Evaluation of the High-Resolution CMORPH Satellite Rainfall Product Using Dense Rain Gauge Observations and Radar-Based Estimates

EMAD HABIB
Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, Louisiana

ALEMSEGED TAMIRU HAILE
Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, Louisiana, and African Climate Policy Center, UNECA, Addis Ababa, Ethiopia

YUDONG TIAN
Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

ROBERT J. JOYCE
NOAA Climate Prediction Center, Camp Springs, Maryland

(Manuscript received 25 January 2012, in final form 15 May 2012)

ABSTRACT

This study focuses on the evaluation of the NOAA–NCEP Climate Prediction Center (CPC) morphing technique (CMORPH) satellite-based rainfall product at fine space–time resolutions (1 h and 8 km). The evaluation was conducted during a 28-month period from 2004 to 2006 using a high-quality experimental rain gauge network in southern Louisiana, United States. The dense arrangement of rain gauges allowed for multiple gauges to be located within a single CMORPH pixel and provided a relatively reliable approximation of pixel-average surface rainfall. The results suggest that the CMORPH product has high detection skills: the probability of successful detection is ~80% for surface rain rates $>2$ mm h$^{-1}$ and probability of false detection <3%. However, significant and alarming missed-rain and false-rain volumes of 21% and 22%, respectively, were reported. The CMORPH product has a negligible bias when assessed for the entire study period. On an event scale it has significant biases that exceed 100%. The fine-resolution CMORPH estimates have high levels of random errors; however, these errors get reduced rapidly when the estimates are aggregated in time or space. To provide insight into future improvements, the study examines the effect of temporal availability of passive microwave rainfall estimates on the product accuracy. The study also investigates the implications of using a radar-based rainfall product as an evaluation surface reference dataset instead of gauge observations. The findings reported in this study guide future enhancements of rainfall products and increase their informed usage in a variety of research and operational applications.

1. Introduction

Spatial rainfall information represents one of the main inputs to hydrologic, climatologic, and agricultural studies. Rain gauges represent the most direct way for measuring surface rainfall. However, rain gauges are limited by their near-point observations and lack of spatial coverage over many parts of the world. Weather radars represent a viable alternative to rain gauges because of their continuous spatial coverage; however, establishing a world network of radar stations is rather difficult because of accessibility and financial limitations. Considering limitations of gauge and radar networks, satellite sensors remain the only viable options to capture rainfall fields over most parts of the globe.

The two most widely used satellite techniques for rainfall estimation are based on observations from passive microwave (PMW) and thermal infrared (IR) sensors with the former providing more direct observations...
of rainfall than IR sensors. However, observations from PMW sensors have temporal and spatial gaps. This has led to the emergence of algorithms that combine more direct observations from PMW sensors and frequently available observations from IR sensors (e.g., Joyce et al. 2004; Hsu et al. 1997; Sorooshian et al. 2000; Turk and Miller 2005; Huffman et al. 2007). One such algorithm is the National Oceanic and Atmospheric Administration’s (NOAA) Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004). The CMORPH rainfall product is available since December 2002 at various spatial and temporal resolutions (e.g., 8 × 8 km², 0.25° × 0.25°; 30 min, 3 hourly, and daily) for regions that are situated between 60°N and 60°S. Because of its near-real-time availability and high temporal and spatial resolution, this product can be very useful in a variety of hydrologic and water resources applications.

There are several sources of uncertainties that affect the accuracy of satellite rainfall products, which can be grouped into two categories: sampling and retrieval (e.g., Gebremichael and Krajewski 2004; Durden et al. 1998). Therefore, it is necessary to assess the uncertainties in satellite rainfall products before being used in operational applications (e.g., Adler et al. 2001; Turk et al. 2008; Sapiano and Arkin 2009). The reliability of the uncertainty assessment depends on the availability of a fairly accurate knowledge of the “true” surface rainfall distribution and intensity.

Joyce et al. (2004) validated the CMORPH rainfall product (at a scale of daily, 0.25° × 0.25°) over Australia and the United States using gauge-based gridded rainfall data. They reported that CMORPH outperforms blended IR–PMW rainfall estimation techniques that use IR-derived estimates of rainfall when PMW data are not available. Similarly, Dinku et al. (2010) showed that CMORPH slightly underestimates 10-day rainfall (at a scale of 0.25° × 0.25°) while performing better than blended IR–PMW as well as IR-only estimation algorithms over the mountains of Ethiopia. Tian et al. (2007) reported that CMORPH has large positive bias in summer and negatives bias in winter over the southeast United States (at daily, 0.25° × 0.25°). Sapiano and Arkin (2009) reported that the correlation between 3-hourly 0.25° × 0.25° CMORPH rainfall amounts and the corresponding gauge data reaches as high as 0.7. They also reported that CMORPH has a general tendency to overestimate warm season rainfall over the central United States and to underestimate rainfall over the tropical Pacific Ocean. Ebert et al. (2007) reported that at daily and 0.25° × 0.25° scales, CMORPH mostly underestimates the number of rainy days in Australia although its daily rainfall depth has the highest correlation with respect to other satellite rainfall products.

Gauge-based gridded rainfall data usually have coarser resolutions than that of CMORPH and therefore cannot serve as a reference for evaluating the 8-km, 30-min high-resolution version of CMORPH. The density of operational gauges is often too poor to satisfactorily capture rainfall characteristics at subdaily scales and fine spatial resolutions (e.g., Haile et al. 2009; Habib et al. 2009). In addition, the methods used to interpolate gauge observations onto a regular grid usually introduce some degree of smoothing that may not always be acceptable (Ebert et al. 2007).

Experimental studies showed that more than a single gauge is required within a grid pixel of a certain satellite product to satisfactorily capture subgrid variability of rainfall (e.g., Huff 1970; Habib and Krajewski 2002). The difference between area-rainfall and gauge-point rainfall estimates must be considered in interpreting the satellite–gauge difference statistics. For instance, Habib et al. (2009) showed that using observations from a single gauge as a reference to assess radar-based 4 × 4 km²-scale hourly product can result in an unrealistic inflation of the actual product estimation error by 120%–180%. Area-point differences are not always known and consequently results of most validation studies of high-resolution satellite rainfall products can be rather inconclusive.

Considering such limitations, Ciach et al. (2007) advocated the need for experimental gauges that provide sufficiently accurate approximations of the true area-averaged rainfall. Rainfall records from experimental gauge networks are quality controlled and have high resolution. Therefore these rainfall records, when available, can provide an opportunity to benchmark the accuracy of satellite rainfall products to high-resolution surface data (see Anagnostou et al. 2010). Such benchmarking will provide useful information to hydrologists and other users that are interested in applying satellite rainfall products at high spatiotemporal resolution. The first objective in the present study is to evaluate the accuracy of CMORPH at its finest spatial resolution using a high-quality dense experimental gauge network in southern Louisiana, United States.

Experimental rain gauge networks cover only small areas; therefore, evaluation studies used radars as a source of reference data to evaluate satellite rainfall products (e.g., Zeweldi and Gebremichael 2009; Ebert et al. 2007; AghaKouchak et al. 2011). Because of their large spatial coverage, radar networks provide a highly promising reference data source that can be potentially superior to operational gauge networks for validation of satellite rainfall products. Using the Oklahoma Mesonet gauge network (interpolated to 25 × 25 km² at 3 hourly), Anagnostou et al. (2010) reported that the radar-based Stage IV rainfall product has significantly lower
bias and root-mean-square error than CMORPH. Zeweldi and Gebremichael (2009) applied radar-based rainfall product to evaluate the accuracy of CMORPH at fine space–time scales (8 × 8 km² and hourly). Such validation efforts, however, may not provide the true characteristics of the errors associated with satellite rainfall products, since radar-rainfall products have their own uncertainties. For example, Habib et al. (2009) reported that the U.S. National Weather Service (NWS) radar-based Multisensor Precipitation Estimator (MPE) product overestimates low rain rates up to 60%–90% for rates lower than 0.5 mm h⁻¹, and underestimates high rain rates up to −20% for rates higher than 10 mm h⁻¹. A thorough review of radar-rainfall uncertainties is available in Villarini and Krajewski (2010). With these uncertainties in mind, the second objective in the present study is to evaluate implications of using radar-based rainfall estimates instead of gauge data as a source of reference data in evaluation of high-resolution satellite products.

The present study focuses on evaluation of the CMORPH rainfall product at its highest spatial resolution (8 × 8 km²) and at different temporal scales (hourly, event, and daily). The performance accuracy of CMORPH is evaluated using reference rainfall information from a dense gauge cluster and a radar-based product. The study uses a suite of summary error metrics that provide hydrologically relevant information about the accuracy of satellite rainfall products. Conditional versions of these statistical error metrics are estimated by conditioning the CMORPH rainfall estimates on varying thresholds of reference rain rates. The analysis addresses other relevant product attributes, such as effect of availability of PMW estimates with respect to the target analysis time. The results of the current study will guide future enhancements of satellite rainfall algorithms and increase their informed usage in a variety of research and operational applications. The results will also guide future evaluation studies that rely more, and exclusively, on radar-based products as an evaluation reference because of the lack of adequate rain gauge networks.

2. Study area and datasets

a. CMORPH satellite product

The satellite rainfall product under evaluation in our study is CMORPH (Joyce et al. 2004). CMORPH combines rainfall estimates from multiple PMW sensors, which include the Advanced Microwave Sounding Unit B (AMSU-B), the Special Sensor Microwave Imager (SSM/I), the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), and the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), respectively. To fill the time and space gap in the combined PMW-based rainfall estimates, the algorithm used cloud motion vectors derived from spatial lag correlation of successive geostationary satellite IR images. The cloud motion vectors are further adjusted to rainfall propagation motion using radar-rainfall motion. These adjusted vectors are used to propagate the PMW-based rainfall features for time periods between two successive PMW overpasses. The shape and intensity of the rainfall patterns is then morphed through linear interpolation using weights that are obtained from forward advection (previous to most current PMW overpass) and backward advection (most current to previous overpass) of rainfall features. The CMORPH product is available starting since December 2002 at a finescale of 30 min in time and 8 × 8 km² in space.

b. Rain gauge cluster

The CMORPH product will be evaluated using surface observations from a dense cluster of rain gauges located within the Isaac-Verot watershed (~35 km²) in the city of Lafayette, Louisiana (Fig. 1). The gauge network is composed of 13 tipping-bucket rain gauge sites, with each site having two gauges located side by side in a dual-gauge setup. To ensure high quality and continuity of gauge observation and data collection, a series of quality control and maintenance procedures were implemented, such as frequent site maintenance and data downloading for early detection of malfunctions. The gauges record the time of occurrence of successive 0.254-mm (0.01 in.) tips, which can then be used to estimate hourly rainfall intensities (or accumulations) to match the resolution of the CMORPH and the radar-based MPE products. It should be noted that tipping-bucket gauges are subject to significant levels of random errors at low rainfall intensities (Habib et al. 2001; Ciach 2003). Based on the network’s spatial configuration, 7 of the 13 gauges are situated within one CMORPH 8 × 8 km² pixel. Within this pixel, the intergauge distances are on the order of 1–2 km, which is smaller than the correlation distance of hourly rainfall in this area (Habib et al. 2009). Therefore, these seven dual-gauge sites will be used to provide a reliable estimate of the true surface rainfall over the 8-km scale of the CMORPH pixel.

Based on the availability of the gauge observations from the experimental network, a period of 28 months (August 2004–December 2006) was selected for this study. This period represents average annual rainfall conditions (about 1300 mm or 50 in.) and included more than 130 rainfall events. In this study, an event is defined as continuous raining period interrupted by no longer than 6 h of no rain and with a rainfall depth of at least
10 mm. The study period witnessed two major events: Tropical Storm Matthew in October 2004 and Hurricane Rita in September 2005, where each of the two events generated more than 250 mm of rainfall over the network site. Frequent occurrences of sustained intense rain in excess of 25 mm h\(^{-1}\) are observed during the study period (see Habib et al. 2009 for a detailed description of the dataset).

c. Radar-based MPE product

The radar-based product used in this study comes from the NWS MPE algorithm, which is routinely run at the NWS River Forecasting Centers (RFCs). The MPE produces hourly rainfall estimates on the approximately 4 \(\times\) 4 km\(^2\) Hydrologic Rainfall Analysis Project (HRAP) grid. The algorithm is based on combining radar-based estimates from the Next Generation Weather Radar (NEXRAD) Weather Surveillance Radar-1988 Doppler (WSR-88D) radars, with data from operational rain gauges. The algorithm uses different routines for bias correction and optimal merging of radar-only estimates with gauge observations, and produces as many as seven precipitation products (Fulton et al. 2002) based on different combinations of radar-gauge adjustment and merging (Seo et al. 2010, and references therein). The forecasters at the RFC have the flexibility to choose whichever product to be eventually used for their hydrologic operations.

A gauge-only analysis (GAGEONLY) is first produced by the MPE algorithm and is based on optimal interpolation of hourly rain gauge observations within the service area of the RFC. All of the subsequent MPE products start from the hourly gridded Digital Precipitation Product (DPA; Klazura and Imy 1993). The DPA is generated by the WSR-88D Precipitation Processing System (PPS), where a power law \(Z-R\) relationship is applied to the raw reflectivity data, which are then integrated over time to produce hourly accumulation.

A radar-only mosaic (RMOSAIC) product is produced by mosaicking the DPA products without any use of gauge observations. A simple mosaicking technique is used in areas of radar coverage overlap where the precipitation estimate is obtained from the radar with the lowest unblocked beam. The RMOSAIC product may have significant biases that may vary by radar site, season, and precipitation regime. To adjust for these biases, the MPE algorithm has two alternative methods: the first is a mean-field bias-correction method and is based on applying a radar-specific, time-varying but spatially uniform multiplicative adjustment factor to each pixel within the effective coverage of the radar in the DPA product (Seo et al. 1999). The result is a new product called the mean-field bias-adjusted radar mosaic (BMOSAIC). This correction is intended to adjust for systematic, spatially uniform biases that may result from inappropriate \(Z-R\) relationships or radar calibration problems.
The second alternative bias-adjustment method in the MPE algorithm is a local-bias correction that corrects spatially nonuniform biases in the RMOSAIC field (e.g., biases due to bright band and the vertical profile of reflectivity effects) and results in a new product called LMOSAIC. A merging procedure that uses a cokriging-like optimal estimation technique is then applied to both of the two bias-adjusted products, BMOSAIC and LMOSAIC, to produce two multisensor products, MMOSAIC and LMOSAIC, respectively. Finally, the MPE algorithm gives the RFC forecaster the option to choose in real time amongst the different MPE products described above. The final selection, known as the XMRG product, is not a new product in itself, but is driven by the forecaster’s experience and assessment of the current situation and often includes manual corrections made by the forecaster. Thus it is considered the “best” MPE multisensor product since it combines the automatic data processing via the algorithm and the experienced input by the human forecasters. As such, the current study will use the XMRG product from the MPE algorithm.

The MPE–XMRG data during the study period were obtained from the MPE archives of the NWS Lower Mississippi River Forecast Center (LMRFC). The MPE product over the study site is primarily derived from the WSR-88D KLCH radar in Lake Charles, Louisiana, which is approximately 120 km away from the study site. At such a distance, the height of the axis of the lowest radar beam is about 1.82 km above the ground surface. It should be noted that at such a range, beam partial filling and overshooting of lower cloud bases and shallower precipitation might become a concern. Beam overshooting may result in low probability of detection, especially with light rain. Relative degradation of the beam sampling resolution increases the likelihood that rainfall fills only part of the beam and may result in underestimation of the rainfall intensity due to sample volume averaging of the received power. It is also noted that the 13 rain gauges used in this study were not used by LMRFC in developing the MPE products over the study site.

3. Methods

In the present study, a three-way intercomparison analysis is performed: 1) the CMORPH product is compared against observations from the dense rain gauge network, 2) the CMORPH product is compared against the MPE product, and 3) the MPE product is compared against observations from the dense rain gauge network. The first comparison serves the first objective of the study on evaluation of the CMORPH product using a reference dataset that is accurate and representative of surface-area-rainfall amounts at the scale of the CMORPH product. The second and third comparison analyses will assess the use of a radar-based product as an alternative reference dataset that provides consistent and spatially available surface rainfall information.

The CMORPH and the radar-based products have $8 \times 8$ km$^2$ and $4 \times 4$ km$^2$ spatial resolution, respectively, while gauges represent point-rainfall observations. Therefore, the first step needed is to convert these different datasets into a consistent spatial resolution for reliable assessment and intercomparison. The resolution of CMORPH will serve as the basis. Since most of the gauges in the present study are situated in a single pixel of CMORPH (Fig. 1), the analysis is limited to this grid pixel. The Thiessen polygon interpolation method is used to estimate $8 \times 8$ km$^2$ area-averaged hourly rainfall ($G_{avg}$) from individual rain gauges at the spatial resolution of CMORPH. It is noted that while these gauges are not uniformly distributed within the pixel, the use of seven gauges, in contrast to a single gauge per pixel as typically available from other observational networks, is likely to provide an improved estimate of the areal rainfall over the pixel size. The $4 \times 4$ km$^2$ MPE hourly rainfall estimates are averaged to the CMORPH $8 \times 8$ km$^2$ pixel by weighting based on partial fractions of the MPE pixel areas located within the CMORPH pixel. For simplicity, these spatially averaged estimates will also be referred to as MPE.

A suite of graphical representations and statistical metrics will be used to evaluate the accuracy of CMORPH. In the following formulation, $R_{CMORPH}$, $G_{avg}$, and $R_{MPE}$ are used to denote rain rates, or accumulations, over the CMORPH $8 \times 8$ km$^2$ pixel size assembled from the CMORPH product, the average of rain gauge observations within the CMORPH pixel (Fig. 1), and the area-weighted average of MPE that intersect with the CMORPH pixel, respectively. The parameter $R_{REF}$ is used to denote the reference rainfall rate, or accumulation, that can be either $G_{avg}$ or $R_{MPE}$. As discussed earlier, the statistical evaluation is performed using three-way comparisons ($R_{CMORPH}$ versus $R_{Gavg}$, $R_{CMORPH}$ versus $R_{MPE}$, or $R_{MPE}$ versus $R_{Gavg}$). The rainfall samples are based on paired datasets excluding hours where no rainfall was recorded by any of the three datasets. The analysis will focus on assessing both the systematic and random errors using statistical metrics that have been proposed and commonly used in the satellite rainfall literature (Hossain and Huffman 2008) including bias, standard deviation, and measures of association. The exceedance probability distributions of the three datasets will also be compared. The detection skill of the CMORPH product is assessed through categorical measures such as probabilities of
detection (POD) and false detection (POFD). First, these metrics will be calculated based on the entire sample to obtain summary statistics, also referred to as unconditional statistics. Next, different conditioning mechanisms on the selected metrics will be applied to obtain conditional statistics that give a detailed insight about how the CMORPH product performs at different rainfall rates.

a. Summary statistics

The most fundamental evaluation metric is the bias ($B$), which can be defined as the mean difference between the rainfall estimates under evaluation ($R_{\text{CMORPH}}$ or $R_{\text{MPE}}$) and the reference values ($R_{\text{avg}}$ or $R_{\text{MPE}}$). For example, the bias of the CMORPH product can be expressed as

$$B = R_{\text{CMORPH}} - R_{\text{REF}}; \quad \text{RB} = \frac{R_{\text{CMORPH}} - R_{\text{REF}}}{R_{\text{REF}}}.$$ (1)

The formula to the left expresses the bias in the same units of the estimate, while the one on the right is a relative bias (RB) that is normalized by the mean of the reference. Similar formulas for the bias (and for the following statistical metrics) can be written for the evaluation of $R_{\text{MPE}}$ when using $R_{\text{avg}}$ as a reference but will not be shown here for brevity.

The random error of the product can be characterized using the standard deviation statistic ($\sigma$) in absolute and relative units:

$$\sigma(R_{\text{CMORPH}} - R_{\text{REF}}); \quad \frac{\sigma(R_{\text{CMORPH}} - R_{\text{REF}})}{R_{\text{REF}}}.$$ (2)

The detection capability of CMORPH will be quantified using POD and POFD. POD is estimated as the ratio of the number of correct hourly rainfall detections to the total number of rainfall occurrences in the reference dataset. The POFD is calculated as the ratio of the number of correct hourly rainfall detections to the total number of actual nonrainfall hours. Both POD and POFD range from 0 to 1, with 1 being a perfect POD and 0 being a perfect POFD. The study will calculate the conditional volumes of rainfall correctly and incorrectly identified by the rainfall product as an indication of the significance of the calculated POD and POFD values.

b. Conditional statistics

To provide further insight into how the CMORPH product performs at different ranges of reference rainfall rates, a parallel set of the summary statistics described above can be developed in a conditional sense (called conditional statistics), where each statistic is calculated for varying ranges of rainfall rates within the available sample. The POD and POFD will be recalculated for rain rates higher than preselected threshold values that cover the full range of the sample.

The total bias as estimated by Eq. (1) is a summary statistic aggregated over the entire dataset. Assessment of sources of the total bias helps to gain additional insight into performance accuracy of CMORPH. Therefore, following Tian et al. (2009) and Habib et al. (2009), we decomposed the total bias into three components: hit bias (HB), missed-rain bias (MB), and false-rain bias (FB). HB refers to the total difference between the remote sensing rainfall product and the corresponding reference when both detect rainfall. MB refers to the total rainfall depth reported in the reference dataset when the rainfall product does not report rainfall. FB refers to the total amount of falsely detected rainfall by the rainfall product. The three bias components add up to the total bias:

$$\text{HB} = \sum (R_{\text{CMORPH}} - R_{\text{REF}})/(R_{\text{CMORPH}} > 0 \& R_{\text{REF}} > 0),$$ (3)

$$\text{MB} = \sum R_{\text{REF}} \mid (R_{\text{CMORPH}} = 0 \& R_{\text{REF}} > 0),$$ and (4)

$$\text{FB} = \sum R_{\text{CMORPH}} \mid (R_{\text{CMORPH}} > 0 \& R_{\text{REF}} = 0).$$ (5)

The bias components have opposite signs and can cancel each other when added up into the total bias, which may mask out their individual contributions. Thus, use of total bias may be misleading in verifying the accuracy of rainfall products unless it is jointly interpreted along with its components.

The mean and the standard deviation of the estimation error can also be formulated by conditioning on a certain reference value ($R_{\text{REF}}$). If the estimation error is defined as the difference between the remotely sensed–based estimate and the reference value $[e = (R_{\text{CMORPH}} - R_{\text{REF}})]$, then the conditional error can be expressed as $[e \mid (R_{\text{REF}} = r_j)]$. The conditional mean [also referred to as conditional bias (CB)] and the conditional standard deviation of the error can be formulated as

$$\mu_e(r_j) = E(e \mid R_{\text{REF}} = r_j) \quad \text{and}$$ (7)

$$\sigma_e(r_j) = \sqrt{E((e - \mu_e)^2 \mid R_{\text{REF}} = r_j)}.$$ (8)

In these notations, the uppercase and lowercase letters denote random variables and their experimental values, respectively. Following Ciach et al. (2007), a kernel regression approach was used to obtain a nonparametric
estimate of the two conditional statistics, \( \mu_c(r_c) \) and \( \sigma_c(r_c) \), using a moving-window averaging formula [see Ciach et al. 2007, their Eqs. (6) and (9)]. The window size is proportional to the conditioning \( R_{\text{REF}} \) value, and as such, the conditional sample size within the selected window decreases significantly as the center of the window moves to larger \( R_{\text{REF}} \) values. Therefore, no results are reported on the conditional statistics beyond \( R_{\text{REF}} = 25 \text{ mm h}^{-1} \) to maintain acceptable sizes of the conditional samples.

c. Bootstrap distributions

Bootstrap sampling distributions (Efron and Tibshirani 1993) of the conditional statistics will be computed to assess the sample size impact on the estimated statistics, especially at higher range of \( R_{\text{REF}} \), and to assess statistical significance of the observed differences in the statistics, if any, that are derived using either \( G_{\text{avg}} \) or the MPE as an evaluation reference. The bootstrap distributions will be obtained by random drawing with replacement from each conditional sample and repeating the calculations of the specific conditional statistic for each of these bootstrap pseudosamples. This resampling procedure will be repeated 500 times for each conditional sample to achieve stable uncertainty bounds for the derived conditional statistic. The bootstrap distributions are summarized by presenting their 2.5 and 97.5 percentiles.

4. Results

a. Analysis of rainfall distributions

The analysis starts by examining scatterplots of CMORPH against each of MPE and \( G_{\text{avg}} \) at hourly, daily, and event scales (plots are not shown for space limitation). Significant scatter between CMORPH and \( G_{\text{avg}} \) was evident especially at the hourly scale where differences of 15–30 mm h\(^{-1}\) were not uncommon. Several instances of failed or false detection by CMORPH up to 30–40 mm h\(^{-1}\) were observed. As expected, the daily and event scales show less scatter, suggesting a moderate improvement in the relationship; however, significant differences, as high as 50–75 mm for a given event, were still observed. The degree of scatter in the CMORPH–\( G_{\text{avg}} \) pairs is overall similar to that of the CMORPH–MPE.

Figure 2 shows the probability of exceedance defined as the probability that a rain rate exceeds a certain rainfall rate threshold. This analysis provides information that is relevant for studying extreme hydrologic events. The probabilities derived from the three rainfall datasets show a very good agreement for rain intensities between \( \sim 3–8 \text{ mm h}^{-1} \) that correspond to 3.5%–9% exceedance probabilities. Compared to the \( G_{\text{avg}} \) reference dataset, CMORPH and MPE have higher exceedance probabilities for small rain intensities (0–3 mm h\(^{-1}\) thresholds) and lower probabilities for large rain intensities (\( > 8 \text{ mm h}^{-1} \)). However, it is noted that the accuracy of these probabilities of exceedance at extreme rain rates is possibly affected by the sample size, which decreases with an increase in threshold value. Also, some level of dampening in high rain rates may be caused by the spatial averaging from the near-point gauge observations and the 4 × 4 km\(^2\) MPE estimates to the coarser 8 × 8 km\(^2\) CMORPH scale.

b. Analysis of total and conditional bias

The total bias [Eq. (1)] and its three subcomponents [Eqs. (3)–(5)] were calculated over the entire time period of the study. The gauge estimates (\( G_{\text{avg}} \)) served as a reference to evaluate the bias of both CMORPH and MPE, but the analysis also assesses the bias of CMORPH when MPE is used as a reference. As explained earlier, the three bias components (hit, missed, and false) may have opposite signs and therefore may cancel each other when added up into a total bias value. The percentages shown above or below the bars in Figure 3 represent the bias components relative to the total rainfall depth. The total bias of CMORPH estimates is only 0.04% when \( G_{\text{avg}} \) served as a reference. The hit bias of CMORPH is very small but the missed-rain and false-rain biases are significant and reach \( \sim 20\% \) of the total rainfall amount. Therefore, the small total bias of this satellite product is due to the fact that the values of its components canceled each other. This
indicates the use of total bias without Considering the magnitude of its components can be misleading. MPE has small total bias that is comparable to the magnitude of its components. The use of MPE as a reference instead of \( G_{\text{avg}} \) affected the magnitudes of hit bias and false rain of CMORPH but it only slightly affected its total bias and missed rain. The use of MPE as reference value resulted in more favorable performance accuracy of the satellite product in terms of false rain than its actual accuracy (17.16% instead of 21.83%). However, it resulted in less favorable accuracy in terms of hit bias (6.24% instead of −0.69%). The overall low bias reported with the MPE algorithm is probably due to its own bias-correction schemes that use surface observations from operational gauges (but not the same gauges used in the current study).

Next, we assess the bias at an event scale (Fig. 4). While CMORPH had a fairly negligible bias when assessed for the entire sample (Fig. 3), it has significant biases on an event scale. The histogram of the CMORPH relative bias [\( RB; \text{ Eq. (1)} \)] has skewed distribution with 59% of the events having a negative, or underestimation, bias. For some events RB reaches up to 1, indicating a bias as high as the event rain depth. We note that there are seven events with RB values that exceed a value of 1.0 (not shown in the plot). The bias of MPE on an event scale is between −0.5 and 0.7 of \( G_{\text{avg}} \). RB values of MPE are slightly less skewed than those of CMORPH since only 53% of the events have negative biases. Using MPE as reference value instead of \( G_{\text{avg}} \) did not significantly change the event-scale RB distribution of CMORPH.

The total bias estimated previously provides information over the entire sample without distinguishing between the accuracy of CMORPH for intense and light rain rates. To evaluate the accuracy of the rainfall product over a range of rain intensities, the conditional bias [\( m_r(r_s); \text{ Eq. (7)} \)] was estimated by conditioning the bias on a certain value of \( r_s \) in the reference \( G_{\text{avg}} \) or MPE (Fig. 5). The bootstrap-derived sampling distributions of the conditional bias are used to evaluate the effect of decreasing sample sizes with the increase in the conditioning intensity threshold (because of space limitation, only one distribution is shown for the CMORPH product conditioned on \( G_{\text{avg}} \); the two other conditional sampling distributions showed similar bounds). Figure 5 indicates that both of the CMORPH and MPE products have a consistently negative bias for rain rates that exceed \( \sim 3-4.0 \text{ mm h}^{-1} \). The MPE underestimation deteriorates almost linearly and can be estimated as \( \sim 20\% \) of the surface rainfall rate. The conditional bias of CMORPH is more alarming over the entire range of rain rates and can be estimated as \( \sim 50\% \) of the surface rainfall rate. For intensities \( < 10 \text{ mm h}^{-1} \), the CB of CMORPH is not significantly affected by the selected source of reference data (MPE versus \( G_{\text{avg}} \)) but there is large
difference for higher rain intensities with more favorable accuracy obtained when MPE serves as a reference. However, such differences do not appear statistically significant since they are fairly comparable to the 95% bounds of the bootstrap distribution.

c. Analysis of conditional detection

Next, the detection capability of CMORPH is assessed in terms of POD and POFD (Figure 6). The two probabilities were calculated by conditioning CMORPH on various thresholds of rain intensities ($G_{avg}$). The discussion of the detection analysis is focused on thresholds starting from 0.5 mm h$^{-1}$ since smaller thresholds are subject to discrepancy in each sensor’s sensitivity at very light rain. The analysis is repeated to evaluate the MPE product. CMORPH was also evaluated by conditioning on MPE to gain insight about the effect of using radar-based products as a reference data source. POD and POFD values do not provide the amount of rainfall missed or falsely detected. As such, we kept track of the missed and the falsely detected rainfall depths that correspond to the calculated POD and POFD values. We only show results for thresholds up to >5 mm h$^{-1}$ since the probabilities reach asymptotic values afterward.

First, consider the case when CMORPH is conditioned on $G_{avg}$. The POD value for rain intensities higher than 0.5 mm h$^{-1}$ is 65%, which results in a missed-rainfall amount of about 18%. The POD increases to reach an asymptotic value of 80% at a threshold of 2.0 mm h$^{-1}$ with a missed-rain depth of $\approx$15%. This indicates that rain intensities exceeding 2.0 mm h$^{-1}$ are detected by CMORPH 80% of the time. Using MPE as reference data instead of $G_{avg}$ resulted in similar POD results. While the probability of false-rain detection by CMORPH seems negligible, its contribution in terms of rainfall depth is relatively significant. The amount of falsely detected rain by CMORPH is about 20% for small thresholds but maintains a value of 10%–15% even with the increase of the conditioning.
threshold. These results are consistent with the conditional bias results presented earlier in the paper. The use of MPE as reference instead of $G_{\text{avg}}$ resulted in smaller falsely detected rain depth, suggesting a better performance accuracy of CMORPH than its actual accuracy.

d. Analysis of random error

The random error of the CMORPH product [defined earlier as $e = (R_{\text{CMORPH}} - R_{\text{REF}})$] was characterized using its standard deviation ($\sigma_e$) based on either $R_{\text{Gavg}}$ or $R_{\text{MPE}}$ as a surrogate of $R_{\text{REF}}$. The $\sigma_e$ was found to be rather high with values of 8.4 and 7.6 times the mean surface rainfall rate when $R_{\text{Gavg}}$ or $R_{\text{MPE}}$ were used as a reference, respectively. Similar to the conditional bias, the standard deviation of the CMORPH random error in a conditional sense [$\sigma_e(r_s)$, Eq. (8)] was assessed by conditioning on a certain value of $r_s$ in the reference $G_{\text{avg}}$ or MPE (Figure 5). A clear dependence of the CMORPH random error on the surface rainfall rate is evident. The standard deviation of the CMORPH error is significant and increases from 2 mm h$^{-1}$ at small rain intensities to $>6$ mm h$^{-1}$ for intensities higher than 10 mm h$^{-1}$.

The conditional random error of MPE is significantly smaller than that of CMORPH. The bootstrap-derived sampling distributions of the conditional standard deviation of the random error (only one example is shown in Fig. 5) indicated that the CMORPH error has a sampling distribution that is noticeably wider than that of the MPE product. This indicates the need for larger sample size requirements when dealing with characterization of the CMORPH errors. It is also observed that using MPE as reference instead of $G_{\text{avg}}$ does not significantly change the error standard deviation of CMORPH.

The above analysis indicated fairly high levels of the random error in the CMORPH product ($\sigma_e \sim 7$–$8$ times

---

Fig. 6. (a) POD of CMORPH and MPE conditioned on threshold value of $G_{\text{avg}}$ rain rate and (c) its corresponding missed-rain depth; (b) POFD by CMORPH and MPE conditioned on threshold value of $G_{\text{avg}}$ rain rate and (d) its corresponding falsely detected rain depth. “CMORPH on MPE” refers to the statistics when CMORPH is conditioned on threshold value of MPE.
the mean surface rainfall rate). However, at such fine scales of the product (8 km and hourly), the CMORPH estimates might be significantly dominated with errors due to temporal and spatial navigation and sampling problems. For example, significant differences exist between the cloud-top emission and scattering processes that make the PMW-based estimates, and the actual rainfall that reaches the surface. Time and spatial offsets may exist in the PMW and IR fields and could contribute significantly to the finescale random errors. The fact that it may take about 15–20 min for the rain near cloud top to reach the surface makes it possible that the precipitation could fall into a different $8 \times 8$ km$^2$ pixel (see Turk et al. 2009 for a detailed discussion). Most of these errors should diminish if the CMORPH product is aggregated in time and/or space. It should also be noted that the CMORPH product uses different available microwave instruments (AMSR-E, SSM/I, and AMSU) with different spatial resolutions and temporal availability. Joyce and Xie (2011) showed that the accuracy of PMW estimates that make the final CMORPH product depends largely on the particular PMW instrument and on temporal propagation distance between the sensor overpass and the current estimation time. These factors may actually explain the high levels of random errors observed in this study.

To quantify such potential improvements in the CMORPH accuracy, the random error of the CMOPRH was reassessed at a range of aggregated temporal and spatial scales that start from 1 to 24 h, and 8 km ($\theta = 0.0833^\circ$) to 2$^\circ$, respectively (Figure 7). In addition to calculating the error of standard deviation ($\sigma_e$), the correlation coefficient between the CMORPH estimates and the corresponding surface reference rainfall ($R_{G_{avg}}$) was calculated. As expected, the accuracy of CMORPH improved significantly with the temporal aggregation scale. The correlation between CMORPH and $R_{G_{avg}}$ increases from $\sim 0.42$ at an hourly scale to $\sim 0.65$ at a daily scale. Similarly, the relative standard deviation of error $[\sigma_{(R_{CMORPH}-R_{G_{avg}})}]$ normalized with mean of $R_{G_{avg}}$
decreases from 8.4 times at an hourly scale and reaches its asymptotic value (~2.6) when the aggregation time is increased beyond 24 h. A significant portion of the improvement (~50%) is achieved across the first 6 h of temporal aggregation.

Similarly, the CMORPH product was assessed at different spatial aggregation scales (Fig. 7) using a constant daily temporal scale. A significant gain in performance is observed when going from an 8-km scale (the finest scale of the operationally available CMORPH product) to the next level of product availability (0.25°×8). The correlation between daily CMORPH and RGavg increases from 0.65 to about 0.8, and the relative standard deviation of CMORPH–RGavg differences decreases from ~2.6 to ~1.25 times the mean rainfall rate (RGavg). Further improvement in the product accuracy is still observed beyond the 0.25°×8 scale but at a much reduced rate. Besides temporal or spatial aggregation, applying a time offset between CMORPH and the reference rainfall may also help in gaining improved accuracy if the finescale resolution of the product were to be maintained (Turk et al. 2009).

e. Temporal autocorrelation

Next, the study assesses how the CMORPH product can reproduce the underlying temporal organization present in the surface rainfall. To do so, the temporal autocorrelation coefficients of the CMORPH estimates for different temporal lags were calculated and compared to the corresponding autocorrelations of the surface observation (Gavg) (Figure 8). The figure also shows the corresponding autocorrelation function calculated for the MPE product. As expected, the autocorrelation function decreases rapidly as the time lag increases (Fig. 8). The CMORPH hourly rainfall rates have stronger self-correlation compared to Gavg and MPE for lag times up to 4 h. This behavior might be related to the lifetimes of the underlying clouds, but can be also attributed to the interpolation, or advection, procedure adopted in the CMORPH algorithm to estimate rain rates between two PMW satellite overpass times.

f. Effect of PMW availability

As explained earlier, the CMORPH algorithm is based on propagation and interpolation of most recently available PMW rainfall estimates. Therefore, it is expected that the accuracy of this process will depend on the temporal availability of PMW estimates. Besides rainfall estimates, the 30-min 8-km CMORPH product also contains information on the timestamp of the nearest PMW scan from either the forward or backward propagated PMW rainfall field to current analysis time. The availability of this information, and the high resolution of the product estimates and the reference rain gauge data, allows an examination of the CMORPH product using one of its most useful associated parameters—the timestamp.

To perform such analysis, the full CMORPH–Gavg sample was stratified according to the timestamp parameter into three categories: 1) timestamp 0, which represents instantaneous PMW observations, and timestamp 1, which represents PMW information only 30 min far; 2) timestamps of 2 and 3, which represent CMORPH estimates that use PMW information that is not instantaneous, however not that far either; and 3) timestamps of 4 and greater, which guarantee that the nearest PMW information used in CMORPH is at least 2 h or longer. For each category, we reevaluated some of the performance measures that were considered earlier. For space limitations, the stratified-based analyses are only represented for the three components of the conditional bias (Fig. 9) calculated using Eqs. (3)–(5), and the POD and POFD (Figure 10).

The results of this timestamp-conditioned analysis reveals some interesting features. The behavior of the first two categories (timestamp of 0 or 1, and 2 or 3) is fairly similar to each other and to the overall sample before stratification. The most striking difference is reported for the timestamp ≥4, which corresponds to PMW scans that are 2 h or more from the target analysis time. The total and hit biases of this category reach up to 40% and 20% compared to less than 5% for the other two categories. The missed-rain and false-rain biases reach up to ~30% and 55%, respectively, which are about 1.5 to 2 times worse than the other two categories. The inferior behavior of the timestamp ≥4 is also

![Fig. 8. Temporal autocorrelation of hourly rainfall rate derived for Gavg, MPE, and CMORPH.](image-url)
reiterated in the conditional POD and POFD analysis (Fig. 10) where the deterioration in the CMORPH accuracy persists with the increase of the surface rainfall threshold.

Similar behavior by the stratified samples was also observed for the other statistical metrics used earlier in this study. The large overall and conditional bias values indicate the problematic behavior of the CMOPRH estimates at processing times when PMW scans are not available within 2 h. The significant improvement in the CMORPH accuracy at shorter temporal distances (0.5–1.5) hours is promising for future CMORPH developments in the Global Precipitation Mission (GPM 2008) era with the expected increase in availability and temporal sampling frequency of PMW sensors.

5. Summary and conclusions

The main objective of this study was to provide the user community and algorithm developers with some insight about the accuracy of the CMORPH rainfall product at high resolutions (8 × 8 km² and hourly, daily, and event scales). Observations from a dense cluster of rain gauges over a single 8 × 8 km² CMORPH pixel are used as an evaluation reference. The implications of using a radar-based rainfall product as reference data source for evaluation of CMORPH are also examined. Several error metrics were utilized, such as bias, random error, rain rate probability distribution, and rain detection skill. These error metrics were estimated for the entire sample and by applying various conditioning mechanisms to undertake conditional-based evaluation. The results of the analysis can be summarized as follows:

1) When evaluated over the entire 28-month sample, the CMORPH product shows negligible overall bias (<1%). However, when evaluated on an event temporal scale, the CMORPH product shows significant under- and overestimation biases for about 60% and 40% of the events, respectively. The underestimation bias is more dominant and reaches 100% and more for several events.

2) The CMORPH product has a successful detection skill of 80% for hourly rainfall of 2 mm h⁻¹ and higher. This corresponds to a missed-rainfall bias of about 21% of the total rainfall volume over the entire

![Fig. 9](image1.png)

**Fig. 9.** Stratification of the CMORPH total bias and its three components as a function of PMW nearest future or past overpass timestamp (expressed in increments of half-hour periods).

![Fig. 10](image2.png)

**Fig. 10.** Stratification of the CMORPH missed and falsely detected rain depth as a function of PMW nearest future or past overpass timestamp (expressed in increments of half-hour periods).
study period. Although false detection of rain occurred only <3.5% of the time, it resulted in significant false-rainfall amounts (22%). The significant event-scale biases and the false and missed-rainfall volumes are probably the most alarming CMORPH performance attributes and thus warrant further product development support.

3) The CMORPH bias shows strong dependence on the surface rainfall rate condition. It has a positive bias for surface rainfall rates smaller than 3 mm h\(^{-1}\) and a negative bias for larger rain rates. The underestimation increases almost linearly with the increase of the surface rainfall rate and can be estimated as −20% of the surface rainfall rate; for example, conditional bias is about −5 and −10 mm h\(^{-1}\) when the surface rainfall rate is 10 and 20 mm h\(^{-1}\), respectively.

4) Similar to the conditional bias, the standard deviation of the CMORPH random error shows a clear dependence on the surface rainfall rate condition. It increases from 4 mm h\(^{-1}\) when the surface rate is 2.5 mm h\(^{-1}\) and reaches asymptotic value of 6–7 mm h\(^{-1}\) for surface rate of 10 mm h\(^{-1}\) and larger.

5) A strong scatter between the CMORPH and surface rainfall observations is observed. At an hourly scale, the correlation coefficient is about 0.42 and the standard deviation of the CMORPH random error is more than 8 times the hourly mean rain rate. A rapid gain in accuracy is attained when the CMORPH estimates start to be aggregated in time or space. This indicates the significant contribution of factors such as temporal offsets and spatial navigation problems to the overall random errors of the product, especially when used at its finest resolution.

6) The CMORPH rainfall fields are relatively “smooth” as reflected in the stronger self-correlation levels compared to the surface gauge observations. While this behavior could be related to the underlying cloud lifetimes, it may also be caused by the interpolation procedure used in the CMORPH algorithm to estimate rain rates between any two successive fields of PMW rainfall estimates.

7) The error of the radar-based MPE product should be small enough if it is to be used as reference data to evaluate CMORPH in absence of adequate gauge data. In our study area, and when evaluated against dense rain gauge observations at the same spatial and temporal scale of the CMORPH product (8 km and hourly), the MPE product has a high correlation coefficient (0.9), a high detection skill (99%), and small overall bias levels (<3%). However, on an event scale, the MPE product has significant biases (20%–40% of total event depth). When conditioned on different surface rain rates, the conditional bias and conditional random error of the MPE product are fairly significant; however, they are smaller (about \(\frac{1}{5}\) and \(\frac{1}{2}\), respectively) than those of the CMORPH product. These results should be taken into consideration whenever the MPE product, or other similar products, are used as a reference for evaluating satellite-based products.

8) The use of MPE as reference value instead of gauge data resulted in more favorable performance of CMORPH than its actual accuracy in terms of false-rain detection and less favorable performance in terms of total bias and hit bias. The effect of the selected reference data (MPE versus \(G_{\text{avg}}\)) was not statistically significant in terms of the correlation coefficient, conditional bias, and conditional error standard deviation.

9) Being based on propagation and interpolation of PMW-based rainfall estimates, the performance of the CMORPH product was found to be highly dependent on the temporal availability of the PMW fields. The performance was fairly low when the temporal distance between the analysis time and the most recent available PMW field was 2 h or longer. The performance is much better for the case of 0.5–1.5 h, indicating potential enhancement when the overall availability and temporal frequency of PMW sensors is expected to improve under future satellite missions.

Acknowledgments. The authors acknowledge support from the National Science Foundation (NSF) Grant OISE-0914618, the NASA EPSCoR/BoR DART2 program, and the Louisiana LaSPACE program. The authors thank John Janowiak and Pingping Xie for their efforts in producing and providing CMORPH data and Hiroko Kato Beaudoin and Yelena Yarosh for assistance with data access. Comments and suggestions from anonymous reviewers are acknowledged.

REFERENCES


