Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of Stage IV toward CPC Gauge-Based Analysis

DINGCHEN HOU,* MIKE CHARLES,*,+,## YAN LUO,*,# ZOLTAN TOTH,@ YUEJIAN ZHU,*
ROMAN KRZYSZTOFOWICZ,& YING LIN,* PINGPING XIE,** DONG-JUN SEO,++,@@
MALAQUIAS PENA,*,# AND BO CUI*,#

* *Environmental Modeling Center/NCEP/NWS/NOAA, College Park, Maryland*
+ *Science Application International Corporation, Camp Springs, Maryland*
# *I. M. Systems Group, Inc., Camp Springs, Maryland*
@ *Global System Division/ESRL/OAR/NOAA, Boulder, Colorado*
& *University of Virginia, Charlottesville, Virginia*
** *Climate Prediction Center/NCEP/NWS/NOAA, College Park, Maryland*
++ *Office of Hydrologic Development/NWS/NOAA, Silver Spring, Maryland*

(Manuscript received 28 October 2011, in final form 31 August 2012)

ABSTRACT

Two widely used precipitation analyses are the Climate Prediction Center (CPC) unified global daily gauge analysis and Stage IV analysis based on quantitative precipitation estimate with multisensor observations. The former is based on gauge records with a uniform quality control across the entire domain and thus bears more confidence, but provides only 24-h accumulation at $\frac{1}{8}°$ resolution. The Stage IV dataset, on the other hand, has higher spatial and temporal resolution, but is subject to different methods of quality control and adjustments by different River Forecasting Centers. This article describes a methodology used to generate a new dataset by adjusting the Stage IV 6-h accumulations based on available joint samples of the two analyses to take advantage of both datasets. A simple linear regression model is applied to the archived historical Stage IV and the CPC datasets after the former is aggregated to the CPC grid and daily accumulation. The aggregated Stage IV analysis is then adjusted based on this linear model and then downscaled back to its original resolution. The new dataset, named Climatology-Calibrated Precipitation Analysis (CCPA), retains the spatial and temporal patterns of the Stage IV analysis while having its long-term average and climate probability distribution closer to that of the CPC analysis. The limitation of the methodology at some locations is mainly associated with heavy to extreme precipitation events, which the Stage IV dataset tends to underestimate. CCPA cannot effectively correct this because of the linear regression model and the relative scarcity of heavy precipitation in the training data sample.

1. Introduction

At the National Centers for Environmental Prediction (NCEP), some postprocessing procedures, namely bias correction and downscaling, are applied to numerical weather prediction (NWP) products such as temperature

Current affiliation: Climate Prediction Center/NCEP/NWS/NOAA, College Park, Maryland.
@@ Current affiliation: Department of Civil Engineering, University of Texas at Arlington, Arlington, Texas.

Corresponding author address: Dr. Dingchen Hou, Environmental Modeling Center/NCEP/NWS/NOAA, 5830 University Research Court, College Park, MD 20740.
E-mail: dingchen.hou@noaa.gov

and wind from the Global Ensemble Forecasting System (GEFS) and the North America Ensemble Forecasting System (NAEFS). These techniques have demonstrated significant benefits in improving local forecasts over the contiguous United States (CONUS) domain (Cui et al. 2012; B. Cui et al. 2012, unpublished manuscript). The application of the same procedures to precipitation is hindered by the lack of a satisfying precipitation dataset. The required dataset should be our best estimate of 6-h accumulation, on an approximately 5×5 km^2 grid such as the National Digital Forecast Database (NDFD) grid, and it should be accurate and quality controlled.

Atmospheric scientists and hydrologists have been studying the behavior of precipitation over a wide range of spatial and temporal scales. However, because of paucity of data and the intermittency of precipitation, especially

precipitation associated with cumulus convection, analysis of observed rainfall distributions is often compromised as a tradeoff between spatial and temporal resolution: for example, hourly fields at catchment scales (Onof and Wheater 1996) versus monthly means at global scales (Chen et al. 1996).

Objective techniques have been developed and applied to construct analyzed fields of precipitation over global land areas from surface gauge observations (e.g., Xie and Arkin 1996; Chen et al. 2002). Spaceborne measurements of precipitation, with continuous developments and refinements of retrieval algorithms, have yielded operational precipitation products based on satellite observations of infrared (Arkin and Meisner 1987; Susskind et al. 1997; Xie and Arkin 1998), passive microwave (Wilheit et al. 1991; Spencer 1993; Ferraro 1997), and spaceborne precipitation radar (Kummerow et al. 2000). Although combining information from multiple satellite sensors as well as gauge observations and numerical model outputs yields analyses of global precipitation with stable and improved quality (e.g., Huffman et al. 1997; Xie and Arkin 1997; Xie et al. 2003), the merged precipitation products have one deficiency—that is, their quantitative uncertainty over land (e.g., Nijssen et al. 2001; Fekete et al. 2004).

Among the individual inputs used to define the combined precipitation analyses, both the satellite estimates and the model predictions are indirect in nature and need to be calibrated or examined using the gauge observations (e.g., Ebert and Manton 1998; McCollum et al. 2002). Therefore, gauge observations (Seo and Breidenbach 2002) play a critical role in constructing precipitation analyses over land. Gauge-based monthly precipitation analysis has been constructed over the global land domain (e.g., Xie and Arkin 1996; Dai et al. 1997; New et al. 2000; Chen et al. 2002). Similar analyses on submonthly time scales are relatively new because of limited accessibility of corresponding station observations from many countries. Nevertheless, the NCEP Climate Prediction Center (CPC) unified global daily gauge analysis (P. Xie et al. 2012, unpublished manuscript) has generated products over global land areas. This analysis is defined by interpolating quality controlled gauge reports at ~30 000 stations over the global land areas (about 12 000 over CONUS). For the purpose mentioned earlier, this CPC dataset bears more confidence but it provides only 24-h accumulation at 0.125° spatial resolution.

During the last decade or so, the advent of high spatial- and temporal-resolution precipitation analysis over CONUS made tremendous progress by combining gauge and radar observations. Currently, each of the 12 River Forecast Centers (RFCs) of the National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) routinely produces a precipitation analysis over its own domain and 9 of them use radar data. All RFCs but NWRFC (from which hourly analysis is not available) produce hourly as well as 6-hourly analyses. These analyses from individual RFCs are mosaicked at NCEP into a national product, called the NCEP Stage IV (Lin and Mitchell 2005).

The Stage IV precipitation dataset is used in the construction of precipitation statistics at scales that are sufficiently fine for many hydrologic applications (Kursinski and Mullen 2008). It is also used as input to hydrological models (Chen et al. 2007) and as truth for model verification (Zhao and Jin 2008). As it has a spatial resolution nearly equal to NDFD grid and a temporal resolution of 6 h, it is an excellent candidate to be used as the truth for bias correction and downscaling of precipitation forecast products. However, the product is subject to different methods of quality control and adjustments by different River Forecasting Centers. Although the implementation of Doppler radar at the national level has greatly improved precipitation estimates, serious limitations still exist. Despite its fine spatiotemporal resolution, caution must be employed when analyzing Stage IV data because of the uncertainty of radar retrievals in regions of complex terrain or melting hydrometeors. For this reason, some users restrict their analysis to the region east of 105°W and place highest confidence east of 100°W (Kursinski and Mullen 2008).

To provide a better proxy of the truth for the precipitation field over CONUS, it is apparently advantageous to combine the higher climatological reliability of the CPC dataset and the higher temporal and spatial resolution of the Stage IV dataset. We describe the development of such a new dataset by combining the two available datasets for this purpose. This paper is organized as follows: Section 2 provides descriptions of the Stage IV and CPC datasets used in the study; section 3 describes the methodology, including the statistical algorithm and the related application procedures; while the implementation of the new product and generation of the historical dataset are given in section 4. Qualitative and quantitative evaluations of the methodology and the new dataset are presented in section 5 and concluding remarks and further discussions are offered in section 6.

2. Input datasets

The CPC unified gauge-based analysis is constructed using the same interpolation algorithm described in Xie et al. (2007). First, gridded fields of daily precipitation climatology are constructed. This is done by interpolating the station daily precipitation climatology using

the inverse-distance technique of Shepard (1968) and then adjusting the gridded fields against the monthly climatology of Parameter-Elevation Regressions on Independent Slopes Model (PRISM). By doing this, seasonal evolution of precipitation can be captured very well with consideration of orographic effects. Daily precipitation analysis is defined by interpolating the ratio of the observed daily totals to the daily climatology through the optimal interpolation (OI) algorithm of Gandin (1965) and multiplying the interpolated ratio to the daily precipitation climatology.

For the CONUS land domain, the CPC analysis provides 24-h precipitation accumulations each day (1200–1200 UTC) at a 0.125° latitude–longitude mesh. CPC continuously collects gauge observations, performs basic quality control, and conducts a temporal version of the daily analysis on a regular basis. This "real time" CPC analysis uses slightly over 8000 daily gauges processed (often with quality evaluation) and sent to NCEP by individual RFCs. Additional nonreal-time daily gauge data are received from the National Climatic Data Center (NCDC), and this makes a full repository of typically over 12 000 station reports. A final version of the analysis can be generated after comprehensive quality control. For this study, CPC provided the final analysis for the period from 1 January 2000 to 31 December 2006. As the final analysis for 2007 and later years was not available, the temporal version is used for the period from 1 January 2007 through 6 November 2009. By mixing the two different versions, the sample size is increased with assumption that the statistical difference between the temporal and the final version of the CPC analysis can be neglected. For simplicity, this dataset is referred to as CPC.

NCEP's Stage IV precipitation analysis, mosaicked from the regional analyses produced at the 12 CONUS RFCs, provides area-averaged, 6-hourly estimates of precipitation on a 4-km pixel over CONUS. Each RFC produces a regional analysis over its own domain (see Fig. 1). The three western RFCs (NWRFC, CNRFC, and CBRFC) use the PRISM/Mountain Mapper approach to produce gauge-based analyses. The nine RFCs east of the continental divide use a multisensor approach to produce analyses using precipitation estimates from Weather Surveillance Radar-1988 Doppler (WSR-88D) radars, hourly gauges, and sometimes satellite data (when radar data are unavailable). Readers are referred to the online document on NWS/Advanced Hydrologic Prediction Service (AHPS) (NWS/AHPS 2011), which provides a reference on this issue.

The procedure has been operationally running and the products being archived since 1 January 2002. The data used in this study consist of 6-hourly (1200–1800, 1800–0000, 0000–0006, and 0600–1200 UTC) accumulations
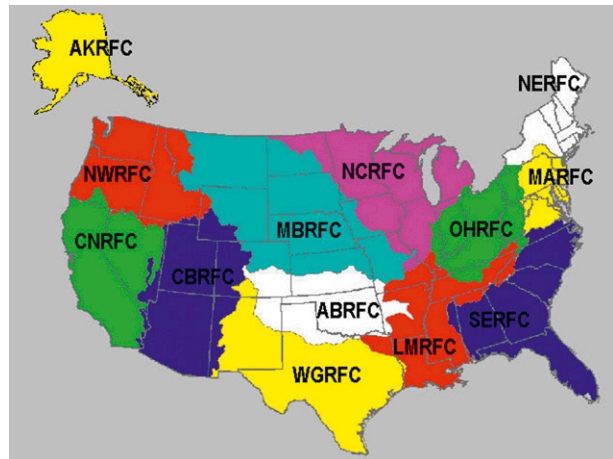

FIG. 1. The domains of the 13 NOAA RFCs. Note that the Stage IV analysis covers the 12 RFCs over CONUS.

on the 4-km Hydrologic Rainfall Analysis Project (HRAP) grid and span more than 8 years from 2002 to 2009 with only a few files corrupt or missing. Unlike the CPC analysis, which is defined only over the CONUS land area, the Stage IV analysis extends beyond the CONUS coast and political boundaries and covers some offshore areas and some bordering regions of Canada and Mexico (Fig. 2), though one must be more cautious when using data outside of RFC domains, as RFCs normally apply much more rigorous quality control inside of their own domains proper. In this paper, ST4 is used sometime to refer to the Stage IV data.

It should be pointed out that the CPC daily analysis and the Stage IV 6-hourly analysis are not completely independent to each other. There is considerable overlap between the gauge data used in the two datasets. For Stage IV, RFCs generally use all gauge data available to them. Hourly gauges available throughout the CONUS include the Hydrometeorological Automated Data System (HADS) gauges, the hourly Automated Surface Observing System (ASOS), and the Automated Airport Weather Stations (AWOS) reports. Locally, Snowpack Telemetry (SNOTEL) reports are used by western RFCs, and Integrated Flood Observing and Warning Systems (IFLOWS) are used by eastern RFCs. Additional gauges used by some RFCs include those from the municipal Automated Local Evaluation in Real Time (ALERT) and from the Limited Area Remote Collectors (LARC). This is likely not a comprehensive list, as the usage of gauge data and quality control process varies by each RFC. Our cross comparison of gauge station IDs shows that over 90% of the ~8000 daily reports used by the CPC analysis also appear in the HADS/ASOS/AWOS gauge lists. For gauges that provide both daily and more frequent reports, the daily reports used by CPC
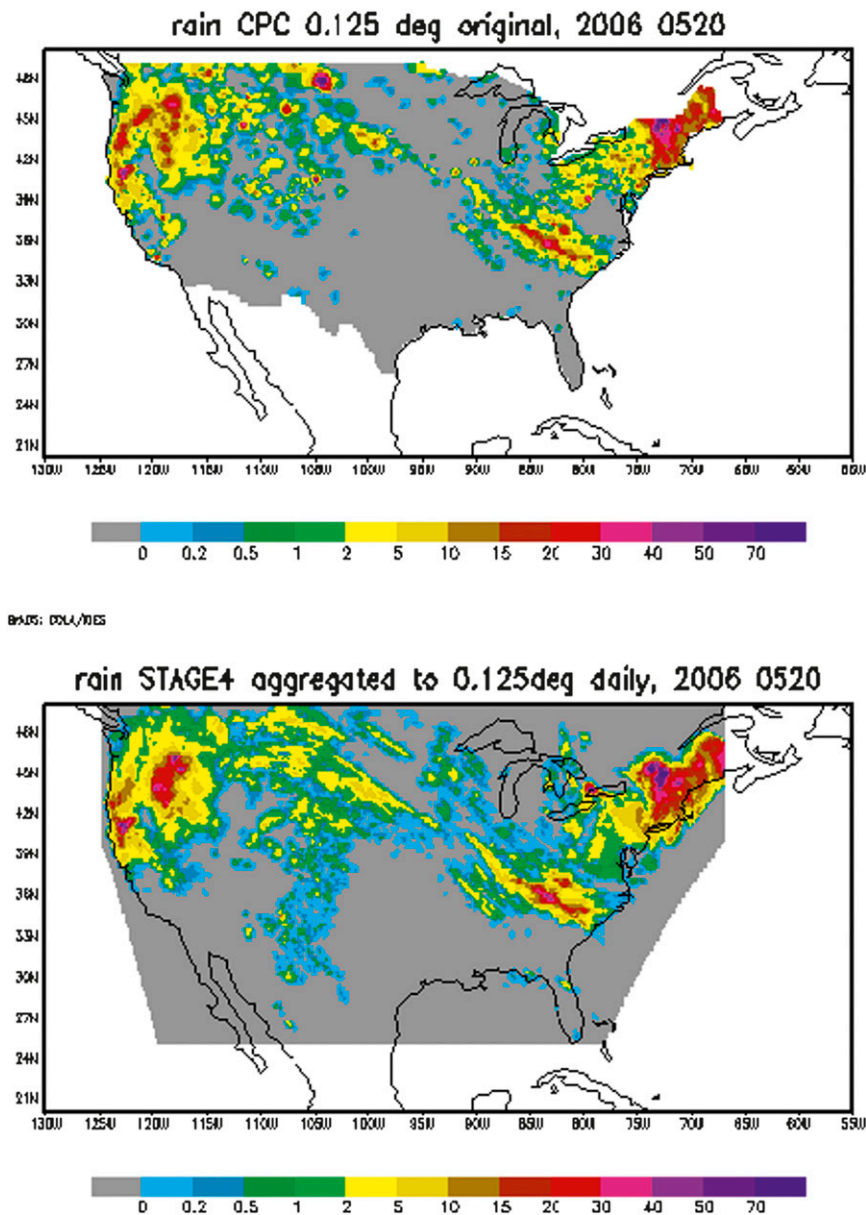
FIG. 2. (top) The 24-h precipitation from the 0.125° CPC analysis for 20 May 2006 and
(bottom) the Stage IV analysis aggregated to the same grid and accumulation time period.

are likely to be more reliable than their individual shorter-time readings (used by Stage IV), as the RFCs often perform some quantitative evaluation on the daily totals (Tollerud et al. 2005). The final version of the CPC analysis likely contains many more gauges that are independent from those used for Stage IV, since it employed ~4000 additional daily gauge reports not available in real time. Since the Stage IV analysis west of the Continental Divide (from NWRFC, CNRFC, and CBRFC) are based on PRISM/Mountain Mapper algorithm similar to that in the CPC analysis, Stage IV is closer to the CPC analysis in this region. To the east of the Continental Divide, the use of

radar precipitation estimates in the multisensor approach leads to greater independence between the two datasets.

Figure 2 provides a comparison of the two datasets for a randomly selected but typical example with ST4 aggregated to the same resolution and accumulation period as in CPC. While the two analyses have similar patterns of daily rainfall distribution, there exist significant differences. In this example, ST4 has more trace precipitation (<0.5 mm) and extreme values (>40 mm). CPC recorded a significant area of heavy precipitation (>20 mm) over northern Montana–Minnesota, which is missed in ST4.

3. Methodology and algorithms

The statistical adjustment of ST4 toward CPC has to be performed at the CPC grid of lower resolution. Therefore, the 6-hourly accumulations in the Stage IV dataset are first aggregated to the resolution of the CPC—that is, daily accumulation over 0.125° latitude–longitude grid boxes. As the next step, a statistical relationship is established between the two datasets at the CPC resolution and used to adjust the aggregated Stage IV data to make its climatology look like the CPC dataset. Finally, the adjusted Stage IV data are downscaled back to their original resolution to recover the highly desired variability in time and space.

a. Aggregation of 6-hourly Stage IV to daily accumulation at CPC grid

To adjust the Stage IV data toward the climatology of CPC, one needs to establish a statistical relationship between the two analyses. As CPC has lower resolution in both space and time, the first step of the procedure is to aggregate the Stage IV data to the resolution of the CPC dataset. The four 6-h accumulations from 1200 UTC of a specific day to 1200 UTC the next day are first aggregated into a single 24-h accumulation for the day in consideration. Mathematically, this is written as

$$\mathrm{ST4}_{H,24\mathrm{h}} = \sum_{n=1}^{4} \mathrm{ST4}_{H,6\mathrm{h}}^{n}, \qquad (1)$$

where the subscript H indicates the higher spatial resolution associated with the original HRAP grid of the Stage IV data (abbreviated as ST4). The next step is interpolating $\mathrm{ST4}_{H,24\mathrm{h}}$ to the CPC 0.125° grid (indicated by subscript L), symbolically written as

$$\mathrm{ST4}_{H,24\mathrm{h}} \rightarrow \mathrm{ST4}_{L,24\mathrm{h}}, \qquad (2)$$

with the arrow representing an interpolation/extrapolation scheme.

In this study, a carefully designed budget interpolation method, used operationally at the NCEP Environmental Modeling Center (EMC), is applied to remap precipitation from the finer ST4 grid to the coarser CPC grid, and vice versa. To a desired degree of accuracy, the procedure conserves the precipitation amount in the input grid. The algorithm simply computes weighted averages of bilinearly interpolated points arranged in a square box centered around each output grid point and stretching nearly halfway to each of the neighboring grid points. Some details of the algorithm can be found in Mesinger (1996) and Accadia et al. (2003).

Despite the precipitation conservation of the budget interpolation algorithm, Eq. (2) is not reversible. In other words, when $\mathrm{ST4}_{L,24\mathrm{h}}$ is interpolated from the CPC grid back to the original ST4 grid, the resulting $\mathrm{ST4}_{L\text{-}H,24\mathrm{h}}$ is different from $\mathrm{ST4}_{H,24\mathrm{h}}$ because of the information loss in the interpolation. Here the subscript L-H represents an interpolation from the lower-resolution grid to the higher-resolution grid. Nevertheless, the ratio between the two fields at each high-resolution grid point—that is,

$$\frac{\mathrm{ST4}_{H,24\mathrm{h}}}{\mathrm{ST4}_{L-H,24\mathrm{h}}} \qquad (3)$$

—can be used to recover the lost information of the fine-scale patterns.

b. Linear regression for each day of the year and each CPC grid box

Applying statistical adjustment to precipitation is difficult because of the noncontinuous nature (i.e., it is nonnegative) and nonsymmetric probability distribution of the variable, and no method is satisfactory for all applications. The probability distribution function (PDF) matching method is widely used in calibration of forecasts against observations but its application requires well-defined PDFs of the input datasets, which, in turn, requires a sufficiently large sample size.

The purpose of the statistical adjustment of the Stage IV data is to make its climatology close to that of the CPC data. Because of the existence of complicated geographic patterns and orographic features in space and domination of annual cycle or seasonal variation in the precipitation observation and analysis (Chen et al. 2002; Xie et al. 2007), climatology is better defined at each grid box for each day of the year. With just 7 years of training data, the sample size may not be sufficiently large for the PDF matching method. On the other hand, visual examination of many scatterplots for various locations in the analysis domain indicated that the relationship between the CPC and Stage IV analysis are generally very strong, partly owing to the overlapping information content discussed in section 2. This suggests that the conventional linear regression method can be employed. This is not the optimal approach but it provides a straightforward solution to the problem and more sophisticated algorithms, including PDF matching, will be tested in the future when larger training datasets are available.

Even this linear regression approach requires increased sample size. A 61-day window is used by including the 30 days before and after the day in consideration with a maximum sample size of 427. This choice of the width of the sampling window is the result of a compromise between a reasonable actual sample size (wet days) at most grid points
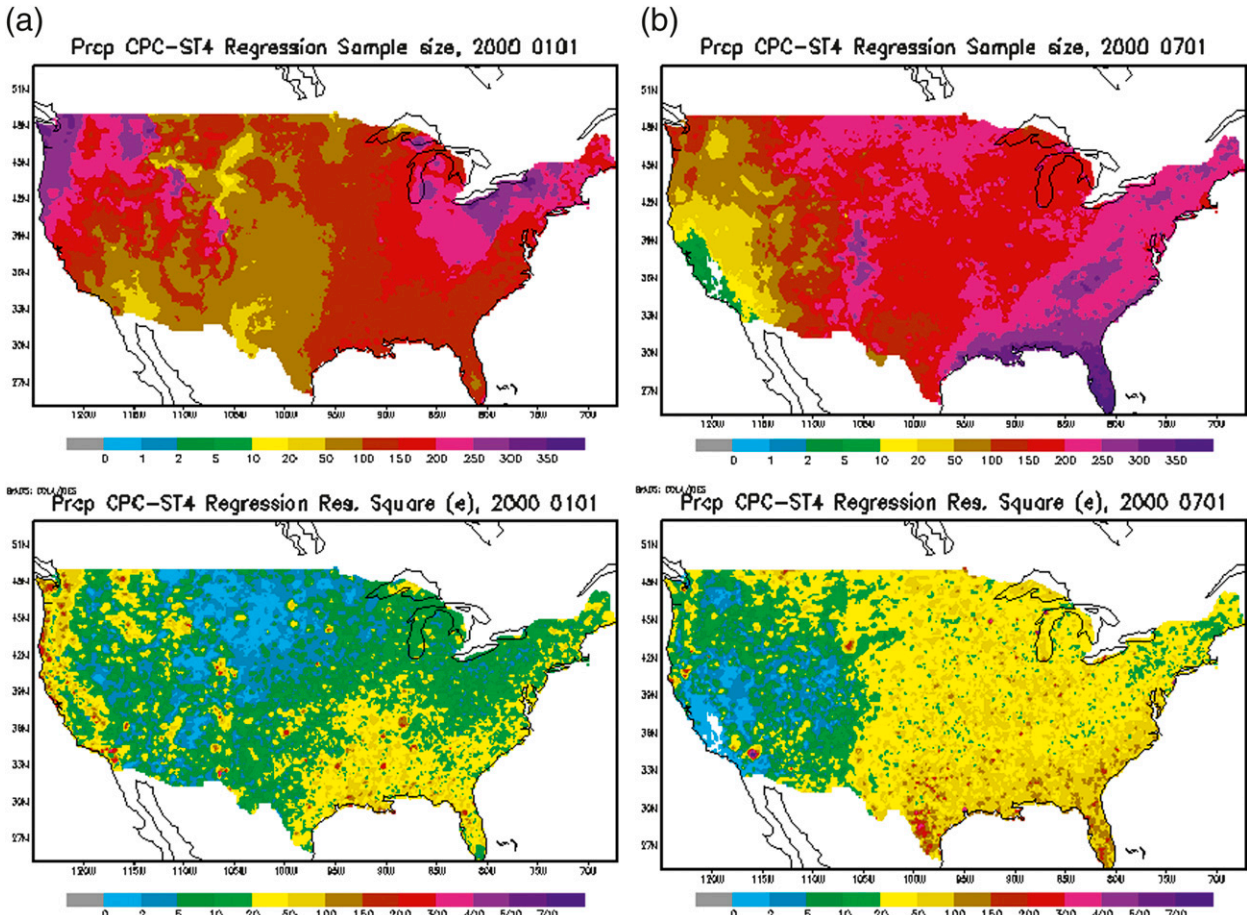
FIG. 3. (a) The (top) actual sample size and (bottom) regression residual square error (mm) for 1 Jan. (b) As in (a), but for 1 Jul.

for most days and a relatively uniform sample for each day. The actual sample size is dependent on the geographic location of the grid box and the season because only the days with recordable precipitation indicated by Stage IV are counted to reduce the impact of the large number of "dry" days. Figure 3 shows the actual sample size for 1 January and 1 July, as two examples. With this 61-day window, the actual sample size in the eastern United States is over 200 for most cases, while an empty sample is encountered over the southwestern desert region during the summer months (May–September). Note that a spatial window is avoided to retain the higher spatial resolution of the Stage IV data.

For each CPC grid box and each day of the year, a regression of CPC on the Stage IV precipitation is established; that is,

$$CPC = aST4_{L,24h} + b, \tag{4}$$

where a and b are the slope and intercept, respectively. As noted earlier, the subscripts of ST4 are used to emphasize the difference in the spatial and temporal

resolution between the original data and the aggregated value.

The regression coefficients, a and b, are estimated with the least squares solution and the fitting error, or the average of squared residuals [residual sum of squares (RSS) divided by the sample size] is shown in Fig. 3 for two special dates. The fitting error is generally less than 100 mm^2 except in isolated small areas along the West Coast and in the South, but larger in the rainy season (winter in the West Coast and summer to the east of the Rockies). The quality of the fitting tends to be higher over the northern part of the Great Plains where the fitting error is less than 20 mm^2 in the winter. On the other hand, the lower RSS error over the Southwest may indicate overfitting with very small sample sizes.

c. Gap filling and temporal smoothing of the regression coefficients

Figure 4 displays the maps of the regression coefficients a and b for 1 August, estimated from the data sample. Both parameters show significant spatial variations with
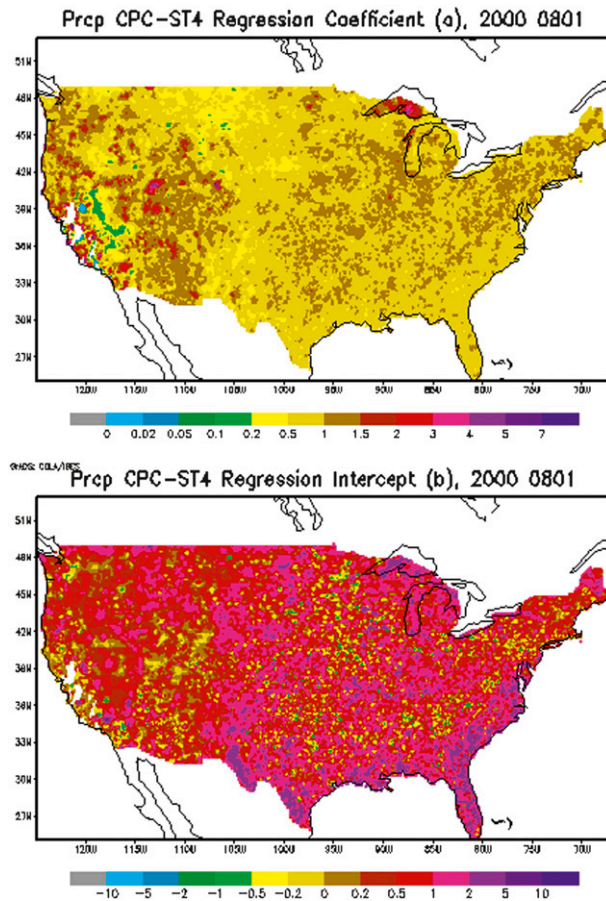
FIG. 4. Regression coefficients (top) *a* and (bottom) *b*, for 1 Aug, estimated using 7-yr historical datasets and a 61-day time window.

much larger amplitudes and much more finescale patterns over the West, especially in the mountainous areas. This is consistent with the analysis of Kursinski and Mullen (2008) on the quality of the Stage IV dataset and suggests that the adjustment by the regression is working in the correct direction by removing the impact of the terrain.

There are some issues to be addressed with these coefficients. First, as discussed earlier and shown in Fig. 3, the regression is invalid in some areas of the southwestern United States during the summer months, with the actual sample size close to 0. Second, there are some abnormally large values of the slope *a* at a few grid points, with the maximum being over 100 (see Fig. 5). A careful investigation of these rare cases suggests that these extreme slopes occur over the Pacific coastal areas where Stage IV precipitation is systematically smaller than CPC by one or two orders of magnitude. This may be caused by systematic differences in the algorithms of the two analyses and remains to be further investigated but out of the focus of the current study. Finally, despite the clear seasonal variations, the time series of both coefficients are characterized by irregular jumps and drops (Figs. 5 and 6).

To deal with the above-mentioned problems, two steps are taken to modify and refine the slope and intercept separately. First, an interpolation algorithm is applied independently to the *a* and *b* fields for each day of the year to fill the spatial gaps and replace the unreliable values from very small samples. For any grid box where the coefficients are missing (sample size $N < 2$) or unreliable (sample size less than a threshold, $N < N_{min}$), the gap is filled by a weighted average of the same coefficient at the grid boxes in its vicinity. This type of algorithm is widely used in the analysis of irregularly spaced data (Shepard 1968) such as precipitation (Xie et al. 2007). The second step is to smooth the 366-day time series of each coefficient through Fourier truncation in which the raw time series is replaced by the accumulation of the first three harmonic components. Xie et al. (2007) used this method to remove the high-frequency noises in the daily climatology of precipitation in their gauge-based precipitation analysis over East Asia. Experiments with the number of the harmonic components suggest that three is a reasonable choice for the coefficients, in contrast to six used by Xie et al. (2007) for precipitation of daily climatology. As shown in Figs. 5 and 6, the smoothed time series well represent the annual cycle and long-term trend.

The choice of N_{min} may affect this refinement of the coefficients. However, carefully checking some examples of time series suggests that interpolated values during the summer season usually have poor continuity with the raw values before and after the dry season. As a result, these unrepresentative values are more subtly changed in the temporal smoothing step and the results are not sensitive to the choice of N_{min}. Experiments with $N_{min} = 2$ and $N_{min} = 10$ showed little difference and thus $N_{min} = 2$ is used. The refined regression coefficients, while having smoother temporal variations, have little difference from the raw values on daily maps. The example of 1 August is shown in Fig. 7.

d. Application of the regression to the aggregated Stage IV data

The refined regression coefficients *a* and *b* (without changing symbols) are applied to adjust the aggregated Stage IV data $ST4_{L,24h}$; that is,

$$ST4'_{L,24h} = aST4_{L,24h} + b, \qquad (5)$$

where the prime represents the results of the adjustment. From Figs. 4 to 7, the slope *a* is between 0.5 and 1.5 and intercept *b* in the range of −2 to 6 mm for most of the 0.125° grid boxes, especially over the eastern two-thirds of the CONUS domain. Over the mountainous
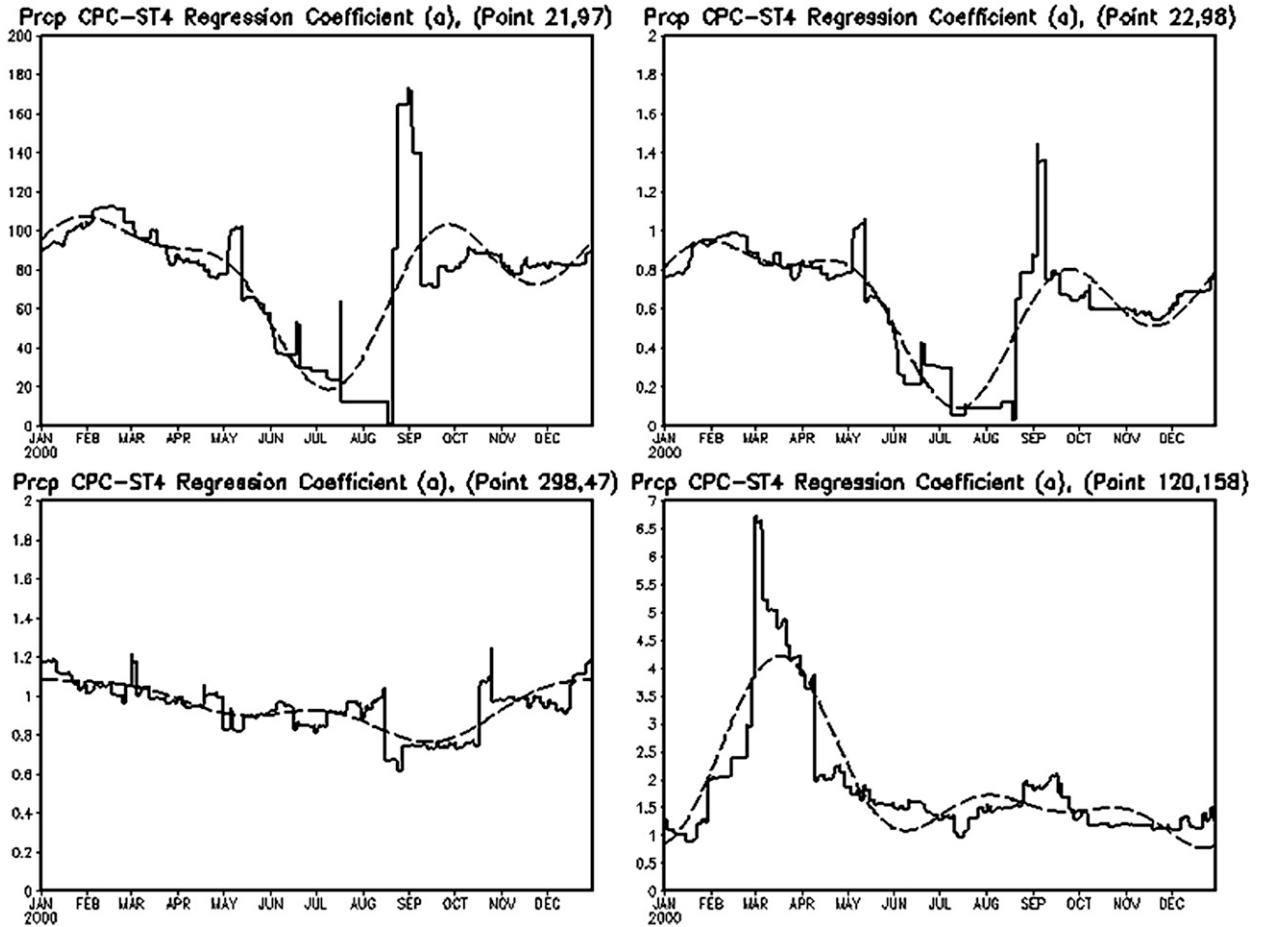
FIG. 5. Time series of regression coefficient *a* for four grid points. The gap-filled raw value (solid) and the smoothed version (dashed) are shown.

West, the slope has a larger range, but it is still between $1/7$ and 7 except for a very few points. As discussed earlier, even a value as large as 100 is consistent with the inputs into the regression analysis. As a result, no extreme precipitation is produced because of the abnormally large value of coefficient *a*. To avoid possible occurrence of this unexpected situation in the real-time application, an upper limit of 500 mm is set for $ST4'_{L,24h}$. On the other hand, the relatively more frequent cases of negative intercept *b* can result in negative precipitation in the adjusted value. Whenever this happens, $ST4'_{L,24h}$ is set to be zero.

e. Downscaling the adjusted precipitation in space and time

The final step of the process is to downscale the adjusted Stage IV analysis $ST4'_{L,24h}$ in both space and time. As discussed in section 1, the goal of adjusting the Stage IV data is to retain the higher spatial and temporal resolution while adjusting its climatology toward that of

the CPC. As discussed in section 3b, a simple interpolation of $ST4'_{L,24h}$; that is,

$$ST4'_{L,24h} \rightarrow ST4'_{L-H,24h} \tag{6}$$

even using the budget interpolation scheme water volume conservation scheme, will not achieve this goal because of the irreversibility of (2). Nevertheless, $ST4'_{L-H,24h}$ can be used as the basis for the spatial downscaling. Recall that the ratio (3) at each high-resolution grid box can be used to "remember" the finescale structures, which in turn can be posed to $ST4'_{L-H,24h}$. In other words,

$$ST4'^{*}_{L-H,24h} = ST4'_{L-H,24h} \times \frac{ST4_{H,24h}}{ST4_{L-H,24h}} \tag{7}$$

can be used to recover the finescale structures with the asterisk (*) indicating the spatial downscaling.

The temporal downscaling is conceptually straightforward: $ST4'^{*}$ is disaggregated into the four successive,
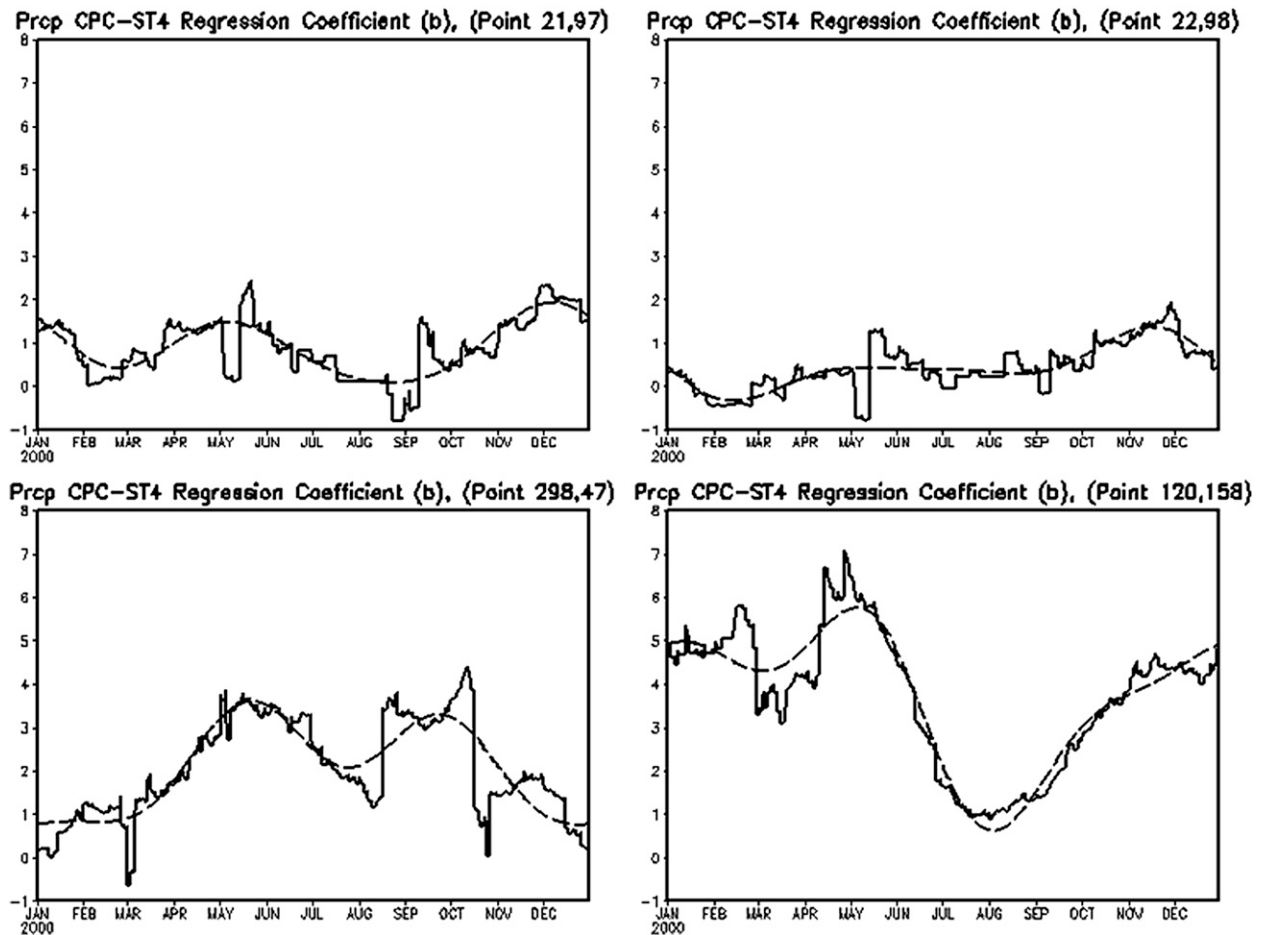
FIG. 6. As in Fig. 5, but for regression coefficient *b*.

nonoverlapping 6-h periods with the same proportion as that in the raw Stage IV records. Mathematically, this is the inversion of Eq. (1) except applied to ST4′* instead of ST4; that is,

$$\text{ST4}'^{*k}_{H,6h} = \text{ST4}'^{*}_{L-H,24h} \times \frac{\text{ST4}^{k}_{H,6h}}{\sum\limits_{n=1}^{4} \text{ST4}^{n}_{H,6h}} \quad \text{for} \quad k = 1, 2, 3, 4.$$

(8)

In (7) and (8), the grid value for the left-hand side is automatically set to zero when the denominator is zero.

4. Operational implementation and dataset status

A software package has been developed to implement the algorithm described in section 3. The first component of the package determines the regression coefficients *a* and *b*, following steps (a), (b), and (c) described in section 3 and using the historical datasets of CPC and stage IV for the period from 1 June 2002 to 31 July 2009. The chosen training period provides an identical potential sample size of 427 ($=61 \times 7$) for all 366 days of the year, with a 61-day window and exactly 7 years of data. In the second component of the package, the gap-filled and temporally smoothed version of the regression coefficients are retrieved from an archive and applied to adjust real-time and historical Stage IV 6-h accumulated precipitation analyses, following steps (a), (d), and (e) of the algorithm.

The later component of the software package, named Climatology-Calibrated Precipitation Analysis (CCPA), was implemented in NCEP's production suite on 13 July 2010. Since then, CCPA has been running on a real-time basis to process the Stage IV 6-h precipitation data. Following the data flow and schedule of Stage IV products, the first version of the CCPA dataset for the 24-h period ending 1200 UTC each day is available shortly after 1600 UTC and it will be updated 8, 32, and 40 h later to reflect the updates in stage IV data. For most days, the final version is available with the first
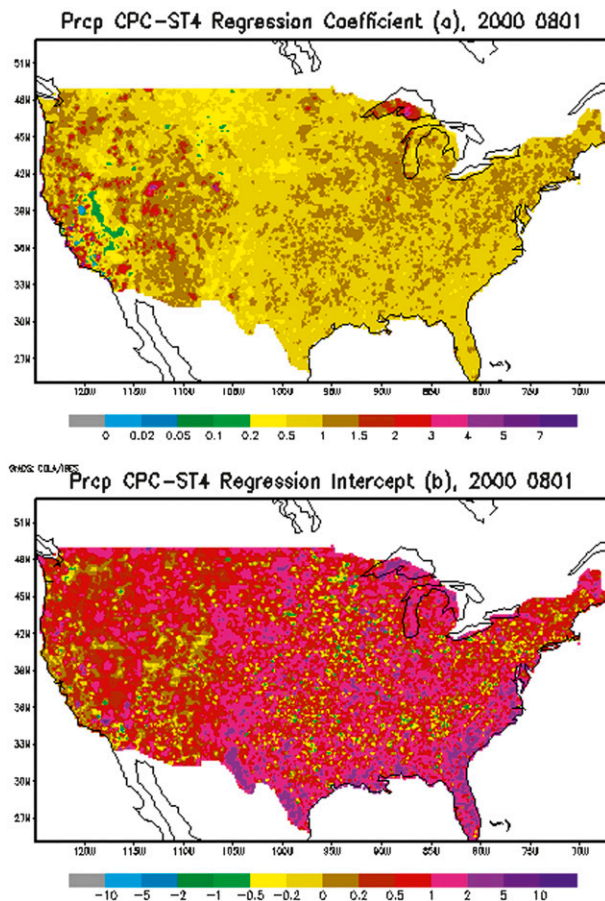
FIG. 7. Regression coefficients (top) *a* and (bottom) *b* for 1 Aug after gap filling and smoothing.

update shortly after 0000 UTC the next day. This product is also called CCPA.

To take advantage of this new CCPA product, its historical archive was generated at NCEP/EMC for the period from 1 January 2002 to 13 July 2010. This historical archive, combined with the real-time output, is available to the meteorological/hydrological community and the general public. As a calibrated version of the Stage IV 6-h precipitation analysis, CCPA can be used in evaluation and calibration of precipitation forecasts. Using the budget interpolation algorithm described in section 3a, CCPA is converted to the 5-km NDFD grid and latitude–longitude grids at 0.125°, 0.5°, and 1.0° resolution, all covering the CONUS domain.

5. Evaluation

To evaluate the methodology and dataset described in this paper, CCPA and Stage IV can be directly compared at the original resolution of the latter, with 6-h accumulations. Figure 8 displays the example for the 6-h

period starting at 1800 UTC 30 December 2009. The quantitative difference between the two fields is visible, especially in the shape and extent of the 10-mm contours in the lower Mississippi states and Utah. This is the impact of the CCPA adjustment. However, the general precipitation patterns in CCPA (Fig. 8a) and ST4 (Fig. 8b) are almost identical. In fact, the spatial pattern correlation coefficient between the two fields are always well above 0.99. Therefore, there is no doubt that the finescale structures are retained by the CCPA methodology.

As described in section 2, the central piece of the CCPA methodology is the application of linear regression between ST4 and CPC at 0.125° resolution and 24-h accumulation. Therefore, evaluation of the CCPA methodology and dataset is focused on this aspect. For this purpose, CCPA and the original Stage IV data (ST4) are aggregated to 0.125° resolution and 24-h accumulation periods, and compared with CPC. As the data sample for regression analysis is relatively small, there is a need to test the robustness of the methodology. Following an approach similar to Xie et al. (2007), cross validation is performed with a data holding technique. Regression slope and intercept are reestimated with the same sample pool as described in sections 3b, 3c, and 4, except that the data for a particular 1-yr period (1 July to 30 June next year) are excluded, and the analysis for the same period is reproduced from these new regression coefficients using (5). The same procedure is repeated for each of the seven exclusive and nonoverlapping periods and the dataset reproduced is referred to as the cross-validation analysis (CVA).

As a single point example, Fig. 9 shows the time series of the three analyses (CPC, ST4, and CCPA) for a typical grid point located at 42°N, 102°W over the warm and wet seasons of July–August 2008 and May–June 2009. Overestimation in ST4 compared with CPC for most precipitation events and phase difference between the two are clearly shown. As expected, CCPA generally follows the variation of ST4 and brings it toward CPC in most cases. However, exceptions do exist and the most striking one is the 15 August 2008 event. For that day, ST4 missed the heavy rainfall (about 25 mm) documented by CPC and CCPA did little to adjust it. This issue will be addressed in section 6. CVA (not shown) generally follows the variation of CCPA.

Objective evaluation of the CCPA methodology and dataset requires comparing some statistic scores averaged over some extended periods. For this purpose, an annual average is used in this paper as a 1-yr period is long enough to smooth out random variations and short enough to include CVA in the comparison. Figure 10a displays the 24-h accumulation from the CPC dataset, averaged over the 1-yr period from 1 July 2008 through
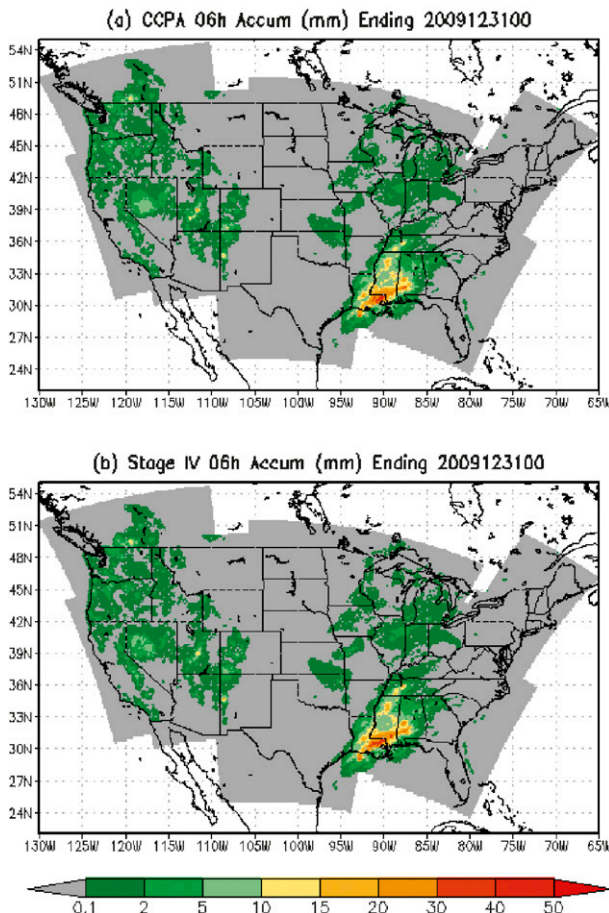
FIG. 8. The 6-h precipitation from (a) CCPA at 4-km HRAP grid accumulated for the period from 1800 UTC 30 Dec to 0000 UTC 31 Dec 2009 and (b) Stage IV analysis at the same grid and accumulation time period.

FIG. 9. Time series of 24-h precipitation at (42°N, 102°W) from 0.125° CPC (short dashed line), Stage IV (dotted and dashed line), and CCPA (solid line) for two periods: (a) 1 Jul–31 Aug 2008 and (b) 1 May–30 Jun 2009.

30 June 2009. Since the emphasis here is to evaluate how much CCPA is closer to CPC in contrast to Stage IV, the differences of Stage IV, CVA, and CCPA with respect to CPC are shown in Figs. 10b, 10c, and 10d, respectively. Without calibration, Stage IV has larger differences with either negative or positive values, ranging from −5.3 to 5.6 mm (Fig. 10b). Although the spatial patterns are patchy, the positive "bias" is apparently larger over the Missouri Basin River Forecast Center (MBRFC) area (Figs. 1 and 10b) than other areas, exceeding 5 mm at many points. In fact, the time series in Fig. 9 are from this area. As discussed in the introduction, different RFCs use different quality control and algorithms and this leads to different statistical properties of the ST4 analysis.

Since CCPA and CVA are statistical adjustments to Stage IV in magnitude, their "biases" preserve the finescale features of high spatial variability in Stage IV and have much smaller amplitudes, with the absolute
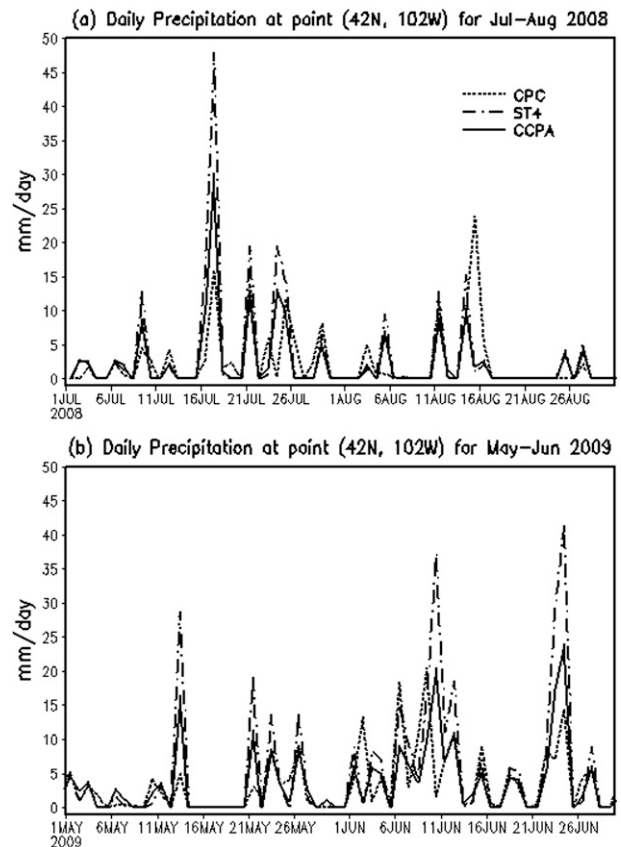
value being less than 1 mm for almost all grid points. Particularly, the widespread large positive bias over MBRFC is significantly reduced. In addition, the improvements in CVA and CCPA are very similar to each other, suggesting that the methodology is reasonably robust and the sample size is (though marginally) sufficient for the regression analysis. The positive impact of the CCPA statistical adjustment is even more appealing in terms of root-mean-square error (RSME). From Fig. 11a, it can be seen that RMSE of ST4 is roughly following the pattern of annual mean precipitation (Fig. 10a) while the maxima along the West Coast and over the southern plains can be as large as 20 mm. The percentage of RMSE reduction by the CCPA procedure, displayed in Fig. 11b, is positive everywhere except for a few scattered local spots in the Southwest desert, Great Lakes, and other remote regions. CCPA reduces the RMSE by 40% or more over a large portion of CONUS.

In addition to the long-term mean, another aspect of the precipitation climatology is the standard deviation. From Eq. (5) and the least squares solution to linear
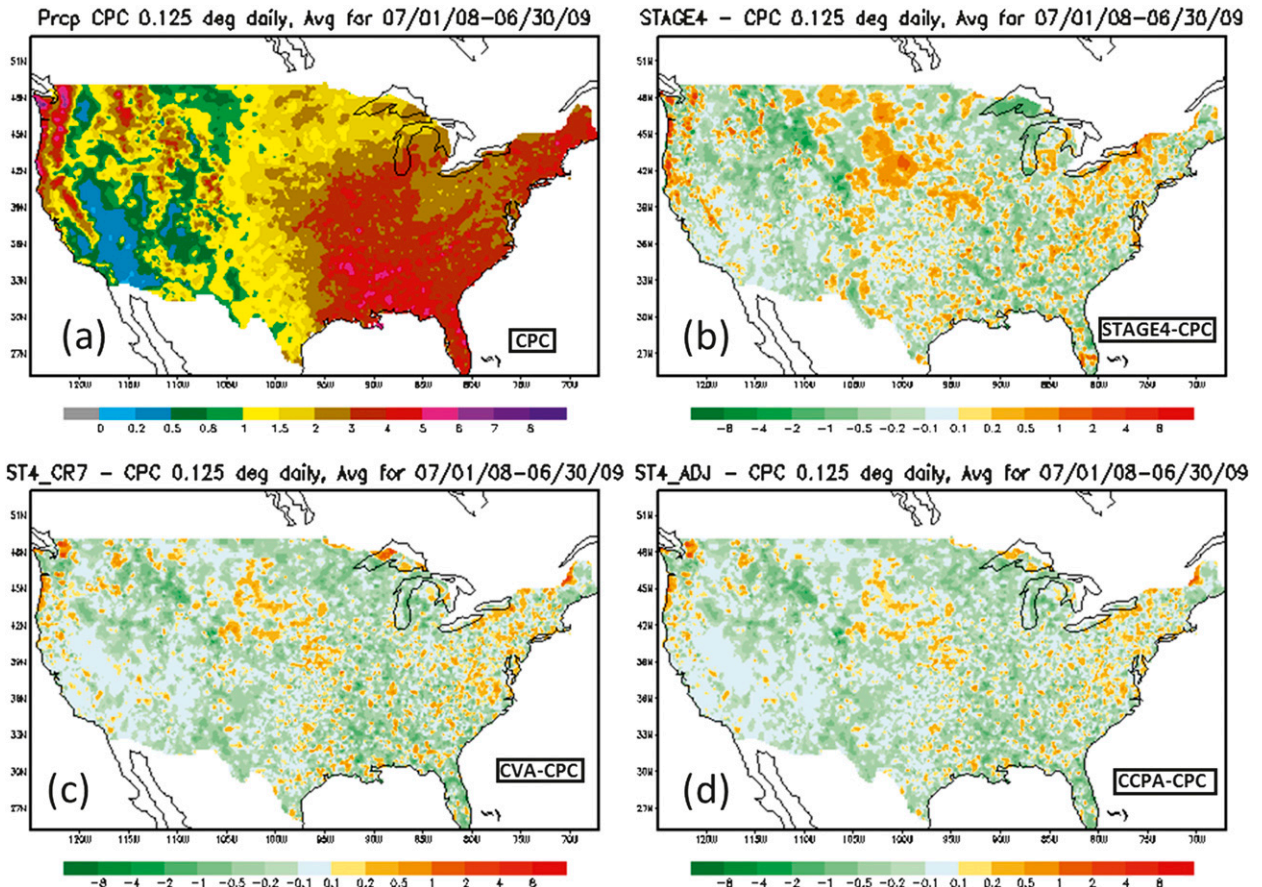
FIG. 10. (a) The 24-h precipitation from 0.125° CPC averaged between 1 Jul 2008 and 30 Jun 2009 and the differences of (b) Stage IV, (c) CVA, and (d) CCPA with respect to CPC. Stage IV, CVA, and CCPA are aggregated to the same grid and accumulation time period as CPC.

regression (4), it is straightforward that when evaluated with the 7-yr dependent sample, CCPA would have a mean equal to that of CPC and its standard deviation multiplied by the regression slope a. This adjustment of the mean toward CPC can be seen for a partial sample, as demonstrated in Fig. 10. The change in standard deviation can be displayed by comparing the cumulative distribution function (CDF) of the daily precipitation of the three datasets. Figure 12 shows such a comparison for the same grid point as shown in Fig. 9 for the 1-yr period from July 2008 to June 2009. The difference in the CDF between the ST4 and CPC is clear: ST4 has higher probability of precipitation (POP) compared with CPC (0.42 versus 0.36); the extra precipitation events in ST4 are concentrated in very light rain [0–1 mm (0.16 versus 0.11)] and heavy rainfall [>30 mm (0.04 versus 0.01)]. Since the regression slope is around unity and mainly in the (0.5, 1.0) range and the intercept is mainly positive and small (less than 5 mm), the adjustment in CDF is basically moving the extreme cases toward the middle range. As expected, the CDF of CCPA is very close to that of CPC except near the dry end, where the negative

values in CCPA are artificially set to 0 (section 3d). To quantify this improvement, the 7-yr historical dataset is used to construct the CDF at each grid point. CPC is assumed as the observation and continuous ranked probability scores (CRPS) are calculated by integrating over the range between 1 and 29 mm. The difference in this partial form of CRPS between CCPA and ST4 is shown in Fig. 13 and the error reduction in CPPA is clearly shown for most grid boxes. Significant increases in CRPS are found only over limited areas, with concentrations in the South, especially along the coast of the Atlantic and Gulf of Mexico.

To further examine the impact of the statistical adjustment and the quality of the CCPA dataset, CCPA, ST4, and CVA are verified against a different dataset, estimated from NOAA RFC rain gauge network (Zhu 2007). These daily rain gauge reports are box averaged to 0.125° and defined as RFC rain gauge analysis (RFC). As CPC final analysis uses about 4000 extra daily gauge reports, RFC is considerably different from CPC, despite the overlap in information content. For this reason, verification against RFC provides an alternative, but
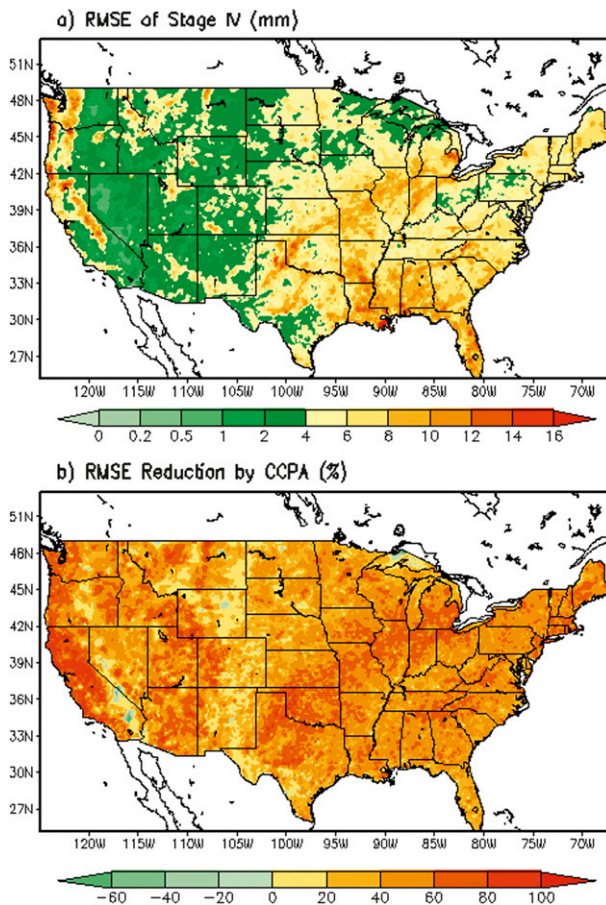
FIG. 11. (a) RMSE of Stage IV daily precipitation at 0.125° CPC grid, calculated for the period between 1 Jul 2008 and 30 Jun 2009. (b) Percentage of RMSE reduction due to the CCPA procedure; that is, the difference in RMSE between Stage IV and CCPA, normalized by the RMSE of Stage IV shown in (a).
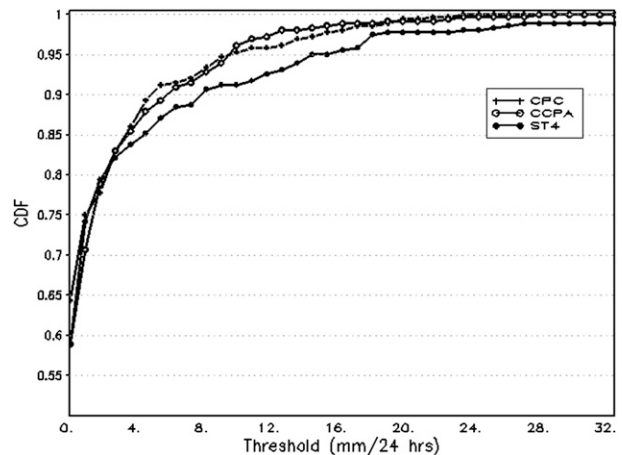

FIG. 12. An example of CDF of CPC (cross), ST4 (closed circle), and CCPA (open circle), constructed using daily accumulation data from 1 Jul 2008 to 30 Jun 2009, for the 0.125° grid.

nonindependent, test to the CCPA methodology. RMSE and absolute mean error (ABSE) of 24-h precipitation are calculated against the RFC rain gauge analysis as a function of precipitation threshold. For each threshold value, all of the grid points where the RFC rain gauge analysis is less than the value are excluded in the calculation. These statistical scores are averaged over each of the seven exclusive and nonoverlapping 1-yr periods. The results for the last period are shown in Fig. 14. When calculated with all grid points with observed precipitation (threshold 0.0 or 0.2 mm), there are clear improvements in terms of RMSE in both CCPA and CVA over the raw stage IV data, with a RMSE reduction of 0.27 mm for CCPA and 0.11 mm for CVA. The improvements are still evident for higher thresholds up to 15 mm. For large precipitation amounts (thresholds of 35 and 50 mm), the improvement by CCPA is less impressive. For thresholds greater than 20 mm, CVA is not as good as ST4. The statistics in terms of ABSE, though showing less

difference among the three lines, still supports similar conclusions. The deterioration of CVA at higher thresholds suggests that the 61-day window in the sample pool of the regression analysis is necessary with the relatively short ST4 archive and caution must be practiced in the application of CCPA with heavy precipitation events.

6. Summary and further discussions

A simple linear regression method is employed to statistically calibrate the multisensor Stage IV precipitation analysis over the CONUS domain and makes its climatology closer to that of the rain gauge–based estimate of the CPC unified global daily gauge analysis. Available archived historical datasets of the two analyses within a 7-yr period are used to estimate the regression coefficients for each Julian day of the year and each grid point in the CPC grid mesh (0.125° latitude–longitude) over the CONUS. After gap filling and temporal smoothing, these coefficients are then applied to the spatially and temporally aggregated Stage IV data to generate an adjusted analysis, which is then downscaled in space and time to the original grid and accumulation period of Stage IV. The procedure, referred as Climatology-Calibrated Precipitation Analysis (CCPA), was implemented in the NCEP production suite on 13 July 2010 to routinely process incoming stage IV 6-h precipitation analyses and generate corresponding CCPA files. It was also used to process the historical Stage IV dataset covering the period from 1 January 2002 to 13 July 2010 to form a complete archive of CCPA.

Subjective and objective methods are used to evaluate the CCPA methodology and the corresponding dataset. CCPA retains the fine spatial and temporal structures of ST4 analysis; it has a climatological distribution (mean
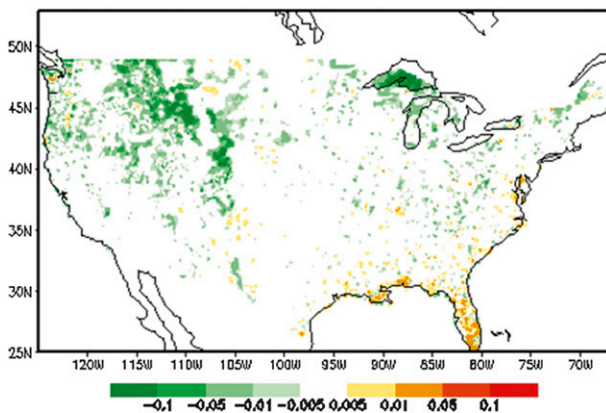
FIG. 13. Difference in CRPS between CCPA and ST4. CRPS is calculated from the corresponding CDFs from 7-yr historical datasets with the corresponding CDF of CPC as the reference. Negative values indicate CRPS reduction owing to the adjustment of ST4 using CCPA methodology.
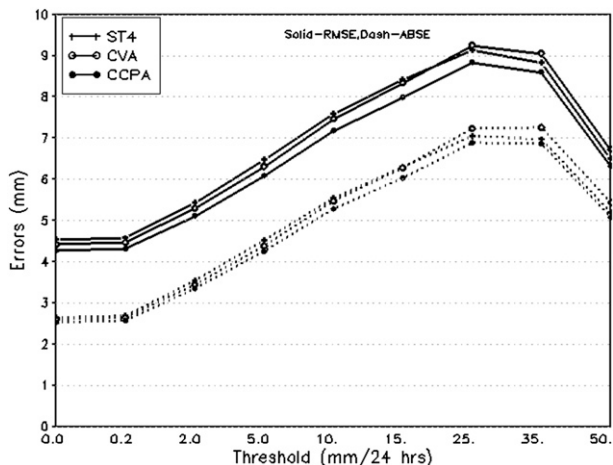

FIG. 14. RMSE (solid line) and ABSE (dotted line) of 24-h precipitation from Stage IV (cross), CVA (open circle), and CCPA (closed circle) verified against RFC rain gauge analysis at the 0.125° grid, as a function of precipitation threshold.

and cumulative distribution function) closer to that of the CPC analysis for daily analysis at 0.125° resolution; when compared with CPC or a similar dataset from gauge observations, CCPA shows significant improvement in terms of root-mean-square error (RMSE). Cross-validation analysis generated by the same method but with partially held training data has similar properties but slightly less improvement, suggesting that the CCPA methodology is robust and further improvement can be expected as the training datasets become larger.

In addition to the choice of the linear regression approach discussed in section 3b, the methodology used in this study is also subject to other limitations. The estimation of regression coefficients only used data points with ST4 positive and those samples with positive CPC value but zero ST4 record are neglected. Various artificial rules are set in the application of CCPA: the adjustment is applied only when daily precipitation is greater than zero and all negative values resulting from the regression will be set to zero in the adjusted version. As a result, the probability of precipitation (POP) is not adjusted to the CPC value (see Fig. 12 and its discussions).

When compared with CPC or a similar dataset based on gauge observations, CCPA shows a less impressive impact for higher threshold values. The reason for this nonuniform behavior can be revealed by scatterplots of Stage IV analysis against CPC values in the sample of regression development for each grid point and each Julian day. Four selected examples for 1 July are shown in Fig. 15 and the corresponding CCPA/CPC points are also plotted. In an ideal case (Fig. 15a), all data points are concentrated in a narrow band along the regression line close to the diagonal and the adjustment is trivial. A less ideal case, shown in Fig. 15b, has a less concentrated

band along a regression line away from the diagonal and hence larger adjustment by CCPA. For the cases shown in Figs. 15c and 15d, the data points are widely scattered and the slope of the apparent "regression" curve is different for the lower and higher precipitation ranges. As heavy precipitation events are scarce, a linear regression is always dominated by the lower precipitation points. While the mean is brought closer to CPC by the CCPA procedure, some points with large precipitation in CPC but much smaller ST4 value could be taken further away from the diagonal (and contributes to larger RMSE). This scenario is seen in Fig. 15d where ST4 missed the only few heavy precipitation records in CPC despite its overall overestimation for light and medium rain.

In summary, the merit of the methodology used in this paper is limited by the weakness of the simple linear regression model, an inadequate sample of high-precipitation events, and the noncontinuous nature of precipitation. For some grid points, Stage IV data tend to underestimate the frequency of heavy precipitation and the CCPA procedure may not be able to correct it. Cross validation suggests that the current estimation of regression coefficients is fairly robust with the current 7 years of archived data over all precipitation events, but may be inadequate for heavy rainfall amounts. Therefore, the users of the CCPA dataset should take extra caution with heavy precipitation, especially the extreme events.

The quality of CCPA should be improved in the future by increasing the length of the archived CPC and Stage IV datasets and using a more realistic statistical adjustment method. There is a plan at NCEP/EMC to perform periodic updating of the regression coefficients with increased sample size.
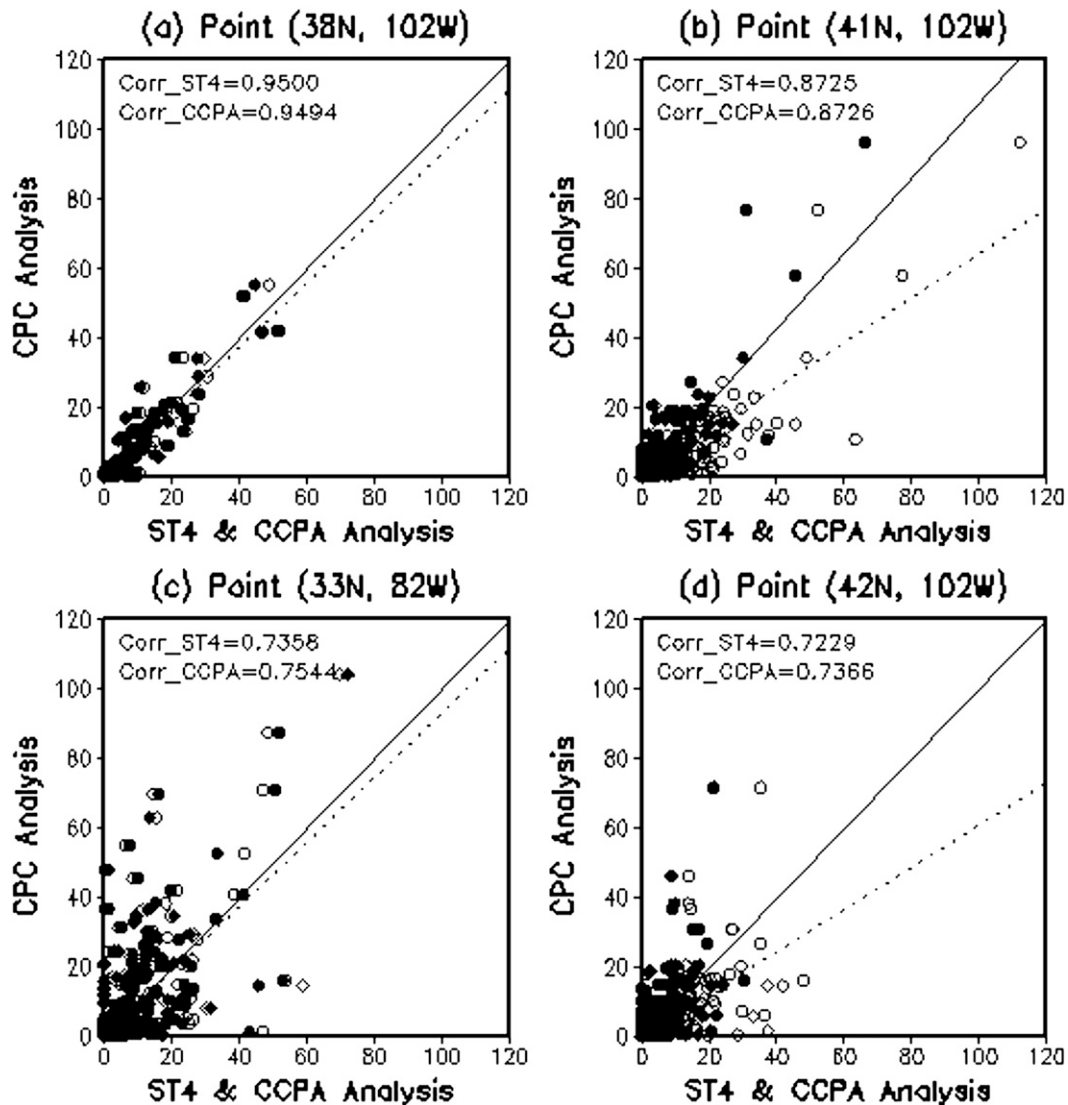
FIG. 15. Scatterplots of CPC and ST4 sample data (open circle) used for the estimation of regression coefficients for 1 Aug, at four grid points. The CCPA vs CPC data (closed circle), the diagonal (solid line), and the regression line (dotted line) are also shown.

REFERENCES

Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2003: Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average on high-resolution verification grids. *Wea. Forecasting,* **18,** 918–932, doi:10.1175/1520-0434(2003)018<0918:SOPFSS>2.0.CO;2.

Arkin, P. A., and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud cover over the Western Hemisphere during 1982–1984. *Mon. Wea. Rev.,* **115,** 51–74, doi:10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2.

Chen, F., and Coauthors, 2007: Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *J. Appl. Meteor. Climatol.,* **46,** 694–713, doi:10.1175/JAM2463.1.

Chen, M., R. E. Dickinson, X. Zeng, and A. N. Hahmann, 1996: Comparison of precipitation observed over the continental

United States to that simulated by a climate model. *J. Climate,* **9,** 930–951.

——, P. Xie, and J. E. Janowiak, 2002: Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeor.,* **3,** 249–266, doi:10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.

Cui, B., Z. Toth, Y. Zhu, and D. Hou, 2012: Bias correction for global ensemble forecast. *Wea. Forecasting,* **27,** 396–410, doi:10.1175/WAF-D-11-00011.1.

Dai, A., I. Y. Fung, and A. D. del Genio, 1997: Surface observed global land precipitation variations during 1900–1988. *J. Climate,* **10,** 2943–2962, doi:10.1175/1520-0442(1997)010<2943:SOGLPV>2.0.CO;2.

Ebert, E. E., and M. J. Manton, 1998: Performance of satellite rainfall estimation algorithms during TOGA COARE. *J. Atmos. Sci.,* **55,** 1537–1557, doi:10.1175/1520-0469(1998)055<1537:POSREA>2.0.CO;2.

Fekete, B. M., C. J. Vorosmarty, J. O. Roads, and C. J. Willmott, 2004: Uncertainties in precipitation and their impacts on runoff estimates. *J. Climate,* **17,** 294–304, doi:10.1175/1520-0442(2004)017<0294:UIPATI>2.0.CO;2.

Ferraro, R. R., 1997: Special sensor microwave image derived global rainfall estimates for climatological applications. *J. Geophys. Res.,* **102** (D14), 16 715–16 735, doi:10.1029/97JD01210.

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields.* Israel Program for Scientific Translations, 242 pp.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center ATLAS 7, 40 pp.

Huffman, G. L., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.,* **78,** 5–20, doi:10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Kummerow, C., and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.,* **39,** 1965–1982, doi:10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

Kursinski, A., and S. Mullen, 2008: Spatiotemporal variability of hourly precipitation over the eastern contiguous United States from Stage IV multisensory analysis. *J. Hydrometeor.,* **9,** 3–21, doi:10.1175/2007JHM856.1.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. Preprints, *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.]

McCollum, J., W. F. Krajewski, R. R. Ferraro, and M. B. Ba, 2002: Evaluation of biases of satellite rainfall estimation algorithms over the continental United States. *J. Appl. Meteor.,* **41,** 1065–1081, doi:10.1175/1520-0450(2002)041<1065:EOBOSR>2.0.CO;2.

Mesinger, F., 1996: Improvements in quantitative precipitation forecasting with the Eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.,* **77,** 2637–2649, doi:10.1175/1520-0477(1996)077<2637:IIQPFW>2.0.CO;2.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate,* **13,** 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.

Nijssen, B., G. M. O'Donnel, and D. P. Lettenmaier, 2001: Predicting the discharge of global rivers. *J. Climate,* **14,** 3307–3323, doi:10.1175/1520-0442(2001)014<3307:PTDOGR>2.0.CO;2.

NWS/AHPS, cited 2011: About the precipitation analysis pages. [Available online at http://water.weather.gov/precip/about.php.]

Onof, C., and H. S. Wheater, 1996: Analysis of the spatial coverage of British rainfall fields. *J. Hydrol.,* **176,** 97–113, doi:10.1016/0022-1694(95)02770-X.

Seo, D.-J., and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor.,* **3,** 93–111, doi:10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2.

Shepard, D., 1968: A two dimensional interpolation function for irregularly spaces data. *Proc. 23rd National Conf. of the Association for Computing Machinery,* Princeton, NJ, ACM, 517–524.

Spencer, R. W., 1993: Global oceanic precipitation from MSU during 1979–91 and comparisons to other climatologies. *J. Climate,* **6,** 1301–1326, doi:10.1175/1520-0442(1993)006<1301:GOPFTM>2.0.CO;2.

Susskind, J., P. Piraino, L. Rokkle, and A. Mehta, 1997: Characteristics of the TOVS pathfinder Path A dataset. *Bull. Amer. Meteor. Soc.,* **78,** 1449–1472, doi:10.1175/1520-0477(1997)078<1449:COTTPP>2.0.CO;2.

Tollerud, E., R. S. Collander, Y. Lin, and A. Loughe, 2005: On the performance, impact, and liabilities of automated precipitation gage screening algorithms. Preprints, *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction,* Washington, DC, Amer. Meteor. Soc., P1.42. [Available online at https://ams.confex.com/ams/WAFNWP34BC/techprogram/paper_95173.htm.]

Wilheit, T. J., A. T. C. Chang, and L. S. Chiu, 1991: Retrieval of monthly rainfall indices from microwave radiometric measurements using probability distribution functions. *J. Atmos. Oceanic Technol.,* **8,** 118–136, doi:10.1175/1520-0426(1991)008<0118:ROMRIF>2.0.CO;2.

Xie, P., and P. A. Arkin, 1996: Analysis of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858, doi:10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2.

——, and ——, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

——, and ——, 1998: Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. *J. Climate,* **11,** 137–164, doi:10.1175/1520-0442(1998)011<0137:GMPEFS>2.0.CO;2.

——, J. E. Janowiak, P. A. Arkin, R. Adler, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate,* **16,** 2197–2214, doi:10.1175/2769.1.

——, A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.,* **8,** 607–626, doi:10.1175/JHM583.1.

Zhao, Q., and Y. Jin, 2008: High-resolution radar data assimilation for Hurricane Isabel (2003) at landfall. *Bull. Amer. Meteor. Soc.,* **89,** 1355–1372, doi:10.1175/2008BAMS2562.1.

Zhu, Y., 2007: Objective evaluation of global precipitation forecast. *Extended Abstracts, Int. Symp. on Advances in Atmospheric Science and Information Technology,* Beijing, China, 3–8.