CGMS IPWG

International Precipitation Working Group

Proceedings

1st Workshop

**Madrid, Spain
23-27 September 2002**

NOAA INM EUMETSAT

VERIFICATION EXPERIMENT OF THE GLOBAL PRECIPITATION CLIMATOLOGY PROJECT 1 DEGREE DAILY PRODUCT

Mark L. Morrissey[1], J. S. Greene[1], Mike D. Klatt[1], and J. Ensworth[1]

Abstract

A verification exercise was conducted comparing satellite algorithm rainfall estimates from the Global Precipitation Climatology Project (GPCP) with several dense rain gauge networks within Florida, US. The networks are operated by the United States Florida Water Management Districts. The gauge data from 1997 through present has been analyzed and are archived by the NASA Tropical Rainfall Measuring Mission (TRMM) Ground Validation program (GV). The GPCP daily 1 x 1 degree latitude, longitude rainfall product was compared with daily box averages of the gauge data from 1997 to 2001. Results indicate a significant difference in the mean values of the GPCP product with the gauge data. The difference indicates that the 1DD is approximately 50% higher on the averaged than the corresponding box-averaged gauge data on high rainfall days. The bias is conditional on rainfall. A positive bias for low rainfall rates and the relatively small difference in long-term averaged rainfall, shown by an intercept of 1.2 mm d^{-1}, indicates that the overall bias in the 1 DD product is close to zero.

Further examination of the gauge data with United States Cooperative Observer rain gauge records indicate that the Water Management District gauge data were biased low by about 20%.

Comparisons of the Water Management District gauge data with the Global Precipitation Climatology Center (GPCC) gauge analysis from the same areas indicate that the GPCC analysis produces area estimates approximately 10% higher than the reference Water Management gauge data. The GPCC product is used to adjust the accumulated satellite rainfall estimates to the monthly amounts observed in the GPCC analysis.

Further statistical analysis confirmed that the observed gauge and satellite biases were consistent with rainfall amount indicating a systematic conditional bias in the gauges and/or the GPCP 1DD product.

Assuming the Water Management reference gauge rainfall estimates are biased low by the maximum amount of 20%, this still leaves the GPCP satellite estimates with an unexplained conditional bias of approximately 30% (at high rain rates). In addition, the

---

[1] Environmental Verification and Analysis Center, 3200 Marshall Ave., Suite 110, University of Oklahoma, Norman, OK 73072 USA

GPCP 1 DD product underestimates rainfall at very low rainfall amount. This was indicated by a consistent box to box intercept value of approximately 37 mm month^{-1}.

1.   Introduction

The Global Precipitation Climatology (GPCP) project (WCRP, 1986; Huffman et al., 1997) began during the 1980's with the initial objective of producing a global satellite-gauge rainfall product on a monthly, 2.5 x 2.5 degree, latitude-longitude scale to be utilized as input and assessment tools for global climate models. The need for such a project is outlined in the document by the World Meteorological Organization (WMO, 1985).

During the experimental phase (1986-1993 roughly), various satellite-based algorithms were tested for reliability and accuracy. It was quickly realized that as technology changes with time, so do the satellite algorithms. Thus, different and improved algorithms have been developed over time with the GPCP continually producing global satellite estimates and updating the past records. This, of course, necessitates a continual effort at testing new algorithm performance.

Due to the request from the modeling community for rainfall estimates as accurate as possible, it was decided during the initiation of the GPCP to establish the Global Precipitation Climatology Center (GPCC) at the Deutscher Wetterdienst (the German Weather Service), Offenbach, Germany. The GPCC produces a rain gauge-only analysis which is then merged with the satellite product to produce the final GPCP global product. The gauge analysis scheme is described by Rudolf (1993).

During the last 5 years the global modeling community recognized a need for global rainfall estimate at higher time and space scales than that produced by the GPCP. Thus, GPCP scientists began experimenting with different satellite algorithms with the goal of producing an operational one degree latitude-longitude, daily product. Starting in 1996, the GPCP began producing a product at this scale (refer to http://precip.gsfc.nasa.gov/gpcp_daily_comb.html for a full description of the product). It is important to note that this product does not direct include the GPCC gauge analysis but is corrected to the monthly totals produced by the GPCP monthly product which does include the GPCC gauge analysis (G. Huffman, personal communication, 2002).

The work presented in this paper is one of the initial and ongoing efforts to assess the accuracy of the product. One common method to assess accuracy is through direct comparison of satellite estimates with dense rain gauge network data representing the same scale. Difficulties to overcome using this method are the assessment of the errors, both systematic and random in the gauge network (hereafter referred to as the reference data) and the matching of scale between the gauge estimate and the reference data. The primary source of random error in gauge network is incomplete spatial sampling, while the main source of random error in the satellite estimates of temporal sampling error.

During the startup phase of the GPCP the need existing for ongoing validation of the continually changing GPCP algorithms was recognized. Thus, the Surface Reference Data Center (SRDC) was established at the National Climatic Data Center (NCDC) in Ashville, NC, US. In the mid-1990's the SRDC was re-located at the University of Oklahoma at the Environmental Verification and Analysis Center (EVAC) (refer to

http://srdc.evac.ou.edu). In addition, the SRDC works in close concert with the University of Iowa (Dr. W. Krajewski) validating GPCP products.

2.   Reference data

For this study, the dense rain gauge networks operated by the Florida Water Management Districts were utilized as reference data. These data were collected and archived by NASA's Tropical Rainfall Measuring Mission (Simpson et al., 2000) (also refer to (TRMM Ground Validation Program, http://trmm-fc.gsfc.nasa.gov/trmm_gv/index.html).
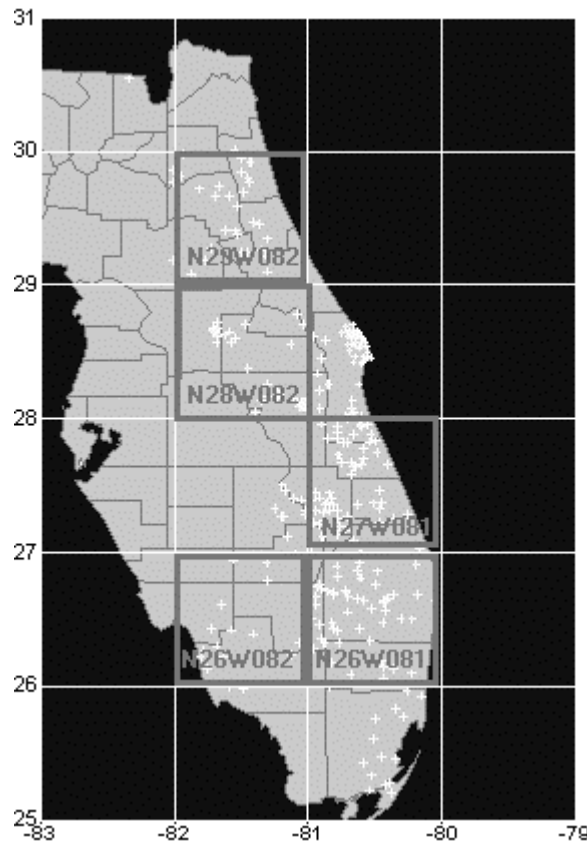

Figure 1. *TRMM Florida Gauge locations and 1 degree boxes used in this study.*

The one minute resolution gauge data were averaged into daily values and boxes selected for this experiment were only those having a sampling error below a specified threshold. The sampling error was computed using a method developed by Morrissey et al. (1995) and boxes selected were those shown by outline in Fig. 1. Five boxes in total were select.

3.   Comparison of gauge network data with GPCP 1 DD estimates

Given the 5 year records of the Florida Water Management Districts gauge data archived by the TRMM GV program, we compared 1DD and gauge records from 1997 through

2001 only. For an initial quick look we simply compared the box averaged gauge values to corresponding 1DD values. The results for one of the boxes are shown in Fig. 2. Table 1 summarizes the findings for the slope and intercept values for each of the 5 boxes.. Of note is the consistent slope values less than one suggesting a high bias in the 1DD product relative to rain accumulation. The intercept values are also consistently greater than one. This suggests a low bias in the 1DD product for days when rainfall amount is very small. Thus, amounts to approximately 37 mm month^{-1} underestimated by the 1DD product on light rain days. Thus, a conditional bias is indicated.
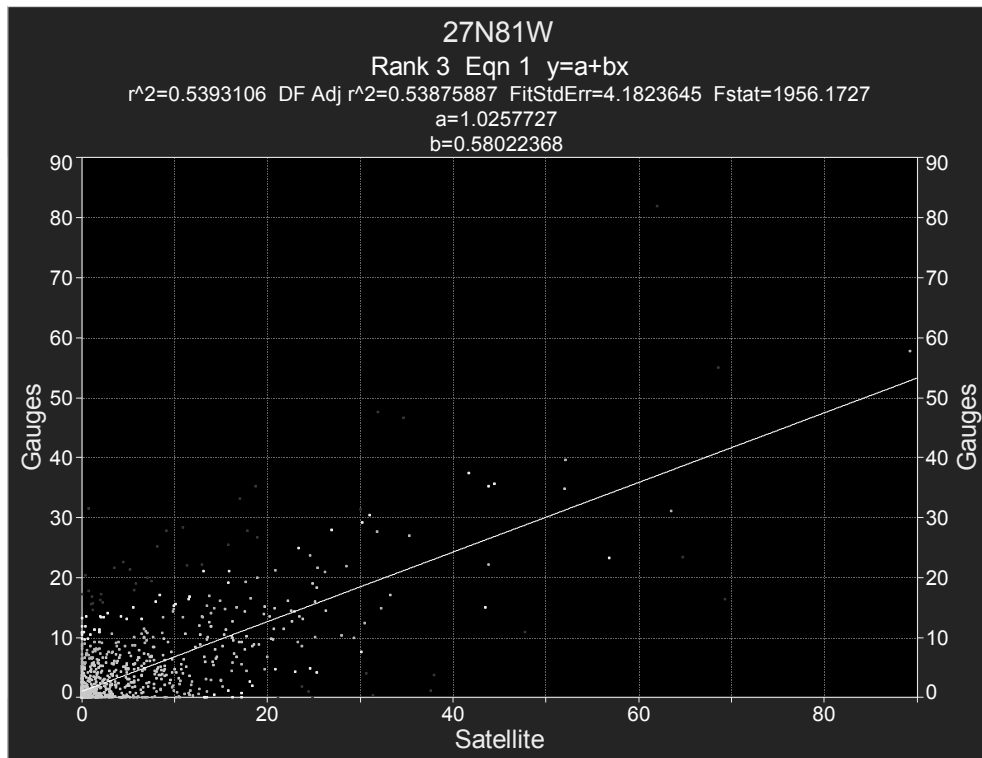

Figure 2. *Representative scatter of 1DD values vs. box average daily gauge estimates.*
Fig. 2 shows the comparison of the Florida 1 degree box with the southwest corner located at 27N, 81W. A simple linear regression analysis produces a slope of 0.58 and an intercept of 1.02.

Slope	Intercept
0.40	1.23
0.60	1.14
0.58	1.03
0.40	1.30
0.35	1.40
0.48	1.22

Table 1. *Summary statistics for FLA boxes.*

These results bring up several questions. For example, although an obvious and consistent difference is observed from box to box between the satellite and gauge estimates, which (or are both?) estimates are truly biased? We attempted to take a closer look at the TRMM gauges. We first selected two boxes located within the north central climate regime of Florida (28N 82W and 28N 81W). We then retrieved an independent set of gauge data from those locals and conducted a comparison between to two gauge data sets. The second gauge data set was daily values from the U.S. Cooperative Station network archive (i.e. Co-op) at NCDC. There were 533 Co-op station-days within the two boxes. There were 2653 station-days for the TRMM gauges within the two boxes.

4. Comparison of gauge networks using the probability density function

To determine the differences in the statistical distributions of the two gauge dataset we use the non-parametric kernel method of fitting a probability density function (pdf) to each set of data (Silverman, 1998). An approximation to the Gaussian function was used as the kernel. The result, after the application of appropriate smoothing methods, were two functions approximating the true pdf for each data set (Figure 3). The figure indicates comparable shapes to the distribution, but some differences can be seen. To illustrate these differences more clearly, we generated the cumulative pdf for each data set through integration of the pdf (i.e. F(x)). We then compared the F(x) values between the two data sets with different values of x.
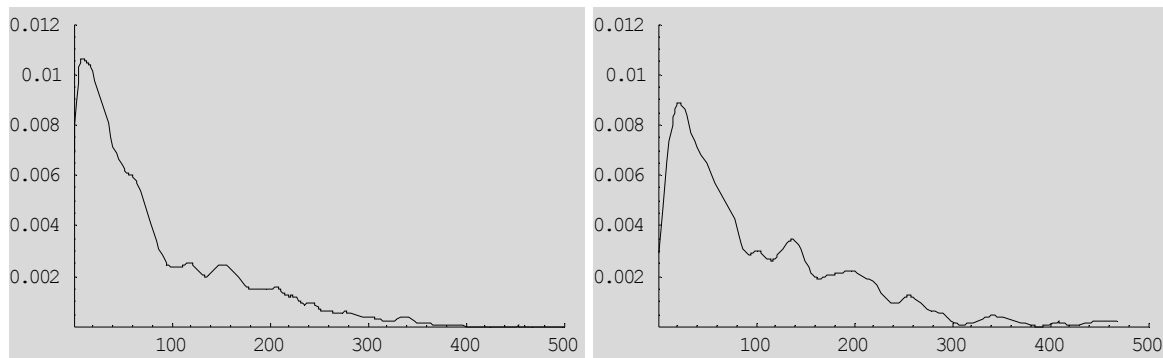

Figure 3. *The estimated pdf's for TRMM (left) and the Co-op (right) data sets in north central Florida.*

The results of this comparison indicate that the TRMM gauge data underestimates daily totals compared to the Co-op data. It also indicates that this underestimation is consistent with daily rainfall amount. Thus, there appears to be a consistent, systematic bias in the TRMM gauge data in Florida. The total difference, however, is at a maximum 20%, which still doesn't fully explain the 50% difference revealed in the TRMM gauge-1DD GPCP product comparison.

Comparisons were also made between the TRMM gauge data and the GPCC analysis. The results indicated an approximately 10% low overall bias in the TRMM gauge values. However, it should be noted that GPCC gauge density in the study region is extremely

low, approximately one gauge per one degree box. Thus, the statistic significance of the difference estimate is quite low as well.
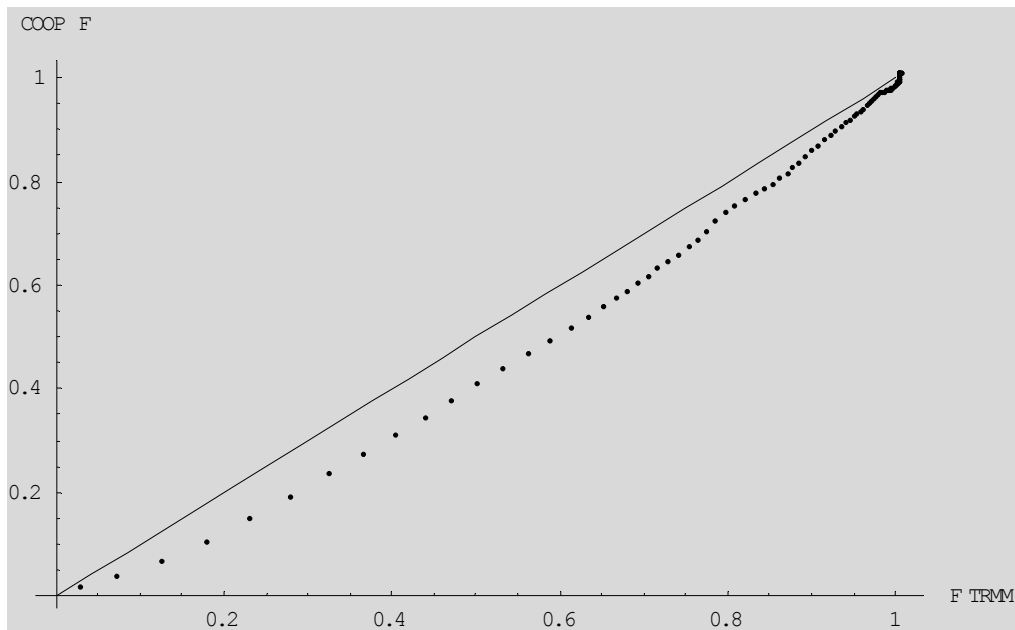

Figure 4. *Comparison of F(x) for Co-op data (y-axis) with TRMM gauge data (x-axis) (the 45 degree line is also shown for comparison).*

5.   Conclusions

A significant conditional bias was observed in the GPCP 1DD product when compared to the Florida Water Management Districts gauge network of tipping bucket gauges (archived by NASA TRMM GV).   A smaller conditional bias was also found using TRMM gauges in the Houston, Texas, US region (not shown).   The difference was approximately 50%. A careful comparison study of two independent gauge networks indicated that only about 20% at the maximum of this bias could be explained by gauge under catch (or calibration error) in the Florida TRMM gauge data set. Thus, a preliminary conclusion is that the 1DD product is conditionally biased (at least over Florida).
Of course, addition comparisons over different areas are required to determine the extent, nature and source of this bias. By working in collaboration with the satellite algorithm developers, the SRDC will continue to provide reasonably objective validation results which the developers use to improve their algorithms.

6.   References

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The Global Precipitation Climatology Product (GPCP) Combined Precipitation Data Set. *Bull. Amer. Meteor. Soc.*, **78**, 5-20.

Morrissey, M. L., J. A. Maliekal, J. S. Greene, and J. Wang, 1995: The uncertainty of simple spatial averages using rain gauge networks, *Wat. Resourc. Res.*, **31**, 2011-2017.

Rudolf, B., 1993: Management and analysis of precipitation data on a routine basis. *Proc. Int. WMO/IAHS/ETH Symp. on Precip. and Evap.*, Slovak Hydromet. Inst., Bratislava, Sept. 1993, 69-76.

————, H. Hauschild, W. Rueth, and U. Schneider, 1994: Terrestrial precipitation analysis: operational method and required density of point measurements. *NATO ASI I/26, Global Precipitations and Climate Change*, M. Desbois and F. Desalmand, Eds., Springer Verlag Berlin, p. 173-186.

Silverman, B. W., 1998: *Density estimation for statistics and data analysis*. Chapman and Hall/CRC, Washington.D.C.US, 175 p.

Simpson, J., C. D. Kummerow, R. Meneghini, A. Hou, R. F. Adler , G. J. Huffman, B. Barkstrom, B. Wielicki, S. J. Goodman, H. Christian, T. Kozu, T. N. Krishnamurti, S. Yang, and B. Ferrier, 2000: The tropical rainfall measuring Mission (TRMM) progress Report. *Earth Observation and Remote Sensing*, **18**, August Issue.

WCRP, 1986: *Report on the workshop on global large scale precipitation data sets for the World Climate Research Program*. WCRP-111, WMO/TD-No. 94, 45pp. (Available from the World Meteorological Organization, P. O. Box 2300, CH-1211 Geneva 2, Switzerland.)

WMO , 1985. *Review of requirements for area-averaged precipitation data, surface-based and space-based estimation techniques, space and time sampling, accuracy and error; data exchange*. WCP-100, WMO/TD-No. 115, 57 pp. and appendices.