Journal of Hydrology 669 (2026) 135122

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# Assessing and optimizing high-resolution global river streamflow estimates with triple collocation analysis

Mingze Sun[a,b], Natthachet Tangdamrongsub[c,*], Yu Sun[a,b,**], Jianzhi Dong[d], Edwin Sutanudjaja[e], Mikhail Smilovic[f,g]

[a] Key Lab of Spatial Data Mining and Information Sharing of Ministry of Education, Fuzhou University, Fujian, China
[b] National Engineering Research Center of Geospatial Information Technology, Fuzhou University, Fujian, China
[c] Water Engineering and Management, Faculty of Civil and Environmental Engineering, Asian Institute of Technology, Pathum Thani, Thailand
[d] Institute of Surface-Earth System Science, Tianjin University, Tianjin, China
[e] Department of Physical Geography, Utrecht University, Utrecht, the Netherlands
[f] Water Security Research Group, International Institute for Applied Systems Analysis (IIASA), Laxenburg 2031, Austria
[g] ETH Zürich, Chair of Hydrology and Water Resources Management, 8093 Zurich, Switzerland

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Gerald Augusto Corzo, Associate Editor

*Keywords:*
Global streamflow
Hydrology model
Triple collocation

ABSTRACT

Accurate river discharge data are essential for effective water resource management, flood forecasting, drought mitigation, and ecological sustainability. However, the global river gauge network remains sparse, outdated, and often inaccessible, particularly in developing or remote regions, limiting our ability to monitor and manage hydrological systems. In response, global hydrological models have emerged as critical tools for estimating streamflow, yet their accuracy and applicability require systematic evaluation. This study assesses the performance of three global hydrological models—Community Water Model (CWatM), PCRaster Global Water Balance (PCR-GLOBWB), and H08—at both high (5 arcmin) and low (30 arcmin) spatial resolutions. Observed discharge data from 1,707 Global Runoff Data Centre (GRDC) stations and 62 stations in Thailand were used for evaluation. Additionally, two data fusion methods—Triple Collocation (TC) and simple averaging—were employed to enhance streamflow simulations by integrating outputs from the three models. Results show that CWatM, particularly at high resolution, consistently outperforms PCR-GLOBWB and H08, while all models demonstrate stronger performance in Europe than in other regions. Both TC and averaging methods improve simulation accuracy compared to individual models, with TC offering the highest overall performance, especially when using H08 as the reference system. Resolution analysis reveals that high-resolution models yield more accurate estimates in most regions. This study highlights the value of multi-model integration and resolution optimization in enhancing global streamflow simulation. The findings provide practical guidance for model selection and data fusion strategies, particularly in improving hydrological assessments for ungauged or poorly monitored regions.

## 1. Introduction

River discharge simulation lies at the heart of hydrology and plays a vital role in understanding and managing global water resources, as it provides essential information for future planning and decision-making. Accurate simulation of river discharge enhances the prediction and management of water-related disasters such as floods and droughts, thereby helping to mitigate their impacts and safeguard lives and property (Sahani et al., 2019; Tan et al., 2020; Kron et al., 2021). It also facilitates the rational allocation and management of water resources across agriculture, industry, and domestic sectors (Wang et al., 2016; Sun et al., 2018a; Ye et al., 2018; Liu et al., 2020). However, the global streamflow observation network remains uneven, with many regions, particularly in developing countries and remote areas, lacking consistent or sufficient in situ measurements (Akpoti et al., 2024). The scarcity of streamflow gauge observations hampers accurate assessments of water availability, detection of hydrological extremes, and validation of model simulations, posing significant challenges for water resource management and disaster preparedness (Sheffield et al., 2018). Hydrological modelling, therefore, plays a particularly critical role by

* Corresponding author.
** Corresponding author at: Key Lab of Spatial Data Mining and Information Sharing of Ministry of Education, Fuzhou University, Fujian, China.
*E-mail addresses:* natthachet@ait.asia (N. Tangdamrongsub), jade.yusun@outlook.com (Y. Sun).

https://doi.org/10.1016/j.jhydrol.2026.135122
Received 22 July 2025; Received in revised form 27 January 2026; Accepted 9 February 2026
Available online 15 February 2026

providing spatially and temporally distributed streamflow estimates that can supplement direct measurements, especially in data-sparse regions (McCabe et al., 2017; Wada et al., 2017).

In response to these challenges, numerous global hydrology models have been developed to extend streamflow information into regions with limited observational data. However, their accuracy often correlates with proximity to regions with dense observation networks, such as data centers, raising concerns about their performance in underrepresented areas (e.g., Tangdamrongsub, 2023). A systematic evaluation of model-generated streamflow estimates is thus essential to assess their reliability before applying them in target regions. Although in situ discharge observations provide the most reliable reference for evaluation, establishing and maintaining dense monitoring networks remains costly and infeasible in many regions (Bonnema and Hossain, 2019; Jin and Jin, 2020). Consequently, large-scale hydrological models are widely used to provide spatially and temporally continuous streamflow information for forecasting and water-resources management (Krabbenhoft et al., 2022).

Several widely used global hydrological models include the Community Water Model (CWatM) (Burek et al., 2020), PCRaster Global Water Balance (PCR-GLOBWB) (Sutanudjaja et al., 2018), and H08 (Hanasaki et al., 2018). These models have been applied to assess water supply–demand balance, predict hydrological responses under climate change scenarios, and support water resource management (Guillaumot et al., 2022; Raghav and Eldho, 2023; Tangdamrongsub et al., 2017). By simulating surface water, groundwater, and human water use processes, they serve as practical tools for analyzing water scarcity, floods, droughts, and other extreme events (Ntona et al., 2022). However, the limited availability of high-quality observational data, especially in developing regions, poses challenges for proper model calibration and validation (Miao et al., 2022). To address this, modelers often calibrate models using limited regional observations, improving performance where data are available while reducing computational costs. This practice, however, raises concerns about the models' transferability and accuracy in regions where they have not been calibrated (Yoshida et al., 2022). Evaluating multiple models under diverse hydrological conditions is, therefore, essential to quantify uncertainties, identify the most reliable models, and inform their use in data-sparse areas (Horton et al., 2022). However, systematic comparisons among CWatM, PCR-GLOBWB, and H08 remain scarce, particularly across diverse spatial resolutions and globally distributed river basins. Comprehensive evaluations of this kind are crucial for identifying the relative strengths and limitations of each model and for assessing their transferability under varying hydrological and geographic conditions.

Low-resolution models, commonly used in large-scale hydrological simulations, tend to smooth spatial variability in surface parameters such as river networks, soil types, and vegetation. While this approach reduces computational demands, it often oversimplifies surface conditions—particularly in basins with complex terrain or rapidly changing land use—leading to potential errors in flow simulation (Wang et al., 2019; Guo et al., 2020). In contrast, high-resolution models offer finer representations of surface features, improving the accuracy of hydrological simulations by capturing spatial heterogeneity more effectively (Becker et al., 2019; Yang et al., 2021; Belvederesi et al., 2022). These models can also better represent hydrological connectivity and river network structure, which may improve the spatial realism of runoff generation and thus benefit streamflow estimation (Jin et al., 2021; Loritz et al., 2021). The growing need for local-scale accuracy, together with advancements in computing power, has driven the trend toward higher-resolution hydrological modeling (Hakala Assendelft et al., 2020; Sidle, 2021). Incorporating high-resolution inputs—such as detailed topography and precipitation data—potentially further enhances model accuracy (Huang et al., 2019; Rocha et al., 2020).

Uncertainties in streamflow simulations arise from model structure, parameterization, and meteorological forcing, and they are often amplified in data-scarce regions where calibration and validation are limited (Her et al., 2019; Moges et al., 2021; Herrera et al., 2022). To mitigate these uncertainties, a practical strategy is to combine multiple model-based estimates, thereby leveraging complementary strengths and reducing both random variability and systematic biases in individual simulations (Nourani et al., 2021; Sheikh and Coulibaly, 2024; Singh et al., 2025). Such multi-model fusion has been increasingly adopted in hydrological applications because no single model is consistently optimal across diverse hydroclimatic conditions and data environments. However, many fusion approaches still rely on empirical tuning or subjective weighting choices, which can limit transferability and interpretability when applied at large scales (Di Curzio et al., 2021; Gonzales-Inca et al., 2022; Jia et al., 2023). This motivates statistically grounded fusion methods that can estimate relative error characteristics and derive weights from the datasets themselves, providing a more transparent basis for global streamflow optimization.

Against this background, a promising method is Triple Collocation (TC), which provides a statistically robust approach to estimate and minimize error without requiring knowledge of the true values (Dodet et al., 2022). TC has been successfully applied to optimize hydrometeorological variables such as wind speed (Stoffelen, 1998; Tan et al., 2024), precipitation (Alemohammad et al., 2015; Lyu et al., 2021), and soil moisture (Gruber et al., 2016; Ming et al., 2022), but has yet to be applied to global streamflow estimation. This study represents the first application of Triple Collocation (TC) for global-scale river discharge evaluation across multiple hydrological models and spatial resolutions. TC has been shown to outperform conventional methods in various environmental applications, reducing uncertainty and improving accuracy in multi-source data fusion (Duan et al., 2021). Incorporating TC into streamflow simulations may, therefore, enhance the reliability of predictions by leveraging complementary strengths across multiple models.

In this study, we first evaluate the accuracy of global streamflow estimates from three hydrological models, CWatM, PCR-GLOBWB, and H08, with two different spatial resolutions, 30 arcmin (~50 km) and 5 arcmin (~10 km), and assess the benefits of optimizing these estimates using both simple averaging and TC. Model accuracy is evaluated using streamflow observations from the Global Runoff Data Centre (GRDC) for global assessments and from the Royal Irrigation Department (RID) of Thailand for regional evaluation between 2000 and 2022. We aim to address the following questions: (1) How accurate are current global hydrological models in simulating river discharge? (2) What are the benefits of using high-resolution models for streamflow estimation? (3) Can optimization techniques such as TC significantly improve global streamflow predictions? Building on these questions, we make three contributions beyond the current state of the art: (i) the first global-scale application of TC for multi-model streamflow optimization, extending TC for hydrological variables; (ii) we quantify how much moving from 30 arcmin to 5 arcmin improves streamflow simulation, and whether such gains remain after applying TC-based optimization; (iii) we demonstrate the practical usefulness of the approach by testing it against two complementary observation datasets (global and local). By systematically evaluating model performance across different resolutions and optimization methods, this study contributes to improving global streamflow estimates, informs model selection under varying hydrological conditions, and advances data fusion approaches in hydrological modelling.

## 2. Study areas

The uneven distribution of GRDC stations leads to significant differences in the density of observation stations across regions. To comprehensively evaluate the performance of hydrological models, it is necessary to consider not only global simulation accuracy but also to conduct an in-depth analysis at the regional scale. In this study, we focused specifically on regions with high densities of GRDC stations, which were North America, South America, Europe, South Africa, and

Australia (Fig. 1a).

We selected Thailand as a case study to comprehensively evaluate the performance of different hydrological models at the local scale (Fig. 1b). Thailand was chosen due to its complex hydrological regime and the need for accurate streamflow simulations to support effective water resource management (Khaikham and James, 2019; Noy et al., 2021). A detailed local evaluation is critical because global-scale model assessments may not fully represent region-specific hydrological characteristics. Thus, evaluating model accuracy at the national level can significantly enhance Thailand's hydrological modeling strategies and decision-making processes.

## 3. Global streamflow simulations

The CWatM is a large-scale, grid-based hydrological model designed to simulate the full terrestrial water cycle under both natural and human-influenced conditions. It calculates potential evapotranspiration using the Penman–Monteith method by default and distinguishes rainfall and snowfall based on elevation and temperature, with snowmelt estimated through the degree-day approach. The model separately computes water balance components for multiple land cover types, including forest, irrigated, and urban areas, using the Xinanjiang method for infiltration and the Van Genuchten formulation for soil moisture redistribution in three soil layers. Groundwater is represented as a linear reservoir, and surface runoff is routed using the kinematic wave approximation of the Saint-Venant equations. A key strength of CWatM lies in its explicit representation of reservoir regulation and human water use. Reservoirs and lakes, derived from the HydroLakes database, are dynamically integrated into the river routing network and categorized as global or local systems depending on their spatial scale. Reservoir operations follow the LISFLOOD scheme, using defined thresholds for conservation, normal, and flood storage to simulate controlled releases that balance flood protection, ecological flow, and hydropower demand. On the demand side, CWatM quantifies water withdrawals, consumption, and return flows for irrigation, domestic, industrial, and livestock sectors, linking daily water demand to climatic, agricultural, and socioeconomic factors. Irrigation demand is computed dynamically based on crop water requirements, distinguishing between paddy and non-paddy systems. Groundwater and surface water are used jointly to meet demand, constrained by environmental flow limits and local water availability. Through this integrated treatment of reservoir operation and human water management, CWatM provides a comprehensive framework for assessing the coupled natural–human water cycle and evaluating water resources sustainability under changing climate and socio-economic conditions. Through its modular structure and coupled treatment of both physical hydrological and human water management processes, CWatM enables integrated assessments of water availability, demand, and infrastructure operation at global to regional scales, and can be linked with other IIASA models to support analyses of the water–energy–food nexus and hydro-economic decision-making under climate change (Burek et al., 2020; Guillaumot et al., 2022).

PCR-GLOBWB is a global, grid-based hydrological and water resources model designed to simulate the full terrestrial water balance and anthropogenic water management (Sutanudjaja et al., 2018). It represents surface and subsurface processes through snow and interception stores, as well as soil and groundwater ones, capturing various hydrological fluxes, such as precipitation, evaporation and transpiration, infiltration, snowmelt, and runoff. The latter is routed through river networks, as well as lakes and reservoirs. The model incorporates more than 6,000 reservoirs from the Global Reservoir and Dam Database (GRanD), each operated using a simple rule-based scheme that considers construction year and storage capacity (see also Section 2.4.1 of Steyaert et al., 2025). Human water use is fully integrated within the hydrological cycle, estimating daily water demand, withdrawal, consumption, and return flows for irrigation, industrial, domestic, and livestock sectors. Water can be abstracted dynamically from surface water, groundwater (renewable and non-renewable), or desalinated sources based on local availability and pumping capacity, with sectoral allocation and feedback to river discharge. Irrigation demand is computed following FAO guidelines and MIRCA2000 crop calendars, distinguishing between paddy and non-paddy systems, while industrial and domestic demands vary seasonally with socioeconomic indicators such as GDP and population. Through this integrated coupling of reservoir operation, water withdrawal, and return flow, PCR-GLOBWB provides a comprehensive framework for assessing human–water interactions and the impacts of regulation and consumption on global hydrological dynamics.

The H08 model is an integrated global hydrological model designed to simulate both natural and human-influenced water cycles. It represents the terrestrial water balance by incorporating groundwater recharge and abstraction, surface and subsurface runoff, evapotranspiration, and river routing, while explicitly simulating human interventions such as aqueduct water transfer, desalination, and return

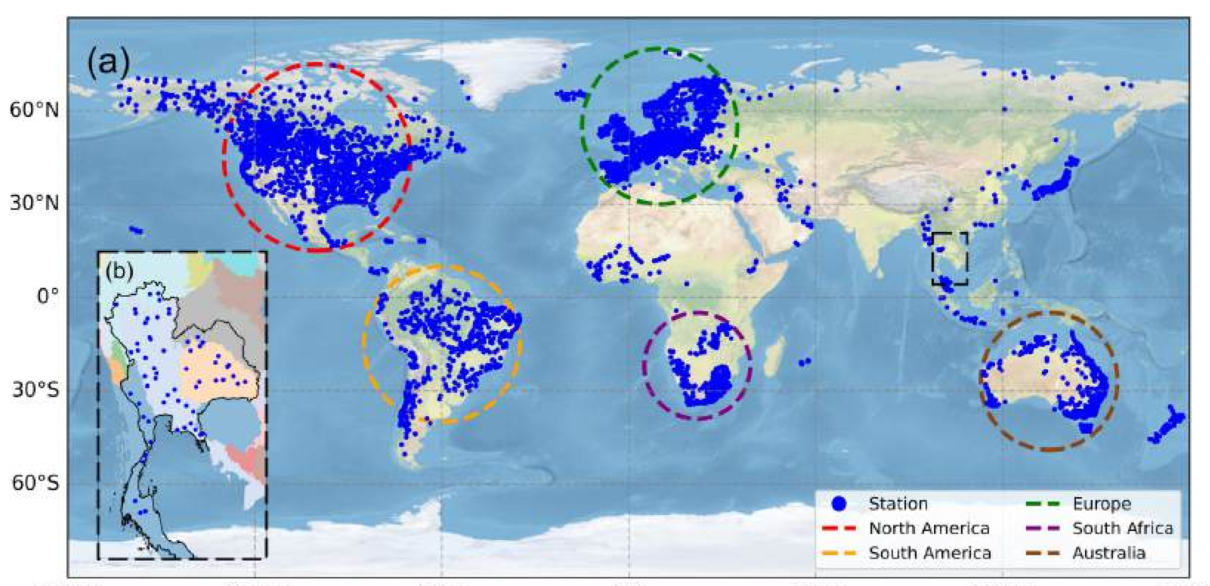


**Fig. 1.** (a) Global distribution of GRDC gauge stations (blue dots) that provide data between 2000 and −2022. The focused study areas are shown in dotted circles with different colors. (b) The inset (at the bottom-left corner) shows the distribution of the measurement sites in Thailand (not included in GRDC).

flows. Reservoirs are divided into global and local types—global reservoirs, typically large dams, directly regulate river discharge along the main channels, whereas local reservoirs act within tributaries to store runoff and mitigate water shortages. Human water use in H08 is comprehensively modeled through dynamically coupled modules that estimate daily demands from irrigation, domestic, industrial, and livestock sectors. Water withdrawal is sourced from surface water, renewable and nonrenewable groundwater, and desalinated water, considering physical constraints such as groundwater pumping capacity and aqueduct network distribution. The model also simulates return flow and delivery losses, reflecting the fraction of water re-entering river systems and losses through evaporation or inefficiency. Aqueduct transfer schemes enable inter-basin water redistribution, while the desalination module accounts for freshwater production in coastal and arid regions. Together, these coupled human–natural processes allow H08 to assess how infrastructure operations, water allocation, and groundwater dependence jointly shape global and regional water availability under changing climatic and socioeconomic conditions (Hanasaki et al., 2018).

Three open-source global hydrological models—CWatM, PCR-GLOBWB, and H08—were employed to simulate monthly river streamflow from January 2000 to December 2022 (Guillaumot et al., 2022; Sutanudjaja et al., 2018; Hanasaki et al., 2018). Each model was implemented at two spatial resolutions: 30 arcmin (~50 km) and 5 arcmin (~10 km), hereafter referred to as ModelName_30arcmin and ModelName_05arcmin, respectively. Simulations were performed using daily meteorological forcing, with all three models requiring similar meteorological inputs (see Sect. 4.3). The use of consistent input datasets helps ensure that performance differences across models are attributable to model-specific formulations and parameterizations rather than discrepancies in meteorological data. A key advantage of these models is the availability of predefined or derivable land surface and hydrological parameters, which allows for immediate application. This characteristic is particularly beneficial in large-scale studies, where detailed land surface data may be challenging to obtain. Although higher-resolution (1 km) model versions have been developed for regional applications (Huang et al., 2019; Fersch et al., 2020), they remain computationally intensive and are currently unsuited for global-scale assessments. The extended descriptions of CWatM, PCR-GLOBWB, and H08 can be found in Guillaumot et al. (2022), Sutanudjaja et al. (2018), and Hanasaki et al. (2018), respectively.

CWatM, PCR-GLOBWB, and H08 are executed with given model parameters and default settings to simulate monthly streamflow between 2000 and 2022. The model is spun up for 460 years to ensure that all storages reach an equilibrium state. The spin-up process reruns the model between 2000 and 2022 20 times, using the state variables of the last time step of each revolution as the initial state of the next simulation. The simulation output is the monthly global streamflow of two spatial resolutions (30 and 5 arcmin), which are evaluated against ground observations. Monthly evaluations help reduce the influence of short-term fluctuations and measurement noise, offering clearer insights into long-term trends and seasonal patterns. Additionally, monthly observational records tend to be more complete and reliable than daily records, minimizing data gaps and uncertainty. Many hydrological models are also calibrated/validated at monthly scales due to the temporal resolution of input datasets and model structure (Guan et al., 2020; Lerat et al., 2021).

## 4. Data

### 4.1. Global runoff data Centre (GRDC)

GRDC is an international data repository dedicated to collecting, storing, and disseminating river discharge data from over 10,000 stations across major river basins worldwide (Pavelsky et al., 2014). The GRDC ensures high data quality through rigorous quality control procedures aimed at maintaining accuracy and consistency. For this study, two criteria were applied to select stations from the GRDC dataset: (1) the data must cover the simulation period from 2000 to 2022, and (2) each station must contain at least 10 consecutive years of discharge records without significant data gaps. After applying these filters, a total of 1,707 GRDC stations were selected for evaluation (Fig. 2).

### 4.2. Thailand gauge data

River discharge observations in Thailand were obtained from the RID, which operates over 100 telemetry stations across the country's major river basins. The data are recorded at a daily time step and were aggregated to a monthly resolution to match the temporal scale of the model simulations. Consistent with the GRDC data selection criteria, only stations with continuous data from 2000 to 2022 were included in the analysis. After applying this criterion, 62 valid RID stations were retained for evaluation.

### 4.3. Forcing data

The meteorological forcing datasets required for model simulations include precipitation, air temperature, humidity, wind speed, and radiation. These forcing variables were obtained from the ERA5 and ERA5-Land reanalysis products (Hersbach et al., 2020; Muñoz-Sabater et al., 2021), which provide hourly data with a latency of approximately one week. To align with the model time step, the data were temporally aggregated from hourly to daily resolution. The ERA5 product, with a native spatial resolution of approximately 25 km, was spatially averaged to 30 arcminutes and used as input for the 30 arcmin model simulations. In contrast, ERA5-Land provides meteorological fields at a higher resolution of approximately 10 km. These data were regridded to match the corresponding model grid before being used to run the 5 arcmin resolution simulations.

## 5. Methods

### 5.1. Triple Collocation approach

TC combines three different hydrological model outputs and generates an improved streamflow estimate. The combined streamflow can be expressed as a weighted sum of three model-based simulations:

$$U = weight_x \cdot x_{t,l} + weight_y \cdot y_{t,l} + weight_z \cdot z_{t,l} \quad (1)$$

where $x_{t,l}$, $y_{t,l}$, and $z_{t,l}$ represent the streamflow results at the time $t$ and location $l$ from CWatM, PCR-GLOBWB, and H08, respectively. In the triple collocation framework, $z_{t,l}$ is treated as the reference, and other datasets are compared relative to it. Where $U$ represents the combined streamflow estimated by the TC method, $x_{t,l}$, $y_{t,l}$, and $z_{t,l}$ represent simulated streamflow results from the three hydrological models, respectively, and $weight_x$, $weight_y$, and $weight_z$ represent their corresponding optimal weights derived from the TC. To assess TC performance relative to the selected reference, each model (H08, PCR-GLOBWB, and CWatM) is used in turn as the reference when applying the TC method. Unless otherwise specified, H08 is used as the default reference system for comparative analysis.

It should be noted that TC relies on two key assumptions: (i) the three model-based discharge estimates are approximately linearly related to each other, and (ii) the corresponding errors are mutually independent across the three systems. However, in global hydrological modelling, shared meteorological forcing and structural similarities may induce inter-model error dependence, so TC-derived weights and performance gains should be interpreted as conditional on the prevailing error structure rather than as a universal correction. Sensitivity analysis can be performed by repeating the TC analysis, with each model serving as

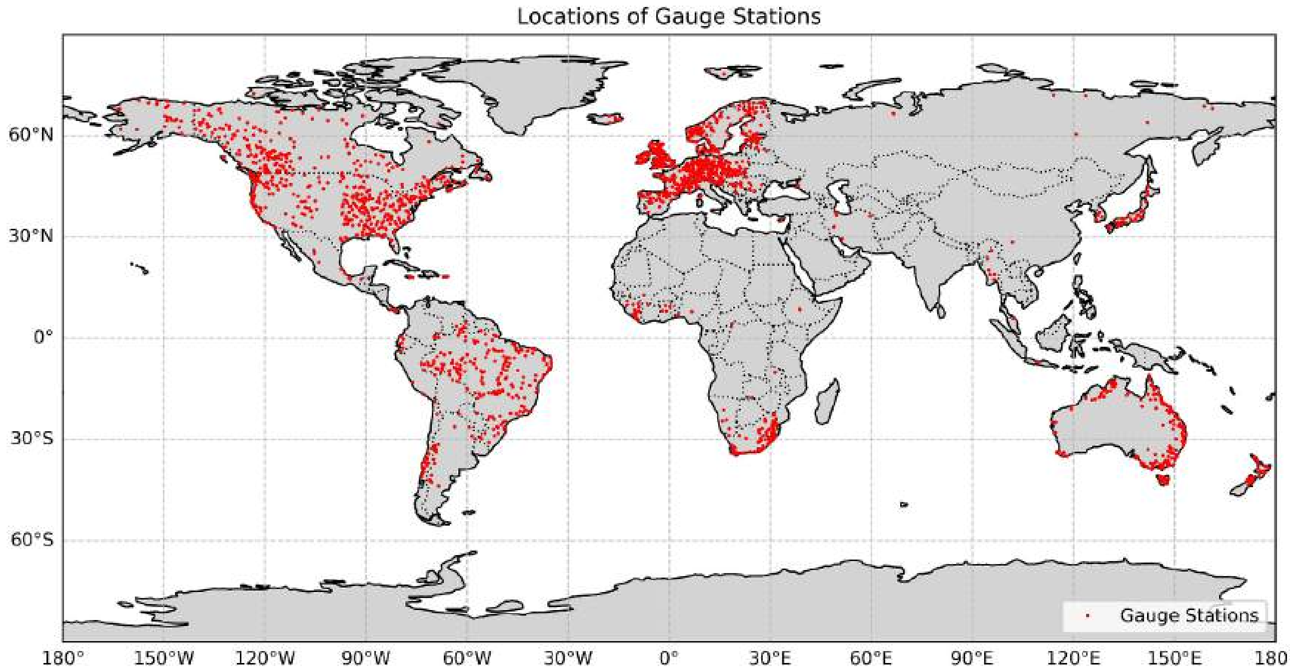


**Fig. 2.** The distribution of GRDC stations that are used in this study.

the reference system in turn to evaluate the robustness of the results to reference-system selection (see Sect. 6.3).

Consistent with the zero-mean error requirement, for each model, the mean streamflow over the period 2000–2022 is first computed and removed to ensure that all three datasets have a zero-mean before combination. Removing the mean ensures that the combination process focuses on deviations from the average behavior of each model, rather than absolute magnitudes (Dodet et al., 2022; Vogelzang and Stoffelen, 2021):

$$x_{diff} = x_{t,l} - \overline{x} \tag{2}$$

$$y_{diff} = y_{t,l} - \overline{y} \tag{3}$$

$$z_{diff} = z_{t,l} - \overline{z} \tag{4}$$

where $\overline{x}$, $\overline{y}$, and $\overline{z}$ represent the means of $x$, $y$, and $z$.

In the covariance and sensitivity calculations (Eq. (5)–(11), the reference system z (H08) is used to determine the relative signal sensitivity and noise level of the other two systems. The covariance matrix is calculated based on the differences between the three model results:

$$\sigma_{hat} = \begin{bmatrix} \frac{\sum\left(x_{diff}^2\right)}{n_1} & \frac{\sum\left(x_{diff} \bullet y_{diff}\right)}{n_1} & \frac{\sum\left(x_{diff} \bullet z_{diff}\right)}{n_1} \\ \frac{\sum\left(x_{diff} \bullet y_{diff}\right)}{n_1} & \frac{\sum\left(y_{diff}^2\right)}{n_1} & \frac{\sum\left(y_{diff} \bullet z_{diff}\right)}{n_1} \\ \frac{\sum\left(x_{diff} \bullet z_{diff}\right)}{n_1} & \frac{\sum\left(y_{diff} \bullet z_{diff}\right)}{n_1} & \frac{\sum\left(z_{diff}^2\right)}{n_1} \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} & \sigma_{XZ} \\ \sigma_{YX} & \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{ZX} & \sigma_{ZY} & \sigma_Z^2 \end{bmatrix} \tag{5}$$

where $n_1$ represent the number of effective samples. The diagonal elements of the covariance matrix represent the variances of the models, and the off-diagonal elements represent covariances between them.

After obtaining the covariance matrix, the next step is to estimate the error variances of each measurement system. The error variances are used to measure the magnitude of noise in each system, that is, the degree to which they deviate from the true signal (Tsamalis, 2022). The formula is as follows:

$$\sigma_{\varepsilon_X}^2 = \sigma_X^2 - sensitivity_x \tag{6}$$

$$\sigma_{\varepsilon_Y}^2 = \sigma_Y^2 - sensitivity_y \tag{7}$$

$$\sigma_{\varepsilon_Z}^2 = \sigma_Z^2 - sensitivity_z \tag{8}$$

where $\sigma_X^2$, $\sigma_Y^2$, $\sigma_Z^2$ represent the diagonal elements of the estimated covariance matrix $\sigma_{hat}$ (see Eq. (5). Next, the sensitivity is used to measure the ability of each measurement system to respond to changes in the true signal. The stronger the system's response to the true signal, the greater its ability to capture the true signal (Wang et al., 2022). The sensitivity is computed as follows:

$$sensitivity_x = \frac{\sigma_{XY} \bullet \sigma_{XZ}}{\sigma_{YZ}} \tag{9}$$

$$sensitivity_y = \frac{\sigma_{YX}\sigma_{YZ}}{\sigma_{XZ}} \tag{10}$$

$$sensitivity_z = \frac{\sigma_{ZX} \bullet \sigma_{ZY}}{\sigma_{XY}} \tag{11}$$

Finally, the weights of each model are calculated using the standard deviation, and normalization is applied to ensure that the sum of the weights across different systems equals 1 (Badar et al., 2024; Chen et al., 2022; Yilmaz and Crow, 2014). This allows the combined result to reflect each system's contribution:

$$iweight_x = \frac{1}{\sigma_{\varepsilon_X}^2} \tag{12}$$

$$iweight_y = \frac{1}{\sigma_{\varepsilon_Y}^2} \tag{13}$$

$$iweight_z = \frac{1}{\sigma_{\varepsilon_Z}^2} \tag{14}$$

$$weight_x = \frac{iweight_x}{iweight_x + iweight_y + iweight_z} = \frac{\sigma^2_{\varepsilon_Y}\sigma^2_{\varepsilon_Z}}{\sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Y} + \sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Z} + \sigma^2_{\varepsilon_Y}\sigma^2_{\varepsilon_Z}} \tag{15}$$

$$weight_y = \frac{iweight_y}{iweight_x + iweight_y + iweight_z} = \frac{\bullet\sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Z}}{\sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Y} + \sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Z} + \sigma^2_{\varepsilon_Y}\sigma^2_{\varepsilon_Z}} \tag{16}$$

$$weight_z = \frac{iweight_z}{iweight_x + iweight_y + iweight_z} = \frac{\sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Y}}{\sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Y} + \sigma^2_{\varepsilon_X}\sigma^2_{\varepsilon_Z} + \sigma^2_{\varepsilon_Y}\sigma^2_{\varepsilon_Z}} \tag{17}$$

### 5.2. Performance assessment

In this study, we employed five statistical metrics—Kling-Gupta Efficiency (KGE), Nash–Sutcliffe Efficiency (NSE), correlation coefficient (R2), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE) to comprehensively evaluate model performance. KGE, as a composite metric, integrates correlation, variability, and bias components, providing an overall measure of model accuracy and consistency (Gupta et al., 2009; Lee and Choi, 2022; dos Reis et al., 2022). NSE and $R^2$ focus on the degree of agreement and correlation between observed and simulated values (Duc and Sawada, 2023; Melsen et al., 2025; Williams, 2025). RMSE and NRMSE quantify the magnitude of average deviation (Bringeland and Fotopoulos, 2024; Hodson, 2022; Wu et al., 2024). Collectively, these complementary metrics offer a robust and multi-dimensional assessment of the accuracy, reliability, and systematic bias of streamflow simulations.

The KGE is computed as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{18}$$

where $r$ measures the linear relationship between observed and simulated data, $\beta$ measures the ratio of the mean of the simulated values to the mean of the observed values, and $\gamma$ measures the ratio of the standard deviations of the simulated and observed data. As recommended by Knoben et al. (2019), a KGE score of $-0.4$ is used as a benchmark, which is the score of the observational mean; i.e., KGE $> -0.4$ represents acceptable performance.

The NSE is computed as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_{sim,i} - Q_{obs,i}\right)^2}{\sum_{i=1}^{n}\left(Q_{obs,i} - \overline{Q_{obs,i}}\right)^2} \tag{19}$$

where $Q_{sim,i}$ and $Q_{obs,i}$ denote the simulated and observed discharge at the time step $i$, $\overline{Q_{obs}}$ is the mean observed discharge, and $n$ is the number of observations measures how well the simulated time series matches the observed dynamics. NSE ranges from $-\infty$ to 1, with 1 indicating a perfect fit, 0 representing performance equal to the mean observed flow, and values below 0 indicating poorer performance. Generally, NSE $>$ 0 denotes acceptable model performance.

The $R^2$ is computed as follows:

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(Q_{sim,i} - \overline{Q_{sim}}\right)\left(Q_{obs,i} - \overline{Q_{obs}}\right)\right]^2}{\sum_{i=1}^{n}\left(Q_{sim,i} - Q_{sim}\right)^2\sum_{i=1}^{n}\left(Q_{obs,i} - Q_{obs}\right)^2} \tag{20}$$

where $R^2$ ranges from 0 to 1, with higher values indicating better model performance.

The RMSE is computed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(Q_{sim,i} - Q_{obs,i}\right)^2} \tag{21}$$

where $n$ is the number of data points used in the calculation. RMSE ranges from 0 to $+\infty$, with smaller values indicating higher model accuracy and better agreement between simulated and observed flows.

The NRMSE is computed as follows:

$$NRMSE = \frac{RMSE}{\overline{Q_{obs}}} \tag{22}$$

where NRMSE ranges from 0 to $+\infty$, with smaller values indicating better model performance and higher consistency across basins with different flow magnitudes.

### 5.3. Selecting gauges for comparison

Streamflow simulation time series were extracted for each in situ gauge location. Due to differences in spatial resolution, a single model grid cell may encompass multiple gauge stations (Lehner et al., 2022). Additionally, some gauge stations are located on both main and tributary streams, whereas model simulations may only represent one of these—typically the main tributary. As a result, directly matching model outputs to gauge locations can lead to duplicated or inconsistent time series. Therefore, a data screening process was implemented to ensure the uniqueness and validity of comparisons (Gao et al., 2021; Wheater et al., 2022). To address this, we first identified model grid cells that contained multiple gauge stations. For each of these stations, the KGE value was computed by comparing observed and simulated streamflow. The station with the highest KGE value within each grid cell was retained for further analysis. We also visually inspected the time series of each selected gauge to ensure that the observed and simulated data captured consistent hydrological signals. Furthermore, the same set of gauge stations was used to evaluate both the 30 arcmin and 5 arcmin model versions. This approach serves three key purposes: (1) it increases the likelihood that the selected gauge station corresponds to the tributary represented in the model, (2) it ensures fairness and consistency in comparisons among different model resolutions, and (3) it enables a direct performance assessment of the 5 arcmin model relative to the 30 arcmin model. The gauges that best align with the 30 arcmin simulation serve as benchmarks; improvements in KGE when using the 5 arcmin model indicate enhanced performance due to higher spatial resolution.

## 6. Results

### 6.1. Global evaluation

A comprehensive assessment was carried out to evaluate performance differences across models and spatial resolutions, focusing on both the accuracy of streamflow simulations and their spatial distribution. The spatial evaluation highlights the non-uniformity in simulation accuracy across different models and regions (Fig. 3). Models tend to perform well in areas such as North America, Europe, and parts of Asia, where dense observational data support accurate representation of basin hydrodynamics (Ali et al., 2023; Roure et al., 2009). In contrast, model performance is substantially lower in data-scarce regions, such as Southern Africa and Australia, where KGE values frequently fall below $-0.4$, because these stations are in drier climates, where precisely estimating streamflow is inherently challenging (Akhtar et al., 2022; Andrade et al., 2024; Quichimbo et al., 2023; Wanzala et al., 2022). Overall, high-resolution models consistently outperform their low-resolution counterparts (e.g., Fig. 3a vs. Fig. 3b). Furthermore, integrating multiple datasets through methods such as simple averaging and TC enhances simulation accuracy beyond that of any individual model. This is particularly evident in the improved spatial consistency and elevated performance metrics observed in the combined model outputs (Fig. 3g–j).

A global comparison of model performance reveals pronounced differences in accuracy, consistency, and sensitivity to spatial resolution across all models and metrics (Fig. 4, Figs. S1 – S3, and Table S1). Among the individual models, CWatM_05arcmin consistently exhibits the best overall performance, achieving the highest global KGE (0.49 [0.47, 0.50]), NSE (0.21 [0.19, 0.22]), and the lowest NRMSE (0.17 [0.17,

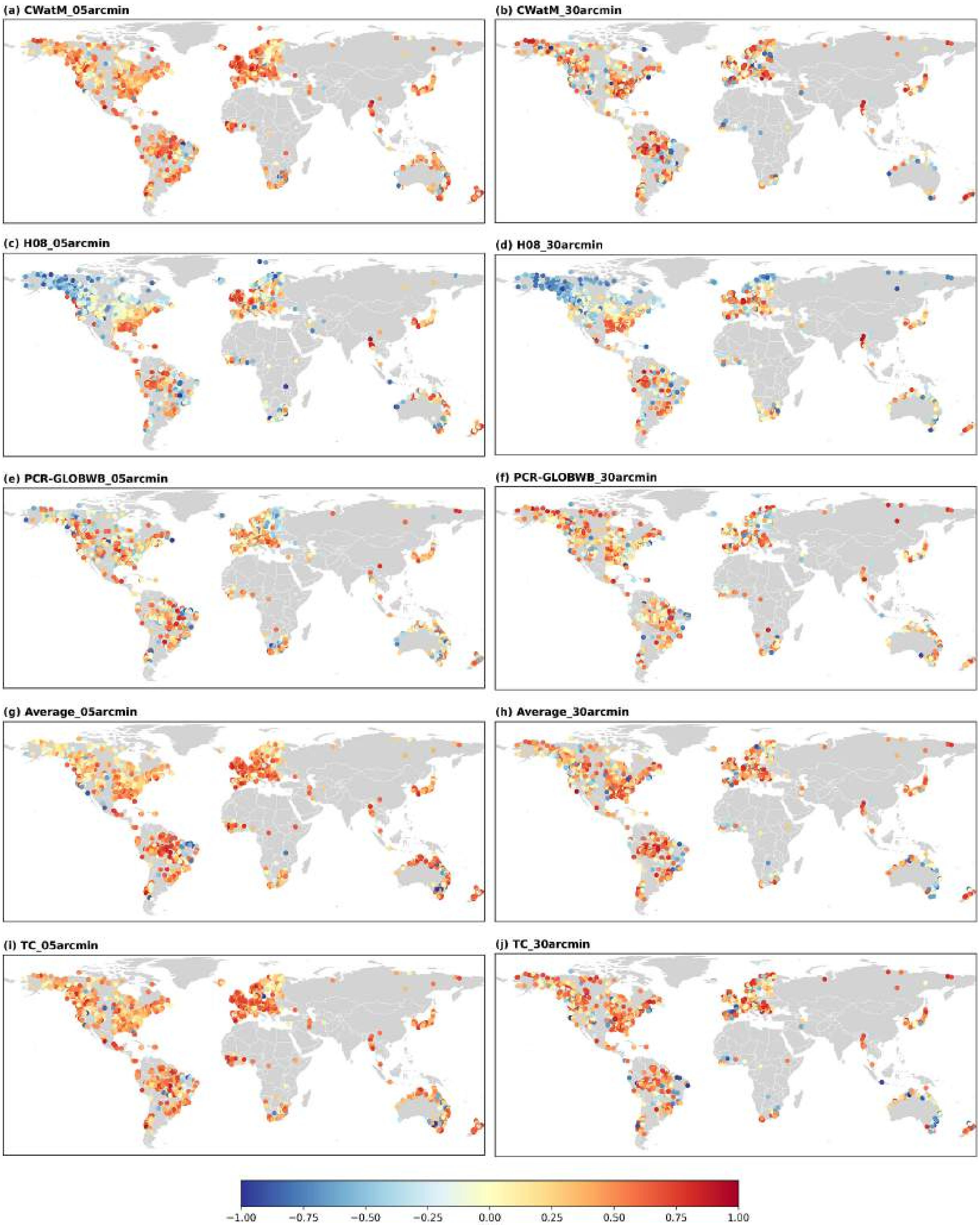


**Fig. 3.** Global distribution of KGE values under different simulation scenarios, (a) CWatM (05 arcmin), (b) CWatM (30 arcmin), (c) H08 (05 arcmin), (d) H08 (30 arcmin), (e) PCR-GLOBWB (05 arcmin), (F) PCR-GLOBWB (30 arcmin), (g) Average (05 arcmin), (h) Average (30 arcmin), (i) TC (05 arcmin), (j) TC (30 arcmin).

0.18]) with a corresponding RMSE of 239.06 m3/s. This confirms that finer spatial resolution substantially enhances CWatM's ability to capture hydrological variability. However, when downgraded to 30 arcmin resolution, its KGE decreases to 0.38 (and RMSE nearly doubles to 465.02 $m^3/s$), demonstrating the loss of accuracy associated with coarser spatial representation. By contrast, H08 and PCR-GLOBWB show relatively low and inconsistent accuracy across all metrics, particularly the H08 model, where NSE values drop below zero, and $R^2$ decreases markedly. This reflects the limited ability of these models to capture flow variability and extremes at coarse resolutions.

The TC method demonstrates consistently superior performance across both spatial resolutions, outperforming not only all individual models but also the Average method (Fig. 4, Figs. S1 – S3, and Table S1). At the finer 05 arcmin resolution, TC_05arcmin achieves the highest

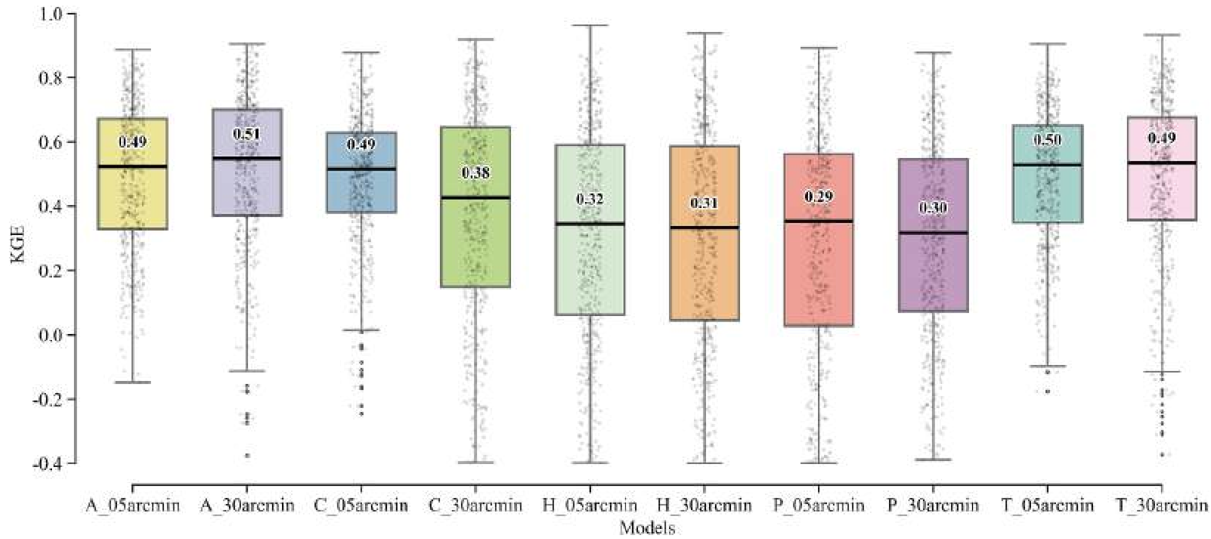


**Fig. 4.** Boxplots of model performance metrics (KGE) (see NRMSE, NSE, and $R^2$ in supplements) for different hydrological models and spatial resolutions in the global. The label format is Model_Resolution, where A, C, H, P, and T represent Average, CWatM, H08, PCR-GLOBWB, and TC.

overall accuracy, with KGE = 0.50 [0.48, 0.51], NSE = 0.26 [0.24, 0.28], and $R^2$ = 0.59, surpassing CWatM_05arcmin (KGE = 0.49) and Average_05arcmin (KGE = 0.49). The narrower interquartile range and higher median in the boxplots indicate that TC not only improves the mean performance but also enhances stability across global stations. Even at the coarser 30 arcmin resolution, TC_30arcmin maintains strong predictive capability (KGE = 0.49, NSE = 0.23), exceeding the best single model (CWatM_30arcmin, KGE = 0.38) by approximately 29%, and remaining close to the Average_30arcmin (KGE = 0.51). The consistent performance of TC across both resolutions highlights its strength in effectively synthesizing information from multiple model outputs, thereby compensating for the limitations and biases of individual models. Importantly, TC's ability to maintain relatively high KGE values at coarser resolutions suggests that integrating multiple models can significantly improve global streamflow simulations. This is achieved by capturing a broader range of hydrological processes while mitigating individual model errors. Similarly, the simple averaging method also enhances performance, indicating that even basic multi-data combination approaches can contribute to more accurate streamflow simulations.

### 6.2. Regional analysis

To assess regional variations in model performance and the effectiveness of data fusion methods, a comparative analysis was conducted across five major global regions: North America, South America, Europe, Southern Africa, and Australia (Fig. 5). Among the individual models, CWatM_05arcmin consistently achieves the highest KGE and NSE values

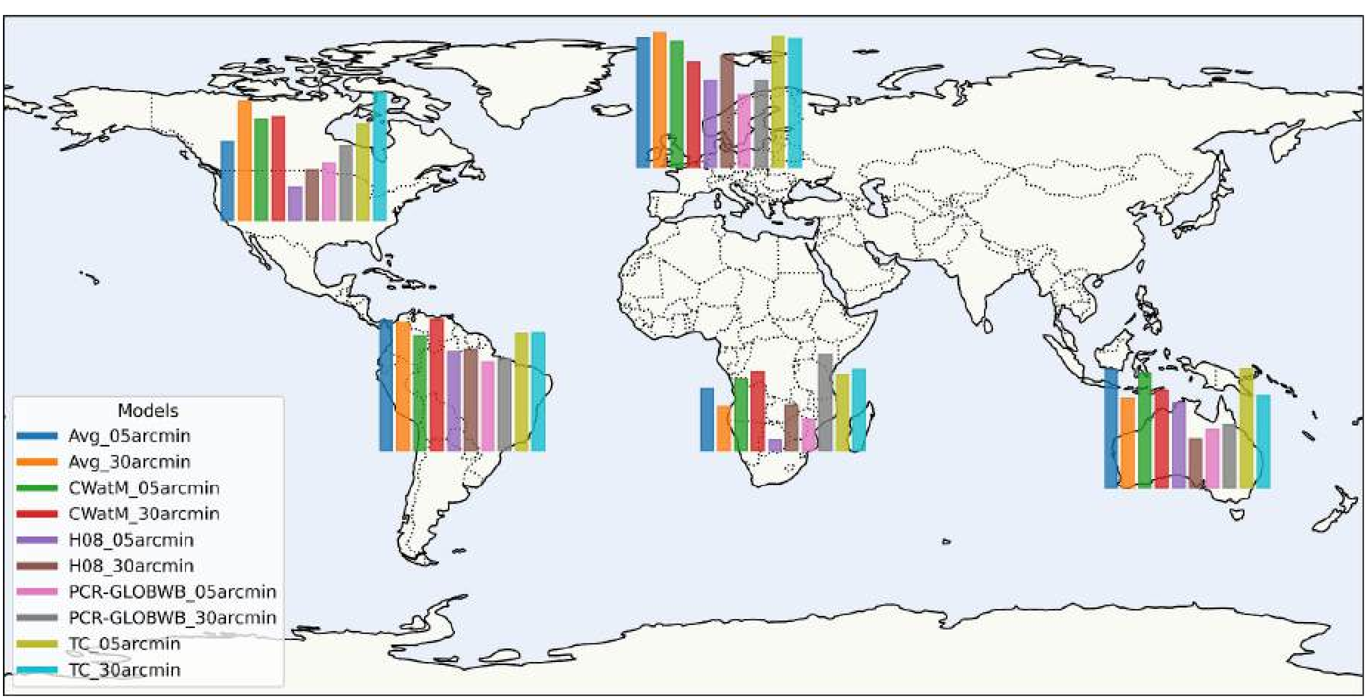


**Fig. 5.** Regional-average KGE computed from different model simulation scenarios. The bar chart represents the normalized KGE values, allowing relative performance comparisons by bar length.

across most regions, particularly in North America and Europe, underscoring the importance of fine spatial resolution for capturing basin-scale hydrological variability. However, its performance decreases at coarser resolution (CWatM_30arcmin), especially in Southern Africa and Australia, where sparse observations and complex hydrological regimes amplify model uncertainty. The PCR-GLOBWB model performs moderately well, with relatively balanced performance across regions, but without clear dominance in any particular area. In contrast, H08 shows the weakest and most unstable results.

Regarding data fusion approaches, the TC method demonstrates the most consistent and spatially superior performance across all regions and metrics (Fig. 5 and Fig. S4). At fine resolution, TC_05arcmin achieves the highest or near-highest values of KGE (as well as NSE and $R^2$) in nearly all regions, confirming its ability to effectively integrate multi-model information and capture hydrological variability. Since all statistical metrics yield similar conclusions, this section focuses on KGE (Fig. 5), with other metrics presented in Fig. S4. In Europe, it shows outstanding performance (KGE = 0.57) with the lowest errors (NRMSE = 0.16, RMSE = 50 $m^3/s$), indicating both high accuracy and stability. In North America, TC_05arcmin (KGE = 0.40) also surpasses Average_05arcmin (KGE = 0.37), suggesting better process representation under strong flow variability. Similarly, in Australia, TC_05arcmin yields relatively strong results across all metrics (KGE = 0.62), highlighting its robustness in arid and highly seasonal regions. At 30 arcmin resolution, although the accuracy of all models declines, TC_30arcmin still maintains the highest regional averages in most metrics—particularly in Southern Africa (KGE = 0.62) —demonstrating stronger resilience to spatial degradation than the Average method. In North America, TC_30arcmin achieves a higher correlation (KGE = 0.50), showing that it effectively balances bias correction and dynamic consistency even in large basins with complex hydrology.

To further assess model performance in capturing large-scale hydrological dynamics, three major river basins were selected for individual analysis (Fig. 6 and Table S3). This regional-scale evaluation provides insights into model behavior under distinct hydrological conditions influenced by basin size, climatic variability, and human interventions such as flow regulation—factors that may not be fully captured in global-scale assessments (Munier and Decharme, 2022; Tu et al., 2024). The Danube, Mississippi, and Rhine River basins were chosen due to their geographic diversity, contrasting watershed characteristics, and significant hydrological importance. The Danube River, which spans multiple climate zones and landforms, plays a critical role in water resource management across Central and Eastern Europe (Tian et al., 2025; Wojkowski et al., 2024). The Mississippi River basin, encompassing a wide range of climatic regions, is a major water source for the United States and has experienced frequent extreme hydrological events, making it a suitable case for evaluating model robustness (Dommo et al., 2024). The Rhine River basin, a transboundary river system, is not only vital for water management across several European countries but also holds significant economic importance for the region (Jalink and Dieperink, 2024; Krapesch et al., 2024). Collectively, these three rivers represent a broad spectrum of climatic, geomorphological, and anthropogenic influences, providing a comprehensive framework for evaluating hydrological model performance under varying environmental and management conditions.

High-resolution models, particularly TC_05arcmin, demonstrate

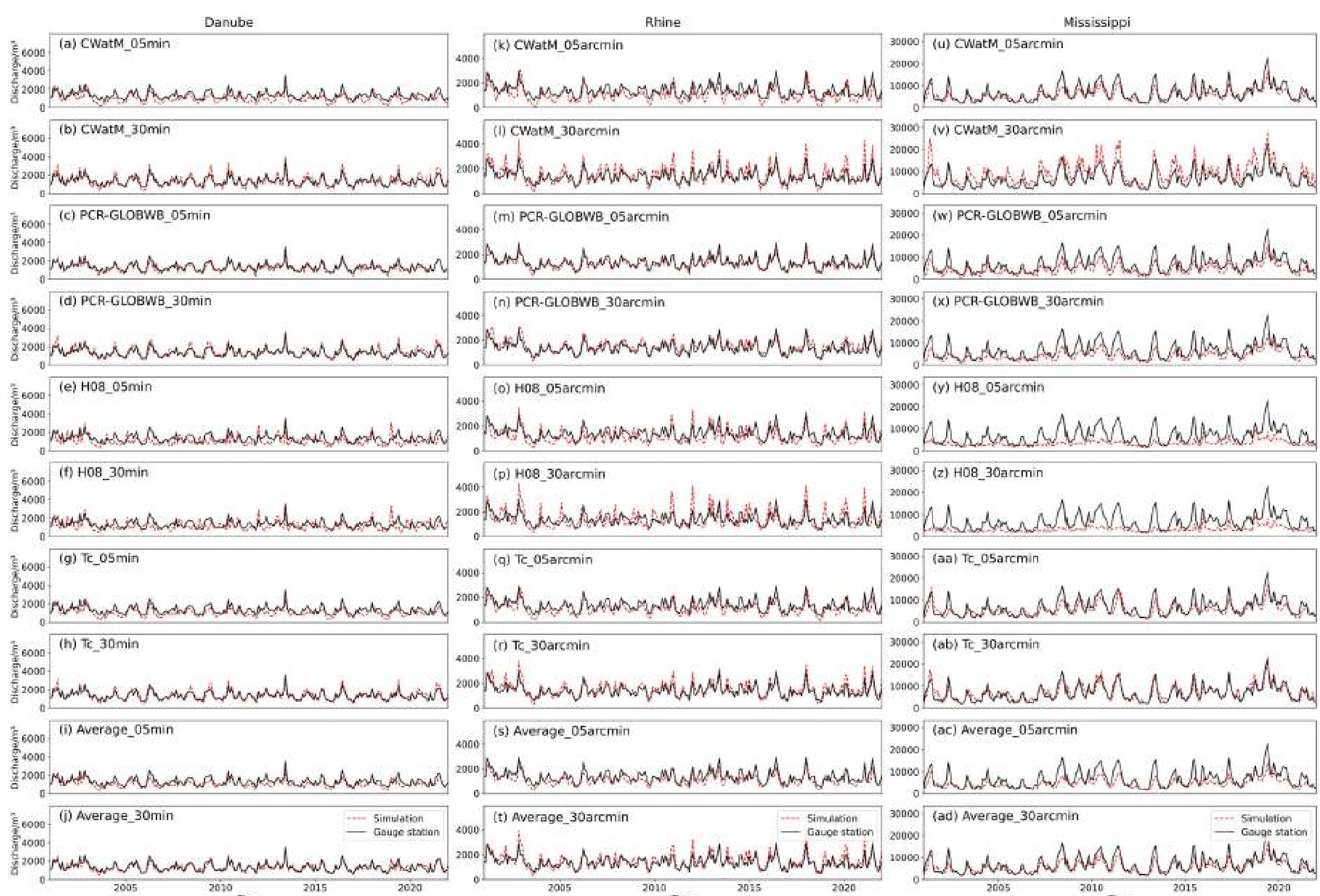


**Fig. 6.** Streamflow estimates from different model simulation scenarios (red) and in situ data (black) over three river basins, the Danube River (left column), Rhine River (middle column), and Mississippi River (right column).

superior performance across all three river basins (Fig. 6g, q, and aa). For example, in the Mississippi River, TC_05arcmin effectively captures both peak flows and low-flow periods (Fig. 6aa), showcasing its strong capability in simulating dynamic hydrological conditions. In the Danube and Rhine Rivers, high-resolution models also accurately reproduce the timing and magnitude of flood peaks (Fig. 6g and q), suggesting their potential utility in flood monitoring and early warning systems. In contrast, low-resolution models—such as CWatM_30arcmin, H08_30arcmin, and PCR-GLOBWB_30arcmin—perform considerably poorer, particularly in the Rhine and Mississippi basins (Fig. 6l, n, p, v, x, and z). These models frequently underestimate the magnitude of peak flows and exhibit unrealistic recession rates, deviating substantially from observed discharge trends. For instance, in the Mississippi River, low-resolution models fail to reproduce the fluctuating discharge patterns during high-flow periods. Interestingly, despite their coarser resolution, TC_30arcmin and Ave_30arcmin outperform the individual low-resolution models (Fig. 6h, j, r, t, ab, and ad), indicating the effectiveness of data fusion in improving simulation accuracy even at lower resolutions. These findings suggest that individual hydrological models may have limited capability in representing localized watershed characteristics, especially in large and complex basins. Data fusion approaches such as Triple Collocation and multi-model averaging can enhance streamflow simulation by integrating complementary information from multiple models, thereby compensating for individual model weaknesses (Arathy Nair et al., 2024; Himeur et al., 2022; Jiang et al., 2024; Zhang et al., 2023).

### 6.3. Triple collocation performance with different references

The performance of the TC method can vary considerably depending on the choice of reference system. Analyzing the TC results generated from different reference systems is essential for assessing their impact and selecting the best reference system necessary for achieving the most accurate simulation results. In this experiment, CWatM, PCR-GLOBWB, and H08 were selected as the reference systems for the TC method, and their effects on the simulation performance were investigated. Note that only H08 has been used as the reference system so far.

From Fig. 7, the TC method using H08 as the reference system exhibits higher KGE values across many global regions, particularly in North America, South America, and Europe (Fig. 7e and f). In contrast, the TC methods using PCR-GLOBWB and CWatM as reference systems show relatively weaker performance in most regions, notably northwestern North America (Fig. 7a, b, c, and d). In general, the high-resolution model performs better except for TC_C_05arcmin, where the average KGE is about 19% lower than its coarser resolution (Fig. 8).

When examining TC experiments using different reference systems,

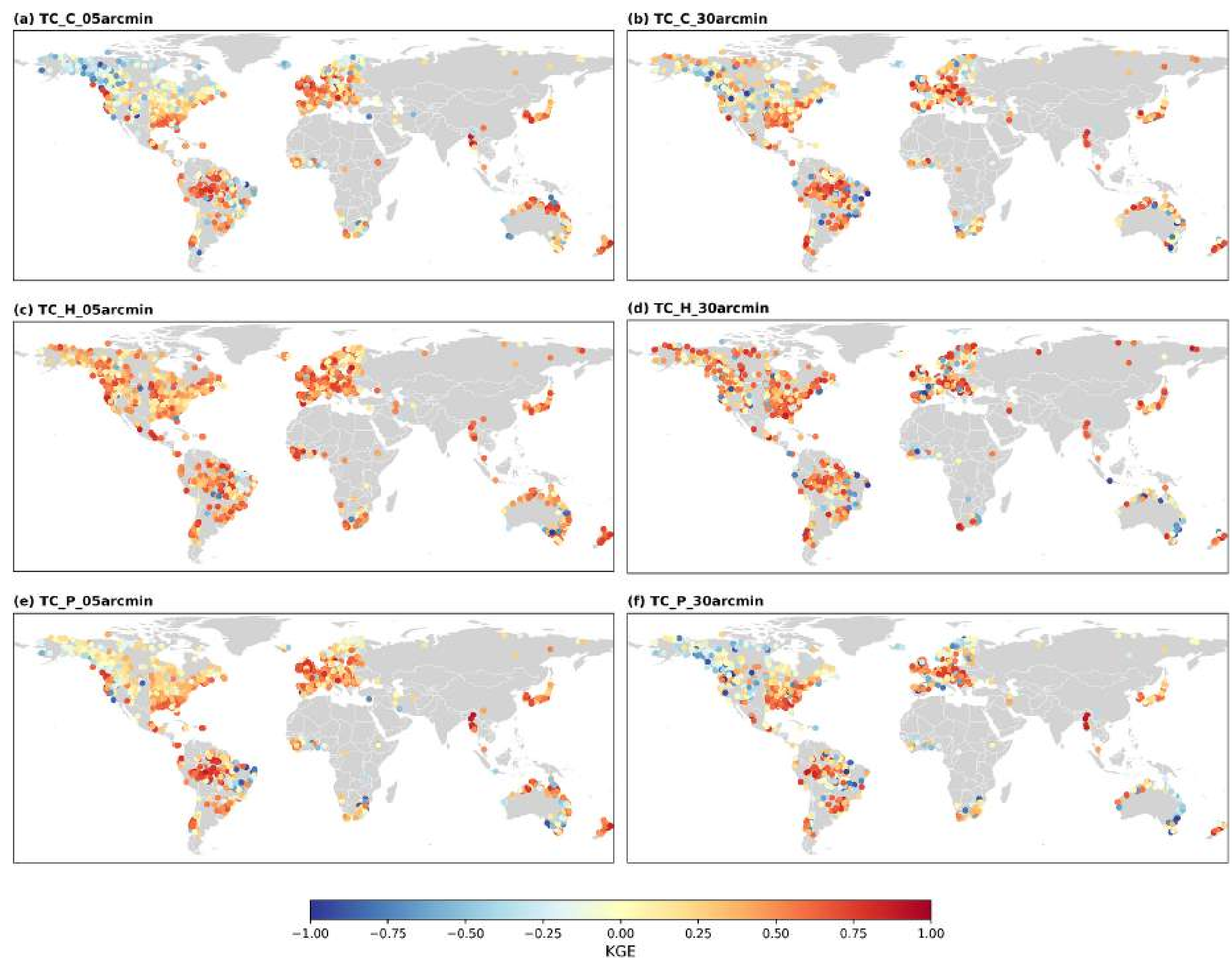


**Fig. 7.** Global distribution of KGE values under the TC method with different reference systems and spatial resolutions. The color scale indicates the range of KGE values, with red representing low KGE values and blue representing high KGE values. The label format TC_REF_Resolution, where TC represents Triple Collocation, REF represents the reference (C: CWatM, P: PCR-GLOBWB, and H: H08), and Res indicates spatial resolution.

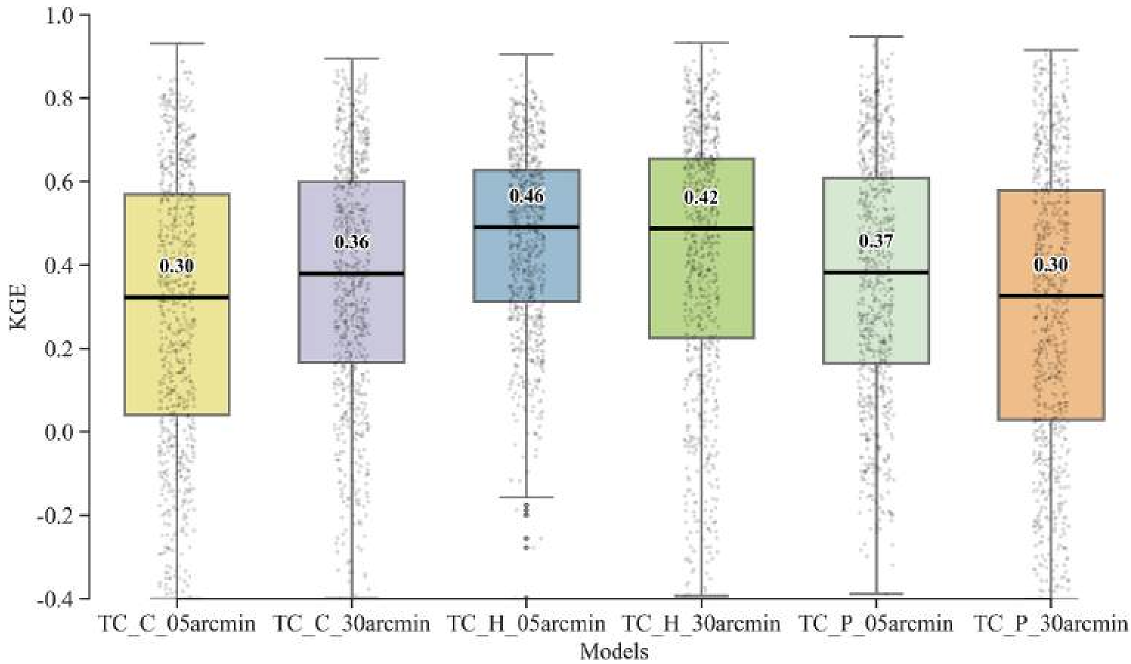


**Fig. 8.** Boxplots of model performance metric (KGE) (see NRMSE, NSE, and $R^2$ in supplements) from TC results with different references. The label format TC_REF_Resolution, where TC represents Triple Collocation, REF represents the reference (C: CWatM, P: PCR-GLOBWB, and H: H08), and Res indicates spatial resolution.

clear differences emerge in model behavior and fusion efficiency (Fig. 8, Figs. S5–S7, and Table S2). When H08 serves as the reference system, the resulting TC_H_05arcmin achieves the best overall performance, with KGE = 0.46 (as well as NSE = 0.26, NRMSE = 0.17, and R2 = 0.57), substantially outperforming all other TC configurations. Even at the coarser resolution (TC_H_30arcmin), the model maintains its accuracy, indicating robust behavior under spatial degradation. This suggests that although H08 exhibits weaker standalone performance, its relatively independent error structure makes it a suitable baseline for error decomposition, allowing TC to effectively balance and integrate complementary model strengths. By contrast, when CWatM is used as the reference, the results show lower performance (e.g., TC_C_05arcmin: KGE = 0.30) despite CWatM's strong standalone accuracy. This reduction occurs because the TC framework projects all model errors relative to the reference system, and if the reference's error pattern correlates with those of other models (as in CWatM, which shares meteorological forcings with PCR-GLOBWB and H08), the assumption of error independence is partially violated, leading to weaker performance. Similarly, using PCR-GLOBWB as the reference yields intermediate results (TC_P_05arcmin: KGE = 0.37), reflecting moderate but less stable fusion outcomes. Overall, these findings demonstrate that the choice of reference system critically determines TC performance. A weaker but error-independent model (such as H08) can yield superior fusion results, whereas a strong but highly correlated model (such as CWatM) can limit the benefits of the TC approach. This underscores the need to carefully diagnose inter-model error correlations before selecting a reference system to ensure robust, unbiased triple collocation results.

### 6.4. *Performance of streamflow simulation in Thailand*

While streamflow simulations perform well against the global database, their performance remains unclear in regions without GRDC station data. This is the case for Thailand, which was excluded from both the global (Sect. 6.1) and regional (Sect. 6.2) analyses due to the absence of GRDC observations, limiting the applicability of the previous findings to the country. Moreover, given that Thailand frequently experiences severe flood events, this additional analysis provides important context for evaluating the model's capability in flood-prone environments (Darnkachatarn and Kajitani, 2025; Munpa et al., 2022; Suwannachai et al., 2025).

The simulation results of each model and data fusion method in Thailand were first analyzed for the period from 2001 to 2022. Individual models generally perform better in northern and central Thailand, while stations in southern and eastern regions often exhibit lower or even negative KGE values (Fig. 9). CWatM_05arcmin demonstrates relatively strong performance, achieving an average KGE = 0.36 as well as NSE = 0.25 and NRMSE = 0.18 (Table S4). In contrast, CWatM_30arcmin shows the weakest overall performance (e.g., KGE = 0.08), confirming that coarse spatial resolution significantly limits the model's ability to capture regional hydrological variability (Meema et al., 2025; Pokavanich et al., 2024). Similarly, the PCR-GLOBWB model exhibits moderate skill at high resolution (e.g., KGE = 0.29) but degrades considerably at coarse resolution (e.g., KGE = 0.12), reinforcing that spatial detail is critical for improving simulation accuracy.

Among all approaches, the TC method shows clear improvement over individual models, producing higher KGE values across most stations in northern, central, and northeastern Thailand (Fig. 9). The fine-resolution TC_05arcmin achieves the best overall performance (e.g., KGE = 0.40), while at 30 arcmin the accuracy decreases (e.g., KGE = 0.16), suggesting that the benefits of multi-data integration may be partially offset by the coarse spatial representation of hydrological features but remains comparable to or slightly better than the best low-resolution individual models (Mourad et al., 2024). The Average method also performs well, especially at fine resolution (Average_05arcmin, e.g., KGE = 0.39), closely matching TC_05arcmin, whereas the coarse-resolution version (Average_30arcmin, e.g., KGE = 0.27) remains superior to most low-resolution individual models. This confirms that even a simple ensemble averaging approach can substantially improve the stability and reliability of streamflow simulations. Overall, both data-fusion methods outperform individual models in Thailand, and fine-resolution ensembles (5 arcmin) consistently yield more accurate and spatially coherent discharge estimates than their coarse-resolution counterparts, emphasizing the importance of spatial

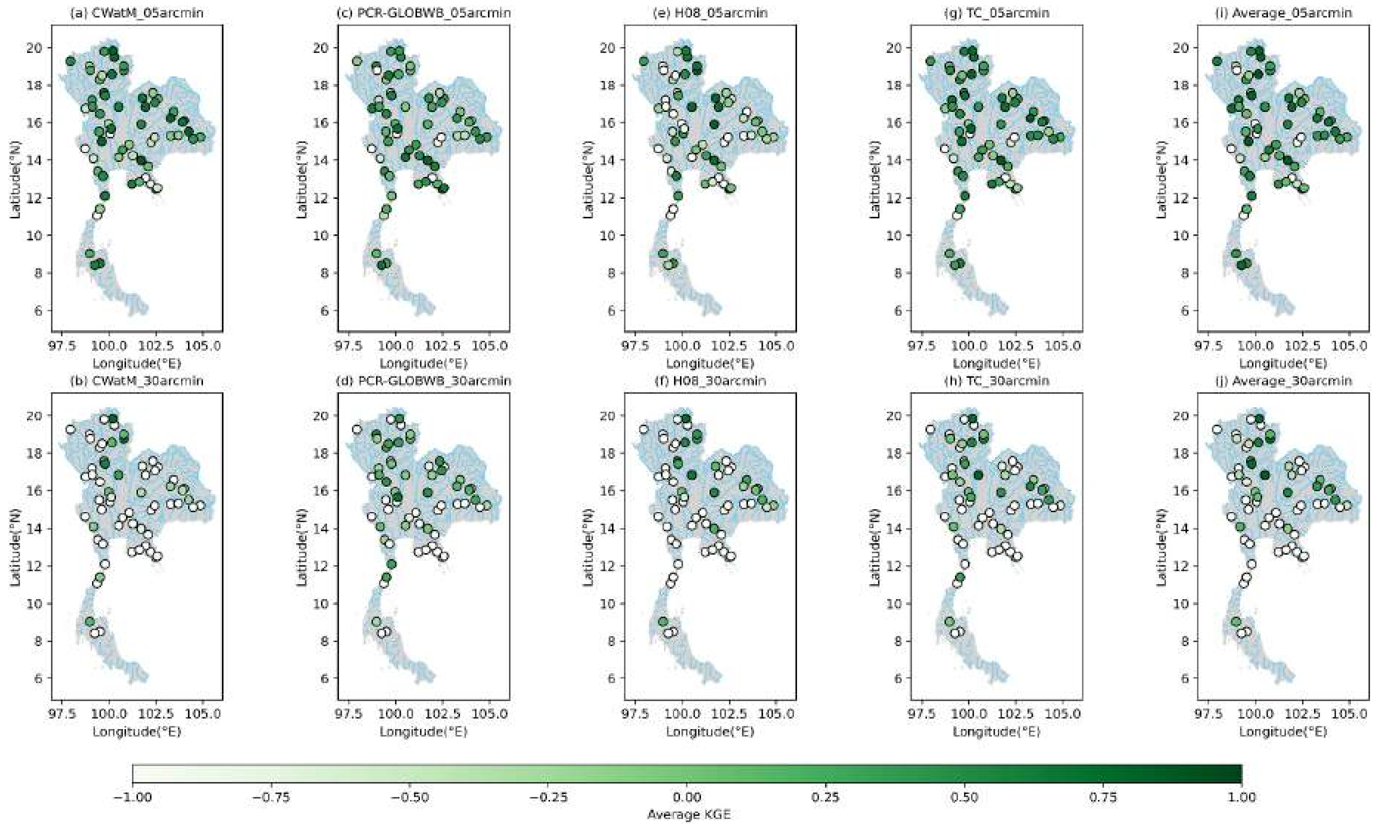


**Fig. 9.** Spatial distribution of KGE values for different models and resolutions in Thailand.

detail in hydrological modeling (Alcantara, 2023; Alcantara and Ahn, 2024).

## 7. Discussion

This discussion highlights the key findings and contributions of this study in the context of global streamflow estimation. First, we provide the first global-scale application of TC for multi-model streamflow optimization, extending TC beyond its common use for hydrometeorological variables. Second, by comparing two widely used global resolutions (30 arcmin vs 5 arcmin), we clarify—in a way that is directly interpretable—whether finer resolution actually leads to better discharge estimates and whether such gains persist after applying optimization approaches (TC and simple averaging). Third, by combining a global GRDC with an independent regional RID, we demonstrate that the main conclusions are not driven by a single dataset and remain relevant for practical applications. The following subsections are organized accordingly: Section 7.1 summarizes baseline model performance; Section 7.2 examines the impact of improved spatial resolution on model results; Section 7.3 discusses the effectiveness of the TC framework under model dependency; Section 7.4 interprets the influence of reservoir regulation and water management; Section 7.5 evaluates the benefits of streamflow optimization; and Section 7.6 highlights practical implications using the Thailand case study.

### 7.1. Individual model performances

Our evaluation indicates significant regional differences in the performance of the three hydrological models (CWatM, PCR-GLOBWB, and H08), which aligns with previous research findings. Specifically, the CWatM model exhibits superior performance in Western Europe, consistent with the study by Burek et al. (2020), who also validated the CWatM model using GRDC data and reported high KGE values primarily concentrated in Western Europe. This further confirms the accuracy of the CWatM model in simulating river discharge in this region. On the other hand, we observed that the PCR-GLOBWB model generally performs better in Europe and North America, a finding supported by earlier research, e.g., Sutanudjaja et al. (2018). Additionally, in Western Europe, the PCR-GLOBWB improvements in the 5 arcmin simulation compared to the 30 arcmin simulation are primarily observed in the Alps and Norwegian mountains, likely due to the better representation of terrain and snow dynamics at higher resolutions (Sutanudjaja et al., 2018). However, model performances in some parts of Africa were relatively poor, particularly in the Niger River basin, likely attributed to an overestimated groundwater decay coefficient and insufficient simulation of evaporation in inland delta regions (Sutanudjaja et al., 2018). This suggests a need for further adjustments to model parameters or processes in Africa. According to our analysis, the H08 model performed poorly in northwestern North America, central South America, and southern Africa, while its performance in Western Europe was relatively acceptable. The relatively poor performance of H08 in South Africa can likely be attributed to the model's substantial overestimation of dry-season streamflow, a known limitation of global hydrological models in semi-arid and arid climates (Hanasaki et al., 2018; Zaherpour et al., 2018). Such regions typically exhibit pronounced seasonality in precipitation and evapotranspiration, which complicates the accuracy of streamflow simulations, particularly under low-flow conditions. Conversely, the relatively acceptable performance of H08 in Western Europe aligns with previous studies that noted improved performance of global hydrological models in wetter climates at mid-northern latitudes (Kumar et al., 2022; Zaherpour et al., 2018). The clearly defined seasonal cycles of precipitation and runoff processes in these areas enable more accurate representation by the model, especially regarding snowmelt processes and soil moisture dynamics.

### 7.2. Impact of spatial resolution on model performance

Model resolution is a critical factor influencing the performance of hydrological models. Higher spatial resolution is generally expected to better capture basin-scale heterogeneity, leading to improved simulation accuracy. In our analysis, models with a 5 arcmin resolution typically outperformed those with a 30 arcmin resolution, reflecting the benefits of higher resolution in representing localized hydrological processes. However, there are cases where 30 arcmin models demonstrated better performance. This counterintuitive outcome may be attributed to the smoothing effect inherent in coarser-resolution models, where larger grid cells average out localized variability and potential errors (Müller et al., 2024), thus improving the overall agreement with observed data. Ren et al. (2023) suggest that at a coarse resolution, the simplified river network structure and aggregated basin response processes may better align with the spatial scale of input data, contributing to improved model stability and reliability. Additionally, the aggregation of observational data over larger areas at coarser resolution may help offset localized errors, resulting in stronger agreement between model outputs and observations in some areas, despite the slightly lower average KGE values for low-resolution models (Khosh Bin Ghomash et al., 2025; Liu et al., 2024).

These findings suggest that while increasing model resolution can enhance simulation accuracy, its effectiveness depends on several factors, including the quality of input data, the capacity for parameter calibration, and available computational resources (Herrera et al., 2022). In some contexts, coarser-resolution models may offer a practical and reliable alternative (Haddeland et al., 2002; Shrestha et al., 2006). Future research should aim to improve the performance of high-resolution models by developing more efficient calibration techniques, increasing the spatial resolution and accuracy of input datasets, and optimizing model structures for computational efficiency (Bierkens et al., 2015; Huang et al., 2019; Sun et al., 2020, Sun et al., 2018b). Moreover, the selection of model resolution should be tailored to specific research objectives, study areas, and resource constraints to ensure that simulation accuracy is aligned with the intended application (Fatichi et al., 2016; Nath et al., 2024; Singh et al., 2025).

### 7.3. Effectiveness of the TC framework under model dependency

In this study, we explicitly acknowledge the key assumptions of the TC framework and the main limitations relevant to our global streamflow application, to facilitate transparent interpretation of the results. The three global hydrological models share common meteorological forcings, reflecting the shared goal of achieving accurate hydrological simulations. This interdependence supports fair performance evaluation, as all models are assessed under similar conditions, allowing differences in performance primarily to reflect variations in model processes (Brunner et al., 2021; Herrera et al., 2022; Tang et al., 2023). At the same time, the use of a common forcing dataset constitutes an important limitation for TC because it can introduce error dependencies across models, and replacing it is difficult due to the scarcity of alternative long-term, high-resolution datasets (Mankin et al., 2025; Shuai et al., 2022). ERA5-Land currently provides the most comprehensive and consistent meteorological record for global hydrological simulations, ensuring optimal model performance and comparability (Dutta and Markonis, 2024; Muñoz-Sabater et al., 2021; Xu et al., 2022).

With respect to the TC framework, the most critical assumption in this context is the independence of simulation errors among the models. However, beyond the forcing data, most hydrological models, including those used here, share similar conceptual frameworks and rely on common datasets for parameterization, making full independence difficult to achieve (Beven, 2024; Refsgaard et al., 2022; Singh et al., 2025). As a result, their simulation errors may not be entirely independent, potentially violating one of the TC framework assumptions (Alemohammad et al., 2015; Qiao et al., 2022; Zhu et al., 2023). While this is a known limitation, complete independence among models is rarely attainable in real-world applications (Koster et al., 2021). Previous TC studies have shown that partial violations of this assumption are acceptable for practical purposes (Su et al., 2014; Yilmaz and Crow, 2014; Zwieback et al., 2012). Despite these limitations, the TC-derived discharge estimates effectively fulfill the study objectives, providing improved results compared to individual models. These improvements should not be interpreted as a universal correction of model structural deficiencies. Rather, the sensitivity to the reference system indicates that TC performance is conditional on the underlying error-correlation structure and the extent to which TC assumptions are approximately satisfied in practice. This demonstrates that even when strict assumptions are not fully satisfied, the TC framework remains robust, yielding meaningful improvements and consistent comparative insights by leveraging the relative consistency among models.

Finally, it should be acknowledged that our evaluation is conducted at a monthly time scale, which is appropriate for global long-term assessment but may smooth short-term extremes and operational signals that occur at sub-monthly scales. Future research could explore alternative or multi-source forcing datasets to enhance model independence and, where suitable global forcing and discharge records become available, extend the TC-based analysis to finer temporal resolutions; however, obtaining comprehensive variables from unified sources comparable to ERA5 remains a major challenge for global-scale hydrological applications.

### 7.4. Influence of reservoir regulation and water management

Reservoir operation and water management represent critical yet complex factors influencing the consistency between simulated and observed streamflow (Abeshu et al., 2023; Dash et al., 2023). Many hydrological stations used for validation are located in river basins where discharge is heavily regulated by reservoirs, irrigation withdrawals, and other anthropogenic activities (Asadi et al., 2024; Ha et al., 2023). Consequently, discrepancies between model and observed flows may not solely reflect model structural limitations but also the effects of human interventions that are either simplified or absent in the model frameworks (Alam et al., 2022; Jafarzadegan et al., 2023).

In our analysis, the three models—CWatM, PCR-GLOBWB, and H08—exhibit substantial differences in how they represent reservoir operation and water management, which directly affect their consistency with observed discharge during validation. CWatM provides the most comprehensive representation by dynamically integrating reservoirs and lakes from the HydroLakes database into the routing network, simulating their operation through the LISFLOOD scheme, while also quantifying water withdrawals, consumption, and return flows across multiple sectors. This detailed coupling enables CWatM to realistically reproduce the effects of regulated flows, such as the attenuation of flood peaks and augmentation of dry-season discharge (Burek et al., 2020; Guillaumot et al., 2022). In contrast, PCR-GLOBWB includes more than 6,000 reservoirs from the GRanD database and fully integrates human water use, but its simplified rule-based operation does not fully capture short-term regulation or reservoir-specific dynamics. As a result, it can reproduce the overall discharge magnitude but may show temporal mismatches, such as delayed flood recession or underestimated baseflow under strong regulation (Sutanudjaja et al., 2018). H08, meanwhile, adopts a conceptual storage–release scheme and differentiates between global and local reservoirs but lacks dynamic control rules. Although it accounts for aqueduct transfers, desalination, and return flows, the simplified structure leads to naturalized discharge behaviour with higher variability than observations, especially in strongly regulated basins like Southeast Asia and North America (Hanasaki et al., 2018). Consequently, the differences among these models in representing human regulation partly explain the observed discrepancies during validation—CWatM tends to yield more realistic hydrographs in managed basins, while PCR-GLOBWB and H08 often produce greater variability, reflecting the absence or simplification of anthropogenic regulation mechanisms.

This divergence highlights that part of the model–observation mismatch arises from inconsistent treatment of human impacts rather than from hydrological process uncertainty alone (Worden et al., 2025; Zheng et al., 2024). For example, an apparently high bias in simulated discharge during flood periods could result from unmodeled flood control releases, while underestimation during dry seasons may be linked to unrepresented irrigation withdrawals. Therefore, interpreting validation results without accounting for anthropogenic regulation can overstate model deficiencies (Pool et al., 2021). These findings emphasize the need for more consistent integration of human water management processes in large-scale hydrological modelling. Incorporating detailed reservoir operation rules, irrigation water use, and return flows could help reduce the systematic biases caused by simplified or missing regulation mechanisms (Alam et al., 2022; Ferreira et al., 2023; Zipper et al., 2022). Future research should also explore the use of hybrid data assimilation or enhanced human-influence parameterization frameworks to improve the representation of regulated basins, thereby enhancing the reliability of multi-model comparisons and fusion-based streamflow estimates (Mohammadi, 2024; Slater et al., 2023). However, implementing such improvements remains a major challenge for most hydrological models due to the difficulty in obtaining restricted national data on reservoir operations and water management, which are often unavailable for public use.

### 7.5. Benefit of streamflow optimization

Combining streamflow outputs from CWatM, PCR-GLOBWB, and H08 using TC and simple averaging methods significantly improved simulation accuracy. Both approaches outperformed individual model outputs, demonstrating notable enhancements across five major global regions as well as in Thailand. In particular, the TC method, which explicitly accounts for systematic and random errors among models, yielded more accurate streamflow estimates (Kim et al., 2023; Xu et al., 2024; Zhou et al., 2021). However, it is important to note that the choice of reference system within the TC framework significantly influences the results. In this study, using H08 as the reference yielded the best overall streamflow estimates. Since the TC method is limited to three input datasets, the selection of participating hydrological models becomes critical. Given that the available models often exceed this number, the fusion results may vary depending on which models are chosen, especially when the assumption of mutual independence among them is not fully met (Lyu et al., 2021; Pan et al., 2015). Future work could incorporate a broader range of hydrological models to systematically assess the impact of model selection on the fusion results.

While the more straightforward averaging method achieved slightly lower performance compared to TC, it still consistently outperformed individual models. Due to its computational simplicity and ease of implementation, averaging remains a practical and effective strategy, particularly in applications where resources are limited or rapid deployment is required.

### 7.6. Streamflow optimization in Thailand

The analysis explicitly conducted for Thailand yielded conclusions consistent with those from global and regional evaluations. Despite using different observational datasets (RID for Thailand and GRDC for global-scale), the relative performance ranking of the hydrological models (CWatM, PCR-GLOBWB, and H08) and the effectiveness of the data fusion methods (TC and Avg methods) remained consistent. This consistency can be primarily attributed to the applicability of the three hydrological models and the robustness of the TC method in effectively addressing systematic and random errors inherent across different scales (Li et al., 2018; Zhao et al., 2024).

Thailand can substantially benefit from the insights presented in this study, especially given the limited availability of high-quality, country-scale streamflow simulations (Alcamo et al., 2003; Mateo et al., 2014; Padiyedath Gopalan et al., 2021, 2022a, Padiyedath Gopalan et al., 2022b; Zhang et al., 2020). Enhanced streamflow predictions obtained through data fusion methods have significant implications for improving national flood forecasting, drought management strategies, and overall water resources planning (Badar et al., 2024; VanRheenen et al., 2004). Future research efforts in Thailand may leverage hydrological models and these advanced data fusion approaches to achieve more precise and reliable hydrological forecasts, thereby supporting informed decision-making and sustainable water resource management (Hlaing et al., 2024; Suwannachai et al., 2025; Thaisiam et al., 2024; Waqas et al., 2024).

The findings of this study have several practical implications for global streamflow applications, particularly in data-scarce and regulated basins. First, the results suggest that data fusion can provide a robust pathway to improve discharge estimates when extensive calibration is infeasible or observations are limited because it leverages complementary information across models rather than relying on a single model configuration. Second, the resolution comparison indicates that higher spatial resolution does not automatically translate into better streamflow performance; therefore, decisions to adopt finer-resolution products should be guided by demonstrated gains in skill and by basin characteristics, rather than by resolution alone. Third, in strongly regulated basins, part of the model–observation mismatch may reflect incomplete representation of reservoir operation and human water use, implying that validation results should be interpreted with caution and, where possible, complemented by information on regulation intensity. Overall, these implications provide actionable guidance for selecting model resolution and applying fusion strategies to produce more reliable global streamflow estimates under realistic constraints of data availability and human influence.

## 8. Conclusions

This study provides a comprehensive evaluation of global river discharge simulations using three widely used hydrological models—CWatM, PCR-GLOBWB, and H08—alongside two multi-data fusion approaches: Triple Collocation (TC) and simple averaging. By benchmarking model outputs against observed streamflow data from the GRDC and Thailand's RID networks, we assessed model accuracy, regional performance variability, and the added value of data fusion and spatial resolution enhancements. Among individual models, CWatM demonstrated the most reliable performance, particularly at high resolution (5 arcmin). Regional analysis revealed that all models performed best in Europe, where dense observational networks likely contributed to improved calibration and validation. In contrast, performance declined in data-sparse regions such as Southern Africa and Australia.

Both data combination methods—TC and averaging—consistently outperformed individual models across all five major global regions, demonstrating their potential to enhance simulation accuracy by integrating complementary strengths from different models. Notably, the TC method, which explicitly accounts for systematic and random errors, produced the best results when using H08 as the reference system, despite H08′s weaker standalone performance. This underscores the importance of reference system selection in TC applications and suggests that underutilized models can still play a valuable role within data fusion frameworks. These improvements should be interpreted as conditional on the TC assumptions and inter-model error characteristics, rather than as a universal correction of model structural deficiencies. Moreover, the sensitivity of TC to reference system selection highlights the importance of making careful methodological choices in multi-data integration.

Resolution-based analysis confirmed that high-resolution simulations substantially improve streamflow estimates, particularly for CWatM, which showed a clear performance gain over its low-resolution version. However, for PCR-GLOBWB, H08, and the fusion methods, increasing resolution did not lead to substantial improvements. This suggests that the advantages of higher resolution may vary depending on the model structure, the quality of input data, and the model's calibration capability.

Overall, the findings emphasize the importance of strategic model selection, integrating multiple model outputs, and considering spatial resolution in hydrological modeling efforts. These insights provide practical guidance for enhancing streamflow simulation in both data-rich and data-sparse regions, laying the foundation for future research that expands the capabilities of high-resolution simulations and refines multi-data fusion methods.

## CRediT authorship contribution statement

**Mingze Sun:** Writing – review & editing, Writing – original draft, Visualization, Data curation. **Natthachet Tangdamrongsub:** Writing – review & editing, Writing – original draft, Validation, Data curation, Supervision. **Yu Sun:** Writing – review & editing, Writing – original draft, Visualization. **Jianzhi Dong:** Writing – review & editing, Writing – original draft, Methodology. **Edwin Sutanudjaja:** Writing – review & editing. **Mikhail Smilovic:** Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This research was supported by the National Key Research and Development Program of China (2025YFE0102700), the National Natural Science Foundation of China (grant 42171426, 42374041 and 42304099), and the Thailand's Agricultural Research Development Agency (ARDA), PRP6705031850. We thank Dr. Naota Hanasaki for supporting the H08 configuration and suggestions. Finally, we thank the editor and reviewers for their significant effort in assessing the manuscript and for their insightful suggestions, which greatly improved the paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2026.135122.

## Data availability

Code and data availability are shown in Sects. 3 – 4. The internet links are as follows:

| | |
|---|---|
| CWatM | https://cwatm.iiasa.ac.at/ |
| PCR-GLOBWB | https://globalhydrology.nl/research/models/pcr-globwb-2-0/ |
| H08 | https://h08.nies.go.jp/h08/ |
| GRDC | https://grdc.bafg.de/ |
| Thailand gauge data | http://hydro-6.rid.go.th/ |

## References

Abeshu, G.W., Tian, F., Wild, T., Zhao, M., Turner, S., Chowdhury, A.F.M.K., Vernon, C. R., Hu, H., Zhuang, Y., Hejazi, M., Li, H.-Y., 2023. Enhancing the representation of water management in global hydrological models. Geosci. Model Dev. 16, 5449–5472. https://doi.org/10.5194/gmd-16-5449-2023.

Akhtar, F., Borgemeister, C., Tischbein, B., Awan, U.K., 2022. Metrics assessment and streamflow modeling under changing climate in a data-scarce heterogeneous region: a Case study of the Kabul River Basin. Water 14, 1697. https://doi.org/10.3390/w14111697.

Akpoti, K., Mekonnen, K., Leh, M., Owusu, A., Dembélé, M., Tinonetsana, P., Seid, A., Velpuri, N.M., 2024. State of continental discharge estimation and modelling: challenges and opportunities for Africa. Hydrol. Sci. J.

Alam, M.F., McClain, M., Sikka, A., Pande, S., 2022. Understanding human–water feedbacks of interventions in agricultural systems with agent based models: a review. Environ. Res. Lett. 17, 103003. https://doi.org/10.1088/1748-9326/ac91e1.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrol. Sci. J. https://doi.org/10.1623/hysj.48.3.317.45290.

Alcantara, A. L. and Ahn, K.-H.: Advance: Amplifying data validation and analysis through time-varying quadruple collocation for enhanced precipitation error estimation and integration, https://doi.org/10.2139/ssrn.4586440, 28 September 2023.

Alcantara, A.L., Ahn, K.-H., 2024. Time-varying quadruple collocation for enhanced satellite and reanalysis precipitation data error estimation and integration. Int. J. Appl. Earth Obs. Geoinf. 127, 103692. https://doi.org/10.1016/j.jag.2024.103692.

Alemohammad, S.H., McColl, K.A., Konings, A.G., Entekhabi, D., Stoffelen, A., 2015. Characterization of precipitation product errors across the United States using multiplicative triple collocation. Hydrol. Earth Syst. Sci. 19, 3489–3503. https://doi.org/10.5194/hess-19-3489-2015.

Ali, M.H., Popescu, I., Jonoski, A., Solomatine, D.P., 2023. Remote Sensed and/or Global Datasets for distributed Hydrological Modelling: a Review. Remote Sens. (Basel) 15, 1642. https://doi.org/10.3390/rs15061642.

Andrade, J.M., Ribeiro Neto, A., Nóbrega, R.L.B., Rico-Ramirez, M.A., Montenegro, S.M. G.L., 2024. Efficiency of global precipitation datasets in tropical and subtropical catchments revealed by large sampling hydrological modelling. J. Hydrol. 633, 131016. https://doi.org/10.1016/j.jhydrol.2024.131016.

Arathy Nair, G.R., Adarsh, S., El-Shafie, A., Ahmed, A.N., 2024. Enhancing hydrological data completeness: a performance evaluation of various machine learning techniques using probabilistic fusion imputer with neural networks for streamflow data reconstruction. J. Hydrol. 639, 131583. https://doi.org/10.1016/j.jhydrol.2024.131583.

Asadi, S., Mousavi, S.J., López-Ballesteros, A., Senent-Aparicio, J., 2024. Analyzing hydrological alteration and environmental flows in a highly anthropized agricultural river basin system using SWAT+, WEAP and IAHRIS. J. Hydrol.: Reg. Stud. 52, 101738. https://doi.org/10.1016/j.ejrh.2024.101738.

Badar, Z., Almazah, M.M.A., Raza, M.A., Hussain, I., Al-Duais, F.S., Al-Rezami, A.Y., 2024. Integration of three drought indices based on triple collocation and multi-scalar weighted amalgamated drought index. Stoch Environ. Res. Risk Assess 38, 1179–1195. https://doi.org/10.1007/s00477-023-02623-w.

Becker, R., Koppa, A., Schulz, S., Usman, M., aus der Beek, T., Schüth, C., 2019. Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data. J. Hydrol. 577, 123944. https://doi.org/10.1016/j.jhydrol.2019.123944.

Belvederesi, C., Zaghloul, M.S., Achari, G., Gupta, A., Hassan, Q.K., 2022. Modelling river flow in cold and ungauged regions: a review of the purposes, methods, and challenges. Environ. Rev. 30, 159–173. https://doi.org/10.1139/er-2021-0043.

Beven, K., 2024. A brief history of information and disinformation in hydrological data and the impact on the evaluation of hydrological models. Hydrol. Sci. J. 69, 519–527. https://doi.org/10.1080/02626667.2024.2332616.

Bierkens, M.F.P., Bell, V.A., Burek, P., Chaney, N., Condon, L.E., David, C.H., De Roo, A., Döll, P., Drost, N., Famiglietti, J.S., Flörke, M., Gochis, D.J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R.M., Reager, J.T., Samaniego, L., Sudicky, E., Sutanudjaja, E.H., Van De Giesen, N., Winsemius, H., Wood, E.F., 2015. Hyper-resolution global hydrological modelling: what is next?: "everywhere and locally relevant,. Hydrol. Process. 29, 310–320. https://doi.org/10.1002/hyp.10391.

Bonnema, M., Hossain, F., 2019. Assessing the potential of the surface water and ocean topography mission for reservoir monitoring in the Mekong River Basin. Water Resour. Res. 55, 444–461. https://doi.org/10.1029/2018WR023743.

Bringeland, S., Fotopoulos, G., 2024. Analysis of gap filling techniques for GRACE/GRACE-FO terrestrial water storage anomalies in Canada. J. Hydrol. 630, 130644. https://doi.org/10.1016/j.jhydrol.2024.130644.

Brunner, M.I., Slater, L., Tallaksen, L.M., Clark, M., 2021. Challenges in modeling and predicting floods and droughts: a review. WIREs Water 8, e1520.

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., Smilovic, M., Guillaumot, L., Zhao, F., Wada, Y., 2020. 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. 13, 3267–3298. https://doi.org/10.5194/gmd-13-3267-2020.

Chen, C., He, M., Chen, Q., Zhang, J., Li, Z., Wang, Z., Duan, Z., 2022. Triple collocation-based error estimation and data fusion of global gridded precipitation products over the Yangtze River basin. J. Hydrol. 605, 127307. https://doi.org/10.1016/j.jhydrol.2021.127307.

Darnkachatarn, S., Kajitani, Y., 2025. Flood damage assessment model of industrial sectors in a Megacity: Derivation from business survey data in the Bangkok Metropolitan Region. Int. J. Disaster Risk Reduct. 118, 105221. https://doi.org/10.1016/j.ijdrr.2025.105221.

Dash, S.S., Sahoo, B., Raghuwanshi, N.S., 2023. An integrated reservoir operation framework for enhanced water resources planning. Sci. Rep. 13, 21720. https://doi.org/10.1038/s41598-023-49107-z.

Di Curzio, D., Castrignanò, A., Fountas, S., Romić, M., Viscarra Rossel, R.A., 2021. Multi-source data fusion of big spatial-temporal data in soil, geo-engineering and environmental studies. Sci. Total Environ. 788, 147842. https://doi.org/10.1016/j.scitotenv.2021.147842.

Dodet, G., Abdalla, S., Alday, M., Accensi, M., Bidlot, J., and Ardhuin, F.: Error characterization of significant wave heights in multidecadal satellite altimeter product, model Hindcast, and in situ measurements using the triple collocation technique, https://doi.org/10.1175/JTECH-D-21-0179.1, 2022.
Dommo, A., Aloysius, N., Lupo, A., Hunt, S., 2024. Spatial and temporal analysis and trends of extreme precipitation over the Mississippi River Basin, USA during 1988–2017. J. Hydrol.: Reg. Stud. 56, 101954. https://doi.org/10.1016/j.ejrh.2024.101954.
Duan, Z., Duggan, E., Chen, C., Gao, H., Dong, J., and Liu, J.: Comparison of Traditional Method and Triple Collocation Analysis for Evaluation of Multiple Gridded Precipitation Products across Germany, https://doi.org/10.1175/JHM-D-21-0049.1, 2021.
Duc, L., Sawada, Y., 2023. A signal-processing-based interpretation of the Nash–Sutcliffe efficiency. Hydrol. Earth Syst. Sci. 27, 1827–1839. https://doi.org/10.5194/hess-27-1827-2023.
Dutta, R., Markonis, Y., 2024. Does ERA5-land capture the changes in the terrestrial hydrological cycle across the globe? Environ. Res. Lett. 19, 024054. https://doi.org/10.1088/1748-9326/ad1d3a.
Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., Tarboton, D., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. J. Hydrol. 537, 45–60. https://doi.org/10.1016/j.jhydrol.2016.03.026.
Ferreira, A., Rolim, J., Paredes, P., do R. Cameira, M., 2023. Methodologies for water accounting at the collective irrigation system scale aiming at optimizing water productivity. Agronomy 13, 1938. https://doi.org/10.3390/agronomy13071938.
Fersch, B., Senatore, A., Adler, B., Arnault, J., Mauder, M., Schneider, K., Völksch, I., Kunstmann, H., 2020. High-resolution fully coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation. Hydrol. Earth Syst. Sci. 24, 2457–2481. https://doi.org/10.5194/hess-24-2457-2020.
Gao, S., Chen, M., Li, Z., Cook, S., Allen, D., Neeson, T., Yang, T., Yami, T., Hong, Y., 2021. Mapping dynamic non-perennial stream networks using high-resolution distributed hydrologic simulation: a case study in the upper blue river basin. J. Hydrol. 600, 126522. https://doi.org/10.1016/j.jhydrol.2021.126522.
Gonzales-Inca, C., Calle, M., Croghan, D., Torabi Haghighi, A., Marttila, H., Silander, J., Alho, P., 2022. Geospatial artificial intelligence (GeoAI) in the integrated hydrological and fluvial systems modeling: review of current applications and trends. Water 14, 2211. https://doi.org/10.3390/w14142211.
Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., Wagner, W., 2016. Recent advances in (soil moisture) triple collocation analysis. Int. J. Appl. Earth Obs. Geoinf. 45, 200–211. https://doi.org/10.1016/j.jag.2015.09.002.
Guan, X., Zhang, J., Yang, Q., Tang, X., Liu, C., Jin, J., Liu, Y., Bao, Z., Wang, G., 2020. Evaluation of precipitation products by using multiple hydrological models over the upper Yellow River Basin, China. Remote Sensing 12, 4023. https://doi.org/10.3390/rs12244023.
Guillaumot, L., Smilovic, M., Burek, P., de Bruijn, J., Greve, P., Kahil, T., Wada, Y., 2022. Coupling a large-scale hydrological model (CWatM v1.1) with a high-resolution groundwater flow model (MODFLOW 6) to assess the impact of irrigation at regional scale. Geosci. Model Dev. 15, 7099–7120. https://doi.org/10.5194/gmd-15-7099-2022.
Guo, X., Fu, Q., Hang, Y., Lu, H., Gao, F., Si, J., 2020. Spatial Variability of soil moisture in relation to land use types and topographic features on hillslopes in the black soil (Mollisols) area of northeast China. Sustainability 12, 3552. https://doi.org/10.3390/su12093552.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.
Ha, L.T., Bastiaanssen, W.G.M., Simons, G.W.H., Poortinga, A., 2023. A new framework of 17 hydrological ecosystem services (HESS17) for supporting river basin planning and environmental monitoring. Sustainability 15, 6182. https://doi.org/10.3390/su15076182.
Haddeland, I., Matheussen, B.V., Lettenmaier, D.P., 2002. Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. Water Resour. Res. 38, 29–1-29–10. https://doi.org/10.1029/2001WR000854.
Hakala Assendelft, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., and Seibert, J.: Hydrological modeling of climate change impacts, https://doi.org/10.1002/9781119300762.wsts0062, 2020.
Hanasaki, N., Yoshikawa, S., Pokhrel, Y., Kanae, S., 2018. A global hydrological simulation to specify the sources of water used by humans. Hydrol. Earth Syst. Sci. 22, 789–817. https://doi.org/10.5194/hess-22-789-2018.
Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., Seong, C., 2019. Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions. Sci. Rep. 9, 4974. https://doi.org/10.1038/s41598-019-41334-7.
Herrera, P.A., Marazuela, M.A., Hofmann, T., 2022. Parameter estimation and uncertainty analysis in hydrological modeling. WIREs Water 9, e1569.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorolog. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.
Himeur, Y., Rimal, B., Tiwary, A., Amira, A., 2022. Using artificial intelligence and data fusion for environmental monitoring: a review and future perspectives. Inf. Fusion 86–87, 44–75. https://doi.org/10.1016/j.inffus.2022.06.003.
Hlaing, P.T., Humphries, U.W., Waqas, M., 2024. Hydrological model parameter regionalization: Runoff estimation using machine learning techniques in the Tha Chin River Basin. Thailand, MethodsX 13, 102792. https://doi.org/10.1016/j.mex.2024.102792.
Hodson, T.O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci. Model Dev. 15, 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022.
Horton, P., Schaefli, B., Kauzlaric, M., 2022. Why do we have so many different hydrological models? A review Based on the Case of Switzerland. WIREs Water 9, e1574.
Huang, Y., Bárdossy, A., Zhang, K., 2019. Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data. Hydrol. Earth Syst. Sci. 23, 2647–2663. https://doi.org/10.5194/hess-23-2647-2019.
Jafarzadegan, K., Moradkhani, H., Pappenberger, F., Moftakhari, H., Bates, P., Abbaszadeh, P., Marsooli, R., Ferreira, C., Cloke, H.L., Ogden, F., Duan, Q., 2023. Recent advances and new frontiers in riverine and coastal flood modeling. Rev. Geophys. 61, e2022RG000788. https://doi.org/10.1029/2022RG000788.
Jalink, H., Dieperink, C., 2024. Towards design principles for good multi-level drought risk governance: some lessons from the rhine basin. Water 16, 879. https://doi.org/10.3390/w16060879.
Jia, L., Yen, N., Pei, Y., 2023. Spatial and temporal water quality data prediction of transboundary watershed using multiview neural network coupling. IEEE Trans. Geosci. Remote Sens. 61, 1–16. https://doi.org/10.1109/TGRS.2023.3334291.
Jiang, A., Zhang, W., Liu, X., Zhou, F., Li, A., Peng, H., Wang, H., 2024. Improving hydrological process simulation in mountain watersheds: Integrating WRF model gridded precipitation data into the SWAT model. J. Hydrol. 639, 131687. https://doi.org/10.1016/j.jhydrol.2024.131687.
Jin, X., Jin, Y., 2020. Calibration of a distributed hydrological model in a data-scarce Basin based on GLEAM datasets. Water 12, 897. https://doi.org/10.3390/w12030897.
Jin, Z., Zhao, Q., Qin, X., Zhang, J., Zhang, H., Qin, J., Qin, Y., Li, H., Chen, J., Liu, Y., Li, Y., Wang, L., 2021. Quantifying the impact of landscape changes on hydrological variables in the alpine and cold region using hydrological model and remote sensing data. Hydrol. Process. 35, e14392. https://doi.org/10.1002/hyp.14392.
Khaikham, L. and James, H.: Socio-Political Transformation After the 2011 Floods in Thailand, in: Population, Development, and the Environment: Challenges to Achieving the Sustainable Development Goals in the Asia Pacific, edited by: James, H., Springer Nature, Singapore, 227–253, https://doi.org/10.1007/978-981-13-2101-6_14, 2019.
Khosh Bin Ghomash, S., Yeste, P., Apel, H., Nguyen, V.D., 2025. Monte Carlo-based sensitivity analysis of the RIM2D hydrodynamic model for the 2021 flood event in western Germany. Nat. Hazards Earth Syst. Sci. 25 (975–990). https://doi.org/10.5194/nhess-25-975-2025.
Kim, H., Crow, W., Li, X., Wagner, W., Hahn, S., Lakshmi, V., 2023. True global error maps for SMAP, SMOS, and ASCAT soil moisture data based on machine learning and triple collocation analysis. Remote Sens. Environ. 298, 113776. https://doi.org/10.1016/j.rse.2023.113776.
Knoben, W.J.M., Freer, J.E., Woods, R.A., 2019. Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23, 4323–4331. https://doi.org/10.5194/hess-23-4323-2019.
Koster, R.D., Liu, Q., Reichle, R.H., Huffman, G.J., 2021. Improved estimates of Pentad Precipitation through the Merging of Independent Precipitation Data Sets. Water Resour. Res. 57, e2021WR030330. https://doi.org/10.1029/2021WR030330.
Krabbenhoft, C.A., Allen, G.H., Lin, P., Godsey, S.E., Allen, D.C., Burrows, R.M., DelVecchia, A.G., Fritz, K.M., Shanafield, M., Burgin, A.J., Zimmer, M.A., Datry, T., Dodds, W.K., Jones, C.N., Mims, M.C., Franklin, C., Hammond, J.C., Zipper, S., Ward, A.S., Costigan, K.H., Beck, H.E., Olden, J.D., 2022. Assessing placement bias of the global river gauge network. Nat Sustain. 5, 586–592. https://doi.org/10.1038/s41893-022-00873-0.
Krapesch, M., Klösch, M., ten Brinke, W., Habersack, H., 2024. The Rhine catchment: a review of sediment-related knowledge, monitoring, and a future research perspective. Water 16, 1121. https://doi.org/10.3390/w16081121.
Kron, W., Tingsanchali, T., Loucks, D.P., Renaud, F.G., Bogardi, J.J., Fekete, A., 2021. Water-related hazard and risk management. https://doi.org/10.1007/978-3-030-60147-8_22. In: Bogardi, J.J., Gupta, J., Nandalal, K.D.W., Salamé, L., van Nooijen, R.R.P., Kumar, N., Tingsanchali, T., Bhaduri, A., Kolechkina, A.G. (Eds.), Handbook of Water Resources Management: Discourses, Concepts and Examples. Springer International Publishing, Cham, pp. 675–734.
Kumar, A., Gosling, S.N., Johnson, M.F., Jones, M.D., Zaherpour, J., Kumar, R., Leng, G., Schmied, H.M., Kupzig, J., Breuer, L., Hanasaki, N., Tang, Q., Ostberg, S., Stacke, T., Pokhrel, Y., Wada, Y., Masaki, Y., 2022. Multi-model evaluation of catchment- and global-scale hydrological model simulations of drought characteristics across eight large river catchments. Adv. Water Resour. 165, 104212. https://doi.org/10.1016/j.advwatres.2022.104212.
Lee, J.S., Choi, H.I., 2022. A rebalanced performance criterion for hydrological model calibration. J. Hydrol. 606, 127372. https://doi.org/10.1016/j.jhydrol.2021.127372.
Lehner, B., Messager, M.L., Korver, M.C., Linke, S., 2022. Global hydro-environmental lake characteristics at high spatial resolution. Sci. Data 9, 351. https://doi.org/10.1038/s41597-022-01425-z.
Lerat, J., Thyer, M., McInerney, D., Kavetski, D., 2021. Advantages of calibrating a daily rainfall-runoff model to monthly streamflow data. EGU21-6913 doi, 10.5194/egusphere-egu21-6913.

Li, C., Tang, G., Hong, Y., 2018. Cross-evaluation of ground-based, multi-satellite and reanalysis precipitation products: Applicability of the Triple Collocation method across Mainland China. J. Hydrol. 562, 71–83. https://doi.org/10.1016/j.jhydrol.2018.04.039.

Liu, C., Zhang, Z., Xu, C., Zhang, W., 2024. Reconstructing long-term, high-resolution groundwater storage changes in the Songhua river basin using supplemented GRACE and GRACE-FO Data. Remote Sens. (Basel) 16, 4566. https://doi.org/10.3390/rs16234566.

Liu, J., Qin, K., Zhen, L., Xiao, Y., Xie, G., 2020. How to allocate interbasin water resources? A method based on water flow in water-deficient areas. Environ. Dev. 34, 100460. https://doi.org/10.1016/j.envdev.2019.100460.

Loritz, R., Hrachowitz, M., Neuper, M., Zehe, E., 2021. The role and value of distributed precipitation data in hydrological models. Hydrol. Earth Syst. Sci. 25, 147–167. https://doi.org/10.5194/hess-25-147-2021.

Lyu, F., Tang, G., Behrangi, A., Wang, T., Tan, X., Ma, Z., Xiong, W., 2021. Precipitation merging based on the triple collocation method across Mainland China. IEEE Trans. Geosci. Remote Sens. 59, 3161–3176. https://doi.org/10.1109/TGRS.2020.3008033.

Mankin, K.R., Mehan, S., Green, T.R., Barnard, D.M., 2025. Review of gridded climate products and their use in hydrological analyses reveals overlaps, gaps, and the need for a more objective approach to selecting model forcing datasets. Hydrol. Earth Syst. Sci. 29, 85–108. https://doi.org/10.5194/hess-29-85-2025.

Mateo, C.M., Hanasaki, N., Komori, D., Tanaka, K., Kiguchi, M., Champathong, A., Sukhapunnaphan, T., Yamazaki, D., Oki, T., 2014. Assessing the impacts of reservoir operation to floodplain inundation by combining hydrological, reservoir management, and hydrodynamic models. Water Resour. Res. 50, 7245–7266. https://doi.org/10.1002/2013WR014845.

McCabe, M.F., Rodell, M., Alsdorf, D.E., Miralles, D.G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N.E.C., Franz, T.E., Shi, J., Gao, H., Wood, E.F., 2017. The future of Earth observation in hydrology. Hydrol. Earth Syst. Sci. 21, 3879–3914. https://doi.org/10.5194/hess-21-3879-2017.

Meema, T., Wattanasetpong, J., Wichakul, S., 2025. Integrating machine learning and zoning-based techniques for bias correction in gridded precipitation data to improve hydrological estimation in the data-scarce region. J. Hydrol. 646, 132356. https://doi.org/10.1016/j.jhydrol.2024.132356.

Melsen, L.A., Puy, A., Torfs, P.J.J.F., Saltelli, A., 2025. The rise of the Nash-Sutcliffe efficiency in hydrology. Hydrol. Sci. J. 70, 1248–1259. https://doi.org/10.1080/02626667.2025.2475105.

Miao, C., Gou, J., Fu, B., Tang, Q., Duan, Q., Chen, Z., Lei, H., Chen, J., Guo, J., Borthwick, A.G.L., Ding, W., Duan, X., Li, Y., Kong, D., Guo, X., Wu, J., 2022. High-quality reconstruction of China's natural streamflow. Science Bulletin 67, 547–556. https://doi.org/10.1016/j.scib.2021.09.022.

Ming, W., Ji, X., Zhang, M., Li, Y., Liu, C., Wang, Y., Li, J., 2022. A Hybrid Triple Collocation-Deep Learning Approach for improving Soil Moisture Estimation from Satellite and Model-based Data. Remote Sens. (Basel) 14, 1744. https://doi.org/10.3390/rs14071744.

Moges, E., Demissie, Y., Larsen, L., Yassin, F., 2021. Review: sources of Hydrological Model Uncertainties and advances in their Analysis. Water 13, 28. https://doi.org/10.3390/w13010028.

Mohammadi, B., 2024. Improving Hydrological Modelling In Cold Regions Using Satellite Remote Sensing And Machine Learning Techniques. Doctoral Thesis (compilation). Lund University, Lund.

Mourad, R., Schoups, G., Bastiaanssen, W., Kumar, D.N., 2024. Expert-based prior uncertainty analysis of gridded water balance components: Application to the irrigated Hindon River Basin, India. J. Hydrol.: Reg. Stud. 55, 101935. https://doi.org/10.1016/j.ejrh.2024.101935.

Müller, O.V., McGuire, P.C., Vidale, P.L., Hawkins, E., 2024. River flow in the near future: a global perspective in the context of a high-emission climate change scenario. Hydrol. Earth Syst. Sci. 28, 2179–2201. https://doi.org/10.5194/hess-28-2179-2024.

Munier, S., Decharme, B., 2022. River network and hydro-geomorphological parameters at 1/12° resolution for global hydrological and climate studies. Earth Syst. Sci. Data 14, 2239–2258. https://doi.org/10.5194/essd-14-2239-2022.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Munpa, P., Kittipongvises, S., Phetrak, A., Sirichokchatchawan, W., Taneepanichskul, N., Lohwacharin, J., Polprasert, C., 2022. Climatic and hydrological factors affecting the assessment of flood hazards and resilience using modified UNDRR indicators: Ayutthaya, Thailand. Water 14, 1603. https://doi.org/10.3390/w14101603.

Nath, N.K., Das, P., Mishra, L.R., Kumar, A., Suryawanshi, S.L., Gautam, V.K., 2024. Hydrological modeling and simulation for water resource assessment. https://doi.org/10.1007/978-3-031-62079-9_3. In: Yadav, A.K., Yadav, K., Singh, V.P. (Eds.), Integrated Management of Water Resources in India: A Computational Approach: Optimizing for Sustainability and Planning. Springer Nature Switzerland, Cham, pp. 43–58.

Nourani, V., Gökçekuş, H., Gichamo, T., 2021. Ensemble data-driven rainfall-runoff modeling using multi-source satellite and gauge rainfall data input fusion. Earth Sci. Inform 14, 1787–1808. https://doi.org/10.1007/s12145-021-00615-4.

Noy, I., Nguyen, C., Patel, P., 2011. Floods and spillovers: households after the 2011 Great Flood in Thailand. Econ. Dev. Cult. Chang. 2021. https://doi.org/10.1086/703098.

Ntona, M.M., Busico, G., Mastrocicco, M., Kazakis, N., 2022. Modeling groundwater and surface water interaction: an overview of current status and future challenges. Sci. Total Environ. 846, 157355. https://doi.org/10.1016/j.scitotenv.2022.157355.

Padiyedath Gopalan, S., Hanasaki, N., Champathong, A., Tebakari, T., 2021. Impact assessment of reservoir operation in the context of climate change adaptation in the Chao Phraya River basin. Hydrol. Process. 35, e14005. https://doi.org/10.1002/hyp.14005.

Padiyedath Gopalan, S., Champathong, A., Sukhapunnaphan, T., Nakamura, S., Hanasaki, N., 2022a. Inclusion of flood diversion canal operation in the H08 hydrological model with a case study from the Chao Phraya River basin: model development and validation. Hydrol. Earth Syst. Sci. 26, 2541–2560. https://doi.org/10.5194/hess-26-2541-2022.

Padiyedath Gopalan, S., Champathong, A., Sukhapunnaphan, T., Nakamura, S., Hanasaki, N., 2022b. Potential impact of diversion canals and retention areas as climate change adaptation measures on flood risk reduction: a hydrological modelling case study from the Chao Phraya River Basin, Thailand. Sci. Total Environ. 841, 156742. https://doi.org/10.1016/j.scitotenv.2022.156742.

Pan, M., Fisher, C.K., Chaney, N.W., Zhan, W., Crow, W.T., Aires, F., Entekhabi, D., Wood, E.F., 2015. Triple collocation: beyond three estimates and separation of structural/non-structural errors. Remote Sens. Environ. 171, 299–310. https://doi.org/10.1016/j.rse.2015.10.028.

Pavelsky, T.M., Durand, M.T., Andreadis, K.M., Beighley, R.E., Paiva, R.C.D., Allen, G.H., Miller, Z.F., 2014. Assessing the potential global extent of SWOT river discharge observations. J. Hydrol. 519, 1516–1525. https://doi.org/10.1016/j.jhydrol.2014.08.044.

Pokavanich, T., Worrawatanathum, V., Phattananuruch, K., Koolkalya, S., 2024. Seasonal dynamics and three-dimensional hydrographic features of the eastern gulf of thailand: insights from high-resolution modeling and field measurements. Water 16, 1962. https://doi.org/10.3390/w16141962.

Pool, S., Francés, F., Garcia-Prats, A., Puertes, C., Pulido-Velazquez, M., Sanchis-Ibor, C., Schirmer, M., Yang, H., Jiménez-Martínez, J., 2021. Hydrological modeling of the effect of the transition from flood to drip irrigation on groundwater recharge using multi-objective calibration. Water Resour. Res. 57, e2021WR029677. https://doi.org/10.1029/2021WR029677.

Qiao, D., Li, Z., Zeng, J., Liang, S., McColl, K.A., Bi, H., Zhou, J., Zhang, P., 2022. Uncertainty characterization of ground-based, satellite, and reanalysis snow depth products using extended triple collocation. Water Resources Research 58, e2021WR030895. https://doi.org/10.1029/2021WR030895.

Quichimbo, E.A., Singer, M.B., Michaelides, K., Rosolem, R., MacLeod, D.A., t. Asfaw, D., Cuthbert, M.O., 2023. Assessing the sensitivity of modelled water partitioning to global precipitation datasets in a data-scarce dryland region. Hydrological Processes 37. https://doi.org/10.1002/hyp.15047.

Raghav, P., Eldho, T.I., 2023. Investigations on the hydrological impacts of climate change on a river basin using macroscale model H08. J. Earth Syst. Sci. 132, 87. https://doi.org/10.1007/s12040-023-02102-4.

Refsgaard, J.C., Stisen, S., Koch, J., 2022. Hydrological process knowledge in catchment modelling – Lessons and perspectives from 60 years development. Hydrol. Process. 36, e14463. https://doi.org/10.1002/hyp.14463.

dos Reis, A.A., Weerts, A., Ramos, M.-H., Wetterhall, F., dos S. Fernandes, W., 2022. Hydrological data and modeling to combine and validate precipitation datasets relevant to hydrological applications. J. Hydrol.: Reg. Stud. 44, 101200. https://doi.org/10.1016/j.ejrh.2022.101200.

Ren, W., Li, X., Zheng, D., Zeng, R., Su, J., Mu, T., Wang, Y., 2023. Enhancing Flood Simulation in Data-Limited Glacial River Basins through Hybrid Modeling and Multi-Source Remote Sensing Data. Remote Sensing 15, 4527. https://doi.org/10.3390/rs15184527.

Rocha, J., Duarte, A., Silva, M., Fabres, S., Vasques, J., Revilla-Romero, B., Quintela, A., 2020. The importance of high resolution digital elevation models for improved hydrological simulations of a mediterranean forested catchment. Remote Sens. (Basel) 12, 3287. https://doi.org/10.3390/rs12203287.

Roure, F., Cloetingh, S., Scheck-Wenderoth, M., Ziegler, P.A., 2009. Achievements and Challenges in Sedimentary Basin Dynamics: A Review. In: New Frontiers in Integrated Solid Earth Sciences. Springer, Dordrecht, pp. 145–233. https://doi.org/10.1007/978-90-481-2737-5_5.

Sahani, J., Kumar, P., Debele, S., Spyrou, C., Loupis, M., Aragão, L., Porcù, F., Shah, M.A. R., Di Sabatino, S., 2019. Hydro-meteorological risk assessment methods and management by nature-based solutions. Sci. Total Environ. 696, 133936. https://doi.org/10.1016/j.scitotenv.2019.133936.

Sheffield, J., Wood, E.F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., Verbist, K., 2018. Satellite remote sensing for water resources management: potential for supporting sustainable development in data-poor regions. Water Resour. Res. 54, 9724–9758. https://doi.org/10.1029/2017WR022437.

Sheikh, M.R., Coulibaly, P., 2024. Review of recent developments in hydrologic forecast merging Techniques. Water 16, 301. https://doi.org/10.3390/w16020301.

Shrestha, R., Tachikawa, Y., Takara, K., 2006. Input data resolution analysis for distributed hydrological modeling. J. Hydrol. 319, 36–50. https://doi.org/10.1016/j.jhydrol.2005.04.025.

Shuai, P., Chen, X., Mital, U., Coon, E.T., Dwivedi, D., 2022. The effects of spatial and temporal resolution of gridded meteorological forcing on watershed hydrological responses. Hydrol. Earth Syst. Sci. 26, 2245–2276. https://doi.org/10.5194/hess-26-2245-2022.

Sidle, R.C., 2021. Strategies for smarter catchment hydrology models: incorporating scaling and better process representation. Geosci. Lett. 8, 24. https://doi.org/10.1186/s40562-021-00193-9.

Singh, S., Mishra, K., Chavan, R., Tiwari, H.L., 2025. Advancements and challenges in hydrological modeling: a comprehensive review. https://doi.org/10.1007/978-981-97-7474-6_32. In: Hydrology and Hydrologic Modelling, Singapore, pp. 423–442.

Slater, L.J., Arnal, L., Boucher, M.-A., Chang, A.-Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R.L., Wood, A., Zappa, M., 2023. Hybrid forecasting: blending climate predictions with AI models. Hydrol. Earth Syst. Sci. 27, 1865–1889. https://doi.org/10.5194/hess-27-1865-2023.

Steyaert, J.C., Sutanudjaja, E., Bierkens, M., Wanders, N., 2025. A data derived workflow for reservoir operations for simulating reservoir operations in a global hydrologic model. Egusphere 1–38. https://doi.org/10.5194/egusphere-2024-3658.

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. J. Geophys. Res. Oceans 103, 7755–7766. https://doi.org/10.1029/97JC03180.

Su, C.-H., Ryu, D., Crow, W.T., Western, A.W., 2014. Beyond triple collocation: applications to soil moisture monitoring. J. Geophys. Res. Atmos. 119, 6419–6439. https://doi.org/10.1002/2013JD021043.

Sun, J., Li, Y.P., Zhuang, X.W., Jin, S.W., Huang, G.H., Feng, R.F., 2018a. Identifying water resources management strategies in adaptation to climate change under uncertainty. Mitig. Adapt Strateg. Glob Change 23, 553–578. https://doi.org/10.1007/s11027-017-9749-9.

Sun, R., Hernández, F., Liang, X., Yuan, H., 2020. A calibration framework for high-resolution hydrological models using a multiresolution and heterogeneous strategy. Water Resour. Res. 56, e2019WR026541. https://doi.org/10.1029/2019WR026541.

Sun, W., Fan, J., Wang, G., Ishidaira, H., Bastola, S., Yu, J., Fu, Y.H., Kiem, A.S., Zuo, D., Xu, Z., 2018b. Calibrating a hydrological model in a regional river of the Qinghai–Tibet plateau using river water width determined from high spatial resolution satellite images. Remote Sens. Environ. 214, 100–114. https://doi.org/10.1016/j.rse.2018.05.020.

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., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wisser, D., Bierkens, M.F.P., 2018. PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geosci. Model Dev. 11, 2429–2453. https://doi.org/10.5194/gmd-11-2429-2018.

Suwannachai, L., Phumiphan, A., Kuntiyawichai, K., Supakosol, J., Sriworamas, K., Sivanpheng, O., Kangrang, A., 2025. Integrating hydrological models for improved flash flood risk assessment and mitigation strategies in Northeastern Thailand. Water 17, 345. https://doi.org/10.3390/w17030345.

Tan, M.L., Gassman, P.W., Yang, X., Haywood, J., 2020. A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes. Adv. Water Resour. 143, 103662. https://doi.org/10.1016/j.advwatres.2020.103662.

Tan, S.-Q., Guo, H.-F., Liao, C.-H., Ma, J.-H., Tan, W.-Z., Peng, W.-Y., Fan, J.-Z., 2024. Collocation-analyzed multi-source ensembled wind speed data in lake district: a case study in Dongting Lake of China. Front. Environ. Sci. 11. https://doi.org/10.3389/fenvs.2023.1287595.

Tang, G., Clark, M.P., Knoben, W.J.M., Liu, H., Gharari, S., Arnal, L., Beck, H.E., Wood, A.W., Newman, A.J., Papalexiou, S.M., 2023. The impact of meteorological forcing uncertainty on hydrological modeling: a global analysis of cryosphere basins. Water Resour. Res. 59, e2022WR033767. https://doi.org/10.1029/2022WR033767.

Tangdamrongsub, N., 2023. Comparative analysis of global terrestrial water storage simulations: assessing CABLE, Noah-MP, PCR-GLOBWB, and GLDAS Performances during the GRACE and GRACE-FO Era. Water 15, 2456. https://doi.org/10.3390/w15132456.

Tangdamrongsub, N., Steele-Dunne, S.C., Gunter, B.C., Ditmar, P.G., Sutanudjaja, E.H., Sun, Y., Xia, T., Wang, Z., 2017. Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model. Hydrol. Earth Syst. Sci. 21, 2053–2074. https://doi.org/10.5194/hess-21-2053-2017.

Thaisiam, W., Yomwilai, K., Wongchaisuwat, P., 2024. Utilizing sequential modeling in collaborative method for flood forecasting. J. Hydrol. 636, 131290. https://doi.org/10.1016/j.jhydrol.2024.131290.

Tian, F., Wu, B., Zeng, H., Zhang, M., Zhu, W., Yan, N., Lu, Y., Li, Y., 2025. GMIE: a global maximum irrigation extent and central pivot irrigation system dataset derived via irrigation performance during drought stress and deep learning methods. Earth Syst. Sci. Data 17, 855–880. https://doi.org/10.5194/essd-17-855-2025.

Tsamalis, C., 2022. Clarifications on the equations and the sample number in triple collocation analysis using SST observations. Remote Sens. Environ. 272, 112936. https://doi.org/10.1016/j.rse.2022.112936.

Tu, T., Wang, J., Zhao, G., Zhao, T., Dong, X., 2024. Scaling from global to regional river flow with global hydrological models: Choice matters. J. Hydrol. 633, 130960. https://doi.org/10.1016/j.jhydrol.2024.130960.

VanRheenen, N.T., Wood, A.W., Palmer, R.N., Lettenmaier, D.P., 2004. Potential implications of PCM climate change scenarios for Sacramento–San Joaquin river basin hydrology and water resources. Clim. Change 62, 257–281. https://doi.org/10.1023/B:CLIM.0000013686.97342.55.

Vogelzang, J., Stoffelen, A., 2021. Quadruple collocation analysis of in-situ, scatterometer, and NWP winds. J. Geophys. Res. Oceans 126, e2021JC017189. https://doi.org/10.1029/2021JC017189.

Wada, Y., Bierkens, M.F.P., de Roo, A., Dirmeyer, P.A., Famiglietti, J.S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T.J., van Dijk, A.I.J.M., van Emmerik, T., Van Huijgevoort, M.H.J., Van Lanen, H.A.J., Vörösmarty, C.J., Wanders, N., Wheater, H., 2017. Human–water interface in hydrological modelling: current status and future directions. Hydrol. Earth Syst. Sci. 21, 4169–4193. https://doi.org/10.5194/hess-21-4169-2017.

Wang, D., Yu, X., Jia, G., Wang, H., 2019. Sensitivity analysis of runoff to climate variability and land-use changes in the Haihe Basin mountainous area of north China. Agr Ecosyst Environ 269, 193–203. https://doi.org/10.1016/j.agee.2018.09.025.

Wang, H., Yuan, Q., Zhao, H., Xu, H., 2022. In-situ and triple-collocation based assessments of CYGNSS-R soil moisture compared with satellite and merged estimates quasi-globally. J. Hydrol. 615, 128716. https://doi.org/10.1016/j.jhydrol.2022.128716.

Wang, J., Hou, B., Jiang, D., Xiao, W., Wu, Y., Zhao, Y., Zhou, Y., Guo, C., Wang, G., 2016. Optimal allocation of water resources based on water supply security. Water 8, 237. https://doi.org/10.3390/w8060237.

Wanzala, M.A., Ficchi, A., Cloke, H.L., Stephens, E.M., Badjana, H.M., Lavers, D.A., 2022. Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: a case study of Kenya. J. Hydrol.: Reg. Stud. 41, 101105. https://doi.org/10.1016/j.ejrh.2022.101105.

Waqas, M., Humphries, U.W., Hlaing, P.T., Wangwongchai, A., Dechpichai, P., 2024. Advancements in daily precipitation forecasting: a deep dive into daily precipitation forecasting hybrid methods in the Tropical climate of Thailand. MethodsX 12, 102757. https://doi.org/10.1016/j.mex.2024.102757.

Wheater, H.S., Pomeroy, J.W., Pietroniro, A., Davison, B., Elshamy, M., Yassin, F., Rokaya, P., Fayad, A., Tesemma, Z., Princz, D., Loukili, Y., DeBeer, C.M., Ireson, A. M., Razavi, S., Lindenschmidt, K.-E., Elshorbagy, A., MacDonald, M., Abdelhamed, M., Haghnegahdar, A., Bahrami, A., 2022. Advances in modelling large river basins in cold regions with Modélisation Environmentale Communautaire—Surface and Hydrology (MESH), the Canadian hydrological land surface scheme. Hydrol. Process. 36, e14557. https://doi.org/10.1002/hyp.14557.

Williams, G.P., 2025. Friends don't let friends use Nash-Sutcliffe Efficiency (NSE) or KGE for hydrologic model accuracy evaluation: a rant with data and suggestions for better practice. Environ. Model. Software 194, 106665. https://doi.org/10.1016/j.envsoft.2025.106665.

Wojkowski, J., Wałęga, A., Lepeška, T., Młyński, D., Radecki-Pawlik, A., 2024. Rich north, poor south - Regionalization of European water retention: the landscape hydric potential concept. Sci. Total Environ. 954, 176494. https://doi.org/10.1016/j.scitotenv.2024.176494.

Worden, M.A., Bilir, T.E., Bloom, A.A., Fang, J., Klinek, L.P., Konings, A.G., Levine, P.A., Milodowski, D.T., Quetin, G.R., Smallman, T.L., Bar-On, Y.M., Braghiere, R.K., David, C.H., Fischer, N.A., Gentine, P., Green, T.J., Jones, A., Liu, J., Longo, M., Ma, S., Magney, T.S., Massoud, E.C., Myrgiotis, V., Norton, A.J., Parazoo, N., Tajfar, E., Trugman, A.T., Williams, M., Worden, S., Zhao, W., Zhu, S., 2025. Combining observations and models: a review of the CARDAMOM framework for data-constrained terrestrial ecosystem modeling. Glob. Chang. Biol. 31, e70462. https://doi.org/10.1111/gcb.70462.

Wu, Z., Cui, N., Zhang, W., Liu, C., Jin, X., Gong, D., Xing, L., Zhao, L., Wen, S., Yang, Y., 2024. Estimating soil moisture content in citrus orchards using multi-temporal sentinel-1A data-based LSTM and PSO-LSTM models. J. Hydrol. 637, 131336. https://doi.org/10.1016/j.jhydrol.2024.131336.

Xu, J., Ma, Z., Yan, S., Peng, J., 2022. Do ERA5 and ERA5-land precipitation estimates outperform satellite-based precipitation products? A comprehensive comparison between state-of-the-art model-based and satellite-based precipitation products over Mainland China. J. Hydrol. 605, 127353. https://doi.org/10.1016/j.jhydrol.2021.127353.

Xu, Z., Liu, J., Wu, Z., Guo, X., 2024. Enhancing streamflow simulation accuracy in ungauged catchments via parameter calibration with triple collocation-based merged evapotranspiration and streamflow features. J. Hydrol. 639, 131627. https://doi.org/10.1016/j.jhydrol.2024.131627.

Yang, H., Xiong, L., Liu, D., Cheng, L., Chen, J., 2021. High spatial resolution simulation of profile soil moisture by assimilating multi-source remote-sensed information into a distributed hydrological model. J. Hydrol. 597, 126311. https://doi.org/10.1016/j.jhydrol.2021.126311.

Ye, Q., Li, Y., Zhuo, L., Zhang, W., Xiong, W., Wang, C., Wang, P., 2018. Optimal allocation of physical water resources integrated with virtual water trade in water scarce regions: a case study for Beijing, China. Water Res. 129, 264–276. https://doi.org/10.1016/j.watres.2017.11.036.

Yilmaz, M.T., Crow, W.T., 2014. Evaluation of assumptions in soil moisture triple collocation. Analysis. https://doi.org/10.1175/JHM-D-13-0158.1.

Yoshida, T., Hanasaki, N., Nishina, K., Boulange, J., Okada, M., Troch, P.A., 2022. Inference of parameters for a global hydrological model: identifiability and predictive uncertainties of climate-based parameters. Water Resour. Res. 58, e2021WR030660. https://doi.org/10.1029/2021WR030660.

Zaherpour, J., Gosling, S.N., Mount, N., Schmied, H.M., Veldkamp, T.I.E., Dankers, R., Eisner, S., Gerten, D., Gudmundsson, L., Haddeland, I., Hanasaki, N., Kim, H., Leng, G., Liu, J., Masaki, Y., Oki, T., Pokhrel, Y., Satoh, Y., Schewe, J., Wada, Y., 2018. Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts. Environ. Res. Lett. 13, 065015. https://doi.org/10.1088/1748-9326/aac547.

Zhang, X., Qu, Y., Ma, M., Liu, H., Su, Z., Lv, J., Peng, J., Leng, G., He, X., Di, C., 2020. Satellite-based operational real-time drought monitoring in the transboundary Lancang–Mekong River Basin. Remote Sens. (Basel) 12, 376. https://doi.org/10.3390/rs12030376.

Zhang, Z., Zeng, Y., Huang, Z., Liu, J., Yang, L., 2023. Multi-source data fusion and hydrodynamics for urban waterlogging risk identification. Int. J. Environ. Res. Public Health 20, 2528. https://doi.org/10.3390/ijerph20032528.

Zhao, H., Montzka, C., Vereecken, H., Franssen, H.-J.-H., 2024. A comparative analysis of remote sensing soil moisture datasets fusion methods: novel LSTM approach versus widely used triple collocation technique. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 17, 16659–16671. https://doi.org/10.1109/JSTARS.2024.3455549.

Zheng, F., Yin, H., Zhang, J., Duan, H.-F., Gupta, H.V., 2024. A Bayesian deep learning approach for video-based estimation and uncertainty quantification of urban rainfall intensity. J. Hydrol. 640, 131706. https://doi.org/10.1016/j.jhydrol.2024.131706.

Zhou, J., Crow, W.T., Wu, Z., Dong, J., He, H., Feng, H., 2021. A triple collocation-based 2D soil moisture merging methodology considering spatial and temporal non-stationary errors. Remote Sens. Environ. 263, 112509. https://doi.org/10.1016/j.rse.2021.112509.

Zhu, L., Li, W., Wang, H., Deng, X., Tong, C., He, S., Wang, K., 2023. Merging Microwave, Optical, and Reanalysis Data for 1 Km Daily Soil Moisture by Triple Collocation. Remote Sens. (Basel) 15, 159. https://doi.org/10.3390/rs15010159.

Zipper, S.C., Farmer, W.H., Brookfield, A., Ajami, H., Reeves, H.W., Wardropper, C., Hammond, J.C., Gleeson, T., Deines, J.M., 2022. Quantifying streamflow depletion from groundwater pumping: a practical review of past and emerging approaches for water management. JAWRA J. Am. Water Resour. Assoc. 58, 289–312. https://doi.org/10.1111/1752-1688.12998.

Zwieback, S., Scipal, K., Dorigo, W., Wagner, W., 2012. Structural and statistical properties of the collocation technique for error characterization. Nonlinear Processes Geophys. 19, 69–80. https://doi.org/10.5194/npg-19-69-2012.