The National Solar Radiation Database (NSRDB) for CSP Applications

Preprint

Manajit Sengupta,[1] Aron Habte,[1] Yu Xie,[1] Anthony Lopez,[1] and Christian A. Gueymard[2]

[1] National Renewable Energy Laboratory
[2] Solar Consulting Services

The National Solar Radiation Database (NSRDB) for CSP Applications

Preprint

Manajit Sengupta,[1] Aron Habte,[1] Yu Xie,[1] Anthony Lopez,[1] and Christian A. Gueymard[2]

[1] National Renewable Energy Laboratory
[2] Solar Consulting Services

NOTICE

The National Solar Radiation Database (NSRDB) for CSP Applications

Manajit Sengupta[1, a)], Aron Habte[2], Yu Xie[2], Anthony Lopez[2], and Christian A. Gueymard[3]

[1]*PhD., Chief Scientist, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden CO 80401, USA.*
[2]*National Renewable Energy Laboratory, Golden, CO 80401.*
[3]*Solar Consulting Services, Colebrook, NH, USA*

a)Corresponding author: Manajit.Sengupta@nrel.gov

Abstract. This study examines the differences in direct normal irradiance (DNI) in two versions of the National Solar Radiation Database (NSRDB) produced by the National Renewable Energy Laboratory (NREL). NSRDB V3 of the NSRDB 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 less 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 method includes high-quality ground-measured data for seven National Oceanic and Atmospheric Administration Surface Radiation Budget Network (SURFRAD) stations for 1998–2015.

INTRODUCTION

Measured and modeled solar resource data are essential for quantifying the available energy for concentrating solar power (CSP) projects. The desired accuracy and spatiotemporal resolution of these data sources depend on the requirement of various project phases. Although financing and policy decisions mostly require annualized information, the required accuracy and resolution might increase significantly for day-to-day operations. Generally, the availability of measured data is limited because deploying ground-based measurement stations are expensive; requires adequate knowledge of the devices used to measure solar resources; and demands personnel for routine maintenance, data quality control, and analysis [1]. These requirements increase the time and financial investment required to maintain and operate ground-based measurement stations. On the other hand, satellite-derived models provide long-term solar resource data and high spatial coverage. Well-maintained ground-based measurements are highly accurate and often can be used to validate satellite-based solar resource data or to correct long time series of such data through the process referred to as "site adaptation" [2]. Therefore, the two sources of solar resource information (modeled and measured) are complementary and important to delivering high-quality and high-resolution data. The National Renewable Energy Laboratory (NREL) developed the gridded National Solar Radiation Database (NSRDB) using a physics-based two-step model called the Physical Solar Model (PSM) [3]. The model uses cloud information physically retrieved from the Geostationary Operational Environment Satellites (GOES) that cover North and South America. This 4-km, half-hourly database covers regions from Canada in the north to Brazil in the south (latitudes 60°N to 21°S) and is freely available for the 1998–2016 period from https://nsrdb.nrel.gov. Recently, NREL deployed NSRDB V3 of the NSRDB, which includes significant changes from NSRDB V2, both in various parts of the model and in the inputs it uses, most importantly those related to clouds and aerosols. This study introduces the NSRDB, summarizes the differences between the two model versions, particularly regarding direct normal irradiance (DNI), and discusses the resulting changes in the available solar resource.

METHODOLOGY

NSRDB Processing

Over the years, the NSRDB has been updated to meet the growing demand for solar resource data at a high spatiotemporal resolution for solar conversion systems. Such resource information is needed from the conceptual phase to routine solar power plant operation. Recently, NREL released the gridded NSRDB (1998–2016) based on NSRDB V3 of the PSM. The NSRDB data sets contain gridded solar irradiance— DNI, global horizontal irradiance, and diffuse horizontal irradiance—at a 4-km by 4-km spatial resolution and half-hourly temporal resolution covering 19 years. Details on the model and the data set are available at the NSRDB website (https://nsrdb.nrel.gov). Additional details about the development of the NSRDB are also available in [4], [5]. In the current version of the NSRDB (developed using PSM V3), NREL

implemented major changes in the meteorological input and processing. Hourly variables, such as aerosol optical depth (AOD) or precipitable water vapor (PWV), are now extracted from NASA's Modern Era Retrospective Analysis for Research and Applications NSRDB V2 (MERRA-2) (Fig. 1) [6]. These variables are particularly important to correctly evaluate DNI under clear-sky conditions. Additionally, downscaling methodologies to match the lower-resolution MERRA-2 data (0.5x0.625°) to the high-resolution 4-km NSRDB grid were redesigned to improve the accuracy in the ancillary variables such as PWV, AOD, temperature, pressure, and relative humidity. In parallel, the information from GOES-East satellite data requires being shifted in time to produce data at the top and middle of the hour. Whereas PSM V2 (used in NSRDB V2) shifted the solar radiation directly using a parametric model, PSM V3 (used in NSRDB V3) shifts the cloud products derived from GOES-East and uses those properties to compute solar radiation at the correct time. More importantly for DNI calculations, NSRDB V2 used monthly averaged AOD, whereas NSRDB V3 uses hourly AOD from MERRA-2.
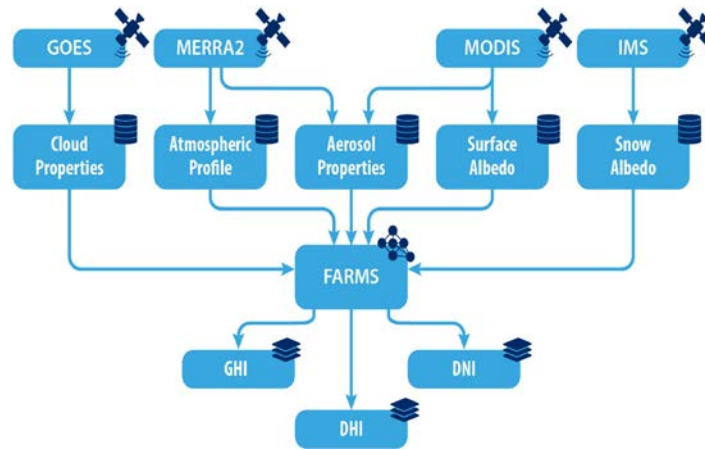

Figure 1. PSM flowchart for NSRDB.

NSRDB DNI Comparison and Uncertainty Analysis

As stated, both versions of the NSRDB contain some differences in ancillary input data and downscaling methodologies of the input parameters used in the NSRDB processing. Therefore, certain criteria needed to be implemented in the comparison and uncertainty estimation methodologies. Seven ground measurement stations from the National Oceanic and Atmospheric Administration's Surface Radiation Budget Network (SURFRAD) stations (Fig. 2) were used for this comparison representing various climatic regions [7], [8].

The conditions used in the comparison include:

a) Solar zenith angles must be less than 80°.

b) Irradiance must be strictly greater than zero.

c) Data records with missing values from the surface measurements are excluded from both the surface measurements and NSRDB data sets.

d) A sky clarity index is determined to detect clear-sky periods from the surface measurement and the NSRDB data using this definition:

$$Sky\ clarity\ index = \frac{surface\ all-sky\ GHI}{NSRDB\ clear-sky\ GHI} \tag{1}$$

Approximately clear-sky situations are assumed when this clarity index is greater than 0.8. This identifies half-hourly clear periods using only NSRDB predictions and corresponding surface measurements.

e) Cloud masking from the satellite-based retrievals were used to determine cloudy-sky conditions.

2

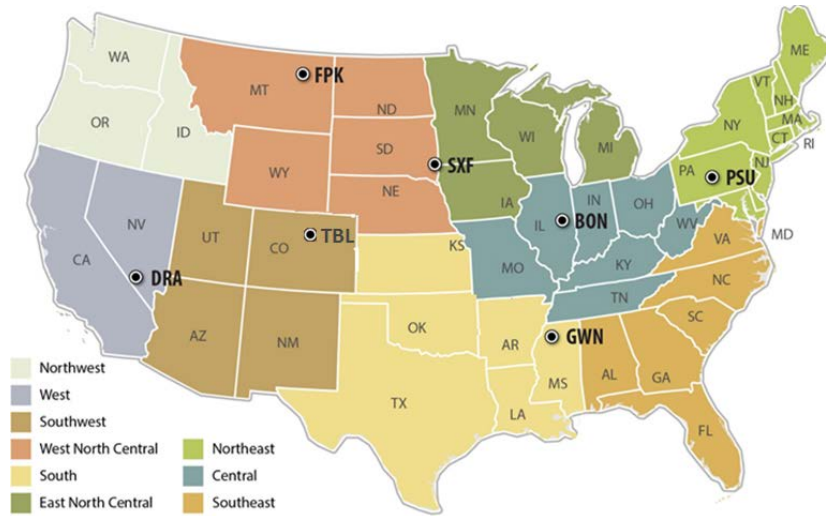These selection criteria were applied to both versions of the NSRDB (V2 and V3) and to the surface measurements.

Figure 2. SURFRAD Network locations overlaid on U.S. climatic regions. Image modified from [4]. More information about the stations: https://www.esrl.noaa.gov/gmd/grad/surfrad/sitepage.html

As stated in [7], *The Guide to the Expression of Uncertainty in Measurement* (usually referred to as the GUM method) is employed to estimate the uncertainty in the NSRDB data set. This uncertainty includes that in the surface measurement as well as the mean bias error (MBE) and root mean square error (RMSE) statistics of the modeled estimates. Overall, the following equation is used to calculate the NSRDB uncertainty for a 95% confidence interval (with coverage factor ~2, assuming a normal distribution):

$$U_{95} = 2 * \pm \sqrt{\left(\frac{U_{meas}}{2}\right)^2 + \left(\frac{MBE}{2}\right)^2 + \left(\frac{RMSE}{2}\right)^2} \qquad (2)$$

RESULTS

Understanding the impacts of the changes from PSM V2 to V3 on the modeled solar resource and quantifying the uncertainty of these changes constitute essential information for accurately designing utility-scale CSP systems. An example, shown in Fig. 3, demonstrates that DNI, which is the essential resource information for CSP applications, presents significant differences between the two NSRDB versions, especially over the southeastern United States. A preliminary assessment attributes these differences to both the input data set and the interpolation methodologies. Because the increased resource predicted by NSRDB PSM V3 makes a significant difference in the economics of power generation from CSP technologies, it is important that these differences be carefully examined and explained. To assess these differences, a traceable and standardized uncertainty characterization method is applied. It is meant to provide confidence in the data set for use by financiers, developers, and site operators of solar energy conversion systems and ultimately to reduce deployment costs.
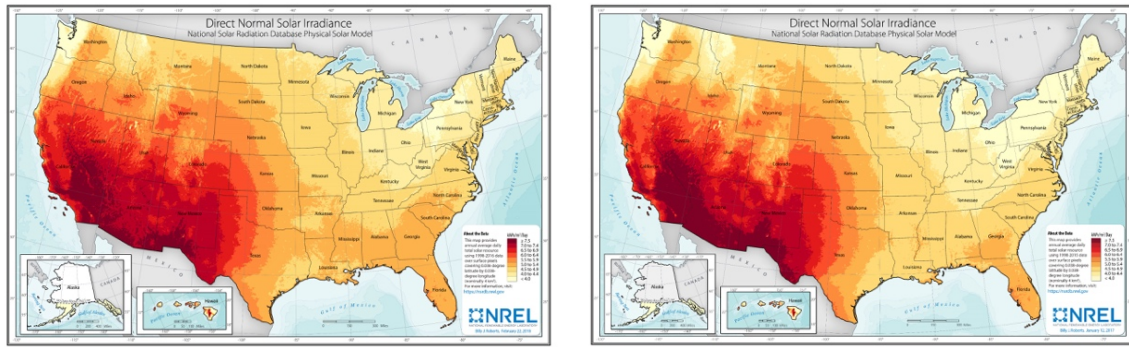
Figure 3. DNI differences between NSRDB V3 (left) and NSRDB V2 (right).

The analysis in Fig. 4 is focused on clear-sky conditions, which are critical for CSP applications. Clear-sky periods were selected using the selection criteria mentioned in Section 2. At almost all SURFRAD locations, the NSRDB V3 demonstrates significant improvement or reduction in uncertainty compared to V2, as shown in Fig. 4. 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.
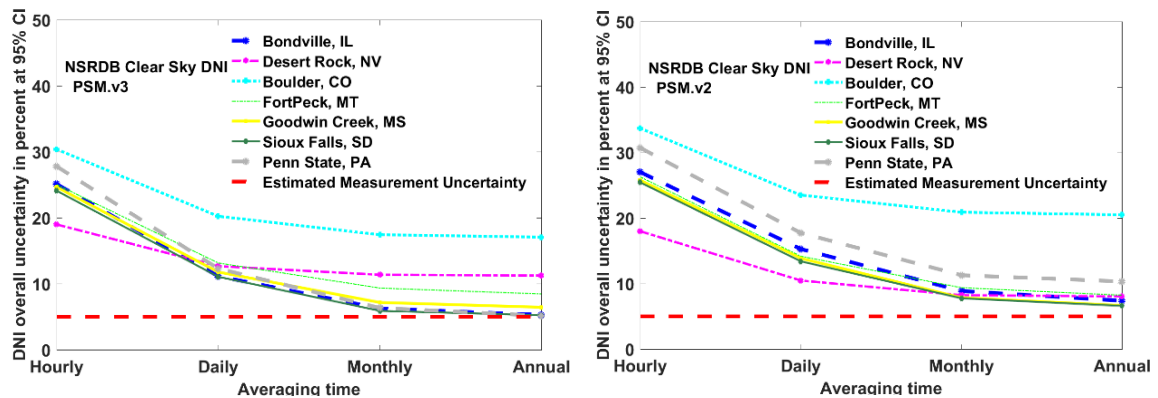

Figure 4. DNI uncertainty from NSRDB V3 (left) and V2 (right). Years: 1998–2015.

Conversely and unexpectedly, however, the two western locations of Desert Rock (DRA) and Fort Peck (FPK) do not show reduction 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 higher 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 DNI is significantly higher at all timescales.

The detailed statistics in Fig. 5 of the sources of uncertainty, such as MBE and RMSE, provide some insights in explaining the overall uncertainty estimation described in Fig. 4. The MBE (Fig. 5 top panel) shows a negative bias at all locations in the NSRDB V2 but only at the three western locations in V3. In contrast, Bondville and Goodwin Creek in the East show very small negative and positive bias, respectively, whereas Penn State and Sioux Falls demonstrate practically no bias in V3. A tentative explanation is that the changes implemented in V3 improved the clear-sky estimates.
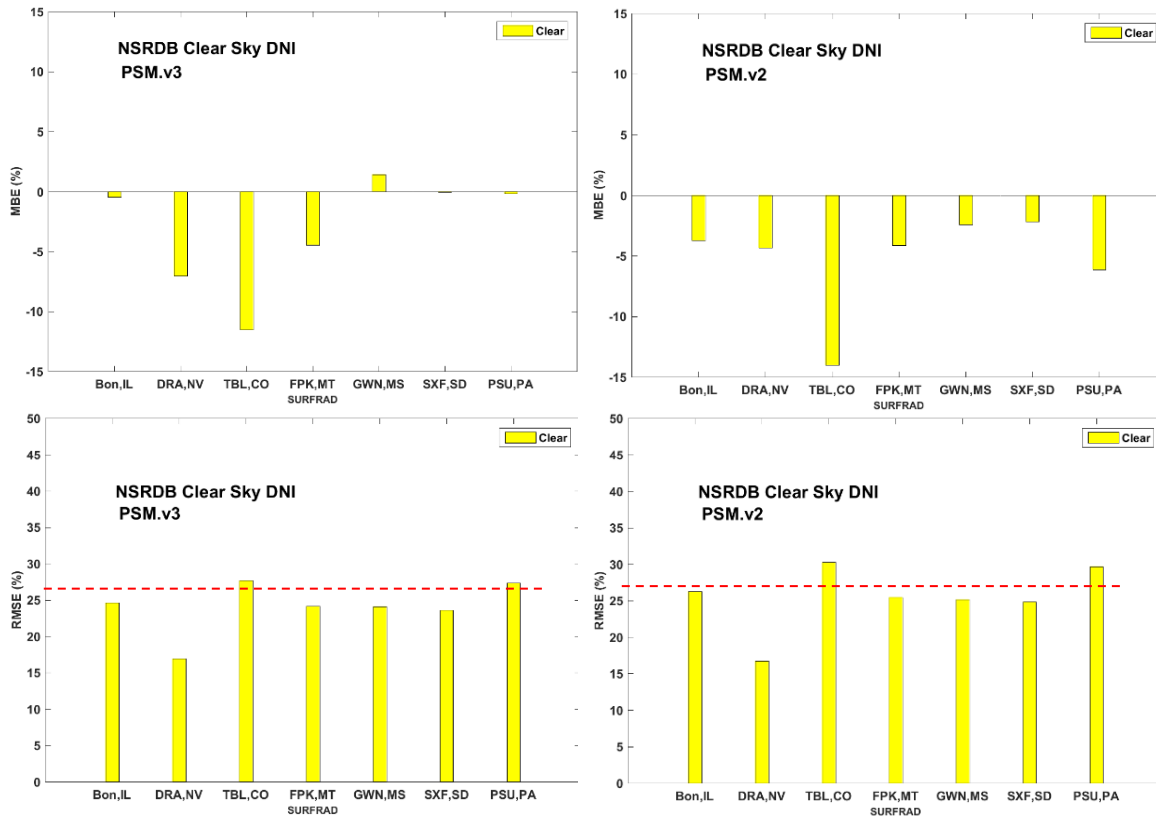
4

Figure 5. Top: Percentage MBE for NSRDB V3 (left) and V2 (right). Bottom: 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 in Fig. 4. 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 Fig. 4 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.

Changes in DNI Prediction Caused by the Aerosol Data Timescale

The MERRA-2 AOD is a fairly new data set, so there is a need to further evaluate the accuracy of this data set, especially in conditions where the AOD is low, which generally occurs in locations in the West such as Desert Rock and Fort Peck. Another important challenge in DNI estimation occurs in situations where the cloud-masking algorithm misses subpixel cloudiness. Note that the cloud masking uses the visible channel of the GOES satellites, which have a 1-km resolution. In the case of the NSRDB, only one pixel (the Northwest corner) is used for cloud detection.

It is well established (see, e.g., [9]) that the clear-sky DNI is a strong function of atmospheric turbidity. This is usually characterized by two main quantities: AOD at a specific wavelength (usually 550 nm) and the Ångström exponent, which represents the rate of change of AOD with wavelength. The NSRDB V2 used daily interpolations of monthly-mean aerosol information, as described in [10]. In contrast, V3 is now using hourly aerosol information from MERRA-2—with, however, some bias correction after topographic regridding [10]. This higher temporal resolution is expected to result in a much better frequency distribution of DNI at any site [11], which in turn can improve the evaluation of the potential of CSP plants [12]. This effect is caused by variations in the shape of AOD's lognormal distribution over different timescales [13].

As an illustration, Fig. 6 shows the frequency distributions of AOD and clear-sky DNI at Boulder for three different cases of AOD input data: monthly (as used in the NSRDB V2), hourly from MERRA-2 (raw), and hourly MERRA-2 after bias correction and topographic regridding (as used in the NSRDB V3). The displacement of the DNI frequencies is uniquely caused by variations in AOD input. Figure 7 is similar to Fig. 6, but for Desert Rock, where AOD is generally lower than at Boulder but where a stronger bias

5

correction was necessary in the hourly data. At Boulder and various other sites, the distribution of the predicted DNI tends to shift to the right (higher frequency of high DNI) as a consequence of using hourly rather than monthly AOD data, which might be significant in CSP applications.
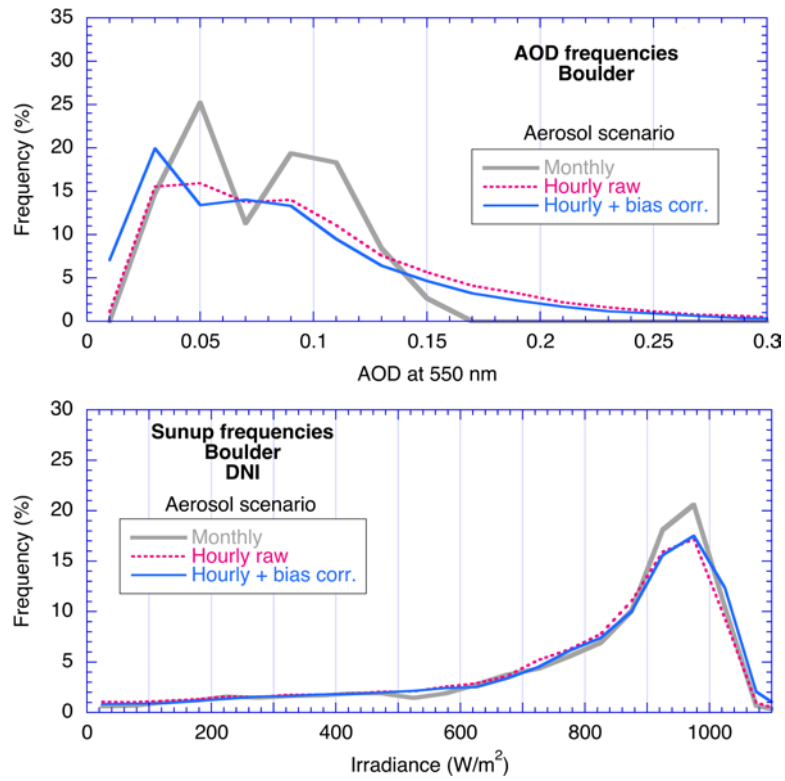

FIGURE 6. Top: Frequency distribution of AOD at Boulder using monthly data (NSRDB V2), raw hourly from MERRA-2, and hourly MERRA-2 after bias correction and topographic regridding (NSRDB V3)
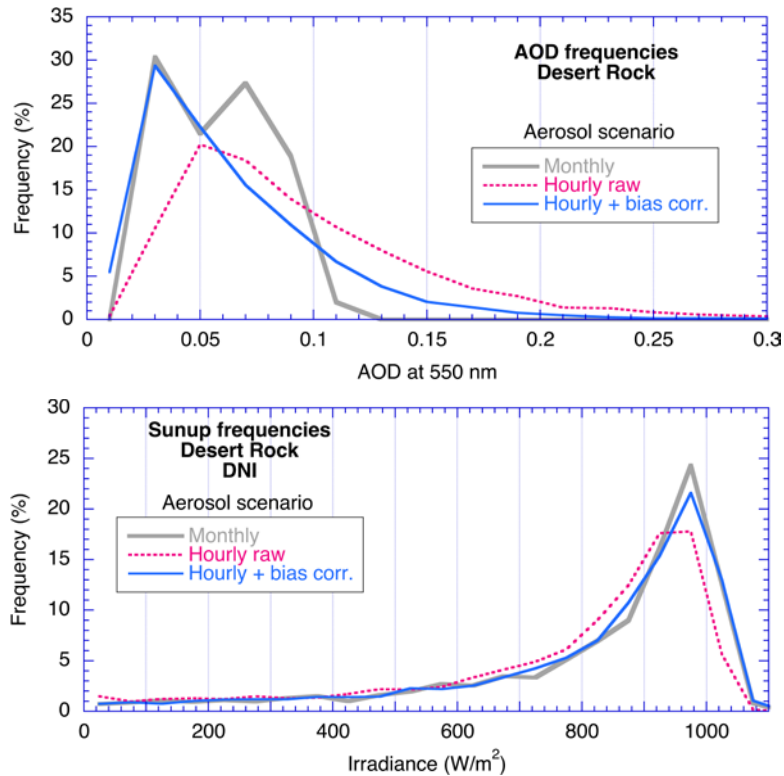
FIGURE 7. Top: Frequency distribution of AOD at Desert Rock using monthly data (NSRDB V2), raw hourly from MERRA-2, and hourly MERRA-2 after bias correction and topographic regridding (NSRDB V3)

CONCLUSION

NSRDB V3 demonstrated reduced uncertainties in DNI estimation compared to the previous version, NSRDB V2, especially for the GOES-EAST satellite. The improvement is assumed to result from a combination of factors, including (i) better downscaling methodologies, particularly in the interpolation and extrapolation used to align the multiple data sets to the same grid, and (ii) the use of hourly values of aerosol information (principally AOD), and surface albedo instead of interpolated monthly averages. Using hourly precipitable water vapor from MERRA-2 might have marginally contributed to the improvement too. The western regions do not show similar improvement, and this might be because of accuracies in the AOD from MERRA-2 in cleaner locations with significantly low AOD. Future improvements in the NSRDB include the use of next-generation GOES-16 data at a 2-km resolution with the cloud-masking algorithm making use of all visible pixels that are available at a 1-km resolution. It is also expected that cloud masking will improve in GOES-16 because the additional visible channels will enable cloud snow discrimination, which currently does not exist for the older four-channel GOES satellites.

Further validation efforts will look more closely into the delicate issue of precisely defining "clear-sky" periods because part of the uncertainty reported in clear-sky DNI might be caused by unsuspected cloud contamination in the data.

ACKNOWLEDGEMENTS

REFERENCES

[1] M. Sengupta, A. Habte, C. Gueymard, S. Wilbert, and D. Renné. (2017). Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Second Edition.

[2] J. Polo, S. Wilbert, J.A. Ruiz-Arias, R. Meyer, C. Gueymard, M. Súri, L. Martín, T. Mieslinger, P. Blanc, I. Grant, J. Boland, P. Ineichen, J. Remund, R. Escobar, A. Troccoli, M. Sengupta, K.P. Nielsen, D. Renne, N. Geuder, and T. Cebecauer, Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Solar Energy 132, 25–37 (2016).

[3] Xie, Y., M. Sengupta, and J. Dudhia. 2016. "A Fast All-Sky Radiation Model for Solar Applications (FARMS): Algorithm and Performance Evaluation." Solar Energy 135: 435–445.

[4] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin and J. Shelby, "The National Solar Radiation Data Base," Renew. Sustain. Energy Rev. (2018)

[5] Y. Xie, M. Sengupta and J. Dudhia, "A Fast All-Sky Radiation Model for Solar Applications (FARMS): Algorithm and performance evaluation," Solar Energy 135: 435–445 (2016).

[6] Gelaro R, McCarty W, Suárez M, Todling R, Molod A, Takacs L, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Climate. 2017;30:5419-54.

[7] A. Habte, M. Sengupta and A. Lopez, "Evaluation of the National Solar Radiation Database (NSRDB): 1998-2015," NREL/TP-5D00-67722, 2017.

[8] M. Sengupta, A. Weekley, A. Habte, A. Lopez, C. Molling and A. Heidinger, "Validation of the National Solar Radiation Database (NSRDB) (2005–2012): Preprint," Paper presented at the European PV Solar Energy Conference and Exhibition, Hamburg, Germany, September 14–18, 2015.

[9] C.A. Gueymard, Temporal variability in direct and global irradiance at various time scales as affected by aerosols, Solar Energy 86, 3544–3553 (2012).

[10] C.A. Gueymard, A. Habte and M. Sengupta, Reducing Uncertainties in large-scale solar resource data: The impact of aerosols, J. Photovoltaics (in press).

[11] J.A. Ruiz-Arias, C.A. Gueymard, F.J. Santos-Alamillos, S. Quesada-Ruiz and D. Pozo-Vázquez, Bias induced by the AOD representation time scale in long-term solar radiation calculations. Part 2: Impact on long-term solar irradiance predictions, Solar Energy 135, 625–632 (2016).

[12] J.A. Ruiz-Arias, C.A. Gueymard, F.J. Santos-Alamillos and D. Pozo-Vázquez, Worldwide impact of aerosol's time scale on the predicted long-term concentrating solar power potential, Nature Scientific Reports 6:30546, DOI: 10.1038/srep30546.

[13] J.A. Ruiz-Arias, C.A. Gueymard, F.J. Santos-Alamillos, S. Quesada-Ruiz and D. Pozo-Vázquez, Bias induced by the AOD representation time scale in long-term solar radiation calculations. Part 1: Sensitivity of the AOD distribution to the representation time scale, Solar Energy 137, 608–620 (2016).