Probabilistic and Ensemble Representations of the Uncertainty in an IR/Microwave Satellite Precipitation Product

TIMOTHY J. BELLERBY

Department of Geography, The University of Hull, Hull, United Kingdom

JIZHONG SUN

Department of Physics, The University of Hull, Hull, United Kingdom

(Manuscript received 15 November 2004, in final form 12 May 2005)

ABSTRACT

While current satellite techniques are theoretically capable of producing precipitation estimates to image pixel resolutions, significant uncertainty is present in such high-resolution products. This uncertainty is frequently difficult to characterize using scalar measures of additive error. This paper describes the development of a methodology to more fully represent the uncertainty in satellite precipitation retrievals. The methodology derives conditional probability distribution functions of rainfall on a pixel-by-pixel basis. This array of distribution functions is then combined with a simple model of the spatiotemporal covariance structure of the uncertainty in the precipitation field to stochastically generate an ensemble precipitation product. Each element of the ensemble represents an equiprobable realization of the precipitation field that is consistent with the original satellite data while containing a random element commensurate with the uncertainty in that field. The technique has been tested using data from the Tropical Rainfall Measuring Mission (TRMM) Texas and Florida Underflight Experiment (TEFLUN-B) field campaign.

1. Introduction

A wide range of applications require precipitation data at small spatial and temporal scales. For the many regions of the globe, dense rain gauge or surface radar networks are not available and satellite instruments provide the only available source of such data. Current multiplatform satellite precipitation algorithms (Adler et al. 2000; Bellerby et al. 2000; Bellerby 2004; Huffman et al. 2003; Joyce et al. 2004; Kidd et al. 2003; Kwon et al. 2000; Sorooshian et al. 2000; Todd et al. 2001; Marzano et al. 2004; Nicholson et al. 2003a, 2003b; Xu et al. 1999) are potentially capable of producing estimates up to the spatial and temporal resolution of geostationary weather satellite imagery (typically ~5 km and 15 min, respectively). However, when operating at the maximum possible spatial and temporal resolution, these techniques remain primarily dependent on the physically indirect relationship between geostationary cloud indices and rainfall rates and consequently experience a high degree of uncertainty.

For a wide range of climate monitoring and hydrological modeling applications, the above limitations may be overcome by aggregating the satellite estimates to lower spatiotemporal scales. Such aggregation may be based on a regular grid or on the local catchment geometry, the latter scheme potentially taking advantage of the natural rainfall-runoff scaling properties of many river basins (Wood et al. 1990). However, this process inevitably discards a significant portion of the spatial information present in the original satellite imagery. In addition, many hydrological modeling applications require precipitation to be estimated at or close to the full spatial and temporal resolution of geostationary image data. In particular, it may not always be feasible to scale a hydrological model to operate at very low spatial resolutions. In such cases, it may be preferable to drive the hydrological model using uncertain high-resolution precipitation products than to employ more certain low-resolution products. If high-resolution satellite products are to be used to drive opera-

Corresponding author address: Timothy J. Bellerby, Department of Geography, The University of Hull, Hull HU6 7RX, United Kingdom.
E-mail: T.J.Bellerby@hull.ac.uk

tional hydrological models then it is essential that the uncertainty present in the measured precipitation field is quantified as fully as possible. This must be achieved in such a manner as to allow uncertainties in precipitation measurements to be straightforwardly translated into corresponding uncertainties in runoff and other hydrological model outputs. The quantitative propagation of error from satellite estimate to hydrological model has the potential to decouple the choice of spatial resolution from concerns regarding retrieval uncertainty.

Hydrological modeling is not the only potential benefactor of a more complete representation of satellite precipitation retrieval uncertainties. Meteorological model validation and data assimilation currently tend to eschew the use of satellite precipitation retrieval products in favor of the forward modeling of at-sensor brightness temperatures. The primary reason for this approach is the difficulty in characterizing satellite retrieval uncertainties (Chevallier and Bauer 2003). The development of improved quantitative representations of retrieval uncertainties is thus a prerequisite to the widespread adoption of satellite precipitation products for this type of application.

The error in a satellite precipitation product is typically quantified using the variance of the point-by-point difference between a "best estimate" field and ground truth (e.g., Gebremichael et al. 2003). While this represents an effective approach for lower-resolution products, at near-pixel resolutions the most significant errors in individual retrievals tend to be associated with mistakes in categorizing and precisely locating precipitation regimes. For example, in many IR techniques, a nonprecipitating cirrus cloud with a cold IR signature may frequently be confused for a strongly precipitating cumulonimbus. Other IR techniques attempt to separate convective and stratiform precipitation (e.g., Adler and Negri 1988), again with nontrivial rates of misclassification and the potential for mislocating convective cores due to parallax and shear effects (Vicente et al. 2002). This type of error is difficult to characterize effectively using a point-by-point scalar measure, especially when error propagation through the final application is important. A modest error in the location of a small area of intense convective rainfall within the same catchment may be associated with a high point-precipitation error but a relatively modest change in the magnitude and timing of local runoff.

A more general representation of the uncertainty in a satellite precipitation product is provided by the conditional probability distribution of precipitation with respect to a given satellite measurement. If this distribution possesses a single modal value that is reasonably close to its expected value then it is reasonable to talk in terms of a best estimate of precipitation from the satellite data. However, if the conditional distribution function takes a more complex form or has a modal value that is essentially unrelated to its expected value then the provision of a best estimate and an associated scalar error value becomes significantly less useful to the end user. In these cases it may be preferable to attempt a more direct representation of the conditional distribution function itself.

This paper describes a methodology that produces a field of conditional cumulative probability density functions of precipitation from satellite data. Such an approach acknowledges that while geostationary satellite data may have difficulty in determining actual precipitation rates to a pixel resolution, they may be effective at determining the local statistical properties of the precipitation field. Some applications, such as meteorological model validation and data assimilation, will be able to make direct use of such a field. However, other applications, such as deterministic rainfall-runoff models, need to be driven by concrete precipitation data. To accommodate such applications, the conditional distribution field is combined with a simple model of the spatiotemporal covariance structure of the uncertainty in the precipitation field to generate a stochastic ensemble product. Each component of the ensemble is a precipitation field consistent with the original satellite data while containing a random element constrained by the conditional distribution field and the covariance model. Feeding each member of the ensemble, in turn, into a deterministic rainfall-runoff model will map out the distribution of possible hydrological responses that may be associated with the original sequence of satellite imagery.

2. Methodology

a. Background

Geostationary satellite precipitation algorithms make use of a variety of cloud indices to estimate surface rainfall rates. These include cloud areas and durations, IR brightness temperatures, the locations of minima in the brightness temperature field, cloud shapes and textures, and estimated rates of cloud growth (Arkin and Meisner 1987; Adler and Negri 1988; Bellerby 2004; Griffith et al. 1978; Todd et al. 2001; Wu et al. 1985). To improve their performance, these techniques are frequently calibrated using data from passive and active microwave instruments carried by satellites in low earth orbits (LEO). Instruments em-

ployed for this purpose include the Special Sensor Microwave Imager (SSM/I) on the U.S. Defense Meteorological Satellite Program (DMSP) series of spacecraft (Wilheit et al. 1994) and the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) and Precipitation Radar (PR) on the National Aeronautics and Space Administartion (NASA)–National Space Development Agency of Japan (NASDA) TRMM satellite (Kummerow et al. 1998, 2000). Passive and active microwave data display significantly more physically direct relationships to precipitation then do geostationary cloud indices. However, the spacecraft carrying such instruments typically pass over a given geographical location only once or twice per day (somewhat less than this in the case of TRMM) rendering their data unsuitable for monitoring accumulated rainfall volumes (Morrissey and Janowiak 1996). Merging microwave data from multiple LEO platforms improves this situation, and the planned Global Precipitation Measurement (GPM) program aims to develop a constellation of satellites with a 3-h return time for most points on the earth's surface (Smith 2003). Combined LEO/geostationary satellite precipitation monitoring techniques aim to provide the highest possible spatial and temporal resolutions by dynamically calibrating geostationary algorithms against precipitation estimates from available microwave instruments.

In most combined LEO/geostationary techniques, the region being monitored is divided into a number of subregions and the satellite data further divided into equal time steps. A geostationary precipitation algorithm is independently calibrated within each spatially and temporally bounded *calibration domain* through a comparison of microwave precipitation estimates with coincident geostationary imagery. These calibrations are used to produce precipitation estimates from geostationary imagery over the full spatial and temporal extent of the domain for which they were calculated. Domain sizes vary significantly between different algorithms (Kidd et al. 2003; Todd et al. 2001; Xu et al. 2000) and the calibration process may take many forms including the training of neural networks (Bellerby et al. 2000; Sorooshian et al. 2000; Tapiador et al. 2004) or the initialization and termination of a sequential advection-related precipitation-field morphing process (Joyce et al. 2004). Techniques intended for operational use may additionally incorporate synoptic station measurements and other relevant data (Adler et al. 2000; Huffman et al. 2003). The final combined product may either be based on the geostationary estimates alone or may use a combination of the geostationary estimate and near-polar-satellite and surface instrument data when available.

A number of combined algorithms (e.g., Todd et al. 2001; Marzano et al. 2004) are based on the concept of probability matching (Rosenfeld et al. 1993). Such techniques derive a monotonic relationship between an estimator variable (T) and rainfall rates (R) in a manner that ensures the probability density function (pdf) of estimated rainfall derived from the calibration dataset equals the pdf of "observed" rainfall within the same dataset. This is achieved by solving a discrete approximation to the following integral equation:

$$\int_{T'=0}^{T'=T} p(T')\,dT' = \int_{R'=R}^{\infty} p(R')\,dR'. \qquad (1)$$

Note that (1) assumes an inverse relationship between T and R. If the relationship between estimator and estimated variables is weak, as will be the case if the former corresponds to pixel-resolution geostationary IR brightness temperatures, this technique may still reproduce local precipitation frequency distributions even if specific events are misidentified or mislocated. Such a product, employed at full resolution, may be useful in some hydrological modeling situations where identifying the presence of high-precipitation events may be more important than precisely determining their locations with the catchment. However, while these techniques make allowance for the uncertainty in the T/R relationship, they do not directly quantify that uncertainty.

The determination of information on precipitation probability distributions directly from the scalar values of estimating variables is an established procedure that forms the basis of statistical downscaling models (Wilby et al. 1998; Wilby and Wigley 2000; Xu 1999). Such models relate aspects of the probability distribution of a variable such as precipitation at a high spatiotemporal resolution to the value of one or more predictor variables available at a lower resolution. If this relationship has been determined using coincident examples of actual predictor and predicted values (Widmann et al. 2003) the resulting probability distribution will represent a combination of two different uncertainties: one resulting from the difference in spatiotemporal scales, the other representing the uncertainty in the relationship between the predictor and predicted variables. If the resolution of the downscaled product were to be reduced toward that of the original model outputs, the former would vanish and the latter predominate. When the two resolutions become equal, the output field would represent the conditional probability distribution of the predicted variable with respect to its predictors.

Stochastic interpolation techniques are closely re-

lated to statistical downscaling approaches (Lanza et al. 2001). In particular, it is possible to conditionally simulate a range of possible rainfall patterns that are consistent with a given pattern of synoptic station data (Lanza 2000; Shah et al. 1996a). Such simulations employ a statistical model to generate an ensemble of random precipitation fields, where the parameters governing the generated precipitation at any given point in the field are determined from the available synoptic data and the local correlation structure of precipitation. The resulting ensemble may be used to drive a deterministic rainfall runoff, mapping out the possible range of hydrological responses to the uncertain precipitation input (Shah et al. 1996b).

b. Algorithm

The error estimation methodology was developed around a straightforward two-satellite precipitation-retrieval technique in which geostationary infrared brightness temperatures were used as the sole predictor of precipitation and a combined passive microwave/precipitation radar product from the TRMM satellite provided the calibration data. This base algorithm is less sophisticated than current operational techniques such as Adler et al. (2000) or Huffman et al. (2003), particularly in terms of the quantity and variety of input data used. However, the intention was to develop a methodology to represent algorithm uncertainty rather than an improved retrieval algorithm. The simplicity of the base algorithm has a number of advantages in this respect. In particular, the use of a single LEO satellite product for calibration avoids introducing issues of intersatellite and intersensor uncertainties and differences in sensor footprint sizes. The TRMM satellite is particularly suitable for this type of study, since it is not in a sun-synchronous orbit and thus monitors the full range of the diurnal cycle. The approach detailed below may be generalized to any scalar indicator of rainfall, including deterministic precipitation estimates obtained using complex satellite algorithms.

As noted above, the uncertainty in the relationship between IR brightness temperatures $T(x, t)$ and underlying precipitation rates $R(x, t)$ may be represented using the cumulative conditional probability distribution function of R with respect to T: $F(R;T)$. This function may be derived by comparing IR brightness temperatures with coincident precipitation data, $R^*(x, t)$, derived from a secondary source such as passive microwave or precipitation radar satellite data. To model the uncertainty in the precipitation field fully, it would be necessary to possess an error model for $R^*(x, t)$. However, uncertainty in the T/R relationship is significantly higher than that associated with most other types of precipitation measurement, and, in the absence of a suitable error model, an assumption that the secondary data represent a reasonable approximation to ground truth may be expected to yield an acceptable representation of overall uncertainty. The conditional distribution function is most readily derived if it is assumed to remain constant within a given calibration domain. In this case, the changing effects of orography, surface type, the diurnal cycle, and other spatially and temporally varying forcing factors within each bounded spatiotemporal domain will be incorporated into the overall uncertainty in the T/R relationship for that domain.

Once $F(R;T)$ has been determined for a given calibration domain, an ensemble of possible realizations of the precipitation field may be stochastically generated using

$$R_i(\mathbf{x}, t) = \begin{cases} F^{-1}\{N[z_i(\mathbf{x}, t)];T(\mathbf{x}, t)\} & N[z_i(\mathbf{x}, t)] \geq F[0;T(\mathbf{x}, t)] \\ 0 & N[z_i(\mathbf{x}, t)] < F[0;T(\mathbf{x}, t)] \end{cases}, \tag{2}$$

where $N(z)$ is the cumulative standard normal distribution function and $z_i(x, t)$, $i = 1, 2, \ldots N$, are a set of standard normal random fields that display the same spatiotemporal structure as the uncertainty in the precipitation field.

If precipitation is assumed to constitute a second-order-stationary random spatial field then its spatiotemporal structure may be described using a normalized covariance function of the form

$$C_R(\delta\mathbf{x}, \delta t) = \frac{\langle R(\mathbf{x} + \delta x, t + \delta t)R(\mathbf{x}, t)\rangle - \langle R(\mathbf{x} + \delta x, t + \delta t)\rangle\langle R(\mathbf{x}, t)\rangle}{\langle R(\mathbf{x}, t)^2\rangle - \langle R(\mathbf{x}, t)\rangle^2}, \tag{3}$$

where $\langle R\rangle$ denotes the expected value of R. The structure of each field, $R_i(x, t)$, in the ensemble will be determined by a superposition of the structures of $T(x, t)$ and $z_i(x, t)$, the exact combination depending on the pertinent information content of $T(x, t)$. To ensure that the covariance function of each field, $R_i(x, t)$, in the ensemble matches that possessed by the secondary precipitation data, $R^*(x, t)$, the covariance function for the standard normal random fields, $z_i(x, t)$, must be set equal to the covariance function for $z_{R^*}(x, t)$ where

$$Z_{R^*}(\mathbf{x}, t) = \begin{cases} N^{-1}\{F[R^*(\mathbf{x}, t);T(\mathbf{x}, t)]\} & \text{if} \quad R^*(\mathbf{x}, t) > 0 \\ \text{Undefined} & \text{if} \quad R^*(\mathbf{x}, t) = 0 \end{cases}. \tag{4}$$

Since it is impossible to assign a unique value to z_{R^*} for zero rainfall values, $C_{z_{R^*}}(\delta x, \delta t)$ must be calculated us-

ing data from raining points only. Note that Eq. (4) describes the spatiotemporal structure of the uncertainty in the precipitation field rather than the structure of the precipitation field itself.

c. Generating three-dimensional standard normal random fields

To implement the above algorithm, it is necessary to generate three-dimensional random fields with a given spatial and temporal covariance structure. The turning bands method (Mantoglou and Wilson 1982; Mantoglou 1987) is a commonly used method for the generation of synthetic random fields for stochastic hydrological processes (Tompson et al. 1989; Shah et al. 1996a). The technique facilitates the generation of second-order stationary, isotropic, multidimensional random fields with standard normal distributions and specified covariance structures. The method transforms a multidimensional simulation problem into the sum of a finite series of one-dimensional processes. At each point y in the field, a synthetic value $z(y)$ is determined from the weighted sum of corresponding values drawn from a set of one-dimensional processes, assumed to occur along a collection of radial lines, $\mathbf{n}_i$, $i = 1 \ldots L$, uniformly distributed on the unit circle or sphere:

$$z(\mathbf{y}) = \frac{1}{L^{0.5}} \sum_i z_i'(\mathbf{n}_i \cdot \mathbf{y}). \tag{5}$$

Here $z_i'(\xi)$ represents a one-dimensional normally distributed random line process with zero mean and a given covariance function $C_1(\xi)$. The covariance function of the one-dimensional process is related to that of the multidimensional field by

$$C(\mathbf{y}_1, \mathbf{y}_2) = \frac{1}{L} \sum_{i=1}^{L} C_1[\mathbf{n}_i \cdot (\mathbf{y}_2 - \mathbf{y}_1)]. \tag{6}$$

For large L, this relationship approximates to

$$C(\mathbf{y}_1, \mathbf{y}_2) = \frac{1}{|\mathbf{y}_2 - \mathbf{y}_1|} \int_0^{|\mathbf{y}_2 - \mathbf{y}_1|} C_1(\xi)\, d\xi \tag{7}$$

(Mantoglou and Wilson 1982).

To generate the standard normal random fields required to implement the satellite precipitation retrieval algorithm, a three-dimensional turning bands method was used to simulate random fields $z_i(\mathbf{y})$, $y = (x, t)$ in two spatial dimensions and one temporal dimension using 200 lines randomly distributed on the unit sphere (Tompson et al. 1989).

3. Implementation

a. Dataset

The new methodology was tested using data from the Texas and Florida Underflight Experiment (TEFLUN-B) TRMM field validation campaign that took place from 1 August to 30 September 1998 (NASA 2003). The test dataset consisted of surface precipitation estimates from the TRMM 2B31 combined rain-profiling algorithm and Geostationary Operational Environmental Satellite (GOES) band-4 (10.7 μm) infrared imagery. The 2B31 product is based on a combination of TMI and PR data. Polarized 85-GHz data are used to estimate the amount of ice overlying the rain. A suite of rainfall–radiance relationships are then tested to select a PR-derived rain profile that best fits the TMI data (Kummerow et al. 2000). In addition to the satellite data, 2-km spatial resolution instantaneous rainfall-rate products were obtained for the ground-based precipitation radar at Melbourne, Florida (TRMM validation product 2A53). The TRMM precipitation estimates and GOES image data were remapped to a common 0.1° latitude–longitude grid covering the study area (Fig. 1). The radar data were remapped to the same spatial resolution as the TRMM and GOES data and temporally aggregated to a 15-min time step that matched, as closely as possible, the sampling interval of the GOES imagery.

b. Calibration

The implementation of the ensemble retrieval methodology requires the determination of two properties of the precipitation field: the conditional distribution of rainfall rates with respect to IR brightness temperatures, $F(R;T)$, and the covariance function for the uncertainty in the precipitation field, $C_{z_R*}(\delta x, \delta t)$. To evaluate a satellite-only procedure, both of these properties were determined using coincident TRMM 2B31 data and GOES band-4 imagery. The selection of optimal calibration domain sizes for combined satellite algorithms is the subject of ongoing study (Kidd et al. 2003; Todd et al. 2001; Xu et al. 2000). While smaller domains are better at capturing local meteorology, they may constitute less complete samples of the precipitation field, particularly in terms of the representation of infrequent high-rainfall events. This study adopted the more cautious approach of treating the entire TEFLUN-B dataset as a single calibration domain. The coincident dataset was derived for the total 7° latitude by 9° longitude area shown in Fig. 1 over the period from 1 August to 30 September 1998. A number of TRMM overpasses within the study period were unusable due to the lack of a coincident IR image. The

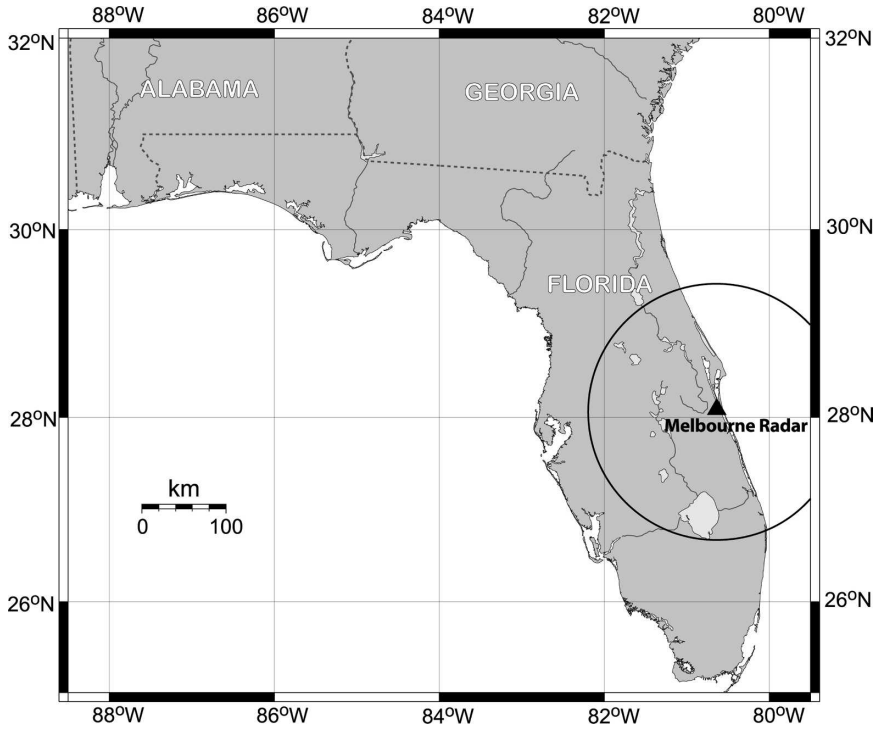

FIG. 1. A map of the area covered by the TEFLUN-B dataset used in this study. The coverage of the Melbourne radar is also shown. Axes are labeled with latitudes and longitudes in °N and °E, respectively.

remaining data comprised 104 overpasses encompassing 33 994 raining points.

The conditional distribution function was determined using a form of surface fitting. The TRMM data were divided into 12 categories, each category corresponding to a 10° range of coincident GOES band-4 IR brightness temperatures. Cumulative distributions for selected categories are shown in Fig. 2. An earlier study (Bellerby and Montaseri 2001) had examined a range of possible functions to model this type of conditional distribution, including normal, two-parameter gamma, two-parameter lognormal, three-parameter lognormal, log-Pearson type III, and Gumbel. An intercomparison using linear L-moment ratio diagrams (Hosking 1990) and probability plot correlation coefficients (Filliben 1975; Loucks et al. 1981) favored the use of a two-parameter gamma distribution function for the raining part of the distribution. A two-parameter gamma distribution function was fitted to the data in each category using a maximum likelihood analysis (Essenwanger 1976). The mean and shape parameters computed for each category were then fitted to empirical functions of T, creating a distribution function that varied continuously with brightness temperature:

$$F(R;T) = \begin{cases} P_o(T) + [1 - P_o(T)] \times P[R;\mu(T),\kappa(T)] & \text{if} \quad R > 0 \\ P_o(T) & \text{if} \quad R = 0 \end{cases} \quad \text{where} \quad \begin{aligned} \mu(T) &= A \times e^{-aT} \\ \kappa(T) &= B \times \text{erf}(b + cT) \\ P_o(T) &= f + \frac{(1 - f)}{2} \text{erf}(g + hT) \end{aligned} . \tag{8}$$

Here $P(X;\mu, \kappa)$ is the two-parameter cumulative gamma distribution function with μ as the distribution mean and κ the shape parameter; $P_o(T)$ is the probability of zero rainfall. Figure 3 shows the fitted data. The fit is most problematic for colder brightness temperatures and higher rainfall rates, particularly with respect

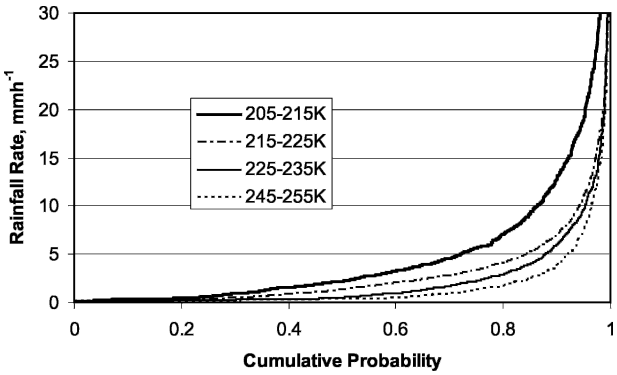

FIG. 2. Cumulative frequency histograms of rainfall rates from TRMM data categorized according to coincident GOES band-4 brightness temperatures.

to the shape parameter. This difficulty appears to be partially the result of the relatively low number of data points corresponding to these categories in the coincident dataset.

Coincident TRMM 2B31 and IR brightness temperature data were additionally used to determine the covariance function $C_{z_R*}(\delta x, \delta t)$. Covariance values were computed for a range of spatial and temporal displacements (δx and δt) within the dataset and fitted to an exponential function of the form

$$C_{z_R}(\delta\mathbf{x},\delta t) = e^{-\sqrt{a\delta x^2 + b\delta t^2}}. \quad (9)$$

With the conditional distribution and covariance functions determined, the turning bands method was used to generate 500 standard normal random fields that were then used to create 500 simulated precipitation fields covering the period 1 August 1998 to 30 September 1998.

4. Results

The ensemble satellite precipitation product was designed to meet two specific criteria. First, that the conditional rainfall probability distributions generated for each geographical location and time should represent both the information content and uncertainty of the satellite precipitation retrieval for that location. Second, that the spatiotemporal structure of each field in the ensemble should resemble that displayed the true precipitation field. These criteria formed the basis of the algorithm validation. Validation was performed using 15-min accumulated 2A53 ground radar data for the

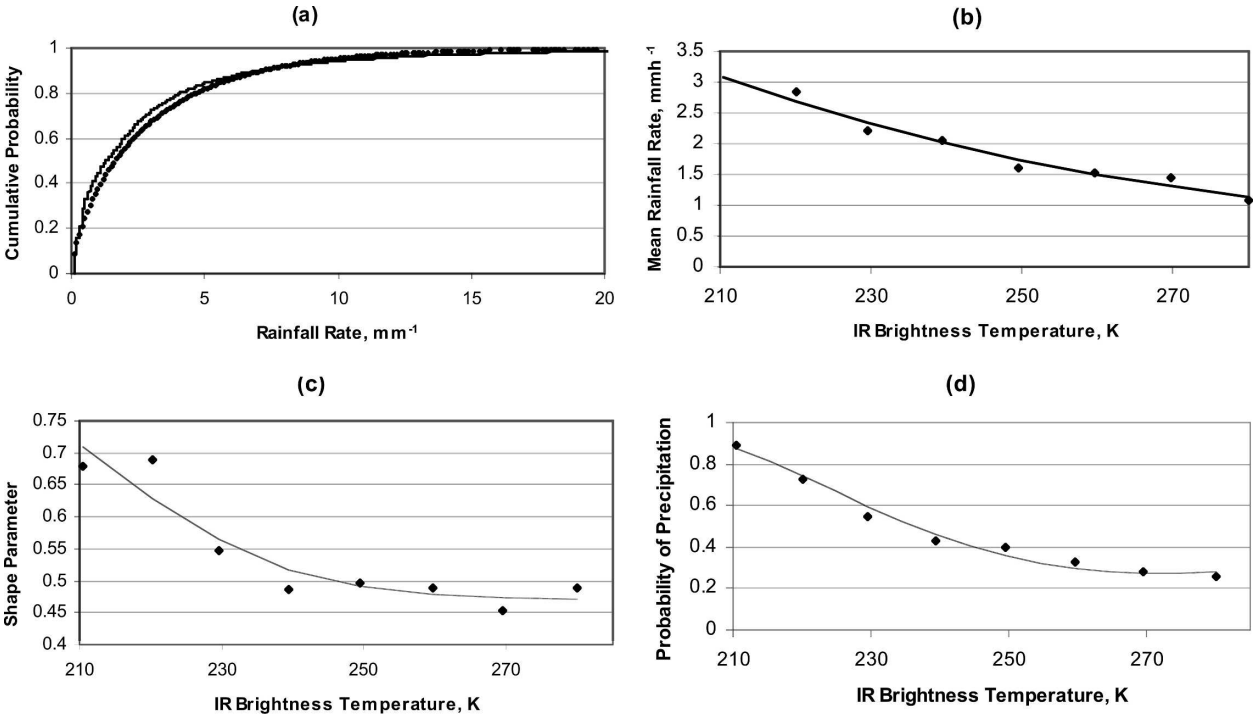

FIG. 3. Modeling the conditional probability distribution of precipitation as a continuous function of IR brightness temperature. (a) A two-parameter gamma distribution function fitted to the frequency histogram for category 215–225 K. (b) Mean rainfall as a function of IR brightness temperature. (c) The gamma-distribution shape parameter as a function of IR brightness temperature. (d) The probability of rainfall as a function of IR brightness temperature.

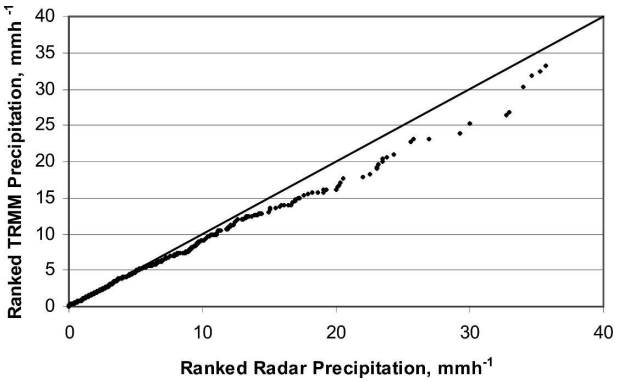

FIG. 4. Ranked precipitation data from the Melbourne radar plotted against ranked TRMM 2B31 precipitation data drawn from the same coincident dataset.

period 1 August–30 September 1998. The validation area was defined by the 150-km range of the Melbourne radar shown in Fig. 1. The TRMM 2B31 and ground radar data were compared to locate and statistical discrepancies. The intersection of these two datasets was quite small, totaling 13 109 coincident points. The TRMM data posses a bias of 0.003 mm with respect to the ground radar, a correlation of 0.61, and a root-mean-squared error of 1.5 mm. Figure 4 plots ranked data from the two datasets to compare probability distributions. The TRMM data appear to underestimate the probability of rainfall over 5 mm as compared to the ground data. However, given the small number of points in the coincident dataset, this conclusion has to remain tentative. The ground radar data themselves, of course, represent an imperfect representation of ground truth. An assessment of uncertainty in the TEFLUN-B ground radar dataset is given in Habib and Krajewski (2002).

Some care must be exercised in interpreting the ensemble precipitation product. The retrieval methodology generates a set of possible realizations of the significant small-scale variability present in the precipitation field. Localized precipitation structures for the same location and time can differ significantly from field to field within the inherent uncertainty of the retrieval process, and individual realizations should not be viewed as consisting of a base estimate to which an error field has been applied. A point-by-point comparison of an individual pixel-resolution field in the ensemble would be expected to yield a very low correlation to observed data at the same resolution. As spatial or temporal resolutions are reduced however, the field should display an increasing correlation to ground truth, especially when the resolution becomes significantly greater than the correlation distance of the precipitation uncertainty. Table 1 provides standard validation statistics [correlation, root-mean-square error (rmse), and bias] for three arbitrary members of the ensemble against the surface radar data, computed for four different spatial resolutions: 0.1°, 0.3°, 0.7°, and 1.4°. The table also includes validation statistics for a probability-matched product, derived by solving Eq. (1) for the coincident TRMM/GOES dataset and applying the resulting brightness-temperature/precipitation relationship to the full coverage of the GOES imagery. The

TABLE 1. Validation of three arbitrary fields from the ensemble product against coincident ground radar data. Equivalent validation results for a deterministic probability-matched product are shown for comparison.

	Ensemble field 1	Ensemble field 2	Ensemble field 3	Probability-matched product
		0.1°		
Correlation	0.01	0.02	0.02	−0.003
Rmse, mm h^{-1}	8.32	7.41	8.10	15.1
Bias, mm h^{-1}	0.80	0.63	0.77	2.6
		0.3°		
Correlation	0.12	0.12	0.12	0.04
Rmse, mm h^{-1}	2.73	2.62	2.74	8.34
Bias, mm h^{-1}	0.01	0.002	0.04	0.68
		0.7°		
Correlation	0.24	0.20	0.20	0.07
Rmse, mm h^{-1}	1.98	2.03	2.03	4.69
Bias, mm h^{-1}	−0.06	−0.04	−0.03	0.11
		1.4°		
Correlation	0.69	0.58	0.68	0.67
Rmse, mm h^{-1}	0.61	0.71	0.62	0.81
Bias, mm h^{-1}	−0.07	−0.06	−0.04	−0.08

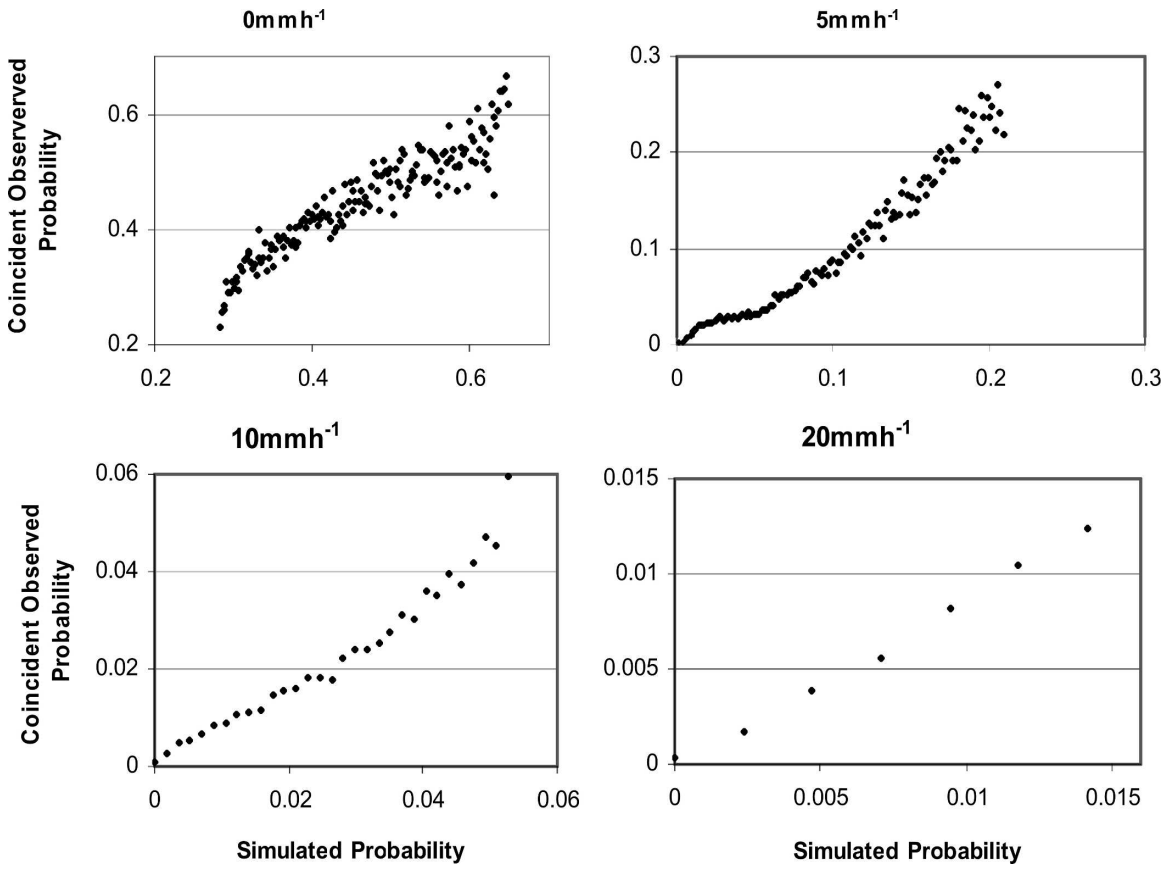

FIG. 5. Simulated exceedence probabilities computed using the ensemble satellite product plotted against observed exceedence probabilities computed using coincident ground radar data for points associated with the given simulated exceedence probability.

probability-matched product and the ensemble field display very similar error characteristics, with an essentially zero correlation displayed at 0.10° resolution and the error characteristics improving rapidly for resolutions above the spatial correlation distance: *1/a* in Eq. (9), determined to be 0.53° by the calibration procedure. At lower resolutions, the ensemble fields display slightly better correlations to ground truth than does the deterministic product.

At full pixel resolution, the ensemble product is designed to reproduce the conditional distribution of precipitation with respect to the satellite data. These conditional pdfs were validated using a comparison of simulated exceedence probabilities and observed exceedence probabilities computed only for points associated with the given simulated exceedence probability. When applied to a range of rain-rate thresholds, this comparison amounts to a validation of the conditional cumulative probability distributions represented by the ensemble product. The *simulated exceedence probability* for each location and time in the grid was defined to be the fraction of fields in the ensemble for which the rainfall rate exceeded a given threshold. These simulated probabilities were binned into 500 intervals (50 for a 0 mm h^{-1} threshold) that were used in turn to categorize coincident ground radar data. The *observed exceedence probability* for each category was then defined to be the fraction of radar measurements in that category exceeding the rain-rate threshold. Figure 5 plots the result of this exercise for four different rain-rate thresholds: 0, 5, 10, and 20 mm h^{-1}. For the 0 mm h^{-1} threshold, there is some tendency to overestimate probabilities at the lower end, causing the algorithm to introduce some additional very light rainfall into the ensemble product. This effect may be traced to differences in the discernment of very low rainfall rates by the 2B31 retrieval algorithm and the ground radar product. For the next two thresholds, the relationship between simulated and observed probabilities is both reasonably linear and relatively unbiased. For higher rainfall rates, 20 mm h^{-1} and above, a more significant bias (13%) is introduced into the simulated probabilities. This bias may be attributed to the low occurrence of higher rainfall rates in the TRMM 2B31 dataset. As

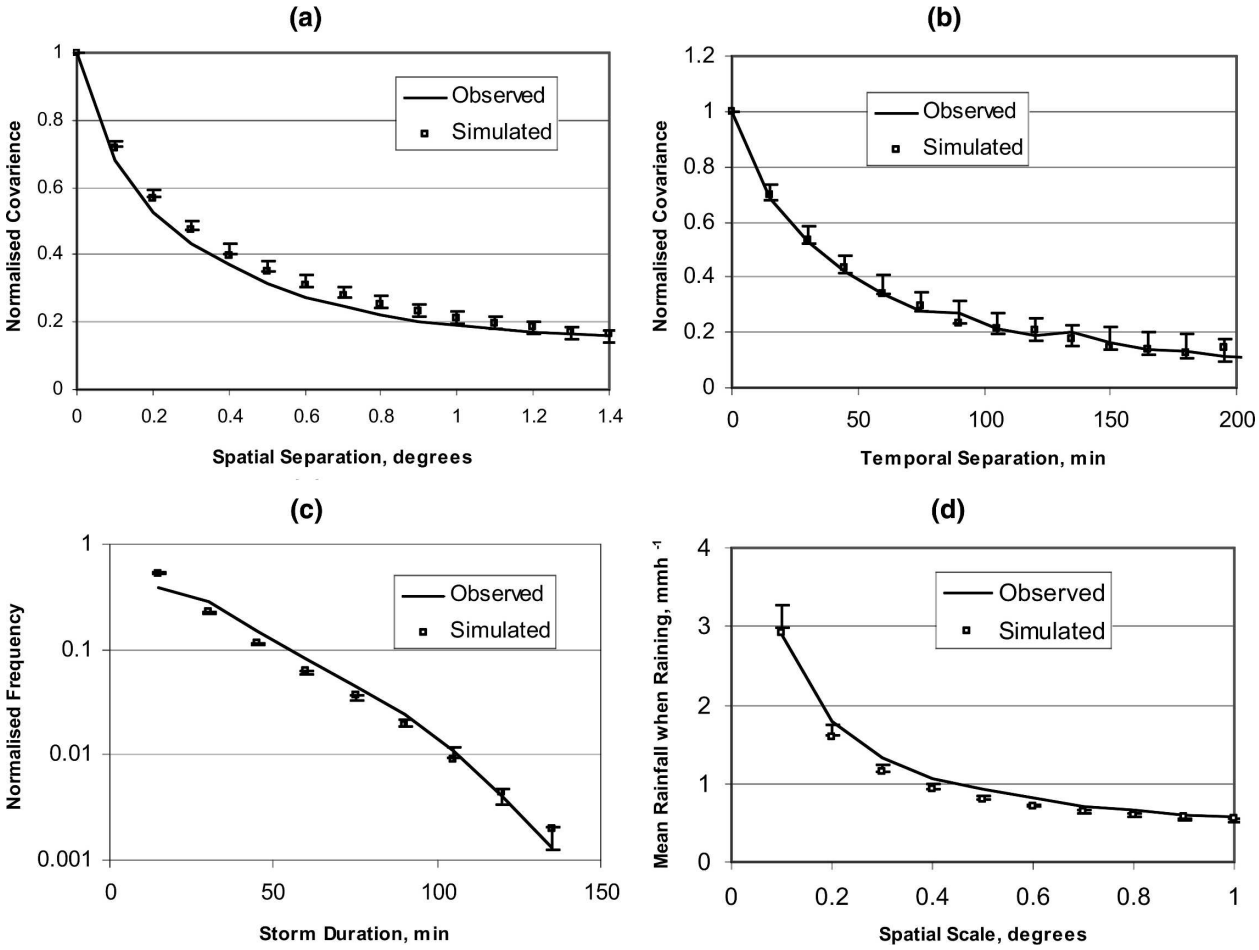

FIG. 6. A comparison of the structural characteristics of an arbitrary component of the ensemble product to those displayed by coincident ground radar data. In each case, the error bars on the simulated data indicate variability across the ensemble (ensemble 10th and 90th percentiles). (a) Covariance of the square root of rainfall rate as a function of spatial separation. (b) Covariance of the square root of rainfall rate as a function of temporal separation. (c) Frequency distribution of storm durations. (d) Mean rainfall while raining as a function of spatial resolution.

noted above, the infrequency of high-rain-rate examples makes the upper end of the conditional distribution function difficult to characterize. This is a fundamental limitation of this type of approach. Equivalent difficulties are shared by deterministic combined satellite precipitation algorithms. However, characterizing the probability distribution of high rainfall rates does appear to be easier than attempting to relate high-rainfall events to specific cloud features (cf. Bellerby et al. 2000). Validation above 20 mm h^{-1} was not possible due to a lack of sufficient data points to construct robust probability estimates. This difficulty could be overcome to some extent using a larger validation dataset. However, there will always be an upper limit on rainfall rates whose probabilities may be validated.

The spatiotemporal structures of individual fields in the ensemble were validated using a number of statistical measures. Figures 6a and 6b compare spatial and temporal covariance functions for an arbitrary member of the ensemble product to those for ground radar data. In each case, covariance values were computed using the square root of rainfall rate in an attempt to alleviate the effects of skewed precipitation distributions. The error bars associated with each simulated value represent variation across the ensemble product (ensemble 10% and 90% percentiles).

The simulated precipitation fields were compared to ground radar data using two additional statistics considered relevant to the utility of the retrieval algorithm in a hydrological modeling context. These were the frequency distribution of storm durations (Fig. 6c) and the mean rainfall rate while raining as a function of spatial scale (Fig. 6d). In the former plot, a storm is defined as contiguous period of nonzero rainfall. These distribu-

tions match each other remarkably well, given the extreme sensitivity of this parameter to rain/no-rain thresholding and the precise temporal structure of the precipitation field. There is a large discrepancy in the number of 15-min (one image) storms that may be attributed to an admixture of low-rain-rate point noise in the algorithm output noted above and a general overestimation of longer storm durations. The latter may be attributed to limitations of the covariance model (9).

To generate Fig. 6d, the mean rainfall while raining was computed for full-resolution, 10-km data and for data aggregated to lower spatial resolutions down to 120 km. The observed and simulated curves display a stronger correspondence than do those for storm duration. This is to be expected since this statistic is less sensitive to rain/no-rain thresholding.

5. Conclusions

A methodology has been developed to quantify the uncertainty present in high-resolution satellite precipitation estimates by generating probabilistic and ensemble representations of the measured precipitation field rather than a single best estimate. The technique has been implemented using GOES channel-4 and TRMM 2B31 data from the TEFLUN-B study and validated against coincident 2A53 data from the Melbourne radar. A comparison of simulated exceedence probabilities and their associated coincident observed exceedence probabilities for a range of instantaneous rainfall rates has been used to demonstrate the effectiveness of the algorithm at generating point conditional probability distributions of precipitation. The algorithm was shown to provide a reasonable representation of the conditional pdf of the precipitation field, although some bias was present in simulated probabilities for rainfall rates at and above 20 mm h^{-1}. Future work is required to confirm that the procedure applies equally well to other rainfall regimes, and an extensive multiyear study would be required to extend the validation to higher rainfall rates, although there will always be an upper limit on the rainfall rates covered by any such validation.

Individual fields randomly chosen from the ensemble have been validated against observed rainfall at a range of spatial resolutions, displaying similar error characteristics to a deterministic precipitation product generated using probability matching. In addition, the spatiotemporal structures of these fields have been shown to match those displayed by the radar data to a reasonable extent. This aspect of the retrieval algorithm could be improved through the introduction of a more complex covariance model that eschewed the isotropy and stationarity assumptions made by the current model.

The algorithm presented here makes no allowance for the uncertainty present in its calibration data. Nor did the validation procedure allow for the uncertainty present in the ground radar data. Given the fact that the relationship between IR brightness temperatures and precipitation is significantly more uncertain than the equivalent relationships for either TRMM 2B31 or ground radar data, the assumption that the latter datasets represented reasonable approximations to ground truth was deemed reasonable in the context of this initial analysis and may prove sufficient for many operational applications. A more complete quantification of uncertainty in the measured precipitation field would require the addition of an error model for the calibration data. If such a model were available, it would be a relatively straightforward task to incorporate it into the methodology outlined in this paper. Operational satellite retrieval algorithms commonly employ multiple polar satellites to supply the calibration data. In such a context multiple uncertainty models would be required.

The further development of this type of algorithm involves the consideration of many questions also raised in the development of deterministic satellite algorithms, including the determination of optimal calibration domain sizes and the implementation of more sophisticated, physically based geostationary cloud indices. The latter issue includes the development of multispectral indices designed to take advantage of new sensors such as the Spinning Enhanced Visible and Infrared Imager (SEVERI) instrument on the Meteosat Second Generation series of spacecraft (Bellerby et al. 2000; Ba and Gruber 2001; Marzano et al. 2004). In addition, the adoption of an ensemble-product approach may make it possible to utilize cloud indices that are not directly applicable to generating rainfall rates but may be more applicable to the determination of local rainfall regimes. An example of such an index is the output from automatic cloud classification techniques such as Tian et al. (1999).

Acknowledgments. A part of this study was funded by the United Kingdom Engineering and Physical Sciences Research Council (EPSRC) Grant GR/N32426/01. Data for this study were obtained from the NASA Tropical Rainfall Measuring Mission Science Data and Information System and from the National Climatic Data Center. Particular thanks are due to Axel Graumann, Dr. Martin Todd of University College London, and Dr. Chris Kidd of the University of Birmingham. Dr. Peter Bauer of the European Centre for Medium-Range Weather Forecasting and Dr. R. Nawaz of the University of Leeds provided useful information on

model validation and downscaling, respectively. Anonymous referees are thanked for constructive suggestions.

REFERENCES

Adler, R. F., and A. J. Negri, 1988: A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor.,* **27,** 30–38.

——, G. J. Huffman, D. T. Bolvin, S. Curtis, and E. J. Nelkin, 2000: Tropical rainfall distributions determined using TRMM combined with other satellite and gauge information. *J. Appl. Meteor.,* **39,** 2007–2023.

Arkin, P. A., and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982–84. *Mon. Wea. Rev.,* **115,** 51–74.

Ba, M. B., and A. Gruber, 2001: GOES Multispectral Rainfall Algorithm (GMSRA). *J. Appl. Meteor.,* **40,** 1500–1514.

Bellerby, T. J., 2004: A feature-based approach to satellite precipitation monitoring using geostationary IR imagery. *J. Hydrometeor.,* **5,** 910–921.

——, and M. Montaseri, 2001: A statistical modelling approach to satellite precipitation monitoring for numerical hydrological models. *Advanced Remote Sensing Techniques in Hydrology,* Bristol, United Kingdom, British Historical Society.

——, M. Todd, D. Kniveton, and C. Kidd, 2000: Rainfall estimation from a combination of TRMM Precipitation Radar and GOES multispectral satellite imagery through the use of an artificial neural network. *J. Appl. Meteor.,* **39,** 2115–2128.

Chevallier, F., and P. Bauer, 2003: Model rain and clouds over oceans: Comparison with SSM/I observations. *Mon. Wea. Rev.,* **131,** 1240–1255.

Essenwanger, O., 1976: *Applied Statistics in Atmospheric Science, Part A: Frequencies and Curve Fitting.* Elsevier, 412 pp.

Filliben, J. J., 1975: The probability plot correlation coefficient test for normality. *Technometrics,* **17** (1), 111–117.

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

Griffith, C. G., W. L. Woodley, P. G. Grube, D. W. Martin, J. Stout, and D. N. Sikdar, 1978: Rain estimation from geosynchronous satellite data—Visible and infrared studies. *Mon. Wea. Rev.,* **106,** 1153–1171.

Habib, E., and W. F. Krajewski, 2002: Uncertainty analysis of the TRMM ground validation radar: Applications to the TEFLUN-B field campaign. *J. Appl. Meteor.,* **41,** 558–572.

Hosking, J. R. M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc.,* **52,** 105–124.

Huffman, G. J., R. F. Adler, E. F. Stocker, D. T. Bolvin, and E. J. Nelkin, 2003: Analysis of TRMM 3-hourly multi-satellite precipitation estimates computed in both real and post-real time. Preprints, *12th Conf. on Satellite Meteorology and Oceanography,* Long Beach, CA, Amer. Meteor. Soc., CD-ROM, P4.11.

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

Kidd, C., D. R. Kniveton, M. C. Todd, and T. J. Bellerby, 2003: Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeor.,* **4,** 1088–1104.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817.

——, and Coauthors, 2000: The status of the Tropical Rainfall Monitoring Mission (TRMM) two years after orbit. *J. Appl. Meteor.,* **36,** 1965–1982.

Kwon, Y. C., M. H. Ahn, B. J. Sohn, H. S. Chung, and H. S. Park, 2000: Rainrate estimation using the GMS-5 IR data and verification. Preprints, *10th Conf. on Satellite Meteorology and Oceanography*, Long Beach, CA, Amer. Meteor. Soc., 99–101.

Lanza, L. G., 2000: A conditional simulation model of intermittent rainfall fields. *Hydrol. Earth Syst. Sci.,* **4,** 173–183.

——, J. A. Ramirez, and E. Todini, 2001: Stochastic rainfall interpolation and downscaling. *Hydrol. Earth Syst. Sci.,* **5,** 139–143.

Loucks, D. P., J. R. Stedinger, and D. A. Haith, 1981: *Water Resource Systems Planning and Analysis*. Prentice-Hall, 559 pp.

Mantoglou, A., 1987: Digital simulation of multivariate two- and three- dimensional stochastic processes with a spectral turning bands method. *Math Geol.,* **19,** 129–149.

——, and J. L. Wilson, 1982: The turning bands method for simulating random fields using line generation by a spectral method. *Water Resour. Res.,* **18,** 1379–1394.

Marzano, F. S., M. Palmacci, D. Cimini, G. Giuliani, and F. J. Turk, 2004: Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale. *IEEE Trans. Geosci. Remote Sens.,* **42,** 1018–1032.

Morrissey, M. L., and J. E. Janowiak, 1996: Sampling-induced conditional biases in satellite climate-scale rainfall estimates. *J. Appl. Meteor.,* **35,** 541–548.

NASA, cited 2003: Tropical Rainfall Monitoring Mission: Field campaigns. [Available online at http://trmm-fc.gsfc.nasa.gov/trmm_gv/.]

Nicholson, S. E., and Coauthors, 2003a: 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.

——, and Coauthors, 2003b: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part II: Validation of TRMM rainfall products. *J. Appl. Meteor.,* **42,** 1355–1368.

Rosenfeld, D., D. B. Wolff, and D. Atlas, 1993: General probability-matched relations between radar reflectivity and rain rate. *J. Appl. Meteor.,* **32,** 50–72.

Shah, S. M. S., P. E. O'Connell, and J. R. M. Hosking, 1996a: Modelling the effects of spatial variability in rainfall on catchment response. 1: Formulation and calibration of a stochastic rainfall field model. *J. Hydrol.,* **175,** 67–88.

——, ——, and ——, 1996b: Modelling the effects of spatial variability in rainfall on catchment response. 2: Experiments with distributed and lumped models. *J. Hydrol.,* **175,** 89–111.

Smith, E., 2003: NASA's scientific agenda for GPM mission. *Proc. Third GPM Workshop: Consolidating the Concept,* Noordwijk, Netherlands, ESTEC, S101. [Available online at http://www.estec.esa.nl/conferences/03C06/.]

Sorooshian, S., K. Hsu, X. Gao, H. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satel-

lite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.,* **81,** 2035–2046.

Tapiador, F. J., C. Kidd, V. Levizzani, and F. S. Marzano, 2004: A neural networks–based fusion technique to estimate half-hourly rainfall estimates at 0.1° resolution from satellite passive microwave and infrared data. *J. Appl. Meteor.,* **43,** 576–594.

Tian, B., M. A. Siakh, M. R. Azimi-Sadjadi, T. H. Vonder Haar, and D. L. Reinke, 1999: A study of cloud classification with neural networks using spectral and textural features. *IEEE Trans. Neural Networks,* **10,** 138–151.

Todd, M., C. Kidd, D. R. Kniveton, T. J. Bellerby, and D. Kilham, 2001: A combined satellite infrared and passive microwave technique for the estimation of small scale rainfall. *J. Atmos. Oceanic Technol.,* **18,** 742–755.

Tompson, A. F. B., R. Ababou, and L. W. Gelhar, 1989: Implementation of the three dimensional turning bands random field generator. *Water Resour. Res.,* **25,** 2227–2243.

Vicente, G. A., J. C. Devenport, and R. A. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution rainfall rate distribution. *Int. J. Remote Sens.,* **23,** 221–230.

Widmann, M., C. Bretherton, and E. P. Salathé, 2003: Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J. Climate,* **16,** 799–816.

Wilby, R. L., and T. M. I. Wigley, 2000: Precipitation predictors for downscaling: Observed and general circulation model relationships. *Int. J. Climatol.,* **20,** 641–661.

——, H. Hassan, and K. Hanaki, 1998: Statistical downscaling of hydrometeorological variables using general circulation model output. *J. Hydrol.,* **205,** 1–19.

Wilheit, T., and Coauthors, 1994: Algorithms for retrieval of rainfall from passive microwave measurements. *Remote Sens. Rev.,* **11,** 163–194.

Wood, E. F., M. Sivapalan, and K. Beven, 1990: Similarity and scale in catchment storm response. *Rev. Geophys.,* **28,** 1–18.

Wu, R., J. A. Weinmann, and R. T. Chin, 1985: Determination of rainfall rates from GOES satellite images by a pattern recognition technique. *J. Atmos. Oceanic Technol.,* **2,** 314–330.

Xu, C. Y., 1999: From GCMs to river flow: A review of downscaling methods and hydrologic modelling approaches. *Prog. Phys. Geogr.,* **23,** 229–249.

Xu, L., X. Gao, S. Sorooshian, and P. A. Arkin, 1999: A microwave infrared threshold technique to improve the GOES precipitation index. *J. Appl. Meteor.,* **38,** 569–579.

——, ——, ——, and B. Imam, 2000: Parameter estimation of the GOES precipitation index at different timescales. *J. Geophys. Res.,* **105** (D15), 20 131–20 143.