Check for updates

*Hydrological Processes*

WILEY

RESEARCH ARTICLE OPEN ACCESS

# Comparative Evaluation of Gridded Precipitation Datasets in Capturing Hydrological Extremes in a Mesoscale Heterogeneous Catchment in Austria

Zryab Babker[1] | Mohammed Basheer[2] | Tim G. Reichenau[1] | Jürgen Komma[3] | Oscar M. Baez-Villanueva[4] | Morteza Zargar[1] | Karl Schneider[1]

[1]Institute of Geography, Faculty of Mathematics and Natural Sciences, University of Cologne, Cologne, Germany | [2]Department of Civil and Mineral Engineering, University of Toronto, Toronto, Ontario, Canada | [3]Institute of Hydraulic Engineering and Water Resources Management, Faculty of Civil Engineering, Vienna University of Technology, Vienna, Austria | [4]Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium

**Correspondence:** Zryab Babker (zryab.babker@uni-koeln.de; zryabbabker@gmail.com)

**Received:** 22 July 2025 | **Revised:** 27 October 2025 | **Accepted:** 29 November 2025

**Keywords:** Austria | elevation zones | extreme precipitation | gridded precipitation products | Kamp catchment | point-to-pixel evaluation | streamflow simulation | SWAT+ model

## ABSTRACT

Accurate representation of extreme precipitation is crucial for assessing flood hazards and developing risk mitigation strategies. For such applications, gridded Precipitation Products (PPs) can be a promising alternative to traditional point measurements, especially in regions where such measurements are sparse or non-existent. However, the accuracy of PPs in representing extreme precipitation should be evaluated before use. In this study, we evaluate the performance of four PPs (SPARTACUS v2.1, IMERG-F v07, CHIRPS v2.0, and ERA5-Land) against 33 precipitation gauges at a daily time scale over the Kamp catchment in Austria for the period 1998–2020. The hydrological response in the catchment is influenced not only by the intensity of extreme precipitation events but also by antecedent soil moisture and seasonal conditions. Continuous and categorical performance metrics are used to evaluate the performance of the PPs at gauge locations. Additionally, the Soil and Water Assessment Tool Plus (SWAT+) model is used to assess the reliability of PPs when used as forcings for hydrological modelling. The results reveal that while most evaluated products can detect no-rain events, their ability to capture extreme precipitation events varies notably. SPARTACUS v2.1 exhibited the best ability to detect extremes at gauge locations, resulting in streamflow simulation that closely matched the observed data. IMERG-F v07 demonstrated moderate performance in both extreme precipitation detection and corresponding peak flow generation. In contrast, CHIRPS v2.0 and ERA5-Land showed poor performance in representing extreme precipitation, resulting in underestimated high flows and lower reliability in simulating flood-related hydrological processes. These findings highlight the importance of evaluating the ability of PPs in capturing extreme precipitation to ensure reliable simulation of flood peaks and hydrological extremes. We conclude that catchment-specific validation linking precipitation extremes to hydrological responses is essential for selecting appropriate precipitation forcings for hydrological applications.

## 1 | Introduction

Precipitation is a key component of the hydrological cycle, influencing river flows and the occurrence of hydrological extremes such as floods and droughts. The availability of accurate, long-term, and consistent precipitation data at suitable spatio-temporal resolutions is key for water resources management, infrastructure design, and assessment of hydrological extremes,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Hydrological Processes* published by John Wiley & Sons Ltd.

among other applications (Zhang et al. 2017; Ruhí et al. 2018; Yuan et al. 2018; Fang et al. 2019). Although meteorological gauge stations provide direct and reliable data at specific locations, characterizing spatial precipitation patterns based solely on gauges is challenging, especially where gauge networks are sparse or do not cover high elevations (Xue et al. 2013; Baez-Villanueva et al. 2018; Sharifi et al. 2018). Typically, hydrological impact studies examine hydrological processes at scales and locations where point gauge stations may not be representative (Mankin et al. 2025).

Gridded precipitation products (PPs) offer complementary information by providing insights into the spatial and temporal variability of precipitation. Over the past few decades, several ground-based, satellite-based, reanalysis-based, and multisource merged PPs have been developed at high spatial and temporal resolutions. Ground-based PPs (e.g., E-OBS, Haylock et al. 2008; SPARTACUS, Hiebl and Frei 2018; GPCC, Adler et al. 2018) provide gridded precipitation estimates leveraging information from point measurements, but their accuracy is highly dependent on the density and configuration of the gauge network (Hiebl and Frei 2018; Ma et al. 2025). In contrast, satellite-based products (e.g., PERSIANN-CDR-CCS, Hong et al. 2004; CMORPH, Joyce et al. 2004; CHIRP, Funk et al. 2014; IMERG v07 Final Run, Huffman et al. 2024) offer near-global coverage and high temporal resolution, which is valuable in data-scarce regions. However, they may underestimate light precipitation and tend to underperform in snow- and ice-covered regions and in complex terrains (Joyce et al. 2004; Kidd et al. 2012; Huffman et al. 2019; Beck et al. 2019; Tang et al. 2020; Sadeghi et al. 2021). Reanalysis-based products (e.g., ERA5, Hersbach et al. 2020; ERA5-Land, Muñoz-Sabater et al. 2021) combine numerical weather prediction models with assimilated observations, offering temporally consistent, physically coherent, and long-term datasets, but their spatial resolution is often coarse, and local extremes may be smoothed due to model biases (Hersbach et al. 2020; Muñoz-Sabater et al. 2021). Multisource (blended/merged) products (e.g., CHIRPS, Funk et al. 2015; MSWEP, Beck et al. 2019; GMCP, Ma et al. 2025) integrate satellite, reanalysis, and gauge networks, leveraging the strengths of multiple sources. These products can better represent precipitation extremes and local variability. However, their performance depends on the density and quality of the input data (Rauthe et al. 2013; Funk et al. 2015; Cornes et al. 2018; Beck et al. 2019; Ma et al. 2025). For a comprehensive review of the advantages and limitations of precipitation data from these different sources, the reader is referred to Ma et al. (2025).

The performance of PPs has been widely evaluated across different regions worldwide (Beck et al. 2017; Sun et al. 2018; Nguyen et al. 2018; Mankin et al. 2025), and results consistently show that their reliability depends on climate, terrain, and gauge density (Isotta et al. 2015). The inclusion of a high density of station data in PPs typically results in enhanced accuracy and a better representation of extreme events, especially in regions with heterogeneous precipitation distributions. In recent years, several efforts have been made to evaluate the performance of different PPs over Europe (e.g., Zolina et al. 2004; Skok et al. 2016; Kidd et al. 2012; Navarro et al. 2019; Lledó et al. 2024).

Despite the benefits of using PPs for diverse applications, particularly in data-scarce or ungauged regions, they often contain biases that can limit their utility. Therefore, these products require a comprehensive evaluation to assess their performance and determine their suitability before they are used in specific regions. In hydrological applications, the accuracy of precipitation data directly impacts the reliability of hydrological simulations, particularly for extreme events (Skøien and Blöschl 2006; Towner et al. 2019; Nester et al. 2012), which are key in assessing flood risk (Nester et al. 2012). In recent decades, the intensity and frequency of extreme precipitation events and the occurrence probability and magnitude of extreme flood events have increased in many regions, causing socioeconomic losses (IPCC 2012; Shi et al. 2017; Fang et al. 2019; Eingrüber and Korres 2022).

Austria is increasingly exposed to frequent riverine floods and droughts, often resulting in severe socioeconomic losses (Ucal 2017; Leitner et al. 2020). However, a comprehensive analysis of PPs' performance in capturing the intensity, frequency, and spatial distribution of such events, as well as their suitability for hydrological modelling in Austria, is lacking. Only a few studies have focused explicitly on the performance of PPs in Austria (Kann et al. 2015; O et al. 2017; Hiebl and Frei 2018; Sharifi et al. 2018, 2019; Ghaemi et al. 2021; Peßenteiner et al. 2021; Haas et al. 2024; Probst and Mauser 2022), and a limited number of these studies have explicitly evaluated PPs' performance in capturing extreme precipitation events (Sharifi et al. 2019; Haas et al. 2024). Table 1 provides an overview of previous studies on the evaluation of precipitation products in Austria. A systematic and comprehensive evaluation is necessary to robustly select suitable products for specific applications and to understand uncertainties. Such evaluation is particularly necessary in mountainous and remote areas, where precipitation exhibits high spatial and temporal variability. In these regions, flood occurrence is often influenced by antecedent soil moisture and other preconditions, leading to nonlinear hydrological responses in which precipitation-induced runoff may be amplified under saturated conditions or suppressed when soils are dry and highly absorptive (Haslinger et al. 2025).

This study focuses on four PPs (CHIRPS v2.0, IMERG-F v07, ERA5-Land, and SPARTACUS v2.1) applied to the Kamp catchment from 1998 to 2020. We acknowledge that several recently developed global precipitation datasets not considered in this study could provide high spatiotemporal resolution and offer valuable insights when applied in Austria. In this study, however, we deliberately focused on the four selected products for their complementary characteristics. To ensure a balanced comparison, we selected one representative product from each major data source: the satellite-based IMERG v07 Final Run, the reanalysis-based ERA5-Land, the ground-based SPARTACUS v2.1, and the merged CHIRPS v2.0. These products also represent different spatial domains, ranging from global to regional and national scales, and all consistently cover the timeframe of our analysis (1998–2020). Additionally, they offer varying spatial resolutions, enabling us to investigate potential scale effects on their performance. CHIRPS v2.0 and IMERG-F v07 are widely recognized global datasets; however, their performance has not been previously evaluated in Austria. Our study, therefore, provides new insights into their applicability in this region. ERA5-Land is



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 1** | Overview of previous studies on the evaluation of precipitation products in Austria.

| No | Author(s) | Spatial extent | Temporal resolution | Evaluated product name(s) | Study period | Study objective | Major findings |
|---|---|---|---|---|---|---|---|
| 1 | Sharifi et al. (2018) | Northeast Austria (square covering primarily Niederoesterreich and Wien) | 1,3,6,12-hourly and daily | IMERG-Final Run (FR) IMERG-Real-Time (RT) | ~10 Months Mid-March 2015—end of January 2016 | Evaluation of precipitation products against a network of 62 precipitation stations from the Zentralanstalt für Meteorologie und Geodynamik | Satellite Precipitation Estimate (SPE) products detect more heavy and extreme precipitation events but capture fewer precipitation occurrences for light precipitation (excluding IMERG RT) than the station datasets. For the entire spectrum of precipitation rates ($p \geq 0.1$ mm), 1, 3, 6-hourly, IMERG FR did not show a clear improvement of the Bias over IMERG-RT. For 12-hourly and daily precipitation, the bias in IMERG-FR has improved compared to IMERG-RT.<br>IMERG-FR shows an improvement in RMSE compared to IMERG-RT. IMERG-FR systematically underestimates moderate to extreme precipitation and overestimates light precipitation for all time scales against rain gauges.<br>From bias, RMSE, and correlation coefficients, IMERG-FR outperforms IMERG-RT, particularly for 6-hourly, 12-hourly, and daily precipitation. |
| 2 | O et al. (2017) | Southeast Austria (two IMERG $0.1^{o} \times 0.1^{o}$ grid cells) | Half hourly | IMERG Early (E) IMERG Late (L) IMERG Final (F) | ~6 months × 2 April–October for two years | Evaluation of IMERG products against gauge-based gridded rainfall data from the WegenerNet Feldbach region (WEGN) high-density climate station network (40 and 39 stations) | IMERG-F rainfall estimates show the best overall agreement with WEGN data, followed by IMERG-L and IMERG-E, particularly during the hot season. Insufficient Passive Microwave (PMW) sources and errors in motion vectors can result in significant discrepancies. |

(Continues)



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 1** | (Continued)

| No | Author(s) | Spatial extent | Temporal resolution | Evaluated product name(s) | Study period | Study objective | Major findings |
|---|---|---|---|---|---|---|---|
| 3 | Sharifi et al. (2019) | Austria | Daily and monthly | MSWEP-V2.2<br>IMERG-VO5B<br>IMERG-VO6A<br>IMERG-VO5-RT<br>ERA5<br>SM2RAIN-ASCAT | 17 months June 2014–December 2015 | Evaluation of satellite-based precipitation products over mountainous terrain, compared to precipitation data from 782 stations provided by BMNT. | All products are capable of reproducing the spatial precipitation distribution across the country. MSWEP-V2.2, IMERG-V05B, and V06A exhibit the strongest agreement with observations and perform better in terms of spatial patterns and statistical metrics.<br>IMERG and ERA5 products have systematic precipitation overestimation at the monthly timescale. MSWEP performs the best in terms of Probability of Detection (POD) and False Alarm Ratio (FAR). ERA5 and MSWEP exhibit robust average POD and FAR values for light to moderate precipitation. Products often struggle to capture the effects of topography on precipitation. Performance was lower in the peripheries of the study area, where most stations are situated in the mountainous area, and higher in the low-altitude regions. MSWEP-V2.2, followed by IMERG-V06A and V05B, are the most suitable for driving hydro-meteorological, agricultural, and other models over mountainous terrain. |
| 4 | Sharifi et al. (2018) | Northeast Austria | Daily | IMERG | March 15th 2015–January 31st 2016 | Presentation of an error adjustment approach based on statistical differences between satellite precipitation products and in situ observations using additive and multiplicative error models. | IMERG precipitation estimates improved after error adjustment compared to the original product. Copula-based simulations associated with the additive error model performed better than those of the multiplicative error model. |
| 5 | Peßenteiner et al. (2021) | Southeastern Austria (Woelzerbach and Raab catchments in Styria) | Daily, monthly | ÖKS15<br>EURO-CORDEX<br>SPARTACUS | 1961–2005 | Assessment of the use of ÖKS15 for hydrological impact studies. | ÖKS15 hydro-meteorological indicators are comparable to observations and therefore considered appropriate for hydrological impact studies. Some biases remain despite bias correction. Precipitation is overestimated by the Gridded Precipitation Analysis for the Real-time and the Daily scale dataset version 1 (GPARD1) data at the grid point level; however, some catchment areas sometimes experience underestimation. Differences between the GRRD and station data are largest in summer. |

(Continues)



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 1** | (Continued)

| No | Author(s) | Spatial extent | Temporal resolution | Evaluated product name(s) | Study period | Study objective | Major findings |
|---|---|---|---|---|---|---|---|
| 6 | Kann et al. (2015) | Southeastern Austria (Styria) | 5 min<br>30 min | Rapid-INCA<br>WegenerNet | 2011 | Evaluating the skill of very high-resolution and frequently updated precipitation analyses (rapid-INCA) using WegenerNet weather stations. | Rapid-INCA generally underestimates precipitation amounts. The capability of the analysis to assess the convective precipitation distribution depends on the representativeness of the stations under prevailing convective conditions. |
| 7 | Ghaemi et al. (2021) | Southeastern Austria (small study area near Raab) | Annual | INCA<br>WegenerNet | 12 years<br>2007–2018 | Assessment of the Integrated Nowcasting through Comprehensive Analysis (INCA) precipitation against WegenerNet weather stations. | INCA overestimated the annual mean precipitation from 2007–2011, except in 2009, and underestimated precipitation in grid cells farther from the meteorological stations. INCA overestimates the annual mean precipitation even more severely from 2012 to 2014 (35%) – potentially due to errors in newly installed radars. Overestimation remains dominant from 2015 onwards, but becomes less severe (12.5%). INCA performs better in wet seasons. False events happen less frequently, closer to the stations. The differences between INCA and WegenerNet are more noticeable at the extremes. INCA underestimated event peak intensity until mid-2012, then overestimated it. Relative differences in peak intensity in extreme convective short-duration events range from approximately −40% to 40%. |
| 8 | Hiebl and Frei (2018) | Austria | Daily | SPARTACUS<br>WegenerNet<br>GPARD1<br>E-OBs<br>Daymey<br>StartClim<br>INCA | 1961–1990 | Presentation of a grid dataset of daily precipitation for modelling and planning tasks – SPARTACUS. | The dataset is “at least as plausible as previously existing datasets” but presents some differences. There is an overestimation of light and moderate precipitation intensities and an underestimation of intense and heavy precipitation – a bias up to 20%. Random errors in grid point estimates are typically within a factor of 1.5 of the observation, but are larger in data-sparse or high-elevation regions, or in summer. |

(Continues)



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 1** | (Continued)

| No | Author(s) | Spatial extent | Temporal resolution | Evaluated product name(s) | Study period | Study objective | Major findings |
|---|---|---|---|---|---|---|---|
| 9 | Haas et al. (2024) | South-eastern Austria | Hourly, daily | ERA5-Land<br>WegenerNet<br>INCA<br>SPARTACUS | 1961–2022 | Analysis of warm-season short-duration extreme convective precipitation events with INCA and Wegener Net to assess the potential climate-change-induced amplification of sub-daily precipitation extremes.<br>Preparation of the short-duration extreme convective precipitation events classification and threshold exceedance amount metrics. | Results of the different precipitation products are largely comparable. ERA5 Land data show no significant trend in the ratio between short-duration maximum hourly precipitation and the average hourly precipitation of daily sums, whereas WEGN and INCA show a tendency for an increase in the ratio over the last two decades, although this remains without statistical significance.<br>Modelling maximum hourly precipitation with SPARTACUS and ERA5 land data revealed more information about climate amplification. |
| 10 | Probst and Mauser (2022) | The Danube Basin | Hourly, daily, monthly | ERA5<br>WFDE5 | 1980–2006 | Evaluation of meteorological forcing data for hydrological modelling (PROMET).<br>Uncorrected products and products bias corrected with WorldClim2 or PRISM climatology are compared. | ERA5 and WFDE5 were found to be well-suited for hydrological modelling, but bias correction was essential. Bias correction with GLOWA and PRISM performed best in the Alps. The study recommends regional high-resolution precipitation climatology as an alternative to global data sets. |

a widely used reanalysis product in hydrological modeling studies across Europe, offering physically consistent precipitation and land-surface estimates. SPARTACUS v2.1 is a high-density, gauge-based national reference dataset with a fine spatial resolution (1 km).

The objectives of this study are to (1) comprehensively evaluate the ability of four PPs in capturing daily precipitation patterns and precipitation extremes in the Kamp Catchment, (2) evaluate the ability of PPs in capturing the effects of topography on extreme precipitation, and (3) assess the hydrological utility of PPs in simulating streamflow and historical flood events when used as forcing for the Soil and Water Assessment Tool (SWAT+) model. The Kamp catchment provides a suitable case study, as it has experienced severe flooding events in the recent past (Viglione et al. 2010).

This study is done in the context of the EU-funded DISTENDER project. DISTENDER develops actionable strategies for climate change mitigation and adaptation. The assessment of strategies to address the effects of climate change is based on well-tested hydrological models. Against this background, an assessment of the effects of different PPs upon modelled streamflow is essential.

## 2 | Data and Methods

### 2.1 | Study Area

The Kamp catchment is located in northern Austria, approximately 120 km northwest of Vienna (Viglione et al. 2010). Figure 1 shows the extent and general characteristics of the catchment. The catchment has an area of approximately 1729 $km^2$ and encompasses a wide range of elevations, from around 220 to 1035 m above sea level (m.a.s.l.). The Kamp River is the longest in the "Waldviertel" region of Austria, with ecological, social, and economic importance. The river originates in the Weinsberg Forest and flows for around 160 km into the Danube east of Krems at Altenwörth at an elevation of 180 m.a.s.l. (Christoph et al., n.d.).

Three reservoirs are within the Kamp catchment: Ottenstein, Dobra, and Thurnberg, which are used for flood control, hydropower generation, and fishing.

The climate in the catchment is classified as Temperate Oceanic Climate (Cfb) according to the Köppen-Geiger climate classification (Beck et al. 2018). The annual precipitation in the catchment typically ranges from approximately 600 to 900 mm, with an average of around 711 mm (Fick and Hijmans 2017). Precipitation, temperature, and streamflow exhibit clear seasonal behaviours in the catchment, with the highest precipitation and temperature occurring in the summer (June to August) and the highest streamflow observed in the spring, particularly during March and April (Figure S1). The runoff response in the Kamp is known to be sensitive to the interplay of precipitation timing and antecedent moisture, with high flows often occurring only after sustained wet periods or snowmelt inputs (e.g., Komma et al. 2007; Nester et al. 2012). Extreme floods are usually observed in the catchment due to intense rainfall events from June through August



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

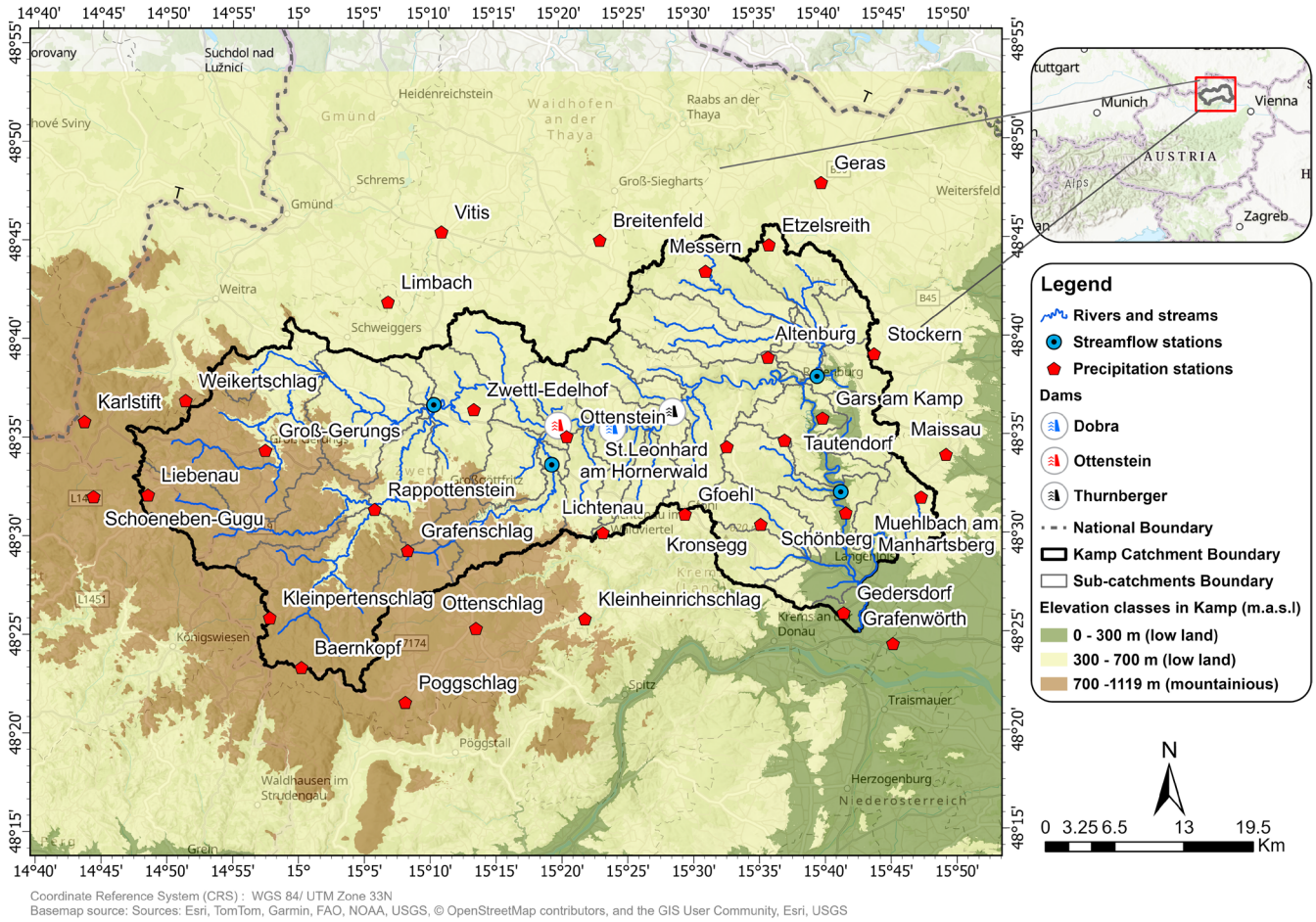


**FIGURE 1** | Overview map of the study area, including catchment boundary, major rivers and streams, dams, topography, and the locations of the precipitation and streamflow gages used in this study.

and snowmelt from March to April (Brigode et al. 2015). The average annual flow of the Kamp River is approximately 9.5 $m^3/s$ (~300 million $m^3$/year), measured at the Stiefern gauge, located around 20 km upstream of the catchment outlet. Several devastating flood events occurred in the catchment, as shown in Table S1.

The Kamp catchment is characterised by a mix of natural and human-modified landscapes, reflecting varied topography and land use practices. Forests dominate the upper regions, while agriculture is more prominent in the lower areas, with water bodies and diverse land uses supporting the local economy and ecosystem.

## 2.2 | Data Used

### 2.2.1 | Ground-Based Precipitation Data

Daily precipitation measurements for the period 1998–2020 from 33 ground-based stations, distributed within and around the Kamp catchment, were obtained from the electronic Hydrography platform (eHYD), operated by the Hydrographic Service of Austria (available online at: https://ehyd.gv.at/?g_suche=208520) and from GeoSphere Austria—Dataset (available online at: https://www.geo-sphere.at/dataset/). Figure 1 highlights the uneven distribution of the stations' network. Sparser data coverage is evident over the upper part of the catchment compared to low-elevation regions. Only 2.5% (9 stations) of the gauge stations are located at altitudes above 700 m.a.s.l. The stations' daily time series were subjected to quality control and checked for data gaps. Only the stations with less than 15% missing daily data were included in the analysis. Table S2 presents the percentage of missing data for the 33 selected gauge stations.

### 2.2.2 | Evaluated Gridded and Reanalysis Precipitation Products

The performance of four PPs was evaluated over the Kamp catchment for the period 1998–2020. The PPs are the Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS v2; Funk et al. 2015), the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM) mission Final (IMERG-F v07; Huffman et al. 2023, 2024), the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis dataset (ERA5-Land; Muñoz-Sabater 2019; Muñoz-Sabater et al. 2021), and the Spatiotemporal Reanalysis Dataset for Climate in Austria (SPARTACUS; Hiebl and Frei 2018). Table 2 provides an overview



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of the main characteristics of the PPs used in this study, including their spatial and temporal resolutions. More information about the evaluated products can be found in Section S1.

#### 2.2.3 | Data for Hydrological Modelling

Different datasets, including topography, soil types, land use/land cover, and meteorological data (precipitation, maximum and minimum air temperature, relative humidity, wind speed, and solar radiation), are required as inputs for the SWAT+ model (Bieger et al. 2017). Additionally, measured streamflow data are required for model calibration and validation. In the present study, the required datasets were collected from different sources (Table 3). Precipitation data from the PPs were used to to drive SWAT+ and identify the best-performing product for hydrological simulation.

### 2.3 | Methodology

We applied point-to-pixel and hydrological modelling evaluation techniques to assess the performance of PPs in the study area. A point-to-pixel approach (Bai and Liu 2018; Baez-Villanueva et al. 2018; Zhang et al. 2018; Basheer and Elagib 2019; Basheer et al. 2019; Gumindoga et al. 2019; McNamara et al. 2021; Khatakho et al. 2024) was applied to compare the time series of 33 precipitation gauges with the corresponding pixel values of the PPs at a daily timescale. For hydrological modelling, we utilized the restructured version of SWAT (SWAT+ rev 61.0.1; Arnold et al. 2018) for the Kamp catchment. SWAT+ is used to assess the effects of different PPs on the streamflow simulation. SWAT+ has been tested and applied extensively worldwide (Krysanova and Arnold 2008; Arnold and Fohrer 2005; Gassman et al. 2007; Akoko et al. 2021).

The methodology used in this study is presented in Figure 2.

#### 2.3.1 | Definition and Evaluation of Extreme Events

In this study, we used the P95th and P99th indices to identify daily extreme and very extreme precipitation events, respectively. These indices are defined as the daily precipitation amounts that exceed the 95th or 99th percentile of precipitation amounts on wet days within the evaluation period (1998–2020). A day is considered wet when the precipitation amount is greater than or equal to 1 mm ($p \geq 1$ mm) and is considered dry when the precipitation amount is less than 1 mm ($p < 1$ mm). Many studies have used the 95th and 99th percentiles of days with $p \geq 1$ mm as thresholds for evaluating precipitation extremes (Sharifi et al. 2019; Cavalcante et al. 2020; Keikhosravi-Kiany and Balling 2024). The 95th and 99th percentiles of daily precipitation on wet days were calculated based on observations from the precipitation gauges over the evaluation period. Next, we identified all days when observed daily precipitation exceeded these thresholds at each station. For these days, corresponding daily precipitation values were extracted from the PPs for direct comparison. To evaluate the ability of the PPs to detect the sequence of extreme precipitation events, such as dry and wet spells, we used the maximum annual Consecutive Wet Days (CWD) and Consecutive Dry Days (CDD) indices, which were developed by the Expert Team on Climate Change Detection and Indices (ETCCDI; Zhang et al. 2011). The selected indices for precipitation extremes are reported in Table 4.

#### 2.3.2 | Point-to-Pixel Evaluation of Precipitation Products

**2.3.2.1 | Performance Metrics and Ranking of Precipitation Products.** In this study, we performed a statistical evaluation using seven metrics to identify the best-performing PPs for the study area. The different performance metrics address different aspects of performance. We used: the Root Mean Square Error (RMSE), the Mean Bias Error (MBE), the Willmott index of agreement ($d$), and the modified Kling-Gupta Efficiency (KGE) and its three components: the linear correlation coefficient ($r$), the bias ratio (Beta; $\beta$), and the variability ratio (Gamma; $\gamma$). Moreover, to evaluate the ability of PPs to detect different precipitation events (Table 4), we used three categorical metrics: Probability of Detection (POD), False Alarm Ratio (FAR), and Frequency of Bias Index (FBI). Table S3 summarizes the performance metrics and their formulas. To draw overall conclusions on the performance of PPs and select the best-performing precipitation product in the study area, a multi-step ranking approach was applied. Details on the PPs ranking methodology can be found in Section S2.

#### 2.3.3 | Hydrological Evaluation of Precipitation Products

**2.3.3.1 | SWAT+ Hydrological Modelling Framework.** We used SWAT+ to assess the performance of PPs as a boundary condition for hydrological modelling. SWAT is a process-based, semi-distributed eco-hydrological model developed by the USDA Agricultural Research Service to simulate the impact of environmental changes on hydrological systems with varying soils, land use, and management conditions over extended periods (Arnold et al. 2012). The latest version of SWAT+ (SWAT+ rev 61.0.1; Arnold et al. 2018) was used to develop a daily hydrological model for the Kamp catchment. SWAT+ enables the simulation of streamflow across a wide range of spatial and temporal scales, environmental conditions, land management practices, and land use and climate change scenarios (Bieger et al. 2017). It utilizes the deterministic paradigm. Thus, the same set of model drivers acting on the same set of model parameters and state variables will produce the same model results. As a result of this paradigm, spatial units with the same characteristics can be merged into Hydrologic Response Units (HRU). HRUs are areas characterized by unique combinations of land use, soil, slope class, and management practices (Neitsch et al. 2005). SWAT+ is driven by generally available input data: meteorological data, topography data, land use, and soil information. It uses empirical and physically based equations to represent the hydrological processes within the catchment by subdividing it into subbasins and calculating the vertical fluxes in each HRU and the lateral fluxes (Abbaspour et al. 2013; Neitsch et al. 2011). Bieger et al. (2017) provide a general description of data preparation, watershed configuration, delineation of HRUs, Landscape units (LSUs), and other SWAT+ model components.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 2** | Summary of gridded precipitation products used in this study and their spatial and temporal resolutions and coverages.

| Dataset | Type | Method | Spatial resolution | Temporal resolution | Spatial coverage | Temporal coverage | Source | References |
|---|---|---|---|---|---|---|---|---|
| CHIRPS v2.0 | Grided dataset | Combines satellite-derived precipitation estimates (infrared Cold Cloud Duration data) with in situ ground-based rain gauge observations | 0.05° (~5 km) | Daily, pentadal, monthly | Global (50° S to 50° N) | 1981-present | Produced by Climate Hazards Center at the University of California, Santa Barbara (UCSB) and partners. | Funk et al. (2015) |
| ERA5-Land | Reanalysis dataset | Derived from downscaling ERA5 (same source as ERA5) | 0.1° (~10 km) | Hourly | Global (land areas only) | 1950-present | Produced by ECMWF. | Muñoz-Sabater (2019); Muñoz-Sabater et al. (2021) |
| IMERG Final v07 | Grided dataset | It integrates data from multiple satellite-based sensors (e.g., the Global Precipitation Measurement (GPM) mission) and incorporates limited ground-based rain gauge data for calibration. | 0.1° (~10 km) | Half-hourly, daily, monthly | Global (60° S to 60° N) | 1998-present | Produced by NASA's GPM team. | Huffman et al. (2023); Huffman et al. (2024) |
| SPARTACUS v2.1 | Grided dataset | It is based on in situ ground station observations across Austria, using interpolation techniques to produce consistent spatial coverage. | 0.01° (1 km) | Daily | Austria | 1991-present | Produced by GeoSphere Austria | Hiebl and Frei (2018) |



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 3** | Summary of the data used in SWAT+ hydrological modeling.

| Data type | Source | Resolution | References |
|---|---|---|---|
| Digital surface model (DSM) | ALOS World 3D (AWED30) | 30 m | Japan Aerospace Exploration Agency (2021) |
| Landuse/land cover | Corine 2018 | 100 m | EEA (2020) |
| Soil data | Digital Soil World Map (DSWM) | 5 km | FAO (2003) |
| Temperature | SPARTACUS v2.1 | 0.01° (1 km) | Hiebl and Frei (2018) |
| Precipitation | Gauge stations data from the electronic Hydrography platform (eHYD), operated by the Hydrographic Service of Austria, and GeoSphere Austria Datahub | Points | Hydrographic Service of Austria (2022); GeoSphere Austria (2024). |
| | SPARTACUS v2.1 | 0.01° (1 km) | Hiebl and Frei (2018) |
| | IMERG Final v07 | 0.1° (~10 km) | Huffman et al. (2023, 2024) |
| | CHIRPS v2.0 | 0.05° (~5 km) | Funk et al. (2015) |
| | ERA5-Land | 0.1° (~10 km) | Muñoz-Sabater (2019); Muñoz-Sabater et al. (2021) |
| Evapotranspiration, wind speed, relative humidity, and solar radiation | ERA5-Land (ECMWF) | 0.1° (~10 km) | Muñoz-Sabater (2019); Muñoz-Sabater et al. (2021) |
| Streamflow data | Bundesmeldenetz (BMN) and electronic Hydrography platform (eHYD), operated by the Hydrographic Service of Austria | Point | Hydrographic Service of Austria (2022) |

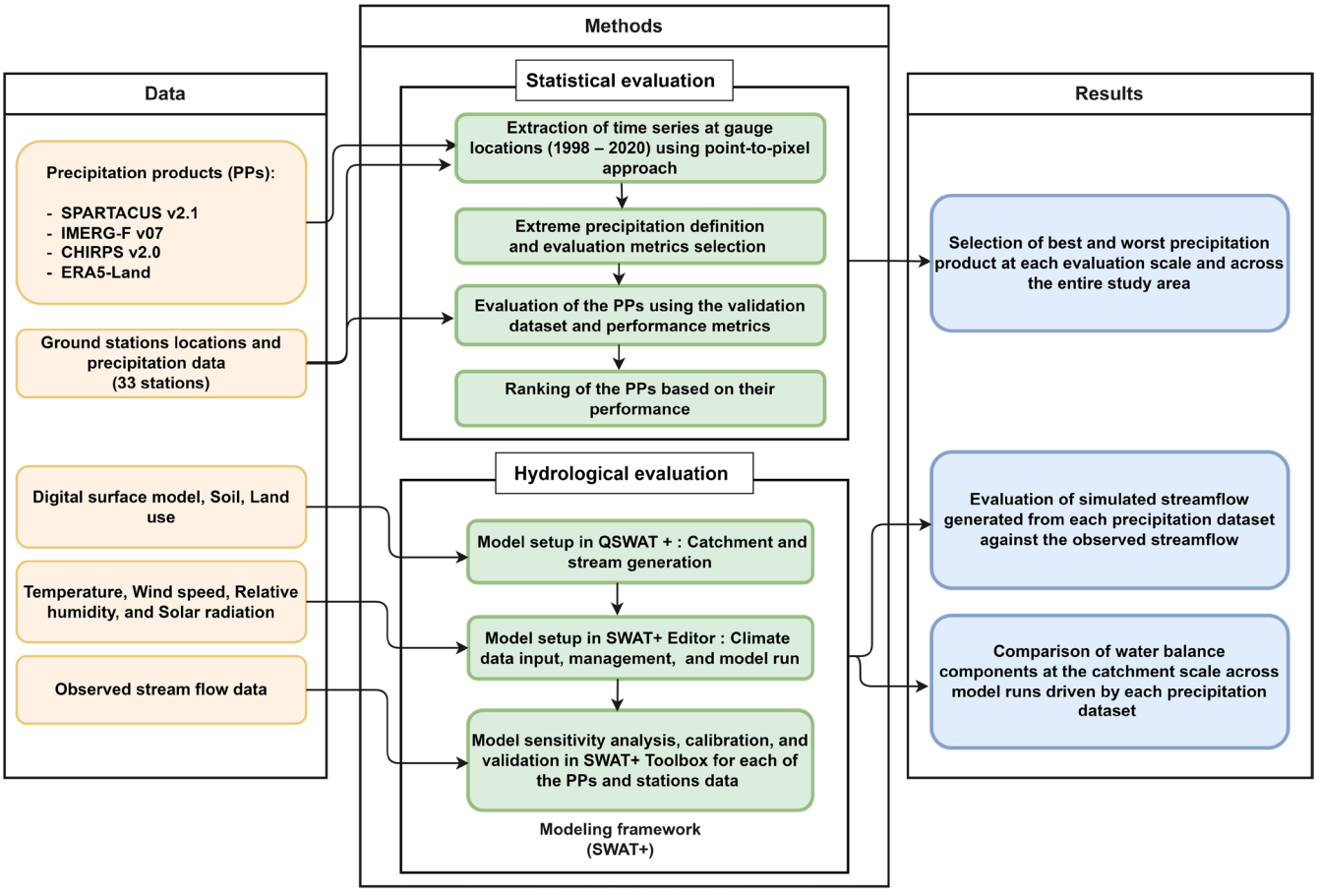


**FIGURE 2** | Methodological framework applied in this study.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3 provides an overview of the model input data used in this study. A 30 m Digital Surface Model (DSM) was used to generate the streams, delineate the watershed and its subbasins, and define three slope classes (0%–5%, 5%–10%, and > 10%). HRUs are defined by adopting thresholds of 0% for land use, 5% for soil types, and 5% for slope. Consequently, 7171 homogeneous HRUs are determined, and 664 LSUs are created by grouping HRUs that fall within the same spatial zone (upland, floodplain, or terrace). Figure S2 illustrates the delineated watershed, comprising 41 subbasins and the created LSUs, along with the derived stream network, the locations of the three reservoirs, the generated slope classes, and the utilised soil and land use classes. The model was parameterized in the SWAT+ Editor version 3.0.8 to incorporate crop and reservoir management parameters. The Penman-Monteith method was employed to estimate potential evapotranspiration (PET), and the Muskingum method was used as a flow routing method. The PPs data were extracted at virtual stations distributed within each subbasin and around the catchment to overcome the problem of uneven distribution of ground stations, especially in the upper part of the catchment. SWAT+ considers the closest virtual station to the subbasin centroid in the simulation. To assess the influence of selecting PPs as forcing in the model, SWAT+ was independently calibrated for each PP, resulting in five calibrated models that were used to simulate streamflow. Calibration, validation, and sensitivity analysis were performed for each of the five SWAT+ model runs.

For this study, 25 SWAT+ model parameters that targeted runoff, evapotranspiration, soil infiltration, lateral flow, percolation to groundwater, and snow were identified and used for calibrating each model (Table S6). To find the most sensitive model parameters according to each PP, a Global Sensitivity Analysis (GSA) of the selected parameters was performed in SWAT+ version 3.0.6 using the Sobol method (Sobol 2001) with 1040 simulations (samples) and the Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe 1970), as the objective function.

**TABLE 4** | The selected indices for precipitation extremes.

| Index name | Index definition | Unit |
|---|---|---|
| Daily extreme precipitation (P95th = $p$ > 95th percentile of wet days) | Daily extreme precipitation events when $p$ > 95th percentile of wet days during the period of analysis. | mm |
| Daily very extreme precipitation (P99th = $p$ > 99th percentile of wet days) | Daily extreme precipitation events when $p$ > 99th percentile of wet days during the period of analysis. | mm |
| Dry days/No rain | Days when $p$ < 1 mm. | mm |
| Consecutive wet days (CWD) | Maximum annual number of consecutive wet days (when $p \geq 1.0$ mm). | days |
| Consecutive dry days (CDD) | Maximum annual number of consecutive dry days (when $p$ < 1.0 mm). | days |

Each model was calibrated to the observed streamflow data at the Zwettl Bahn Brücke gauge station (see Figure 1), considering the period from 2006 to 2018, and was validated for the period from 2000 to 2003 with a two-year warm-up period (2004 and 2005). Latin-hypercube Sampling Iteration (CALSI), an automatic calibration method, was used to calibrate the selected parameters with 3000 model simulations.

**2.3.3.2 | SWAT+ Model Performance Evaluation.** The Nash–Sutcliffe Efficiency (NSE), modified Kling-Gupta Efficiency (KGE), Root Mean Square Error (RMSE), and Percent Bias (Pbias) were selected to evaluate the performance of the SWAT model, driven by each precipitation product, in simulating streamflow during the calibration and validation periods. The equations of these selected evaluation metrics are shown in Table S3.

To further assess the hydrological utility of the PPs in simulating flood event dynamics, we investigated their ability to reproduce streamflow peaks and associated hydrological responses within the SWAT+ model framework. Specifically, we examined how different precipitation forcings influenced (i) event-based total runoff volume, (ii) peak flow generation, (iii) event-based runoff coefficient ratios, and (iv) modifications in the state variables of the model through parameter calibration.

We selected four historical flood events in the Kamp catchment (6–10 August 2002; 11–15 August 2002; 5–15 July 2005; and 14–19 August 2005) that represent different conditions in terms of season, precipitation characteristics, and antecedent soil moisture states, as identified by Komma et al. (2007). The characteristics of these flood events are summarized in Table S4.

The hydrological performance of PPs in the selected flood events was evaluated using five diagnostic statistics. The Root Mean Square Error (RMSE, $m^3/s$) was employed to quantify the magnitude of deviations between simulated and observed streamflow hydrographs. The relative bias in precipitation amount ($BIAS_p$, %) was calculated to assess systematic over- or underestimation of precipitation input by each product, while the relative bias in total runoff volume ($BIAS_v$, %) was used to evaluate the cumulative effect of precipitation biases on event-scale runoff generation. The relative bias in peak streamflow ($BIAS_{peak}$, %) was applied to specifically evaluate the accuracy of simulated flood magnitudes, whereas the peak timing error (PTE, days) was used to detect temporal mismatches between simulated and observed flood peaks. Collectively, these metrics provide a comprehensive evaluation of the ability of the different precipitation forcings to reproduce both the magnitude and timing of extreme flood events in the catchment.

To further investigate the effect of precipitation forcings on the catchment's hydrological response, we analyzed runoff coefficients and their relative biases ($BIAS_{RC}$). The runoff coefficient (RC) quantifies the fraction of precipitation that is transformed into direct runoff during an event. For the reference values, we used estimates from Komma et al. (2007), who derived event-based direct runoff depths and runoff coefficients by separating baseflow from observed event hydrographs over the same catchment of this study.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For each event, the simulated runoff coefficient was computed as:

$$RC_{event} = \frac{\sum_{i=1}^{n} Q_i direct}{\sum_{i=1}^{n} P_i} \quad (1)$$

where $RC_{event}$ is the runoff coefficient for a given flood event; $Q_i direct$ is the daily direct runoff (surface runoff + lateral flow, mm) on day i simulated by the SWAT+ model forced by a PP; $P_i$ is the daily precipitation (mm) on day i; and *n* is the number of days in the flood event period.

This formation enabled a direct comparison of simulated runoff coefficients from each PP with the reference values, thereby quantifying the relative biases introduced by the precipitation forcings.

## 3 | Results

### 3.1 | Statistical Evaluation of the Precipitation Products

#### 3.1.1 | Ability to Capture Daily Precipitation Patterns

First, we calculated some basic diagnostic statistics to systematically compare the products with the ground-based data. The comparison results are summarized in Table 5. All precipitation products exhibit small differences in mean daily precipitation values compared to the observed data. CHIRPS v2.0 shows higher values in the standard deviation and maximum daily precipitation values, while other products show lower values than those shown by ground stations. The number and percentage of wet days are higher with SPARTACUS v2.1, IMERG-F v07, and ERA5-Land compared to the results from ground-based precipitation, while CHIRPS v2.0 shows the opposite pattern.

Figure 3 presents box plots of the six metrics used to evaluate the performance of PPs with ground precipitation data at the daily scale from 1998 to 2020. The results show that, based on the median performance metrics of the 33 stations, SPARTACUS v2.1 outperforms the other products, achieving the highest KGE value of 0.83 and the lowest errors, with an MBE of −0.02 mm/day and an RMSE of 2.53 mm/day. IMEGR-F v07 and ERA5-Land follow, with KGE values of 0.63 and 0.44, MBE values of 0.19 and 0.31 mm/day, and RMSE values of 3.81 and 3.78 mm/day, respectively. CHIRPS v2.0 demonstrates the lowest overall performance, with a KGE of 0.40 and an RMSE of 6.13 mm/day. The results of MBE reveal varying over- and under-estimation of daily precipitation as recorded by the station observations. All PPs except SPARTACUS v2.1 are characterized by daily precipitation overestimation, with the highest average MBE of 0.31 mm/day recorded in ERA5-Land.

#### 3.1.2 | Ability to Detect Precipitation Extremes

Figure 4 shows the results of PPs' detection ability for no rain (dry days), extreme ($p > 95$th percentile), and very extreme ($p > 99$th percentile) daily precipitation intensity classes using three categorical indices (i.e., POD, FBI, and FAR). All PPs show a high detection skill for dry days or precipitation intensity less than 1 mm/day, with the lowest POD median value of 0.78 (ERA5-Land) and the highest value of 0.95 (SPARTACUS v2.1). However, their detection skills are relatively poor for extreme and very extreme precipitation. SPARTACUS v2.1, followed by IMERG-F v07, exhibited better detection abilities for extreme (POD values of 0.63 and 0.40, respectively) and very extreme (POD values of 0.38 and 0.34, respectively) precipitation than the other products. Considering the FBI results, CHIRPS v.20 exhibits a systematic overestimation in all precipitation intensities (FBI > 1). In contrast, IMERG-F v07, followed by SPARTACUS v2.1, demonstrates better performance in no rain, extreme, and very extreme precipitation. Additionally, SPARTACUS v2.1, followed by ERA5-Land, exhibited fewer false alarms in all precipitation intensity classes analyzed, with FAR values ranging from 0.02 to 0.52.

#### 3.1.3 | Ability to Detect Sequences of Precipitation Extremes

Figure 5 shows the evaluation results of PPs in capturing the longest sequences of dry days (quantified by CDD) and longest sequences of wet days (quantified by CWD) in the Kamp catchment. Based on MBE results, CHIRPS v2.0 overestimates the CDD and underestimates the CWD, with median MBE values of 5.17 and −1.43, respectively. In contrast, ERA5-land exhibits the opposite pattern, with median MBE values of −5.43 and 2.83, respectively. SPARTACUS v2.1 shows a minor underestimation of CDD and an overestimation of CWD (MBE = −0.09 and 0.13,

**TABLE 5** | Basic diagnostic statistics of the daily precipitation in Kamp Catchment over the period 1998–2020: Mean, standard deviation (SD), maximum precipitation, number (percentage) of wet days ($p \geq 1$ mm/day), and number (percentage) of dry days ($p < 1$ mm/day).

| Precipitation source | Mean (mm) | SD (mm) | Max precipitation (mm) | Number (percentage) of wet days | Number (percentage) of dry days |
|---|---|---|---|---|---|
| Stations data | 1.90 | 5.01 | 94.91 | 2312 (27.6%) | 5969 (71.3%) |
| SPARTACUS v2.1 | 1.88 | 4.67 | 86.61 | 2410 (28.7%) | 5905 (70.3%) |
| IMERG Final v07 | 2.05 | 4.98 | 79.08 | 2667 (31.7%) | 5731 (68.2%) |
| ERA5-Land | 2.22 | 4.19 | 68.48 | 3239 (38.6%) | 5039 (59.9%) |
| CHIRPS v2.0 | 2.02 | 6.52 | 115.15 | 1577 (18.8%) | 6824 (81.2%) |



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

respectively). Based on the d and RMSE performance metrics results, SPARTACUS v2.1 performs the best in capturing the CDD and CWD, as it is the closest product to the optimal metrics value (vertical dotted red line), followed by IMERG-F v07.

#### 3.1.4 | Evaluation Based on Topographical Gradient

Figure 6 illustrates the association between topographical gradient and the performance of PPs in capturing extreme precipitation. PPs show varying performance across different elevation zones. All PPs underestimate extreme precipitation events, as indicated by the negative MBE results. In addition, all products have low correlation with ground stations (R values ranging from 0.56 to 0.14). The results of KGE, as shown in Figure 6, reveal that SPARTACUS v2.1 performs better in the mountainous regions of the catchment (between 700 and 1119 m.a.s.l.), with a median KGE value of 0.42, compared to lower elevations. In contrast, the performance of IMERG-F v07 in the lowland area (elevation 0–300 m.a.s.l.) is relatively better than in higher elevations, with a median KGE value of 0.17 compared to 0.14 and 0.08 in elevation zones of 300–700 and 700–1119 m.a.s.l., respectively. Compared to SPARTACUS v2.1 and IMERG-F v07, ERA5-Land and CHIRPS v2.0 performed worse in capturing the extreme precipitation events in all elevation zones.

#### 3.1.5 | Ranking of Precipitation Products

Following the steps outlined in Section S2, Table S5 presents the results of the Composite performance score (S) and the final ranking of PPs in the study area. The results show that SPARTACUS v2.1 demonstrates superiority among the four products across all the evaluation scales, with the highest S value and the lowest overall ranking. IMERG-F v07 comes second, followed by ERA5-Land and CHIRPS v2.0. Therefore, overall, SPARTACUS v.2.1 was found to be the best-performing product, followed by IMERG-F v07, ERA5-Land, and then CHIRPS V 2.0.

### 3.2 | Evaluation of Precipitation Products as Inputs for Streamflow Simulation

#### 3.2.1 | Model Parameters' Sensitivity

Sobol's global sensitivity analysis method analysed all 25 calibration parameters used in each PPs data forced model. The sensitive parameters are then ranked based on their effect on the model results using the NSE objective function. Table 6 exhibits the results of the twelve most sensitive SWAT+ model parameters from each of the PP-based models. While the snowmelt base temperature (snomelt_tmp) and the minimum snowmelt rate on December 21 (winter), when solar radiation is lowest (snomelt_min), were the most sensitive parameters for SPARTACUS v2.1 and ERA5-Land, respectively, IMERG-F v07 and CHIRPS v2.0 mainly were sensitive to the moist bulk density of the soil (bd). In general, soil, groundwater, and snow parameters were identified as the most influential parameters in the catchment, as also shown by other studies done in the catchment (e.g., Viglione et al. 2010). This suggests that runoff generation is highly influenced by subsurface processes and antecedent conditions, which can lead to non-linear responses depending on soil saturation and snowmelt timing.

#### 3.2.2 | Model Calibration and Validation

Table S6 presents the automatic-calibration analysis results for the model parameters, reported as the fitted values obtained using the four different PPs. The results show that the final fitted values for the calibration parameters vary from one precipitation dataset to another, resulting in different model performances when comparing simulated streamflow with observed

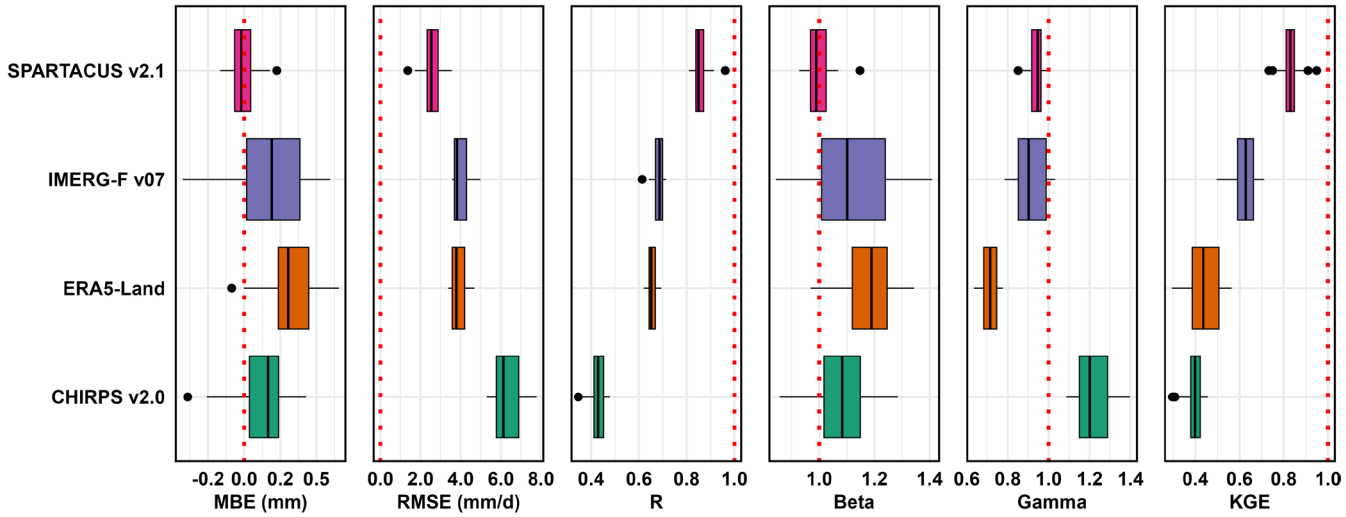


**FIGURE 3** | Comparison of the selected precipitation products with the ground-based precipitation estimation at the daily scale, using (from left to right) Mean Bias Error (MBE; mm/day), Root Mean Square Error (RMSE; mm/day), Linear correlation (R), Bias ratio (Beta), Variability ratio (Gamma), and modified Kling-Gupta Efficiency (KGE). In the boxplots, the vertical continuous line in the middle represents the median, the left and right ends of the box indicate the 25th and 75th percentiles, respectively, and the whiskers represent the range of the performance metric values. Outliers are shown as black dots. The vertical dotted red line indicates the optimal values for each performance metric.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

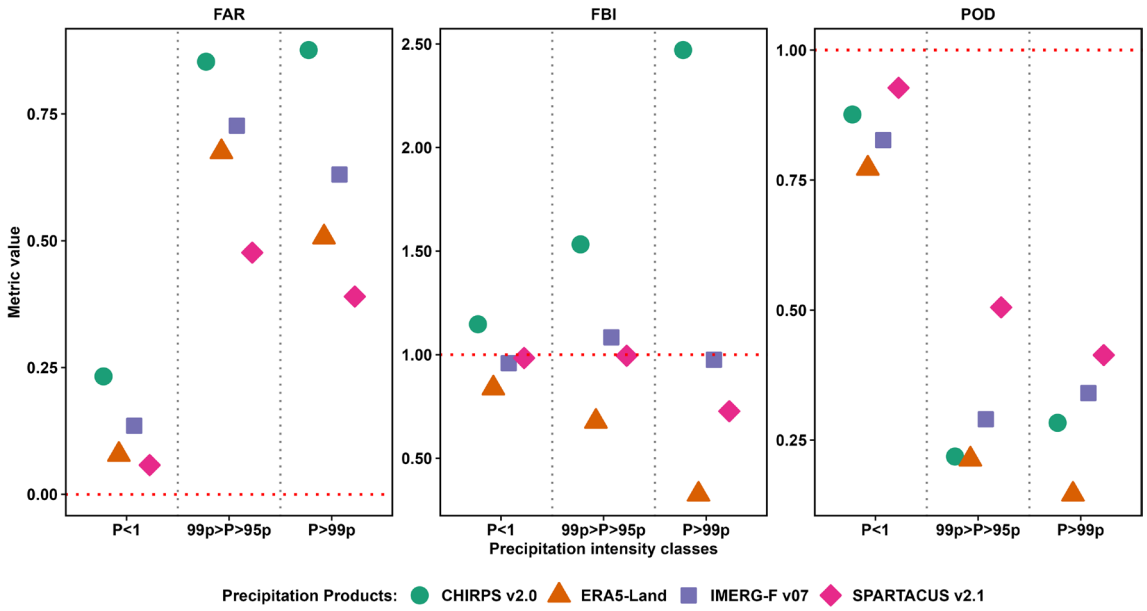


**FIGURE 4** | Evaluation of the precipitation products' detection ability for no rain (dry days), extreme daily precipitation ($p > 95$th percentile of wet days), and very extreme daily precipitation ($p > 99$th percentile of wet days) intensity classes, using the median values (for the 33 evaluation location-say) of three categorical metrics: False Alarm Ratio (FAR), Frequency of Bias Index (FBI), and Probability of Detection (POD). $p < 1$ = The intensity of the No rain/dry days class when precipitation is less than 1 mm; $99p > p > 95p$ = The precipitation intensity class when precipitation is greater than extreme daily precipitation and less than very extreme daily precipitation; $p > 95p$ = The precipitation intensity class when precipitation is greater than every extreme daily precipitation.

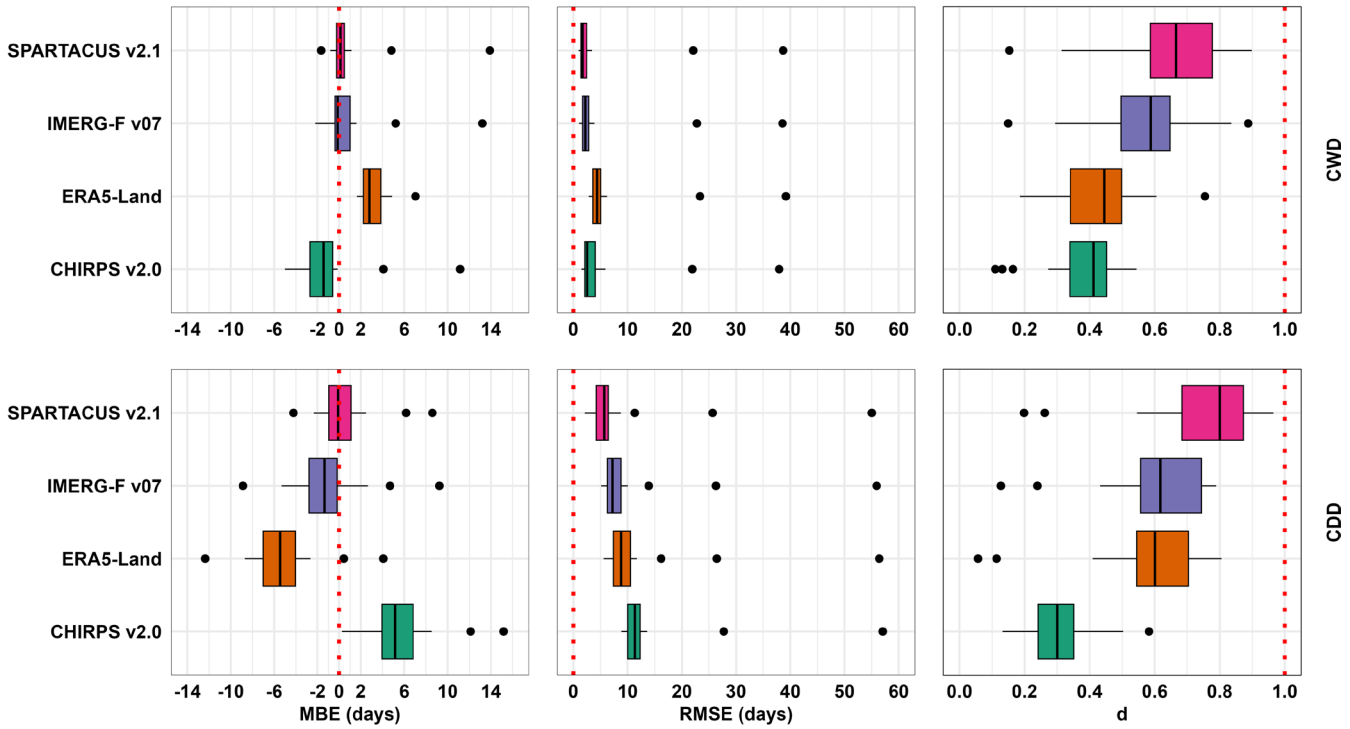


**FIGURE 5** | Comparison of the selected precipitation products with the ground-based observations in capturing the sequence of consecutive wet days (CWD, top) and consecutive dry days (CDD, bottom) in the Kamp catchment, using (from left to right) MBE (days), RMSE (days), and Willmott index (d). In the boxplots, the vertical continuous line in the middle represents the median, the left and right ends of the box indicate the 25th and 75th percentiles, respectively, and the whiskers represent the range of the performance metric values. Outliers are shown as black dots. The vertical dotted red line indicates the optimal values for each performance metric.

streamflow. Figure 7 illustrates the observed and simulated streamflow at the Zwettl Bahn Brücke gauge station during the calibration and validation periods, utilizing data from the four PPs. For a closer visual examination of model outputs, the simulated and observed hydrographs for one year in the calibration period (2006) and the validation period (2002) are presented in



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

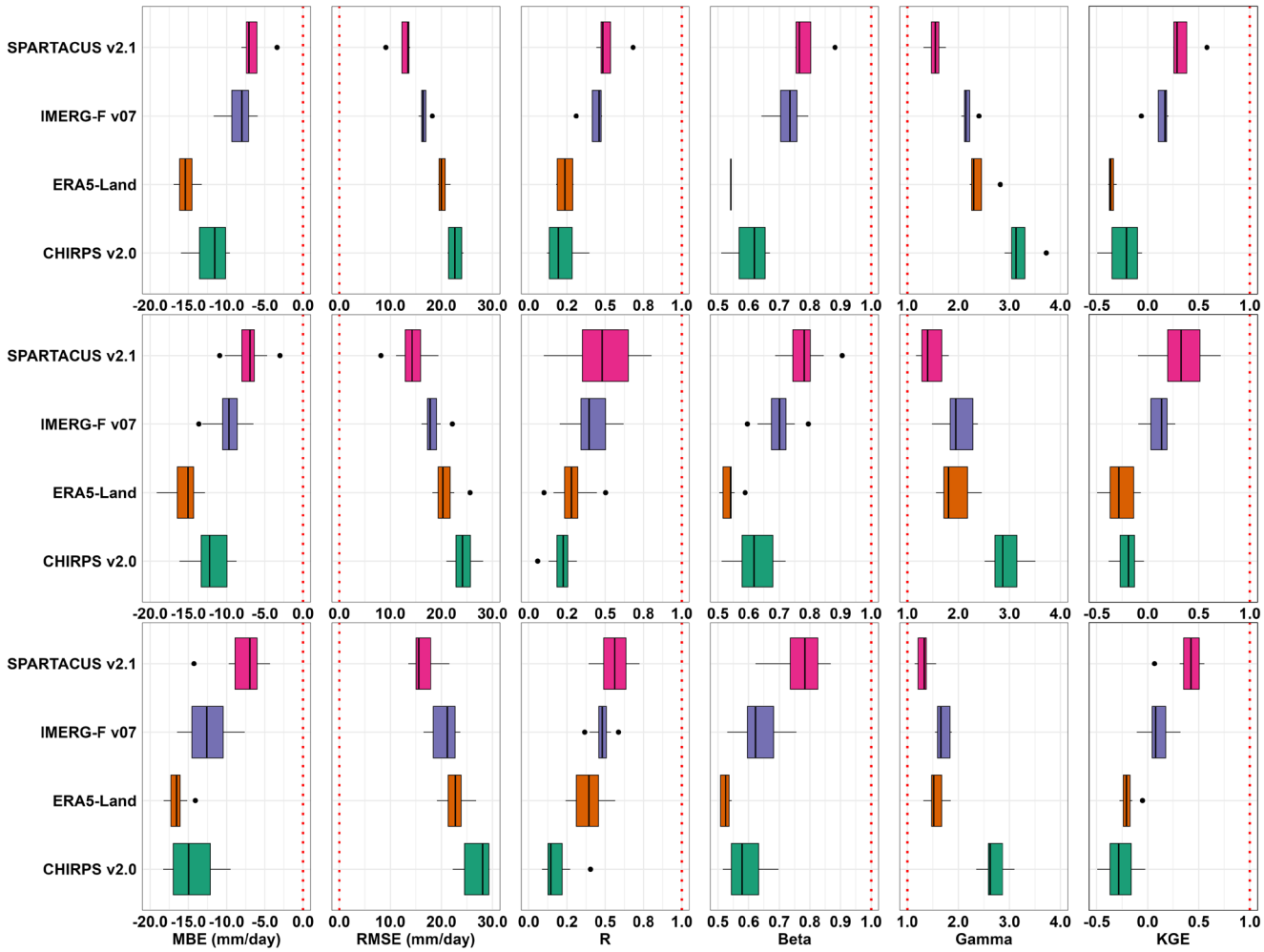


**FIGURE 6** | Comparison of the selected precipitation products with the ground-based observations in capturing extreme precipitation ($p > 95$th percentile of wet days) in different elevation zones in the catchment, using (from left to right) MBE (mm/day), RMSE (mm/day), Linear correlation (R), Bias ratio (Beta), Variability ratio (Gamma), and modified Kling-Gupta Efficiency (KGE). In the boxplots, the vertical continuous line in the middle represents the median, the left and right ends of the box indicate the 25th and 75th percentiles, respectively, and the whiskers represent the range of the performance metric values. Outliers are shown as black dots. The vertical dotted red line indicates the optimal values for each performance metric.

Figure 8. Except for the ERA5-Land-based model, the simulated streamflow corresponds well with the observed hydrological pattern in both the calibration and validation periods for all PPs. However, the peak streamflow is underestimated with all PPs. The consistent underestimation of peak flows could indicate that some products miss key aspects of event build-up, such as wet antecedent conditions that precede large runoff responses in this catchment.

The results of the performance metrics used for the model (Table 7) enable a comparison of the effects of various PPs. The IMERG-based simulations exhibit good model performance during calibration and validation, as assessed by the performance categories proposed by Moriasi et al. (2007), with NSE values of 0.70 in calibration and 0.60 in validation. IMERG-F v07 resulted in less underestimation (Pbias = −1.99%) during calibration; however, it led to an overestimation of streamflow during the validation period, with a Pbias value of 13.06%. The CHIRPS v.20 model tended to overestimate the simulated streamflow by 2.72% during calibration and by 15.21% during validation, as indicated by the Pbias results. Furthermore, CHIRPS v2.0 yielded good model performance during the calibration period (NSE = 0.64), but its performance dropped to unsatisfactory levels during the validation years, with an NSE value of 0.37. ERA5-Land caused the lowest NSE values of −0.14 and −0.01 and the highest underestimation of streamflow (Pbias values of −26.36% and −36.20%) during calibration and validation, respectively, indicating unsatisfactory model performance. The SPARTACUS v2.1 model demonstrates superior performance compared to other datasets, with improved NSE, KGE, and RMSE, and exhibits lower bias in both calibration and validation periods. It achieved the highest NSE values of 0.75 and 0.65, with less overestimation of simulated streamflow compared to other datasets, as indicated by Pbias values of 3.23% and 4.83% during calibration and validation, respectively.

### 3.2.3 | Event-Based Evaluation of the Hydrological Response of the Model Under Extreme Flood Events

Figure 9 shows the observed and simulated daily hydrographs of the selected historical flood events at the Zwettl Bahn Brücke



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 6** | Results of the global sensitivity analyses: Selection and ranking of the twelve most sensitive model parameters[a] using the Sobol method in SWAT+ Toolbox.

| Sensitivity ranking in descending order | SWAT+ parameter | | | |
|---|---|---|---|---|
| | SPARTACUS v2.1 | IMERG Final v07 | ERA5-Land | CHIRPS v2.0 |
| 1 | snomelt_tmp | bd | snomelt_min | bd |
| 2 | revap_co | surlag | perco | cn2 |
| 3 | snomelt_min | chk | cn3_swf | snomelt_tmp |
| 4 | cn2 | k | snomelt_max | perco |
| 5 | canmx | awc | surlag | snomelt_min |
| 6 | perco | z | snomelt_lag | canmx |
| 7 | esco | bf_max | esco | awc |
| 8 | surlag | revap_min | awc | surlag |
| 9 | z | snomelt_max | z | k |
| 10 | cn3_swf | alpha | bd | snofall_tmp |
| 11 | bd | snomelt_lag | canmx | snomelt_lag |
| 12 | snofall_tmp | epco | k | esco |

[a]The description of the parameters mentioned in this table is provided in Table S5.

streamflow station using the SPARTACUS v2.1-, IMERG-F v07-, ERA5-Land-, and CHIRPS v2.0-based precipitation data inputs. The results of the hydrological performance, as quantified using RMSE, $BIAS_p$, $BIAS_v$, $BIAS_{peak}$, and PTE, are presented in Table 8. Across all events (Figure 9; Table 8), the four PPs produced distinct streamflow magnitudes. However, the SWAT+ model systematically underestimated high flows when forced with the selected PPs. SPARTACUS v2.1, to some extent, provided a better representation of these flood events compared to other datasets, with an RMSE ranging from 11 to 83 $m^3/s$ and a PTE of ≤1 day, followed by the IMERG-driven model, which reproduced event timing but underestimated flood magnitudes. ERA5-Land- and CHIRPS v2.0-based simulations exhibited lower performance, with substantial underestimation of both runoff volumes and flood peaks, particularly under dry or moderate antecedent soil moisture conditions.

The total runoff volume bias ($BIAS_v$) quantifies the "missing water" during each event compared to observations. SPARTACUS v2.1-based simulation reproduced approximately 60%–80% of observed event runoff volumes (BIASv = −21% to −41%, or 2.2–21.4 $Mm^3$ deficit), IMERG-F v07- and CHIRPS v2.0- based models underestimated by 26%–47% (2.8–24.6 $Mm^3$ deficit), and 71%–87% (45.5–60 $Mm^3$ deficit), respectively. The ERA5-Land-based simulation exceeded −85%. Similarly, SPARTACUS v2.1-based simulation underestimated peak discharge by 41%–56%, IMERG-F v07- and CHIRPS v2.0- based models by 49%–87%, and ERA5-Land-based simulation by up to −98%. Peak timing error (PTE) results were minimal for SPARTACUS v2.1- and IMERG-F v07-based simulation (≤1 day), but reached ±2 days for the ERA5-Land-based simulation, confirming that rainfall timing errors directly influenced the hydrological response.

Event-to-event comparisons highlight how antecedent moisture modulated the precipitation-runoff relationship. As shown in Table 8, Event 1 (dry initial soil; ~1000-year return period) produced an extreme flood with high infiltration potential, as also reported by Komma et al. (2007). SPARTACUS v2.1-based simulation reproduced around 59% of the observed total runoff volume; CHIRPS v2.0- and IMERG-F v07-based models reproduced 19% and 53%, whereas ERA5-Land-forced simulation underestimated the runoff due to its inability to capture localized convective rainfall (Terblanche et al. 2022; Gomis-Cebolla et al. 2023) and reproduced only 4% of the observed runoff volume. Event 2 (highly wet antecedent conditions; ~500-year return period) generated the highest observed runoff coefficient (0.49; Table S4), showing efficient rainfall-runoff conversion. SPARTACUS v2.1 and IMERG-F v07 overestimated precipitation by 40% and 35%, respectively, while CHIRPS v2.0 remained close to the observed precipitation; yet all underestimated flood peaks ($BIAS_{peak}$ = −55% to −98%), suggesting that surplus precipitation was compensated by higher simulated infiltration and evapotranspiration losses during calibration. ERA5-Land, with a precipitation bias of −86%, fails to reproduce the hydrograph shape or magnitude. The 2005 floods were less intense and hydrologically more moderate. Event 3 (moderately wet antecedent soil) was captured to some extent by SPARTACUS v2.1- and IMERG-F v07-based simulations, both of which achieved a PTE of approximately 0 days and a $BIAS_v$ of approximately −38%, whereas Event 4 (dry antecedent soil) exhibited the strongest dependence on initial moisture conditions: ERA5-Land- and CHIRPS v2.0-based simulations captured only 13%–25% ($BIAS_v$ ~ 87%–75%) of the observed total runoff, while SPARTACUS v2.1- and IMERG-F v07-based models reproduced 75%–80%.

The event runoff coefficient ($RC_{event}$) results closely reflected the antecedent soil moisture and characteristics of each event (Tables 9 and S4). SPARTACUS v2.1 reproduced $RC_{event}$ values between 0.16 and 0.5, with deviations from observations ranging from −30% to +3%, providing the most



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

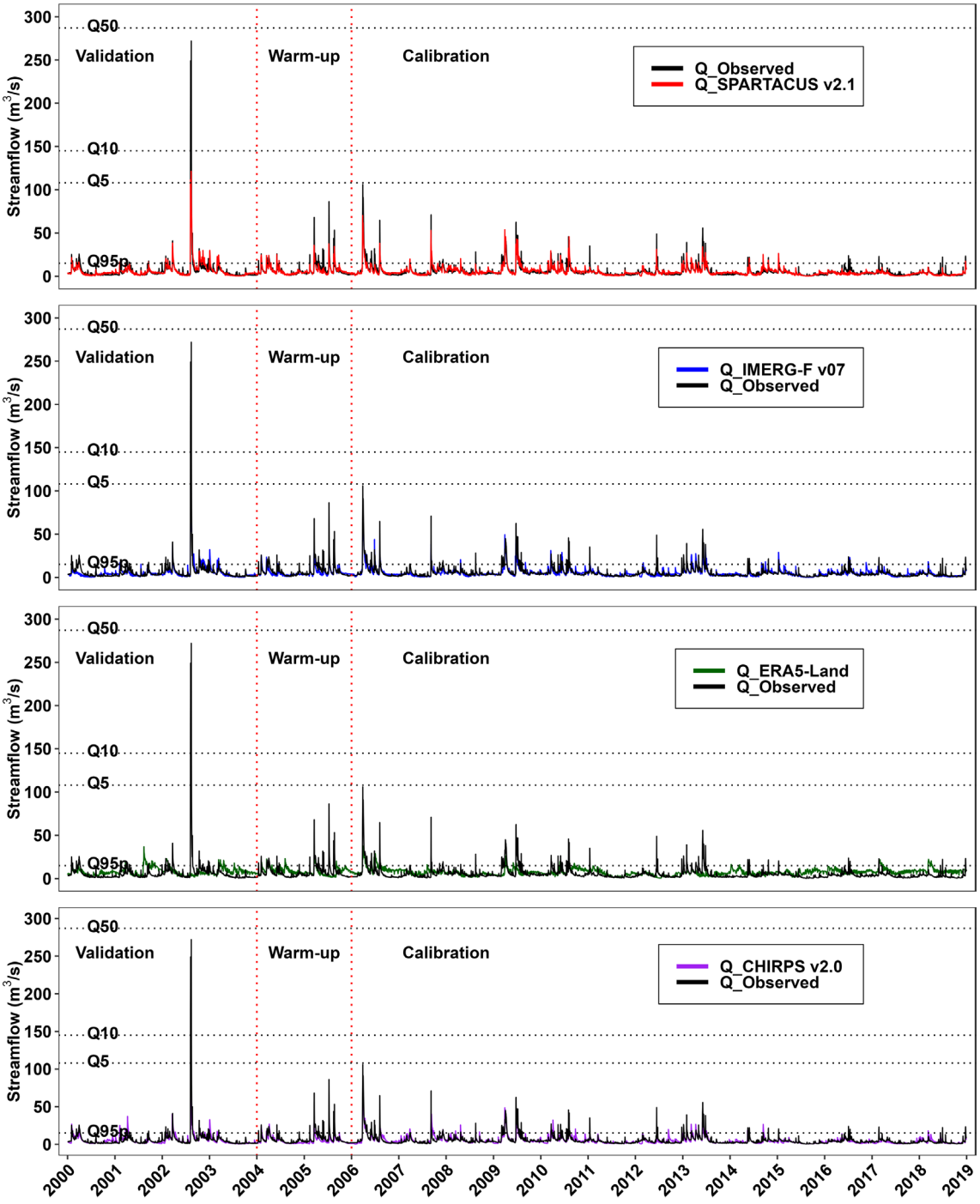


**FIGURE 7** | Comparison of the daily hydrographs of the observed and simulated streamflow at Zwettl Bahn Brücke gauge station in the calibration and validation periods using precipitation data from top to bottom SPARTACUS v2.1, IMERG-F v07, ERA5 Land, and CHIRPS v2.0. The simulated streamflow was based on the parameter sets that achieved the highest NSE values during calibration. Q_observed = observed streamflow; and Q_PP name = simulated streamflow of each PP; Q5,10,50 = 5-year, 10-year, 50-year return periods flood, respectively; Q95p = the streamflow value that is exceeded 95% of the time or very low flow (drought-prone).

consistent match across events. IMERG-F v07 showed similar behaviour ($RC_{event}$ = 0.2–0.53; $BIAS_{RC}$ = −27% to +10%), while CHIRPS v2.0 systematically underestimated runoff generation ($RC_{event}$ = 0.06–0.37; $BIAS_{RC}$ = −25% to −66%): ERA5-Land showed erratic behaviour, producing both extreme overestimation in Event 1 (+1505%) and strong underestimation (−59% to −61%) in Events 2 and 3, reflecting its poor representation of convective precipitation. The results confirm that both precipitation accuracy and pre-event soil moisture strongly control the modulated runoff generation in the hydrological model, as reported by Komma et al. (2007).

### 3.2.4 | Event-Scale Water Balance Partitioning During Extreme Floods

The analysis of the four selected major flood events in the Kamp catchment demonstrated that the event-scale water balance



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

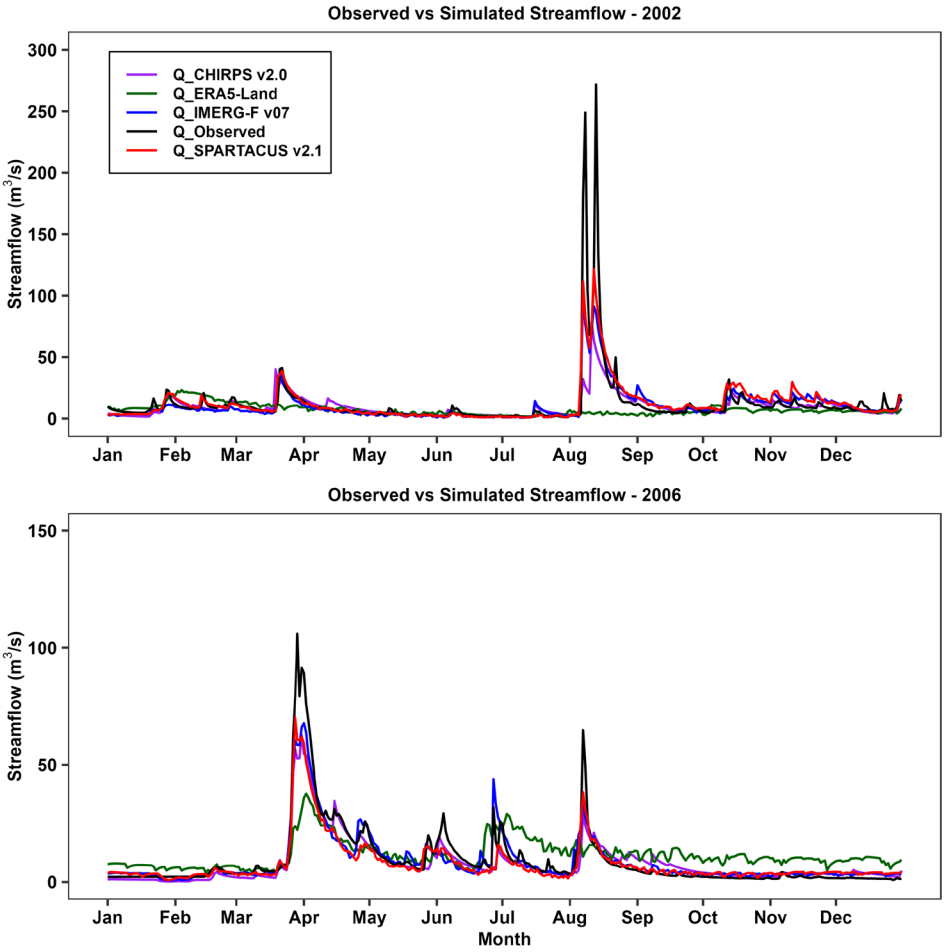


**FIGURE 8** | Comparison of the daily hydrographs of the observed and simulated streamflow at Zwettl Bahn Brücke streamflow station extracted for one year in the validation period (2002; top panel) and another year during the calibration period (2006; bottom panel).

(WB) simulated by SWAT+ was strongly affected by the precipitation forcing and antecedent moisture conditions (Figure 10). Across the four events, WB components varied substantially among PPs, except for evapotranspiration, where the model produced only minor differences.

The SPARTACUS v2.1- and IMERG-F v07-based model produced the most realistic partitioning compared to other products, with direct runoff of 25–55 mm and actual evapotranspiration of 5–25 mm, resulting in a total event water yield of 30–60 mm, which is relatively closer to the observed flood magnitudes. In contrast, ERA5-Land- and CHIRPS v2.0-based simulations produced smaller total direct runoff, less than 30 mm, and total water yield less than 40 mm, reflecting an underestimation of precipitation and weaker flood responses.

Antecedent soil moisture strongly modulates these dynamics. Under wet conditions (e.g., August 2002), precipitation was efficiently converted to quick streamflow, with SPARTACUS v2.1- and IMERG-F v07-based models yielding higher direct runoff (approximately 50–55 mm). Under dry conditions (e.g., August 2005), infiltration dominated (direct runoff $<20$ mm), enhancing subsurface percolation and attenuating peak flows.

Overall, high-resolution national datasets, such as SPARTACUS v2.1, produced a relatively more realistic WB partitioning and streamflow, whereas coarser products, like ERA5-Land and CHIRPS v2.0, shifted the hydrological balance toward percolation and ET losses, resulting in noticeable flow underestimation of all flood events.

### 3.2.5 | Effect of Precipitation Products Input Data on Long-Term Water Balance Components

Figure S3 presents a comparison of the long-term main water balance (WB) components derived from the SWAT+ model, calibrated using daily precipitation data from the selected products in the catchment, over the period 1998–2020. ERA5-Land showed the highest mean annual precipitation (866 mm) in the catchment, followed by IMERG-F v07 (745 mm) and CHIRPS v2.0 (729 mm), which overestimated it by 8% and 6%, respectively. SPARTACUS v2.1 slightly underestimated precipitation by ~1.4% (681 mm). These changes in precipitation magnitude significantly impacted surface runoff, water yield, percolation, and average soil water content. ERA5-Land and CHIRPS v2.0-based models yielded the highest values of surface runoff and water yield (runoff: 214 and 164 mm; water yield: 241 and 202 mm, respectively). In contrast, the SPARTACUS v2.1-based model produced the lowest surface runoff and water yield, followed by the IMERG-F v07-based model.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 7** | SWAT+ model performance metrics in the calibration and validation periods at Zwettl Bahn Brücke streamflow station (Lat. 48.608, Long. 15.162).

| | | Daily model | |
|---|---|---|---|
| **Precipitation products** | **Performance metric** | **Calibration (2006–2018)** | **Validation (2000–2003)** |
| SPARTACUS v2.1 | NSE | 0.75 | 0.64 |
| | KGE | 0.82 | 0.61 |
| | RMSE ($m^3/s$) | 3.23 | 7.81 |
| | Pbias % | 3.23 | 0.97 |
| IMERG Final v07 | NSE | 0.70 | 0.63 |
| | KGE | 0.76 | 0.61 |
| | RMSE ($m^3/s$) | 3.57 | 7.91 |
| | Pbias % | −1.99 | 0.91 |
| ERA5-Land | NSE | −0.14 | −0.15 |
| | KGE | −0.17 | −0.21 |
| | RMSE ($m^3/s$) | 12.82 | 14.01 |
| | Pbias % | −26.36 | −35.56 |
| CHIRPS v2.0 | NSE | 0.64 | 0.37 |
| | KGE | 0.73 | 0.37 |
| | RMSE ($m^3/s$) | 3.87 | 10.41 |
| | Pbias % | 2.72 | 4.12 |

Abbreviations: KGE, modified Kling-Gupta efficiency; NSE, Nash-Sutcliffe coefficient of efficiency; Pbias, Percent bias; RMSE, root mean square error.

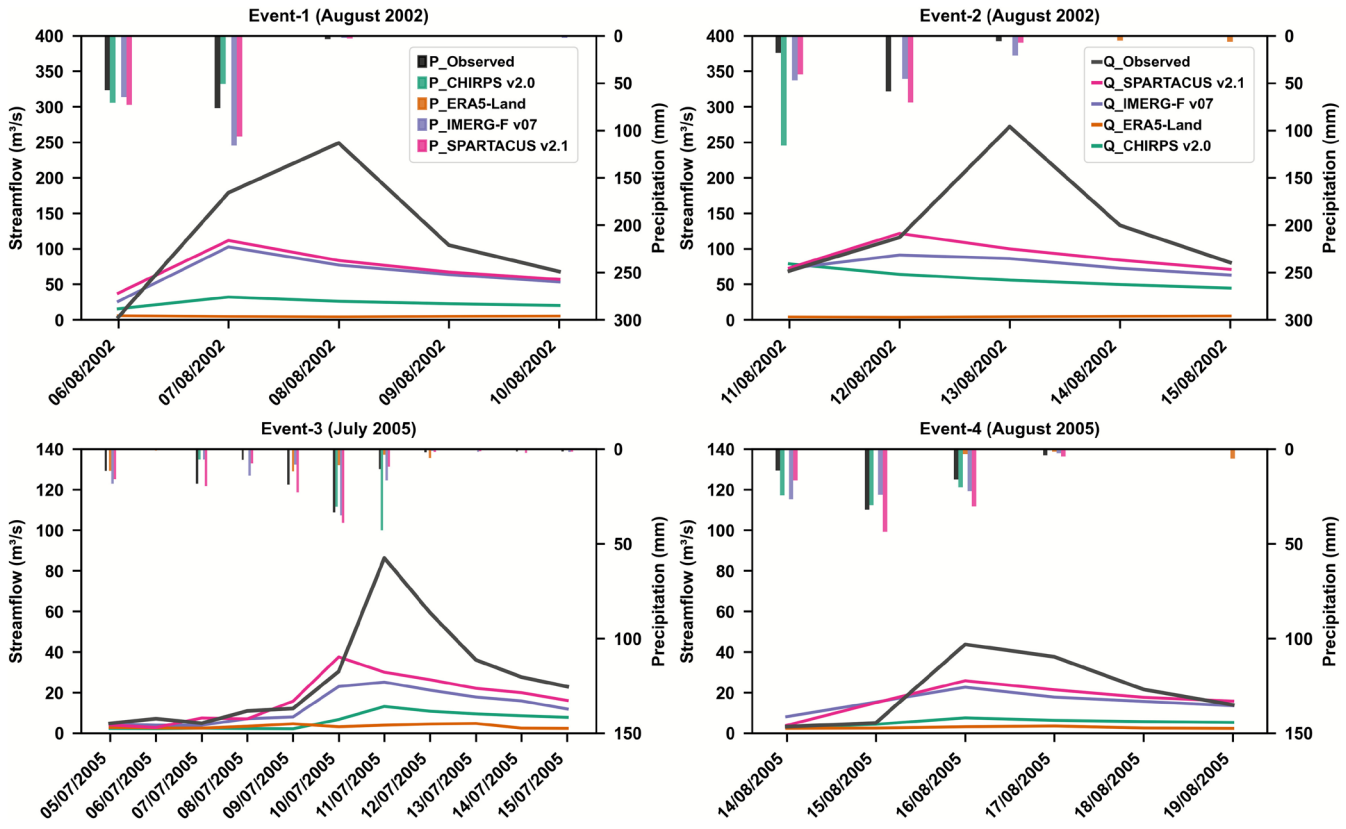


**FIGURE 9** | Comparison of the observed and simulated daily hydrographs of four historical flood events at the Zwettl Bahn Brücke streamflow station using the SPARTACUS v2.1, IMERG-F v07, ERA5 Land, and CHIRPS v2.0-based precipitation data inputs.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 8** | Performance of the simulated flood events at Zwettl Bahn Brücke streamflow station catchment (596.23 km$^2$) under different precipitation forcings.

| Flood event | Precipitation product | Precipitation (mm) | Runoff total volume (Mm$^3$) | RMSE (m$^3$/s) | PTE (day) | BIASp (%) | BIASv (%) | BIASpeak (%) |
|---|---|---|---|---|---|---|---|---|
| Event–1 (06/08/2002–10/08/2002) | Observed | 139.2 | 52.3 | — | — | — | — | — |
| | SPARTACUS V2.1 | 182.7 | 30.8 (21.5) | 83.2 | −1 | 31.2 | −41.0 | −55.2 |
| | ERA5-Land | 0.5 | 2.1 (50.2) | 144.5 | −2 | −99.7 | −95.9 | −97.7 |
| | IMERG-F v07 | 186.8 | 27.9 (24.4) | 87.0 | −1 | 34.2 | −46.7 | −58.8 |
| | CHIRPS v2.0 | 121.8 | 10.0 (42.3) | 127.0 | −1 | −12.5 | −80.8 | −87.1 |
| Event-2 (11/08/2002–15/08/2002) | Observed | 84.5 | 57.9 | — | — | — | — | — |
| | SPARTACUS V2.1 | 118.4 | 38.8 (19.1) | 80.2 | −1 | 40.1 | −33.0 | −55.3 |
| | ERA5-Land | 12.1 | 1.9 (56.0) | 148.7 | 2 | −85.6 | −96.7 | −98.0 |
| | IMERG-F v07 | 113.7 | 33.3 (24.6) | 88.5 | −1 | 34.5 | 42.5 | −66.5 |
| | CHIRPS v2.0 | 115.9 | 25.3 (32.6) | 107.5 | −2 | 37.1 | 56.3 | −71.1 |
| Event-3 (05/07/2005–15/07/2005) | Observed | 102.6 | 26.1 | — | — | — | — | — |
| | SPARTACUS V2.1 | 120.5 | 16.3 (9.8) | 20.6 | −1 | 17.5 | −37.6 | −56.5 |
| | ERA5-Land | 41.1 | 3.2 (22.9) | 34.0 | 2 | −59.9 | −87.9 | −94.5 |
| | IMERG-F v07 | 101.7 | 12.3 (13.8) | 23.1 | 0 | −0.9 | −37.6 | −70.9 |
| | CHIRPS v2.0 | 79.4 | 5.8 (20.3) | 29.8 | 0 | −22.6 | −77.7 | −84.8 |
| Event-4 (14/08/2005–19/08/2005) | Observed | 62.6 | 10.8 | — | — | — | — | — |
| | SPARTACUS V2.1 | 94.6 | 8.6 (2.2) | 10.8 | 0 | 51.1 | −20.5 | −41.0 |
| | ERA5-Land | 9.4 | 1.4 (9.4) | 23.5 | 1 | −85.0 | −87.1 | −92.1 |
| | IMERG-F v07 | 75.3 | 8.0 (2.8) | 12.9 | 0 | 20.2 | −25.6 | −48.2 |
| | CHIRPS v2.0 | 74.2 | 2.7 (8.1) | 20.9 | 0 | 18.6 | −74.5 | −82.8 |

*Note:* Values in the brackets of the "Runoff total volume" column of the table denote the deviation of the simulated runoff volume using each PP compared to the observed runoff volume of the catchment at Zwettl Bahn Brücke. Positive values indicate underestimation (runoff deficit), while negative values indicate overestimation (runoff excess).

**TABLE 9** | Runoff coefficients for different historical flood events in the Kamp catchment at Zwettl Bahn Brücke streamflow station.

| Precipitation dataset | Event-1 (06/08/2002–10/08/2002) | Event-2 (11/08/2002–15/08/2002) | Event-3 (05/07/2005–15/07/2005) | Event-4 (14/08/2005–19/08/2005) |
|---|---|---|---|---|
| SPARTACUS V2.1 | 0.27 (−30%) | 0.50 (3%) | 0.22 (−29%) | 0.16 (−15%) |
| ERA5-Land | 6.20 (1504%) | 0.20 (−59%) | 0.12 (−61%) | 0.18 (−1%) |
| IMERG-F v07 | 0.31 (−21%) | 0.53 (8%) | 0.20 (−27%) | 0.20 (10%) |
| CHIRPS v2.0 | 0.14 (−64%) | 0.37 (−25%) | 0.13 (−59%) | 0.06 (−66%) |
| Observed | 0.39 | 0.49 | 0.30 | 0.18 |

*Note:* Values in the brackets of the events column denote the relative error of the runoff coefficient compared to the calculated values based on the observed precipitation and direct runoff reported in Komma et al. (2007).

Notably, the highest mean annual percolation was observed in the ERA5-Land- and IMERG-F v07-based models' results (193 and 144 mm, respectively), while the SPARTACUS v2.1-based model produced the lowest percolation (107.21 mm). The CHIRPS v2.0- and SPARTACUS v2.1-based models presented the lowest and highest average soil water content, respectively



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(273 and 237 mm). Actual evapotranspiration (ET) values remain fairly constant across precipitation datasets, ranging from 416 mm (stations) to 433 mm (ERA5-Land-based model), suggesting that ET is less sensitive to precipitation forcings compared to other state variables.

## 4 | Discussion

### 4.1 | Performance of the Evaluated PPs

The results indicated that SPARTACUS v2.1 is the best-performing precipitation product in the study area. Conversely, CHIRPS v2.0 is the least accurate of the four products in the study area. IMERG-F v0.7 ranked second, followed by ERA5-Land. The superior performance of SPARTACUS v2.1 relative to other products is likely due to the use of Austrian gauge data, in addition to its high spatial resolution of 1 km (Hiebl and Frei 2018), which translates to better detection of the occurrence and magnitudes of precipitation events. Although this is expected, information on whether some of the ground stations used in this study for evaluation had already been used in the product's development was not available, and therefore, it was not possible to exclude them from our analysis. Typically, in data-scarce regions, obtaining an independent ground-based validation dataset is challenging as many products rely on the available stations for further corrections (Baez-Villanueva et al. 2018). The relatively good performance of IMERG-F v0.7 could be associated with the integration of multiple satellite sensors (e.g., GPM's GMI, TRMM's TMI), passive microwave sensors (e.g., AMSR-2, SSMIS, and infrared data from geostationary satellites), and recent improvements in the merging algorithm (Huffman et al. 2023, 2024).

SPARTACUS v2.1 shows a slight underestimation of CDD and an overestimation of CWD (Figure 5). Despite the good performance of SPARTACUS v2.1, representing extreme precipitation events accurately remains a challenge for many products (Li et al. 2021; Yu et al. 2021). Although SPARTACUS v2.1 exhibits strong detection capability for precipitation intensity below 1 mm/day, it has a relatively low accuracy in representing extreme daily precipitation (Figure 4). This can be due to increased spatial variability at higher precipitation intensities, which exacerbates representativeness errors in point-to-pixel evaluations, as the fundamental assumption that a point measurement accurately reflects the precipitation of the corresponding grid cell is often invalid under such conditions, as highlighted by Hiebl and Frei (2018) and Risser and Wehner (2020). Our results align with those of Hiebl and Frei (2018), who evaluated SPARTACUS v2.1 around the town of Feldbach in south-eastern Styria using experimental gridded rainfall data from the WegenerNet high-resolution observations (WegenerNet; Kirchengast et al. 2014) and reported that users comparing grid cells with point measurements must expect systematic overestimation for light precipitation and underestimation for heavy precipitation.

Several studies in diverse regions have reported correlation between IMERG estimates and ground-based precipitation observations (O et al. 2017; Anjum et al. 2018; Sharifi et al. 2019; Amjad et al. 2020; Zargar et al. 2025), which are in agreement with our findings. IMERG-F v07 shows a slight overestimation of daily precipitation and an underestimation of the maximum annual number of CDD, while it shows better ability in capturing the CWD. Additionally, it exhibits high detection skills for no rain, as shown by the POD performance metric, which is confirmed by other studies (i.e., Kim et al. 2017; Tan and Santo 2018; Mahmoud et al. 2021). However, IMERG-F v07 showed poor performance in detecting extreme precipitation, with high FAR values (Figure 4). Our results contrast with those of O et al. (2017), who reported a higher percentage of no rain and the limitation of IMERG products in detecting very low precipitation intensities when assessing the performance of IMERG v3 early, late, and final runs in the Feldbach region (Austria) using gauge-based gridded rainfall data from WegenerNet (WEGN). However, our results align with theirs in overestimating IMERG for high precipitation rates. Furthermore, several studies have assessed the performance of satellite-based IMERG across diverse climatic and topographic conditions in South America (Benítez et al. 2024; Rozante and Rozante 2024), highlighting its challenges in accurately estimating precipitation during winter. In Asia, satellite-based PPs such as GPM IMERG and PERSIANN have been extensively tested across diverse hydroclimatic settings (Liu et al. 2016; Tang et al. 2020; Khatakho et al. 2021). GPM IMERG outperformed other satellite and reanalysis precipitation products in China, while it underestimated snowfall compared with gauge and reanalysis data (Tang et al. 2020). The lower detection capability of IMERG-F v07 for extreme precipitation intensities may be due to the inclusion of gauge data at the monthly scale (i.e., from GPCC or CPC Unified Gauge-Based Analysis) to bias-correct its estimates (Huffman et al. 2019, 2023, 2024).

Compared to SPARTACUS v2.1 and IMERG-F v07, CHIRPS-v2.0 overestimated the CDD (also showing a high percentage of dry days) and underestimated the CWD, whereas ERA5-Land exhibited the opposite pattern. Moreover, both products are poor at detecting extreme precipitation. CHIRPS v.2.0 generally showed a low probability of detection for precipitation extremes, whereas it performed better in capturing dry days, as indicated by the POD results, with a slight overestimation. As this study was the first attempt to evaluate CHIRPS v2.0 in Austria, it was not possible to compare its performance with that of previous studies. However, similar studies carried out in neighbouring countries or other regions with similar climates and/or topography provide consistent insights (Hatzianastassiou et al. 2016; Zambrano-Bigiarini et al. 2017; Cavalcante et al. 2020; Nashwan et al. 2020; Khatakho et al. 2021). In East Africa, the satellite-based CHIRP and merged CHIRPS were shown to capture seasonal variability but required bias correction to improve reliability for agricultural and drought monitoring (Dinku et al. 2018). The underestimation of extreme precipitation events by CHIRPS v2.0 could be associated with the use of GPCC gauge data in the calibration of the product. GPCC has a limited local ground stations from eastern and central Europe (including Austria) compared to the tropics or Africa (Funk et al. 2015).

In general, ERA5-Land demonstrated better ability in capturing daily precipitation patterns compared to CHRIPS v2.0 despite its coarser spatial resolution (~10 km); however, it also exhibits biases, particularly in extreme precipitation events. It showed the worst detection ability for precipitation extremes (Figure 4). It



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

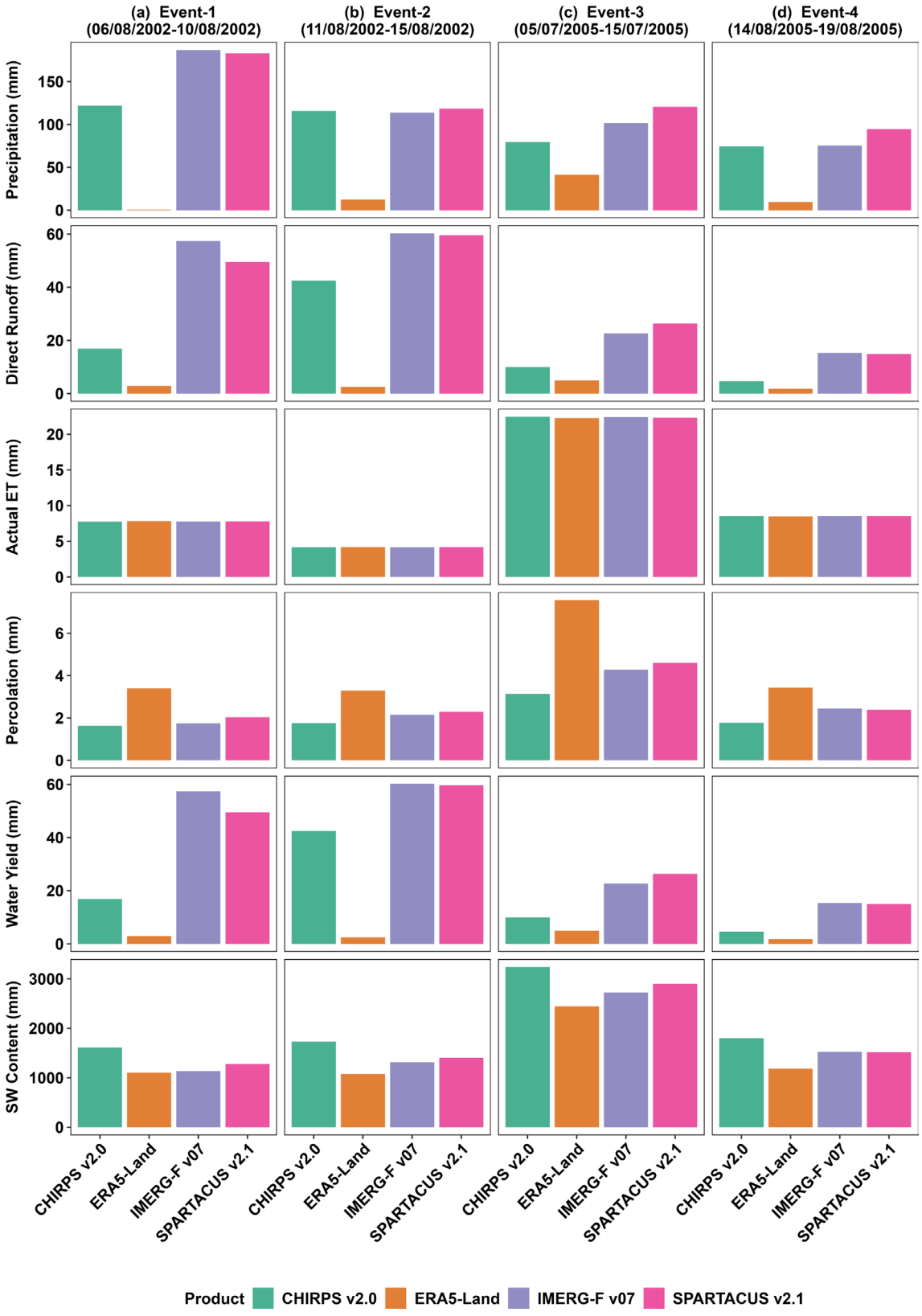


**FIGURE 10** | Events-based daily water balance components over the Kamp catchment at Zwettl Bahn Brücke streamflow station obtained by forcing the SWAT+ model with different PPs for (a) Event-1, (b) Event-2, (c) Event-3, (d) Event-4. The total precipitation for the event was added for comparison purposes.

also exhibits the highest percentage of wet days among all products and fewer extreme precipitation events than the ground stations, as shown in Table 5 and Figure 5. Furthermore, it exhibits the highest overestimation of CWD and underestimation of CDD in the study area. The aforementioned underestimation of extreme precipitation and the overestimation of wet days



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

by ERA5-Land are consistent with previous studies in other regions (e.g., Gomis-Cebolla et al. 2023; Guo et al. 2024; Moccia et al. 2025; Zargar et al. 2025). One possible explanation for this performance is the comparison of gauging station records with the closest grid-based data. The coarser resolution of the ERA5-Land data may have resulted in a mismatch between the ground station location and the corresponding pixel data.

The influence of elevation on precipitation patterns is a key factor that must be considered when applying precipitation products across Austria's varied terrain. We evaluated the impact of topographical gradient on the accuracy of the selected precipitation products in capturing extreme precipitation in the study area by dividing the catchment into three elevation zones (0–300 m, 300–700 m, 700–1119 m). Although all the products performed poorly in capturing extreme precipitation events as recorded by ground stations, they showed differences in performance across elevation zones. For example, SPARTACUS v2.1 performs better in mountainous regions (zone 700–1119 m) compared to lower elevations. In contrast, the performance of IMERG-F v07 in the lowland area (elevation 0–300 m) is relatively better than in higher elevations. Many studies highlight challenges in accurately representing the effects of topography on precipitation patterns and extreme precipitation estimation by SPARTACUS (Hiebl and Frei 2018), IMERG (Navarro et al. 2019; Sharifi et al. 2019), ERA5-Land (Abbas et al. 2024), and CHIRPS (Zambrano-Bigiarini et al. 2017). The discrepancies highlighted in these studies underscore the challenges of representing precipitation extremes in complex terrain using these datasets. The ranking of precipitation products (Table S5) indicates that SPARTACUS v2.1 performs best at capturing elevation effects, followed by IMERG-F v07. CHIRPS v2.0 shows better performance than ERA5-Land in the lowest elevation zone (0–300 m), but exhibits the worst performance across all higher elevation zones.

### 4.2 | Which Precipitation Product Performs Best in Simulating Daily Streamflow?

The model performance metrics indicate that SPARTACUS v2.1 has the best hydrological performance for both calibration and validation periods, followed by IMERG-F v07. Streamflow simulations using ERA5-Land exhibited weaker performance, with negative NSE and KGE values, underestimation of streamflow, and a poor representation of hydrograph recession curves. This can be attributed to biases in precipitation, especially in the detection of extreme events. Such biases are critical given the nonlinear nature of hydrological processes (Su et al. 2008; Khatakho et al. 2021).

Some of the most extreme floods in the catchment—most notably the August 2002 flood, which was triggered by an extended Vb-type cyclone rainfall event—as reported by Gutknecht et al. (2002), Mudelsee et al. (2004), and Viglione et al. (2010), could not be adequately captured by the precipitation products (see Figure 9 and Table 8). Floods like the August 2002 event were not only driven by rainfall intensity, but also by the accumulation of soil moisture over preceding days, leading to a nonlinear hydrological response once soils became saturated, as reported also by Komma et al. (2007). The underrepresentation of extreme rainfall by the PPs directly leads to a corresponding underestimation of simulated streamflow during such events, compromising the model's ability to accurately reproduce flood dynamics.

### 4.3 | Model Compensation and Parameter Sensitivity Under Different Precipitation Forcings

The event-scale diagnostics revealed that differences in precipitation magnitude and distribution affected the SWAT+ model calibration parameters. Under wet antecedent conditions (e.g., August 2002), the model generated rapid hydrological responses, using SPARTACUS v2.1 and IMERG-F v07 to attempt to reproduce observed peaks by adjusting the internal water balance primarily through enhanced surface and lateral flow contributions, as well as reduced infiltration and evapotranspiration losses. Conversely, under dry conditions (e.g., August 2005), ERA5-Land and CHIRPS v2.0 produced weaker flood responses dominated by infiltration and soil water storage, reflecting their lower precipitation inputs and the greater soil moisture deficit prior to the events.

Variation in model performance across different PPs underscores the critical importance of choosing an appropriate precipitation product for hydrological modelling.

The use of a biased or unsuitable PP can lead to misleading estimates of streamflow and other water balance components, particularly under extreme conditions; ultimately affecting the reliability of downstream impact assessments, such as flood risk analysis and water resource planning. Moreover, our results reveal that SWAT+ required substantially different parameter values with different PPs to align streamflow simulations with observations, consistent with the equifinality concept (Beven 2006), where different parameter sets may yield similar model results, yet for the wrong reasons.

To compensate for systematic biases in precipitation inputs, SWAT+ model calibration adjusts parameters controlling infiltration, baseflow, snowmelt, and evapotranspiration in an attempt to reproduce the observed streamflow as closely as possible (Table S6). This compensation mechanism, consistent with findings by Baez-Villanueva et al. (2021), allowed the model to reproduce, to some extent, observed streamflow despite biases in precipitation magnitude or timing, thereby forcing errors. For instance, during the analyzed flood events, under precipitation underestimation (ERA5-Land, CHIRPS v2.0), the model increased parameters controlling percolation (*perco*) and reduced the curve number (*cn2*), enhancing infiltration and groundwater contributions to sustain streamflow. Conversely, for PPs with positive precipitation bias (SPARTACUS v2.1, IMERG-F v07), calibration favored reduced soil water capacity (*awc*) and lower evapotranspiration efficiency (*esco, epco*) to limit water losses and maintain realistic runoff responses.

Parameter sensitivity rankings (see Table 6) show that snowmelt, soil, and runoff parameters were most responsive to precipitation characteristics. SPARTACUS v2.1 and IMERG-F v07 exhibited dominant sensitivity to *snomelt_tmp*, *revap_co*, *bd*, and



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

*surlag*, reflecting their influence on event-scale quick streamflow and delayed subsurface responses. In contrast, ERA5-Land and CHIRPS v2.0 showed high sensitivity to *bd, cn2*, *snomelt_tmp*, and *perco*, highlighting the model's reliance on infiltration and percolation pathways to offset precipitation deficiencies.

Overall, these results emphasise that the apparent performance of SWAT+ under different precipitation forcings can conceal substantial internal variability in the catchment's hydrological processes. The calibration process attempted to compensate for precipitation biases by tuning infiltration and evapotranspiration parameters, resulting in acceptable statistical fits to observations, even at the expense of compromised internal hydrological realism. This behaviour is illustrated in Figure 10 and Table S6, where the redistribution of runoff and other WB components across PPs demonstrates how calibration modified the event WB to offset precipitation biases. Similar behaviour is shown in the corresponding variation in the mean monthly water balance components (Figure S4). Across the four events, water balance components showed substantial variation among the precipitation products, except for actual evapotranspiration (AET), where the model simulated only minor differences. This limited variability in AET indicates that evapotranspiration was mainly controlled by available energy and soil moisture conditions rather than by differences in precipitation inputs, consistent with the soil moisture-buffering effect described by Seneviratne et al. (2010).

While SWAT+ is a robust, physically based model widely used for long-term water balance and landuse impact assessments, it may not be optimal for simulating short-duration, high-intensity events due to limitations in its rainfall-runoff response and channel routing schemes (Bieger et al. 2017). In addition, we acknowledge that the model of the Kamp catchment has uncertainties related to input data (low resolution of soil data, simplification of land use classes, etc.), the simplicity of the selected objective function (NSE) for model calibration, model parameterization, and the temporal extent of climate data used for calibration and validation. Thus, a model calibrated using a biased PP may appear to perform well but could fail to accurately simulate hydrological processes, especially extreme flows under different climatic or hydrological conditions, as demonstrated by the results. This limits the robustness of its predictions under current and future conditions. Addressing these uncertainties requires careful consideration of input data quality and its influence on parameter identifiability and model structural adequacy.

Combining statistical and hydrological evaluations enables a comprehensive understanding of the suitability of PPs for a specific purpose in a particular region. Although the findings of this study indicate very good performance for SPARTACUS v2.1, highlighting its potential to improve the reliability of hydrological simulations in the catchment, it still exhibits some bias in simulating peak flows. In order to enhance the reliability of hydrological simulations, bias correction of PPs is necessary.

### 4.4 | Future Work

Future work should examine the performance of PPs across different temporal and spatial scales. We recommend expanding the assessment to include more PPs and covering other areas with different characteristics within Austria. For example, the recently developed merged PP GMCP employs a unified merging and calibration strategy, resulting in a comprehensive global daily precipitation dataset with enhanced accuracy across diverse climates, particularly in regions with limited gauge observations (Ma et al. 2025). Also, satellite-based PPs such as PERSIANN-CDR were found to be a reliable dataset and an alternative information source for a limited gauge network, enabling the simulation of streamflow in the upper Yellow and Yangtze River basins on the Tibetan Plateau (Liu et al. 2016) and in the Himalayan River basin (Khatakho et al. 2021). Global-scale evaluations highlighted that the merged PP MSWEP consistently outperformed satellite-only and reanalysis-only products for hydrological simulations, especially in regions with dense gauge coverage (Beck et al. 2017). To strengthen the robustness of hydrological simulations, particularly under extreme flow conditions and during flood events, it would be valuable to test the performance of PPs using other hydrological models, specifically those designed to better capture high-flow dynamics, peak discharge, and flood wave propagation.

Given the sensitivity of streamflow simulation to the bias in the PPs, bias correction is often recommended to ensure adequate hydrological modelling results. While it is true that hydrological models can, to some extent, compensate for systematic biases in precipitation through calibration, this compensation has limits, especially when precipitation products poorly capture seasonal patterns or have low detection skill for precipitation events. Ultimately, hydrological models can redistribute and modulate available water but cannot compensate for errors in the timing or occurrence of precipitation. In future studies, incorporating local measurements and their associated uncertainties to constrain the parameter space, along with employing multi-objective parameter calibration approaches that simultaneously consider multiple performance metrics (e.g., streamflow magnitude, timing, and variability), could help enhance model robustness and partially mitigate the residual effects of input data biases.

## 5 | Summary and Conclusions

Evaluation of four PPs, namely, SPARTACUS v2.1, IMERG-F v07, CHIRPS v2.0, and ERA5-Land against 33 precipitation stations in the Kamp catchment has been carried out for the period 1998 to 2020. The evaluation employed a set of statistical and hydrological metrics to assess the accuracy of each product in capturing observed precipitation and streamflow characteristics. The key findings and their implications are as follows:

1. SPARTACUS v2.1 consistently outperformed the other products on a daily scale, followed by IMERG-F v07 and ERA5-Land. CHIRPS v2.0 exhibited the worst overall performance. This suggests that high-resolution national datasets, such as SPARTACUS, offer distinct advantages for daily-scale hydroclimatic analyses in Austria.
2. All products demonstrated reasonable accuracy in detecting dry days (precipitation $< 1$ mm/day). However, their ability



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to capture extreme (above the 95th percentile) and very extreme (above the 99th percentile) precipitation events was generally limited. These shortcomings might have critical implications for risk assessments and early warning systems targeting floods and other extreme weather events.

3. The performance of the products in capturing extreme precipitation events varied with topography. SPARTACUS v2.1 showed comparatively better accuracy in mountainous areas, while IMERG-F v07 performed better in lowland regions. These spatial discrepancies highlight the need for elevation-based bias correction or blending approaches to improve the utility of PPs in topographically diverse catchments. This is particularly important in catchments where runoff generation is sensitive to the timing of precipitation relative to existing wetness conditions, leading to nonlinear flood responses.
4. When the different datasets are used to drive the SWAT+ model for streamflow simulation, SPARTACUS v2.1 provided the most reliable streamflow estimates, followed by IMERG-F v07. CHIRPS v2.0 and ERA5-Land showed weaker performance in replicating observed discharge patterns.
5. The event-based analysis demonstrated that the choice of precipitation forcing exerts a dominant control on simulated high flow conditions and calibrated SWAT+ model parameters. SPARTACUS v2.1 and IMERG-F v07 provided relatively realistic representations of streamflow dynamics and water balance across wet and dry conditions, whereas the precipitation biases of ERA5-Land and CHIRPS v2.0 propagated into runoff underestimation or overestimation.
6. The SWAT+ model calibration was able to compensate, to some extent, for the differences between PPs inputs by adjusting the model parameters in an attempt to reproduce the observed streamflow as closely as possible and, therefore, adjusting other water balance components (e.g., baseflow contribution, soil water capacity, surface runoff coefficients, or evapotranspiration losses).

No single precipitation product is universally optimal. Our study confirms that various PPs exhibit different skills, even across different locations within the same catchment area, which is consistent with the findings previous studies. Validation specific to each catchment or region remains essential for selecting an appropriate precipitation dataset for hydrological use or other applications. In addition, the findings of this study highlight the need for rigorous evaluation of precipitation data prior to hydrological applications, and caution against overreliance on calibration performance alone when interpreting modelled hydrological processes. This is particularly important for evaluating hydroclimatic extremes, planning water resources, and developing climate adaptation strategies. Moreover, this study highlights the importance of improving the quality and spatial resolution of global and regional precipitation products, particularly in heterogeneous and mountainous regions where sparse observation networks and terrain-induced biases are prevalent.

We recognise that the outcomes of this validation cannot be directly transferred to other catchments; nonetheless, this study provides a robust methodology that can be utilised in diverse regions. By identifying the most reliable products for simulating precipitation and streamflow dynamics, the outcomes of this research will contribute to addressing the identified gaps, thereby improving the accuracy and reliability of hydroclimatic assessment studies and hydrological modelling in Austria, particularly in heterogeneous catchments like Kamp. This ultimately facilitates informed decision-making about effective water resource management and resilience to hydroclimatic extremes, such as floods and droughts.


## Acknowledgements

This research was conducted within the framework of the DISTENDER project (DevelopIng STratEgies by integrating mitigatioN, aDaptation, and participation to climate changE Risks), funded by the EU Horizon (ID 101056836). We are grateful to the providers of all the data used in this study. Open Access funding enabled and organized by Projekt DEAL.

## Funding

This work was supported by DISTENDER project. DISTENDER has received funding from the European Union's Horizon EU research and innovation programme under grant agreement No. 101056836.


## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

The hydrometeorological data were collected from the Hydrographic Service of Austria (eHYD available online at: https://ehyd.gv.at/?g_suche=208520) and GeoSphere Austria—Dataset (available online at: https://www.geo-sphere.at/dataset/). The DSM data were obtained from https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.112016.4326.2, soil from https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/, and land use from https://land.copernicus.eu/pan-european/corine-land-cover/clc2018. We obtained SPARTACUS v2.1 from https://data.hub.geosphere.at/dataset/spartacus-v2-1d-1km, CHIRPS v2.0 from https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_daily/netcdf/p05/, IMERG-F v07 from https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_07/summary?keywords=IMERG, and EAR5 Land from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview. The other raw data supporting the conclusions of this article will be made available by the authors upon request and authorization of the data providers.

## References


Abbas, H., W. Song, Y. Wang, et al. 2024. "Validation of CRU TS v4.08, ERA5-Land, IMERG v07B, and MSWEP v2.8 Precipitation Estimates Against Observed Values Over Pakistan." *Remote Sensing* 16, no. 24: 4803. https://doi.org/10.3390/rs16244803.

Abbaspour, K. C., E. Rouholahnejad, S. Vaghefi, R. Srinivasan, H. Yang, and B. Kløve. 2013. "A Continental-Scale Hydrology and Water Quality Model for Europe: Calibration and Uncertainty of High-Resolution Large-Scale SWAT Model." *Journal of Hydrology* 524: 733–752. https://doi.org/10.1016/j.jhydrol.2015.03.027.

Adler, R. F., M. R. P. Sapiano, G. J. Huffman, et al. 2018. "The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation." *Atmosphere* 9, no. 4: 138. https://doi.org/10.3390/ATMOS9040138.

Akoko, G., T. H. Le, T. Gomi, and T. Kato. 2021. "A Review of SWAT Model Application in Africa." *Water* 13, no. 9: 1313. https://doi.org/10.3390/w13091313.





10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Amjad, M., M. T. Yilmaz, I. Yucel, and K. K. Yilmaz. 2020. "Performance Evaluation of Satellite- and Model-Based Precipitation Products Over Varying Climate and Complex Topography." *Journal of Hydrology* 584: 124707. https://doi.org/10.1016/j.jhydrol.2020.124707.

Anjum, M. N., Y. Ding, D. Shangguan, et al. 2018. "Performance Evaluation of Latest Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) Over the Northern Highlands of Pakistan." *Atmospheric Research* 205: 134–146. https://doi.org/10.1016/j.atmosres.2018.02.020.

Arnold, J., K. Bieger, M. White, R. Srinivasan, J. Dunbar, and P. Allen. 2018. "Use of Decision Tables to Simulate Management in SWAT+." *Water* 10, no. 6: 713. https://doi.org/10.3390/w10060713.

Arnold, J. G., and N. Fohrer. 2005. "SWAT2000: Current Capabilities and Research Opportunities in Applied Watershed Modelling." *Hydrological Processes* 19, no. 3: 563–572. https://doi.org/10.1002/hyp.5611/.

Arnold, J. G., D. N. Moriasi, P. W. Gassman, et al. 2012. "SWAT: Model Use, Calibration, and Validation." *Transactions of the ASABE* 55, no. 4: 1491–1508. https://doi.org/10.13031/2013.42256.

Baez-Villanueva, O. M., M. Zambrano-Bigiarini, P. A. Mendoza, et al. 2021. "On the Selection of Precipitation Products for the Regionalisation of Hydrological Model Parameters." *Hydrology and Earth System Sciences* 25: 5805–5837. https://doi.org/10.5194/hess-25-5805-2021.

Baez-Villanueva, O. M., M. Zambrano-Bigiarini, L. Ribbe, A. Nauditt, J. D. Giraldo-Osorio, and N. X. Thinh. 2018. "Temporal and Spatial Evaluation of Satellite Rainfall Estimates Over Different Regions in Latin-America." *Atmospheric Research* 213: 34–50. https://doi.org/10.1016/j.atmosres.2018.05.01.

Bai, P., and X. Liu. 2018. "Evaluation of Five Satellite-Based Precipitation Products in Two Gauge-Scarce Basins on the Tibetan Plateau." *Remote Sensing* 10, no. 8: 1316. https://doi.org/10.3390/rs10081316.

Basheer, M., and N. A. Elagib. 2019. "Performance of Satellite-Based and GPCC 7.0 Rainfall Products in an Extremely Data-Scarce Country in the Nile Basin." *Atmospheric Research* 215: 128–140. https://doi.org/10.1016/j.atmosres.2018.08.028.

Basheer, M., R. Sulieman, and L. Ribbe. 2019. "Exploring Management Approaches for Water and Energy in the Data-Scarce Tekeze-Atbara Basin Under Hydrologic Uncertainty." *International Journal of Water Resources Development* 37, no. 2: 182–207. https://doi.org/10.1080/07900627.2019.1591941.

Beck, H., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, and E. F. Wood. 2018. "Present and Future Köppen-Geiger Climate Classification Maps at 1-Km Resolution." *Scientific Data* 5: 180214. https://doi.org/10.1038/sdata.2018.214.

Beck, H. E., N. Vergopolan, M. Pan, et al. 2017. "Global-Scale Evaluation of 22 Precipitation Datasets Using Gauge Observations and Hydrological Modeling." *Hydrology and Earth System Sciences* 21, no. 12: 6201–6217. https://doi.org/10.5194/hess-21-6201-2017.

Beck, H. E., E. F. Wood, M. Pan, et al. 2019. "MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment." *Bulletin of the American Meteorological Society* 100, no. 3: 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1.

Benítez, V. D., F. P. Forgioni, M. A. Lovino, L. Sgroi, M. E. Doyle, and G. F. Müller. 2024. "Capability of Satellite Data to Estimate Observed Precipitation in Southeastern South America." *International Journal of Climatology* 44, no. 3: 792–811. https://doi.org/10.1002/joc.8356.

Beven, K. 2006. "A Manifesto for the Equifinality Thesis." *Journal of Hydrology* 320, no. 1: 18–36. https://doi.org/10.1016/J.JHYDROL.2005.07.007.

Bieger, K., J. G. Arnold, H. Rathjens, et al. 2017. "Introduction to SWAT+, a Completely Revised Version of the Soil and Water Assessment Tool." *Journal of the American Water Resources Association* 53, no. 1: 115–130. https://doi.org/10.1111/1752-1688.12482.

Bollmeyer, C., J. D. Keller, C. Ohlwein, et al. 2015. "Towards a High-Resolution Regional Reanalysis for the European CORDEX Domain." *Quarterly Journal of the Royal Meteorological Society* 141, no. 686: 1–15. https://doi.org/10.1002/qj.2486.

Brigode, P., E. Paquet, P. Bernardara, et al. 2015. "Dependence of Model-Based Extreme Flood Estimation on the Calibration Period: Case Study of the Kamp River (Austria)." *Hydrological Sciences Journal* 60, no. 7–8: 1424–1437. https://doi.org/10.1080/02626667.2015.1006632.

Cavalcante, R. B. L., D. B. d. S. Ferreira, P. R. M. Pontes, R. G. Tedeschi, C. P. W. da Costa, and E. B. de Souza. 2020. "Evaluation of Extreme Rainfall Indices From CHIRPS Precipitation Estimates Over the Brazilian Amazonia." *Atmospheric Research* 238: 104879. https://doi.org/10.1016/j.atmosres.2020.104879.

Christoph, H., S. Bernhard, and H. Helmut. n.d. "Evaluation Functionality and Sustainability of River Widenings at the Kamp River/Austria Concerning Flood Protection and Aquatic Ecology Including A Numerical Sensitivity Test [Unpublished technical paper]."

Cornes, R. C., G. van der Schrier, E. J. M. van den Besselaar, and P. D. Jones. 2018. "An Ensemble Version of the e-OBS Temperature and Precipitation Data Sets." *Journal of Geophysical Research: Atmospheres* 123: 9391–9409. https://doi.org/10.1029/2017JD028200.

Dinku, T., C. Funk, C. Funk, et al. 2018. "Validation of the CHIRPS Satellite Rainfall Estimates Over Eastern Africa." *Quarterly Journal of the Royal Meteorological Society* 144: 292–312. https://doi.org/10.1002/QJ.3244.

Eingrüber, N., and W. Korres. 2022. "Climate Change Simulation and Trend Analysis of Extreme Precipitation and Floods in the Mesoscale Rur Catchment in Western Germany Until 2099 Using Statistical Downscaling Model (SDSM) and the Soil & Water Assessment Tool (SWAT Model)." *Science of the Total Environment* 830: 155775. https://doi.org/10.1016/j.scitotenv.2022.155775.

European Environment Agency (EEA). 2020. "CORINE Land Cover (CLC) 2018, Version 2020_20u1. Copernicus Land Monitoring Service." https://land.copernicus.eu/pan-european/corine-land-cover/clc2018.

Fang, J., W. Yang, Y. Luan, J. Du, A. Lin, and L. Zhao. 2019. "Evaluation of the TRMM 3B42 and GPM IMERG Products for Extreme Precipitation Analysis Over China." *Atmospheric Research* 223: 24–38. https://doi.org/10.1016/j.atmosres.2019.03.001.

FAO-UNESCO. 2003. *The Digital Soil Map of the World, Version 3.6*. Land and Water Development Division.

Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1 Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37, no. 12: 4302–4315. https://doi.org/10.1002/joc.5086.

Funk, C. C., P. J. Peterson, M. F. Landsfeld, et al. 2014. "A Quasi-Global Precipitation Time Series for Drought Monitoring. U.S. Geological Survey Data Series 832." *U.S. Geological Survey*: 4. https://doi.org/10.3133/ds832.

Funk, C., P. Peterson, M. Landsfeld, et al. 2015. "The Climate Hazards Infrared Precipitation With Stations—A New Environmental Record for Monitoring Extremes." *Scientific Data* 2: 150066. https://doi.org/10.1038/sdata.2015.66.

Gassman, P. W., M. R. Reyes, C. H. Green, and J. G. Arnold. 2007. "The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions." *Transactions of the ASABE* 50, no. 4: 1211–1250. https://doi.org/10.13031/2013.23637.

GeoSphere Austria. 2024. "GeoSphere Austria Datahub [Precipitation data]. GeoSphere Austria." https://data.hub.geosphere.at/.

Ghaemi, E., U. Foelsche, A. Kann, and J. Fuchsberger. 2021. "Evaluation of Integrated Nowcasting Through Comprehensive Analysis (INCA) Precipitation Analysis Using a Dense Rain-Gauge Network in Southeastern Austria." *Hydrology and Earth System Sciences* 25, no. 8: 4335–4356. https://doi.org/10.5194/hess-25-4335-2021.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Gomis-Cebolla, J., V. Rattayová, S. Salazar-Galán, and F. Francés. 2023. "Evaluation of ERA5 and ERA5-Land Reanalysis Precipitation Datasets Over Spain (1951–2020)." *Atmospheric Research* 284: 106606. https://doi.org/10.1016/j.atmosres.2023.106606.

Gumindoga, W., T. H. M. Rientjes, A. T. Haile, H. Makurira, and P. Reggiani. 2019. "Performance Evaluation of CMORPH Satellite Precipitation Product in the Zambezi Basin." *International Journal of Remote Sensing* 40, no. 20: 7730–7749. https://doi.org/10.1080/01431161.2019.1602791.

Guo, C., N. Ning, H. Guo, Y. Tian, A. Bao, and P. De Maeyer. 2024. "Does ERA5-Land Effectively Capture Extreme Precipitation in the Yellow River Basin?" *Atmosphere* 15, no. 10: 1254. https://doi.org/10.3390/atmos15101254.

Gutknecht, D., C. Reszler, and G. Blöschl. 2002. "Das Katastrophenhochwasser Vom 7. August 2002 Am Kamp – Eine Erste Einschätzung (The August 7, 2002 – Flood of the Kamp – A First Assessment)." *Elektrotechnik Und Informationstechnik* 119, no. 12: 411–413.

Haas, S., G. Kirchengast, and J. Fuchsberger. 2024. "Exploring Possible Climate Change Amplification of Warm-Season Precipitation Extremes in the Southeastern Alpine Forelands at Regional to Local Scales." *Journal of Hydrology: Regional Studies* 56: 101987. https://doi.org/10.1016/j.ejrh.2024.101987.

Haslinger, K., K. Breinl, L. Pavlin, et al. 2025. "Increasing Hourly Heavy Rainfall in Austria Reflected in Flood Changes." *Nature* 639: 667–672. https://doi.org/10.1038/s41586-025-08647-2.

Hatzianastassiou, N., C. D. Papadimas, C. J. Lolis, et al. 2016. "Spatial and Temporal Variability of Precipitation Over the Mediterranean Basin Based on 32-Year Satellite GPCP Data, Part I: Evaluation and Climatological Patterns." *International Journal of Climatology* 36, no. 15: 4741–4754. https://doi.org/10.1002/joc.4666.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. Jones, and M. New. 2008. "A European Daily High-Resolution Gridded Data Set of Surface Temperature and Precipitation for 1950-2006." *Journal of Geophysical Research* 113, no. 20. https://doi.org/10.1029/2008JD010201.

Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146: 1999–2049. https://doi.org/10.1002/qj.3803.

Hiebl, J., and C. Frei. 2018. "Daily Precipitation Grids for Austria Since 1961—Development and Evaluation of a Spatial Dataset for Hydroclimatic Monitoring and Modelling." *Theoretical and Applied Climatology* 132, no. 1: 327–345. https://doi.org/10.1007/s00704-017-2093-x.

Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao. 2004. "Precipitation Estimation From Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System." *Journal of Applied Meteorology* 43, no. 12: 1834–1853. https://doi.org/10.1175/JAM2173.1.

Huffman, G. J., D. T. Bolvin, R. Joyce, et al. 2023. "Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation." https://gpm.nasa.gov/sites/default/files/2023-07/IMERG_TechnicalDocumentation_final_230713.pdf.

Huffman, G. J., D. T. Bolvin, R. Joyce, et al. 2024. "IMERG V07 Release Notes." https://gpm.nasa.gov/sites/default/files/2024-12/IMERG_V07_ReleaseNotes_241126.pdf.

Huffman, G. J., E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan. 2019. "GPM IMERG Final Precipitation L3 Half Hourly 0.1 Degree × 0.1 Degree V06." Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed August 01, 2025. https://doi.org/10.5067/GPM/IMERG/3B-HH/07.

Hydrographic Service of Austria. 2022. "eHYD: Hydrological Data Portal of Austria [Data set]. Federal Ministry of Agriculture, Regions and Tourism (Austria)." Accessed November 08, 2022. http://ehyd.gv.at.

IPCC. 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, edited by C. B. Field, V. Barros, T. F. Stocker, et al. Cambridge University Press.

Isotta, F. A., R. Vogel, and C. Frei. 2015. "Evaluation of European Regional Reanalyses and Downscalings for Precipitation in the Alpine Region." *Meteorologische Zeitschrift* 24, no. 1: 15–37. https://doi.org/10.1127/metz/2014/0584.

Japan Aerospace Exploration Agency. 2021. "ALOS World 3D 30 Meter DEM. V3.2, Jan 2021. OpenTopography." Accessed November 08, 2022. https://doi.org/10.5069/G94M92HB.

Joyce, R., J. E. Janowiak, P. A. Arkin, and P. Xie. 2004. "CMORPH: A Method That Produces Global Precipitation Estimates From Passive Microwave and Infrared Data at High Spatial and Temporal Resolution." *Journal of Hydrometeorology* 5, no. 3: 487–503. https://doi.org/10.1175/1525-7541(2004)005<>2.0.CO;2.

Kann, A., I. Meirold-Mautner, F. Schmid, et al. 2015. "Evaluation of High-Resolution Precipitation Analyses Using a Dense Station Network." *Hydrology and Earth System Sciences* 19, no. 3: 1547–1559. https://doi.org/10.5194/hess-19-1547-2015.

Keikhosravi-Kiany, M. S., and R. C. Balling. 2024. "Evaluation of GPM IMERG Early, Late, and Final Run in Representing Extreme Rainfall Indices in Southwestern Iran." *Remote Sensing* 16: 2779. https://doi.org/10.3390/rs16152779.

Khatakho, R., A. Firoz, N. A. Elagib, and M. Fink. 2024. "Hydrological Modelling Using Gridded and Ground-Based Precipitation Datasets in Data-Scarce Mountainous Regions." *Hydrological Processes* 38, no. 12: e70024. https://doi.org/10.1002/hyp.70024.

Khatakho, R., R. Talchabhadel, and B. R. Thapa. 2021. "Evaluation of Different Precipitation Inputs on Streamflow Simulation in Himalayan River Basin." *Journal of Hydrology* 599: 126390. https://doi.org/10.1016/j.jhydrol.2021.126390.

Kidd, C., P. Bauer, J. Turk, et al. 2012. "Intercomparison of High-Resolution Precipitation Products Over Northwest Europe." *Journal of Hydrometeorology* 13, no. 1: 67–83. https://doi.org/10.1175/JHM-D-11-042.1.

Kim, K. Y., J.-M. Park, J. Baik, and M. Choi. 2017. "Evaluation of Topographical and Seasonal Feature Using GPM IMERG and TRMM 3B42 Over Far-East Asia." *Atmospheric Research* 187: 95–105. https://doi.org/10.1016/j.atmosres.2016.12.007.

Kirchengast, G., T. Kabas, A. Leuprecht, C. Bichler, and H. Truhetz. 2014. "WegenerNet: A Pioneering High-Resolution Network for Monitoring Weather and Climate." *Bulletin of the American Meteorological Society* 95, no. 2: 227–241. https://doi.org/10.1175/BAMS-D-11-00161.1.

Komma, J., C. Reszler, G. Blöschl, and T. Haiden. 2007. "Ensemble Prediction of Floods – Catchment Non-Linearity and Forecast Probabilities." *Natural Hazards and Earth System Sciences* 7, no. 4: 431–444. https://doi.org/10.5194/NHESS-7-431-2007.

Krysanova, V., and J. G. Arnold. 2008. "Advances in Ecohydrological Modelling With SWAT — A Review." *Hydrological Sciences Journal* 53, no. 5: 939–947. https://doi.org/10.1623/hysj.53.5.939.

Leitner, M., P. Babcicky, T. Schinko, and N. Glas. 2020. "The Status of Climate Risk Management in Austria: Assessing the Governance Landscape and Proposing Ways Forward for Comprehensively Managing Flood and Drought Risk." *Climate Risk Management* 30: 100246. https://doi.org/10.1016/j.crm.2020.100246.

Li, R., K. Wang, and D. Qi. 2021. "Event-Based Evaluation of the GPM Multisatellite Merged Precipitation Product From 2014 to 2018 Over China: Methods and Results." *Journal of Geophysical Research: Atmospheres* 126, no. 1. https://doi.org/10.1029/2020JD033692.

Liu, X., X. Liu, T. Yang, K. Hsu, C. Liu, and S. Sorooshian. 2016. "Evaluating the Streamflow Simulation Capability of PERSIANN-CDR Daily Rainfall Products in Two River Basins on the Tibetan Plateau."



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

*Hydrology and Earth System Sciences* 21, no. 1: 169–181. https://doi.org/10.5194/HESS-21-169-2017.

Lledó, L., T. Haiden, and M. Chevallier. 2024. "An Intercomparison of Four Gridded Precipitation Products Over Europe Using an Extension of the Three-Cornered-Hat Method." *Hydrology and Earth System Sciences* 28: 5149–5162. https://doi.org/10.5194/hess-28-5149-2024.

Ma, Z., Y. Chen, G. Tang, et al. 2025. "GMCP: A Fully Global Multisource Merging-And-Calibration Precipitation Dataset (1-Hourly, 0.1°, Global, 2000-The Present)." *Bulletin of the American Meteorological Society* 106, no. 2: E260–E275. https://doi.org/10.1175/BAMS-D-24-0051.1.

Mahmoud, M., S. A. Mohammed, M. A. Hamouda, and M. M. Mohamed. 2021. "Impact of Topography and Rainfall Intensity on the Accuracy of IMERG Precipitation Estimates in an Arid Region." *Remote Sensing* 13, no. 1: 13. https://doi.org/10.3390/RS13010013.

Mankin, K. R., S. Mehan, T. R. Green, and D. M. Barnard. 2025. "Review of Gridded Climate Products and Their Use in Hydrological Analyses Reveals Overlaps, Gaps, and the Need for a More Objective Approach to Selecting Model Forcing Datasets." *Hydrology and Earth System Sciences* 29: 85–112. https://doi.org/10.5194/hess-29-85-2025.

McNamara, I., O. M. Baez-Villanueva, O. M. Baez-Villanueva, et al. 2021. "How Well Do Gridded Precipitation and Actual Evapotranspiration Products Represent the Key Water Balance Components in the Nile Basin." *Journal of Hydrology: Regional Studies* 37: 100884. https://doi.org/10.1016/J.EJRH.2021.100884.

Moccia, B., L. Buonora, C. Bertini, E. Ridolfi, F. Russo, and F. Napolitano. 2025. "What Is Our Pick? Assessment of Satellite and Reanalysis Precipitation Datasets Over Italy." *Journal of Hydrology: Regional Studies* 60: 102487. https://doi.org/10.1016/j.ejrh.2025.102487.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. 2007. "Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations." *Transactions of the ASABE* 50, no. 3: 885–900. https://doi.org/10.13031/2013.23153.

Mudelsee, M., M. Mudelsee, M. Börngen, G. Tetzlaff, and U. Grünewald. 2004. "Extreme Floods in Central Europe Over the Past 500 Years: Role of Cyclone Pathway 'Zugstrasse Vb'." *Journal of Geophysical Research* 109. https://doi.org/10.1029/2004JD005034.

Muñoz-Sabater, J. 2019. "ERA5-Land Hourly Data From 1950 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)." https://doi.org/10.24381/cds.e2161bac.

Muñoz-Sabater, J., E. Dutra, A. Agustí-Panareda, et al. 2021. "ERA5-Land: A State-Of-The-Art Global Reanalysis Dataset for Land Applications." *Earth System Science Data* 13: 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting Through Conceptual Model. Part 1—A Discussion of Principles." *Journal of Hydrology* 10: 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nashwan, M. S., S. Shahid, A. Dewan, T. Ismail, and N. Alias. 2020. "Performance of Five High Resolution Satellite-Based Precipitation Products in Arid Region of Egypt: An Evaluation." *Atmospheric Research* 236: 104809. https://doi.org/10.1016/j.atmosres.2019.104809.

Navarro, A., E. García-Ortega, A. Merino, J. L. Sánchez, C. D. Kummerow, and F. J. Tapiador. 2019. "Assessment of IMERG Precipitation Estimates Over Europe." *Remote Sensing* 11, no. 21: 2470. https://doi.org/10.3390/rs11212470.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, J. R. Williams, and K. W. King. 2005. "SWAT Theoretical Documentation. Soil and Water Research Laboratory."

Neitsch, S. L., J. R. Williams, J. G. Arnold, and J. R. Kiniry. 2011. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*. Texas Water Resources Institute.

Nester, T., J. Komma, A. Viglione, and G. Blöschl. 2012. "Flood Forecast Errors and Ensemble Spread: A Case Study." *Water Resources Research* 48, no. 10. https://doi.org/10.1029/2011WR011649.

Nguyen, P., M. Ombadi, S. Sorooshian, et al. 2018. "The PERSIANN Family of Global Satellite Precipitation Data: A Review and Evaluation of Products." *Hydrology and Earth System Sciences* 22, no. 11: 5801–5816. https://doi.org/10.5194/hess-22-5801-2018.

Sungmin, O., 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." *Hydrology and Earth System Sciences* 21, no. 12: 6559–6572. https://doi.org/10.5194/hess-21-6559-2017.

Peßenteiner, S., C. Hohmann, G. Kirchengast, and W. Schöner. 2021. "High-Resolution Climate Datasets in Hydrological Impact Studies: Assessing Their Value in Alpine and Pre-Alpine Catchments in Southeastern Austria." *Journal of Hydrology: Regional Studies* 38: 100962. https://doi.org/10.1016/j.ejrh.2021.100962.

Probst, E. N., and W. Mauser. 2022. "Evaluation of ERA5 and WFDE5 Forcing Data for Hydrological Modelling and the Impact of Bias Correction With Regional Climatologies: A Case Study in the Danube River Basin." *Journal of Hydrology: Regional Studies* 40: 101023. https://doi.org/10.1016/j.ejrh.2022.101023.

Rauthe, M., H. Steiner, U. Riediger, A. Mazurkiewicz, and A. Gratzki. 2013. "A Central European Precipitation Climatology – Part I: Generation and Validation of a High-Resolution Gridded Daily Data Set (HYRAS)." *Meteorologische Zeitschrift* 22, no. 3: 235–256. https://doi.org/10.1127/0941-2948/2013/0436.

Risser, M. D., and M. Wehner. 2020. "The Effect of Geographic Sampling on Evaluation of Extreme Precipitation in High-Resolution Climate Models." *Atmospheric Science Letters* 6, no. 2: 115–139. https://doi.org/10.5194/ASCMO-6-115-2020.

Rozante, J. R., and G. Rozante. 2024. "IMERG V07B and V06B: A Comparative Study of Precipitation Estimates Across South America With a Detailed Evaluation of Brazilian Rainfall Patterns." *Remote Sensing* 16, no. 24: 4722. https://doi.org/10.3390/rs16244722.

Ruhí, A., A. Ruhí, M. L. Messager, and J. D. Olden. 2018. "Tracking the Pulse of the Earth's Fresh Waters." *Nature Sustainability* 1, no. 4: 198–203. https://doi.org/10.1038/S41893-018-0047-7.

Sadeghi, M., P. Nguyen, M. Rahnamay Naeini, K. Hsu, D. Braithwaite, and S. Sorooshian. 2021. "PERSIANN-CCS-CDR, a 3-Hourly 0.04° Global Precipitation Climate Data Record for Heavy Precipitation Studies." *Scientific Data* 8, no. 1: 157. https://doi.org/10.1038/S41597-021-00940-9.

Seneviratne, S. I., T. Corti, E. Davin, et al. 2010. "Investigating Soil Moisture-Climate Interactions in a Changing Climate: A Review." *Earth-Science Reviews* 99, no. 3: 125–161. https://doi.org/10.1016/J.EARSCIREV.2010.02.004.

Sharifi, E., J. Eitzinger, and W. Dorigo. 2019. "Performance of the State-Of-The-Art Gridded Precipitation Products Over Mountainous Terrain: A Regional Study Over Austria." *Remote Sensing* 11, no. 17: 2018. https://doi.org/10.3390/rs11172018.

Sharifi, E., B. Saghafian, and R. Steinacker. 2019. "Copula-Based Stochastic Uncertainty Analysis of Satellite Precipitation Products." *Journal of Hydrology* 570: 739–754. https://doi.org/10.1016/j.jhydrol.2019.01.034.

Sharifi, E., R. Steinacker, and B. Saghafian. 2018. "Multi Time-Scale Evaluation of High-Resolution Satellite-Based Precipitation Products Over Northeast of Austria." *Atmospheric Research* 206: 46–63. https://doi.org/10.1016/J.ATMOSRES.2018.02.020.

Shi, P., X. Bai, F. Kong, et al. 2017. "Urbanization and Air Quality as Major Drivers of Altered Spatiotemporal Patterns of Heavy Rainfall in China." *Landscape Ecology* 32, no. 8: 1723–1738. https://doi.org/10.1007/s10980-017-0538-3.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Skøien, J. O., and G. Blöschl. 2006. "Catchments as Space-Time Filters – A Joint Spatio-Temporal Geostatistical Analysis of Runoff and Precipitation." *Hydrology and Earth System Sciences* 10, no. 5: 645–662. https://doi.org/10.5194/HESS-10-645-2006.

Skok, G., N. Žagar, L. Honzak, R. Žabkar, J. Rakovec, and A. Ceglar. 2016. "Precipitation Intercomparison of a Set of Satellite- and Raingauge-Derived Datasets, ERA Interim Reanalysis, and a Single WRF Regional Climate Simulation Over Europe and the North Atlantic." *Theoretical and Applied Climatology* 123, no. 1: 217–232. https://doi.org/10.1007/s00704-014-1350-5.

Sobol', I. M. 2001. "Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates." *Mathematics and Computers in Simulation* 55, no. 1–3: 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6.

Su, F., Y. Hong, and D. P. Lettenmaier. 2008. "Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin." *Journal of Hydrometeorology* 9, no. 4: 622–640. https://doi.org/10.1175/2007JHM944.1.

Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K. Hsu. 2018. "A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons." *Reviews of Geophysics* 56, no. 1: 79–107. https://doi.org/10.1002/2017RG000574.

Tan, M. L., and H. Santo. 2018. "Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR Satellite Precipitation Products Over Malaysia." *Atmospheric Research* 202: 63–76. https://doi.org/10.1016/j.atmosres.2017.11.006.

Tang, G., M. P. Clark, S. M. Papalexiou, Z. Ma, and Y. Hong. 2020. "Have Satellite Precipitation Products Improved Over Last Two Decades? A Comprehensive Comparison of GPM IMERG With Nine Satellite and Reanalysis Datasets." *Remote Sensing of Environment* 240: 111697. https://doi.org/10.1016/j.rse.2020.111697.

Terblanche, D., A. Lynch, Z. Chen, and S. Sinclair. 2022. "ERA5-Derived Precipitation: Insights From Historical Rainfall Networks in Southern Africa." *Journal of Applied Meteorology and Climatology* 61, no. 10: 1473–1484. https://doi.org/10.1175/JAMC-D-21-0096.1.

Towner, J., H. Cloke, E. Zsoter, et al. 2019. "Assessing the Performance of Global Hydrological Models for Capturing Peak River Flows in the Amazon Basin." *Hydrology and Earth System Sciences* 23, no. 7: 3057–3080. https://doi.org/10.5194/HESS-23-3057-2019.

Ucal, M. 2017. "Economic Evaluation of Climate Change Impacts Development of a Cross-Sectoral Framework and Results for Austria." *Environment, Development and Sustainability* 20, no. 4: 1901–1903. https://doi.org/10.1007/S10668-017-9968-Y.

Viglione, A., G. B. Chirico, J. Komma, R. Woods, M. Borga, and G. Blöschl. 2010. "Quantifying Space-Time Dynamics of Flood Event Types." *Journal of Hydrology* 394, no. 1: 213–229. https://doi.org/10.1016/j.jhydrol.2010.05.041.

Xue, X., Y. Hong, A. S. Limaye, et al. 2013. "Statistical and Hydrological Evaluation of TRMM-Based Multi-Satellite Precipitation Analysis Over the Wangchu Basin of Bhutan: Are the Latest Satellite Precipitation Products 3B42V7 Ready for Use in Ungauged Basins?" *Journal of Hydrology* 449: 91–99.

Yu, L., G. Leng, A. Python, and J. Peng. 2021. "A Comprehensive Evaluation of Latest GPM IMERG V06 Early, Late and Final Precipitation Products Across China." *Remote Sensing* 13, no. 6: 1208. https://doi.org/10.3390/rs13061208.

Yuan, F., B. Wang, C. Shi, et al. 2018. "Evaluation of Hydrological Utility of IMERG Final Run V05 and TMPA 3B42V7 Satellite Precipitation Products in the Yellow River Source Region, China." *Journal of Hydrology* 567: 696–711. https://doi.org/10.1016/j.jhydrol.2018.06.045.

Zambrano-Bigiarini, M., A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe. 2017. "Temporal and Spatial Evaluation of Satellite-Based Rainfall Estimates Across the Complex Topographical and Climatic Gradients of Chile." *Hydrology and Earth System Sciences* 21, no. 2: 1295–1320. https://doi.org/10.5194/hess-21-1295-2017.

Zargar, M., A. Bronstert, T. Francke, et al. 2025. "Comparison and Hydrological Evaluation of Different Precipitation Data for a Large Tropical Region: The Blue Nile Basin in Ethiopia." *Frontiers in Water* 7: 1536881. https://doi.org/10.3389/frwa.2025.1536881.

Zhang, C., X. Chen, H. Shao, et al. 2018. "Evaluation and Intercomparison of High-Resolution Satellite Precipitation Estimates—GPM, TRMM, and CMORPH in the Tianshan Mountain Area." *Remote Sensing* 10, no. 10: 1543. https://doi.org/10.3390/rs10101543.

Zhang, M., Y. Chen, Y. Shen, and Y. Li. 2017. "Changes of Precipitation Extremes in Arid Central Asia." *Quaternary International* 436: 16–27. https://doi.org/10.1016/J.QUAINT.2016.12.024.

Zhang, X., L. V. Alexander, G. C. Hegerl, et al. 2011. "Indices for Monitoring Changes in Extremes Based on Daily Temperature and Precipitation Data." *Wiley Interdisciplinary Reviews: Climate Change* 2, no. 6: 851–870. https://doi.org/10.1002/WCC.147.

Zolina, O., A. Kapala, C. Simmer, and S. Gulev. 2004. "Analysis of Extreme Precipitation Over Europe From Different Reanalyses: A Comparative Assessment." *Global and Planetary Change* 44, no. 1: 129–161. https://doi.org/10.1016/j.gloplacha.2004.06.009.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Average monthly: (a) precipitation over the catchment and temperature at Gars am Kamp station, and (b) streamflow at Stiefern streamflow station. **Figure S2:** SWAT+ model inputs for the Kamp Catchment: (a) Delineated catchment map with derived streams, subbasins, outlets, and reservoirs' location; (b) Classified soil map; (c) Classified land use map; and (d) Classified slope map. **Figure S3:** Mean annual results of the water balance components over the Kamp catchment, obtained by forcing the SWAT+ model with different precipitation products. Mean annual precipitation was added for comparison purposes. **Figure S4:** Mean monthly water balance components over the Kamp catchment, obtained by forcing the SWAT+ model with different precipitation products. Mean monthly precipitation was added for comparison purposes. **Table S1:** Historical flood events in Kamp catchment measured at the Stiefern streamflow station, ~20 km upstream of the catchment outlet. **Table S2:** List of precipitation gauge stations used for PPs evaluation. **Table S3:** Summary of the continuous and categorical statistical performance metrics applied in evaluating PPs and/or the simulated streamflow. **Table S4:** Characteristics of flood events analysed in the Kamp catchment at Zwettl Bahn Brücke streamflow station. **Table S5:** Ranking (Composite performance score) of the evaluated precipitation products in the study area. **Table S6:** SWAT+ model parameters and their final fitted values after calibration using the ground stations and the four PPs as input. **Table S7:** USDA soil texture classification for the top soil layer of selected Digital Soil World Map (DSWM) units used in the model based on sand, silt, and clay fraction.



10991085, 2025, 12, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hyp.70359 by Bibl. der Universitat zu Koln, Wiley Online Library on [23/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License