PERSIANN-MSA: A Precipitation Estimation Method from Satellite-Based Multispectral Analysis

ALI BEHRANGI, KUO-LIN HSU, BISHER IMAM, AND SOROOSH SOROOSHIAN

Center for Hydrometeorology and Remote Sensing, The Henry Samueli School of Engineering, Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California

GEORGE J. HUFFMAN

NASA Goddard Space Flight Center Laboratory for Atmospheres, and Science Systems and Applications, Inc., Greenbelt, Maryland

ROBERT J. KULIGOWSKI

NOAA/NESDIS/Center for Satellite Applications and Research, Camp Springs, Maryland

ABSTRACT

Visible and infrared data obtained from instruments onboard geostationary satellites have been extensively used for monitoring clouds and their evolution. The Advanced Baseline Imager (ABI) that will be launched onboard the Geostationary Operational Environmental Satellite-R (GOES-R) series in the near future will offer a larger range of spectral bands; hence, it will provide observations of cloud and rain systems at even finer spatial, temporal, and spectral resolutions than are possible with the current GOES. In this paper, a new method called Precipitation Estimation from Remotely Sensed information using Artificial Neural Networks–Multispectral Analysis (PERSIANN-MSA) is proposed to evaluate the effect of using multispectral imagery on precipitation estimation. The proposed approach uses a self-organizing feature map (SOFM) to classify multidimensional input information, extracted from each grid box and corresponding textural features of multispectral bands. In addition, principal component analysis (PCA) is used to reduce the dimensionality to a few independent input features while preserving most of the variations of all input information. The above method is applied to estimate rainfall using multiple channels of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) satellite. In comparison to the use of a single thermal infrared channel, the analysis shows that using multispectral data has the potential to improve rain detection and estimation skills with an average of more than 50% gain in equitable threat score for rain/no-rain detection, and more than 20% gain in correlation coefficient associated with rain-rate estimation.

1. Introduction

The accurate estimation of the amount, and temporal and spatial distribution of precipitation is critical to a wide range of applications from global climate modeling to local weather and flood forecasting. Precipitation is a key component of the earth's hydrological cycle and it has great effects on human lives and property. At regional to global scales, the existing ground-based precipitation observation networks are insufficient; thus, satellites provide a viable and attractive alternative. Clearly, the continuing development of satellite-based precipitation retrieval algorithms that provide progressively better estimates of precipitation has been a growing area of research because of the opportunities and the challenges it entails.

On the basis of the assumption that colder cloud tops are correlated with higher rain rate, cloud-top infrared (IR; ~11 μm) data from geosynchronous earth-orbiting (GEO) satellites have been frequently used to provide high spatial and temporal resolution rain retrievals. In general, these indirect approaches give good results when rain estimates are integrated over larger time and

Corresponding author address: Ali Behrangi, The Henry Samueli School of Engineering, Department of Civil and Environmental Engineering, E.4130, Engineering Gateway, Irvine, CA 92697-2175. E-mail: abehrang@uci.edu

space scales, but they tend to provide high uncertainty for instantaneous rainfall estimates at small scales (Arkin and Meisner 1987). In contrast, passive microwave (PMW) data, particularly over oceans, capture hydrometeor-relevant information and facilitate more accurate instantaneous rainfall estimates because they are more directly related to rain rates. To date, the main disadvantage of PMW sensors has been that they are carried onboard low Earth-orbiting (LEO) satellites; therefore, they are limited in their temporal resolution. The Global Precipitation Measurement (GPM) mission will coordinate the collection of higher-quality PMW-based global rainfall observations at ~3-h average revisit time and ~10-km pixel resolution (Hou et al. 2008). Although this represents a significant improvement, the current and future network of PMW sensors will continue to lack the spatial and temporal resolutions that are required by some applications. Among such applications are flash floods caused by extreme convective storms, where the life of the storm from initiation to dissipation can occur within an hour or less and is confined to a highly localized area.

Many recent algorithms show encouraging results by combining data from IR and PMW sensors (Adler et al. 1993: Bellerby et al. 2000, 2009; Hsu et al. 1997, 2009; Huffman et al. 2001, 2007; Joyce et al. 2004; Kidd et al. 2003; Kuligowski 2002; Kummerow and Giglio 1995; Levizzani et al. 1996; Miller et al. 2001; Sorooshian et al. 2000; Todd et al. 2001; Turk et al. 2000, 2003; Xu et al. 1999). Arguably, GEO satellite observations will continue to play an important role in rain estimation (Huffman et al. 2007), and the challenge is to provide more accurate estimates. One possible direction of improving the utility of GEO satellites is to explore the benefit of multispectral imagers for rain detection and estimation. This is consistent with the increasing number of spectral bands on recent and future GEO-based instruments, along with higher time and space resolution. For example, the operational Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument, onboard Meteosat Second Generation (MSG), has 12 spectral bands and currently scans the earth's surface every 15 min, with a pixel size of 3 km at the subsatellite point (Schmetz et al. 2002). Similarly, the future Geostationary Operational Environmental Satellite-R (GOES-R) series, planned to be launched in early 2015, will carry the Advanced Baseline Imager (ABI), which will provide images in 16 spectral bands ranging from 0.47 to 13.3 μm, with improved temporal and spatial resolutions (Schmit et al. 2005).

A great deal of research has been devoted to precipitation-relevant multispectral studies, either using multispectral bands to characterize clouds and understand some of the near-cloud-top microphysical processes associated with precipitation processes (Arking and Childs 1985; Levizzani and Setvák 1996; Pilewskie and Twomey 1987; Rosenfeld and Gutman 1994; Turk and Miller 2005) or to delineate the areal extent of precipitation (Behrangi et al. 2009; Cheng et al. 1993; Inoue and Aonashi 2000; Lensky and Rosenfeld 2003; Lovejoy and Austin 1979; Tsonis 1984). Efforts to estimate rain rate from a combination of channels have been also investigated, mainly limited to using only a single visible and a single infrared band together (Bellerby et al. 2000; Griffith et al. 1978; Hsu et al. 1999; King et al. 1995; Negri et al. 1984; O'Sullivan et al. 1990). A few studies have used multispectral bands (more than two channels) for rain-rate estimation. Kurino (1997) argued that areas with brightness temperature difference $(\text{Tb}_{11\mu\text{m}} - \text{Tb}_{12\mu\text{m}})$ equal or greater than 3K correspond to cirrus clouds with no rain, and areas with $(\text{Tb}_{11\mu\text{m}} - \text{Tb}_{6.7\mu\text{m}}) < 0$ correspond to deep convective cloud with heavy rain. Using these three channels and composite digital radar data, he developed a three-dimensional (3D) lookup table of rain probability and mean rain rate (MRR) to estimate both deep and shallow precipitation. In another effort, Ba and Gruber (2001) developed the GOES Multispectral Rainfall Algorithm (GMSRA), which uses five spectral bands—0.65, 3.9, 6.7, 11, and 12 μm—to estimate rainfall. GMSRA builds on a series of previously developed techniques: raining clouds are identified using spatial temperature gradients (Adler and Negri 1988) and time changes (Vicente et al. 1998), and during the daytime, effective radii of cloud particles(Rosenfeld and Gutman 1994). Rainfall rates are derived from 11-μm brightness temperatures and are also adjusted for subcloud moisture in a manner similar to Vicente et al. (1998). The self-calibrating multivariate precipitation retrieval (SCaMPR; Kuligowski 2002) is another multispectral approach, but it adds the dimension of being calibrated in real time against rain rates from PMW sensors: regression techniques are used to select the optimal predictors (selected from GOES bands 3–6 plus some derived quantities) and coefficients for rain/no-rain discrimination and rainfall rate estimation.

In this paper we develop a multidimensional rain estimation technique called Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Multispectral Analysis (PERSIANN-MSA) to estimate rain rate using multispectral data plus textural features. Textural features represent various aspects of each channel's value across several neighboring grid boxes in a satellite image. The results are also compared to evaluate the value of such multidimensional data. PERSIANN-MSA employs the ANN-based self-organizing feature map (SOFM; Kohonen 1982) to

cluster multidimensional input information in a manner that facilitates the assessment of their individual and combined utility in rain estimation. By clustering input features into localized maps, SOFM has the advantage of facilitating analysis capabilities, and by extension, the ability to interpret the nonlinear output resulting from ANN models (Behrangi et al. 2009; Hsu et al. 1997, 1999; Tapiador et al. 2004) and to shed light on the value of ANN in rain estimation. In section 2, we describe the datasets used in the study and the study location. Algorithm development and scenarios of combined channel selection are described in section 3. The results of these combination scenarios are then compared in section 4 using both statistical scores and visual analyses. Lastly, the conclusions of this study are presented in section 5.

2. Description of the dataset

As part of the algorithm development efforts for GOES-R, the GOES-R Algorithm Working Group (AWG) has archived datasets for development and validation work, including SEVIRI data for the first nine days from each of January, April, July, and October of 2005 at the full temporal (15 min) and spatial (3 km at subpoint) resolution of the instrument. Of the 12 spectral channels for each SEVIRI image, 10 channels in the present study were used, centered at 0.635 (hereafter referred to as VIS0.6), 0.81 (VIS0.8), 1.64 (NIR1.6), 6.25 (WV6.2), 7.35 (WV7.3), 8.70 (IR8.7), 9.70 (IR9.7), 10.80 (IR10.8), 12.0 (IR12.0), and 13.4 μm (IR13.4).

Notice that despite its demonstrated value, the 3.9-μm channel is not considered in this study. There are several studies in support of using the 3.7-μm (and by extension 3.9-μm) channel to retrieve cloud particle properties and rainfall (Arking and Childs 1985; Rosenfeld and Gutman 1994; Rosenfeld and Lensky 1998; Ba and Gruber 2001), retrieval of cirrus cloud parameters (Rao et al. 1995), and observation of convective storm tops and plumes (Levizzani and Setvák 1996; Setvák et al. 2003), among others. During the daytime, the 3.7/3.9-μm channel radiance contains both solar reflection and thermal emission. To consider the effect of solar zenith angle (SZA) on the reflection component of this channel, the solar component must be quantified and uncoupled, which in itself has been a subject of several investigations in the past (Rao et al. 1995; Rosenfeld and Lensky 1998; Setvák and Doswell 1991). However, in practice, most of these studies are subject to some necessary simplifications and assumptions, such as zero transmissivity of clouds, which is only acceptable for optically thick clouds. Some other simplifications have been addressed by Setvák et al. (2003). Therefore, despite of its importance, the SEVIRI 3.9-μm chan-

nel is not considered in the present comparative study, preventing any misinterpretation due to the necessary simplifications.

In this study, the first five days of each month were allocated for algorithm training and calibration, with the remaining four days set aside for verification and testing. A study region covering longitudes 30°W–0° and latitudes 15°S–15°N was selected because 1) it only includes the ocean as background, where the PMW rain estimates used for training and calibration purposes are more accurate; 2) it includes the equatorial region, where a large number of major precipitation events is expected; 3) it minimizes the effects of parallax on the cloud locations because the SEVIRI images are near nadir; and 4) it maximizes the sample size for daylight-only bands by avoiding higher latitudes, where the length of daylight is relatively short in January (Northern Hemisphere) or July (Southern Hemisphere) and SZA becomes significant.

The first three channels (VIS0.6, VIS0.8, and NIR1.6) are highly affected by SZA and must be normalized. We tested two different normalization methods suggested in the literature (i.e., Cheng et al. 1993; King et al. 1995; Minnis and Harrison 1984; Tsonis and Isaac 1985): 1) the inverse of cos(SZA) and 2) the inverse square root of cos(SZA). Our experiments show that, for SZA < 60°, using 1/cos(SZA) results in more reasonable correction (Behrangi et al. 2009), and thus the former method was chosen and applied for grid boxes with SZA < 60°.

Because of the limited amount of available ground observations of rainfall in the study area, a set of combined intercalibrated normalized PMW-derived rain-rate estimates, provided by the National Oceanic and Atmospheric Administration's Climate Prediction Center (NOAA/CPC; Joyce et al. 2004), was used for training, calibration, and validation purposes. The combined PMW data includes rain-rate estimates from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), the Special Sensor Microwave Imager (SSM/I), and Advanced Microwave Sounding Unit/Microwave Humidity Sounder (AMSU/MHS) sensors. The precedent order of TMI, AMSR-E, SSM/I, and AMSU/MHS is used for sensor overlaps in a 30-min period (Joyce et al. 2004). Both the combined PMW rain estimates (with a grid size of about 0.07° × 0.07°) and the SEVIRI multispectral images (with the nominal grid size of 3 km) were remapped onto common 0.08° latitude–longitude grids and approximately coincident pairs were used. Because of the scan time of the SEVIRI images (15 min) and the temporal precision of the combined PMW rain maps (30 min), a 30-min time lag (at worst) between the pairs

TABLE 1. Representation of scenarios used in the present work to assess the role of multispectral and textural features in precipitation retrievals.

Scenario index	Band/bands' wavelength (μm)	Number of textural features per channel	Dimension of input features (pixel value + textural features)	Dimension of input features (after PCA)
1	10.8	1	1	N/A
2	10.8	5	5	N/A
3	6.2 + 7.3 + 10.8 + 12	1	4	N/A
4	6.2 + 7.3 + 10.8 + 12	5	20	4
5	6.2 + 7.3 + 8.7 + 9.7 + 10.8 + 12 + 13.4	5	35	4
6	0.6	1	1	N/A
7	0.6	5	5	N/A
8	0.6 + 10.8	1	2	N/A
9	0.6 + 10.8	5	10	N/A
10	0.6 + 6.2 + 7.3 + 10.8 + 12	1	5	N/A
11	0.6 + 6.2 + 7.3 + 10.8 + 12	5	25	6
12	0.6 + 1.6 + 6.2 + 7.3 + 8.7 + 9.7 + 10.8 + 12 + 13.4	5	50	8

can occur, which to some extent could insert uncertainties in both training and evaluation processes.

3. Methodology

a. Scenario development

To assess the role of multispectral data in improving satellite rainfall estimation, we selected 12 combination scenarios (Table 1). A fair comparison is ensured by distinguishing "anytime" from "daytime" scenarios: "anytime" scenarios are those excluding both visible and near-infrared bands (VIS06, VIS08, and NIR1.6). This results in five scenarios (scenarios 1–5, hereafter group 1) that are applicable to both daytime and nighttime hours, whereas the remainder (scenarios 6–12, hereafter group 2) are only applicable during the daytime. As described in section 2, the daytime hours consist of grid boxes with SZA < 60°, allowing more reliable normalization of visible channels during daytime. Although "anytime" scenarios can be used during daytime and nighttime, the study is restricted to only daytime images to provide a direct comparison of the various approaches. Notice that scenarios 3, 4, 10 and 11 contain GOES 8–11 imager-like channels, and thus their results are arguably extendable to current GOES as well.

b. Input feature extraction and compression

Local textural features around each satellite grid box are useful in rainfall estimation (i.e., Wu et al. 1985). In this study, we extract five features similar to those used in the PERSIANN (Sorooshian et al. 2000) algorithm. These include the grid-box value itself along with the means and standard deviations of 3×3 and 5×5 windows of pixels centered on each grid box. Clearly, using both multispectral images and textural features significantly increases the number of inputs (see Table 1) and imposes significant computational demand on the algorithm. However, the high interband correlations, which reflect redundancy among the spectral bands, can be employed to reduce the dimensionality of the problem. One common way to obtain such reduction is by using principal component analysis (PCA). PCA uses an orthogonal linear transformation of the data to a new coordinate system, such that the first coordinate (principal component) contains the greatest data variance, the second coordinate contains the second largest data variance, and so on (Jolliffe 2002). Depending on the field of application, PCA is also known as Karhunen–Loeve transform (KLT), empirical orthogonal functions (EOFs), or the Hotelling transformation (Anderson 2003; Jolliffe 2002). Briefly, let X be the data matrix containing N variables. The covariance matrix ($\mathbf{C}$) of $\mathbf{X}$ is calculated as follows:

$$\mathbf{C} = \text{Cov}(\mathbf{X}) = E[(\mathbf{X} - \overline{X})^{\text{T}}(\mathbf{X} - \overline{X})], \quad (1)$$

where $\overline{X}$ is the mean value of $\mathbf{X}$. The linear transformation of $\mathbf{X}$ to orthogonal matrix $\mathbf{Y}$ is

$$\mathbf{Y} = (\mathbf{X} - \overline{X})\mathbf{V}, \quad (2)$$

where $\mathbf{Y}$ is the matrix of principal components and $\mathbf{V}$ is the eigenvector matrix of the covariance matrix $\mathbf{C}$. The transformed components in $\mathbf{Y}$ are uncorrelated to each other, and the covariance matrix of principal components is

$$\mathbf{D} = \text{Cov}[\mathbf{Y}] = E[\mathbf{Y}^{\mathrm{T}}\mathbf{Y}] = E[\mathbf{V}^{\mathrm{T}}(\mathbf{X} - \overline{X})^{\mathrm{T}}(\mathbf{X} - \overline{X})\mathbf{V}]$$

$$= \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_N \end{bmatrix}, \tag{3}$$

where $\mathbf{D}$ is diagonal matrix of eigenvalues of $\mathbf{C}$. The total variance of the data matrix can be represented as $\text{Tv} = \sum_{i=1}^{n}\lambda_i$, where the λ_i values are the eigenvalues ranked from the largest to the smallest variance. For relatively high-dimensional scenarios in our study, as shown in the last column in Table 1, the number of principal components selected for the developmental data preserve 99% of the total variance.

c. Classification of the input features

In this study, the ANN-based SOFM technique (Kohonen 1982) is used to classify input vectors into a number of classes (hereafter clusters), each presenting a unique combination of input features. Rain observations are not used to train the SOFM because experience shows that uncertainties in the assumed-true rain rate tend to degrade the result. The cluster centers are arranged into a two-dimensional discrete map, randomly initialized near the center of the feature space, and trained by introducing input vectors one by one. After training, the clusters are spread out in the input feature space, and each cluster center represents neighboring input vectors with similar properties (Behrangi et al. 2009). By using the SOFM technique, features within each cluster retain the same order in which they were introduced to the network. This property helps to visualize the input features as described in section 3d.

The process of training the SOFM, which is further described in Kohonen (1982) and Hsu et al. (1999), consists of introducing input vectors one by one from the training dataset to the network. All input feature must be standardized so that they become comparable to one another in magnitude. The distance d between each standardized input vector (x_i, $i = 1, \ldots n_0$) and the center of each SOFM cluster is calculated as follows:

$$d_j = \left[\sum_{i=1}^{n_0}(x_i - w_{ij})^2\right]^{1/2}, \quad j = 1, \ldots, n_1, \tag{4}$$

where n_1 is the total number of clusters and w_{ij} is the cluster-center vector (weight vector) of the SOFM that connects input feature i to the specified cluster j.

The best-matching SOFM cluster c (winning node) is the one corresponding with the minimum distance (d_c) between the input feature vector and the SOFM connection weights w_{ij} as follows:

$$d_c = \min(d_j), \quad j = 1, \ldots, n_1. \tag{5}$$

Through a recursive process of competitive cluster selection and weight adjustment, the locations of the cluster centers are incrementally shifted in the N-dimensional vector space until they become stable. Thereafter, the trained SOFM has the ability to assign any arbitrary input feature vector x_i to an SOFM cluster (with fixed weights) according to its minimum distance.

d. Calculating mean rain rate for each cluster

Given the classification described in section 3c, each SEVIRI grid box is assigned to one of the SOFM clusters, bringing with it the corresponding coincident PMW rain estimate. By processing the entire algorithm development dataset through the SOFM network, the MRR for each cluster is calculated as follows:

$$\text{MRR}_C = \frac{\sum \text{RR}_C}{\text{Nr}_C + \text{Nn}_C}, \tag{6}$$

where MRR_C is the mean rain rate for cluster c, RR_C is the rain rate (including zero rain rate) of every single grid box of the cluster c, and Nr_C and Nn_C are the numbers of rain and no-rain grid boxes within the cluster c, respectively.

Because the SOFM preserves the topological order of the input features (Hsu et al. 1999; Kohonen 1982), a distinct map of input features and MRR can be displayed for each cluster. Figures 1a and 1b show these input maps for scenario 8, which has only two input features. The grids in Fig. 1 represent clusters arranged in a 2D network with a size of 20×20. Comparison of albedo (Fig. 1a), brightness temperature (Fig. 1b), and MRR (Fig. 1c) maps demonstrates that brighter and colder cloud (similar to convective clouds) is associated with higher rain rates (zone B), and almost no precipitation is expected from low-reflectance regions, either with high temperature (zone D, similar to clear-sky condition) or low temperature (zone C, similar to thin cirrus clouds).

e. Precipitation estimation

The power-law regression (e.g., Martin et al. 1990; Vicente et al. 1998; Kuligowski 2002) and the histogram matching (e.g., Huffman et al. 2007; Kidd et al. 2003; Todd et al. 2001; Turk et al. 2003; Hong et al. 2004) techniques have been commonly used in estimating rain rate using brightness temperature of clouds obtained
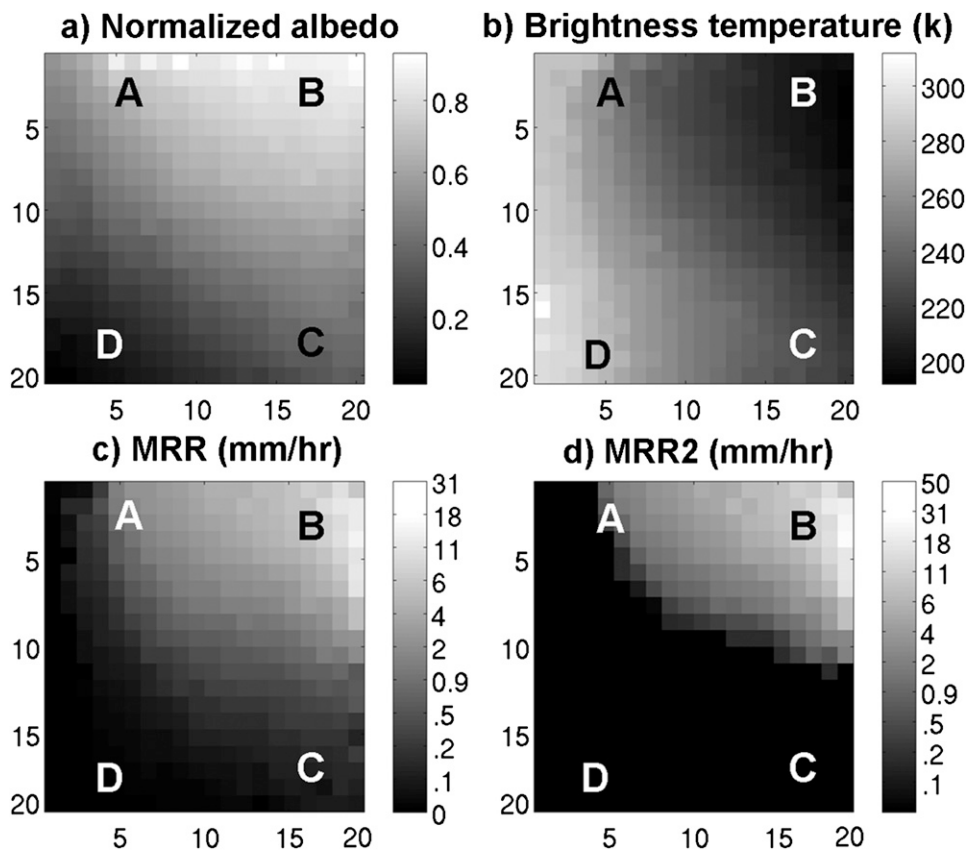
FIG. 1. Visualizing the SOFM weights and clusters' mean rain rates before and after redistribution for scenario 8: (a) albedo (VIS06), (b) Tb (IR10.8), (c) MRR, and (d) MRR2. The 400 SOFM cluster centers (weights) are arranged in a 20 × 20 cluster structure, as shown on the axes of each panel.

from a single IR band. Although the latter techniques may reduce systematic errors from spatially mismatched cloud features and surface rainfall (Kidd et al. 2003), they are based on the assumption that rainfall rate monotonically increases as cloud-top temperature decreases (Scherer and Hudlow 1971: Scofield 1987). However, as illustrated in Fig. 1, this assumption does not always hold: a relatively warm and thick cloud (high albedo) can produce substantial rainfall (zone A), whereas in some cases, almost no precipitation is observed when cold clouds have low reflectance (i.e., cirrus cloud, zone C).

The classification of multispectral data to discrete clusters and the subsequent calculation of MRR for each cluster (Fig. 1c) provide the basis to populate the rain histogram using clusters (which lumps many input features together) as opposed to the IR Tb alone. In other words, using multispectral (or textural feature) classification and statistics, the notion that "more intense rainfall belongs to colder temperature group" is modified to "more intense rainfall belongs to a group of higher MRR." This view suggests applying the histo-

gram matching technique to redistribute the data to the clusters based on the MRR of each cluster.

As such, the MRR for each cluster is first calculated. The cluster (containing N_1 samples) that presents the highest MRR is given the first rank and the cluster (containing N_2 samples) that demonstrates the second highest MRR is labeled the second rank. This continues for all clusters, and the number of samples in each cluster is recorded. Afterward, the first-ranked cluster will be assigned the N_1 highest rain-rate values, the second highest MRR cluster will be given the next N_2 highest rain-rate values, and so on for all of the clusters in the map, and then a new mean rate (MRR2) is calculated for each cluster. A comparison of MRR2 (Fig. 1d) with MRR (Fig. 1c) shows that MRR2 is increased in clusters having high MRR and decreased in clusters containing low MRR.

4. Evaluation and comparison of results

The PMW rain estimates are taken to be the "observation" and are used to build the contingency table with

TABLE 2. Overall statistics for the studied scenarios. Notice that negative numbers for FAR and RMSE gains represent improvement in skill.

Scenario index	ETS	ETS gain (%)	BIAS$_a$	POD	POD gain (%)	FAR (%)	FAR gain (%)	CORR	CORR gain (%)	RMSE (mm h^{-1})	RMSE gain (%)	BIAS$_v$
1	0.301	0	0.946	0.467	0	0.506	0	0.531	0	0.849	0	0.984
2	0.316	5.0	0.973	0.491	5.1	0.496	−2.0	0.549	3.4	0.820	−3.4	0.998
3	0.322	7.1	0.890	0.476	1.9	0.466	−8.1	0.541	1.9	0.835	−1.6	0.964
4	0.335	11.3	0.940	0.502	7.5	0.466	−8.0	0.555	4.6	0.808	−4.8	0.981
5	0.339	12.7	0.962	0.512	9.8	0.467	−7.7	0.562	5.7	0.800	−5.7	0.990
6	0.407	35.1	0.912	0.565	21.1	0.380	−24.9	0.534	0.6	0.863	1.65	0.956
7	0.434	44.4	0.924	0.595	27.4	0.357	−29.6	0.607	14.3	0.775	−8.7	0.930
8	0.428	42.1	0.913	0.585	25.3	0.359	−29.0	0.620	16.7	0.769	−9.3	0.943
9	0.441	46.7	0.925	0.601	28.8	0.350	−30.9	0.625	17.8	0.752	−11.4	0.949
10	0.435	44.4	0.912	0.591	26.6	0.352	−30.5	0.622	17.2	0.778	−8.4	0.970
11	0.458	52.3	0.901	0.608	30.3	0.325	−35.9	0.642	20.8	0.748	−11.9	0.967
12	0.462	53.4	0.900	0.611	30.9	0.321	−36.6	0.645	21.4	0.742	−12.5	0.960

concurrent SEVIRI-based rain estimates, which represent "model predictions." The construction of the contingency table is based on identifying binary (0/1 or yes/no) events. This is accomplished by selecting a threshold (0.1 mm h^{-1}) above which a rain event would be considered to have occurred. This approach yields information on the algorithm's ability to delineate rain/no-rain areas and enables us to compute evaluation statistics that include the following:

$$\text{probability of detection: POD} = \frac{H}{H + M}, \qquad (7)$$

$$\text{false-alarm ratio: FAR} = \frac{F}{H + F}, \qquad (8)$$

$$\text{bias estimate (area): BIAS}_a = \frac{H + F}{H + M}, \quad \text{and} \quad (9)$$

$$\text{equitable threat score: ETS} = \frac{H - G}{(H + M + F) - G}, \qquad (10)$$

where

$$G = \frac{(H + M)(H + F)}{(H + M + F + Z)},$$

H (hits) = number of grid boxes correctly classified as rain,

M (misses) = number of grid boxes incorrectly classified as no rain,

F (false alarm) = number of grid boxes incorrectly classified as rain, and

Z (correct no rain) = number of grid boxes correctly classified as no rain.

The ETS allows the scores to be compared "equitably" across different regimes (Schaefer 1990) and is insensitive to being influenced by systematic over- or underforecasting.

In addition to the categorical evaluation statistics, the following continuous evaluation statistics were computed: 1) correlation coefficient (CORR), 2) root-mean-square error (RMSE) and 3) volume bias (BIAS$_v$). These statistics are used to evaluate the skill of each scenario in estimating rain intensity. Unlike the bias in areal coverage (BIAS$_a$), which is computed using the contingency table, BIAS$_v$ represents the ratio of the average estimate to the average observation:

$$\text{BIAS}_v = \frac{\frac{1}{N} \sum_{k=1}^{N} \text{RR}_{\text{est}}(k)}{\frac{1}{N} \sum_{k=1}^{N} \text{RR}_{\text{obs}}(k)}, \qquad (11)$$

where N is the total number of grid boxes, $\text{RR}_{\text{est}}(k)$ is the estimated rain-rate value for grid box k using PERSIANN-MSA, and $\text{RR}_{\text{obs}}(k)$ is the PMW rain-rate value for grid box k, treated as a rain-rate observation in this study. As a side note, both area and volume bias give only an overall comparison of rain magnitude and its areal extent; they do not measure the magnitude of the errors. For fair evaluation, all statistics should be considered simultaneously, which in some cases is not an easy task, particularly when they show improvements in some aspects but degradation in others.

Table 2 shows the overall statistics for all 12 scenarios listed in Table 1. To facilitate cross comparisons between the scenarios, the statistics in Table 2 are plotted in Fig. 2. Notice that these statistics are computed from all coincident PMW SEVIRI images for which at least
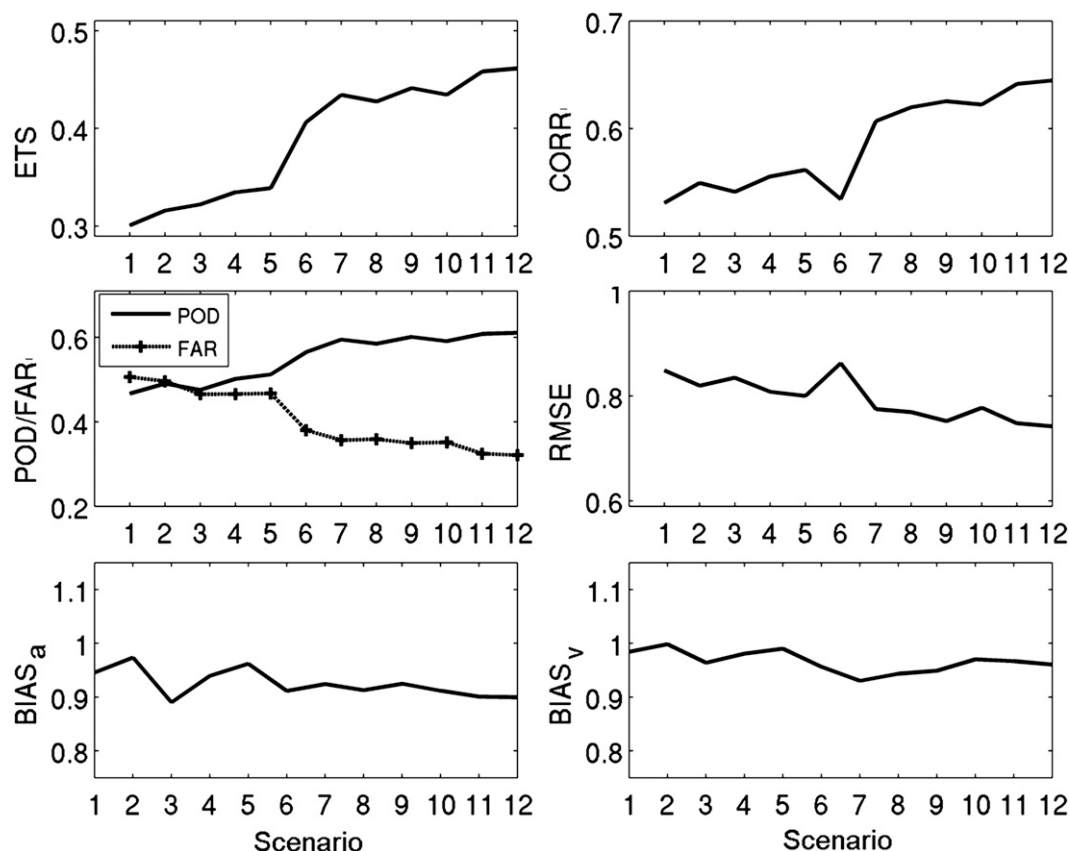
FIG. 2. Visual comparison of overall statistics for all 12 scenarios. (left) Skills for rain and no-rain delineation, and (right) how well scenarios are estimating the precipitation rate.

two percent of the total grid boxes in the image contain rain.

Comparing group 1 with group 2, significant improvement in skill is evident whenever the VIS06 channel is used in combination. Similar results are achieved by replacing the VIS06 channel with VIS08 (not shown here). Scenario 6 (only VIS06 without any textural features) results in substantially better rain/no-rain statistical scores than any of the scenarios in group 1, in which no visible channel is included. However, for rain-rate estimation, using albedo alone does not produce significant improvement over the remaining combinations, including the single IR channel (scenario 1). The role of textural features is highlighted by comparing scenario 7 with scenario 6. Scenario 7 adds extracted textural features to the albedo channel in scenario 6 and shows significant improvements over the IR-only scenario 1, particularly for rain/no-rain detection. This implies that textural features extracted from VIS06 add information about cloud type, and thus improve the classification and result in more robust estimation of rain intensity. Arguably, the important role of the visible channel for

precipitation retrieval can be linked to its utility in removing thin cirrus clouds. In addition, using IR-only data may result in screening out precipitation from relatively warm but dense raining clouds, which is an error that can be avoided by adding the VIS channel. Our assessment of the value of visible channels for rain retrieval is consistent with some previous studies (e.g., Ba and Gruber 2001; Tsonis and Isaac 1985).

Comparison of statistics within group 1 shows that scenarios containing more spectral channels generally produce better skill for both rain detection and estimation. However, the improvement is more pronounced when comparisons are made for the rain/no-rain delineation problem. This observation is also valid for the scenarios in group 2. Note that the improvements within each group (groups 1 and 2) are less significant than between groups, highlighting the important role of the visible channel. Although including more spectral bands was found to be effective, the statistics are systematically better when textural features are added to gridbox values for each spectral image (see scenarios 2, 4, 6, 8, and 10 as compared to scenarios 1, 3, 5, 7, 9, and 11 in Table 2).

Comparison of scenarios 4 and 5 (group 1) and scenarios 11 and 12 (group 2) with other scenarios demonstrates that PCA is a useful technique that compresses input features while preserving precipitation-relevant information (see Tables 1 and 2). The correlated features were compressed to only four independent variables from the original 20 and 35 correlated features in scenarios 4 and 5, respectively, while maintaining 99% of the variance in the data. Similarly, the 25 and 50 correlated features (in scenarios 11 and 12) were compressed into six and eight independent features, respectively. The effectiveness of the PCA technique to compress the original set of input features into a fewer number of independent features reduces the computational burden and speeds up the rain estimation process. This is particularly important when the algorithm is used for operational purpose.

The comparison of all scenarios, except 5 and 12, with scenario 1 (IR only) indicates that using the common multispectral channels (visible, water vapor, and thermal IR channels) mostly available from current GOES series satellites can lead to various levels of improvement, particularly for precipitation detection. Although only one water vapor channel exists in most of the current GOES series, experiments have shown that even a single water vapor channel adds information relevant to precipitation (Ba and Gruber 2001; Behrangi et al. 2009; Desbois et al. 1982; Kurino 1997; Shenk et al. 1976). Following Behrangi et al. (2009), we define a relative gain/loss metric by referencing all scenarios to the performance associated with scenario 1—the commonly used IR ($\sim$11 μm) channel. Each element of the gain/loss column shown in Table 2 is calculated as follows:

$$\% \text{ Gain/Loss}(S, i) = \frac{S_i - S_1}{S_1} \times 100, \quad (12)$$

where S and i represent the performance metric (ETS, POD, FAR, CORR, or RMSE) and scenario index ($i = 1, \ldots 12$), as shown in Tables 1 and 2. For example, in Table 2, the gain in ETS for scenario 2 is computed as follows:

$$\% \text{ Gain/Loss(ETS, 2)} = \frac{\text{ETS}_2 - \text{ETS}_1}{\text{ETS}_1} \times 100$$
$$= \frac{0.316 - 0.301}{0.301} \times 100 = +5\%.$$

Whether the previous index is considered as a gain or a loss depends on whether an increase or decrease of the value of performance measure is better or worse. For example, considering FAR and POD, the former is said to have gained if Eq. (12) yields a negative number,
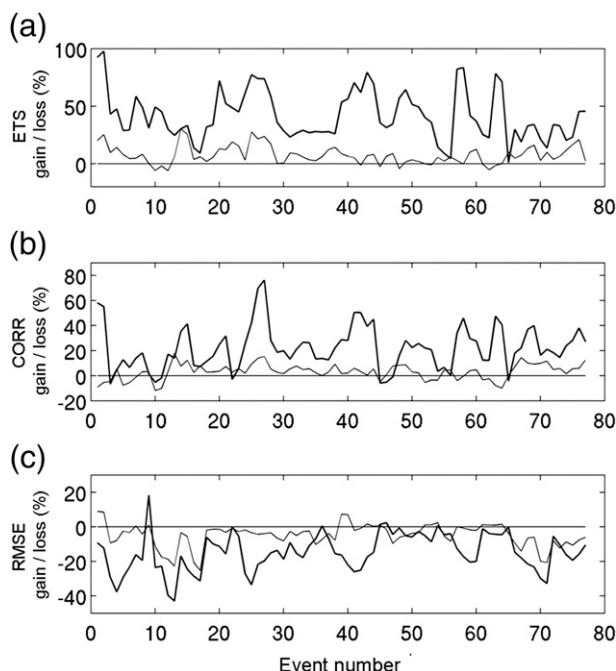

FIG. 3. Gain/loss percent of scenarios 5 (thin lines) and 12 (thick lines) over scenario 1 (IR only), calculated from Eq. (12) for (a) ETS, (b) CORR, and (c) RMSE. Notice that negative numbers for RMSE gain represent improvement in skill over scenario 1.

whereas the latter gains when Eq. (12) produces a positive value.

Employing the additional SEVIRI channels (i.e., NIR1.6, WV8.7, IR9.7, and IR13.4) results in only marginal improvements. Although more detailed study is needed to assess the role of each of these spectral bands, the complementary role of NIR1.6 for rain detection has been reported in previous studies (e.g., Capacci and Conway 2005; Inoue and Aonashi 2000). Within group 1, the all-infrared channel combination along with individual channel textural features (scenario 5) results in 12.7%, 9.8%, 7.7%, 4.6%, and 4.8% gain in ETS, POD, FAR, CORR, and RMSE, respectively. Including the remaining "daytime spectral bands" (scenario 12) leads to further skill improvements of 53.4%, 30.9%, 36.6%, 21.4%, and 12.5% respectively.

To further examine gains and losses due to using multispectral data and textural features, we selected 77 precipitation events from the subset of coincident SEVIRI PMW overpass images, containing at least 3000 rainy grid boxes. This criterion was set to narrow the comparison to relatively extensive rainfall events. For each event, Eq. (12) was applied to the evaluation statistics of scenarios 5 and 12. Figure 3 summarizes the gain/loss values associated with all 77 events. Except for a few events, both scenarios 12 and 5 consistently outperform scenario 1 in terms of rain/no-rain detection
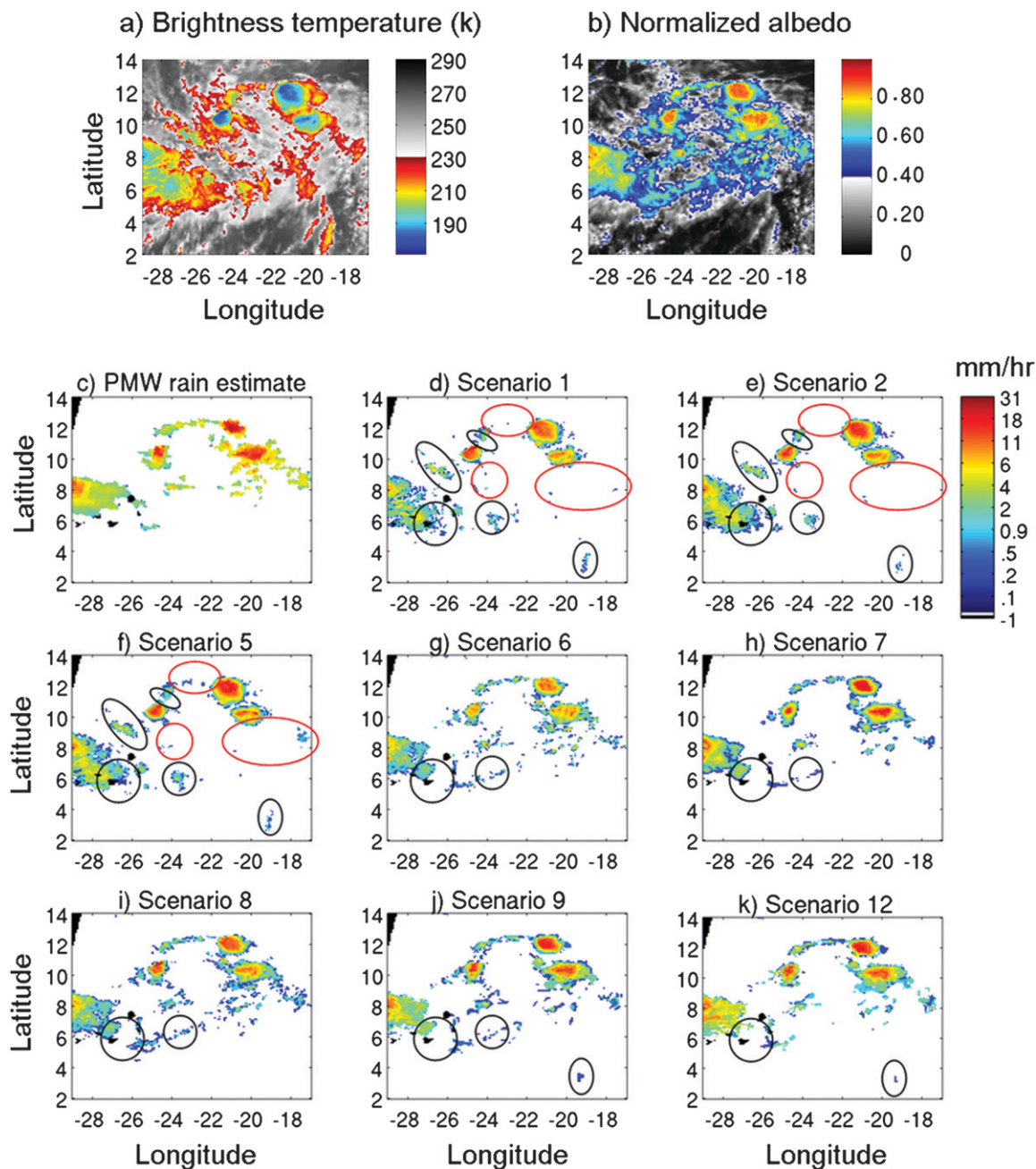
FIG. 4. Comparison of selected combination scenarios for a precipitation event at 1157 UTC 9 Jul 2005: (a) Tb (10.8 μm), (b) normalized albedo, (c) PMW rain rate, and (d)–(k) rain rate estimates for scenarios 1, 2, 5, 6, 7, 8, 9, and 12. Blacked-out blocks have no data.

skill, as indicated by the ETS (Fig. 3a). With respect to rain intensity, the correlation coefficient (Fig. 3b) and root-mean-square error (Fig. 3c) also demonstrate the superior skill of scenario 12. The skill of scenario 5, which employs all "anytime" spectral bands with their textural features, was not as consistent as scenario 12, and in some cases it was below that of scenario 1. Similar results are observed for scenario 4 (not shown), in-

dicating that infrared multispectral data are not as useful for rain-rate estimation as for rain detection.

A visual event–scale exploration of a precipitation event at 1157 UTC 9 July 2005 is shown in Fig. 4. Brightness temperature (IR10.8), normalized albedo (from VIS06) and PMW rain-rate maps are shown in Figs. 4a–4c, respectively. The rain-rate estimates from scenarios 1, 2, 5, 6, 7, 8, 9, and 12 are shown in Figs. 4d–4k.

TABLE 3. Statistics for a precipitation event at 1157 UTC 9 Jul 2005, shown in Fig. 4.

Statistics index	Scenario 1	Scenario 2	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 12
ETS	0.296	0.314	0.344	0.495	0.510	0.504	0.503	0.525
POD	0.476	0.521	0.557	0.677	0.680	0.727	0.699	0.722
FAR	0.441	0.446	0.423	0.275	0.255	0.307	0.284	0.275
$BIAS_a$	0.852	0.941	0.965	0.934	0.910	1.049	0.976	0.996
CORR	0.603	0.634	0.659	0.826	0.837	0.813	0.854	0.857
RMSE (mm h^{-1})	1.613	1.571	1.555	1.200	1.054	1.153	1.044	1.071
$BIAS_v$	0.635	0.748	0.773	0.522	0.674	0.621	0.709	0.827

The ellipses in the scenario figures highlight noticeable regions of false detection (black) and misses (red). The performance measures for these eight scenarios are shown in Table 3.

From Figs. 4d–4f it is clear that scenarios 1, 2, and 5, which do not contain daytime-only channels, have relatively low skill in screening cold and thin cloud, and detecting rainfall from warm and relatively thick clouds. Introducing the daytime-only channels (Figs. 4g–4k) results in significant improvement in evaluation statistics, and even using albedo without any textural features (Fig. 4g) results in satisfactory performance. However, scenario 6 does not show good skill for estimation of intense rain rates, resulting in substantial underestimation of total rain amount. Again, the addition of textural features improves rain retrieval performance, as shown in Table 3 and discerned by comparing Figs. 4e, 4h, and 4j with Figs. 4d, 4g, and 4i, respectively. Consequently, from Fig. 4 and Table 3, one can also infer that using additional spectral bands can lead to important improvements, particularly in delineation of the areal extent of precipitation.

For the above rainfall event and using scenario 8 (bispectral scenario: Tb and albedo), Fig. 5 illustrates a classified cloud map and PMW rain coverage (Fig. 5a) as well as the corresponding map of the SOFM clusters (Fig. 5b). The figure geographically maps the SOFM clusters associated with scenario 8 and is constructed by assigning the corresponding cluster ID to each image pixel. The resulting cluster image (Fig. 5a) is then colored using the 2D color bar in Fig. 5b, which reflects the cluster position in SOFM map described in Fig. 1. In Fig. 5b, similar to Fig. 1, the rain rate fades in all directions from its maximum value in the upper-right corner (zone B) toward no rain in the three remaining corners (zones A, C, and D), and the rain/no-rain boundary denoted by the white line is identical to the rain/no-rain boundary in Fig. 1d. Examining the rain area delineated by the PMW rain estimates (area inside the black line in Fig. 5a), the SOFM clusters help to capture the areal extent of rainfall reasonably well. The colors red and yellow are related to the image pixels of

low IR temperature and high albedo (also see Fig. 1), which correspond well to the intense PMW rain area. The shades of purple are associated with low IR temperature and medium albedo. Clouds with such properties are thin-layer cold clouds or cirrus and are mainly consistent with observed no rain in Fig. 5a. Similarly, the clusters in green shades are warm clouds with medium reflectance. Clusters with dark greens appear to be thick clouds with a potential to give warm rain. Lastly, image pixels associated with the clusters near the bottom-left corner of Fig. 5b are warm and have low albedo, where no rain is observed. Although the figure constructed here is based on VIS06 and IR10.8, similar figures can be constructed for any spectral combination regardless of the number of bands.

The evaluation against the PMW performed so far in this section provides a measure of how well PERSIANN-MSA is able to "fit" the target PMW data. However, it does not provide a measure of the absolute skill of the algorithm, and it does not provide an independent validation dataset. Consequently, validation was also performed on a selected set of VIS/IR scenarios using NASA's TRMM level 2A near-surface rain-rate estimate (TRMM-PR 2A25) as the "truth" rainfall rate. The precipitation radar (PR) onboard TRMM is generally considered to be the most accurate source of rain information for the study area, which is over the open tropical ocean. The original horizontal resolution of the PR rain estimate product is 5 km, and it was remapped to 30-min images of 0.08° × 0.08° (latitude × longitude) grid boxes prior to evaluation.

Table 4 presents the overall PR-based evaluation statistics for the three most distinct scenarios: IR only (scenario 2), all IR-based channels (scenario 5), and all studied channels (scenario 12). In addition, the table includes the same PR-based statistics to evaluate PMW rain estimates, which are exclusively from TMI, given the requirement for coincident PR data. The evaluation statistics are computed for the entire study period, including both calibration and validation periods. The statistics support the previous results, indicating that including spectral channels helps to improve the rain

a. Cluster image
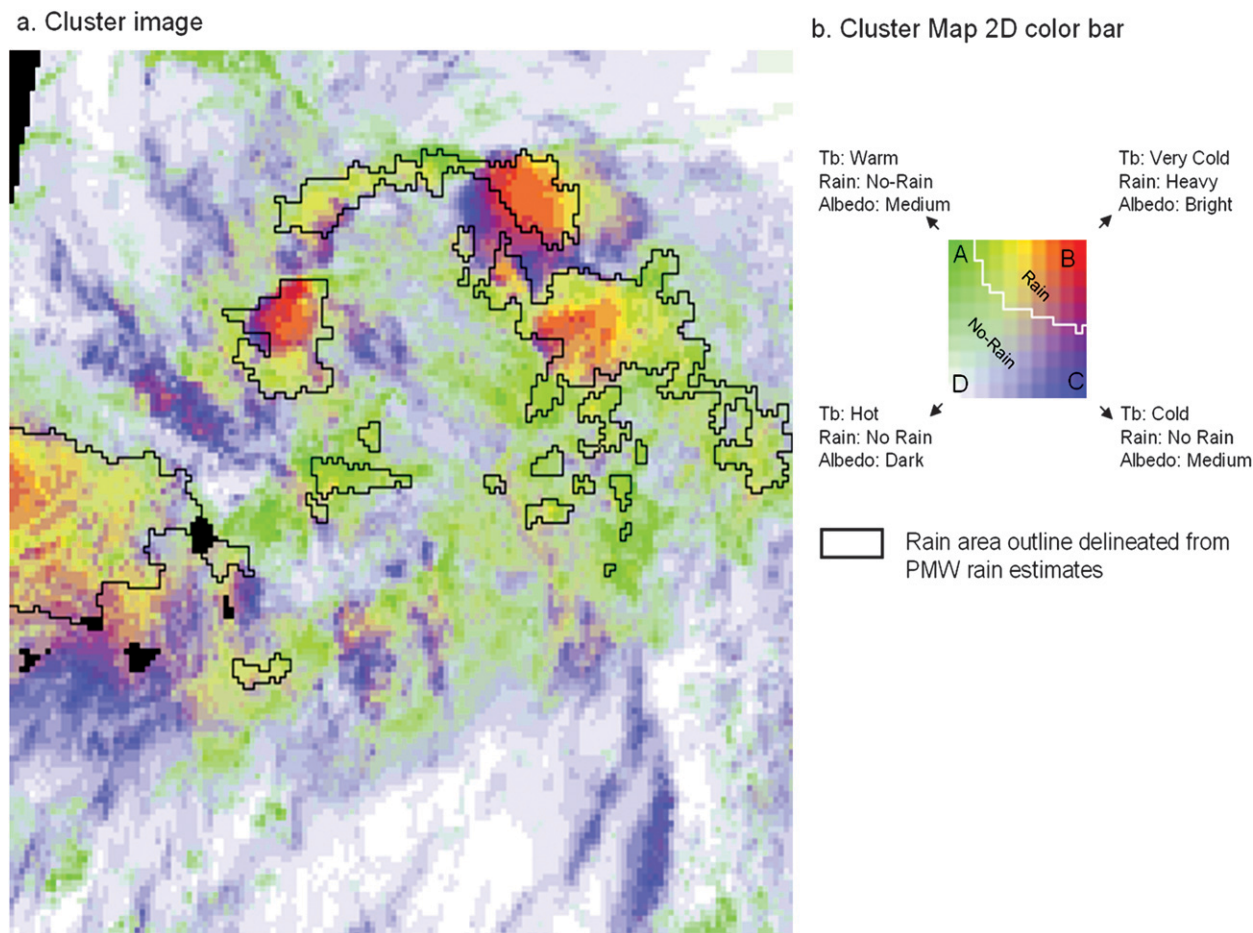

b. Cluster Map 2D color bar

FIG. 5. Example of bidirectional mapping of (a) the 1157 UTC 9 Jul 2005 rain event displayed in Fig. 4 to (b) SOFM cluster centers. The example displays scenario 8 (albedo and Tb). Blacked-out blocks have no data. We recommend viewing this image in conjunction with Fig. 1 and Fig. 4(i).

estimation. For visual demonstration, Fig. 6 and Table 5 document a precipitation event at 1342 UTC 1 July 2005. Maps of brightness temperature (from IR10.8), normalized albedo (from VIS06), and PR rain estimate are shown in the top row of Fig. 6. Scenarios 2, 5, and 12, and the PMW rain estimate are displayed in the bottom row. PMW (Fig. 6g) performs the best followed by scenarios 12 (Fig. 6f), 5 (Fig. 6e), and 2 (Fig. 6d). The superior performance of the PMW estimates relative to the PERSIANN-MSA scenarios is not at all surprising; given that PERSIANN-MSA trained against PMW, the best that could be hoped for would be a statistical tie. Although the TRMM PR is expected to provide the most reliable rain-rate estimates over the study area, it has a poor temporal sampling, and thus its use requires a much longer record of SEVIRI data to achieve any statistically robust conclusion. Even pooling the dependent and independent days as we did here cannot provide as many samples as the PMW does during the

independent days. Consequently, more definite results will only come from a much longer record of SEVIRI data.

5. Conclusions

In this paper, an algorithm called PERSIANN-MSA was developed to estimate precipitation rate from

TABLE 4. Overall PR-based evaluation of the estimated rainfall from scenarios 2, 5, 12 and the combined PMW. Notice that the evaluation is performed within the whole period of this study.

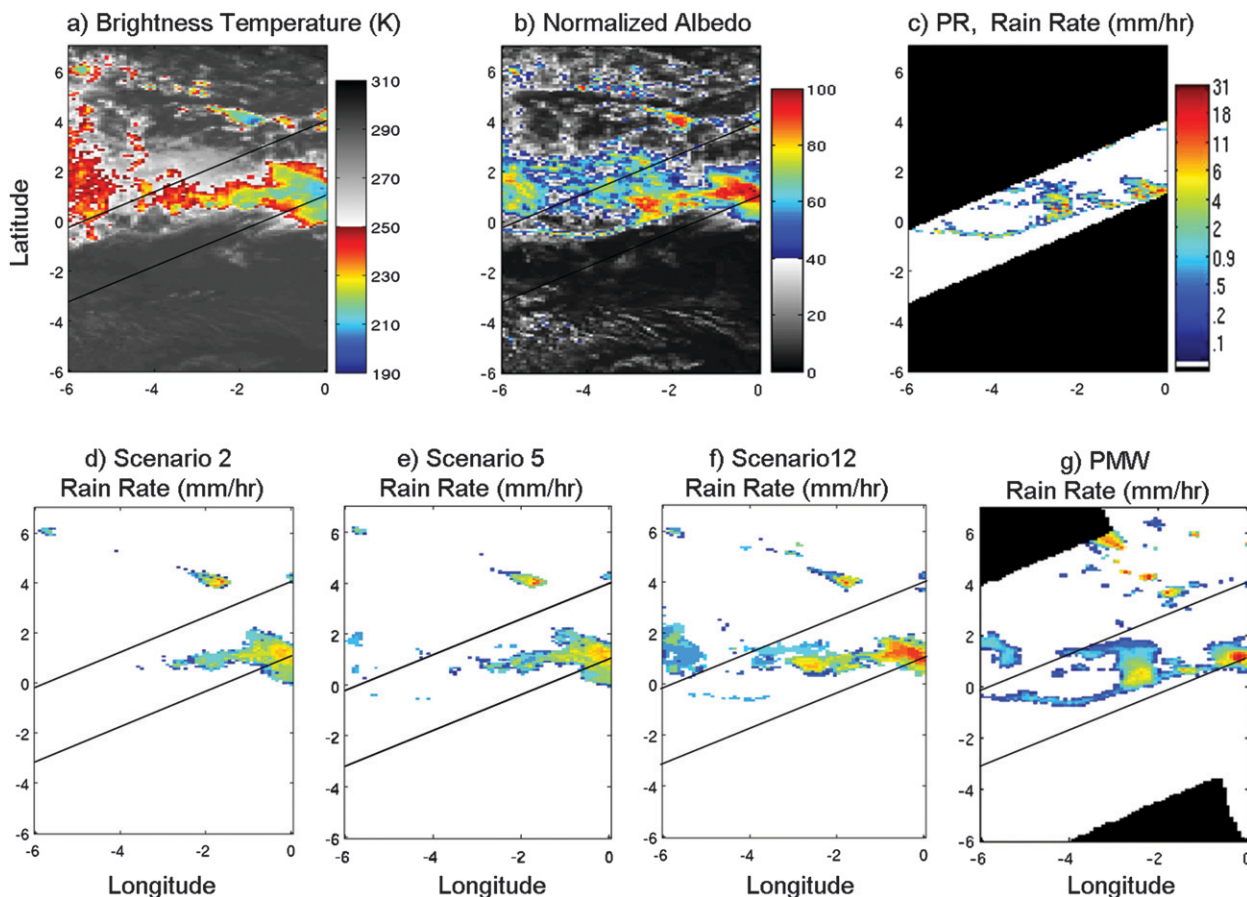| Statistics index	Scenario 2	Scenario 5	Scenario 12	PMW
ETS	0.295	0.313	0.397	0.470
POD	0.420	0.442	0.530	0.711
FAR	0.460	0.443	0.355	0.394
$BIAS_a$	0.778	0.793	0.822	1.173
CORR	0.344	0.365	0.513	0.711
RMSE (mm h^{-1})	1.074	1.079	0.981	0.781
$BIAS_v$	0.856	0.874	0.897	0.973

FIG. 6. Visual demonstration of a precipitation event at 1342 UTC 1 Jul 2005. (a) Tb (10.8 μm), (b) normalized albedo, (c) PR rain rate, and (d)–(g) rain rate estimates for scenarios 2, 5, 12, and combined PMW; (d)–(g) share the color bar in (c). Blacked-out blocks have no data, and the thin black lines delimit the PR swath's location.

multidimensional inputs. The proposed algorithm was tested over the equatorial Atlantic Ocean west of Africa, using 10 SEVIRI spectral bands. An unsupervised feature classification technique, SOFM, was used to classify input spectral features to a predetermined number clusters. The mean rain rate (MRR) for each SOFM cluster was calculated based on time/space-matched PMW rain rate and SEVIRI data. Finally, a probability matching method was used to assign grid boxes with higher rainfall values to clusters with higher MRR. This technique extends the IR-only histogram matching concept to multiple dimensions.

The role of multispectral data and textural features in improving rain estimation, and rain and no-rain detection skills was investigated by defining 12 input scenarios and calculating performance metrics using PMW precipitation estimate as the reference precipitation. Our results indicate the following:

1) Visible channels add significant information, mainly regarding cloud thickness. Including at least one

visible channel (either VIS06 or VIS08) can significantly improve both rain-rate estimation and rain area detection.

2) Other spectral channels were also found useful for improving the algorithm's skill for both rain detection and estimation. However, these improvements were not as significant when VIS information is excluded, as is the case at night. For anytime (day

TABLE 5. Statistics for a precipitation event at 1342 UTC 1 Jul 2005, shown in Fig. 6. Notice that the estimated rain from TRMM PR serves as the reference rain rate.

Statistics index	Scenario 2	Scenario 5	Scenario 12	PMW
ETS	0.209	0.226	0.386	0.463
POD	0.373	0.421	0.667	0.880
FAR	0.456	0.467	0.384	0.402
BIAS$_a$	0.687	0.790	1.081	1.470
CORR	0.382	0.404	0.507	0.641
RMSE (mm h^{-1})	1.490	1.464	1.518	1.254
BIAS$_v$	0.727	0.737	1.228	1.061

and night) scenarios, adding the water vapor channel was found to be effective. Further studies are needed to determine the effectiveness of each single channel.

3) The textural features, as defined herein, provide information about the gridbox neighborhood and can improve statistics. Therefore, extraction of textural features in conjunction with multispectral bands is once more demonstrated as a valuable source of information for precipitation retrieval.

4) PCA is an effective tool for the compression of high-dimension multispectral data. Our analysis demonstrates that the first few principal components are sufficient to extract the majority of the independent information, with the benefit of a significant reduction of the input dimensionality as well as the computational cost of multispectral rain estimation. However, PCA can lead to a loss of information; one must proceed cautiously, because there is no guarantee that the selected directions of maximum variances include the best features for rain retrieval. Plus, a low-order truncation might fail to adequately capture specific rare—but meteorologically important—situations.

Global observations in the visible, water vapor, and thermal IR bands are currently available with relatively high temporal and spatial resolution from existing GEO satellites. As more spectral bands become available from recently launched (e.g., SEVIRI) and future (e.g., ABI) sensors, more emphasis on the analysis and development of multispectral precipitation retrieval algorithms is expected. In this paper, we investigated the value of spectral bands in improving precipitation retrieval using the proposed method. More detailed studies are needed, but we are hopeful that these efforts in conjunction with the anticipated launch and operation of NASA's Global Precipitation Measurement (GPM) mission will help to provide more consistent, higher quality global rainfall observations.

REFERENCES

Adler, R. F., and A. J. Negri, 1988: A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor.,* **27,** 30–51.

——, ——, P. R. Keehn, and I. M. Hakkarinen, 1993: Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data. *J. Appl. Meteor.,* **32,** 335–356.

Anderson, T. W., 2003: *An Introduction to Multivariate Statistical Analysis.* 3rd ed. Wiley, 752 pp.

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.

Arking, A., and J. D. Childs, 1985: Retrieval of cloud cover parameters from multispectral satellite images. *J. Climate Appl. Meteor.,* **24,** 322–333.

Ba, M. B., and A. Gruber, 2001: GOES multispectral rainfall algorithm (GMSRA). *J. Appl. Meteor.,* **40,** 1500–1514.

Behrangi, A., K. Hsu, B. Imam, S. Sorooshian, and R. J. Kuligowski, 2009: Evaluating the utility of multispectral information in delineating the areal extent of precipitation. *J. Hydrometeor.,* **10,** 684–700.

Bellerby, T., 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.

——, K. Hsu, and S. Sorooshian, 2009: LMODEL: A satellite precipitation methodology using cloud development modeling. Part I: Algorithm construction and calibration. *J. Hydrometeor.,* **10,** 1081–1095.

Capacci, D., and B. J. Conway, 2005: Delineation of precipitation areas from MODIS visible and infrared imagery with artificial neural networks. *Meteor. Appl.,* **12,** 291–305.

Cheng, M., R. Brown, and C. G. Collier, 1993: Delineation of precipitation areas using Meteosat infrared and visible data in the region of the United Kingdom. *J. Appl. Meteor.,* **32,** 884–898.

Desbois, M., G. Seze, and G. Szejwach, 1982: Automatic classification of clouds on Meteosat imagery: Application to high-level clouds. *J. Appl. Meteor.,* **21,** 401–412.

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

Hong, Y., K. L. Hsu, S. Sorooshian, and X. G. Gao, 2004: Precipitation Estimation from Remotely Sensed Imagery using an Artificial Neural Network Cloud Classification System. *J. Appl. Meteor.,* **43,** 1834–1852.

Hou, A., G. S. Jackson, C. Kummerow, and J. M. Shepherd, 2008: Global precipitation measurement. *Precipitation: Advances in Measurement, Estimation, and Prediction,* S. Michaelides, Ed., Springer, 131–170.

Hsu, K., X. G. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.,* **36,** 1176–1190.

——, H. V. Gupta, X. G. Gao, and S. Sorooshian, 1999: Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resour. Res.,* **35,** 1605–1618.

——, T. Bellerby, and S. Sorooshian, 2009: LMODEL: A satellite precipitation methodology using cloud development modeling. Part II: Validation. *J. Hydrometeor.,* **10,** 1096–1108.

Huffman, G. J., R. F. Adler, M. 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.

——, and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 38–55.

Inoue, T., and K. Aonashi, 2000: A comparison of cloud and rainfall information from instantaneous visible and infrared scanner and precipitation radar observations over a frontal zone in East Asia during June 1998. *J. Appl. Meteor.,* **39,** 2292–2301.

Jolliffe, I. T., 2002: *Principal Component Analysis.* Springer Series in Statistics, Vol. 29, Springer, 487 pp.

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.

King, P. W. S., W. D. Hogg, and P. A. Arkin, 1995: The role of visible data in improving satellite rain-rate estimates. *J. Appl. Meteor.,* **34,** 1608–1621.

Kohonen, T., 1982: Self-organized formation of topologically correct feature maps. *Biol. Cybern.,* **43,** 59–69.

Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeor.,* **3,** 112–130.

Kummerow, C., and L. Giglio, 1995: A method for combining passive microwave and infrared rainfall observations. *J. Atmos. Oceanic Technol.,* **12,** 33–45.

Kurino, T., 1997: A satellite infrared technique for estimating "deep/shallow" precipitation. *Adv. Space Res.,* **19,** 511–514.

Lensky, I. M., and D. Rosenfeld, 2003: A night-rain delineation algorithm for infrared satellite data based on microphysical considerations. *J. Appl. Meteor.,* **42,** 1218–1226.

Levizzani, V., and M. Setvák, 1996: Multispectral, high-resolution satellite observations of plumes on top of convective storms. *J. Atmos. Sci.,* **53,** 361–369.

——, F. Porcú, F. S. Marzano, A. Mugnai, E. A. Smith, and F. Prodi, 1996: Investigating a SSM/I microwave algorithm to calibrate Meteosat infrared instaneous rainrate estimates. *Meteor. Appl.,* **3,** 5–17.

Lovejoy, S., and G. L. Austin, 1979: The delineation of rain areas from visible and IR satellite data for GATE and mid-latitudes. *Atmos.–Ocean,* **17,** 77–92.

Martin, D. W., B. Goodman, T. J. Schmit, and E. C. Cutrim, 1990: Estimation of daily rainfall over the Amazon basin. *J. Geophys. Res.,* **95,** 17 043–17 050.

Miller, S. W., P. A. Arkin, and R. Joyce, 2001: A combined microwave/infrared rain rate algorithm. *Int. J. Remote Sens.,* **22,** 3285–3307.

Minnis, P., and E. F. Harrison, 1984: Diurnal variability of regional cloud and clear-sky radiative parameters derived from GOES data. Part I: Analysis method. *J. Climate Appl. Meteor.,* **23,** 993–1011.

Negri, A. J., R. F. Adler, and P. J. Wetzel, 1984: Rain estimation from satellites: An examination of the Griffith–Woodley technique. *J. Climate Appl. Meteor.,* **23,** 102–116.

O'Sullivan, F., C. H. Wash, M. Stewart, and C. E. Motell, 1990: Rain estimation from infrared and visible GOES satellite data. *J. Appl. Meteor.,* **29,** 209–223.

Pilewskie, P., and S. Twomey, 1987: Cloud phase discrimination by reflectance measurements near 1.6 and 2.2 μm. *J. Atmos. Sci.,* **44,** 3419–3420.

Rao, N. X., S. C. Ou, and K. N. Liou, 1995: Removal of the solar component in AVHRR 3.7-μm radiances for the retrieval of cirrus cloud parameters. *J. Appl. Meteor.,* **34,** 481–499.

Rosenfeld, D., and G. Gutman, 1994: Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data. *Atmos. Res.,* **34,** 259–283.

——, and I. M. Lensky, 1998: Satellite-based insights into precipitation formation processes in continental and maritime convective clouds. *Bull. Amer. Meteor. Soc.,* **79,** 2457–2476.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting,* **5,** 570–575.

Scherer, W. D., and M. D. Hudlow, 1971: A technique for assessing probable distributions of tropical precipitation echo lengths for X-band radar from Nimbus 3 HRIR data. *BOMEX Bull.,* **10,** 63–68.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.,* **83,** 977–992.

Schmit, T. J., M. M. Gunshor, W. P. Menzel, J. J. Gurka, J. Li, and A. S. Bachmeier, 2005: Introducing the next-generation Advanced Baseline Imager on GOES-R. *Bull. Amer. Meteor. Soc.,* **86,** 1079–1096.

Scofield, R. A., 1987: The NESDIS operational convective precipitation estimation technique. *Mon. Wea. Rev.,* **115,** 1773–1792.

Setvák, M., and C. A. Doswell, 1991: The AVHRR channel 3 cloud top reflectivity of convective storms. *Mon. Wea. Rev.,* **119,** 842–847.

——, R. M. Rabin, C. A. Doswell III, and V. Levizzani, 2003: Satellite observations of convective storm tops in the 1.6, 3.7 and 3.9 μm spectral bands. *Atmos. Res.,* **67–68,** 589–627.

Shenk, W. E., R. J. Holub, and R. A. Neff, 1976: A multispectral cloud type identification method developed for tropical ocean areas with Nimbus-3 MRIR measurements. *Mon. Wea. Rev.,* **104,** 284–291.

Sorooshian, S., K. L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.,* **81,** 2035–2046.

Tapiador, F. J., C. Kidd, K. L. Hsu, and F. Marzano, 2004: Neural networks in satellite rainfall estimation. *Meteor. Appl.,* **11,** 83–91.

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

Tsonis, A. A., 1984: On the separability of various classes from the GOES visible and infrared data. *J. Climate Appl. Meteor.,* **23,** 1393–1410.

——, and G. A. Isaac, 1985: On a new approach for instantaneous rain area delineation in the midlatitudes using GOES data. *J. Climate Appl. Meteor.,* **24,** 1208–1218.

Turk, F. J., and S. D. Miller, 2005: Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.,* **43,** 1059–1069.

——, J. Hawkins, E. A. Smith, F. S. Marzano, A. Mugnai, and V. Levizzani, 2000: Combining SSM/I, TRMM and infrared geostationary satellite data in a near-real time fashion for rapid

precipitation updates: Advantages and limitations. *Proc. EUMETSAT Meteorological Satellite Data Users' Conf.,* Bologna, Italy, EUMETSAT, 705–707.

——, E. E. Ebert, H. J. Oh, B. J. Sohn, V. Levizzani, E. A. Smith, and R. R. Ferraro, 2003: Validation of an operational global precipitation analysis at short time scales. Preprints, *12th Conf. on Satellite Meteorology and Oceanography,* Long Beach, CA, Amer. Meteor. Soc., JP1.2. [Available online at http://ams. confex.com/ams/annual2003/techprogram/paper_56865.htm.]

Vicente, G. A., R. A. Scofield, and W. P. Menzel, 1998: The operational GOES infrared rainfall estimation technique. *Bull. Amer. Meteor. Soc.,* **79,** 1883–1898.

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