Rainfall Estimation in the Sahel. Part II: Evaluation of Rain Gauge Networks in the CILSS Countries and Objective Intercomparison of Rainfall Products

Abdou Ali, Abou Amani, Arona Diedhiou, Thierry Lebel

Rainfall Estimation in the Sahel. Part II: Evaluation of Rain Gauge Networks in the CILSS Countries and Objective Intercomparison of Rainfall Products

ABDOU ALI

IRD, LTHE, Grenoble, France, and Centre AGRHYMET, Niamey, Niger

ABOU AMANI

Centre AGRHYMET, Niamey, Niger

ARONA DIEDHIOU AND THIERRY LEBEL

IRD, LTHE, Grenoble, France

(Manuscript received 27 January 2005, in final form 14 June 2005)

ABSTRACT

This study investigates the accuracy of various precipitation products for the Sahel. A first set of products is made of three ground-based precipitation estimates elaborated regionally from the gauge data collected by Centre Regional Agrometeorologie–Hydrologie–Meteorologie (AGRHYMET). The second set is made of four global products elaborated by various international data centers. The comparison between these two sets covers the period of 1986–2000. The evaluation of the entire operational network of the Sahelian countries indicates that on average the monthly estimation error for the July–September period is around 12% at a spatial scale of 2.5° × 2.5°. The estimation error increases from south to north and remains below 10% for the area south of 15°N and west of 11°E (representing 42% of the region studied). In the southern Sahel (south of 15°N), the rain gauge density needs to be at least 10 gauges per 2.5° × 2.5° grid cell for a monthly error of less than 10%. In the northern Sahel, this density increases to more than 20 gauges because of the large intermittency of rainfall. In contrast, for other continental regions outside Africa, some authors have found that only five gauges per 2.5° × 2.5° grid cell are needed to give a monthly error of less than 10%. The global products considered in this comparison are the Climate Prediction Center (CPC) merged analysis of precipitation (CMAP), Global Precipitation Climatology Project (GPCP), Global Precipitation Climatology Center (GPCC), and Geostationary Operational Environmental Satellite (GOES) precipitation index (GPI). Several methods (scatterplots, distribution comparisons, root-mean-square error, bias, Nash index, significance test for the mean, variance, and distribution function, and the standard deviation approach for the kriging interval) are first used for the intercomparison. All of these methods lead to the same conclusion that CMAP is slightly the better product overall, followed by GPCC, GPCP, and GPI, with large errors for GPI. However, based on the root-mean-square error, it is found that the regional rainfall product obtained from the synoptic network is better than the four global products. Based on the error function developed in a companion paper, an approach is proposed to take into account the uncertainty resulting from the fact that the reference values are not the real ground truth. This method was applied to the most densely sampled region in the Sahel and led to a significant decrease of the raw evaluation errors. The reevaluated error is independent of the gauge references.

1. Introduction

Various rainfall products, widely available through the Internet, are increasingly used as standard input to run hydrological and crop models or to validate atmospheric model outputs. In general, these products combine information from satellites—infrared imagers (IR), microwave imagers [such as the Special Sensor Microwave Imager (SSM/I)], radars [i.e., Tropical Rainfall Measuring Mission (TRMM)]—and/or rain gauges. In recognition of the necessity to evaluate these rainfall products as rigorously as possible, several studies (e.g., Morrissey and Greene 1993; Chiu et al. 1993; Xie and Arkin 1995; Ebert and Manton 1998; Bell et al. 2001;

Corresponding author address: Abdou Ali, B.P. 11416, IRD, Niamey, Niger.
E-mail: a.ali@agrhymet.ne

Adler et al. 2001; Ha et al. 2002) have assessed the quality of products from different parts of the world and have shown that the performance of algorithms and the added value of rainfall products may depend on the region concerned (Kummerow et al. 2004).

For at least three main reasons, such a validation is especially needed in West Africa: 1) the severe drought that started in the beginning of the 1970s has made precipitation more irregular and more difficult to estimate, 2) after reaching a peak in the early 1980s, in situ networks are now becoming less dense, and 3) West Africa is one of the world's most sensitive regions in terms of precipitation variability because its economy is still largely dependent on rain-fed agriculture.

Several papers that address rainfall estimation in Africa were published in recent years, such as those of Thorne et al. (2001) for South Africa and McCollum et al. (2000) in equatorial Africa. In West Africa, Laurent et al. (1998) compared and discussed the validation of estimates resulting from five satellite methods, and Ramage et al. (2000) compared rainfall products with different resolutions in Niger. Jobard (2001) reviewed satellite-based rainfall retrieval using multisource data. However, a comprehensive evaluation of various rainfall products remains to be undertaken for Africa. The work presented here may be considered as one step in that direction. Focusing on the Sahel, it adds to previous work in several ways. First, it explicitly addresses the problem that the reference values are not the real ground truth (section 2). Second, as explained in section 3, this study covers a longer period (1986–2000) than any previous work carried out in West Africa and thus addresses a much larger range of climatological conditions. Third, based on the error function developed in a companion paper (Ali et al. 2005), this study examines both regional products—three products elaborated regionally from the gauge data collected by Centre Regional Agrometeorologie–Hydrologie–Meteorologie (AGRHYMET)—and global products—elaborated by various international data centers (sections 4 and 5). Last, a way of incorporating the uncertainty associated with reference rain fields into the assessment of the quality of rainfall estimates is tested in section 6.

2. Problem statement and method

a. Generalities

Satellite-based rainfall estimates have been an important research topic since the first operational methods were proposed in the early 1980s [e.g., from the initial review of Barrett and Martin (1981) to the more recent analyses of Petty (1995) that address satellite rainfall estimation over land and Levizzani et al. (2001) that look at future perspectives]. Two important points must be considered with respect to satellite-based estimate evaluation and intercomparison. First, satellite-based estimates are areal values whereas ground-based references are point values, raising the question of the significance of statistics computed using a reference estimate instead of the unknown true value. The second concern is linked to the objectives of the evaluation and intercomparison, raising the question of which statistics should be used—in other words, which products give the best estimates and in what ways are these estimates better? Several statistics will be used here to substantiate this analysis. Estimate evaluation has traditionally laid emphasis on bias (correspondence between mean values), cofluctuation (the strength of the linear relationship between the estimate and a reference value), and accuracy (the point-by-point level of agreement between the estimate and the reference). However, other performance characteristics are important in assessing the quality of an estimate. For example, a difference in the frequency of low values for two products having similar mean statistics is of great importance for agricultural applications. On the other hand, the high extreme values are of greater importance in hydrology.

b. Relationship between statistics computed using a reference value and those of the unknown true value

The most usual statistics used in rainfall product evaluation relate to the bias b, the root-mean-square error (rmse), and the correlation coefficient r. For more details and discussion of these standard statistics, see, for example, Stanski et al. (1989). These statistics are, in general, computed from an estimated reference value (b_{SR}, rmse_{SR}, and r_{SR}). In the following, the bias and the root-mean-square error are computed using the true reference value (b_{ST} and rmse_{ST}). Because this true reference is unknown, the statistics will be expressed as a function of those derived from the estimated reference value.

Considering that R_S is the estimate under evaluation, R_T is the unknown true reference value, and R_R is the estimated reference value, b_{ST} will be the bias of R_S with respect to R_T, b_{SR} will be the bias of R_S with respect to R_R, and b_{RT} will be the bias of R_R with respect to R_T. Also, rmse_{ST}, rmse_{SR}, and rmse_{RT} will be the respective root-mean-square errors. Thus,

$$b_{ST} = E(R_S - R_T) = E(R_S - R_R) + E(R_R - R_T) = b_{SR} + b_{RT}. \quad (1)$$

So, if the reference value is unbiased, the bias of the estimate R_S with respect to this reference value is the true bias. One can denote the following to express rmse_{ST}:

$$\Delta_{SR} = R_S - R_R,$$
$$\Delta_{ST} = R_S - R_T,$$
$$\Delta_{RT} = R_R - R_T,$$
$$\sigma^2_{\Delta SR} = E(R_S - R_R)^2 = (\text{rmse}_{SR})^2,$$
$$\sigma^2_{\Delta ST} = E(R_S - R_T)^2 = (\text{rmse}_{ST})^2, \quad \text{and}$$
$$\sigma^2_{\Delta RT} = E(R_R - R_T)^2 = (\text{rmse}_{RT})^2.$$

Also, because $R_S - R_R = (R_S - R_T) - (R_R - R_T)$, we have

$$\sigma^2_{\Delta SR} = E(R_S - R_T)^2 - 2\,\text{Cov}(R_S - R_T, R_R - R_T) + E(R_R - R_T)^2 - 2E(R_S - R_T)E(R_R - R_T) \tag{2}$$

or

$$\sigma^2_{\Delta ST} = \sigma^2_{\Delta SR} - \sigma^2_{\Delta RT} + 2\,\text{Cov}(\Delta_{ST}, \Delta_{RT}) + 2b_{ST}b_{RT}. \tag{3}$$

The relationships above show how the bias and the random error of the reference value can have a significant influence on the computation of the raw statistics. The evaluation of the reference value is thus a crucial part of any objective validation and intercomparison exercise.

Using these relationships, two approaches are possible. First, one can directly base the analysis on the raw statistics computed using the estimated reference values, without taking into account their uncertainty. This means that more confidence is given to the reference value than to the evaluated products. This condition is acceptable if the reference value error is much smaller than the expected error of the evaluated product.

This is, for example, the implicit assumption of Nicholson et al. (2003) in their comparison of rain products, which does not include a quantification of the possible error of the reference. This makes it difficult to assess the degree of significance of the statistics obtained. Thorne et al. (2001), on the other hand, evaluate the reference uncertainty to draw attention to the significance of their statistics but do not use it directly in their calculation.

The second approach is more objective but also more complicated. The statistics (b_{ST}, rmse_{ST}) are related to the true unknown values. With consideration of Eqs. (1) and (3), this approach first requires estimation of the following associated statistics: b_{RT}, rmse_{RT}, and $\text{Cov}(\Delta_{ST}, \Delta_{RT})$. Correction of the raw rmse (rmse_{SR}) has been attempted by some authors. Barnston (1991), Ciach and Krajewski (1999), and Gebremichael et al. (2003), for instance, have used Eq. (3) in which they assume $\text{Cov}(\Delta_{ST}, \Delta_{RT})$, b_{ST} and b_{RT} to be equal to zero. The method has been called error variance separation (EVS), and Eq. (3) becomes

$$\sigma^2_{\Delta ST} = \sigma^2_{\Delta SR} - \sigma^2_{\Delta RT}. \tag{4}$$

In others words, the estimate product errors and reference product errors are assumed to be independent. Because the products under evaluation here often consist of a merging of rain gauge data and satellite information, the reference rain gauge and regional/global products have some common information, which implies that $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ may not be equal to zero and its influence may be significant in Eq. (3). The importance of $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ may depend on the relation between the gauges used for validation and the gauges that are included in the regional and global products. This situation is a critical point for large-scale validation exercises of global products in West Africa, because the available rain gauge network generally used to create the reference value is often not dense and is not very different from the network used to create the global products. Moreover, the non-Gaussianity of rain fields accentuates the effects of the computing bias. The application of the assumption that the error covariance is equal to zero must be considered with much attention. We analyze below how to account for ground-based sampling error in the evaluation error of rainfall products, that is, how rmse_{ST} can be computed by application of Eq. (3) without neglecting the covariance. The results obtained with and without taking into account the reference uncertainty will be compared.

c. Estimation of the terms of Eq. (3)

To assess the rmse_{ST} from Eq. (3), the terms b_{RT}, rmse_{RT}, and $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ are estimated as follows.

1) ESTIMATION OF THE TERM RMSE_{RT}

The question of computing rmse_{RT} has been treated in Ali et al. (2005, hereinafter Part I). A cross-validation procedure has been used to investigate different kriging methods designed to improve the calculation of the ground-based error (rmse_{RT}). A relative error function [Eq. (5) below] has been established and will be used here for the computation of rmse_{RT}:

$$e(A, N_g, K_T, P_T) = \frac{1.05}{\sqrt{N_g}\sqrt{K_T}}\left(\frac{P_T}{K_T}\right)^{-0.2} \times \left[0.28 + 0.17 \log\left(\frac{A}{N_g}\right)\right], \quad (5)$$

where A is the grid area in square kilometers, N_g is the number of rain gauges in A, K_T is the number of events during the considered period, P_T is the cumulative rainfall in millimeters over the same period, and e is the relative estimation error of the mean areal rainfall.

A major variable concerning the error variability in this error model is the number of events, which is the most important explanatory factor of the Sahelian rainfall variability (Le Barbé et al. 2002). Several approaches to determine the number of events on a regional scale, where only daily data are available, are discussed and validated in Part I.

2) ESTIMATION OF THE TERMS b_{RT} AND COV(Δ_{ST}, Δ_{RT})

(i) The bias: b_{RT}

Geostatistical analysis is an unbiased interpolation method. The evaluation procedure in the companion paper shows that the bias is practically equal to zero. So, if the point measurement is assumed to represent the "true" rainfall, the areal rainfall bias can be assumed to be equal to zero. Otherwise the bias of the areal rainfall is equal to the point bias.

(ii) Cov(Δ_{ST}, Δ_{RT})

Because a direct computation is not possible here, we analyze on the one hand the possibility of neglecting this term. This option consists of an application of EVS [Eq. (4)]. On the other hand, an area with dense network coverage D will be chosen for the assessment of this covariance. The estimation R_D from this dense network can be assimilated to R_T; then Cov(Δ_{ST}, Δ_{RT}) will be estimated as Cov(Δ_{SD}, Δ_{RD}). In section 6 the importance of the estimated covariance will be analyzed and the two approaches will be compared.

3. Sahelian rainfall products: Regional and global

a. Regional products

The AGRHYMET rain gauge database (see Part I for details) is used here to study various regional ground-based products. Three kinds of networks (Fig. 1) are considered: 1) the whole operational network (Comité Inter-Etats de Lutte Contre la Sécheresse au Sahel; average of 650 rain gauges with data available at the end of the rainy season), hereinafter referred to as "CILSS," 2) the monitoring network (about 250 rain gauges with data available on a 10-day basis at the AGRHYMET Regional Center), hereinafter referred to as "CRA," and 3) the synoptic network [80 rain gauges with data available on a daily basis through the Global Telecommunications System (GTS)], hereinafter referred to as "SYNOP." The CRA network is included in the CILSS network, and the SYNOP network is included in the CRA network. In accordance with the resolution of the global products presented below, regression kriging is used to create a regional rainfall product based on a 2.5° × 2.5° grid for these three networks over the period of 1950–2002.

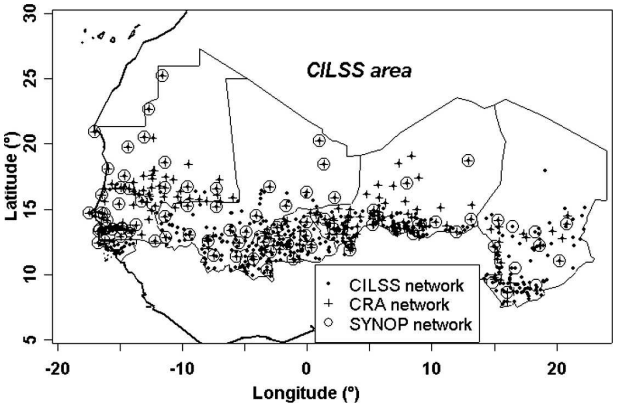

FIG. 1. The operational networks for the Sahelian region. The CILSS network (650 gauges on average) is available at the end of the rainy season at the AGRHYMET Center, the CRA network (280 gauges) is available every 10 days, and the SYNOP network (80 gauges) is available daily.

b. Global products

1) THE GPCC AND GPCP OPERATIONAL DATASETS

The Global Precipitation Climatology Center (GPCC) produces monthly gridded 2.5° × 2.5° rainfall estimates based on gauge-only data from the GTS and other records (Rudolf 1993); this product is referred to as GPCC. The Global Precipitation Climatology Project (GPCP) rainfall product combines different information (gauges and satellite IR and microwave observations). The GPCP product is described in Huffman et al. (1997). The primary GPCP product is a monthly analysis based on a global 2.5° latitude × 2.5° longitude grid spanning the period from January 1979 up to the present date. A second product is a 5-day global analy-

sis (Xie et al. 2003), and a third is a daily 1° latitude × 1° longitude analysis from January 1997 to the present date (Huffman et al. 2001).

2) THE CMAP DATASET

The Climate Prediction Center merged analysis of precipitation (CMAP) was constructed by Xie and Arkin (1997). The temporal and spatial scales are similar to those of GPCP, except that there is no daily CMAP product. The first step is to combine the various satellite estimates based on IR, SSM/I, and Microwave Sounding Unit data using a maximum likelihood approach in which weighting coefficients are inversely proportional to the squares of the individual random errors. The resulting satellite-based analyses provide the field shape, which is then adjusted to the rain gauge data (where available).

3) THE GPI

The Geostationary Operational Environmental Satellite precipitation index (GPI; Arkin et al. 1994) assumes that a 3-mm rainfall occurs every hour of cold-cloud duration over a pixel. Because of its simplicity, GPI is still widely used for climatological studies of global precipitation. This is the only *pure* satellite product among the four global products studied here.

4. Evaluation of the CILSS rain gauge networks

a. Sahelian operational network error analysis

Several authors (e.g., Rudolf et al. 1994; Morrissey et al. 1995; Xie and Arkin 1995) suggest that five rain gauges over a 2.5° × 2.5° cell are sufficient to guarantee an estimation error of less than 10% (percentage error is calculated with respect to the mean of the variable considered) for the areal monthly rainfall over the cell. This result does not hold for the Sahelian region, as may be seen from the study of Lebel and Amani (1999). In fact, as clearly shown in Eq. (5), the estimation error depends both on the monthly total and the number of recorded rain events in this total. The greater the number of events for a given total is, the greater is the smoothing of the event-scale spatial variability and, consequently, the smaller is the estimation error.

This is illustrated well in Fig. 2a, showing the dispersion of estimation errors computed for the same number of gauges available over a 2.5° × 2.5° cell using Eq. (5). For five gauges over a 2.5° × 2.5° grid, the error varies from 5% to 30%, depending on the area and month considered. The error is maximum for dry months in the northern part of the domain and is minimum in the south of the domain where, on average, a larger number of rain events are recorded. This dispersion factor is amplified by the fact that, as seen in Fig. 1, the network is far from being homogeneous in terms of density. Even when limiting the computation to the 3 months of the core of the rainy season [July–September (JAS)] so as to obtain a more homogenous sample, the dispersion of errors over a total sample of 1287 values (39 cells × 3 months × 11 yr) is still significant, with 80% of the 2.5° × 2.5° errors in the range of 5%–22%. Averaged over the whole period of study and JAS, the mean monthly error is 12% for the entire Sahel but only 9% for latitudes below 15°N. A similar computation performed for 1° × 1° grids shows that 80% of the errors at that resolution range from 8% to 28%.

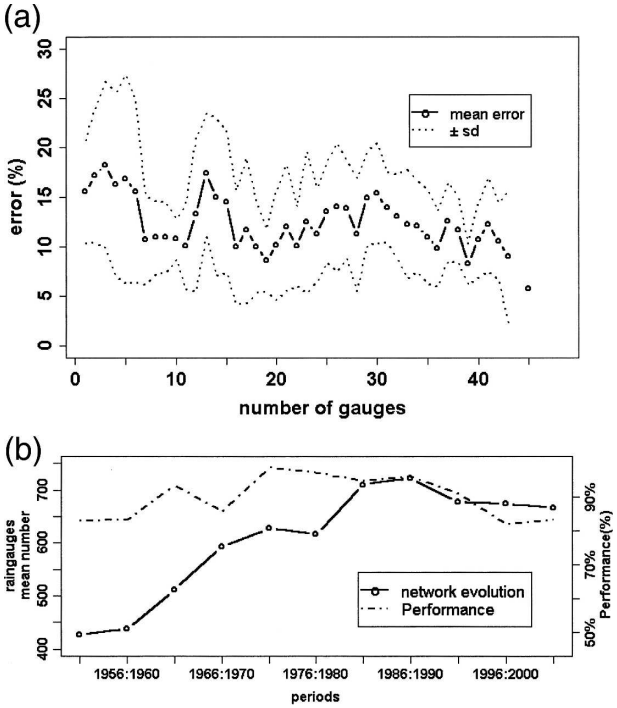

FIG. 2. (a) Relationship between the estimation error and the number of operational gauges. There is a significant dispersion for the same number of gauges (e.g., for five gauges over 2.5° × 2.5°, the error varies from 5% to 30%). (b) Mean number of gauges and the performance (percentage of cells for which the monthly estimation error over 2.5° × 2.5° is less than 20%) of the network for the period of 1950–2002. Note that the rapid increase in the number of gauges during the dry period (1970–90) did not involve a better performance (smaller estimation error), because in dry years rainfall is more variable in space and thus requires a comparatively larger number of gauges to guarantee a given level of accuracy.

The overall performance of the CILSS network is presently decreasing, as shown in Fig. 2b. This decrease is the result of both the decrease in the number of

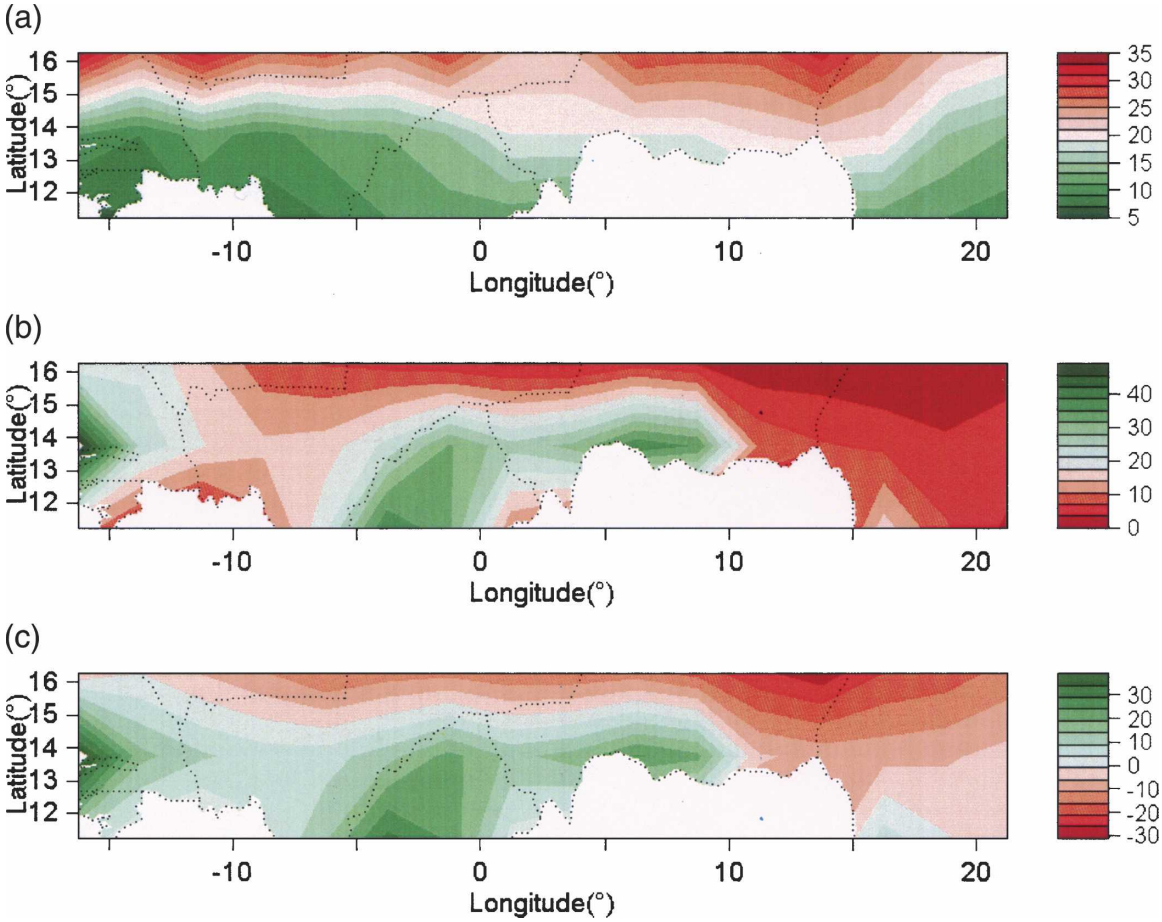

FIG. 3. (a) Mean optimal number of gauges per grid cell (2.5° × 2.5°), which provide the estimation error of less than 10% at the monthly scale for the period 1990–2000. (b) Actual mean number of gauges of the CILSS operational network for the same period. (c) The difference between (a) and (b). The CILSS operational network guarantees a monthly error smaller than 10% for 42% of the grid cells, located south to 15°N and west of 11°E.

gauges since the optimum of the mid-1980s and the decrease in the monthly rainfall of August and September that characterize the drought.

b. *Optimal networks*

One application of the error function is the determination of optimal networks. The optimal network concept depends on the scale of interest [time K_T and space A in Eq. (5)]. The criterion of optimality corresponds to the accepted error threshold. For a fixed error, the inversion of Eq. (5) makes it possible to compute the minimum number N_g^* of stations needed to remain below the error threshold. The error threshold e considered here is 10%. The analysis will focus on the 2.5° × 2.5° resolution [$A = 75\,000$ km^2 in Eq. (5)] and monthly rainfall corresponding to the resolution of the most widely used operational rainfall products.

For the JAS period, it is found that for the period of 1990–2000 the mean optimal number of rain gauges over a 2.5° × 2.5° cell is 14.6, corresponding to a total of 569 gauges (Fig. 3a). By comparison, the network available over this whole period (shown in Fig. 3b) comprises 561 gauges. However, this rough comparison is not really meaningful given the strong heterogeneity in the statistics. The optimal number of gauges per 2.5° × 2.5° cell is six in the southwest (south of 13°N and west to 5°W), 15 for the central part, and 35 in the north (north of 16°N). This variability in the optimal number of rain gauges is linked to the variability in the number of events $\overline{K}_T$ as may be seen in Table 1. The network evaluation of Fig. 3c shows that the only area that is in fact optimally covered is the one located south of 15°N and west of 11°E, representing 42% of the CILSS region. For example, at 11°N, the mean optimal number of rain gauges is 8.3, with the actual number being 18.3. At 16°N, the optimal number is 20.6 and the actual number is 5.9.

TABLE 1. Details on the characteristics of optimal networks and the evaluation of the CILSS operational network for different areas and for two scales: monthly values over 2.5° × 2.5° grid cells and 10-day values over 1° × 1° grid cells. The only area over which the CILSS network is optimal everywhere for the 2.5° × 2.5° resolution is the area lying south of 15°N and west of 11°E; N_g^* is the average number of gauges per cell satisfying the optimality criterion, and N_g^{real} is the average number of gauges per cell of the CILSS network over the subarea. Optimality corresponds to grid cells for which the error is less than 10%. Italics denote values for which the existing network is less than the optimum number of gauges.

Area	$\overline{K}_T$	N_g^*	N_g^{real}	$(N_g^{\text{real}} - N_g^*)/N_g^{\text{real}}$	No. of optimal cells
Monthly values (JAS) 2.5° × 2.5°					
Total	11.8	14.6	14.4	*−1%*	42%
10°–12.5° lat	16.3	8.2	18.3	55%	72%
12.5°–15° lat	12.1	11.2	21.1	47%	69%
15°–17.5° lat	7	20.6	5.9	*−248%*	0%
<15° lat and west of 11°E	11.27	9.9	27.1	82%	100%
10-day values 1° × 1°					
Total	5.6	16.7	2.6	*−538%*	1.5%
10°–13°lat	7.33	11	3.4	*−223%*	2%
13°–15° lat	6	14.25	3.6	*−296%*	2.8%
15°–18° lat	3.9	23.4	1.2	*−1900%*	0%
<15° lat and west of 11°E	6.8	12.2	4.5	*−169%*	3.2%

At the smaller scales of 10 days and 1° × 1°, the number of required stations over the whole region increases to 3559. To comply with the 10% error criterion only for the area located south of 15°N, one would need 1736 stations. Only 1.5% of the cells have a number of rain gauges above the optimum number (see also the Table 1 for details).

A rough rule of thumb concerning times scales can be proposed: on average, a network producing a 20% error on a monthly basis provides estimates with a 30% error on a 10-day basis and a 10% error on a seasonal basis at 1° × 1° scale.

5. Global product evaluation: The CILSS product as a reference value

This section uses the classical approach to evaluate global products. The optimal estimates from the CILSS network are used as reference values, not taking into account their uncertainty. The scale considered is monthly and has a 2.5° × 2.5° spatial resolution.

a. Interannual and intraseasonal homogeneity analysis

Before combining any monthly or annual data for global analysis, we carry out a systematic trend analysis to check for homogeneity. This section aims at determining whether there is a seasonal cycle in the various statistics used for the evaluation and also whether there are systematic differences between wet and dry years. The period of 1986–99 is considered, that is, the period over which the GPCC product was available. Figures 4a(1), 4a(2), and 4a(3) show that there is indeed a seasonal cycle for all three indicators. The rmse and the bias are minimum in the middle of the rainy season and, inversely, the correlation is maximum in the middle of the season. These results show that, for the sake of homogeneity, the core and the margin of the season must be separated. Figures 4a(1), 4a(2), and 4a(3) also show that CMAP has a good score in the core of the season whereas GPCC and GPCP are best in the margins. GPI has the poorest score for the whole season. In view of these results, the analysis will be restricted to the months of JAS. Figures 4b(1), 4b(2), and 4b(3) show that there are no systematic differences between wet and dry years, which indicates that wet and dry years can be mixed in a global analysis covering the whole 1990–99 period. Another result of this analysis is that the rmse has increased in recent years for all products. The nonexistence of this trend for the bias suggests that this deterioration is due to an increased sampling error. The decrease of the network density observed in Fig. 2 for recent years supports this assumption. We can also note a degrading trend in the correlation coefficient after 1992 for all products, which can be explained by the same reason as that for the rmse.

b. Global analysis of the JAS period

1) SCATTERPLOTS AND HISTOGRAMS

Figure 5 is a joint representation of the linear relationship between the global products and the CILSS product and the experimental distribution of monthly areal rainfall for different products. Note that for all products, the regression line is under the bisecting line for high values, suggesting an underestimation with respect to the ground estimate. Also, the regression line tends to be over the bisecting line for low values, suggesting an overestimation. Note that the GPI distribution is far less asymmetric than the other distributions because of the way it is computed.

2) COMPARISON OF DISTRIBUTIONS

Apart from GPI, for which the 25% quartile is roughly equal to the 50% quartile of the other products, the box plots (Fig. 6a) of the different products are

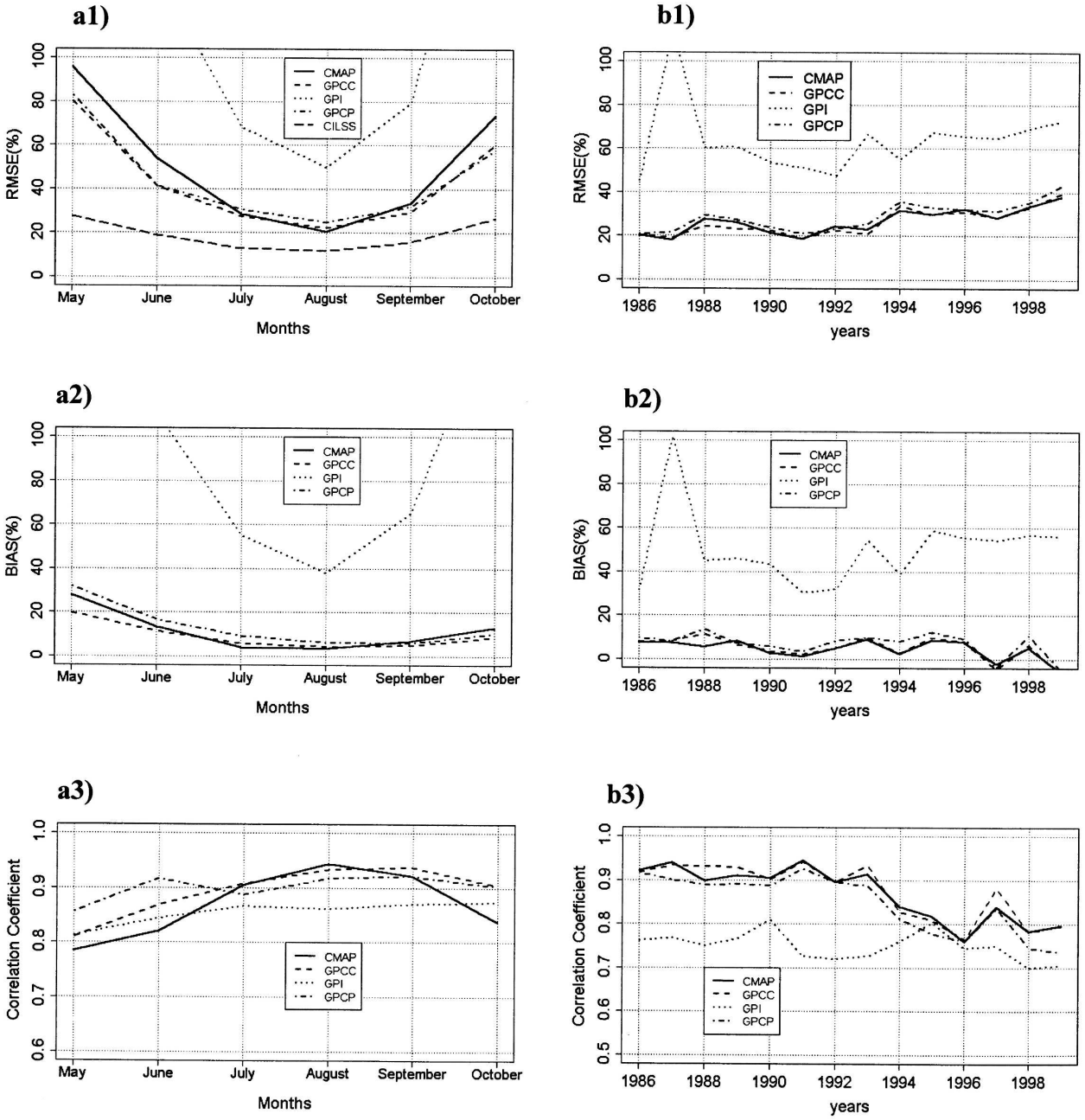

FIG. 4. (a) Monthly mean values for the three statistics used to compare each global product with the CILSS reference [(a1) rmse, (a2) correlation coefficient, and (a3) bias]. The minimum rmse and bias and the maximum correlation occur in the core of the rainy season (i.e., Aug). All three statistics show a systematic intraseasonal pattern. (b) Annual mean values computed from JAS monthly values only for the three statistics used to compare each global product with the CILSS reference [(b1) rmse, (b2) correlation coefficient, and (b3) bias]. In general, the statistics do not show any systematic difference between wet and dry years. However, the rmse shows an increase in recent years.

similar. However, the 25% quartiles of CMAP, GPCC, and GPCP are slightly higher than that for CILSS. The interquartile intervals (25%–75%) of CILSS and CMAP are very close and that of GPCC is close to that of GPCP. For the empirical distributions (Fig. 6b), all of the products underestimate the frequency of the low values, as compared with CILSS, and they overestimate the frequency of medium values. The proportion of values below the 25% quartile is 17% for CMAP, 13% for GPCC, 13% for GPCP, and 6% for GPI. This under-

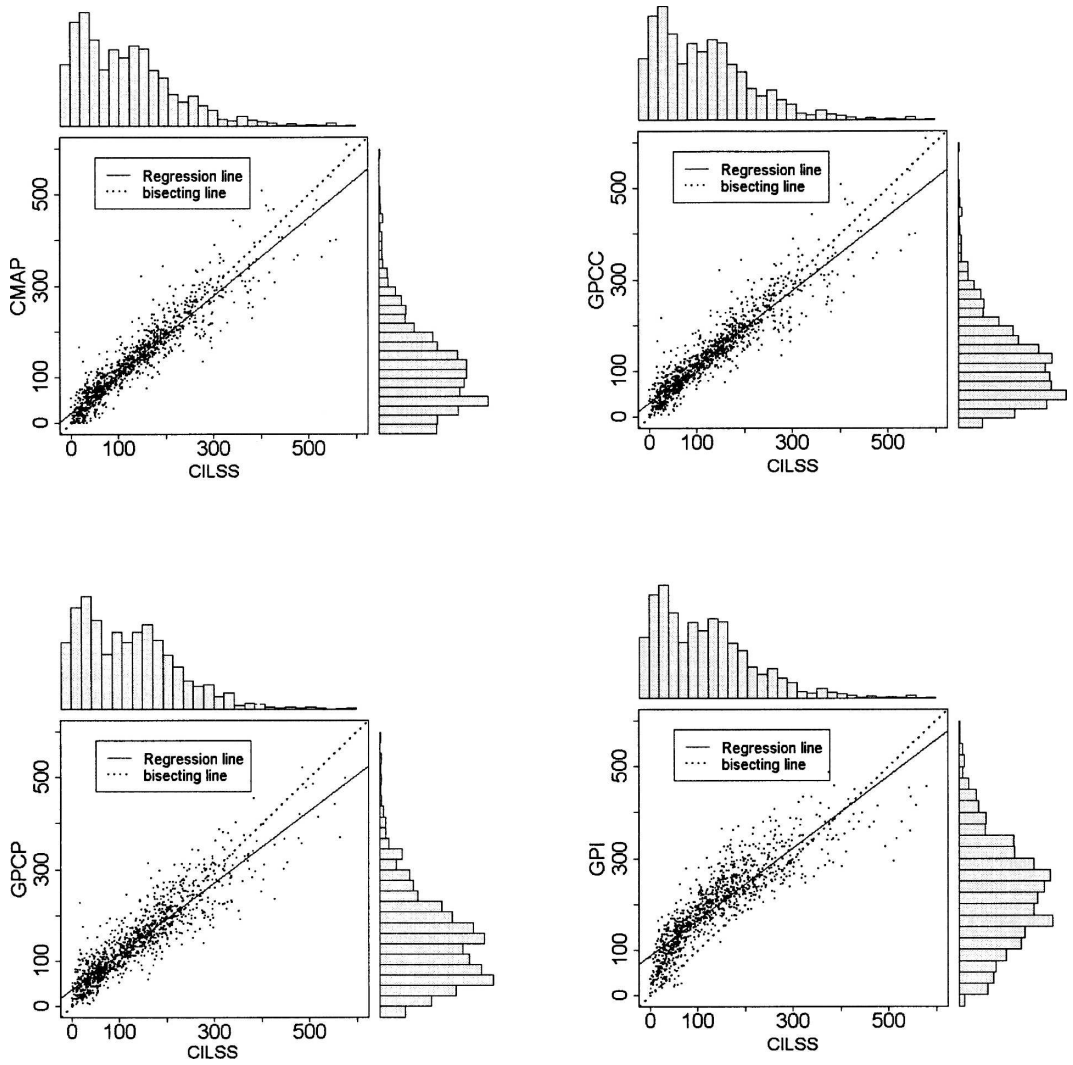

FIG. 5. A joint analysis of scatterplots and histograms for the four global products (CMAP, GPCC, GPCP, GPI) with respect to CILSS. The values represent the JAS monthly rainfall between 1990 and 1999.

estimation of the frequency of low values is due to the inability of satellites to correctly estimate intermittency. For the extremes values, apart from GPI, which continues to be biased, the product distributions are similar. For example, the probability of exceeding 320 mm, which is the CILSS's 95% quartile, is 5%, 4%, 3.7%, and 3.3% for CILSS, CMAP, GPCC, and GPCP, respectively. For GPI, 11% of its values are over the CILSS's 95% quartile.

Figure 7 shows that the major discrepancy between the distribution of the global product and that of the CILSS is observed for the northern part of the domain (north of 15°N), with the GPI distribution being completely unrealistic. In this area the ground-based estimate is also highly uncertain (theoretical kriging error equal to 34%).

3) NUMERICAL CRITERIA: RMSE$_{SR}$, b_{SR}, AND THE NASH INDEX

For this analysis, the Nash index I_{SR} presented below is considered as an additional statistic to the rmse, bias, and correlation criteria, because it is not influenced by the bias, as is the correlation:

$$I_{SR} = 1 - \frac{\text{rmse}_{SR}^2}{\sigma_R^2}. \tag{6}$$

The Nash index is equal to 1 for a perfect estimate [i.e., $(R_S)_i = (R_R)_i$ for all i] and is equal to 0 if, for all i, $(R_S)_i = \overline{R}_R$.

Table 2 shows that, in fact, CMAP, GPCC, and GPCP have a very low bias (less than 5%), while GPI

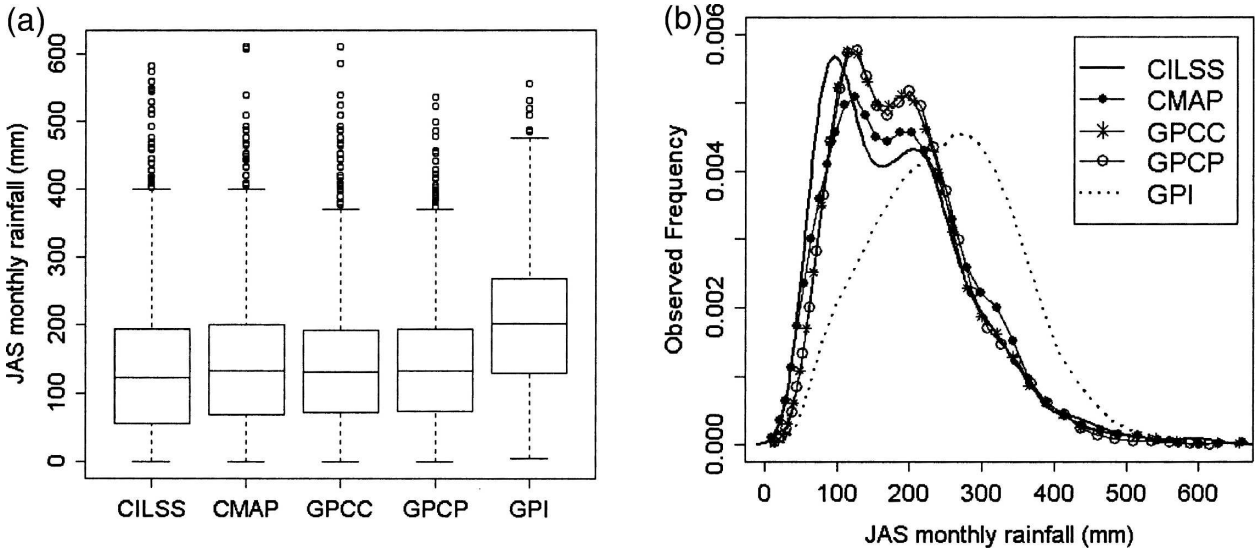

FIG. 6. Comparison of (a) box plots and (b) observed probability distribution function for the CILSS reference and the four global products for the JAS period over the entire CILSS region. The global products overestimate the frequency of the median while the frequency of low values is underestimated.

is highly biased (43%). This, of course, is related to the fact that CMAP, GPCC, and GPCP incorporate ground observations. Because of this low bias the other statistics computed for these products give similar results, with CMAP being slightly better, with 2% and 25% in terms of b_{SR} and rmse_{SR}, followed by GPCC (4% and 27%), GPCP (5% and 28%), and GPI (43% and 55%).

In analyzing these statistics along latitudes, the conclusion is the same as that for the analysis of distributions: the quality of the products is far worse in the northern part of the domain. The minimum errors (rmse_{SR}) are obtained for the medium strip (12.5°–15°N): 16% for CMAP, 17% for GPCC, 21% for GPCP, and 53% for GPI, against 24%, 22%, 23%, and 44%, respectively, for the southern strip (10°–12.5°N). For the northern strip (15°–17.5°N), the errors are very high: 55%, 69%, and 80% for CMAP, GPCC, and GPCP, respectively, and 130% for GPI.

c. *Significance test for the statistics*

Because the statistical distribution of rain fields is not Gaussian, nonparametrical statistical tests are used to verify the equality of the means, variances, and cumulative distribution functions (CDF) of global products and the reference regional product (CILSS). The tests are the Wilcoxon *U* test (abbreviated as *W*; Hollander and Wolfe 1973) for the equality of the means, the Ansari–Bradley (Hollander and Wolfe 1973) and Mood (Conover 1971) tests for the equality of the variances, and the Kolmogorov–Smirnov (KS) test for the equality of the CDFs.

A p value of 10% is used to accept or reject the

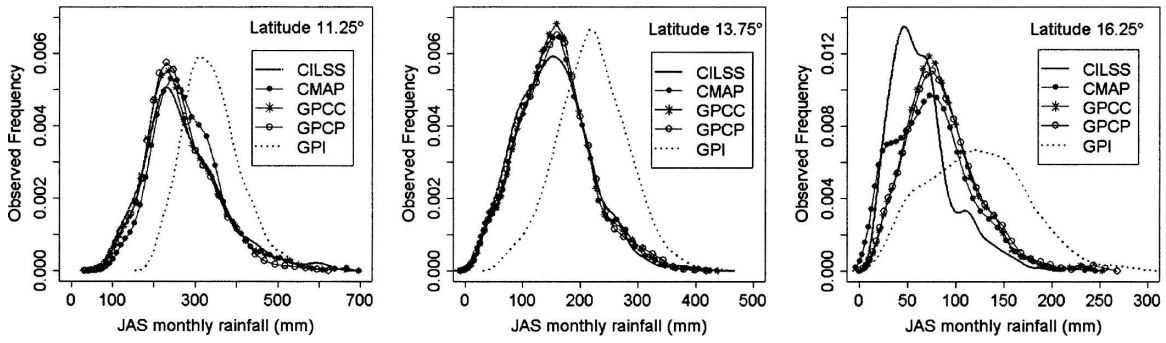

FIG. 7. The observed probability distributions as a function of latitude. At 11.25° and 13.75°N, apart from the GPI product, the other distributions are similar and display the same mode. At 16.25°N all of the distributions become very different from the reference CILSS distribution.

TABLE 2. Average statistics over the entire Sahel when using the CILSS rainfall as a reference, not taking into account its uncertainty.

	CMAP	GPCC	GPCP	GPI
Rmse (%)	25	27	28	55
Bias (%)	2	4	5	43
I	0.88	0.86	0.76	0.42
R^2	0.88	0.86	0.85	0.77

hypothesis tested. The calculations are performed so as to test the homogeneity of statistics with respect to time [the means are computed over 3 (months) $\times$ 14 (years) samples with 39 cells both for CILSS and the product under consideration] and with respect to space [the means are computed over 39 (cells) samples of 3 (months) $\times$ 14 (years)].

The results, summarized in Table 3, show that the temporal fields are more homogeneous than the spatial fields. The percentage of cases accepted is larger for the temporal fields than for the spatial fields, except for the variance. The equality of means and CDFs is accepted for more than 95% of the temporal fields for CMAP, GPCC, and GPCP. CMAP also performs very well for the equality of variances, both in time and space, confirming that it is the closest product to the CILSS reference. Except for the variance, it can be concluded that the global products do a better job in reproducing the seasonal cycle than in accounting for the spatial variability of the monthly rainfall. GPI is an interesting case to analyze because, as already seen previously, its CDF and mean are very different from the CILSS references. However, GPI performs as well as the other products—and even better than GPCC and GPCP—in terms of the equality of the variances. This means that products using only the satellite information may correctly estimate the rain field variances at the scale tested here, confirming the finding of Mathon et al. (2002), that satellite is able to distinguish rain events; because in the Sahelian region the variability of rain fields is mainly due to the variability in the number of events.

d. Interval of kriging standard deviation for performing the analysis: Nonlinear correlation

The computation performed here counts the number of times the estimate of a given rainfall product falls into the theoretical one or two standard deviation interval. The theoretical standard deviation considered is the kriging standard deviation (ksd) of estimation error of the reference. The probability of belonging to this interval must be found experimentally because rain fields are not Gaussian. A cross-validation procedure performed on the reference CILSS fields shows that this probability is 73% on a monthly basis. This approach has already been used by Thorne et al. (2001) and is a useful complement to linear statistics because it makes it possible to quantify the nonlinear cofluctuation between the products and the rain gauge–based reference:

$$P = \text{the number of times that } -\text{ksd} \le R_S - R_R \le +\text{ksd or } -2\text{ksd} \le R_S - R_R \le +2\text{ksd}.$$

In accordance with this criterion P, Table 4 shows that CMAP is still the best product, followed by GPCC, GPCP, and GPI. These results indicate some consistency between the different criteria used in this study and show the robustness of the statistics used. Figure 8 also shows that the distribution of the discrepancy between CMAP and CILSS is very close to the discrepancy between GPCC and CILSS.

6. Taking into account the ground-based error in statistics

This section addresses the implementation of a solution to the problem that was theoretically addressed in sections 2b and 2c. The objective is to analyze the degree of uncertainty associated with the classical approach used in section 5 and to obtain a supposedly closer to reality estimation of rmse_{ST}. Three analyses are performed: 1) the whole CILSS network is the reference and the gridded estimates from the two other

TABLE 3. Percentage of null hypothesis H0 assumption accepted with regard to the tests applied month by month or cell by cell for a 10% acceptance level.

	Equality of the CDF: KS test		Equality of the means: W test		Equality of the variances: Ansari–Bradley test		Equality of the variances: Mood test	
Products	Months	Cell	Months	Cell	Months	Cell	Months	Cell
CMAP	100	76	100	72	98	98	98	95
GPCC	95	71	100	74	74	93	70	88
GPI	5	9	2	2	90	88	92	90
GPCP	95	67	97	62	74	90	71	90

TABLE 4. The number of times the discrepancy D between the products and CILSS reference is within 1 or 2 times the standard kriging deviation.

P	CMAP	GPCC	GPI	GPCP
Percent ($-\text{ksd} \leq \Delta \leq \text{ksd}$)	40	38	8	37
Percent ($-2\text{ksd} \leq \Delta \leq 2\text{ksd}$)	71	69	21	66

networks (CRA, SYNOP) are considered as operational ground-based products to be compared with the global products (CMAP, GPCC, GPI, and GPCP), 2) the CRA and SYNOP networks are considered as possible alternate references and the fluctuation of rmse_{SR} as a function of the reference used is analyzed, and 3), using a densely instrumented subregion, an empirical study is done to estimate the different terms of Eq. (3) to assess rmse_{ST} as compared with rmse_{SR}.

a. Comparing regional and global products with respect to CILSS

In this section, for stationarity and homogeneity purposes, only the area south of 15°N (southern Sahel) is considered, representing 23 grid cells with an average density of 19.2 gauges per cell. As may be seen from the CILSS column of Table 5, CRA and SYNOP have a smaller error than the global products when the optimal CILSS rain fields are taken as the references. On the other hand, when neglecting the drift in the computation of the ground products (the "no drift" subcolumn

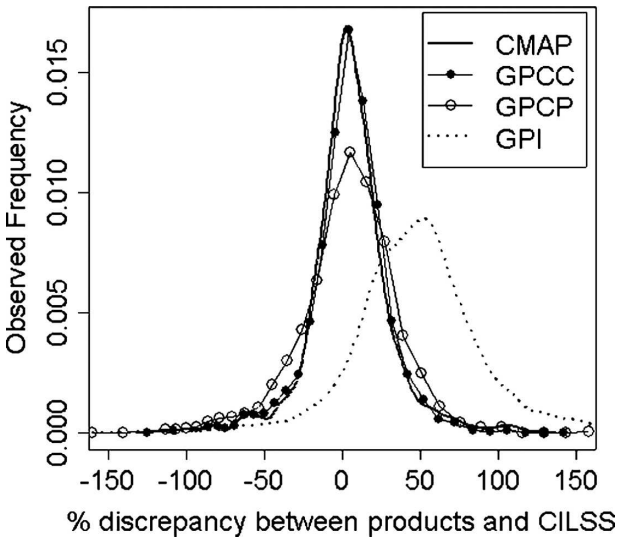

FIG. 8. Probability distributions of the percentage difference between the four global products and the CILSS reference. The distributions for CMAP and GPCC are very similar, whereas the distribution of GPCP is more dispersed. GPI shows significant differences from the other products.

TABLE 5. Estimation errors (rmse_{SR}) when using different references for the area lying south to 15°N. Values are in percent. The estimation error associated with each reference (rmse_{RT}) is 8.2% for CILSS, 12% for CRA, and 13.6% for SYNOP.

	Values of rmse_{SR} for different references			
	CILSS reference			
Products	Drift	No drift	CRA reference	SYNOP reference
CRA	14.1	18		
SYNOP	14.4	24.5		
CMAP	15.9	21	18	16
GPCC	16	25	18	15.9
GPCP	18	27	19.5	19.1
GPI	42	45	44	44.5

in Table 5), SYNOP performs worse than CMAP. This result means that 1) a blended satellite–ground product, because of the continuous space coverage of the satellite information, is able to restitute better the latitudinal gradients than can a nonoptimally interpolated ground product and 2) it is necessary to use interpolation algorithms incorporating the drift in their interpolation scheme. In terms of the best spatialization method, the ranking of the different products is thus CRA, SYNOP, CMAP, GPCC, GPCP, and GPI.

b. Comparing global products with CRA and SYNOP as references

Changing the reference to either CRA or SYNOP (columns CRA or SYNOP in Table 5) produces contrasted results. On the one hand, the CRA reference is associated with larger rmse_{SR} values of the global products, indicating that references obtained from less dense networks will tend to produce larger values of rmse_{SR}. On the other hand, for CMAP and GPCC, the rmse_{SR} values decrease when shifting from the CRA reference to the SYNOP reference (for CMAP, rmse_{SR} = 15.9% when CILSS is used as reference, 18% when CRA is considered to be the reference, and 16% when the reference is SYNOP). A likely explanation is that CMAP and GPCC include ground data provided by synoptic stations only (not all the synoptic stations, however). This clearly illustrates how the choice of the reference (both the network used and the interpolation algorithm) may influence the results of an intercomparison exercise.

c. Assessment of rmse_{ST}

The sensitivity of rmse_{SR} to the reference used underlines the necessity to evaluate the neglected terms of Eq. (3) to obtain a more credible assessment of the

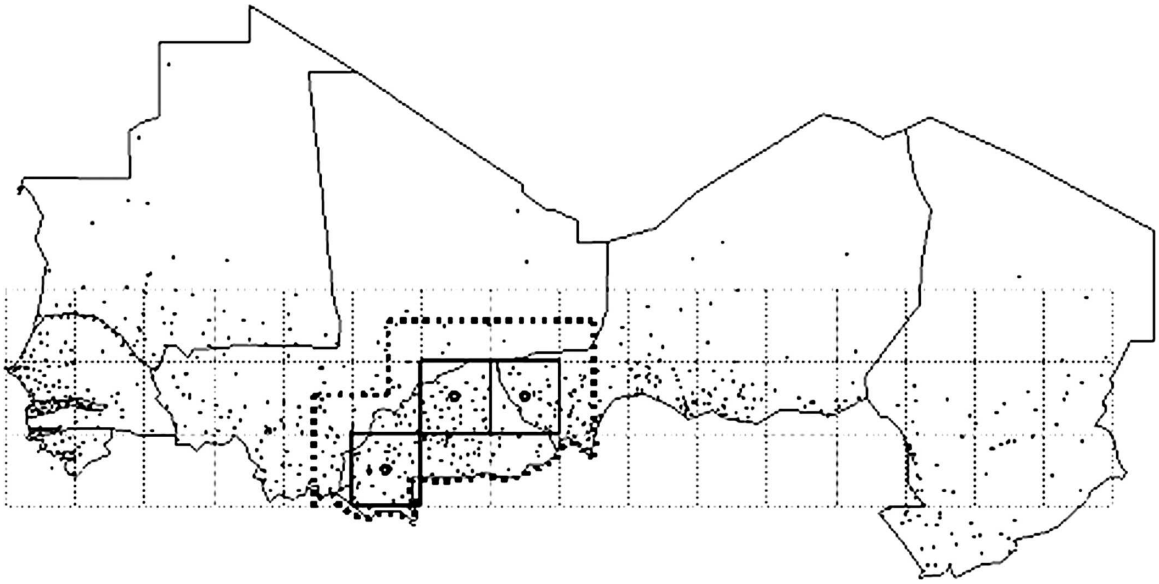

FIG. 9. The target area (thick dashed line) for the subsampling approach in the estimation of the covariance of errors. A total of 217 gauges are covering this area. Various subsampled networks are used to create ground references of varying accuracy over the three central grid meshes (solid line). The objective is to evaluate the influence of the reference on the computation of both the rmse_{SR} and rmse_{ST}.

rainfall product errors. If one assumes that the reference rain fields are unbiased, two terms remain to be evaluated as addressed in section 2c: rmse_{RT} and $\text{Cov}(\Delta_{ST}, \Delta_{RT})$. The theoretical computation of rmse_{RT} is straightforward in the statistical context used here (the results of this computation are given in the caption of Table 5 for the three references tested here). The EVS method presented in Eq. (4) only makes use of rmse_{RS} and rmse_{RT} to compute rmse_{ST}, assuming the covariance to be negligible in comparison of these terms. Because there is no straightforward way to compute $\text{Cov}(\Delta_{ST}, \Delta_{RT})$, this assumption is rarely tested. For the case treated here, the EVS computation leads to surprising results, with a CMAP EVS error equal to 13.6% when CILSS is used as reference and 8.4% when SYNOP is used. There is a clear reference dependency in these results, pointing out the need to assess how the neglected term $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ could influence the value of rmse_{ST}. Because the gauge network used in CMAP is closer to the SYNOP network than to the CILSS network, one can reasonably assume that $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ is larger when SYNOP is the reference than when CILSS is the reference.

A way to test how $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ might influence the overall assessment of rmse_{ST} is to work on a more densely instrumented area where it is possible to build better proxys of R_T than when considering the whole Sahel. Over the target area delimited by the thick dashed line in Fig. 9, a total of 217 gauges is available, corresponding to 2 times the density of the CILSS network over the southern Sahel (19.2 gauges per 2.5° × 2.5° grid cell). To study the effects of using various references on the computation of rmse_{SR} and rmse_{ST}, a subsampling approach is used to create different reference networks of increasing density, by choosing randomly between 5 and 217 gauges with increments of 10 gauges. For each number of gauges, 21 subsample networks are created. The comparison is then carried out for the three 2.5° × 2.5°cells located at the center of this target area.

As a first step, rmse_{SR} is computed for the different reference networks, leading to the results shown in Fig. 10 [the computations are made with a sample of 14 years (1986–99) × 3 cells × 3 months × 21; i.e., 2646 values]. Except for GPI, which is less dependent on the reference used, rmse_{SR} decreases significantly until reaching a density of approximately 10 gauges over a 2.5° × 2.5° grid cell). This density is one-half of the CILSS density over the southern Sahel.

The second step is to use the denser network of 217 gauges to provide a supposedly better proxy of R_T to estimate the covariance $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ in Eq. (3). It then becomes possible to compute rmse_{ST} for all of the references of step 1. The computation of rmse_{ST} taking into account an estimation of $\text{Cov}(\Delta_{ST}, \Delta_{RT})$ is denoted as $\text{rmse}_{ST}^{\text{COV}}$; when the covariance is neglected it is denoted as $\text{rmse}_{ST}^{\text{EVS}}$. The results of the computation of these two rmse_{ST} are shown in Fig. 11

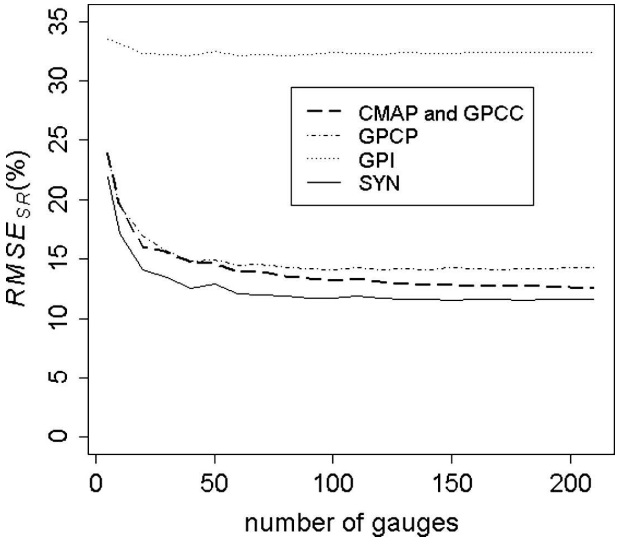

FIG. 10. The raw evaluation error of the global products ($rmse_{SR}$) as a function of the number of gauges used to compute the reference value in the dense coverage area. There is a decrease of the evaluation error for all products when the number of gauges increases. However, the effect on GPCP and, most notable, on GPI is weaker.

for CMAP, GPCC, GPCP, and SYNOP. The comparison of Fig. 10 and Fig. 11 leads to three main conclusions. 1) The behavior of $rmse_{ST}^{EVS}$ depends on the reference and the products under evaluation, whereas the value of $rmse_{ST}^{COV}$ does not depend on the gauge density of the reference and seems to be constant for a given product. 2) For all products, the values of $rmse_{ST}^{EVS}$ and $rmse_{ST}^{COV}$ are lower than those for $rmse_{SR}$. For a reference with a density of three gauges per $2.5° \times 2.5°$, $rmse_{SR}$ is approximately 18% and $rmse_{ST}^{COV}$ is 11% for CMAP and GPCC. These statistics are, respectively, 15% and 11% for a reference network with a density of 25 gauges per $2.5° \times 2.5°$. 3) The $rmse_{ST}^{EVS}$ is highly sensitive to the density of the reference, especially when it is low. This fact leads to some incoherence in the product ranking, which can change with the density of the reference network. For example, the values of $rmse_{ST}^{EVS}$ are, respectively, 11.5% and 10% for CMAP and GPCP for a 20-gauge reference network, changing to, respectively, 11% and 12.5% for a 60-gauge reference network. The behavior of $rmse_{ST}^{COV}$ is more stable and coherent than that of $rmse_{ST}^{EVS}$. For instance, when SYNOP is used as reference in this target area, $rmse_{SR}$ is 12.8%, $rmse_{ST}^{EVS}$ is 4.8%, and $rmse_{ST}^{COV}$ is 11.4% for CMAP. When CILSS is used as a reference, $rmse_{SR}$ is 15.9%, $rmse_{ST}^{EVS}$ is 14%, and $rmse_{ST}^{COV}$ is 11% for CMAP. This is just an illustration of how$rmse_{ST}^{COV}$is less sensitive to the density of the reference used than are the two other rmse, as may be seen from Fig. 11.

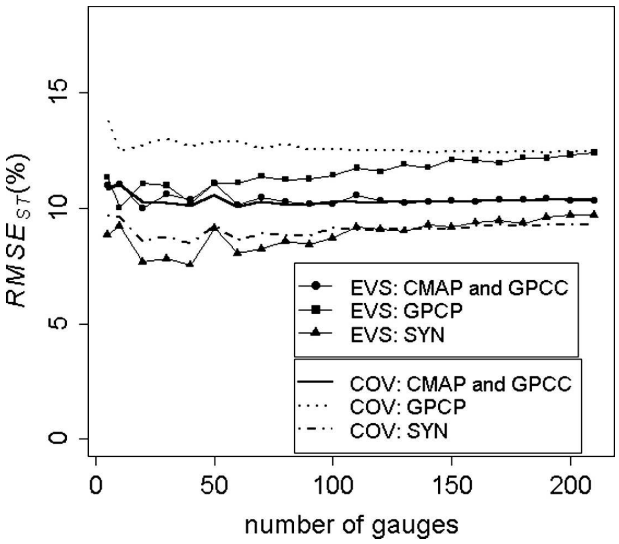

FIG. 11. Correction of the raw evaluation errors shown in Fig. 10. EVS means that the error covariance is neglected ($rmse_{ST}^{EVS}$ in the text). COV (solid line) takes into account the estimation of this covariance ($rmse_{ST}^{COV}$ in the text). Note that the EVS error displays a much larger sensitivity to the gauge density (oscillations in the curves) than does the COV error.

It is important to note that using $rmse_{ST}^{COV}$ does not change the ranking of products obtained from the simple computation of $rmse_{SR}$ but leads to a reevaluation of the estimation error, with $rmse_{ST}^{COV}$ being, in our case—and this cannot be extended to other regions without specific studies—significantly lower than $rmse_{SR}$.

7. Discussion and conclusions

This work presents the use of an error function developed in a companion paper (Part I) to treat three interwoven questions: 1) the definition of an optimal rain gauge network for the ground-based estimation of monthly and 10-day rainfall over the Sahel, optimality being defined as guaranteeing an estimation error below 10% for grid meshes of a given size ($2.5° \times 2.5°$ for monthly rainfall and $1° \times 1°$ for 10-day rainfall), 2) the intercomparison of global products, based on a reference considered as being the ground truth, and 3) the incorporation of the uncertainty of the reference into the quantification of the estimated error of the rainfall products.

Over the Sahel globally (an area of roughly 3 million km^2), the total number of gauges guaranteeing an error smaller than 10% is 568 on the monthly $2.5° \times 2.5°$ scale. This corresponds to a density of 14.5 gauges per $2.5° \times 2.5°$ cell—a number significantly greater than the

value of five gauges per 2.5° × 2.5° cell given by several authors for other regions of the world. This number of 568 gauges compares well globally to the 561 gauges available over the region for the period of 1990–2000. However, the density of the optimal network should increase when moving north because of the gradient of the average number of rain events. The distribution of the current network does not conform to this pattern. For instance, the "optimal" network should comprise 238 gauges south of 15°N—or 10.3 gauges per 2.5° × 2.5° cell—whereas the actual network comprises 465 gauges (19.2 per cell), meaning that the south Sahel is covered well. Over the whole region, however, only 42% of the 39 2.5° × 2.5° meshes satisfy the optimality criterion for monthly rainfall, with the northern and eastern areas being undersampled. North of 15°N, the average estimation error is 34%, which makes the validation of global products very difficult there (the density should be 20.6 gauges per cell in this area). On the 10-day 1° × 1° scale, 1736 gauges would be needed to ensure a good coverage of the area located south of 15°N, as compared with the 465 available. Thus, it is not surprising that the optimality criterion is satisfied for only 3.2% of the 135 cells in this area.

In a second step, four global rainfall products (GPCC, CMAP, GPCP, GPI) were compared, using the rain fields produced by the entire regional rain gauge network (CILSS) as a reference. All of these products (CMAP and GPCP blend satellite and ground information, GPCC is ground-based, and GPI is satellite-based only) highly underestimate the frequency of small rainfall linked to the high spatial intermittency, which is better documented by direct point measurements. Only 17% of the CMAP values, 13% of the GPCC values and GPCP values, and 6% of the GPI values are under the 25% CILSS quantile. Global products also slightly underestimate the frequency of the high values.

Two alternate regional ground-based products were also considered: the CRA monitoring network (about 250 gauges providing data in real time on a 10-day basis) and the synoptic network (about 80 gauges, available daily). Comparison of these two regional ground products with global products has been carried out for the CILSS area with latitudes less than 15°N. At the coarser resolution (1 month, 2.5° × 2.5°), the synoptic network performs very well, with its rmse (14.4%) with respect to the CILSS reference being smaller than the rmse of the four global products. CMAP, which performs slightly better than the others global products, has 15.9% rmse. This result is somewhat surprising because, in principle, the mixed global products incorporating ground information make use of the synoptic information available on the GTS and should therefore perform at least as well as the synoptic-only product. An explanation may reside in the fact that mixed products do not use all of the synoptic stations because not all of them transmit their data on the GTS.

The last part of the paper was devoted to assessing how using a reference value that is not the true value might influence the computation of evaluation errors. This question leads us to take into account two quantities that are usually neglected in the computation: the error of the reference and the covariance between the product-to-truth errors and the reference-to-truth errors. Using a denser instrumented zone, an empirical evaluation of these two quantities was performed, leading to an estimate that the "true" evaluation errors of both the regional and global products might be significantly lower than when these two terms are neglected. For instance, the reevaluated value of rmse for CMAP is around 11%, as compared with the above-mentioned value of 15.9% when the simplified computation is used. This difference is due both to the reference error and to the nonzero covariance existing between the product-to-truth errors and the reference-to-truth errors. Note also that this reevaluation of the rmse does not change the ranking of methods established from the simplified computation. The ground products remain, in this region, superior to the blended products analyzed here.

It is clear that ample room remains for improving the quality of satellite rainfall estimates over the Sahel. The availability of new infrared (e.g., Meteosat Second Generation) and microwave (e.g., TRMM and Global Precipitation Measurement) sensors will undoubtedly lead to more accurate satellite rainfall products. This study shows, however, that a careful and optimal use of ground information is needed to evaluate and intercompare any rainfall products, especially in regions of high rainfall variability.

Acknowledgments. This research was funded by IRD in the framework of the AMMA-CATCH "ORE" program initiated by the French Ministry of Research. Special thanks are given to the Departement Soutien et Formation of IRD for the grant allocated to the first author of this paper. We also acknowledge S. Bonaventure for his efforts in the exploitation of the AGRHYMET Center rainfall database and H. Harder for his help in proofreading the article.

REFERENCES

Adler, R. F., C. Kidd, G. Petty, M. Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products:

The Third Precipitation Intercomparison Project (PIP-3). *Bull. Amer. Meteor. Soc.,* **82,** 1377–1396.

Ali, A., T. Lebel, and A. Amani, 2005: Rainfall estimation in the Sahel. Part I: Error function. *J. Appl. Meteor.,* **44,** 1691–1706.

Arkin, P. A., R. Joyce, and J. E. Janowiak, 1994: IR techniques: GOES precipitation index. *Remote Sens. Rev.,* **11,** 107–124.

Barnston, A. G., 1991: An empirical method of estimating raingage and radar rainfall measurement bias and resolution. *J. Appl. Meteor.,* **30,** 282–296.

Barrett, E. C., and D. W. Martin, 1981: *The Use of Satellite Data in Rainfall Monitoring*. Academic Press, 340 pp.

Bell, T. L., P. K. Kundu, and C. D. Kummerow, 2001: Sampling errors of SSM/I and TRMM rainfall averages: Comparison with error estimates from surface data and a simple model. *J. Appl. Meteor.,* **40,** 938–954.

Chiu, L., A. Chang, and J. Janowiak, 1993: Comparison of monthly rain rates derived from GPI and SSM/I using probability distribution functions. *J. Appl. Meteor.,* **32,** 323–334.

Ciach, J. G., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.,* **22,** 585–595.

Conover, W. J., 1971: *Practical Nonparametric Statistics*. John Wiley and Sons, 234 pp.

Ebert, E. E., and M. J. Manton, 1998: Performance of satellite rainfall estimation algorithms during TOGA COARE. *J. Atmos. Sci.,* **55,** 1537–1557.

Gebremichael, M., W. F. Krajewski, M. Morrissey, D. Langerud, and G. J. Huffman, 2003: Error uncertainty analysis of GPCP monthly rainfall products: A data-based simulation study. *J. Appl. Meteor.,* **42,** 1837–1848.

Ha, E., G. R. North, C. Yoo, and K.-J. Ha, 2002: Evaluation of some ground truth designs for satellite estimates of rain rate. *J. Atmos. Oceanic Technol.,* **19,** 65–73.

Hollander, M., and D. A. Wolfe, 1973: *Nonparametric Statistical Inference*. John Wiley and Sons, 250 pp.

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

——, M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.,* **2,** 36–50.

Jobard, I., 2001: Status of satellite retrieval of rainfall at different scales using multi-source data. *Proc. Second MEGHA-TROPIQUES Scientific Workshop*, Paris, France, ISRO-CNES, 10 pp. [Available online at http://meghatropiques.ipsl.polytechnique.fr/documents/workshop2/proc_s4p03.pdf.]

Kummerow, C., P. Poyner, W. Berg, and J. Thomas-Sthale, 2004: The effects of rainfall inhomogeneity on climate variability of rainfall estimated from passive microwave sensors. *J. Atmos. Oceanic Technol.,* **21,** 624–638.

Laurent, H., I. Jobard, and A. Toma, 1998: Validation of satellite and ground-based estimates of precipitation over the Sahel. *Atmos. Res.,* **47–48,** 651–670.

Le Barbé, L., T. Lebel, and D. Tapsoba, 2002: Rainfall variability in West Africa during the years 1950–1990. *J. Climate,* **15,** 187–202.

Lebel, T., and A. Amani, 1999: Rainfall estimation in the Sahel: What is the ground truth? *J. Appl. Meteor.,* **38,** 555–568.

Levizzani, V., and Coauthors, 2001: EURAINSAT—Looking into the future of satellite rainfall estimations. *Proc. 2001 EUMETSAT Meteorological Satellite Data Users' Conf.*, Antalya, Turkey, EUMETSAT, 375–384.

Mathon, V., H. Laurent, and T. Lebel, 2002: Mesoscale convective system rainfall in the Sahel. *J. Appl. Meteor.,* **41,** 1081–1092.

McCollum, J. R., A. Gruber, and M. B. Ba, 2000: Discrepancy between gauges and satellite estimates of rainfall in equatorial Africa. *J. Appl. Meteor.,* **39,** 666–679.

Morrissey, M. L., and J. S. Greene, 1993: Comparison of two satellite based rainfall algorithms using atoll rain gauge data. *J. Appl. Meteor.,* **32,** 411–425.

——, J. A. Malikekal, J. S. Greene, and J. Wang, 1995: The uncertainty in spatial averages using raingauge networks. *Water Resour. Res.,* **31,** 2011–2017.

Nicholson, S. E., and Coauthors, 2003: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and pre-TRMM satellite and blended products. *J. Appl. Meteor.,* **42,** 1337–1354.

Petty, G. W., 1995: The status of satellite-based rainfall estimation over land. *Remote Sens. Environ.,* **51,** 125–137.

Ramage, K., I. Jobard, T. Lebel, and M. Desbois, 2000: Satellite estimation of 1-day to 10-day precipitation: Comparison and validation over tropical Africa of TRMM, Meteosat and GPCP products. *Proc. 2000 EUMETSAT Meteorological Satellite Data User's Conf.*, Bologna, Italy, EUMETSAT, 363–369.

Rudolf, B., 1993: Management and analysis of precipitation data on a routine basis. *Proc. Int. WMO/IAHS/ETH Symp. on Precipitation and Evaporation,* Bratislava, Slovakia, Slovak Hydrometeorology Institute, 69–76.

——, H. Hauschild, W. Rüth, and U. Schneider, 1994: Terrestrial precipitation analysis: Operational method and required density of point measurements. *Global Precipitations and Climate Change,* M. Dubois and F. Desalmand, Eds., Springer Verlag, 173–186.

Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verification methods in meteorology. Tech. Doc. WMO/TD-No. 358, 114 pp.

Thorne, V., P. Coakeley, D. Grimes, and G. Dugdale, 2001: Comparison of TAMSAT and CPC rainfall estimates with raingauges, for southern Africa. *Int. J. Remote Sens.,* **22,** 1951–1974.

Xie, P., and P. A. Arkin, 1995: An intercomparison of gauge observations and satellite estimates of monthly precipitation. *J. Appl. Meteor.,* **34,** 1143–1160.

——, 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.

——, J. E. Janowiak, P. A. Arkin, R. F. 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.