On an Enhanced PERSIANN-CCS Algorithm for Precipitation Estimation

MAJID MAHROOGHY,^{*,+} VALENTINE G. ANANTHARAJ,[#] NICOLAS H. YOUNAN,^{*,+} JAMES AANSTOOS,⁺ AND KUO-LIN HSU[@]

* Department of Electrical Engineering, Mississippi State University, Mississippi State, Mississippi

⁺ Geosystems Research Institute, Mississippi State University, Mississippi State, Mississippi

[#] National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee

[®] Center for Hydrometeorology and Remote Sensing, Civil and Environmental Engineering, University of California,

Irvine, Irvine, California

(Manuscript received 2 September 2011, in final form 11 January 2012)

ABSTRACT

By employing wavelet and selected features (WSF), median merging (MM), and selected curve-fitting (SCF) techniques, the Precipitation Estimation from Remotely Sensed Imagery using an Artificial Neural Networks Cloud Classification System (PERSIANN-CCS) has been improved. The PERSIANN-CCS methodology includes the following four main steps: 1) segmentation of satellite cloud images into cloud patches, 2) feature extraction, 3) classification of cloud patches, and 4) derivation of the temperature–rain-rate (T–R) relationship for every cluster. The enhancements help improve step 2 by employing WSF, and step 4 by employing MM and SCF. For the study area herein, the results show that the enhanced methodology improves the equitable threat score (ETS) of the daily and hourly rainfall estimates mostly in the winter and fall. The ETS percentage improvement is about 20% for the daily (10% for hourly) estimates in the summer at some rainfall thresholds. In the winter and fall, the area bias is improved almost at all rainfall thresholds for daily and hourly estimates. However, no significant improvement is obtained in the spring, and the area bias in the summer is also greater than that of the implemented PERSIANN-CCS algorithm.

1. Introduction

Rainfall estimation at high spatial and temporal resolutions is beneficial for research and applications in areas such as weather, climate, precipitation forecasting, hydrology, water resources management, flood forecasting, and agriculture (Anagnostou 2004). Flooding from localized intense precipitation is one of the most serious natural disasters. According to the International Strategy for Disaster Reduction (ISDR) program of the United Nations reports, 8 out of the top 10 most deadly natural disasters in 2007 were flood related. Hence, the accurate monitoring of precipitation is imperative for improving flood and operational weather forecasting systems (Hsu et al. 2009; ISDR 2008).

Several high-resolution satellite precipitation estimation (HRSPE) algorithms, based on several different

DOI: 10.1175/JTECH-D-11-00146.1

approaches, are already in routine use in research and applications: the Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004), Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007), Naval Research Laboratory (NRL; Turk and Miller 2005), Precipitation Estimation from Remotely Sensed Imagery using an Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000), and PERSIANN Cloud Classification System (PERSIANN-CCS; Hong et al. 2004). A HRSPE algorithm can be categorized based on the underlying sensors and platform instrumentations. These sensors include active (radar) and passive microwave (MW) measurements obtained from low Earth orbit (LEO), as well as visible (VIS) and infrared (IR) imagery of geostationary (GEO) satellites. Although MW signals can provide the microphysical information about clouds, the temporal resolution from LEO platforms is not sufficient for high temporal applications. On the other hand, the IR sensors on board the GEO platform can provide high temporal observation.

Corresponding author address: Majid Mahrooghy, Box 9627, Geosystems Research Institute, Mississippi State University, Mississippi State, MS 39762. E-mail: mm858@msstate.edu

However, the IR observations are only correlated to the cloud-top temperature, and are not always physically related to the microphysical properties of clouds (Adler et al. 1994). To achieve higher accuracy, many HRSPE algorithms use a combination of IR satellite data with MW observations and/or ground-based observations for further calibration (Adler et al. 1994; Turk and Miller 2005; Huffman et al. 2007).

Rainfall estimation algorithms using infrared data can also be categorized into the following three groups depending on the method used for information extraction from infrared cloud images: (a) pixel-based, (b) local texture-based, and (c) patch-based algorithms (Hong et al. 2004). In pixel-based algorithms, a rain rate (either fixed or variable) is assigned to every pixel of the cloud and just that pixel alone is considered. Cloud local texture-based techniques calculate the pixel rain rates by considering a range of the neighborhood pixel coverage. Cloud patch-based techniques assign a rain rate to each pixel by considering the cloud coverage under either a specified temperature threshold or specific criteria. The original PERSIANN algorithm (Hsu et al. 1997; Sorooshian et al. 2000) uses the local texturebased method and then extends to the PERSIANN-CCS, which uses cloud patch-based techniques. Mainly, the PERSIANN-CCS algorithm uses cloud patch classification to provide 100 representations (groups) of different cloud patch types. For each of them, a temperaturerain-rate (T-R) relationship is obtained (in the training mode). Therefore, to estimate rainfall for each input cloud patch, the cloud patch first is compared to the representations (groups), and the most similar cloudtype representation is selected and the corresponding T-R is exploited to assign rainfall to the patch. Furthermore, the PERSIANN-CCS uses an exponential curve fitting and applies the probability matching method (PMM) (Atlas et al. 1990) to all patches to obtain the T-R for each cloud-type representation. Because the cloud patches are clustered based on cloud-top temperatures, it is possible that some patches, which have similar cloud-top information but different microphysical properties, are grouped in a cluster. It is also possible that some patches are incorrectly classified to a cluster because of imperfect classification and features extraction.

In this paper, an enhanced version of the existing PERSIANN-CCS algorithm is developed by exploiting wavelet and selected features (WSF), median merging (MM), and selected curve-fitting (SCF) techniques. The WSF approach increases the performance of the cloud patch classification and the MM and SCF methods improve the T-R relationship for each group representing different cloud types. Note that by incorporating wavelet features, we obtain more texture information from

cloud-top temperatures, and also by using a feature selection method, the optimal and effective similarity measures (features) are selected to increase the performance of cloud patch clustering. Also, instead of an exponential curve fitting, which is used in the PERSIANN-CCS, a selected curve fitting is utilized to select a proper curve-fitting method to fit the T-Robtained from applying the MM. Furthermore, the impacts of the improper patches for obtaining the T-R relationship are reduced by applying the MM technique; that is, the MM technique improves the T-R relationship for each group. As a result, an increase in the precipitation estimation accuracy is obtained by increasing the performance of cloud patch classification (effective cloudtype representations) and also creating a more accurate T-R for the clusters.

2. Methodology

The PERSIANN-CCS methodology incorporates four main steps to derive precipitation estimates. The first step is to segment the satellite IR cloud images into patches by using a region-growing method (Gonzalez and Woods 2007). The second step is to extract from the segmented cloud patches features such as statistics, geometry, and texture at different brightness temperature thresholds. Note that throughout the terms temperature and brightness temperature are used to refer to the latter. The next step is to categorize the cloud patches into separate clusters using a self-organizing map (SOM), and the final step is to obtain a relationship between the brightness temperature of cloud patches and the rain rate for every cluster by applying PMM and an exponential curve fitting (Hong et al. 2004).

We have augmented the PERSIANN-CCS while including all of the above steps, with the WSF, MM, and SCF techniques. The implemented PERSIANN-CCS with the above-mentioned techniques are being referred to as mCCS-WMS. We have augmented step 2 by employing WSF and step 4 by employing the MM and SCF methods.

Figure 1 shows a block diagram of the mCCS-WMS algorithm in the training and testing modes. The goal in the training mode is to obtain a *T*–*R* relationship for every cluster. In this mode, the radiance measurements from the *Geostationary Operational Environmental Satellite-12 (GOES-12)* are first calibrated into brightness temperature values. A region-growing segmentation method (Gonzalez and Woods 2007) is used to segment the clouds into patches in the segmentation part (step 1). In the segmentation procedure, the minimum brightness temperature (Tb_{min}) of the clouds is first determined, and then used as a seed. Then, Tb_{min} is



FIG. 1. The PERSIANN-CCS WMS algorithm in the training and testing modes.

incremented by 1 K, and a new set of pixels is identified. If these pixels are neighbors of a seed, then they will be considered as the area of that seed; otherwise, those pixels are treated as new seeds. The threshold of the temperature is iteratively increased to a maximum of 255 K. Afterward, a morphological operation is applied to remove/merge the tiny regions (Gonzalez and Woods 2007; Hong et al. 2004).

In the feature extraction part (step 2), for every patch and at Tb thresholds of 220, 235, and 255 K, the PERSIANN-CCS extracts statistical features that can be categorized into three groups: the minimum and mean temperatures (coldness feature), patch area and patch shape index (geometry), and standard deviation (std dev), including the mean of the local standard deviation, standard deviation of the local STD, gradient, and grayimage texture.

In addition to the PERSIANN-CCS features, additional texture features are obtained from applying a wavelet transform. The wavelet transform is one of the powerful tools in texture analysis; it analyzes localized variations of energy and signal within a time series or the content of images. For instance, in a 1D time series wavelet analysis, the time series is decomposed into a time–frequency space, so one can find out both the principal modes of variability and their variations in time (Torrence and Compo 1998). In a 2D image wavelet analysis, a 1D wavelet transform is first performed along the horizontal direction *x*, and then along the vertical direction *y*. In the first level of the decomposition, the given image is decomposed into one low-pass approximation and three added-detail images, which contain high-frequency information of the image in the vertical, horizontal, and diagonal directions. In the next level, the decomposition is repeated and performed in the lowpass approximation subimage resulting from level 1. Using this process, the wavelet analysis provides the decomposition of an image into different frequency subbands while capturing localization information both in the spatial and frequency domains (Burrus et al. 1997).

To obtain the wavelet coefficients and also wavelet features, a wavelet transform, with a Daubechies mother wavelet (Burrus et al. 1997), is applied to each pixel of the original IR image (before segmentation) using a 7×7 sliding window. The sliding window is decomposed into seven wavelet coefficient subbands (horizontal, vertical, and diagonal detail coefficient subbands for levels 1 and 2, along with an approximation coefficient subband). By calculating the mean and standard deviation of the coefficients' energy for each of the seven subbands, 14 corresponding coefficient values are obtained for each pixel of the IR image. The wavelet features for each patch are the average of these values corresponding to



FIG. 2. (a) GOES-12 cloud-top brightness temperatures at 1245 UTC 4 Feb 2008, (b) mean of the wavelet coefficients' energy for the horizontal detail in level 1, (c) the vertical detail in level 2, and (d) the diagonal detail in level 2.

the pixels covered by the patch at different threshold levels. Because we have three threshold levels (220, 235, and 255), the wavelet features are $14 \times 3 = 42$.

Figure 2a shows an example of the brightness temperature of the infrared clouds in the area of study at 1245 UTC 4 February 2008. The corresponding mean of the wavelet coefficients' energy for the horizontal detail in level 1, vertical detail in level 2, and diagonal detail in level 2 are depicted in Figs. 2b–d, respectively. As these figures show, the high wavelet coefficients are related to the high variation in brightness temperature of the clouds (in different orientations). Note that the level 2 of the decomposition provides more high-frequency variations. This figure also shows that more variation occurs around the big patch, especially in level 1. We see some variation in the center of the patch in level 1 that does not exist in level 2; this means that the center of the patch has smooth variation compared to some parts around of the patch. Thus, we can attain different degrees of detailed information and variation of the IR cloud-top temperature by using the wavelet coefficients.

To have effective and optimal features, we apply a feature selection method to the wavelet features together

with the PERSIANN-CCS features. Seven features are selected from the PERSIANN-CCS and four from the wavelet features, that is, a total of 11 features are exploited for the mCCS-WMS algorithm. These wavelet and PERSIANN-CCS selected features are called wavelet and selected features (WSF). Note that the feature selection method is based on the entropy index evaluation, where different feature sets are generated by a genetic algorithm and the best feature set (the one with the lowest entropy index) is selected (Mahrooghy et al. 2011). The entropy index checks if a feature (or a feature set) provides relevant and useful information.

To classify the cloud patches (step 3), a SOM artificial neural network is employed. In our algorithm, a 10×10 unit map with a hexagonal structure is used for clustering (Kohonen 1982). Figure 3 depicts the process of the training and testing modes of the SOM. In the training mode, the input training pattern is used to adjust the weights of the clusters such that the weights of the winner cluster, along with its neighbors (within the hexagonal graph), are updated. In the testing mode, the most similar cluster to the input pattern (with features of the input patch) is selected as a winner cluster, and the 926



FIG. 3. Clusters in the training and testing modes.

corresponding temperature–rain-rate graph for that cluster (which is explained in the next paragraph) is exploited for rainfall estimation.

In step 4, a T-R curve is assigned to every cluster in the training mode. To acquire this curve, first a T-Rrelationship is obtained for each cluster, and then a curve-fitting technique is employed to fit this T-R relationship. To assign a characteristic T-R to each cluster, a MM technique is used. First, a PMM is applied to each individual patch of the cluster, that is, the T-R relationship for each single patch is identified (depicted as the dashed lines in Fig. 4). The median rain rate at each temperature is obtained from all of the patch T-R values in the cluster corresponding to that temperature (dots in Fig. 4). The MM technique reduces the impact of any improper patches that may have been introduced as a result of imperfect classification and features extraction, as well as a lack of enough information of the cloud or other factors. In other words, although the IR observations obtained by IR sensors on board the GEO platform can have high temporal resolutions, they are only correlated to the cloud-top temperature and are not always physically related to the microphysical properties of clouds. Therefore, it is possible that some patch clouds, which have the same cloud-top information (features) but differ in terms of their microphysical properties (i.e., their corresponding T-R relationship would be different) are grouped in one cluster in the training mode even though classifying the patch clouds based on cloud-top temperatures can differentiate many clouds. Note that in the PERSIANN-CCS algorithm, the PMM is applied to the T-R pixel pairs (temperature pixels and the corresponding rain-rate pixels) of all of the patches within the cluster to obtain the T-R relationship. Thus, in the PERSIANN-CCS algorithm, all patches in the cluster impact the creation of the T-R of the clusters, and it is



FIG. 4. Temperature–rain-rate relationship of patches within a cluster (dashed lines) and median merging samples (filled dots) along with curve-fitting techniques (thick black, gray lines) for the cluster.

not necessary to attain the T-R for each patch. However, in our approach, by using the MM technique, the T-R of all of the patches is obtained, and the effect of improper patches is reduced by using medium merging.

Because the T-R relationship resulting from applying the MM method does not cover the entire range of temperature values, a SCF procedure is then applied to fit the *T*–*R* samples to cover the range of temperatures. That is, a 10th-order polynomial curve fitting and a 5thorder exponential curve fitting (Hong et al. 2004) are computed for the MM T-R points (black and gray lines in Fig. 4, respectively), and the curve with the smaller mean squared error (MSE) against the MM T-R points is selected. Figure 4 depicts how the MM and SCF methods provide a T-R relationship for a cluster. The dashed lines show the T-R relationship for the patches belonging to the cluster. First, the MM approach is applied to the T-R of the patches to merge the rain rates to one value for each temperature. The MM is carried by computing the median of the patch rain rate for temperatures having at least one corresponding patch rain rate for each cluster. These T-R-merged values are shown as filled dot marks in Fig. 4. Two types of curve fitting [polynomial (black) and exponential (gray)] are applied to the T-R-merged samples to cover all of the temperatures between 200 and 255 K. The SCF selects the curve-fitting approach that has a smaller MSE (fitting error) to the median samples. In this example, the SCF selects the polynomial curve fitting because it has a smaller MSE value.

3. Verification of results

The study region encompasses 30°-38°N, 95°-85°W of the United States, which covers parts of Louisiana, Arkansas, Kansas, Tennessee, Mississippi, and Alabama. The winter (January and February), spring (March-May), summer (June-and August), and fall (September-November) periods of 2008 are used for testing (approximately 16 000 images of the area of study are utilized for rainfall estimation in the testing mode). To train the SOM and also to obtain the T-R maps for each cluster, we use 1000 patches (as training data) that are randomly selected from 1 month before the respective testing month (e.g., for computing rainfall for the month of July 1000 patch samples are selected from the month of June to update and train the SOM and T-R maps). Note that because the training samples are chosen randomly from 1 month before the respective testing month, the month of December 2007 is not considered as part of the 2008 winter season testing (January and February). The IR data are obtained from GOES-12 (channel 4) with 30min interval images covering the entire area of study. It also has a nominal spatial resolution of $4 \text{ km} \times 4 \text{ km}$. The Next-Generation Weather Radar (NEXRAD) stage IV precipitation products, which are available on an hourly and also a daily basis at spatial resolutions of $4 \text{ km} \times 4 \text{ km}$, are used for training and validation (Lin and Mitchell 2005).

Figure 5 shows an example of the hourly precipitation estimate (1 out of 16 000 estimates) at 1300 UTC 4 February 2008 (the precipitation estimates are typically derived every 30 min; however, for validating the results against NEXRAD stage IV, we accumulate them in hourly estimates). The mCCS-WMS algorithm estimate is shown in Fig. 5a. The corresponding values of the implemented PERSIANN-CCS (hereafter mCCS) and NEXRAD stage IV data are shown in Figs. 5b,c, respectively. By comparing the two algorithms to NEXRAD's stage IV, it is clearly seen that the mCCS-WMS algorithm provides better rain area detection than the mCCS.

A set of four commonly used verification metrics, which includes the probability of detection (POD), the false-alarm ratio (FAR), equitable threat score (ETS), and area bias ratio (Ebert et al. 2007), are utilized to evaluate the algorithms against the NEXRAD stage IV product at rainfall thresholds of 0.01, 0.1, 1, 2, 5, 15, and 25 mm. The area bias is the ratio of the estimated-to-observed rain areas. An area bias value of 1 indicates that the estimation and observation have identical area coverage.

Figure 6 shows the FAR and POD verification metrics of the daily estimates for the mCCS, implemented



FIG. 5. Estimated hourly rainfall estimates ending at 1300 UTC at the threshold of 1 mm h^{-1} on 4 Feb 2008: (a) PERSIANN-CCS WMS, (b) PERSIANN-CCS, and (c) NEXRAD stage IV.

PERSIANN-CCS using the MM technique with an exponential curve fitting (mCCS-MM), implemented PERSIANN-CCS using MM and SCF methods (mCCS-MMSCF), and mCCS-WMS methods in all seasons of 2008 against the NEXRAD stage IV product at different rainfall thresholds. Except at low rainfall thresholds, the mCCS-MM method has high FAR compared to other algorithms in the seasons. In the spring and summer, the mCCS-WMS, mCCS-MMSCF, and mCCS methods have almost the same FAR performance. In the winter (Fig. 6a), the FAR resulting from the mCCS-WMS method is almost the lowest at large rainfall thresholds (about 10% less than that of the mCCS). Although the mCCS has less FAR in the fall than that of the mCCS-WMS and mCCS-SCFMM, the POD of the mCCS is significantly low at all threshold levels. Figure 6 also shows that the mCCS-WMS and mCCS-MMSCF methods have larger PODs than that of the mCCS almost at



FIG. 6. FAR and POD validation results for 2008 (daily estimates): (a),(b) winter, (c),(d) spring, (e),(f) summer, and (g),(h) fall.

all threshold levels in all seasons. In the winter, the enhanced algorithm (mCCS-WMS) has around 10%–40% POD improvement compared to the mCCS at medium and high rainfall threshold levels. In the spring, the POD of the mCCS-WMS is slightly more at medium rainfall thresholds, and almost similar at other threshold levels. In the summer, the POD percentage improvement obtained by the mCCS-WMS method is around 10% compared to the mCCS at most of the rainfall thresholds. And, in the fall, the mCCS-WMS and mCCS-MMSCF have around 30% POD improvement compared to the mCCS at medium and high rainfall thresholds. Furthermore, although the mCCS-MM method provides larger POD almost in all seasons, its high FAR provides poor performance in total.

Based on Fig. 6, it can be inferred that the large FAR difference between the mCCS-MM and other algorithms shows that in many cases, the exponential curve fitting does not match to the merged T-R resulting from applying the MM technique. In fact, it creates more fitting error [e.g., we see in Fig. 4 that the green line representing the exponential curve fitting does not match to

the blue samples (*T*–*R*-merged sample) and provides more fitting error]. Therefore, it can be inferred that in most cases, the polynomial curve fitting is selected by the SCF method in the mCCS-WMS and mCCS-MMSCF algorithms, and the SCF approach improves precipitation estimation when applied to the MM *T*–*R* samples. In addition, because the mCCS-WMS, mCCS-MMSCF, and mCCS-MM algorithms use the MM method and provide better POD performance compared to the traditional mCCS at medium and high rainfall thresholds, it can be inferred that the MM technique can be an effective method to improve the POD performance.

Figure 7 shows the ETS and area bias performance of the algorithms for the daily estimates for all seasons of 2008. By comparing the mCCS-WMS (and also mCCS-MMSCF) to the mCCS method, we see more improvement in the winter and the fall. As Fig. 7a shows, the mCCS-WMS has about 2%-20% greater ETS performance than that of the mCCS in the winter. The ETS of the mCCS-MMSCF algorithm also outperforms that of the mCCS in most of the rainfall range with approximately 16% improvement at some rainfall thresholds. Furthermore, at medium and large rainfall thresholds, the ETS of the mCCS-WMS is greater than that of the mCCS-MMSCF by approximately 4%. Figure 7b depicts that, at low rainfall thresholds, all algorithms, except the mCCS-MM, have an area bias values higher than 1. However, at medium and large thresholds, the area bias of the mCCS-MM increases abruptly due to increasing the FAR. Furthermore, for these thresholds, the mCCS underestimates and other algorithms overestimate precipitation. In addition, the area bias of the mCCS-WMS is almost less than those obtained from other methods at medium and larger threshold levels. In the spring (Figs. 7c,d), except for the mCCS-MM, other algorithms almost have the same ETS and also the same area bias performance. In the summer (Figs. 7e,f), at low rainfall thresholds, nearly all algorithms have the same ETS performance. However, at medium and high thresholds, the mCCS-WMS provides larger ETS than others with a maximum of 5% more than that of the mCCS. In addition, all algorithms have an area bias higher than 1, where the mCCS has the less area bias at all threshold levels. In the fall (Figs. 7g,h), the mCCS-WMS algorithm has superior ETS and also less area bias than others. By comparing the mCCS-WMS with the mCCS, the ETS percentage improvement is approximately 10% at medium and high rainfall thresholds in the fall. Furthermore, the mCCS-WMS has somewhat greater ETS performance than that of the mCCS-MMSCF at some threshold levels (approximately 3% at a threshold of 5 mm day⁻¹) in the fall.

Figure 8 depicts the ETS and area bias performance of the algorithms for the hourly estimates for all seasons of 2008. It depicts that both the mCCS-WMS and mCCS-MMSCF have greater ETS performance in the winter and the fall than that of the mCCS for the hourly estimates. The percentage of ETS improvement is approximately 10% in the winter and 8% in the fall at some threshold levels. The area bias is also improved by these algorithms (mCCS-WMS and mCCS-MMSCF) at most threshold levels in the winter and fall. However, by comparing the mCCS-WMS and mCCS-MMSCF to the mCCS in the spring and the summer, it is seen that no significant improvement is gained by the mCCS-WMS and mCCS-MMSCF. Note that both the mCCS-WMS and mCCS-MMSCF have almost the same hourly ETS performance. However, the mCCS-WMS has better area bias overall.

Based on Figs. 6-8, it can be inferred that most of the ETS improvement is due to extreme improvement in POD at medium and large rainfall thresholds for the mCCS-WMS, mCCS-MMSCF, and mCCS-MM algorithms. Furthermore, because the traditional mCCS has less ETS than that of the mCCS-WMS and mCCS-MMSCF algorithms at all threshold levels in the daily and hourly estimates for the winter and the fall seasons, it can be concluded that applying the MM and SCF is more effective in these seasons in order to filter out the improper patches in the clusters and also decrease the fitting error (MSE). Figures 7 and 8 also show that the mCCS-WMS provides better ETS performance than that of the mCCS-MMSCF in the winter and fall seasons for the daily estimates and also almost less area bias in the daily and hourly estimates in all seasons. It can be concluded that the wavelet features along with feature selection not only provide effective and useful information from cloud-top temperature, but also decrease the dimensionality and complexity and most probably increase the processing speed.

4. Conclusions and summary

Two main improvements to the PERSIANN-CCS algorithm are incorporated to enhance precipitation estimation. The first one is the use of wavelet features and feature selection (WFS) methods to provide proper and effective features to group and classify cloud patches. The second one is the improvement of the T-R for each cluster (group) using median merging (MM) and selected curve-fitting (SCF) techniques. The wavelet features provide more information about the texture of cloud patches, specifically information of the different details of variation from cloud-top temperatures. In addition to incorporating the wavelet features, a feature selection technique is used



FIG. 7. As in Fig. 6, but for ETS and area bias (BIAS) validation results for 2008 daily estimates.

to effectively find the best similarity measures (features) to group the clouds and reduce redundant and/or irrelevant features. The MM method, which computes the median rain rate at each temperature of the patch T-R values, is also used to diminish the impact of the improper patches, which may result of imperfect classification and

features extraction, as well as a lack of enough information of the cloud or other factors. In addition, a selected curve-fitting (SCF) method is implemented to fit the T-Rsamples resulted from applying the MM method. The SCF technique chooses either a polynomial curve fitting or an exponential curve fitting based on their MSE values.



FIG. 8. As in Fig. 6, but for ETS and area bias (BIAS) validation results for 2008 hourly estimates.

Overall, the use of the WFS, MM, and SCF approaches can enhance the PERSIANN-CCS algorithm and improve precipitation estimation. The results show that the enhanced algorithm (mCCS-WMS), as a result of incorporating the above methods (WFS, MM, and SCF), provides approximately 20% ETS improvement for the daily (10% for hourly) estimate in the winter, 10% for the daily (8% for hourly)

estimates in the fall, and at most 5% for the daily estimates in the summer at some rainfall thresholds. In the spring, no significant improvement is obtained. The area bias is improved at almost all rainfall thresholds for the daily and hourly estimates in the winter and fall. However, the area bias in the summer is greater than that of the implemented PERSIANN-CCS algorithm.

VOLUME 29

Acknowledgments. This research was sponsored by the NASA Applied Sciences Program under Grant NNS06AA98B and the NOAA Office of Atmospheric Research via Grant NA07OAR4170517. We also thank Dr. S. Sorooshian and the PERSIANN group at UC Irvine for the PERSIANN-CCS data and the helpful discussions about their methodology. Dr. Valentine Anantharaj is also supported by the National Center for Computational Sciences, Oak Ridge National Laboratory.

REFERENCES

- Adler, R. F., G. J. Huffman, and P. R. Keehn, 1994: Global tropical rain estimates from microwave-adjusted geosynchronous IR data. *Remote Sens. Rev.*, **11**, 125–152.
- Anagnostou, E. N., 2004: Overview of overland satellite rainfall estimation for hydro-meteorological applications. *Surv. Geophys.*, 25, 511–537.
- Atlas, D., D. Rosenfeld, and D. B. Wolff, 1990: Climatologically tuned reflectivity-rain rate relations and links to area-time integrals. J. Appl. Meteor., 29, 1120–1135.
- Burrus, C. S., R. A. Gopinath, and H. Guo, 1997: Introduction to Wavelets and Wavelet Transforms: A Primer. Prentice Hall, 268 pp.
- Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of nearreal-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, 88, 47–64.
- Gonzalez, R. C., and R. E. Woods, 2007: *Digital Image Processing*. 3rd ed. Prentice Hall, 976 pp.
- 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.
- Hsu, K. L., X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. J. Appl. Meteor., 36, 1176–1190.

- —, T. Bellerby, and S. Sorooshian, 2009: LMODEL: A satellite precipitation Geostationary Operational Environmental Satellite methodology using cloud development modeling. Part II: Validation. J. Hydrometeor., 10, 1096–1108.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeor., 8, 38–55.
- ISDR, 2008: Early warning systems can save lives when cyclones strike. Press Release UNISDR 2008/05, 1 pp. [Available online at http://www.unisdr.org/files/5434_pr200806myanmarcy clonenargis.pdf.]
- 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.
- Kohonen, T., 1982: Self-organized formation of topologically correct features maps. *Biol. Cybern.*, 43, 59–69.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. Preprints, 19th Conf. on Hydrology, San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at http://ams.confex.com/ams/ pdfpapers/83847.pdf.]
- Mahrooghy, M., V. G. Anantharaj, N. H. Younan, and J. Anstoos, 2011: On the enhancement of infrared satellite precipitation estimates using genetic algorithm filter-based feature selection. Proc. 34th Int. Symp. on Remote Sensing of Environment, Sydney, NSW, Australia, ISRSE.
- 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.
- Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. Bull. Amer. Meteor. Soc., 79, 61–78.
- 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.