

Trends in U.S. Total Cloud Cover from a Homogeneity-Adjusted Dataset

MELISSA FREE

NOAA/Air Resources Laboratory, College Park, Maryland

BOMIN SUN

NOAA/NESDIS/Center for Satellite Applications and Research, I. M. Systems Group, Rockville, Maryland

(Manuscript received 26 November 2013, in final form 17 March 2014)

ABSTRACT

Cloud cover data from ground-based weather observers can be an important source of climate information, but the record of such observations in the United States is disrupted by the introduction of automated observing systems and other artificial shifts that interfere with our ability to assess changes in cloudiness at climate time scales. A new dataset using 54 National Weather Service (NWS) and 101 military stations that continued to make human-augmented cloud observations after the 1990s has been adjusted using statistical changepoint detection and visual scrutiny. The adjustments substantially reduce the trends in U.S. mean total cloud cover while increasing the agreement between the cloud cover time series and those of physically related climate variables. For 1949–2009, the adjusted time series give a trend in U.S. mean total cloud of $0.11\% \pm 0.22\% \text{ decade}^{-1}$ for the military data, $0.55\% \pm 0.24\% \text{ decade}^{-1}$ for the NWS data, and $0.31\% \pm 0.22\% \text{ decade}^{-1}$ for the combined dataset. These trends are less than one-half of those in the original data. For 1976–2004, the original data give a significant increase but the adjusted data show an insignificant trend from $-0.17\% \text{ decade}^{-1}$ (military stations) to $0.66\% \text{ decade}^{-1}$ (NWS stations). Trends have notable regional variability, with the northwest United States showing declining total cloud cover for all time periods examined, while trends for most other regions are positive. Differences between trends in the adjusted datasets from military stations and NWS stations may be rooted in the difference in data source and reflect the uncertainties in the homogeneity adjustment process.

1. Introduction

Cloud feedbacks are a major source of uncertainty in climate model sensitivity to anthropogenic forcings, and cloudiness is an important factor influencing surface solar radiation and the hydrologic cycle. Despite the critical roles of clouds, our information about past trends in cloudiness is inadequate. Although satellites have provided near-global coverage of cloud amount and characteristics since the early 1980s, long-term satellite cloud datasets are likely to contain inhomogeneities due to factors such as orbital drift and changes in view angle (Evan et al. 2007; Norris 2005).

The records of routine cloud observations by human weather observers extend back many decades before the

beginning of satellite observations and therefore may provide a useful alternative source of information about past cloudiness changes. However, these datasets are also subject to changes in observing and archiving procedures that can make them unsuitable for climate change detection.

The primary widely recognized problem with U.S. cloud observations is the introduction of the Automated Surface Observing System (ASOS) beginning in the early 1990s at most National Weather Service (NWS) stations (Dai et al. 2006). ASOS uses a vertically pointed laser ceilometer to determine sky conditions up to 3.6 km above the instrument, a measurement that is fundamentally different from that made by human cloud observers. The change to ASOS thus introduced a major inhomogeneity to the U.S. cloud cover record at NWS stations that did not continue to make human cloud observations (Warren et al. 1991; Sun 2003; Sun and Groisman 2004; Dai et al. 2006). At a limited number of large airport stations, cloud observations are now

Corresponding author address: Melissa Free, NOAA/Air Resources Laboratory, R/ARL, NCWCP, 5830 University Research Court, College Park, MD 20740.
E-mail: melissa.free@noaa.gov

augmented with human observations that may allow us to extend total cloud cover records beyond the time when ASOS was introduced. Many military weather station observers also continued to make cloud cover observations. The goal of our current work is to use data from these stations to extend the cloud cover record in the United States beyond the mid-1990s.

Like other long-term cloud cover trend studies cited in section 2, this paper is focused on monthly and annual mean total cloud cover rather than on the statistical distribution of daily or hourly sky condition, layer clouds, or cloud types. Section 2 of this paper reviews the previous work using U.S. total cloud observations for trend analysis. Section 3 briefly describes the major changes in the U.S. cloud observing system over time. The NWS and military station datasets used for this analysis are described in section 4. Section 5 documents the methods used to create an improved monthly cloud cover dataset for the United States. Section 6 reports the effects of those methods on time series and trends in cloud cover, and evaluates the quality of the resulting dataset by comparing the results to other physically related climate variables.

2. Results from previous studies

Past assessments of the homogeneity of the U.S. total cloud cover record have been limited. Karl and Steurer (1990) considered data continuity issues for U.S. cloud observations and found no problems after 1950. Angell (1990) compared cloudiness observations to those from sunshine duration recorders for 1950–88, found that the relationship between the two quantities showed inconsistencies, and concluded that the most likely cause was increases in thin cirrus clouds that were visible to cloud observers but were not detected by the sunshine recorders. More recently, Dai et al. (2006) pointed out the disruption in U.S. cloudiness records caused by ASOS and noted that the increase in cloud cover at military stations since 1990 appeared to be inconsistent with the record of diurnal temperature range for that period, but did not attempt to test the homogeneity of the military cloud cover data. Free and Sun (2013) describe homogeneity issues and other data problems that plague the U.S. cloud cover record, especially the change to the aviation routine weather report (METAR) in 1996 discussed in section 3 below.

Trends in total cloud in previous studies have varied with the time period. Several authors have found significant positive trends in U.S. total cloud for periods beginning in or after 1947 and ending before 1996 (e.g., Angell 1990; Plantico et al. 1990; Sun et al. 2001), but Elliott and Angell (1997) found no significant trend in

cloudiness for 1973–93. Sun (2003) indicates increasing trends from about 1950 to 1990 but declining cloud cover from the 1980s to 2000, suggesting that the positive trend might be limited to the period from about 1950 to the 1980s. On the other hand, Dai et al. (2006) used military stations to extend the record of total cloud past 1996 until 2004 and found a statistically significant trend of $\sim 1.4\%$ decade⁻¹ for 1976–2004. Reported trends in low cloud amount for the United States are positive from 1948 to the early 1990s but negative from the 1980s to 2002 (Sun and Groisman 2004). Warren et al. (2007) found no total cloud cover trend in North America for 1971–96, and a slight declining trend for global land areas, using ground-based weather observation data. For 1971–2009, Eastman and Warren (2013) do not report trends for most of North America because of the discontinuity resulting from ASOS, and they find a decline of $\sim 0.7\%$ decade⁻¹ for the remaining land areas between 20° and 50°N.

3. Relevant history of cloud observations in the United States

The most significant change in the U.S. NWS cloud observing system since 1950 was the introduction of ASOS, discussed in the introduction. In addition to the conversion of most NWS stations to the automated ASOS system in the early 1990s, a number of other changes have occurred that may affect our ability to determine trends in cloud cover. Most important for our purposes is the shift to the METAR system of reporting, which occurred on 1 July 1996. Before this date, sky condition was reported as one of four broad categories defined in terms of tenths of sky cover: clear (0), scattered (1–5), broken (6–9), or overcast (10). Under METAR, these were replaced by five categories defined in terms of eighths: clear (0), few (1–2), scattered (3–4), broken (5–7), or overcast (8) (see Free and Sun 2013).

Although military stations continued to report sky condition using human observers after ASOS was introduced at NWS stations, many military stations introduced automated observing systems and discontinued those human observations in the mid-2000s. The military station data are also affected by the shift to METAR in 1996.

4. Data sources

The primary current source for U.S. cloud cover observations is the Integrated Surface Database (ISD) (Lott et al. 2008) at the National Oceanic and Atmospheric Administration's (NOAA's) National Climatic Data Center (NCDC), which contains hourly observations from

various archives. NCDC also archives several other datasets created before the ISD that contain cloud information, including DSI-3280 (NCDC 2005) and DATSAV3 (NCDC 2003) datasets. Cloud information in DSI-3280 has been used in a number of published trend studies and its major characteristics are described in Steurer and Bodosky (2000), Sun et al. (2001), and Free and Sun (2013). The DATSAV3 dataset, created by the U.S. Air Force (USAF) Air Force Weather Agency 14th Weather Squadron, contains surface weather observations from the Global Telecommunications System (GTS) network and other sources, including synoptic, airways, and METAR from both human and automated observations. To our knowledge, DATSAV3 has not been used directly in published cloud trend analysis although this archive was used as one source in the creation of ISD. We examined these three cloud datasets and used data from all three in our total cloud product. Problems with the cloud data from these sources are discussed in more detail in Free and Sun (2013).

Because diurnal temperature range (DTR) and precipitation are physically related to and correlated with cloud cover, we used DTR and precipitation data from several sources for comparisons with total cloud data. For NWS stations, we use DTR computed from monthly-mean adjusted maximum and minimum temperature data in the Historical Climatology Network (HCN) at NCDC (Lawrimore et al. 2011). For military stations, matching HCN station data are not generally available, so we used monthly mean DTR from gridded adjusted HCN data (Lawrimore et al. 2011).

We also used daily precipitation data from NWS cooperative weather stations (COOP) (Groisman et al. 2004) kindly provided by Pasha Groisman of NCDC. To minimize the effects of changes in instrumentation, we computed the number of days with precipitation greater than 0.5 mm day^{-1} for each month and station. If data were not available for our stations, we used precipitation from the same dataset at the closest available station.

5. Dataset production procedures

- 1) We start with hourly data from stations that maintained human cloud observations after the early 1990s. We begin with 1949 because significant changes in reporting occurred in the 1940s that would affect trend analysis (Karl and Steurer 1990) and because fewer data are available for the earlier years. Three NWS stations and 17 military stations were dropped from the analysis because the period of record was too short or the data were otherwise insufficient. The map in Fig. 1 (top) shows the final station set, containing 54 NWS and 101 military stations. The

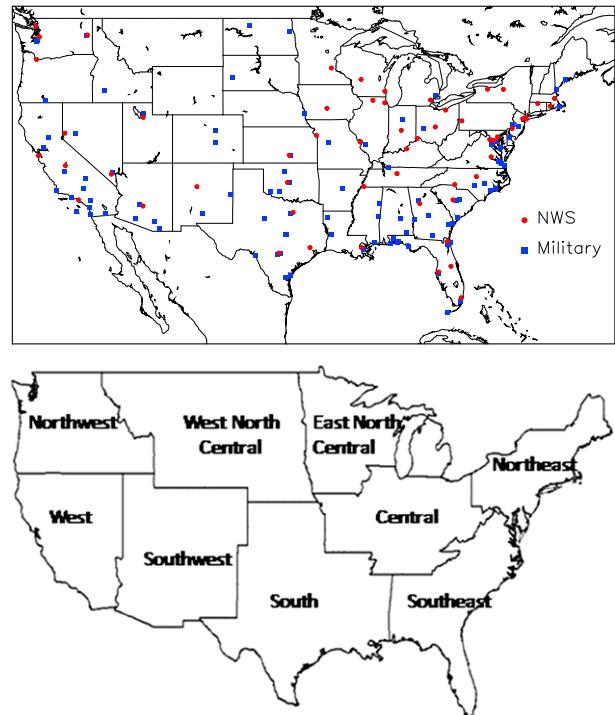


FIG. 1. (top) Locations of stations used in this study. (bottom) Outlines of nine regions used in this study.

station density is relatively high along the west and east coasts but low in the central and inland western regions. After examining the alternative data sources from NCDC described in section 4, we chose to use DSI-3280 and DATSAV3 rather than the current ISD as our primary sources because ISD appears to be less homogeneous (although more complete) than the other archives. The reasons for that choice are described in more detail in Free and Sun (2013).

- 2) NWS station data were taken from DSI-3280, which ends in 2005, and from ISD after 2005. Before ASOS, the data are observations of fractional sky cover reported in tenths. After the introduction of ASOS, cloud information appears in DSI-3280 in the form of descriptive sky condition reports that must be converted to numeric values to compute total cloud in percent. To determine these numeric values, we computed the relative frequency of reported fractional values in the numeric data available between 1973 and the introduction of ASOS at each individual NWS station. For example, since “few” is defined to mean $1/8$ – $2/8$ sky cover, the value assigned to “few” at a given station is a number between $1/8$ and $2/8$ corresponding to the relative frequency of $1/8$ and $2/8$ in the earlier fractional cloud cover reports. Similarly, “scattered” receives a value between $3/8$ and $4/8$, and “broken” is assigned a value between $5/8$ and $7/8$.

The data from ISD used after 2005 were adjusted to match the rest of the dataset by calculating the difference between DSI-3280 and ISD values of the monthly mean cloud cover at each station for January 2000–December 2005, which appears to be systematic, and subtracting that difference from the ISD after 2005.

- 3) While the NWS data before the introduction of ASOS represent fractional sky cover estimates, the military data for all periods are derived from the sky condition reported as one of four or five categories, and those categories changed, as described above, in July 1996. To interpret these records, we therefore used different methods for data before and after the introduction of METAR. Cloud cover observations recorded as 2 in DATSAV3 before METAR were interpreted as scattered. Before 1 July 1996, “scattered” was defined as sky cover between 1/10 and 5/10. We therefore assigned the value 0.3 to those military observations during that time period. Similarly, cloud observations recorded as 7 were assigned the value 0.75, corresponding to “broken,” which is defined before 1996 as 6/10–9/10 sky cover.

In the METAR, observations reported as 2 correspond to “few,” which is defined as 1/8–2/8 sky cover. We therefore assigned the value of 0.19 (1.5/8) to such observations in the DATSAV3 data beginning 1 July 1996. Similarly, we assigned the value 0.44 (3.5/8) to reports coded as 4, or scattered (defined as 3/8–4/8), and 0.75 to reports coded as 6, or broken (5/8–7/8) in METAR data.

- 4) To create monthly means from these hourly observations, we used observations made every 3 h (which are more consistently available than those made every hour). We used only daytime hours (1500, 1800, and 2100 UTC) to avoid difficulties with observing clouds at night (Hahn et al. 1995). We averaged the monthly means for each such hour to get an overall monthly mean. We required a minimum of 30% nonmissing hourly observations at a station to create a monthly mean. [This gives similar data coverage as the requirement of 75 observations per year imposed by Warren et al. (2007) in calculating trends from their cloud data.]
- 5) We examined the monthly mean total cloud data to remove observations that were clearly made with ASOS, without human augmentation. For NWS data we used cloud height information to identify months in which no cloud higher than 3600 m was observed, assumed that the data for those months were not human augmented, and therefore deleted data for those months. If non-ASOS data were available from

the alternative source (ISD) for that month, those data were substituted. This affected primarily years between 1993 and 1996 at ~50 stations. For the military stations, cloud layer information was less complete for the period after automated observing systems were introduced than was the case at NWS stations. We therefore relied primarily on visual examination of the time series to identify and remove physically implausible drops in total cloud (likely due to lack of observations of clouds above the range of the ceilometer) in the mid- and late 2000s at 76 of the 101 military stations.

- 6) The data produced by the steps above constitute version 1. We applied statistical methods to identify change points in frequency time series in the original data and in version 1 total cloud time series. To find change points in the frequency of occurrence of cloud fractional values we applied the standard normal homogeneity test as described in Reeves et al. (2007), but required that the shift be greater than the standard deviation of the frequency time series in order to identify only the largest shifts. We treated any such large and significant change in the frequencies as a metadata event and tested for a shift in the total cloud time series at that time using a Student's *t* test. We also tested all total cloud time series for a shift at 1996 by comparing the means for the five years before and after July 1996 and testing the significance of the difference in means using a Student's *t* test.
- 7) We adjusted the total cloud time series at times when significant shifts were found in both frequency and total cloud, and also at July 1996 if a significant change was found in total cloud at that time. Adjustments are made by adding a constant value to the time series after the time of the change point. We did not adjust the frequency distributions. The results of this step are version 2.
- 8) We visually examined the resulting time series in comparison with precipitation and DTR data, and cloud data from neighboring stations, to accept or reject the adjustments and check for undetected shifts. In a few cases we made adjustments that were not indicated by the statistical tests but were required to make time series from nearby stations visually consistent with each other. The result of this step was version 3.
- 9) We combine station data into nine regions used by NCDC for climate monitoring [shown in Fig. 1 (bottom)] by averaging anomalies from each station and then adding the mean climatological values to give total cloud cover. To create the U.S. mean results, time series of anomalies at the individual stations

were averaged into 2.5° boxes and the grid box means were then averaged with equal weighting. Anomalies are calculated with respect to the mean for 1973–93. Trends are calculated by least squares linear regression and their errors are estimated using a correction for autocorrelation in the time series. For comparison, we also calculated total cloud cover from the ISD using the reported values present in that dataset without any of the conversion procedures described above.

6. Results

a. Effects of the adjustments

A large majority of the statistically based adjustments made at NWS stations appeared to be unwarranted after visual examination. Statistical changepoints were detected at 35 of the 54 stations, but only 10 stations were adjusted in the final versions of the NWS data. Of the 39 statistical changepoints detected in the NWS data, only three were accepted as is and four were accepted with modifications. Five new adjustments were added as a result of the visual assessment. For the military stations, 65 of the stations were adjusted by the statistical methods. Of the total of 82 statistical adjustments at military stations, 30 were accepted and 12 were modified as a result of visual assessment. We added adjustments at 10 military stations based on visual assessment.

Figure 2 (top) shows an example of an adjustment at the NWS station at Albuquerque, New Mexico. Version 1 showed total cloud more than 5% higher than that at a neighboring military station between 2001 and 2007. The statistical adjustment (version 2) gave total cloud for this period that was more consistent with the neighboring station and with the precipitation record. This adjustment was accepted during visual examination. After 2007, we modified the statistically adjusted time series to return to the original to avoid the anomalously low values at the end of the record, giving our version 3 (not shown). Figure 2 (bottom) shows another example, in this case from Shaw Air Force Base, in comparison with adjusted cloud data from the NWS station at Columbia, South Carolina. The statistical adjustment at 1996 produced cloud cover that was much higher than that at the neighboring NWS station and was not consistent with the precipitation data from a nearby COOP station. This adjustment was rejected.

For the NWS stations, the largest effect of our procedures was due to the use of DSI-3280 rather than ISD, the removal of ASOS-contaminated data, and the conversion from descriptive sky condition data to numeric

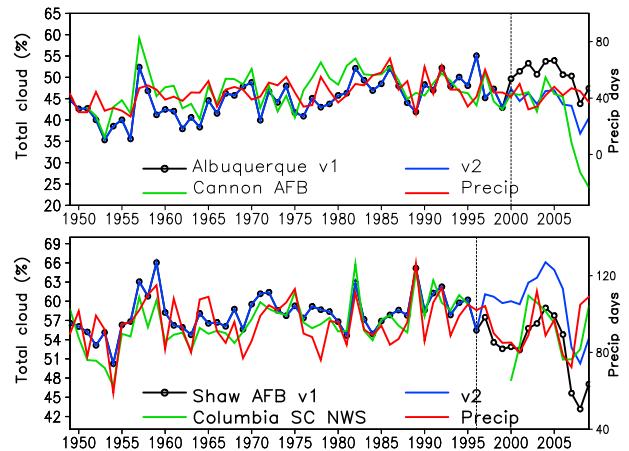


FIG. 2. (top) Annual mean of total cloud at the NWS station at Albuquerque, NM [World Meteorological Organization (WMO) station ID 723650], in versions 1 and 2 compared with that from the neighboring military station at Cannon Air Force Base (AFB) (WMO 722686), and annual mean of days per month with precipitation at Albuquerque. The statistical adjustment in version 2 at 2000 was accepted. The decline at the military station after 2005 is due to the advent of ASOS at that station. (bottom) Annual mean of total cloud at military station Shaw AFB (WMO 747900) in versions 1 and 2 compared with that from a nearby NWS station at Columbia, SC (WMO 138835; 60 km away), and with days per year with precipitation at a nearby COOP station. The statistical adjustment at 1996 was rejected.

data after the introduction of ASOS. Figures 3a and 3b show the annual mean U.S. mean time series of total cloud from NWS stations for the original ISD and three versions of the adjusted product. The low value of total cloud in 1996 in ISD is likely related to contamination of the data with ASOS reports at that time. The primary result of using DSI-3280 data in the U.S. mean is to reduce total cloud cover by $\sim 2\%$ – 3% around 2000 as compared to the results from ISD. In the large-scale mean, effects of the statistical adjustments and visual scrutiny are small in comparison to the effects of using the alternative dataset, although the effect can be large for some individual stations.

Similarly, the data modifications (converting recorded values to be more consistent before and after METAR) described in section 5c have a large effect on the time series after 1996 for military stations (Figs. 3c,d). Before these modifications, the mean total cloud amount for military stations for 1991–95 in DATSAV3 was lower than the mean for 1997–2001 by 1.9%. After the modification, the earlier period was higher than the later period by 2.4%, a change of over 4% total cloud cover. The statistical adjustments and the visual screening of those adjustments have smaller effects on this 1996 shift, tending to reduce the effect of the original modification to the military data, but U.S. mean cloud cover after

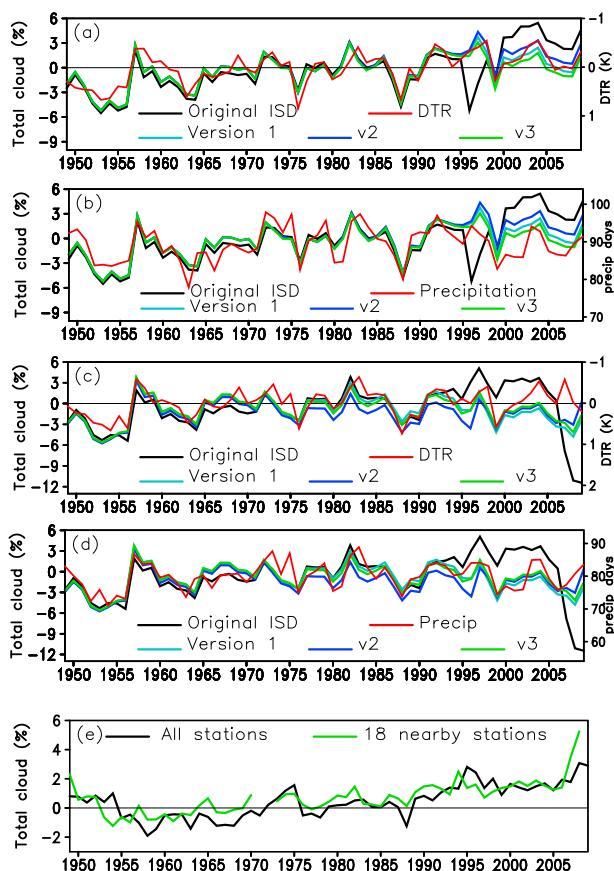


FIG. 3. (a) Annual means of gridded anomalies of U.S. mean total cloud cover from NWS stations in ISD and three versions of the adjusted product, with gridded collocated DTR. Data that were contaminated by ASOS measurements were removed from the adjusted versions but not from ISD. (b) As in (a), but for gridded collocated days per year with precipitation. (c) As in (a), but for data from military stations. (d) As in (c), but for days per year with precipitation. (e) Annual mean anomaly of U.S. mean total cloud from gridded NWS data minus that from gridded military station data using final adjusted dataset (version 3) for all stations and for 18 stations that are within 50 km of each other. For the 18 collocated stations, only months of data present at both the military and NWS stations are used to create the mean.

1996 in the final version of the military data is still less than before that time.

For the NWS data, the time series of U.S. average gridded total cloud with and without the questionable version 2 adjustments that were dropped in version 3 are very similar before 1998, and differ by $\sim 2\%$ cloud cover after that time (Figs. 3a,b). The final version has lower total cloud after 1995 than versions 1 or 2. For the military stations, the statistical adjustments made a more noticeable difference in the mean time series before 1996 (Figs. 3c,d). At these stations, the effect of the statistical adjustments was to lower total cloud between 1976 and 1996. The final version, reflecting

visual scrutiny of the statistical adjustments, gave mean results between the original and statistically adjusted results for most months before 1996. After 1996, the final version 3 shows greater total cloud cover than versions 1 or 2. However, results after 2005 must be viewed with caution because data from many military stations were deleted for that time period because the stations had shifted to ASOS, reducing the number of stations from 101 in 2001, to 86 at the end of 2005, and to 24 in 2009.

b. Trends

Least squares linear trends in total cloud for the United States, calculated from the mean of gridded station data, are lower for the adjusted data than for the original data (Table 1) for the full period (1949–2009), the pre-METAR period 1949–94 used in many past cloud studies, the period 1976–2004 used in Dai et al. (2006), and the satellite period 1984–2008 (e.g., Stubenrauch et al. 2013). For the full time period 1949–2009, the original ISD gave a trend of $1.10\% \pm 0.26\%$ decade⁻¹, but the final adjusted data has a trend of $0.31\% \pm 0.22\%$ decade⁻¹. (For purposes of these trend calculations, we excluded post-2004 ASOS data from the military ISD time series.) Trends in the U.S. mean for 1949–2009 are largest in summer and fall and are significant for summer and the annual mean. However, trends in the adjusted data for some more recent time periods are smaller or even negative (and not significant). For the period 1976–2004 used in Dai et al. (2006), the original ISD had a statistically significant trend of $1.47\% \pm 0.72\%$ decade⁻¹, but the trend in the adjusted data is only $0.05\% \pm 0.62\%$ decade⁻¹. Trends less than zero (but not significant) occur in spring and fall for 1976–2004. For the satellite time period 1984–2008 the trend in the annual mean is $-0.42\% \pm 0.72\%$ decade⁻¹.

The military and NWS station sets give significantly different trends for U.S. means for most time periods, with the adjusted data from military stations showing smaller trends than the NWS stations for all time periods we examined (Table 2). The difference between the U.S. time series of the two subsets of adjusted data shows NWS total cloud increasing relative to the military station data around 1970 and again in the 1990s (Fig. 3e). To see if the trend differences could be caused by differences in spatial sampling, we compared the mean total cloud for a subset of 18 locations where the NWS and military stations were less than 50 km apart. (Although individual clouds can have spatial scales well below 50 km, monthly or annual mean total cloud cover time series for the 18 station pairs we used were highly correlated, with r values from 0.80 to 0.96.) Figure 3e shows the time series of the difference between annual mean total cloud from these 18 military and NWS

TABLE 1. Least squares linear trends in total cloud (% decade⁻¹) and confidence intervals for the original ISD combining NWS and military data, for the adjusted versions of the combined station set, and for the combined final dataset excluding military stations that converted to ASOS in the last decade. Trends statistically significant at the 0.05 level are in boldface. Confidence intervals are based on twice the standard error of the trends. Original ISD trends are for time series with ASOS data after 2004 deleted from military cloud time series. For “non-ASOS version 3” trends, military stations that have ASOS data in the 2000s are excluded completely.

	1949–2009	1949–94	1976–2004	1984–2008
ISD original all stations	1.10 ± 0.26	0.92 ± 0.34	1.47 ± 0.72	1.90 ± 0.86
All stations adjusted version 1	0.31 ± 0.22	0.78 ± 0.36	−0.14 ± 0.62	−0.66 ± 0.72
All stations adjusted version 2	0.44 ± 0.22	0.53 ± 0.34	0.52 ± 0.60	0.49 ± 0.70
All stations adjusted version 3	0.31 ± 0.22	0.69 ± 0.34	0.05 ± 0.62	−0.42 ± 0.72
Non-ASOS version 3	0.43 ± 0.22	0.75 ± 0.34	0.35 ± 0.62	−0.18 ± 0.74

stations. The NWS cloud still shows a similar relative increase over cloud cover at military stations as seen in the full datasets. The trends for the 18-station subsets were similar to those for the complete station sets, and trends for the NWS stations were still larger than those for the military stations by $\sim 0.4\%$ decade⁻¹ or more, suggesting that the differences in the U.S. mean trends are not primarily the result of different spatial sampling.

Because the majority of military stations are excluded after about 2005 because of the loss of human cloud observations, the spatial sampling changes at the end of the record, and this shift in station set could affect our results. If we exclude the military stations that shifted to ASOS in the 2000s, the smaller subset of military stations has larger or less negative trends (see bottom row in Table 2) for most time periods as compared to the full military dataset. Combining this smaller military station set with the NWS stations gives a combined station set that has larger increases in total cloud than the full station set (bottom row in Table 1), with a trend of 0.35% decade⁻¹ for 1976–2004. These trends are, however, still much lower than those in the original ISD or the trend of 1.4% decade⁻¹ found in Dai et al. (2006) for the military station data for 1976–2004.

Trends show noticeable variation among regions (Table 3 and Fig. 4). In the combined dataset, the

TABLE 2. Least squares linear trends in total cloud (% decade⁻¹) for original ISD and final adjusted NWS and military station subsets, and for only the 25 military stations that did not convert to ASOS. Trends statistically significant at the 0.05 level are in boldface. For “non-ASOS” trends, military stations that have ASOS data in the 2000s are excluded completely.

	1949–2009	1949–94	1976–2004	1984–2008
NWS original ISD	1.00	0.86	1.26	2.22
NWS version 3	0.55	0.78	0.66	0.12
Military original ISD	1.07	0.90	1.52	1.61
Military version 3	0.11	0.57	−0.17	− 0.77
Military version 3 non-ASOS only	0.37	0.93	−0.24	− 0.69

Northwest differs from the other regions, with a negative but insignificant trend for 1949–2009 and a significantly negative trend for 1976–2004. The increase in cloud cover in the West region is larger than in the other regions for the full time period. Most regions have negative trends for 1984–2008 and insignificantly positive trends for 1976–2004 in the combined data. Like the U.S. mean trends, the regional trends differ between the military and NWS stations, with the largest trend differences for 1949–2009 seen in the South (more than 1% decade⁻¹) and Southeast (0.65% decade⁻¹). Most regions have positive trends for this period in the NWS data, but for the East North Central and the Southeast regions, the signs of the trends are opposite for the NWS and military station sets. In the NWS data, trends are largest in the South, while they are negative in the Northwest for both station sets.

For some regions, the differences in total cloud cover between the two networks shown in Fig. 4 are caused primarily by their difference in spatial sampling. For example, NWS stations in the Northwest are all located in the northern part of the region [see Fig. 1 (top)],

TABLE 3. Least squares linear trends in total cloud for nine U.S. regions (% decade⁻¹) from NWS, military, and all stations combined, using adjusted data (version 3). Trends statistically significant at the 0.05 level are in boldface. There are no NWS stations in our station set in the West North Central region.

Region	1949–2009			1976–2004		
	NWS	Military	Both	NWS	Military	Both
Central	0.43	0.14	0.37	0.56	0.47	0.53
East North Central	0.47	−0.07	0.31	0.54	−0.38	0.28
Northeast	0.49	0.20	0.39	0.93	0.13	0.59
Northwest	− 0.48	−0.23	−0.33	− 1.60	− 1.16	− 1.28
South	1.15	0.01	0.38	1.52	−0.32	0.11
Southeast	0.45	−0.20	0.04	0.88	0.04	0.26
Southwest	0.56	0.16	0.28	−0.22	− 0.89	−0.76
West	1.08	0.66	0.70	1.08	0.18	0.33
West North Central	—	0.42	0.42	—	−0.07	−0.07

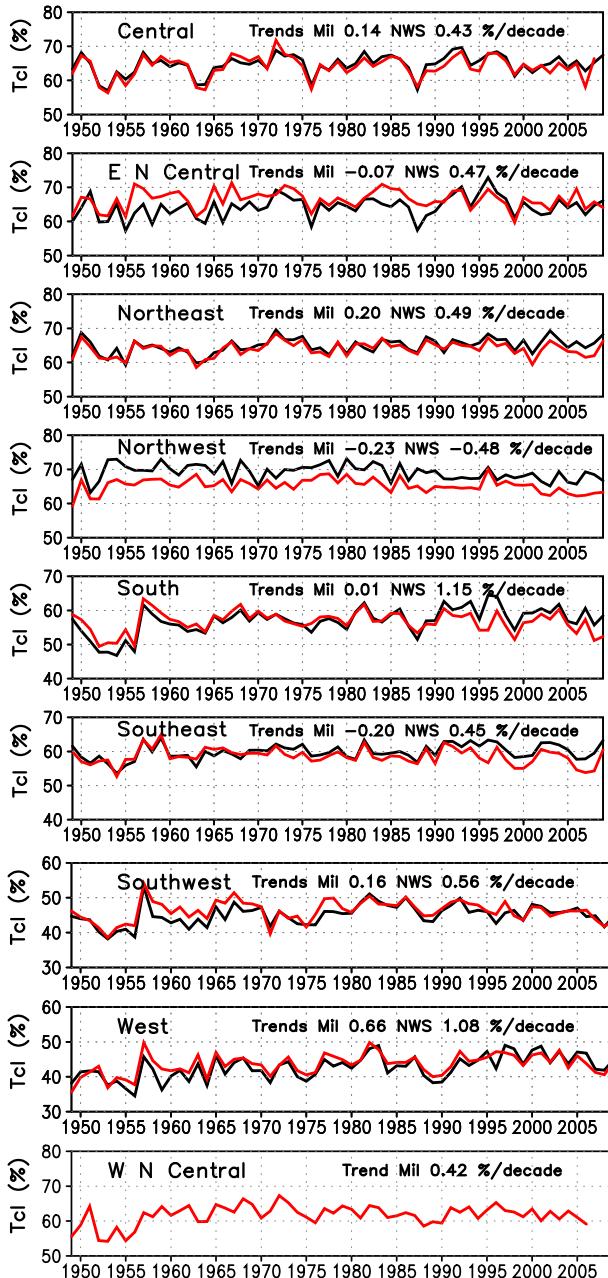


FIG. 4. Annual means of total cloud from nine regions indicated in Fig. 1 (bottom), using version 3 adjusted data from NWS (black) and military (red) stations. The y-axis coordinates differ to accommodate the mean cloud in different regions. There are no NWS stations in our station set in the West North Central region.

where climatological cloud cover is much higher than in the southern part (see <http://www.atmos.washington.edu/CloudMap/>), where some of the military stations are located. This leads to the systematic higher total cloud in the NWS regional average compared to the corresponding military average.

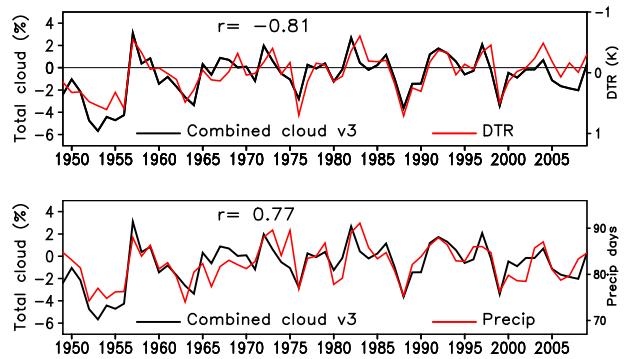


FIG. 5. Gridded combined U.S. military and NWS annual mean anomalies of total cloud from final adjusted dataset (version 3) with gridded (top) DTR and (bottom) precipitation data.

c. Comparisons with other physically related variables

Because precipitation and DTR are generally well correlated with cloud cover, we test our adjusted dataset by comparing it to observations of those variables. Since time series of precipitation and DTR are physically related to clouds but not identical to them, and because the precipitation and DTR data may themselves contain inhomogeneities, these comparisons cannot provide a definitive basis on which to prefer one subset or version of cloud data to the other. That being said, the correlation of either of these variables to cloud cover can be used to detect discontinuity problems in our cloud cover time series. Although DTR and precipitation are expected to be related more closely to cloud cover from all hours (day and night) than to daytime cloud cover alone, we found that the relation to our daytime cloud data was in most cases adequate to make the comparison useful. The correlations might be improved if we used cloud data from all hours, but at the price of potential error from erroneous nighttime cloud reports due to poor illumination.

Figure 5 compares the U.S. mean cloud anomaly time series using gridded data from all stations with the corresponding means for DTR and days with precipitation. Overall, the combined NWS and military station cloud cover matches closely with the mean number of days with precipitation at collocated stations, but appears to show somewhat less cloud cover after about 2004 than would be expected from comparison to DTR data. The correlation coefficient between the combined annual adjusted cloud cover and number of days with precipitation is 0.54 for the original ISD and 0.77 for the final adjusted version. Correlations between total cloud cover and precipitation are largest in fall and smallest in summer, but in all seasons they are better for the final than for the original data. For DTR, correlations with

the combined dataset are -0.73 for the original ISD and -0.81 for the final adjusted version. Correlations are slightly higher for the period before 1994.

Figures 3a–d compare the U.S. mean cloud anomaly time series using the military and NWS station sets with the corresponding means for precipitation and DTR. The decline in the time series of adjusted military cloud after 1995 is present to a lesser extent in the precipitation time series (Fig. 3d), while even after our adjustments, the NWS cloud data still appear to be biased high in comparison to precipitation days in that period (Fig. 3b). This suggests that, to be consistent with precipitation data, the true total cloud after 1996 should fall between our adjusted NWS and military results. On the other hand, the NWS-adjusted data agree reasonably well visually with the collocated DTR data (Fig. 3a) while the military cloud appears much too low after about 2000 in comparison with DTR (Fig. 3c), suggesting that the NWS-adjusted data may be more reliable. The low bias of military cloud cover at the end of the record is more apparent in spring and summer than in fall and winter (plots not shown).

The correlation between the U.S. annual mean total cloud for NWS stations and the mean number of days with precipitation for collocated stations was 0.44 for the original ISD and 0.71 for version 3 of the adjusted data. The corresponding correlations for the military stations were 0.53 in the original data and 0.68 for the final adjusted data. Our final data show better correlations with precipitation than do the original data for all four seasons. For DTR, correlations went from -0.67 in the original NWS data to -0.83 in the final version, and from -0.65 in the original military data to -0.70 in the final. The improvement in correlations with DTR after adjustment occurs in all seasons for NWS data, but only for fall and winter for the military data.

7. Discussion

a. Differences between results from military and NWS stations

The differences between the adjusted military and NWS results are most likely rooted in the differences in the data sources and the methods we used to convert the source data to total cloud cover. They reflect the difficulty of compensating for the inhomogeneities in the data.

To explore whether our procedure introduces a discontinuity in the DATSAV3 cloud cover time series from the transition of four sky condition categories prior to METAR to five categories after METAR, we used the NWS DSI-3280 hourly fractional cloud cover data from 1973 to the start of ASOS to simulate those two sets of

DATSAV3 sky condition categories and then applied the same procedure as described in section 5 to compute cloud cover values for the two versions. The test revealed that mean cloud cover with the pre-METAR version was 0.4% higher than with the post-METAR version, suggesting that the adjusted military data may have a shift of -0.4% at 1996. If a correction of equal size is applied to the DATSAV3 data at 1996, the difference in trend between NWS and DATSAV3 is reduced. For example, for the 18 collocated stations, the trend difference for 1984–2004 is reduced to 0.43% decade⁻¹ from 0.71% decade⁻¹.

This test assumes that the cloud cover definition is the same for the two time periods of interest. This assumption is obviously not true for the NWS data as demonstrated by the abrupt changes in clear-sky and overcast frequencies in 1996 [see Fig. 4 of Free and Sun (2013)]. We believe the method we used to connect the NWS data before and after 1996 (see section 5) is reliable but we cannot be certain that our method does not introduce a discontinuity around 1996 in the NWS dataset.

The shift in DATSAV3 cloud cover in 1996 discussed in the second paragraph of this section, if real, accounts for approximately 40% of the NWS–DATSAV3 trend difference for the data after the early 1980s. The other factors that contribute to the remaining difference between the two datasets around the METAR introduction in 1996 are unknown to us.

In the part of the record before METAR, the time series of differences between cloud cover from NWS and military data at 18 collocated stations shows that the military cloud cover is higher than the NWS cloud (a pattern also seen in the time series from the full station set) (Fig. 3e). As revealed in Free and Sun (2013), military data in the 1970s and 1980s contain more obvious discontinuities in the time series of the frequency of occurrence of fractional cloud cover (from which total cloud cover is computed) than seen in the NWS data, which suggests that the military cloud cover time series may be less reliable than the NWS data for that period.

The visual comparison with DTR and precipitation in Fig. 3 suggests that the true cloud cover time series may lie between the military and NWS results. However, because factors other than cloud cover will affect precipitation and DTR (Lauritsen and Rogers 2012), the relationship between DTR or precipitation and cloud cover can vary in space and time, so we must use these comparisons with caution.

b. Other uncertainties

Sources of uncertainty in our final dataset include the possibility of inhomogeneities that have not been

adjusted, errors that may be introduced in the adjustment process, and spatial sampling issues. Using the difference between trends in the adjusted versions as a measure of the uncertainty in trend associated with the adjustments, we find uncertainties are only $\sim 0.1\%$ or less of total cloud per decade for the longer time periods in the combined dataset, but can be of the same order of magnitude as the trends themselves for the more recent (and shorter) time periods. However, the discrepancy between results for the NWS and military datasets suggests larger uncertainties, on the order of $0.4\% \text{ decade}^{-1}$ for the U.S. mean, for even the longest time period. The possible remaining discontinuity in the military data discussed in [section 7a](#) above implies trend uncertainties of up to 0.3% for the period 1984–2008, and lesser uncertainty for longer time periods.

The statistical changepoint adjustment procedure is designed to correct only the largest shifts, to avoid removing real variability. Artificial shifts below the threshold will not be corrected. Similarly, in our subjective evaluation of the cloud cover time series we tried to remove only shifts that seemed clearly artificial. Small shifts are therefore likely to remain unadjusted. The shifts detected by our statistical adjustment process range from 2.5% to 13% of total cloud cover, substantially larger than the 0.4% estimated post-METAR bias in the military data. Thus, a bias of that size associated with the METAR transition would probably not be corrected by our statistical adjustment procedures. As an example of the possible effects of undetected shifts, an uncorrected shift of 2% could produce an artificial trend of up to $0.5\% \text{ decade}^{-1}$ for 1949–2009, and more for shorter time periods.

Because the changes at 1996 occurred at all stations, use of neighbor stations in the adjustment process can be of only limited help around that time. Although DTR and precipitation are useful for visual comparison, the correlations between these quantities and total cloud at individual stations are not generally high enough to allow their use in a statistical changepoint adjustment scheme. Given these problems, it is reasonable to assume that additional inhomogeneity may exist in the total cloud data for the United States from these human-augmented stations, especially in the 1990s when METAR was introduced. On the other hand, statistical homogeneity adjustments can sometimes reduce or eliminate real trends along with artificial shifts ([Gaffen et al. 2000](#)), so it is possible that our results understate the actual trends in total cloud.

Spatial sampling issues may also be significant sources of error in the U.S. mean trends. The NWS station trends for 1949–2009 vary widely even after adjustment, with a spatial standard deviation of $0.48\% \text{ decade}^{-1}$,

which would give a confidence interval for the mean of the trends of $0.96\% \text{ decade}^{-1}$. The equivalent confidence interval for the military stations is larger, at $1.48\% \text{ decade}^{-1}$. This spread is likely to be a combination of actual spatial variability and uncorrected errors due to inhomogeneities and other data quality issues.

c. Comparison with results from earlier studies

Despite the differences in the data sources and uncertainties in the adjustment process, the adjusted datasets agree more closely with the precipitation and DTR data than did the original cloud data. Both sets indicate substantially smaller trends in U.S. total cloud than seen in other recent studies. Our results for the time period 1976–2004 show much less increase in total cloud than those of [Dai et al. \(2006\)](#). The difference between our results and those of Dai et al. is unlikely to be related to our use of NWS stations in addition to the military stations. The trend in the U.S. mean for our adjusted data for military stations alone is $-0.17\% \text{ decade}^{-1}$ for 1976–2004, compared to 0.66% for the NWS stations (neither of which is statistically significant). Thus, our trend results would be even less consistent with those of Dai et al. if we used only military stations. The primary source of the difference is likely to be our homogeneity adjustments. Dai et al. did not attempt to assess the homogeneity of their data.

Our results are more consistent with those of [Elliott and Angell \(1997\)](#), who found a trend of only $0.2\% \text{ decade}^{-1}$ for 1973–93, and the trend of 0.55% for 1948–99 reported by [Sun \(2003\)](#), but appear inconsistent with those of [Angell \(1990\)](#), who found increases of 2% between 1950–68 and 1970–88, and those of [Plantico et al. \(1990\)](#).

8. Conclusions

Homogeneity adjustments to the U.S. cloud record from human weather observers reduce the trends in U.S. mean total cloud for 1949–2009 from $1.1\% \pm 0.26\% \text{ decade}^{-1}$ to $0.3\% \pm 0.22\% \text{ decade}^{-1}$, and for 1976–2004 from $1.47\% \pm 0.72\% \text{ decade}^{-1}$ to $0.05\% \pm 0.62\% \text{ decade}^{-1}$. The resulting trends are inconsistent with some previous work including the trend of $\sim 1.4\% \text{ decade}^{-1}$ for 1976–2004 in [Dai et al. \(2006\)](#). Most regions still show positive trends, but the majority are not statistically significant. In the Northwest, total cloud cover has declined significantly.

The results for trends at military stations differ significantly from those for NWS stations, and this difference does not appear to be caused by differences in spatial sampling. The discrepancy illustrates the uncertainty in trends of U.S. cloud cover. Comparisons with precipitation and diurnal temperature range data

suggest that the actual cloud cover history for the past 20 years may lie somewhere between the results for the two station sets. Our adjustments improve the agreement between time series of U.S. total cloud data and those of DTR and precipitation, but it is likely that inhomogeneities remain in the adjusted datasets.

Acknowledgments. This work was funded in part by the Climate Monitoring Program of NOAA's Climate Program Office. We thank William Brown at NCDC for providing DSI-3280 and DATSAV3 data and Pasha Groisman at NCDC for providing daily precipitation data, and Dian Seidel, Julian Wang, and anonymous reviewers for helpful comments on the manuscript.

REFERENCES

- Angell, J. K., 1990: Variation in United States cloudiness and sunshine duration between 1950 and the drought year of 1988. *J. Climate*, **3**, 296–308, doi:10.1175/1520-0442(1990)003<0296:VIUSCA>2.0.CO;2.
- Dai, A., T. R. Karl, B. Sun, and K. E. Trenberth, 2006: Recent trends in cloudiness over the United States: A tale of monitoring inadequacies. *Bull. Amer. Meteor. Soc.*, **87**, 597–606, doi:10.1175/BAMS-87-5-597.
- Eastman, R., and S. G. Warren, 2013: A 39-year survey of cloud changes from land stations worldwide 1971–2009: Long-term trends, relation to aerosols, and expansion of the tropical belt. *J. Climate*, **26**, 1286–1303, doi:10.1175/JCLI-D-12-00280.1.
- Elliott, W. P., and J. K. Angell, 1997: Variations of cloudiness, precipitable water, and relative humidity over the United States: 1973–1993. *Geophys. Res. Lett.*, **24**, 41–44, doi:10.1029/96GL03616.
- Evan, A. T., A. K. Heidinger, and D. J. Vimont, 2007: Arguments against a physical long-term trend in global ISCCP cloud amounts. *Geophys. Res. Lett.*, **34**, L04701, doi:10.1029/2006GL028083.
- 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.
- Gaffen, D. J., M. A. Sargent, R. E. Habermann, and J. R. Lanzante, 2000: Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *J. Climate*, **13**, 1776–1796, doi:10.1175/1520-0442(2000)013<1776:SOTAST>2.0.CO;2.
- Groisman, P. Ya., R. W. Knight, T. R. Karl, D. R. Easterling, B. Sun, and J. M. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *J. Hydrometeorol.*, **5**, 64–85, doi:10.1175/1525-7541(2004)005<0064:CCOTHG>2.0.CO;2.
- Hahn, C. J., S. G. Warren, and J. London, 1995: The effect of moonlight on observation of cloud cover at night, and application to cloud climatology. *J. Climate*, **8**, 1429–1446, doi:10.1175/1520-0442(1995)008<1429:TEOMOO>2.0.CO;2.
- Karl, T. R., and P. M. Steurer, 1990: Increased cloudiness in the United States during the first half of the twentieth century: Fact or fiction? *Geophys. Res. Lett.*, **17**, 1925–1928, doi:10.1029/GL017i011p01925.
- Lauritsen, R. G., and J. C. Rogers, 2012: U.S. diurnal temperature range variability and regional causal mechanism, 1901–2002. *J. Climate*, **25**, 7216–7231, doi:10.1175/JCLI-D-11-00429.1.
- Lawrimore, J. H., M. J. Menne, B. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, and J. Rennie, 2011: An overview of the Global Historical Climatology Network monthly mean temperature dataset, version 3. *J. Geophys. Res.*, **116**, D19121, doi:10.1029/2011JD016187.
- Lott, N., R. S. Vose, S. A. Del Greco, T. F. Ross, S. Worley, and J. L. Comeaux, 2008: The Integrated Surface Database: Partnerships and progress. *Extended Abstracts, 24th Conf. on Interactive Information and Processing Systems*, New Orleans, LA, Amer. Meteor. Soc., 3B.5. [Available online at https://ams.confex.com/ams/88Annual/techprogram/paper_131387.htm.]
- NCDC, 2003: Data documentation for data set 9956 (DSI-9956): DATSAV3 global surface hourly data. National Climatic Data Center, 52 pp. [Available online at <ftp://ftp.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9956.pdf>.]
- , 2005: Data documentation for data set 3280 (DSI-3280): Surface airways hourly. National Climatic Data Center, 30 pp. [Available online at <ftp://ftp.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3280.pdf>.]
- Norris, J. R., 2005: Multidecadal changes in near-global cloud cover and estimated cloud cover radiative forcing. *J. Geophys. Res.*, **110**, D08206, doi:10.1029/2004JD005600.
- Plantico, M. S., T. R. Karl, G. Kukla, and J. Gavin, 1990: Is recent climate change across the United States related to rising levels of anthropogenic greenhouse gases? *J. Geophys. Res.*, **95**, 16 617–16 637, doi:10.1029/JD095iD10p16617.
- Reeves, J., J. Chen, X. L. Wang, R. Lund, and Q. Lu, 2007: A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteor. Climatol.*, **46**, 900–915, doi:10.1175/JAM2493.1.
- Steurer, P. M., and M. Bodosky, 2000: Surface airways hourly (TD-3280) and airways solar radiation (TD-3281). National Climatic Data Center, 50 pp.
- Stubenrauch, C. J., and Coauthors, 2013: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel. *Bull. Amer. Meteor. Soc.*, **94**, 1031–1049, doi:10.1175/BAMS-D-12-00117.1.
- Sun, B., 2003: Cloudiness over the contiguous United States: Contemporary changes observed using ground-based and ISCCP D2 data. *Geophys. Res. Lett.*, **30**, 1053, doi:10.1029/2002GL015887.
- , and P. Ya. Groisman, 2004: Variations in low cloud cover over the United States during the second half of the twentieth century. *J. Climate*, **17**, 1883–1888, doi:10.1175/1520-0442(2004)017<1883:VILCCO>2.0.CO;2.
- , —, and I. I. Mokhov, 2001: Recent changes in cloud-type frequency and inferred increases in convection over the United States and the former USSR. *J. Climate*, **14**, 1864–1880, doi:10.1175/1520-0442(2001)014<1864:RCICTF>2.0.CO;2.
- Warren, S. G., J. London, and C. J. Hahn, 1991: Cloud hole over the United States? *Bull. Amer. Meteor. Soc.*, **72**, 237–238.
- , R. Eastman, and C. J. Hahn, 2007: A survey of changes in cloud cover and cloud types over land from surface observations, 1971–96. *J. Climate*, **20**, 717–738, doi:10.1175/JCLI4031.1.