CAMS–OPI: A Global Satellite–Rain Gauge Merged Product for Real-Time Precipitation Monitoring Applications

JOHN E. JANOWIAK

Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, Maryland

PINGPING XIE

Research Data and Systems Corporation, Greenbelt, Maryland

27 August 1998 and 22 March 1999

ABSTRACT

A method has been developed to produce real-time rain gauge–satellite merged analyses of global monthly precipitation. A dataset of these analyses spans the period from January 1979 to the present, which is sufficiently long to allow the computation of reasonably stable base period means from which departures from ''normal'' can be computed. The dataset is used routinely for global precipitation monitoring purposes at the National Oceanic and Atmospheric Administration/National Weather Service/National Centers for Environmental Prediction/Climate Prediction Center, is updated monthly, and is available via the Internet.

1. Introduction

At least two datasets exist that contain global monthly precipitation analyses that are composed of a mix of rain gauge observations and satellite estimates. Two such analyses are the Global Precipitation Climatology Project (GPCP) combined analysis that is described by Huffman et al. (1997) and the Climate Prediction Center Merged Analysis of Precipitation (CMAP) of Xie and Arkin (1997). However, neither of these products is available to meet the needs of real-time global precipitation monitoring efforts at the present time. The reason for this is largely procedural, since the timely production of both products is dependent upon the availability of rain gauge data and satellite estimates that are collected at scheduled times to meet the needs of the GPCP. Since the GPCP is charged with the development of analyses that are intended for research and model validation and not with the production of real-time precipitation analyses, the data that flow into the project are generally received too late for real-time assessments of global precipitation to be made.

The need for real-time monitoring of monthly global precipitation anomalies was underscored by a spectrum of disciplines during the exceptionally strong 1997/98 El Niño. To this end, a method was developed to produce such analyses (called the ''CAMS_OPI''), which is described in this paper. The CAMS_OPI technique is named after the two data sources that are used in the generation of the product, namely, 1) monthly rain gauge totals from the Climate Anomaly Monitoring System (CAMS; Ropelewski et al. 1984), and 2) satellite-based estimates from outgoing longwave radiation (OLR) anomalies that are generated by the OLR Precipitation Index (OPI; Xie and Arkin 1998). As discussed in section 2a, however, the rain gauge reports are presently extracted from a different source. Two features of these data make them attractive for the task at hand. First, both are available in real time. Second, the lengths of record of the datasets (1979–present) enable the computation of base period means of merged precipitation analyses from which reasonably stable assessments of deviations from normal can be computed. These desirable features should not necessarily persuade users to choose the CAMS_OPI over the GPCP or CMAP analyses. Both of the latter techniques incorporate multichannel passive microwave information in addition to higher temporal resolution IR data from geostationary satellites than are presently available for the generation of real-time CAMS_OPI analyses. In addition, the nonreal-time production schedule for those methods affords considerably more time and manpower for quality control of the global rain gauge data than are available for the real-time estimates derived by the CAMS_OPI algorithm. Therefore, we *strongly* suggest the use of the GPCP or CMAP estimates for purposes

Corresponding author address: Mr. John Janowiak, Climate Prediction Center, NCEP/NWS/NOAA, W/NP52, Room 605, WWBG, NOAA, 5200 Auth Road, Camp Springs, MD 20746-4304.
E-mail: wd55jj@sun1.wwb.noaa.gov

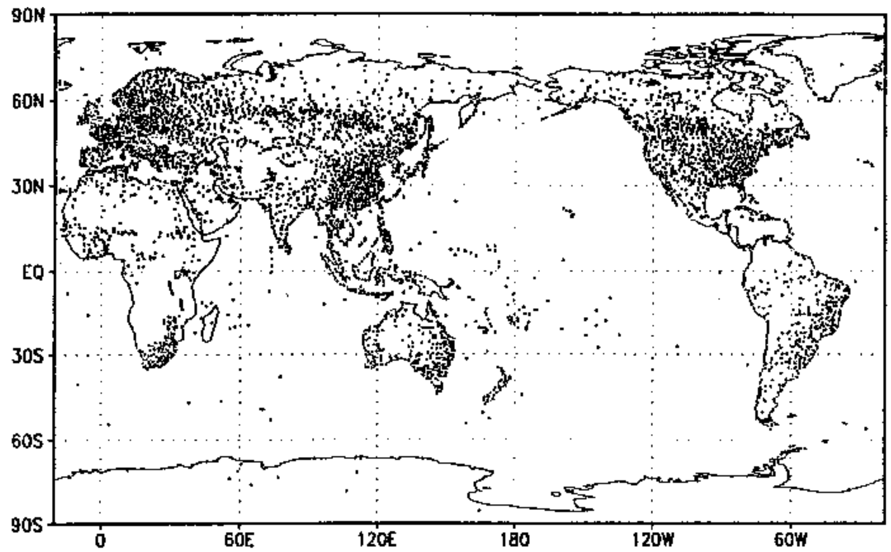

FIG. 1. Typical distribution of rain gauge reports received via GTS at the Climate Prediction Center. This map is for Jun 1998.

other than real-time precipitation monitoring on climatic spatial scales.

2. Description of the data and the merging process

The final merged product is obtained via a two-step process. First, the rain gauge data are objectively analyzed to equally spaced (in latitude and longitude) grid points as described in section 2a. Then the satellite estimates (OPI, described in section 2b) are merged with the gauge analysis as described in section 2c.

a. Rain gauge data

The rain gauge data that are used in this merged product are extracted from the Climate Anomaly Data Base (CADB), which is maintained at the Climate Prediction Center (CPC) and is a compilation of global synoptic weather reports that are received over the Global Telecommunications System (GTS). Reports are generally received from more than 7000 stations in a typical month. Despite this seemingly large number of reports, no observations are available over the oceans (except for isolated island stations that are often unrepresentative of open-ocean conditions due to terrain-induced effects of the islands) and the gauge distribution over land is far from uniform, with sparse coverage in much of Africa, large portions of South America, and parts of Asia (Fig. 1). In the early version of this product, the rain gauge data were extracted from the CAMS database, but a switch was made to the CADB reports beginning with the September 1982 analyses (the beginning of the CADB archive record) in the current version since substantially more rain gauge reports are available from that source than from CAMS. The reason for the difference between CAMS and the CADB is that the CAMS database is limited to reports from stations with sufficiently long historical records while the CADB contains reports for all available stations regardless of the length of the station histories. The CAMS–OPI analysis is redone near the middle of each month when station CLIMAT reports become available.

In preparation for merging the station data with the satellite-derived precipitation estimates, a rain gauge–only analysis is produced by gridding the rain gauge reports to a 2.5° lat–long grid using a modified spherical version of the Shepard (1968) scheme. In this scheme monthly precipitation observations are first gridded at 0.5° lat–long grid resolution by interpolating surrounding gauge observations and adjusting them by weighting coefficients that are inversely proportional to the gauge–grid point distance. The areal mean value for each 0.5° lat–long grid box is then computed as the arithmetic average of the point values at its four corners. Finally, the area-mean precipitation for each 2.5° lat–long grid box is defined as the area-weighted average of the 25 0.5° lat–long boxes. Further details of this procedure are discussed in Willmott et al. (1985). The modification that we made to this method is that the value at each 2.5° grid box is set to a missing value when no rain gauge reports exist within two 2.5° grid boxes (in every direction) of the grid location in question. This change was made to the scheme to avoid interpolations of rain gauge data over large distances where no rain gauge reports are available.

More accurate results are generally obtained by analyzing anomalies (i.e., departures from normal) and then adding these anomalies to the climatological means rather than analyzing mean values directly (New et al. 1997). However, analyzing data in this manner is done at the expense of a considerable reduction in the number

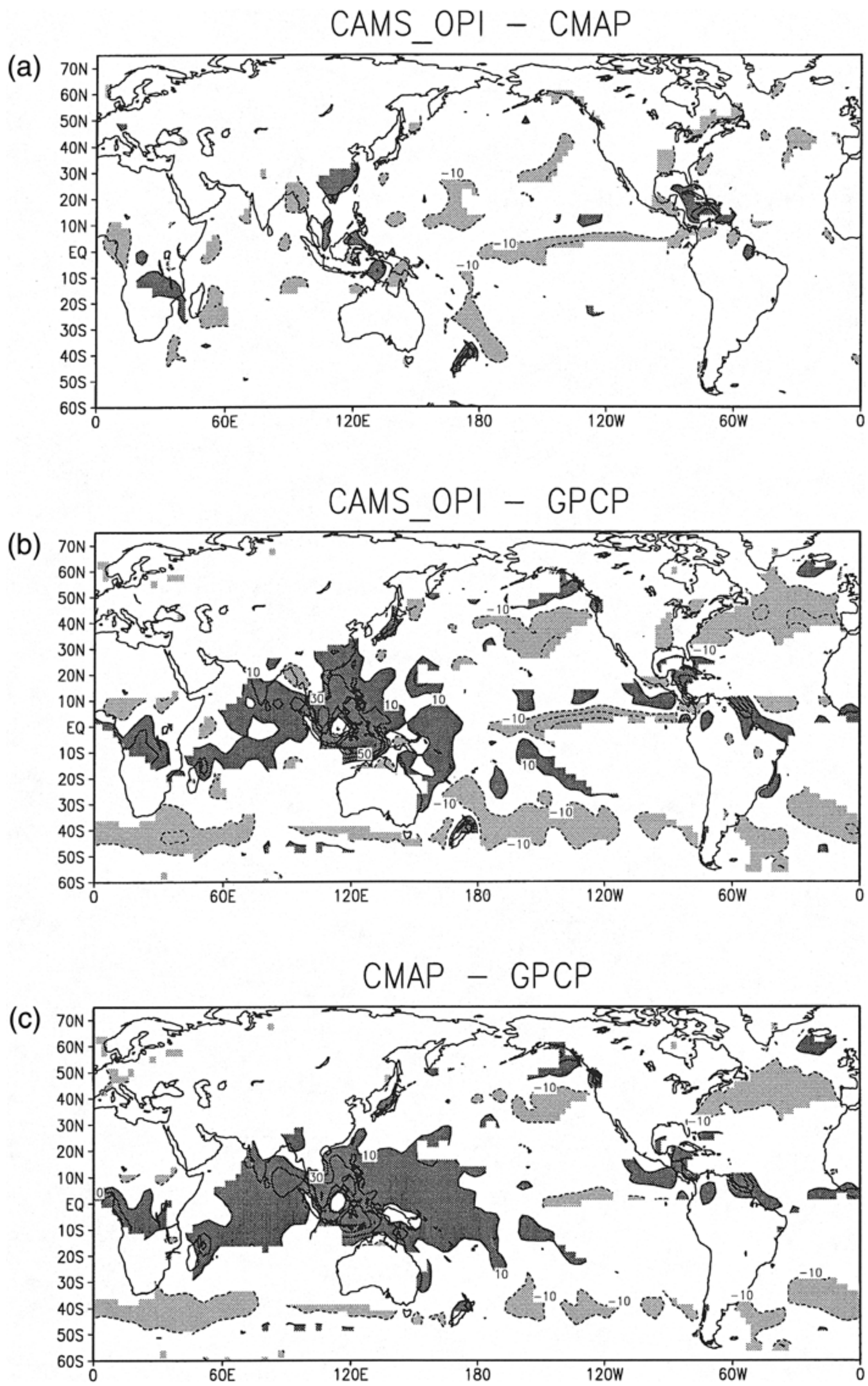

FIG. 2. Mean differences (%) among the CAMS_OPI, CMAP, and GPCP merged analyses of monthly precipitation over the period Jan 1996–Dec 1997. Light (dark) shading indicates negative (positive) differences in excess of 10%. Differences are masked out in regions where the mean monthly precipitation is less than 2 mm day^{-1}. A nine-point smoother has been applied to the data for visual clarity.

of observations, since only those observations with sufficiently long temporal records from which stable climatologies can be computed can be used. For example, if rain gauge *anomalies* were to be analyzed instead of monthly totals, the number of CADB observations would be reduced by a factor of 4, thus we decided to analyze the precipitation totals instead, and users of these analyses should bear this in mind.

The quality of any gauge-based precipitation analysis depends primarily on the gauge network density and the quality of the individual rain gauge reports. Previous investigations have shown that the random error of the gauge-based analysis decreases with increasing gauge network density, while significant bias exists over grid boxes with no gauges (where values are determined by interpolating observations over the surrounding areas). The quality of individual rain gauge reports is assessed by comparing the monthly totals with the climatological value (1979–95) of the GPCP rain gauge–only analysis (provided by the Global Precipitation Climatology Center, Deutscher Wetterdienst, Germany) for the grid location in which the rain gauge location resides, and with the collocated satellite estimate (OPI). Suspicious reports are written to a ''warning file'' that is recorded for each month and then a manual determination is made on whether to keep or discard such reports based on the information in the warning file. Such determinations are not always straightforward, and we have chosen to keep the rain gauge reports unless we can be reasonably confident that they are wrong. Admittedly, we are biased toward the removal of unusually large amounts of precipitation, although it is probably more likely that the majority of errors involve precipitation amounts that are too low.

b. Satellite estimates

The satellite estimates used here are derived from OLR observations from National Oceanic and Atmospheric Administration polar-orbiting satellites using the OPI technique (Xie and Arkin 1998). The OPI algorithm is based on studies that indicate that precipitation anomalies correlate well with OLR anomalies over most of the globe and that the proportional coefficients relating them can be expressed as a globally uniform linear function of the local mean precipitation (Xie and Arkin 1998).

The OPI estimates of monthly precipitation are produced for each grid box in two steps. First, precipitation anomalies are estimated from OLR anomalies for a given month. Second, these anomalies are added to the mean annual cycle of precipitation for that month to obtain an estimate of the monthly precipitation total. The mean annual cycle of the precipitation is defined from the CMAP precipitation analysis of Xie and Arkin (1996) for the 8-yr period from July 1987 (when passive microwave data from the Special Sensor Microwave Imager instrument became available) to June 1995. The reason for using the mean annual cycle information from the CMAP analyses is to train, in a statistical sense, the OPI estimates to conform to the presumably more accurate CMAP estimates, which incorporate information from several microwave sensors and which use higher temporal resolution IR data than OLR. The precipitation anomalies for a given month are then estimated from the OLR anomalies at each location by applying the appropriate coefficients that relate the CMAP mean annual cycle of the precipitation to the OLR anomalies.

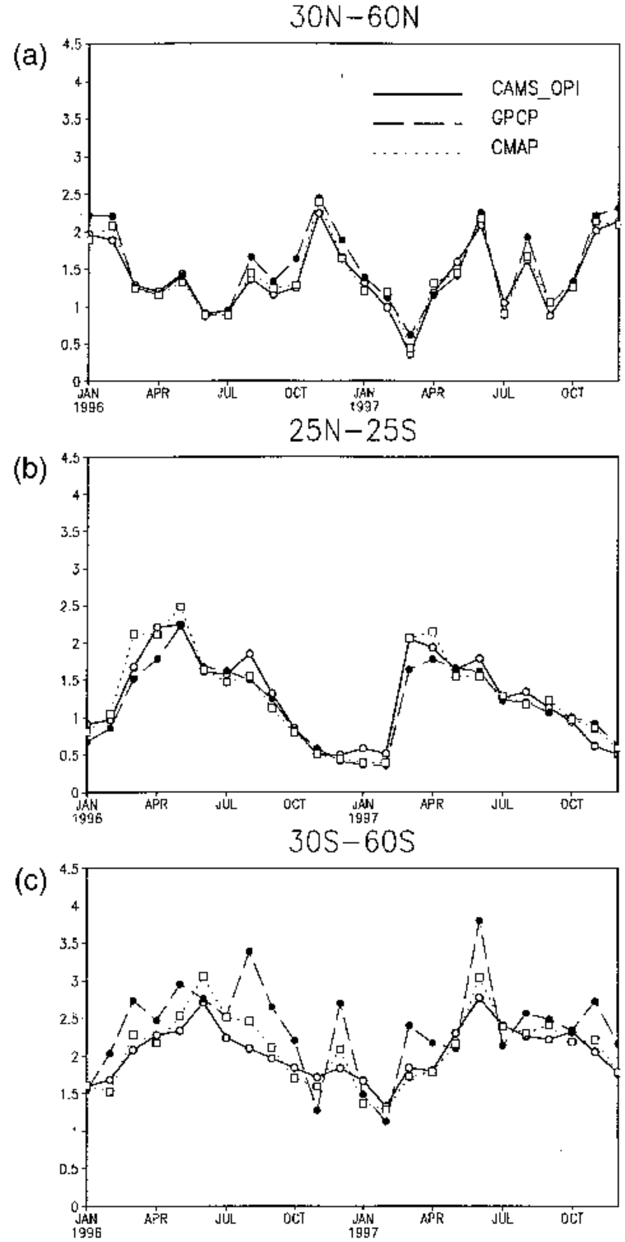

FIG. 3. Time series of zonally averaged monthly precipitation (mm day⁻¹) for CAMS_OPI (thick, solid line), GPCP (dashed line), and CMAP (dotted line) for (a) the 30°–60°N zonal band, (b) 25°N–25°S zonal band, and (c) 30°–60°S zonal band.

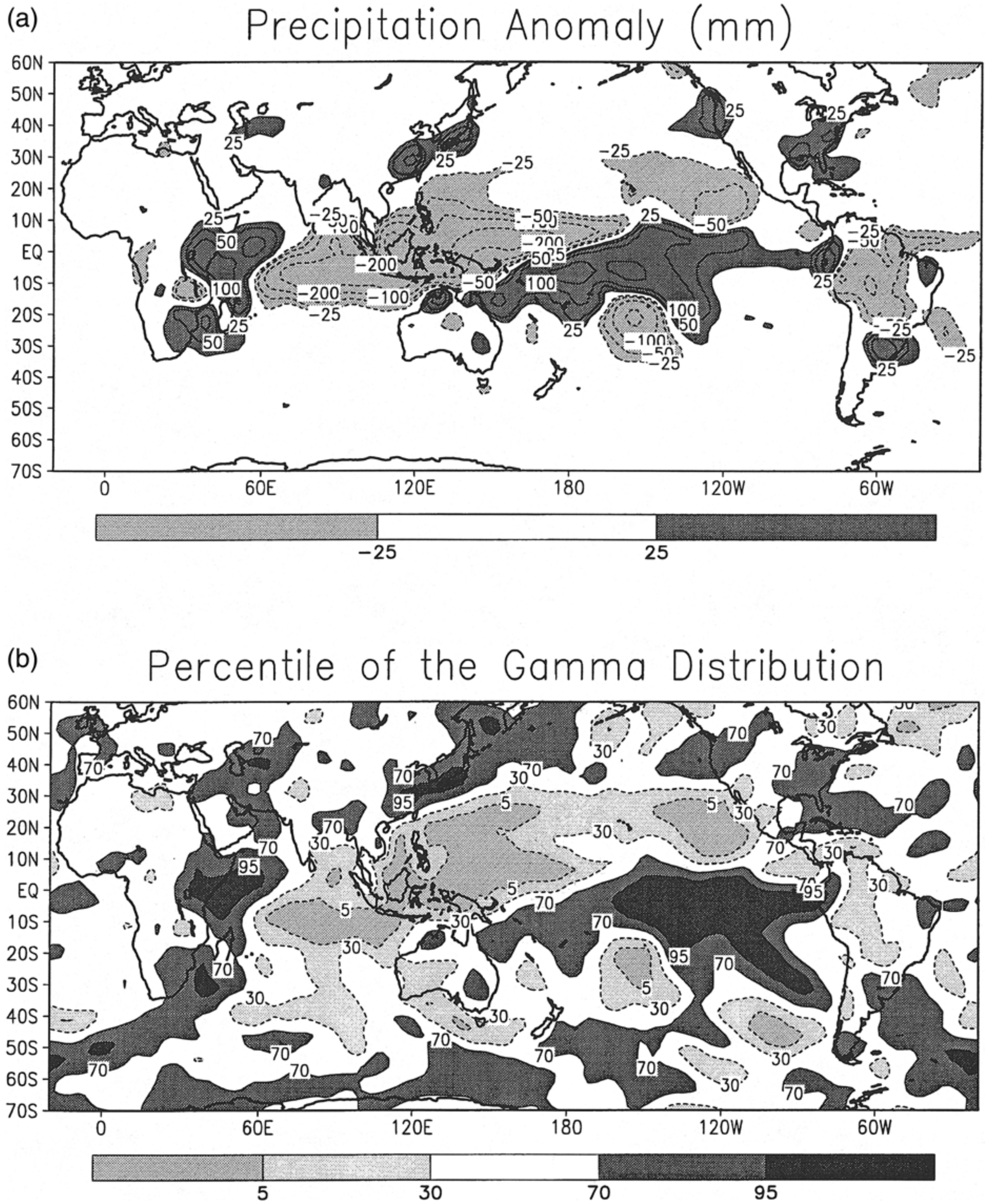

FIG. 4. CAMS–OPI (a) precipitation anomaly (mm) and (b) percentile of the gamma distribution for Jan 1998.

Comparisons of the OPI estimates with independent precipitation observations from rain gauge reports indicate that the temporal correlation between them is generally near 0.60 and that relative random error is near 50% with limited bias (Xie and Arkin 1998).

c. Data merging procedure

The individual rain gauge and OPI gridded analyses are merged to provide complete global coverage. Over the open oceanic areas, where rain gauge observations

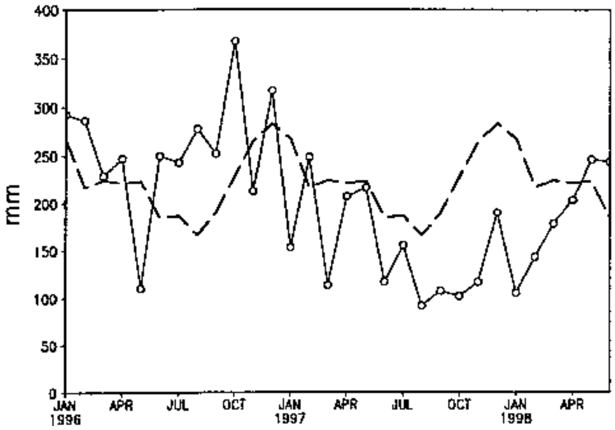

FIG. 5. Time series of CAMS–OPI precipitation over the near-equatorial Indonesia region (5°N–5°S, 95°–130°E). Units are mm. Solid line represents the CAMS–OPI monthly estimates of precipitation for Jan 1996–Jun 1998; dashed line is the 1979–97 CAMS–OPI base period mean.

are unavailable, the merged analysis amounts match the OPI estimates. Over land, the gauge-only analysis values are used in the merged analysis where such values are available. At locations where the rain gauge analysis has a missing value, the OPI estimates are blended with the rain gauge analysis via the method of Reynolds (1988). This blending technique interpolates the gauge-based analysis from nearby grid boxes with the constraint that the shape of the precipitation field, defined by the Laplacian operator, follows that of the OPI estimates. Verification tests have confirmed that the merging procedure described above is able to produce precipitation analyses with almost no systematic error (bias) over grid locations where rain gauge data are not available (Xie and Arkin 1996). These tests were accomplished by comparing the differences (over many iterations) between the analyses based on all available station data with analyses on the same gauge data but with 20% of the stations withheld (randomly selected) from the full set.

3. Comparison with other merged analyses

In Fig. 2 the CAMS_OPI analysis is compared to the GPCP and CMAP analyses over the period January 1996–December 1997, a 24-month period that is *not* part of the developmental sample upon which the OPI estimates were trained against the CMAP product (see section 2b). Over land, where the estimates of all three products are dominated by rain gauge observations, the differences are less than 10% nearly everywhere and >10% only in regions where the rain gauge density is relatively sparse.

More widespread differences (10%–30%) are observed over the oceans, which is expected since it is there that the differences among the precipitation estimation algorithms are largest. The major differences among the estimation methods over the oceans are that 1) the CMAP and GPCP algorithms use passive microwave data that are not used (for timeliness considerations) in the OPI, and 2) in an effort to remove bias in the satellite estimates over the oceans in CMAP the estimates are adjusted to rain gauge reports on atolls in the South Pacific (Morrissey and Greene 1991), which is not done in the GPCP merged analysis.

The smallest difference among the comparisons exists between the CAMS_OPI and CMAP analyses, which is not surprising since the OPI satellite estimation method has been trained on the CMAP analyses. Note that the largest difference occurs between the GPCP and CAMS_OPI analyses and that difference is similar to the difference between the GPCP and CMAP analyses. The differences are largest in the Southern Hemisphere (Fig. 3) which is consistent with the discussion above since most of the global oceanic area is in that hemisphere.

In summary, little difference among the three analyses exist over the land surfaces. Over the oceans, relatively small differences between CAMS_OPI and CMAP (generally <10%) are smaller than the difference between the CMAP and GPCP analyses. These findings suggest that the CAMS_OPI is suitable for real-time climate monitoring purposes, but we reiterate that the GPCP and CMAP analyses should be used whenever possible instead of the CAMS_OPI analysis for reasons discussed in section 1.

4. Real-time monitoring applications

Since the information used by the CAMS–OPI method is available in near–real time, the method lends itself well to the task of real-time monitoring of global precipitation. In addition, the historical record of the OLR and rain gauge data extend back to 1979, therefore reasonably stable monthly base period means can be computed (20 years at the time of this writing) and fairly robust statistical inferences on the monthly departures from ‘‘normal’’ can be computed world-wide. However, users are urged to exercise caution when using these analyses over oceanic regions as the precipitation amounts there are usually derived solely from the satellite-based OPI algorithm.

It is well known that the distribution of precipitation on many timescales is not Gaussian in nature and tends to be skewed (Conrad and Pollack 1962). Thus, the gamma distribution (Thom 1958) has seen relatively wide use and is used at CPC to characterize precipitation anomalies in terms of departure from normal. Parameters of the gamma distribution are computed at each grid location, for each month (or longer), using the entire (presently 20 year) CAMS_OPI record so that local precipitation anomalies can be characterized with respect to the historical record. A 20-yr record is a relatively short period for the computation of these parameters. However, the base period is limited by the fact

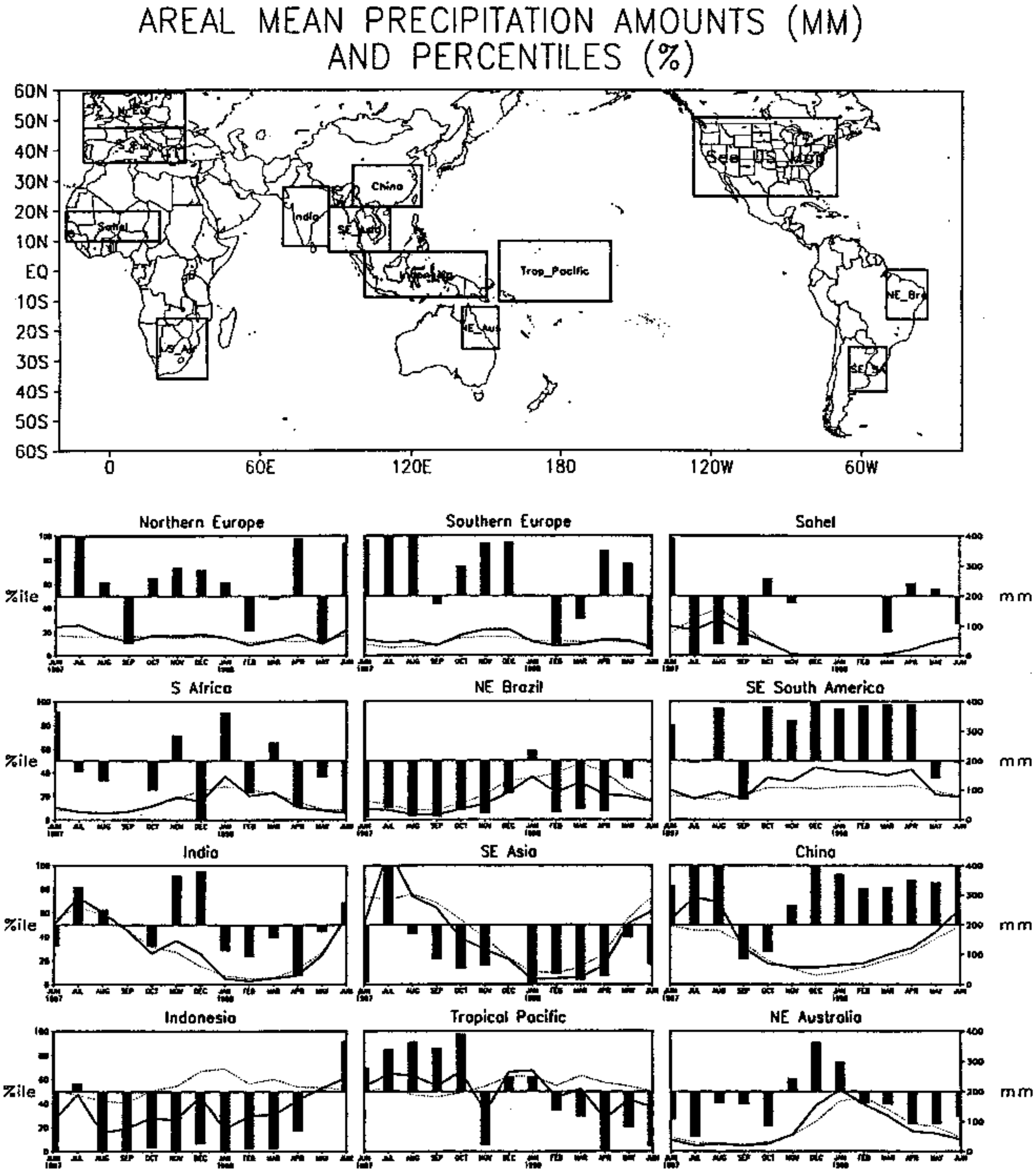

FIG. 6. Regional time series of CAMS–OPI precipitation for Jun 1997–Jun 1998. Bars represent the percentiles of the gamma distribution; solid lines are the monthly precipitation accumulations; dashed lines represent the 1979–95 base period mean. The scales to the left of each box are the percentile ranks of the gamma distribution and apply to the bars; scales to the right are in mm units and apply to the lines.

that stable OLR data are available only back to 1979. Thus, users are cautioned that departures from normal conditions need to be interpreted in this context, particularly the extreme values. An example of the precipitation anomalies and their ranks with respect to the gamma distribution at each grid box for January 1998, when mature El Niño conditions prevailed in the tropical Pacific, is shown in Fig. 4. While the large displacements in precipitation features are apparent from the precipitation anomalies (Fig. 4a), the relative importance of these anomalies are more readily apparent when ranked with respect to the gamma distribution (Fig. 4c). For example, the precipitation deficits in the eastern Pacific to the west of Mexico during January 1998 were near 50 mm, compared to deficits greater than 200 mm in the far western equatorial Pacific and the central Indian Ocean. However, the monthly precipitation totals for each of these locations rank in the lowest 5% over the 1979–98 period.

A time series of precipitation anomalies over Indonesia (Fig. 5) shows the transition from near-normal precipitation to the extended period of abnormally dry conditions there that were associated with the 1997/98 El Niño event. The oscillating nature of the precipitation from August 1996 through May 1997 is a manifestation of the strong Madden–Julian oscillation (Madden and Julian 1971) activity that preceded the onset of the El Niño, with a strong downward trend in the magnitude of the precipitation evident over that period as well. Precipitation remained below normal for 11 straight months (June 1997–April 1998) over Indonesia and the surrounding oceanic area.

Displays of the CAMS–OPI precipitation data such as in Fig. 6, which appear routinely in the CPC Monthly Climate Diagnostics Bulletin, are powerful tools to quickly monitor the evolution of regional precipitation anomalies. For example, the evolution of precipitation anomalies related to the 1997/98 El Niño is evident from quick inspection of the bar graphs for Indonesia, northeast Brazil, and the tropical Pacific. These plots also convey the uncharacteristically small precipitation anomalies in northeastern Australia during this event, including the unusual occurrence (for El Niño years) of above-normal precipitation there during the peak of the wet season.

5. Summary and future plans

Monthly analyses of the CAMS–OPI merged rain gauge–satellite precipitation estimates for the period January 1979 to the present are available from the Climate Prediction Center (http://www.cpc.noaa.gov). At present, the only remotely sensed estimates of precipitation that are used by the method are those generated from OLR by the OPI algorithm. We plan to augment the remotely sensed estimates with the inclusion of microwave sounding unit data, which will provide more physically based estimates of rainfall than can be derived from the OPI method alone. We also plan to explore the possibility of reducing the accumulation period from monthly to pentad, or perhaps to daily.

Acknowledgments. The authors would like to express their appreciation to George Fulwood for helping us to obtain the daily rain gauge data; Gerry Bell, Mike Halpert, and Wayne Higgins for their assistance with the gamma distribution computations and display; and to Vern Kousky, Chet Ropelewski, and Phil Arkin for their insightful comments on the methodology. Appreciation is also extended to two anonymous reviewers, particularly Reviewer ''B,'' whose suggestions greatly improved the content of this note. This work was partially funded by the NASA Tropical Rainfall Measuring Mission (TRMM) Contract No. S-12886-F.

REFERENCES

Conrad, V., and L. W. Pollack, 1962: *Methods in Climatology.* Harvard University Press, 458 pp.

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

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.,* **28,** 702–708.

Morrissey, M. L., and J. S. Greene, 1991: The Pacific atoll raingauge data set. Tech. Report, University of Hawaii, Manoa, Honolulu, Hawaii, 45 pp. [Available from the University of Hawaii, 2444 Dole St., Honolulu, HI 96822.]

New, M., M. Hulme, and P. Jones, 1997: A 1961–1990 mean monthly climatology of global land areas. Tech. Rep., Climatic Research Unit, University of East Anglia, Norwich, United Kingdom, 32 pp.

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate,* **1,** 75–86.

Ropelewski, C. F., J. E. Janowiak, and M. S. Halpert, 1984: The Climate Anomaly Monitoring System (CAMS). Climate Analysis Center, NWS, NOAA, Washington, DC, 39 pp. [Available from the Climate Prediction Center, Camp Springs, MD 20746.]

Shepard, D., 1968: A two dimensional interpolation function for irregularly spaced data. *Proc. 23d National Conf. of American Computing Machinery,* Princeton, NJ, 517–522.

Thom, H. C. S., 1958: A note on the Gamma distribution. *Mon. Wea. Rev.,* **86,** 117–122.

Willmott, C. J., C. M. Rowe, and W. D. Philot, 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer. Cartogr.,* **12,** 5–16.

Xie, P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858.

——, and ——, 1997: Global precipitation 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc,* **78,** 2539–2558.

——, and ——, 1998: Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. *J. Climate,* **11,** 137–164.