Estimates of Root-Mean-Square Random Error for Finite Samples of Estimated Precipitation

GEORGE J. HUFFMAN

NASA/GSFC Laboratory for Atmospheres and Science Systems and Applications, Inc., Greenbelt, Maryland

(Manuscript received 7 August 1996, in final form 11 March 1997)

ABSTRACT

The random errors contained in a finite set $E$ of precipitation estimates result from both finite sampling and measurement–algorithm effects. The expected root-mean-square random error associated with the estimated average precipitation in $E$ is shown to be

$$
\sigma_r = \sqrt{r(H - \bar{p}N)}
$$

where $H$ is a function of the shape of the probability distribution of precipitation (the nondimensional second moment), $p$ is the frequency of nonzero precipitation in $E$, and $N$ is the number of independent samples in $E$. All of these quantities are variables of the space–time-average dataset. In practice $H$ is nearly constant and close to the value 1.5 over most of the globe. An approximate form of $\sigma_r$ is derived that accommodates the limitations of typical monthly datasets, then it is applied to the microwave, infrared, and gauge precipitation monthly datasets from the Global Precipitation Climatology Project. As an aid to visualizing differences in $\sigma_r$ for various datasets, a “quality index” is introduced. Calibration in a few locations with dense gauge networks reveals that the approximate form is a reasonable first step in estimating $\sigma_r$.

1. Introduction

Detailed estimation of the error associated with precipitation estimates is a difficult problem requiring statistical information not generally provided in current global precipitation datasets (North et al. 1991). Nonetheless, some estimate of error is required for both the creation and use of precipitation estimates. From the producers’ perspective, estimates of error for each independent dataset, whether from satellites or rain gauges, provide key information for combining the datasets. This was certainly true for development of the Satellite-Gauge-Model (SGM) combination technique (Huffman et al. 1995) and production of the Global Precipitation Climatology Project (GPCP) Version 1 Combined Precipitation Data Set (Huffman et al. 1997). From the users’ perspective, error estimates allow inferences about the reliability (or even advisability) of comparisons between datasets. Furthermore, both producers and users are best served when the error estimates are spatially and temporally varying, rather than being quoted as single dataset-average values.

This study begins by adopting the usual partitioning of each instantaneous estimate of precipitation into a “true” value, a (constant) bias error, and a random error

$$
r = r + b + e,
$$

where bias is the difference between the estimated values and the “true” values averaged over a very large dataset, and random error is the rest of the error for each estimate. Note that random errors can be correlated in space or time, potentially affecting the estimated correlation structure of the precipitation. Errors arise from two sources. First, the various estimates depend on measurements and algorithms, each of which introduces its own set of bias and random errors. Second, the estimates contain “sampling error” in which a finite set of observations only samples the complete set of events. Classically, the random part of sampling error depends only on the pattern of sampling, while biases arise from systematic problems such as siting rain gauges on land alone. However, this study has a finite set of estimates, so there is also sampling of the measurement and algorithm errors that happen to be present in the estimated values. As a result, the error studied in this paper is a mixture of sampling and measurement–algorithm effects. It is an important exercise to untangle the contributions from each kind of error, but one that is beyond the scope of the current study. It is also left for future study to examine how errors might appear as biases on one set of space–timescales, but as random errors on another. Thus, the current study represents only one stage in characterizing errors.

The random term alone is considered in this paper for several reasons. First, the random error is frequently the dominant term for the sparsely sampled datasets that Huffman et al. (1995, 1997) and many other papers treat. Second, precipitation estimation techniques usually
strive to generate unbiased estimates (“small” \(b\)). Third, even when \(b\) is not small, its analysis is relatively well understood, for example, by computing long-term or large-area averages in which the random errors can be assumed to average out. Note, however, that the lack of calibration data frequently prevents such an analysis for \(b\).

In the following sections the random error is characterized by deriving a functional form for rms fluctuations in instantaneous estimates, and then for rms error in space–time-average estimates. A simplification is introduced to accommodate limitations in the datasets at hand. Finally, the practical application of this work is illustrated with the various precipitation datasets used in the GPCP Version 1 Dataset. The illustration is aided by defining a “quality index” that simplifies the comparison of the rms errors between and within datasets.

2. A functional representation of rms error

The general expression for rms error in time-average estimates is a complicated filtered integral over the space–time domain, as shown by North and Nakamoto (1989). Using the Fourier representation, it is the product of a term depending on the observational locations (in space–time) and a term depending on the structure of the true rain field. Fortunately, North and Nakamoto (1989) also show that for the cases of typical satellite or rain gauge datasets the general expression collapses to simple expressions that depend on the variance of the instantaneous “true” precipitation and the number of independent samples in the observed data. Knowing this result, one can begin by writing the variance of the instantaneous precipitation rates in a set \(E\) of precipitation estimates without explicit reference to the detailed space–time pattern of samples:

\[
\sigma_i^2 = \int_0^\infty f(r)(r - \bar{r})^2 \, dr, \tag{2}
\]

where \(f(r)\) is the probability distribution of precipitation rate \(r\) that occurred in set \(E\) (and which sums to 1 over the entire range of \(r\)), and \(\bar{r}\) is the precipitation rate averaged over set \(E\):

\[
\bar{r} = \int_0^\infty f(r)r \, dr. \tag{3}
\]

The frequency of all nonzero precipitation \(p\) is

\[
p = \int_0^\infty f(r)r \, dr, \tag{4}
\]

where the lower limit does not include zero. There is a delta function in \(f(r)\) of size \((1 - p)\) at zero precipitation rate.

For completeness, (2) is rewritten by splitting off the delta function at \(r = 0\), expanding the squared term, and recognizing that \(\bar{r}\) is not dependent on \(r\):

\[
\sigma_i^2 = (1 - p)\bar{r}^2 + \int_0^\infty f(r)r^2 \, dr - 2\bar{r}^2 \int_0^\infty f(r) \, dr. \tag{5}
\]

The lower limit on the integral in the third term of (5) may be set to zero because \(f(r)r\) is zero at \(r = 0\). Then (3) and (4) can be applied to simplify the last two terms in (5).

The second term in (5) can be simplified as follows. The lower limit of the integral can be set to zero because \(f(r)r^2\) is zero at \(r = 0\). This step reveals that the integral is the second moment of the rain-rate histogram. Define new nondimensional precipitation rate \(r'\) and frequency \(f(r')\) as

\[
r' = \frac{\bar{r}}{p},
\]

and

\[
f(r) = \frac{p^2}{\bar{r}^2} f'(r'). \tag{6}
\]

respectively. Substituting (6) as needed, the second term of (5) becomes

\[
\frac{p^2}{\bar{r}^2} \int_0^\infty f'(r')r'^2 \, dr'. \tag{7}
\]

To simplify the following discussion the integral is labeled as \(H\), the nondimensional second moment of the probability distribution

\[
H = \int_0^\infty f'(r')r'^2 \, dr'. \tag{8}
\]

This is useful because \(H\) is a function of the shape of the probability distribution alone, with stretching due to changes in \(\bar{r}\) or \(p\) isolated in the \(\bar{r}^2/p\) term. Collecting the simplifications to (5),

\[
\sigma_i^2 = (1 - p)\bar{r}^2 + \frac{p^2}{\bar{r}^2} H - 2\bar{r}^2 + p\bar{r}^2,
\]

or

\[
\sigma_i^2 = \bar{r}^2 \left(\frac{H}{p} - 1\right). \tag{9}
\]

How does \(H\) vary? One might expect \(H\) to be well behaved because probability distributions of instantaneous precipitation rate (accumulated over a period such as a month) tend to look similar (i.e., lognormal; Kedem et al. 1990). In fact, the result is much stronger. Analytic examples (shown in Fig. 1 and summarized in Table 1) demonstrate quite stable behavior over a wide range of probability distributions. Turning to a more realistic case, the Goddard Scattering Algorithm, Version 2 (GSCAT2, Adler et al. 1994) was used to generate the necessary monthly variables from Special Sensor Mi-
FIG. 1. Schematics of analytic examples of the probability distribution of precipitation: (a) spike, (b) boxcar, and (c) triangle. All distributions have a delta function of value \((1 - p)\) at \(r = 0\). Both dimensional and nondimensional coordinates are shown.

FIG. 2. Scattergram of \(H\) as a function of \(r\) using GSCAT2 precipitation estimates for August 1987.

TABLE 1. Summary of analytic probability distribution results. The constant \(R\) is a location on the \(r\) axis (Fig. 1).

<table>
<thead>
<tr>
<th>Case</th>
<th>Spike</th>
<th>Boxcar</th>
<th>Triangle</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f(r))</td>
<td>(1 - p; r = 0)</td>
<td>(1 - p; r = 0)</td>
<td>(1 - p; r = 0)</td>
</tr>
<tr>
<td>(p; r = R)</td>
<td>(p/R; r \in (0, R])</td>
<td>(2p/R(1 - r/R); r \in (0, R])</td>
<td></td>
</tr>
<tr>
<td>(0; r \neq 0, R)</td>
<td>(0; r \not\in [0, R])</td>
<td>(0; r \not\in [0, R])</td>
<td></td>
</tr>
<tr>
<td>(H)</td>
<td>1</td>
<td>4/3</td>
<td>3/2</td>
</tr>
</tbody>
</table>
Fig. 3. Estimates of $H$ using GSCAT2 precipitation estimates for August 1987. Regions with zero and no precipitation estimates for the month are denoted by gray and black, respectively.

Fig. 4. Scattergram of $r_c$ as a function of $r_A$ for the GSCAT2 precipitation estimates for August 1987.

\[
\sigma_r^2 = \left( \frac{1}{N_I} \sum_{i=1}^{N_I} \left( r_i - \bar{r} \right)^2 \right),
\]

or

\[
\sigma_r^2 = \frac{\sigma_r^2}{N_I},
\]

where the $I$ subscript emphasizes that $N_I$ is the number of independent samples entering the average, and it is assumed that $N_I$ is much smaller than the number of independent samples in the complete set of events. In general, determining the true number of independent observations is complicated but, as discussed below, it is not necessary to be too precise when empirical fitting is used to approximate constants. Substituting from (9),

\[
\sigma_r^2 = \frac{\bar{r}^2}{N_I} \left( \frac{H}{\bar{r}} - 1 \right).
\]

The expression in (11) is consistent with Eq. 30 in Bell et al. (1990) if the 1 in the parentheses is neglected (as it will be below) and $H$ absorbs the proportionality constant. The current result clarifies the basis for considering $H$ to be nearly constant as Bell et al. (1990) asserted.

3. An approximation to rms error

The exact representation of $\sigma_r$ in (11) presents considerable practical problems when applied to the global monthly precipitation datasets ordinarily available. In particular, this is true in generating error estimates for use in the SGM technique. First, one usually lacks the $\sigma_r$ to calculate $H$ directly. The discussion in section 2 implies that $H$ can be approximated as a global constant with the value 1.5. As long as the GSCAT2 results are somewhat representative, this choice yields rms error...
inaccuracies of about +12% to −6% at low values of \( \bar{r} \) (for \( H \) ranging over 1.2–1.7), and considerably better accuracies at higher values of \( \bar{r} \).

The second problem is more severe. Estimates of the frequency of precipitation often are either questionable (e.g., GSCAT2 tends to produce \( p \) values smaller than climatological values) or entirely missing (most rain gauge analyses). Noting that conditional rain rate \( r_c \) is considered well-behaved or even constant in various studies, in the current study \( p \) is replaced with \( \bar{r} \) and \( r_c \) is parameterized as a function of \( \bar{r} \). This approach was tested using the GSCAT2 dataset for August 1987 (used above). As shown in Fig. 4 there is a tremendous amount of scatter, but \( r_c \) does increase with \( \bar{r} \). Nonlinear fitting of a power law with a nonzero intercept for all GSCAT2 data in 1988 yielded

\[
\bar{r}_c = 72 + 12\sqrt{\bar{r}},
\]

where \( \bar{r} \) is in millimeters per day. Early testing of this formulation against the Surface Reference Data Center (SRDC, described below) calibration data seemed to show that \( r_c \) needed a steeper slope to fit the gauge data, which are presumed to be more realistic than the GSCAT2 data. A range of values for slope and intercept was tested in the calibration (while retaining the square-root power law) and the most consistent results across the range of \( \bar{r} \) seemed to be given by

\[
\bar{r}_c = 24 + 49\sqrt{\bar{r}}.
\]

This result is best viewed as a stopgap until a more detailed study or reliable \( p \) information becomes available.

A third practical issue is that \( \bar{r} \) is most easily estimated by the average over the finite set \( E \) of precipitation estimates. However, (11) produces zero \( \sigma \) any time \( \bar{r} \) is zero, even though \( \bar{r} \) has some uncertainty. This effect is accounted for by adding the constant \( S \) to the leading \( \bar{r} \), where \( S \) is a measure of the standard deviation in calibration when \( \bar{r} = 0 \).

The fourth problem is in estimating \( N \) on a global basis, which requires information on the (spatially and temporally varying) correlation time and distances of the algorithm-estimated precipitation field. These values are affected both by the correlations in the true precipitation and by correlations in the algorithmic errors. For simplicity it is assumed that \( N \) is related to the reported number of samples, \( N \), by the constant multiplier \( I \) as

\[
N_I = IN.
\]

Thereby it is assumed both that the space–time distribution of samples within a dataset always yields about the same degree of oversampling (if any), and that the correlation distances and times within that dataset are relatively constant. Both of these conditions are severe enough that the author believes it mandatory to develop a better model of \( N_I \) in future research.

A final change to (11) for computational simplicity is to neglect the 1 since \( H \gg p \). Substituting all of these approximations into (11), the error estimate becomes

\[
\sigma^2_I \equiv \frac{H(\bar{r} + S)}{I} \left[ \frac{24 + 49\sqrt{\bar{r}}}{N} \right].
\]

where \( \bar{r} \) and \( N \) can be estimated for each grid box with available data, while \( H/I \) and \( S \) are assumed to be global constants. The constant \( I \) certainly has the potential for taking different values for different fields (microwave emission technique, gauge analysis, etc.), depending on the units of \( N \) and the typical sampling and space–time correlation patterns of each. However, the author believes that \( S \) and \( H \) also must be allowed to take different values for different data sources because sampling introduces algorithm dependence, as discussed in the introduction. To be explicit, not only does a weak algorithm yield worse instantaneous results, but average values computed from the weak algorithm should have larger (random) errors and therefore larger \( S \), compared to accurate algorithms. There should not be much change in \( H \), but an inaccurate algorithm might systematically change the shape of the probability distribution of precipitation. Consequently, one must account for \( I \), \( H \), and \( S \) taking different values for each data source. Two practical benefits of using (13) to set constants from data are that \( H \) and \( I \) do not need to be estimated separately, and multiplicative errors in the assumed form of \( r_c \) (the term in square brackets) will automatically be absorbed into the fitted \( H/I \). It will be important to compute \( H \) and \( I \) separately in future studies of datasets that have appropriate information for doing so.

4. Global constants

The individual GPCP datasets for which rms errors are computed in this paper are the SSM/I scattering estimates over land, SSM/I emission estimates over water, microwave-adjusted infrared estimates [the adjusted GOES precipitation index (AGPI)], and the Global Precipitation Climatology Centre (GPCC) gauge analysis (see Huffman et al. 1997 for descriptions). As shown in Table 2 under “units of \( N \)” the datasets have rather different measures of sampling. For the purposes of computing the GPCP Version 1 Dataset several of the units were modified (shown as “modified units of \( N \)” in Table 2). The number-of-samples variable for each SSM/I technique was converted to an approximate number of 55 km × 55 km boxes that would be fully covered by the original number of samples that was accumulated in the 2.5° × 2.5° grid box over the month. The number-of-samples variable for the AGPI was the count of infrared images that covered the 2.5° × 2.5° grid box during the month. The number-of-samples variable for the GPCC rain gauge analysis started as the number of active rain gauges in the 2.5° × 2.5° grid box during the month. Then the number-of-samples variable was modified because the SPHEREMAP technique, which GPCP uses for analysis, considers data outside the box. Following Huffman et al. (1995) this extra information was roughly modeled by adding a fraction (0.125) of
the number of surrounding boxes containing data to the number of gauges in the box.

The scarcity of dense rain gauge networks around the globe prevents one from simply computing the global H/I and S for each data source. Rather, calibration must be done with a few datasets that are believed representative. The primary source for calibration in this study is the rain gauge analysis produced by the GPCP’s SRDC, located at the National Climatic Data Center (NCDC). The SRDC uses the PRISM technique (Daly et al. 1994) to analyze 2.5° × 2.5° grid cells that have reasonably dense rain gauge coverage (around 50 per 2.5° box). The SRDC data are drawn from the years 1987–91 for seven “test sites” (most with more than one 2.5° × 2.5° grid box) around the globe. The Morrissey et al. (1995) collection of tropical Pacific atoll gauge data is also used as a source of calibration. The atolls provide sparse coverage (typically 1–4 stations per 2.5° × 2.5° grid box) throughout the years 1987–95. These data are believed to be approximately representative of open ocean conditions in the tropical Pacific. The atoll reports are analyzed onto the GPCP’s 2.5° × 2.5° grid using simple grid-box averages for this study. Finally, the SSM/I and AGPI estimates are compared to the GPCC gauge analysis as a consistency check.

The distribution of stations contributing to the GPCC rain gauge analysis in a 2.5° × 2.5° grid box varies wildly from place to place, so the author chose to compute the constants for the GPCC gauge analysis using all available months for all of the SRDC test sites that had good coverage at the time of this study, namely, the southeastern United States (SEUS); Canada, Honduras, Thailand, and Australia. On the other hand, satellite data coverage tends to be fairly uniform in tropical and subtropical regions, so the satellite dataset constants were computed for SRDC test sites and months that were judged to best represent the dominant climatic regime. The summertime SEUS was considered to have the best chance of representing the convective regime typical of the moist Tropics, both land and ocean, since its precipitation is dominated by convection, but there is no strong orography. The SSM/I emission estimates are not available at the SRDC test sites, so the Pacific atoll data were used for setting the constants.

The SRDC station data are sufficiently dense that it can be assumed that errors in the grid-box averages are much smaller than errors for the datasets under study. In contrast, the low gauge populations in the atoll data require that random errors be assumed in both the calibration data and the SSM/I emission estimates. Furthermore, there is a difference (bias) between the average values of the SSM/I emission estimates and atoll analyses. Recalling (1) and assuming that the random errors in the two fields are uncorrelated, the expected value of the squared difference between two precipitation estimates is

$$\langle (r_1 - r_2)^2 \rangle = (b_1 - b_2)^2 + \sigma_1^2 + \sigma_2^2,$$  \hspace{1cm} (14)

where \(\sigma\) is the rms error. Adding and subtracting the expected value of \(r_i\) inside the first term on the right and solving for \(\sigma_i^2\)

$$\sigma_i^2 = \langle (r_1 - r_2)^2 \rangle - \langle (r_1 - r_2) \rangle^2 - \sigma_1^2.$$  \hspace{1cm} (15)

Thus, if the mean-square error of field 1 is computed using (13) with previously set H/I and S, all the other terms on the right side of (15) can be computed from the data, and the mean-square error of field 2 can be found. Working backward in (13), H/I and S can be inferred for field 2. The flaw in this scheme is that all of the errors of approximation manifest themselves in \(\sigma_i^2\). In practice, (15) is applied for different ranges of precipitation amount, and the identities of fields 1 and 2 are interchanged to assess how well the inferred rms error functions are performing.

Table 2 contains the values of H/I and S that are computed for the different datasets using the calibration datasets discussed above. In each case a variety of values was specified for H/I and S, and the values that minimized inaccuracies over the whole range of \(r\) were chosen. The final results represent a subjective trade-off because the fits for gauge and SSM/I scattering had to be biased in order to get reasonable checks against the GPCC and atoll gauge analyses. The corresponding performance of the error estimates in the dependent datasets is shown in Fig. 5. Each of the techniques’ points were produced by binning on rain totals with breakpoints at 0.33, 1.67, 3.33, 5.00, 6.67, and 10.00 mm day\(^{-1}\). In each case the observed rms errors have been adjusted for bias, following (15). Over the whole range of rain rates the rms error estimate replicates the general behavior of the observed rms error, although differences indicate that the functional form is only a first approximation.

An example of the rms error estimates is shown in Fig. 6. Curves of constant sample size are plotted to display relative error as a function of precipitation rate. The sample sizes displayed for AGPI and SSM/I emission (Fig. 6b) roughly bound the normal range of population sizes. Overall, the curves agree with previous work. For comparison, the Bell et al. (1996) estimates of relative error for GATE and for SSM/I in the western Tropical Pacific during November 1992–February 1993 are scaled to 2.5° × 2.5° boxes and plotted on Fig. 6b. The GATE results are lower than either SSM/I estimate, perhaps because there is no algorithmic error in the calculation from GATE. The fitted Bell et al. (1996) SSM/I curve is systematically above the SSM/I curves in this study, although the separation is less than their error bars. It is not clear why this difference occurs, although it might be related to regional differences in precipitation.

5. Quality index

A final error-related issue uncovered in Huffman et al. (1997) is the visual representation of error. By sub-
Table 2. Datasets, original and modified dimensional units of \(N\) (see text), and computed values of the constants \(H/I\) and \(S\) used in the GPCP Version 1 Dataset.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Units of (N)</th>
<th>Units of modified (N)</th>
<th>(H/I) (same units as modified (N))</th>
<th>(S) (mm day(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSM/I emission</td>
<td>Low-resolution SSMI pixels</td>
<td>Filled 55 km (\times) 55 km boxes</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>SSM/I scattering</td>
<td>Days with data in an ascending orbit plus days with data in a descending orbit</td>
<td>Filled 55 km (\times) 55 km boxes</td>
<td>3.2</td>
<td>1.0</td>
</tr>
<tr>
<td>AGPI</td>
<td>(2.5° \times 2.5°) images</td>
<td>(2.5° \times 2.5°) images</td>
<td>0.45</td>
<td>0.5</td>
</tr>
<tr>
<td>Rain gauge analysis</td>
<td>Stations in a (2.5° \times 2.5°) box</td>
<td>(N + 0.125\times(\text{number of surrounding boxes with } N &gt; 0))</td>
<td>0.0075</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Fig. 5. Comparison of observed and estimated rms error \(\sigma\) for the calibration datasets used to set constants for the four precipitation estimation techniques: (a) SSM/I scattering technique, with observations in solid and estimates in dotted; and SSM/I emission technique (combined with errors in the atoll gauge analysis), with observations in dash–dotted and estimates in dashed; and (b) gauge analysis, with observations in solid and estimates in dotted; and AGPI technique, with observations in dash–dotted, and estimates in dashed. The observed rms errors are corrected for bias as in (15).

\[
Q = \left(\frac{H}{I}\right) \left(\frac{\bar{r} + S}{\sigma_r^2}\right) \left[24 + 49\sqrt{\bar{r}}\right].
\] (16)

That is, when \(H/I\) and \(S\) are set to the gauge analysis values (subscripted \(g\)), \(Q\) is an estimate of the number of gauges required to produce the \(\sigma\) value that a technique produces in a particular grid box (at the \(\bar{r}\) inferred by the algorithm). Here, \(Q\) has the units of “equivalent gauges.” Qualitatively, low values of \(Q\) (1 or 2) indicate a relatively poor estimate, while higher values (greater than 10) indicate a confident estimate. Even when representations of error become more sophisticated it is...
Fig. 6. Relative rms error $\sigma_f/\bar{f}$ as a function of $\bar{f}$ for three precipitation estimation techniques with various sample sizes: (a) GPCC rain gauge analysis for 0.5, 2, 8, and 32 gauges in a $2.5^\circ \times 2.5^\circ$ grid box, depicted with solid, dotted, dash-dotted, and dashed lines, respectively; and (b) SSM/I emission estimates for 550 and 700 55 km $\times$ 55 km equivalent images in a month, depicted with solid and dotted lines, respectively; and AGPI estimates for 200 and 240 $2.5^\circ \times 2.5^\circ$ images in a month, depicted with dash-dotted and dashed lines, respectively. The GATE and SSM/I estimates from Bell et al. (1996) are plotted in (b) for comparison as a single star and a long-dashed line, respectively. The SSM/I scattering estimates are so close to the SSM/I emission estimates that they are not plotted.

6. An example

The technique developed above was applied in computing the GPCP Version 1 Dataset (Huffman et al. 1997). The combination of SSM/I, infrared, and rain gauge data (the SG estimate) for August 1987 is shown in Fig. 8 together with quality index, rms error, and relative rms error. These images display many of the qualities discussed above. The rms error is larger in regions of higher precipitation, while the relative rms error is smaller. The quality index clearly displays boundaries in data coverage, such as the switch between AGPI and SSM/I at 40°N and 40°S, the gap in geosynchronous infrared in the Indian Ocean, and the addition of gauges (in varying amounts) in land areas. Regions having dense gauge coverage, such as Europe and China, have smaller values of both rms and relative rms error, and higher values of quality index. Notice how large portions of the tropical land areas have quality index values close to the satellite values alone, due to sparse gauge coverage. Although all three error representations are based on the computation of rms error, the different representations provide different insights into the character of the errors.

What do these rms error estimates tell us about the utility of the precipitation field (remembering that this is only the random part of the error)? The original TRMM goal was relative error less than 10% for monthly $5^\circ \times 5^\circ$ estimates (Simpson et al. 1988). Since then, work such as this study has motivated discussions about adopting an absolute floor to prevent the 10% threshold from requiring tiny error limits in light-rain areas that are not operationally meaningful. For the current study (which considers monthly $2.5^\circ \times 2.5^\circ$ data) the TRMM goal becomes 20% and a reasonable, small floor of 0.4 mm day$^{-1}$ (12.4 mm for a 31-day month such as August) is chosen for illustration. In Fig. 8 the subtropical highs and intertropical convergence zone meet these goals, but the large regions with intermediate precipitation rates require additional data. Over land, conscientious assembly of existing, but fragmented rain gauge datasets (as the GPCC is doing) will have beneficial effects in many countries. Over the remainder of the globe, including some remote land areas, the introduction of additional microwave-sensing satellites would easily re-
produce the random error problem. Throughout, improved analysis techniques would also reduce the errors.

7. Concluding remarks

A simple exact expression has been developed for estimating the rms error in space–time-average precipitation estimates using quantities that are variables of the averaged dataset. This approach accounts for rms errors in the sampling of both the true precipitation field and of the measurement–algorithm-induced errors. These errors may have a correlation structure that is not accessible in monthly averages, but can complicate the analysis and should be addressed in the future. The non-dimensional second moment of the precipitation distribution was shown to be a key variable, and analytic and data work showed that it is relatively constant over a wide range of precipitation rates. Several approximations were introduced to allow computation of rms error with the available data. The constants in the functional form were separately set for each estimation technique based on the datasets believed to be representative of rms error for that technique. The resulting rms error fields perform tolerably well in comparison to observed rms errors. The quality index was introduced to make it easier to interpret the error pattern. Inspection of the various error representations makes clear the need for more data in the monthly estimates.

A great deal of work remains to be done to adequately characterize error. The greatest need is to set the constants on a regional basis, instead of having a single global value. A somewhat related issue is that the model for the number of independent samples should be redefined. As well, it would clearly be advantageous to have calibration–validation datasets representative of all the various climate regimes around the globe. Finally, error estimation would be facilitated if groups producing precipitation estimates computed estimates of higher-order statistics. In particular, estimates of the variance of instantaneous precipitation values and the fractional coverage by precipitation would allow the use of the exact expression in (11) rather than the parameterized expression in (13).

Acknowledgments. This research was conducted as part of the prelaunch algorithm development work for the Tropical Rainfall Measurement Mission funded by NASA under Dr. R. Kakar. Scientific discussions with Drs. R. F. Adler and T. L. Bell helped the author shape his understanding of the error estimation problem. Comments by the anonymous reviewers and Dr. I. Polyak improved the final paper.
REFERENCES


Fig. 8. GPCP combined satellite–gauge (SG) precipitation estimate for August 1987 in millimeters per day (top) with three representations of the estimated error field: (upper middle) quality index $Q$ in equivalent gauges, (lower middle) rms error $\sigma$, in millimeters per day, and (bottom) relative rms error $\sigma/r$ in percent. Regions with no estimate are denoted by black.