Comparisons of IMERG Version 06 Precipitation at and between Passive Microwave Overpasses in the Tropics

MANIKANDAN RAJAGOPAL,a EDWARD ZIPSER,a GEORGE HUFFMAN,b JAMES RUSSELL,a AND JACKSON TANb,c

a University of Utah, Salt Lake City, Utah
b NASA Goddard Space Flight Center, Greenbelt, Maryland
c Universities Space Research Association, Columbia, Maryland

(Manuscript received 22 September 2020, in final form 25 June 2021)

ABSTRACT: The Integrated Multisatellite Retrievals for Global Precipitation Measurement Mission (IMERG) is a global precipitation product that uses precipitation retrievals from the virtual constellation of satellites with passive microwave (PMW) sensors, as available. In the absence of PMW observations, IMERG uses a Kalman filter scheme to morph precipitation from one PMW observation to the next. In this study, an analysis of convective systems observed during the Convective Process Experiment (CPEX) suggests that IMERG precipitation depends more strongly on the availability of PMW observations than previously suspected. Following this evidence, we explore systematic biases in IMERG through bulk statistics. In two CPEX case studies, cloud photographs, pilot’s radar, and infrared imagery suggest that IMERG represents the spatial extent of precipitation relatively well when there is a PMW observation but sometimes produces spurious precipitation areas in the absence of PMW observations. Also, considering an observed convective system as a precipitation object in IMERG, the maximum rain rate peaked during PMW overpasses, with lower values between them. Bulk statistics reveal that these biases occur throughout IMERG Version 06. We find that locations and times without PMW observations have a higher frequency of light precipitation rates and a lower frequency of heavy precipitation rates due to retrieval artifacts. These results reveal deficiencies in the IMERG Kalman filter scheme, which have led to the development of the Scheme for Histogram Adjustment with Ranked Precipitation Estimates in the Neighborhood (SHARPEN; described in a companion paper) that will be applied in the next version of IMERG.

KEYWORDS: Tropics; Mesoscale systems; Precipitation; Satellite observations; Field experiments

1. Introduction

Precipitation has direct impacts on human populations through natural disasters (e.g., Bodenreider et al. 2019; Paul et al. 2019) and by influencing the global economy through agriculture (e.g., Zhang et al. 2012; Sohoulande et al. 2019) and commerce (Steinker et al. 2017; Verstraete et al. 2019). Satellites are crucial for observing atmospheric variables, including precipitation, over remote locations such as oceans and mountains, where surface observations are often scarce. The most detailed precipitation measurement from space is possible through the precipitation radar on board the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al. 1998) satellite and its successor, the Global Precipitation Measurement mission (GPM; Hou et al. 2014) Core Observatory satellite. The cloud radar on board the CloudSat satellite is complementary to the precipitation radars since it is sensitive to light snow and rain (Hayden and Liu 2018). These precipitation and cloud radars have extremely sparse revisit times, so long-term global satellite precipitation products use passive microwave (PMW) sensor observations. These long-term precipitation records help us understand the hydrological cycle and climate change impacts at global and regional scales. These records include satellite precipitation data such as the Global Precipitation Climatology Project (GPCP; Huffman et al. 1997; Adler et al. 2018) and the TRMM Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007), which have regularly been used to study recent climate trends (e.g., Gu et al. 2007; Zhou et al. 2015). Despite their value in climate analyses, both GPCP and TMPA have relatively coarse time-space resolution; TMPA, for example, is available every 3 h with a spatial resolution of 0.25° × 0.25° (Huffman et al. 2007). This relatively coarse temporal resolution makes TMPA less suitable for studying important objects in precipitation fields such as mesoscale convective systems (MCSs), which can produce severe weather events that cause damage through strong winds, floods, and hail (Akaeda et al. 1995; McCollum et al. 1995). Also, MCSs contribute more than 50% of annual rainfall in the tropics (Mohr et al. 1999; Nesbitt et al. 2006; Liu 2011).

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-20-0226.s1.

Corresponding author: Manikandan Rajagopal, mani.rajagopal@utah.edu

DOI: 10.1175/JHM-D-20-0226.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
The Integrated Multisatellite Retrievals for GPM (IMERG), the successor to TMPA, is a global precipitation product at a spatial resolution of 0.1° × 0.1°, available every 30 min, and is therefore more suitable to investigate MCSs and better resolve precipitation events. IMERG follows the pioneering TMPA approach, achieving near-global coverage by taking advantage of a virtual constellation (because their orbits are not coordinated) of low-Earth-orbit (LEO) satellites with PMW sensors. At the locations with PMW observations, IMERG uses the Goddard profiling algorithm (GPROF2017) that converts PMW brightness temperature \( T_b \) to precipitation estimates for various PMW sensors (Kummerow et al. 2015; Randel et al. 2020), except the Sounder for Probing Vertical Profiles of Humidity (SAPHIR), for which the Precipitation Retrieval and Profiling Scheme (PRPS) is employed (Kidd 2018). At locations without PMW observations, IMERG uses a time morphing scheme, based on Kalman filter (Joyce and Xie 2011), to estimate precipitation.

For both weather and climate research, it is vital to understand the strengths and limitations of IMERG precipitation. Past validation studies compared GPROF precipitation estimates for various PMW sensors to the GPM Dual-Frequency Precipitation Radar (DPR), rain gauges, or ground radar. They show that GPROF overestimates light precipitation rates and underestimates heavy precipitation rates (Tan et al. 2018; Kidd et al. 2018; You et al. 2020). In addition to differences in estimates, they found that PMW precipitation has a higher frequency of light rain rates and a lower occurrence of heavy precipitation rates than surface observations. The higher frequency of light precipitation rates may be attributed to the Bayesian nature of the GPROF algorithm, where the observed PMW brightness temperatures are matched to the closest values in a priori database. For each pixel, this produces multiple matches, i.e., multiple surface precipitation rates, including zeros. The Bayesian averaging of multiple values at each pixel results in fewer zeros and a higher frequency of light rain rates. IMERG uses GPROF precipitation after removing unreasonably low rain rates through thresholding and histogram adjustment to combined radar–radiometer precipitation. The past validation studies have evaluated GPROF precipitation estimates with observations, but there are fewer studies (Tan et al. 2016; Maranan et al. 2020) that examined IMERG precipitation estimates from the morphing algorithm, which is the focus of this work.

The studies mentioned above and many others (Gaona et al. 2016; Tan et al. 2017; Sungmin et al. 2017; Sharifi et al. 2016; Ason et al. 2017; Watters et al. 2018; Bytheway et al. 2020; Tapiador et al. 2020; Gowan and Horel 2020) involve point-to-point comparisons of GPROF or IMERG precipitation with other observations at different space and time resolutions. As an alternative to the point-to-point comparison, an object-based approach (e.g., Davis et al. 2006; Johnson et al. 2013) has recently been introduced for evaluating reanalyses, model forecasts, and other global gridded products. This entails identifying objects in the fields of interest and establishing biases or errors based on these objects’ properties. However, object-based approaches to validating IMERG are rare. Cui et al. (2020) compared MCSs tracked in IMERG with MCSs defined using ground-based radar precipitation estimates over the continental United States from 2014 to 2016. They found that MCSs in IMERG had systematically larger precipitation areas and higher precipitation volume than radar-observed objects. To our knowledge, there have not been similar object-based approaches to the validation of IMERG algorithmic components over tropical oceans.

In this work, we use an object-based approach and the aircraft data from two case studies from the Convection Processes Experiment (CPEX) to investigate the representation of MCSs in IMERG. These MCSs represented as IMERG precipitation objects occasionally exhibit an unrealistic change in precipitation area and maximum rain rate in the absence of PMW observations. Since these case studies are too few to provide definitive bias characteristics, we use bulk statistics to show the systematic differences between IMERG precipitation estimates from the GPROF and morphing algorithms.

The rest of the article is structured as follows. Section 2 describes the CPEX field program and the DC-8 aircraft instrumentation with the forward camera and pilot’s radarscope used to compare with IMERG precipitation estimates. Then, we summarize the IMERG V06 Final Run procedure to estimate precipitation both during and, with greater complexity, between PMW overpasses. Section 3 presents the subjective tracking of MCSs and their properties inferred from IMERG for two case studies from the CPEX field program. Then a method for constructing precipitation rate distributions over the tropics from 2001 to 2019 is outlined. Results presented in section 4 show specific differences in IMERG precipitation during and between PMW overpasses, both for two case studies and bulk statistics from 2001 to 2019. Section 5 discusses these findings, and section 6 summarizes the conclusions.

2. Data

a. CPEX aircraft data

CPEX was a National Aeronautics and Space Administration (NASA) sponsored field program conducted in May–June 2017. The DC-8 aircraft carried a range of remote sensing instruments to measure vertical profiles of temperature, humidity, wind, and hydrometeors to study the environmental conditions and the convection. Also, the aircraft had a forward camera, a nadir camera, and a pilot’s radarscope that provided visual references. During its 100 flight hours, the DC-8 investigated MCSs over the Gulf of Mexico, the Caribbean Sea, and the western Atlantic. The forward camera and forward-looking pilot’s radar give a holistic view of each MCS and are more valuable for IMERG validation than the downward-pointing aircraft precipitation radar. This is because the aircraft precipitation radar resolution and its swath are smaller than most PMW sensor footprints, rendering their comparisons less useful. The IMERG animation overlaid with the aircraft track, position, and heading are visually compared with the forward camera and pilot’s radar from multiple CPEX missions. Only the evaluation for MCSs observed on 6 and 10 June 2017, are
presented here. They both highlight the differences in IMERG precipitation during and between PMW overpasses.

b. IMERG Version 06

IMERG precipitation products are available at three different latencies, namely, Early Run (~4 h), Late Run (~14 h), and Final Run (~3.5 months), to cater to the time-sensitivities of different applications. In this study, we use the half-hourly precipitation from the Version 06 Final Run, which is a research-quality product. IMERG uses multiple algorithms/procedures for intercalibration, precipitation estimates, and bias corrections. Also, there exist differences in algorithmic steps between different runs. Since we use the half-hourly Final Run in this study, only its algorithmic components will be summarized here. Information on other runs, daily, and monthly products are available in the technical documentation and algorithm theoretical basis document (Huffman et al. 2020a,b).

IMERG uses observations from a virtual constellation of LEO satellites with PMW sensors to create near-global precipitation products. The final IMERG precipitation estimate is a multistep process starting with the intercalibration of \( T_b \) of various PMW sensors against the TRMM Microwave Imager (TMI) or GPM Microwave Imager (GMI), depending on available periods of coincidence. The GPROF (Version 2017) algorithm uses a Bayesian approach to convert the intercalibrated \( T_b \) of multiple PMW frequencies of a sensor to surface precipitation (Kummerow et al. 2015). The resulting PMW precipitation has a resolution that depends on sensor frequency channels (You et al. 2020, their Table 1). These GPROF precipitation estimates are regridded to a 0.1° × 0.1° spatial grid using nearest-neighbor interpolation. The constellation PMW precipitation rates are calibrated to TMI or GMI precipitation rates through climatological histogram matching. This removes most differences in “precipitation rate frequency distribution” that would otherwise contaminate the IMERG inputs. All the PMW precipitation fields are then dynamically calibrated to that would otherwise contaminate the IMERG inputs. All the precipitation rates are calibrated to TMI or GMI precipitation using nearest-neighbor interpolation. The constellation PMW precipitation estimates are regridded to a 0.1° spatial grid and then backward propagated precipitation from the past PMW observation, 2) forward propagated precipitation from the future PMW observation, and 3) IR precipitation from geostationary satellites at the analysis time. The motion vectors used to propagate precipitation in IMERG Version 06 (Final Run) are computed from a time sequence of the total column water vapor (TOV) fields as analyzed by the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2; Gelaro et al. 2017; Tan et al. 2019). This is a change from the earlier versions of IMERG (and CMORPH-KF) that relied on motion vectors diagnosed from geostationary IR data. The IR precipitation input to the Kalman filter is based on the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS) algorithm. This algorithm converts IR cloud-top \( T_b \) gradients into precipitation estimates using neural networks, trained on coincident IR and PMW observations (Hong et al. 2004; Nguyen et al. 2018). IMERG computes the weighted average of all three sources (forward propagated, backward propagated, and IR precipitation), using Kalman filter weights that are proportional to their respective locally estimated error variance. The Kalman weight for propagated PMW precipitation decreases with increasing propagation time since the PMW observation (Joyce and Xie 2011). Tan et al. (2016) and Maranan et al. (2020) defined three groups in the absence of instantaneous PMW observations: “morph,” “IR-only,” and “morph+IR,” based on whether the Kalman filter uses propagated PMW precipitation, IR precipitation, or both. In this article, we follow Joyce and Xie (2011) and use the alternative definition for “morphing” to refer to Kalman weighted average estimates that include all three groups. In other words, “morphing” here refers to output from the Kalman filter that strives to produce a smooth transition from one PMW observation to the next.

Finally, bias corrections computed using monthly rain gauge data from the Global Precipitation Climatology Center (GPCC; Schneider et al. 2014) are applied to merged precipitation fields from PMW and morphing. We use these gauge calibrated precipitation fields to compute bulk statistics; however, using the uncalibrated precipitation yielded similar results.

In addition to the estimated precipitation, each half-hour data file has variables to identify the source of precipitation estimates (Tan et al. 2016). The variable “HQprecipSource” has a PMW sensor index value for each grid cell, including the value zero when there is no PMW observation. The variable “IRkalmanFilterWeight” has Kalman weights for IR precipitation expressed in percentage, and the remaining percentage is the weight assigned to propagated PMW precipitation. In the half-hour interval immediately before and after a PMW observation, the morphing algorithm weights only the forward and backward propagated PMW precipitation. Hence, the “IRkalmanFilterWeight” is zero in such locations. More than 30 min away from a PMW observation, the morphing algorithm weights together forward propagated, backward propagated, and IR precipitation. The complete list of data variables, description, and index values for various PMW sensors are available in the IMERG documentation (Huffman et al. 2020a,b).

3. Methods

a. MCSs: Tracking and properties

The initial focus of our analyses is two MCSs that the DC-8 aircraft investigated during the CPEX field campaign. At 0900 UTC 6 June 2017, an MCS started as an isolated convective system (Fig. 1). Later, around 1300 UTC, the system connected with the adjacent precipitation area of low rain rates (<4 mm h⁻¹), but the MCS remained distinct and subjectively
identifiable. In the next 6 h, it grew upscale to become a large and intense MCS at 1900 UTC. At about 2200 UTC, the system merged with a precipitation band that extended into the western Atlantic. After this merger, the tracking was stopped because the MCS was indistinguishable from the contiguous precipitation band. Figure 1 shows the 2-h snapshots from the subjective tracking with the object boundary marked by a black contour. Though objective tracking is more desirable than a subjective method, most tracking algorithms define MCSs as contiguous precipitation areas. Such a traditional definition is not suitable for IMERG. Figure 2 from the broader region shows multiple MCSs connected by low precipitation rates to form a single contiguous area that extends thousands of kilometers. This contiguous area cannot be defined as a single MCS.

The MCS from the 6 June case study undergoes multiple mergers and splits. We looked at the IMERG animation repeatedly to discern the precipitation area that can be deemed an MCS. If an embedded cell breaks away and merges with a neighboring MCS, we exclude it from the MCS boundary in previous time steps. Similarly, if a cell joins with the MCS under study, we include it within the object boundary in the previous time steps. The IMERG animation is provided as supplemental material for the reader’s reference. Choosing a different MCS boundary by including or excluding small cells will somewhat affect metrics such as area and rain volume; however, our conclusion remains the same.

On 10 June 2017, the DC-8 examined a small MCS in the western Atlantic near 73 W and 25 N. Manual identification of the MCS boundary in IMERG was more straightforward due to its isolated nature. This MCS was tracked until it dissipated on 11 June at 0200 UTC (Fig. 3).

For both MCSs, defined as IMERG objects, we compute the precipitation area, volumetric rain rate, and maximum rain rate at each half-hour from IMERG precipitation. The volumetric rain rate (mm km$^{-2}$ h$^{-1}$) at each pixel is the product of the precipitation rate and the area. For an MCS, it is the sum over all pixels in the MCS of the volumetric rain rate at each pixel. Although rain volume is a time-integrated variable, we use it interchangeably with the volumetric rain rate. Also, despite the technical difference between “precipitation” and “rain,” we use them interchangeably since all surface precipitation in these cases was rain.

b. Constructing precipitation rate distributions

In section 4c, we examine the frequency distribution of precipitation rates for nearly the full length of the IMERG record and global tropics extending from 30°N to 30°S. Following Tan et al. (2016, 2018) and Sungmin et al. (2017), we use a minimum threshold of 0.1 mm h$^{-1}$ to determine if a grid cell is precipitating. We repeated this analysis for different thresholds such as 0.2 and 0.3 mm h$^{-1}$ (figures not shown), and our conclusions are not sensitive to these choices.
The precipitating grid cells are categorized into different sources based on variables “HQprecipSource” and “IRkalmanFilterWeight,” as discussed in section 2b. Then, a precipitation rate distribution is computed for each source as a probability density function (PDF) (Lamb and Verlinde 2011). The resulting PDF will have the area under the curve equal to 1.0.

In some analyses, IMERG sources are combined into two broad categories—“All PMW” and “Morphing.” The “All PMW” category includes instantaneous precipitation estimates from all the PMW sensors and represents precipitation at locations with PMW observations. The “Morphing” category represents precipitation estimates from the Kalman filter at locations without PMW observations, which could be either propagated PMW or IR, or propagated PMW + IR. We also compute the precipitation rate distributions for these two broad categories.

The long-term statistics are essential to remove temporal variations and detect the systematic differences between PMW precipitation and morphing estimates. To reduce computation, we randomly selected 19,000 half-hour IMERG data files from 2001 to 2019 (with 1000 files per year). The PDF computed for each half-hour file is averaged over 19,000 files to produce a mean distribution. To test if 19,000 files are an adequate sample size to capture the systematic biases, we used a different random seed to select a new set of 19,000 half-hour files. The new sample set had PDFs similar to the original sample set with negligible differences (figure not shown), implying that 19,000 half-hour data were sufficient.

c. Computing precipitation occurrence (%)

In the global tropics, only a percentage of grid cells receive precipitation in a 30-min interval. The tropical precipitation

![Figure 2](image_url)  
**Fig. 2.** The MCS from the CPEX case study on 6 June 2017, marked by a black contour, is connected to the neighboring MCSs by low precipitation rates. If the traditional object definition of contiguous area is used, the MCS will span thousands of kilometers.

![Figure 3](image_url)  
**Fig. 3.** The MCS observed on 10 June 2017, during the CPEX field program, is tracked as the IMERG precipitation object from 1600 to 0200 UTC the next day. The multipanel figure shows 2-h snapshots with MCS boundary marked by a black contour.
occurrence varies temporally, and for the current climate, it will follow a certain distribution with a spread that represents the temporal variability. For each half-hour data, locations between 30°N and 30°S with and without PMW observations are grouped as “PMW” and “Morphing” grid cells, respectively. For each category, we determine precipitation occurrence, expressed as “percentage of grid cells with precipitation.”

We compute precipitation occurrence for randomly selected 19,000 half-hour files from 2001 to 2019 (with 1000 files per year). The histogram of precipitation occurrence for 19,000 half-hour data exhibits the variability in precipitation occurrence and the systematic bias between PMW precipitation and morphing estimates.

IMERG has precipitation rates below $10^{-3}$ mm h$^{-1}$, so the precipitation occurrence depends on the choice of minimum threshold. We apply various subjective thresholds 0.1, 0.5, 1.0, and 2.0 mm h$^{-1}$ to understand their effects on precipitation occurrence.

4. Results

The following subsections describe differences in IMERG precipitation during and between PMW overpasses. First, we demonstrate the differences through two case studies investigated during the CPEX field campaign. Following the CPEX case studies, we use bulk statistics to confirm that the biases are systematic by comparing precipitation rate distributions between locations with and without PMW observations.

a. Case study 1: 6 June 2017

The DC-8 aircraft mission on 6 June 2017 investigated a precipitation system extending from the central Gulf of Mexico northeastward to Florida’s west coast. Figures 4b and 4d show the aircraft’s location, just before making its first penetration of the convective system, about 200 km off Florida’s west coast. The aircraft track, position, and heading are overlaid on IR $T_b$ and IMERG precipitation fields. We use the merged-IR product (Janowiak et al. 2001) available every 30 min centered on the hour and half-hour. For example, in Fig. 4b, the IR $T_b$ at 1900 UTC is from the observation period 1845–1915 UTC. However, the IMERG precipitation (Fig. 4d) is for the period 1830–1900 UTC. We use the IR data at 1900 UTC and not 1830 UTC because they are closest to the aircraft observation at 1848 UTC and PMW overpass at 1853 UTC.

The forward camera has a minimum and maximum horizontal field of view (FOV) of 88° and 115°, respectively. They are marked on IMERG and IR images to reference the area being viewed. The forward camera (Fig. 4a) clearly shows vigorous mature convective towers extending well above the aircraft altitude of approximately 32,000 ft (9800 m), which matches with heavily precipitating grid cells in IMERG. The pilot’s radar also shows a long convective line oriented from
southwest to northeast, similar to the precipitation band on IMERG. Table 1 provides the pilot’s radar’s reflectivity scale and corresponding rainfall rates. IMERG precipitation for this period was retrieved from the Advanced Microwave Scanning Radiometer (AMSR-2), a PMW sensor, with an overpass around 1853 UTC, which is close to the aircraft observation time at 1848 UTC. These aircraft observations indicate that IMERG represents the observed convection qualitatively well at this PMW observation time. However, upon examination of the time series of this MCS, we note discrepancies between the PMW overpasses.

The MCS is subjectively tracked in IMERG and its properties such as precipitation area, rain volume, and maximum rain rate are diagnosed as described in section 3a. Figure 5 presents the time series of these MCS properties with PMW sensor names annotated on the x axis if there is an overpass in that half-hour period. The time series show that the MCS’s maximum precipitation rate peaks near 80 mm h$^{-1}$ several times when there are PMW overpasses. However, it “sags” to about 40 mm h$^{-1}$ shortly before and after PMW overpasses. If we extend the MCS boundary to include adjacent cells, our conclusion does not change because most intense cells in the region are present within the current object boundary. The sagging of the maximum precipitation rate suggests a discrepancy between IMERG estimates from PMW observations and time morphing. If the MCS’s maximum precipitation rate is linearly interpolated between two PMW observations, it would be a straight line connecting the peaks. However, it appears that the Kalman weighted averaging of propagated PMW and IR precipitation somehow creates this sag between the PMW overpasses, a concept that we will explore in section 5.

b. Case study 2: 10 June 2017

On 10 June 2017, the DC-8 aircraft investigated a small MCS located over the western Atlantic, approximately 100 km in each direction (east–west and north–south). The system was probably in its most active growth phase when the DC-8 aircraft approached it at 1914 UTC, at an altitude of 10 km. For the next 3 h, the aircraft penetrated and circumnavigated the MCS as it rapidly matured and entered its decay phase. Figure 6d shows the location of the DC-8 at 2030 UTC, flying at an altitude of 10 km, and its track from the previous hour overlaid on IMERG precipitation. At 2030 UTC, according to IMERG, the aircraft was about to penetrate a mesoscale region of light precipitation after the turn. However, a careful inspection of the forward camera’s cloud photo shows no deep or stratiform clouds but only scattered shallow cumulus. The forward camera has a minimum and maximum diagonal FOV of 100° and 126°, respectively. This is more relevant than horizontal FOV since the aircraft was making a turn. These diagonal FOVs are marked on the IMERG and IR images (Figs. 6b,d) as a reference to the area that the camera is observing. In the cloud photo, the horizon is approximately 356 km away from the aircraft’s location for the flight altitude of 10 km, limiting the area viewed by the forward camera. Though the camera’s FOV is limited, it is evident from the cloud photo that the area ahead does not have widespread precipitation. The 1-min camera video (provided as supplemental material) covers a much larger area that clearly shows the scattered shallow cumulus over this region as the camera’s FOV swept through it when the DC-8 made the turn. Also, the pilot’s radar (Fig. 6c) shows no widespread reflectivity other than a small echo ahead of the aircraft, possibly from an isolated cumulus. Since reflectivity values below 20 dBZ would mean very light rain or none (Table 1), we cannot rely solely on the radar. On the IR image (Fig. 6b) from 2030 UTC, which is closest to the aircraft observation time, we see no cold cloud top to the west of the MCS, where IMERG shows light precipitation. These pieces of evidence strongly suggest that the widespread light precipitation rates in IMERG on the west side of the MCS are spurious.

For this small MCS, Fig. 7 displays the time series of the precipitation area, rain volume, and maximum rain rate diagnosed from IMERG. In the half-hour period from 2000 to 2030 UTC, the IMERG precipitation area jumped to 3 times its previous value. This sudden growth comes from a spurious precipitation area, identified as such from the aircraft photo, pilot’s radar, and IR. To trace the origin of this false precipitation, we examined the source of IMERG precipitation estimates. Figure 8 presents the MCS evolution from 1930 through

---

**Table 1. DC-8 aircraft pilot’s radar reflectivity codes (for signal gain set at calibration) based on IntuVue RDR-4000 Weather Radar Pilot’s Guide. N/A indicates “not applicable.”**

<table>
<thead>
<tr>
<th>Color</th>
<th>Returns</th>
<th>Reflectivity (dBZ)</th>
<th>Rainfall rate (mm h$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>Very light, or none</td>
<td>&lt;20</td>
<td>&lt;0.7</td>
</tr>
<tr>
<td>Green</td>
<td>Light</td>
<td>20–30</td>
<td>0.7–4</td>
</tr>
<tr>
<td>Yellow</td>
<td>Medium</td>
<td>30–40</td>
<td>4–12</td>
</tr>
<tr>
<td>Red</td>
<td>Strong</td>
<td>≥40</td>
<td>&gt;12</td>
</tr>
<tr>
<td>Magenta</td>
<td>Turbulence</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

---
2300 UTC with the sources annotated in blue font. There were PMW observations at 1930 and 2300 UTC from a Microwave Humidity Sensor (MHS) and Special Sensor Microwave Imager/Sounder (SSMIS), respectively. At the time of the aircraft observation (1830 UTC), the precipitation estimates were from the weighted average of propagated PMW and IR. Hence this spurious precipitation is likely an artifact from the morphing algorithm, as will be discussed in section 5.

FIG. 6. Images from the CPEX mission on 10 June 2017 around 2030 UTC: (a) the DC-8 aircraft’s forward camera, (b) Merged-IR data at 1830 UTC, (c) the DC-8 aircraft pilot’s radar with range circles at 50 and 100 n mi. Note that the color bar in Fig. 6c differs from Fig. 4c due to different ambient lighting when recording them, and (d) IMERG precipitation for the half-hour period 2030–2100 UTC. Infrared and IMERG images are overlaid with the DC-8 track (red and black line, respectively). The red/black dot represents the DC-8 aircraft location, the arrow points in the direction of the aircraft heading, and the two dotted lines represent the minimum (100°) and maximum (126°) diagonal field of view (FOV) for the forward camera.

FIG. 7. Time series of the IMERG precipitation area, rain volume, and maximum rain rate of the MCS observed on the 10 June 2017 CPEX mission. PMW sensor names are annotated on the x axis at their respective half-hour overpass interval.
The precipitation area for the CPEX case study on 6 June (Fig. 5) also shows a sudden growth during half-hours, starting at 1600 and 2100 UTC, immediately after PMW overpasses. If we chose a different MCS boundary by including or excluding small neighboring cells, it affects the metrics. However, the area would still grow unrealistically, as clearly noticeable at 1600 UTC in IMERG animation (provided as supplemental material). For both MCSs, if the precipitation areas were linearly interpolated between PMW overpasses, these sudden growths would not exist, and the MCS size would appear to grow more reasonably.

In both CPEX case studies, the unrealistic growth is not always present after a PMW observation. Similarly, the maximum precipitation rate for the 10 June case (Fig. 7) does not sag between PMW observations from 1800 to 2000 UTC. This contrasts with the CPEX case on 6 June 2017, where the maximum precipitation rate peaked during most PMW observations and sagged between overpass times (Fig. 5). Thus, these examples also suggest that the morphing issues uncovered are not uniformly present.

c. Precipitation rate distribution: PMW versus morphing

Given the apparent limitations in IMERG Version 06 revealed by this analysis of two MCSs from the CPEX field program, we explore the systematic biases in IMERG precipitation using bulk statistics. Specifically, we look at the precipitation rate distribution (as a probability density function) of PMW observations and morphing over a wider region from 30°N to 30°S and for an extended period of 19 years. Our initial analysis showed that TMI and GMI exhibit differences in their precipitation rate distribution, so the data are grouped into TRMM (2001–13) and GPM (2015–19) periods. In the year 2014, PMW sensors are intercalibrated to TMI for dates before June 1st and to GMI for the rest of the year, so 2014 is excluded to avoid mixing the calibrations. The PDF differences between TMI and GMI are not analyzed further, since this is beyond the scope of this study.

The precipitation rate distributions from both the TRMM and GPM periods (Fig. 9) show that differences between PMW sensors are very small as expected, since they are intercalibrated to TMI or GMI histograms. However, differences between distributions for PMW and morphed precipitation stand out compared to differences between individual sensors. We can infer from Fig. 9 that the morphing estimates have a higher frequency of light precipitation rates and a lower frequency of heavy precipitation rates relative to PMW estimates. This conclusion is similar to the result from the CPEX mission on 10 June 2017, where the small MCS observed at 2030 UTC had spurious precipitation areas with light precipitation rates.

Another important conclusion from the CPEX case study on 6 June 2017 (Fig. 5) was that the MCS’s maximum precipitation rate decreased further and further as the time interval from the PMW overpass increased. Figure 10 presents precipitation rate distributions for time intervals away from the PMW overpass (both past and future). The time interval “30 min away” represents time morphing estimates from the past PMW observations propagated forward to the next half-hour and future PMW observations propagated backward to the previous half-hour. Similar logic can be extended to other time intervals. The distributions show that the frequency of light precipitation increases slightly with increasing time intervals. In contrast, there is a prominent decrease in the frequency of heavy
precipitation rates, as one moves further away from the PMW overpass times. This conclusion is consistent with the observed sag in the maximum precipitation rate away from PMW overpasses in the CPEX case study on 6 June 2017.

d. Precipitation occurrence: PMW versus morphing

Figure 11 displays histograms of PMW and morphing precipitation occurrence from 19,000 half-hour data for different minimum thresholds (0.1, 0.5, 1.0, and 2.0 mm h\(^{-1}\)). The spatial coverage of PMW sensors changes with time as they move over various regions on Earth. This introduces spatial variability and sampling variability in the tropical precipitation occurrence. Therefore, the spread in the histogram represents all three variabilities: temporal, spatial, and PMW coverage. The statistic “precipitation occurrence” also can be interpreted as the sum of frequencies of various precipitation rate bins in Fig. 9 expressed as a fraction of total grid cells in each category. This statistic is more intuitive and highlights the net effect of morphing on tropical precipitation occurrence.

For a minimum threshold of 0.1 mm h\(^{-1}\), Fig. 11a shows that locations with PMW observations have a 6.56% median precipitation occurrence, whereas locations with estimates from morphing have an 8.86% median precipitation occurrence. This difference is considerable and implies that morphing creates 35% more precipitation for a same-sized region. Using a nonparametric significance test (permutation test; Wilks 2011), we found that the difference in median precipitation occurrence between PMW and morphing is statistically significant at a 99% confidence interval. This suggests that the difference in precipitation occurrence between PMW and morphing is substantially larger than the variabilities or spread in the histogram. The higher precipitation occurrence in morphing estimates implies that it produces spurious precipitation cells, consistent with the CPEX case study from 10 June 2017.

When the minimum threshold is increased to 0.5 mm h\(^{-1}\), the absolute and relative difference in precipitation occurrence between PMW and morphing decreases sharply (Fig. 11b).
Results across the range of minimum thresholds (0.1, 0.5, 1.0, and 2.0 mm h\(^{-1}\)) show that light rain rates contribute to the significant difference in precipitation occurrence between PMW and morphing.

5. Discussion

The specific instances of spurious precipitation in the CPEX case can be traced back to the time morphing algorithm. In general, a precipitation system observed at two different times will often have different sizes and shapes of precipitation areas. At intermediate times, the morphing algorithm propagates these PMW observations in the forward and backward direction using motion vectors with “little change” in their shape and size. Even if the motion vectors place the centroid of a forward and backward propagated system on top of each other, the weighted averaging will create a larger area if the shapes are different. In addition, the increase in precipitation area may be exacerbated if the motion vectors have biases that offset the centroids of the forward and backward propagated systems.

The overlapping of different sizes and shapes from two observation times can explain the unrealistic growth sometimes observed in CPEX case studies. The precipitation area created from overlapping shapes is larger at intermediate times and suddenly changes to a smaller area at the closest PMW observation. This might show up as a sudden decay or growth in precipitation area immediately before or after PMW observations, as observed in the CPEX case studies.

Though overlaps produce larger areas at intermediate times, the propagated precipitation rates are reduced by their Kalman weights that are roughly inversely proportional to propagation time (Joyce and Xie 2011). Some grid cells in the intersecting region of forward and backward propagated precipitation areas receive estimates from both. However, other cells receive estimates from only one. Though the morphing algorithm combines propagated precipitation with IR precipitation, the Kalman averaging will more likely produce light rain in grid cells with the contribution from only one input source. For example, say that the weights for forward propagated, backward propagated, and IR in a particular grid cell are 0.26, 0.34, and 0.40, respectively. If the grid cell’s only nonzero contribution is forward propagated precipitation of 2 mm h\(^{-1}\), the Kalman averaging would estimate a value of 0.52 mm h\(^{-1}\) (0.26 \times 2 + 0.34 \times 0 + 0.40 \times 0).

Figure 8 shows the MCS from the CPEX case study on 10 June 2017, exhibiting different shapes and sizes during PMW observations at 1930 and 2300 UTC from an MHS and an SSMIS, respectively. At the intermediate times, the MCS takes a shape and size that is an overlap of these “PMW observations” propagated to the respective analysis time. At 2030 UTC, the west side of the MCS must be provided by the backward propagated precipitation because the IR image (Fig. 6b) shows no deep convection, and the forward propagated PMW observation from 1930 UTC has no rainfall over this region. The intensity of backward propagation of PMW precipitation fell from 2 to 0.5 mm h\(^{-1}\) on the west side of the MCS due to decreasing Kalman weights and lack of other sources. The aircraft photos and IR image in Fig. 6 show that the IMERG precipitation on the west side of the MCS is indeed spurious at 2030 UTC.

IMERG, with its high temporal resolution, is more suitable than the past gridded precipitation data to track MCSs.
as precipitation objects and study their upscale growth. Understanding the IMERG biases presented in this article might help define MCSs and their upscale growth, more appropriately. This study shows that using rain volume to define the growth and decay of an MCS yields stable results, based on the CPEX case studies. In contrast, the use of precipitation area to define the MCS life cycle is sometimes contaminated by spurious precipitation. This can be remedied to some extent by using a low-end cutoff precipitation threshold. Figures 9 and 11 suggest that a precipitation map thresholded at or above 0.5 mm h$^{-1}$ has minimal spurious rain areas. However, one should use the cutoff cautiously because it merely sets the low end of the PDF to zero, but it cannot recover the underestimation at the high end. Taken together, a cutoff will cause the overall mean and rain volume of the morphed estimates to be underestimated. The problem is more fundamental and affects the entire PDF.

This line of reasoning was developed in conversations with the IMERG team, leading them to develop a new algorithm called Scheme for Histogram Adjustment with Ranked Precipitation Estimates Neighborhood (SHARPEN) that resolves some of the biases reported here. This algorithm restores the precipitation rate distribution of morphing estimates to a distribution similar to the PMW estimates using a quantile mapping based approach. SHARPEN will most likely be implemented in the upcoming release of IMERG Version 07, and is described in the companion paper (Tan et al. 2021). Preliminary testing (not shown) demonstrated that SHARPEN reduced the spurious precipitation considerably for the 10 June CPEX case study.

6. Conclusions

In this study, an object-oriented approach is used to compare MCS properties, such as precipitation area, rain volume, and maximum precipitation rate during and between PMW observations. Two case studies are used to investigate and illustrate potential issues with the time morphing in IMERG before these issues are confirmed to be prevalent throughout IMERG using bulk statistics.

In the case studies, IMERG precipitation is evaluated against DC-8 aircraft observations from the NASA CPEX field program. The aircraft photos, pilot radar, and IR image indicate that at times of PMW observations, IMERG precipitation area is similar to aircraft observations. However, in the absence of PMW observations, IMERG appeared to have spurious precipitation areas. For the small MCS of 10 June 2017, the precipitation area grew unrealistically in the next 30 min after the PMW observation at 1930 UTC. We found that this sudden growth and the spurious precipitation originates from the morphing algorithm. Another finding was that for the intense, large MCS of 6 June 2017, the maximum precipitation rate peaked at most PMW overpass times, but sagged between those observation times. Although these are limited case studies, the evidence from CPEX guided us to investigate systematic biases in the IMERG’s time morphing algorithm. The analysis of precipitation rate distribution and precipitation occurrence shows that there are systematic differences between precipitation estimates from PMW observations and time morphing. When PMW observations are absent, IMERG has a higher frequency of light precipitation and a lower occurrence of heavy precipitation rates. The high occurrence of light precipitation rates may sometimes produce large precipitation areas.

In our work, we highlight biases in time morphed precipitation estimates compared to PMW precipitation. However, PMW precipitation estimates from GPROF have their own biases as highlighted in past studies. An ideal validation would be to compare morphing precipitation estimates with surface observations. Nevertheless, in the absence of global surface observations, comparing IMERG precipitation estimates for PMW observation and time morphing is still valuable to bring out their algorithmic differences. Identifying these morphing biases helped the IMERG team improve the next version of IMERG by bringing the morphing estimates more into line with the PMW (and PMW-calibrated IR) estimates. The end users of IMERG Version 06 should be cautious about the higher occurrence of light precipitation and lower precipitation intensity in the absence of PMW observations.

Acknowledgments. We appreciate the skill and dedication of the pilots, flight crews, instrument teams, and NASA supporting staff that made the CPEX field program successful. We are grateful for discussions on statistical techniques with Prof. Courtney Strong and Prof. Gannet Hallar at the Department of Atmospheric Sciences, University of Utah. We also thank the three anonymous reviewers whose comments and suggestions helped refine this manuscript. We appreciate the computing and storage support from the Center for High-Performance Computing (CHPC) at the University of Utah. This research was supported by NASA Grant NNX17AG74G (Univ. of Utah) and the GPM mission (G. Huffman) under the leadership of Ramesh Kakar and Gail Jackson.

Data availability statement. The aircraft track from the NASA’s CPEX field program flights is accessible at https://tcis.jpl.nasa.gov/data/cpex/trck/. The aircraft’s forward camera and pilots’ radar videos from the CPEX program are available at https://asp-archive.arc.nasa.gov/CPEX/Video/. IMERG is a NASA’s global satellite precipitation product available at different time resolution and latencies. IMERG Version 06 (Final Run) used in this study can be downloaded from the https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary?keywords=%22IMERG%20final%22. The web interface allows a user to choose the appropriate period (2001–19) and the domain (30°N–30°S).

REFERENCES


