RESEARCH ARTICLE
10.1029/2020EA001518

Key Points:
• We set up a rainfall-runoff model using hourly rainfall data and validated it against the hourly water level for the West Rapti River Basin, Nepal.
• We evaluated the performance of different satellite-based rainfall estimates in capturing an extreme precipitation event.
• We corrected poor-performed products with respect to gauge data and filled gauge data gaps using the information of other well-performed products.

Correspondence to:
R. Prajapati and R. Talchabhadel, rocky.ioe@gmail.com

Citation:
Talchabhadel, R., Nakagawa, H., Kawaike, K., Yamanoi, K., Musumari, H., Adhikari, T. R., & Prajapati, R. (2021). Appraising the potential of using satellite-based rainfall estimates for evaluating extreme precipitation: A case study of August 2014 event across the West Rapti River Basin, Nepal. Earth and Space Science, 8, e2020EA001518. https://doi.org/10.1029/2020EA001518

Received 18 OCT 2020
Accepted 29 JUN 2021

Abstract
Heavy precipitation events are recurrently occurring in Nepal, affecting lives and properties every year, especially in the summer monsoon season (i.e., June-September). We investigated an extreme (heavy) precipitation event of August 2014 over the West Rapti River (WRR) Basin, Nepal. First, we forced a rainfall-runoff model with ground-based (gauge) hourly rainfall data of nine stations. Second, we validated against hourly water level at an outlet of the WRR Basin. This study then evaluated the performance of different satellite-based rainfall estimates (SREs) in capturing an extreme precipitation event. We considered the use of half-hourly data of Integrated Multi-satellite Retrievals for GPM (IMERG) (Early, Late, and Final versions), spatial resolution (10 km), and hourly data of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), spatial resolution (25 km), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System (PERSIANN-CCS), spatial resolution (4 km). Also, we used 3 h data of Tropical Multi-satellite Precipitation Analysis (TMPA) product real-time (3B42RT), spatial resolution (25 km). In general, we find that all selected SREs depicted a similar pattern of extreme precipitation as shown by the gauge data on a daily scale. However, we find these products could not replicate precisely on a sub-daily scale. Overall, IMERG and TMPA showed a better performance than PERSIANN and PERSIANN-CCS. Finally, we corrected poor-performed SREs with respect to gauge data and also filled data gaps of gauge rainfall using the information of good-performed SREs. Our study reveals that there is a great challenge in local flood simulation employing SREs at high-temporal resolution in Nepal.

Plain Language Summary
We assessed a heavy precipitation event of August 2014, where hourly rainfall data were applied in a hydrologic model in the WRR basin, Nepal. We evaluated the performance of different SREs and found all selected SREs demonstrated a similar tendency compared to gauge data on a daily scale. However, they failed to replicate on a sub-daily scale. Finally, we corrected poor-performed SREs (PERSIANN family) with respect to gauge data and also filled data gaps of gauge rainfall using the information of good-performed SREs (IMERG family and TMPA). Thus, we find there is a great challenge in local flood simulation using SREs at high-temporal resolution in Nepal.

1. Introduction
Heavy to very heavy precipitation (from now on extreme precipitation) events are recurrently occurring on global and regional scales. Nepal, lying along the southern slopes of the central Himalayas, is highly susceptible to the development of extreme precipitation and subsequent flooding (Karki et al., 2018). Several incidences of extreme precipitation are observed affecting lives and properties, especially during the summer monsoon season (SMS) in Nepal. For instance, a multi-day cloudburst of June 13–19, 2013 in the north-western mountainous region near the Nepal-India border resulted in large flash floods and massive landslides. A huge loss of life and property was reported (Karki et al., 2017; Paudel et al., 2013). An extreme precipitation event of August 14–15, 2014, is another seriously destructive event of recent time in which...
a maximum precipitation amount of 528 mm in 24 h (08:45 Local Time, i.e., 03UTC August 14–15, 2014) was observed at Chisapani station (location shown in Figure 1a), which is the highest ever recorded 24 h accumulated precipitation in Nepal (Karki et al., 2018). Similarly, severe floods were seen in many regions across the country in the SMS in recent times, like 2015, 2017, and 2019. Recently, a new pattern of a flood is observed across the entire range (i.e., from east to west) of the southern plain of the country during the SMS at the same time every year. The frequency of occurrences of rainfall >100 mm/day shows a significant increasing trend in most of the southern part of Nepal (<2,000 m asl), which was not common before 2000 (Pokharel et al., 2019).

Interestingly, two significant rainfall peaks appear over the southern slope of the central Himalayas due to its unique topographical setting, the first peak appears along 500–700 m above sea level (asl), and the second peak appears along 2,000–2,200 m asl (D. Shrestha et al., 2012; Talchabhadel et al., 2018). The former is attributed to fewer but heavy rainfall events, whereas the latter is weak but frequent and persistent (D. Shrestha et al., 2012). Therefore, a higher amount of seasonal precipitation is observed along 2,000–3,000 m asl compared to the southern plain. Still, when it comes to extreme precipitation, the southern plain receives intermittent high-intensity extreme precipitation. During extreme precipitation, warm moist monsoonal air
primarily strikes the first southern lower hill ranges (Churia range), resulting in a heavy downpour (Talchabhadel et al., 2018). Karki et al. (2017), Talchabhadel, Aryal, Kawaike, Yamanoi, and Nakagawa, 2021, and Bohlinger and Sorteberg (2018) reported that heavy precipitation related indices; maximum 1-day precipitation amount (RX1day), maximum 5-days precipitation amount (RX5day), and precipitation of days in >95th percentile (R95p), are rising significantly in lower hills of the western region of the country.

During the latitudinal swaying of monsoon trough to the north of normal position, meaning shifts to foothills of Himalayas, break monsoon occurs. As a result, westerly winds prevail, and rainfall amounts in Central India are small. In contrast, rainfall amounts enhance across foothills of the Himalayas, including the southern plains of Nepal. One or two major breaks in monsoon are observed frequently in a year. However, sporadically mid-latitude cyclonic systems coincide with the monsoonal system during the break monsoon, resulting in extreme precipitation in the form of cloudburst over the Himalayas (A. Shrestha, 2016). In general, this type of interaction provides a suitable synoptic condition to converge massive moisture from both the Arabian Sea and the Bay of Bengal, affecting the country’s western region (Talchabhadel et al., 2018). Bohlinger et al. (2017) systematically investigated synoptic conditions and moisture sources triggering extreme precipitation in Nepal during 1979–2010, employing FLEXible PARTICle (FLEXPART), a Lagrangian transport and dispersion model of the atmospheric transport process. Rigorous exploration is needed case-wise to investigate the complex dynamics of any extreme precipitation. Responses of watershed/river basins on different cases of extreme precipitation are different. Observation constraints (sub-daily resolution) make such exploration difficult.

Space-borne products like satellite-based rainfall estimates (SREs) are helpful on many occasions, but we need to take care of underlying biases. Currently, there is the availability of several SREs, and a comprehensive review of SREs’ algorithm is beyond the scope of this paper. Please see Sun et al. (2018) for detailed information on the global precipitation data set, including SREs. Products based on multiple-satellite and merging satellite and gauge data have been considered to improve the accuracy of SREs. Verification of SREs is essential for any practical application. In general, we can perform an error evaluation of SREs if sufficient gauge data exist. After then, systematic biases can be reduced. There are many studies (Aghakouchak et al., 2011; Almazroui, 2011; Hong et al., 2007; Mahbod et al., 2019; Pakoksung & Takagi, 2016; Pombo & de Oliveira, 2015; Talchabhadel, Aryal, Kawaike, Yamanoi, Nakagawa, Bhatta, et al., 2021; Tarek et al., 2017; Yoshimoto & Amarnath, 2017) dealing with error evaluation of SREs and their applications but most of these studies are carried out on a daily, monthly, seasonal or annual scale. Here, we tried to assess the potential of using SREs during an extreme precipitation event. We chose the extreme precipitation event of August 14–15, 2014 across the WRR basin of Nepal and conducted our analysis for the study period August 10–19, 2014.

The torrential rain seriously affected the southern plain of west Nepal, and water levels in most rivers increased above warning and danger levels. These warning and danger levels are set based on hydrodynamic modeling and inundation scenarios in downstream settlements. The warning level is anticipated flood flow representing the riverbank’s full stage and the beginning of inundation in the settlement areas. In contrast, the danger level is that level of flow in which the floodwater enters the settlements affecting people and properties, or the inundation depth remains within one meter (Gautam & Dulal, 2013). In Babai River, adjacent to the WRR basin shown in Figure 1a, the early warning system (EWS) was completely washed away by the flood, and no EWS worked, which caused human deaths and huge property loss (Reliefweb, 2014). Such a scenario highlights the importance of SREs at the time of failure of gauge data. The rain gradually reduced from the afternoon of August 15, 2014 even though floodwaters took some days to recede from inundated villages. For a numerical analysis of such temporarily dynamic processes, sub-daily (e.g., hourly hydrometeorological) data are required. There are four hydrometric and nine precipitation stations with sub-daily temporal resolution (shown in Figure 1c) across the WRR basin. We selected the WRR basin for our analysis because other adjacent basins of the affected region (i.e., Babai, Karnali) do not have sufficient sub-daily data.

In this study, we applied a kinematic wave flow model on hill slopes for simulating surface runoff. A shallow-water unsteady flow model was used to simulate the inundation propagation. First, the hydrologic model was validated against hourly gauge river discharge. This study evaluated the performance of SREs in capturing an extreme precipitation event. We used (a) NASA Global Precipitation Measurement (GPM)
Integrated Multi-satellite Retrievals for GPM (IMERG) - Early, Late and Final versions, (b) Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), (c) PERSIANN-Cloud Classification System (PERSIANN-CCS), and (d) Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) products. In this study, we proposed corrections of SREs with respect to (wrt) gauge data, and at the same time, we also filled data gaps of gauge data using the information of SREs.

The paper is organized as follows. First, the study area is described in Section 2. Next, we describe precipitation data (gauge and SREs) in Section 3. This section also includes a description of the error evaluation of SREs and numerical simulation of the response of extreme precipitation in the study area. Next, Section 4 presents our results, and we offer some discussions. Finally, Section 5 concludes the overall study.

2. Materials and Methods

2.1. Study Area

The study area (Figure 1a), a catchment area of 6,368 km² up to the Nepal-India border, has a diverse elevation range (shown in Figure 1b) varying from 100 to 3,600 m above-sea-level (asl) within a small latitude extent of nearly 1°. The response time of extreme precipitation events is very quick due to the steep topography. The mean annual precipitation (MAP) is about 1,500 mm, and more than 80% of it occurs during the SMS. Figure 1c shows a spatial distribution of total precipitation of the selected period (August 10–19, 2014). A localized nature of heavy precipitation, almost 45% of the MAP, is observed around station 3. Stations 1–5 (from left) are located in the lower elevation (0–500 m asl), and stations 6–9 (from left) are located in the upper region of the study area.

2.2. Hydrometeorological Data

We used hourly rainfall data of nine automatic precipitation stations and water level data of four automatic hydrologic stations distributed across the WRR basin maintained by the Department of Hydrology and Meteorology (DHM), Government of Nepal. Hourly discharges at stations were estimated employing the stage-discharge relation. However, the developed stage-discharge association was based on the limited number of streamflow measurements. Also, peak discharge is challenging to measure; therefore, it is estimated using different empirical approaches. Generally, daily average water levels (i.e., the arithmetic average of three recordings at 8 a.m., 12 noon, and 4 p.m.) are used to estimate daily average streamflow. Thus, the estimation of hourly discharge has some uncertainty, and their quantification is beyond this study’s scope.

The automatic network includes tipping bucket-type precipitation gauges. Please see Talchabhadel et al. (2017) for detailed information about precipitation measurement in Nepal. We used half-hourly data of IMERG (Early, Late, and Final versions), resolution (10 km), hourly data of PERSIANN, resolution (25 km), and PERSIANN-CCS, resolution (4 km), and 3-h data of TMPA (3B42RT), resolution (25 km).

IMERG evaluates rainfall from the various passive microwave (PMW) sensors comprising GPM constellation are computed employing Goddard Profiling Algorithm (GPROF2017). Detailed information could be found in Huffman et al. (2019). We used a current version, that is, 06B version, of Early, Late, and Final products of IMERG. Early product is accessible after ~4 h of observation time, and Late product after ~14 h of observation time without an ultimate calibration, whereas Final product is available after ~3.5 months of observation time after the ultimate calibration based on monthly gauge analysis. TMPA algorithm (Huffman et al. 2007) offers rainfall estimates in the TRMM regions using microwave-calibrated infrared approximations. We used a 3B42RT version7 TMPA product in this study. The TMPA products ended on December 31, 2019. New products to supersede the TMPA datasets are being produced under the GPM umbrella with the IMERG algorithm.

PERSIANN system is developed by the Center for Hydrometeorology and Remote Sensing (CHRS), which employs a neural network to estimate rainfall rate using the infrared brightness temperature image of geostationary satellites (Hsu et al., 1997). The system has an updating adaptive training feature of the parameters upon the availability of independent rainfall estimates, for instance, TRMM. PERSIANN-CCS is
a real-time global high-resolution satellite product that permits the categorization of cloud-patch features based on cloud information, including height, areal extent, and variability of texture estimated from satellite imagery (Hong et al., 2004). Detailed information of the PERSIANN family is presented in Nguyen et al. (2018).

### 2.3. Performance Analysis of Satellite-Based Rainfall Estimates

Inter-comparisons of different data sources were conducted with gauged data at nine stations. At first, the gauge data was checked for any data gaps. Then, at the station without any data gaps, rainfall detection abilities were evaluated using four statistical indices, namely (a) probability of detection (POD), (b) critical success index (CSI), (c) false alarm ratio (FAR), and (d) frequency bias index (FBI). Table 1 shows a detail of the calculation of indices. POD denotes the proportion of the number of correctly detected rainfall events of SRE wrt the number of rainfall occurrences observed by gauge data. Similarly, CSI displays an overall ratio of rainfall events correctly identified by SRE, whereas FAR indicates the proportion of false status. POD, CSI, and FAR vary from 0 to 1: 0 being an ideal FAR and 1 being a perfect POD and CSI. FBI represents a simple bias between gauge data and SRE. Detection indices were calculated on hourly, 3 h, 6 h, 12 h, 18, and 24 h scales.

For a magnitude-based evaluation of SREs, mean difference (MD), mean absolute difference (MAD), root mean square error (RMSE), and percentage bias (PBIAS) wrt gauge data were computed.

\[
MD = \frac{1}{N} \sum_{i} (SRE_i - \text{Gauge}_i) 
\]

\[
MAD = \frac{1}{N} \sum |SRE_i - \text{Gauge}_i| 
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum (SRE_i - \text{Gauge}_i)^2} 
\]

\[
PBIAS = \frac{\sum (SRE_i - \text{Gauge}_i)}{\sum \text{Gauge}_i} \times 100 
\]

where \(i\) represents the station, \(N\) represents the number of stations (i.e., nine in this study). The pixel value at the station location of SREs was used for the comparison, following the point-to-pixel approach. Negative values of MD and PBIAS represent an underestimation, and positive values represent an overestimation of SRE wrt gauge data. A lower value is ideal for MAD and RMSE. This study applied a simple linear scaling correction to SREs on an event basis at stations without any data gaps. The gauge stations with data gaps were filled using the SREs having good performances.

### Table 1: Statistical Indices for Rainfall Detection Evaluation

| Status | Satellite | Gauge | Detection indices | Equation | Range | Ideal value |
|--------|-----------|-------|-------------------|----------|-------|-------------|
| True   | Rainy     | Rainy | Probability of detection | \(T/(T + M)\) | 0 to 1 | 1           |
| False  | Non-Rainy | Non-Rainy | Critical success index | \(T/(T + M + F)\) | 0 to 1 | 1           |
| Miss   | Non-Rainy | Rainy | False alarm ratio | \(F/(T + F)\) | 0 to 1 | 0           |
| Null   | Non-Rainy | Non-Rainy | Frequency bias index | \((T + F)/(T + M)\) | 0 to \(\infty\) | 1           |

*Note.* Rainfall less than 0.2 mm in an hour is treated as a non-rainy event.

Rain threshold = 0.2 mm per hr.

Abbreviations: F, False; M, Miss; T, True; X, Null.
2.4. Hydrologic Modeling

An RRI model (Sayama et al., 2012), developed by the International Center for Water Hazard and Risk Management (ICARM), was used employing a 2D kinematic wave model for hillslope and a 1D river-routing model for river channel for simulating rainfall-runoff processes. We used the topography data of HydroSHEDS (Lehner et al., 2008) with 30-s (approximately 900-m) resolution, provided by the U.S. Geological Survey. Governing equations are based on the conservation of mass and momentum.

\[
\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = r
\]  

(5)

\[
\frac{\partial q_x}{\partial t} + \frac{\partial u q_x}{\partial x} + \frac{\partial v q_x}{\partial y} = -gh \frac{\partial H}{\partial x} - \frac{gn^2 u^2 + v^2}{h^{1/3}}
\]  

(6)

\[
\frac{\partial q_y}{\partial t} + \frac{\partial u q_y}{\partial x} + \frac{\partial v q_y}{\partial y} = -gh \frac{\partial H}{\partial x} - \frac{gn^2 u^2 + v^2}{h^{1/3}}
\]  

(7)

where \( h \) is water depth, \( q_x \) and \( q_y \) are the unit width discharges in \( x \)- and \( y \)-directions respectively, \( r \) is rainfall intensity, \( u \) and \( u \) are flow velocities in \( x \)- and \( y \)-directions respectively, \( g \) is the gravitational acceleration, \( n \) is Manning’s roughness parameter, and \( H \) is the height of the water from the datum.

Under kinematic wave approximation, local acceleration, convective acceleration, and pressure terms are neglected, meaning the gravity and friction forces balance each other. Detailed information is described in Sayama et al. (2012) and the RRI model manual (ICARM, 2015). The RRI model uses flow accumulation and direction to locate river channels but not for flood routing (Sayama et al., 2015). River width and depth are estimated using the following functions of the upstream contributing area.

\[
W = C_W A^{0.5}
\]  

(8)

\[
D = C_D A^{0.5}
\]  

(9)

where, \( W \) and \( D \) are river width and river depth respectively in meters, \( C_W, S_W, C_D, \) and \( S_D \) are geometry parameters. These coefficients are used to estimate the river width and depth at all computation grids. In this study, we used the default values (ICARM, 2015): \( C_W = 5.0, S_W = 0.35, C_D = 0.95, \) and \( S_D = 0.2. \) Based on observed data of river width and river depth at different cross-sections, these parameters could be precisely estimated in a local context. Since we do not have observed cross-sections at river channels upstream of Kusum station, we limited this study with the default values. However, we pay attention to calibrating the model by changing the roughness coefficient of river and hill slopes. In this study, the geometry is assumed to be a rectangle for simplicity. These geometric details are sensitive to hydrodynamic routing.

We first calibrated the hydrologic model on the flood event of August 2012, one of the big floods in WRR during the recent time after implementing the automatic hydrometeorological station. The WRR crossed the danger level set by the DHM for 19 h during that event (Talchabhadel & Sharma, 2014). After checking the model applicability, it was applied for the case of August 2014. The hydrologic model was fed with (a) gauge data, (b) SREs data, (c) bias-corrected SREs, and (d) gauge data after filling data gaps. We assessed the performance of hydrologic simulation using Nash-Sutcliffe Efficiency (NSE), percentage bias (PBIAS), and square of correlation coefficient \( R^2 \) by comparing simulated and observed streamflow on an hourly timescale for calibration, validation, and all scenarios. NSE is a normalized index that tells the magnitude of residual variance compared to observed variance (Nash & Sutcliffe, 1970). NSE ranges from -\( \infty \) to +1. The ideal value of NSE is 1. PBIAS informs the volume deviation compared to observed values. Here, negative PBIAS represents an underestimation, and positive PBIAS represents an overestimation. \( R^2 \) describes the correlation between observed and simulated streamflow. A value close to zero represents poor correlation, whereas a value close to one represents a strong correlation. Please see Krause et al. (2005) and Moriasi et al. (2007) for detailed information on these performance metrics.
3. Results and Discussion

3.1. Performance Analysis of Satellite-Based Rainfall Estimates

Figure 2 shows basin-averaged rainfall computed using the Thiessen weightage method during the selected period (August 10–19, 2014) based on gauged data and different satellite-based rainfall estimates. The basin averaged rainfall is computed using Thiessen’s weightage method.

Figure 2. Temporal distribution of basin-averaged hourly rainfall of the study area during August 10–19, 2014 based on gauged data and different satellite-based rainfall estimates. The basin averaged rainfall is computed using Thiessen’s weightage method.

In general, PERSIANN and PERSIANN-CCS showed an underestimation, whereas IMERG (Early, Late, and Final versions) and TMPA showed an overestimation wrt to gauge data on a spatially average scale. All SRE products depicted clear information on heavy precipitation from the evening of 14 August to the afternoon of 15 August. However, they showed rain on the morning of 14 August, which was not depicted by the gauge data. Also, detection and magnitude-based indices were used for the evaluation of the performance of SREs. Before evaluating the performance of any SREs, we first need to understand the reliability and continuity of the gauge data during the selected period. In our case, station 2 did not record any rainfall throughout the study period. Similarly, stations 1 and 9 had data gaps more than 35% of the time, and station 4 had data gaps almost 60%
The remaining five stations had uninterrupted data during the selected period. Detection and magnitude-based indices were calculated at these five stations for the evaluation of the performance of SREs. Table 2 summarizes total rainfall during the selected period at different stations based on gauge data and different SREs. A linear correction factor (CF) on a 10 days event scale was computed for the stations without any data interruption. And the CF was assumed to be 1 (meaning no corrections) for the stations with data gaps. Out of five stations, four stations showed underestimation for PERSIANN and PERSIANN-CCS. Other SREs showed overestimation most of the time except for station 3 for IMERG-Final and TMPA. Thus, we have two main challenges in using such data: (a) filling the gauge data gap and (b) correcting SREs.

Figure 3 shows the performance of different SREs on different hourly, 3, 6, 12, 18, and 24 h scales for the uncorrected SREs. Rainfall detection indices on a coarser temporal resolution showed better values (i.e., toward ideal value shown in Table 1), indicating that SREs perform quite well on a daily scale. However, all SREs had a considerable false alarm, and FAR values were almost similar. On an hourly scale, FAR values were around 0.5, and on a daily scale, they were around 0.25. PERSIANN performed better in FAR and FBI; however, PERSIANN was inferior to IMERG (Early, Late, and Final versions) in terms of POD. PERSIANN-CCS was the most poor-performed SRE.

Table 3 shows the error values of different SREs wrt to the gauge data on an hourly scale for uncorrected SREs. PERSIANN and PERSIANN-CCS highly underestimated rainfall, indicating a requirement of bias correction. IMERG Early and IMERG Late had relatively higher positive biases. IMERG Final version showed significantly improved performance than Early and Late versions. TMPA had the lowest mean deviation and PBIAS; however, RMSE and MAD values were larger, indicating a mismatch between SRE and gauge data on an hourly scale.

### 3.2. Hydrologic Modeling

We conducted seven cases of hydrologic simulation using gauge data and six individual SREs. Since PERSIANN and PERSIANN-CCS underestimated rainfall significantly, we corrected them using above...
mentioned CFs and simulated two more cases. We simulated additional four cases after filling data gaps in gauge data from TMPA and IMERG (Early, Late, and Final versions). Figure 4a shows a comparison plot between simulated and observed hourly discharge at three outlets: (a) Upstream, (b) Midstream, and (c) Downstream during the calibration period of August 1–10, 2012. Though the model did not precisely represent base flow, peaks were in an acceptable range. Since our simulation period is only for an event and we simulated for only 10 days, the initial water and moisture properties are difficult to replicate. Therefore, the baseflow is not mimicked with the current simulation. A long-term hydrologic simulation, incorporating spatially different land use and soil types, is needed for a base flow and water balance analysis. The calibrated roughness coefficients for slopes and river channels were 0.1 and 0.018, respectively. They are analogous with Musumari et al. (2019).

Figure 3. Performance of different satellite-based rainfall estimates with respect to gauge data on different time scales (hourly, 3, 6, 12, 18, and 24 h) in the study area during August 10–19, 2014. POD: probability of detection, CSI: critical success index, FAR: false alarm ratio, and FBI: frequency bias index.

Table 3
Error Evaluation of Different Satellite-Based Rainfall Estimates With Respect to Gauge Data Before Applying Correction During August 10–19, 2014

| Magnitude based indices | TMPA | IMERG early | IMERG late | IMERG final | PERSIANN | PERSIANN-CCS |
|-------------------------|------|-------------|------------|-------------|----------|--------------|
| MD (mm/h)               | 0.07 | 0.27        | 0.35       | 0.13        | −0.51    | −0.43        |
| MAD (mm/h)              | 1.38 | 1.58        | 1.61       | 1.32        | 1.10     | 1.39         |
| RMSE (mm/h)             | 4.07 | 5.19        | 5.32       | 4.15        | 3.91     | 4.67         |
| PBIAS (%)               | 6.48 | 23.66       | 30.65      | 11.49       | −44.82   | −37.52       |

Note. Uniform hourly rain is used for TMPA 3 h resolution.
Abbreviations: MD, mean difference; MAD, mean absolute difference; PERSIANN, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks; PERSIANN-CCS, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System; RMSE, root mean square error; RMSE (refer to Equations 1–4), percentage bias; TMPA, Tropical Multi-satellite Precipitation Analysis.
The developed model was then used for the study period (August 10–19, 2014). Figure 4b shows the hydrologic response of the extreme precipitation event of the selected period. Our result showed that the peak values did not match accurately, but the hydrograph's tendency was congruous. Moreover, a double-peak in observed streamflow could not be replicated by the simulation. More localized and sub-hourly rainfall details may be required to mimic such local features. However, since gauge data suffered some data gaps in a few stations, spatial interpolation from the limited stations might bring underestimation. Therefore, we tried to fill gauge data gaps with different SREs to evaluate this issue. Another possible reason is also associated with uncertainty related to the stage-discharge relationship at high peaks. In addition, tipping bucket type automatic rainfall measurement normally underestimates by about 10% compared with manual rain gauges (Talchabhadel et al., 2017). Moreover, though it records about 10% underestimation on a daily scale, it is again hard to quantify at the sub-daily or hourly level. Therefore, one possible reason for underestimating simulated discharge could be underestimating precipitation estimates by the tipping bucket method.

In the WRR basin, out of nine stations, only three stations are collocated with automatic and manual rain gauges. Therefore, station-wise comparisons were not straightforward possible. Figure 5 shows the spatial distribution of station-wise total precipitation for the selected period based on gauge data of DHM. Figure 5a presents aggregated hourly precipitation for the chosen period, and Figure 5b illustrates 10 days total precipitation based on manual 24 h recorded precipitation. As mentioned earlier, four precipitation stations had data gaps hours (gap values and locations shown in Figure 5a). Apart from those four stations with data gaps, tipping bucket type automatic rainfall recorded almost similar amounts of manually observed 24 h accumulated precipitation during the selected period, indicating the good performance of automatic gauges wrt to manual gauges. The maximum total station-wise precipitation for the chosen period is 620 mm for automatic and 597.5 mm for manual. 24 h accumulated precipitation data from more stations, in total >20 stations (inside and outside the basin), are crucial information for better understanding spatial variability. We need to keep in mind that though these SREs show similar results on a daily scale or accumulated precipitation amounts for a longer period, they have high hour-to-hour fluctuations. These fluctuations result in discrepancies between observed and simulated streamflow at a sub-daily level. If the disintegration of 24 h accumulated precipitation can be performed on an hourly scale, such data would be used in numerical simulation of the basin’s response to extreme precipitation.

A sample of correction of SREs for PERSIANN and PERSIANN-CCS is shown in Figure 6 for station 3 (Table 2 shows CFs for all stations). Since we used a temporally static CF, a prior precipitation increase could
be found due to false alarm values. Therefore, for this study, we employed an event-based linear CF for PERSIANN and PERSIANN-CCS.

Figures 7–9 show hydrologic responses of the study area based on different data sources. Figure 7a shows a comparison plot between observed and simulated discharge from the gauge, PERSIANN, and PERSIANN-CCS. As mentioned earlier, both PERSIANN and PERSIANN-CCS underestimated rainfall, due to which simulated discharges were also highly underestimated. PERSIANN and PERSIANN-CCS show over 40% underestimation of the streamflow. However, the $R^2$ and NSE values are fairly good, indicating that the tendency of variation is well replicated. Figure 7b shows a comparison after bias correction. We found peak discharge was increased to some extent but was still underestimated. PBIAS improved from over $-40\%$ to about $-10\%$. NSE and $R^2$ were also enhanced by a noticeable amount. The bias-corrected products show a good agreement between observed and simulated peaks. However, due to false alarms before a peak rain, there was a double peak nature of the hydrograph, and the false alarms are pronounced after the bias correction. It suggests using a temporally dynamic bias correction on a sub-daily to daily scale rather than using a static CF based on an event.

Figures 8a and 8b show a comparison for IMERG (Early, Late, and Final) and TMPA. They delivered superior results than PERSIANN and PERSIANN-CCS. They show NSE values over 0.65. However, all three IMERG products and TMPA provide an overestimation. All three IMERG products offer an $R^2$ value over 0.74, whereas TMPA shows $R^2$ to be approximately 0.7. The double peak was still visible, but it was comparatively lesser in IMERG Final. By looking into simulation results and performance metrics, IMERG Final proved better in hydrograph nature even though it showed an overestimation (PBIAS = 26%). Similarly, looking at the magnitude of peak value, IMERG Late and TMPA performed better. However, they still possessed significant false alarms before peak rainfall events.

Figures 9a and 9b show a comparison plot between observed and simulated discharge from the gauge and gauge gaps filled using the information of IMERG (Early, Late, and Final) and TMPA. NSE and $R^2$ values improved to over 0.85, and PBIAS is in the range of 5% deviation. We find the simulations were better than gauge rainfall, indicating that SREs are important to fill data gaps of gauge rainfall. And at the complete wash
away of the hydrometeorological monitoring system (though not desired), SREs play a vital role in understanding the hydrologic response of the catchment.

Overall, our results reveal the prospective of SREs’ application for the hydrologic analysis. Based on the presented results, we recommend using IMERG products for the numerical simulation of extreme precipitation. However, this study is based on a single event and does not draw a general conclusion. Therefore, further research is necessary to evaluate the performance of SREs. Furthermore, gauge data are always valuable information. Thus, the number of automatic rainfall gauges should be increased, and these stations should be taken care of with proper continuity and regular maintenance. This study uses a stage-discharge relationship developed based on limited streamflow measurements. In general, qualified hydrologists only measure about 10–15 measurements in a year during low, mid, and high flows (Talchabhadel, Aryal, Kawaike, Yamanoi, & Nakagawa, 2021). Since the streamflow measurement during the extreme instances
is challenging, extreme hourly streamflow has uncertainty. However, for a mean daily discharge, these stage-discharge relations have a good application. We need special measuring techniques to measure the streamflow during such extreme instances. In addition, the introduction of an X-band radar is quite demanding for a precise understanding of the dynamics of heavy precipitation events and numerical simulation of catchment response on such extreme precipitation events. Because, X-band radar has very short wavelengths (2.5–4 cm), the product has a high spatiotemporal resolution (for instance, 250 m and 1 min). However, the observation range is limited to less than 100 km (Talchabhadel, Ghimire, et al., 2021).

4. Conclusions and Recommendation

This study evaluates six SRE products’ performance in capturing an extreme precipitation event using detection and magnitude-based indices and hydrologic simulation. We find that IMERG (Early, Late, and Final versions) and TMPA showed a better result than PERSIANN and PERSIANN-CCS in capturing the magnitude of extreme precipitation. Furthermore, for the selected study period in the WRR basin, this study finds that the IMERG Late version showed a better result than others. Importantly, all SREs replicate the general tendency of precipitation patterns. However, they all possess equally false alarms. These false alarms are critical in operational flood forecasting and early warning dissemination.

This study faced two major challenges related to the rainfall data. One is data gaps of gauge rainfall, and another is biases contaminated with SREs. By comparing gauge and SREs on an hourly scale, this study corrected poor-performed SREs. Since PERSIANN showed a significant underestimation, we applied bias correction and got improved results. At the same time, we also filled the data gaps in ground-based observations using well-performed SREs (IMERG Early, Late, and Final versions and TMPA) and found more agreeable results than gauges only. For a precise replication of the local rainfall scenario and likely river response, we recommend the inclusion of bias correction of SREs on a daily or sub-daily basis. At the same time, the performance of gauge data should be checked properly by comparing it with other data sources. Therefore, we aim to incorporate a dynamic bias correction of SREs in the coming days.

Data Availability Statement

Gauge hydrometeorological data are available in DHM and can be purchased. Space-borne rainfall estimates are freely available on their websites, respectively. Data used in the study will be provided on request.
Acknowledgments
The Japan Society supported this work for the Promotion of Science (JSPS) Postdoctoral Fellowship Program (grant in aid P19052). We want to thank the Department of Hydrometeorology (DHM) for providing the gauge hydrometeorological data.

References
Aghakouchak, A., Behrangi, A., Soroshian, S., Hsu, K., & Amitai, E. (2011). Evaluation of satellite-retrieved extreme precipitation rates across the central United States. Journal of Geophysical Research, 116(2), 1–11. https://doi.org/10.1029/2010JD014741
Almazroui, M. (2011). Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009. Atmospheric Research, 99(3–4), 400–414. https://doi.org/10.1016/j.atmosres.2010.11.006
Bohlinger, P., & Sorteberg, A. (2018). A comprehensive view on trends in extreme precipitation in Nepal and their spatial distribution. International Journal of Climatology, 38(4), 1833–1845. https://doi.org/10.1002/joc.5299
Böhlinger, P., Sorteberg, A., & Sodemann, H. (2017). Synoptic conditions and moisture sources actuating extreme precipitation in Nepal. Journal of Geophysical Research: Atmospheres, 122(23), 12653–12671. https://doi.org/10.1002/2017JD027543
Gautam, D. K., & Dulal, K. (2013). Determination of threshold runoff for flood early warning in Nepalese Rivers. IDRIM Journal, 3(1), 126–136. https://doi.org/10.5595/idrim.2013.0061
Hong, Y., Gochis, D., Cheng, I., Hsu, K., & Soroshian, S. (2007). Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. Journal of Hydrometeorology, 8(3), 469–482. https://doi.org/10.1175/JHM574.1
Hong, Y., Hsu, K.-L., Soroshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. Journal of Applied Meteorology, 43(12), 1834–1853. https://doi.org/10.1175/1520-0450(2004)043<1834:PEFRSM>2.0.CO;2
Hsu, K., Gao, X., Soroshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. Journal of Applied Meteorology, 36(9), 1176–1190. https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSM>2.0.CO;2
Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Kidd, C., et al. (2019). Algorithm Theoretical Basis Document (ATBD) Version 06 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Available at https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf
Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1), 38–55. https://doi.org/10.1175/JHM560.1
ICHARM. (2015). Rainfall-runoff-inundation (RRI) model user’ manual, international center for water hazard and Risk management (ICHARM). Public Works Research Institute (PWRI). Available online https://www.icharm.pwri.go.jp/research/rrri_top.html
Karki, R., Hasson, S. ul, Gerlitz, L., Talchabhadel, R., Schenk, E., Schickhoff, U., et al. (2018). WRF-based simulation of an extreme precipitation event over the Central Himalayas: Atmospheric mechanisms and their representation by microphysics parameterization schemes. Atmospheric Research, 214, 21–35. https://doi.org/10.1016/j.atmosres.2018.07.016
Karki, R., Hasson, S. ul, Schickhoff, U., Scholten, T., & Böhmer, B. (2017). Rising precipitation extremes across Nepal. Climate, 5(1), 4. https://doi.org/10.3390/climate5010004
Karki, R., Talchabhadel, R., Aalto, J., & Baiydi, S. K. (2016). New climatic classification of Nepal. Theoretical and Applied Climatology, 125(3–4), 799–808. https://doi.org/10.1007/s00704-015-1549-0
Krause, P., Boyle, D. P., & Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5, 89–97. https://doi.org/10.5194/adgeo-5-89-2005
Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union, 90(10), 93. https://doi.org/10.1029/2008EO100003
Mahbod, M., Shirvani, A., & Veronesi, F. (2019). A comparative analysis of the precipitation extremes obtained from tropical rainfall-measuring mission satellite and rain gauges datasets over a semi-arid region. International Journal of Climatology, 39(1), 495–515. https://doi.org/10.1002/joc.5824
Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3), 895–900. https://doi.org/10.13031/2013.23153
Musumari, H., Nakagawa, H., Kawaike, K., & Talchabhadel, R. (2019). Numerical simulation of Fluvial inundation considering gate operation numerical simulation of Fluvial inundation considering gate operation. Journal of Japanese Society for Natural Disaster Science, 38(3), 97–108https://doi.org/10.24762/jjuds.38.S06.97
Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology, 10(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6
Nguyen, P., Ombadi, M., Soroshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., et al. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. Hydrology and Earth System Sciences, 22(11), 5801–5816. https://doi.org/10.5194/hess-22-5801-2018
Pakokkung, K., & Takagi, M. (2016). Effect of satellite based rainfall products on river basin responses of runoff simulation on flood event. Modeling Earth Systems and Environment, 2(3), 1–14. https://doi.org/10.1007/s40808-016-0200-0
Puadl, P., Regnem, shital B., & Upadhyay, S. N. (2013). Overview of June 2013 flood and landslides with focus on Darchula disaster. Hydro Nepal(13), 57–63.
Pokharel, B., Wang, S.-Y. S., Meyer, J., Marahatta, S., Bhakta, B., Chikamoto, Y., & Gillies, R. (2019). The east–west division of changing precipitation in Nepal. International Journal of Climatology, 40(July), 3348–3359. https://doi.org/10.1002/joc.6401
Pombo, S., & de Oliveira, R. P. (2015). Evaluation of extreme precipitation estimates from TRMM in Angola. Journal of Hydrology, 523, 663–679. https://doi.org/10.1016/j.jhydrol.2015.02.014
ReliefWeb. (2014). Floods and Landslides DREF operation MDRNP007; Emergency Plan of Action (EPoA). Retrieved January 14, 2020, from https://reliefweb.int/report/nepal/nepal-floods-and-landslides-dref-operation-mdrnp007-emergency-plan-action-epoa
Sayama, T., Ozawa, G., Kawakami, T., N abesaka, S., & Fukushima, K. (2012). Rainfall–runoff–inundation analysis of the 2010 Pakistan flood in the Kabul River basin. Hydrological Sciences Journal, 57(2), 298–312. https://doi.org/10.1080/02626667.2011.644245
Sayama, T., Tatebe, Y., Iwami, Y., & Tanaka, S. (2015). Hydrologic sensitivity of flood runoff and inundation: 2011 Thailand floods in the Chao Phraya River basin. Natural Hazards and Earth System Sciences, 15(7), 1617–1630. https://doi.org/10.5194/nhess-15-1617-2015
Shrestha, A. (2016). Cloudbursts in the Nepal Himalayas: Interaction between the Indian monsoon and extratropics. University of Wisconsin-Madison.
Shrestha, D., Singh, P., & Nakamura, K. (2012). Spatiotemporal variation of rainfall over the central Himalayan region revealed by TRMM Precipitation Radar. Journal of Geophysical Research, 117(D22), D22106. https://doi.org/10.1029/2012JD018140
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Soroshian, S., & Hsu, K. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Reviews of Geophysics, 56(1), 79–107. https://doi.org/10.1029/2017RG000574
Talchabhadel, R., Aryal, A., Kawaike, K., Yamanoi, K., & Nakagawa, H. (2021). A comprehensive analysis of projected changes of extreme precipitation indices in West Rapti River basin, Nepal under changing climate. *International Journal of Climatology, 41*(S1), E2581–E2599. https://doi.org/10.1002/joc.6866

Talchabhadel, R., Aryal, A., Kawaike, K., Yamanoi, K., Nakagawa, H., Bhatta, B., et al. (2021). Evaluation of precipitation elasticity using precipitation data from ground and satellite-based estimates and watershed modeling in Western Nepal. *Journal of Hydrology: Regional Studies, 33*, 100768. https://doi.org/10.1016/j.ejh.2020.100768

Talchabhadel, R., Ghimire, G. R., Sharma, S., Dahal, P., Panthi, J., Baniya, R., et al. (2021). Weather radar in Nepal: Opportunities and challenges in a mountainous region. *Weather*, 1–5. https://doi.org/10.1002/wea.3994

Talchabhadel, R., Karki, R., & Parajuli, B. (2017). Intercomparison of precipitation measured between automatic and manual precipitation gauge in Nepal. *Measurement, 106*, 264–273. https://doi.org/10.1016/j.measurement.2016.06.047

Talchabhadel, R., Karki, R., Thapa, B. R., Maharjan, M., & Parajuli, B. (2018). Spatio-temporal variability of extreme precipitation in Nepal. *International Journal of Climatology, 38*(11), 4296–4313. https://doi.org/10.1002/joc.5669

Talchabhadel, R., & Sharma, R. (2014). Real time data analysis of West Rapti River Basin of Nepal. *Journal of Geoscience and Environment Protection*, 2(05), 1–7. https://doi.org/10.4236/gep.2014.25001

Tarek, M. H., Hassan, A., Bhattacharjee, J., Choudhury, S. H., & Badruzzaman, A. B. M. (2017). Assessment of TRMM data for precipitation measurement in Bangladesh. *Meteorological Applications, 24*(3), 349–359. https://doi.org/10.1002/met.1633

Yoshimoto, S., & Amarnath, G. (2017). Applications of satellite-based rainfall estimates in flood inundation modeling—A case study in Mundeni Aru River Basin, Sri Lanka. *Remote Sensing, 9*(10), 998. https://doi.org/10.3390/rs9100998