IMPROVING MODELED SOLAR IRRADIANCE HISTORICAL TIME SERIES: WHAT IS THE APPROPRIATE MONTHLY STATISTIC FOR AEROSOL OPTICAL DEPTH?

Christian A. Gueymard
Solar Consulting Services
P.O. Box 392
Colebrook, NH 03576
Chris@SolarConsultingServices.com

Manajit Sengupta
National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401
Manajit.Sengupta@nrel.gov

ABSTRACT

Under clear-sky conditions, direct normal irradiance (DNI) is particularly sensitive to the aerosol optical depth (AOD). It is found that the daily AOD is nearly always log-normally distributed, resulting in significant differences between the monthly mean, median and mode values. Monthly AOD data are still used extensively in radiation modeling. This study thus investigates whether the monthly mean AOD should still be selected as the proper statistic to provide the necessary aerosol input to solar radiation models that are used for hourly irradiance predictions.

To solve this issue, daily AOD data from nine sunphotometric sites in various climates are analyzed, and their monthly frequency distributions are obtained to derive the usual statistics. Simulations of cloudless-sky DNI are obtained with the REST2 radiative model, alternatively using individual daily data (providing reference results) and monthly-aggregated statistics of AOD as inputs.

Results show that the monthly mean AOD yields too low predicted DNI, whereas the monthly mode results in too high DNI values. The median value is found close to the optimum in general. Some scatter exists, however, so that this finding may not be valid at all sites or for all months. The ratio between the monthly median and mean AOD also varies largely, with a typical average of ≈0.8.

The practical outcome of this study is that the existing DNI predictions made by NREL in the past may be too low over the USA (where monthly mean AOD values were used), and may be too high over India (where the monthly mode was used).

1. INTRODUCTION

Modeled time series of solar irradiance are an essential aspect of any solar resource assessment. Currently, such time series provide hourly or sub-hourly irradiance data over periods as long as 10–20 years, and are usually based on satellite imagery for the necessary information on clouds.

Under cloudless-sky conditions, direct normal irradiance (DNI) is particularly sensitive to aerosols [1, 2]. The most important atmospheric variable that conditions DNI is then the aerosol optical depth (AOD). This means that accurate predictions of DNI can only be obtained if good AOD data are available. Over the last two decades, high-quality AOD data have become available from ground stations equipped with sunphotometers, but only a few hundred sites provide such data over the world. Since solar irradiance time series may be necessary for solar projects at almost any possible site in the world, satellite-based retrievals of AOD are the most frequent source of AOD data for the prediction of historical solar irradiance time series. On a daily basis, the geographical coverage of such data is very limited: most satellites cannot observe all possible areas during any particular day, and cannot retrieve AOD over cloudy scenes, and often over highly reflective surfaces. Most irradiance time series are thus modeled using satellite-derived AOD datasets that are spatially aggregated, supplemented wherever/whenever needed, and averaged on a monthly basis.

Gridded daily AOD data recently became available [3], and their use appears to improve the daily accuracy of modeled DNI and of its frequency distribution over various areas [4, 5]. However, such daily AOD datasets are largely modeled with a low spatial resolution, and their absolute accuracy is not known over all continents, due to lack of validation studies. Preliminary tests show that such daily AOD data may not be appropriate for the southwest USA, for instance.

Recently, a few informal reports have surfaced among the solar resource community that the predicted DNI could be systematically too low over many areas, particularly under clear skies. This is particularly the case with NREL's National Solar radiation Data Base (NSRDB) over the USA

(Pers. comm. with Richard Perez, 2009). Since the NSRDB has been developed with monthly mean (climatological) AOD statistics [6] rather than the desirable daily AOD data, one possible explanation is that the monthly *mean* may not be the correct statistics to use for the AOD inputs to radiative models when *hourly* irradiance results are to be obtained. In parallel, a recent investigation has demonstrated that large DNI biases could exist in satellite-derived data series over hazy areas, such as Northern Africa and the Middle East, and that a bias correction of the underlying AOD data could significantly improve the quality of these DNI datasets [7].

This study delves into the issue of selecting the most appropriate monthly AOD statistic. The methodology is based on a combination of coincident AOD and DNI measured data from a few sites with largely different climatic conditions, used along with a high-performance radiative model for accurate DNI predictions.

2. AEROSOL DATA AND STATISTICS

High-quality AOD data from NASA's Aeronet worldwide network are used here (http://aeronet.gsfc.nasa.gov/). Each Aeronet site is equipped with a multiwavelength sunphotometer that monitors the irradiance over a few specific narrow spectral bands, usually within the interval 340–1020 nm. This interval corresponds to the peak of the solar spectrum, and also to the largest aerosol extinction effects (mostly scattering). After reduction of the raw sunphotometric data, AOD is separately derived for each channel of the instrument. The unattended instrument tracks the sun, and has a 15-minute sampling rate. A cloud-screening algorithm is used *a posteriori* to filter out all periods when the sun was most likely obscured by clouds. This normally ensures that the "Level-2" processed AOD data does not suffer from cloud interference, and thus really corresponds to pure aerosol extinction. Assuming there are at least two working channels at any given valid time t, it is then possible [1, 2] to use Ångström's Law to evaluate three important optical characteristics of the total aerosol column that existed over the site at time t: the AOD at 550 nm, τ_{a550}, the AOD at 1000 nm, β, and the Ångström exponent, α. These three quantities are related by:

$$\tau_{a550} = \beta\, 0.55^{\alpha}. \qquad (1)$$

The AOD at 550 nm is now of particular interest in global studies because it the quantity provided by recent sources of AOD data. Historically, however, β has been in continuous use since its introduction by Ångström in 1929. In what follows, most results are reported in terms of β.

A specific channel (around 940 nm) of sunphtometers is exclusively devoted to the retrieval of precipitable water (PW), w. After AOD, PW is the next most important source of extinction of radiation under cloudless skies. Other sources of extinction also exist, but their effect on irradiance is small [8], so that approximate or interpolated data can be used for them without significant degradation of performance.

It has been shown recently [2, 9] that it is possible to obtain very accurate values of cloudless irradiances when sunphotometric data of AOD and PW are known. This modeling capability is used here at its fullest, as detailed in Section 4.

A compilation of the observed AOD data available from Aeronet or similar sources shows that the time distribution of AOD is generally not Gaussian. For instance, Fig. 1 shows the frequency distribution of the daily β from the Aeronet data of Sede Boker, Israel. An obvious skewness is visible, which is best described by a log-normal distribution. Indeed, tests show that the frequency distribution of $\ln\beta$ (or of $\ln\tau_{a550}$ for that matter) is Gaussian.

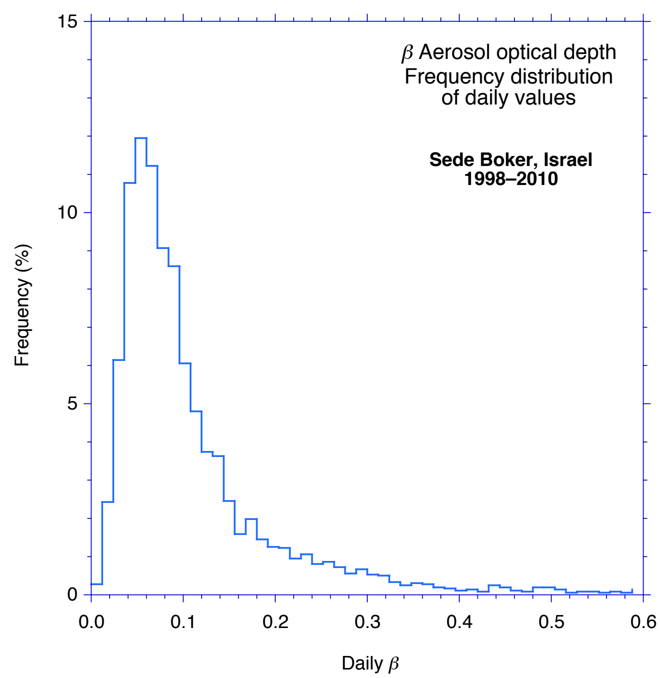

Fig. 1: Frequency distribution of β at Sede Boker.

A direct consequence of the log-normality of the AOD distribution is that, contrary to a normal distribution, the mean value differs from other statistics, such as the median or mode. For the example of Fig. 1, based on 3582 daily β data points between 1998 and 2010, the mean, median, and mode values are 0.116, 0.082 and 0.054, respectively. Similar differences are obtained when considering frequency distributions for each month at a specific site, rather than for the whole year, as in Fig. 1. However, monthly distributions are not as smoothly behaved as the one in Fig.1, because of the much smaller number of data points. This is demonstrated in Fig. 2 for two widely different aerosol conditions.

If using the mean monthly AOD translates into too low DNI, as seems to be the case, which other statistic should rather be used? The possible choices are the median, the mode, or some intermediate percentile value.

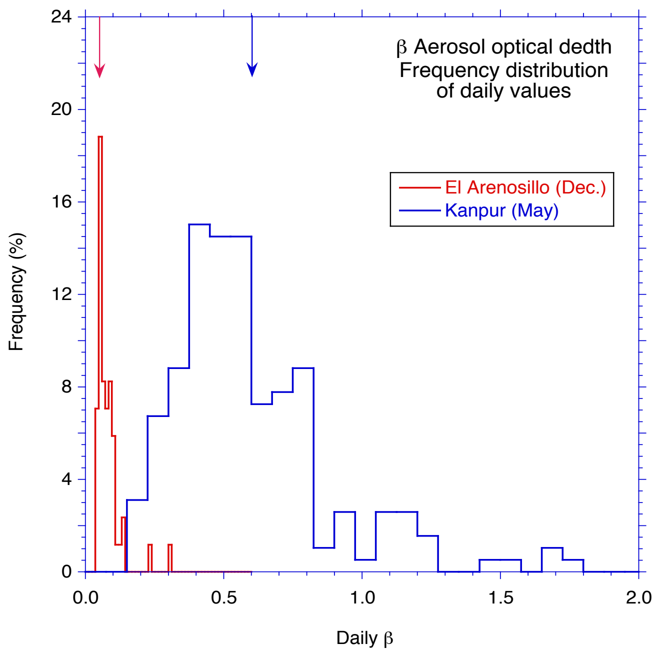

Fig. 2: Frequency distribution of measured daily β during low-AOD and high-AOD months. For each site, the monthly mean β is indicated by an arrow.

Before this important question could be resolved, the latest solar resource maps and datasets for India developed by NREL (http://www.nrel.gov/international/ra_india.html) were prepared in 2010 under the assumption that the mode should be used rather the mean to generate the necessary monthly gridded AOD datasets [10]. The validity of this assumption is discussed here to advance the issue, before the planned update of these maps is undertaken. For solar resource applications in hazy areas like India, the selection of the proper AOD statistics is critical, since a large absolute difference may exist between the mean, median and mode.

3. IRRADIANCE SENSITIVITY TO AOD

It is now established that the solar irradiance components are more sensitive to an absolute difference in AOD than to a relative difference [2]. This indicates that the choice of the proper monthly AOD statistic is critical, and might have large consequences on the modeling of irradiance over hazy areas, since the absolute difference between each statistic is then much larger than under very clean atmospheric conditions.

Reference [2] offers a sensitivity analysis that describes how errors in instantaneous AOD affect DNI, based on a parameterization of the functional dependence between DNI and AOD. In that analysis, an absolute error in β, $\Delta\beta$, translates into a relative error in DNI, $\Delta E_{bn}/E_{bn}$, according to:

$$\Delta E_{bn}/E_{bn} = \Delta T_a/T_a \qquad (2)$$

where T_a is the broadband aerosol transmittance, which can be conveniently expressed as:

$$T_a = \exp(-m\beta\lambda_e^{-\alpha}) = \exp(-\gamma\beta) \qquad (3)$$

where:

$$\gamma = m\lambda_e^{-\alpha}. \qquad (4)$$

In Eqs. (3) and (4), m is the air mass and λ_e is the effective wavelength, itself an intricate function of m, w, and α.

From Eqs. (2) and (3), it can be construed that the effect of aerosol extinction on DNI is a direct function of the product $\gamma\beta$, rather than a direct function of β alone. Moreover, the frequency distribution of E_{bn} under cloudless conditions is directly affected by how the exponential function and γ both affect the underlying frequency distribution of β. To study the sensitivity of DNI to AOD on a monthly—rather than instantaneous—basis, an analytic derivation similar to that developed in [2] is desirable. Currently, this appears a daunting task, because γ, the multiplicative term of β in Eq. (3), is a strong function of m and α, which, in turn, are both functions of time. Therefore, to perform a sensitivity analysis on a monthly time scale, the statistical relationships between β, α, and m would have to be determined and expressed in simple terms, which currently does not seem possible. To avoid this critical problem, a pragmatic approach, similar to that described in [2] is followed here.

4. MONTHLY SENSITIVITY ANALYSIS

For this study, nine Aeronet stations with long periods of record (9 years or more) have been selected as representative of widely different climatic and geographic conditions. Detailed information on these stations appears in Table 1. The mean annual measured β is provided for further reference. Figures 3 and 4 compare the frequency distribution of the daily τ_{a550} at two of these sites during months of intense dust activity, with therefore very high AOD. Under such circumstances, the difference between the monthly mean, median and mode is significant, as mentioned earlier. Note, however, that this difference is not the same for the two situations illustrated in the figures. At Kanpur in June, the mean, median and mode are 0.7721, 0.6691 and 0.6380, respectively. At Banizoumbou in March, these numbers become 0.8877, 0.6960 and 0.4880, respectively. This spread of results under similar *average* aerosol loads implies that some aspect of the local aerosol climatology might be different at the two sites, and might be a determinant factor in general. Since these high-AOD conditions constitute the worst-case scenario (as far as statistical effects on irradiance are concerned), the emphasis is deliberately put here on moderately hazy to very hazy environments.

TABLE 1. SUNPHOTOMETER STATIONS PROVIDING AOD DATA FOR THIS STUDY

Station	Country	Lat.	Long.	Elev. (m)	Annual β
Bonanza Creek	USA	64.74	-148.32	150	0.060
Moscow	Russia	55.70	37.51	192	0.090
Belsk	Poland	51.84	20.79	190	0.087
El Arenosillo	Spain	37.11	6.73	0	0.083
FORTH-Crete	Greece	35.33	25.28	20	0.103
Sede Boker	Israel	34.78	30.86	480	0.116
Kanpur	India	26.51	80.23	123	0.353
Solar Village	Saudi Arabia	24.91	46.40	764	0.268
Banizoumbou	Niger	13.54	2.67	250	0.447

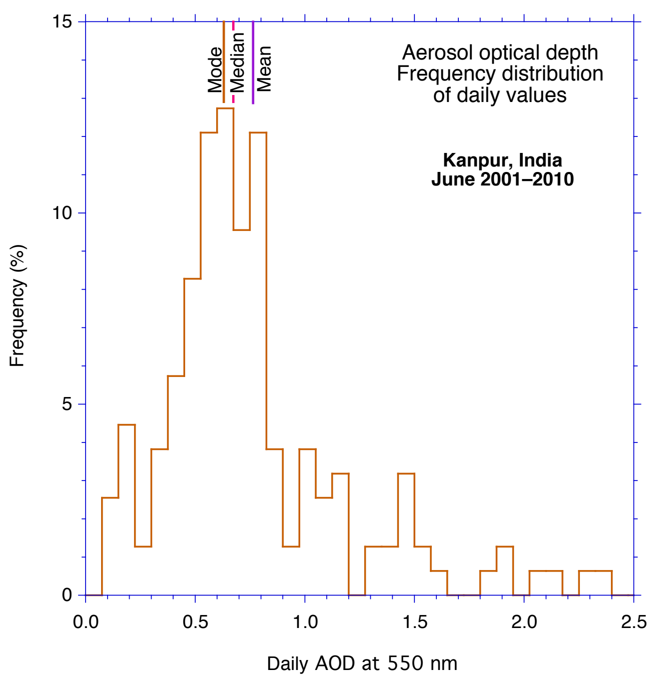

Fig. 3: Frequency distribution of daily AOD at Kanpur, and corresponding monthly statistics: mean, median and mode.

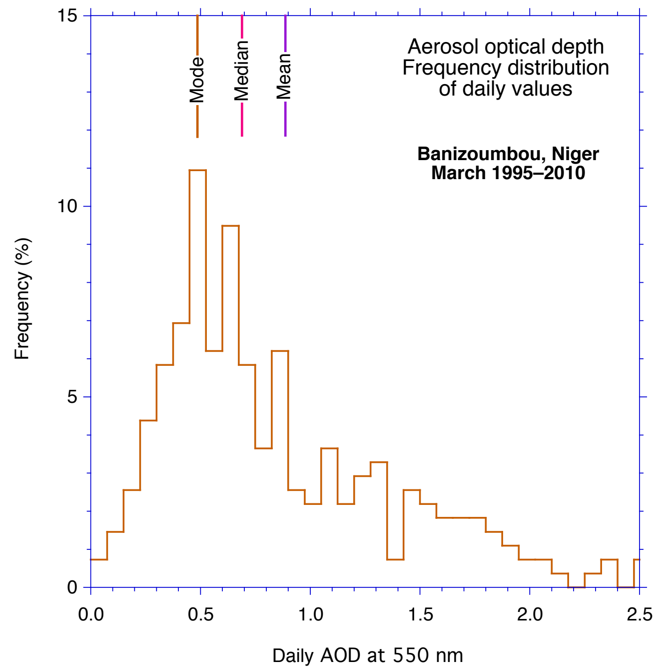

Fig. 4: Frequency distribution of daily AOD at Banizoumbou, and corresponding monthly statistics: mean, median and mode.

Daily AOD data from the nine sites of Table 1 is analyzed on a monthly basis, for each month with sufficient data points. (Due to its high latitude, there are no AOD observations at Bonanza Creek during winter, when the sun remains too low on the horizon.) The monthly mean, median and mode of the measured daily β values can thus be obtained.

For each individual day for which measured values of AOD (reduced here to β and α) and PW are available, the sub-hourly and daily irradiances are calculated with the high-performance REST2 radiative model [1], using also reasonable values of ozone, nitrogen dioxide, and ground albedo, derived from existing climatologies. For this analysis, is assumed here that the mean measured β, α and w values for each day are constant throughout that day. This cannot be perfectly realistic because there are always short-term variations in these quantities. However, the use of sub-daily values is not practical since such records are never complete: Aeronet instruments cannot operate beyond a maximum air mass of 7, and there might be some cloud passages with no data.

Comparisons between the predicted hourly or daily irradiances and their measured counterpart at the BSRN site of Sede Boker (where nearly all days are cloudless in summer) show good agreement, thus confirming the method's potential (Fig. 5). Slight differences in peak DNI are caused by the simplification described above. A significant daily variability in DNI is obvious there, and is essentially the direct result of that in AOD [2].

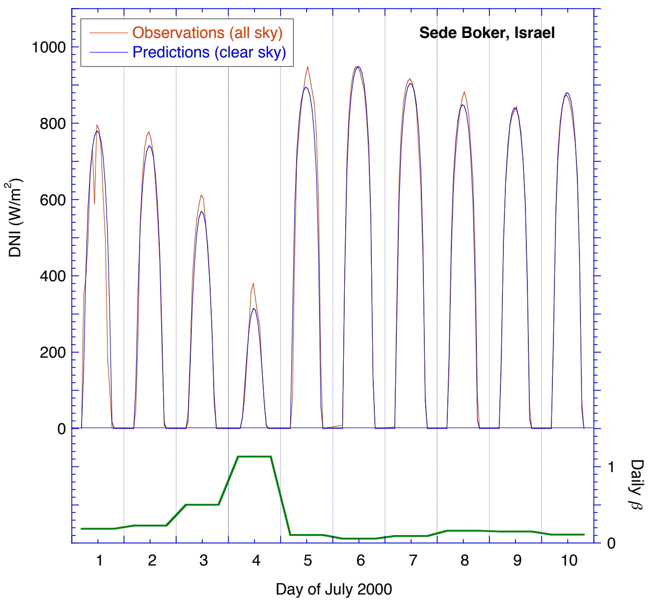

Fig. 5: Modeled vs. measured DNI at Sede Boker. The measured mean daily value of β appears in the bottom plot.

Based on these daily calculations, twelve long-term daily cloudless-sky DNI averages can be calculated for each month and each station. These are used here to represent the reference mean daily cloudless DNI. Such reference results are obtained for 102 station-months (12 months for each of the 9 stations, except Bonanza Creek, where only 6 months could be used).

In a second step, the process is repeated three more times, but this time alternatively replacing the actual daily value of β by monthly-aggregated statistics of AOD—long-term monthly mean, median or mode, respectively—while keeping all other daily inputs the same as in the initial run. These three alternate calculations are "static" since they do not take the daily variability of β into account, in contrast with the "dynamic" calculation of the reference DNI results described above.

5. RESULTS

The monthly-average DNI values derived from the monthly mean β data are found systematically lower than those obtained with the reference daily "dynamic" data (Fig. 6). This suggests that the monthly mean β is not the appropriate statistic to use for the problem at hand. Conversely, the monthly DNI values obtained when using the monthly modal values of β are systematically too high (Fig. 6). A better agreement (even though not perfect) is found when using the monthly median value of β (Fig. 6). To find the "effective" monthly values of β that would yield a nearly perfect agreement between the static and dynamic methods, linear interpolation in the overall DNI results is used. The relationship between these effective β values and their monthly mean and median counterparts is shown in Fig. 7. These results show some scatter, which means that, depending on station and month, the ideal statistic to use may vary.

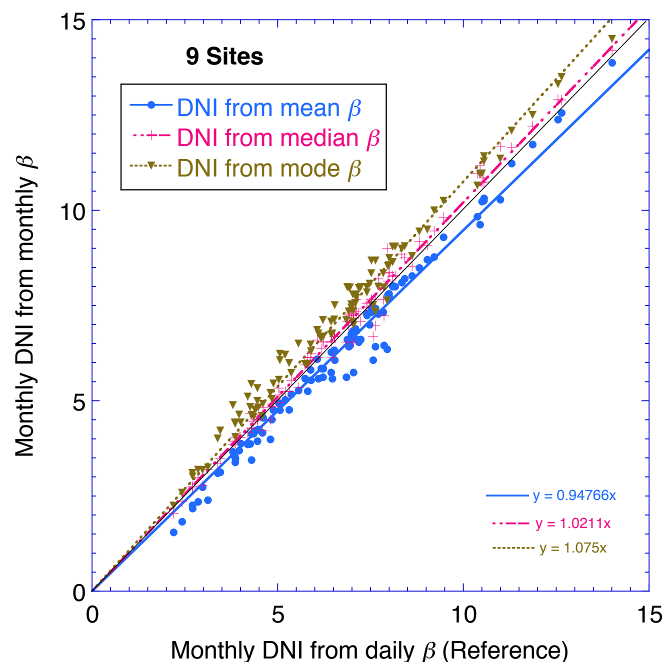

Fig. 6: Monthly DNI (kWh/m^2) obtained with various monthly statistics of β vs. reference values based on daily values.

It is clear from Fig. 7 that, on average, the effective β is close to its monthly median value. A more detailed investigation, which would have to consider more stations, a larger spread of climatic and aerosol conditions, and a range of percentiles in addition to the three usual statistics used here, is desirable. It is likely that the best results will be obtained close to the 50th percentile, but there might be climatic or seasonal overtones, as alluded to above. These could explain the scatter in Figs. 6 and 7.

6. MEDIAN vs. MEAN AOD

The three monthly statistics that have been discussed thus far are directly calculated from measured daily AOD values. There are many applications for which such daily AOD values are not accessible, and only the monthly mean AOD is available. This is for example the case of the irradiance datasets that have been developed at NREL until now: their underlying AOD values were derived from monthly gridded satellite data, and were thus monthly means. In such a case, the only way to evaluate the effective monthly β (e.g., the median) is from its monthly mean value. It is therefore important to understand the general relationship between the median and the mean.

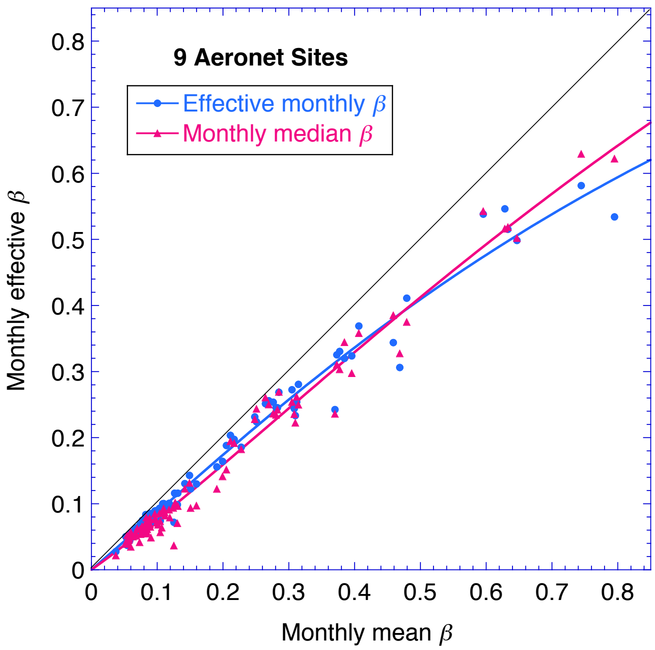

Fig. 7: Effective monthly β derived from the reference monthly DNI, and median β, as a function of the monthly mean β. Best-fit quadratics are also shown.

Using daily AOD data from 124 Aeronet stations over the world, the mean and median β for each calendar month has been obtained. Only months with at least 60 days of data over the whole record period have been considered, to avoid undersampling issues. The ratio of the monthly median and mean β for the 1345 station-month data points appears in Fig. 8 as a function of the mean β. A value ≈ 1 for this ratio suggests a normal distribution. It is found that the median/mean ratio varies largely, between 0.25 and 1.05. The scatter is large, but by far all station-months exhibit a skewed, log-normal behavior, with a typical ratio close to 0.8. Different color codes are used in Fig. 8 to separately show ratio values from sites in North Africa or Middle East (regions dominated by dust and smoke aerosols) and from sites in Asia (where dust and pollution are the dominant types of aerosol). No clear pattern emerges from this geographical sorting. Further research is thus necessary to investigate the causes of the observed scatter, and ultimately obtain an accurate determination of the median β when only the monthly mean β is known.

7. CONCLUSION

This investigation addressed a very specific issue, with which solar radiation modelers are confronted when evaluating hourly solar irradiance based on monthly-average aerosol data, which is currently a very common situation. Of particular importance in the present context, this is the case of NREL's NSRDB over the USA (for which the monthly *mean* AOD was used), or of its resource assessment data for India, released in 2010 (for which AOD's monthly *mode* was rather used).

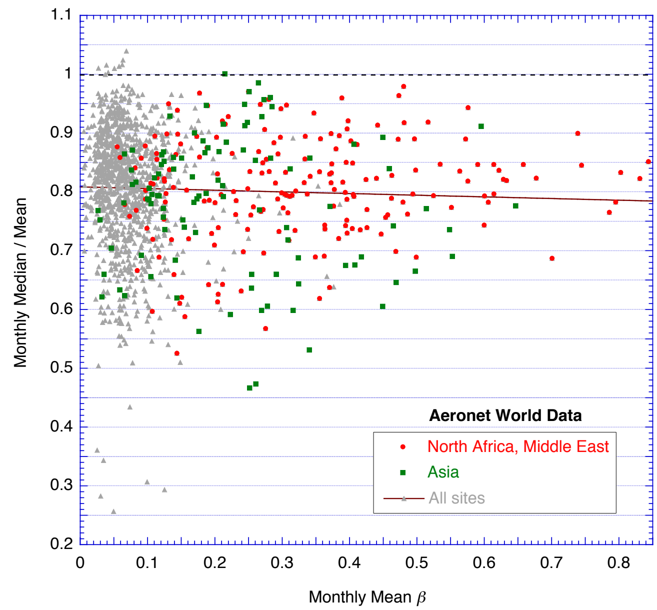

Fig. 8: Ratio of monthly median and mean β as a function of the mean β. The grey triangles indicate sites outside of North Africa, Asia or the Middle East. The brown line indicates the best linear fit for all 124 sites.

Using high-quality aerosol (spectral AOD) data at various sites in the world, as well as ancillary atmospheric data, a method has been devised to compare the predictions of cloudless direct irradiance (DNI) when alternatively using the monthly mean, median or mode AOD, to reference results obtained from individual daily AOD data.

Based on this analysis, it is found that the use of the monthly mean AOD tends to produce underestimated DNI values, whereas the use of the monthly mode AOD generally produces overestimated DNI values. This has practical consequences on the accuracy of the above-mentioned solar resource modeled data for the USA and India: a likely low bias in DNI can be expected for the former, and a likely high bias for the latter.

It is also found that, in general, the best simple monthly statistic to use is the median. However, the results provided here should be considered preliminary because they have been obtained at only 9 sites, and, even with this limited dataset, a lot of scatter is apparent. This suggests that some local or climatic dependence is likely, and possibly influenced by the specific aerosol climatology at each site and for each season. This factor (whose nature remains to be investigated) could also be the cause of the large scatter observed when comparing the monthly mean and median AOD over the world. At many sites in hazy areas, such as North Africa, the Middle East and Asia, the median AOD is typically ≈20% (or more) lower than the mean, so that using the monthly mean AOD may results in significantly underestimated daily DNI over these regions.

Further research is recommended to evaluate whether the prediction of hourly DNI from monthly AOD data can be improved by using appropriate, possibly location-dependent monthly statistics, or whether this approach should be abandoned in favor of using daily AOD data.

8. ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Laboratory. The AERONET and BSRN staff and participants are thanked for their successful effort in establishing and maintaining the various sites whose data were advantageously used in this investigation.

9. REFERENCES

1. Gueymard C.A., REST2: High performance solar radiation model for cloudless-sky irradiance, illuminance and photosynthetically active radiation—Validation with a benchmark dataset. *Solar Energy*. **82**: 272-285, 2008.

2. Gueymard C.A., Temporal variability in direct and global irradiance at various time scales as affected by aerosols. *Solar Energy*, in press, 2012.

3. Cebecauer T. and Suri M., Accuracy improvements of satellite-derived solar resource based on GEMS reanalysis aerosols. Proc. *SolarPACES Conf.,* Perpignan, France, 2010.

4. Cebecauer T., Perez P., and Suri M., Comparing performance of SolarGIS and SUNY satellite models using monthly and daily aerosol data. Proc. *Solar World Congress*, Kassel, Germany, ISES, 2011.

5. Cebecauer T., Suri M., and Gueymard C.A., Uncertainty sources in satellite-derived direct normal irradiance: how can prediction accuracy be improved globally? Proc. *SolarPACES 2011 Conf.,* Granada, Spain, 2011.

6. Wilcox S., Anderberg M., George R., Marion W., Myers D., Renné D., Lott N., Whitehurst T., Beckman W., Gueymard C., Perez R., Stackhouse P., and Vignola F., Completing poduction of the updated National Solar radiation Database for the United States. Proc. *Solar 2007 Conf.,* Cleveland, OH, American Solar Energy Soc., 2007.

7. Gueymard C.A., Uncertainties in Modeled Direct Irradiance Around the Sahara as Affected by Aerosols: Are Current Datasets of Bankable Quality? *J. Sol. Energ-T. ASME*. **133**: 031024-13, 2011.

8. Gueymard C.A., Direct solar transmittance and irradiance predictions with broadband models. Pt 2: Validation with high-quality measurements. *Solar Energy*. **74**: 381-395, 2003. Corrigendum: *Solar Energy* **76**, 515, 2004.

9. Gueymard C.A., Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Solar Energy*, in press, 2012.

10. Gueymard C.A. and George R., Gridded aerosol data for improved direct normal irradiance modeling: The case of India. Proc. *Solar 2011 Conf.* Raleigh, NC, American Solar Energy Soc., 2011.