P1.6

PROCESSING DAILY RAIN-GAUGE PRECIPITATION DATA FOR THE AMERICAS AT THE NOAA CLIMATE PREDICTION CENTER

W. Shi[1,*], E. Yarosh[1], R.W. Higgins[2] and R. Joyce[1]

[1]RS Information Systems, Inc., McLean, VA 22102
[2]Climate Prediction Center/NCEP/NWS/NOAA, Camp Springs, MD 20746

1. INTRODUCTION

Accurate and complete estimates of precipitation are critical to a wide variety of problems ranging from understanding the water budget to improved monitoring and prediction of climate. Most areas of the globe are not adequately sampled, either by in situ or remote sensing. The conterminous U.S. is covered by a relatively dense array of in situ (hourly and daily) rain-gauge data. Precipitation over the U.S. can also be estimated using satellite data and radar data that is archived at high temporal and spatial resolution. These resources allow us to focus on improving the quality of the analysis of precipitation in the U.S. over a range of space and time scales.

Improving the analysis of precipitation requires careful consideration of the quality of the input observations. In general the quality control (QC) of gauge precipitation analyses has not been emphasized enough. In the recent years, we have developed an improved QC system of rain-gauge data for the U.S. at the NOAA/Climate Prediction Center (CPC). The CPC routinely produces quality controlled gauge-only precipitation analyses for the U.S. as part of its effort to monitor current and past conditions and to provide improved climate forecasts for the U.S.

The CPC has undertaken a comprehensive program to improve the analysis of gauge-based precipitation over the Americas on a range of space and time scales. The goal is to develop near-real-time and historical precipitation analyses for all of the Americas. The approach has been incremental, by first focusing on the U.S. and then by expanding this effort to include the remainder of North, Central and South America.

* *Corresponding author address*: Dr. Wei Shi, Climate Prediction Center/NCEP/NWS/NOAA, 5200 Auth Road, Room 605, Camp Springs, MD 20746, USA; e-mail: Wei.Shi@noaa.gov

In this presentation, we focus on our activities on the U.S. rain-gauge data QC, precipitation analyses and products. A flow chart that summarizes how U.S. rain-gauge data is processed at CPC is shown in Figure 1. The procedures of gauge data processing for the remainder of the Americas is more or less similar to this.

2. NEAR-REAL-TIME DAILY PRECIPITATION ANALYSIS

2.1 Background

The spatial coverage and accuracy of precipitation observations by “first order” stations in the U.S. is decreasing. Many in the climate community fear that the continued deterioration in surface observations will jeopardize our ability to perform real-time climate monitoring, forecasting and forecast verification. The problem is compounded by emerging requirements for daily (and even hourly) precipitation analyses.

There are several potential sources of precipitation data that could be used for precipitation analyses including 24-hr “first order” WMO GTS sites (near-real-time), 24-hr SHEF-encoded precipitation reports received via AFOS from the River Forecast Centers (near-real-time), hourly GOES/DCP and CADAS precipitation reports (near-real-time), hourly and 24-hr NCDC cooperative reports (non-real-time), and perhaps many other sources (e.g. SNOTEL data, HADS data). In December of 1996 the CPC organized a Precipitation Working Group to examine this problem in detail, in particular to inventory potential sources of suitable near-real-time precipitation data and to make appropriate intercomparisons to address issues of spatial coverage, reliability, and availability.

The group recommended the development of a near-real-time “U.S. Precipitation Quality Control (QC) System and Analysis” whose input was raingauge data. Such a system was built in early 1997 and has been undergoing continuous development and improvement since that time. A fully automated script was implemented to control data acquisition, run the QC (see section 3.0), prepare the analysis, archive the data and disseminate analysis products on the CPC Web Site (http://www.cpc.ncep.noaa.gov). Current products include daily accumulated precipitation, monthly and seasonal precipitation monitoring products, forecast verification products and drought / flood potential products. The current suite of precipitation products for the U.S. also supports ongoing efforts in the CPC to deliver Climate Services, including the US Threats Assessment, a US Drought Forecast System, the Palmer Drought Index, and a Soil Moisture Forecasting System. Our products are also used by many external research projects, including the Land Data Assimilation System (LDAS), and the NCEP Regional Reanalysis Project.

2.2 Characteristics of the Daily Analysis

The daily analyses are gridded at a horizontal resolution of (lat,lon)=(0.25˚x0.25˚) over the domain 140° W - 60° W, 20° N - 60° N using a Cressman (1959) scheme with modifications (Glahn et al. 1985; Charba et al. 1992). An intercomparison of precipitation analyses produced by Cressman (Cressmann 1959), Barnes (Barnes 1964), Shepard (Shepard 1968) and OI (Gandin 1963) schemes (not shown) revealed only minor differences in the analyses, presumably due to sufficient data density over the U.S. The input dataset for the near-real-time analysis is the CPC Cooperative dataset (24-hr “first order” WMO GTS sites and 24-hr SHEF-encoded precipitation reports received via AFOS from the River Forecast Centers). The analysis on Day 1 is valid for the 24-hour window from 1200Z on day 0 to 1200Z on day 1; a typical station distribution and daily precipitation analysis are shown in Figs. 2a and 2b respectively.

Several types of QC are currently applied to the gauge data: (1) A “duplicate station check” which eliminates duplicates and key punch errors from the rain-gauge reports; (2) A “buddy check” to eliminate extreme values; (3) a standard deviation check, which compares the daily rain-gauge data against a gridded daily climatology; and (4) NEXRAD radar QC of the gauge data to eliminate spurious zeroes; some details of the first 3 of these are discussed in section 3.1 while the fourth one is discussed in section 3.2. In addition, a fifth type of QC is included in the operational analysis involving satellite QC of the NEXRAD data (see section 3.3). All QC flags are inserted back into the gauge data archive for future reference. Station dictionaries are updated routinely to ensure proper elimination of duplicates as part of the QC procedure. It is anticipated that these QC initiatives will ultimately benefit radar-only and multi-sensor analyses.

Currently the daily analysis is available within ~16 hrs of real time.

3. QUALITY CONTROL INITIATIVES

While the raw rain-gauge datasets are undergoing continuous development and improvement, there are nevertheless many problems with the resulting precipitation analysis despite the QC steps already in place; these problems are due to a combination of instrument error, bad rain-gauge reports that remain undetected and errors in the analysis scheme. High resolution radar and satellite-derived precipitation estimates offer potential for additional improvements to the QC of rain-gauge data.

3.1 Standard Quality Control

There are three standard QC steps currently applied in our analysis system: (1) A “duplicate station check” which eliminates duplicates and key punch errors from the rain-gauge reports; (2) A “buddy check” to eliminate extreme values from the dataset and (3) a standard deviation check, which compares the daily rain-gauge data against a gridded daily climatology. The “buddy check” examines the absolute value of the difference between the current station and all stations within a one-degree grid box. If more than 50% exceed a specified threshold, then the current station is tossed. For the standard deviation check we currently use a daily climatology derived from the Unified Rain-gauge Dataset (Shi et al. 2002). The observations are compared to the nearest gridpoint value from the climatology. The current observation must be within 5 standard deviations (10 for hurricane events) of the daily climatology.

3.2 Radar Quality Control

One serious problem in the CPC Cooperative Dataset is the number of incorrect reports of zero precipitation in the 24-hour SHEF-encoded “RFC” precipitation data (~6000-7000 reports daily). This problem is illustrated in Fig. 3, which shows the number of stations in the southeastern U.S. with no precipitation, less than 2 inches of precipitation, less than 4 inches of precipitation and greater than ten inches of precipitation for the period January-

March 1998. The 1997/1998 El Niño event was characterized by heavy rain in the southeast US during this period, so it is clear that stations reporting no precipitation are in error.

While reports with erroneous large values are easy to detect and eliminate (i.e. via extreme value checks, buddy checks, etc), reports with erroneous zero (or small) values are hard to detect. A solution (currently implemented in our analysis system) is to eliminate spurious zeros from the rain-gauge data prior to analysis using hourly radar estimates of precipitation. One advantage of such a QC step is that it counters the tendency for underestimating the observed rainfall as is typically the case in gridded analyses.

The approach is as follows:

(1) Accumulate hourly radar precipitation estimates to 24-hr values (1200Z-1200Z).

(2) Compare all daily raingauge reports against the nearest gridpoint in the 24-hr radar estimate of precipitation (technically similar to the procedure used in our standard deviation check). The high horizontal resolution of the radar data (4-km) works to our advantage since it ensures that the radar estimate and raingauge report are reasonably close to each other (i.e. within about 2-km).

(3) Eliminate rain-gauge reports below suitable thresholds in the radar estimates. A careful examination of the bias in the radar data (Fig. 4) suggested that a threshold of 2 mm day^{-1} was suitable; this value is currently implemented in our analysis system.

(4) Insert the QC flags back into the CPC Cooperative Dataset. The QC information can be used to investigate the origin of these reports.

3.3 Satellite Quality Control

This QC step incorporates satellite based estimates of precipitation into the QC System. Radar estimates of precipitation are biased due to radar-radar calibration differences (when a single Z-R relationship is used), differences in precipitation rate between the radar scan level and the ground, and anomalous propagation of the beam (Fig. 4 shows two examples). In the past the QC of radar data has often been performed with information from other sensors (i.e. rain-gauge, satellite) and a number of investigators continue to examine this (Smith et al. 1997; Fulton et al. 1997; Seo et al. 1997; Ahnert et al. 1986; Office of Hydrology 1992). Recently, we developed an algorithm that uses satellite data to remove bias in radar estimates of precipitation before other QC steps are invoked. This algorithm was developed by porting software / experience from an earlier study by Joyce et al. (1998) into our QC system.

Basically, the algorithm uses high resolution GOES-IR data (currently ½ hour on a ½^{o} x ½^{o} latitude-longitude grid) to screen out heavy hourly radar precipitation estimates when collocated IR temperatures before, during and after the hour in question are warmer than a set threshold. The closest 4 km GOES 8 or GOES 10 IR pixel is collocated with the midpoint of each 0.1˚ (lat, lon) hourly radar precipitation estimate using 30 minute IR images. The coldest pixel is determined from all three IR images for 3 spatial extensions: (1) the exact IR pixel collocated to the radar estimate; (2) all pixels within 25 km of the collocation; and (3) all pixels within 50 km of the collocation. The statistics are further separated by stratifying the collocations into categories of radar precipitation from 0 to > 25 mm hr^{-1} for classes every 5 mm hr^{-1}. The mean of the coldest IR pixel found is computed for all radar rainfall cases and the three spatial extensions. Standard deviation of the coldest IR pixel about the mean for each rainfall class is then computed, from which frequency maps of occurrences in classes of 0.5 sigma from the mean are computed. The mean coldest IR temperature (within 25 km of the radar estimate) for the class of no radar precipitation was the warmest at 262.9 K (Fig. 5a) with a standard deviation of 22.2 K (Fig. 5b). This quickly drops to 230.2 K with a standard deviation of 15.9 K for a radar rainfall of 0-5 mm hr^{-1}. For cases of radar rainfall > 25 mm hr^{-1} the mean coldest IR temperature was 208.2 K with a standard deviation of 8.2 K.

The distribution of the cases of coldest IR pixel about the mean coldest IR pixel (Figs. 6a and 6b) reveals that for 80% of the no radar rainfall cases (May 1999) the coldest IR pixel is warmer than 240.0 K, or one standard deviation below the mean of 262.9 K. In almost 100% of the cases of radar rainfall greater than 25 mm hr^{-1}, the coldest IR pixel is colder than 237.0 K, or 3.5 standard deviations above the mean (Figs. 6c and 6d), 240.0 K for rainfall greater than 20 (but less than 25) mm hr^{-1} . This gives considerable utility in eliminating incorrect radar estimates.

From the statistics previously described, an IR temperature threshold of 3.5 standard deviations warmer than the mean coldest IR pixel (within 25 km of the radar estimate) is used for the radar rainfall classes. If the coldest collocated IR pixel within 25 km of the radar location for the image before, during, and after the radar rainfall estimate is not colder than this threshold, then the estimate is regarded as false. Results for screening radar rainfall cases in this way have been very encouraging. Moderate and heavy radar rainfall cases are screened very well in virtually all cases. Light radar rainfall cases (0-5 mm hr^{-1}) are the most difficult to screen.

4. PRECIPITATION ANALYSIS PRODUCTS AND APPLICATIONS

The near-real-time daily precipitation analysis has been used to develop a number of additional products and applications. Some of these are described in the following subsections.

4.1 Products

Precipitation analysis products developed at CPC include a daily precipitation analysis and associated station map; precipitation monitoring maps that highlight hydrologic anomalies over the conterminous U.S. for the previous 30 days and 90 days (Fig. 7), a series of products that verify precipitation forecasts from the operational MRF and ensembles, and a drought/flood potential product that highlights expected changes in observed precipitation anomalies.

All of these products are disseminated on a daily basis via the CPC Web Site (http://www.cpc.ncep.noaa.gov/products/precip/realtime). These products are undergoing continuous development and improvement and have benefitted significantly from the extended QC initiatives described in section 3.

We have applied our rain-gauge data QC system and analysis scheme with adjustments to the data collected from other countries of the Americas. Several gridded daily analysis products for the Americas are currently available:

(1) Near-Real Time Analyses

--United States (1996-present; daily)
--South America (1999-present; daily)
–US_Mexico Merged (2001-present; daily)

(2) Historical Reanalysis

--United States (1948-1998; daily)
--United States (1948-1999; hourly)
--Mexico (1948-present; daily)
--Brazil (1961-1997; daily)
--Canada (1961-1999; daily)

4.2 Applications

The near-real-time precipitation analysis is used by several other CPC projects, including:

(1) The U.S. National Threats Assessment http://www.cpc.ncep.noaa.gov/products/expert_assessment/threats.html

(2) The U.S. Drought Assessment http://www.cpc.ncep.noaa.gov/products/expert_assessment/drought_assessment.html

(3) The Palmer Drought Index http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/drought.html

(4) The CPC Soil Moisture Forecast project http://www.cpc.ncep.noaa.gov/soilmst/forecasts.html

The analyses are also used by external projects, including:

(1) The Land Data Assimilation System (LDAS) http://ldas.gsfc.nasa.gov/index.shtml

(2) The NCEP Regional Reanalysis http://wesley.ncep.noaa.gov/reanalysis.html

In the case of LDAS, an early analysis (based only on RFC data) is provided on a daily basis.

As a result of these collaborations, we are frequently required to adapt the content, design and availability of our precipitation products as well as to respond to changing user requirements.

References

Ahnert, P., W. Krajewski, and E. Johnson, 1986: Kalman filter estimation of radar-rainfall field bias. *Preprints 23rd Conference on Radar Meteorology*, Snowmass, Co., Amer. Meteor. Soc., JP33-JP37.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. App. Met.*, **3**, 396-409.

Charba, J. P., A. W. Harrell III and A. C. Lackner III, 1992: A monthly precipitation amount climatology derived from published atlas maps: Development of a digital data base. *TDL Office Note 92-7*, NOAA, U.S. Department of Commerce, 20 pp.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367-374.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, and D. A. Miller, 1997: The WSR-88D rainfall algorithm. *Wea. Forecasting*

Gandin, L. S., 1963: Objective analysis of meteo-rological fields. Gidrometeor. Isdat., Leningrad. [Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.]

Glahn, H. R., T. L. Chambers, W. S. Richardson, and H. P. Perrotti, 1985: Objective map analysis for the local AFOS MOS Program. *NOAA Technical Memorandum NWS TDL 75*, NOAA, U. S. Department of Commerce, 34 pp.

Joyce, R., P. Arkin and W. Higgins, 1998: Estimation of orographic precipitation over western North America using geostationary satellite observations. 9th Conference on Satellite Meteorology and Oceanography, 25-29 May, Paris, France. pp 61-64.

Office of Hydrology, 1992: Stage II precipitation processing. National Weather Service, NOAA, U.S. Department of Commerce, 16 pp.

Seo, D.-J., R. Fulton, J. Breidenbach, and E. Johnson, 1997: Real-time estimation of mean field bias in radar rainfall data -- a review of current techniques in the National Weather Service and a proposed alternative. *J. Atmos. Ocean. Technol.*

Shepard, D., 1968: A two-dimensional interpolation function for irregularly spaced data. ACM National Conference. 512-524.

Shi, W., R.W. Higgins and E. Yarosh, 2002: A Unified Raingauge Dataset and Multi-year Daily Precipitation Reanalysis for the United States. (*Submitted to Journal of Geophysical Research*)

Smith, J. A., M. L. Baeck, and M. Steiner, 1997: Final report: Hydrometeorological assessments of the NEXRAD rainfall algorithms. Submitted to Office of Hydrology, National Weather Service, NOAA, U.S. Department of Commerce.

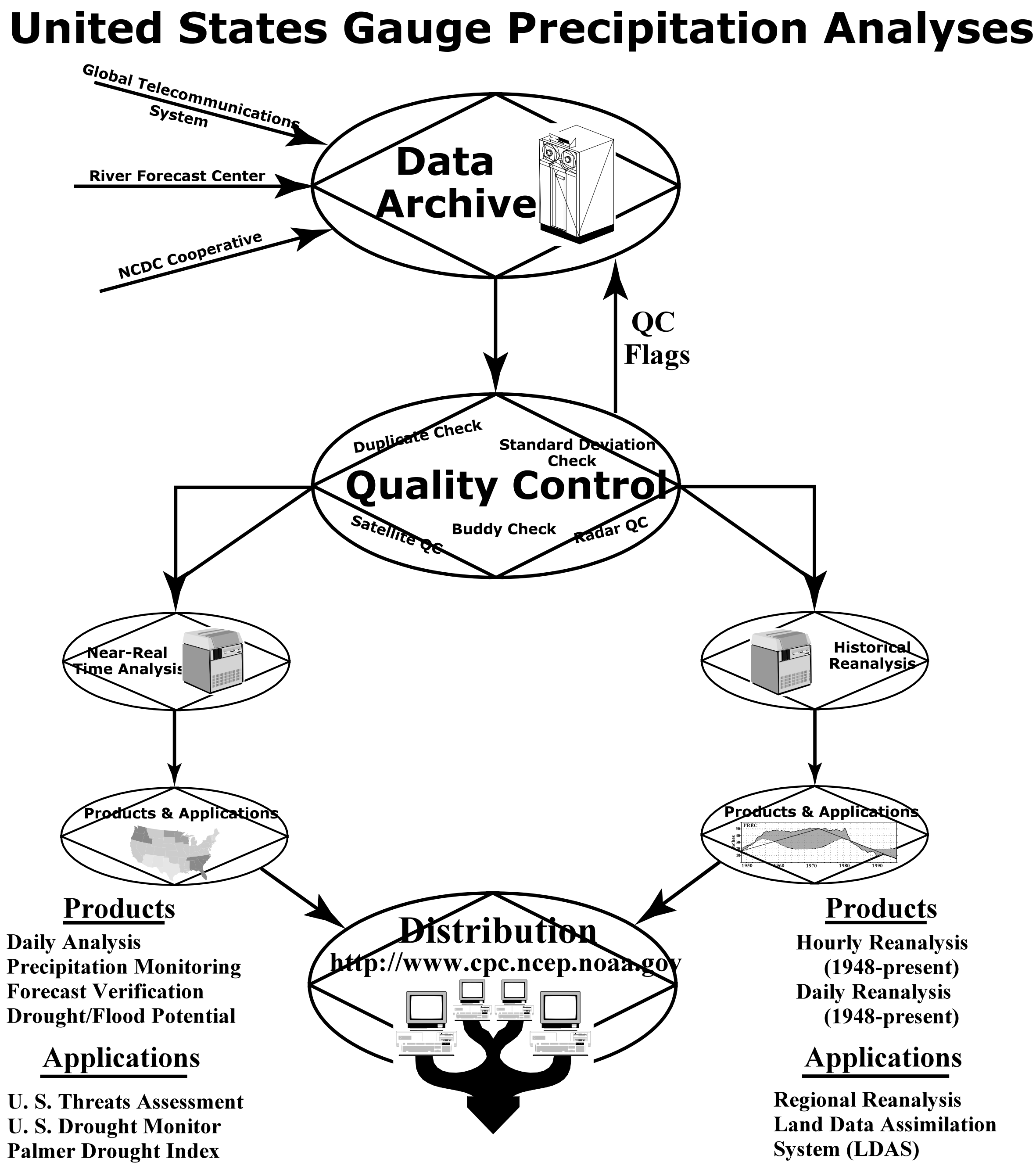

Figure 1. Flow chart summarizing how U.S. raingauge data is processed at the Climate Prediction Center.

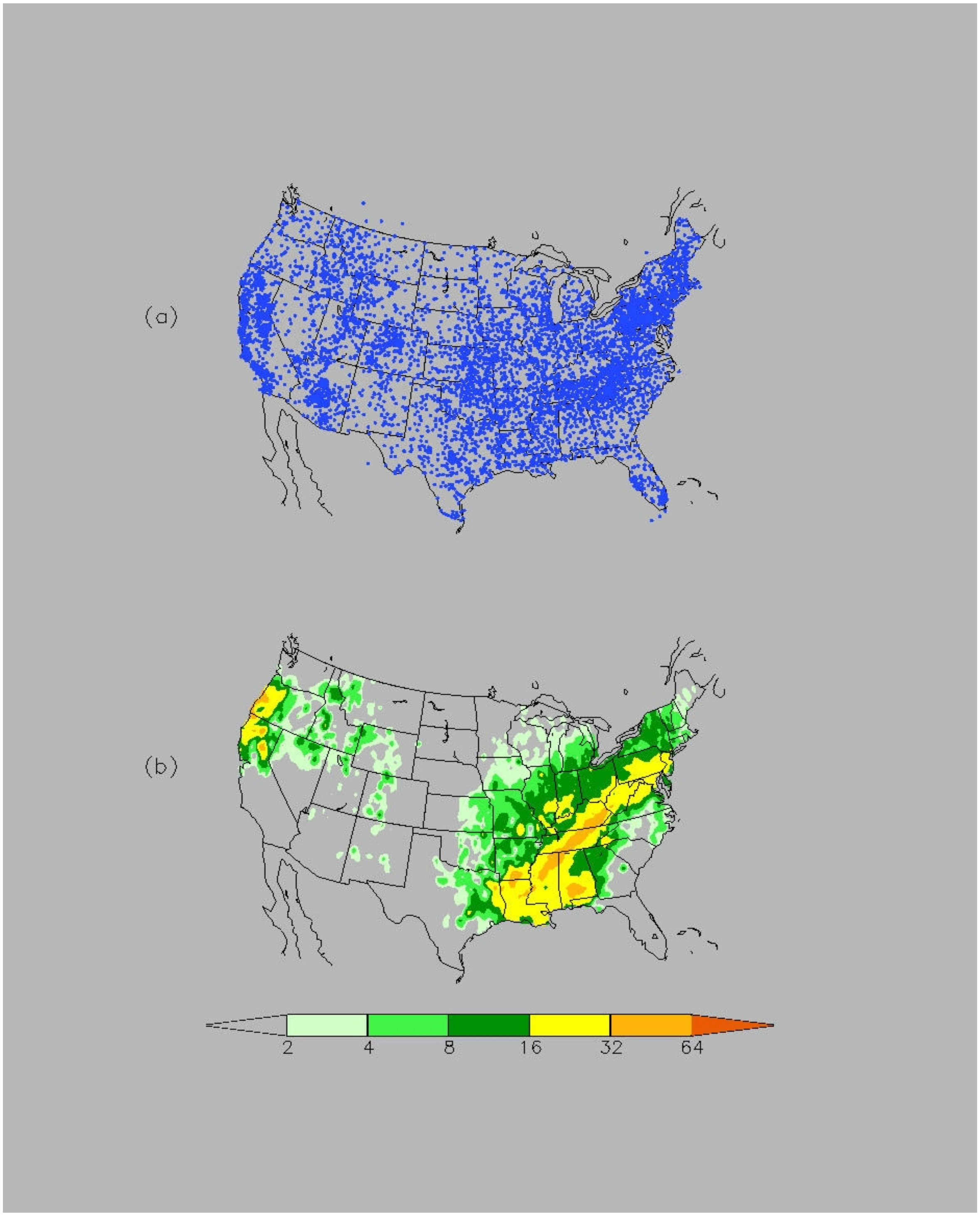

Figure 2. (a) Typical station distribution for daily reporting stations in the United States from the CPC Cooperative Dataset. The dataset consists of reports gathered by the River Forecast Centers (~6000-7000 daily reports) and the Climate Anomaly Data Base (~400-500 daily reports). (b) Daily precipitation analysis (Units: mm) based on 24-hr accumulations for the period from 1200Z December 13, 2000 - 1200Z December 14, 2000. The analysis is gridded at a horizontal resolution of 0.25 degrees.

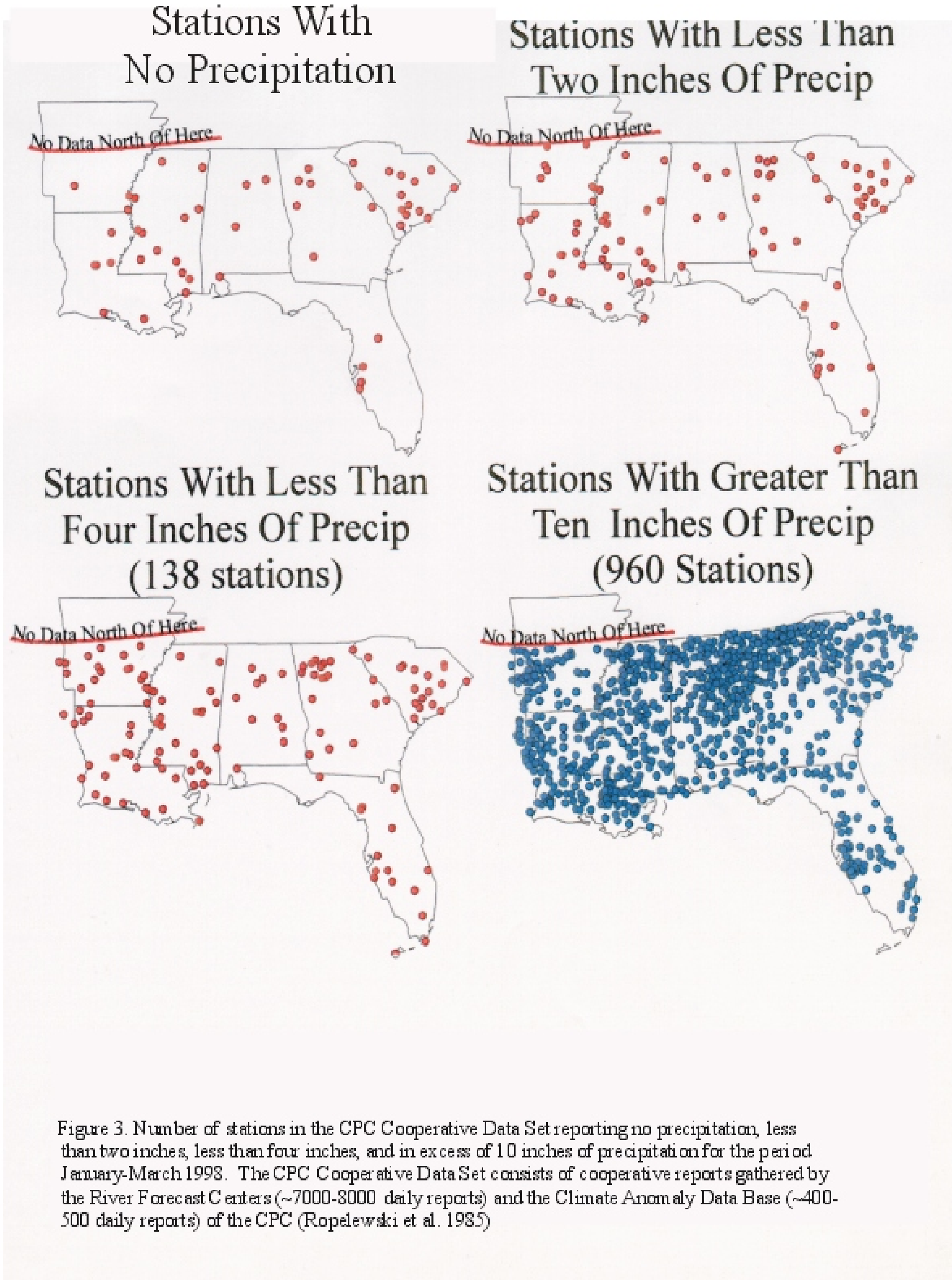

Figure 3. Number of stations in the CPC Cooperative Dataset reporting no precipitation, less than two inches, less than four inches, and in excess of 10 inches of precipitation for the period January-March 1998.

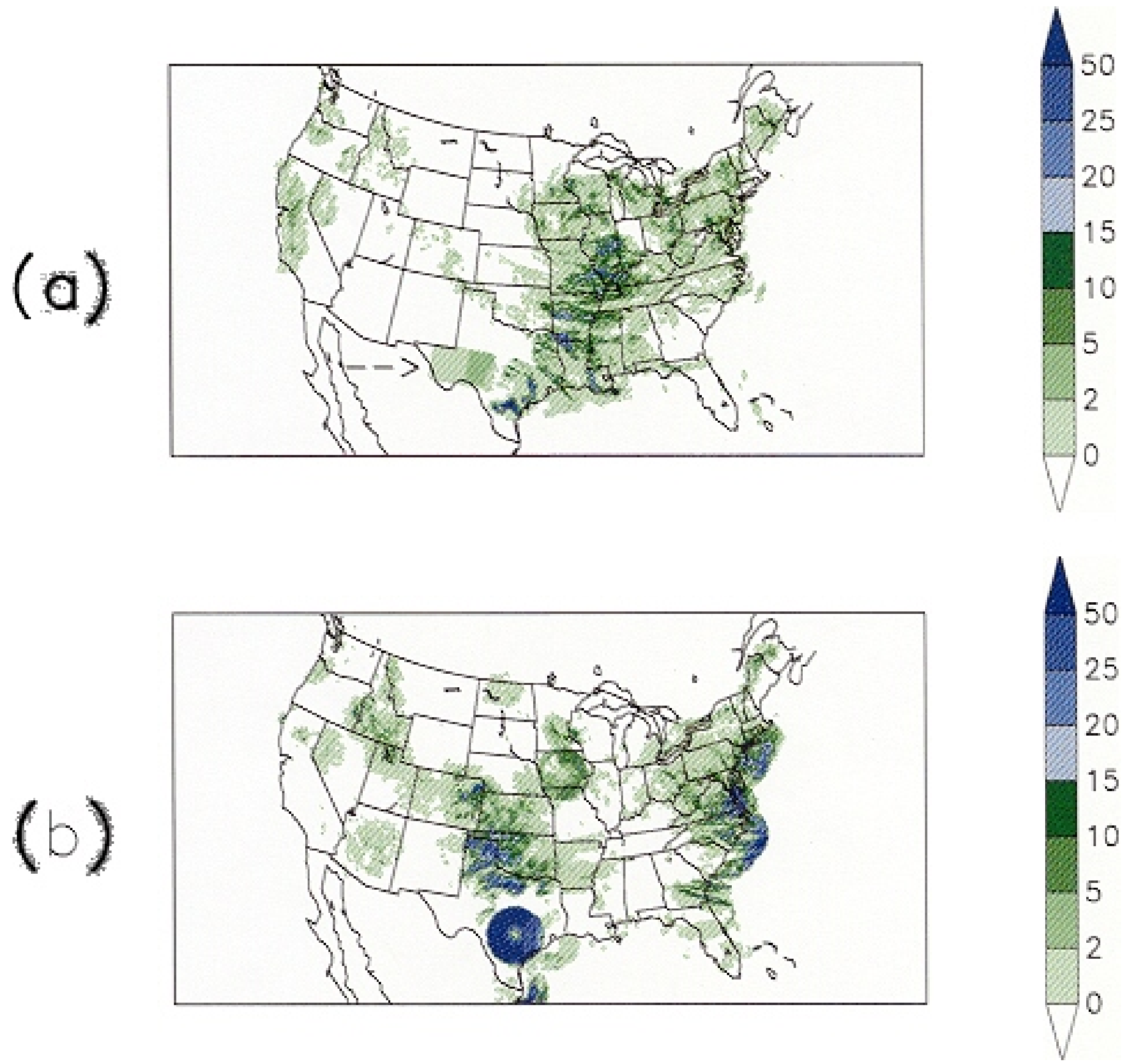

Figure 4. Typical examples of bias in radar data (Units: mm). For these examples the hourly biased data were used to produce 24-hr accumulations on a grid with a horizontal resolution of 4-km. The 24-hr accumulations are valid for (a) 1200Z May 2, 1998 - 1200Z May 3, 1998 and (b) 1200Z May 8, 1998 - 1200 Z May 9, 1998.

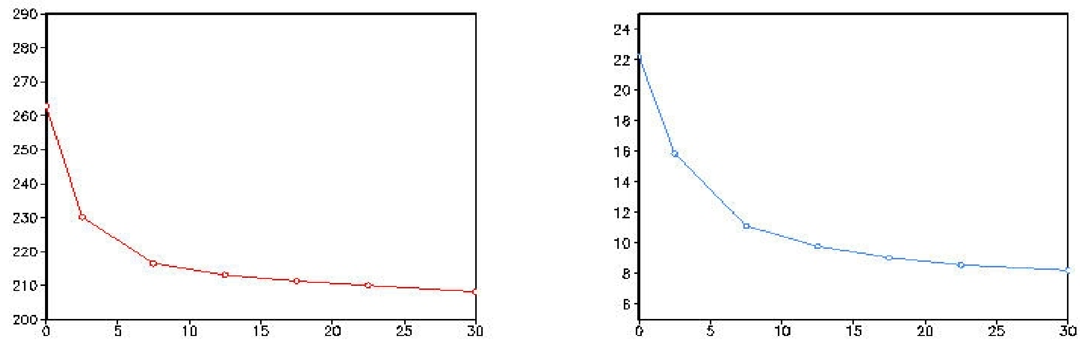

Figure 5 a. Mean of coldest satellite IR pixel (K) found within 0.25° of 0.1° Hourly Radar Precip (mm/hr), abscissa is hourly radar rainfall, May 3–31, 1999. b. Standard deviations in sigma (K).

Figure 5. (a) Mean of the coldest satellite IR pixel (K) found within 0.25˚ of the 0.1˚ hourly radar precipitation estimates (units: mm hr^{-1}). Abscissa is the hourly radar precipitation estimate, based on May 1999. (b) Standard deviation of (a).

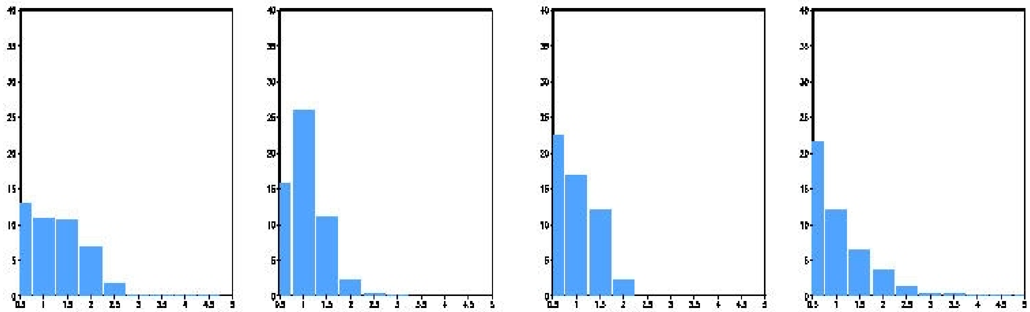

Figure 6 a. Distribution (%) of cases of coldest IR pixel (within 0.25° of 0.1° Hourly Radar Precip) in standard deviations (sigma = 22.2 K) colder than the coldest pixel mean of 262.9 K for collocated zero radar rainfall. b. Same as (a) except for warmer than the coldest pixel mean. c. Distribution of cases of coldest IR pixel in standard deviations (sigma = 8.2 K) colder than the coldest pixel mean of 208.2 K for collocated radar rainfall > 25 mm/hour d. Same as (c) except for warmer than the coldest pixel mean.

Figure 6. (a) Distribution (%) of cases of coldest IR pixel (within 0.25° of the 0.1° hourly radar precipitation estimates) in standard deviations (sigma=22.2K) colder than the mean coldest pixel of 262.9 K for collocated zero radar rainfall. (b) Same as (a) except warmer than the mean coldest pixel. (c) Distribution of cases of coldest IR pixel in standard deviations (sigma=8.2K) colder than the mean coldest pixel of 208.2 K for collocated radar rainfall > 25 mm hr^{-1}. Same as (c) except warmer than the mean coldest pixel.

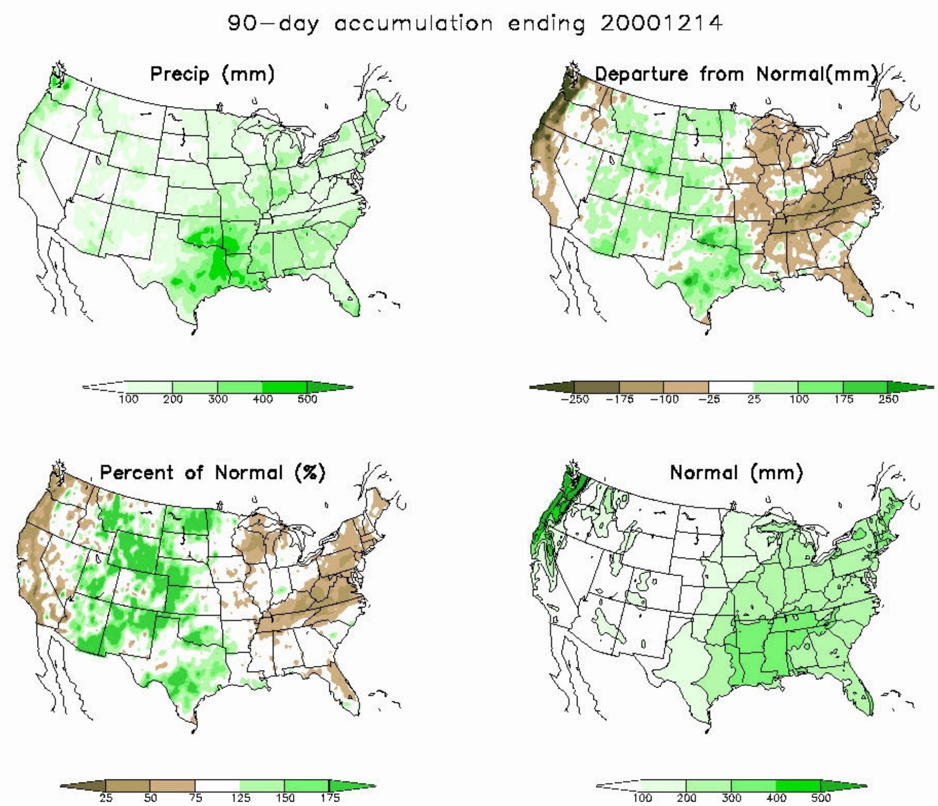

Figure 7. Observed precipitation (upper left), departure from normal (upper right), percent of normal (lower left) and normal precipitation (lower right) for the 90-day period ending 30 June, 1998. Results are based on CPC's daily precipitation analysis which is produced by the U.S. Precipitation QC System and Analysis.