Comparison between Total Cloud Cover in Four Reanalysis Products and Cloud Measured by Visual Observations at U.S. Weather Stations

MELISSA FREE

NOAA/Air Resources Laboratory, College Park, Maryland

BOMIN SUN

I. M. Systems Group at NOAA Center for Satellite Application and Research, College Park, Maryland

HYE LIM YOO

Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

(Manuscript received 4 September 2015, in final form 7 December 2015)

ABSTRACT

A homogeneity-adjusted dataset of total cloud cover from weather stations in the contiguous United States is compared with cloud cover in four state-of-the-art global reanalysis products: the Climate Forecast System Reanalysis from NCEP, the Modern-Era Retrospective Analysis for Research and Applications from NASA, ERA-Interim from ECMWF, and the Japanese 55-year Reanalysis Project from the Japan Meteorological Agency. The reanalysis products examined in this study generally show much lower cloud amount than visual weather station data, and this underestimation appears to be generally consistent with their overestimation of downward surface shortwave fluxes when compared with surface radiation data from the Surface Radiation Network. Nevertheless, the reanalysis products largely succeed in simulating the main aspects of interannual variability of cloudiness for large-scale means, as measured by correlations of 0.81–0.90 for U.S. mean time series. Trends in the reanalysis datasets for the U.S. mean for 1979–2009, ranging from -0.38% to -1.8% decade^{-1}, are in the same direction as the trend in surface data (-0.50% decade^{-1}), but further effort is needed to understand the discrepancies in their magnitudes.

1. Introduction

Clouds are a critical element in numerical climate and forecast models since they strongly affect the radiation balance and are also a major factor in determining climate sensitivity. Unfortunately, the recent evolution of cloud amount on decadal and longer scales is unclear because of uncertainties in both satellite (e.g., Norris and Evan 2015; Sun et al. 2015) and ground-based observational records (e.g., Free and Sun 2013). Given these uncertainties, it may be tempting to use reanalysis outputs to assess trends in cloud cover. However, although reanalyses are created to facilitate climate monitoring and climate change assessment, they often do not accurately represent long-term climate trends for precipitation and temperature due primarily to changes in observational inputs over time and to problems with spatial distribution of observations (e.g., Lorenz and Kunstmann 2012). Since cloud is not assimilated into these reanalyses but instead is predicted by the models, cloud cover from reanalyses is also subject to errors related to physical parameterizations used in the model. Because of these problems, it is important to assess the reanalyses against long-term time series of observations.

Previous studies have compared climatological cloud amount in operational models or reanalysis datasets to satellite cloud data or ground-based remote sensing cloud data. Some have found underestimates of cloud amount in various large-scale regions including midlatitude oceans (Bedacht et al. 2007) and midlatitude zonal means (Jakob 1999) in previous versions of the ECMWF and NCEP reanalyses, although overestimates may also occur, as, for example, in the Arctic during the

---

Corresponding author address: Melissa Free, NOAA/Air Resources Laboratory, 5830 University Research Court, College Park, MD 20740.
E-mail: melissa.p.free@gmail.com

DOI: 10.1175/JCLI-D-15-0637.1

cold season (Clark and Walsh 2010). In the United States, several papers have found cloud cover at the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) site to be underestimated by ERA-Interim, NCEP–NCAR reanalyses (e.g., Wu et al. 2012), and MERRA (Kennedy et al. 2011). However, only a few papers have compared the evolution of cloudiness on interannual or decadal scales in models or reanalyses to those in observations. Jakob (1999) found that the previous ECMWF reanalysis (ERA-40) was able to capture the interannual variability of cloud cover seen in ISCCP satellite data although it underestimated extratropical cloud cover and overestimated trade wind cumulus. On the other hand, Bedacht et al. (2007) found the time series of cloud cover from visual observations over the ocean were not well correlated with those in the NCEP–NCAR reanalyses.

Here we use a homogeneity-adjusted dataset of total cloud cover from weather stations in the contiguous United States to assess cloud cover in four current reanalysis products from 1979 to 2009 with particular attention to the interannual variability and trends in the datasets. We also examine the relation of biases in cloud cover to biases in surface solar radiation in the reanalysis products.

2. Data used

a. Weather station visual observations

The total cloud cover dataset from U.S. weather stations, described in detail in Free and Sun (2014), consists of data from 54 NWS stations and 101 military weather stations in the contiguous United States that continued to make visual observations of clouds after the introduction of automated systems at many NWS stations in the 1990s. Human observers record cloud fraction, in eighths or tenths of sky cover, or sky condition (e.g., clear, scattered, overcast, etc.). To avoid problems with nighttime visual cloud observations, we used reports from 3-hourly daytime observations to construct the monthly mean cloud cover dataset (Free and Sun 2014). The dataset has been adjusted for temporal homogeneity as described in Free and Sun (2014) and was used to evaluate satellite cloud datasets in Sun et al. (2015).

The weather station observer is looking at the sky from a single point, which will likely give a different result than that from a satellite or a reanalysis gridbox value. For example, the single-point view of the whole sky dome tends to give a larger cloud amount for areas close to the horizon in the presence of broken clouds because the observer sees the sides as well as the bottoms of clouds. Wu et al. (2012) estimated that this effect could reduce the reanalysis cloud fraction by 2%–6% from that seen by whole sky–viewing ground-based instruments (which have a similar point of view to that of a human observer). This difference in field of view may contribute to some extent to the climatological biases shown in that paper and in the present work.

b. Reanalyses

1) CFSR

The Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) was produced by NOAA using NCEP's Climate Forecast System and covers 1979–2009 with a high-resolution coupled atmosphere–ocean–land–sea ice system. Inputs include time-varying CO_2 and stratospheric aerosols from volcanoes. We used "regular" monthly means, which combine values from all four initialization times for each day in the month. The values used are forecasts of 6-h averaged total cloud cover, with horizontal resolution of ½° latitude × ½° longitude (downloaded from http://rda.ucar.edu/datasets/ds093.2/#!access).

2) MERRA

The Modern-Era Retrospective Analysis for Research and Applications (MERRA) was created by NASA using the Goddard Earth Observing System Model, version 5 (GEOS-5), and designed to focus especially on the hydrological cycle (Rienecker et al. 2011). It uses historical CO_2 values and analyzed ozone but climatological values for aerosols and other trace gases. We used the time-averaged monthly means of total cloud from four daily initialization times obtained from ''MATMNXRAD History'' files, with a horizontal resolution of ⅔° longitude × ½° latitude (found at ftp://goldsmr2.sci.gsfc.nasa.gov/data/s4pa/MERRA_MONTHLY/).

3) ERA-INTERIM

The European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011) is designed as a bridge between ERA-40 (from September 1957 to August 2002) and the next-generation reanalysis. Improvements in ERA-Interim compared to ERA-40 include four-dimensional variational data assimilation, improvements in radiative transfer modeling, changes in cloud parameterization, and bias correction of satellite data. This reanalysis uses climatological rather than time-evolving CO_2 and aerosols. We used monthly regular means (average of four daily analyses at 0000, 0600, 1200, and 1800 UTC) with horizontal resolution of ½° latitude × ½° longitude (provided at http://apps.ecmwf.int/datasets/data/interim-full-moda/).

4) JRA-55

The Japanese 55-year Reanalysis Project (JRA-55) is the second version of the Japanese Reanalysis Project produced by the Japan Meteorological Agency and covers 55 years from 1958 to 2013. JRA-55 shows an improvement over the previous product, the Japanese 25-year Reanalysis Project (JRA-25), with increased spatial resolution, a new radiation scheme, and four-dimensional variational data assimilation (Kobayashi et al. 2015). It uses time-varying well-mixed greenhouse gases. We use monthly mean total cloud cover data from JRA-55 monthly mean model-resolution two-dimensional instantaneous diagnostic fields, produced from the output hours at 0000, 0600, 1200, and 1800 UTC (data are available at http://rda.ucar.edu/datasets/ds628.1/#!access).

c. *SURFRAD*

We previously used surface radiation data from the seven Surface Radiation Network (SURFRAD) stations (Augustine et al. 2005) in the United States to evaluate cloud cover variations in satellite cloud products (Sun et al. 2015). In this study, this dataset is employed to investigate whether biases in cloud cover cause biases in surface solar radiation in the reanalysis products. The stations are located at Desert Rock, Nevada; Bondville, Illinois; Table Mountain, Colorado; Fort Peck, Montana; Goodwin Creek, Michigan; Pennsylvania State University, Pennsylvania; and Sioux Falls, South Dakota. The available time periods for each site vary between the stations, beginning no earlier than 1995. The SURFRAD data are monthly, accessed from Global Monitoring Division, NOAA/Earth System Research Laboratory (ftp://aftp.cmdl.noaa.gov/data/radiation/surfrad/).

3. Methods

We extract regular monthly mean total cloud cover data (averaged from all four initialization times) from the reanalyses at grid boxes that contain the locations corresponding to our weather stations. For comparison, we also extracted data for 1800 UTC (which is closer to the observation times for our surface cloud data) for the three datasets (CFSR, ERA-Interim, and MERRA) that had that data available. Results presented here are for the regular rather than the 1800 UTC mean unless stated otherwise.

We use the time period 1979–2009, the longest period for which all cloud datasets listed in section 2 are available. We compute mean biases of total cloud cover between reanalysis and station data for each station and the U.S. mean of those biases. The annual cycle is removed to create anomalies using the reference period of 1979–2009. Pearson correlation coefficients are computed between monthly mean anomaly time series. Trends are calculated using least squares linear regression, and confidence intervals for the trends are derived using a correction for autocorrelation in the time series. We create U.S. mean time series by combining time series of anomalies in 2.5° grid boxes so as to reduce spatial sampling errors. To test the significance of the difference between trends from two products, we examine the significance of the trend in the time series of the difference between monthly means for the two products. To evaluate the cloud cover bias with surface solar radiation data, we selected the reanalysis grid box closest to the U.S. SURFRAD stations.

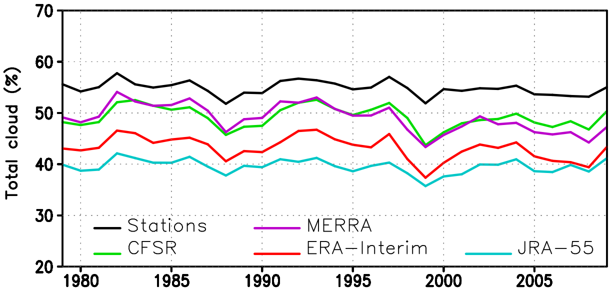
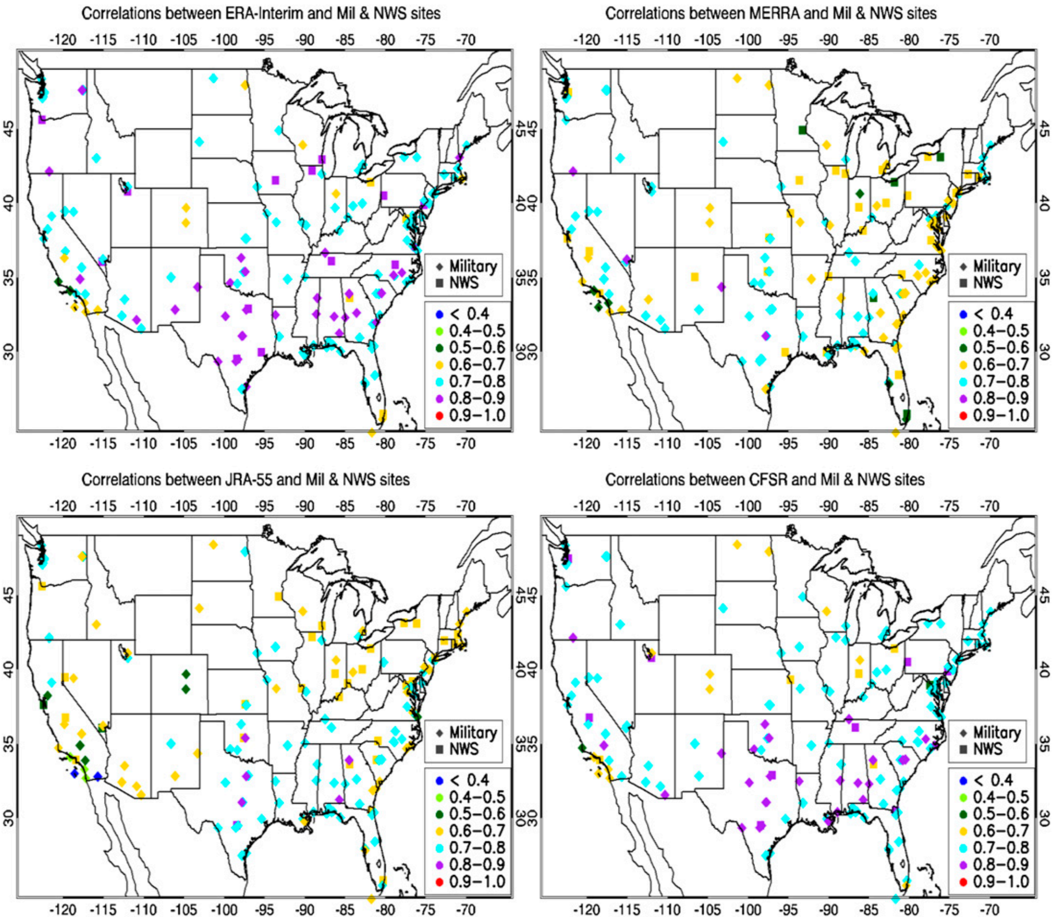


FIG. 1. Annual-mean total cloud cover from four reanalyses and weather station data averaged over 154 U.S. locations using gridded station and reanalysis data.

4. Results

a. *Climatological means and biases*

All reanalyses have lower cloud cover amount than the weather station observations for all regions and for the U.S. mean (Fig. 1), suggesting that there may be a bias in cloud-related physics shared by those climate models. The U.S. mean difference is ~9% for MERRA and CFSR, ~15% for ERA-Interim, and ~19% for JRA-55. Reanalysis biases of cloud cover at individual stations vary from slightly positive to −27%. ERA-Interim biases are negative at all stations; the other three reanalyses have negative biases at all but a handful of stations. Biases for 1800 UTC diurnal means from reanalysis data are generally within 3% of those for the entire 24-h cycle, and are sometimes larger despite the observations being more closely matched in time, especially for MERRA. CFSR, ERA-Interim, and JRA-55 tend to give larger annual cycles than those in the station data, while MERRA gives a slightly smaller cycle in the U.S. mean and does not reproduce the summer minimum in cloud cover seen in the station data (not shown). Biases for the U.S.


Correlations between ERA-Interim and Mil & NWS sites

Correlations between MERRA and Mil & NWS sites

Correlations between JRA-55 and Mil & NWS sites

Correlations between CFSR and Mil & NWS sites

FIG. 2. Correlations between reanalysis total cloud cover time series and weather station data for U.S. locations for 1979–2009.

mean are largest in summer for all but MERRA, which has much smaller biases in July and August than in the rest of the year. Root-mean-square differences between monthly mean reanalysis cloud cover and that from station data are smallest for CFSR (~12%) and largest for JRA-55 (~20%), indicating that CFSR does a better job in representing absolute total cloud cover seen from the weather stations.

The underestimate of cloud cover in the United States is similar to findings in a number of earlier studies for other regions. We used reanalysis cloud cover for all hours of the day, while our station dataset is limited to daytime hours. However, since results for 1800 UTC are generally similar to those for whole-day means, we conclude that the mismatch of observation times is not a significant source of differences between reanalyses and station data or between different reanalysis products. Another potential reason for biases may stem from the difference between the top-down reanalysis viewpoint and that of a weather observer, as described in section 3a. However, based on estimates from Wu et al. (2012), this difference does not appear to be enough to explain the biases shown here.

b. *Correlations between time series*

Pearson correlation coefficients between monthly anomaly time series from reanalyses and weather stations for individual stations (Fig. 2) range from 0.31 for JRA-55 at San Clemente, California, to 0.88 for ERA-Interim at Klamath Falls, Oregon. For all but MERRA, there is a tendency for the best correlations to occur at stations in the Northwest, in Texas, and in the Southeast and the worst in California. For U.S. mean monthly reanalysis time series, correlations with station data are highest for ERA-Interim (0.90) and lowest for MERRA (0.81) (Table 1), with CFSR (0.89) very close to ERA-Interim. The high correlation between the reanalysis products and surface data could be attributed to reliable interannual signals from assimilated observations including temperature and moisture that are used in cloud

TABLE 1. Pearson correlation coefficients between monthly U.S. mean reanalysis total cloud cover and that from weather station data, with least squares trends (% decade^{-1}) in U.S. mean total cloud cover and their confidence intervals (twice the standard error of the trend) for 1979–2009. The U.S. means are derived from gridded data. All correlations are significant at the 0.05% level.

Dataset	Correlation	Trend
CFSR	0.89	0.72 ± 0.74
ERA-Interim	0.90	1.09 ± 0.73
JRA-55	0.85	0.38 ± 0.51
MERRA	0.81	1.84 ± 0.76
Stations		0.50 ± 0.50

parameterization. Military and NWS station subsets do not differ strongly in correlations with the reanalyses, and correlations for the 1800 UTC diurnal mean from reanalysis outputs are not generally better than those for the full 24-h means.

All reanalysis products show the best correlations with station data in the fall or winter and the worst correlations in summer (Table 2). MERRA has much lower correlations in summertime with both military and NWS stations, and this appears to be the primary reason for its lower overall correlation. The poorer correlations in summer and in California, which are also shown in satellite cloud products (see Sun et al. 2015), could be related to greater small-scale spatial or temporal variability in cloud cover captured by the weather stations, or to issues with the representation of specific cloud types that are more common in summer and in certain regions. Specifically, small cumulus clouds tend to be more frequent in the summer, and those clouds are more likely to be ‘‘seen’’ differently by observers than by a model or satellite, whereas the stratus clouds that are common in the winter are more likely to produce similar estimates from both top-down and ground observers.

Examination of low and high cloud cover or individual cloud types might help indicate the reasons for differences between reanalyses and station data. However, individual cloud types similar to those recorded by weather observers are not generally available from reanalyses, and the availability of low cloud and other cloud type information from the U.S. weather stations is much more limited since the 1990s than before then, making such comparisons difficult.

c. Trends

Table 3 shows the trends in the U.S. mean time series from the reanalyses and the station data. Like the satellite products compared in Sun et al. (2015), reanalyses typically show more negative trends than the station data; for the U.S. mean and the period 1979–2009, this difference is greatest for MERRA (~1.4% decade^{-1}), while the trend in JRA-55 is within 0.12% decade^{-1} of that in the station data. The MERRA and ERA-Interim trends are significantly different from those in the station data while the JRA-55 and CFSR trends are not. In both station and reanalysis data, trends seem to be more negative in winter than in spring or summer (Table 3). For MERRA and CFSR, the difference from trends in station data is greatest for summer. Trends for JRA-55 are within 0.3% decade^{-1} of those in station data for all seasons.

Trends for 1800 UTC reanalysis output for the U.S. mean, where available, are within 0.5% decade^{-1} of those for all hours. For CFSR and ERA-Interim, 1800 UTC output has a more negative trend than that for all hours, while the reverse is true for MERRA. Despite these small changes, the trend for MERRA in the 1800 UTC output is still significantly different from the station trend, and those for CFSR and ERA-Interim are not.

d. Relation to surface solar radiation

We expect that an underestimate of cloudiness will lead to an overestimate of surface solar radiation. Such overestimates have been found in many GCMs and reanalyses (Wild 2008; Boilley and Wald 2015). To test this, we examined the surface radiative fluxes from the reanalyses and compared these to the data from SURFRAD stations using scatterplots (Fig. 3). As expected, the surface solar radiation in the reanalyses is generally overestimated. An exception occurs at Desert Rock for ERA-Interim, which generates too little surface solar radiation at that location although cloud cover is underestimated rather than overestimated in that area. [This can be seen as a roughly linear collection of points in Fig. 3 (top left).] Overall, the overestimate of surface solar radiation is less severe for ERA-Interim

TABLE 2. Pearson correlation coefficients between U.S. mean time series of total cloud cover from reanalysis and weather station data by season. All correlations except that for MERRA in June–August are significant at the 0.05% level.

	December–February	March–May	June–August	September–November
CFSR	0.92	0.80	0.63	0.97
ERA-Interim	0.94	0.92	0.61	0.96
JRA-55	0.94	0.80	0.66	0.93
MERRA	0.93	0.89	0.26	0.90

TABLE 3. Least squares linear trends for 1979–2009 by season, from gridded U.S. means. None of the seasonal trends is significant at the 0.05 level.

	December–February	March–May	June–August	September–November
CFSR	−0.99	−0.92	−0.12	−0.89
ERA-Interim	−1.62	−1.20	−0.09	−1.48
JRA-55	−0.71	−0.34	0.44	−0.89
MERRA	−1.69	−1.31	−1.75	−2.59
Stations	−0.83	−0.61	0.51	−0.91

than for the other reanalyses, although cloud cover is not less biased in that product than in the other reanalyses.

We calculated correlations between time series of cloud cover bias (reanalysis minus station values) and those of surface shortwave radiation bias (reanalysis minus SURFRAD values). The correlation coefficients ranged from −0.82 to 0.16, with a mean of −0.34. Bias correlations are lower for ERA-Interim (mean of −0.23

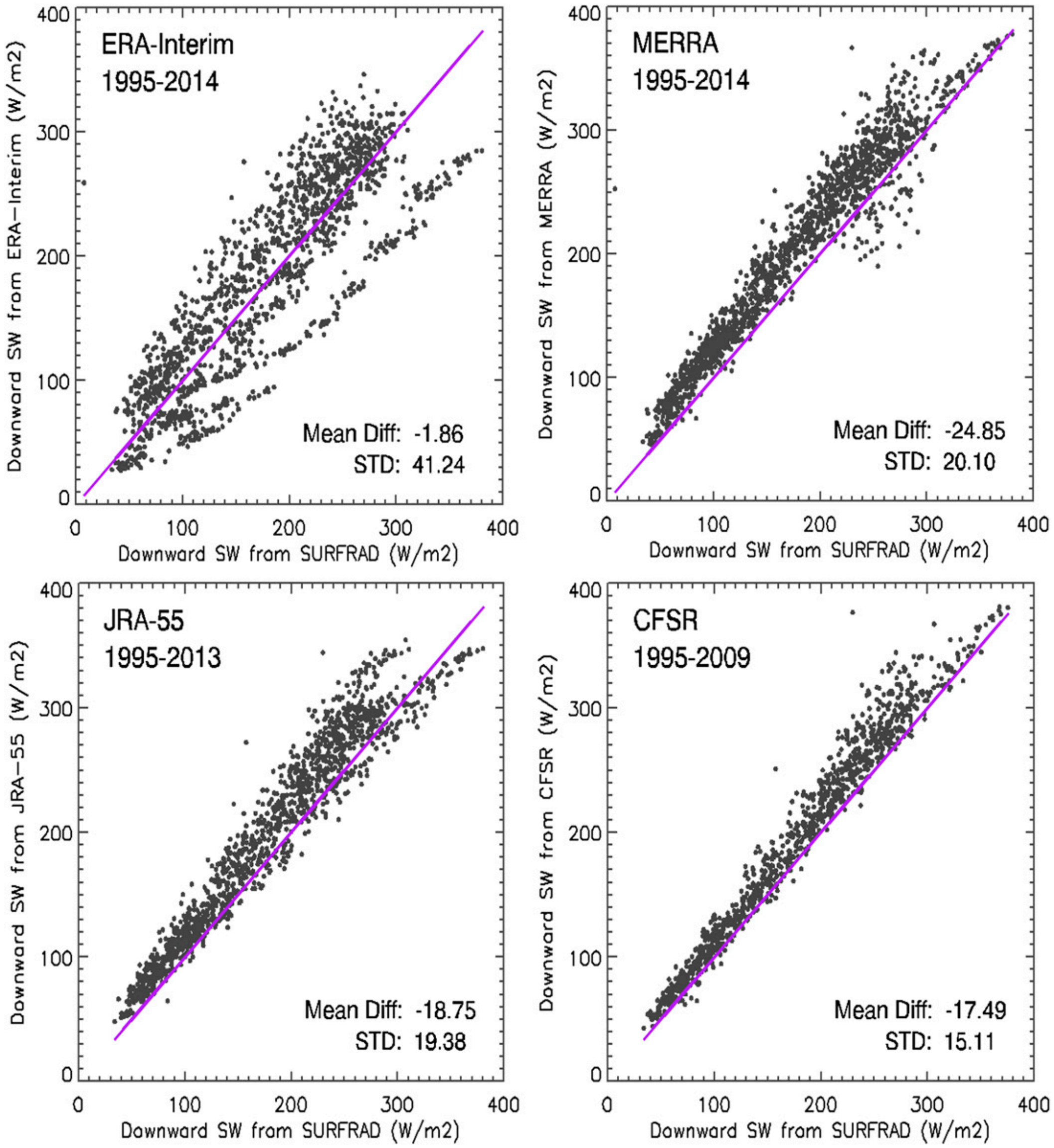


FIG. 3. Scatterplots of monthly mean solar radiation from SURFRAD (x axes) and total cloud cover from reanalyses (y axes) for 7 U.S. locations, with mean difference (site value minus reanalysis value) and standard deviation of the differences (STD).

for all sites) and for the Desert Rock site (mean −0.03 for all reanalyses). The relatively low bias correlations could be due in part to the fact that the SURFRAD stations and the weather stations are not exactly collocated, as well as to the existence of other factors influencing surface solar radiation. The lower correlations for ERA-Interim than for the other reanalyses suggest that other aspects of the radiation calculations may be more important sources of error than cloud cover biases for that reanalysis product. Other studies have found underestimates of surface shortwave radiation for clear-sky conditions in the ECMWF general circulation model used for the IPCC Fourth Assessment Report (Wild 2008). Similar problems could explain the large root-mean-square errors and severe underestimates of surface shortwave radiation at sites such as Desert Rock for ERA-Interim.

5. Conclusions

We compared long-term cloud cover trends and interannual variations from current reanalysis products and surface measurements, with the following main results:

1) Most reanalyses examined here do a reasonable job of simulating the interannual variations in U.S. cloud cover, but all underestimate the mean climatological cloudiness from 9% to almost 20% for the U.S. mean, and more for some individual stations.
2) The bias in total cloud cover is greatest for JRA-55 and smallest for MERRA and CFSR. This bias is reflected in a high bias for surface shortwave radiation at most SURFRAD locations. However, biases in surface solar radiation and those in total cloud cover are not well correlated at some locations, especially for ERA-Interim.
3) Overall, ERA-Interim and CFSR have the best correlations with station data, followed by JRA-55, with MERRA showing the lowest correlation. The U.S. mean total cloud time series from reanalyses accounts for ~65%–81% of the variance in the U.S. mean station data.
4) For the U.S. mean time series, trends in cloud cover are all negative, ranging from −0.38% decade^{-1} for JRA-55 to −1.8% decade^{-1} for MERRA, compared to −0.50% decade^{-1} for the station data. The trends in U.S. means from ERA-Interim and MERRA are significantly different from those in the station data.

Acknowledgments. This work was funded in part by NOAA's Climate Program Office. We thank the reanalysis teams for access to the cloud cover time series and Junye Chen and two anonymous reviewers for helpful comments on the paper.

REFERENCES

Augustine, J. A., G. B. Hodges, C. R. Cornwall, J. J. Michalsky, and C. I. Medina, 2005: An update on SURFRAD—The GCOS surface radiation budget network for the continental United States. *J. Atmos. Oceanic Technol.*, **22**, 1460–1472, doi:10.1175/JTECH1806.1.

Bedacht, E., S. K. Gulev, and A. Macke, 2007: Intercomparison of global cloud cover fields over oceans from the VOS observations and NCEP/NCAR reanalysis. *Int. J. Climatol.*, **27**, 1707–1719, doi:10.1002/joc.1490.

Boilley, A., and L. Wald, 2015: Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface. *Renew. Energy*, **75**, 135–143, doi:10.1016/j.renene.2014.09.042.

Clark, J. V., and J. E. Walsh, 2010: Observed and reanalysis cloud fraction. *J. Geophys. Res.*, **115**, D23121, doi:10.1029/2009JD013235.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Free, M., and B. Sun, 2013: Time-varying biases in U.S. total cloud cover data. *J. Atmos. Oceanic Technol.*, **30**, 2838–2849, doi:10.1175/JTECH-D-13-00026.1.

——, and ——, 2014: Trends in U.S. total cloud cover from a homogeneity-adjusted dataset. *J. Climate*, **10**, 4959–4969, doi:10.1175/JCLI-D-13-00722.1.

Jakob, C., 1999: Cloud cover in the ECMWF reanalysis. *J. Climate*, **12**, 947–959, doi:10.1175/1520-0442(1999)012<0947:CCITER>2.0.CO;2.

Kennedy, A. D., X. Dong, B. Xi, S. Xie, Y. Zhang, and J. Chen, 2011: A comparison of MERRA and NARR reanalyses with the DOE ARM SCP data. *J. Climate*, **24**, 4541–4557, doi:10.1175/2011JCLI3978.1.

Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, doi:10.2151/jmsj.2015-001.

Lorenz, C., and H. Kunstmann, 2012: The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis. *J. Hydrometeor.*, **13**, 1397–1420, doi:10.1175/JHM-D-11-088.1.

Norris, J., and A. T. Evan, 2015: Empirical removal of artifacts from the ISCCP and PATMOS-x satellite cloud records. *J. Atmos. Oceanic Technol.*, **32**, 691–702, doi:10.1175/JTECH-D-14-00058.1.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.

Sun, B., M. Free, H. Yoo, M. J. Foster, A. Heidinger, and K. Karlsson, 2015: Variability and trends in U.S. cloud cover: ISCCP, PATMOS-x, and CLARA-A1 compared to homogeneity-adjusted weather observations. *J. Climate*, **28**, 4373–4389, doi:10.1175/JCLI-D-14-00805.1.

Wild, M., 2008: Short-wave and long-wave surface radiation budget in GCMs: A review based on the IPCC-AR4/CMIP3 models. *Tellus*, **60A**, 932–945, doi:10.1111/j.1600-0870.2008.00342.x.

Wu, W., Y. Liu, and A. K. Betts, 2012: Observationally based evaluation of NWP reanalyses in modeling cloud properties over the Southern Great Plains. *J. Geophys. Res.*, **117**, D12202, doi:10.1029/2011JD016971.