water

MDPI

Article

Evaluation of Four Large-Scale and Model-Based River Discharge Datasets Against Gauged Observations in the Aras Basin, 2012–2025

Kebir Emre Saraçoğlu

Civil Engineering Department, Istanbul University-Cerrahpaşa, Istanbul 34320, Türkiye; kebir.saracoglu@iuc.edu.tr

Abstract

The Aras Basin, located in the northeastern part of Türkiye, has a sparse and insufficient river discharge observation network, mainly due to its mountainous terrain and transboundary basin characteristics. The lack of long-term and gap-free river discharge records in the Aras Basin has demonstrated the urgent need to evaluate alternative river discharge datasets for the basin. With this motivation, the performances of three global river discharge datasets, namely, the GloFAS v4.0 Hydrological Reanalysis, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service, and one regional river discharge dataset, namely, EFAS Historical v5.0, were evaluated for the Aras Basin in terms of (1) the monthly, seasonal, and water-year performance assessments, (2) performance assessment across low-, moderate-, and high-flow conditions based on flow duration curves, (3) the performance in detecting high-flow events, and finally (4) hydrological drought characterization using the Streamflow Drought Index (SDI). In these analyses, daily mean river discharges from six gauging stations operated by the General Directorate of State Hydraulic Works in the Aras Basin were used to cover the period from 1 October 2012 to 30 September 2025. The monthly, seasonal, water-year-based assessments indicated that GRADES-hydroDL generally performed better than the other river discharge products, with KGE values of 0.34–0.83 at four of the six stations, whereas the other products yielded negative KGE values at five stations. The model-based products exhibited performance differences in detecting low-, moderate-, and high-flow conditions based on flow duration curves. The GRADES-hydroDL dataset again demonstrates a better capability in detecting high-flow conditions, particularly at the D24A042-Hasankale Ç. Eğirmez gauging station. When it comes to high-flow event detection, GRADES-hydroDL is again the outstanding river discharge dataset, with CSI values of 0.41–0.63 at four stations compared with at most 0.37 for the other products. GRADES-hydroDL was able to identify two of the three main hydrological drought periods (2014–2015 and 2016–2018), whereas GloFAS v4.0 could predict two main dry periods (2014–2015 and 2021–2024) in the Aras Basin. However, only GRADES-hydroDL retained its skill when the SDI was standardized with the observed distribution. GRADES-hydroDL is therefore recommended for high-flow detection, drought monitoring, and monthly to seasonal water balance applications in the basin, whereas GloFAS v4.0 is suitable only for the timing of high flows and dry periods.

Keywords: model-based river discharge products; performance evaluation; flow duration curve; high-flow event detection; streamflow drought index (SDI); Aras Basin

Check for updates

Academic Editors: Giuseppe Oliveto and Bommanna Krishnappan

Received: 18 August 2026
Revised: 14 November 2026
Accepted: 25 November 2026
Published: 28 November 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

River discharge, which is one of the most important components of the hydrological cycle, is affected by climate change, basin characteristics, anthropogenic interventions, and so on [1]. River discharge records are used as a fundamental input in a wide range of research and implementation areas, from the assessment of climate change impacts to water resource planning and management, as well as flood and drought analyses [2–9]. River discharge is also essential for environmental flow management, which aims to provide the flow quantity and quality required to support aquatic species and their habitats [10]. River discharge records can be directly obtained from streamflow gauging stations. However, establishing and operating a dense streamflow monitoring network that can adequately represent the spatiotemporal variability of a basin requires substantial effort and financial funding. Therefore, global- or regional-scale model-based river discharge estimates and runoff estimates are an important alternative in regions where the gauging network is insufficient or sparse or unavailable [11]. The need for river discharge estimates is growing day by day, particularly in mountainous and transboundary basins where gauged observations are also spatiotemporally limited. In mountainous and transboundary basins, challenges of establishing and operating gauging networks and sharing of data due to geopolitical constraints escalate the need for alternative datasets [12].

Such alternative datasets of river discharge and runoff are commonly derived from global hydrological models (GHMs), land surface models (LSMs), reanalysis products, etc. The aim of producing such alternative global or regional datasets is to represent the long-term spatial and temporal behavior of hydrological variables in basins [12]. Recent advances in modeling studies, such as GHMs, LSMs, and reanalysis products enable these alternative datasets to be generated at higher resolutions. This has become highly accessible as publicly available, high-resolution river discharge and runoff datasets that can provide spatiotemporal representations of hydrological variability in river basins quite well. However, it is recommended that the applicability of these datasets should be tested by comparing them with gauged observations.

To this end, studies in the literature focusing on evaluating the performance of global river discharge and runoff products against gauged observations have increased in recent years. Lin et al. [7] introduced publicly available high-resolution river flow series called GRADES (Global Reach-Level A Priori Discharge Estimates for SWOT). GRADES was developed by VIC-based runoff simulations with the RAPID routing model across 2.94 million river reaches. The resulting 35-year daily and monthly river discharge estimates were evaluated using observations obtained from more than 14,000 gauging stations worldwide. GRADES-hydroDL, the advanced version of GRADES, is a global river discharge estimate that can produce daily river discharges. It was trained using grid long short-term memory over the MERIT-Basins river network in 4215 basins [13]. Compared with the process-based VIC model used in the original GRADES, the Grid LSTM approach improved the median KGE by about 0.15–0.22, with the largest gains in correlation reported for snow-dominated regions such as the Rocky Mountains and eastern Europe [13]. However, GRADES-hydroDL has no explicit process representation of snow accumulation and melt, reservoirs, or other human alterations, and its routing component was not tuned against discharge observations [13]. In addition, its training basins are concentrated in North America, Europe, and Australia, so that its skill in data-scarce regions such as eastern Türkiye, where no training gauges are available, has not been independently verified.

Sikder et al. [12] analyzed the runoff outputs from different land surface models by focusing on the two most populated Asian transboundary basins, the Ganges–Brahmaputra–Meghna (GBM) and the Mekong. In this case, runoff outputs produced by different land surface models, including CLM, Mosaic, Noah, and VIC within the GLDAS

framework, as well as the ECMWF ERA-Interim/Land product, were evaluated after being routed into river discharge using the RAPID model. Alfieri et al. [3] described the procedure in which the GloFAS-Reanalysis dataset v3.0 was developed and calibratedfor 1226 river sections using ERA5 meteorological data as input. The GloFAS-Reanalysis dataset v3.0 is a free dataset that covers daily discharges spanning 1980–2018. Liu et al. [14] evaluated multiple land surface products, including Noah-GL1, VIC-GL1, CLM-GL1, MOS-GL1, Noah-GL2.0, Noah-GL2.1, ERA-Interim/Land, PGFv2, and NRD-YRCC datasets, not only in terms of runoff performance but also in terms of water budget. The reason is that these datasets include not only runoff but also precipitation, ET, etc. It was stated that the datasets other than ERA-Interim/Land contained systematic biases in terms of runoff and ET. Koukoula et al. [15] evaluated 18 publicly available Global Water Resources Reanalysis products in terms of precipitation, river discharge, evapotranspiration, and terrestrial water storage in the Upper Blue Nile River Basin. In this evaluation, the results of the study showed that the river discharges, which were the outputs of the reanalysis products, were overestimated. Another study focused on the Upper Blue Nile Basin in Ethiopia; Lakew et al. [16] evaluated two different versions of water resources reanalysis runoff products. These versions, WRR1 and WRR2, with a spatial resolution of 0.5° (WRR1) and 0.25° (WRR2), were produced under the EU-funded eartH2Observe project, and evaluation indicates that the WRR2 (0.25°) UniK product shows better performance than WRR1 (0.5°). Miao and Wang [17] evaluated routed-runoff datasets from LSM/GHM by CaMa-Flood (VIC-CN05.1, CLM-CFSR, CLM-ERAI, CLM-MERRA, and CLM-NCEP) and three reanalysis datasets, namely, ERAI/Land, JRA55, and MERRA-2, against river discharges in major river basins in China spanning 1980–2008. The results show that all products except for MERRA-2 catch the interannual pattern quite well, whereas all products mainly underestimate monthly discharges.

Harrigan et al. [18] described the way to produce the GloFAS-ERA5 river discharge reanalysis dataset released as a part of GloFAS v2.1. GloFAS-ERA5 was generated by routing ERA5-based surface and sub-surface runoff through the LISFLOOD routing model. GloFAS-ERA5 provides daily global river discharge estimates at a spatial resolution of 0.1° from 1979 to near real time. Hales et al. [19] provide conceptual information on the development of the GEOGloWS ECMWF Streamflow Service model, which provides 15-day ensemble river discharge forecasts and river discharges from 1979 to the present. Chuphal and Mishra [11] developed a reconstructed monthly streamflow dataset coupling the H08 hydrological model and the mizuRoute routing model for Indian subcontinent river basins covering the years 1951–2021. The reconstructed streamflow dataset showed good performance against streamflow observations for 85 gauging stations. Ghimire et al. [20] introduced a new reconstructed daily and monthly streamflow dataset (Dayflow) by integrating VIC (Variable Infiltration Capacity) and RAPID (Routing Application for Parallel computatIon of Discharge) models for the period 1980–2015, and the new dataset was tested at 7526 US Geological Survey sites. Bain et al. [21] evaluated the performance of three different ECMWF runoff datasets, including ERA-Interim, ERA5 Reanalysis, and ERA5-Land, by routing with RAPID to generate daily and monthly discharges in the Mississippi River Basin. The results show that the model performance of each runoff product against gauged observations varies spatiotemporally. Tu et al. [2] aimed to evaluate the performances of five global hydrological models, CaMa-Flood, GloFAS-reanalysis, GRFR, PCR-GLOBWB, and WaterGAP, on a global scale. In this context, comparisons made between 1980 and 2014 showed that the datasets provided high correlation with the observations.

Swain et al. [22] attempted to reveal the distribution of hydrological drought by the Streamflow Drought Index (SDI) in India for the 1980–2020 period, using the outputs of GloFAS v3.1. Swain et al. [22] reach the conclusion that the SDI-3, SDI-6, SDI-9, and SDI-12 derived from observed discharges are in good agreement with those calculated from

GloFAS v3.1 in the Narmada River Basin, India. Kim et al. [23] developed a 9 km resolution river discharge dataset in the Ganges–Brahmaputra–Meghna river basins by routing ERA5-Land runoff through the VIC-River Routing Model for the period 1951–2023. The new dataset, which presents daily river discharges, was then evaluated against observations obtained from three gauging stations in Bangladesh. The results showed that the dataset exhibited a good correlation with the observations. Parmar and Karmakar [24] evaluated the performances of different runoff forcings such as ANU W3, ERA5 Reanalysis HTESSEL, JULES, LISFLOOD, ORCHIDEE, PCR-GLOBWB, and WaterGAP by generating multiple flood scenarios in terms of the flood hazard assessment framework using the CaMa-Flood model over India. It was concluded that the ERA5 Reanalysis dataset exhibited the best overall performance. Patidar et al. [25] evaluated the performance of GloFAS v3.1 in terms of daily and monthly scales during monsoon and non-monsoon periods. In addition, the Streamflow Drought Index (SDI-12) was considered in the hydrological drought assessments. Doğan [26] evaluated the performance of EFAS and GloFAS datasets in terms of performance metrics (NSE, KGE, and R^2) and hydrological indices against streamflow records with varying record lengths in Türkiye.

Previous studies in the literature have underlined the huge need for evaluation of global and regional model-based river discharge products and runoff datasets, particularly in regions where gauged observations are insufficient or sparse or unavailable. This study aims to present the performance of four different river discharge datasets (GloFAS v4.0 Hydrological Reanalysis as a part of the Global Flood Awareness System, EFAS Historical v5.0 as a part of the European Flood Awareness System, GRADES-hydroDL, and GEOGloWS ECMWF Streamflow Service) in the Aras Basin in Türkiye by comparing them against river discharge observations spanning 2012–2025, due to the lack of long-term and gap-free so-called continuous observations. The evaluation of the framework in this study includes (1) the monthly, seasonal, and water-year-based performance assessments; (2) performance assessment across low-, moderate-, and high-flow conditions based on flow duration curves; (3) the performance in detecting high-flow events; and (4) hydrological drought characterization using the Streamflow Drought Index (SDI).

Observation-based evaluations of GRADES-hydroDL and the GEOGloWS ECMWF Streamflow Service are still limited, particularly in terms of high-flow event detection and hydrological drought characterization. To address this gap, four river discharge products, namely, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and GEOGloWS ECMWF Streamflow Service, which are based on different modelling approaches, are evaluated together under identical conditions in a mountainous and transboundary basin in Türkiye, with a focus on these two aspects in addition to overall accuracy. The results are also used to guide the selection of suitable products for different applications in the study area.

2. Materials and Methods

2.1. Study Area

The study area, Aras Basin, is situated in the northeastern part of Türkiye, extending between 41°20′30″–42°36′36″ N and 41°10′12″–44°49′38″ E. The Aras Basin is a transboundary basin that shares borders with Georgia, Armenia, and Iran. The surface area of the basin is approximately 27,921.30 km^2 [27]. The reasons for focusing on the Aras Basin in this study are: (1) it is a mountainous and transboundary basin, (2) it does not have a sufficient streamflow gauging network, and (3) long-term and complete or continuous river discharge records with no missing values are not available. This indicates that long-term and gap-free discharge records are urgently needed to support future studies regarding

hydrology, climate change, water resource management, etc. Figure 1 illustrates the location of the Aras Basin and, in addition, the locations of the gauging stations within the basin.

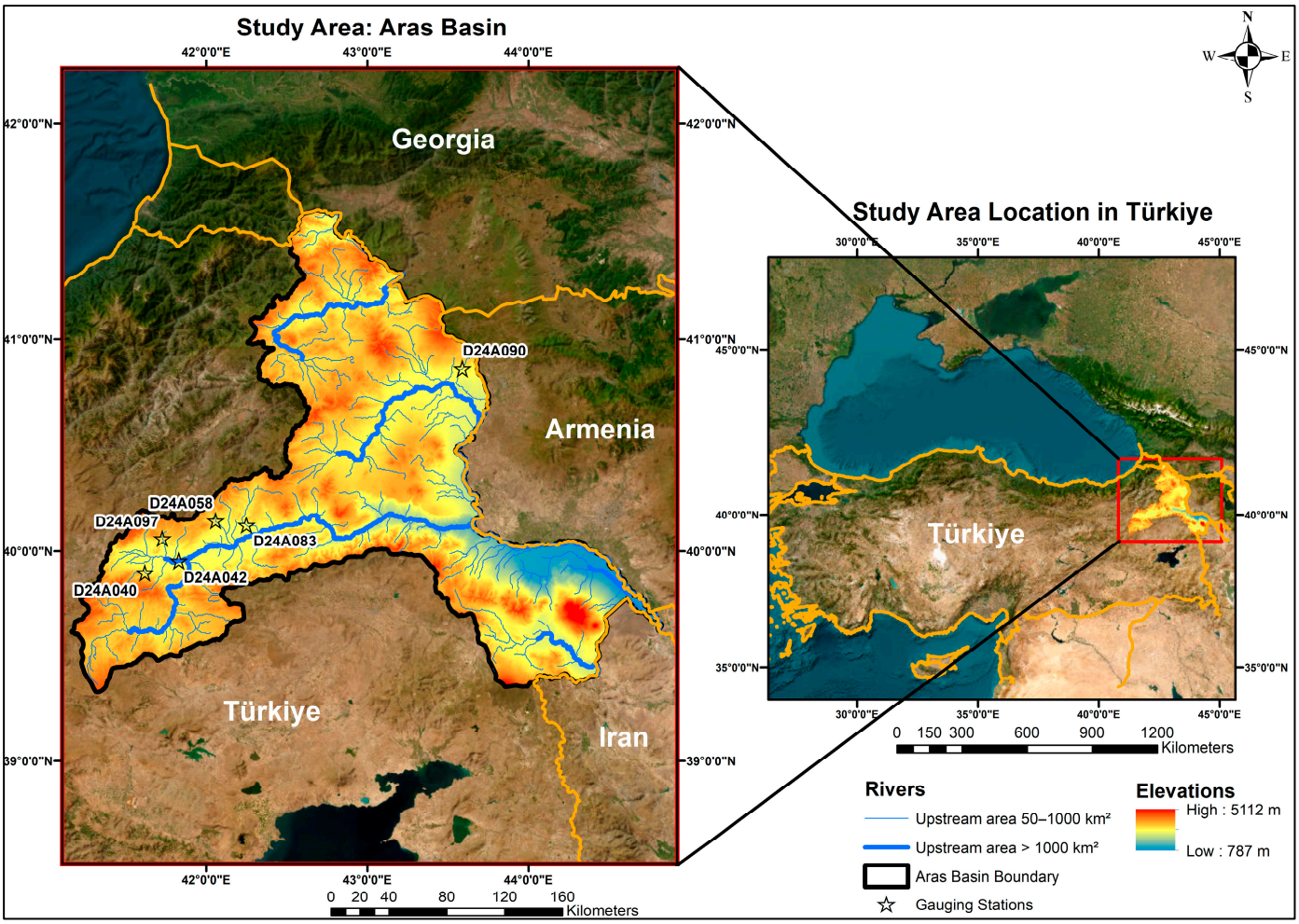


Figure 1. The location of the Aras Basin and gauging stations.

The spatial distribution of the selected gauging stations, which record gap-free observations spanning 2012–2025, indicates that gauging stations are mainly concentrated on the southwestern part of the basin. Five of the six stations (D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, D24A058-Kışla D. Çayırdüzü, D24A083-Süngütaş D. Çermik, and D24A097-Tımar Ç. Tımar) are located on the upper Aras sub-basin in the west and southwest, whereas the D24A090-Karahan Ç. Taşdere gauging station is located on the northeastern part of the Aras Basin. No gauging station with a gap-free record for the 2012–2025 period was available in the eastern part of the basin or in the transboundary downstream areas; therefore, the findings of this study apply to the upstream part of the basin within Türkiye. This uneven gauge network underlines the importance of global, regional, and model-based discharge data sets for the Aras Basin as well.

The Aras Basin has a continental climate: the mean temperature is below 0 °C in all winter months in the provinces of Kars and Ardahan, the mean annual precipitation ranges from 258.8 mm (Iğdır) to 550.8 mm (Ardahan), and heavy snowfall occurs in the upper parts of the basin [28]. Most of the land in the provinces of the basin remains covered with snow for 5–6 months [27], and the flow regime in the headwaters of the Aras River is dominated by snowmelt, with approximately 76.5% of the total annual runoff volume occurring during the spring and early-summer snowmelt season [29]. Consistent with this

regime, the observed discharges at all six gauging stations in this study reach their highest monthly mean values in April and May (see Section 3.1).

2.2. Datasets

In this study, the performance of four different global-regional and model-based datasets, including the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES hydroDL, and GEOGloWS ECMWF Streamflow Service, was evaluated. The reference dataset covers the period of 1 October 2012–30 September 2025 and consists of daily observed discharges obtained from the gauging stations operating by the General Directorate of State Hydraulic Works, while the global-regional and model-based datasets are the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service, respectively. The GloFAS v4.0 Hydrological Reanalysis is a model-based and daily river discharge product generated by the LISFLOOD model forcing ERA5 meteorological reanalysis products. In addition to providing daily river discharge estimates, the GloFAS v4.0 also includes hydrological variables such as soil wetness index, snow water equivalent, and runoff water equivalent. However, only river discharge outputs of the GloFAS v4.0 were used within the scope of this study.

The EFAS Historical v5.0 river discharge dataset is a part of the European Flood Awareness System (EFAS). The EFAS Historical dataset covers not only 6-hourly river discharge but also important hydrological variables, including volumetric soil moisture, snow water equivalent, soil wetness index, and runoff water equivalent. After extracting 6-hourly river discharge time series from the EFAS Historical NetCDF files, the 6-hourly time series were converted into daily time series in order to compare the evaluated datasets easily. Like GloFAS v4.0, EFAS Historical v5.0 was produced using the LISFLOOD model, forcing observational meteorological datasets such as temperature and precipitation.

GRADES-hydroDL, another global and model-based river discharge database, is the upgraded version of GRADES (Global Reach-level A priori Discharge Estimates for SWOT). Unlike GloFAS v4.0 and EFAS Historical v5.0, GRADES-hydroDL provides river discharge estimates for millions of river reaches all around the world, spanning 1980 to the present. GRADES-hydroDL was developed using a long short-term memory model instead of the VIC Land Surface Model with a resolution of 0.25° and routing by RAPID.

Lastly, the GEOGloWS ECMWF Streamflow Service is a global and model-based river discharge dataset, which was developed using Hydrologic Modeling as a Service (HMaaS) to produce a runoff dataset. In an effort to generate a river discharge dataset from a runoff dataset, the GEOGloWS ECMWF Streamflow Service routed it through the global river networks, forcing ERA-5 retrospective data. The river discharge outputs are presented as a web application.

Table 1 gives the information regarding observations obtained from gauging stations operated by the General Directorate of State Hydraulic Works, whereas Table 2 presents the details regarding the model-based datasets.

Table 1. Information regarding the gauging stations and observations.

Station No.	Station Name	Longitude (Degree)	Latitude (Degree)	Elevation (m)	Drainage Area (km^2)	Data	Observation Period
D24A040	Masat D. Bulkasım	41.618943	39.897861	1780	95.50	Daily Mean	October 2012–September 2025
D24A042	Hasankale Ç. Eğirmez	41.829242	39.954565	1597	1470	Daily Mean	October 2012–September 2025
D24A058	Kışla D. Çayırdüzü	42.0575	40.1475	1760	77.10	Daily Mean	October 2012–September 2025
D24A083	Süngütaş D. Çermik	42.248527	40.126688	1636	486.30	Daily Mean	October 2012–September 2025
D24A090	Karahan Ç. Taşdere	43.588979	40.860143	1745	152.50	Daily Mean	October 2012–September 2025
D24A097	Tımar Ç. Tımar	41.726548	40.061429	1801	97.52	Daily Mean	October 2012–September 2025

Table 2. Information regarding the model-based river discharge products.

	GloFAS v4.0 Hydrological Reanalysis	EFAS Historical v5.0	GRADES-hydroDL	GEOGloWS ECMWF Streamflow Service	Reference
Hydrological model	LISFLOOD	LISFLOOD	LSTM (Long Short-Term Memory)	ERA5-derived runoff (HTESSEL land-surface scheme)	[13,30–37]
Meteorological forcing	ERA5 Meteorological Reanalysis	Gridded Station Observations	MSWEP v2.8 precipitation; other meteorological fields from ERA5	ERA5 Reanalysis Runoff	[13,30,32,34,35,38]
Spatial resolution of Output	0.05° × 0.05°	1 × 1 arcmin	River Reach Level Output	River Reach Level Output	[13,30–35]
Snow scheme	Modification of Degree-day Method with elevation zones	Modification of Degree-day Method with elevation zones	No explicit process-based snow module (Data-driven LSTM)	Snow processes represented by HTESSEL within ERA5	[13,35,37,39,40]
Routing	Kinematic wave routing	Kinematic wave routing	RAPID routing model	Muskingum routing method	[13,34,35,41,42]
Calibration/training	Calibrated at 1995 gauges with a minimum drainage area of 500 km^2	Calibrated at 1903 gauges with a minimum drainage area of 150 km^2	Trained on daily discharge from 4215 gauged basins worldwide	No station-by-station calibration	[13,35,42–44]
Temporal coverage/step	1979–present, daily	1992–present, 6-hourly	1980–present, daily	1940–present, hourly	[13,30,32,34,35]

2.3. Statistical Evaluation Metrics

The performance of the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service was evaluated using statistical metrics, including the Root Mean Square Error (RMSE), Nash–Sutcliffe Efficiency (NSE), Pearson Correlation Coefficient (r), Kling–Gupta Efficiency (KGE), and finally Percent Bias (PBIAS). Equations given below (Equations (1)–(5)) present the formulations of the aforementioned statistical metrics. In addition, Table 3 provides ideal value and reference for each evaluation metric. In these equations, "Obs" represents the observational discharge, and "Mod" shows the river discharge of model-based products. On the other hand, $\overline{\text{Obs}}$ and $\overline{\text{Mod}}$ are the mean observed discharge and mean model-based discharge, respectively. The subscript sd in the KGE formula represents the standard deviation. In other words, Obs_{sd} and Mod_{sd} denote the standard deviations of the observed and model-based discharges, respectively. RMSE and PBIAS values approaching zero indicate a higher level of agreement between observed and model-based discharges. Higher NSE, r, and KGE values particularly close to one show quite good agreement between observed and model-based discharges.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\text{Obs}_i - \text{Mod}_i)^2}{n}} \tag{1}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(\text{Obs}_i - \text{Mod}_i)^2}{\sum_{i=1}^{n}\left(\text{Obs}_i - \overline{\text{Obs}}\right)^2} \tag{2}$$

$$r = \frac{\sum_{i=1}^{n}\left(\text{Obs}_i - \overline{\text{Obs}}\right)\left(\text{Mod}_i - \overline{\text{Mod}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\text{Obs}_i - \overline{\text{Obs}}\right)^2}\sqrt{\sum_{i=1}^{n}\left(\text{Mod}_i - \overline{\text{Mod}}\right)^2}} \tag{3}$$

$$\text{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\text{Mod}_{\text{sd}}}{\text{Obs}_{\text{sd}}} - 1\right)^2 + \left(\frac{\overline{\text{Mod}}}{\overline{\text{Obs}}} - 1\right)^2} \tag{4}$$

$$\text{PBIAS} = \frac{\sum_{i=1}^{n}(\text{Mod}_i - \text{Obs}_i)}{\sum_{i=1}^{n}(\text{Obs}_i)} \times 100 \tag{5}$$

Table 3. Information regarding the statistical evaluation metrics.

Statistical Metric	Abbreviation	Ideal Value	Reference
Root Mean Square Error	RMSE	0	[45]
Nash–Sutcliffe Efficiency	NSE	1	[46]
Pearson Correlation Coefficient	r	1	[47]
Kling–Gupta Efficiency	KGE	1	[48]
Percent Bias	PBIAS	0	[49]

2.4. Flow Duration Curve

The flow duration curve (FDC) is used to rank river discharges from high-flow to low-flow, giving a relationship between the probability of exceedance and river discharges. In this method, daily observed river discharges obtained from D24A040, D24A042, D24A058, D24A083, D24A090, and D24A097 stations operated by the General Directorate of State Hydraulic Works and daily discharges extracted from the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service were used to create flow duration curves for each station and dataset. In order to do that, the first step is to sort river discharges in descending order. In the second step, the highest river discharge is assigned to the first rank, and the lowest river discharge is

assigned to the last one. Then, the following formulation is used to calculate the probability of exceedance for each river discharge value.

$$P_i = \frac{m}{n+1} \times 100 \quad (6)$$

In Equation (6), the term P_i represents the probability of exceedance of river discharge i, m stands for the river discharge rank number, and n symbolizes the total number of data in the river discharge time-series. $m = 1$ denotes the largest river discharge value in the time series, while the last m value represents the lowest river discharge value in the same time series. The flow duration curve is represented graphically, and the vertical axis of this graph shows the river discharge values in m^3/s, while the horizontal axis shows the probability of exceedance values of the discharges. In this study, the performance of model-based discharge outputs was evaluated using flow duration curves under different flow regimes such as low-flow, median-flow, and high-flow regimes.

The high-flow conditions are categorized as Q5 (the river discharge corresponding to a 5% exceedance probability) and Q10 (the river discharge corresponding to a 10% exceedance probability), whereas the median-flow conditions are classified as Q25 (the river discharge corresponding to a 25% exceedance probability), Q50 (the river discharge corresponding to a 50% exceedance probability), and Q75 (the river discharge corresponding to a 75% exceedance probability). The low-flow conditions in terms of the flow duration curve are described as Q90 (the river discharge corresponding to a 90% exceedance probability) and Q95 (the river discharge corresponding to a 95% exceedance probability).

2.5. Performance Metrics of High-Flow Event Detection

One of the important aspects of evaluating the performance of model-based outputs, namely, river discharge outputs, is the detection of high-flow events. In this study, the Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI) were calculated in order to detect high-flow events using both ground observations and the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service [16,25]. These metrics are a reflection of the performance of model-based outputs in capturing flood events. Before calculating these metrics, a station-specific threshold is defined for each gauging station based on the observed daily river discharge. The station-specific threshold is defined as the 90th percentile of the observed daily river discharge data. When this threshold value is exceeded, the event is referred to as a high-flow event.

The POD, also known as the Hit Rate, indicates how well high-flow events can be predicted by hydrological models. The POD is calculated using the following equation (Equation (7)).

$$POD = \frac{H}{H+M} \quad (7)$$

In Equation (7), the term "H," the abbreviation of "Hit," indicates that the hydrological model is able to hit the high-flow events in the river discharge records. On the other hand, the term "M" in Equation (7), the abbreviation of "Miss," denotes that the high-flow events could not be detected by the model-based products accurately. The ideal value for the POD is 1.

Another categorical metric that can be used for a similar purpose is the False Alarm Ratio (FAR). FAR represents the hydrological model-detected high-flow events, which do not exist in the observed river discharge time series. The formulation used to calculate the FAR metric is presented below (Equation (8)).

$$FAR = \frac{F}{H+F} \quad (8)$$

Here, "F," the abbreviation of False Alarm, stands for the case where high-flow events are presented in the model-based outputs, although high-flow events do not exist in the observed time series. The ideal value for the FAR is 0.

The final metric is the CSI (Critical Success Index), which defines the proportion of high-flow events hit by the river discharge products to the total number of high-flow events in the observed time series. The following equation (Equation (9)) is used for this purpose.

$$CSI = \frac{H}{H + M + F} \tag{9}$$

The terms in Equation (9) correspond to "H" as "Hit," "M" as "Miss," and "F" as "False Alarm" as defined in the previous equations, Equations (7) and (8). The ideal value for the CSI is 1 as well.

2.6. Streamflow Drought Index (SDI)

Drought is a natural phenomenon which affects many areas such as human life, ecosystems, habitats, economic activities, etc. Drought has adverse effects not only on precipitation regimes but also on river discharges. Hydrological drought is one of the types of drought that refers to a decrease in the amount of water within the land phase of the hydrological cycle, especially in discharges, reservoirs, groundwater, etc. [50]. The Streamflow Drought Index (SDI) is one of the drought indices developed to evaluate hydrological drought conditions using river discharges. One of the methods to be used in evaluating the performance of model-based outputs (GloFASv4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service) is the ability to determine drought periods accurately using the SDI. In this drought approach, cumulative streamflow volumes ($V_{i,k}$) are initially estimated using monthly discharge volumes ($Q_{i,j}$). The cumulative streamflow volumes are derived using Equation (10). Here, $V_{i,k}$ is derived from the monthly discharge volumes in terms of the i^{th} hydrological year and the k^{th} reference period.

$$V_{i,\,k} = \sum_{j=1}^{3k} Q_{i,\,j} \qquad i = 1,\ 2,\ 3,\ \ldots \quad and\ j = 1,\ 2,\ 3,\ \ldots,\ 12\ \ and\ k = 1,\ 2,\ 3,\ 4 \tag{10}$$

In Equation (10), i denotes the hydrological year, j marks the month within the hydrological year ($j = 1$ represents October and $j = 12$ symbolizes September), and finally k represents the reference period ($k = 1$ as the October–December period, $k = 2$ as the October–March period, $k = 3$ as the October–June period, and $k = 4$ as the October-September period).

In the next step, the SDI is determined from the cumulative streamflow volumes. In order to do that, Equation (11) will be used.

$$SDI_{i,\,k} = \frac{V_{i,\,k} - \overline{V}_k}{S_k} \qquad i = 1,\ 2,\ 3,\ \ldots \quad and\ k = 1,\ 2,\ 3,\ 4 \tag{11}$$

Here, $\overline{V}_k$ and S_k are the mean cumulative streamflow volume and the standard deviation of the cumulative streamflow volume, respectively.

Based on the SDI values, drought conditions can be classified, ranging from non-drought to extreme drought, in terms of severity levels from 1 to 5. Table 4 gives information regarding the classification of hydrological drought conditions in detail.

Table 4. Hydrological drought conditions in terms of the Streamflow Drought Index [50].

Condition No.	Drought Condition	SDI
1	Non-drought	SDI ≥ 0
2	Mild Drought	$-1 \leq$ SDI < 0
3	Moderate Drought	$-1.5 \leq$ SDI < -1
4	Severe Drought	$-2 \leq$ SDI < -1.5
5	Extreme Drought	SDI < -2

3. Results

3.1. Evaluation of Monthly, Seasonal, and Water-Year-Based Results

Daily mean river discharge time series from six gauging stations operated by the General Directorate of State Hydraulic Works in the Aras Basin were used for the analyses presented in this section. In addition, daily mean river discharge and six-hourly mean river discharge time series were extracted from the GloFAS v4.0 Hydrological Reanalysis and the EFAS Historical v5.0, respectively. Then, the six-hourly mean discharge time series extracted from the EFAS Historical v5.0 were converted into daily mean time series to allow direct comparison. Regarding the GRADES-hydroDL and GEOGloWS ECMWF Streamflow Service datasets, the time series of these datasets are provided for the river reaches where the stations are located. The GRADES-hydroDL and GEOGloWS ECMWF Streamflow Service datasets also produce daily mean river discharge estimates. The water-year-based analyses cover the period from 1 October 2012 to 30 September 2025. The water-year-based time series of the D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, D24A058-Kışla D. Çayırdüzü, D24A083-Süngütaş D. Çermik, D24A090-Karahan Ç. Taşdere, and D24A097-Tımar Ç. Tımar gauging stations are presented in Figures 2–7, respectively. In each figure, the comparisons of (a) GloFAS v4.0 Hydrological Reanalysis vs. Observed River Discharge, (b) EFAS Historical v5.0 vs. Observed River Discharge, (c) GRADES-hydroDL vs. Observed River Discharge, and (d) GEOGloWS ECMWF Streamflow Service vs. Observed River Discharge are presented.

In addition to the time series plots presented in Figures 2–7, statistical evaluation metrics were calculated for each gauging station and each model-based streamflow product to quantify their performance. The results are presented as a heatmap in Figure 8. Thus, the performance of the model-based river discharge products was compared in terms of both their time series pattern and statistical evaluation metrics.

Regarding D24A040-Masat D. Bulkasım (see Figure 2), it can be said that the hydrological model showing the best overall performance is GRADES-hydroDL. The EFAS Historical v5.0 showed the weakest performance among the evaluated model-based discharge products. The EFAS Historical v5.0 dataset provides very low river discharge estimates, almost close to zero. The GloFAS v4.0 Hydrological Reanalysis stands out for its ability to capture peak river discharges at D24A040-Masat D. Bulkasım. In Figure 3, the GloFAS v4.0 Hydrological Reanalysis again captured the peak discharges much better than the other products. The EFAS Historical v5.0 also predicted similar river discharges, as in the previous station D24A040. The GloFAS v4.0 Hydrological Reanalysis and GRADES-hydroDL stand out among the evaluated model-based river discharge products in capturing peak discharges at D24A042-Hasankale Ç. Eğirmez. Figure 4 shows the river discharge time series for D24A058-Kışla D. Çayırdüzü. GRADES-hydroDL clearly follows the general pattern of the observed river discharges. However, a similar performance for GRADES-hydroDL is not observed in terms of peak discharge behavior. The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 estimated considerably low river discharges. The GEOGloWS ECMWF Streamflow Service, on the other hand, exhibits a highly responsive behavior to

river discharge fluctuations. In the plots presented in Figure 5 for D24A083-Süngütaş D. Çermik, the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 provide a clear improvement in their river discharge estimates. However, these two hydrological models produced discharges during some periods, which were not observed in either the observed discharge records or the other model-based discharge products. Again, it can be said that GRADES-hydroDL exhibits a superior performance to other model-based products at the D24A083 station. It is observed that all model-based datasets exhibit a limited capability to capture peak flows. It should be noted that the RMSE values are quite high for all model-based datasets at the D24A083. Figure 6 presents both the observed river discharge records and model-based river discharge outputs for D24A090-Karahan Ç. Taşdere. The river discharge records show lower discharges during the period of March–April–May 2013 compared with the river discharges observed at the other gauging stations. This resulted in evident improvement in the performance of the GloFAS v4.0 Hydrological Reanalysis dataset. In fact, on certain dates, the GloFAS v4.0 Hydrological Reanalysis predicts higher river discharges than the observations. The other model-based products that stand out in terms of performance are GRADES-hydroDL and the GEOGloWS ECMWF Streamflow Service for D24A090-Karahan Ç. Taşdere. Unfortunately, all model-based datasets are not sufficient in capturing peak discharges at this gauging station. For the D24A097-Tımar Ç. Tımar gauging station, a similar pattern was also observed (see Figure 7). For this station, one of the datasets that determined river discharge values close to the discharge observation is again GRADES-hydroDL. The GloFAS v4.0 Hydrological Reanalysis, on the other hand, overestimates the peak discharges at D24A097 in most water years. However, its performance in estimating the low-flow conditions is relatively poor. EFAS Historical v5.0 maintains its consistency in predicting low river discharge conditions, meaning that EFAS Historical v5.0 underestimates river discharges again. In the water-year-based analyses (covering the period from 1 October 2012 to 30 September 2025), it can be easily said that the model-based product that clearly stands out is GRADES-hydroDL once again. However, it is also clear that its performance in estimating the peak discharge values remained limited.

The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 are provided in NetCDF format. Station-specific river discharge data were extracted from the grid cell where the stations are located. Just in case, the river discharge time series in the neighboring grid cells were also extracted again for the stations where the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 provided considerably low discharge conditions. For example, for the EFAS Historical v5.0 dataset, the distances of the D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, and D24A058-Kışla D. Çayırdüzü stations to the nearest neighboring grid cells are 195 m, 331 m, and 278 m, respectively. The neighboring grid cells reduced the volume bias of EFAS Historical v5.0 at these stations, but its KGE values remained negative; therefore, the EFAS Historical v5.0 series were not changed. For GloFAS v4.0, the grid cell located on the southwest of the D24A097-Tımar Ç. Tımar gauging station was used instead of the station grid cell. This situation is thought to be related to the upstream drainage area of the stations. This issue is examined in detail in the discussion section.

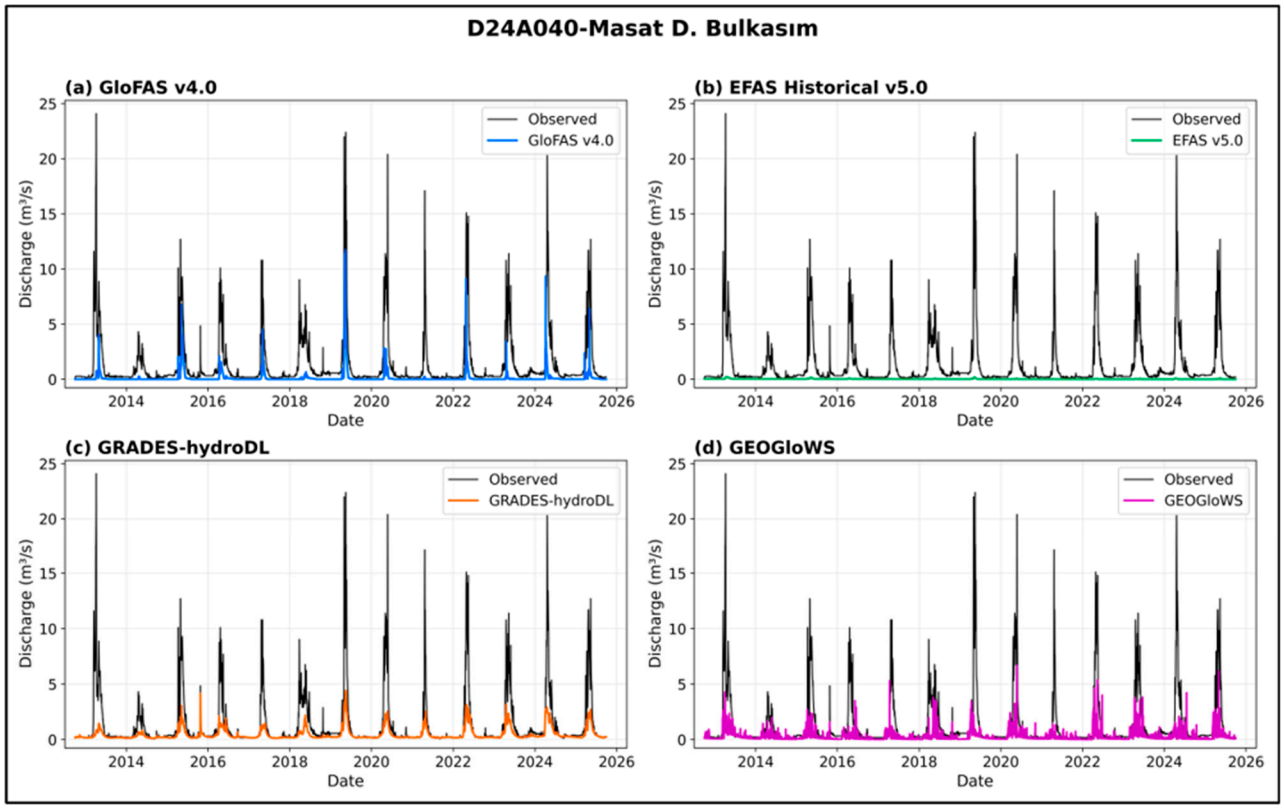


Figure 2. The water-year-based time series comparisons for the D24A040-Masat D. Bulkasım gauging station covering the period from 1 October 2012 to 30 September 2025.

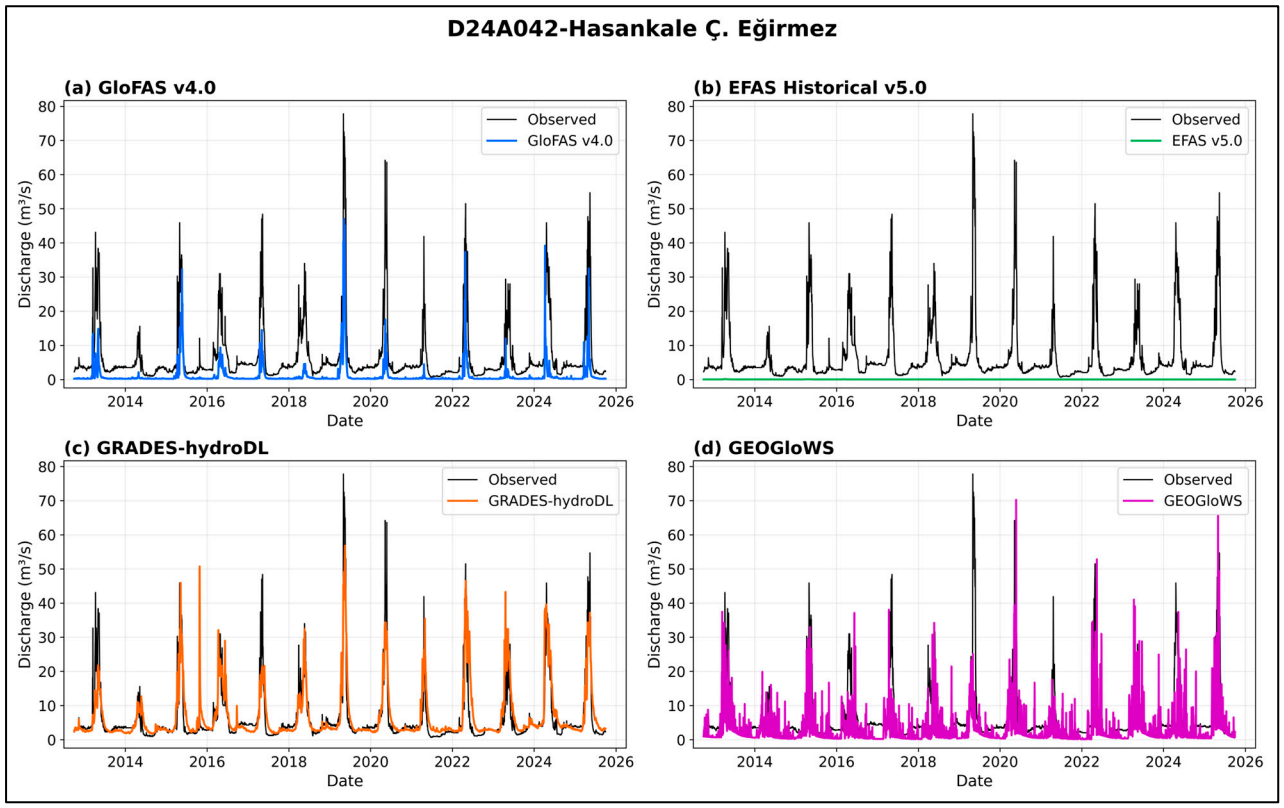


Figure 3. The water-year-based time series comparisons for the D24A042-Hasankale Ç. Eğirmez gauging station covering the period from 1 October 2012 to 30 September 2025.

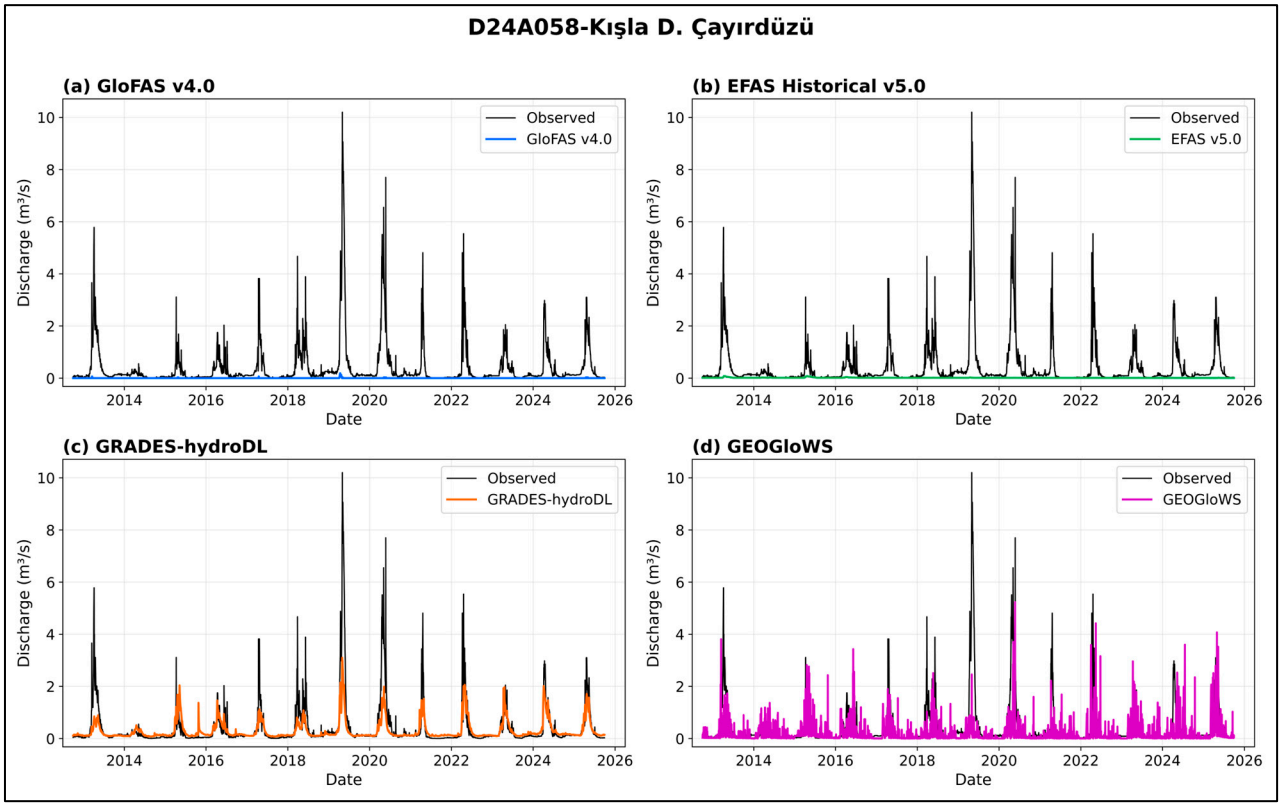


Figure 4. The water-year-based time series comparisons for the D24A058-Kışla D. Çayırdüzü gauging station covering the period from 1 October 2012 to 30 September 2025.

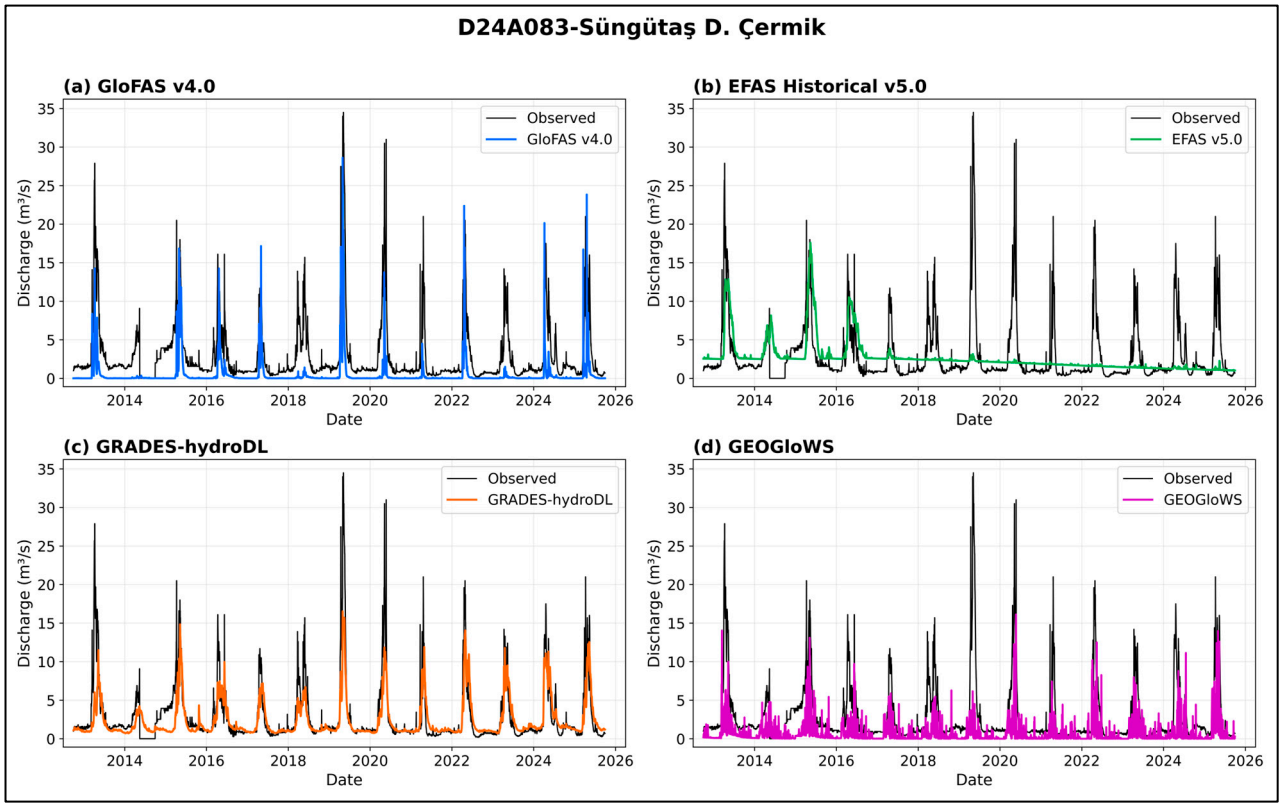


Figure 5. The water-year-based time series comparisons for the D24A083-Süngütaş D. Çermik gauging station covering the period from 1 October 2012 to 30 September 2025.

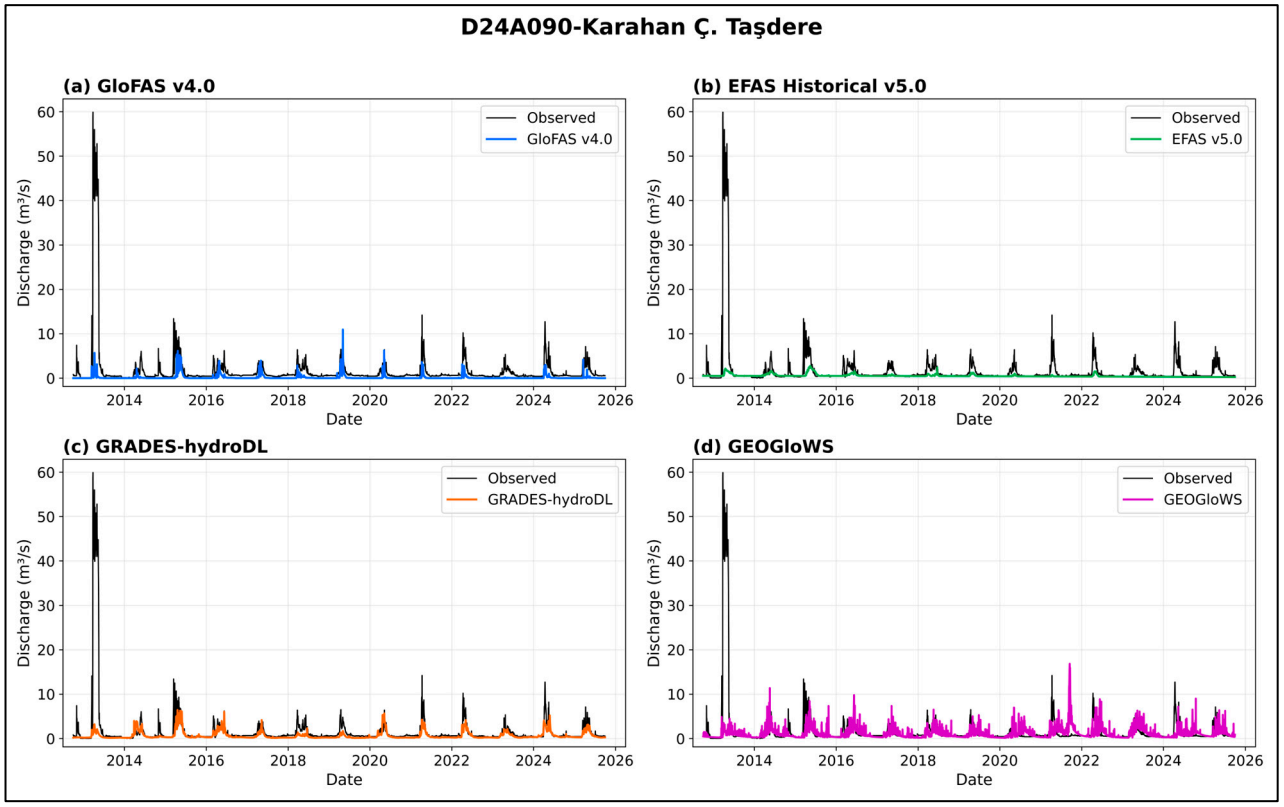


Figure 6. The water-year-based time series comparisons for the D24A090-Karahan Ç. Taşdere gauging station covering the period from 1 October 2012 to 30 September 2025.

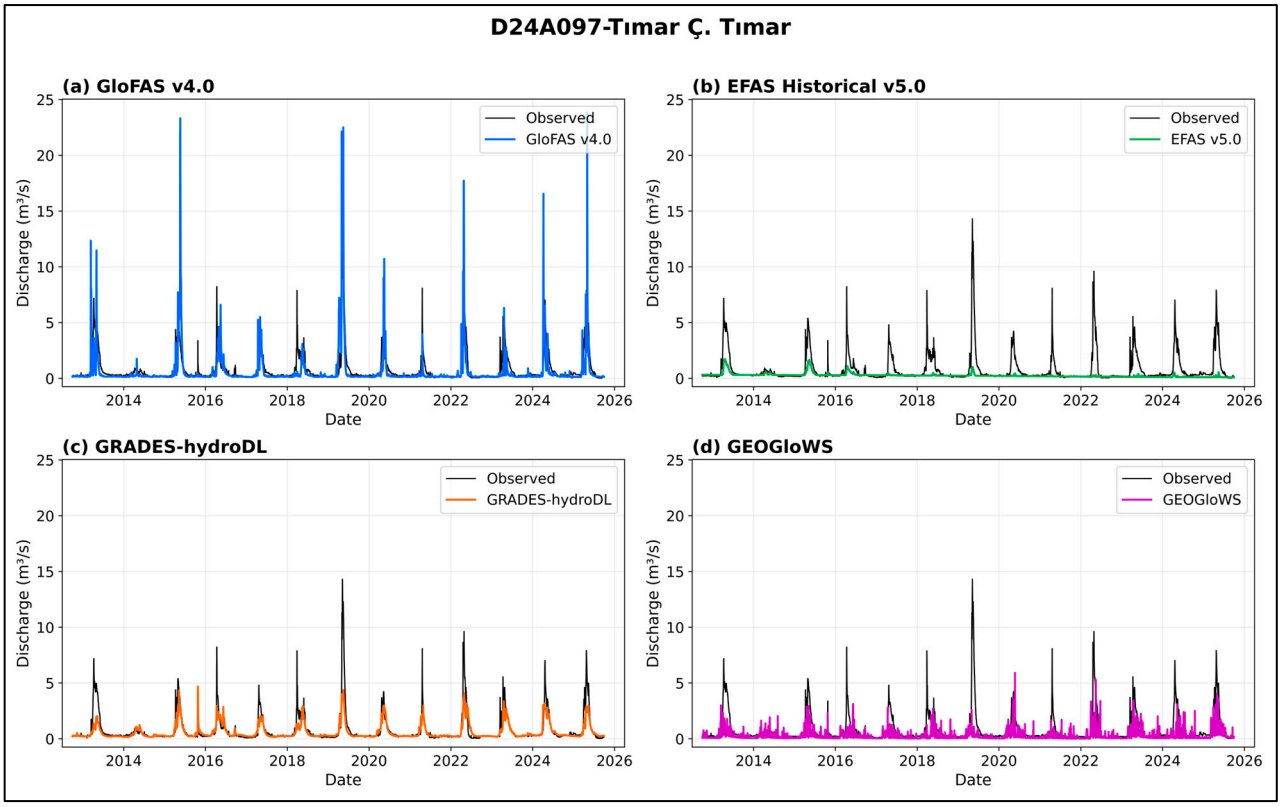


Figure 7. The water-year-based time series comparisons for the D24A097-Tımar Ç. Tımar gauging station covering the period from 1 October 2012 to 30 September 2025.

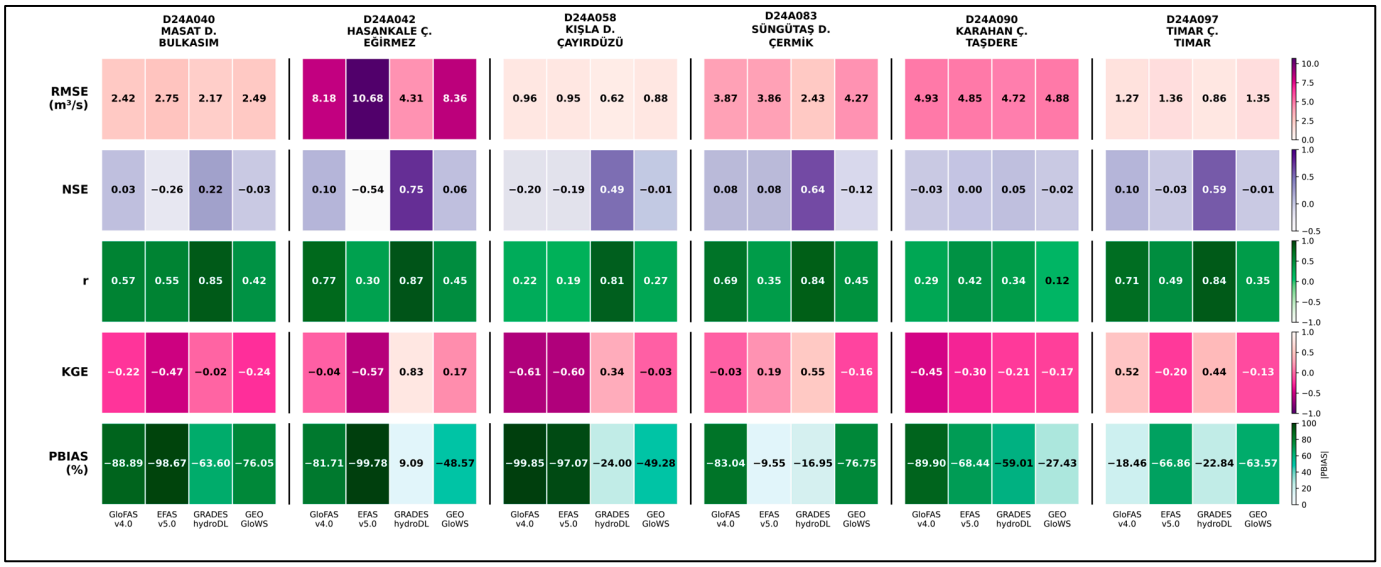


Figure 8. Results of statistical evaluation metrics reflecting the performance of model-based river discharge products against observed river discharge (1 October 2012–30 September 2025).

It is thought that considering only the water-year performance of model-based discharge products in terms of daily mean discharge for the period from 1 October 2012 to 30 September 2025 would not be sufficient. Therefore, it is considered that seasonal and monthly performances of model-based products should also be investigated. Figure 9 presents the seasonal mean river discharges of observations and outputs of model-based products for each station for the period from October 2012 to September 2025. At the same time, Figure 10 presents the monthly mean river discharges for the same period. Thus, the performance of GloFAS v4.0, EFAS v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service can be easily examined in terms of monthly and seasonal variations.

In Figure 9a, the seasonal mean river discharges for the D24A040-Masat D. Bulkasım station are presented. Based on the results plotted in Figure 9a, in autumn, it is seen that GRADES-hydroDL and the GEOGloWS ECMWF Streamflow Service estimate river discharges close to the observed mean discharges. The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0, in contrast, provide considerably low river discharges. In other words, the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 cannot reflect the real river discharge conditions. A similar river discharge pattern is also valid for the same station during the winter. In spring, the observed mean river discharges exceed 3.5 m^3/s. However, all model-based products predict lower mean discharge conditions than observations in spring at D24A040-Masat D. Bulkasım. Nevertheless, the output of the model-based discharge dataset that provides the highest mean discharge values is GRADES-hydroDL. It yields a mean river discharge of approximately 1 m^3/s in the spring. In contrast to the other seasons, namely, autumn and winter, the GloFAS v4.0 Hydrological Reanalysis produces distinct mean river discharges in spring. EFAS Historical v5.0 remains behind the other model-based datasets in terms of performance by providing considerably low seasonal mean discharge values. In spring, the reason why the observations present considerably higher mean discharge values than the model-based products, except for GRADES-hydroDL at D24A042 and GloFAS v4.0 at D24A097, is attributed to snowmelt-driven runoff. This indicates that the model-based products evaluated in this study have quite limited capabilities to represent the snowmelt-driven river discharge.

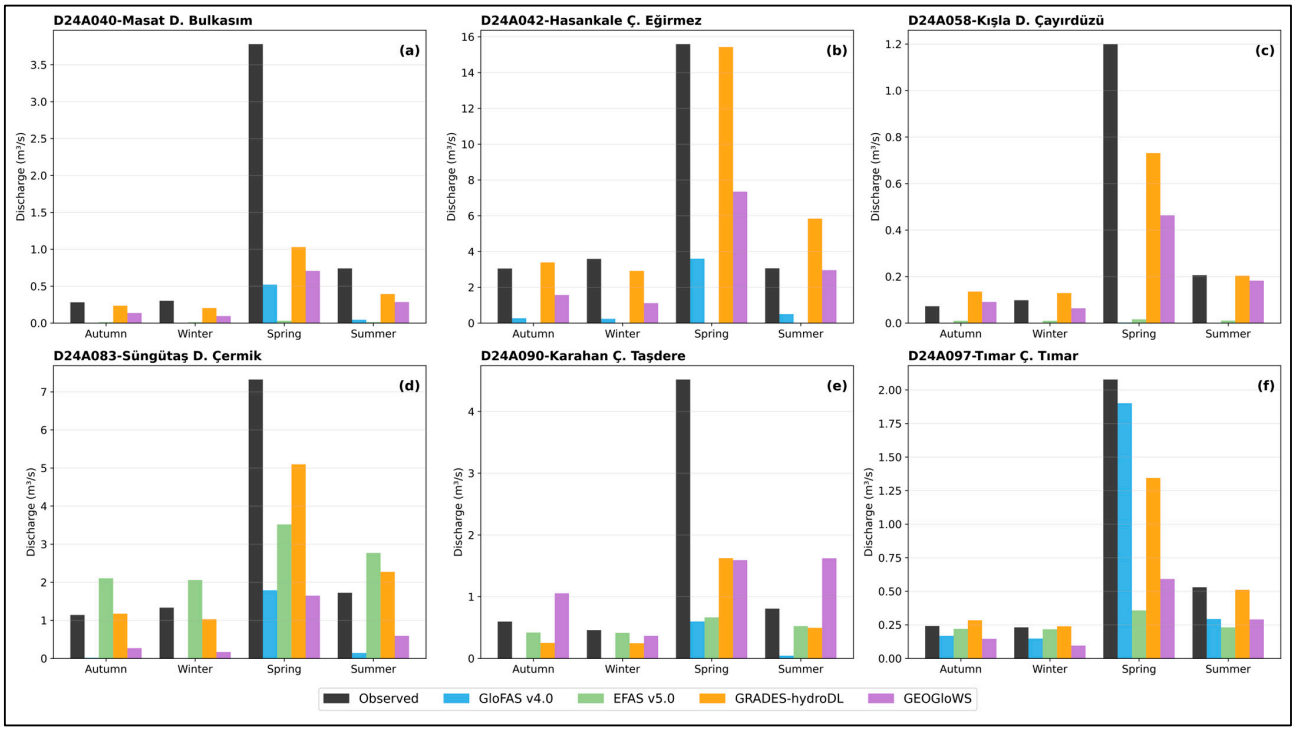


Figure 9. Comparison of the seasonal mean performance of the model-based river discharge products (October 2012–September 2025). (**a**) D24A040-Masat D. Bulkasım; (**b**) D24A042-Hasankale Ç. Eğirmez; (**c**) D24A058-Kışla D. Çayırdüzü; (**d**) D24A083-Süngütaş D. Çermik; (**e**) D24A090-Karahan Ç. Taşdere; (**f**) D24A097-Tımar Ç. Tımar.

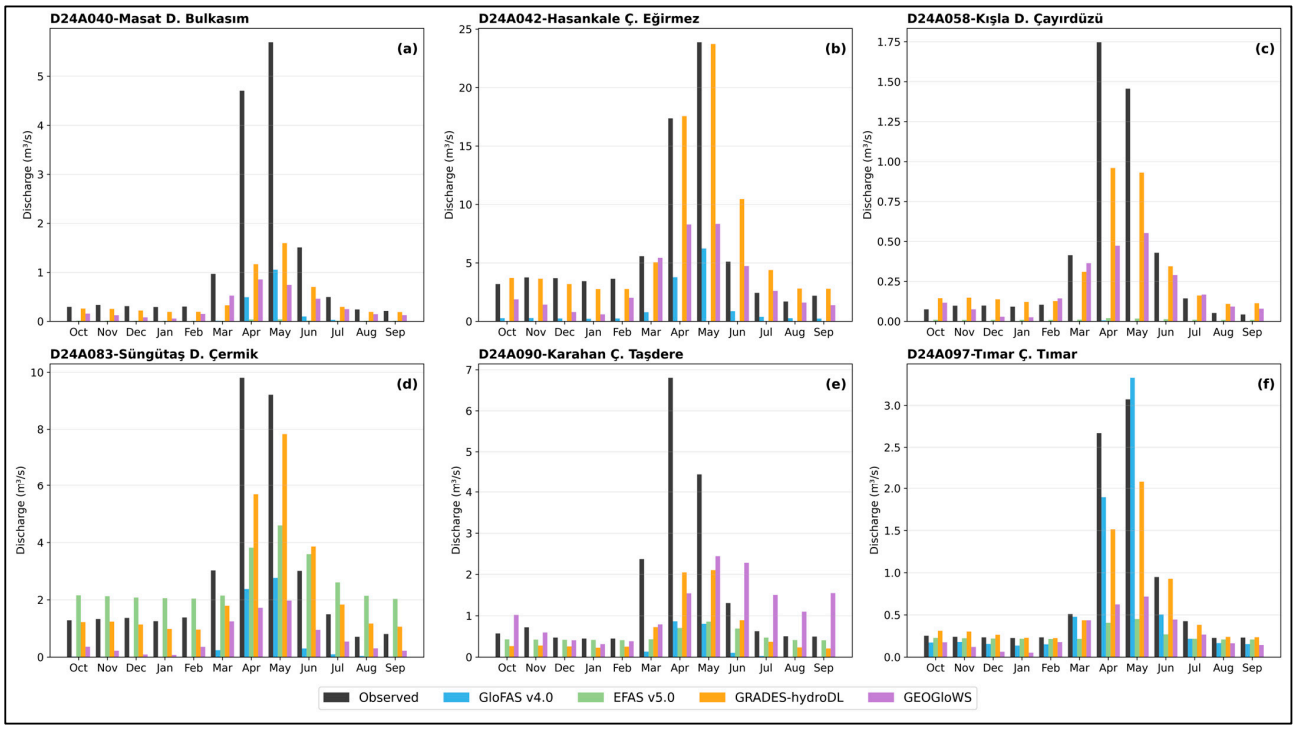


Figure 10. Comparison of the monthly mean performance of the model-based river discharge products (October 2012–September 2025). (**a**) D24A040-Masat D. Bulkasım; (**b**) D24A042-Hasankale Ç. Eğirmez; (**c**) D24A058-Kışla D. Çayırdüzü; (**d**) D24A083-Süngütaş D. Çermik; (**e**) D24A090-Karahan Ç. Taşdere; (**f**) D24A097-Tımar Ç. Tımar.

Figure 9b shows the results of seasonal mean river discharges for D24A042-Hasankale Ç. Eğirmez. This plot indicates that, in autumn and winter, the dataset that produces the river discharges closest to the observations is GRADES-hydroDL. In spring, when snowmelt-driven river discharges are pronounced, seasonal mean observations are also considerably high at this station. At the D24A042-Hasankale Ç. Eğirmez gauging station, the model-based dataset that provides the seasonal mean river discharges closest to the observed mean river discharges is again GRADES-hydroDL. In summer, however, GRADES-hydroDL simulates higher seasonal mean discharges than observations. In summer, the dataset that produces seasonal mean river discharges closest to the seasonal mean observation record is the GEOGloWS ECMWF Streamflow Service. The plot in Figure 9c covers the seasonal mean discharges for the D24A058-Kışla D. Çayırdüzü gauging station. A similar pattern at D24A040-Masat D. Bulkasım and D24A042-Hasankale Ç. Eğirmez is evident in autumn and winter for D24A058-Kışla D. Çayırdüzü in terms of seasonal mean river discharge conditions. In spring, observations remain again higher than the river discharges simulated by the model-based products due to the snowmelt-driven process. In summer, the model-based datasets that clearly stand out are GRADES-hydroDL and the GEOGloWS ECMWF Streamflow Service among the discharge products. In Figure 9d, it is seen that the EFAS Historical v5.0 clearly produces river discharges. EFAS Historical v5.0 was providing considerably low seasonal mean discharge values at the previous stations, namely, D24A040, D24A042, and D24A058. Nevertheless, for the D24A083 station, EFAS Historical v5.0 predicts higher river discharges than the seasonal mean observations in autumn, winter, and summer. This may be associated with the upstream drainage area of D24A083-Süngütaş D. Çermik. This will be evaluated in detail in Section 4. At this station, GRADES-hydroDL again exhibits the best performance among the model-based river discharge products. Regarding D24A090-Karahan Ç. Taşdere (see Figure 9e), this time, the GEOGloWS ECMWF Streamflow Service simulates higher seasonal mean river discharges than the observed seasonal mean river discharges in autumn. GRADES-hydroDL, on the other hand, remains below the seasonal mean values in all seasons. In winter, except for the GloFAS v4.0 Hydrological Reanalysis, the other model-based datasets yield river discharges close to each other. In contrast to the other results, in summer, the river discharges estimated by the GEOGloWS ECMWF Streamflow Service exceed the observed seasonal mean discharges. EFAS Historical v5.0 and GRADES-hydroDL, in summer, produce similar river discharge values. In Figure 9f, the GloFAS v4.0 Hydrological Reanalysis estimates the spring mean discharge close to the observations, whereas it simulates lower discharges in the other seasons. GRADES-hydroDL, at D24A097-Tımar Ç. Tımar, estimates seasonal mean discharge values similar to the observations, except in spring.

The final results under this section are the results related to the comparison of monthly mean discharge values derived from model-based products versus observations. The ability to reflect monthly variations in river discharge conditions is also an important indicator of model-based products' performance. In most months and at most stations, GRADES-hydroDL exhibits the highest agreement with the observed monthly mean discharges. In months with high observed discharge values, all model-based products show limited capabilities in capturing the peak discharge volumes. For D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, and D24A058-Kışla D. Çayırdüzü, the EFAS Historical v5.0 simulates considerably low monthly mean river discharges. At the other stations, namely, D24A083-Süngütaş D. Çermik, D24A090-Karahan Ç. Taşdere, and D24A097-Tımar Ç. Tımar, however, it is able to estimate noticeable monthly mean discharges. It even calculates, in some months, higher monthly mean values than the monthly mean observations. In April and May, it is observed that the monthly mean observations present markedly higher river discharges compared with the outputs of the model-based products, except

for GRADES-hydroDL at D24A042 and GloFAS v4.0 at D24A097. When it comes to the monthly performance ranking, the order is (1) GRADES-hydroDL (the best performance), (2) the GEOGloWS ECMWF Streamflow Service, (3) EFAS Historical v5.0, and (4) the GloFAS v4.0 Hydrological Reanalysis (the worst performance).

Harrigan et al. [18] demonstrated that the GloFAS-ERA5 Reanalysis exhibited different performances depending on location. The GloFAS-ERA5 Reanalysis showed negative KGESS (Kling–Gupta Efficiency Skill Score) in many basins across Africa, northeastern parts of Brazil, Thailand, and southern parts of Spain. This means that the GloFAS-ERA5 Reanalysis produced very poor river discharge estimates. In this study, similarly, the GloFAS v4.0 Hydrological Reanalysis simulates performance differences at the gauging stations in the Aras Basin. By simulating considerably low discharge estimates at some stations, it supports the findings of Harrigan et al. [18].

Hales et al. [19] reported that the GEOGloWS ECMWF Streamflow Service had limited capabilities in capturing the peak discharge and short-duration storms. The results in this study obtained for the GEOGloWS ECMWF Streamflow Service support the finding of Hales et al. [19] regarding peak discharge values.

Yang et al. [13] presented the performance of GRADES-hydroDL in terms of median KGE. In their study, all median KGEs were 0.592 or higher. This indicates that GRADES-hydroDL exhibited quite good performance. The evaluations in this part show that GRADES-hydroDL, similarly, shows quite good performance at most of the gauging stations in the Aras Basin.

The EFAS dataset evaluated in the study of Thielen et al. [51] can be considered the predecessor of the EFAS Historical v5.0 dataset. However, Thielen et al. [51] reported that the EFAS version evaluated in their study simulated considerably low discharge values for the Danube River. A similar issue also exists at some stations in the Aras Basin within the scope of this study. Thielen et al. [51] attributed this issue to the weak precipitation gauge network in the Alps.

Bartholmes et al. [52] reported that the performance of the EFAS decreased significantly for the upstream drainage areas smaller than 4000 km^2. The drainage areas of all gauging stations considered in the Aras Basin are considerably below 4000 km^2. It is thought that the EFAS Historical v5.0 produced considerably low discharges at some stations in the Aras Basin for the same reason.

3.2. Evaluation of the Performance of Model-Based River Discharge Products in Representing Low-, Moderate-, and High-Flow Conditions Using Flow Duration Curves

One of the other objectives of this study is to evaluate the performance of model-based products in terms of low-, moderate-, and high-flow conditions using flow duration curves. In this context, firstly, flow duration curves were plotted for the daily mean observations and the daily mean time series extracted from the model-based products. The resulting flow duration curves for all gauging stations are presented in Figure 11.

In the graphs presented in Figure 11, the horizontal axes show the probabilities of exceedance, whereas the vertical axes indicate the river discharges in m^3/s. The river discharge conditions are categorized into three groups, namely, low-, moderate-, and high-flows. Subsequently, low-, moderate-, and high-flows in m^3/s were derived from the flow duration curves plotted in Figure 11. Exceedance probabilities play a major role in determining discharge values categorized as low-, moderate-, and high-flow conditions. In other words, the low-flow conditions, namely, Q90 and Q95, refer to the river discharges corresponding to 90% and 95% exceedance probabilities, while the moderate-flow conditions, namely, Q25, Q50, and Q75, refer to the river discharges corresponding to 25%, 50%, and 75% exceedance probabilities. At the same time, the high-flow conditions, namely, Q5 and Q10, are also referred to as flow conditions corresponding to 5% and 10% exceedance probabilities.

The low-, moderate-, and high-flows derived from the flow duration curves presented in Figure 11 for each gauging station are summarized in Table 5.

Table 5. Low, moderate, and high flows derived from the flow duration curves.

Gauging Stations	Flow Quantile	Datasets				
		Observed	GloFAS v4.0	EFAS v5.0	GRADES-hydroDL	GEOGloWS
Masat D. Bulkasım	Q5 (5%)	6.7010	0.6875	0.0429	1.7500	1.2000
	Q10 (10%)	3.9200	0.1562	0.0273	1.1900	0.7700
	Q25 (25%)	0.8410	0.0156	0.0205	0.4500	0.3300
	Q50 (50%)	0.3220	0.0000	0.0117	0.2300	0.1200
	Q75 (75%)	0.2210	0.0000	0.0054	0.1900	0.0600
	Q90 (90%)	0.1750	0.0000	0.0029	0.1700	0.0400
	Q95 (95%)	0.1520	0.0000	0.0024	0.1600	0.0200
D24A042 Hasankale Ç. Eğirmez	Q5 (5%)	25.5000	5.4547	0.0293	25.5550	13.1575
	Q10 (10%)	15.6000	2.0656	0.0200	18.4830	8.6000
	Q25 (25%)	4.9200	0.5000	0.0176	6.9800	3.4900
	Q50 (50%)	3.6000	0.2734	0.0127	3.3400	1.3000
	Q75 (75%)	2.3700	0.2266	0.0054	2.7400	0.6000
	Q90 (90%)	1.6300	0.2188	0.0020	2.4400	0.3800
	Q95 (95%)	1.2600	0.2031	0.0010	2.2700	0.2700
D24A058 Kışla D. Çayırdüzü	Q5 (5%)	1.9110	0.0000	0.0242	1.1055	0.8855
	Q10 (10%)	1.0600	0.0000	0.0139	0.7700	0.5900
	Q25 (25%)	0.2580	0.0000	0.0117	0.2900	0.2000
	Q50 (50%)	0.1110	0.0000	0.0107	0.1400	0.0500
	Q75 (75%)	0.0510	0.0000	0.0074	0.1200	0.0100
	Q90 (90%)	0.0220	0.0000	0.0059	0.1000	0.0000
	Q95 (95%)	0.0100	0.0000	0.0054	0.0900	0.0000
D24A083 Süngütaş D. Çermik	Q5 (5%)	11.8100	2.6562	7.2456	8.3610	3.0200
	Q10 (10%)	7.5600	0.7500	3.6019	6.0400	1.9100
	Q25 (25%)	3.1600	0.0938	2.5526	2.5600	0.7000
	Q50 (50%)	1.3200	0.0156	2.2545	1.2600	0.1700
	Q75 (75%)	0.8900	0.0000	1.5766	1.0100	0.0200
	Q90 (90%)	0.5140	0.0000	1.2339	0.9100	0.0000
	Q95 (95%)	0.3660	0.0000	1.1276	0.8500	0.0000
D24A090 Karahan Ç. Taşdere	Q5 (5%)	4.6300	0.9375	1.1307	2.6200	3.5000
	Q10 (10%)	2.8800	0.3438	0.7765	1.7500	2.5700
	Q25 (25%)	0.9975	0.0391	0.4961	0.6600	1.4600
	Q50 (50%)	0.5850	0.0000	0.4517	0.2900	0.7300
	Q75 (75%)	0.4470	0.0000	0.3545	0.2100	0.3700
	Q90 (90%)	0.3900	0.0000	0.2793	0.1800	0.2600
	Q95 (95%)	0.2780	0.0000	0.2514	0.1600	0.2200
D24A097 Tımar Ç. Tımar	Q5 (5%)	3.5500	2.8750	0.5137	2.2455	1.1000
	Q10 (10%)	2.1600	1.1562	0.3167	1.6400	0.7400
	Q25 (25%)	0.5870	0.2812	0.2907	0.6000	0.3100
	Q50 (50%)	0.2820	0.1719	0.2361	0.2800	0.1200
	Q75 (75%)	0.1980	0.1406	0.1510	0.2300	0.0500
	Q90 (90%)	0.1620	0.1250	0.1114	0.2000	0.0400
	Q95 (95%)	0.1180	0.1250	0.1021	0.1900	0.0300

Note: Dark blue: high flows (Q5 and Q10); medium blue: moderate flows (Q25, Q50, and Q75); light blue: low flows (Q90 and Q95).

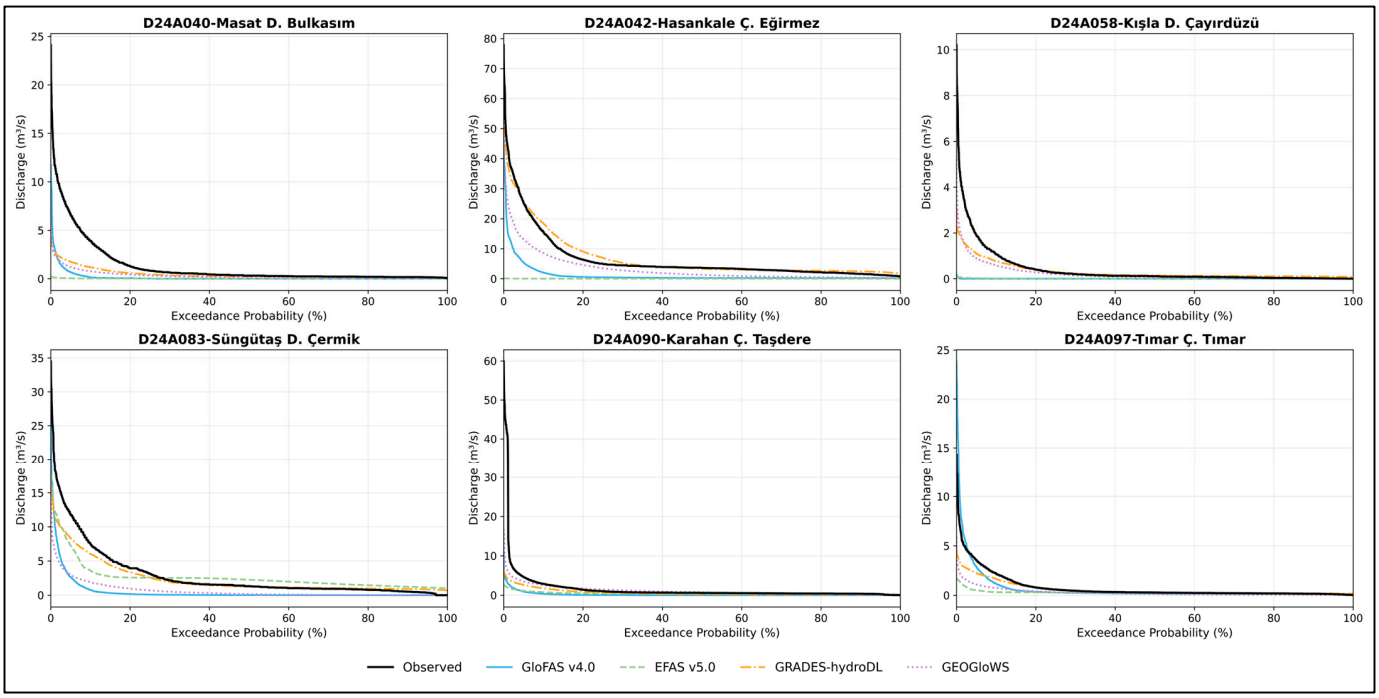


Figure 11. The resulting flow duration curves for all gauging stations.

For D24A040-Masat D. Bulkasım, the observed Q5 and Q10, the high-flow conditions, are 6.7010 m^3/s and 3.92 m^3/s, respectively (see Table 5). At the same time, GRADES-hydroDL produces the high-flow conditions as Q5 (1.75 m^3/s) and Q10 (1.19 m^3/s), whereas the high-flow conditions for the GEOGloWS ECMWF Streamflow Service are Q5 (1.20 m^3/s) and Q10 (0.77 m^3/s). For the same gauging station, the Q25, Q50, and Q75 derived from the observations are 0.8410, 0.3220, and 0.2210 m^3/s, respectively. The GloFAS v4.0 Hydrological Reanalysis yields a discharge condition only for the Q25 due to the low river discharges. That Q25 is also quite low when compared with the Q25 calculated by the observed river discharge time series. Among the moderate-flow conditions, GRADES-hydroDL again simulates the Q25-Q50-Q75 quite well. A similar behavior is also valid for the low-flow conditions. In fact, the low-flow conditions (Q90 and Q95) provided by GRADES-hydroDL are quite similar to the low-flow conditions calculated by the observations. For D24A042-Hasankale Ç. Eğirmez, the high-flow conditions calculated by observed river discharges are 25.50 m^3/s (Q5) and 15.60 m^3/s (Q10), respectively. These are the highest high-flow conditions (Q5 and Q10) among all stations. While GRADES-hydroDL determines the Q5 very well, it provides a higher Q10 than the observation at D24A042. The ability of GRADES-hydroDL to capture the moderate-flow conditions Q25, Q50, and Q75 is also quite high. In terms of low-flow conditions, GRADES-hydroDL determines higher river discharges than the observed low-flow conditions. For the D24A058-Kışla D. Çayırdüzü, in terms of overall performance, GRADES-hydroDL is again the best dataset that is trustworthy. For this station, Kışla D. Çayırdüzü, the GloFAS v4.0 Hydrological Reanalysis does not provide any low-, moderate-, and high-flow conditions. The EFAS Historical v5.0 also calculates considerably low discharge values for all flow categories. At D24A083, GRADES-hydroDL can calculate discharge values close to the observed river discharge values in the high-flow and moderate-flow categories, whereas it overestimates the low-flow conditions (Q90 and Q95). For D24A083-Süngütaş D. Çermik, the most notable finding is that the GloFAS v4.0 Hydrological Reanalysis is able to calculate river discharges reflecting high-flow conditions (Q5 and Q10). Considering all of the stations, the ability of the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 to reflect low-, moderate-, and

high-flow conditions varies from station to station, without a consistent relationship with the magnitude of the observed discharges. In other words, at gauging stations where low observed discharge values are measured, the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 remain insufficient in reflecting low-, moderate-, and high-flow conditions. For D24A090-Karahan Ç. Taşdere, the GEOGloWS ECMWF Streamflow Service provides the flow quantiles closest to the observations, whereas for D24A097-Tımar Ç. Tımar, GRADES-hydroDL and the GloFAS v4.0 Hydrological Reanalysis provide the closest values.

Overall, it is clear that GRADES-hydroDL outperformed in determining low-, moderate-, and high-flow conditions.

Hossain et al. [53] stated that GloFAS v2.1 produced better results than GloFAS v3.1 for the Brahmaputra River, particularly in identifying Q10 and Q5 flow conditions. Hossain et al. [53] clearly reported that both versions, the GloFAS v2.1 and GloFAS v3.1, showed limited skill in identifying Q1 representing extreme flood conditions. Similarly, in this study, GloFAS v4.0 also has limited capacity in detecting low-, moderate-, and high-flow conditions.

3.3. Performance Analysis of the Model-Based Discharge Products in Terms of High-FlowEvent Detection

In order to test the performance of the model-based discharge products, namely, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service, it is considered necessary to also evaluate their behavior in detecting high-flow events. This is because time series analyses, in addition to low-, moderate-, and high-flow conditions, are not sufficient in terms of the overall performance of the model-based products. Motivated by this, this study also evaluated the behavior of model-based products, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service, in detecting high-flow events. To this end, the daily mean discharge time series provided by the observations and model-based products for the period from 1 October 2012 to 30 September 2025 were used. Then, as a first step, the station-specific threshold was considered the 90th percentile of the observed daily mean river discharge time series that corresponds to the Q10. The station-specific Q10s were determined from the flow duration curves plotted in the previous section. The Q10 values for all gauging stations can be obtained from Table 5 presented in the previous section. As noted earlier, the values of station-specific Q10s are the river discharges corresponding to a 10% exceedance probability on the flow duration curves. When the defined threshold (Q10) is exceeded, in this case, the river discharge is referred to as "high-flow," and the corresponding event is also classified as a "high-flow event." The ability of model-based products to detect high-flow events is considered an important tool for identifying possible measures against high-flow events, particularly in flood-prone areas. The determination of the Hit (H), Miss (M), and False Alarm (FA) used in calculating the Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI), so-called performance metrics, is critical. After identifying the Hit (H), Miss (M), and False Alarm (FA), the POD, FAR, and CSI were calculated using Equations (7)–(9) presented in Section 2.5. These performance metrics, in fact, reflect the ability of the model-based products to detect high-flow events in terms of station-specific threshold Q10. All contingency and high-flow metrics are presented in Table 6.

Table 6. Contingency and high-flow metrics for all model-based river discharge datasets.

Station No.	Dataset	Hit (H)	Miss (M)	False Alarm (FA)	Total Days	Total Observed High-Flow Days	POD	FAR	CSI	Performance
D24A040	GloFAS v4.0	28	438	0	4748	466	0.0601	0.0000	0.0601	Poor
	EFAS v5.0	0	466	0			0.0000	N/A	0.0000	No detection
	GRADES-hydroDL	5	461	0			0.0107	0.0000	0.0107	Poor
	GEOGloWS	5	461	4			0.0107	0.4444	0.0106	Poor
D24A042	GloFAS v4.0	47	424	1	4748	471	0.0998	0.0208	0.0996	Poor
	EFAS v5.0	0	471	0			0.0000	N/A	0.0000	No detection
	GRADES-hydroDL	405	66	170			0.8599	0.2957	0.6318	Good
	GEOGloWS	91	380	75			0.1932	0.4518	0.1667	Poor
D24A058	GloFAS v4.0	0	472	0	4748	472	0.0000	N/A	0.0000	No detection
	EFAS v5.0	0	472	0			0.0000	N/A	0.0000	No detection
	GRADES-hydroDL	211	261	44			0.4470	0.1725	0.4089	Moderate
	GEOGloWS	67	405	104			0.1419	0.6082	0.1163	Poor
D24A083	GloFAS v4.0	80	389	2	4748	469	0.1706	0.0244	0.1699	Poor
	EFAS v5.0	104	365	117			0.2217	0.5294	0.1775	Poor
	GRADES-hydroDL	235	234	63			0.5011	0.2114	0.4417	Moderate
	GEOGloWS	13	456	16			0.0277	0.5517	0.0268	Poor
D24A090	GloFAS v4.0	50	421	4	4748	471	0.1062	0.0741	0.1053	Poor
	EFAS v5.0	0	471	0			0.0000	N/A	0.0000	No detection
	GRADES-hydroDL	139	332	49			0.2951	0.2606	0.2673	Poor
	GEOGloWS	104	367	281			0.2208	0.7299	0.1383	Poor
D24A097	GloFAS v4.0	211	261	100	4748	472	0.4470	0.3215	0.3684	Moderate
	EFAS v5.0	0	472	0			0.0000	N/A	0.0000	No detection
	GRADES-hydroDL	214	258	39			0.4534	0.1542	0.4188	Moderate
	GEOGloWS	22	450	21			0.0466	0.4884	0.0446	Poor

The performance of high-flow event detection of the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service is analyzed to cover the period from 1 October 2012 to 30 September 2025. This period covers a total of 4748 days. At D24A040-Masat D. Bulkasım, a total of 466 high-flow events were observed. At this station, the performance of all model-based products except for EFAS Historical v5.0 in determining high-flow events is categorized as poor. The performance of EFAS Historical v5.0 is classified as "No Detection" because EFAS Historical v5.0 had simulated relatively very low river discharges at this station. Overall, the river discharge estimates at D24A040 remained quite poor in determining high-flow events. This means that all river discharge products missed a large proportion of the observed high-flow events. At D24A042-Hasankale Ç. Eğirmez, GRADES-hydroDL exhibits prominent performance by successfully detecting a total of 405 high-flow events among the evaluated discharge estimates. Similar to D24A040-Masat D. Bulkasım, EFAS Historical v5.0 was classified as "no detection." This is because, at the D24A042 gauging station, the EFAS Historical v5.0 dataset also presents considerably low river discharges. When it comes to the D24A058-Kışla D. Çayırdüzü, GRADES-hydroDL is again the best-performing discharge product. The performance of GRADES-hydroDL at this station is classified as "moderate." Since the GloFAS v4.0 Reanalysis and EFAS Historical v5.0 provided considerably low river discharges for the D24A058, the performance of the aforementioned products is classified as "No Detection" in capturing high-flow events. For D24A083-Süngütaş D. Çermik, GRADES-hydroDL is again the best-performing model, with its success in detecting 235 of the 469 high-flow events. Unfortunately, for this station, the other products—namely, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, and the GEOGloWS ECMWF Streamflow Service—are poorly performing discharge products when it comes to tracking high-flow events. Unfortunately, a similar situation is also valid for D24A090-Karahan Ç. Taşdere. Out of the 471 observed high-flow events observed at the D24A090 gauging station, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service detected 50, 0, 139, and 104 high-flow events, respectively. This is considered poor performance. Again, it should also be noted that EFAS Historical v5.0 shows poor performance, ascertaining 0 of 471 high-flow events. The daily mean river discharges measured at D24A090-Karahan Ç. Taşdere are lower than the river discharges observed at the other stations. It was mentioned in the previous sections that EFAS Historical v5.0 and the GloFAS v4.0 Reanalysis performed relatively better at the gauging stations where low river discharges were observed. Notwithstanding this, the performance improvement was not observed in detecting high-flow events. This is a major disadvantage for EFAS Historical v5.0 and GloFAS v4.0. At the last gauging station, D24A097-Tımar Ç. Tımar, improvement in terms of high-flow event performance is observed for GRADES-hydroDL and the GloFAS v4.0 Hydrological Reanalysis, both of which are classified as "moderate". EFAS Historical v5.0 again failed to capture any of the 472 observed high-flow events. The GEOGloWS ECMWF Streamflow Service also presents a low high-flow detection rate. The highest False Alarm Ratio at this station is determined by the GEOGloWS ECMWF Streamflow Service (FAR = 0.4884). It means that the GEOGloWS ECMWF Streamflow Service simulates 21 high-flow events that do not exist in the observed daily river discharge time series.

Overall, all evaluated model-based products have limited abilities in capturing high-flow events in the Aras Basin.

To test the effect of threshold selection on the results, the high-flow detection analysis was repeated in three ways: (1) the Q5, Q10, and Q20 values of the observed series were used as common thresholds; (2) the Q10 value of each dataset was applied as the threshold for that dataset; and (3) consecutive days exceeding the threshold were counted as a single event, and the metrics were calculated on an event basis. The results are summarized in Table 7 in terms of CSI values. With the observation-based thresholds, the CSI values of all products increase as the threshold is relaxed, but the ranking of the products is largely preserved; GRADES-hydroDL shows the best performance in 15 of the 18 station–threshold combinations, and EFAS v5.0 the lowest or tied for the lowest in 15. When the dataset-specific threshold is used, the CSI values of GloFAS v4.0 at the stations where it strongly underestimates the discharge (D24A040, D24A042, D24A083, and D24A090) rise from 0.06–0.17 to 0.42–0.59 and reach a level similar to that of GRADES-hydroDL. This indicates that the low performance of GloFAS v4.0 in Table 6 results largely from its systematic underestimation, whereas the timing of high flows is captured reasonably well. The CSI values of the GEOGloWS ECMWF Streamflow Service, on the other hand, remain at 0.16–0.24 even with its threshold. The event-based evaluation is consistent with the day-based results.

Table 7. Sensitivity of high-flow detection skill (CSI) to the threshold definition.

Station No.	Dataset	Observed Threshold			Dataset-Specific Threshold	Event-Based (Observed Q10)
		Q5	Q10	Q20	Q10	Q10
D24A040	GloFAS v4.0	0.05	0.06	0.14	0.56	0.17
	EFAS v5.0	0.00	0.00	0.00	0.43	0.00
	GRADES-hydroDL	0.00	0.01	0.44	0.62	0.05
	GEOGloWS	0.00	0.01	0.16	0.23	0.09
D24A042	GloFAS v4.0	0.10	0.10	0.22	0.59	0.19
	EFAS v5.0	0.00	0.00	0.00	0.21	0.00
	GRADES-hydroDL	0.50	0.63	0.66	0.63	0.64
	GEOGloWS	0.07	0.17	0.33	0.24	0.34
D24A058	GloFAS v4.0	0.00	0.00	0.00	0.08	0.00
	EFAS v5.0	0.00	0.00	0.00	0.14	0.00
	GRADES-hydroDL	0.14	0.41	0.70	0.52	0.40
	GEOGloWS	0.03	0.12	0.29	0.20	0.20
D24A083	GloFAS v4.0	0.14	0.17	0.18	0.59	0.32
	EFAS v5.0	0.16	0.18	0.23	0.17	0.18
	GRADES-hydroDL	0.17	0.44	0.63	0.51	0.46
	GEOGloWS	0.02	0.03	0.12	0.22	0.16
D24A090	GloFAS v4.0	0.03	0.11	0.17	0.42	0.16
	EFAS v5.0	0.00	0.00	0.13	0.30	0.00
	GRADES-hydroDL	0.08	0.27	0.55	0.43	0.28
	GEOGloWS	0.07	0.14	0.32	0.16	0.17
D24A097	GloFAS v4.0	0.32	0.37	0.54	0.47	0.40
	EFAS v5.0	0.00	0.00	0.14	0.27	0.00
	GRADES-hydroDL	0.12	0.42	0.77	0.55	0.38
	GEOGloWS	0.01	0.04	0.22	0.21	0.21

3.4. Hydrological Drought Characterization of Discharge Estimates by StreamflowDrought Index (SDI)

In this study, the performance of the river discharge products is evaluated from multiple hydrological perspectives, such as time series analyses, flow conditions, etc. In this section, the performance of the GloFAS v4.0 Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service is evaluated in terms of drought analysis using the Streamflow Drought Index (SDI). What is meant by "drought analysis" is to test the ability of the model-based river discharge products to identify hydrological dry and wet periods in the Aras Basin. The SDIs were calculated for the 12-month reference period, and are called SDI-12s. The purpose of selecting this reference period is that SDI-12 can reflect long-term annual hydrological variations. The SDI-12 plots calculated for the 12-month reference period are shown in Figure 12 for each gauging station.

In the plots in Figure 12, the blue bars represent the hydrological wet periods, whereas the red bars denote the hydrological dry periods for each gauging station. Since three global river discharge datasets (GloFAS v4.0 Hydrological Reanalysis, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service) and one regional river discharge dataset (EFAS Historical v5.0) are compared, SDI-12s are presented using line plots. The solid blue lines show the SDI-12 of the GloFAS v4.0 Hydrological Reanalysis, the dashed green lines represent the SDI-12 of the EFAS Historical v5.0, the dash-dotted orange lines denote the SDI-12 for GRADES-hydroDL, and the dotted magenta lines in the plots indicate those of the GEOGloWS ECMWF Streamflow Service. The aim of this section is to quantify the performance of discharge estimates in identifying hydrological dry and wet periods in the Aras Basin. The findings in this section will concurrently reveal the applicability of the mentioned discharge estimates in drought analysis studies. Figure 12a shows the SDI-12 for D24A040-Masat D. Bulkasım. At this station, the GloFAS v4.0 Hydrological Reanalysis and GEOGloWS ECMWF Streamflow Service are the discharge products that generally captured the drought pattern. EFAS v5.0 has truly limited capacity to calculate hydrological dry and wet periods. For example, EFAS Historical v5.0 is able to capture only the May 2021–April 2022 period among the observed hydrological dry periods. In Figure 12b, which shows the SDI-12 for D24A042-Hasankale Ç. Eğirmez, the GloFAS v4.0 is more successful than the other river discharge datasets in capturing hydrological dry and wet periods. This dataset also predicts a large part of the severe hydrological drought periods at the D24A042 gauging station. GRADES-hydroDL, on the other hand, is able to clearly reflect the drought pattern, especially during April 2014–January 2016. In terms of drought performance, the next best-performing river discharge dataset is the GEOGloWS ECMWF Streamflow Service at this station. The SDI-12 calculated for D24A058-Kışla D. Çayırdüzü is seen in Figure 12c. None of the model-based discharge datasets could predict the SDI-12, so-called hydrological dry and wet periods at this station, at a satisfactory level. In fact, in some periods, all evaluated discharge estimates produced quite different results. At D24A083-Süngütaş D. Çermik, the GloFAS v4.0 Reanalysis is able to capture the larger portion of the hydrological drought periods compared with the other discharge products. Regarding drought severity, GloFAS v4.0, EFAS v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service have difficulties in determining drought severities. In Figure 12e, which shows the SDI-12 for D24A090-Karahan Ç. Taşdere, the most notable discrepancy is that the GEOGloWS ECMWF Streamflow Service presents a spurious hydrological wet period from 2021 to 2024. However, the SDI-12 calculated from the observed river discharges shows hydrological dry conditions during a large part of this period. Similarly, the EFAS v5.0 also predicts severe drought conditions, particularly during 2024–2025. At D24A097-Tımar Ç. Tımar, the model-based discharge products that show the best performance in identifying severe drought conditions are GloFAS v4.0 and GRADES-hydroDL. It can be said that these

discharge products also capture the drought conditions very well during 2014–2015 and 2021–2022.

Overall, GloFAS v4.0 and GRADES-hydroDL are successful, particularly in identifying hydrological dry periods in the Aras Basin. However, the same is not valid for identifying the drought severities. In fact, some model-based products produce spurious hydrological dry conditions during certain periods. Lastly, it can be said that EFAS Historical v5.0 has limited capacity in capturing hydrological dry and wet periods among the evaluated river discharge products.

In this study, although GloFAS v4.0 simulated considerably low discharges at some gauging stations in the Aras Basin, it exhibited good performance in identifying the hydrological dry periods. In addition, the findings of Swain et al. [22] for the SDI-12 using GloFAS v3.1 are consistent with the SDI-12 by the GloFAS v4.0 Hydrological Reanalysis in this study.

Patidar et al. [25] reported a correlation of $\rho = 0.58 \pm 0.14$ between the SDI-12 calculated by the observations and the SDI-12 derived from GloFAS v3.1. The correlation, in this study, between the SDI-12 for observations and the SDI-12 derived from GloFAS v4.0 is $\rho = 0.58 \pm 0.13$. Both correlations are quite comparable.

The Streamflow Drought Index is a standardized index. Therefore, even if the river discharges provided by the model-based products are considerably lower than the observations, the calculated SDI can still identify hydrological dry and wet periods. To test this effect, the SDI-12s were recalculated using a second approach. In the second approach, the 12-month volumes of the extracted time series were standardized using the distribution parameters of the observed series. In this way, the low discharge values are reflected in the SDI. The hydrological drought metrics were calculated for both approaches and are presented in Table 8. Unlike the formulations given in Section 2.5, the POD, FAR, and CSI values in Table 8 were calculated on a monthly basis. The criterion for identifying drought months is an SDI value of -1 or below. According to Table 8, in the first approach, GloFAS v4.0 captures most of the drought months at the D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, and D24A097-Tımar Ç. Tımar gauging stations with very few false alarms. In the second approach, however, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, and the GEOGloWS ECMWF Streamflow Service produce drought conditions throughout the entire period at most stations; specifically, the low discharge estimates from these products are interpreted as continuous drought when the index is calculated based on the observations. GRADES-hydroDL, on the other hand, gives similar results in both approaches. Accordingly, the agreement in Figure 12 indicates that GloFAS v4.0 captures the timing of the drought periods, but it does not mean that it correctly represents the drought severity.

Table 8. Hydrological drought metrics based on the SDI-12s for all model-based river discharge datasets.

Station No.	Dataset	Observed Drought Months	Observed Drought Events	Self-Standardized SDI-12 (Figure 12)						Observation-Referenced SDI-12					
				POD	FAR	CSI	*r*	Onset Lag	End Lag	POD	FAR	CSI	*r*	Onset Lag	End Lag
D24A040	GloFAS v4.0	33	4	0.67	0.08	0.63	0.71	−0.5	−0.5	1.00	0.77	0.23	0.21	–	–
	EFAS v5.0			0.24	0.75	0.14	−0.05	3.0	0.0	1.00	0.77	0.23	0.00	–	–
	GRADES-hydroDL			0.42	0.42	0.33	0.77	−1.0	−0.5	1.00	0.77	0.23	0.60	–	–
	GEOGloWS			0.55	0.22	0.47	0.75	−0.5	−1.0	1.00	0.77	0.23	0.13	–	–
D24A042	GloFAS v4.0	25	3	0.88	0.04	0.85	0.74	−0.5	−0.5	1.00	0.83	0.17	0.43	–	–
	EFAS v5.0			0.08	0.94	0.03	−0.04	−7.0	27.0	1.00	0.83	0.17	0.00	–	–
	GRADES-hydroDL			0.48	0.40	0.36	0.64	−8.0	0.0	0.48	0.08	0.46	0.63	−1.0	0.0
	GEOGloWS			0.60	0.44	0.41	0.48	2.5	−1.0	1.00	0.83	0.17	0.40	–	–
D24A058	GloFAS v4.0	22	1	0.00	N/A	0.00	0.56	N/A	N/A	1.00	0.85	0.15	0.00	–	–
	EFAS v5.0			0.00	1.00	0.00	−0.24	N/A	N/A	1.00	0.85	0.15	0.00	–	–
	GRADES-hydroDL			0.55	0.37	0.41	0.53	−5.0	−10.0	0.50	0.08	0.48	0.52	−1.0	−11.0
	GEOGloWS			0.00	1.00	0.00	−0.16	N/A	N/A	0.36	0.85	0.12	−0.16	−1.0	−14.0
D24A083	GloFAS v4.0	31	5	0.61	0.59	0.33	0.68	−0.7	4.0	1.00	0.79	0.21	0.54	–	–
	EFAS v5.0			0.00	1.00	0.00	0.33	N/A	N/A	0.45	0.77	0.18	0.32	−17.0	34.0
	GRADES-hydroDL			0.48	0.32	0.39	0.47	3.3	1.5	0.58	0.40	0.42	0.46	0.3	2.5
	GEOGloWS			0.48	0.35	0.38	0.52	0.7	−1.0	1.00	0.79	0.21	0.33	–	–
D24A090	GloFAS v4.0	13	4	0.38	0.62	0.24	0.36	0.0	7.0	1.00	0.91	0.09	0.11	–	–
	EFAS v5.0			0.31	0.86	0.11	0.39	1.0	25.0	1.00	0.89	0.11	0.24	−40.5	37.8
	GRADES-hydroDL			0.00	1.00	0.00	0.11	N/A	N/A	1.00	0.86	0.14	0.20	−20.5	3.8
	GEOGloWS			0.23	0.87	0.09	−0.07	−8.0	1.0	0.23	0.67	0.16	0.04	−1.0	1.0
D24A097	GloFAS v4.0	24	4	0.92	0.04	0.88	0.80	−0.5	0.0	0.92	0.67	0.32	0.76	−1.5	0.0
	EFAS v5.0			0.00	1.00	0.00	−0.05	N/A	N/A	1.00	0.83	0.17	0.07	–	–
	GRADES-hydroDL			0.88	0.28	0.66	0.66	−3.0	−2.0	0.92	0.52	0.46	0.66	−4.5	1.0
	GEOGloWS			0.42	0.58	0.26	0.37	−1.0	0.0	1.00	0.83	0.17	0.32	–	–

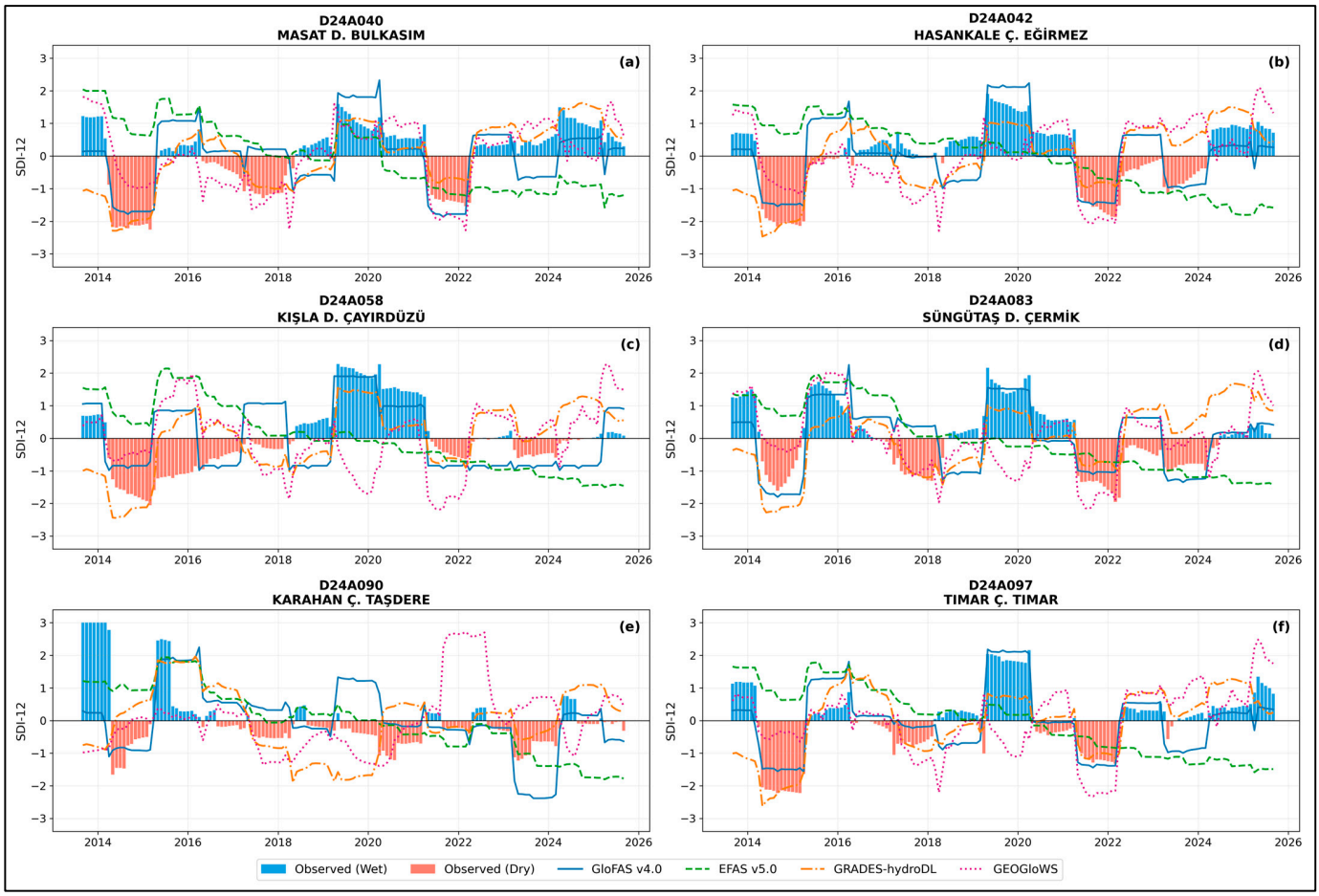


Figure 12. Streamflow Drought Index (SDI) calculated for the 12-month reference period for each gauging station. (**a**) D24A040-Masat D. Bulkasım; (**b**) D24A042-Hasankale Ç. Eğirmez; (**c**) D24A058-Kışla D. Çayırdüzü; (**d**) D24A083-Süngütaş D. Çermik; (**e**) D24A090-Karahan Ç. Taşdere; (**f**) D24A097-Tımar Ç. Tımar.

4. Discussion

In this study, the performance of three global datasets—GloFAS v4.0 Reanalysis, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service—and one regional dataset—EFAS Historical v5.0—were examined in many hydrological aspects using river discharge datasets from six gauging stations operated by the General Directorate of State Hydraulic Works in the Aras Basin in Türkiye. The valuable findings indicate that four evaluated model-based streamflow products exhibit different performance behaviors in terms of hydrological aspects.

Evaluation of the water-year-based time series and, in addition, the monthly and seasonal mean discharge plots, which were presented in the previous section, show that all four model-based products cannot fully represent snowmelt-driven runoff. The gauging stations considered in this study are operated at elevations of 1597–1801 m. Snowmelt-driven runoff is an important tool for gauging stations operated in both mountainous terrain and high elevations. The structural weaknesses of the river discharge products regarding the snowmelt-driven runoff become evident, particularly in spring months. Because the spring months, March–April–May, are the dominant months for the snowmelt-driven surface runoff in the Aras Basin. In other words, during the spring months (March–April–May), the river discharges provided by the model-based products remain markedly below the observed discharges, except for GRADES-hydroDL at D24A042 and GloFAS v4.0 at D24A097. Further improvements for the evaluated river discharge products are urgently needed to

overhaul the ability to calculate snowmelt-driven surface runoff. At this point, the first aspect to be evaluated is the hydrological models of the four model-based river discharge products. Both the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 employ the LISFLOOD hydrological model to generate river discharge estimates. On the other hand, GRADES-hydroDL adopts a long short-term memory (LSTM) model to estimate runoff that is subsequently routed through the river network using RAPID, whereas the GEOGloWS ECMWF Streamflow Service produces river discharge predictions by routing ERA5-derived runoff using the Muskingum method. Meteorological forcing is also one of the important factors affecting the accuracy of snowmelt-driven surface runoff. The meteorological forcing data used in GloFAS v4.0 are ERA5, and in GRADES-hydroDL are MSWEP v2.8 precipitation and other meteorological variables from ERA5. Unlike that, the GEOGloWS ECMWF Streamflow Service takes surface runoff from ERA5 as input to its river routing model. Both the ERA5 Reanalysis and MSWEP are also gridded datasets. Hafizi and Sorman [54] also investigated the performance of gridded precipitation datasets, including ERA5 and MSWEP, and then reported that ERA5 exhibited generally poor performance during the winter season in the Karasu Basin, Türkiye. On the flip side, the MSWEP gridded dataset estimated mean daily precipitation during the winter season close to ground observations. The accurate modeling of snowmelt-driven runoff is a key process that depends mainly on the accurate representation of snow accumulation during winter. This is because accurately capturing the amount and timing of precipitation during the winter season is also essential for simulating snow accumulation and snowmelt-driven runoff during the following spring.

The LSTM model underlying GRADES-hydroDL was trained using daily observed discharge datasets from 4215 gauged basins worldwide. Moreover, GRADES-hydroDL employs a Grid LSTM model instead of the VIC land surface model to generate gridded runoff fields at a spatial resolution of 0.25°. Despite the absence of river discharge observations from gauging stations in Türkiye, the superior performance of GRADES-hydroDL in the Aras Basin may be attributed to its Grid LSTM structure and its MSWEP-based precipitation forcing.

The second important finding of this study is that the GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 simulate considerably low discharge estimates compared with the observations at some gauging stations. The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0, unlike the other model-based river discharge products evaluated in this study, are grid-based products. For GloFAS v4.0 and EFAS v5.0, the river discharge time series are extracted from the grid cell in which the gauging station is located. Initially, it was thought that the issue could be potentially related to the location of the gauging station very close to a neighboring grid cell. Therefore, the time series of the eight neighboring grid cells adjacent to each gauging station cell were re-extracted from the gridded datasets (GloFAS v4.0 and EFAS v5.0). For the GloFAS v4.0, it was observed that only the discharge values extracted from the grid cell located on the southwest of the D24A097-Tımar Ç. Tımar gauging station showed markedly better agreement with the river discharge records than the station grid cell. Accordingly, the discharge time series extracted from the grid cell located on the southwest of the D24A097 were considered in all analyses. It was also considered that this issue could be related to the elevations of the gauging stations. The D24A040, D24A042, D24A058, D24A083, D24A090, and D24A097 gauging stations are located at elevations of 1780, 1597, 1760, 1636, 1745, and 1801 m, respectively (see Table 1). However, no clear correlation could be found. As the next step, it was decided to examine the relationship between the gauging stations and the model-based river discharge products, which were the best-performing estimates. Regarding the performance distribution among stations, GRADES-hydroDL and the GEOGloWS ECMWF Streamflow Service show their best per-

formance at D24A042-Hasankale Ç. Eğirmez. The D24A042 gauging station has the largest drainage area. The drainage area of the D24A042 is 1470 km^2. The gauging stations with the smallest drainage areas in the Aras Basin are D24A058-Kışla D. Çayırdüzü (77.10 km^2) and D24A040-Masat D. Bulkasım (95.50 km^2), respectively. EFAS Historical v5.0 produces considerably low discharges at D24A040-Masat D. Bulkasım, D24A042-Hasankale Ç. Eğirmez, and D24A058-Kışla D. Çayırdüzü, whereas GloFAS v4.0 simulates similar discharges with EFAS Historical at the D24A058-Kışla D. Çayırdüzü gauging station. This suggests that a possible relationship may exist between the drainage area of the gauging stations and the performance of the river discharge products, although it cannot be statistically confirmed with six stations. As mentioned previously, Bartholmes et al. [52] reported a similar issue for the EFAS dataset. In their study, grid cells with an upstream drainage area of less than 4000 km^2 were excluded for this reason. The reason why the GloFAS v4.0 Hydrological Reanalysis and EFAS v5.0 underestimate river discharges more than the observations may also be partly attributed to the calibration procedures adopted for these two model-based products. Gauging stations with drainage areas below 500 km^2 are excluded from the calibration process of GloFAS v4.0 [43]. A similar calibration criterion is also applied to EFAS v5.0. In this case, gauging stations with drainage areas below 150 km^2 are not considered for the calibration process [44]. The drainage areas of the gauging stations in this study range from 77.10 to 1470 km^2. Among the gauging stations, only D24A042-Hasankale Ç. Eğirmez has a drainage area exceeding the 500 km^2 threshold used for the GloFAS v4.0 calibration procedure. The Aras Basin is located near the south-eastern boundary of the EFAS v5.0 modeling domain. Additionally, EFAS v5.0 relies on ground-based observations to generate the gridded meteorological forcing data. Consequently, the performance of EFAS v5.0 may also depend on the density of the meteorological observation network [51]. The low density of the meteorological gauging station network in the region where the Aras Basin is located may have affected the performance of EFAS v5.0.

To examine the dependence of dataset performance on basin characteristics, the gauging stations were grouped by drainage area into three classes: below 100 km^2 (D24A040-Masat D. Bulkasım, D24A058-Kışla D. Çayırdüzü, and D24A097-Tımar Ç. Tımar), 100–500 km^2 (D24A083-Süngütaş D. Çermik and D24A090-Karahan Ç. Taşdere), and above 1000 km^2 (D24A042-Hasankale Ç. Eğirmez). The Spearman correlation between the KGE values and the drainage area is positive for all four products ($\rho = 0.26$–0.60) but not statistically significant ($p > 0.2$), and the pattern is not consistent among the products: GRADES-hydroDL and GEOGloWS reach their highest KGE at the largest catchment (D24A042-Hasankale Ç. Eğirmez), whereas EFAS v5.0 performs best at D24A083 (486 km^2) and GloFAS v4.0 at D24A097 (98 km^2). A grouping by elevation was not meaningful because the station elevations fall within a narrow range (1597–1801 m), and all six stations are located on rural, high-altitude areas with similar land cover, so a grouping by surface type would not differentiate the stations. Therefore, the results of this study should not be extrapolated to the eastern and transboundary parts of the basin, where no validation was possible; in these ungauged areas, particularly along the main stem of the Aras River with much larger drainage areas, the performance of the products may differ from that found here.

Despite underestimating river discharge, GloFAS v4.0 was able to capture hydrological drought periods, particularly at the 12-month timescale (SDI-12). The fundamental reason is that the Streamflow Drought Index (SDI) is a standardized index. This finding suggests that GloFAS v4.0 may be effective in capturing hydrological drought periods in terms of the SDI-12. However, this does not imply that the GloFAS Hydrological Reanalysis can accurately reflect the severity of hydrological droughts. Indeed, SDI values standardized

using the observed streamflow distribution diminish the apparent skill of GloFAS v4.0 in identifying hydrological drought periods (Table 8).

Based on the results obtained at the six gauging stations, the suitability of the four products for typical application scenarios can be summarized as follows. GRADES-hydroDL is the only product that can be recommended for high-flow event detection, hydrological drought monitoring, and monthly to seasonal water balance applications in the study area. The GloFAS v4.0 Hydrological Reanalysis captures the timing of high flows and of dry and wet periods, but its magnitudes cannot be used without correction. The GEOGloWS ECMWF Streamflow Service may be used for monthly to seasonal volumes only after bias correction, and EFAS Historical v5.0 is not recommended for any of the scenarios considered. None of the products is adequate for the spring snowmelt period or for low-flow conditions without correction. It should be noted that the suitability for flood early warning could not be assessed in this study, since the products were evaluated as retrospective simulations rather than forecasts, and no comparison with recorded flood events was made.

5. Conclusions

The Aras Basin, located in the northeastern part of Türkiye, is a transboundary and mountainous basin. This makes it difficult to establish a robust river discharge monitoring network in the basin. Therefore, the number of river discharge gauging stations providing long-term and gap-free discharge observations in the Aras Basin is very limited. The gap-free observations covering the period from 1 October 2012 to 30 September 2025 could only be obtained from six stations operated by the General Directorate of State Hydraulic Works. That period is the longest period for which the longest-gap-free river discharge records could be obtained. Five of six gauging stations are clustered in the southwestern part of the basin. In other words, there is not a sufficient discharge monitoring network in the basin that provides both gap-free and long-term discharge records while also adequately representing the entire basin. This reveals the urgent need for the river discharge record or datasets which present river discharge time series in the Aras Basin.

With this motivation, the performance of three global streamflow datasets—the GloFAS v4.0 Hydrological Reanalysis, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service—and one regional dataset—EFAS Historical v5.0—in the Aras Basin were evaluated from many hydrological aspects using river discharge observations obtained from six gauging stations in the Aras Basin. Thus, the applicability of the four model-based river discharge products in the Aras Basin was tested. This approach is a major step toward overcoming the lack of river discharge observations and a sufficient monitoring network in the Aras Basin.

As a result of the evaluations in this study, the following main conclusions were reached.

1. The water-year-based time-series analysis (covering 1 October 2012–30 September 2025) and the monthly mean and seasonal mean outputs show that GRADES-hydroDL is the best-performing model-based river discharge product among the evaluated products for the Aras Basin.
2. During the spring (March–April–May), the snowmelt-driven runoff increases the river discharge observations. However, it is clear that all four discharge products, namely, the GloFAS v4.0 Hydrological Reanalysis, EFAS Historical v5.0, GRADES-hydroDL, and the GEOGloWS ECMWF Streamflow Service, have limited capacities to reflect the snowmelt-driven surface runoff process.
3. The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 were found to simulate considerably low river discharge estimates against discharge observations at some gauging stations. This issue is more evident at some gauging stations with small drainage areas. This suggests a possible relationship between the drainage area

of the gauging stations and the performance of the model-based discharge products; however, this relationship could not be statistically confirmed with six stations.

4. The low-, moderate-, and high-flow conditions were derived from the flow duration curves for each model-based discharge product and discharge observation. The results indicate that river discharge estimates exhibit performance differences in capturing the low-, moderate-, and high-flow conditions. In detecting the high-flow conditions, the GRADES-hydroDL dataset was quite successful in determining the Q5 at the D24A042-Hasankale Ç. Eğirmez gauging station. On the other hand, the GloFAS v4.0 Reanalysis exhibits poor performance in identifying the low-flow conditions, namely, Q95, at most of the gauging stations.
5. The ability of the four model-based discharge products to detect high-flow events was also tested. In this context, GRADES-hydroDL again shows the highest capability in identifying high-flow events among the evaluated discharge estimates.
6. In the hydrological drought analysis, GRADES-hydroDL is quite successful in identifying the two of the three main drought periods: 2014–2015, 2016–2018, but not 2021–2024, whereas the GloFAS v4.0 predicts the two main hydrological drought periods of 2014–2015 and 2021–2024. This means that all four model-based river discharge products exhibit performance fluctuations in terms of drought characterization.

Overall, GRADES-hydroDL outperforms the other three evaluated discharge products in terms of multiple hydrological aspects. EFAS Historical v5.0, as a regional river discharge product, shows the weakest performance compared to the other global products. Thus, GRADES-hydroDL stands out as an alternative discharge dataset for the Aras Basin. Finally, the findings of this study are valid for the upstream part of the Aras Basin within Türkiye, where the six gauging stations with drainage areas of 77–1470 km^2 and elevations of 1597–1801 m are located. Their extension to the eastern and transboundary ungauged parts of the basin requires additional validation.

Funding: This research received no external funding.

Data Availability Statement: The four model-based river discharge datasets analyzed in this study are publicly available. The GloFAS v4.0 Hydrological Reanalysis and EFAS Historical v5.0 can be obtained from the Copernicus Emergency Management Service through the Early Warning Data Store (https://ewds.climate.copernicus.eu (accessed on 14 November 2026)). The GRADES-hydroDL dataset is available at https://www.reachhydro.org/home/records/grades-hydrodl (accessed on 14 November 2026), and the GEOGloWS ECMWF Streamflow Service data at https://geoglows.ecmwf.int (accessed on 14 November 2026). The observed daily mean river discharge records were obtained from the General Directorate of State Hydraulic Works of Türkiye. The General Directorate of State Hydraulic Works streamflow yearbooks covering the period up to 30 September 2021 are publicly available at https://www.dsi.gov.tr/Sayfa/Detay/744 (access on 14 November 2026), whereas records after this date are provided on a paid basis and therefore cannot be redistributed by the author. The derived time series and performance metrics generated during this study are available from the author upon reasonable request.

Acknowledgments: The author thanks the General Directorate of State Hydraulic Works for providing the observed river discharge records used in this study.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CSI	Critical Success Index
ECMWF	European Centre for Medium-Range Weather Forecasts
EFAS	European Flood Awareness System

FAR	False Alarm Ratio
FDC	Flow Duration Curve
GEOGloWS	Group on Earth Observations Global Water Sustainability
GHM	Global Hydrological Model
GloFAS	Global Flood Awareness System
GRADES	Global Reach-level A priori Discharge Estimates for SWOT
KGE	Kling–Gupta Efficiency
LSM	Land Surface Model
LSTM	Long Short-Term Memory
NSE	Nash–Sutcliffe Efficiency
PBIAS	Percent Bias
POD	Probability of Detection
RAPID	Routing Application for Parallel computatIon of Discharge
RMSE	Root Mean Square Error
SDI	Streamflow Drought Index
SWOT	Surface Water and Ocean Topography
VIC	Variable Infiltration Capacity

References

1. Xu, S.; Sun, H.; Tang, L.; Sun, X. A Global Benchmark of the Vector-Based Routing Model MizuRoute: Similarities and Divergent Patterns in Simulated River Discharge. *Water* **2026**, *18*, 485. [CrossRef]
2. Tu, T.; Wang, J.; Zhao, G.; Zhao, T.; Dong, X. Scaling from global to regional river flow with global hydrological models: Choice matters. *J. Hydrol.* **2024**, *633*, 130960. [CrossRef]
3. Alfieri, L.; Lorini, V.; Hirpa, F.A.; Harrigan, S.; Zsoter, E.; Prudhomme, C.; Salamon, P. A global streamflow reanalysis for 1980–2018. *J. Hydrol. X* **2020**, *6*, 100049. [CrossRef] [PubMed]
4. Burek, P.; Satoh, Y.; Kahil, T.; Tang, T.; Greve, P.; Smilovic, M.; Guillaumot, L.; Zhao, F.; Wada, Y. Development of the Community Water Model (CWatM v1.04)—A high-resolution hydrological model for global and regional assessment of integrated water resources management. *Geosci. Model Dev.* **2020**, *13*, 3267–3298. [CrossRef]
5. Comte, L.; Olden, J.D.; Tedesco, P.A.; Ruhi, A.; Giam, X. Climate and land-use changes interact to drive long-term reorganization of riverine fish communities globally. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2011639118. [CrossRef] [PubMed]
6. Döll, P.; Kaspar, F.; Lehner, B. A global hydrological model for deriving water availability indicators: Model tuning and validation. *J. Hydrol.* **2003**, *270*, 105–134. [CrossRef]
7. Lin, P.; Pan, M.; Beck, H.E.; Yang, Y.; Yamazaki, D.; Frasson, R.; David, C.H.; Durand, M.; Pavelsky, T.M.; Allen, G.H.; et al. Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water Resour. Res.* **2019**, *55*, 6499–6516. [CrossRef] [PubMed]
8. Sutanudjaja, E.H.; van Beek, R.; Wanders, N.; Wada, Y.; Bosmans, J.H.C.; Drost, N.; van der Ent, R.J.; de Graaf, I.E.M.; Hoch, J.M.; de Jong, K.; et al. PCR-GLOBWB 2: A 5 arcmin global hydrological and water resources model. *Geosci. Model Dev.* **2018**, *11*, 2429–2453. [CrossRef]
9. Wing, O.E.J.; Quinn, N.; Bates, P.D.; Neal, J.C.; Smith, A.M.; Sampson, C.C.; Coxon, G.; Yamazaki, D.; Sutanudjaja, E.H.; Alfieri, L. Toward Global Stochastic River Flood Modeling. *Water Resour. Res.* **2020**, *56*, e2020WR027692. [CrossRef]
10. Arthington, A.H.; Bhaduri, A.; Bunn, S.E.; Jackson, S.E.; Tharme, R.E.; Tickner, D.; Young, B.; Acreman, M.; Baker, N.; Capon, S.; et al. The Brisbane Declaration and Global Action Agenda on Environmental Flows (2018). *Front. Environ. Sci.* **2018**, *6*, 45. [CrossRef]
11. Chuphal, D.S.; Mishra, V. Hydrological model-based streamflow reconstruction for Indian sub-continental river basins, 1951–2021. *Sci. Data* **2023**, *10*, 717. [CrossRef] [PubMed]
12. Sikder, M.S.; David, C.H.; Allen, G.H.; Qiao, X.; Nelson, E.J.; Matin, M.A. Evaluation of Available Global Runoff Datasets Through a River Model in Support of Transboundary Water Management in South and Southeast Asia. *Front. Environ. Sci.* **2019**, *7*, 171. [CrossRef]
13. Yang, Y.; Feng, D.; Beck, H.E.; Hu, W.; Abbas, A.; Sengupta, A.; Delle Monache, L.; Hartman, R.; Lin, P.; Shen, C.; et al. Global Daily Discharge Estimation Based on Grid Long Short-Term Memory (LSTM) Model and River Routing. *Water Resour. Res.* **2025**, *61*, e2024WR039764. [CrossRef]

14. Liu, Y.; Yang, Z.; Lin, P.; Zheng, Z.; Xie, S. Comparison and evaluation of multiple land surface products for the water budget in the Yellow River Basin. *J. Hydrol.* **2020**, *584*, 124534. [CrossRef]
15. Koukoula, M.; Nikolopoulos, E.I.; Dokou, Z.; Anagnostou, E.N. Evaluation of Global Water Resources Reanalysis Products in the Upper Blue Nile River Basin. *J. Hydrometeorol.* **2020**, *21*, 935–952. [CrossRef]
16. Lakew, H.B.; Moges, S.A.; Anagnostou, E.N.; Nikolopoulos, E.I.; Asfaw, D.H. Evaluation of Global Water Resources Reanalysis Runoff Products for Local Water Resources Applications: Case Study-Upper Blue Nile Basin of Ethiopia. *Water Resour. Manag.* **2020**, *34*, 2157–2177. [CrossRef]
17. Miao, Y.; Wang, A. Evaluation of Routed-Runoff from Land Surface Models and Reanalyses Using Observed Streamflow in Chinese River Basins. *J. Meteorol. Res.* **2020**, *34*, 73–87. [CrossRef]
18. Harrigan, S.; Zsoter, E.; Alfieri, L.; Prudhomme, C.; Salamon, P.; Wetterhall, F.; Barnard, C.; Cloke, H.; Pappenberger, F. GloFAS-ERA5 operational global river discharge reanalysis 1979-present. *Earth Syst. Sci. Data* **2020**, *12*, 2043–2060. [CrossRef]
19. Hales, R.C.; Nelson, E.J.; Souffront, M.; Gutierrez, A.L.; Prudhomme, C.; Kopp, S.; Ames, D.P.; Williams, G.P.; Jones, N.L. Advancing global hydrologic modeling with the GEOGloWS ECMWF streamflow service. *J. Flood Risk Manag.* **2025**, *18*, e12859. [CrossRef]
20. Ghimire, G.R.; Hansen, C.; Gangrade, S.; Kao, S.-C.; Thornton, P.E.; Singh, D. Insights From Dayflow: A Historical Streamflow Reanalysis Dataset for the Conterminous United States. *Water Resour. Res.* **2023**, *59*, e2022WR032312. [CrossRef]
21. Bain, R.L.; Shaw, M.J.; Geheran, M.P.; Tavakoly, A.A.; Wahl, M.D.; Zsoter, E. Intercomparison of global ERA reanalysis products for streamflow simulations at the high-resolution continental scale. *J. Hydrol.* **2023**, *616*, 128624. [CrossRef]
22. Swain, S.; Mishra, S.K.; Pandey, A.; Srivastava, P.K.; Nandi, S. Characterization and assessment of hydrological droughts using GloFAS streamflow data for the Narmada River Basin, India. *Environ. Sci. Pollut. Res.* **2024**, *31*, 54281–54294. [CrossRef] [PubMed]
23. Kim, B.-H.; Hossain, S.; Kam, J. High-resolution gridded streamflow data for Ganges-Brahmaputra-Meghna River Basins in Bangladesh (1951–2023). *Sci. Data* **2025**, *12*, 673. [CrossRef] [PubMed]
24. Parmar, J.; Karmakar, S. A multi-scenario framework for quantifying flood hazard and exposure accounting for runoff-driven uncertainty in global flood models. *Sci. Total Environ.* **2025**, *1000*, 180381. [CrossRef] [PubMed]
25. Patidar, R.; Pingale, S.M.; Khare, D.; Dayal, D.; Choudhary, S. Evaluating the GloFAS discharge reanalysis product for hydrological drought monitoring in India: Performance across temporal scales and flow regimes. *Hydrol. Sci. J.* **2026**, *71*, 1759–1778. [CrossRef]
26. Doğan, E. Türkiye'de Küresel Akım Veri Setlerinin Tutarlılığının Değerlendirilmesi [Evaluation of the Consistency of Global Streamflow Datasets in Türkiye]. Master's Thesis, Department of Geography, Institute of Social Sciences, Bursa Uludağ University, Bursa, Türkiye, 2024. (In Turkish)
27. Ministry of Agriculture and Forestry, General Directorate of Water Management. *Aras Havzası Kuraklık Yönetim Planı Stratejik Çevresel Değerlendirme Kapsam Belirleme Raporu [Aras Basin Drought Management Plan: Strategic Environmental Assessment Scoping Report]; Aras ve Çoruh Havzaları Kuraklık Yönetim Planlarının Hazırlanması Projesi [Project for the Preparation of Drought Management Plans for the Aras and Çoruh Basins]*; SU PEK Proje ve Müşavirlik A.Ş.: Ankara, Türkiye, 2023. (In Turkish)
28. Ministry of Agriculture and Forestry, General Directorate of Water Management. *Aras Havzası Taşkın Yönetim Planı Yönetici Özeti [Aras Basin Flood Management Plan: Executive Summary]*; SU PEK Proje ve Müşavirlik A.Ş.: Ankara, Türkiye, 2019. (In Turkish)
29. Sorman, A.A.; Tas, E.; Dogan, Y.O. Comparison of hydrological models in upper Aras Basin. *Pamukkale Univ. Muh. Bilim. Derg.* **2020**, *26*, 1015–1022. [CrossRef]
30. Copernicus Emergency Management Service. River Discharge and Related Historical Data from the Global Flood Awareness System. Available online: https://ewds.climate.copernicus.eu/datasets/cems-glofas-historical (accessed on 14 November 2026).
31. Copernicus Emergency Management Service. GloFAS v4 Hydrological Model. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/340755422/GloFAS+v4+hydrological+model (accessed on 14 November 2026).
32. Copernicus Emergency Management Service. River Discharge and Related Historical Data from the European Flood Awareness System. Available online: https://ewds.climate.copernicus.eu/datasets/efas-historical (accessed on 14 November 2026).
33. Copernicus Emergency Management Service. EFAS v5.0-Hydrological Model. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/340784752/EFAS+v5.0+-+Hydrological+Model (accessed on 14 November 2026).
34. ReachHydro. GRADES-hydroDL. Available online: https://www.reachhydro.org/home/records/grades-hydrodl (accessed on 14 November 2026).
35. GEOGLOWS. River Forecast System: Model Formulation, Hydrography and Retrospective Data. Available online: https://training.geoglows.org (accessed on 14 November 2026).
36. GEOGloWS. GEOGloWS ECMWF Streamflow Service. Available online: https://geoglows.ecmwf.int/ (accessed on 14 November 2026).
37. Cao, B.; Arduini, G.; Zsoter, E. Brief communication: Improving ERA5-Land soil temperature in permafrost regions using an optimized multi-layer snow scheme. *Cryosphere* **2022**, *16*, 2701–2708. [CrossRef]

38. Copernicus Emergency Management Service. CEMS-Flood Meteorological Observation Maps. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/333791599/CEMS-Flood+meteorological+observation+maps (accessed on 14 November 2026).
39. European Commission Joint Research Centre. LISFLOOD Model Documentation: Snow Module. Available online: https://ec-jrc.github.io/lisflood-model/2_04_stdLISFLOOD_snowmelt/ (accessed on 14 November 2026).
40. Lees, T.; Reece, S.; Kratzert, F.; Klotz, D.; Gauch, M.; De Bruijn, J.; Kumar Sahu, R.; Greve, P.; Slater, L.; Dadson, S.J. Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 3079–3101. [CrossRef]
41. Copernicus Emergency Management Service. CEMS-Flood LISFLOOD Model Setup. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/333803476/CEMS-Flood+LISFLOOD+model+setup (accessed on 14 November 2026).
42. Chancay, J.E.; Sánchez-Lozano, J.L.; Valencia, B.G.; Trujillo-Vela, M.G.; Nelson, E.J.; Hales, R.C.; Gutiérrez, A.L. Enhancing GEOGLOWS River Forecast System with a High-Resolution Pre-Processing Approach for Runoff Bias Correction. *Hydrology* **2026**, *13*, 128. [CrossRef]
43. Copernicus Emergency Management Service. GloFAS v4 Calibration Data. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/340755424/GloFAS+v4+calibration+data (accessed on 14 November 2026).
44. Copernicus Emergency Management Service. EFAS v5.0-Calibration Data. Available online: https://confluence.ecmwf.int/spaces/CEMS/pages/340784754/EFAS+v5.0+-+Calibration+Data (accessed on 14 November 2026).
45. Nourani, V.; Kheirieh, S.; Kantoush, S.A.; Huang, J.J. Bias correction of Global Flood Awareness System (GloFAS) data for multi-station river flow prediction by ensemble modelling. *Eng. Appl. Comput. Fluid Mech.* **2026**, *20*, 2665857. [CrossRef]
46. Krause, P.; Boyle, D.P.; Bäse, F. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **2005**, *5*, 89–97. [CrossRef]
47. Gebrechorkos, S.H.; Leyland, J.; Dadson, S.J.; Cohen, S.; Slater, L.; Wortmann, M.; Ashworth, P.J.; Bennett, G.L.; Boothroyd, R.J.; Cloke, H.L.; et al. Global-scale evaluation of precipitation datasets for hydrological modelling. *Hydrol. Earth Syst. Sci.* **2024**, *28*, 3099–3118. [CrossRef]
48. Cinkus, G.; Mazzilli, N.; Jourde, H.; Wunsch, A.; Liesch, T.; Ravbar, N.; Chen, Z.; Goldscheider, N. When best is the enemy of good-Critical evaluation of performance criteria in hydrological models. *Hydrol. Earth Syst. Sci.* **2023**, *27*, 2397–2411. [CrossRef]
49. Dessie, M.; Verhoest, N.E.C.; Pauwels, V.R.N.; Admasu, T.; Poesen, J.; Adgo, E.; Deckers, J.; Nyssen, J. Analyzing runoff processes through conceptual hydrological modeling in the Upper Blue Nile Basin, Ethiopia. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 5149–5167. [CrossRef]
50. Nalbantis, I. Evaluation of a hydrological drought index. *Eur. Water* **2008**, *23/24*, 67–77.
51. Thielen, J.; Bartholmes, J.; Ramos, M.-H.; de Roo, A. The European Flood Alert System-Part 1: Concept and development. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 125–140. [CrossRef]
52. Bartholmes, J.C.; Thielen, J.; Ramos, M.-H.; Gentilini, S. The European Flood Alert System EFAS-Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 141–153. [CrossRef]
53. Hossain, S.; Cloke, H.L.; Ficchì, A.; Gupta, H.; Speight, L.; Hassan, A.; Stephens, E.M. A decision-led evaluation approach for flood forecasting system developments: An application to the Global Flood Awareness System in Bangladesh. *J. Flood Risk Manag.* **2025**, *18*, e12959. [CrossRef]
54. Hafizi, H.; Sorman, A.A. Assessment of 13 Gridded Precipitation Datasets for Hydrological Modeling in a Mountainous Basin. *Atmosphere* **2022**, *13*, 143. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.