The National Solar Radiation Database (NSRDB) Final Report: Fiscal Years 2019-2021

Manajit Sengupta, Yu Xie, Aron Habte, Grant Buster, Galen Maclaurin, Paul Edwards, Haiku Sky, Mike Bannister, and Evan Rosenlieb

National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-82063
Revised September 2022

The National Solar Radiation Database (NSRDB) Final Report: Fiscal Years 2019-2021

Manajit Sengupta, Yu Xie, Aron Habte, Grant Buster, Galen Maclaurin, Paul Edwards, Haiku Sky, Mike Bannister, and Evan Rosenlieb

National Renewable Energy Laboratory

Suggested Citation
Sengupta, Manajit, Yu Xie, Aron Habte, Grant Buster, Galen Maclaurin, Paul Edwards, Haiku Sky, Mike Bannister, and Evan Rosenlieb. 2022. *The National Solar Radiation Database (NSRDB) Final Report: Fiscal Years 2019-2021*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-82063. https://www.nrel.gov/docs/fy22osti/82063.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-82063
Revised September 2022

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Errata

This report, originally published in February 2022, was revised in September 2022 to combine sections with similar topics during the 3-year work.

Acknowledgments

The authors acknowledge Dr. Tassos Golnas and Dr. Guohui Yuan, the technology manager and the program manager, respectively, of the Systems Integration program of the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office, for their support of this project. The authors also acknowledge Dr. Michael Foster from the University of Wisconsin and Dr. Andrew Heidinger of the National Oceanic and Atmospheric Administration for the development of the satellite algorithms and for producing the cloud properties that are integral to the National Solar Radiation Database (NSRDB). Finally, the authors acknowledge Dr. Christian Gueymard for his contribution to the improvement of aerosols in the NSRDB.

List of Acronyms

AOD	aerosol optical depth
ATBD	Algorithm Theoretical Basis Document
AWS	Amazon Web Services
BDN	Bondville
CONUS	contiguous United States
CSP	concentrating solar power
DHI	diffuse horizontal irradiance
DNI	direct normal irradiance
DRA	Desert Rock
FARMS	Fast All-Sky Model for Solar Applications
FARMS-NIT	Fast All-Sky Radiation Model for Solar Applications with Narrowband Irradiances on Tilted Surfaces
FPK	Fort Peck
GHI	global horizontal irradiance
GOES	Geostationary Operational Environmental Satellite
GUM	Guide to the Expression of Uncertainty in Measurement
GWN	Goodwin Creek
HDF5	Hierarchical Data Service 5
HPC	high-performance computing
HSDS	Highly Scalable Data Service
IEEE	Institute of Electrical and Electronics Engineers
MBE	mean bias error
MERRA	Modern Era-Retrospective Analysis for Research and Analysis
NREL	National Renewable Energy Laboratory
NSRDB	National Solar Radiation Database
POA	plane of array
PSM	Physical Solar Model
PSU	Penn State University
PV	photovoltaic
RMSE	root mean square error
SAM	System Advisor Model
SMARTS	Simple Model of the Atmospheric Radiative Transfer of Sunshine
SURFRAD	Surface Radiation Budget
SXF	Sioux Falls
TBL	Table Mountain
TGPY	typical global horizontal irradiance-based plane-of-array year
TMY	typical meteorological year
TPY	typical plane-of-array year
WECC	Western Electricity Coordinating Council

Executive Summary

The National Solar Radiation Database (NSRDB) is the leading public source of high-resolution solar resource data in the United States, with more than 166,000 users annually. This database represents the state of the art in satellite-based estimation of solar resource information and uses a unique physics-based modeling approach that enables improvements in accuracy with the deployment of the next-generation geostationary satellites. Making the highest quality, state-of-the-art, regularly updated data sets available on a timely basis for users reduces costs of solar deployment by providing accurate information for siting studies and system output prediction, and thereby reduces levelized cost of energy. Also, high-resolution information from the NSRDB enables moving beyond levelized cost of energy when valuing the impact of renewables on the grid. Additionally, the NSRDB enables the integration of large amounts of solar on the grid by providing critical information about solar availability and variability that is used to enhance grid reliability and power quality.

Table of Contents

List of Figures

List of Tables

Introduction

This project supports and enables the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office initiative in making solar energy fully cost-competitive with traditional energy sources and facilitates the development of tens of gigawatts of solar power plants. Specifically, this project

- Updates the National Solar Radiation Database (NSRDB) on an annual basis to extend the data to cover the period from 1998–2020
- Updates typical meteorological year (TMY) data sets to include additional years on an annual basis
- Updates the Physical Solar Model (PSM) to provide data of enhanced accuracy from using the next-generation Geostationary Operational Environmental Satellite (GOES)
- Updates the NSRDB website and servers to enable the delivery of the new data sets
- Validates the updated NSRDB data sets with ground measurements to ensure accuracy of the new data
- Delivers complete, long-term, high-resolution data as a public resource for all users.

In the first year of the budget period, we updated the NSRDB data and TMY data with the data from the most recent calendar year (2018) using the newly available next-generation satellite (GOES-16) and PSM Version 3 (PSM V3). The update was validated using ground measurements. A TMY data set was also updated. After the update was complete, a webinar was held to inform stakeholders about the availability of the new data and solicit feedback about the stakeholders' requirements from the NSRDB. This feedback was used to adjust budget periods two and three NSRDB activities (while staying within the scope of the statement of project objectives) to better meet the stakeholders' needs. The following chapters describe the related work conducted under multiple tasks.

1 Task 1: Acquisition of Satellite and Ancillary Data Sets

1.1 Project Approach

- Acquire calibrated and navigated geostationary data for 2018 from two satellites (GOES-15 and GOES-16) and calculate cloud properties using the satellite information.
- Acquire calibrated and navigated geostationary data for 2019 from two satellites (GOES-16 and GOES-17) and calculate the cloud properties using the satellite information.
- Acquire calibrated and navigated geostationary data for 2020 from two satellites (GOES-16 and GOES-17) and calculate cloud properties using the satellite information.
- Acquire ancillary data from the Modern Era-Retrospective Analysis for Research and Analysis (MERRA-2) satellite for aerosols, water vapor, temperature, humidity wind speed, and wind direction.
- Acquire surface albedo data from the University of Massachusetts Boston if updates are available.

1.2 Project Results and Discussion

Under a subcontract, the University of Wisconsin completed the delivery of cloud property data sets (2018-2020) for GOES-West and GOES-East. Further, the MERRA-2 data were acquired for use in the PSM V3 (Figure 1).

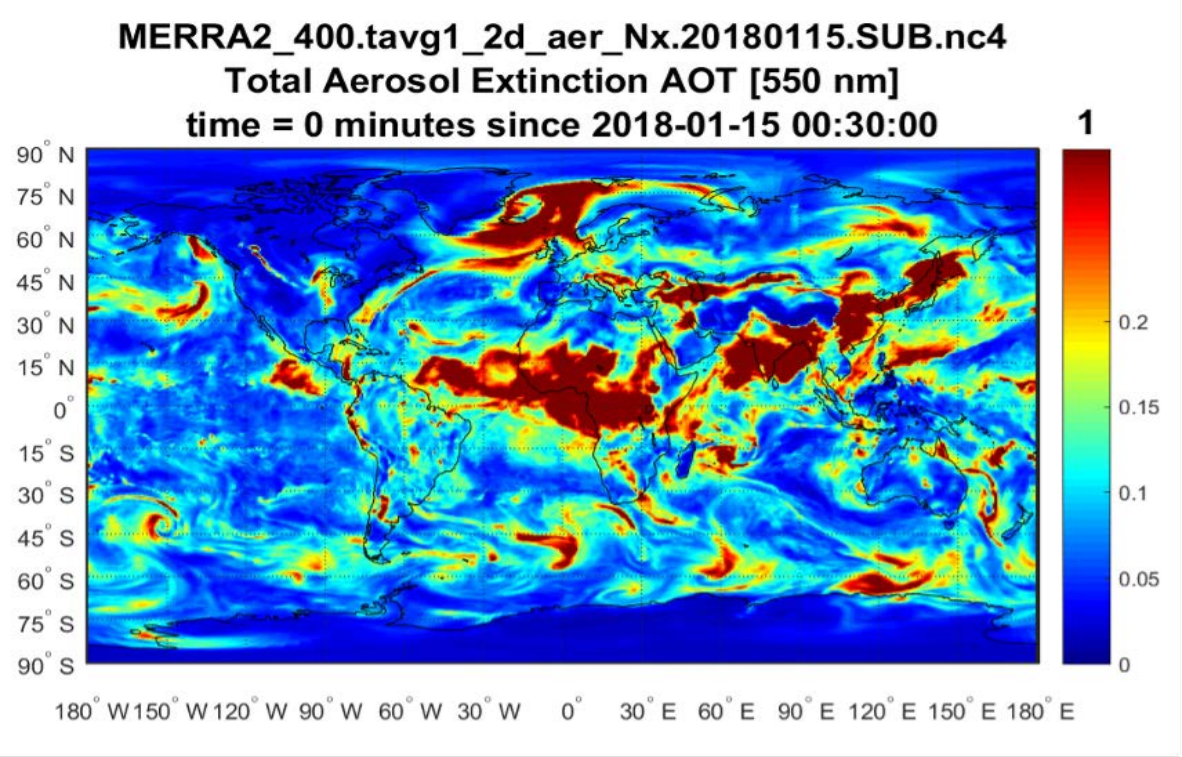

Figure 1. Example of MERRA-2 AOD data set for one time stamp on January 15, 2018

The University of Wisconsin discovered an error with the visible channel calibration for GOES-16 and GOES-17. It involves a doubling of the earth-sun distance modifier and causes a seasonal bias ranging from 0%–3% during the course of the year. It is unclear how large an effect this has on cloud properties. The error has been fixed, and the reprocessing and delivery of the corrected GOES-16/GOES-17 records for 2018–2020 is done. For the NSRDB 2018–2019, NREL continued to use the older data because the evaluation with the new data did not demonstrate significant differences.

The calibrated and navigated geostationary data for 2020 from two satellites (GOES-16 and GOES-17) were acquired and stored in NREL's HPC. The sample data for cloud optical thickness on July 1, 2020, are shown in Figure 2. The spatial and temporal resolutions of the data are 2 km and 5 minutes, respectively.

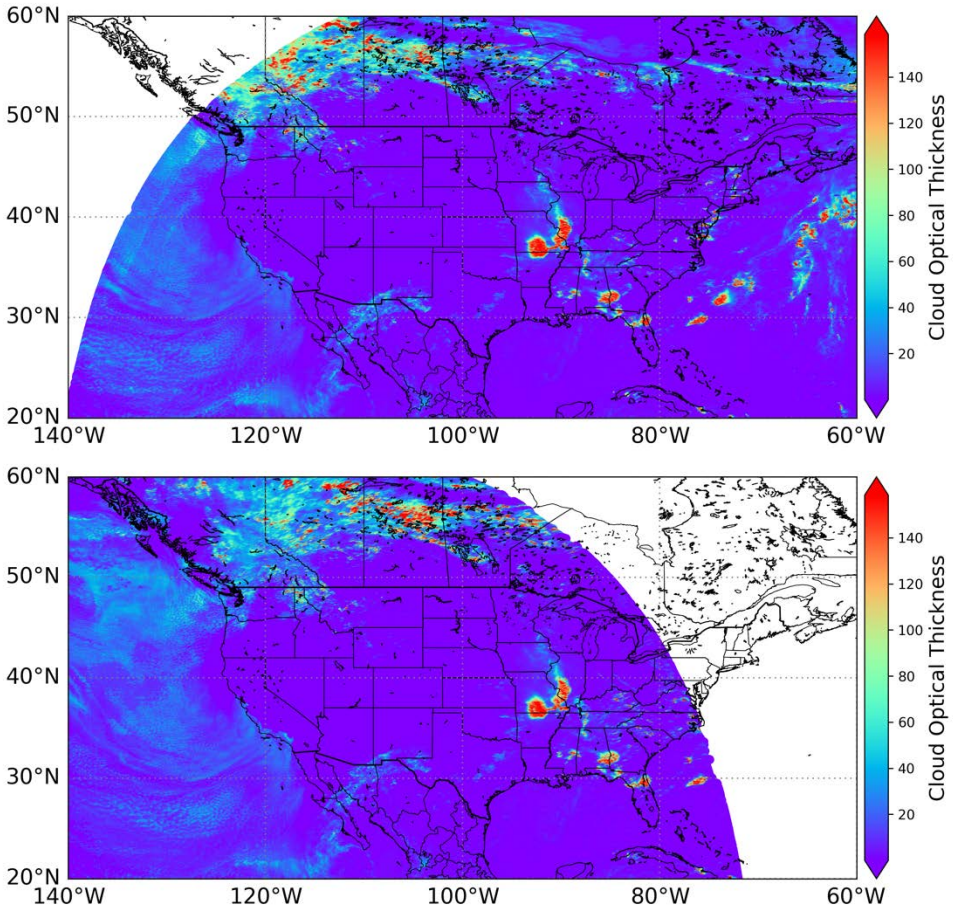

Figure 2. Cloud optical thickness on July 1, 2020, from (a) GOES-16 and (b) GOES-17

The ancillary data for 2020—including aerosols, water vapor, temperature, humidity, wind speed, and wind direction—from the MERRA-2 product were acquired and stored in NREL's HPC. The sample data for AOD and precipitable water vapor on July 1, 2020, are shown in Figure 3. The spatial and temporal resolutions of the data are one-half degree and 1 hour, respectively.

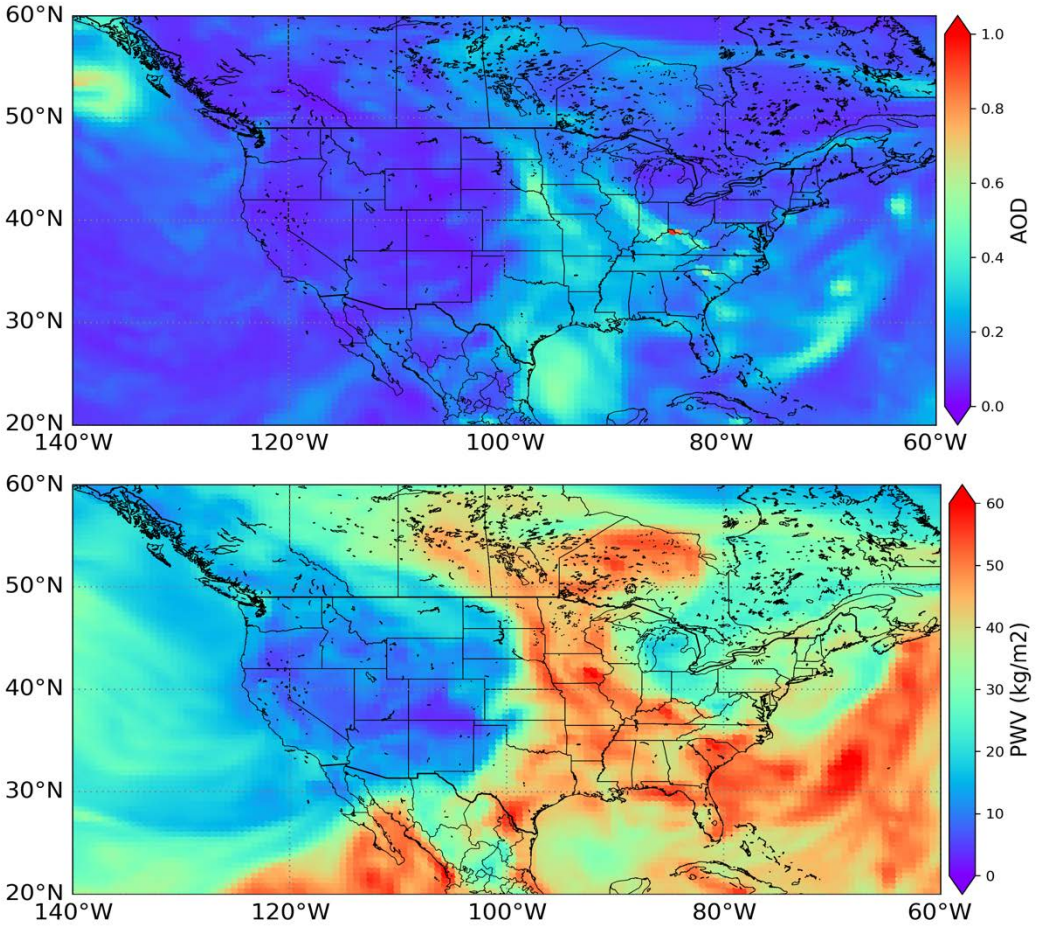

Figure 3. (a) AOD and (b) precipitable water vapor on July 1, 2020

The surface albedo, which contains snow and ice information data for 2020 from the National Ice Center, have been updated and are stored in NREL's HPC. The sample data for the surface albedo on July 1, 2020, are demonstrated in Figure 4. The spatial resolution of the daily data is approximately 1 km.

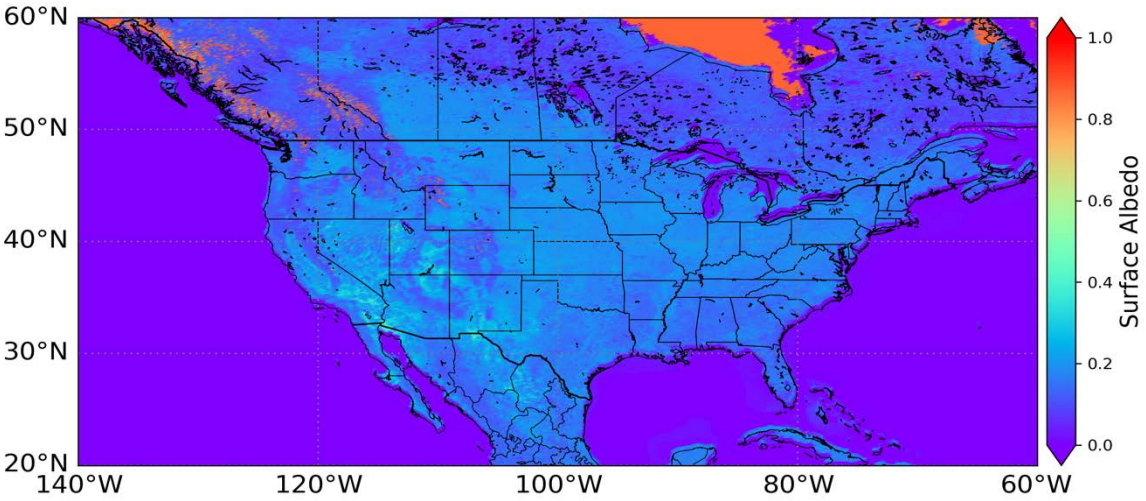

Figure 4. Surface albedo on July 1, 2020

2 Task 2: Development of Solar and Meteorological Data

2.1 Project Approach

- Develop a static, uniform 2-km x 2-km grid to map all GOES-16 satellite data sets.
- Develop an elevation map for the 2-km x 2-km grid to be used to scale aerosol and moisture variables during the MERRA2 downscaling process.
- Map the cloud properties to the static grid.
- Process ancillary data to map to the static grid and produce the meteorological output while maintaining physical consistency of variables.
- Develop high-resolution aerosol data mapped to the grid and calibrate to remove bias.
- Check the physical consistency of the cloud properties and ancillary data sets for developing solar and meteorological products.
- Run the PSM to produce solar radiation, including global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI)
- Fill gaps in solar and meteorological data to produce a serially complete data set.
- Check to ensure that output solar radiation and meteorological products are error free, physically consistent, and serially complete.
- Investigate 5-minute data from GOES-16 for 1 full year (2018) of data. Specifically, develop methods to average the high-resolution information at the 30-minute, 4-km NSRDB resolution.
- Reduce the resolution of the solar data sets to 4 km x 4 km, half-hourly, to match the current NSRDB resolution.
- Blend the east and west satellite data to provide coverage for the whole United States.
- Develop a TMY using all years of data from 1998–2018.
- Update the Algorithm Theoretical Basis Document (ATBD) to reflect changes from older satellites to GOES-16.

2.2 Project Results and Discussion

2.2.1 Re-Grid 2018 Satellite Data Set to Static Grid and Create Individual 2-km x 2-km Pixel Files for Use in the National Solar Radiation Database

Prior to producing a gridded NSRDB as part of our satellite-based mapping project, a static, uniform, latitude-longitude grid (2-km x 2-km nominal resolution) was created for the GOES-16 data set (Figure 5).

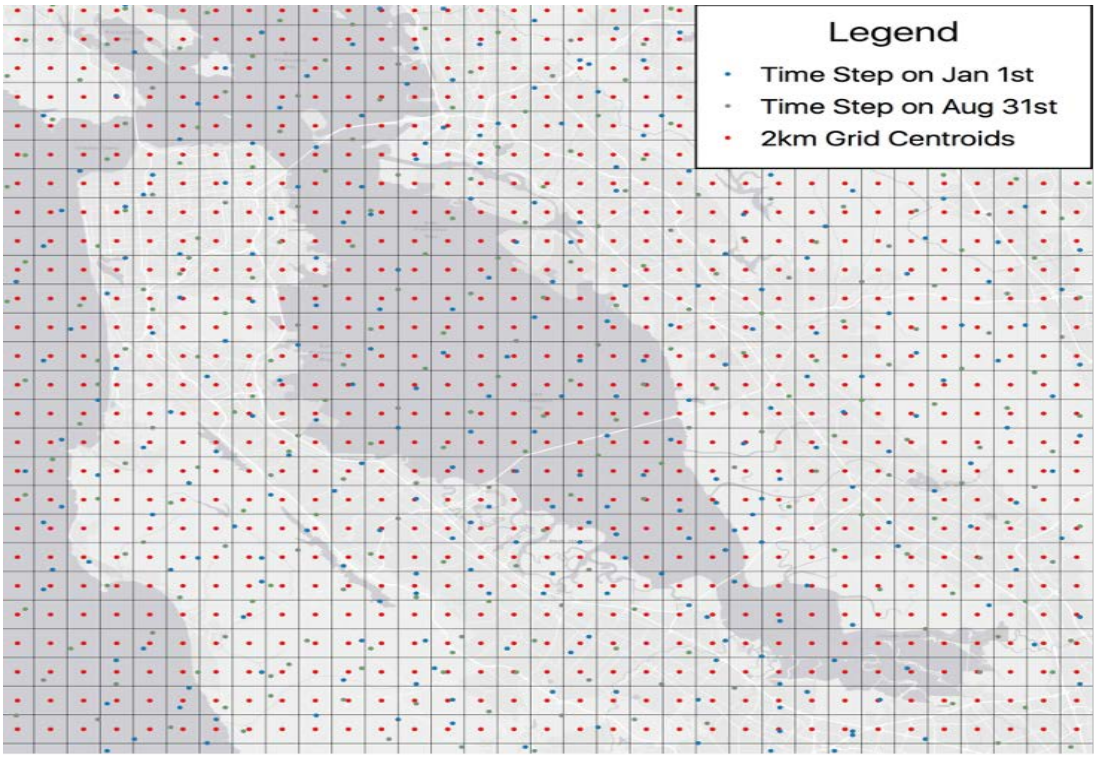

Figure 5. Example of the remapping of a floating satellite grid (green and orange dots) to a static grid (squares)

The static NSRDB grid has the following features:

- Covers all areas covered by the original NSRDB
- Is four times the spatial resolution (2 km instead of 4 km)
- Each old NSRDB point is associated with four new points.
- The spatial coverage is extended to include all of South America.

2.2.2 The National Solar Radiation Database for Concentrating Solar Power Applications

The National Renewable Energy Laboratory (NREL) presented a conference paper titled "The National Solar Radiation Data Base (NSRDB) for CSP Applications."[1] This study examines the differences in DNI in two versions of the NSRDB produced by NREL. NSRDB V3 includes significant changes to various parts of the radiative transfer model and inputs to the model compared to NSRDB V2. The changes in NSRDB V3 resulted in reduced uncertainty than NSRDB V2. The study quantified the uncertainty and the spatial and temporal variability under clear-sky conditions. The uncertainty estimation was performed using a standardized method, the Guide to the Expression of Uncertainty in Measurement (GUM). The evaluation of the accuracy of the NSRDB was conducted using high-quality ground measurements for seven National Oceanic and Atmospheric Administration Surface Radiation Budget (SURFRAD) observing network stations for 1998–2015.

[1] See https://www.nrel.gov/docs/fy19osti/72310.pdf.

The analysis shown in Figure 6 focused on clear-sky conditions, which are critical for concentrating solar power (CSP) applications. Half-hourly, clear-sky periods were selected using the criteria of 90% clear points in 1-minute ground data. At almost all SURFRAD locations, the NSRDB V3 demonstrates significant improvement or reduction in uncertainty compared to the NSRDB V2, as shown in Figure 6. These improvements are clearly noticeable for the eastern locations, such as Bondville (BND), Goodwin Creek (GWN), Sioux Falls (SXF), and Penn State University (PSU). Further, the observed reduction in uncertainty occurs across all averaging timescales. On an annual basis, remarkably, the uncertainty is very close to the measurement uncertainty. Desert Rock (DRA) and Fort Peck (FPK) do not show reductions in uncertainty in the NSRDB V3; instead, they performed even worse—their uncertainty tends to increase for some averaging timescales. The Table Mountain (TBL) station, in Boulder, Colorado, shows some reduction in uncertainty in the NSRDB V3; however, this location also shows increased uncertainty in both NSRDB versions compared to the other locations. This could be attributed to the closeness of the site to the Rocky Mountains, resulting in shading in the observations during afternoon hours. The NSRDB does not include terrain impacts, so the calculated DNI is significantly higher at all timescales when there is possible horizon shading.

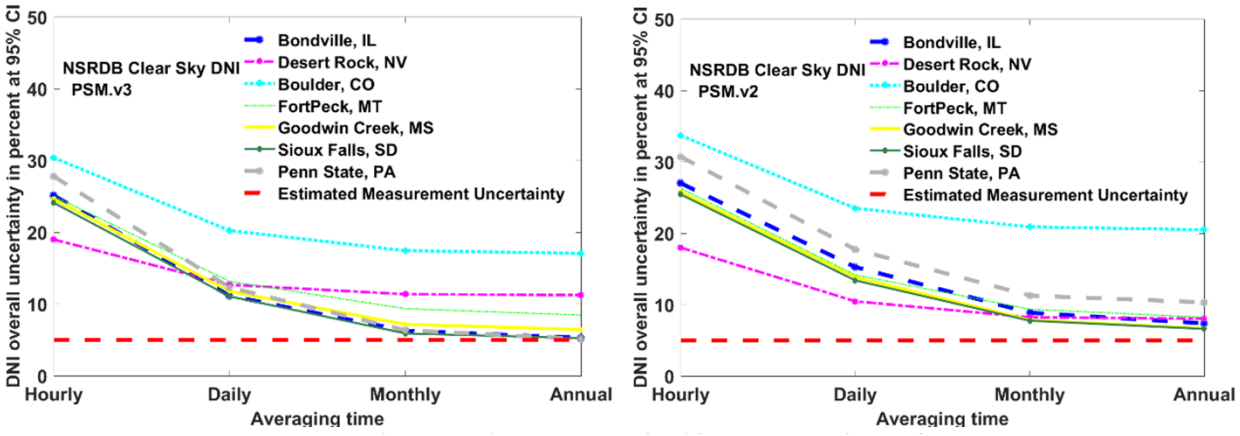

Figure 6. DNI uncertainty from (left) NSRDB V3 and (right) NSRDB V2. Years: 1998–2015

The detailed statistics in Figure 7 of the sources of uncertainty, such as mean bias error (MBE) and root mean square error (RMSE), provide some insights in explaining the overall uncertainty estimation described in Figure 3. The MBE (Figure 4, top panel) shows a negative bias at all locations in the NSRDB V2 but only at the three western locations in the NSRDB V3. In contrast, Bondville and Goodwin Creek in the east show very small negative and positive bias, respectively, whereas Penn State University and Sioux Falls demonstrate practically no bias in the NSRDB V3. A tentative explanation is that the changes implemented in the NSRDB V3 improved the clear-sky estimates.

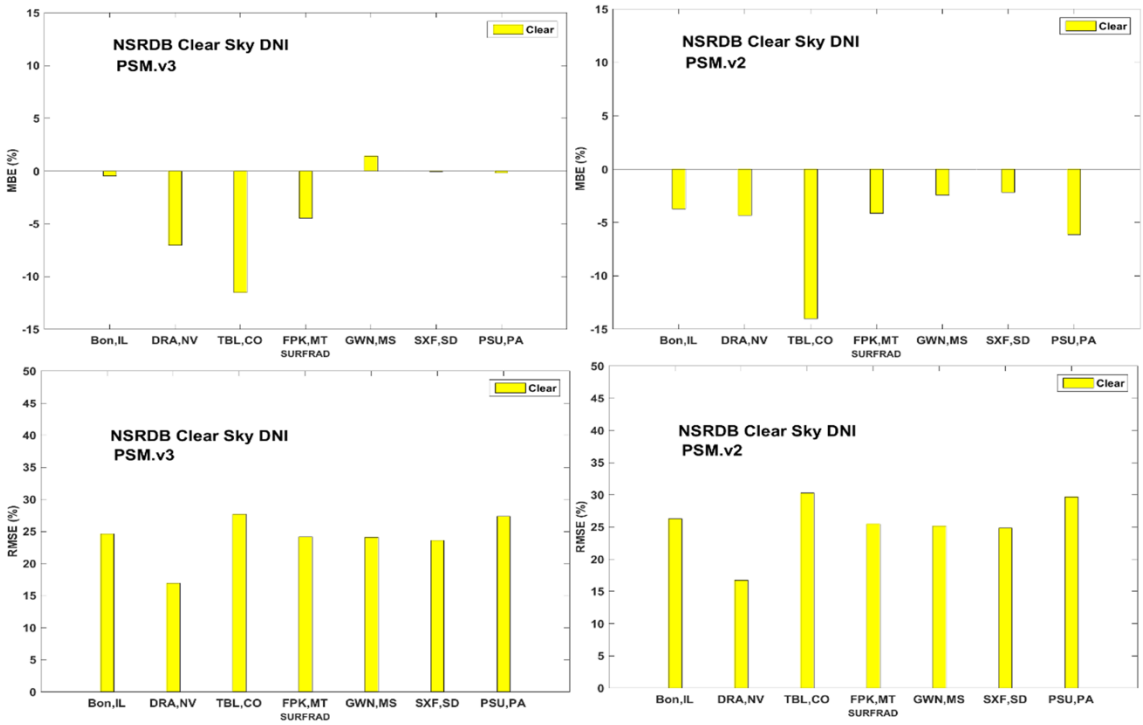

Figure 7. Top: Percentage MBE for (left) NSRDB V3 and (right) NSRDB V2. Bottom: The same but for the percentage RMSE. Years: 1998–2015

Desert Rock and Fort Peck in the west show slightly higher MBE but similar RMSE (a measure of scatter or randomness) in the NSRDB V3, which is consistent with the results shown in Figure 3. The MBE does not change with timescale average; however, the RMSE or scatter decreases as the averaging time increases. That is why these two locations did not demonstrate decreased uncertainty in Figure 3 with the increase in averaging time. In other words, the MBE becomes dominant in the uncertainty estimation as the averaging time increases from hourly to annual.

2.2.3 Solving Issues in the Model and Data

During the quality assurance process, the NREL NSRDB team discovered some issues that result in some variables occasionally producing unrealistic values under certain conditions. Some of these issues include:

1. *Unrealistic, high aerosol optical depth (AOD) values from MERRA-2.* Previously, the code did not have set values/ranges for maximum and minimum AOD (e.g., Figure 8). With the new update, realistic minimum and maximum value thresholds for all variables were put into place to check for erroneous values.

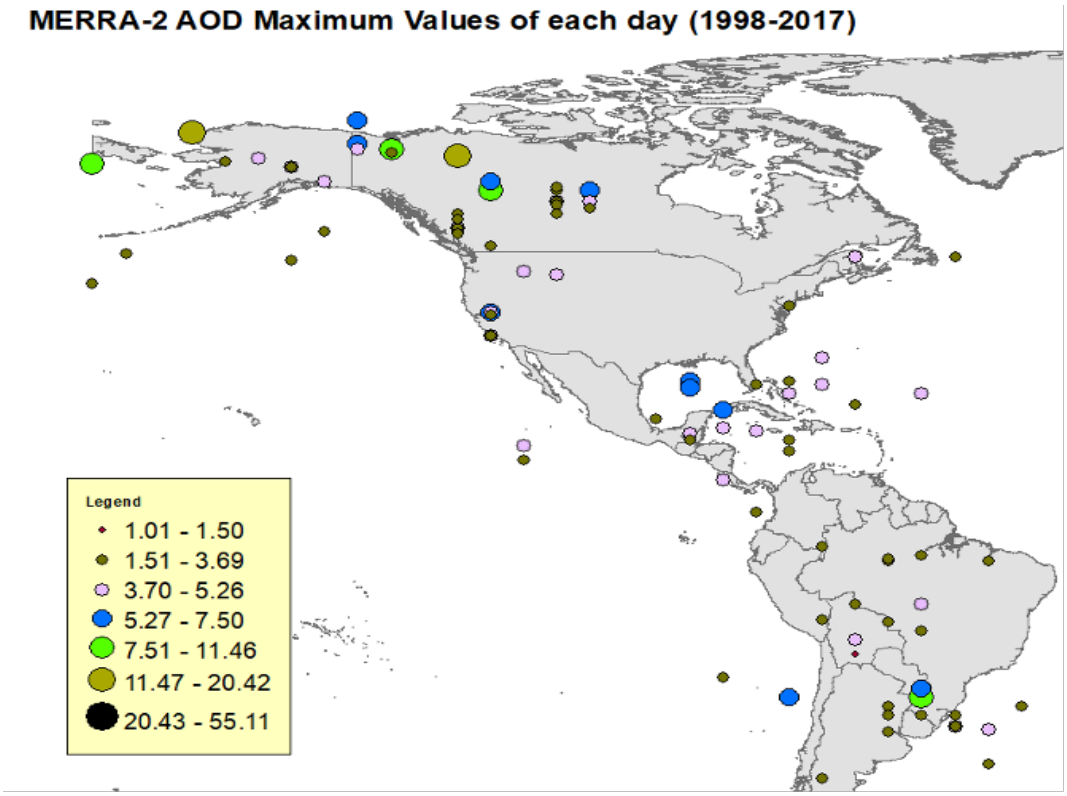

Figure 8. Extreme erroneous AOD values from MERRA-2

2. *Cloud optical depth and particle size radius were not available occasionally even if clouds were detected by the satellite algorithms.* The cloud optical depths and effective radius were filled when missing values were identified. Every time step of missing cloud property (optical depth or particle size) at a given NSRDB grid cell was filled from the nearest time step at that grid cell with valid data and the same cloud category (ice, water). For example, if grid cell 1 was missing cloud properties for an ice cloud at noon on January 1, and the temporally nearest valid ice cloud property data are on the morning of January 1, the missing cloud properties are filled from the ice cloud properties from the morning of January 1.
3. *Resolution issues for the wind speed and direction data.* These values were also recalculated and updated on the server.

TMY 2016 and TMY 2017 were recalculated to reflect these changes. Although the 20 years of NSRDB irradiance data sets were not changed, the TMY calculation (TMY 2016 and TMY 2017) takes wind speed into account, so the TMY data (and TMY irradiance) were updated to reflect the recomputed wind speed. Note that TMY 1998–2016 refers to TMY generated using 1998–2016 data. Similarly, TMY 2017 was generated using 1998–2017data.

2.2.4 National Solar Radiation Database Processing

All-sky irradiance models were refactored and are used for 2018 runs (Figure 9). This approach facilitates the processing time. Using this new approach, subsets of data (e.g., a months' worth of data) can be processed, whereas previously only full-year datasets could be processed. This was a big improvement because GOES-16 and then later GOES-17 have high-temporal and high-

spatial-resolution data; therefore, processing subsets of the data set was the only feasible approach. Ancillary variable processing (MERRA-2, solar zenith angle, asymmetry) was also refactored for the 2018 runs.

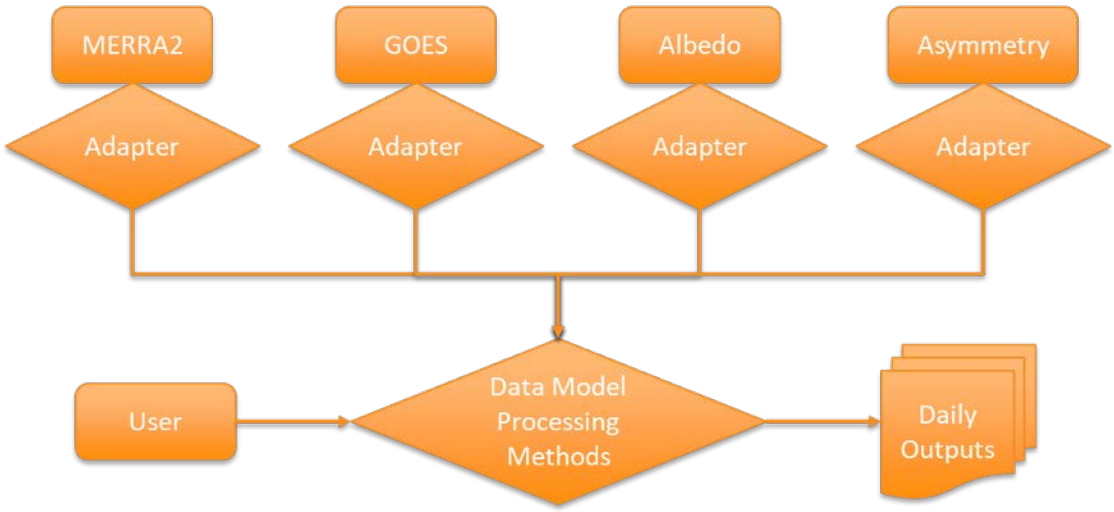

Figure 9. Flowchart of the NSRDB data

The 2018 NSRDB data set were processed for the east, west, and contiguous United States (CONUS). All data sets were combined into an aggregated 4-km, 30-minute NSRDB file (matching the format and metadata from previous years) (Figure 10). The 2018 TMY for the NSRDB was developed using 1998–2018 data.

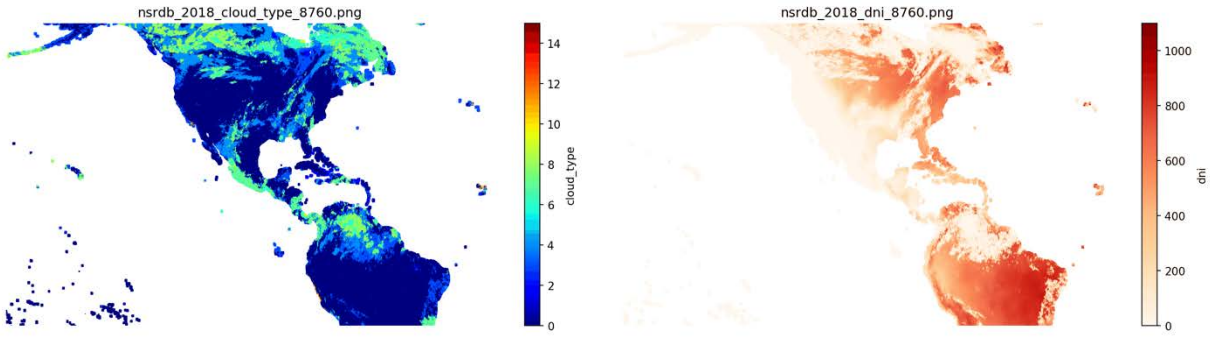

Figure 10. Example of the 2018 NSRDB

2.2.5 Creating a Latitude-Longitude Grid

Prior to producing a gridded NSRDB as part of our satellite-based mapping project, a static, uniform, latitude-longitude grid (2-km x 2-km nominal resolution) was created for the GOES-17 data set (Figure 11) for the new GOES-17 West data set.

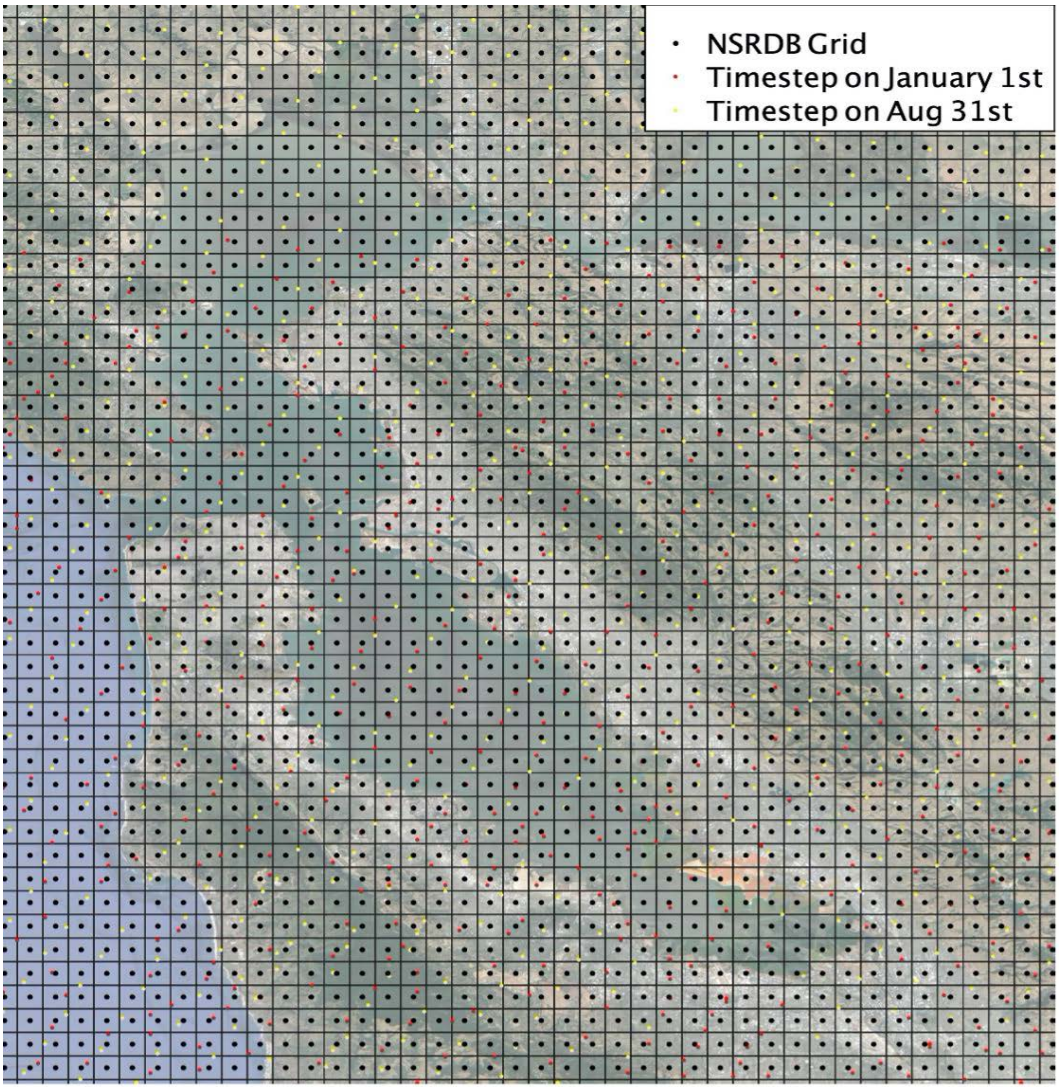

Figure 11. Example of the remapping of a floating satellite grid (yellow and orange dots) to a static grid (squares)

This static NSRDB grid has the following features:

- Covers all areas covered by the original NSRDB
- Is four times the spatial resolution (2 km x 2 km, as opposed to 4 km x 4 km)
- Each old NSRDB point is associated with four new points.
- The extent is extended south to include all of South America.

Elevations and time zones have been associated with points; thus, all inputs crucial to the actual calculation were completed. All 2-km metadata were associated with the points that we have for the 4-km grid so that the metadata is consistent with the 4-km metadata. Moreover, all ancillary meteorological data sets were mapped to the static NSRDB grid by applying various physically consistent interpolation techniques. After processing the 2019 data set, data from the two satellites were merged to create seamless 30-minute solar irradiance.

NREL also developed a method to produce ultraviolet data sets from the NSRDB. A simple model was developed to produce a global horizontal ultraviolet data set. The simple model was validated, and the results showed good agreement with measurements, which provides confidence about the accuracy of the model (Figure 12). The model typically under- or overestimates the measured ultraviolet irradiance on average by only ±2W/m^2. The data are disseminated through the NSRDB viewer to the public.

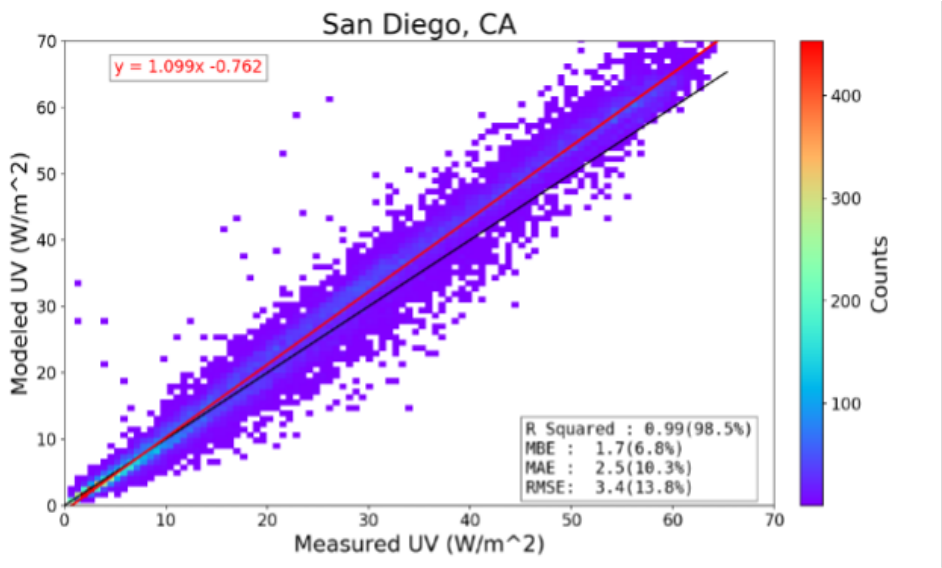

Figure 12. Modeled versus measured global ultraviolet irradiance under all-sky conditions and zero tilt. The red dashed line is a 1:1 line, and the black line is a regression line.

2.2.6 Resume National Solar Radiation Database 2020 Processing

The calibrated and navigated geostationary data for 2020 from two satellites (GOES-16 and GOES-17) were acquired and stored in NREL's HPC. The ancillary data for 2020—including aerosols, water vapor, temperature, humidity, wind speed, and wind direction—from the MERRA-2 product were also acquired and stored in NREL's HPC. The spatial and temporal resolutions of the data are one-half degree and 1 hour, respectively. The surface albedo, which contains snow and ice information data for 2020 from the National Ice Center, have been updated and are stored in NREL's HPC. The spatial resolution of the daily data is approximately 1 km.

We used the data to run the PSM and produce the solar radiation data for 2020, including GHI, DNI, and DHI. Figure 13 and Figure 14 show the sample data of the GHI, DNI, cloud optical thickness, and surface albedo on January 1, 2020, and July 1, 2020, respectively.

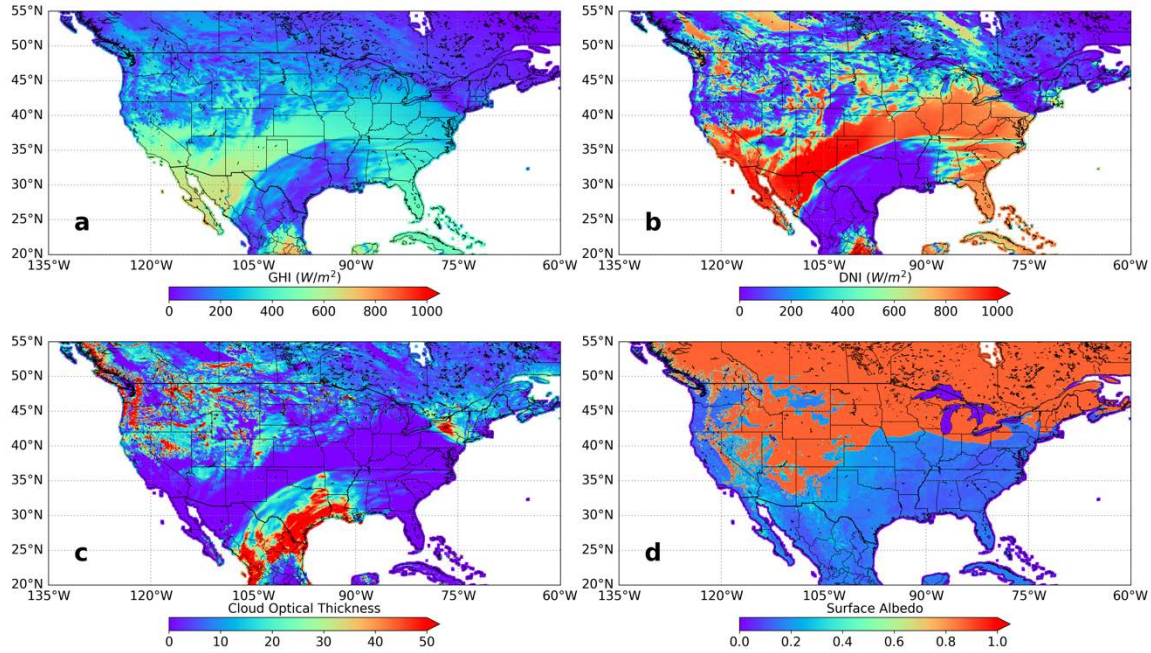

Figure 13. GHI, DNI, cloud optical thickness, and surface albedo on January 1, 2020, at 10 UTC

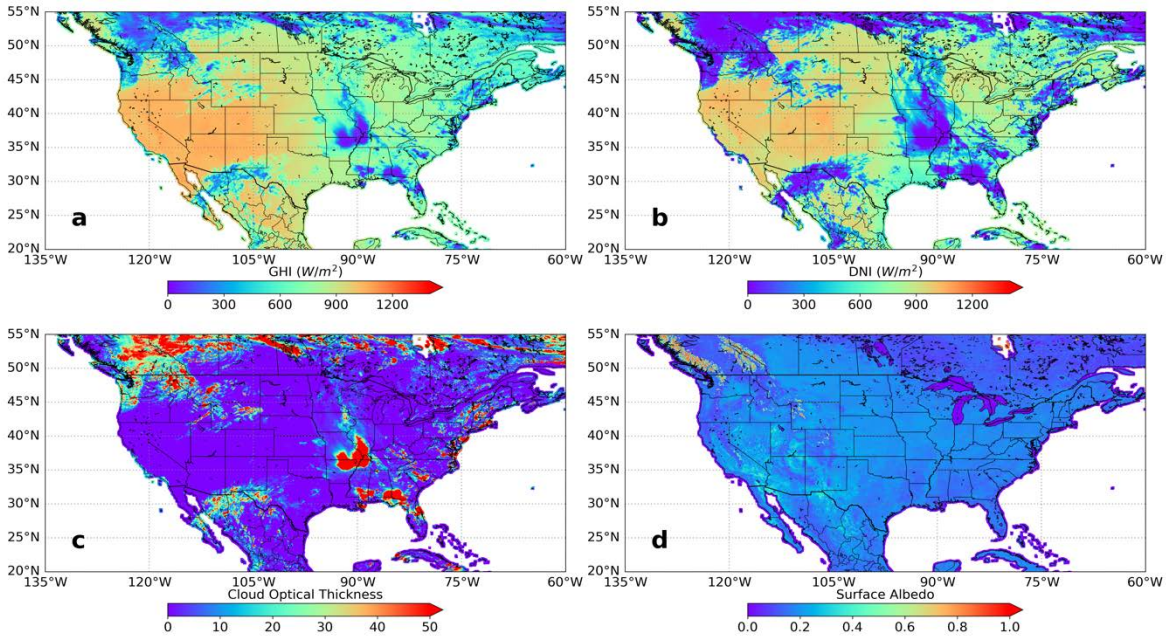

Figure 14. GHI, DNI, cloud optical thickness, and surface albedo on July 1, 2020 at 10 UTC

2.2.7 Evaluate the Impact of Terrain on Resource Availability

Many improvements have been made to the topographic shading algorithm in the past quarter. Previously, the algorithm was written to read in the Digital Elevation Model in a geographic projection and calculate shading factors based on the projected elevation model; however, upon initial examination of results, it became clear that this algorithm is uniquely susceptible to biases introduced by geographic projection. The need to compare the azimuth of topographic features to

the azimuth of the sun at any given time step means that the preservation of absolute and relative angles is necessary. Because the horizon angle of topographic features is dependent on the distance of the feature to the point of perspective, it is important that distances are represented correctly in every direction. It is impossible to choose any one projection that does not introduce these biases for more than one point. Large map-scale projections can provide a small bias for a small area; however, this is not a computationally efficient option for processing shading throughout the entire NSRDB domain. Therefore, the algorithm was modified to read unprojected elevation data and then project each model in a custom azimuthal equidistant projection centered around the point of perspective. This improvement means that topographic features are modeled with better accuracy than had been previously shown in the literature; however, importantly, this also means that we will be more easily able to provide shading-corrected irradiance values for the entire NSRDB domain available through the NSRDB viewer.

Although there has been literature, dating back to at least the 1980s, on the use of geographic information systems to estimate topographic shading, until now there has been no literature that directly estimated the effect on PV system performance. This was identified as a gap in the literature that we can target for publication that furthers the core research mission of NREL and the objective of the NSRDB. Although the algorithm is quite fast—resolving the topographic shading of any given location in only a few seconds—the calculation for the entire NSRDB domain and its millions of points still represents a significant computational undertaking. Therefore, to understand whether topographic shading is likely to cause appreciable amounts of PV shading loss in any scenario, the PV performance accounting for shading was first tested on two small subsets of data:

- A database compiled by Lawrence Berkeley National Laboratory of all existing utility-scale PV plants in the United States. When limited to CONUS, the total is 780 real-world PV installations. The PV installations do not represent a random spatial sampling of CONUS—for example, there are many more in certain states with friendly policy regimes; however, the data set contains locations of varied topography and latitude. Because the data set represents real existing installations, where assumedly topographic shading was considered during siting, the presence of appreciable shading at multiple locations would suggest the possibility that potential topographic shading losses are being underestimated in national PV capacity expansion models.
- Additionally, a small sample of community solar sites in Colorado was assessed. Because community solar PV has less freedom of location than utility PV, it was hypothesized that it would be more likely to show shading losses than utility PV. Unfortunately, without a nationwide data set of community solar sites like those with utility PV sites, it is not possible to run a similar analysis. In case utility PV sites did not show a high degree of topographic shading, it was still deemed worthwhile to examine some community sites because the presence of shading at existing community solar sites could still imply that there are real-world scenarios where correction for topographic shading was important. The 17 community sites run were verified using satellite imagery, and because of their presence along the Front Range and in the mountains of Colorado, they are certainly biased toward exhibiting higher topographic shading.

The sky view model (a virtual representation of the horizon from any given point) was calculated for all utility PV and community solar sites. Using this sky view model, it is possible to compare

between the horizon angle at the azimuth and the horizon angle of the sun to determine if the direct beam irradiance is shaded at a time step. Because this calculation is highly sensitive to temporal resolution, it was calculated at a 5-minute temporal resolution. Additionally, the sky view model can be used to calculate the sky view factor, a calculation of the amount of the sky that topographic features block and therefore the amount of diffuse radiation that topographic features block. With this, it is possible to produce corrected DNI and DHI values, and therefore PV performance can be assessed considering topographic shading. A summary of the results for the two data sets is shown in Table 1.

Table 1. Summary of Data Sets

	Utility PV	Community
Mean	0.33%	1.62%
Minimum	0.0%	0.23%
Q1	0.05%	0.26%
Median	0.22%	0.54%
Q3	0.45%	3.11%
Maximum	4.13%	6.47%

The results for the utility PV sites show that although a very large majority of the sites have negligible shading, at least some sites show appreciable loss from topographic shading. Again, although the community sites are expected to be more biased toward shadier locations than the average community PV site, at the very least they reinforce that topographic shading can have very real impacts in real-world deployment scenarios.

A plot of the percentage shading losses against the sky view factor for every utility PV site shows that there is indeed a tight correlation (a Pearson's correlation coefficient of -.947, although certain points likely exhibit high leverage) (Figure 15).

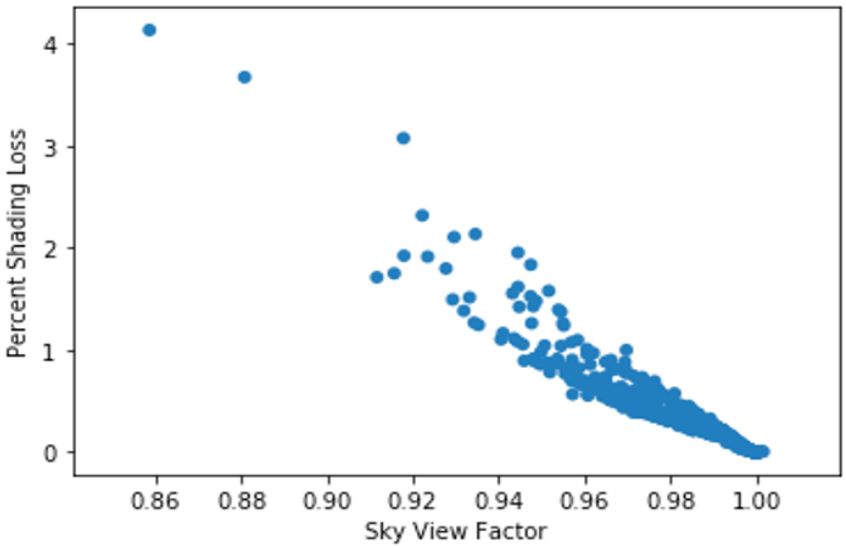

Figure 15. Percentage shading loss as a function of sky view factor

This does seem to imply that the sky view factor could be used as a reasonable proxy for actual shading losses; however, to make a stronger inference, this would need to be tested on a larger selection of sites that are more representative of the population. These results seem to show that topographic shading can cause measurable PV losses, but the scope and scale on a nationwide level is impossible to assess on this subset. Thus, it was decided that for this study it would be much more impactful to perform the calculations for all NSRDB 2-km points within CONUS over 2 million points. The literature review and methods parts of the paper on this study are written in draft form, and the completion of the paper is pending the computation to measure and characterize the shading effects of PV production for all of CONUS. Although processing is continuing, we can examine one preliminary output of the calculation. The sky view models and sky view factors for all NSRDB points within the Western Electricity Coordinating Council (WECC) region have been calculated (Figure 16). The WECC region extends from Canada to Mexico including Alberta, British Columbia, the northern portion of Baja California, Mexico, and the most area of the 14 western states of the United States. A summary of the sky view factor of these points is shown in Table 2.

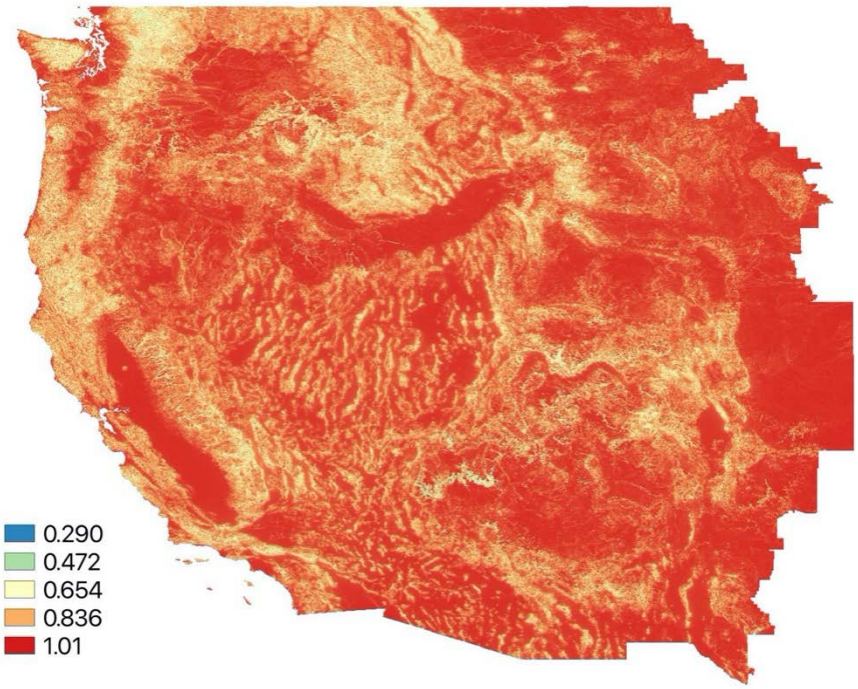

Figure 16. Sky view factor in the WECC region.

Table 2. Summary of the Sky View Factor Points

Mean	.914
Minimum	.290
Q1	.873
Median	.943
Q3	.976
Maximum	1.019

Although PV performance has not been completed yet for these points, using the correlation observed between the sky view factor and the PV performance for the utility PV sites can hint at the level of shading that we might expect to see in the NSRDB points in WECC. Among the utility PV sites, every site that has a sky view factor equal to or less than the median value of .943 found in the WECC NSRDB point has PV losses of at least 1.1% or more. It seems likely, therefore, that approximately half of the land in WECC will have topographic shading losses of at least 1%. This underscores that topographic shading could be a missing aspect to our solar irradiance and PV performance modeling that, though not large, might not be insignificant either.

3 Task 3: Validation of the National Solar Radiation Database

3.1 Project Approach

- Acquire ground data from SURFRAD, Atmospheric Radiation Measurement, and NREL locations for 2018, 2019, and 2020.
- Perform data quality assessment on the ground data.
- Evaluate 2018, 2019, and 2020 data from the NSRDB using ground-based information.
- Publish a report on the evaluation of the NSRDB data sets.

3.2 Project Results and Discussion

3.2.1 Hourly Data Validation

A validation of the performance of the NSRDB V3 1998–2020 was conducted to quantify the accuracy and spatial and temporal variability of the solar radiation data. Comparisons of the NSRDB V3 estimates with selected ground-measured data were conducted under both clear- and cloudy-sky conditions and covered the period from 1998–2020 for seven SURFRAD stations.

The comparison demonstrates that the difference between the NSRDB and surface measurements was within 5% (Figure 17). The MAE of the NSRDB was within 10%. There was significant reduction in the RMSE. The improvement is assumed to result from a combination of factors, including better downscaling methodologies, particularly in the interpolation and extrapolation used to align the multiple data sets to the same grid; the use of hourly values of aerosol information; and the use of surface albedo instead of interpolated monthly averages.

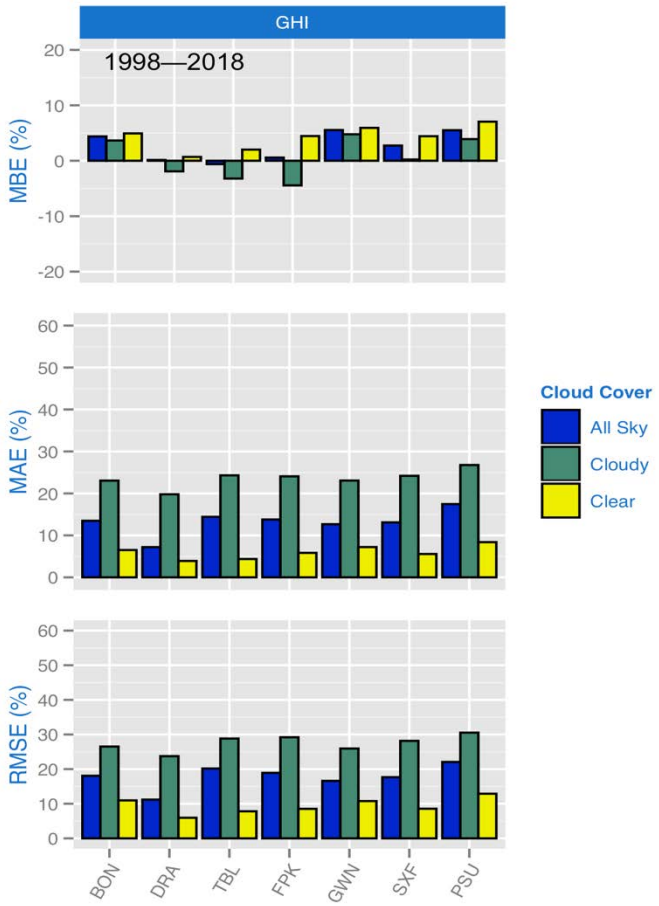

Figure 17. GHI statistical result for seven SURFRAD locations (1998-2018).

The distribution of irradiance from the NSRDB V3 1998–2018 versus ground measurements under varying sky conditions (Figure 18) showed good agreement. The clear-sky conditions demonstrated a close relationship with the measurements.

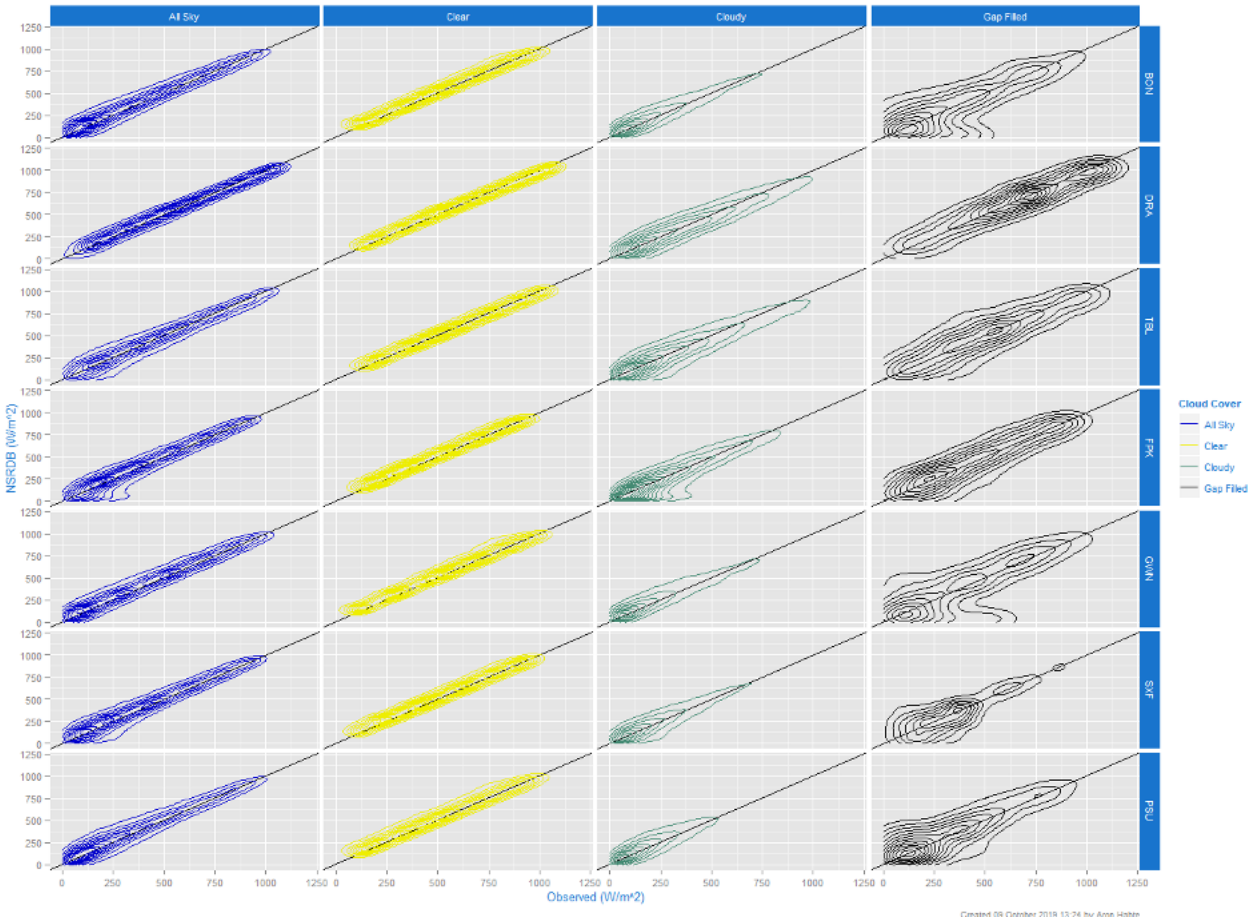

Figure 18. Irradiance distribution for the seven stations under varying sky conditions (1998–2018)

3.2.2 Five-Minute 2018 National Solar Radiation Database Data Validation

The original GOES-16 data are at 2-km, 5-minute resolution, and the solar radiation was generated at this resolution. These data are disseminated through Amazon Web Services (AWS). Prior to distribution, the high-resolution data were validated using high-quality ground measurement data. The results showed less than 5% bias under all sky conditions (Figure 19) but a positive bias (Figure 19) for clear-sky conditions. (See research section for the plan forward to solve this issue.) Moreover, compared to the 1998–2017 data, the overall uncertainty is significantly reduced for the 2018 high-resolution data set in the daily, monthly, and annual time averages (Figure 20).

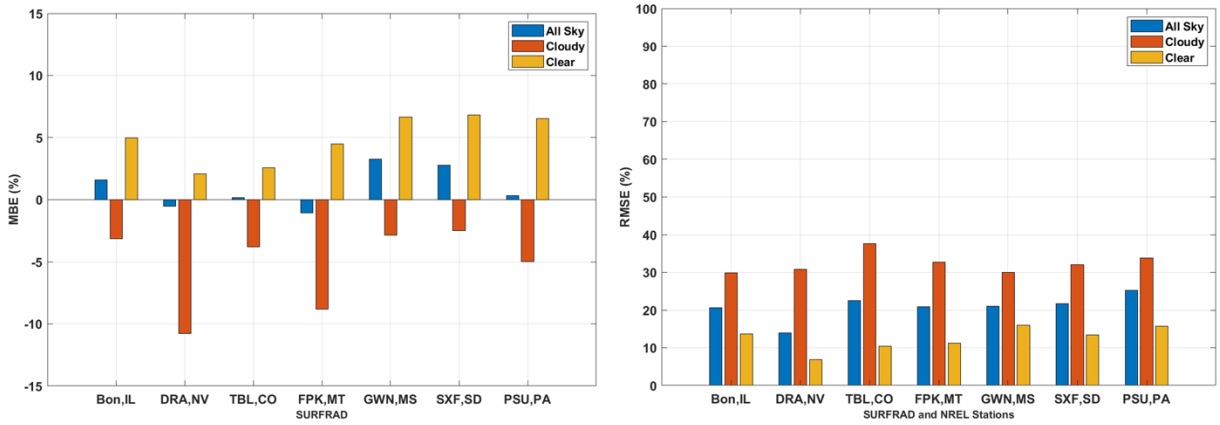

Figure 19. GHI MBE and RMSE in the 5-minute NSRDB 2018 data.

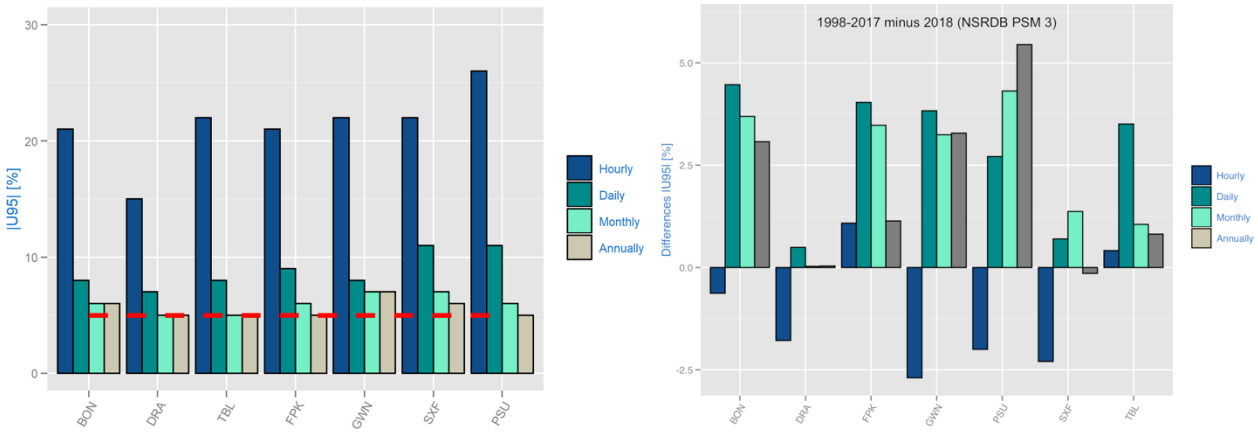

Figure 20. Overall GHI uncertainty comparison. (Left) NSRDB 2018 uncertainty and (right) the differences between NSRDB 1998-2017 and NSRDB 2018 data.

3.2.3 Validation of the NSRDB V3

A validation of the performance of the NSRDB V3 1998–2019 was conducted to quantify the accuracy and spatial and temporal variability of the solar radiation data. Comparisons of the NSRDB V3 1998–2019 and PSM V3 estimates with selected ground-measured data were conducted under both clear- and cloudy-sky conditions and covered the period from 1998–2019 for seven SURFRAD stations.

The comparison demonstrates that the difference between the NSRDB and surface measurements is within ±5% for GHI (Figure 21). Similar to the 1998–2018 data set, the data showed improvement, which could be the result of a combination of factors, including better downscaling methodologies, particularly in the interpolation and extrapolation used to align the multiple data sets to the same grid; the use of hourly values of aerosol information; and the use of surface albedo instead of interpolated monthly averages.

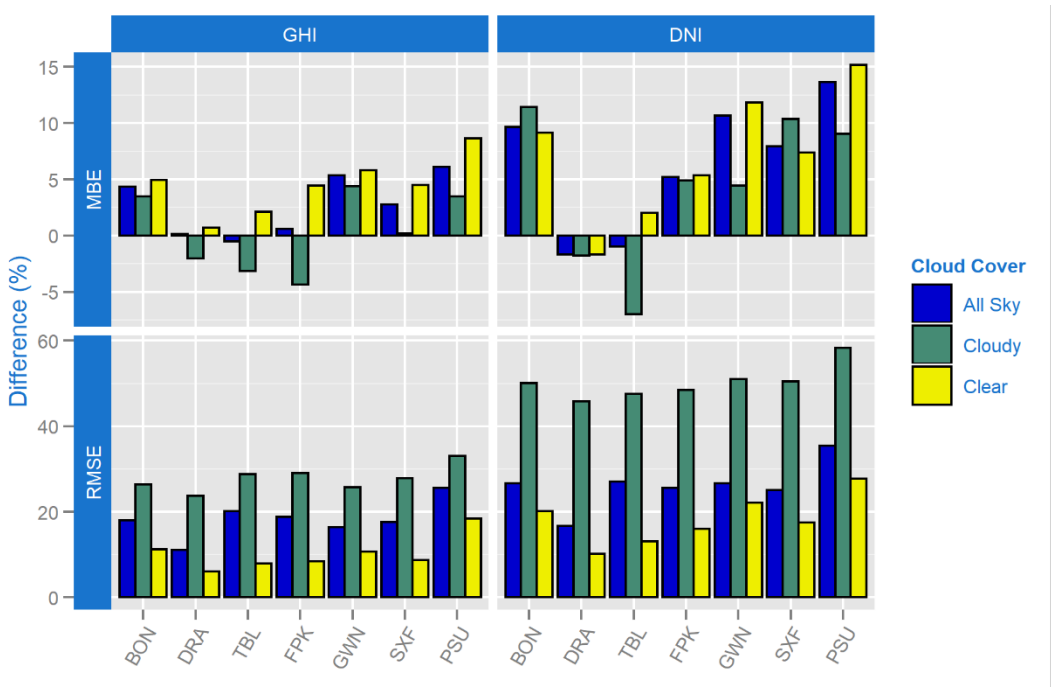

Figure 21. GHI and DNI statistical results for seven SURFRAD locations (1998–2019)

The original GOES-16 and GOES-17 data were at 2-km, 5-minute resolution, and solar radiation was generated at this resolution. Currently, these data are disseminated through various ways. Prior to distribution, the high-resolution data were validated using good quality ground measurement data. The results showed less than 5% bias under all sky conditions (Figure 22) but a positive bias for clear-sky conditions.

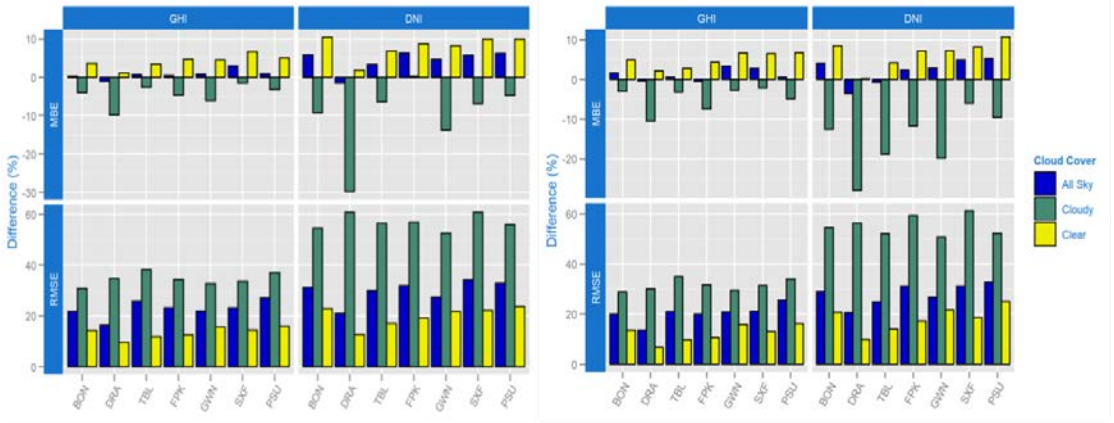

Figure 22. Five-minute MBE and RMSE for both GHI and DNI for (left) 2019 and (right) 2018

The comparison between the 2018 and 2019 data demonstrated similar results, which provides confidence in the quality of the 2019 data set. The uncertainty of both the 2018 and 2019 5-minute data sets was also compared (Figure 23), and the results were similar.

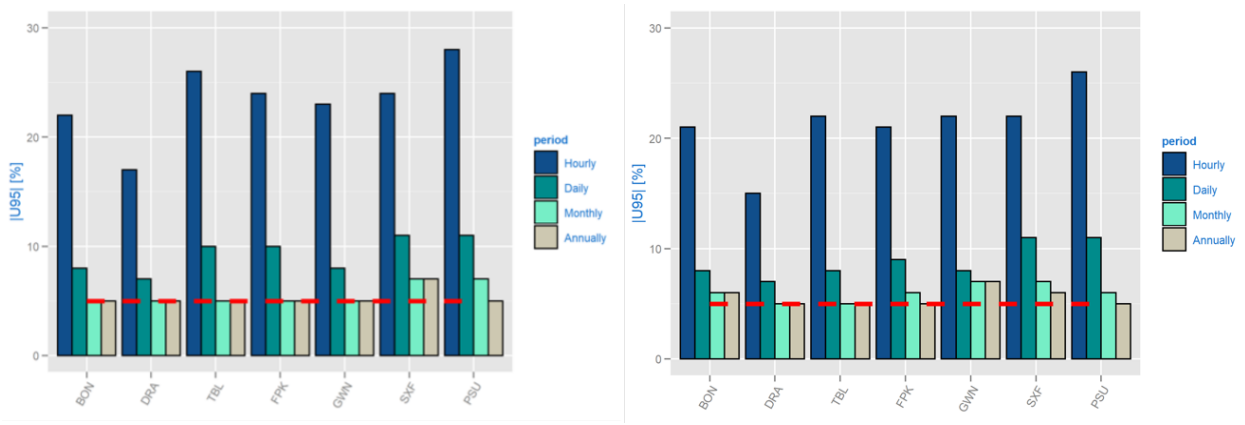

Figure 23. Five-minute uncertainty for GHI for the (left) 2019 and (right) 2018 5-minute data sets under all sky conditions

3.2.4 Validation of the 2-km, 5-minute NSRDB

A validation was carried out using the 2-km, 5-minute resolution NSRDB solar radiation data for 2018–2020. Currently, these data are disseminated through various ways. Prior to distribution, the high-resolution data were validated using good quality ground measurement data from the National Oceanic and Atmospheric Administration SURFRAD (Figure 24). The spatial and temporal differences between the ground-measured and satellite-derived data sets were analyzed. The ground measurements were averaged to 5-minute values centered at the satellite time stamp. This served two purposes: (1) to convert a point measurement to a representation of a finite area covered by a satellite pixel and (2) to provide a 5-minute average estimate that the satellite data is meant to represent. Further, investigating the differences and setting a uniform benchmark is essential to improving the existing satellite-derived data or creating other satellite-based methods to improve the underlying uncertainties.

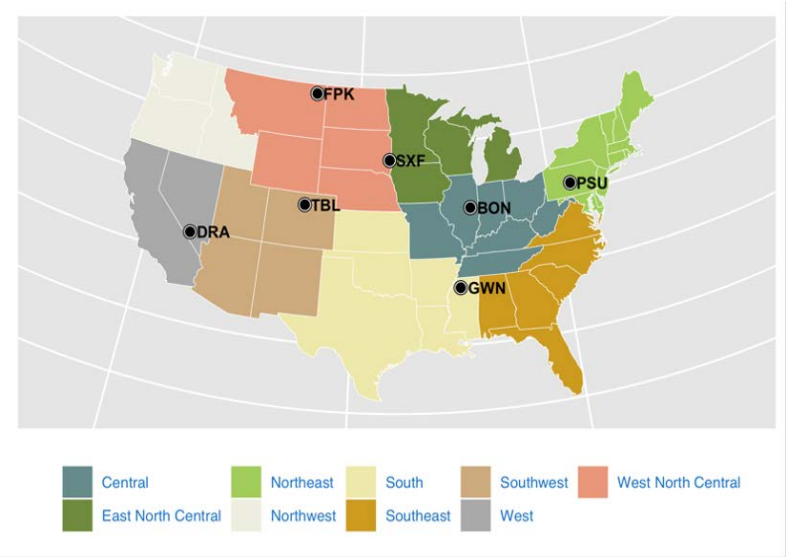

Figure 24. Ground measurement locations for the SURFRAD network

To conduct a valid and reasonable data comparison, data-filtering techniques were implemented for both the reference ground measurements and the NSRDB (2018–2020). The following criteria were used to produce a quality-controlled data set for this validation study:

- Solar zenith angles must be less than 80 degrees.
- GHI and DNI irradiances must be greater than 50 W/m^2.
- Data records with missing values from the surface measurements were excluded from both the surface measurements and the NSRDB data sets.
- Cloud types from the satellite data were used to determine the sky conditions.

Figure 25 shows the differences when the ground-measured data were compared to the NSRDB satellite-derived data. The results demonstrate that the clear-sky conditions showed better statistical metrics than the cloudy conditions; however, the DNI difference between the measured and NSRDB data set was higher than the GHI, especially under cloudy-sky conditions. Further, in most cases, the clear-sky biases tend to be positive, and the reverse is true for the cloudy conditions.

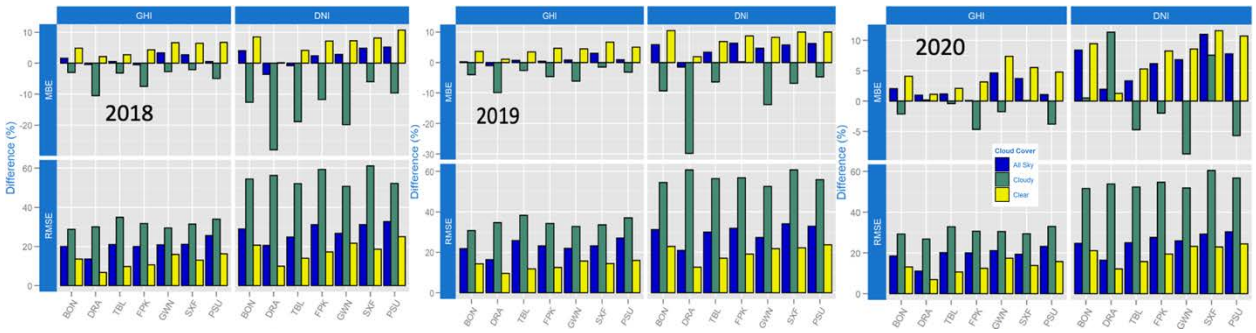

Figure 25. GHI and DNI MBE in percent for 2018-2020 for the seven SURFRAD sites.

The MBE in percentage for the years from 2018–2020 for all seven stations is shown in Figure 25. Under all sky conditions, the MBE, on average, is approximately 5% for GHI and 7% for DNI. Particularly for DNI, the MBE values under cloudy conditions sometimes reach less than 10%, but these could be related to the field-of-view difference between the satellite pixel and the point measurement of the station ground-measuring radiometer, which has an approximate 5-degree field of view, but the satellite DNI represents the 4-km by 4-km area. This could be coupled with the presence of clouds away from the ground measurement field of view, which can trigger higher deviations. Further, these could also be a result of the misidentification of bright surfaces, snow, and fog as clouds by the satellite retrievals. The methodology used in the satellite retrievals to distinguish clouds from snow or vice versa continues to present challenges. Future improvement of the NSRDB will address these situations.

4 Task 4: Data Dissemination

4.1 Project Approach

- Develop the capability to provide additional data through the NSRDB website by enhancing/updating servers.
- Update the web interface for delivering new data sets.
- Update the website content to reflect changes and add new content.
- Maintain the operational capability of the NSRDB website and servers and enhance the storage/server capability to provide new data sets.
- Provide user support for NSRDB users.
- Develop a documentation and methodology for serving data through AWS.
- Process data for upload to AWS.
- Investigate the feasibility of providing the high-resolution (2-km x 2-km, 5-minute) data through AWS and NREL servers.
- Provide a webinar to inform users of the update.
- Enhance the capability to store and deliver high-resolution GOES-16 and GOES-17 data sets.

4.2 Project Results and Discussion

4.2.1 National Solar Radiation Database Website and Server Update

With the next-generation satellite now operational, the data volumes increased by a factor of 24, and there was an urgent need to develop a capability to handle these large data volumes and deliver NSRDB updates on time. Due to these requirements, the group had to build and support its own storage and processing system that can move data from the project's space on NREL's high-performance computing (HPC) environment for permanent storage and ongoing analysis. The team worked with NREL's HPC center and information technology services to upgrade the system. Further, the team worked on updating the NSRDB website content to reflect the changes.

4.2.2 Amazon Web Services Updates

The NSRDB V3 was loaded into a Highly Scalable Data Service (HSDS). An example of accessing data using HSDS in Python was written and distributed. NSRDB V3 files in Hierarchical Data Service 5 (HDF5) format were also uploaded to AWS (s3://nrel-pds-nsrdb/hdf5-source-files-v3/).

NREL provided a webinar on "AWS Data Informational Session: NREL's National Solar Radiation Database (NSRDB)" (https://nsrdb.nrel.gov/about/announcements).

4.2.3 The Case for Custom Typical Meteorological Years

NREL presented a paper at the 46th Institute of Electrical and Electronics Engineers (IEEE) Photovoltaic Specialists Conference titled "The Case for Custom TMY's: Examples Using the NSRDB."[2] TMYs were primarily developed for building simulations, and the meteorological

[2] See https://www.nrel.gov/docs/fy19osti/74063.pdf.

variables are weighted to represent their influence on energy use in buildings. Multiyear, monthly, weighted data are used to identify the most representative month. The most representative months are then concatenated to construct a single-year TMY. For modeling solar generation from different technologies, various special cases of TMY can be constructed. For example, a typical GHI year or a typical DNI year can be constructed using only GHI or DNI, respectively, by weighting the other variables to 0. TMYs or typical GHI years are widely used in the System Advisor Model (SAM), PVWatts®, PVSyst, and PlantPredict to simulate generation from photovoltaic (PV) plants. The paper demonstrated that a plane-of-array (POA) TMY generated by selecting median months from a multiyear POA irradiance time-series data set produces significantly different results than a POA TMY generated by transposing a TMY data set constructed from horizontal data. The study used the Perez transposition model on long-term NSRDB time-series data for six locations to generate POA irradiance data for fixed-latitude tilt and single-axis tracking (east-to-west tracking) orientations because they are widely used in PV deployments. This method generates a typical POA year (TPY) by setting the weight for all variables except POA to 0. For comparison, we use the same Perez transposition model to generate a POA irradiance typical GHI year for the same pixel. For this paper, we call this the typical GHI-based POA year (TGPY) (currently commonly used). The TPYs generated from fixed-latitude tilt and single-axis tracking orientations are then compared to the corresponding TGPYs. For each TPY and the TGPY, the irradiance are averaged on a monthly basis. A percentage difference of the TPY compared to the TGPY for each month is calculated at each location using (TPY – TGPY) / TGPY x 100.

The comparison results point to the need for generating TMYs using POA irradiance time series representing the orientation at which PV panels will be deployed.

4.2.4 National Solar Radiation Database Website and Server Update

NREL updated the mechanism to deliver data through updating physical servers and/or AWS. Concurrently, the web interface was updated to accommodate these changes. We developed a capability to transfer 100% of the incoming traffic to AWS whenever needed. Moreover, this project continues to support its own storage and processing system that can move data from the project's space on the NREL HPC for permanent storage and ongoing analysis. The team worked with NREL's HPC center and information technology services to upgrade the systems (see Figure 26 for the architectural layout of the NREL system), including:

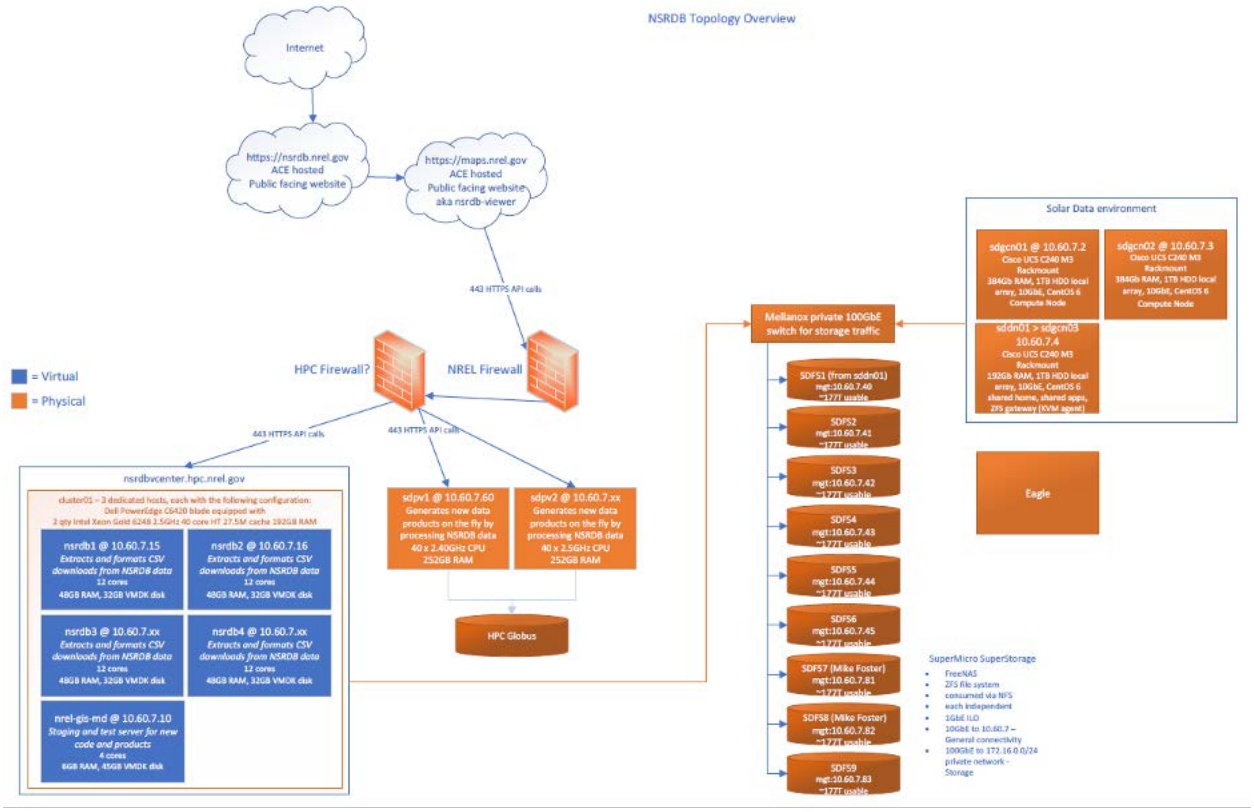

Figure 26. NSRDB servers and HPC assimilation architecture

- Doubled the amount of server power for the NSRDB and spectral processing using on-site hardware purchases
- Migrated all NSRDB data set requests to AWS S3
- Deployed the suite of dedicated HSDS (on AWS) servers to offer highly reliable and high-performing downselecting of all data sets
- Completed the implementation of the architectural overhaul of the entire suite of download application programming interfaces for NSRDB data. The new system leverages asynchronous job queues, HSDS services, data processing swarms, and a combination of both on-site and cloud-hosted infrastructure to achieve a dramatically improved level of performance, reliability, and flexibility.
- Updated the NSRDB viewer to use the new application programming interfaces
- Updated the documentation on data access available both via the NSRDB static site and developer.nrel.gov.

Further, the team worked on updating the NSRDB website content to reflect the changes. The NSRDB website now shows a flow diagram demonstrating the provenance of the original data sets and their manipulation for the generation of the NSRDB data (Figure 27). Old NSRDB data, including TMY2 and TMY3, are now located on the NSRDB website. The RReDC website was decommissioned.

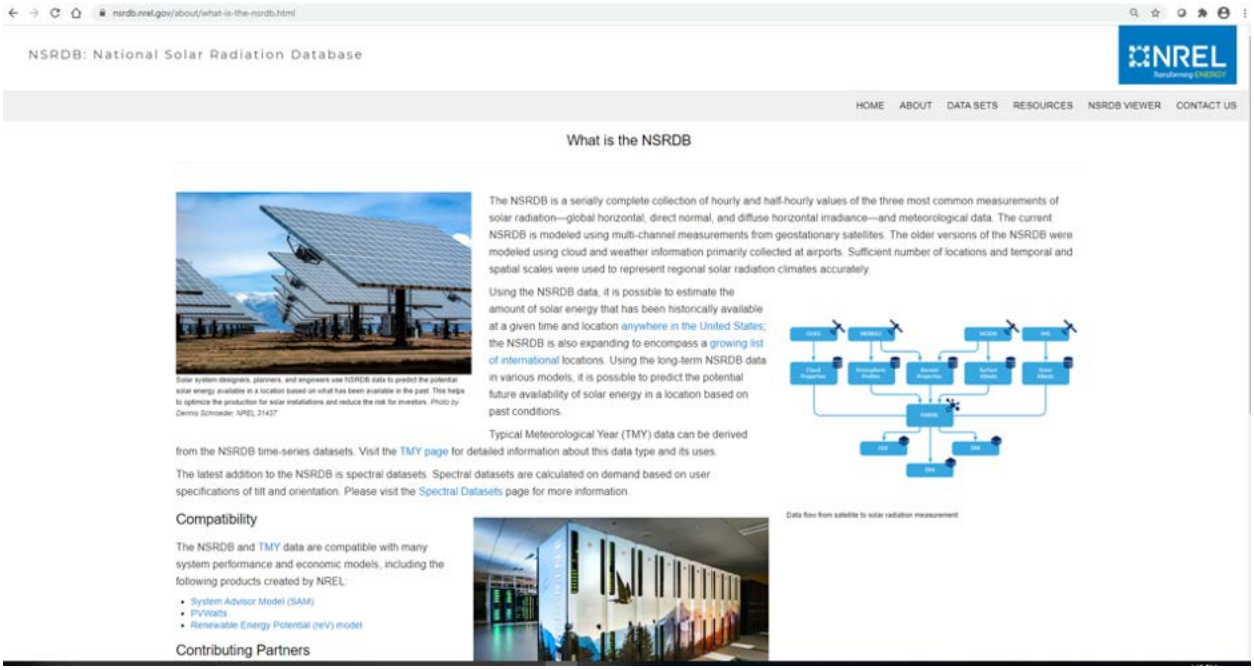

Figure 27. Flowchart of the NSRDB PSM

4.2.5 National Solar Radiation Database Webinar

The NSRDB team held a successful NSRDB webinar on October 6, 2020. Approximately 2,000 participants registered for the webinar; however, only 500 were able to attend because of the webinar application limit. A recording of the webinar and supporting materials are available at https://nsrdb.nrel.gov/about/announcements.html and https://sam.nrel.gov/weather-data/weather-data-videos.html.

4.2.6 Performance of the Fast All-Sky Model for Solar Applications-Physical Solar Model Under Thin Cirrus Clouds

NREL retrieved cloud properties for thin cirrus clouds using the GOES 1.375-um channel. We considered thin cirrus clouds in the computation of GHI using the Fast All-Sky Model for Solar Applications (FARMS). The new computation has better agreement with NREL's Solar Radiation Research Laboratory observations (Figure 28). Additional investigation using 1-year data over the Solar Radiation Research Laboratory is required to understand the long-term impact.

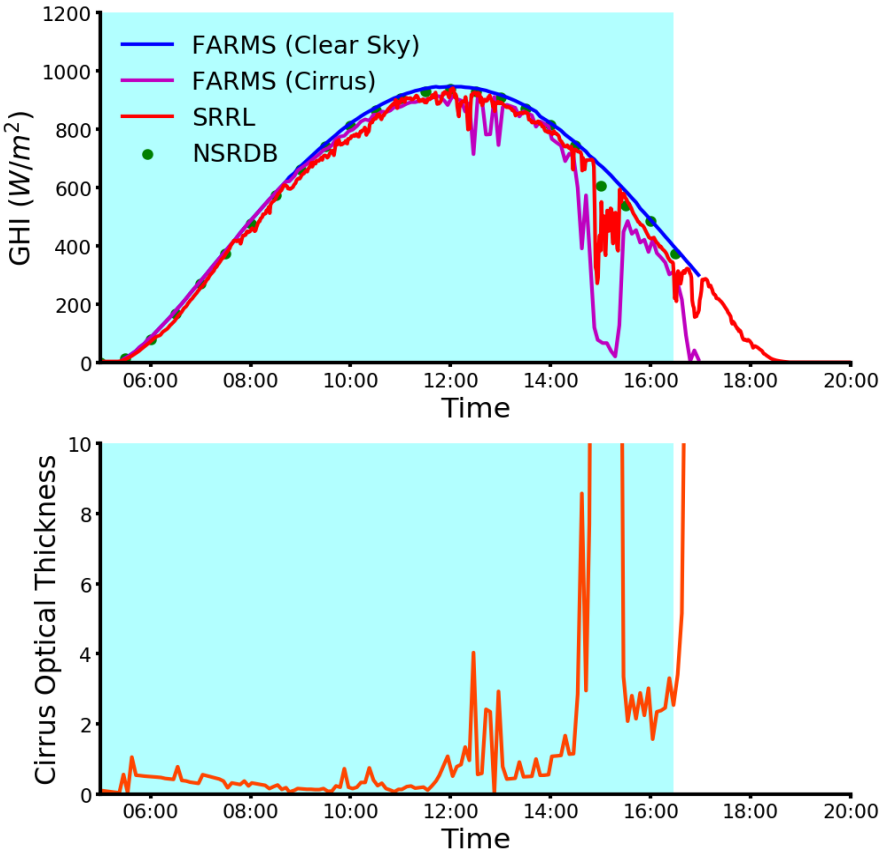

Figure 28. Comparison of FARMS with measurements at NREL's Solar Radiation Research Laboratory

4.2.7 Surface Albedo

NREL prepared a journal article, in collaboration with Solar Consulting Services, titled "Surface Albedo Spatial Variability: Local Measurements vs. Gridded Data" in *Solar Energy*.[3] The paper compared the NSRDB albedo data to ground and other sources of albedo data. The comparison was done for a single year (2015). Using both ground-based measurements and satellite-derived data at 36 test sites in North America, the spatial variability of 500-m and 1-km Moderate Resolution Imaging Spectroradiometer-based albedo products was evaluated within the corresponding 4-km NSRDB pixel. The 36 sites represent a wide variety of latitude, climate, terrain, and environmental (land use) situations.

The paper also described that in addition to the seasonal variations, it is always possible that interannual and long-term trends modify the results over time. In particular, rainy years tend to induce a lower albedo, whereas snowy years have a higher albedo. Moreover, each pixel or measurement station can be affected by changes in vegetation, soil structure, urbanization, etc. Further, the paper demonstrated the variability of albedo values within the NSRDB pixel (Figure 29). The findings confirm the issues that affect the current NSRDB albedo data because of its treatment of albedo in snow situations (fresh snow albedo is used when snow cover is detected),

[3] See https://doi.org/10.1016/j.solener.2021.05.012.

both in terms of magnitude and temporal variation. It is recommended that a more general and physical algorithm be developed to resolve these issues.

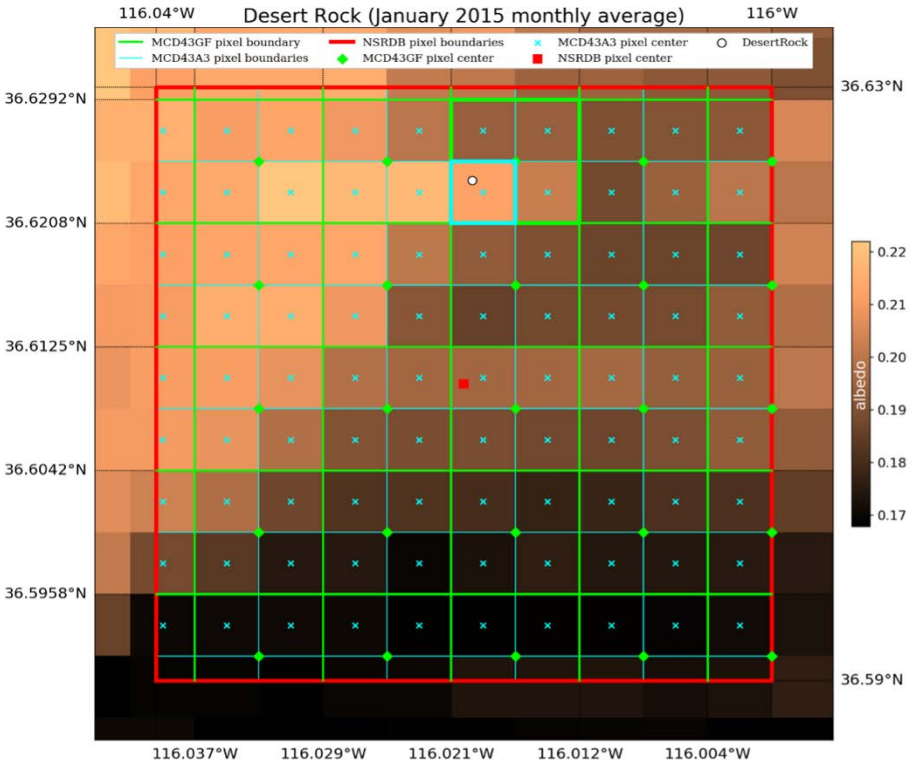

Figure 29. NSRDB, MCD43A3, and MCD43GF pixels relative to the Desert Rock (DRA) site (white dot) using the mean annual MCD43A3 albedo as background

4.2.8 Long-Term Solar Resource Variability

NREL published a journal article titled "Long-Term Spatial and Temporal Solar Resource Variability Over America Using the NSRDB Version 3 (1998–2017)" in *Renewable & Sustainable Energy Reviews*.[4] The NSRDB is used to assess long-term variability. Over the years, the NSRDB has evolved from a point-source solar resource to a pixel-based, high-temporal and high-spatial-resolution database. As of August 2019, NSRDB V3 contains data from 1998–2017 on a half-hourly and 4-km x 4-km temporal and spatial resolution and gets updated with a 1-year lag each year. This study provides the spatial and temporal trend of the GHI and DNI using long-term, 20-year NSRDB data (see Figs. 4-10 in Habte et al., 2020). For each pixel, the temporal variability's coefficient of variation (COV) is obtained by calculating the annual mean daily total irradiance of each individual year (between 1998 and 2017) as well as the mean daily total irradiance over the 20-year period. The COV was used to analyze the spatiotemporal interannual and seasonal variabilities. The variability was assessed using the coefficient of variation for the individual pixel. The highest temporal variability is found over northern Canada, the northern and southern United States, and northwestern South America. On the other hand, the central part of the United States reveals less variability. The spatial variability

[4] See https://doi.org/10.1016/j.rser.2020.110285.

was analyzed by comparing the center pixel to neighboring pixels. The spatial variability result showed a higher coefficient of variation as the number of neighboring pixels increased (Figure 30).

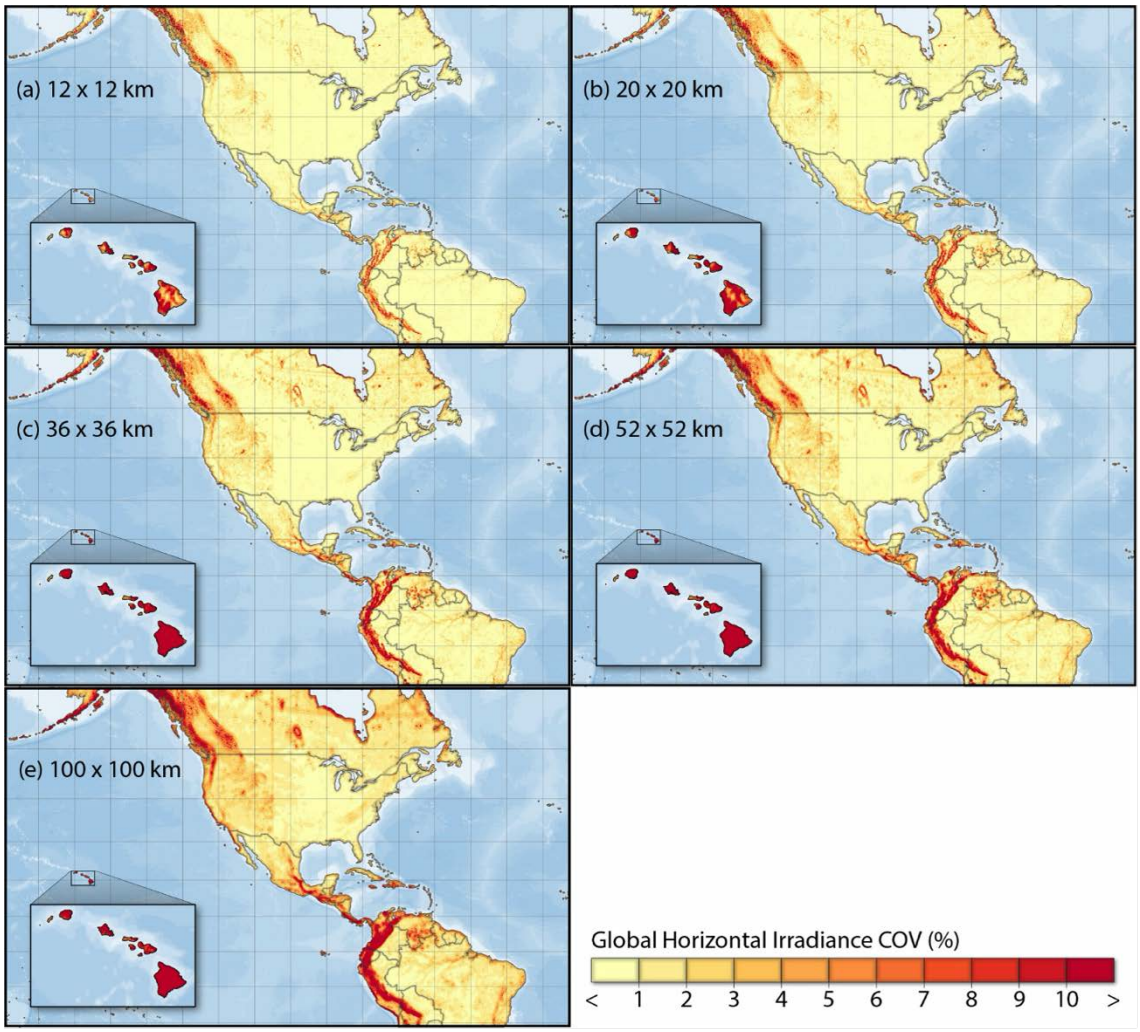

Figure 30. GHI spatial variability for a multiple-pixel matrix (area) for the 3-km x 3-km, 5-km x 5-km, 9-km x 9-km, 13-km x 13-km, and 25-km x 25-km areas

4.2.9 Solar Resource Anomaly

NREL presented a paper at the 47th IEEE Photovoltaic Specialists Conference titled "Annual Solar Irradiance Anomaly Features Over the USA During 1998–2017 Using NSRDB V3."[5] It is known that annual solar irradiance anomalies (or departures from the long-term mean annual value) have a direct impact on various phases of solar energy projects, from prefeasibility studies to technical deployment decisions. Anomalies can happen because of normal climate variability or exceptional weather patterns. This study investigated such anomalies for both GHI and DNI using the NSRDB V3 and surface irradiance measurements at eight U.S. locations. At each site,

[5] See https://www.nrel.gov/docs/fy20osti/76858.pdf.

the annual anomaly for each specific year from 1998–2017 was analyzed by evaluating the irradiance deviation from the long-term average. A positive/negative anomaly indicated that the solar resource was higher/lower than the long-term average during that specific year. The results showed that in most cases the anomaly was within ±5% for GHI and ±10% for DNI using either ground-based irradiance measurements or modeled data from the NSRDB (Figure 31).

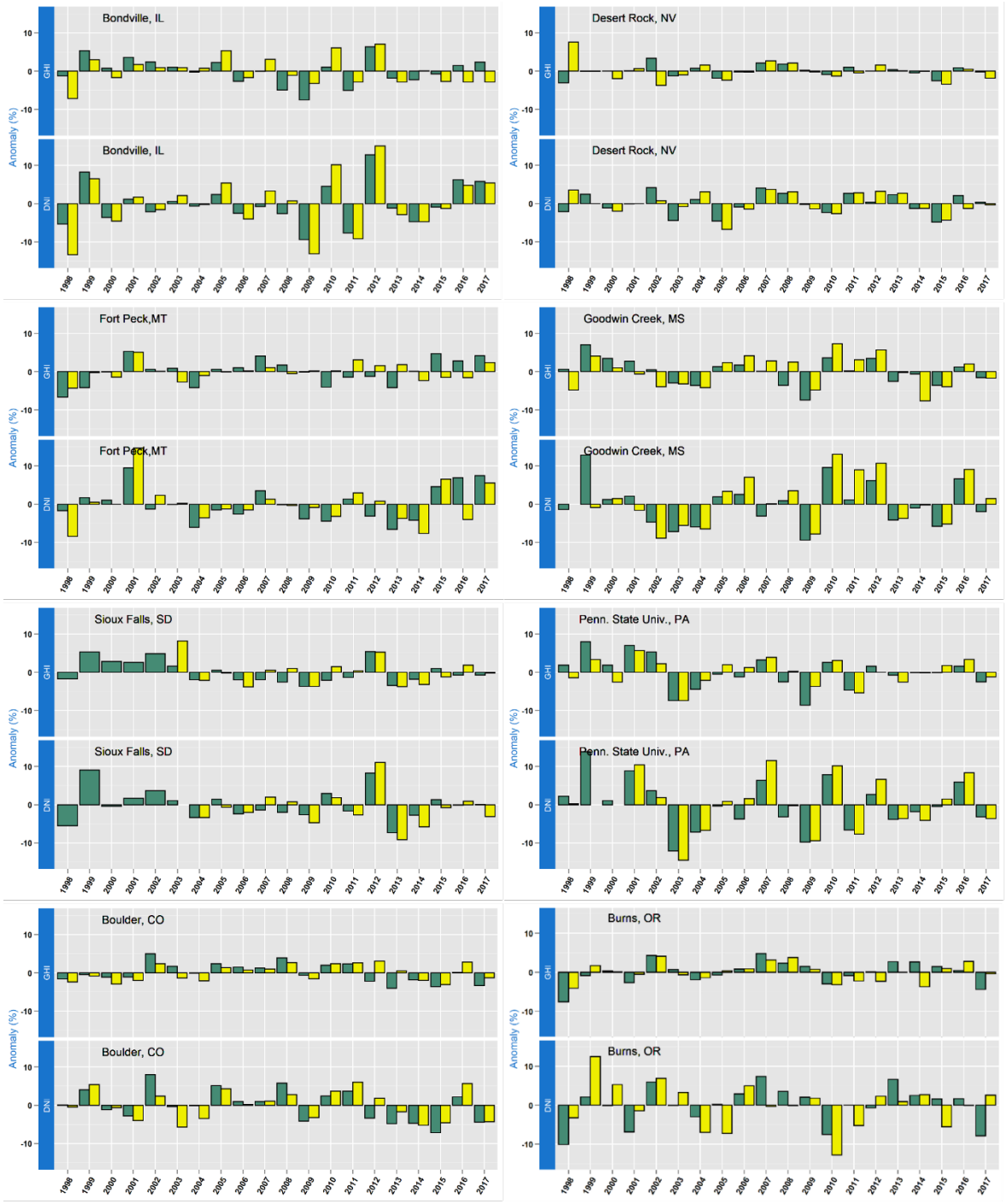

Figure 31. Percentage annual anomaly of ground measurements (yellow bar) and NSRDB (green bar) for GHI (top row) and DNI (bottom row) at eight locations from 1998–2017 (X-axis)

4.2.10 Evaluation of FARMS-NIT

In our previous work supported by this project, we developed a comprehensive capability to provide spectrally resolved solar resource data from the NSRDB, effectively providing 20 years of half-hourly data for all of the United States at a 4-km by 4-km spatial resolution. We integrated the advantages of the current models and developed an innovative radiative transfer

model, the Fast All-Sky Radiation Model for Solar Applications with Narrowband Irradiances on Tilted Surfaces (FARMS-NIT), to efficiently compute irradiances on inclined PV panels for 2002 narrow-wavelength bands from 0.28 μm–4.0 μm. For clear-sky conditions, the Simple Model of the Atmospheric Radiative Transfer of Sunshine (SMARTS) is employed to rapidly provide the optical properties of a given clear-sky atmosphere. The clear-sky radiances in the narrow-wavelength bands are computed by considering three paths of photon transmission and solving the radiative transfer equation with the single-scattering approximation. The bidirectional transmittance distribution function of aerosols is given by their single-scattering phase function with a correction using a two-stream approximation. For cloudy-sky conditions, FARMS-NIT uses cloud reflectance of irradiance and the bidirectional transmittance distribution function from a precomputed lookup table by the LibRadtran model with a 32-stream discrete ordinates radiative transfer. The cloud reflectance and bidirectional transmittance distribution function are combined with the clear-sky properties to efficiently compute spectral radiances on the land surface and POA irradiances.

These data are now available free to users directly through a geographic information system-based Web interface (https://nsrdb.nrel.gov) as well as through an application programming interface. Users of these data can conduct more accurate prefeasibility studies and assess multiple PV technologies. To promote the use of the solar resource data by widely used models—such as PVSyst; NREL's SAM; and PlantPredict, designed by First Solar—it is important to understand the performance and accuracy of the data as evaluated by surface-based observations. We analyzed long-term observations (2013–2018) from six surface sites operated by First Solar to evaluate the solar resource data provided by the NSRDB. Figure 32 illustrates the locations of the surface sites at Picture Rocks, Arizona; Neenach, California; Deming, New Mexico; Calipatria, California; Sarnia, Canada; and London, Canada. Note that the blue spots represent observations from locations where fixed-tilt measurements are available. The red spots represent locations where 1-axis tracking measurements are available. The green spots indicate locations where both fixed-tilt and 1-axis tracking observations are available. The GHI are measured by all sites using a number of Kipp & Zonen CMP11 pyranometers. Figure 33 compares GHI between surface observations and solar resource data from the NSRDB. The blue bands represent the variations in surface observations by different pyranometers. It is clear that the solar resource data have reasonable agreement over all sites under both clear-sky and cloudy-sky conditions; however, the 30-minute resolution solar resource data are insufficient to present the rapid variations in solar radiation that are shown in the 1-minute observations (e.g., in Figures 33d and 33f). This bias should be reduced using the next generation of satellite data with an improved temporal resolution. Further, the observational uncertainty given by different instruments might affect the accuracy of the evaluation, as shown in Figure 33b.

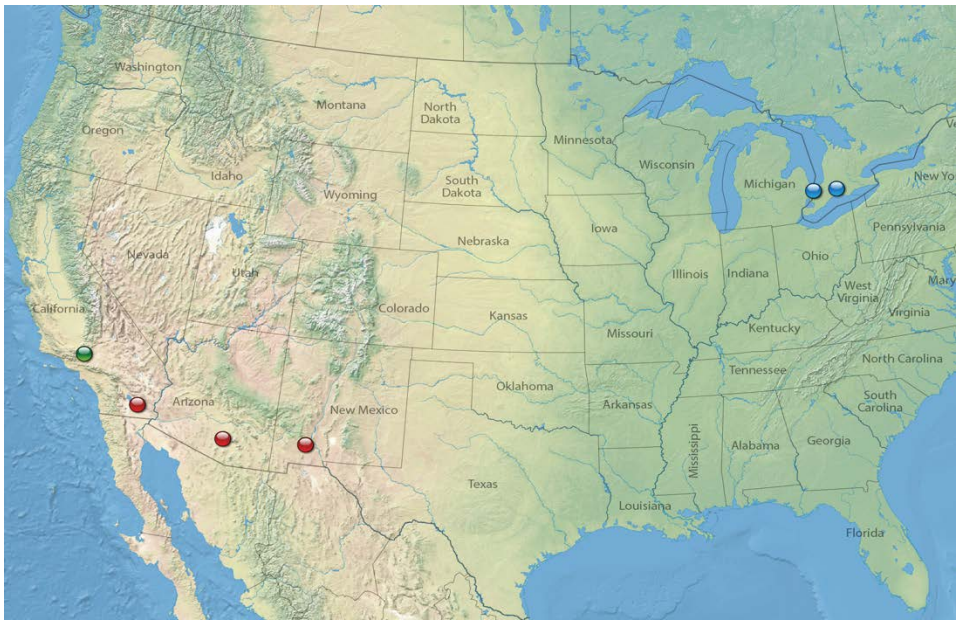

Figure 32. Locations of the six surface sites used in the evaluation. The blue spots represent that observations from fixed-tilt surfaces are available. The red spots represent that those from 1-axis trackers are available. The green spots indicate that both fixed-tilt and 1-axis tracking observations are available.

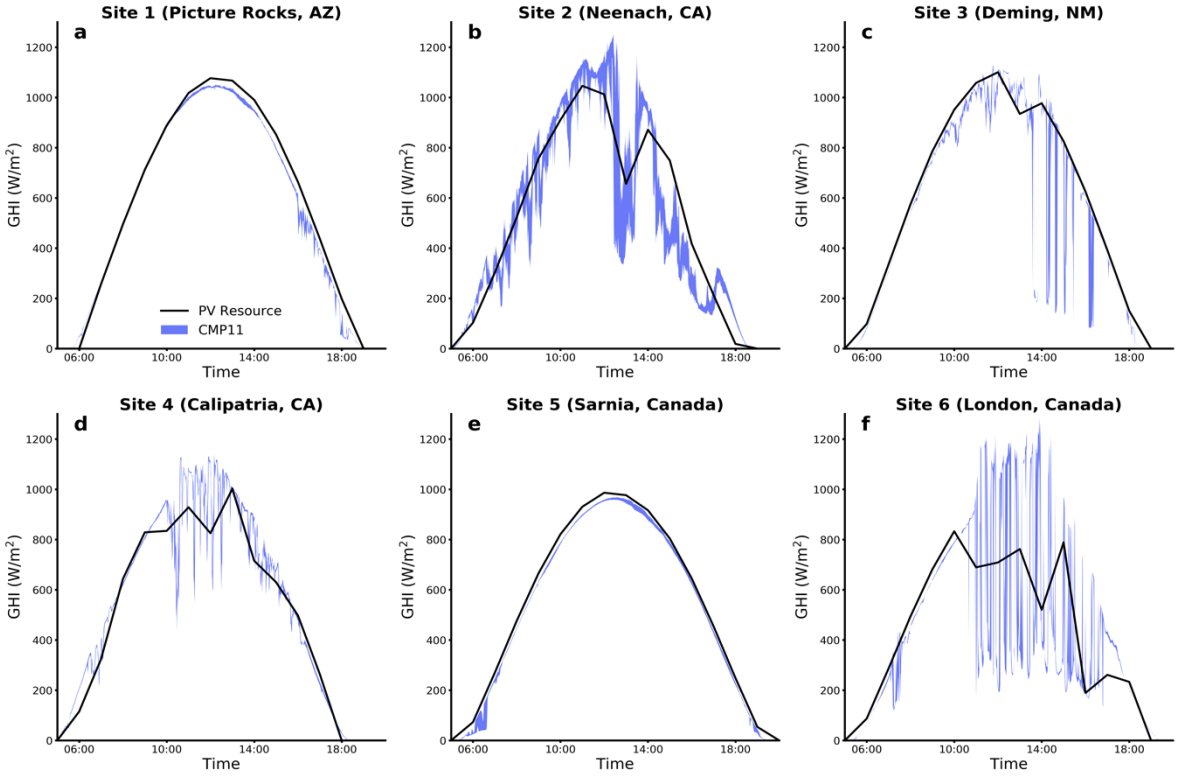

Figure 33. Comparisons of GHI between surface observations (by CMP11) and the NSRDB data on May 1, 2015

4.2.11 Data Dissemination

NSRDB data continue to be disseminated through the NSRDB viewer by the AWS mechanism. Efforts are underway to ascertain whether using a lambda-backed HSDS service can provide cost savings to these operational expenses.

The NSRDB users continue to increase as shown in Figure 34. Industry, labs, and academia outside of NREL also use the NSRDB.

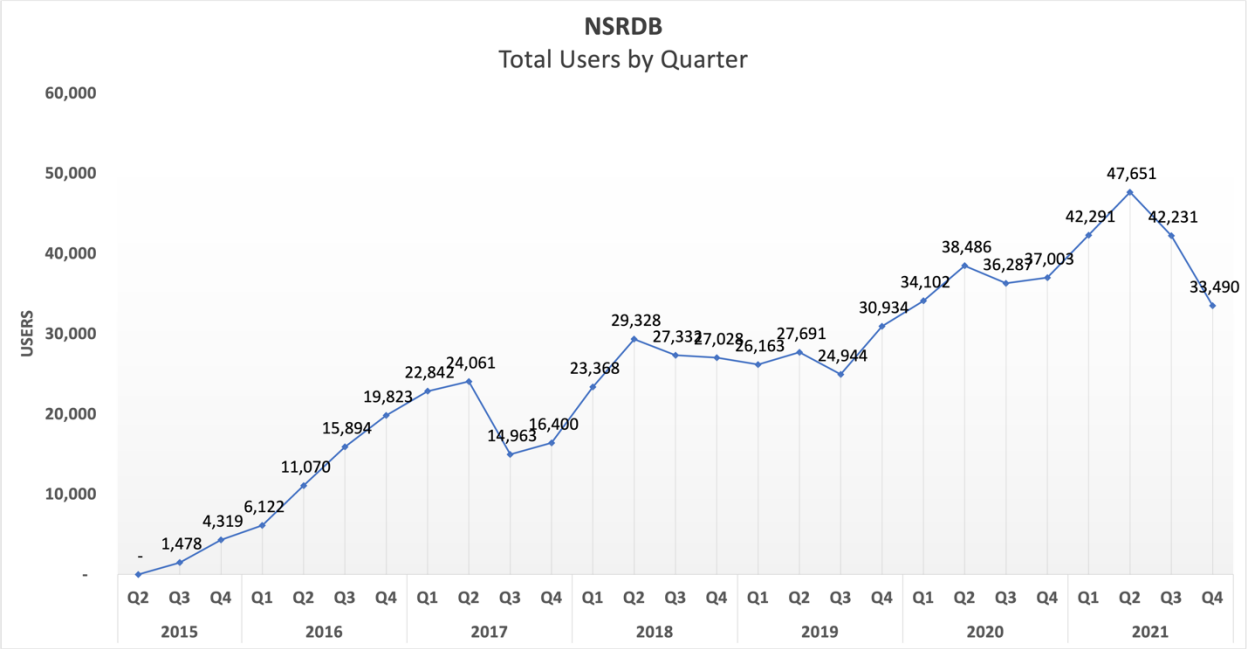

Figure 34. User statistics for the NSRDB website

NREL tracks visits to the NSRDB website.[6] Figure 34 shows user statistics for the quarter.

4.2.12 Development of Typical Plane-of-Array Year

It is known that TMY data sets provide industry-standard resource information for building designers and are commonly used by the solar industry to estimate PV and CSP system performance. They are the basis for system performance and economic models, such as PVWatts[7] and SAM[8]; however, the current application of this data set contains some limitations that affects the fidelity of this data set.

The TMY data are created by concatenating 12 typical meteorological months from statistically analyzed and selected individual months from the entire set of available years of GHI, DNI, and additional ancillary meteorological data sets. This TMY data set is used in solar energy modeling

[6] See https://nsrdb.nrel.gov.
[7] See http://www.nrel.gov/rredc/pvwatts/.
[8] See https://sam.nrel.gov/.

by transposing the data set to the desired tilt to assess the available solar energy in a location. NREL presented a paper at the IEEE Photovoltaic Specialists Conference to demonstrate the limitation of the TMY data set for solar energy. The conference paper was for only a few locations; however, it demonstrated the analysis at the regional scale and at a higher spatiotemporal resolution, and NREL is performing the analysis using a similar method for the full NSRDB domain (1998–2018) and by using a simulated energy yield. Default system information was used—for example, the DC-to-AC ratio was assumed to be 1 for both array types to reduce clipping effects.

POA irradiance data for both fixed-tilt and single-axis orientations were used to generate a TPY by changing the weights of the irradiance variables. For comparison, we also used the existing TMY for the same pixel. Figure 35 shows the process that was carried out in this study. Then the study analyzed the differences in energy yield predictions, which were simulated using PVWatts. The study also analyzed the differences using a nonparametric statistic metric called the Kolmogorov-Smirnov test integral. This test has the advantage of being nonparametric so it is not dependent on the distribution. It compares the similarity of the two distributions, where a lower value of the Kolmogorov-Smirnov test integral indicates that the two CDFs are similar.

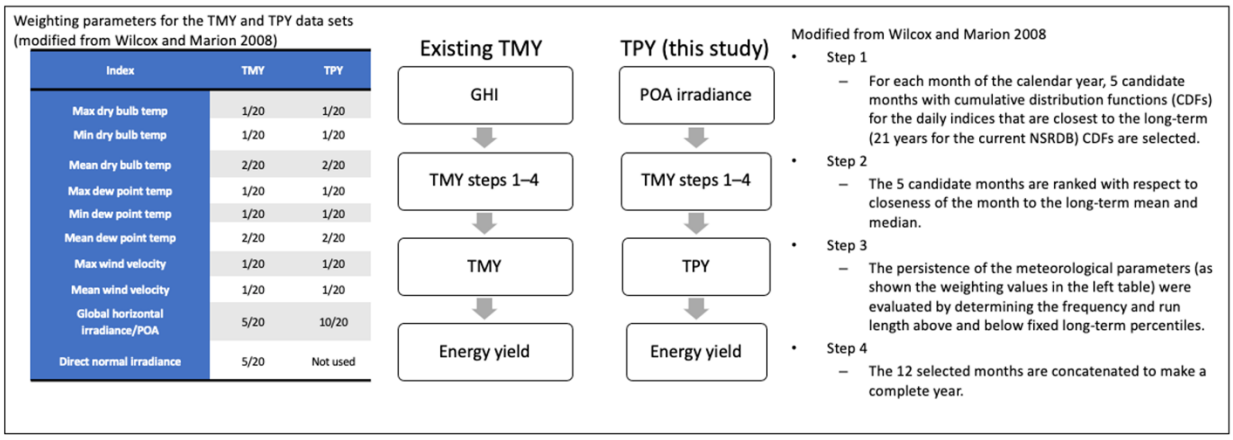

Figure 35. Data development process for both TMY and TPY

The results in Figure 36 are partitioned by Köppen-Geiger climate zones (Table 3) to provide insight into interpreting the results. For example, the preliminary result in Figure 36 is for the fixed-tilt orientation for CONUS in the summer and winter months (January and July). The figures demonstrate that the existing TMY data set overestimates energy yield by approximately 30%–35%, as shown by the MBE statistic, compared to TPY for most climate zones; however, the snow or continental climate zones, labeled D**, demonstrated fewer differences in the winter months and slightly more differences in the summer months. These climate zones, which normally occur north of 40°N latitude, are depicted by an average temperature warmer than 10°C in the warmest months and colder than −3°C in the coldest month. On the other hand, the single-axis array comparison demonstrated a relatively small difference between TMY and TPY. The detailed results will be included in an upcoming paper.

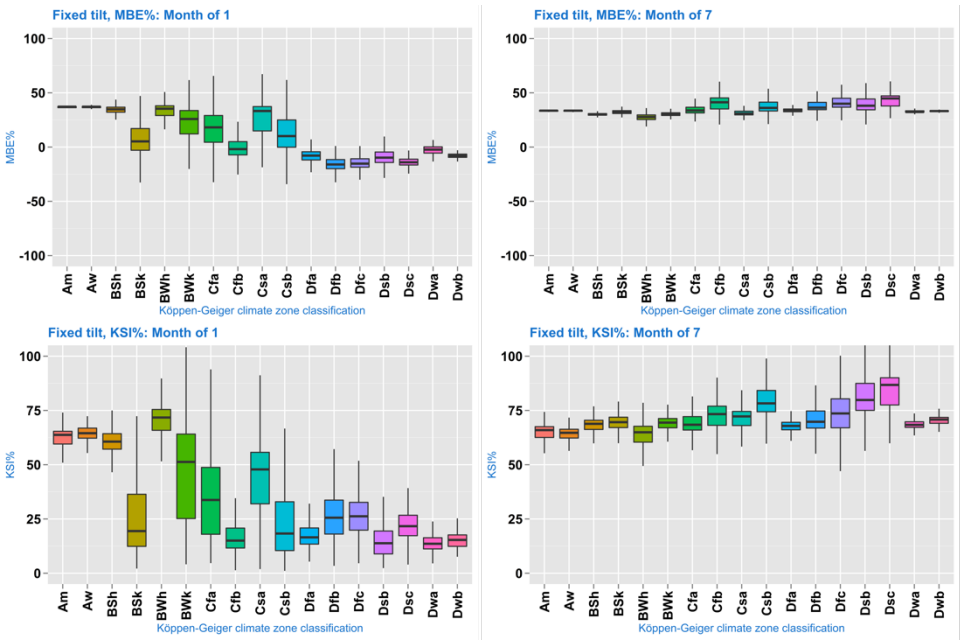

Figure 36. Differences between TMY and TPY. (Top row) MBE percentage and (bottom row) Kolmogorov-Smirnov test integral percentage classified by climate zones for (left) January and (right) July. Note: The statistics are based on the daily average capacity factors.

Table 3. Climate Zones Based on Köppen-Geiger

Main Climates	Precipitation	Temperature
A: Equatorial/tropical	W: Desert	h: Hot arid
B: Arid and semiarid	S: Steppe	k: Cold arid
C: Warm temperate	f: Fully humid	a: Hot summer
D: Snow	s: Summer dry	b: Warm summer
E: Polar	w: Winter dry	c: Cool summer
	m: Monsoonal	d: Extremely continental
		f: Polar frost
		T: Polar tundra

Publications

1. Gueymard, Christian A., Vincente Lara-Lanego, Manajit Sengupta, and Yu Xie. 2019. "Surface Albedo and Reflectance: Review of Definitions, Angular and Spectral Effects, and Intercomparison of Major Data Sources in Support of Advanced Solar Irradiance Modeling over the Americas." *Solar Energy* 182: 194–212.

2. Hanrieder, Natalie, Abdellatif Ghennioui, Ahmed Alami Merrouni, Stefan Wilbert, Florian Wiesinger, Manajit Sengupta, Luis Zarzalejo, and Alexander Schade. 2019. "Atmospheric Transmittance Model Validation for CSP Tower Plants." *Remote Sensing* 11. https://doi.org/10.3390/rs11091083.

3. Xie, Yu, and Manajit Sengupta. 2018. "A Fast All-Sky Radiation Model for Solar Applications with Narrowband Irradiances on Tilted Surfaces (FARMS-NIT): Part I. The Clear-Sky Model." *Solar Energy* 174: 691–702.

4. Gueymard, Christian A., Aron Habte, and Manajit Sengupta. 2018. "Reducing Uncertainties in Large-Scale Solar Resource Data: The Impact of Aerosols." *IEEE Journal of Photovoltaics* 8: 1732–37.

5. Rosseler, Olivier, Aron Habte, Manajit Sengupta, Christian A. Gueymard, and David M. Burns. 2020. "13 - A Simple Estimation of UV Irradiance Under Clear-Sky Conditions." In *Plastics Design Library, Service Life Prediction of Polymers and Coatings*. Edited by Christopher C. White, Mark E. Nichols, and James E. Pickett, 257–66. William Andrew Publishing. ISBN 9780128183670.

6. Xie, Yue, Manajit Sengupta, Yangang Liu, Hai Long, Qilong Min, Weijia Liu, and Aron Habte. 2020. "A Physics-Based DNI Model Assessing All-Sky Circumsolar Radiation." *iScience* 23 (3): 100893. https://doi.org/10.1016/j.isci.2020.100893.

7. Habte, Aron, Manajit Sengupta, Christian Gueymard, Anastasios Golnas, and Yue Xie. 2020. "Long-Term Spatial and Temporal Solar Resource Variability over America Using the NSRDB Version 3 (1998–2017). *Renewable and Sustainable Energy Reviews* 134. ISSN 1364-0321.

8. Peng, Jinqing, Jinyue Yan, Zhiqiang Zhai, Christos N. Markides, Eleanor S. Lee, Ursula Eicker, Xudong Zhao, Tilmann E. Kuhn, Manajit Sengupta, and Robert A. Taylor. 2020. "Solar Energy Integration in Buildings." *Applied Energy*: 114740.

9. Buster, Grant, Michael Rossol, Galen Maclaurin, Yue Xie, and Manajit Sengupta. 2021. "A Physical Downscaling Algorithm for the Generation of High-Resolution Spatiotemporal Solar Irradiance Data." *Solar Energy* 216: 508–17.

10. Liu, Wiejia, Yangang Liu, Xin Zhou, Yue Xie, Yongxiang Han, Shinjae Yoo, and Manajit Sengupta. 2021. "Use of Physics to Improve Solar Forecast: Physics-Informed Persistence Models for Simultaneously Forecasting GHI, DNI, and DHI." *Solar Energy* 215: 252–65.

11. Sengupta, Manajit, Aron Habte, Yu Xie, Anthony Lopez, and Christian A. Gueymard. 2019. "The National Solar Radiation Database (NSRDB) for CSP Applications: Preprint." Presented at the 2018 Solar Power and Chemical Energy Systems Conference (SolarPACES), Casablanca, Morocco, October 2–5, 2018. https://www.nrel.gov/docs/fy19osti/72310.pdf.

12. Sengupta, Manajit, Anthony Lopez, Aron Habte, and Yu Xie. 2019. "Improving the Accuracy of the National Solar Radiation Database (1998–2016): Preprint." Presented at the 2018 European PV Solar Energy Conference and Exhibition (EU PVSEC), Brussels, Belgium, September 24–28, 2018. https://www.nrel.gov/docs/fy19osti/72410.pdf.

13. Xie, Yu, Manajit Sengupta, Mike Dooraghi, and Aron Habte. 2019. *Reducing PV Performance Uncertainty by Accurately Quantifying the PV Resource*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-73377. https://www.nrel.gov/docs/fy19osti/73377.pdf.

14. Xie, Yu, and Manajit Sengupta. 2019. "Modeling of Spectral Irradiance in the POA Using GOES Satellite Data". NREL/PR-5D00-73075.

15. Sengupta, Manajit, Aron Habte, Yu Xie, Anthony Lopez, James Shelby, Christian Gueymard, Michael Foster, and Andrew Heidinger. 2019. "The National Solar Radiation Data Base: Version 3 (1998–2017)". NREL/PR-5D00-73076.

16. Sengupta, Manajit, Yu Xie, Aron Habte, and Christian Gueymard. 2018. "Advances in Solar Measurement and Modeling at NREL." NREL/PR-5D00-72896.

17. Sengupta, Manajit, Aron Habte, Yu Xie, and Galen Maclaurin. 2019. "Spectral and Broadband Data Sets from The National Solar Radiation Database (NSRDB)." Presented at the International Conference Energy Meteorology, Lyngby, Denmark, June 24–27, 2019.

18. Habte, Aron, Manajit Sengupta, Peter Gotseff, and Christian A. Gueymard. 2020. "Annual Solar Irradiance Anomaly Features Over the USA During 1998–2017 Using NSRDB V3: Preprint." Presented at the 47th IEEE Photovoltaic Specialists Conference (PVSC 47), June 15–August 21, 2020. NREL/CP-5D00-76858. https://www.nrel.gov/docs/fy20osti/76858.pdf.

19. Habte, Aron, and Manajit Sengupta. 2020. "Modeling of Ultraviolet Irradiance from Total Irradiance: A Simplified Approach." Presented at the 2020 PV Reliability Workshop, Lakewood, Colorado, February 25–27, 2020.

20. Sengupta, Manajit, Aron Habte, Yu Xie, and Grant Buster. 2020. "Improving the Accuracy of the National Solar Radiation Database (NSRDB) Using High-Resolution Data." Presented at the 2020 American Meteorological Society Meeting, Boston, Massachusetts, January 12–16, 2020. NREL/PR-5D00-75814. https://www.nrel.gov/docs/fy20osti/75814.pdf.

21. Kumler, Andrew, Yu Xie, Yingchen Zhang, Rui Yang, Xin Jin, Manajit Sengupta, and Yangang Liu. 2020. "Integration of Total-Sky Imager Data with a Physics-Based Smart Persistence Model for Intra-Hour Forecasting of Solar Radiation." Presented at the 2020 American Meteorological Society Meeting, Boston, Massachusetts, January 12–16, 2020. NREL/PR-5D00-75755. https://www.nrel.gov/docs/fy20osti/75755.pdf.

22. Lin, Chin-An, Yimin Zhang, Garvin Heath, Daven K. Henze, and Manajit Sengupta. 2020. "Improvement of Aerosol Optical Depth Data for Localized Solar Forecasting." Presented at the 2020 American Meteorological Society Meeting, Boston, Massachusetts, January 12–16, 2020. NREL/PR-5D00-75816. https://www.nrel.gov/docs/fy20osti/75816.pdf.

23. Xie, Yu, and Manajit Sengupta. 2020. "Evaluation of the PV Resource Dataset in Central and North America: Preprint." Presented at the 37th European Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC 2020), September 7–11, 2020. NREL/CP-5D00-77790. https://www.nrel.gov/docs/fy20osti/77790.pdf.

24. Xie, Yu, Manajit Sengupta, Yangang Liu, Hai Long, and Aron Habte. 2020. "Progress on the National Solar Radiation Data Base (NSRDB): A New DNI Computation: Preprint." Presented at the 47th IEEE Photovoltaics Specialists Conference (PVSC-47), June 15–August 21, 2020. https://www.nrel.gov/docs/fy20osti/75911.pdf.

25. Buster, Grant, Michael Rossol, Galen Maclaurin, and Manajit Sengupta. 2020. "A Physical Downscaling Algorithm for the Generation of High-Resolution Spatiotemporal Solar Irradiance Data: Preprint." Presented at the 2019 International Solar Energy Society (ISES) Solar World Congress, Santiago, Chile November 4–7, 2019. https://www.nrel.gov/docs/fy20osti/74386.pdf.

26. Buster, Grant, Mike Bannister, Aron Habte, Dylan Hettinger, Galen Maclaurin, Michael Rossol, Manajit Sengupta, and Yu Xie. 2021. "Physics-Guided Machine Learning for Prediction of Cloud Properties in Satellite-Derived Solar Data: Preprint. Presented at the 48th IEEE Photovoltaic Specialists Conference, June 20–25, 2020. https://www.nrel.gov/docs/fy21osti/79705.pdf.

27. Grue, Nick, Grant Buster, Andrew Kumler, Yu Xie, Manajit Sengupta, and Murali Baggu. 2021. *Quantifying Solar Energy Resource for Puerto Rico*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-75524. https://www.nrel.gov/docs/fy21osti/75524.pdf.

28. Habte, Aron, Manajit Sengupta, Grant Buster, Yu Xie, Michael Rossol, Paul Edwards, Mike Bannister, Haiku Sky, Evan Rosenlieb, Galen Maclaurin, and Billy McCall. 2021. *Estimating Surface Solar Irradiance Using Meteosat-8 Satellite for India and Surrounding Regions (2017–2019)*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-77883. https://www.nrel.gov/docs/fy21osti/77883.pdf.

29. Xie, Yue, Aron Habte, Manajit Sengupta, and Frank Vignola. 2021. *An Evaluation of the Spectral Irradiance Data from the NSRDB*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-80439. https://www.nrel.gov/docs/fy21osti/80439.pdf.

References

Habte, Aron, Manajit Sengupta, and Anthony Lopez. 2017. *Evaluation of the National Solar Radiation Database (NSRDB): 1998–2015*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-67722. https://www.nrel.gov/docs/fy17osti/67722.pdf.

Lopez, Anthony, Galen Maclaurin, Billy Roberts, and Evan Rosenlieb. 2017. *Capturing Inter-Annual Variability of PV Energy Production in South Asia*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-68955. http://www.nrel.gov/docs/fy17osti/68955.pdf.

Sengupta, Manajit, Aron Habte, Christian Gueymard, Stefan Wilbert, Dave Renne, and Thomas Stoffel. 2017. *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Second Edition*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-68886. ISBN 978-0-692-08922-4. https://www.nrel.gov/docs/fy18osti/68886.pdf.

Sengupta, Manajit, Yu Xie, Anthony Lopez, Aron Habte, Galen Maclaurin, and James Shelby. 2018. "The National Solar Radiation Data Base (NSRDB)." *Renewable and Sustainable Energy Reviews* 89: 51–60. https://doi.org/10.1016/j.rser.2018.03.003.

Xie, Yu, and Manajit Sengupta. 2018. "Assessing the Performance of the Fast All-sky Radiation Model for Solar Applications with Narrowband Irradiances on Tilted Surfaces (FARMS-NIT): Preprint." Presented at the 2018 World Conference on Photovoltaic Energy Conversion (WCPEC-7), Waikoloa, Hawaii, June 10–15, 2018. NREL/CP-5D00-71595. https://www.nrel.gov/docs/fy18osti/71595.pdf.

Xie, Yu, and Manajit Sengupta. 2018. "Assessment of Error Sources in the Modeling of POA Irradiance Under Clear-Sky Conditions: Preprint." Presented at the 2018 European PV Solar Energy Conference and Exhibition (EU PVSEC), Brussels, Belgium, September 24–28, 2018. NREL/CP-5D00-71104. https://www.nrel.gov/docs/fy18osti/71104.pdf.

Xie, Yu, Manajit Sengupta, and Jimy Dudhia. 2016. "Fast All-Sky Radiation Model for Solar applications (FARMS): Algorithm and Performance Evaluation." *Solar Energy* 135: 435–45. http://dx.doi.org/10.1016/j.solener.2016.06.003.

Habte, Aron, Manajit Sengupta, Christian Gueymard, Anastasios Golnas, Yu Xie, "Long-term spatial and temporal solar resource variability over America using the NSRDB version 3 (1998-2017)." *Renewable and Sustainable Energy Reviews* 134: 110285. https://doi.org/10.1016/j.rser.2020.110285