SATELLITE BASED ASSESSMENT OF THE NSRDB SITE IRRADIANCES AND TIME SERIES FROM NASA AND SUNY/ALBANY ALGORITHMS

Paul W. Stackhouse, Jr[1], Taiping Zhang[2], William S. Chandler[2], Charles H. Whitlock[2], James M. Hoell[2]
David J. Westberg[2], Richard Perez[3], Steve Wilcox[4]
[1]NASA Langley Research Center, 21 Langley Boulevard, Hampton, VA 23681-2199
[2]SSAI, One Enterprise Parkway, Suite 300, Hampton, VA 23666-5845
[3]State University of New York at Albany, 1400 Washington Avenue, Albany, NY 12222
[4]National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO 80401

ABSTRACT

In April, 2007, the National Solar Radiation Database (NSRDB) of the National Renewable Energy Laboratory was updated for the period from 1991 to 2005. NSRDB includes monthly averaged summary statistics from 221 Class I sites spanning the entire time period with least uncertainty. In 2008, the NASA GEWEX Surface Radiation Budget (SRB) project updated its satellite-derived solar surface irradiance to Release 3.0. This dataset spans July 1983 to June 2006 at a 1°x1° resolution. In this paper, we compare the NSRDB data monthly average summary statistics to NASA SRB data that has been validated favorably against the BSRN, SURFRAD, WRDC and GEBA datasets. The SRB-NSRDB comparison reveals reasonably good agreement of the two datasets.

1. INTRODUCTION

In 1992, the National Renewable Energy Laboratory (NREL) released its earliest version of the National Solar Radiation Database (NSRDB) for the time span from 1961-1990. In 2003, NREL investigated the feasibility of updating the NSRDB, and eventually updated the database for the period from 1991 - 2005 [1].

The update includes data sets from two models, the Meteorological-Statistical (METSTAT) model [2] and the State University of New York (SUNY) model [3]. The 1991-2005 NSRDB has data from 1,454 ground sites that are divided into three categories: Class I with 221 sites; Class II with 637 sites; and Class III with 596 sites. Class I is of the highest quality and we focus on Class I sites only in this paper. The NSRDB Class I data are temporally continuous, and over 99% of its data are modeled and less than 1% are from instrumental measurements. Only the modeled data are used in this analysis.

The NASA Global Energy and Water-cycle Experiment (GEWEX) Surface Radiation Budge (SRB) project has released its Version 3.0 dataset. This dataset provides radiation data, which include downward and upward shortwave radiative fluxes among a number of other variables, at the Earth's surface and the top of the atmosphere (TOA). The dataset is available to users at a 1° by 1° resolution covering the entire globe, and the time-span of the data covers 23 full years from July of 1983 to June of 2006. Temporally, the data are given as 3-hourly, 3-hourly-monthly, daily and monthly means. The algorithm used to produce the dataset is largely based on the physical processes in the atmosphere and at the Earth's surface. For inputs, the ISCCP-DX data provide cloud and surface properties [4]; the GEOS-4 reanalysis provides temperature and humidity information [5]; and a composite of TOMS, TOVS and assimilated SBUV-2 datasets provides column ozone information.

An important part of the GEWEX SRB project is the validation of its products through comparison with ground-based observations. SRB estimates of surface shortwave fluxes are compared against measurements from the Baseline Surface Radiation Network (BSRN), the World Radiation Data Centre (WRDC) and the Global Energy Balance Archive (GEBA) data for the validation purposes.

As of early 2008, the BSRN has provided the highest quality radiation data at 39 ground sites on all seven continents starting from 1992, and has 3734 site-months of data in archive. The overall bias and RMS difference between the SRB dataset and the BSRN dataset is -3 W m^{-2} (-1.8%) and 20 W m^{-2} (12.1%) respectively. Between 60° N and 60° S, these differences are reduce to -1.2 W m^{-2} (0.6%) and 14.4 (7.7%).

The GEBA data set contains monthly means of shortwave radiation at 797 ground sites from 1983 to 2003. These data are not considered as accurate as the BSRN data sets but have undergone stringent quality control testing. The overall bias and RMS differences of SRB compared to these measurements are 6.2 W m^{-2} (3.9%) and 23.0 W m^{-2} (14.6%) respectively.

Lastly, WRDC has daily mean shortwave data over 474 ground sites from 1983 to 1993. These sites are mostly the same sites as GEBA above but with different quality control. In comparison to SRB, mean and RMS differences relative to monthly averages of these measurements are 6.0 W m^{-2} (3.7%) and 28.2 W m^{-2} (17.7%).

These results give the satisfying result that on the whole SRB data agree better with high quality measurements, particularly between the latitudes of 60° N and S where the agreement with BSRN is less than 8%. In this paper, the monthly mean surface shortwave downward fluxes are compared between SRB and NSRDB Class I locations from the Continental US and Hawaii. A subset of the BSRN network, the SURFRAD Network [6] is used to help provide a benchmark for the continental US.

It is found that NSRDB Class I statistical summary data are generally in excellent agreement with SRB except at a limited set of sites in Hawaii and the US. The overall agreement and the variability of the data sets in time are assessed.

2. <u>15-Year NSRDB and NASA SRB Comparisons</u>

The 221 NSRDB sites have continuous records from January, 1991 to December, 2005, and thus have a total of 39,780 data points for monthly means. Here, the monthly averaged values from the statistical summaries are compared to GEWEX SRB values. The statistical summaries are comprised of METSTAT model estimates from 1991 to 1997 and SUNY model from 1998 to 2005.

Figures 1 and 2 show the bias and RMS of the SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux comparison on a site-by-site basis. These figures show the locations of the 221 sites as well. Note that there are no Class I sites in Alaska. The four (4) Hawaiian sites show both large bias and RMS the same time. Several other coastal and mountainous sites show large differences, although there are multiple sites in such regions that give significantly better agreement. However, the plots clearly show that most sites agree with SRB to within +/- 10 W m^{-2} and < 14 W m^{-2} for bias and RMS differences respectively.

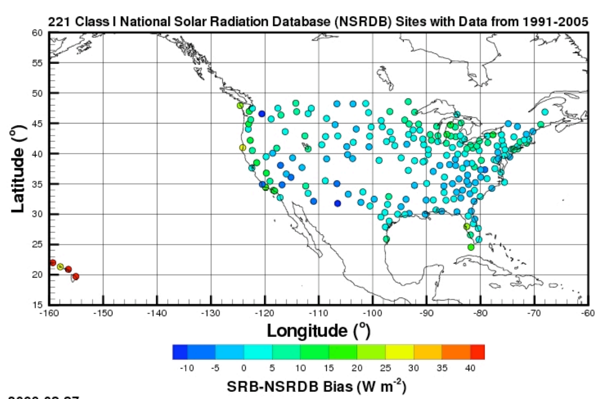

Fig. 1. SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux comparison bias on a site-by-site basis.

Figure 3 shows the SRB(V3.0)-NSRDB agreement in the form of a simple scatter plot. The overall agreement is very good with bias and RMS differences of 3.2 W m^{-2} (1.8%) and 15.1 W m^{-2} (8.7%) respectively. The r^2 correlation coefficient between the two data sets 0.96. The largest density of points lie clearly along the line of one-to-one

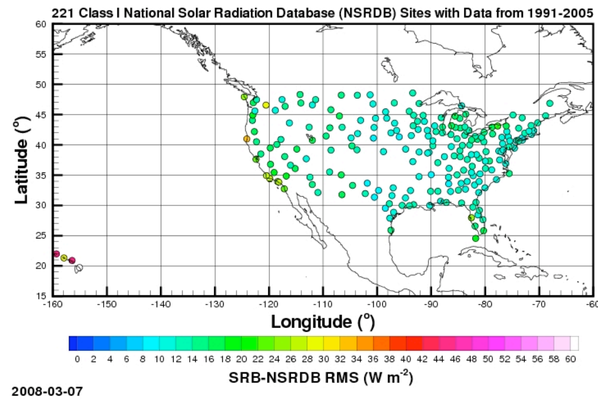

Fig. 2. SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux comparison RMS on a site-by-site basis.

agreement shown for reference but there are some clusters of outlier points.

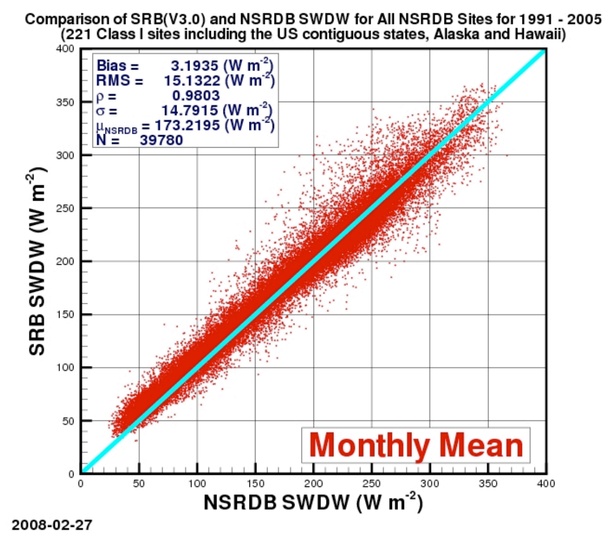

Fig. 3. Scatter plot of SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux comparison.

To gain some insight into those outliers, Figure 4 provides the statistical agreement in absolute bias and RMS for each site. The plot clearly shows that most sites agree within the bias of -5 to 10 W m^{-2} and RMS of 10 - 15 W m^{-2}. It also clearly shows several sites that agree much more poorly.

These sites include the four (4) Hawaiian sites and 11 other sites most of which are on the West Coast of the US. Interestingly, it appears that most of sites numbered 1-110 have biases that range from +/- 5 W m^{-2} and sites from about 111-221 are characterized by a range of 0-10 W m^{-2}. The sites are numbered such that the larger the number the further northward the site. Sites > 110 are located at 40^{o} degrees N and higher. The exception is the 4 Hawaiian sites that correspond to numbers 218-221. There is no SRB discontinuity at the 40^{o} latitude so this phenomenon will have to be studied further.

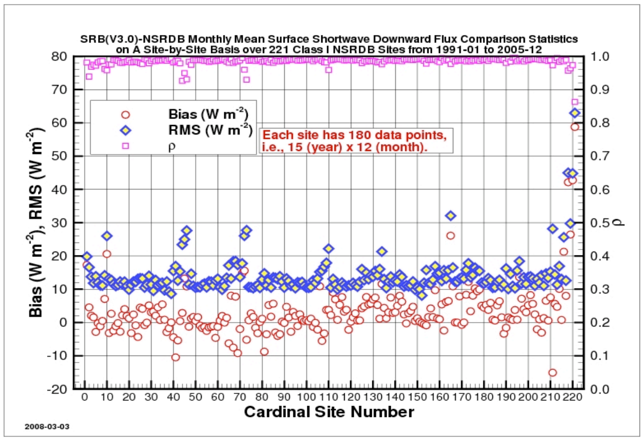

Fig. 4. SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux bias, RMS and correlation coefficient on a site-by-site basis.

Figure 5 provides a simple scatter plot with the 15 outlier sites excluded. The bias and RMS are reduced to 2.1 W m^{-2} (1.2%) and 12.9 W m^{-2} (7.5%) respectively for the remaining sites. The correlation increases to 0.973, thus indicating the extent to which the sites of poor agreement affect the overall agreement.

Figure 6 shows the SRB(V3.0)-NSRDB mean bias, RMS and correlation coefficient on a month-by-month basis for the entire ensemble of Class I sites. Biases vary +/- 10 W m^{-2} with RMS differences ranging between 10 - 20 W m^{-2}. The period between 1996 and 1998 seems to be an exception to this pattern in that RMS differences for several of those months exceeded 20 W m^{-2}. Several explanations such as a change in the satellite calibration and/or viewing and changes in the cloud observations for the NSRDB METSTAT model will be evaluated as potential explanations for these changes, but no further explanation is presently available.

Figure 6 also shows a tendency for a positive bias in fall/winter months followed by a change to a negative bias in summer months. This is illustrated more clearly in Figure 7 providing the frequency distribution of bias for the months January through December as aggregated over the entire 15 year data set. In this figure, each month has 221 (site) x 15 (year) = 3315 data points. The month-by-month distributions clearly show the tendency for positive peaks

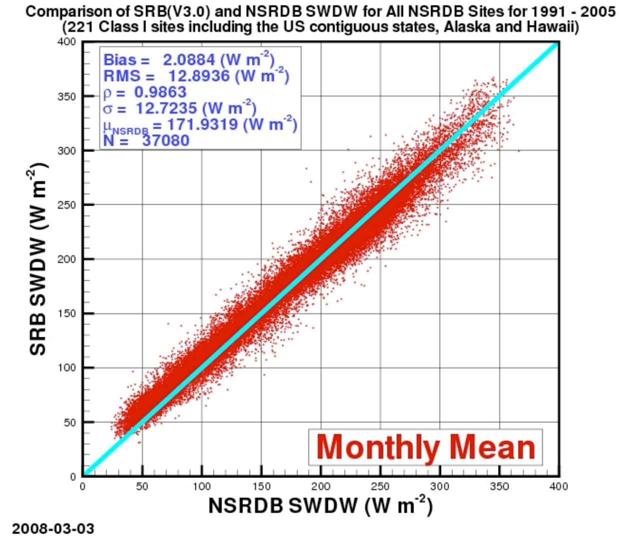

Fig. 5. Same as Figure 3 except four (4) Hawaiian sites and other 11 sites are excluded.

in bias distributions for the months of October to December in the 5-15 W m^{-2} bin. The peak negative bias occurs in June and is in the -15 - -5 W m^{-2}. We also note the bias distributions are much more peaked and narrow in the month of October through December than the May and June indicating larger noise in those months.

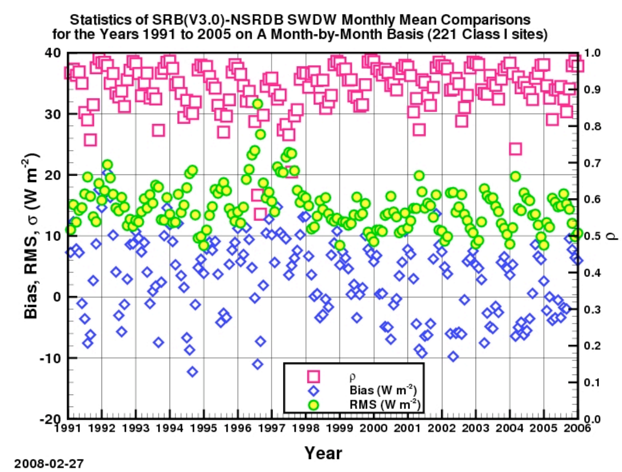

Fig. 6. SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux bias, RMS and correlation coefficient on a month-by-month basis.

The box-whisker depiction of the bias and variability of those monthly distributions shows the characteristics more quantitatively in Figure 8. For this figure, the yellow cross is the MEDIAN of the set. The lower boundary of the solid box is the lower quartile (Q1), the upper boundary of the solid box is the upper quartile (Q3) and the difference, Q3 -

Q1, is called the InterQuartile Range (IQR). The lower red bar (in thin red line) is (Q1-1.5xIQR) and the upper red bar (in thin red line) is (Q3+1.5xIQR). The data points that are between (Q1-1.5xIQR) and (Q1-3.0xIQR), or between (Q3+1.5xIQR) and (Q3+3.0xIQR), are called "mild" outliers, and these are solid blue dots. The data points that are below (Q1-3.0xIQR) or above (Q3+3xIQR) are called "extreme" outliers, and these are in green circles. Figure 8 clearly shows that the median biases shift from greater than zero in the months of September to March to less than zero from April through August. The reasons from the shifting are not entirely clear. It is also evident that the summer months give distributions with larger IQR and contain far more outlier points. It is important to note that the 15 sites with poorer agreement have not been removed from this analysis.

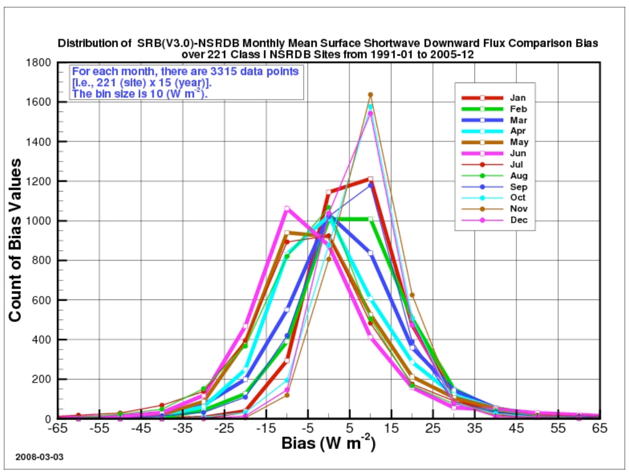

Fig. 7. Distribution of bias January through December. Each month has 3,315 data points, i.e., 221 sites time 15 years.

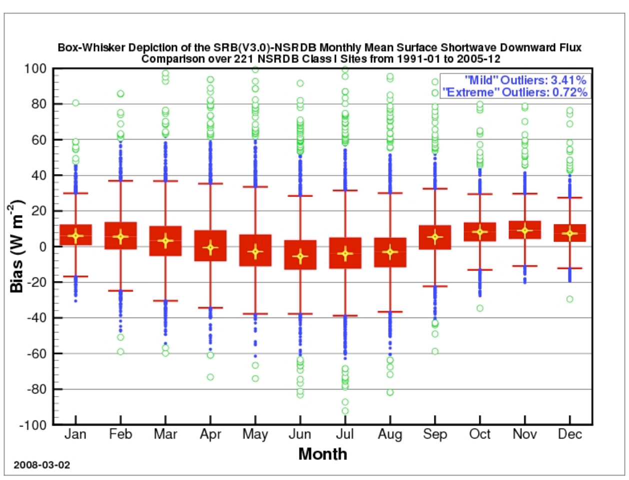

Fig. 8. A box-whisker presentation of the SRB(V3.0)-NSRDB monthly mean surface downward shortwave flux comparison bias for January through December.

3. <u>10-Year SURFRAD and NASA SRB Comparisons</u>

The results of the last section show that the relative difference between SRB and NSRDB changes exhibits an annual cycle. The reasons for this will become the subject of continued study and evaluation. Here we evaluate the SRB relative to the SURFRAD network of sites to show whether the same patterns of differences are consistent with those observed between NSRDB and SRB. The SURFRAD sites were initiated with 5 sites in 1995 with additional sites added over the years. Figure 9 shows the overall scatter plot for the SURFRAD and SRB monthly averages for all sites and years (roughly 5 sites and 10 years). The bias and RMS are 5.1 W m^{-2} (2.8%) and 18.9 W m^{-2} (10.6%) respectively with a 0.946 r^2 value.

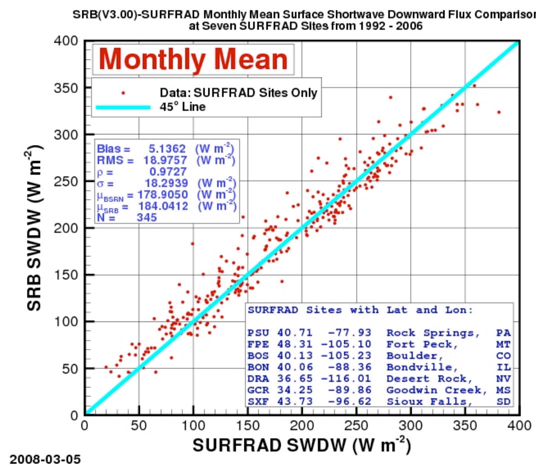

Fig. 9. Scatter plot of the monthly averaged solar irradiance measurement versus the NASA SRB estimate for the 1^{o}x1^o grid box containing the site. The sites are listed on the plot.

Figure 10 provides the box whisker plot in the same format as Figure 8, but for the SURFRAD data sets. This figure shows that SRB and SURFRAD show a similar bias, distribution width and outlier pattern as that with NSRDB. Although there are differences, probably due to the changing sites during the time period under evaluation, the similarity in patterns may imply that the bias differences could more due to uncertainties in the SRB algorithm than in the observations or in NSRDB.

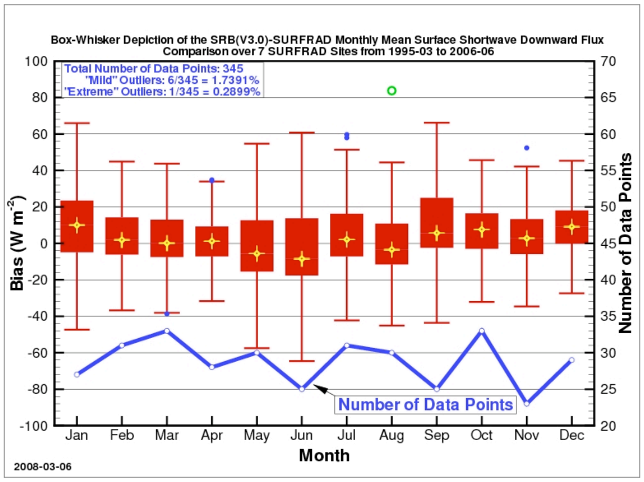

Fig. 10. A box-whisker presentation similar to Figure 8 but this time comparing monthly averaged SRB(V3.0) to measurements.

4. Temporal NSRDB and NASA SRB Comparisons

We next examine the long-term variability of the ensemble of the NSRDB and SRB datasets. To do this, we perform 12-month running averages on the data. Figure 11 shows the time–series of both the SRB(v3.0) and NSRDB data and their difference. Error bars are based on their respective standard deviations. The red circles give the NSRDB ensemble average of the running means and the green line show the same for SRB. The blue line gives the SRB - NSRDB difference as a function of time. The most interesting features of the ensemble differences are that the biases between 1991 and 1999 fluctuate between about 1 and 9 W m^{-2} with a mean of about 5 W m^{-2} (2.9%) during the period but then transition to a bias near 0 from 2000 through 2005. This corresponds with the switch in the NSRDB summary statistics from METSTAT model values through 1997 to SUNY model derived values beginning in

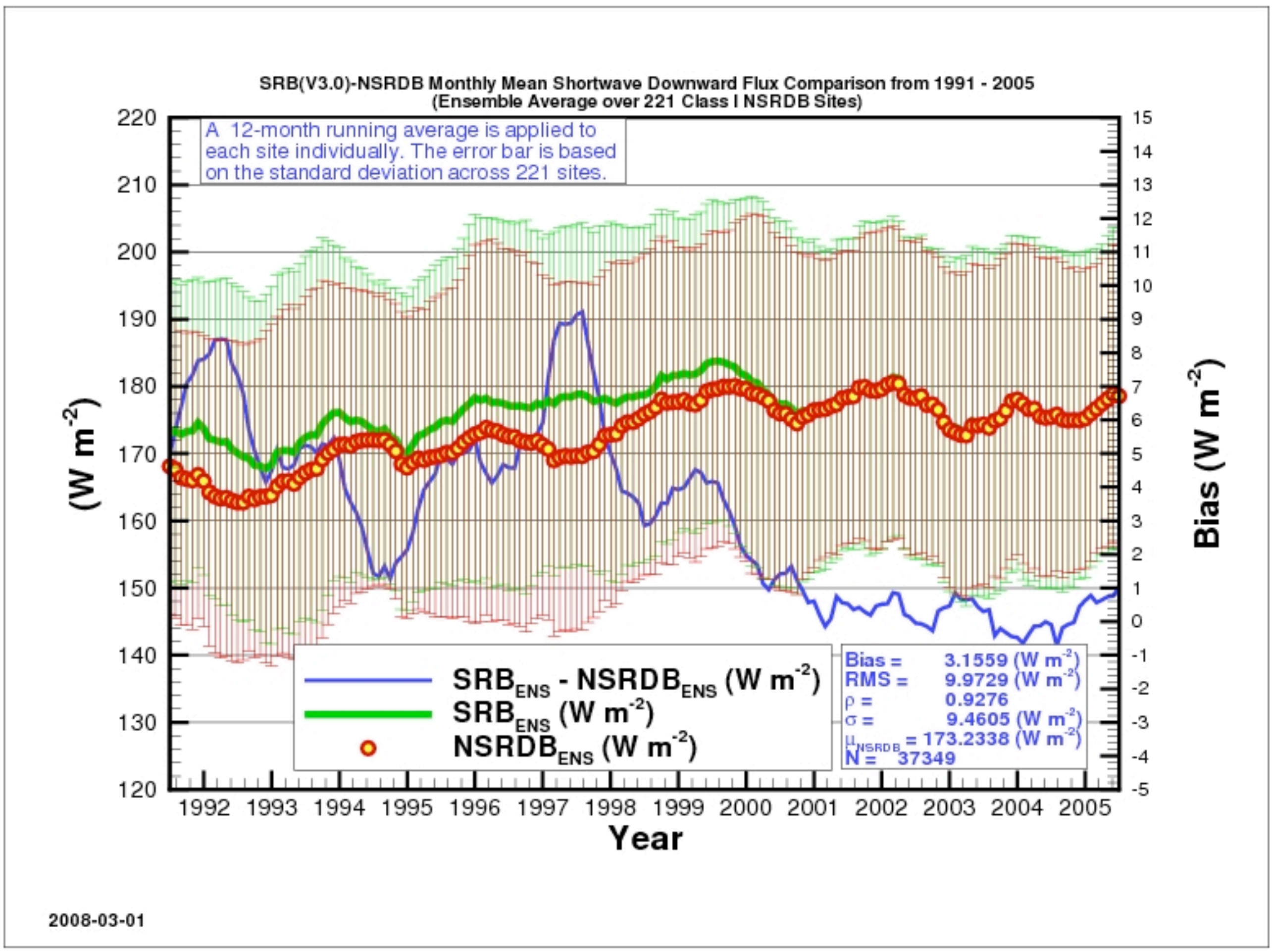

Fig. 11. A 12-month running average is applied to each site individually and then an ensemble average across the 221 sites is performed on both SRB(V3.0) and NSRDB data. The error bars are based on standard deviation of each set respectively. Note that the first five (5) months in 1991 and last six (6) months of 2005 disappear from the picture after the 12-month running average. Also, note that METSTAT model estimates comprise the nearly all the NSRDB values from 1991-1997 and SUNY model estimates from 1998-2005.

1998. The large fluctuation in the summers of 1992, 1994 and 1997 correspond to times of an active El Nino. It is not yet known whether this is just a coincidence or implies physical significance. The 12-month running mean reduces the overall RMS difference to 10.0 W m^{-2} (5.8%) a reduction of about 33%.

Lastly, we note the both NSRDB and SRB show an increase surface solar irradiance from 1991 to 2000 at which point both appear to stabilize. The change in NSRDB over this period of time is estimated at 12-15 W m^{-2}. The SRB change over this time is estimated at 8-10 W m^{-2}. These trends will be studied and quantified in future work. It is noted that the NSRDB change is steeper than the change indicated by SRB and that it appears that SRB shows indication of a decrease owing to higher values in 2000 than NSRDB. The effect of the 1998 transition from METSTAT to SUNY is still being assessed.

5. RESULTS AND CONCLUSIONS

A preliminary analysis of the comparisons between the newly released NASA/GEWEX SRB (v3.0) solar irradiance fluxes and the Class I NSRDB site values from the summary statistics covering 1991-2005 was presented. The overall comparison yielded excellent results with RMS differences over all sites less than 9% for monthly averages. This level of quality is well within the agreement attained for SRB comparisons to true surface measurement networks. Out of the 221 Class I NSRDB sites, SRB(V3.0) and NSRDB disagree appreciably at just 15 sites including four (4) Hawaiian sites. The proximity of these sites to mountainous or coastal areas indicates that the resolution of SRB could be most significant reason for the poor agreement at those sites, indicating a need for more investigation.

Additional month-by-month comparisons showed that the SRB-NSRDB differences have a seasonal dependence such that SRB had positive biases during the fall/winter months and negative biases during the summer months. The fluctuations in median bias were +/- 7 Wm^{-2}. Since similar patterns of differences were found between NOAA SURFRAD sites in the continental US and SRB, it is preliminarily concluded that the source of the differences may be due to the SRB data set with more investigation needed.

Lastly, 12-month running averages significantly reduced the RMS differences by about 33% between SRB and NSRDB. There is a correspondence between the relative shift in bias between NSRDB and GEWEX SRB and the change from predominantly METSTAT model derived estimates in 1997 to SUNY model derived from 1998 onward. The 12-month running averages shown in Figure 11 smooth the transition. The month-to-month changes indicated that both SRB and NSRDB show an increase in solar irradiance over the US from 1991 - 1999 and a stabilization thereafter. However, the change from METSTAT to SUNY makes any trend inference subject to considerably higher uncertainty despite apparent agreement with the SRB satellite derived measurements.

Future work aims to better quantify and explain these differences from hourly METSTAT and SUNY values. Since the METSTAT model derived values span the entire NSRDB data set, a three way comparison between GEWEX SRB, SUNY and METSTAT derived irradiances will be performed. This will better help clarify the preliminary conclusions reached here.

6. ACKNOWLEDGEMENTS

This research is funded through the NASA Applied Sciences program within the Earth Science Directorate. NASA has MOU with NREL which has facilitated a continued collaboration in the updating and evaluation of the National Solar Radiation Database.

7. REFERENCES

1. National Renewable Energy Laboratory (NREL), 2007. National Solar Radiation Database 1991-2005 Update. Technical Report, NREL/TP-581-41364. April, 2007.
2. Perez, R., et al., 2002. A New Operational Satellite-to-Irradiance Model. Solar Energy, 73(5), 307-317.
3. Maxwell, E.L., 1998. METSTAT-The Solar Radiation Model Used in the Production of the National Solar Radiation Database (NSRDB). Solar Energy, 62(4), 263-279.
4. Rossow W. B, Schiffer, R. A., (1999). "Advances in understanding clouds from ISCCP", *Bull. Am. Met. Soc.,* 80: 2261-2287.
5. Bloom., S. A. da Silva, and D. Dee, (2005). Documentation and Validation of the Goddard Earth Observing System (GEOS) Data Assimilation System – Version 4. NASA Technical Report Series on Global Modeling and Data Assimilation, Max J. Suarez, Editor, NASA/TM-2005-104606, **26**.
6. Augustine J. A., J. J. DeLuisi, and C. N. Long, 2000: SURFRAD—a national surface radiation budget network for atmospheric research. Bull. Amer. Meteor. Soc., 81, 2341–2357.