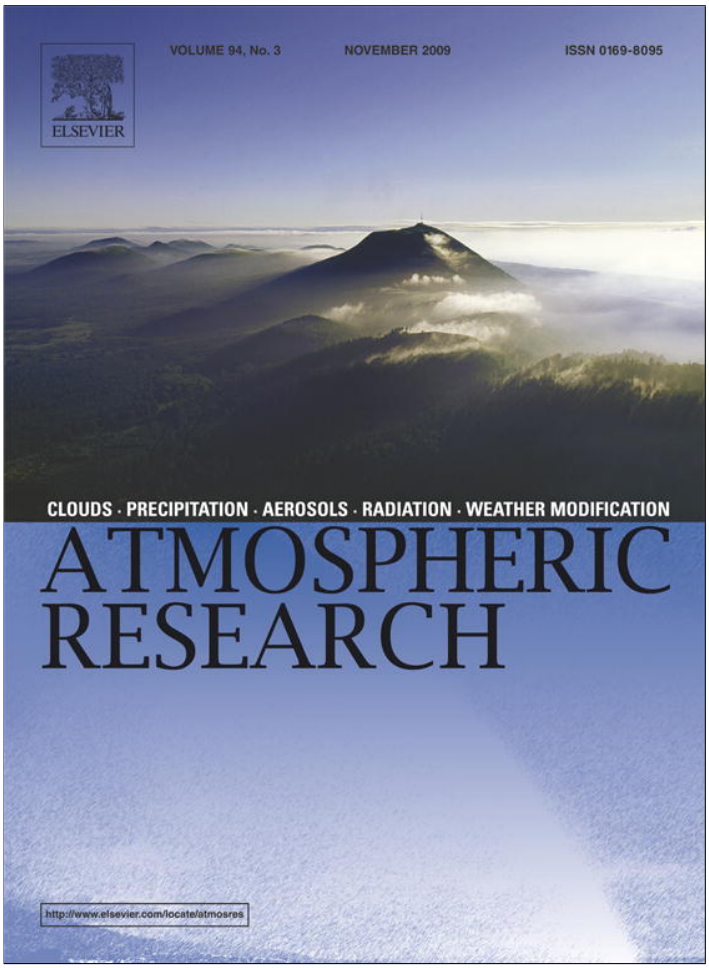
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Atmospheric Research 94 (2009) 373–388

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA

Emad Habib [a,*], Amy Henschke [a], Robert F. Adler [b]

[a] Department of Civil Engineering, University of Louisiana at Lafayette, P.O. Box 42991, Lafayette, LA, 70504, USA
[b] University of Maryland-College Park, MD and NASA/Goddard Space Flight Center, Laboratory for Atmospheres, Mail Code 613.1, Greenbelt, Maryland, 20771, USA

ARTICLE INFO

Article history:
Received 12 February 2009
Received in revised form 21 May 2009
Accepted 29 June 2009

Keywords:
Satellite rainfall
Statistical evaluation
Tropical storms
TRMM

ABSTRACT

This study focuses on the evaluation of 3-hourly 0.25°×0.25° satellite-based rainfall estimates produced by the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA). The evaluation is performed during six heavy rainfall events that were generated by tropical storms passing over Louisiana, United States. Two surface-based rainfall datasets from gauge and radar observations are used as a ground reference for evaluating the real-time (RT) version of the TMPA product and the post-real-time bias adjusted research version. The evaluation analysis is performed at the native temporal and spatial scales of the TMPA products, 3-hourly and 0.25°×0.25°. Several graphical and statistical techniques are applied to characterize the deviation of the TMPA estimates from the reference datasets. Both versions of the TMPA products track reasonably well the temporal evolution and fluctuations of surface rainfall during the analyzed storms with moderate to high correlation values of 0.5–0.8. The TMPA estimates reported reasonable levels of rainfall detection especially when light rainfall rates are excluded. On a storm scale, the TMPA products are characterized by varying degrees of bias which was mostly within ±25% and ±50% for the research and RT products, respectively. Analysis of the error distribution indicated that, on average, the TMPA products tend to overestimate small rain rates and underestimate large rain rates. Compared to the real-time estimates, the research product shows significant improvement in the overall and conditional bias, and in the correlation coefficients, with slight deterioration in the probability of detecting rainfall occurrences. A fair agreement in terms of reproducing the tail of the distribution of rain rates (i.e., probability of surface rainfall exceeding certain thresholds) was observed especially for the RT estimates. Despite the apparent differences with surface rainfall estimates, the results reported in this study highlight the TMPA potential as a valuable resource of high-resolution rainfall information over many areas in the world that lack capabilities for monitoring landfalling tropical storms.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Flooding associated with extreme rainfall events, such as hurricanes, can cause significant losses in human life and economic infrastructure in the United States (Rappaport 2000) and over many parts of the world (Negri et al. 2004). Accurate representation of spatial and temporal rainfall fields associated with such events is critical for the successful development of flood forecasting techniques and early warning systems. The continuous temporal availability and global coverage of satellites provide a valuable resource for rainfall monitoring especially over regions that lack adequate surface-based measuring techniques (e.g., rain gauges or radar). However, the use of satellite-based rainfall estimates in operational flood prediction and forecasting has not been fully achieved in the past (Ebert et al., 2007) due to the following reasons: (1) satellite estimates have been produced

* Corresponding author. Tel.: +1 337 482 6513.
E-mail address: habib@louisiana.edu (E. Habib).

0169-8095/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.atmosres.2009.06.015

Author's personal copy

at rather coarse resolutions in space (2.5°×2.5° or 1°×1°) and time (daily or monthly), which may not be adequate for hydrologic analysis; (2) information on the perceived uncertainty associated with satellite estimates is often lacking, which hinders the full utilization of such estimates in hydrologic and water resources management applications.

Recent years have witnessed significant developments in the field of satellite rainfall estimation. Of particular interest is the development of estimation algorithms that can generate high-resolution rainfall products by merging infrared (IR) and microwave (MW) satellite observations (e.g., Vicente, 1994; Turk and Miller, 2005; Kidd et al., 2003; Todd et al., 2001; Huffman et al., 2001; Huffman et al., 2007; Bellerby et al., 2000; Joyce et al., 2004; Sorooshian et al., 2000; Kuligowski 2002). The concept behind most of these algorithms is to use the more accurate (but infrequent) MW estimates to calibrate the more frequent (but indirect) IR estimates so that the strengths of individual sensors are maintained and the weaknesses are alleviated (Adler et al. 1994; Huffman et al. 2001). One recent algorithm, which is the focus of the current study, was developed at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and is based on the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA; Huffman et al. 2007). The TMPA dataset consists of fields of 3-hourly rainfall rates over a 0.25°×0.25° grid within the global latitude belt ranging between 50° north and 50° south. The algorithm derives its estimates primarily from MW information from several low-orbiting satellites and uses MW-calibrated IR data from geostationary satellites to fill MW coverage gaps. The TMPA estimates are available in the form of two products, a real-time version (3B42-RT) and a gauge-adjusted post-real-time research version (3B42).

Due to the potential for a wide array of applications of the TMPA and other similar products (Jiang et al. 2008; Curtis et al. 2007; Artan et al. 2007; Hong et al., 2006a; Hong et al. 2007; Su et al., 2008; Harris et al., 2007; Hossain and Anagnostou, 2004), efforts have been recently dedicated to investigate their performance accuracy and quantify the associated estimation uncertainties. Gottschalck et al. (2005) evaluated the TMPA 3B42-RT product over the contiguous United States (CONUS) at daily to seasonal time scales and found that, compared to other products, 3B42-RT estimates did not compare favorably in terms of bias and daily correlations. Tian et al. (2007) evaluated the 3B42 research version and another TRMM-based product against gauge and radar-based reference datasets. They found that TMPA 3B42 estimates have lower bias and random errors at seasonal and annual scales, but the performance is less satisfactory at shorter time scales (daily), especially in terms of successful rain detection. The correlation of the TMPA 3B42 product to ground reference data was better in summer than in winter. Tian and Peters-Lidard (2007) presented a comparative evaluation of the 3B42 product between land and inland water bodies and found increased false detections of rain over water bodies. Villarini and Krajewski (2007) performed an extensive evaluation of the TMPA 3B42 research version at its raw temporal 3-hourly resolution using a high-quality high-resolution rain gauge network in Oklahoma, United States. Their analysis reported that compared to the gauge dataset, the 3B42 product underestimates rainfall by ~10%, has a larger percentage of zeros, and lacks sensitivity to lower values of rainfall. The best performance of the 3B42 was found to be in the hot season. Their study gave insight on the temporal interpretation of the product indicating that its estimates should not be treated as instantaneous rain rates but rather as a 100-minute rainfall accumulation.

Following on these recent efforts, the current study provides an evaluation of the performance accuracy of the TMPA products during six heavy tropical-related rain storms in Louisiana, United States. The assessment will be made at the original resolution of the TMPA products (3-hourly) using a suite of graphical and statistical methods. Two surface-based rainfall datasets based on gauge and radar observations are used as a reference for evaluating the TMPA estimates. Continuous statistical measures (e.g., bias, standard deviation of differences, correlation coefficient) as well as categorical metrics (probability of detection) are used. The sensitivity of these performance metrics to various rain rate thresholds will be assessed. To gain further insight into the source of TMPA estimation errors, the overall bias and mean-square-error will be decomposed into sub-components that measure the contribution of different error sources. The added value from incorporating rain gauge information into the TMPA research version will be assessed by comparing the performance of the research product to the real-time TMPA product. In addition to the individual storm analyses, we will also examine the probability distribution of the TMPA error by pooling data from all six storms and studying the conditional statistics of the combined data. The study is intended to contribute to the increasingly sought efforts on evaluation of multi-satellite high-resolution rainfall products and should provide guidance to both algorithm developers as well as to the end users of such valuable datasets. The results also have implications for research efforts concerned with developing satellite rainfall error models (Hossain and Anagnostou, 2006). After the Introduction section, the analyzed storms and dataset sources are presented. This is followed by a section describing the methodology and statistical metrics used in the TMPA evaluation. Results will be presented including the marginal and conditional statistics and error decomposition. The paper closes with conclusions and final remarks.

2. Datasets and study area

2.1. Datasets

2.1.1. TRMM multi-satellite precipitation analysis (TMPA)

The TMPA precipitation estimates are based primarily on a combination of MW and IR estimates from multiple satellites. Microwave data, which are obtained from low earth orbit (LEO) satellites, are recognized as having a strong relationship to precipitation due to the physical nature of microwave detection; however, these data are only available on coarse temporal and spatial scales as a result of the infrequency of the individual satellite sampling. In contrast, the IR data from geosynchronous earth orbit (GEO) satellites produce estimates on fine time-space scales, but lack the strong relationship to rainfall associated with microwaves due to the relatively poor correlation between cloud-top temperature and precipitation (Arkin and Meisner, 1987). For this reason, microwave and IR data are combined to form the best possible

Author's personal copy

estimate for each location at each time step. The algorithm derives its estimates primarily from microwave information from several low-orbiting satellites and uses MW-calibrated IR data from geostationary satellites to fill gaps in MW coverage. The product is 3-hourly, 0.25°×0.25° gridded dataset covering the region bounded by the 50°N latitude and 50°S latitude lines. In order to populate the gridded dataset with combined data from both types of sensors, the microwave data are used first, in the areas where it is available, while the IR data are used to fill in the remaining areas. For a detailed description of the TMPA data and estimation algorithms, the reader is referred to Huffman et al. (2007).

The 3B42 TMPA dataset is available in two versions: a research-quality product (3B42) released 10-15 days after each month and a near-real-time product (3B42 RT), which is released approximately 9 h after real-time. The main difference between the two versions is the use of the rain gauge data for bias reduction, which are unavailable in real-time. The gauge data used in the TMPA algorithm are based on the Global Precipitation Climatology Project (GPCP) monthly rain gauge analysis (Rudolf, 1993). The gauge adjustment process involves aggregating both the gauge and the 3-hourly 3B42 estimates to a monthly scale and then applying the ratio of the 3B42/gauge monthly totals to each 3-hourly time step.

2.1.2. NEXRAD multisensor precipitation estimates (MPE)

Validation of the 3B42 and 3B42 RT data was performed using a radar-based multisensor precipitation product that is based on the NEXRAD system of the National Weather Service (NWS). This product is developed regionally at the NWS River Forecast Centers (RFCs) for operational hydrologic forecasting purposes. The area of the current study is fully within the boundaries of the Lower Mississippi River Forecast Center (LMRFC). The multisensor estimates are produced by combining data from several Weather Surveillance Radar-1988 Doppler (WSR-88D) radars with real-time surface rain gauge observations. Prior to August 2003, the multisensor estimates were available as a Stage III product of hourly rainfall accumulations over a grid of approximately 4 km×4 km (known as the Hydrologic Rainfall Analysis Project, HRAP, grid). As of 2003, the LMRFC began using a new algorithm called the multisensor precipitation estimator (MPE) as a replacement to the original Stage III processes (Fulton, 2002). In a final stage of processing, the regional Stage III or the MPE estimates are mosaicked from all the RFCs into a national product known as the Stage IV hourly precipitation dataset, which can be obtained from the National Center for Environmental Prediction (NCEP). For further details on the WSR-88D estimation and MPE processing algorithms, the reader is referred to literature resources such as Fulton et al. (1998), Seo et al. (1999), Breidenbach and Bradberry (2001) and Fulton (2002), among others.

2.1.3. HADS gauge dataset

The source of rain gauge data used in the evaluation of TMPA estimates is the Hydrometeorological Automated Data System (HADS). HADS is a real-time data acquisition and distribution system operated by the NWS Office of Hydrologic Development (OHD) from over 13,000 in-situ rain and stream gauges over the Continental U.S. (http://www.nws.noaa.gov/oh/hads/WhatIsHADS.html). Raw data are continuously obtained via geostationary satellites from Data Collection Platforms operated by the NWS as well as other national, state, and local offices. The HADS data used in this study are a reprocessed product obtained from the National Climatic Data Center (NCDC). The reprocessing effort includes consolidating sub-hourly accumulations into hourly precipitation values, distinguishing instances of no rain values that were initially categorized as "missing", and checking for erroneously high precipitation values and other quality-control checks (Kim et al., 2006). Since the LMRF uses original HADS gauge data (prior to reprocessing) for bias correction of MPE estimates, the gauge and MPE datasets used in this study cannot be considered independent of each other. However, the HADS gauge data are not used in developing the TMPA estimates.

2.2. Study area and storms

The area of focus in this study (Fig. 1) is the state of Louisiana (LA) in the lower Mississippi River Valley, which is frequently exposed to tropical cyclones due to its proximity to the Gulf of Mexico. Six rainfall events, of either tropical storm or hurricane intensity, which made landfall on or near the Louisiana coastline from 2002 to 2005 (Fig. 2), were selected for the study period, namely: Hurricane Lili, Tropical Storm Bill, Hurricane Ivan, Tropical Storm Matthew, Hurricane Katrina, and Hurricane Rita. The durations of the storms as they passed over the study area varied from 2 days (Hurricane Katrina) to 5 days (Tropical Storm Matthew).

The first storm, Hurricane Lili, made landfall near Intercoastal City, LA on October 3, 2002. Lili originated as a tropical depression in the Atlantic on September 21 and became a hurricane on September 30 while passing through the Caribbean. After moving across the Gulf of Mexico, Lili strengthened to a category 4 hurricane with wind speeds of 125 kt before being downgraded to a category 1 hurricane at landfall 13 h later. The rainfall totals amounted to 4–8 in. (100–200 mm) throughout southern Louisiana.

Tropical Storm Bill became a tropical cyclone at 1200 UTC on June 29, 2003 in the Gulf of Mexico. It made landfall while at its maximum intensity of a tropical storm with 50 kt winds on June 30. Bill came ashore near Cocodrie, LA and brought 5–10 in. (120–250 mm) of rain throughout the southern U.S. as it continued its path through Mississippi, Alabama, and Georgia as a tropical depression.

Hurricane Ivan was a long-lived storm that developed on September 2, 2004 as a tropical depression in the Atlantic. It later made landfall as a category 3 hurricane on September 16 in Gulf Shores, AL. After crossing through the southern U.S. and returning to the Atlantic, Ivan took a southern turn and passed over Florida as an extratropical low. By September 21, it began to gain strength over the Gulf waters and made landfall again in the U.S. along the southwestern Louisiana coast as a tropical storm on the 24th of September, 22 days after is original development. Ivan brought 3–7 in. (75–175 mm) of rain to much of the southern states as well as much of the east coast.

Also occurring in 2004 was Tropical Storm Matthew, which reached tropical depression conditions in the Gulf of Mexico on October 8. Matthew strengthened to a tropical storm 6 h later and made landfall on October 10 near Cocodrie, LA bringing 4–8 in. (100-200 mm) of rain to parts of Louisiana.

Author's personal copy

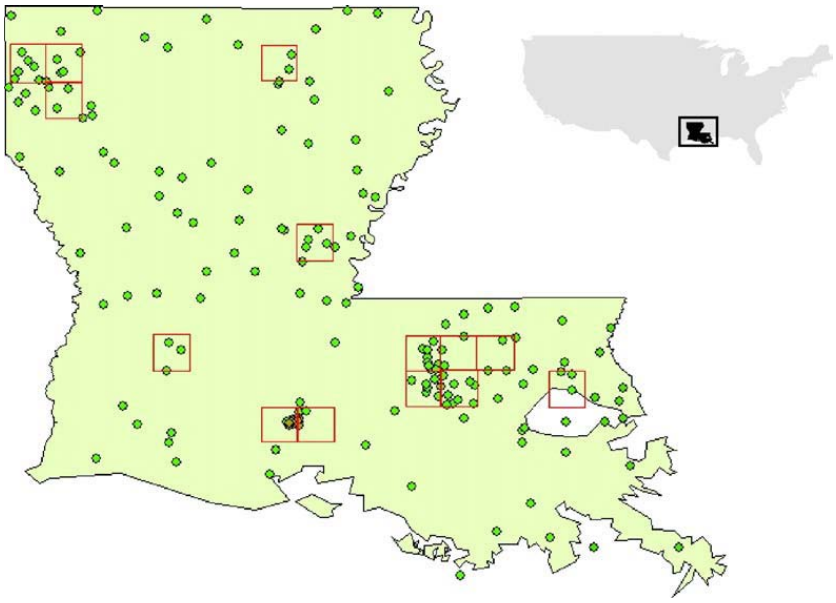


Fig. 1. Study site, Louisiana, US, with the TMPA 0.25°×0.25° pixels (squares) and the HADS rain gauges (filled circles) selected for evaluation of TMPA products.

The 2005 hurricane season marked the year of Hurricanes Katrina and Rita. Occurring less than one month apart, Katrina and Rita brought much rain, flooding, and destruction to the Louisiana coast. Katrina made landfall as a category 3 hurricane on August 29 near Buras, LA, bringing a total of 8–10 in. (200–250 mm) of rainfall. In 2005 Katrina was the costliest hurricane to hit the U.S. and one of the five deadliest.

By September 18, Rita had formed as a tropical depression in the Caribbean. It strengthened to hurricane status by September 20 and reached its peak intensity on September 22 as a category 5 hurricane. Total rainfall accumulations reached 5–9 in. (120-230 mm) throughout Mississippi, Louisiana, and Texas.

3. Methodology

The TMPA performance accuracy is assessed using two reference datasets: the MPE Stage IV product and the HADS gauge dataset. Both TMPA versions, 3B42 and 3B42 RT, are considered for evaluation. For each of the above-mentioned storms, between 3 and 6 TMPA pixels located near the track of the storm were chosen for the evaluation analysis (Fig. 1). The gauge and MPE observations were aggregated in time and space to match the scales of TMPA estimates (3-h and 0.25°). This was done by averaging rain rates from the gauges within each TMPA pixel spatially, over the area of the pixel, and temporally, over a window of 3 h. Similar aggregation was performed on the several 4×4 km^2 MPE pixels located within each TMPA pixel. The sample size for each two-way comparison, which is based upon paired datasets where either the reference value or the TMPA value is greater than zero, varies for each storm depending on the number of pixels analyzed and on the duration of the storm. On average, the sample size for each storm was in the range of 36-42 for Bill, 55-64 for Lili, 63-70 for Ivan, 147-175 for Matthew, 72-85 for Rita, and 63-74 for Katrina.

We acknowledge that both reference datasets (gauge and MPE) are not error-free and are subject to uncertainties that are rather difficult to characterize. Gauges are limited by their lack of areal representativeness and the effect of sub-pixel variability. Ideally, a large number of gauges should be available and distributed within a TMPA pixel area. For example, Villarini and Krajewski (2007) relied on 23 gauges located within a single TMPA pixel in Oklahoma. This rather ideal sampling situation is not available in the current study area. Therefore, in an attempt to alleviate the effect of gauge uncertainties, the pixels considered for our analysis were those containing a minimum of 3 rain gauges. Unlike the gauge data, MPE estimates are available at a high spatial resolution (4×4 km^2) and therefore do not suffer from spatial under-sampling over the scale of a single TMPA pixel. However, MPE products are subject to errors associated with rainfall estimation from radar observations. Recent studies have examined the accuracy of Stage IV MPE products (e.g., Westcott et al., 2008; Young and Brunsell, 2008; Wang et al., 2007; Habib et al., 2009) and reported varying degrees of both over- and under-estimation biases, and rather significant random errors at the hourly 4-km resolution. While it is not fully understood how MPE uncertainties scale up at the TMPA resolution (3-hourly and 0.25°), it is expected that the random component of the MPE error will be smoothed out significantly when a large number of MPE pixels (more than 50 in this study) are spatially averaged to produce an estimate over the TMPA pixel.

Several metrics and techniques have been proposed to evaluate and verify satellite rainfall estimates and forecasts (e.g., Gebremichael et al., 2005, Hong et al., 2006b, Ebert et al. 2007, Hossain and Anagnostou 2006, Hossain and Huffman 2008). In this study, we assess the performance of the two TMPA product versions using both graphical and statistical methods. Graphical comparisons include time series plots of rainfall rates for gauge, MPE, 3B42 and 3B42-RT datasets over the duration of each storm. Scatter plots are generated to visually inspect how TMPA estimates compare against gauge and MPE estimates. The probability distributions of TMPA rainfall estimates are compared to those of MPE and gauge by analyzing the probability of exceedance of each dataset. To quantify systematic and random differences between TMPA and the reference datasets, we applied a suite of statistical techniques that include continuous

Author's personal copy

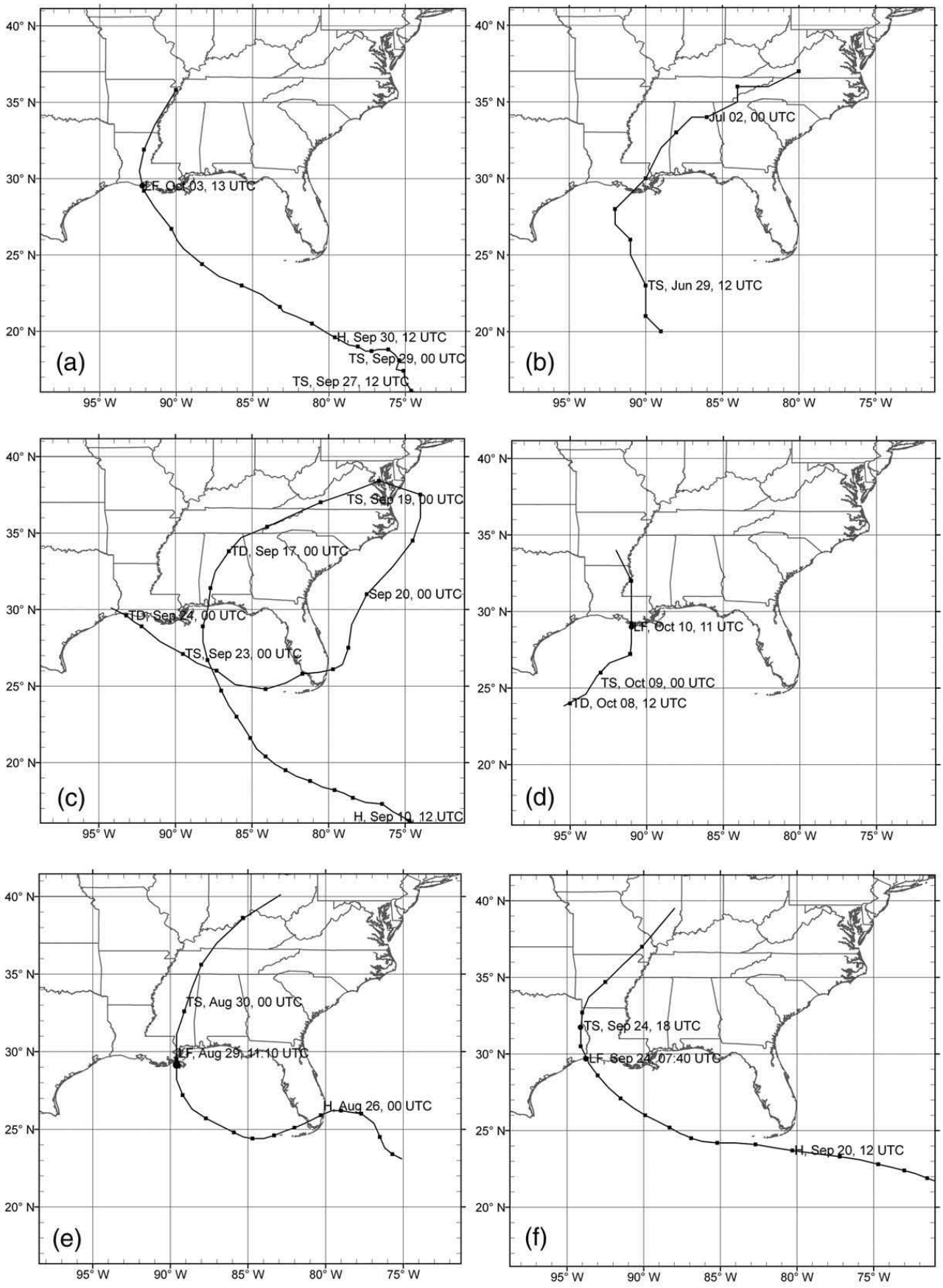


Fig. 2. Tracks of the six tropical storms used in the evaluation of the TMPA products (a) Hurricane Lili, (b) Tropical Storm Bill, (c) Hurricane Ivan, (d) Tropical Storm Matthew, (e) Hurricane Katrina, and (f) Hurricane Rita. The storm tracks were redrawn using information from the National Hurricane Center Tropical Prediction Center web page. The intensities of the storms are categorized as hurricane (H), tropical storm (TS), and tropical depression (TD), with the land fall location indicated by LF.

Author's personal copy

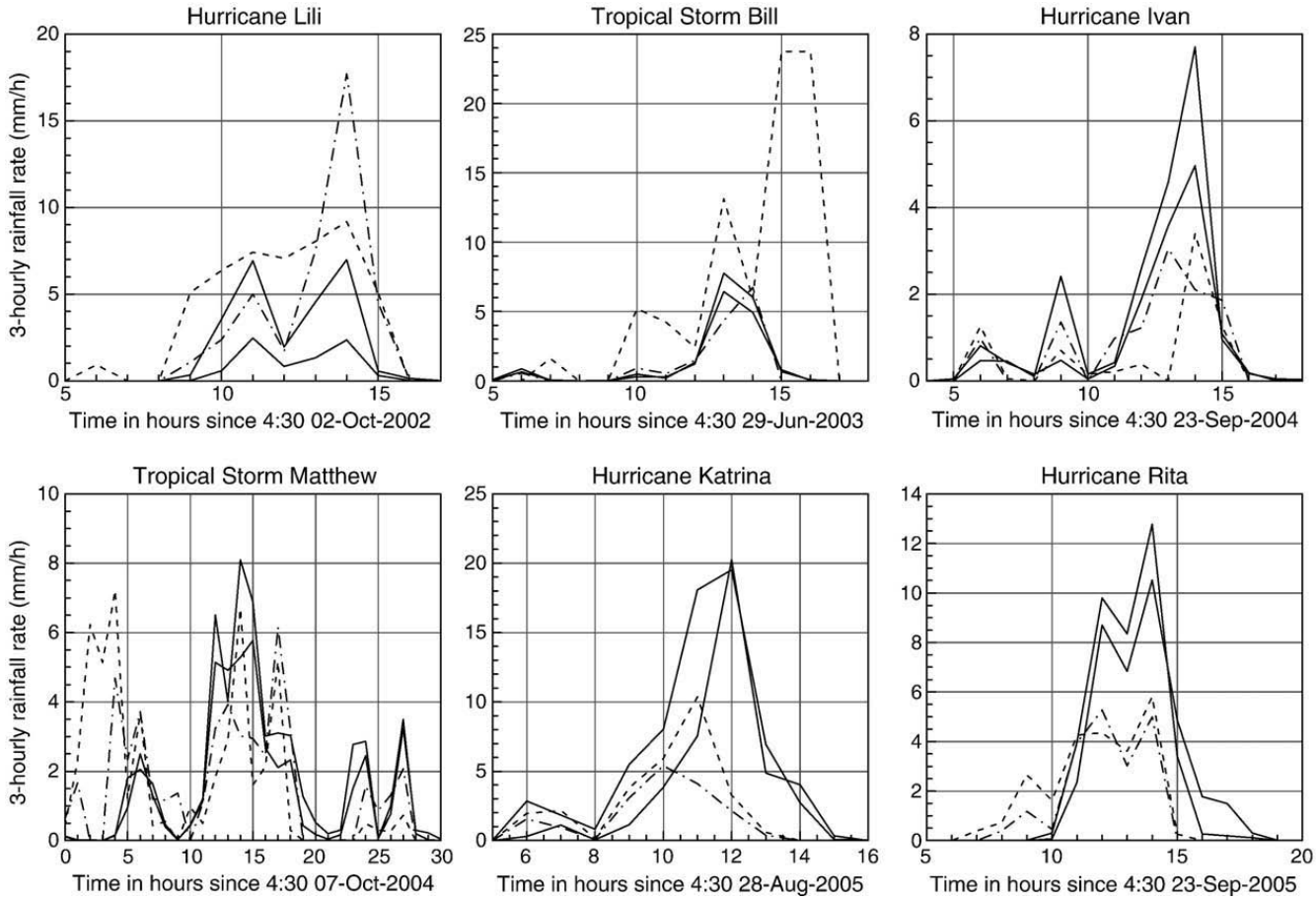


Fig. 3. Time series of 3-hourly rain rates of HADS gauges (thin solid line), MPE (thick solid line), TMPA-Research (dash-dotted line) and TMPA-RT (dashed line). The plotted MPE and HADS values represent 3-hourly rain rates averaged over the size of a certain 0.25°×0.25° TMPA pixel selected for every storm.

and categorical measures, unconditional and conditional metrics, decomposition of bias and mean squares of errors, and an analysis of TMPA error distributions.

3.1. Categorical statistics

The categorical statistics (Wilks 2006) used in the current analysis are the probability of detection (POD) and the probability of false detection (POFD). The POD represents the ratio of the number of correct identifications of rainfall to the total number of rainfall occurrences, as indicated by the reference, while the POFD represents the ratio of the number of false identifications of rainfall to the total number of non-rainfall occurrences, as indicated by the reference. Both POD and POFD range from 0 to 1, with 1 being a perfect POD and 0 being a perfect POFD. The POD and POFD are calculated both on a storm-by-storm basis, as well as on an overall combined-storm basis. The combined-storm POD and POFD are further broken down through conditioning on various rain thresholds of the reference data. The volumes of rainfall correctly and incorrectly detected by TMPA are also observed as an indication of the significance of either probability in the overall performance of the TMPA.

3.2. Continuous statistics

The continuous statistics used to quantify the differences between the TMPA and the references datasets are the mean difference and relative mean difference (a measure of the bias), and the relative standard deviation of differences (a measure of the random error):

$$\text{Mean difference} = (\overline{R}_S - \overline{R}_R) \tag{1}$$

$$\text{Relative mean difference} = \frac{(\overline{R}_S - \overline{R}_R)}{\overline{R}_R} \tag{2}$$

$$\text{Relative standard deviation of difference} = \frac{\sigma_{(R_S - R_R)}}{\overline{R}_R}, \tag{3}$$

where R_S is the rainfall value of the satellite-based TMPA dataset and R_R is the corresponding gauge or MPE reference value. The overbar and σ symbols denote the mean and standard deviation, respectively, estimated over the available sample size. We also computed the Pearson's correlation coefficient as a measure of linear association. These statistics will be presented on a storm-by-storm basis and for all storms combined.

3.3. Error decomposition

The bias and random error measures defined in Eqs. (1) and (2) can be further decomposed as follows:

(a) Bias decomposition.

The bias calculated using Eq. (1) is based on aggregation of differences in rainfall depth over the sample of each storm

Author's personal copy

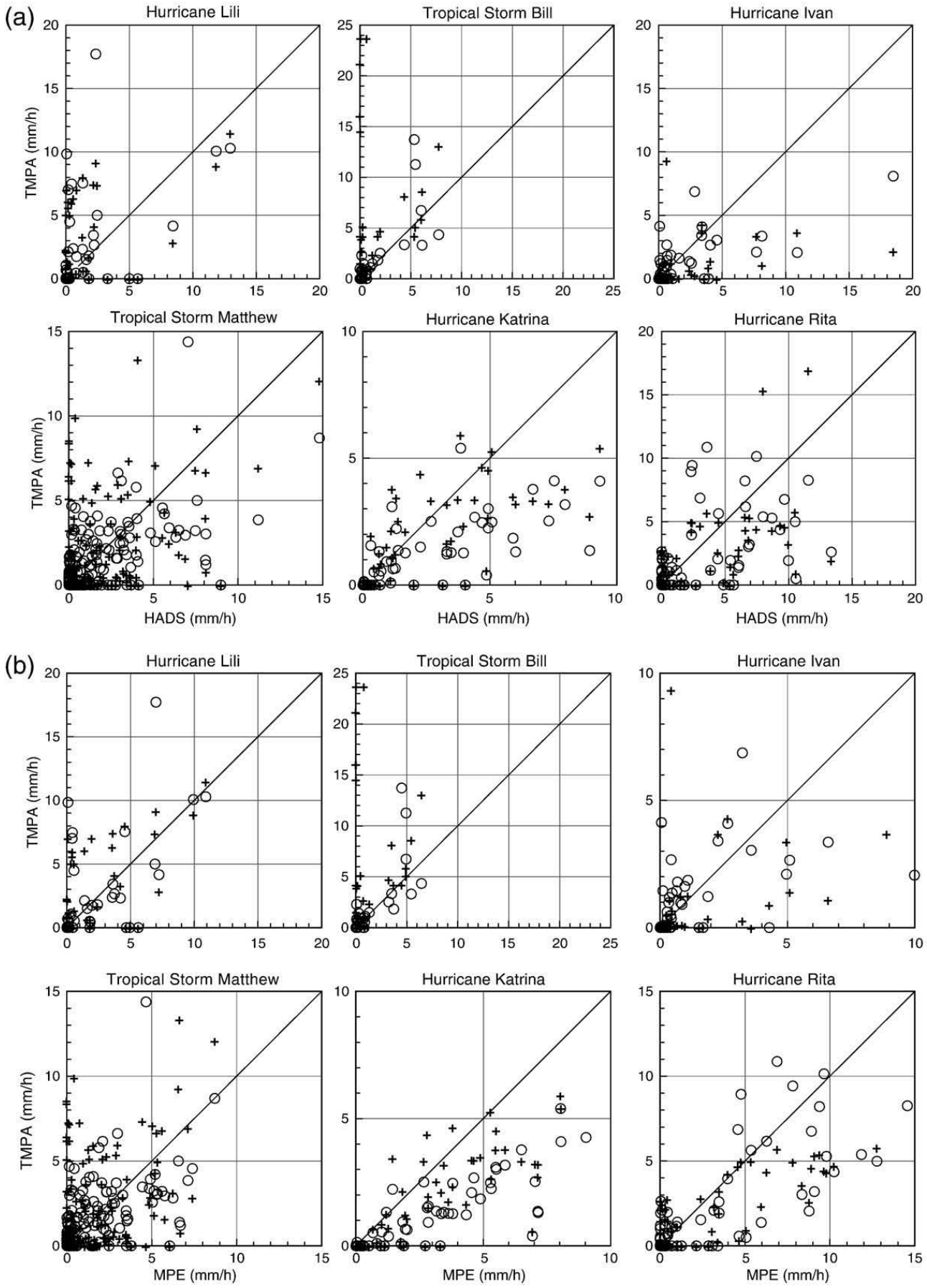


Fig. 4. (a) Scatter plots of 3-hourly 0.25°×0.25° rain rates of the TMPA research product (circle symbol) and the TMPA-RT product (cross symbol) against the corresponding average measurements from HADS gauges within all analyzed TMPA pixels. (b) Same as Fig. 4(a) but using MPE as the reference.

Author's personal copy

and does not provide information on the source of such differences. Therefore, it is desirable to break down the total bias into three components consisting of the bias during successful detections (hits), bias due to rainfall misses, and bias due to false detections:

$$\text{Hit bias (HB)} = \sum(R_S(R_S > 0 \& R_R > 0) - R_R(R_S > 0 \& R_R > 0)) \quad (4)$$

$$\text{Missed rain bias (MB)} = \sum R_R(R_S = 0 \& R_R > 0) \quad (5)$$

$$\text{False rain bias (FB)} = \sum R_S(R_S > 0 \& R_R = 0). \quad (6)$$

Such decomposition can distinguish among the three possible bias sources, whose values can be cancelled by opposite signs if only the total bias is evaluated. The summation of these three components adds up to the total bias ($TB = \sum(R_S - R_R)$). The proportion of total bias that is attributed to each bias source can be described by the ratio of the particular bias component to the total bias (e.g., HB/TB, MB/TB, and FB/TB), with the three ratios adding up to 1.

(b) Decomposition into systematic and random components.

Besides using the mean bias (Eq. (1)) and standard deviation (Eq. (2)), the differences between TMPA estimates and reference values can be alternatively (and sometimes more commonly) described by the mean-square difference (MSD), which provides a measure of the average difference. According to Willmott (1981), MSD can be decomposed into a systematic component (MSD_s) and a random component (MSD_u):

$$MSD = MSD_s + MSD_u \quad (7)$$

$$\frac{\sum_{i=1}^{n}(R_{S,i} - R_{R,i})^2}{n} = \frac{\sum_{i=1}^{n}(\hat{R}_{S,i} - R_{R,i})^2}{n} + \frac{\sum_{i=1}^{n}(R_{S,i} - \hat{R}_{S,i})^2}{n} \quad (8)$$

Where $\hat{R}_S$ is defined by the least-square linear regression relationship $\hat{R}_S = a + bR_R$, with a and b as the intercept and slope. The advantage of this decomposition is that the proportion of MSD that is attributed to systematic errors can be described by (MSD_s/MSD) while the random contribution can be described by (MSD_u/MSD) or ($1 - MSD_s$/MSD).

3.4. Error Distribution

This analysis examines the distribution of the TMPA estimation error by constructing histograms of the differences ($R_S - R_R$). To investigate whether the TMPA error depends on rainfall magnitudes, the error distribution and its statistical properties will be assessed conditionally for various ranges of the reference rainfall. To maintain a reasonable sample size for constructing error distribution, this analysis was conducted based on combining all storms into one large sample.

4. Results and discussions

Time series of rainfall rates during the lifespan of each storm are plotted for selected pixels (Fig. 3) where it is noticed that TMPA estimates track reasonably well the overall pattern of MPE and HADS gauges for most of the time. However, significant deviations are noticed in the magnitudes of 3-hourly rainfall rates (e.g., the later part of storm Bill and the earlier part of storm Matthew). This behavior was variable across different pixels within the same storm (not shown). It is likely that spurious or incorrect data were present during the TMPA-RT processing of Bill (indicated by two observations of very high values, which were probably replaced when data were re-processed after the storm). Therefore, these two time steps will be excluded in the calculation of the statistics considered in this study. It is also noticed that MPE and HADS show reasonable agreement to each other for most of the storms, which may be attributed partly to the fact that HADS data were used in the development MPE products. Point-by-point comparisons of TMPA estimates versus MPE and HADS are shown in the form of scatter plots (Fig. 4). Each point represents the average rainfall rate over a 3-h duration and a spatial domain of 0.25°. These graphs are generated by pooling data pairs from all pixels considered in each storm into one scatter plot. Large scatter exists between the TMPA products and the corresponding reference, whether it is MPE or HADS. Similar degrees of scatter were reported in Villarini and Krajewski (2007) in their analysis of the TMPA research version over Oklahoma. When aggregated to daily scale, the standard deviation of differences was reduced to 76% and 90% for the research and the RT products, respectively. It should be noted that the significant scatter observed at the native TMPA temporal resolution is partly attributed to the fact that the TMPA do not actually represent 3-hourly average rain rates (Villarini and Krajewski, 2007), but rather a sort of quasi-instantaneous rain rate at some point during the 3-h period (at the variable time of satellite overpass).

The marginal distribution of the TMPA products compared to the reference datasets is examined by calculating the probability of exceedance (Fig. 5) of each dataset. The probability of exceedance (sometimes referred to survival function) is defined and calculated as the probability that an estimate exceeds a certain threshold. For the lower range of

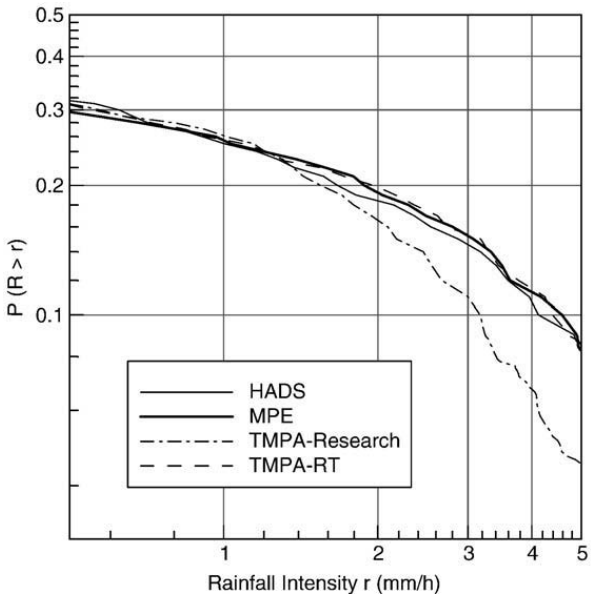


Fig. 5. Probability of exceedance plots for the TMPA products compared with those of the two reference datasets, MPE and HADS gauges.

Author's personal copy

the distributions (rain rate threshold of 1.5 mm/h and lower) both products have exceedance probabilities that are close to that of the reference. For larger thresholds and at the extreme tail of the distribution, the RT product still follows the reference but the research version has smaller frequencies of high rain rates compared to the reference. It is possible that, although, in general, the bias removal technique using monthly gauges reduces the overall bias, in the case of some heavy rain situations, it may lead to reducing high rain rates. If this effect is proven persistent in other validation cases, the procedure may need to be revised to fix this undesirable behavior.

4.1. Analysis of bias

The overall total bias between the TMPA products and the reference rain is quantified by calculating the difference between their arithmetic means, expressed in absolute and relative units (Eq. (1) and (2)); Fig. 6. For the research product, the relative bias showed signs of both underestimation and overestimation, but was mostly bounded by ±25% based on using either MPE or HADS gauges as a reference. The RT product shows higher bias levels than the research product for all storms (except Katrina).

As described earlier, the total bias can be decomposed into three components: hit bias (HB), missed-rain bias (MB), and false-rain bias (FB); Fig. (7). Comparison with HADS gauges showed similar behavior and is not presented. It should be noted that these volumes are combined totals from the individual pixels analyzed for every storm. It is clear that HB is the dominant component that contributes the most to the overall bias during all storms. The HB is significantly reduced in TMPA-research compared to TMPA-RT (except in Katrina). The next significant component is the MB which shows similar magnitudes in both versions with slightly higher values for the TMPA-research product. This indicates that

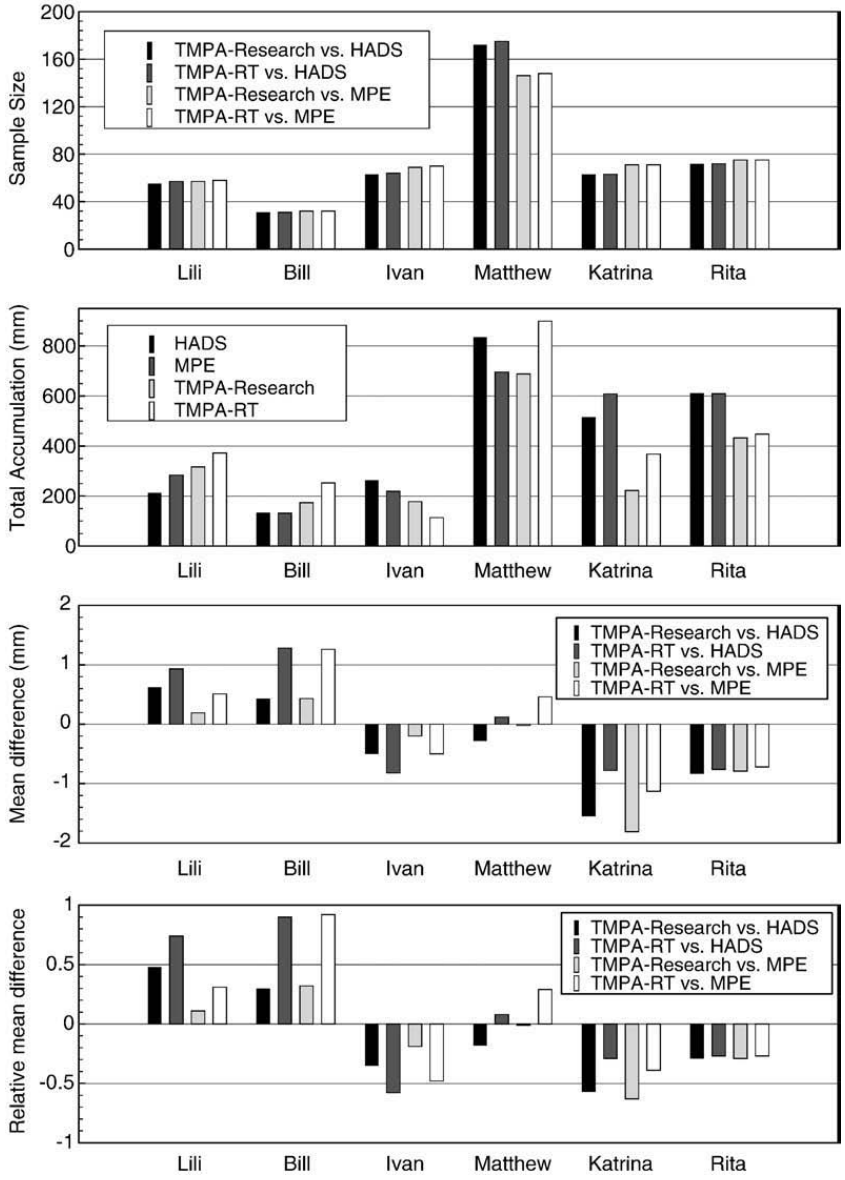


Fig. 6. Summary statistics of the six analyzed storms (from top to bottom): size of joint TMPA-reference samples, total rainfall accumulation for each dataset, TMPA bias, and TMPA relative bias. The relative bias is calculated by normalizing with the mean of the reference. Statistics are based on combined data from all individual pixels analyzed for every storm.

Author's personal copy

during the bias-adjustment process of the research product, some rainy intervals were set to zero-rain rates. The average value of missed rainfall was ~0.5 mm/h with about 85% of the values less than 1 mm/h, 10% between 1 and 3 mm/h and ~5% larger than 3 mm/h but not exceeding 6 mm/h. The contribution of the false-rain bias (FB) is negligible for the research product, indicating the successful removal of false alarms in the TMPA research algorithm. The FB was also minimal for MPA-RT, except during Matthew, where the mean of falsely detected rain rate was 0.26 mm/h and with few rates ranging between 2 to 5 mm/h.

4.2. Analysis of TMPA detection

To further characterize the detection limitations of TMPA products we examine two probabilities: probability of detection, POD, and probability of false detection, POFD. The use of the HADS gauges as a reference in detection assessment is problematic due to the limited number of gauges within each pixel; therefore, only MPE data were used as a reference in calculating POD and POFD. The two probabilities are assessed for both TMPA versions on an individual storm basis and for all storms combined (Fig. 8). Consider first the POD results. Conditioned on MPE larger than zero, the two TMPA products show POD levels in the range of 0.4 to 0.7. When pooled together, the six storms have a POD of slightly lower than 0.6 for both products. It is interesting to note that POD values of TMPA-RT are higher than those of the TMPA-research product, which in turn results in larger amounts of missed rain for the research product (see results on bias components in Fig. 7). This reiterates the earlier observation about setting what seem to be actual rainy intervals into zero-rain rates when going from the TMPA-RT to the research product.

Results on the POFD (Fig. 8) show that the problem of false detection of rainfall by TMPA products is rather minimal, especially for the research product, which reported POFD values of zero for all storms except Matthew. Even during Matthew where POFD was ~15%, it appears that most of these false detections were associated with very light rain as reflected in the small amount of falsely detected rain (~1.5 % of the total storm rainfall depth). The TMPA-RT product showed low POFD values (zero for three storms and 3–6% for two storms) with minimal false rainfall depths except for storm Matthew (~17% of the total storm rainfall depth). The differences in POFD statistics between the two products indicate a successful removal of false rainfall detections from the real-time product. To analyze the detection sensitivity of TMPA estimates, the POD and PODF were recalculated by conditioning on various rain rate thresholds after pooling all storms together (Fig. 9). The results indicate that the rather low POD values in both products are caused by lack of detection of small rainfall intensities in the MPE reference dataset (POD increases to 0.8 when MPE exceeds 0.2 mm/h). Similarly, the POFD values decline rapidly, especially for the research product, with the increase of the threshold considered as a false detection.

4.3. Analysis of Agreement and Disagreement Statistics

The agreement between TMPA and the reference rainfall is assessed using the Pearson's correlation coefficient (Fig. 10). Correlation results based on using either MPE or HADS gauges as a reference are quite similar, except for Rita, which shows correlations to MPE are higher by about 0.2. Overall, the TMPA research product shows moderate to high correlation values for most storms (0.5 to 0.8). Lower correlation values are reported with the RT product in Ivan and Matthew. The correlation values are higher than those reported in previous studies (Villarini and Krajewski, 2007) and reflect a reasonable skill for the TMPA products in reproducing the temporal fluctuations of the reference rainfall during the analyzed tropical-related storms. However, it should be noted that the

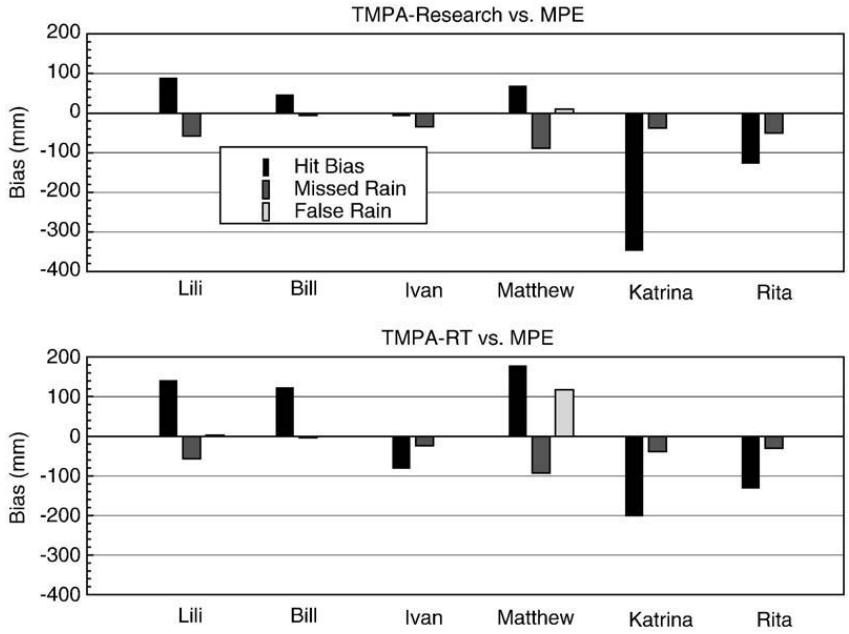


Fig. 7. Decomposition of the total bias of the TMPA products (using MPE as a reference): bias during successful hits, bias due to missed rain, and bias due to false detections. Bias values are based on combined data from individual pixels.

Author's personal copy

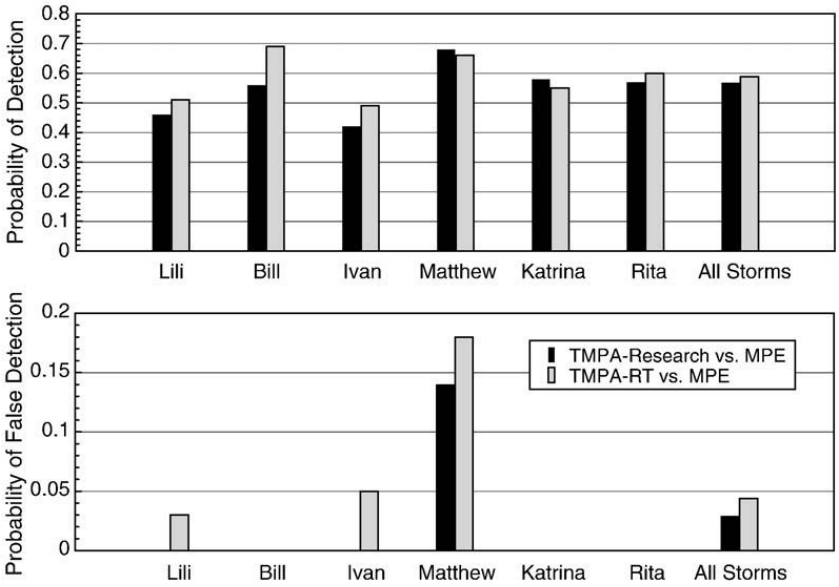


Fig. 8. Probabilities of detection (successful, top panel, and false, lower panel) of the two TMPA products using MPE as a reference.

Pearson's correlation coefficient can be significantly affected by the presence of extreme rain rates, which can be dominant in the case of tropical storms.

Now we turn to statistics that focus on assessing disagreement between TMPA and reference rainfall using the standard deviation of differences expressed in absolute units (mm/h) and also as a ratio relative to mean rainfall rate (Fig. 10). For most storms, the range of this statistic was 1 to 2 times of the average rain rate during each storm. The research product has noticeably lower values during two storms (Ivan and Matthew), with the other storms showing comparable results to the real-time product. We also assessed the average difference using the mean of squares of differences (MSD) and its two components (systematic, MSD_s, and unsystematic or random, MSD_u) as defined in Eqs. (7) and (8). The decomposition is performed on the TMPA-MPE sample but includes only successful hit pairs in the sample. The resulting values of the two components are presented in relative units (MSD_s/MSD and MSD_u/MSD); Table 1. In comparing the two components for either the TMPA-research or the RT product, it appears that the random component is clearly dominant in three storms (Lili, Bill and Matthew), the systematic component is more dominant in two other storms (Ivan and Katrina), and the two components are about the same for Rita. In comparing the two TMPA products, the research product has a smaller (larger) systematic (random) component during storms Lili and Bill and a larger (smaller) systematic (random) component during Matthew and Katrina, with the two other storms (Ivan and Rita) showing similar values for both MSD_u and MSD_s.

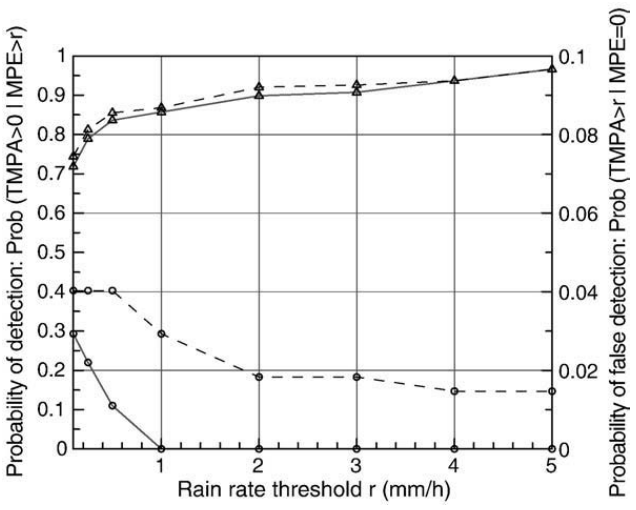


Fig. 9. Conditional probability of detection (lines with triangles; left *y*-axis) and probability of false detection (lines with circles; right *y*-axis) for the TMPA-Research product (solid line) and the TMPA-RT product (dashed line).

4.4. Analysis of TMPA error distribution

Now we turn to the distribution of the differences between each TMPA product and the reference dataset (for space limitations we only present results based on using MPE as the reference). Fig. 11 shows frequency histograms (plotted using lines for clarity) of TMPA-reference differences calculated for various ranges of the difference. The histogram is symmetrical around the central bin, indicating that most differences are within the range of -0.5 to 0.5 mm/h. The distribution of the error of the research product has thinner tails, especially on the positive side (>2 mm/h), which indicates that several of the large errors were reduced to smaller values as evident by the higher frequency at the central bin. To examine whether the error distribution is dependent on the rain rate magnitude, the differences were plotted in the form of scatter points as a function of the MPE rain rate (Fig. 12). The figure clearly shows the mean and variance of the TMPA-MPE difference vary with the MPE rain rate. To quantify such dependence, the full sample was divided into sub-samples based on different ranges of the MPE rain rate and the conditional mean, standard deviation and two quantiles (0.1 and 0.9) were calculated and plotted (Fig. 13). Despite the relatively small sample size of some of the categories (in the order of 100 points), some useful remarks can be made. It is clear that the error statistics

Author's personal copy

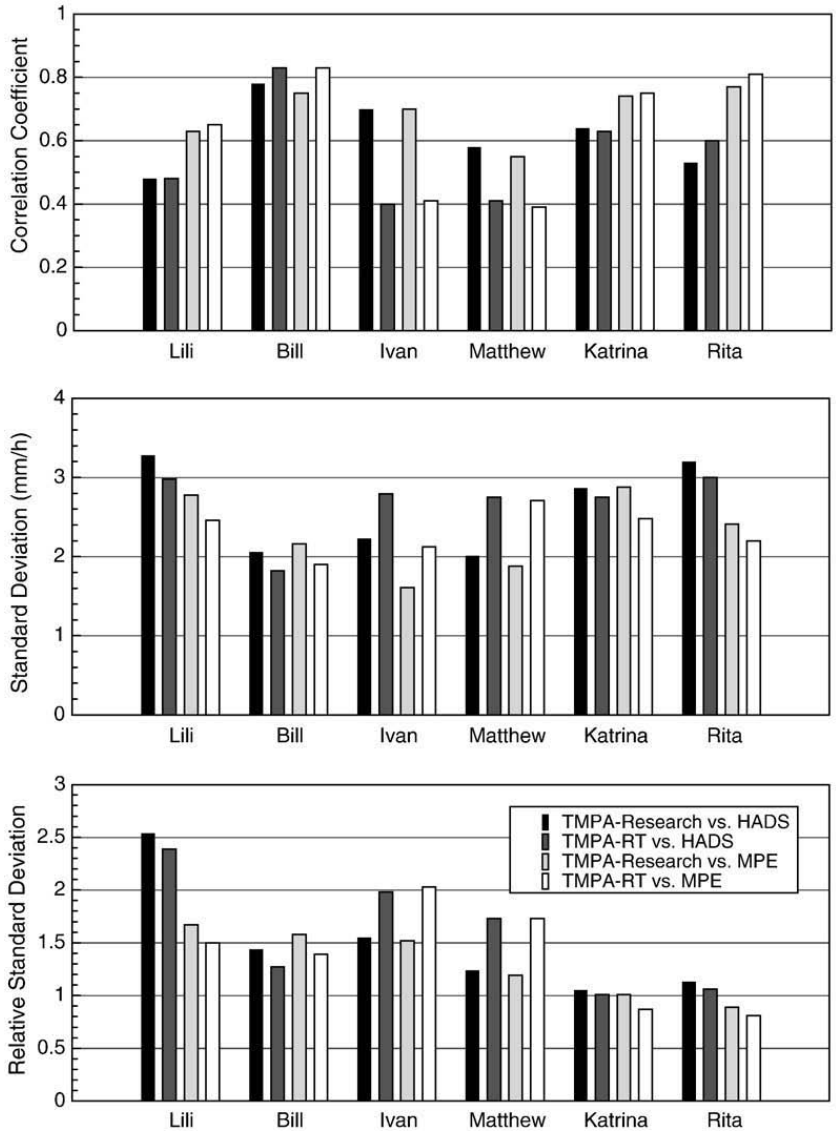


Fig. 10. Correlation coefficient (top panel) and standard deviation of differences (lower two panels) between TMPA products and the reference datasets. The relative standard deviation is calculated by normalizing with the mean of the reference.

depend on the magnitude of the rain rate. The error tends to be positive for the lower range of MPE rates and becomes negative for higher MPE rates, which indicates that, on average, both products tend to overestimate small rain rates and underestimate large rain rates. With the exception of the last category (MPE>5 mm/h), the degree of underestimation or overestimation is higher for the real-time product than for the research product. It is also observed that the RT errors have a wider distribution for all ranges of the MPE rain rate, especially on the positive side of the error as evident in the calculated 0.90 quantiles.

5. Summary and conclusions

This paper evaluated a high-resolution satellite rainfall product (TMPA-3B42) during six tropical-related rainfall events with the objective of providing the user community and the algorithm developers with some insight on its

Table 1
Decomposition of the mean square difference (MSD) between TMPA and MPE estimates into systematic (MSD_s) and unsystematic (MSD_u) components.

	Lili		Bill		Ivan		Matthew		Katrina		Rita	
	3B42	3B42 RT	3B42	3B42 RT	3B42	3B42 RT	3B42	3B42 RT	3B42	3B42 RT	3B42	3B42 RT
MSD	13.58	9.53	8.27	7.38	5.36	9.14	4.56	8.55	19.24	12.74	10.19	8.48
MSD_s/MSD	0.13	0.37	0.10	0.49	0.63	0.66	0.32	0.17	0.95	0.70	0.53	0.45
MSD_u/MSD	0.87	0.63	0.90	0.51	0.37	0.34	0.68	0.83	0.05	0.30	0.47	0.55

Results are based on TMPA–MPE pairs classified as successful hits.

Author's personal copy

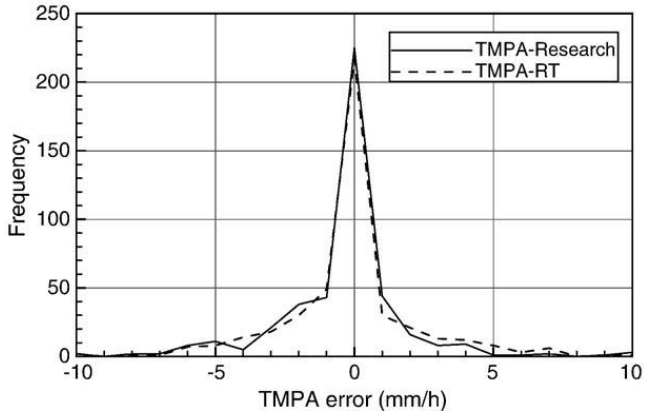


Fig. 11. Frequency distribution of TMPA errors (calculated as differences between TMPA and the reference dataset, MPE in this plot).

accuracy in comparison to surface reference rainfall. Two versions of the TMPA product were considered for evaluation: a real-time version and gauge-adjusted post-real-time research version. Due to the anticipated hydrologic potential of TMPA data, the evaluation was conducted at the native resolution of the product (3-hourly and 0.25°) during six tropical-related heavy rainfall storms over Louisiana, United States. Two surface-rainfall datasets (gauge-based and radar-based) were used as independent reference for evaluating the satellite product. However, given the spatial under-sampling of gauge-based estimates more emphasis is given to the radar data as a reference. A suite of statistical measures and techniques were implemented to characterize the differences and agreement between TMPA estimates and the reference values. Based on the results of this study, the following conclusions and remarks about the two TMPA products can be made:

(1) The TMPA estimates reported reasonable levels of rainfall detection for both products (~0.7 for all storms combined). The detection probability increases to 80-90% when light rain cases (<1 mm/h) are excluded.

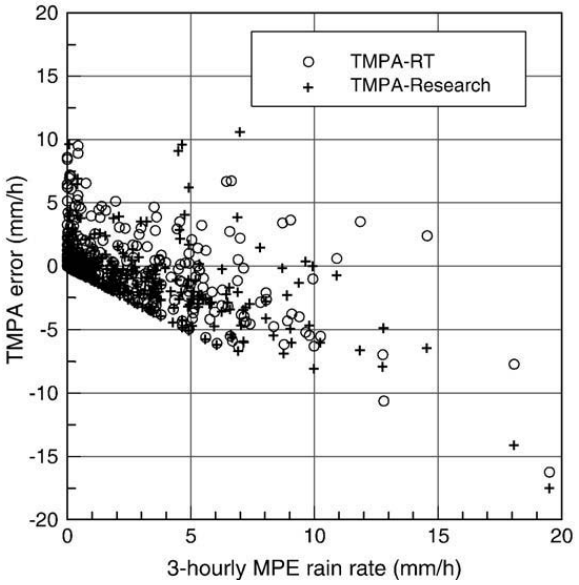


Fig. 12. Scatter plot of the TMPA errors (calculated as differences between TMPA and the reference dataset, MPE in this plot) as a function of MPE rain rates.

(2) The TMPA products are characterized with varied degrees of bias across the different analyzed storms (within ±25% for the research product and ±50% for the RT version). Decomposition of the total bias indicated that the bias during successful detections is the dominant component, with the next significant component being the missed-rain bias. The research product was associated with slightly worse missed-rain bias than the RT product which indicates that some rainy intervals were incorrectly set to zero-rain rates during the bias-adjustment process of the research product. False detections by the RT estimates are observed in one storm only, and are completely eliminated in the research version estimates, indicating the successful removal of false alarms in the TMPA research product.

(3) Overall, the TMPA products track reasonably well the temporal evolution and fluctuations of surface rainfall during the tropical-related storms analyzed in this study as reflected in the moderate to high correlation values during most storms (0.4-0.8 and 0.6-0.8 for the RT and research products, respectively).

(4) Point-by-point comparison between TMPA and the reference estimates shows a significant degree of scatter, which resulted in rather large values of the relative standard deviation of the differences (100-150% for the research product and up to 200% for the RT estimates). The standard deviation was reduced was reduced to 76% and 90% for the research and the RT products, respectively, when aggregated to a daily scale. We note that the significant scatter at the 3-hourly native resolution may be partly attributed to the ambiguity associated with defining the temporal representativeness of "instantaneous" TMPA estimates (Villarini and Krajewski, 2007). Further analysis may

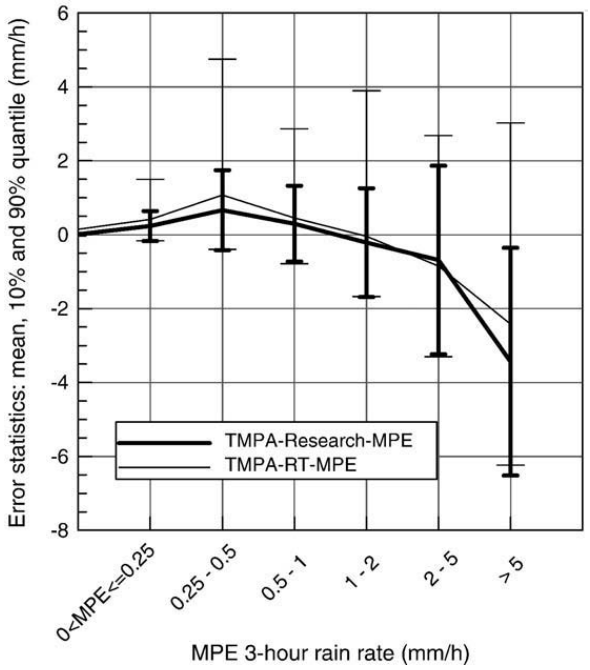


Fig.13. Conditional distribution of the TMPA errors as a function of MPE rain rates. The lines represent mean of the error and the bars represent the 10 and 90% quantiles.

Author's personal copy

be required to quantify the contribution and effect of differences in temporal sampling characteristics between TMPA and reference estimates used in the validation assessment.

(5) The TMPA RT product agrees reasonably well with the reference estimates in terms of the probability of exceedance of various surface rainfall thresholds. Overall, the research version of the product tends to show lower exceedance probabilities compared to the reference estimates, especially at high threshold values. This may be attributed to effects of the monthly bias removal procedure which may lead to reducing high rain rates. If this effect is shown in additional studies, the procedure may need to be revised to prevent this from happening.

(6) Examination of the conditional distribution of TMPA errors indicates that, on average, both products tend to overestimate low rain rates and underestimate high rain rates. The estimation errors associated with the research product have a narrower conditional distribution (especially on the positive side of the error) for all ranges of the reference rain rate.

The statistical metrics used in the product evaluation showed mixed behavior across the different storms analyzed. For example, bias was reported as both overestimation and underestimation. The improvement in the research product performance over that of the real-time product was more noticeable in certain storms than in others, with some storms showing deterioration in performance for the research product. At the current level of our analysis, these inter-storm differences are considered random. Physical explanation of such varying behavior across storms, if it exists, will require more analysis on the rainfall characteristics during these tropical storms and also on the product development details for each storm separately, which is beyond the scope of this study. The analysis reported in this study can be further extended to examine other performance aspects of the TMPA products. For example, it should be interesting to examine how the performance varies when the tropical storms move from over the Gulf towards inland (i.e., water versus land pixels). Some physical insight can be gained by keeping track of the product accuracy during various growth and decay stages of each tropical storm.

It is noted that, overall, the real-time version of the TMPA product performed less favorably than the research-quality post-real-time version. While the research product (available several days after the end of the month) may be suitable for post-type analysis, it is the real-time product (available few hours after observation time) that is of most interest to end users especially during the occurrence of extreme events and their associated hazards. Examples of TMPA real-time desired applications include global landslide detection system (e.g., Hong et al., 2007) and operational flood prediction and forecasting systems (e.g., Li et al., 2008). In this regard, it should be noted that there are on-going efforts to improve the RT version so that it has statistical characteristics closer to the research product. For example, a recent upgrade of the real-time version (Huffman et al., in press) incorporates several additional satellite data sources and employs monthly climatological adjustments to approximate the bias characteristics of the research quality product. This improvement may alleviate some of the limitations noted in this study for the real-time version of the satellite precipitation data.

An outstanding challenge in satellite rainfall validation efforts that has not been addressed in the current study is related to uncertainties in the datasets used as a ground reference. In this study, and in other recent related investigations, radar-based rainfall datasets (e.g., Stage IV product) were used as a validation reference due to their full coverage of the scale of a typical satellite pixel. An implicit assumption is that the greater part of deviations of satellite estimates from the radar dataset is attributed to the satellite error. Such an assumption is not fully justified given the inherent uncertainty of radar-based estimates. The MPE Stage IV product is relatively new and, to the best of the authors' knowledge, the literature does not include validation studies on assessing the performance of Stage IV estimates during the landfall of tropical storms. Using an independent small-scale dense rain gauge network, Habib et al. (2009) found that the MPE product was biased with respect to the total surface rainfall volume during tropical storms Matthew and Rita by -8% and 18%, respectively. Such event-scale biases were obtained over one MPE pixel ($4\times4\ km^2$) and may not fully represent the bias over the TMPA pixel scale; however, they indicate the potential problems associated with using MPE-like datasets to validate TMPA and other similar products. In a recent study, Villarini et al. (2009) presented an approach to account for radar-rainfall conditional and unconditional biases and indicated that the presence of errors in ground-based radar could significantly affect the results of satellite evaluations. Gebremichael et al. (2005) presented a framework for addressing the contribution of the reference error and separating its effect. Given the increasing interest in using radar-based datasets as a reference for validating the satellite products, further investigations are required to better understand and characterize the behavior of the radar error and how it scales up when aggregated to the satellite pixel scale.

With all this kept in mind, the analysis shows that the TMPA data set, especially the research version, is potentially useful for examining rainfall events and statistics for land-falling tropical storms, especially in areas–most of the world–that do not have coastal radars. The apparent advantage of TMPA products is the quasi-global coverage at relatively fine scales; however, the analysis reported in this paper indicated that this may come at the expense of relatively large errors. In this regard, users of the TMPA-like datasets can anticipate more improvements in future satellite-based precipitation estimates. Besides the improvements in the real-time TMPA version s noted above, an upgrade (expected in late 2009) for the research product is designed to provide a variety of improvements in input datasets and to address several algorithmic issues. Other future enhancements for the TMPA product (Huffman et al., in press) are expected to focus on extension to higher latitudes and a shift to a Lagrangian time interpolation scheme. On the long term, the anticipated Global Precipitation Mission (GPM) (Smith et al., 2007), which is a successor to TRMM, is designed to provide high-resolution (~10 km) measurements of global precipitation from a deployed constellation of remote sensing satellites and will have 3-h average revisit time over 80% of the globe. More importantly, GPM will also provide a TRMM-like core satellite

Author's personal copy

to better calibrate all of the microwave estimates on a continuous basis, which will eventually lead to more accurate precipitation estimates that may address the relatively large errors observed in the current study. The analysis presented in this paper provides a baseline for assessment of such anticipated improvements in future satellite-based products and their potential for hydrologic applications.

Acknowledgement

The authors would like to acknowledge the following sources of support: the Louisiana Board of Regents Support Fund under contract number NASA/LEQSF (2005-2010)-LaSPACE and NASA grant number NNG05GH22H and the Research Competitiveness Subprogram of the Louisiana Board of Regents Support Fund, provided to the first author, and the LaSPACE Undergraduate Research Assistantship provided to the second author. The authors also acknowledge Dr. Yang Hong for his input during the early stage of this study and Dr. Dongsoo Kim for providing the quality-controlled version of the HADS gauge dataset.

References

Artan, G., Gadain, H., Smith, J.L., Asante, K., Bandaragoda, C.J., Verdin, J.P., 2007. Adequacy of satellite derived rainfall data for stream flow modeling. Nat Hazards 43 (2), 167.

Adler, R.F., Huffman, G.J., Keehn, P.R., 1994. Global tropical rain estimates from microwave-adjusted geosynchronous IR data. Remote Sens. Rev. 11, 125–152.

Arkin, P.A., Meisner, B.N., 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982–1984. Mon. Wea. Rev. 115, 51–74.

Bellerby, T., Todd, M., Kniveton, D., Kidd, C., 2000. Rainfall estimation from a combination of TRMM precipitation radar and GOES multi-spectral satellite imagery through the use of an artificial neural network. J. Appl. Meteor. 39, 2115–2128.

Breidenbach, J.P., Bradberry, J.S., 2001. Multisensor precipitation estimates produced by National Weather Service River Forecast Centers for hydrologic applications. Proceedings of the 2001 Georgia Water Resources Conf., March 26-27, 2001, Institute of Ecology. InUniversity of Georgia, Athens.

Curtis, S, Crawford, T.W., Lecce, S.A, 2007. A comparison of TRMM to other basin-scale estimates of rainfall during the 1999 Hurricane Floyd flood. Nat Hazards 43 (2), 187–198. doi:10.1007/s11069-006-9093-y.

Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Amer. Meteor. Soc. 88, 47–64.

Fulton, R.A., Breidenbach, J.P., Seo, D., Miller, D.A., O'Bannon, T., 1998. The WSR-88D Rainfall Algorithm. Wea. and Forecasting 13 (2), 380–391.

Fulton, R.A, 2002. Activities to improve WSR-88D radar rainfall estimation in the National Weather Service. 2nd Federal Interagency Hydrologic Modeling Conference, Las Vegas, NV, July 28 – August 1, 2002.

Gebremichael, M., Krajewski, W.F., Morrissey, M.L., Huffman, G.J., Adler, R.F., 2005. A detailed evaluation of GPCP one-degree daily rainfall estimates over the Mississippi River Basin. J Appl Meteorol 44 (5), 665–681.

Gottschalck, J., Meng, J., Rodell, M., Houser, P., 2005. Analysis of multiple precipitation products and preliminary assessment of their impact on Global Land Data Assimilation System land surface states. J. Hydrometeor. 6, 573–598.

Habib, E., Larson, B.F., Graschel, J., 2009. Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in South Louisiana". J. Hydrol. 373 (3–4), 463–478. doi:10.1016/j.jhydrol.2009.05.010.

Harris, A., Rahman, S., Hossain, F., Yarborough, L., Bagtzoglou, A.C., Easson, G., 2007. Satellite-based flood modeling using TRMM-based rainfall products. Sensors 7, 3416–3427.

Hong, Y., Hsu, K., Moradkhani, H., Sorooshian, S., 2006a. Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. Water Resour Res 42, W08421. doi:10.1029/2005WR004398.

Hong, Y., Adler, Robert, Huffman, George, 2006b. Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. Geophys Res Lett 33, L22402. doi:10.1029/2006GL028010.

Hong, Y., Adler, R.F., Negri, A., Huffman, G.J., 2007. Flood and landslide applications of near realtime satellite rainfall products, *Nat*. Hazards 43 (2), 285–294.

Hossain, F., Anagnostou, E.N., 2004. Assessment of current passive-microwave and infrared-based satellite rainfall remote sensing for flood prediction D07102 J. Geophys. Res. 109, 1–14.

Hossain, F., Anagnostou, E.N., 2006. Assessment of a multi-dimensional satellite rainfall error model for ensemble generation of satellite rainfall data. IEEE Geosciences and Remote Sensing Letters 3 (3), 419–423. doi:10.1109/LGRS.2006.873686.

Hossain, F., Huffman, G.J., 2008. Investigating error metrics for satellite rainfall at hydrologically relevant scales. J. Hydrometeorology 9, 563–575.

Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin, in press. The TRMM multi-satellite precipitation analysis (TMPA). Chapter in *Satellite Applications for Surface Hydrology*, F. Hossain and M. Gebremichael, Eds. Springer Verlag. Accepted PDF available at ftp://meso.gsfc.nasa.gov/agnes/huffman/papers/TMPA_hydro_rev.pdf.

Huffman, G., Adler, R.F., Morrisey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., Susskind, J., 2001. Global precipitation at one-degree daily resolution from multi-satellite observations. J. Hydrometeorology 2, 36–50.

Huffman, G., Adler, R., Bolvin, D., Gu, G., Nelkin, E., Bowman, K., Stocker, E., Wolff, D., 2007. The TRMM Multi-satellite precipitation analysis: quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. J. Hydrometeorology 8, 38–55.

Jiang, H., Halverson, J.B., Simpson, J., Zipser, E.J., 2008. Hurricane "rainfall potential" derived from satellite observations aids overland rainfall prediction. J. Appl. Meteor. Climatol. 47, 944–959.

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

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

Kim, D., Nelson, B., Cedrone, L., 2006. Reprocessing of historic hydrometeorological automated data system (HADS) precipitation data. 86th AMS Annual Meeting, 29 January - 2 February 2006, Atlanta, Georgia, combined preprints CD-ROM. InAmerican Meteorological Society, Boston, MA. 10 IOAS 8.2), 6 pp. (January 2006).

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

Li, L., Hong, Y., Wang, J., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Korme, T., Okello, L., 2008. Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria. Africa. Natural Hazards. doi:10.1007/s11069-008-9324-5.

Negri, A, Burkardt, N., Golden, J.H., Halverson, J.B., Huffman, G.J., Larsen, M.C., Mcginley, J.A., Updike, R.G., Verdin, J.P, Wieczorek, J.F., 2004. The hurricane-flood-landslide continuum. Bull. Amer. Meteor. Soc. 86 (9). doi:10.1175/BAMS-86-9-1241.

Rappaport, E.N., 2000. Loss of life in the United States associated with recent Atlantic tropical cyclones. Bull. Amer. Meteor. Soc. 81, 2065–2074.

Rudolf, B., 1993. Management and analysis of precipitation data on a routine basis. In: Sevruk, B., Lapin, M. (Eds.), *Proceedings of International Symposium on Precipitation and Evaporation*, Vol.1. InSlovak Hydrometeorology Institution, pp. 69–76.

Smith, E, et al., 2007. The International Global Precipitation Measurement (GPM) program and mission: An overview. In: Levizzani, V., Turk, F.J. (Eds.), *Measuring Precipitation from Space: URAINSAT and the Future*. Springer, pp. 611–653.

Seo, D.J., Briedenbach, J.P., Johnson, E.R., 1999. Real-time estimation of mean field bias in radar rainfall data. J. Hydrol. 223, 131–147.

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

Su, F., Hong, Y., Lettenmaier, D.P., 2008. Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. J. Hydrometeor. 9, 622–640.

Tian, Y., Peters-Lidard, C.D., 2007. Systematic anomalies over inland water bodies in satellite-based precipitation estimates. Geophys Res Lett 34, L14403. doi:10.1029/2007GL030787.

Tian, Y., Peters-Lidard, C.D., Choudhury, B., Garcia, M., 2007. Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications December 2007 J. Hydromet 8 (6), 1165–1183.

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

Author's personal copy

Turk, F.J., Miller, S.D., 2005. Toward improving estimates of remotely-sensed precipitation with MODIS/AMSR-E blended data techniques. IEEE Trans. Geosci. Rem. Sens. 43, 1059–1069.

Vicente, G., 1994. Hourly retrieval of precipitation rate from the combination of passive microwave and IR satellite radiometric measurements. Ph.D. dissertation, University of Wisconsin. 127 pp.

Villarini, G., Krajewski, W.F., 2007. Evaluation of the research-version TMPA three-hourly 0.25° × 0.25° rainfall estimates over Oklahoma. Geophys Res Lett 34, L05402. doi:10.1029/2006GL029147.

Villarini, G., Krajewski, W.F., Smith, J.A., 2009. New paradigm for statistical validation of satellite precipitation estimates: application to a large sample of the TMPA 0.25-degree three hourly estimates over Oklahoma, Journal of Geophysical Research 114, D12106. doi:10.1029/2008JD011475.

Wang, X., Xie, H., Sharif, H., Zeitler, J., 2007. Validating NEXRAD MPE and Stage III precipitation products for uniform rainfall on the Upper Guadalupe River Basin of the Texas Hill Country. J Hydrol 348, 73–86.

Westcott, N.E., Knappb, H.V., Hilberg, S.D., 2008. Comparison of gage and multisensor precipitation estimates over a range of spatial and temporal scales in the Midwestern United States" 351(1), 1-12 J Hydrol 348, 73–86. doi:10.1016/j.jhydrol.2007.10.057.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, 2nd edition. Burlington, MA, Academic Press. 627 pp.

Willmott, J.R., 1981. On the validation of models. Phys Geogr 2, 184–194.

Young, C.B., Brunsell, N.A., 2008. Evaluating NEXRAD estimates for the Missouri River Basin: analysis using daily rain gauge data. Journal of Hydrologic Engineering 13 (7), 549–553.