3.3 Quality Control of Daily Precipitation Reports at NOAA/CPC

Mingyue Chen*, and Pingping Xie, and
CPC Quality Control Working Group
Climate Prediction Center (CPC)/NCEP/NOAA
*Email: Mingyue.Chen@noaa.gov

1. Introduction

Despite the rapid progress achieved in the last two decades in estimating precipitation from radar and satellite observations and in simulating precipitation through numerical models, gauge observations continue to play a critical role in documenting the characteristics of precipitation over global land areas (Huffman et al. 1997; Xie and Arkin 1997; Adler et al. 2003). At the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC), gauge observations have long been utilized for various applications in climate research and operations (Kumar et al. 2007). One important source of information used by NOAA/CPC and many other meteorological agencies for weather and climate applications is the gauge-observed precipitation reports transmitted through the Global Telecommunications System (GTS). Together with station reports from other national and international sources, the GTS gauge data are used to monitor and assess the climate at station locations (Ropelewski and Halpert 1987) and as inputs to define analyzed fields of global and regional precipitation (Xie et al. 1996).

Quality problems, however, exist in the GTS gauge reports due to human mistakes and transmission errors occurred through the process. These quality problems, often yielding inaccurate or even unrealistic definition of precipitation, may severely compromise our ability to monitor and document the climate and its variability. Quality control (QC) of the daily precipitation reports has been a challenging task due primarily to the combined effects of the large natural variability of precipitation and the lack of independent information of precipitation with reasonable time / space resolution. Recent progress in satellite-based remote sensing and numerical model techniques has yielded global precipitation fields with reasonable quantitative accuracy at a high resolution, making it possible to perform QC for daily precipitation reports over the global land areas.

At NOAA's Climate Prediction Center, a project was launched to develop a set of automated procedures to perform quality control for the GTS daily precipitation reports through comparisons with historical gauge records, concurrent observations at nearby stations, satellite estimates and numerical model forecasts. The objective of this article is to describe the QC procedures, examination tests results and the applications of the quality controlled station data.

2. The GTS Gauge Reports and the QC Procedures

2.1 The GTS Daily Gauge Reports

The Global Telecommunication System (GTS) is the coordinated global system of telecommunication facilities and arrangements for the rapid collection, exchange, and distribution of observed and processed meteorological information within the framework of the World Weather Watch (WWW). Station reports of precipitation, together with those of many other physical variables, are exchanged routinely among the World Meteorological Organization (WMO) member countries through the GTS network. At NOAA / CPC, these precipitation reports are received and processed to form a database of GTS gauge-based daily precipitation. Starting from October 1977, the GTS daily gauge database is updated on a real-time basis. On average, daily reports are available from about 6000 GTS stations (fig.1). The GTS gauge network is relatively dense over United States, Western Europe, and east coasts of Australia and China, while it is very sparse over several regions including equatorial Africa and Amazon.

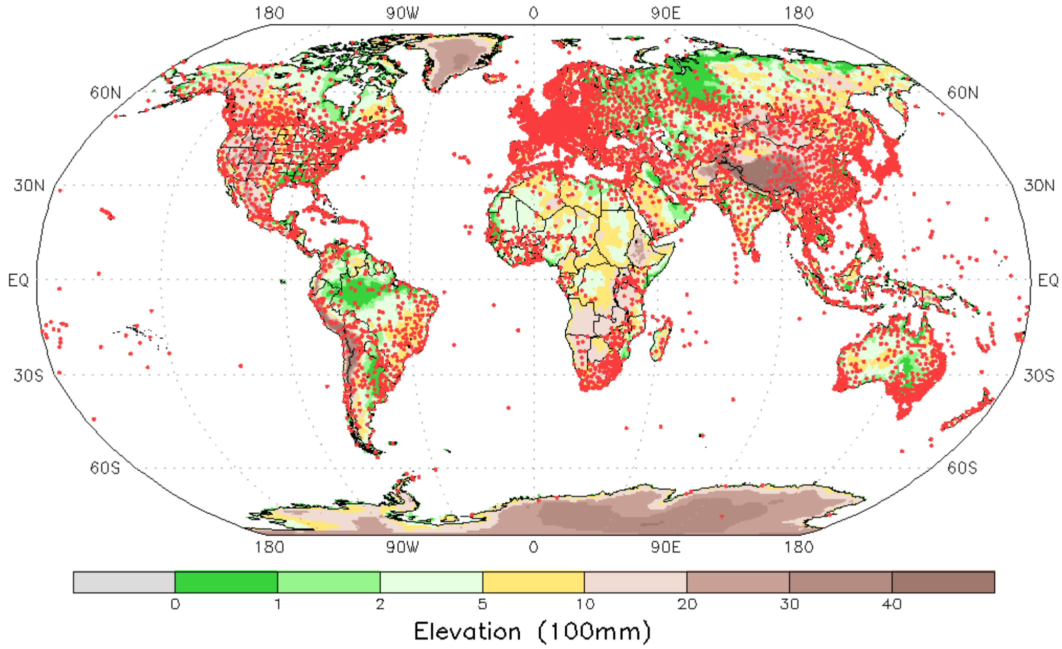

Figure 1: GTS gauge locations for July 1, 2005, plotted on the background of elevation.

2.2 Satellite Estimates and Numerical Model Forecasts

While quality problems may occur with GTS daily reports of any precipitation amounts, our quality control (QC) procedures are designed to target two categories of situations most often observed in daily operations: reports with '0' and extremely large values. Preliminary inspection of the raw GTS daily data found that some of the reports with '0' value are sent out as missing code while reports with large rainfall amounts may be a result of human mistakes or instrumental / transmission errors in the processes of observation, recording, data transmission and processing.

High-resolution precipitation fields derived from satellite observations and numerical model forecasts are employed to examine the GTS daily precipitation reports.

The satellite-based precipitation estimates are produced by the CPC Morphing technique (CMORPH, Joyce et al. 2004). Gridded fields of precipitation is defined on a resolution of 30-minutes and 8 km x 8km by interpolating the microwave estimates of instantaneous rain rates in the time-space domain through the advection vectors of the cloud / precipitation systems computed from consecutive infrared images. In this study, CMORPH data of original resolution is binned into 0.25^{o}lat/lon grid boxes over a 24-hour accumulation period consistent with that for the GTS daily station data.

The numerical model-based precipitation fields used here are those of NOAA National Environmental Prediction Center (NCEP) Global Forecast System (GFS). GFS six-hourly precipitation forecasts are post-processed onto a global grid of 1.0^{o}lat/lon and used as independent information to examine the GTS daily reports. Previous research showed that satellite estimates provide precipitation coverage of reliable quality over tropical and sub-tropical areas while model forecasts perform well over high latitudes especially during cold seasons (Ebert et al. 2007).

2.3 Quality Control Procedure for Reports of ‘0’ Value

Quality control for the GTS daily reports of ‘0’ values is performed through calculating the probability of the target reports with quality problems through comparisons with historical records, current observations at nearby stations and precipitation fields from the satellite estimates and model forecasts.

First, daily precipitation reports for each station are examined for a 26-year period from 1979 to 2004 to count the number of 3-month seasons during which no daily raining events is reported. Stations with five or more no-raining seasons during the 26-year period are identified and included into a ‘black list’ with a probability (P_1) of ‘blackness’ defined as the ratio of the number of seasons with no raining events to the total number of seasons (26) over the data period.

The second component of the QC involves comparisons with the historical records at the target station. Probability of raining (P_2) at the target station location on the target date is computed as the percentage of days with rainfall larger than 1 mm over a 15-day moving window centering at the target date over the 26-year period.

The daily precipitation reports are then compared against those at stations within 300 km of the target gauge (buddy check). The probability of a ‘0’ value GTS daily report being suspicious (P_3) is computed as the percentage of the nearby stations reporting rainfall larger than 1 mm/day.

Additional checks are performed using the satellite estimates and model forecasts of precipitation. Percentage of grid boxes with rainfall amount of 1 mm/day or larger is computed for the CMORPH estimates (P_4) and GFS forecasts (P_5), respectively, over an area of 4^o lat x 4^olon centering at the target station and used as an index of likelihood that the rainfall over the target station is raining.

Probability of a ‘0’ value GTS report being erroneous ($\mathbf{P_{no\text{-}rain}}$) is defined as the weighted mean of the five individual probabilities described above. The weighting for probabilities from individual examinations is defined based on the relative reliability of the independent information used. For example, the weighting coefficient for the CMORPH satellite estimates is a function of the serial correlation between the satellite estimates and the station precipitation calculated from historical records.

A risk index is finally assigned to each GTS report with ‘0’ value by multiplying 10 to the weighted probability ($\mathbf{P_{no\text{-}rain}}$). Ranging from 0 to 10, a risk index of 0 indicates a clean report of 100% confidence, while an index of 5 or higher implies high possibility of quality problems.

2.4 Quality Control Procedure for Reports of Large Rainfall Values

Quality control (QC) for GTS daily precipitation reports with large values is conducted in a similar way as that for reports with ‘0’ value. Daily reports with rainfall amount of 25 mm or larger are identified and subjected to the QC procedure. Instead of the probability, ratio of the reported station rainfall amount to a reference large precipitation amount is calculated to define the risk index.

First, the ratio is computed against the reference large precipitation amount determined from 1) historical records at the station; 2) concurrent observations at nearby stations; 3) CMORPH satellite estimates; and 4) GFS model forecasts. In the comparison with the historical records, the reference is defined as the 95 percentile value of the daily records over a 15-day moving window centering at the target date over the 26-year recording period. The reference is the 95 percentile value of precipitation data over stations within 300 km in the buddy check and values at grid boxes within over a 4^o lat/lon regions centering at the target station in comparisons with satellite estimates and model forecasts, respectively.

Weighted mean of the individual ratios is then calculated using the same weights as those used in the QC for ‘0’ value reports. A risk index for the GTS daily precipitation report is assigned for each daily precipitation report with rainfall amount of 25mm or higher based on the weighted mean ratio (fig.2).

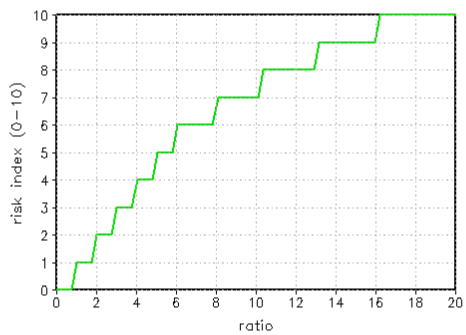

Figure 2: Definition of the risk index for GTS reports of large values based on the weighted mean of the ratio between the daily report amount and reference large values determined from individual sources.

The quality control procedures described above represent an improvement to those developed by Higgins et al. (2000) for the daily precipitation reports over the United States and Mexico. Modifications of QC procedures are implemented to account for different situations for the QC of GTS observations from sparse gauge networks over regions with no radar coverage available.

3. Simulation Tests

Simulation tests are conducted to examine the performance of the QC procedures described in section 2. First, 10% of daily precipitation reports with non-zero values are selected randomly over a four-year period from 2004 to 2007. Rainfall amount of the selected daily reports is then assigned as ‘0’. Quality control is then performed for the station reports with erroneous ‘0’ value through comparisons with historical records, concurrent observations at near by stations, satellite estimates and model forecasts.

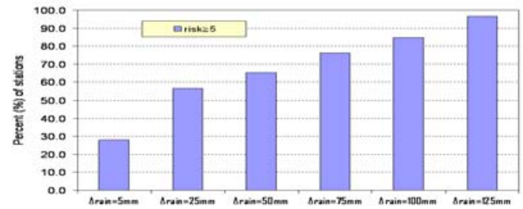

Figure 3: Detection rate of erroneous GTS reports of ‘0’ value as a function of the real rainfall amount.

Our QC procedure works quite well in detecting simulated erroneous GTS reports with ‘0’ value. The detection rate increases with the real rainfall amount at the station. More than half of the wrong reports are picked out over stations with rainfall amount of 25mm or higher.

We then performed the simulation tests for the QC procedure for GTS reports of large values. Similarly, daily precipitation reports at 10% randomly selected stations are re-assigned by adding extra amount of rainfall to the original values. Fig.4 presents the percentage of daily reports with erroneous large rainfall amount detected by our QC system, plotted as a function of real rainfall amount and rainfall amount added on.

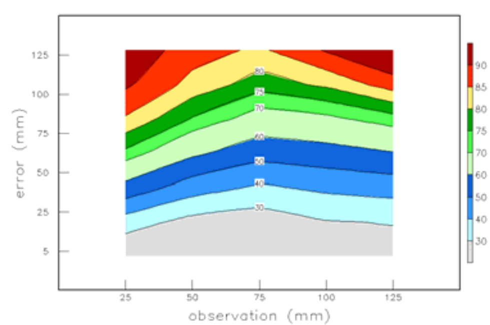

Figure 4: Detection rate of erroneous GTS reports of large value as a function of the real rainfall amount (x-axis) and erroneous amount added on (Y-axis).

Our QC system is capable of detecting the erroneous GTS daily reports with large values reasonable well. Overall, the detection rate is higher for reports of larger error amount. About half of the erroneous reports can be identified for cases with error of 50mm or larger.

4. Applications

The objective procedures described in this article are applied to perform quality control for daily GTS precipitation reports for a 6-year period from 2003 to 2007. Figure 5 shows an example of the quality control results over the globe for November 26, 2005, while figure 6 presents the precipitation fields from gauge data, CMORPH satellite estimates and GFS model forecasts over Africa for the same day.

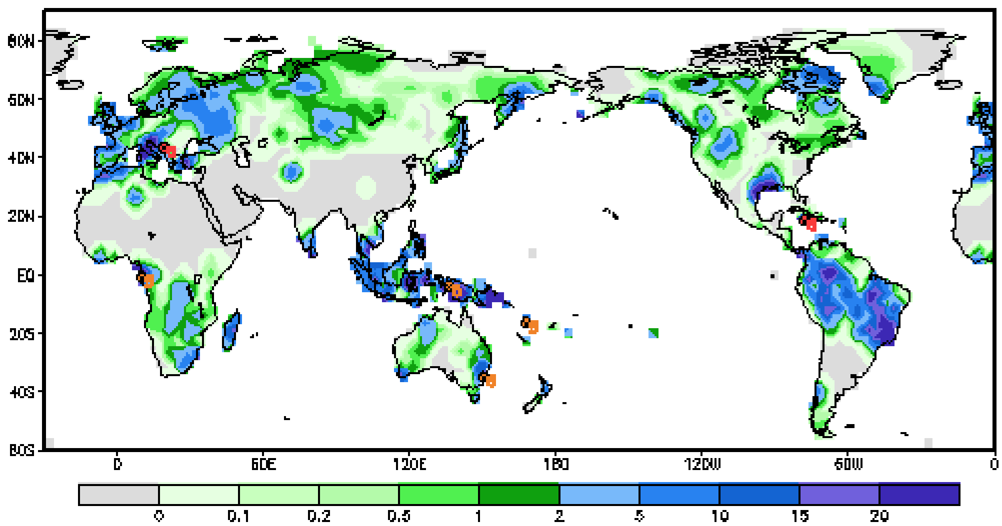

Figure 5: Station location and risk index of detected suspicious GTS daily reports for November 26, 2005, plotted on the background of analyzed field of precipitation based on gauge data.

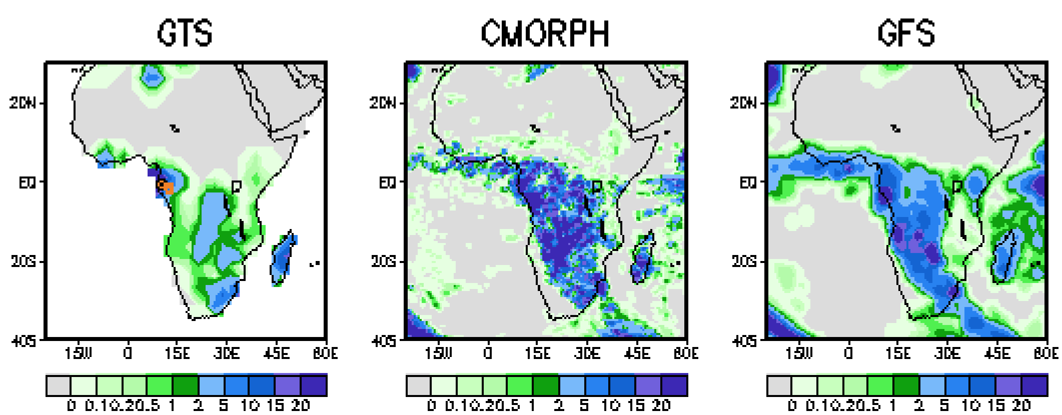

Figure 6: Precipitation distribution over Africa for November 26, 2005, based on 1) GTS gauge data (left), 2) CMORPH satellite estimates (middle), and 3) GFS model forecasts (right). The orange dot on the left panel indicates location of the station with suspicious daily report of '0' rainfall for the day.

Detection of daily reports with suspicious rainfall amounts seems reasonable. For example, a daily report of '0' value over the equatorial western African coastal region (Gabon) is assigned with a high risk index value of 5 (fig. 6, left). The region is covered with heavy rainfall as seen from the satellite estimates (fig.6, middle) and the model forecasts (fig.6, right). In total, about 0.03% and 0.05% of GTS daily reports are identified with a risk index of 5 or higher for suspicious '0' and large value reports, respectively, during the 6-year period from 2003 to 2007 (fig.7).

Quality controlled GTS daily station data are combined with station data from other sources over United States, Mexico, Brazil and Australia to form a unified database of quality controlled gauge precipitation over the globe. Analyses of global land precipitation are produced by interpolating the unified station data through the objective algorithm of Xie et al. (2007) selected based on a comprehensive inter-comparison of objective techniques (Chen et al. 2008). As clear from fig.8, the gauge-based analysis presents reasonable distribution of daily precipitation with fine structures over regions with dense gauge networks.

5. Summary

Objective procedures have been developed for the quality control of the GTS daily precipitation reports through comparisons with historical records at the station, concurrent observations at nearby stations, and precipitation fields from satellite estimates and numerical model forecasts.

Simulation tests have been conducted to examine the performance of the quality control procedures and the results showed reasonable ability of our system to detect simulated daily reports with erroneous rainfall amounts.

An automated operational system has been established at NOAA' Climate Prediction Center (CPC) to perform the quality control for the GTS daily gauge reports on a real-time basis. The quality controlled GTS data is combined with daily reports from other sources to form a unified database of daily precipitation which is used to produce analyzed fields of daily precipitation over the global land areas.

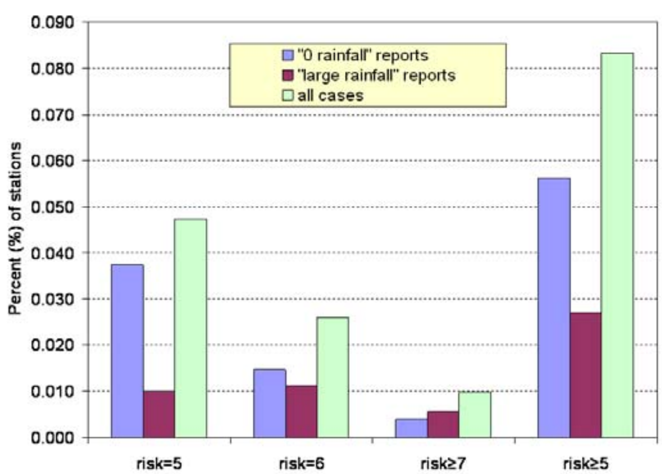

Figure 7: Percentage of daily GTS reports identified with a risk index of 5 or higher for suspicious '0' and large values over a six year period from 2003 to 2007.

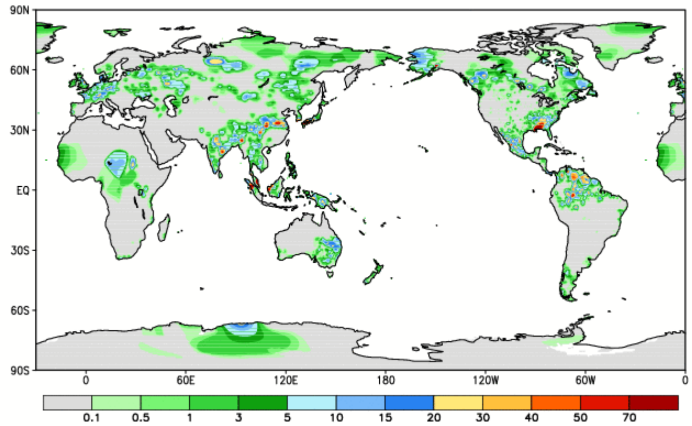

Figure 8: Analysis of daily precipitation (mm) for July 1, 2003, derived by interpolation of quality controlled station reports.

REFERENCES:

Adler, R.F., and co-authors, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979 – present). *J. Hydrometeor.*, **4**, 1147 – 1167.

Chen, M., W. Shi, P. Xie, V.B.S. Silva, V.E. Kousky, R.W. Higgins, and J.E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophy. Phys*. *(in press)*

Ebert, E.E., J.E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**, 47 – 64.

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 No. 7.*

Huffman, G.J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5-20.

Joyce, R.J., J.E. Janowiak, P.A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487 – 503.

Kumar, A., 2007: An overview of operational activities at Climate Prediction Center. (submitted to *Weather and Forecasting*.)

Ropelewski, C.F., and M.S. Halpert, 1987: Global and regional scale precipitation patterns associated with El Nino / Southern Oscillation . *Mon. Wea. Rev.*, **115**, 1606 – 1626.

Xie, P., B. Rudolf, U. Schneider, and P.A. Arkin, 1996: Gauge-based monthly analysis of global land precipitation from 1971 to 1994. *J. Geophy. Res.*, **101D14**, 19023-19034.

Xie, P., and P.A. Arkin, 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.

Xie, P., 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.*, **3**, 249-266.