertainty at the 95% confidence level. The World Climate Research Program estimates solar ground stations can have inaccuracies of 6–12% on the instantaneous irradiance values. Specialized high-quality research sites, such as those from the Baseline Surface Radiation Network, are possibly more accurate by a factor of two.[7] These constraints make direct comparisons between solar radiation datasets difficult, but it is still possible to estimate the relative accuracy if the same reference observations are used. Vaisala did basic quality control of the data from each observation station, and anomalous stations from each network were removed from the comparisons. The statistics presented in the following sections were computed using only daytime irradiance values, which provide a better indication of the accuracy and value of the dataset for use in resource estimation.

Global Validation Statistics

Whenever Vaisala releases a new version of the irradiance dataset, an extensive validation is performed and released publicly. It is extremely important to Vaisala that the integrity of the validation process be unquestionable. To that end, we cultivate an extensive database of public ground station data that is reserved for use exclusively in the validation process and is not allowed to influence the dataset's creation in any way. Additionally, private client data is not allowed to be used in the public validation process except by explicit permission. Our validation results represent the accuracy of our irradiance dataset for a concurrent period of time with independent ground stations not used in the calibration process. Validation of the latest versions of the dataset was carried out in 2019. Results in the tables provided in the Appendices provide a list of statistical metrics. The computed statistics include those most commonly used in the solar industry, such as mean bias error (MBE), mean absolute error (MAE), and hourly root mean square error (RMSE). Mean bias error (MBE) provides information about the average difference in the mean over the entire dataset when compared against observations. Mean absolute error (MAE) measures the average magnitude of the deviation between the ground station and the models. Root mean square error (RMSE) also measures the average magnitude of the deviation, but uses quadratic weighting, which results in large errors carrying more weight. A smaller RMSE value means that the dataset more closely tracks observations on an hour-by-hour basis. Together MBE, MAE, and hourly RMSE can be used to assess the accuracy of a solar dataset compared to observations. Comparison statistics were calculated for GHI based on the overall bias at each location, both regionally and globally. The spatial distribution of GHI bias around the globe is shown in the World GHI Appendix and additional figures are provided in regional appendices. In order to have global representation in the results, GHI data from 196 measurement stations in high quality measurement networks were used in the study. Each site had at least one complete year of measured data.

Globally, Vaisala GHI values show an MBE standard deviation of 4.1% – 4.4% depending on the model (Table A-1). Regionally, the different GHI models show varying results largely tied to the aerosol datasets. The varying accuracy of the aerosol products with geography is one of the reasons we provide multiple options, so that the best data is available locally.

For example, in the North American region, the MBE standard deviations for the Modified Kasten based models (~3.4%) are lower than those for the REST2 based models (4.2% and 4.7%) (Table A-5). However, in East Asia and Oceania, the opposite is true, with the Modified Kasten based model MBE standard deviations being higher (3.8% – 3.9%) than the REST2 based model values (~2.6%) (Table A-3). It should be noted that, in every case, the mean errors are within the standard deviation of the bias of observations, as determined by the World Climate Research Program.

References

[1]Gueymard, C., Wilcox, S., Spatial and Temporal Variability in the Solar Resource. Buffalo, NY: American Solar Energy Society, May 2009. Print.

[2]Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F. "A New Operational Satellite-to-Irradiance Model." Solar Energy 73.5 (2002): 307–317. Print.

[3]Zelenka, A., Perez, R., Seals, R., Renné, D. "Effective Accuracy of Satellite Derived Irradiance." Theoretical and Applied Climatology. 62.3–4 (1999): 199–207. Print.

[4]Marion, W. and Wilcox. S. Users Manual for TMY3 Data sets. Golden: National Renewable Energy Laboratory, 2008.

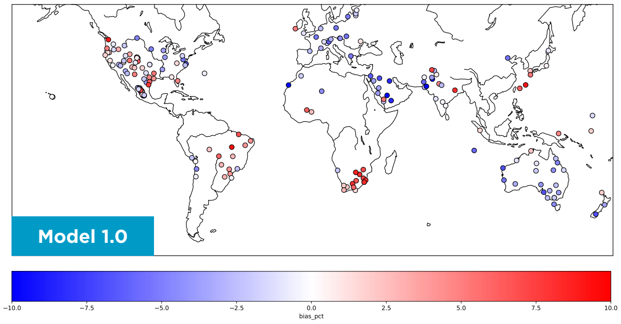
[5]Venkataraman, S., D'Olier-Lees, T. "Key Credit Factors." Standard and Poor's Solar Credit Weekly 29.42 (2009), 29–30, 49-50. Earticle.

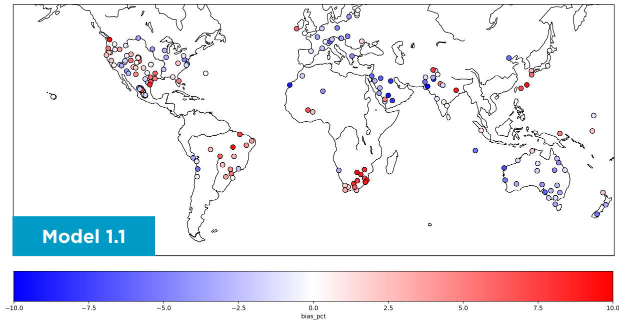
[6]Gueymard, Christian A. REST2: highperformance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation– validation with a benchmark dataset, Solar Energy, vol. 82.3 pp. 272–285, 2008

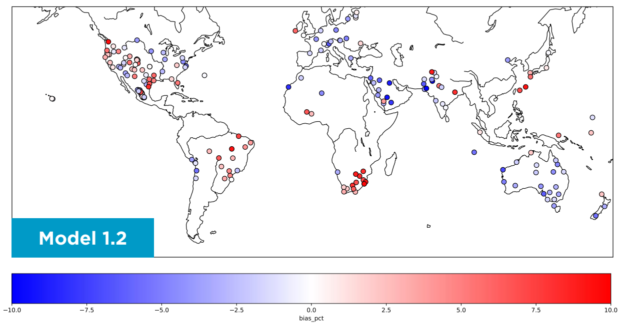
[7]"SSE Release 6.0 Methodology." www.nasa.gov. National Aeronautics and Space Administration. 2010. Downloaded 28 Sept. 2010.

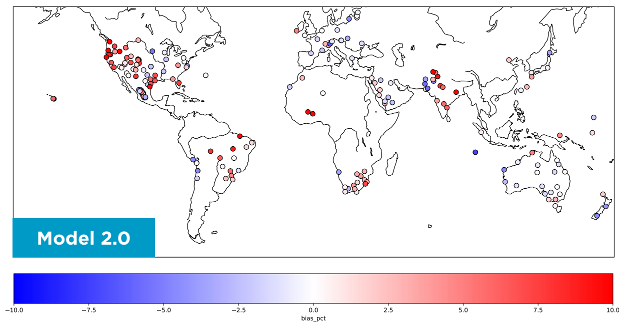
Appendix: Regional Variations

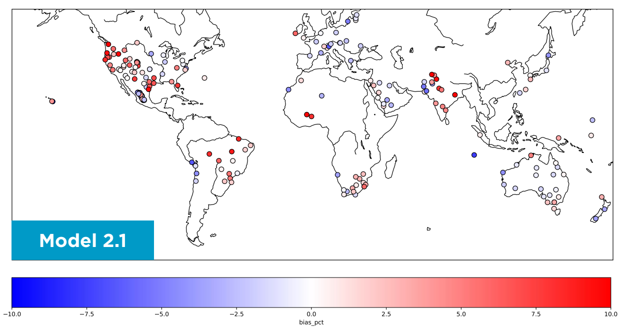
World GHI











Overall Statistics

Table A-1. Worldwide GHI comparison statistics for each of the five Vaisala models. All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	-0.19	4.37	-0.99	20.78	13.63	196
1.1	0.17	4.41	-0.68	20.78	13.62	196
1.2	0.12	4.43	-0.82	20.77	13.61	196
2.0	1.65	4.43	0.85	20.15	12.96	196
2.1	1.20	4.09	0.65	19.94	12.79	196

[1]Mean Bias Error
[2]RMSE = Root Mean Squared Error
[3]MAE = Mean Absolute Error
[4]N = Number of Comparison Locations

Africa and the Middle East GHI

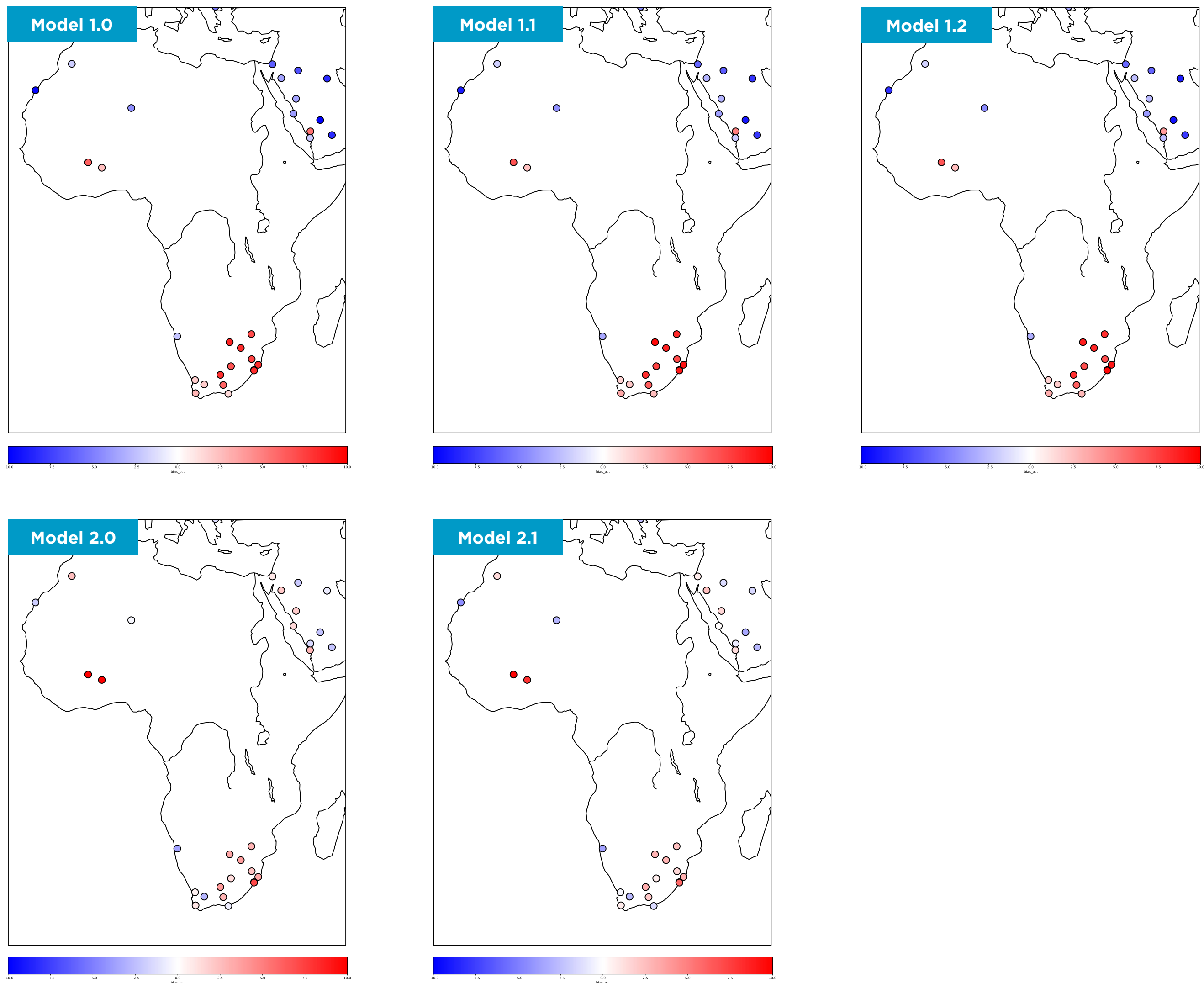


Table A-2. Africa and the Middle East: Regional GHI comparison statistics for each of the five Vaisala models. All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	1.10	6.56	1.94	16.15	10.29	30
1.1	1.47	6.61	2.32	16.22	10.36	30
1.2	1.41	6.61	2.30	16.06	10.20	30
2.0	2.08	4.14	1.75	14.04	8.37	30
2.1	1.23	3.75	1.07	13.79	8.07	30

[1]Mean Bias Error
[2]RMSE = Root Mean Squared Error
[3]MAE = Mean Absolute Error
[4]N = Number of Comparison Locations

East Asia and Oceania GHI

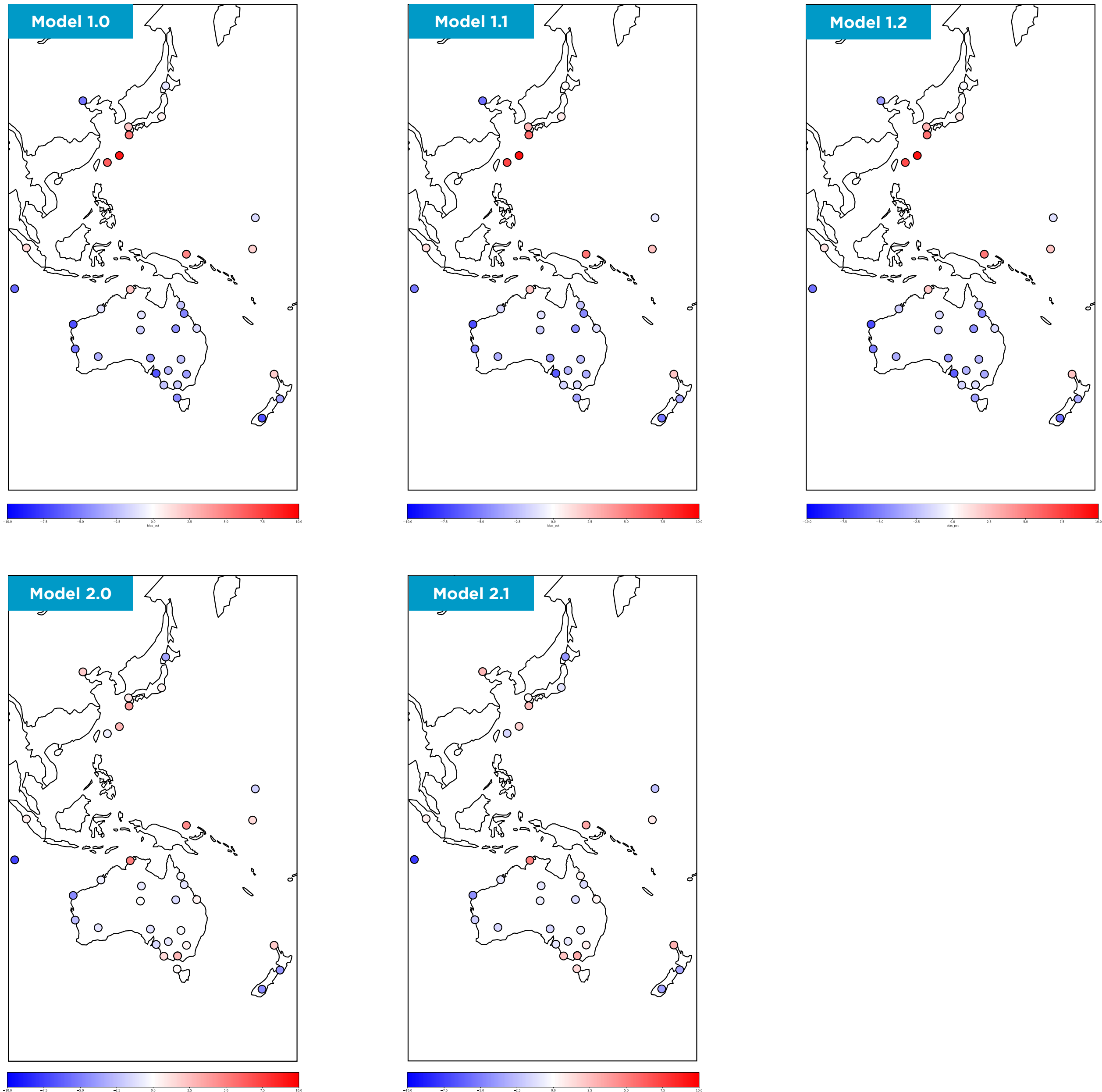
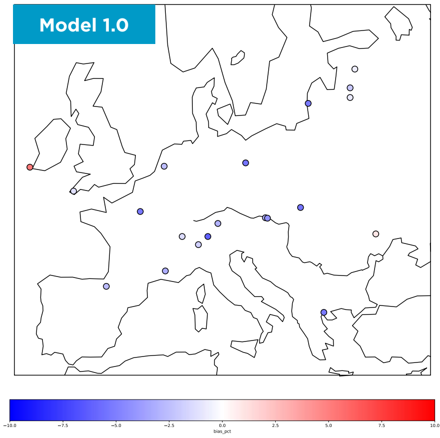


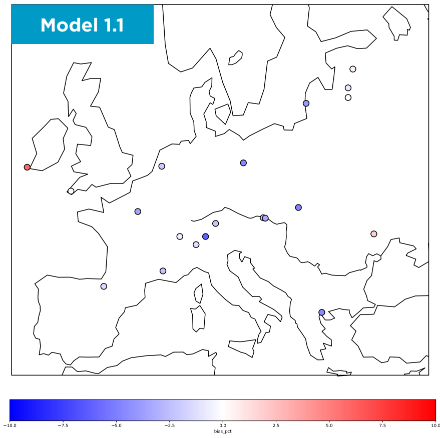
Table A-3. East Asia and Oceania: Regional GHI comparison statistics for each of the five Vaisala models. All values are percent.

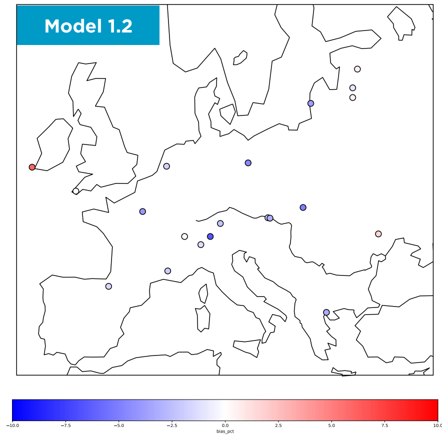
Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	-1.51	3.88	-2.33	22.42	14.92	35
1.1	-1.19	3.93	-1.95	22.42	14.90	35
1.2	-1.22	3.85	-1.98	22.37	14.79	35
2.0	-0.19	2.67	-0.16	21.53	13.63	35
2.1	-0.27	2.64	-0.63	21.50	13.63	35

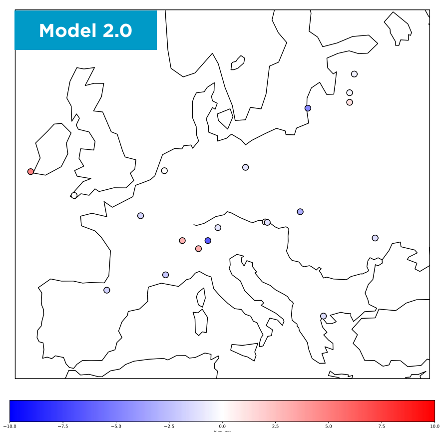
[1]Mean Bias Error
[2]RMSE = Root Mean Squared Error
[3]MAE = Mean Absolute Error
[4]N = Number of Comparison Locations

Europe GHI









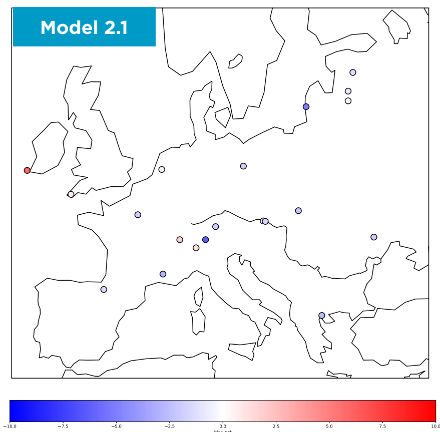


Table A-4. Europe: Regional GHI comparison statistics for each of the five Vaisala models. All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	-2.66	2.69	-2.78	30.44	21.78	20
1.1	-1.80	2.75	-1.99	30.32	21.67	20
1.2	-1.68	2.75	-1.73	30.44	21.68	20
2.0	-0.69	2.57	-1.02	29.30	21.27	20
2.1	-1.12	2.57	-1.46	29.29	21.24	20

[1]Mean Bias Error

[2]RMSE = Root Mean Squared Error

[3]MAE = Mean Absolute Error

[4]N = Number of Comparison Locations

North America GHI

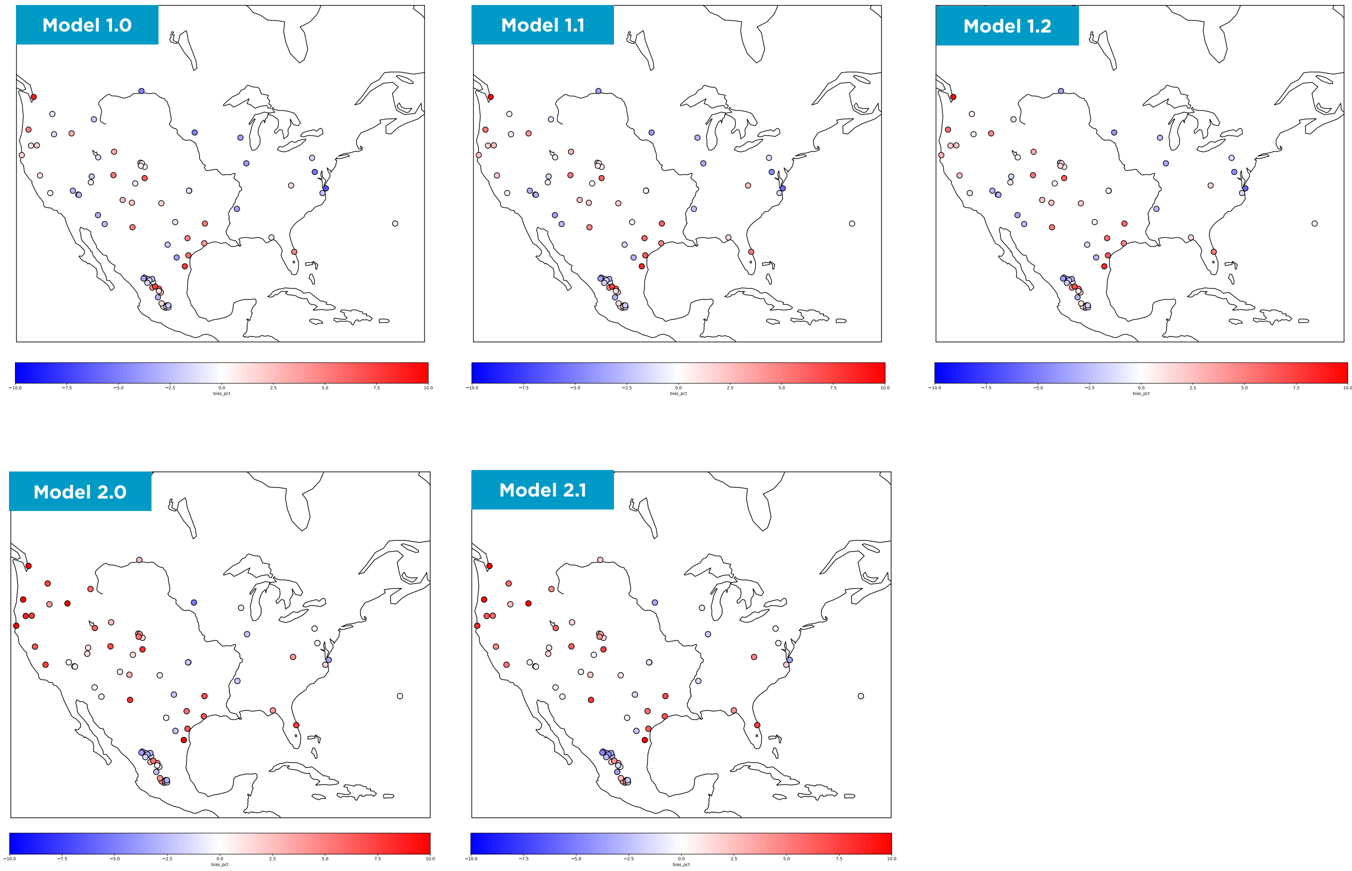


Table A-5. North America: Regional GHI comparison statistics for each of the five Vaisala models. All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	0.22	3.37	-0.25	19.42	12.19	78
1.1	0.53	3.36	0.10	19.40	12.17	78
1.2	0.58	3.35	0.21	19.50	12.27	78
2.0	2.44	4.74	1.72	19.45	12.20	78
2.1	1.88	4.20	1.50	19.14	11.97	78

[1]Mean Bias Error
[2]RMSE = Root Mean Squared Error
[3]MAE = Mean Absolute Error
[4]N = Number of Comparison Locations

South America GHI

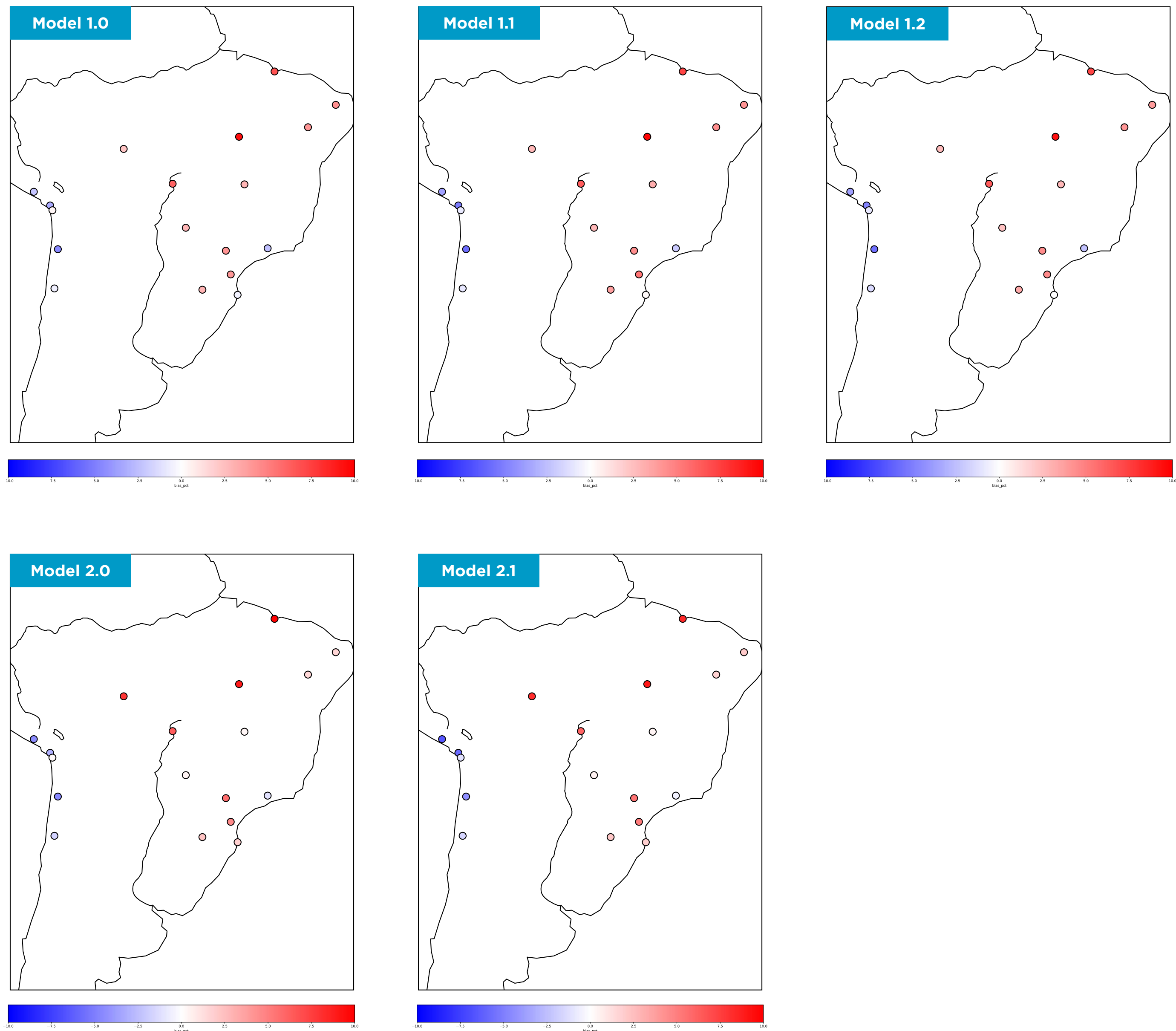


Table A-6. South America: Regional GHI comparison statistics for each of the five Vaisala models. All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	2.02	3.77	2.85	19.85	12.79	18
1.1	2.01	4.34	2.95	19.99	12.95	18
1.2	1.79	4.22	2.69	19.98	12.92	18
2.0	2.06	4.47	1.54	19.99	12.49	18
2.1	1.72	4.70	1.79	20.00	12.53	18

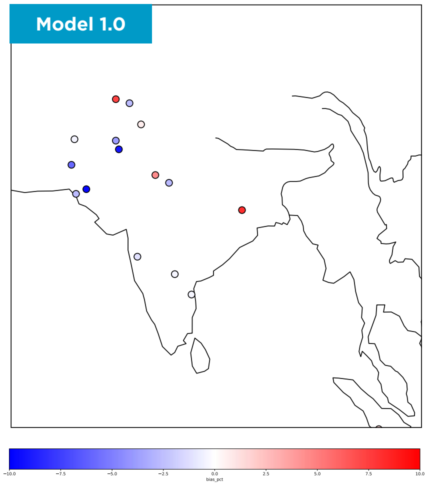
[1]Mean Bias Error

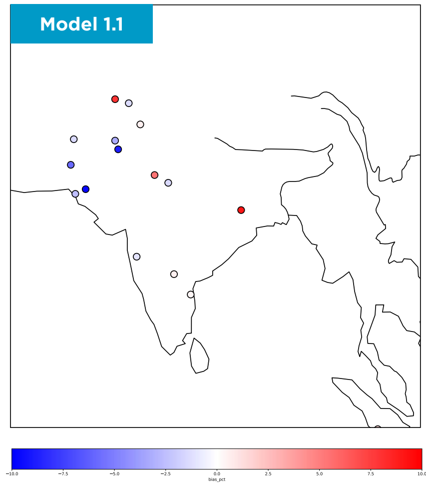
[2]RMSE = Root Mean Squared Error

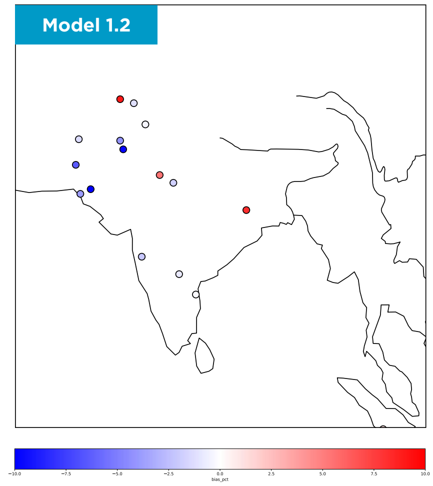
[3]MAE = Mean Absolute Error

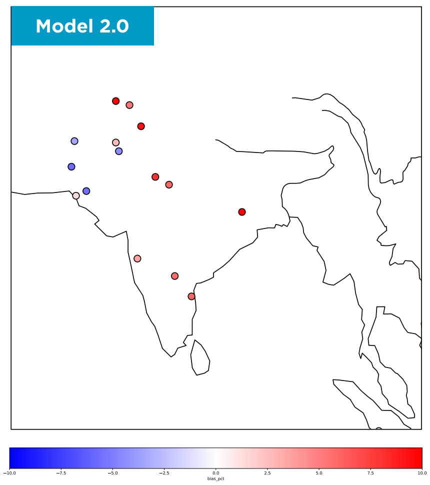
[4]N = Number of Comparison Locations

South Central Asia GHI









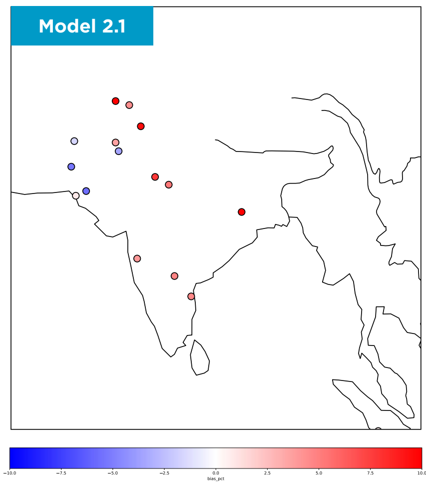


Table A-7. South Central Asia: Regional GHI comparison statistics for each of the five Vaisala models.*
All values are percent.

Vaisala Model	Mean MBE[1]	MBE Std. Dev	Median MBE	Mean RMSE[2]	Mean MAE[3]	N[4]
1.0	-1.20	5.13	-1.12	21.56	14.96	15
1.1	-0.76	5.33	-1.33	21.44	14.81	15
1.2	-1.42	5.73	-1.38	21.18	14.68	15
2.0	3.66	6.21	5.46	20.77	14.03	15
2.1	3.51	5.66	4.42	20.20	13.55	15

[1]Mean Bias Error
[2]RMSE = Root Mean Squared Error
[3]MAE = Mean Absolute Error
[4]N = Number of Comparison Locations

*We are aware that the MERRA2 aerosol data backing the 2.1 model has been shown to have a bias in the India region. NASA does not have plans to fix it at this time.

VAISALA
DIGITAL

www.vaisala.com