Hydrometeorological Assessment of Satellite and Model Precipitation Products over Taiwan

PIN-LUN LI,^a LIAO-FAN LIN,^b AND CHIA-JENG CHEN^a

^a *National Chung Hsing University, Taichung, Taiwan*

^b *Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado*

(Manuscript received 30 June 2021, in final form 31 August 2021)

ABSTRACT: Satellite and model precipitation such as the Global Precipitation Measurement (GPM) data are valuable in hydrometeorological applications. This study investigates the performance of various satellite and model precipitation products in Taiwan from 2015 to 2017, including data derived from the Integrated Multisatellite Retrievals for GPM Early and Final Runs (IMERG_E and IMERG_F), Global Satellite Mapping of Precipitation in near-real time (GSMaP_NRT), and the Weather Research and Forecasting (WRF) Model. We assess these products by comparing them against data collected from 304 surface stations and gauge-based gridded data. Our assessment emphasizes factors influential in precipitation estimation, such as season, temperature, elevation, and extreme event. Further, we assess the hydrological response to each precipitation product via continuous flow simulation in two selected watersheds. The results indicate that the performance of these precipitation products is subject to seasonal and regional variations. The satellite products (i.e., IMERG and GSMaP) perform better than the model (i.e., WRF) in the warm season and vice versa in the cold season, most apparently in northern Taiwan. For selected extreme events, WRF can simulate better rainfall amount and distribution. The seasonal and regional variations in precipitation estimation are also reflected in flow simulations: IMERG in general produces the most rational flow simulation, GSMaP tends to overestimate and be least useful for hydrological applications, while WRF simulates high flows that show accurate time to the peak flows and are better in the southern watershed.

SIGNIFICANCE STATEMENT: Precipitation data derived from satellites or numerical weather prediction models are valuable resources since they can provide comprehensive information regarding areal precipitation over a specific region. However, such precipitation products of varying degrees of accuracy may hinder their usefulness for hydrological and other applications. Understanding which precipitation products to use under various circumstances requires knowledge accrued from scrutinizing the relative performance of these products. This study shows that over Taiwan, the performance of satellite and model precipitation is contingent upon season, region, and event. The best-performing precipitation products can thus generate most rational flow simulations, suggesting possibilities for more emerging applications.

KEYWORDS: Precipitation; Hydrology; Satellite observations; Hydrologic models; Numerical weather prediction/forecasting

1. Introduction

Precipitation is an essential variable in the hydrological cycle. Obtaining accurate precipitation data can crucially aid climate and hydrological research and applications, such as effective allocation of water resources and reducing the risk of flooding disasters. Currently, the most accurate precipitation data still rely on surface observation; however, establishing surface stations is often confined by geographical locations and topographical considerations, resulting in gaps in spatial precipitation information (Qin et al. 2014). Moreover, surface stations are not demonstrative of precipitation processes over a larger area, especially with topographically complex mountain

regions (Zhang et al. 2013). Owing to modern-day scientific progress, precipitation products have become increasingly diverse, and the aforementioned information gaps can be filled using remote sensing technologies, such as rainfall radars and satellites. Among the mentioned products, radar precipitation products boast superior performance in precipitation estimation, but their installation is still constrained by topographical complexity and funding limitations (Behrangi et al. 2011; Zhang et al. 2018). By contrast, satellite precipitation products are capable of monitoring all areas across the globe, thereby compensating the inadequate distribution of radars and surface stations.

In recent years, satellites of the Global Precipitation Measurement (GPM; Hou et al. 2014) mission have played an important role in the information transmission of Earth's water and energy cycles. Built upon the success of the Tropical Rainfall Measuring Mission (TRMM), the GPM mission launched its *Core Observatory* satellite, carrying pioneering radar and radiometer systems, to measure precipitation from space (Huffman et al. 2019). GPM currently

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-21-0130.s1>.

Corresponding author: Chia-Jeng Chen, cjchen@nchu.edu.tw

employs two main satellite precipitation products: the Global Satellite Mapping of Precipitation Project (GSMaP) developed by the Japan Aerospace Exploration Agency (JAXA) and the Integrated Multisatellite Retrievals for GPM (IMERG) developed by the U.S. National Aeronautics and Space Administration (NASA). There has been a proliferation of comparative studies into the precipitation estimation performance of a multitude of satellite precipitation products based on critical factors associated with precipitation estimation, including location (e.g., latitude), landform, rain pattern, and density of surface stations (e.g., [Prakash et al. 2016, 2018](#); [Bharti et al. 2016](#); [He et al. 2017](#); [Yong et al. 2016](#); [Kim et al. 2017](#); [Tian et al. 2018](#); [Huang et al. 2018](#); [Tan et al. 2018](#); [Zhang et al. 2018](#); [Wang and Yong 2020](#)). Since GPM began releasing its precipitation product around 2015, it has attracted global users to perform studies assessing the pertinent precipitation over various-scale regions in different countries, such as the United States in North America and China, India, South Korea, and Japan in Asia. Similar inquiries into Taiwan should stimulate another scientific interest since Taiwan presents its uniqueness of geophysical distributions and weather patterns. Several recent studies have been conducted to assess different satellite precipitation products in depicting rainfall variations over Taiwan (e.g., [Huang et al. 2021a,b](#)), yet these studies have not fully addressed the usefulness of satellite precipitation products, especially in the case of hydrological applications. In Taiwan, precipitation patterns, as well as hydrological responses, show sharp variations in different seasons, suggesting a need for a seasonal assessment of various precipitation products and their usefulness for hydrological modeling.

Precipitation data can also be obtained by running complex calculations (e.g., cumulus and convection parameterization) through numerical weather prediction (NWP) models ([Zhang et al. 2013](#)). In [Zhang et al. \(2018\)](#), an assessment of precipitation products in the Midwest U.S. indicated that the Weather Research and Forecasting (WRF) Model yields better estimates than IMERG in cases of the cold season and vice versa when it comes to the warm season. In light of their finding, this study also includes WRF in the assessment to examine whether or not such seasonal or temperature correlation also applies to such geographically distinct region as Taiwan. Numerous studies have examined the performance of WRF in simulating precipitation over Taiwan. On some occasions the examination of WRF was focused on the ensemble simulation of intensive typhoon precipitation based on specific model configurations (e.g., [Fang et al. 2011](#); [Hong et al. 2015](#)). Other studies emphasized more on the depiction of the influence of different model physics on simulation ([Tao et al. 2011](#)) or the WRF-based ensemble prediction system ([Li et al. 2020](#)). An event-based assessment of flood simulation in response to rainfall forecasted by WRF was also of interest ([Hsiao et al. 2013](#)). The above useful references not only inform us of a reasonable WRF configuration that should be used in our study, but also show a research opportunity that a concordant assessment of both satellite and model precipitation products should be conducted.

Effective utilization of water resources has been a long-standing challenge in Taiwan because of uneven distribution of rainfall over time and space. Many studies have applied

hydrological modeling to the assessment of the short- and long-term variations in surface and subsurface water resources in Taiwan (e.g., [Yu et al. 2002](#); [Yu and Wang 2009](#); [Li et al. 2009](#); [Shih et al. 2019](#)). To better forecast surface runoff in response to heavy precipitation events, several studies coupled hydrological modeling with precipitation forecasts generated by NWP models (e.g., [Hsiao et al. 2013](#); [Yang and Yang 2014](#); [Wu and Lin 2017](#)). While event-based assessments with model precipitation have been conducted, there exists a research gap in a continuous assessment of hydrological modeling driven by satellite precipitation for Taiwan. We thus take a step further to apply selected satellite and model precipitation products in question toward continuous rainfall-runoff simulation, offering an extended assessment of each product's usefulness in terms of hydrological modeling.

Using observation data collected between 2015 and 2017, this study first analyzes the performance of IMERG, GSMaP, and WRF in producing daily rainfall estimates in Taiwan, together with the different factors that might affect a product's performance, such as season, temperature, elevation, and extreme event. Furthermore, each precipitation dataset is applied to the Hydrologic Modeling System (HMS; [USACE 2000](#)) for the simulation of daily flows at two selected watersheds, in order to examine how accurate the simulated flows are in response to different precipitation products. HMS is a widely used, reliable platform that provides an assortment of methods for modeling different components of hydrologic cycles ([Darbandsari and Coulibaly 2020](#)). It can be used for both continuous and event-based simulations, and successful applications can be found in many regions including Taiwan ([Yang and Yang 2014](#)).

The rest of this paper is arranged as below: [section 2](#) introduces the study areas and satellite precipitation and observed data being used. [Section 3](#) describes the configuration of WRF and hydrologic modeling as well as the performance metrics for assessing precipitation products. [Section 4](#) presents results and discussion of the point, gridded and hydrological analyses. Last, [section 5](#) provides the conclusions and recommendations.

2. Materials

a. Study area

The targeted area of this study is the main island of Taiwan, located in East Asia ([Fig. 1](#)). The main island has an area of $\approx 36\,000\text{ km}^2$ and features rapid variations in topography, with high mountain ranges ($>3500\text{ m}$) north-south elongated over the central region and low-lying plains mostly over the western region. In Taiwan, through the regular four seasons, there are, however, five distinct rainy seasons/patterns being identified ([Chen and Chen 2003](#)) and categorized as spring rain (March–April), plum rain or mei-yu (May–June), summer rain (July–September), autumn rain (October–November), and winter rain (December–February). Due to strong synoptic systems or typhoons, extreme precipitation events usually occur during the plum rain and summer rain seasons, respectively. To further analyze hydrological responses to different rainfall products at a smaller scale, two watersheds at which the predominant rainfall types in Taiwan can be found are selected: one is the

FIG. 1. (a) WRF's three-level nested grid modeling domains, and (b) the main island of Taiwan and its elevation map, labeled with the Feitsui Reservoir watershed and Bajhang River watershed.

Feitsui Reservoir watershed located at northern Taiwan, and the other is the Bajhang River watershed in southwestern Taiwan (Fig. 1b). The Feitsui and Bajhang watersheds bear comparable size in drainage area (303 versus 475 km²); however, due to different geographic patterns/locations, these two watersheds possess distinct seasonal rainfall patterns, as the latter has shown less coherent spatial distribution of rainfall than the former does. These two watersheds have also been extensively studied in terms of hydrological applications with a calibrated/validated HMS model in Tseng et al. (2020).

b. Integrated Multisatellite Retrievals for GPM (IMERG) Early and Final Run

This and the following subsections provide a high-level overview of satellites precipitation data used in this study. Associated details regarding the derivation algorithms are not within our scope, so for better elaboration of such details please refer to officially released documents (e.g., Aonashi et al. 2009; Huffman et al. 2015, 2019).

As far as remote sensing has advanced, passive microwave (PMW) sensors provide a relatively accurate precipitation estimation via low-Earth-orbit (LEO) satellites, and IMERG combines data from all PMW sensors in the GPM constellation satellites. Over regions with limited LEO satellite data, IMERG uses infrared (IR) radiation in the geosynchronous Earth orbit (GEO) to fill the spatial/temporal gaps of PWM precipitation estimates. IMERG employs various algorithms to enhance precipitation estimates, including the morphing Kalman

filter Lagrangian time interpolation and the PERSIANN-Cloud Classification System (Huffman et al. 2019). Moreover, IMERG also introduces precipitation data from observation stations for posterior processing and enhancement of precipitation estimates. The source of station data is mainly from the Global Precipitation Climatology Centre (GPCC) of Deutscher Wetterdienst (DWD) in Germany, which gathers monthly weather reports from 7000 to 8000 observation stations around the globe [e.g., from the National Oceanic and Atmospheric Administration/Climate Prediction Center (NOAA/CPC), the Japan Meteorological Agency (JMA), and the Meteorological Office of the United Kingdom] before sorting and organizing them into collected datasets.

IMERG data cover areas between 60° north and south of the equator, with high spatial resolution (0.1°) and temporal resolution (half-hourly, hourly, and daily) and come in three different dataset versions: Early Run, Late Run, and Final Run. The three datasets differ in latency and accuracy: IMERG Early Run is forward morphed with a 4-h latency, the Late Run is both forward and backward morphed with a 12-h latency, while the Final Run, besides having 3.5 months of latency and being both forward/backward morphed, uses a climatological adjustment that incorporates GPCC's gauge data to derive a fixed "ratio multiplier" for each month. This study has applied IMERG data of the latest Version 06. Several studies (e.g., O et al. 2017; Huang et al. 2021a) indicated that the difference between IMERG Early Run and Late Run is insignificant, so this study emphasizes the Early Run and Final Run (Huffman et al. 2019), which have spatial and temporal resolutions of 0.1° and half-hourly, respectively. All the above data can be found

TABLE 1. List of intensity categories and meteorological parameter values associated with the three selected typhoon events examined in this study.

Name	Life period	Intensity category near land ^a	Min central pressure (hPa)	Max wind speed (kt)	Largest radius of storm wind (km)	Largest radius of gale wind (km)
Haitang	28 Jul 2017–31 Jul	3	985	45	0	440
Nepartak	3 Jul 2016–9 Jul	5	900	110	110	440
Soudelor	1 Aug 2015–11 Aug	5	900	115	370	650

^a Based on ESCAP/WMO Typhoon Committee's intensity scale.

on GPM's official website (<https://pmm.nasa.gov/data-access/downloads/gpm>).

c. Global Satellite Mapping of Precipitation (GSMaP) project

Similar to the IMERG products, GSMaP's core algorithms include PMW precipitation retrieval (Shige et al. 2009) and microwave-infrared merged algorithms (Ushio et al. 2009). In addition, the algorithms make use of globally merged full-resolution IR data from $\sim 11\text{-}\mu\text{m}$ IR channels from geostationary satellites of JMA, the European Meteorological Satellites Organization, and NOAA (Janowiak et al. 2001). Other ancillary input data to the algorithms include atmospheric conditions provided by JMA Global Analysis (GANAL) and Merged satellite and in situ data Global Daily Sea Surface Temperatures (MGDSST) (Kubota et al. 2007, 2020).

GSMaP's data (available at <https://sharaku.eorc.jaxa.jp/GSMaP/>) cover the region from 60°N to 60°S with spatial resolutions of 0.25° or 0.1° and temporal resolutions of hourly or daily. Currently, the available versions include v6 and v7, both of which have derivative products such as GSMaP_Now, GSMaP_NRT (near-real time), GSMaP_MVK (microwave), GSMaP_NRL (reanalysis), GSMaP_Gauge (gauge calibration), and GSMaP_RNC (RIKEN Nowcast). GSMaP_NRT with a similar short latency was chosen for comparison against IMERG Early Run. The GSMaP v6 was adopted since v7 was still unavailable when this study was completed.

d. Station data and Taiwan Climate Change Projection and Adaptation Information Platform (TCCIP) data

Station data applied in this study are obtained from hourly precipitation data posted by Taiwan's Central Weather Bureau (CWB), which collects data through its 600 weather stations in Taiwan's main island. To ensure the quality of station data, this study selects stations with less than 10% of missing values during the time period investigated, which total 304 stations. Because station (point) data often face uneven spatial/temporal distribution, the Taiwan Climate Change Projection and Adaptation Information Platform (TCCIP) has integrated historical observation records from various sources including the CWB, Civil Aeronautics Administration, Water Resources Agency, Environmental Protection Administration, and Taiwan Power Company to create a gridded climate (mainly daily precipitation and temperature) database that is reliable over the long term (Weng and Yang 2018). TCCIP first models daily precipitation series as a latent Gaussian variable (LGV) in which the parameters are fitted using maximum likelihood

estimation. The LGV, a reversible monotonic function, is then decomposed using empirical orthogonal functions (EOFs). If missing values in daily precipitation series are present, TCCIP incorporates the spatial optimal interpolation (Simolo et al. 2010) into the incomplete EOF decomposition (Smith et al. 1996) to correct the associated eigenvectors. Last TCCIP applies natural neighbor interpolation (Sibson 1981; Watson 1994) to the corrected eigenvectors to produce gridded eigenvectors, and then reverses the LGV function to obtain gridded precipitation. The current available TCCIP dataset is at a 5-km spatial resolution and daily temporal resolution, and accessible for the years of 1960–2017 (when this study was completed; it is now up to 2019). As the best available observational gridded data for Taiwan (Chen et al. 2020; Henny et al. 2021), the TCCIP precipitation dataset is referred to as the benchmark data for gridded analysis of satellite and model precipitation products. This study also employs the TCCIP average temperature dataset to analyze the precipitation–temperature relationship, which is presented as follows.

e. Selected heavy rainfall events

In our assessment we also focus on several heavy rainfall events; two of these events are induced by plum rain fronts in 2015 and 2017, and the other three events are induced by typhoons. We acquire typhoon-related data from Digital Typhoon (<http://agora.ex.nii.ac.jp/digital-typhoon/>) to provide a list of typhoon parameter values (Table 1) as well as track information (Fig. 2) for the three selected typhoon events. We also examine possible impact on the performance of precipitation estimates involving/due to typhoon characteristics (please see discussions in section 4b).

3. Methodology

Precipitation data assessment begins with using surface station and TCCIP data as the benchmark for point and gridded analyses, followed by hydrological analyses that apply each precipitation product toward hydrologic modeling for performance evaluation of rainfall–runoff simulation. Zhang et al. (2018) assessed the quality of precipitation products over the central United States using IMERG and WRF data; similarly, the WRF modeled data are also assessed versus IMERG the satellite data in this study focusing on the Taiwan region. We determine the period of the assessment from 2015 to 2017 due to data availability and accessibility. We use the data from 2015 to mainly ensure the highest consistency in IMERG precipitation estimates, as IMERG in earlier years (2000–14) is based

FIG. 2. Detailed track information acquired from Digital Typhoon (available at <http://agora.ex.nii.ac.jp/digital-typhoon/>) for (a) Typhoon Haitang in 28–31 Jul 2017, (b) Typhoon Nepartak in 3–9 Jul 2016, and (c) Typhoon Soudelor in 1–11 Aug 2015. Different colors indicate tropical cyclone categories based on ESCAP/WMO (Economic and Social Commission for Asia and the Pacific/World Meteorological Organization) Typhoon Committee's intensity scale, which are 1) blue: tropical depression (category 2), 2) green: tropical storm (category 3), 3) yellow: severe tropical storm (category 4), 4) red: typhoon (category 5), and 5) magenta: extratropical cyclone. The number in each circle indicates the date in the pertinent month.

on the operation of the “TRMM” satellite while from 2015 IMERG is based on the “GPM” satellite. To unify precipitation data coming from different sources, we apply several preprocessing steps to each precipitation product as follows. We utilize bilinear interpolation to ensure each product to follow the same spatial resolution as 0.1° (~ 10 km). Moreover, to align satellite precipitation products with Taiwan local time, we compute daily precipitation data by accumulating hourly data beginning from 1600 UTC. In the following subsections, we briefly describe the WRF configuration, precipitation assessment process, and hydrologic modeling setup.

a. WRF configuration

The numerical model precipitation product adopted in this study is produced by the WRF Model (version 3.9) with the Advanced Research WRF (ARW) solver. We use the Final Reanalysis (FNL) data at the resolutions of 6 h and 1° , generated by the National Centers for Environmental Prediction (NCEP) to create the initial and boundary conditions for WRF. Other geographical information (e.g., elevation) is set to WRF's default, except that the default land use data have been replaced with local land use data released by National Land Surveying and Mapping Center (NLSC) in Taiwan. The

replacement with land use data was suggested in Chen et al. (2020) that using the NLSC land use data can significantly improve WRF simulation for Taiwan locally.

The WRF configuration follows the three-level nested domains, used by Chen et al. (2020) and is shown in Fig. 1a: The horizontal resolutions of domains are respectively 18, 6, and 2 km (D1, D2, and D3) with a total of 45 vertical levels. The Betts–Miller–Janjić scheme (Janjić 1994) is applied for cumulus parameterization except for the second and third nested domains, which do not undergo cumulus parameterization. Additional parameterization schemes include the Rapid Radiative Transfer Model (RRTM) longwave radiation scheme (Mlawer et al. 1997), the Goddard shortwave radiation scheme (Chou and Suarez 1994), WRF Single-Moment 5-class (WSM5) microphysics scheme (Hong et al. 2004), and Yonsei University (YSU) boundary layer scheme (Hong et al. 2006). To calculate surface processes (e.g., soil moisture and heat fluxes), this study adopts the Noah land surface model (Livneh et al. 2011). It should be noted that while the discussions of Chen et al. (2020) focused on central Taiwan, their simulations were conducted over the whole Taiwan. The above parameterization schemes also follow those commonly used in (the whole) Taiwan (Fang et al. 2011; Tao et al. 2011; Hsiao et al. 2013; Hong et al. 2015; Li et al. 2020). To obtain

model results for a specified day, the simulation begins at 1200 UTC of the previous day and ends at 0000 UTC of the next day, yielding a total simulation time of 36 h. In each simulation, the first 12 h are for model spinup, and the following 24-h precipitation outputs from D3 are used to produce daily precipitation data. This simulation procedure is repeated to obtain all daily precipitation data from 2015 to 2017.

b. Precipitation data assessment process

Given that surface station data produce best precipitation estimates at the point scale, we first adopt station data as the benchmark for point analysis. We use Euclidean distance to sort out the closest grid point to the target station for point analysis, and then repeat this process for all the stations. After mapping each precipitation product's gridded data to the stations, we adopt Pearson's correlation coefficient (CC) as the performance metric in evaluating the correlation between the gridded and station time series:

$$CC = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (1)$$

where X represents station data and Y represents gridded data of TCCIP, IMERG_E, IMERG_F, GSMaP, or WRF. Additional metrics that evaluate quantitative differences such as bias or mean absolute deviation (MAD) introduced below are not used for point-scale analysis since precipitation amount measured at a point/station by definition is not comparable to mean areal/gridded precipitation.

The second part of precipitation assessment is gridded analysis, for which we adopt the TCCIP dataset as the benchmark. In addition to Pearson's correlation, the performance metrics used in gridded analysis include bias [Eq. (2)] and MAD [Eq. (3)]:

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n (X_i^e - X_i^r), \quad (2)$$

$$\text{MAD} = \frac{1}{n} \sum_{i=1}^n |X_i^e - X_i^r|, \quad (3)$$

where n stands for number of grids, X_i^e refers to gridded data of IMERG_E, IMERG_F, GSMaP, or WRF, and X_i^r refers to the TCCIP gridded data. The calculation approach to correlation coefficient is the same as with point analysis, with the benchmark data replaced by the TCCIP gridded data.

c. Hydrologic modeling setup and flow-related metrics

Regarding the hydrological analysis, we adopt the HMS model, developed by the Hydrologic Engineering Center (HEC) of the U.S. Army Corps of Engineering (USACE 2000). HEC-HMS is composed of three main components: basin model, meteorological model, and control specifications. Basin model, in charge of calculating rainfall-runoff, is used to construct hydrologic elements within the river basin (e.g., subbasins, reaches, and junctions) and to configure associated parameters at each element (e.g., precipitation loss, direct runoff transformation, and baseflow). Meteorological model serves to provide the

meteorological information needed for hydrologic modeling, such as precipitation and evapotranspiration. We apply Thiessen's polygon method on station point data to distribute precipitation to each basin. By contrast, for gridded datasets, precipitation is directly distributed based on the grid/basin area ratio and then inputted in HEC-HMS. Last, control specifications are used to configure the start/end time and the time step of the simulation run. It is worthy of note that HMS has the temperature index as the only snowmelt method, and uses the fixed monthly estimated potential evapotranspiration (PET); however, in our case, snowmelt is inactive, and our well-calibrated results demonstrate that the fixed monthly PET only yields insignificant influence on low-flow simulation.

Before running HMS with different precipitation products, we have followed the standard practice in hydrological modeling (Smith et al. 2004; Behrangi et al. 2014; Shah and Mishra 2016) to calibrate and validate the model using measured precipitation and flow gauge data at the two watersheds. The Nash-Sutcliffe efficiency (NSE) values (Nash and Sutcliffe 1970) for both the Feitsui Reservoir and Bajhang River watersheds are higher than 0.8 in the validation periods, indicating the well-calibrated parameters. For detailed explanation on the calibration and validation processes, please refer to our previous study (Tseng et al. 2020). After calibrating and validating HMS, we are thus able to run the model with different precipitation products to simulate corresponding flows, which are then compared against the measured precipitation-based simulated flows (hereafter referred to as "Gauge"). The calibration of HMS based on the most accurate gauge data, which determines the best model parameters, can better illustrate relative flow differences only propagated from the differences in precipitation products instead of model parameters.

To further evaluate the usefulness of each precipitation product for continuous hydrological modeling, we introduce five additional performance metrics (Nash and Sutcliffe 1970; Razavi and Coulibaly 2017; Darbandsari and Coulibaly 2020). These metrics are NSE, NSE based on flows more than 90 percentile (NSE90), NSE based on flows less than 10 percentile (NSE10), NSE based on squared transformed data (NSES), and absolute volume error (VE), which can be expressed as below:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (Q_{S_i} - Q_{O_i})^2}{\sum_{i=1}^N (Q_{O_i} - \bar{Q}_O)^2}, \quad (4)$$

$$\text{NSE90} = 1 - \frac{\sum_{i=1}^N (Q_{S90_i} - Q_{O90_i})^2}{\sum_{i=1}^N (Q_{O90_i} - \bar{Q}_{O90})^2}, \quad (5)$$

$$\text{NSE10} = 1 - \frac{\sum_{i=1}^N (Q_{S10_i} - Q_{O10_i})^2}{\sum_{i=1}^N (Q_{O10_i} - \bar{Q}_{O10})^2}, \quad (6)$$

FIG. 3. Results of point analysis from 2015 to 2017: (a)–(e) spatial distribution of correlation coefficient for each precipitation product; (f)–(j) mean precipitation at stations over the northern region of Taiwan; (k)–(o), (p)–(t), (u)–(y) as in (f)–(j), but for the central, southern, and eastern regions, respectively.

$$\text{NSES} = 1 - \frac{\sum_{i=1}^N (Q_{S_i}^2 - Q_{O_i}^2)^2}{\sum_{i=1}^N (Q_{O_i}^2 - \bar{Q}_O^2)^2}, \quad (7)$$

$$\text{VE} = \frac{\sum_{i=1}^N |Q_{S_i} - Q_{O_i}|}{\sum_{i=1}^N Q_{O_i}}, \quad (8)$$

where Q_{S_i} and Q_{O_i} denote the simulated and observed (Gauge in our case) flows for day i , respectively, Q_{S90_i} and Q_{O90_i} represent the simulated and observed more than 90th percentile, respectively, and Q_{S10_i} and Q_{O10_i} represent the simulated and observed less than 10th percentile, respectively. All the NSE-related metrics range from $-\infty$ to 1 (perfect), and VE ranges from 0 (perfect) to ∞ . Among these metrics, NSE and VE can assess the overall, long-term performance of the simulation. NSE90 and NSES can reflect the accuracy of high flows, while NSE10 is the indicator of low-flow accuracy.

4. Results and discussion

a. Point analysis

We first calculate the correlation between the station data and each of the gridded precipitation products (i.e., TCCIP, IMERG_E, IMERG_F, GSMaP, and WRF). Figure 3 is an illustration of the spatial distribution of correlation coefficients of each precipitation product. We can see high correlation coefficients (>0.9) between the station and TCCIP data, justifying the application of TCCIP as the benchmark for succeeding gridded analysis. Figures 3b–e demonstrate apparent regionality to the spatial distribution of correlation coefficients of IMERG_E, IMERG_F, GSMaP, and WRF. In particular, IMERG products display comparatively higher correlation coefficients in the southwestern region. To further examine regional features, we divide the entire domain into four main regions, i.e., north, central, south, and east in Taiwan, and calculate each region's average daily precipitation and average correlation coefficient for the respective products (Figs. 3f–y and Table 2). The results demonstrate that, in the northern region, WRF performs slightly better than the satellite products in terms of temporal variations but tends to overestimate rainfall. Over the other three regions, IMERG products outperform GSMaP and WRF in terms of correlation coefficients, and the latter two still exhibit some level of overestimation.

We further perform the point analysis within five rainy seasons (see section 2a) and show the result in Fig. 4. Based on the rainfall of each region, it is clear that northern and eastern Taiwan receive stable amount of rainfall ($>5 \text{ mm day}^{-1}$) throughout the seasons. By contrast, central and southern Taiwan display clear patterns of dry and wet seasons. Additionally, correlation coefficients of all precipitation products seem to coincide with changes in seasonal precipitation; that is, when rainfall increases, the correlation coefficient increases and vice versa. The results further suggest that IMERG_E and IMERG_F bear similar

TABLE 2. Regional average precipitation (avg.) of each of the five precipitation products, and its respective mean correlation coefficient (cc.) associated with the precipitation at stations, over the four regions of Taiwan. The closest average precipitation to the gauge value (with the least negative bias) and the highest correlation coefficient among the four precipitation products (excluding the benchmark TCCIP) for each region are highlighted in bold.

	TCCIP	IMERG_E	IMERG_F	GSMaP	WRF
Product avg.					
North	6.37	4.91	6.40	7.52	10.51
Central	4.87	5.07	5.96	6.54	8.09
South	6.15	5.81	5.96	10.49	9.19
East	6.61	4.49	5.53	8.49	14.04
Product cc.					
North	0.93	0.60	0.60	0.58	0.65
Central	0.96	0.70	0.68	0.58	0.62
South	0.96	0.79	0.78	0.69	0.64
East	0.95	0.71	0.73	0.63	0.66

performance. From the figure, IMERG_F shows a clearer advantage over IMERG_E in 2015, regardless of the rainy seasons. Since 2016, the distinction between IMERG products has become less distinct, except the winter rain in 2016. The reason IMERG_F exhibits no clear advantage in Taiwan could be attributed to its use of its use of station data from a lower number of gauges for the adjustment process. We preliminarily compared the gauge coverage in Taiwan with that within the United States and found that the coverage in Taiwan is approximately 4–5 times less on average (not shown). In addition, the result also indicates that satellite precipitation is superior to model precipitation in the central, south, and east.

Wang and Yong (2020) suggested that the performance of satellite precipitation estimates varies according to elevation at the global scale, partly due to inaccurate IR rainfall algorithms at high-elevation regions with relatively warm clouds. Their finding thus motivates our following analysis of the dependency of performance metrics on elevation in Taiwan. Figure 5 shows the average correlation coefficient of each product at different elevation intervals from 2015 to 2017. The number of stations used in each elevation interval is also shown in the figure. The result indicates that, in low-altitude regions (30–230 m), the performance of IMERG and GSMaP deteriorates as elevation increases; however, in intermediate to high altitudes ($>400 \text{ m}$), the satellite products perform better as elevation increases. Our finding falls largely in line with that of Wang and Yong (2020), who also indicated that satellite precipitation products perform better in the elevation range from 400 to 1500 m based on their quasi-global evaluation. We find that WRF exhibits an increasing correlation with elevation that seemingly implies WRF's precipitation being better simulated with higher elevation, especially toward the highest. We also find no significant discrepancy in the relative performance of different precipitation products at different elevation intervals, with IMERG products still holding the best performance, followed by WRF and GSMaP. The above findings suggest a need for further investigations (e.g., O and Kirstetter 2018) into how satellite rainfall

FIG. 4. Average daily rainfall by seasons (bars, associated with the left y axis) and the average correlation coefficient between station data and each precipitation product (lines, associated with the right y axis) in the four major regions of Taiwan. Winter is not shown in 2017 due to the unavailable January and February of 2018.

algorithms or WRF precipitation physics are sensitive to different rainfall mechanisms in Taiwan.

The well-performing satellite precipitation over high-elevation regions can be related to the inclusion of the elevation indicator

as well as microwave sensors in the development of datasets (El Kenawy et al. 2015; Liu et al. 2020), despite the difficulty of IR rainfall algorithms in detecting precipitation at high-elevation regions, especially for light rain (Hirpa et al. 2010;

FIG. 5. Average correlation coefficient between station data and each precipitation product at different elevation intervals; in each of these intervals at least eight stations are used for calculation.

FIG. 6. Boxplots of monthly data for (a) TCCIP average daily precipitation and (b)–(d) performance metrics pertaining to each precipitation product, including bias, MAD, and correlation. Circles, thick lines (boxes), and thin lines indicate mean values, interquartile ranges, and total ranges, respectively.

Maggioni et al. 2016). In Taiwan, rainfall amount and intensity are typically subjective to orographic enhancement. In our assessment of satellite precipitation products, heavier rains have more often been identified at relatively higher elevations, which thereby shows a consistent positive correlation. Regarding the likewise increasing correlation of WRF precipitation with elevation, we believe the better resolved topography based on our high-resolution configuration (i.e., 2 km in D3) is one of the main reasons for WRF to simulate orographically enhanced rainfall well (Fang et al. 2011). The WSM5 scheme used in this study could be another factor in improving WRF simulation at high-elevation regions, owing to the better handled ice-phase precipitation.

b. Gridded analysis

Ensured by the point analyses, TCCIP is then referenced as the benchmark for gridded analysis of different precipitation

products. The first step was to assess the performance of monthly precipitation estimation in Fig. 6. Figure 6a shows that rainfall is heavier between March and September as the impact of spring rain, plum rain, and typhoon rain weighs in. Figure 6b details the bias of each precipitation product against TCCIP, and it is noticeable that WRF produces evident overestimation while IMERG_E, IMERG_F, and GSMaP present a slight negative bias. Figure 6c shows that WRF possesses higher errors and variations (i.e., MADs) than its satellite counterparts, especially in summer; however, it improves considerably in winter. Such contrasting performance is also reflected in the correlation coefficient (Fig. 6d). Compared to WRF, the satellite products demonstrate a higher correlation from April to September (warm season) and a lower correlation from November to February (cold season).

The relative performance in the warm/cold season could be explained with the capability of WRF as well as satellite

FIG. 7. Average correlation coefficient between TCCIP and each precipitation product at different temperature intervals; in each of these intervals more than 1000 samples are ensured for calculation.

algorithms. WRF, as many other mainstream NWP models, can handle synoptic systems well (Sperber and Palmer 1996). These are exactly the systems that the majority of winter rainfall in Taiwan is associated with. In addition, the selection of an improved microphysics scheme [e.g., WSM5 over Ferrier, as suggested by Chien et al. (2005)], which can help unravel the correct characteristics of clouds, certainly plays an important role in resolving more accurate precipitation over cold seasons. Another case study for Taiwan conducted by Tao et al. (2011) also indicated that WRF is more sensitive to different microphysics than other model physics (e.g., PBL). In contrast, cold seasons are the most challenging period for satellite algorithms to resolve more accurate precipitation. Frontal systems and orographic lifting of moist air prevailing in wintertime cannot be detected by the IR algorithms (Ebert et al. 2007). While the IMERG or GSMaP products include microwave data, they could still be hindered by complexities in the retrieval due to cloud microphysics (Gottschalk et al. 2005). Our results thus reveal that the complementary behavior of satellite and model precipitation products (i.e., satellite precipitation outperforms model precipitation in the warm season and vice versa in the cold season) is not only valid in the central United States (Zhang et al. 2018) but also in Taiwan.

To further explore the relationship between temperature and the performance of precipitation products, we calculate the average correlation coefficients of each precipitation product at different temperature intervals, graphed in Fig. 7. To ensure that results are paradigmatic, we include only datasets with more than a thousand entries for that specific temperature interval. From the figure, we can see that when temperature falls below 9°C, WRF has a slight advantage over all the satellite precipitation products. When temperature rises above 20°C, the opposite is true, showcasing the relative performance. As temperature continues to rise, all products exhibit declines in correlation coefficients. When temperature is between 9° and 20°C, there is no significant distinction among

the precipitation products. To clarify what the samples are below 9°C, we divide the data into two types of subsets: 1) <1300 m versus >1300 m and 2) summer (JJA) versus winter (DJF). According to these subsets, the sample size in each temperature interval is calculated (not shown), and we can tell that winter samples account for approximately 50% of the total samples in each temperature interval from 3° to 9°C. The low temperature can thus occur in either cold seasons (at any elevations) or at high-elevation regions (in any seasons). In other words, the slight advantage of WRF over satellite precipitation in the low temperature range is generally valid regardless of season or location.

Compared to station data, gridded data can provide comprehensive spatial information of rainfall, and the accuracy of such information can be critical in the moments of extreme precipitation. Past studies (e.g., He et al. 2017) have also pointed out that the performance of satellite precipitation products might vary under different rainfall intensities. With this in mind, we select five “heavy rainfall events” from 2015 to 2017 with daily accumulated rainfall greater than 80 mm (as defined by Taiwan’s Central Weather Bureau) and plot the spatial distribution of each precipitation product in Fig. 8. The first event is Typhoon Haitang in 2017, which is a mild event making landfall in southwestern Taiwan as a category 3 storm (Table 1 and Fig. 2). The rainfall pattern of Typhoon Haitang follows its landfall track to produce significant rainfall amount over southwestern Taiwan, where IMERG_E and IMERG_F actually perform decently in capturing the rainfall pattern but underestimate rainfall amount (in line with Huang et al. 2021a). On the other hand, GSMaP shows clear overestimation, and WRF simulates the main rainfall area with a northward displacement bias. The second event is a plum rain event in 2015, and its main rainfall area is also over southwestern Taiwan. IMERG_E and IMERG_F also underestimate rainfall amount over heavy rainfall areas, with the latter showing milder underestimation. The performance of GSMaP or WRF mimics that in the first event. The third event is another plum

FIG. 8. Spatial distribution of five extreme precipitation events for each precipitation product with TCCIP daily accumulated rainfall greater than 80 mm from 2015 to 2017.

rain event in 2017. According to TCCIP, the main rainfall areas cover the northern, central, and southern mountain regions. In terms of spatial distribution and magnitude, GSMaP and WRF perform better than IMERG. The fourth and fifth events are Typhoons Nepartak and Soudelor in 2016 and 2015, respectively. In contrast to Typhoon Haitang, these two are stronger typhoons (both category 5, see Table 1) when making landfall in southeastern and eastern Taiwan. Because of the slight discrepancy in landfall locations, the associated rainfall patterns of the two typhoons are drastically different. Typhoon Nepartak induces a main rainfall area over eastern Taiwan, while Typhoon Soudelor that penetrates the Central Mountain Range produces a meridional dipole-like rainfall pattern. The peculiar pattern of Typhoon Soudelor subject to complex terrain is very challenging for the satellite precipitation algorithms and WRF to resolve the correct rainfall amount. As a result, among all the events, the largest biases for IMERG_E, GSMaP, and WRF and the largest MAD for GSMaP and WRF can be observed (Table 3). Similar to the previous three events, GSMaP and IMERG respectively show mostly overestimation and underestimation. Last, in the third and fourth events in particular, IMERG_F, compared to IMERG_E, shows no expected improvement in performance. The ineffective climatological adjustment could be attributed to the poor spatial coverage of GPCC stations over Taiwan, as indicated earlier.

Based on the performance in the above five extreme events (Table 3), this study concludes that both satellite and model precipitation products are capable of identifying the main rainfall areas but may varyingly overestimate or underestimate rainfall. Our finding echoes Prakash et al. (2016) and Zhang et al. (2019), who respectively analyzed heavy rainfall events in India and China's Guangdong region and concluded that satellite precipitation products are more effective in estimating spatial distribution than absolute rainfall amount. From Table 3 we can also find that GSMaP has high bias and MAD for the typhoon events, but dominates with the least bias for 2 June 2017 plum rain event. However, if looking into the spatial distribution of this event, we can attribute the least bias of GSMaP to an offset of strong overestimation in northern Taiwan against underestimation in central and southern Taiwan. In general, IMERG that tends to have lower MADs is still better than GSMaP, in consistent with the previous studies (e.g., Huang et al. 2020). In this study, the general effectiveness of WRF in predicting the five events, as demonstrated by the leading performance metrics (e.g., highest CC in the second, fourth, and fifth events). Our finding suggests that WRF is able to produce the reasonable rainfall patterns for more intense, synoptic-scale events (e.g., plum rain fronts or tropical cyclones) even during the warm season, yet a further tune-up of the model can be performed to reduce simulation biases. As for whether WRF will remain as effective in real-time weather prediction, that is another issue that requires further examination.

To verify whether the findings in this study are subject to the selected years, we additionally produce Figs. S1–S3 in the online supplemental material, which are the same as Figs. 3, 4, and 6 but using the newer and available data in 2018/19. Between these two sets of figures, we have noticed quantitative differences in the rainfall amounts that have led to some

TABLE 3. Performance metrics for five selected heavy-rainfall events. Best-performing metric values for each event are bold.

Product	IMERG_E	IMERG_F	GSMaP	WRF
Event 1	Typhoon Haitang, 31 Jul 2017			
CC	0.85	0.85	0.80	0.72
Bias	−2.35	3.95	308.30	−1.13
MAD	34.80	35.13	308.30	46.11
Event 2	24 May 2015 (plum rain)			
CC	0.68	0.68	0.71	0.79
Bias	−37.55	−14.25	49.61	33.05
MAD	44.77	35.32	79.90	41.93
Event 3	2 Jun 2017 (plum rain)			
CC	0.57	0.52	0.49	0.54
Bias	−19.89	−29.14	0.52	−18.53
MAD	49.81	50.21	58.43	63.07
Event 4	Typhoon Nepartak, 8 Jul 2016			
CC	0.57	0.53	0.56	0.77
Bias	−14.06	−49.59	251.06	45.15
MAD	82.58	87.15	256.60	82.38
Event 5	Typhoon Soudelor, 8 Aug 2015			
CC	0.50	0.55	0.57	0.70
Bias	−40.14	−38.18	319.11	90.87
MAD	76.95	77.57	325.06	119.78

changes in the performance of each precipitation product. For instance, in Fig. S3 and Fig. 6, the average rainfall amount in August of 2018/19 is shown considerably higher than its counterpart in August 2015–17. A clear increase has been found in the respective bias (more negative) and MAD for all the precipitation products in 2018/19. Nonetheless, the relative performance, regional patterns, and seasonal patterns are generally similar between these two temporal periods.

c. Hydrological analysis

This section further shows hydrological analysis by applying each precipitation product via rainfall–runoff modeling and then evaluating the flow simulation. First, each precipitation product is used as forcing data in the calibrated HEC-HMS model over the Feitsui and Bajhang watersheds for monthly average flow calculation (Fig. 9). Regarding the Feitsui watershed, with Gauge as the benchmark, although the precipitation products present some difference in the flow simulation, they can well reflect seasonal variations. As expected, the TCCIP-based flow is the closest to Gauge while the WRF- and GSMaP-based flows are overestimated, especially during the warm season. Regarding IMERG, the simulated flows are underestimated, except the flow from IMERG_E in June. In addition, there is no clear difference between the results of Early Run and Final Run. Regarding the Bajhang watershed, all products are capable of reflecting seasonal flow variations, but GSMaP leads to considerable overestimation in the warm season. Meanwhile, compared to the overestimated flows in the Feitsui watershed, WRF yields mild underestimates of flows in the Bajhang watershed. Last, even though IMERG-based flows are slightly underestimated, they showed overall the most accurate seasonal variations.

FIG. 9. Monthly average flow of the (a) Feitsui Reservoir and (b) Bajhang River watersheds simulated by each precipitation product.

To further examine the detailed correspondence between heavy rainfall and high-flow events, we select several high-flow events (>500 cm) to perform daily hydrological analysis. In regard to the Feitsui watershed (Fig. 10), the high-flow events in August and September 2015, caused by Typhoons Soudelor and Dujuan, are first inspected. Figure 10a illustrates how TCCIP and Gauge generally coincide in the simulation results for this period, IMERG products result in underestimated flows, GSMaP generates obvious overestimates, and WRF's overall performance was superior to the satellite precipitation products. In the case of Typhoon Megi in September 2016 (Fig. 10b), IMERG products are able to generate flow simulations with a close agreement with Gauge, GSMaP are significantly overestimated, while WRF shows slight overestimation yet marked performance in capturing the flow peaks. As for the third selected event, induced by Typhoon Khanun's peripheral circulation and northeast monsoon in October 2017 (Fig. 10c), IMERG results in the occurrence of peak flow lagged a day behind,

GSMaP causes overestimation, and WRF, despite slight overestimation, performs reasonably especially for the time to the peak flow.

For the high-flow events of the Bajhang watershed, during the first event in September 2016 (Typhoon Megi from 26 September to 1 October; Fig. 11a), GSMaP is the only product showing simulated flows closer to Gauge, unlike IMERG and WRF both showing notable underestimates. However, in several preceding events, GSMaP still causes some overestimation. When examining several high-flow events caused by plum rain and Typhoon Haitang from early June to August 2017 (Fig. 11b), GSMaP returns to clear overestimation. IMERG, by contrast, shows simulated flows comparatively close to Gauge but at the same time lagged in predicting the time to the peak flows of the plum rain event. Last, WRF produces underestimated flows for the plum rain event but the relatively accurate peak with slight overestimation for the event of Typhoon Haitang.

To better illustrate our findings based on the continuous rainfall-runoff simulation, we calculate the performance

Feitsui Reservoir (High Flow)

FIG. 10. Selected high-flow hydrographs simulated by each precipitation product for the Feitsui Reservoir watershed.

metrics indicated in [section 3c](#) for the entire simulations for both the Feitsui and Bajhang watersheds, and then show the corresponding results in [Table 4](#). Further, we rank different precipitation products based on these metrics to produce the spider charts in [Fig. 12](#). It is clear that for both watersheds, TCCIP as the benchmark produces the best simulation (ranking first in all the metrics), while GSMaP shows the poorest performance in most of the metrics, indicating its poor usefulness for hydrological applications regardless of region. In the Feitsui watershed (northern Taiwan), TCCIP is followed by IMERG_F in all the metrics, suggesting the marked usefulness of satellite precipitation for hydrological modeling. The performance of IMERG_E is inferior to IMERG_F, and its low-flow simulation shows the lowest accuracy (reflected through NSE10) among all the precipitation products. Rainfall-runoff simulation based on WRF in the Feitsui watershed shows barely acceptable accuracy, ranking third in the NSE90 and NSE10 (both positive); the worst performance in VE is carried over from the overestimation of the monthly average flow ([Fig. 9](#)). In the Bajhang watershed (southern Taiwan), IMERG_E performs slightly better than IMERG_F in terms of NSE and VE, but its low-flow simulation is still the least accurate. In fact, IMERG_E, IMERG_F, and WRF bear similar performance in rainfall-runoff simulation in

the Bajhang watershed, suggesting quasi-equal usefulness for hydrological applications.

5. Conclusions and recommendations

Satellite and model precipitation products can compensate at specific research areas where data are inaccessible due to factors such as topographic complexity. This study analyzed two main satellite precipitation products from the GPM Mission, namely, IMERG and GSMaP, as well as the numerical prediction model WRF in terms of the performance of precipitation estimation over Taiwan. To verify the usefulness for hydrological applications, this study further applied the HEC-HMS to conduct flow simulation assessment. Our comprehensive examination reaches the following conclusions:

- 1) The satellite precipitation products performed better than WRF in the warm season and vice versa in the cold season, a phenomenon most prominent in the northern region of Taiwan. As for southern Taiwan, the satellite products were overall better performing than the model precipitation products. This finding suggests that the performance of satellite precipitation is not only influenced

Bajhang River (High Flow)

FIG. 11. As in Fig. 10, but for the Bajhang River watershed.

by seasonal variations but also by regional precipitation patterns.

- 2) Upon examining the correlation between elevation and the performance of precipitation estimation, we found that, at the 30–170-m elevation interval, the satellite precipitation products showed a pattern of deteriorating performance, while at 400–2200 m, all the precipitation products improved in performance as elevation increases.
- 3) Upon examining the correlation between temperature and the performance of precipitation estimation, we found that, when temperature dropped below 9°C (rose above 20°C), the WRF estimates were overall superior (inferior) to the satellite precipitation products.
- 4) In reviewing the extreme precipitation events, we revealed that all products were capable of accurately reflecting the spatial distribution of precipitation. However, IMERG products tended to underestimate rainfall, and the climatological adjustment to IMERG_F seemed ineffective. GSMaP tended to overestimate rainfall, and WRF was capable of correctly simulating both rainfall amount and distribution.
- 5) Seasonal and regional differences in precipitation estimates were carried over into flow simulation. In both the Feitsui Reservoir and Bajhang River watersheds, IMERG achieved overall the most accurate simulation in terms of monthly average flow. GSMaP, by contrast, often overestimated in the warm season.
- 6) Assessment of the high-flow events suggested that IMERG produced reasonable flow estimates but tended to underestimate or delay the peak flows. Meanwhile, GSMaP tended to overestimate high flows and showed poor usefulness for hydrological applications.

According to our analysis and results, we tailor the following recommendations for future endeavors. Since IMERG_E, IMERG_F, and WRF fared equally well in precipitation estimation in Taiwan under certain circumstances, it would be possible to adopt IMERG_E or WRF for near-real-time precipitation applications, such as issuing flood warnings. Further, in Taiwan's rainy season, heavy rainfall events often occur with a short duration, so a similar assessment at hourly or finer time scale should be conducted. As more and more data become available, another topic worth further investigation is the assessment of the influence of climate oscillations (e.g., El Niño or La Niña) on satellite/model precipitation estimates. Last but not least, this study has unequivocally pointed out

TABLE 4. Performance metrics of the HMS simulation using different precipitation products for the Feitsui and Bajhang watersheds.

Product	NSE	VE	NSE90	NSES	NSE10
Feitsui Watershed					
TCCIP	0.90	0.22	0.83	0.80	0.86
IMERG_E	0.30	0.59	0.06	0.17	−0.79
IMERG_F	0.50	0.48	0.21	0.22	0.35
GSMaP	−1.29	0.75	−3.45	−16.11	−0.06
WRF	0.26	0.77	0.18	−0.02	0.25
Bajhang Watershed					
TCCIP	0.83	0.34	0.75	0.84	0.77
IMERG_E	0.60	0.52	0.39	0.48	0.68
IMERG_F	0.51	0.60	0.43	0.45	0.72
GSMaP	−4.60	1.07	−9.54	−174.58	0.72
WRF	0.41	0.60	0.45	0.19	0.74

FIG. 12. Rank of different precipitation products based on five performance metrics derived from the HMS simulation for the Feitsui and Bajhang watersheds.

satellite precipitation products show room for improvement, especially for extreme precipitation estimates. Thus, the rationality of applying products such as IMERG to extreme precipitation-related studies will rely on the development of algorithms for restructuring gridded precipitation data (e.g., Chen et al. 2017).

Acknowledgments. This study was funded by Taiwan's Ministry of Science and Technology (MOST) under Grants MOST 108-2621-M-005-008-MY3 and MOST 109-2221-E-005-001-MY3.

Data availability statement. Data used in this study are available from the corresponding author (cjchen@nchu.edu.tw) upon reasonable request.

REFERENCES

- Aonashi, K., and Coauthors, 2009: GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation. *J. Meteor. Soc. Japan*, **87A**, 119–136, <https://doi.org/10.2151/jmsj.87A.119>.
- Behrangi, A., B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian, 2011: Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.*, **397**, 225–237, <https://doi.org/10.1016/j.jhydrol.2010.11.043>.
- , K. Andreadis, J. B. Fisher, F. J. Turk, S. Granger, T. Painter, and N. Das, 2014: Satellite-based precipitation estimation and its application for streamflow prediction over mountainous western US basins. *J. Appl. Meteor. Climatol.*, **53**, 2823–2842, <https://doi.org/10.1175/JAMC-D-14-0056.1>.
- Bharti, V., C. Singh, J. Ettema, and T. A. R. Turkington, 2016: Spatiotemporal characteristics of extreme rainfall events over the Northwest Himalaya using satellite data. *Int. J. Climatol.*, **36**, 3949–3962, <https://doi.org/10.1002/joc.4605>.
- Chen, C.-J., S. U. S. Senarath, I. M. Dima-West, and M. P. Marcella, 2017: Evaluation and restructuring of gridded precipitation data over the Greater Mekong Subregion. *Int. J. Climatol.*, **37**, 180–196, <https://doi.org/10.1002/joc.4696>.
- , C. C. Chen, M. H. Lo, J. Y. Juang, and C. M. Chang, 2020: Central Taiwan's hydroclimate in response to land use/cover change. *Environ. Res. Lett.*, **15**, 034015, <https://doi.org/10.1088/1748-9326/ab68aa>.
- Chen, C.-S., and Y.-L. Chen, 2003: The rainfall characteristics of Taiwan. *Mon. Wea. Rev.*, **131**, 1323–1341, [https://doi.org/10.1175/1520-0493\(2003\)131<1323:TRCOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<1323:TRCOT>2.0.CO;2).
- Chien, F., B. Jou, P. Lin, and J. Hong, 2005: A real-time MM5/WRF forecasting system in Taiwan. *5th WRF/14th MM5 Users' Workshop*, Boulder, CO, NCAR, P.5, <https://www2.mmm.ucar.edu/mm5/workshop/ws04/PosterSession/Chien.Fang-Ching.pdf>.
- Chou, M. D., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp., <https://ntrs.nasa.gov/api/citations/19950009331/downloads/19950009331.pdf>.
- Darbandsari, P., and P. Coulibaly, 2020: Inter-comparison of lumped hydrological models in data-scarce watersheds using different precipitation forcing data sets: Case study of Northern Ontario, Canada. *J. Hydrol. Reg. Stud.*, **31**, 100730, <https://doi.org/10.1016/j.ejrh.2020.100730>.
- Ebert, E., J. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**, 47–64, <https://doi.org/10.1175/BAMS-88-1-47>.
- El Kenawy, A., J. Lopez-Moreno, M. McCabe, and S. Vicente-Serrano, 2015: Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Global Planet. Change*, **133**, 188–200, <https://doi.org/10.1016/j.gloplacha.2015.08.013>.
- Fang, X., Y. Kuo, and A. Wang, 2011: The impacts of Taiwan topography on the predictability of Typhoon Morakot's record-breaking rainfall: A high-resolution ensemble simulation. *Wea. Forecasting*, **26**, 613–633, <https://doi.org/10.1175/WAF-D-10-05020.1>.
- Gottschalk, J., J. Meng, M. Rodell, and P. Houser, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. *J. Hydrometeorol.*, **6**, 573–598, <https://doi.org/10.1175/JHM437.1>.
- He, Z., L. Yang, F. Tian, G. Ni, A. Hou, and H. Lu, 2017: Intercomparison of rainfall estimates from TRMM and GPM multisatellite products over the upper Mekong River basin. *J. Hydrometeorol.*, **18**, 413–430, <https://doi.org/10.1175/JHM-D-16-0198.1>.

- Henny, L., C. Thorncroft, H. Hsu, and L. Bosart, 2021: Extreme rainfall in Taiwan: Seasonal statistics and trends. *J. Climate*, **34**, 4711–4731, <https://doi.org/10.1175/JCLI-D-20-0999.1>.
- Hirpa, F., M. Gebremichael, and T. Hopson, 2010: Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteor. Climatol.*, **49**, 1044–1051, <https://doi.org/10.1175/2009JAMC2298.1>.
- Hong, J., C. Fong, L. Hsiao, Y. Yu, and C. Tzeng, 2015: Ensemble typhoon quantitative precipitation forecasts model in Taiwan. *Wea. Forecasting*, **30**, 217–237, <https://doi.org/10.1175/WAF-D-14-00037.1>.
- Hong, S. Y., J. Dudhia, and S. H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120, [https://doi.org/10.1175/1520-0493\(2004\)132<0103:ARATIM>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2).
- , Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Hsiao, L., and Coauthors, 2013: Ensemble forecasting of typhoon rainfall and floods over a mountainous watershed in Taiwan. *J. Hydrol.*, **506**, 55–68, <https://doi.org/10.1016/j.jhydrol.2013.08.046>.
- Huang, W.-R., Y.-H. Chang, and P.-Y. Liu, 2018: Assessment of IMERG precipitation over Taiwan at multiple timescales. *Atmos. Res.*, **214**, 239–249, <https://doi.org/10.1016/j.atmosres.2018.08.004>.
- , P.-Y. Liu, Y.-H. Chang, and C.-Y. Liu, 2020: Evaluation and application of satellite precipitation products in studying the summer precipitation variations over Taiwan. *Remote Sens.*, **12**, 347, <https://doi.org/10.3390/rs12030347>.
- , —, —, and C.-A. Lee, 2021a: Evaluation of IMERG Level-3 products in depicting the July to October rainfall over Taiwan: Typhoon versus non-typhoon. *Remote Sens.*, **13**, 622, <https://doi.org/10.3390/rs13040622>.
- , —, J. Hsu, X. Li, and L. Deng, 2021b: Assessment of near-real-time satellite precipitation products from GSMaP in monitoring rainfall variations over Taiwan. *Remote Sens.*, **13**, 202, <https://doi.org/10.3390/rs13020202>.
- Huffman, G. J., D. T. Bolvin, and E. J. Nelkin, 2015: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code 612 Tech. Doc., 48 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf.
- , and Coauthors, 2019: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 6, 34 pp., https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf.
- Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, [https://doi.org/10.1175/1520-0493\(1994\)122<0927:TSMECM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2).
- Janowiak, J. E., R. J. Joyce, and Y. Yarosh, 2001: A real-time global half-hourly pixel-resolution infrared dataset and its applications. *Bull. Amer. Meteor. Soc.*, **82**, 205–218, [https://doi.org/10.1175/1520-0477\(2001\)082<0205:ARTGHH>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<0205:ARTGHH>2.3.CO;2).
- Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.*, **187**, 95–105, <https://doi.org/10.1016/j.atmosres.2016.12.007>.
- Kubota, T., and Coauthors, 2007: Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE. Trans. Geosci. Remote Sens.*, **45**, 2259–2275, <https://doi.org/10.1109/TGRS.2007.895337>.
- , and Coauthors, 2020: Global Satellite Mapping of Precipitation (GSMaP) products in the GPM era. *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 67, Springer, 355–373, https://doi.org/10.1007/978-3-030-24568-9_20.
- Li, C., J. Berner, J. Hong, C. Fong, and Y. Kuo, 2020: The Taiwan WRF Ensemble Prediction System: Scientific description, model-error representation and performance results. *Asia-Pac. J. Atmos. Sci.*, **56**, 1–15, <https://doi.org/10.1007/s13143-019-00127-8>.
- Li, M., W. Tien, and C. Tung, 2009: Assessing the impact of climate change on the land hydrology in Taiwan. *Paddy Water Environ.*, **7**, 283–292, <https://doi.org/10.1007/s10333-009-0175-9>.
- Liu, C., P. Aryastana, G. Liu, and W. Huang, 2020: Assessment of satellite precipitation product estimates over Bali Island. *Atmos. Res.*, **244**, 105032, <https://doi.org/10.1016/j.atmosres.2020.105032>.
- Liveh, B., P. J. Restrepo, and D. P. Lettenmaier, 2011: Development of a unified land model for prediction of surface hydrology and land-atmosphere interactions. *J. Hydrometeor.*, **12**, 1299–1320, <https://doi.org/10.1175/2011JHM1361.1>.
- Maggioni, V., P. Meyers, and M. Robinson, 2016: A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *J. Hydrometeor.*, **17**, 1101–1117, <https://doi.org/10.1175/JHM-D-15-0190.1>.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, <https://doi.org/10.1029/97JD00237>.
- Nash, J., and J. Sutcliffe, 1970: River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.*, **10**, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- O, S., and P. E. Kirstetter, 2018: Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Quart. J. Roy. Meteor. Soc.*, **144** (Suppl. 1), 270–281, <https://doi.org/10.1002/qj.3218>.
- , U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, and W. A. Petersen, 2017: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.*, **21**, 6559–6572, <https://doi.org/10.5194/hess-21-6559-2017>.
- Prakash, S., A. K. Mitra, D. S. Pai, and A. AghaKouchak, 2016: From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water. Resour.*, **88**, 1–7, <https://doi.org/10.1016/j.advwatres.2015.11.008>.
- , —, A. AghaKouchak, Z. Liu, H. Norouzi, and D. S. Pai, 2018: A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.*, **556**, 865–876, <https://doi.org/10.1016/j.jhydrol.2016.01.029>.
- Qin, Y., Z. Chen, Y. Shen, S. Zhang, and R. Shi, 2014: Evaluation of satellite rainfall estimates over the Chinese mainland. *Remote Sens.*, **6**, 11 649–11 672, <https://doi.org/10.3390/rs6111649>.
- Razavi, T., and P. Coulibaly, 2017: An evaluation of regionalization and watershed classification schemes for continuous daily streamflow prediction in ungauged watersheds. *Can. Water Resour. J.*, **42**, 2–20, <https://doi.org/10.1080/07011784.2016.1184590>.
- Shah, H. L., and V. Mishra, 2016: Uncertainty and bias in satellite-based precipitation estimates over Indian subcontinental basins: Implications for real-time streamflow simulation and flood prediction. *J. Hydrometeor.*, **17**, 615–636, <https://doi.org/10.1175/JHM-D-15-0115.1>.

- Shige, S., and Coauthors, 2009: The GSMaP precipitation retrieval algorithm for microwave sounders—Part I: Over-ocean algorithm. *IEEE. Trans. Geosci. Remote Sens.*, **47**, 3084–3097, <https://doi.org/10.1109/TGRS.2009.2019954>.
- Shih, D., C. Chen, M. Li, C. Jang, C. Chang, and Y. Liao, 2019: Statistical and numerical assessments of groundwater resource subject to excessive pumping: Case study in Southwest Taiwan. *Water*, **11**, 360, <https://doi.org/10.3390/w11020360>.
- Sibson, R., 1981: *A Brief Description of Natural Neighbour Interpolation*. John Wiley and Sons, 21–36.
- Simolo, C., M. Brunetti, M. Maugeri, and T. Nanni, 2010: Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. *Int. J. Climatol.*, **30**, 1564–1576, <https://doi.org/10.1002/joc.1992>.
- Smith, M. B., D.-J. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong, 2004: The distributed model intercomparison project (DMIP): Motivation and experiment design. *J. Hydrol.*, **298**, 4–26, <https://doi.org/10.1016/j.jhydrol.2004.03.040>.
- Smith, T., R. Reynolds, R. Livezey, and D. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, **9**, 1403–1420, [https://doi.org/10.1175/1520-0442\(1996\)009<1403:ROHSST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1403:ROHSST>2.0.CO;2).
- Sperber, K., and T. Palmer, 1996: Interannual tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project. *J. Climate*, **9**, 2727–2750, [https://doi.org/10.1175/1520-0442\(1996\)009<2727:ITRIVIG>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2727:ITRIVIG>2.0.CO;2).
- Tan, J., W. Peterson, G. Kirchengast, D. Goodrich, and D. Wolff, 2018: Evaluation of Global Precipitation Measurement rainfall estimates against three dense gauge networks. *J. Hydrometeorol.*, **19**, 517–532, <https://doi.org/10.1175/JHM-D-17-0174.1>.
- Tao, W., and Coauthors, 2011: High-resolution numerical simulation of the extreme rainfall associated with Typhoon Morakot. Part I: Comparing the impact of microphysics and PBL parameterizations with observations. *Terr. Atmos. Ocean. Sci.*, **22**, 673–696, [https://doi.org/10.3319/TAO.2011.08.26.01\(TM\)](https://doi.org/10.3319/TAO.2011.08.26.01(TM)).
- Tian, F., S. Hou, L. Yang, H. Hu, and A. Hou, 2018: How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? *J. Hydrometeorol.*, **19**, 339–349, <https://doi.org/10.1175/JHM-D-17-0161.1>.
- Tseng, S. C., C. J. Chen, and S. U. S. Senarath, 2020: Evaluation of multi-site precipitation generators across scales. *Int. J. Climatol.*, **40**, 4622–4637, <https://doi.org/10.1002/joc.6480>.
- USACE, 2000: Hydrologic modeling system HEC-HMS: Technical reference manual. Hydrologic Engineering Center, U.S. Army Corps of Engineering, 138 pp., [https://www.hec.usace.army.mil/software/hech-hms/documentation/HEC-HMS_Technical%20Reference%20Manual_\(CPD-74B\).pdf](https://www.hec.usace.army.mil/software/hech-hms/documentation/HEC-HMS_Technical%20Reference%20Manual_(CPD-74B).pdf).
- Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Japan*, **87A**, 137–151, <https://doi.org/10.2151/jmsj.87A.137>.
- Wang, H., and B. Yong, 2020: Quasi-global evaluation of IMERG and GSMaP precipitation products over land using gauge observations. *Water*, **12**, 243–258, <https://doi.org/10.3390/w12010243>.
- Watson, D., 1994: *Nngidr: An Implementation of Natural Neighbor Interpolation*. Watson, 170 pp.
- Weng, S. P., and C. D. Yang, 2018: The construction and verification of daily gridded rainfall dataset (1960–2015) (in Chinese with English abstract). *Taiwan Water Conserv.*, **66** (4), 33–52.
- Wu, M., and G. Lin, 2017: The very short-term rainfall forecasting for a mountainous watershed by means of an ensemble numerical weather prediction system in Taiwan. *J. Hydrol.*, **546**, 60–70, <https://doi.org/10.1016/j.jhydrol.2017.01.012>.
- Yang, S., and T. Yang, 2014: Uncertainty assessment: Reservoir inflow forecasting with ensemble precipitation forecasts and HEC-HMS. *Adv. Meteor.*, **2014**, 581756, <https://doi.org/10.1155/2014/581756>.
- Yong, B., J. Wang, L. Ren, Y. You, P. Xie, and Y. Hong, 2016: Evaluating four multisatellite precipitation estimates over the Diaoyu Islands during Typhoon seasons. *J. Hydrometeorol.*, **17**, 1623–1641, <https://doi.org/10.1175/JHM-D-15-0165.1>.
- Yu, P., and Y. Wang, 2009: Impact of climate change on hydrological processes over a basin scale in northern Taiwan. *Hydrol. Processes*, **23**, 3556–3568, <https://doi.org/10.1002/hyp.7456>.
- , T. Yang, and C. Wu, 2002: Impact of climate change on water resources in southern Taiwan. *J. Hydrol.*, **260**, 161–175, [https://doi.org/10.1016/S0022-1694\(01\)00614-X](https://doi.org/10.1016/S0022-1694(01)00614-X).
- Zhang, A., L. Xiao, C. Min, S. Chen, M. Kulie, C. Huang, and Z. Liang, 2019: Evaluation of latest GPM-era high-resolution satellite precipitation products during the May 2017 Guangdong extreme rainfall event. *Atmos. Res.*, **216**, 76–85, <https://doi.org/10.1016/j.atmosres.2018.09.018>.
- Zhang, J., L. F. Lin, and R. L. Bras, 2018: Evaluation of the quality of precipitation products: A case study using WRF and IMERG data over the central United States. *J. Hydrometeorol.*, **19**, 2007–2020, <https://doi.org/10.1175/JHM-D-18-0153.1>.
- Zhang, X., E. N. Anagnostou, M. Frediani, S. Solomos, and G. Kallos, 2013: Using NWP simulations in satellite rainfall estimation of heavy precipitation events over mountainous areas. *J. Hydrometeorol.*, **14**, 1844–1858, <https://doi.org/10.1175/JHM-D-12-0174.1>.