
Mengye Chen 1, Soumaya Nabih 1,2, Noah S. Brauer 3, Shang Gao 1,* 4, Zhen Hong 1, Randall L. Kolar 1 and Yang Hong 1,*

1 Hydrometeorology and Remote Sensing Laboratory, School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, OK 73109, USA; mchen15@ou.edu (M.C.); Soumaya.Nabih-1@ou.edu (S.N.); Zhen.Hong-1@ou.edu (Z.H.); kolar@ou.edu (R.L.K.)
2 Department of Georesources and Environment, Sidi Mohamed Ben Abdellah University, Fez 32000, Morocco
3 Advanced Radar Research Center, School of Meteorology, University of Oklahoma, Norman, OK 73109, USA; nbrauer@ou.edu
4 NOAA National Severe Storm Laboratory, Norman, OK 73109, USA; jj.gourley@noaa.gov
* Correspondence: shang.gao@ou.edu (S.G.); yanghong@ou.edu (Y.H.);
Tel.: +1-949-202-9827 (S.G.); +1-405-996-8128 (Y.H.)

Received: 28 December 2019; Accepted: 29 January 2020; Published: 1 February 2020

Abstract: A new generation of precipitation measurement products has emerged, and their performances have gained much attention from the scientific community, such as the Multi-Radar Multi-Sensor system (MRMS) from the National Severe Storm Laboratory (NSSL) and the Global Precipitation Measurement Mission (GPM) from the National Aeronautics and Space Administration (NASA). This study statistically evaluated the MRMS and GPM products and investigated their cascading hydrological response in August of 2017, when Hurricane Harvey brought historical and record-breaking precipitation to the Gulf Coast (>1500 mm), causing 107 fatalities along with about USD 125 billion worth of damage. Rain-gauge observations from Harris County Flood Control District (HCFCD) and stream-gauge measurements by the United States Geological Survey (USGS) were used as ground truths to evaluate MRMS, GPM and National Centers for Environmental Prediction (NCEP) gauge-only data by using statistical metrics and hydrological simulations using the Ensemble Framework for Flash Flooding Forecast (EF5) model. The results indicate that remote sensing technologies can accurately detect and estimate the unprecedented precipitation event with their near-real-time products, and all precipitation products produced good hydrological simulations, where the Nash–Sutcliffe model efficiency coefficients (NSCE) were close to 0.9 for both the MRMS and GPM products. With the timeliness and seamless coverage of MRMS and GPM, the study also demonstrated the capability and efficiency of the EF5 framework for flash flood modeling over the United States and potentially additional international domains.

Keywords: Hurricane Harvey; MRMS and IMERG; EF5 hydrological simulation

1. Introduction

Floods are believed to be among the most hazardous and frequent natural disasters to human society [1–4]. Flooding can generally damage infrastructure, cost lives, and even cause further water contamination as well as waterborne diseases [5]. In particular, floods over urbanized area are more likely to cause fatalities and severe economic damage because of the population density and developed...
infrastructure, which leads to the intensification of the meteorological extremes [6,7] and increased surface runoff peaks [5]. Globally, the Gulf Coast of North America is one of many places that is heavily affected by tropical storms and their cascading floods in an urbanized area [8]. On August 25th, 2017, Hurricane Harvey made its first landfall at the northern end of San Jose Island, TX. Since then, Harvey stalled over the greater Houston area and produced over 1500 mm of rain in 4 days, which set the US record of total precipitation since the 1880s, when the reliable rainfall records started [9]. During this event, southeast Texas received 20 to 30 trillion tons of water with a return period exceeding 9000 years at some locations [10], interconnected the Colorado River and San Bernard River overland, and caused unprecedented flooding. Hurricane Harvey was estimated to cause about USD 125 billion worth of damage and 107 fatalities, and 127 flash flood warnings were issued during the event [11]. As much as technology has advanced, society is still searching for tools to improve prediction and mitigate the damage from floods.

Over the past few decades, the scientific community has made great improvements in the capacity of flood modeling by combining climate models, weather models, hydrological models, river models, and hydrodynamic models [12]. The applications of flood modeling vary from flood risk assessment and mapping [13] to flood damage assessment [14], real-time flood forecasting [15], engineering for flood prevention [16], post-flood river system hydrology [17], soil and riverbank erosion [18], catchment hydrology [19] and floodplain ecology [20]. Generally, such applications require considering both the acceptable predictive accuracy and also high spatiotemporal resolution while balancing the computational efficiency for real-time operations. For real-time simulations and forecasts, faster run times, and data assimilation are required to provide reliable results [21]. Recent advances in global and continental hydrological models have shown very promising results thus far, which have provided critical information regarding surface runoff, streamflow, soil moisture, soil infiltration and evapotranspiration [22–28]. One example is the application of the Coupled Routing and Excess Storage model embedded within the Ensemble Framework for Flash Flood Forecasting (CREST-EF5) framework [15,29,30]. The CREST-EF5 framework integrates the distributed hydrological model, CREST, and 1D Kinematic Wave routing to simulate multiple excessive precipitation-triggered flash-flooding events in Oklahoma City and Houston at a continental scale [15]. As errors originally contained in the precipitation forcibly propagate through the hydrological model [31], the accuracy of precipitation datasets is also vitally crucial for hydrological modeling performance.

There are three common precipitation observations in the modern world—rain gauges, weather radars, and satellite-based remote sensing technologies [32]. Rain gauge measurement is traditionally the most straightforward in situ method to estimation the surface precipitation, which, for decades, has been regarded as the closest approximation to the true value at a point [33–36]. Many efforts have been made to interpolate the rain gauge data into a distributed precipitation field, and many versions of the optimal rainfall estimation procedures have been adapted by National Center for Environmental Prediction (NCEP) and National Weather Service (NWS) [37]. However, the rain gauge network density varies spatially and is low over many developing regions [38]. Rather than estimating rainfall at a point in space, weather radar networks provide Quantitative Precipitation Estimates (QPE) covering much larger spatial domains (ranges up to 230 km), at spatial resolutions of the order of 1 km² for each pixel. Since the 1990s, the Next Generation Weather Radar (NEXRAD) WSR-88D system has been improved and utilized, such as the dual-polarization capabilities in 2010 [39], for more advanced precipitation products including the MRMS system. The MRMS system integrates data from over 180 operational NEXRAD radars, over 7000 hourly rain gauges from the Hydrometeorology Automated Data System (HADS), the hourly High Resolution Rapid Refresh model analysis data and precipitation climatology [40]; it seamlessly covers the conterminous United States (CONUS) and Southern Canada at 1 km spatial resolution and a two minute temporal resolution using sophisticated algorithms and supplemental input data from ground gauges and environmental models [41]. However, the MRMS radar network still potentially suffers from radar miscalibration, reduced low-level coverage in mountainous areas and errors in the QPE algorithms. Earth observation satellites provide the potential
Remote Sens. 2020, 12, 445 to estimate precipitation on a global scale [42]. One such unprecedented effort, the NASA Global Precipitation Measurement (GPM) mission was launched in 2014 by building upon the success of previous Tropical Rainfall Measuring Mission (TRMM) from 1997 [43]. To date, the GPM mission has used the Integrated Multisatellite Retrievals for GPM (IMERG) algorithm to generate the quasi-global precipitation products at 0.1 by 0.1 arc-degree spatial resolution and 30 min temporal resolution [44,45].

Today, given the availability of the above mentioned three precipitation data sources, the scientific community has exerted efforts on various precipitation product evaluations and intercomparisons, with particular focus over complex terrains and extreme events including the 2017 Hurricane Harvey event on the Mexico Gulf Coast [9,11,46–48]. Omaranian et al. [49] compared the GPM IMERG final run precipitation estimates with NCEP Stage IV radar QPE and indicated that GPM IMERG could capture the pattern and trace the storm, but significantly overestimated the precipitation amount. Hayatbini et al. [50] investigated the improved method for Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks- Cloud Classification System (PERSIANN-CCS) to match the cloud detection of GPM during Hurricane Harvey, which increases the possibility for PERSIANN-CCS to accurately detect extreme precipitation amount. Kao et al. [51] again studied NCEP Stage IV QPE data for Hurricane Harvey using Probable Maximum Precipitation (PMP) estimation methods, and the study suggested a possible link between the extreme precipitation event and global climate change. Recently, a NASA report presented a precipitation estimations comparison between MRMS QPE and GPM IMERG products for Hurricane Harvey. The results indicate that GPM IMERG had a coherent difference from MRMS QPE during Hurricane Harvey, where it underestimated precipitation in the storm core but overestimated it in the outer rainbands [52]. All studies used different precipitation data as the ground truth and yielded various conclusions. Thus, it is necessary to evaluate the precipitation products based on high-density ground reference network and also further investigate how well they perform in a hydrologic context during this extreme event.

The overarching goal of this study is to investigate which precipitation product can better represent the true surface precipitation during the extreme event and further capture its cascading hydrological responses using a very high-density ground gauge network and an operational hydrological modeling framework. The specific objectives are to (a) statistically compare interpolated rain gauge, MRMS QPE and GPM IMERG precipitation products with the local independently managed rain gauge precipitation observation in Harris County, TX and (b) examine the corresponding hydrological response between interpolated rain gauge data, MRMS QPE and GPM IMERG products as forcing data for the hydrological simulation over the Spring Basin in the northern part of Harris County, TX. The study aims to answer the following research questions: (1) Is this type of extreme event detectable and quantifiable using remote sensing technology? (2) Can the hydrological model capture the extreme responses? (3) Which precipitation product performs better during extreme events?

This paper is organized as follows. Section 2 describes the study area, data used in this study, a short description of Ensemble Framework for Flash Flood Forecasting (EF5) hydrological model, and methodology. Section 3 inter-compares the MRMS, GPM IMERG, NCEP gauge-only precipitation products with Harris county rain gauge data and evaluates the above precipitation products using the EF5 modeling framework. Section 4 concludes the study and proposes future directions.

2. Materials and Methods

2.1. Study Area

Figure 1 displays the impact area of Hurricane Harvey, Harris County, TX, and Spring Creek Basin. Despite the 79,000 square kilometers of impacted area by Hurricane Harvey, Harris County, TX was the most impacted, as almost half of the casualties from this event were from this area [53]. It is also the third most populated county in the USA, with 4.65 million people, and has an area of 4602 km², where the dense population leads to the vulnerability to flood extremes and often relates to fatalities and significant economic loss. There are 147 rain gauges managed by the Harris County
Flood Control District (HCFCD) and Hydrometeorological Automated Data System (HADS) from NWS in Harris County, which provided the 1-hour precipitation accumulation data for the study.

To further study the hydrological responses, Spring Basin was selected as the study watershed. Spring Basin is located at the northern end of Harris County, which contains four major rivers: Spring Creek, Willow Creek, Little Cypress Creek, and Cypress Creek, covering over 1960 km$^2$ before entering Lake Houston. The elevation of Spring Basin ranges from 13 to 136 meters above sea level, with an average of 61 m. The slope varies from 0° to 39°, with an average of 1°. Spring Basin has urban structures concentrated at the southern and eastern part of the basin, agricultural ranches in the southwest portion, and forest in the northwest and middle of the basin along Spring Creek.

2.2. Data

2.2.1. Precipitation Data

Five precipitation products (Table 1) were evaluated by the Harris County-managed HADS and HCFCD rain gauge observation data from 0:00 Central State Time (CDT) August 25th, 2017 to 23:00 CDT August 31st, 2017, with a total of 168 hourly time-steps. For the precipitation product evaluation, the ground rain-gauge should not overlap with those used by NSSL and NWS in developing or correcting the QPE estimates. Therefore, the HADS rain gauges (green dots in Figure 1) were removed from all analysis in this study. After quality control and HADS gauges removal, only 99 out of 147 rain gauge observations remained, where the gauges that had more than 60 out of 168 time-steps with continuous stationary values or NA values were eliminated. We obtained the NCEP and Environmental Modeling Center (EMC) national interpolated rain gauge-only hourly precipitation data from Earth Observing Laboratory (EOL) data archive (https://data.eol.ucar.edu/dataset/21.004) from August 15th to September 18th, 2017. This dataset utilizes measurements of about 3000 rain gauges across the CONUS [54] and uses the optimal estimation of rainfall fields methods to interpolate into 4 by 4 km gridded hourly precipitation data [37].
Table 1. Summary of characteristics of the precipitation products in this study.

<table>
<thead>
<tr>
<th>ID</th>
<th>Products</th>
<th>Spatial Resolution</th>
<th>Temporal Resolution</th>
<th>Median (mm/hr)</th>
<th>Mean (mm/hr)</th>
<th>Maximum (mm/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBS</td>
<td>HCFCD rain gauge observation</td>
<td>Point data</td>
<td>1 h</td>
<td>2.03</td>
<td>7.33</td>
<td>171.70</td>
</tr>
<tr>
<td>NCEP</td>
<td>NCEP hourly gauge only</td>
<td>4 km</td>
<td>1 h</td>
<td>1.00</td>
<td>4.66</td>
<td>65.90</td>
</tr>
<tr>
<td>MRMS QPE</td>
<td>MRMS radar based QPE</td>
<td>0.01°</td>
<td>2 min</td>
<td>2.61</td>
<td>8.07</td>
<td>124.28</td>
</tr>
<tr>
<td>MRMS Corr</td>
<td>MRMS 1-r gauge bias-corrected</td>
<td>0.01°</td>
<td>1 h</td>
<td>1.90</td>
<td>6.70</td>
<td>105.60</td>
</tr>
<tr>
<td>V06AUncal</td>
<td>GPM IMERG satellite-based</td>
<td>0.1°</td>
<td>30 min</td>
<td>3.82</td>
<td>6.55</td>
<td>66.80</td>
</tr>
<tr>
<td>V06ACal</td>
<td>GPM IMERG gauge calibrated</td>
<td>0.1°</td>
<td>30 min</td>
<td>3.34</td>
<td>5.68</td>
<td>59.58</td>
</tr>
</tbody>
</table>

Two MRMS precipitation products from April 1st to September 30th, 2017, radar-based QPE (PCP_RATE), and 1-hour gauge bias-corrected radar precipitation accumulations (Q3GC_SHSR_1H), were obtained from the Iowa Environmental Mesonet NWS data archive (https://mesonet.agron.iastate.edu/nws/). PCP_RATE is the radar-based MRMS product that uses multiple R-Z relationships and is derived from MRMS Seamless Hybrid Scan Reflectivity (SGSR), which has a temporal resolution of 2 min and 0.01 by 0.01 arc-degree spatial resolution in real-time. Q3GC_SHSR_1H is the CoCoRaHS rain-gauge-corrected 1-h radar QPE accumulation using a three steps method, which has the temporal resolution of 1 h and 1 km² spatial resolution with a 1.5-h latency [41].

Two GPM IMERG Version 6 (V6) final-run precipitation products [55] from August 15th to September 15th, 2017, PrecipitationUncal (V06AUncal), and Precipitation Cal (V06ACal), were obtained from the NASA GES DISC data archive (https://disc.gsfc.nasa.gov/). The GPM IMERG system runs twice in near-real-time to produce early run and late run results, where the early run has the morphing scheme only propagated forward, and the late run has the morphing scheme applied both forward and backward [45]. The IMERG final run has a 3.5-month latency, where the uncalibrated precipitation product (V06AUncal) is close to IMERG late run and then calibrated with the local rain gauge data to generate the calibrated precipitation product (V06ACal) [56]. Both IMERG datasets have a 30-min temporal resolution and 0.1 by 0.1 arc-degree spatial resolution.

Before any analysis, all precipitation products were aggregated or interpolated into 4 by 4 km spatial resolution and hourly temporal resolution to make all data comparable. The MRMS product family has a higher spatial and temporal resolution, so MRMS QPE and MRMS Corr data were aggregated using an arithmetic mean and the 30-min data were summed to produce hourly precipitation accumulation. The GPM IMERG product family has lower spatial resolution, so V06AUncal, and V06ACal data were interpolated using the bilinear method and then aggregated into hourly time steps using an arithmetic mean.

2.2.2. Stream Gauge and Geographic Data

Five U.S. Geological Survey (USGS) stream gauges, representing the upper, middle, and downstream branches of Spring and Cypress Creek, were selected to validate and calibrate the hydrological modeling process (Figure 1). The 15-min streamflow data of each gauge from April 1st to September 30th, 2017 were obtained from USGS National Water Information System (https://waterdata.usgs.gov/nwis). A high-resolution (15 arc second) hydrologically conditioned Digital Elevation Model (DEM), Flow Direction (FDR), Flow Accumulation (FAA), and major river network data were obtained from the HydroSHEDS (Lehner et al. 2008, https://www.hydrosheds.org/). The potential evapotranspiration (PET) data used in this study were from USGS Famine Early Warning Systems Network (FEWS NET, https://earlywarning.usgs.gov/fews). The daily 1 by 1 arc-degree PET data were calculated from the Global Data Assimilation System (GDAS) using the Penman-Monteith method [57]. The US landcover data were obtained from Multi-Resolution Land Characteristics Consortium (MRLC, https://www.mrlc.gov/) and the 1 by 1 km EF5 parameters in CONUS were from the previous study by Vergara et al. [58].
2.3. EF5 Modeling Framework and Hydrological Evaluation Method

EF5 is a framework built on multiple hydrological model cores including the Coupled Routing and Excess Storage (CREST) model version 2.0, co-developed by the University of Oklahoma and NASA Applied Science Team [29], and its grid-based water balance component is coupled with the kinematic wave water routing model [59]. EF5 supports multiple water balance methods and comes with an automatic calibration module [30]. EF5 was adapted as an operational tool across the NWS for flash flood forecasting by local NWS Forecast Offices in the Flooded Locations and Simulated Hydrographs Project (FLASH) [15]. The current modeling research was modified from the basic implementation from FLASH and was used here to evaluate the hydrological responses of remotely sensed observations and rain gauge interpolated precipitation products.

2.4. Statistical Metrics

Seven common statistical metrics were used to evaluate the performances of different precipitation products and their performances in the hydrological model (Table 2). The correlation coefficient (CC) represents the degree of agreement between the precipitation estimates and the rain/stream gauge observation as the “ground truth.” Two metrics were selected to discover the error and bias between the precipitation products and observations, which were the relative bias (RB) to describe the systematic bias as a ratio, and the root-mean-square error (RMSE) to measure the average error magnitude. Four additional metrics were calculated to evaluate the hydrological responses of different precipitation products, which were the conventional Nash–Sutcliffe coefficient of efficiency (NSCE), peak flow error (PE), peak time error (PTE), and runoff volume ratio (RR).

<table>
<thead>
<tr>
<th>Statistic Metrics</th>
<th>Equation</th>
<th>Value Range</th>
<th>Perfect Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient (CC)</td>
<td>$CC = \frac{\sum_{n=1}^{N}(f_n - \bar{f})(r_n - \bar{r})}{\sqrt{\sum_{n=1}^{N}(f_n - \bar{f})^2} \sqrt{\sum_{n=1}^{N}(r_n - \bar{r})^2}}$</td>
<td>$-\infty, 1$</td>
<td>1</td>
</tr>
<tr>
<td>Relative bias (RB)</td>
<td>$RB = \frac{1}{N} \sum_{n=1}^{N} \frac{f_n - r_n}{r_n}$</td>
<td>$-\infty, +\infty$</td>
<td>0</td>
</tr>
<tr>
<td>Root-mean-square error (RMSE)</td>
<td>$RMSE = \left( \frac{1}{N} \sum_{n=1}^{N} (f_n - r_n)^2 \right)^{\frac{1}{2}}$</td>
<td>$0, +\infty$</td>
<td>0</td>
</tr>
<tr>
<td>Nash-Sutcliffe coefficient efficiency (NSCE)</td>
<td>$NSCE = 1 - \frac{\sum_{n=1}^{N}(f_n - r_n)^2}{\sum_{n=1}^{N}(r_n - \bar{r})^2}$</td>
<td>$-\infty, 1$</td>
<td>1</td>
</tr>
<tr>
<td>Peak flow error (PE)</td>
<td>$PE = f_{\text{max}} - r_{\text{max}}$</td>
<td>$-\infty, +\infty$</td>
<td>0</td>
</tr>
<tr>
<td>Peak time error (PTE)</td>
<td>$PTE = t(r_{\text{max}}) - t(f_{\text{max}})$</td>
<td>$-\infty, +\infty$</td>
<td>0</td>
</tr>
<tr>
<td>Runoff volume ratio (RR)</td>
<td>$RR = \frac{\sum_{n=1}^{N} f_n r_n}{\sum_{n=1}^{N} r_n}$</td>
<td>$0, +\infty$</td>
<td>1</td>
</tr>
</tbody>
</table>

* Variables: n and N, sample index and a total number of samples, $f$ represents the precipitation estimate products from gauge interpolation, radar, and satellite, $r$ represents the reference observation including the Harris County Flood Control District (HCFCD) rain gauge and United States Geological Survey (USGS) stream gauge observations.

3. Results

3.1. Precipitation Evaluation

Two series of comparisons were performed to evaluate the precipitation products, county-averaged analysis, and grid-based analysis, where the statistic results are listed in Table 3. The precipitation accumulation of Hurricane Harvey from all precipitation products is shown in Figure 2. This figure demonstrates that most precipitation products agreed that the southeastern part of Harris county received the highest precipitation amount, and the post-real-time corrections reduced the amount of precipitation for both MRMS and IMERG. In addition, V06ACal not only reduced the peak precipitation accumulation but also increased the minimum precipitation accumulation from
Contrary to the majority agreement, the NCEP gauge-only product shows the lower accumulative rainfall amount closer to the inner core of Hurricane Harvey (Figure 2).

Table 3. Summary of the statistical evaluations of county-averaged and grid-based comparison of precipitation products at 4 km and hourly resolutions.

<table>
<thead>
<tr>
<th>Precipitation Product ID</th>
<th>County-Averaged Statistics</th>
<th>Grid-Based Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
<td>RB (%)</td>
</tr>
<tr>
<td>MRMS QPE</td>
<td>0.92</td>
<td>19.57</td>
</tr>
<tr>
<td>MRMS Corr</td>
<td>0.92</td>
<td>–7.18</td>
</tr>
<tr>
<td>V06AUncal</td>
<td>0.79</td>
<td>55.97</td>
</tr>
<tr>
<td>V06ACal</td>
<td>0.79</td>
<td>32.80</td>
</tr>
<tr>
<td>NCEP</td>
<td>0.81</td>
<td>–28.88</td>
</tr>
</tbody>
</table>

Figure 2. Accumulated precipitation from 25/08/2017 to 31/08/2017 during the Hurricane Harvey event in Harris County, TX. Hourly data from (a) Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimates (QPE), (b) MRMS Corr, (c) V06AUncal, (d) V06ACal, and (e) NCEP gauge-only.

The county-averaged comparisons were conducted by averaging the precipitation data from different products over the whole Harris county and Spring Basin area. Then, we averaged all 99 valid local HCFCD rain gauge data as the reference of the truth and calculated the statistics. For the grid-based comparisons and evaluations, we extracted all hourly precipitation rate data for the grids in which the 99 rain gauges were located from all five precipitation products and then calculated the statistics. Both local rain gauge corrections reduced the bias of MRMS Corr and V06ACal from positive 20% to negative 7% (positive 20% to negative 11% for grid-based) and from 56% to 33% (80% to 58% for grid-based), respectively, as shown in Table 3. However, the corrections only made minimal reductions in the RMSE, which could be attributed to the positive/negative biases being offset. The post-real-time correction of MRMS products has a slight overcorrection (Table 3 and Figure 3). In addition, the correlation coefficient between V06AUncal and V06ACal is exactly 1, possibly because GPM IMERG uses a fairly simple algorithm to calibrate data [56]. Based on the first level statistical analysis, the MRMS product family has the highest correlation coefficient and the lowest RMSE compared with the local rain gauge observation, then NCEP gauge-only precipitation stands the
second closest, and the GPM IMERG product family is the least accurate. The NCEP precipitation product has a small relative bias value in the grid-based analysis but much greater negative bias in the county-averaged analysis, which is further investigated in this study.

![Figure 3](image-url)

**Figure 3.** County averaged precipitation rate (left) and accumulative rainfall (right), where the MRMS family is in red (solid and dash), Global Precipitation Measurement Mission (GPM) family is in blue (solid and dash), National Centers for Environmental Prediction (NCEP) is in green, and HCFCD is in black.

The county-averaged hourly rainfall rate and accumulated rainfall are shown in Figure 3. The MRMS product family has a strong correlation with the OBS (CC = 0.92), which caught most of the precipitation peaks, except for the highest one. IMERG V06 data overestimated at low precipitation intensities but underestimated at high precipitation intensities, which is consistent with the findings in the recent technical report [52]. NCEP has an obvious unresponsive condition after the midday of 28th August 2017, which could be caused by the malfunctions of rain gauges that the algorithm utilized. This could possibly be the cause of the low precipitation accumulation of NCEP (Figures 2 and 3). Generally speaking, the remotely sensed precipitation products (MRMS and IMERG) performed better than rain gauge interpolated product during Hurricane Harvey, as NCEP had a greater RMSE (Table 3). Before the peak rainfall arrived in Harris county on early 08/27/2017, the OBS (HCFCD rain gauge observation) data had the best match with MRMS Corr, then NCEP, MRMS QPE, V06ACal, and V06AUncal was the last. After the peak rainfall, the ranking changed to MRMS QPE, V06AUncal, MRMS Corr, V06ACal, and NCEP. The uncalibrated precipitation products (MRMS QPE and V06AUncal) performed better than calibrated products (MRMS Corr and V06ACal) after the rainfall intensity picked up.

As shown in Figure 4, the grid-based scatter plot has a similar finding as to the county-averaged analysis, where the MRMS product family performed the best during Hurricane Harvey. The NCEP product performed slightly better than the GPM IMERG product family. This difference indicates that rain gauge interpolation could potentially create large errors or patchiness that reduce the accuracy of the precipitation estimates over large spatial extents. The mean relative bias mathematically tends to generate a large positive value when the observation value is small during overestimations, as the lowest possible value of relative bias is −1 for positive datasets such as precipitation. This mathematical nature resulted in the large positive bias of GPM IMERG products due to their overestimation during the low-intensity precipitation period (Figure 3).
If focusing on the statistical results within the 95% confidence interval of all 99 grid points (Figure 5), one can find that the pattern differs where all precipitation products underestimated the precipitation rate during Harvey except for MRMS QPE. The MRMS Corrected product had a slight underestimation, but it was the closest to the perfect value. As shown in Figure 6, the GPM IMERG products had little temporal agreement with OBS in almost all 99 sites, and most of their higher RMSE values were concentrated close to the storm core, which indicates the accuracy of IMERG decreases as the precipitation intensity increases. NCEP had the same high RMSE concentration close to the storm core, possibly due to rain gauge malfunction. Even though the positive and negative bias offset gave the NCEP gauge-only product the best the average bias value (Table 3), the data quality is spatially inconsistent across the county. Among all the precipitation products, MRMS products shows the lowest RMSE and highest CC; however, the error has minor increases near the storm core.

**Figure 4.** Grid-scale evaluation of hourly precipitation at 99 extracted 4 km grid cells between precipitation products and the rain gauge observations. Data from (a) MRMS QPE, (b) MRMS Corr, (c) V06AUncal, (d) V06ACal, and (e) NCEP.

**Figure 5.** The distribution of the grid-scale evaluation statistics: (a) CC, (b) RB, and (c) RMSE within 95% confidence interval and the vertical lines indicate the median of data.
Figure 6. The grid-based statistic spatial distribution of different precipitation products during Hurricane Harvey.
In general, MRMS precipitation products show good agreement with HCFCD local rain-gauge observations, while IMERG and NCEP gauge-only products are comparable in quality (Table 3). This can be attributed to the following factors: first, Hurricane Harvey was an unprecedented precipitation event that potentially caused failures in the ground instruments that could impact methods relying on rain gauge interpolations and bias adjustments. Second, the GPM multi-satellite algorithm relies on passive sensors from satellites in low-earth and geostationary orbits to obtain high temporal resolution, while the radar QPE algorithm enjoys close proximity to the event and active sensors. Overall, MRMS products can best represent the surface precipitation field according to HCFCD rain gauge comparison, with high correlation coefficient (>0.9) and low RMSE (~3 and ~5 mm per hour) at the county and grid-scale and all the post-real-time correction showed overcorrection for Hurricane Harvey.

3.2. Hydrological Evaluation

The evaluation of hydrological response using MRMS QPE, MRMS Corr, V06AUncal, V06ACal and NCEP products was carried out in the northwestern basin of Harris county, where all products had generally acceptable performances according to the previous section. Two rivers were studied using the EF5 modeling framework in this section (Table 4).

Table 4. Summary of stream gauges selections and calibration results.

<table>
<thead>
<tr>
<th>Stream Name</th>
<th>Location</th>
<th>USGS ID</th>
<th>Drainage Area (km²)</th>
<th>Overbank Flow in Harvey?</th>
<th>NSCE</th>
<th>CC</th>
<th>RB (%)</th>
<th>RMSE (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>Mid-Stream</td>
<td>08068275</td>
<td>483</td>
<td>No</td>
<td>0.99</td>
<td>0.94</td>
<td>8.78</td>
<td>21.23</td>
</tr>
<tr>
<td>Spring</td>
<td>Down-stream</td>
<td>08068500</td>
<td>1059</td>
<td>No</td>
<td>0.99</td>
<td>0.91</td>
<td>−8.34</td>
<td>19.99</td>
</tr>
<tr>
<td>Cypress</td>
<td>Upper-stream</td>
<td>08068720</td>
<td>280</td>
<td>No</td>
<td>0.90</td>
<td>0.90</td>
<td>34.64</td>
<td>5.15</td>
</tr>
<tr>
<td>Cypress</td>
<td>Mid - steam</td>
<td>08068800</td>
<td>540</td>
<td>Yes</td>
<td>0.91</td>
<td>0.87</td>
<td>7.36</td>
<td>4.93</td>
</tr>
<tr>
<td>Cypress</td>
<td>Down-stream</td>
<td>08069000</td>
<td>738</td>
<td>Yes</td>
<td>0.95</td>
<td>0.97</td>
<td>−3.53</td>
<td>20.15</td>
</tr>
</tbody>
</table>

The first river, Spring Creek, has a larger channel network with a total of 179 km of open channels and a large natural floodplain. The Spring Creek watershed remains mostly underdeveloped and natural, except for downstream, where the Woodland Township and the city of Tomball are located. Due to the lack of urban development, the flood risk of Spring Creek is comparatively less, and there were no signs of overbank flow from two USGS stream gauges on this river during Harvey. This river was studied to evaluate the performance of different hydrological responses of MRMS, GPM IMERG, and NCEP precipitation products.

The second river, Cypress Creek, is a smaller river with 137 km of open water channel and a well-developed drainage area. The middle and downstream portion of Cypress Creek has experienced intensive urbanization in the past 20 to 30 years, hosting a population greater than 350,000 according to the 2010 U.S. census, while only the upstream area remains as agricultural land (Figure 1). Perhaps exacerbated by the impacts of urbanization, both the middle and downstream USGS gauges experienced overbank flow during Harvey, despite the effort of a 200,000 m² detention basin and multiple detention ponds located around the upper stream of Cypress Creek. This river was studied to examine the early warning capability of the coupled hydrological modeling system when the ground instruments were damaged or malfunctioning during the extreme event.

3.2.1. Spring Creek

First, the EF5 model was forced by MRMS Corrected precipitation data, considered as the most accurate data source during non-extreme events, to benchmark the model parameters from 1st April 2017 to 15th August 2017, with two 15-minute USGS streamflow observations (midstream and downstream of Spring Creek) using the Differential Evolution Adaptive Metropolis (DREAM)
The model then proceeded to warm up for the same period and the same forcing precipitation data (MRMS Corr) was used during the benchmark generating process to create a uniform initial condition for all precipitation product evaluations. Table 4 displays good agreement between simulated streamflow and the observations during the calibration period, with very high NSCE (0.99) and CC (~0.93), as well as a small RMSE (~20 m³/s), slight overestimation at midstream (8.78%) and underestimation at downstream (~8.34%). The MRMS, GPM IMERG, and NCEP rain gauge interpolated precipitation products were then set as the model forcing data to simulate the hydrological responses from 15/08/2017 to 15/09/2017 separately for Hurricane Harvey. All products were read at their original spatial and temporal resolution by the model, as listed in Table 1.

In the hydrological simulations, all precipitation products generally had acceptable performances with a NSCE greater than 0.8, except for one, and a high CC (Table 5). All products overestimated the streamflow at the midstream of Spring Creek, and MRMS QPE had a large spike of simulated streamflow at the peak due to data sensitivity (Figure 7).

Moreover, the total runoff ratio (RR) of MRMS QPE at the midstream was large (Table 5), which could be due to the error propagation from the MRMS QPE as its errors were mostly located at the upstream of Spring Creek (Figure 6), which could partially cause the spike as well. All products yielded reasonable simulation hydrographs that could capture the flow peak with high CCs (~0.94), and MRMS products were able to simulate the peak with over 30 min of lead time, which is crucial and beneficial for flood early warnings. At the downstream gauge, all products underestimated the streamflow peak and MRMS Corrected, V06AUncal and NCEP products had negative bias compared to the USGS stream gauge data. One observation of the simulation statistics is that the near-real-time remote sensing precipitation products performed almost equally well as the post-real-time corrected products, where the differences of NSCE (<0.1), bias (<0.2), and RMSE (<50 m³/s) between the two groups were minimal (Table 5). We can conclude that MRMS QPE and IMERG V06AUncal (equivalent to IMERG late run) are sufficient to drive the hydrological model to provide flood warning information.

### Table 5. Summary of hydrological simulation of Spring Creek.

<table>
<thead>
<tr>
<th></th>
<th>Spring Creek 08068275</th>
<th>Spring Creek 08068500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRMS QPE</td>
<td>V06AUncal</td>
</tr>
<tr>
<td>NSCE</td>
<td>0.62</td>
<td>0.9</td>
</tr>
<tr>
<td>RB (%)</td>
<td>50.64</td>
<td>23.69</td>
</tr>
<tr>
<td>RMSE (m³/s)</td>
<td>167</td>
<td>84</td>
</tr>
<tr>
<td>Peak Error (m³/s)</td>
<td>1066</td>
<td>97</td>
</tr>
<tr>
<td>Peak Time Error (min)</td>
<td>33</td>
<td>28</td>
</tr>
<tr>
<td>RR</td>
<td>1.51</td>
<td>1.24</td>
</tr>
</tbody>
</table>

**Figure 7.** Comparison of EF5 simulated streamflow with USGS stream gauge observation during Hurricane Harvey at midstream (left) and downstream (right) of Spring Creek.

Remote Sens. 2020, 12, x FOR PEER REVIEW 3 of 4
As the near-real-time products can provide timeliness, the remote sensing technologies can significantly increase the accuracy and reliability of global flood early warning systems. As shown in Figure 7 and the statistic metrics in Table 5, the MRMS product family performed best if not considering the sensitivity effect in the midstream, followed by NCEP gauge only and GPM IMERG, according to the metrics of NSCE, RB, CC, and RMSE.

In Figure 8, the three relatively independent statistics are plotted in a 3-dimensional scatter plot, where the lower-left corner is the perfect point with the correlation coefficient of 1 and no flow peak nor peak time error. As shown in Figure 8, the MRMS products are closer to the perfect point overall, followed by GPM IMERG and then NCEP gauge-only, indicating that the performance of GPM IMERG and NCEP gauge-only are comparable for Hurricane Harvey in Spring Basin. Since products that are not based on rain gauge corrections show hydrological performance superior to the gauge-forced product within Spring Basin, we can infer that the gauge-interpolated precipitation product is reliable only when the gauge network is well functioning. Downstream of Spring Creek, the NCEP gauge-only product yielded a hydrograph which had significant deficit in total runoff (RR=0.85), which shows the disadvantage of using the NCEP product as the underestimation occurred near the storm core and this error propagated to the hydrological simulation.

Figure 8. The 3-dimensional scatter plot using CC, PE, and PTE as variables for all three precipitation product families. The perfect point is the left lower corner.

In summary, MRMS performed the best in capturing the hydrological response in Spring Creek, compared to USGS stream gauge observations during Hurricane Harvey, followed by GPM IMERG and NCEP gauge-only precipitation products which had comparable hydrological responses. The post-real-time corrected remote sensing precipitation products did not provide significant improvement in hydrological responses, which justifies the global real-time operational flood warning system based on the near-real-time products.

3.2.2. Cypress Creek

For Cypress Creek, the EF5 model was calibrated and warmed up using the same methodology as with Spring Creek, enabling comparison with three USGS stream gauges located at upper, middle and downstream reaches. Table 4 shows the statistic results during the warm-up period at Cypress Creek stream gauges, which have high NSCE (>0.9) and CC (>0.87) values for all gauge locations during non-extreme periods, which indicates that the EF5 simulation had good agreement with the stream gauge observations. The CREST-EF5 simulation has a tendency towards slight overestimation at the upstream and gradually changed to underestimation at the downstream, which matches the findings of the previous study [61].

Figure 9 shows the hydrographs at three stream gauges along Cypress Creek and EF5 simulated results using MRMS, IMERG, and NCEP products during Hurricane Harvey. At the upstream location (USGS 08068720), the observed hydrograph has an obvious long and gentle receding limb, which was caused by the 650,000 m² Warren Reservoir located 6.4 km ahead of the upstream gauge. The slow-release and regulation of the water caused the long receding limb and was not considered
in this study. Since the Cypress Creek drainage area has been under rapid development during the past three decades, HCFC reported that overbank flow was the common problem of the region and extensive water detention-related engineering jobs were completed. Additionally, as shown in Figure 9, both the mid and downstream observed hydrographs also have longer receding limbs. However, during this intense event, the engineered infrastructure did not prevent the overbank flow and flooding in the middle and downstream of Cypress Creek, where both observed hydrographs (USGS 08,068,800 and 08069000) have an unnatural plateau. Furthermore, the linear appearance of the rising and falling limbs at the midstream gauge was a result of no data being recorded from noon 27/08/2019 until 01/09/2019 due to gauge malfunction.

Figure 9. Comparison of EF5 simulated streamflow with USGS stream gauges during Hurricane Harvey at upstream (left), midstream (middle), and downstream (right) of Cypress Creek.

Due to the lack of information on Warren Reservoir and its operational details during Hurricane Harvey, EF5 simulation results could not capture the reduction in peak flow magnitude nor the long receding limb from the slow release of detained stormwater. In Figure 9, all simulated hydrographs have a very steep receding limb, which represents a more typical hydrological response in urban areas. Therefore, it is inferred that EF5 simulated the urban hydrological conditions without the reservoir interferences, which is sufficient for flood conditions in the related warning scenarios. For the upstream gauge (USGS 08068720), MRMS QPE could closely capture the first peak almost at the same time as the stream gauge observation, and the mismatch of the second peak could be caused by the structural flood control. IMERG performed poorly for Cypress Creek, as its simulated hydrograph was rather “smooth” when not considering structural flood control, unlike typical urban hydrographs. At this point, we can conclude that MRMS products performed the best in the urban hydrological condition, and we can use MRMS products to regenerate a well-approximated hydrograph when the stream gauge is damaged or over bank flow occurs. The simulated streamflow could be utilized for many applications, including but not limited to flood warning, flood risk analysis, flood inundation calculation, and flood control performance assessment.

In summary, the EF5 modeling framework can simulate close-to-reality streamflow using MRMS precipitation products during the extreme precipitation event, especially when the ground stream gauges are damaged, or no gauge is implemented, or when overbank flow occurs. IMERG products might not be suitable for simulating extreme events in small basins (e.g., <200 km²), but given its global coverage, it is sufficient for major rivers and sub-basins, where most CREST and satellite precipitation studies were found successful when applied to different basins globally [5,36,62]. Overall, the EF5 modeling framework combined with current remote sensing technologies (IMERG and MRMS) can provide a robust flood-early-warning system for real-time operational uses at the regional, continental, and even global scale.
4. Discussion

Three results from the previous section require further explanation. First, the result in Figure 3 indicates that after the peak rainfall intensity occurred, the uncalibrated precipitation products (MRMS QPE and V06AUncal) outperformed the calibrated precipitation products (MRMS Corr and V06ACal). This might be caused by a combination of overcorrection of the underestimation of GPM IMERG during the high-intensity precipitation period, possible rain gauge malfunctions, and post-real-time calibration error. Due to the unprecedented nature of Hurricane Harvey, it could cause errors to the post-real-time correction algorithms at such an intensity. Therefore, the current precipitation correction algorithms for MRMS and GPM IMERG are less valid for unprecedented events like Hurricane Harvey. We also suspect that the insensitive precipitation observation from the NCEP data is caused by the damage to instruments by the flood or objects and the mechanical saturations caused by the intensive rainfall. However, the true causes of the suspected rain gauge malfunction are unknown.

Second, the results from Figure 7 show a spike on the MRMS QPE simulated hydrograph. This could be caused by the sensitivity to the high quality and high temporal resolution of MRMS QPE product with the combination of smaller channels in the upstream. When simulated with MRMS Corr, the resulting streamflow had a dramatically smoother spike as the temporal resolution is hourly.

Third, the Cypress Creek hydrological evaluation yields an untypical hydrograph at the upper stream from IMERG produced simulation. However, as stated in the previous section, EF5 did not consider the impact of flood control infrastructure. This result could be caused by the low spatial resolution of IMERG combining with the smaller drainage area which undermined the precipitation representativeness. It was a fact that there were less than 4 grid cells from IMERG products within the drainage area of the upstream gauge. At the downstream location, the IMERG yielded a more reasonable hydrograph, as there were 12 grid cells covering the drainage area.

5. Conclusions

The results of this study indicate that the remote sensing technologies and gauge interpolation method could all detect the unprecedented extreme rainfall associated with Hurricane Harvey, as well as capturing the cascading hydrological responses. This study first focused on statistically comparing the MRMS QPE, MRMS Corr, IMERG final V06AUncal and V06ACal, as well as NCEP gauge-only interpolated precipitation products with the very dense HCFCD local rain gauges. Then, hydrological responses were evaluated using the EF5 modeling framework in Harris County and Spring Basin of Texas, USA. The findings and results of this study can be potentially applicable to other subtropical zones impacted by tropical cyclones or low-lying flood-prone areas that are similar to southeast Texas, particularly in extreme events.

The main conclusions from the cross-evaluation of MRMS, IMERG, and NCEP gauge-only precipitation products at county- and grid-based scales during Harvey, are summarized below:

1. MRMS precipitation products are the best remote sensing rainfall measurements that perform most comparably to the local dense network rain gauge observations. IMERG and NCEP rain gauge-interpolated precipitation products are comparable to each other statistically, but IMERG has the advantage of global coverage, and is not limited to the national radar and local rain gauge network coverage.

2. The post-real-time corrections for remote-sensing-based precipitation products were not necessarily valid for the unprecedented precipitation event and caused overcorrections to MRMS and IMERG, as overcorrection occurred for both product families.

3. IMERG products tended to overestimate the low–moderate precipitation intensity but underestimate the highest precipitation intensities. The NCEP product showed significant underestimation, especially near the storm core region, due to possible instrumental failure during the record Harvey event, implying its high dependence on the functionality and reliability of the ground instruments during extreme events.
In terms of the hydrological evaluation, Spring Basin was selected because no precipitation product had significant errors within the basin from the above analysis. Two rivers were analyzed separately for their differences in urban and natural hydrological environments. The main conclusions are as follows:

1. Consistent with statistical evaluations, MRMS performs the best, showing comparable simulations with USGS stream gauge observation in Spring Basin during Hurricane Harvey, followed by IMERG and NCEP with acceptable performances.

2. The current remote-sensing-based, near-real-time precipitation products are sufficient to capture the extreme precipitation and its cascading hydrological responses. Providing the advantages of timeliness and vast spatial coverage in national and global scale, the user community is encouraged to integrate the latest remote sensing products into their operational flood disaster warning systems for the public to be informed, and to reduce and mitigate the risk of extreme precipitation events.

3. The EF5 modeling framework can capture the hydrological responses during such unprecedented extreme precipitation events; and, more powerfully, such a system can be integrated with the latest remote sensing forcing data (i.e., MRMS and IMERG) into national and even global modeling frameworks to alternatively compliment the vast ungauged regions.

This study proved the value of MRMS precipitation products for extreme precipitation detectability and accuracy, as well as their capability in hydrological prediction when combined with the EF5 model framework, which further confirmed the success of the FLASH project run by the NOAA National Severe Storm Laboratory (NSSL) and the University of Oklahoma (http://flash.ou.edu). Even though the GPM IMERG’s performance during Hurricane Harvey is second to MRMS, its simulated hydrological response is sufficient to provide flood magnitude and peak-timing warnings and to potentially build an operational flood early warning system at the global scale given satellite products’ global coverage. Furthermore, it is reasonable to expect the evolving IMERG products will keep improving for hydrological and water resource applications. Concurrently, as the recent remote sensing technologies have progressed to accurately capture an unprecedented rainfall event, the EF5 modeling framework will need further improvement to provide not only the streamflow estimation but also the flood inundation extents and even water-depth over inundated urbans. Ultimately, future interdisciplinary building blocks are encouraged to connect extreme rainfall, hydrological responses and consequent flood risk analysis, as well as loss quantifications, in order to maximize the socio-economic value of the latest remote sensing observations for the general public, nationally and globally.

Author Contributions: Conceptualization, M.C. and Y.H.; Data curation, M.C., S.G., J.J.G. and Z.H.; Formal analysis, M.C.; Funding acquisition, R.L.K. and Y.H.; Investigation, M.C. and S.N.; Methodology, M.C. and Y.H.; Project administration, Y.H.; Resources, M.C. and J.J.G.; Software, M.C.; Supervision, R.L.K. and Y.H.; Validation, M.C. and S.N.; Visualization, M.C. and S.G.; Writing—original draft, M.C.; Writing—review & editing, N.S.B., S.G., J.J.G. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study is partially funded by University of Oklahoma and also based upon work supported by the National Science Foundation under Grant No. 1545874.

Acknowledgments: We are grateful to the NOAA/NSSL and NASA science team who are responsible for the development of MRMS and GPM precipitation products. The efforts by HydroSHEDS, IEM, and USGS are also highly appreciated for making the data available for public uses.

Conflicts of Interest: The authors declare no conflict of interest.

References


22. Werner, M.; Reggiani, P.; De Roo, A.; Bates, P.; Sprokkereef, E. Flood Forecasting and Warning at the River Basin and at the European Scale. *Nat. Hazards* 2005, 36, 25–42. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).