www.nature.com/scientificdata

scientific data

OPEN

DATA DESCRIPTOR

High-resolution gridded streamflow data for Ganges-Brahmaputra-Meghna River Basins in Bangladesh (1951–2023)

Byeong-Hee Kim [1,2], Sazzad Hossain [3,4] & Jonghun Kam [1] ✉

A 9-km daily gridded streamflow dataset is generated using the Variable Infiltration Capacity-River Routing Model (VIC-RRM) across the Ganges-Brahmaputra-Meghna River basins over 1951–2023, forced by the ERA5-Land reanalysis data for naturalized streamflow. Physically consistent streamflow forecast data is also generated forced by the ECMWF S2S forecasts. The performance of the dataset is evaluated using observed streamflow data from three gauge stations in Bangladesh along the streams of Ganges, Brahmaputra, and Meghna Rivers, calculating the modified Kling-Gupta Efficiency (mKGE) metric for the 365-day climatology. For Ganges, Brahmaputra, and Meghna Rivers, the mKGE values of reconstructed streamflow data are 0.50, 0.75, and 0.25, respectively. Comparing with the reconstructed streamflow data, the streamflow forecasts show a good agreement with mKGE values of 1.00, 0.97, and 0.91 at three gauge stations, respectively. This dataset provides physically consistent reconstructed and forecasted streamflow data at high resolution, offering a valuable resource for the assessment of climate variability and change and the development of river basin-specific water management strategies in the Ganges-Brahmaputra-Meghna Rivers in Bangladesh.

Background & Summary

Streamflow is a critical element of the hydrological cycle, representing the movement and distribution of water influenced by the regional climate and characteristics of the river basin. This flow of water in streams and rivers is essential not only for supporting human activities such as agriculture, industry, and municipal use, but also for maintaining the health and functionality of ecosystems[1]. Given its importance, the efficient and effective management of water resources depends heavily on comprehensive and long-term assessments of streamflow[2]. Such long-term assessments are indispensable as they provide the necessary data to analyze river variability over time, assess the impacts of climate change, and implement measures for environmental conservation[3]. However, a significant challenge in conducting these assessments is the sparsity of streamflow records, both spatially and temporally. Many regions are lacking of sufficient gauge stations that provide continuous and extensive spatial/temporal coverages, particularly in financially disadvantaged countries. Limited available streamflow data often hinder accurate hydrologic predictions and proactive water resources management.

Bangladesh is located in the Ganges-Brahmaputra-Meghna River Delta (Fig. 1), which is the 7th most disaster-prone country by climate change in the world during 2000–2019[4]. Due to its geographical location, Bangladesh is vulnerable to climate change with many extreme weather events such as floods, cyclones, droughts, and coastal erosion affecting the country almost every year[5]. These extreme weather events have caused significant damage to human lives, livelihoods, crops, livestock, and infrastructure[4]. The geographical location of Bangladesh exposes to the highest risk of flood hazard among South Asia countries. Approximately 20–30% of the country is flooded each year. Over the last decade, a combination of climate change and growing population and human activity on floodplains has made floods more frequent and acute. In a changing climate,

[1]Division of Environmental Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Republic of Korea. [2]Human Resource Program for Sustainable Environment in the 4th Industrial Revolution Society, Pohang University of Science and Technology (POSTECH), Pohang, Republic of Korea. [3]Flood forecasting and Warning Centre, BWDB, Dhaka, Bangladesh. [4]Department of Geography and Environmental Science, University of Reading, Reading, UK. ✉e-mail: jhkam@postech.ac.kr

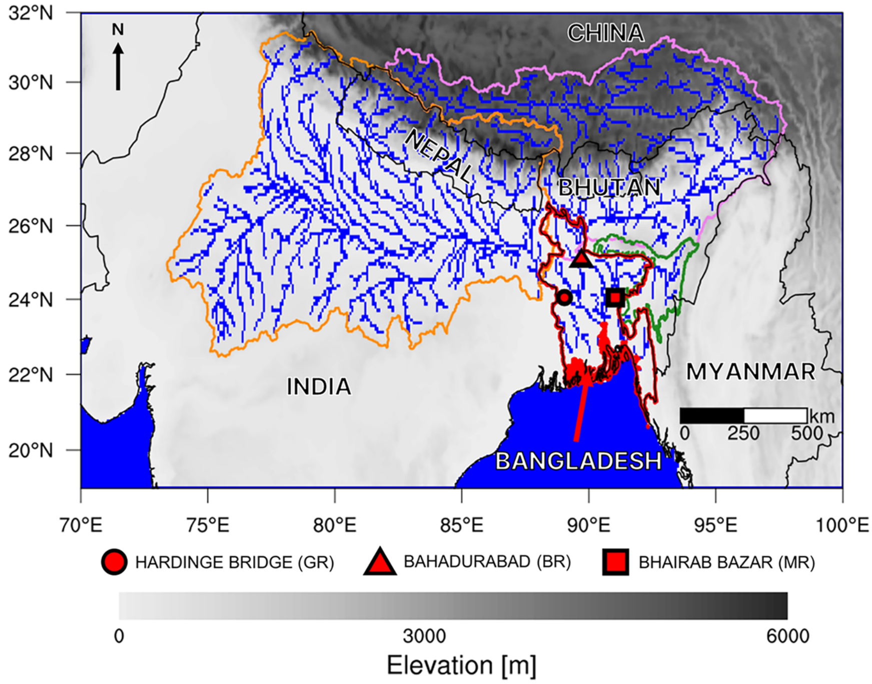

Fig. 1 Domain for the Ganges-Brahmaputra-Meghna River basins and Bangladesh. Upstream basins for Ganges (orange line), Brahmaptura (pink line), and Meghna (green line) where rivers flow into Bangladesh. Blue lines denote river networks within the Ganges-Brahmaptura-Meghna River basins, while red lines indicate the national boundary of Bangladesh. Red markers including circle, triangle, and square are denote gauge stations (Hardinge Bridge, Bahadurabad, and Bhairab Bazar) for Ganges River (GR), Brahmaputra River (BR), and Meghna River (MR) within Bangladesh.

the behavior of flooding in Bangladesh is likely to change and more frequent extreme events are expected with varying intensity. Bangladesh continuously required a hydrologic monitoring and forecasting system for flood control and mitigation.

To mitigate flood damage, Bangladesh has operated the Flood Forecasting and Warning Centre (FFWC) of Bangladesh Water Development Board (BWDB) under the Ministry of Water Resources (MoWR) in 1972. FFWC is fully operated from April to October every year, following the Standing Orders on Disaster (SOD) of the Government of Bangladesh[6]. The FFWC provides short-term (3-days and 5-days) definitive flood forecast, median-term (10-days and 15-days) probabilistic flood forecast, rainfall maps, flood inundation map, flood warning message service. Despite the operation of FFWC, flood damage continues to occur every year since dynamical forecasts in long range are missing, which cannot account for unprecedented floods. Probabilistic/stochastic flood forecasts at a specific location require long-term observational data. Long-term observational records of daily streamflow are available at only three gauge stations at Bahadurabad, Hardinge Bridge, and Bhairab Bazar in Bangladesh (red markers in Fig. 1). Bangladesh has faced limited financial and human resources to build and maintain a national observational network system for hydrologic monitoring. These stations are however at a critical location for data validation because they are located at the downstream of Brahmaputra River (BR), Meghna River (MR) and Ganges River (GR) from neighbor countries including Bhutan, China and India. The observational records of daily streamflow at three gauge stations were provided by BWDB.

We explored other observational data sources for daily streamflow within Bangladesh from the Global Runoff Data Centre (GRDC)[7], which provides the global observational records of streamflow and hydrologic characteristics, such as drainage area and stream lies, and found that GRDC provides only two stations at Bahadurabad and Hardinge Bridge with a shorter temporal coverage than those used in this study (Figure S1). Spatiotemporally limited streamflow records hinder the assessment of long-term change in streamflow and available water resources, which can provide actionable information for water policy makers and resource managers in Bangladesh. To enhance the resiliency of Bangladesh to unprecedented floods in a changing climate, it is essential to reconstruct long-term streamflow data that addresses spatial and temporal limitations in insufficient streamflow monitoring data.

To overcome the spatial and temporal limitations of observations at gauge stations, the water resources community has invested considerable effort in developing advanced hydrologic models[8]. A particularly notable advancement in this field is the development of high-resolution hydrologic modeling[9]. These models operate at very fine spatial and temporal resolutions, providing detailed and accurate simulations of streamflow patterns. High-resolution models can capture the nuances of hydrological processes at scales previously unattainable, offering a powerful tool for researchers and water resource managers. Hydrologic models can estimate the naturalized streamflow data without human activity. Naturalized streamflow can be used as a reference to estimate

Station name	River name	Latitude	Longitude	Time range (YYYY.MM.DD) GRDC	BWDB
Bahadurabad	Brahmaputra	25.16	89.70	1985.04.03-1992.03.31	1994.01.01-2023.12.31
Hardinge Bridge	Ganges	24.06	89.03	1985.04.03-1992.03.31	1994.01.01-2023.12.31
Bhairab Bazar	Meghna	24.05	91.00	—	1994.01.01-2023.12.31

Table 1. Informations for gauge stations in the Ganges-Brahmaputra-Meghna River basins.

the hydrological response to climate impact, assess the ecological status of rivers[10], and determine the available amount of water. In terms of water resource management, naturalized streamflow can be used in various water management scenarios and the impact of these scenarios on the quantity of available water[11,12] including supplying water for various economic needs such as industry, agriculture, and electricity. In the perspective of climate variability and change, naturalized streamflow usually is used a baseline to calibrate hydrological models due to complex to address the issues of climate changes and water-use changes at the same time[13,14]. Hydrological models considered physical process as topographic slope and river network.

Here we aim to reconstruct long-term naturalized streamflow data and generate physically consistent long-range hydrologic forecasts for the Ganges-Brahmaputra-Meghna Rivers in Bangladesh to overcome spatiotemporally limited observational records of daily streamflow. We use the Variable Infiltration Capacity-River routing model (VIC-RRM) that is a physical-based macro-scale hydrological model. Most of macro-scale hydrological models explicitly represent natural processes due to the lacking of data on dams, reservoirs, and land cover changes[15]. As a result, they simulate naturalized streamflow, driven by the long-term meteorological forcing data, which is physically consistent with extended records of precipitation[16,17]. In this study, the VIC-RRM is simulated facilitating the high-resolution (<10 kilometers) digital elevation models (DEMs) that reflect natural topography. The VIC-RRM generates historical long-term naturalized streamflow data using the ERA5-Land analysis product[18] and the forecasted streamflow data using the European Centre for Medium-Range Weather Forecasts (ECMWF) sub-seasonal to seasonal (S2S) forecasts data[19] across the Ganges-Brahmaputra-Meghna River basins.

Methods

Datasets. Long-term observed daily streamflow data (in cubic meters per second) from 1994 to 2023 was collected at the Bahadurabad, Hardinge Bridge, Bhairab Bazar gauge stations on the Brahmaputra River (BR), Ganges River (GR), and Meghna River (MR), with support from the BWDB in Bangladesh. In addition, observed daily streamflow data before 1994 at the Bahadurabad and Hardinge Bridge stations were obtained from the GRDC data download website (https://portal.grdc.bafg.de/applications/public.html?publicuser=PublicUser) by the search-related the interested river names (i.e the BM, GM, and MR). The location information and recorded time range for three gauge stations are shown in Table 1. These observational datasets were utilized to evaluate the performance of the simulated daily streamflow data from the VIC-RRM.

For the simulated streamflow, surface and sub-surface hourly runoff data were derived from the ERA5-Land reanalysis product (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download)[18], covering the period from 1951 to 2023. These hourly runoff datasets were converted to daily datasets to be used as forcing data in the VIC-RRM. The ERA5-Land (hereafter, ERA5-L) reanalysis product, consistent with meteorological data from ERA5, emphasizes the evolution of land variables over several decades and offers an enhanced resolution of 9-km, compared to the 25-km resolution of ERA5.

To generate the daily forecast streamflow data, we obtain the ECMWF S2S total and sub-surface daily runoff forecast data from the 50 ensemble runs with a resolution of 150-km (1.5 degree) during 2016–2023. We calculate the surface runoff forecast data as the difference between total runoff and sub-surface runoff forecast datasets because the VIC-RRM requires daily surface and sub-surface runoff data as hydrologic forcing data. The Subseasonal to Seasonal (S2S) Prediction project[19] has been established by the World Weather Research Programme/World Climate Research Programme to improve and evaluate forecast skill at the subseasonal to seasonal time range. The participant centers provide near-real-time ensemble forecasts up to 60 days using their fully coupled numerical forecast models. Among 11 national centers involved in the S2S project, the ECMWF is is the only center that provides total runoff and sub-surface runoff forecast data with 50 ensemble members from 2016 to the present. Therefore, we choose the ECMWF S2S forecast data in this study. All the S2S data can be accessed by hypertext transport protocol (https://apps.ecmwf.int/datasets/data/s2s-realtime-instantaneous-accum-ecmf/levtype=sfc/type=pf/). Users can register and visit the data portal to browse the contents of the available varaioables, and download the data of interest through the ECMWF web API.

Lastly, we obtain the DEM data from the SRTM 90-meter DEM v4 (https://srtm.csi.cgiar.org)[20] and linear interpolate that to a resolution of 9-km that is consistent with the resolution of the ERA5-Land product.

VIC-River routing model. VIC-RRM is a source-to-sink model based on the linearized Saint-Venant equations[21,22] and uses topography data without consideration of irrigation and reservoir. Simulated streamflow data represents natural flow conditions along a river network caused mainly by rainfall and snow-/glacier-melt runoff due to high temperature and solar radiation effects. The river network is derived from the altitude of each grid within the basin drainage area. Human impacts such as the effect of dam regulation can modulate streamflow and available water resources in a natural river[23]. It have been known that the presence of a dam upstream of a

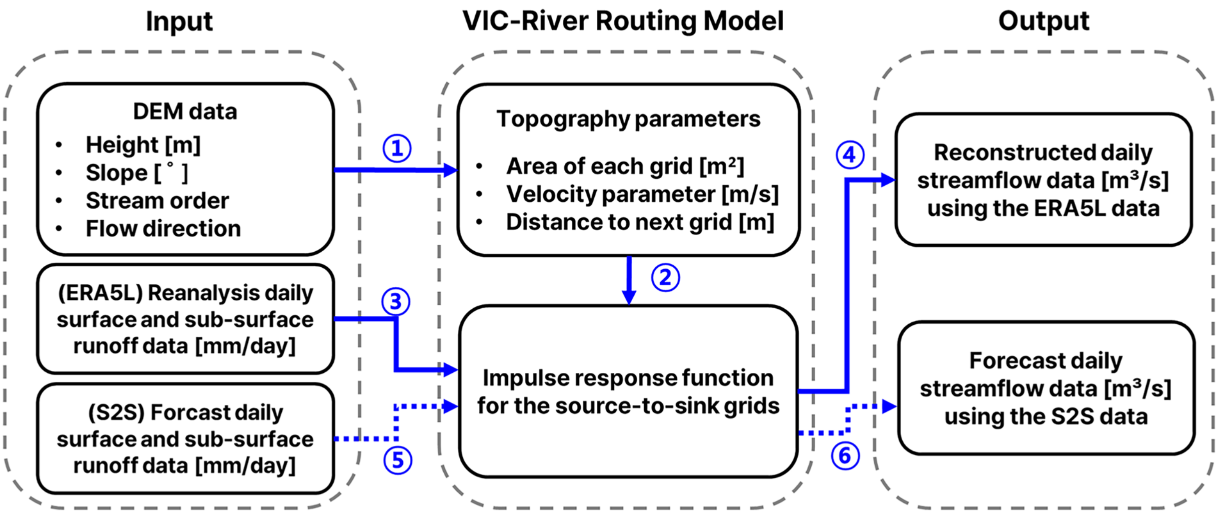

Fig. 2 Schematic overview for the reconstructed and forecast daily streamflow in the Ganges-Brahmaputra-Meghna River basins with the surface and sub-surface runoff data from the reanalysis data (ERA5-L) and forecast data (S2S).

gauge has a greater effect in offsetting natural water flow variability than the presence of a natural lake[16,24]. Some river routing models can assess flood occurrence and impacts at the catchment scale by incorporating watershed characteristics and hydrological calculations that take into account the impacts of dams and reservoirs[25,26]. However, at the regional/continental scales, channel models using floodplain inundation dynamics still have large uncertainties in river deltas as like Bangladesh[27] that originated from lacking of detailed information of hydrologic structures and conditions. Furthermore, impact assessment of local human disturbance, particularly the dam impact[28], remains uncertain, particularly over data sparse regions. In contrast, VIC-RRM is sufficient for this study to reconstruct long-term naturalized streamflow data because this model is physically consistent with precipitation and snow/glacier melt as a natural hydrological response.

VIC-RRM assumes water can exit a grid cell in one direction through at least one river among the eight adjacent grid cells, adding this water to the downstream grid cells in the river network. The model presumes the runoff transport process to be linear and time-invariant, with a non-negative impulse response function (IRF). Thus, the IRF between any source and sink grid points depends only on the horizontal travel time of water within the source grid cell and to the downstream point, including a flow diffusion parameter. The resolution of the DEM determines the spatial resolution of streamflow simulated by VIC-RRM used to generate the geographical information of the watershed in question.

This study used 9-km resolution DEM data over the globe. The first step is that each grid cell's height, slope, stream order, and flow direction are obtained from the DEM, and the topography parameters including each grid area, distance to the next grid, and water velocity are objectively calculated. The second step is that the IRFs are developed for each grid cell based on the 9-km resolution of DEM. If the horizontal resolution of DEM is higher than that of input data (i.e. runoff data), VIC-RRM can initially scale up and consolidate the high-resolution IRF grid to match the resolution of input data. The upscaling process uses the first-order conservative remapping technique[29]. The unit response to input data is maintained due to the preservation of the remapping scheme. If the resolution of the DEM and the runoff data are the same, the remapping scheme did not work and was ignored. In the third step, the routing process is repeated for each specified flow location within the VIC-RRM domain using the input runoff data and the IRFs developed for each grid cell. This process integrates the IRFs to include all watersheds flowing into a grid cell. After developing the IRFs for the high-resolution river network, flow synthesis involves aggregating the flow contributions from all upstream grid cells at each time step, but with delays according to the IRFs. This synthesis method of VIC-RRM accounts for the fact that only a portion of each grid cell's flow reaches the downstream point at each time step, and as the process continues, the outflow reaching the sea in future time steps is added. Finally, the daily streamflow data calculated by the repeated routing process are generated as grid data with the same resolution as the DEM.

Workflow. Figure 2 shows the workflow in this study. We reconstructed the long-term naturalized daily streamflow data for the drainage areas of Ganges-Brahmaputra-Meghna Rivers and Bangladesh (70–100°E, 20–35°N) using the VIC-RRM forced by daily surface and sub-surface runoff data from the ERA5-L product[18]. VIC-RRM was simulated at 9-km resolution based on the high-resolution DEMs for the interesting area. The forced datasets for VIC-RRM are surface and sub-surface daily runoff (units: mm day^{-1}). We obtained the hourly surface and sub-surface runoff data from the ERA5-L product over 1951–2023 and they were calculated the daily cumulative surface and sub-surface runoff data. In the last step, the VIC-RRM simulated long-term daily streamflow in sixty-five thousand grid cells using input datasets including DEM and ERA5-L daily runoff data. In this case, the model was run during the period of ERA5-L (1951–2023).

We compared simulated daily streamflow data against observed daily streamflow records at 3 gauge stations (Bahadurabad, Hardinge Bridge, and Bhairab Bazar) in Bangladesh. We assessed the model's performance using the modified Kling-Gupta Efficiency (mKGE)[30,31] and root-mean-square-error (RMSE) for climatological daily streamflow. The mKGE is composed temporal errors, bias errors, and variability errors:

$$r=\frac{\sum_{t=1}^{n}(X_s(t)-\mu_s)(X_o(t)-\mu_o)}{\sqrt{\sum_{t=1}^{n}(X_s(t)-\mu_s)^2}\sqrt{\sum_{i=1}^{n}(X_s(t)-\mu_s)^2}}, \tag{1}$$

$$\beta=\frac{\mu_s}{\mu_o}, \tag{2}$$

$$\gamma=\frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}, \tag{3}$$

$$\mathrm{mKGE}=1-\sqrt{(r-1)^2+(\beta-1)^2+(\gamma-1)^2}, \tag{4}$$

where r is the Pearson correlation coefficient between simulation (s) and observation (o), β is the bias ratio, γ is the variability ratio, μ is the mean streamflow, and σ is the standard deviation of streamflow. The interpretation of mKGE is easy since the value of mKGE is the lower limit of the three components (r, β, and γ in Eq. 4). The model performance is perfect when the value of mKGE is 1. The RMSE, which allows for an intuitive interpretation of the error magnitude between observed and simulated streamflow, is given below.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{t=1}^{n}(X_o(t)-X_s(t))^2}, \tag{5}$$

The RMSE value is zero if the model has no error, that is, it is a perfect model.

In addition, the VIC-RRM can be used to produce daily streamflow forecast data for more than 10 days using predicted runoff data as input. In this study, the ECMWF S2S (hereafter, S2S) prediction data[19] was used to generate the predicted streamflow data. We used the surface and sub-surface runoff daily forecast data in the S2S forecast from 2016 to 2023 at two-week intervals (14 days) as the input data for VIC-RRM to generate streamflow forecast data. The performance of this forecasted streamflow data was evaluated by comparing it with the reconstructed streamflow data based on the ERA5-L because the generated streamflow datasets from ERA5-L and S2S using VIC-RRM are naturalized streamflow.

Data Records

High-resolution (9-km) historical daily naturalized streamflow datasets by the ERA5-L product over 1951–2023 and forecasted streamflow datasets by the S2S forecast product over 2016–2023 for the Ganges-Brahmaputra-Meghna Rivers areas (70–100°E, 20–35°N) are freely available from Harvard Dataverse (https://doi.org/10.7910/DVN/V2C6G2)[32]. The data was recorded spatially, covering the land area of the Ganges-Brahmaputra-Meghna River areas including China, India, Nepal, Bhutan, and Bangladesh. Grid streamflow datasets were generated daily and saved in netcdf files.

Technical Validation

We evaluated the performance of VIC-RRM for reconstructed streamflow in terms of the 365-day climatology for the 3 gauge stations for BR, GR, and MR in Bangladesh (red markers in Fig. 1) during 1994–2023. Figure 3 shows the 365-day climatological observed and reconstructed streamflow at 3 gauge stations. The reconstructed streamflow using the ERA5-L by VIC-RRM is overestimated compared to observations. All three stations show strong variability of observed and reconstructed daily streamflow from wet months (May through November), rather than dry months (December through April). However, reconstructed data show overestimation of variability of daily streamflow from May through November, indicating large uncertainty in wet months of Bangladesh. The surface and sub-surface runoff are used as forcing data for the VIC-RRM model simulation and are estimated mainly by precipitation. A possible uncertainty source of overestimation in wet months are uncertainty of precipitation from the ERA5 reanalysis product, which requires a further validation assessment of ERA5 precipitation data, particularly during wet months.

For quantified statistical evaluation, we used the mKGE and RMSE analyses to assess the performance of the reconstructed streamflow. The mKGE comprehensively evaluates model performance by assessing the temporal accuracy, bias, and variability of model predictions, providing a holistic assessment of model performance across various aspects. r, β, and γ components in mKGE (i.e. Equation 4) are related to temporal errors, bias erros, and variability errors, respectively. The closer these three components are to 1, the model has good performance. In addition, by comparing the three components, we can identify the dominant errors in the model's performance. The value of RMSE denotes the magnitude of errors between two different datasets.

mKGE values for reconstructed streamflow at the Bahadurabad, Hardinge Bridge, and Bhairab Bazar gauge stations are 0.75, 0.50, and 0.25. This result shows that VIC-RRM can simulate the streamflow at the Bahadurabad station better than at the other stations. In terms of the temporal accuracy (r), three gauge stations have 0.99, 0.88, and 0.95. This means that the temporal errors do not differ significantly between three stations. Values of β are 1.25, 1.46, and 0.91 at three stations, while that of γ are 1.03, 0.83, and 1.75. According to these results, the dominant errors of Bahadurabad and Hardinge Bridge are the bias errors, while that of Bhairab Bazar

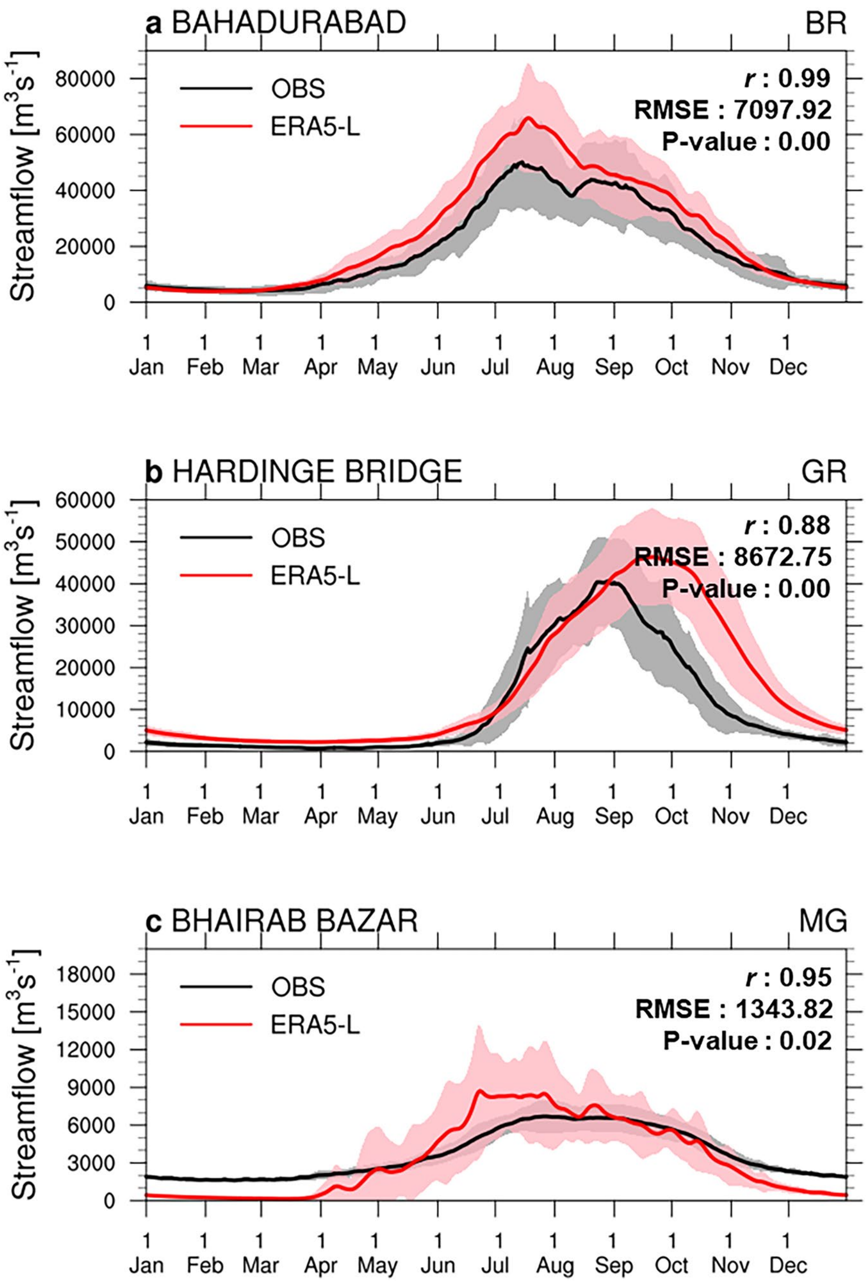

Fig. 3 Daily climatological observed (OBS, black line) and reconstructed (ERA5-L, red line) streamflow over 1994–2023 at Bahadurabad, Hardinge Bridge, and Bhairab Bazar. The correlation coefficient (r) and RMSE between climatological observed and reconstructed streamflow are denoted on the top-right corner in each panel plot including the P-value for the difference of their means using the Student's t-test. Shaded areas in gray and pink colors indicate the $\pm 1\,\sigma$ (one standard deviation) range for OBS and ERA5-L.

is the variability error. Values of RMSE at three stations are 7092.92, 8672.75, and 1343.82. It shows the consistent result that the reconstructed streamflow at Bahadurabad and Harding Bridge have larger errors than Bhairab Bazar in terms of mean bias. The Student's t-test showed the p-values are less than 0.05 for the difference between observed and reconstructed streamflow data at three stations, indicating a statistically significant difference from the observations at a 95% confidence level. The values of mKGE, r, β, γ, and RMSE for the climatological daily reconstructed streamflow are summarized in Table 2. It may mean that the model's performance is not good, but it also implies that the observed streamflow is largely affected by human activities such as irrigation and reservoir storage. The results imply that the impact of human disturbance might be significant at the local scale, which is in line with the findings of a previous study[28].

The bias and variability errors can be affected by human activities. Bangladesh is located at the downstream of three rivers in terms of topography. Most of gauge stations in Bangladesh do not observe natural streamflow due to artificial structures in the upstream areas. As a result, the flow observed at gauge stations in Bangladesh

Station name	mKGE	r	β	γ	RMSE
Bahadurabad	0.75	0.99	1.25	1.03	7092.92
Hardinge Bridge	0.50	0.88	1.46	0.83	8672.75
Bhairab Bazar	0.25	0.95	0.91	1.75	1343.82

Table 2. Values of mKGE, r, β, γ, and RMSE between observed and reconstructed streamflow at three gauge stations during 1994–2023.

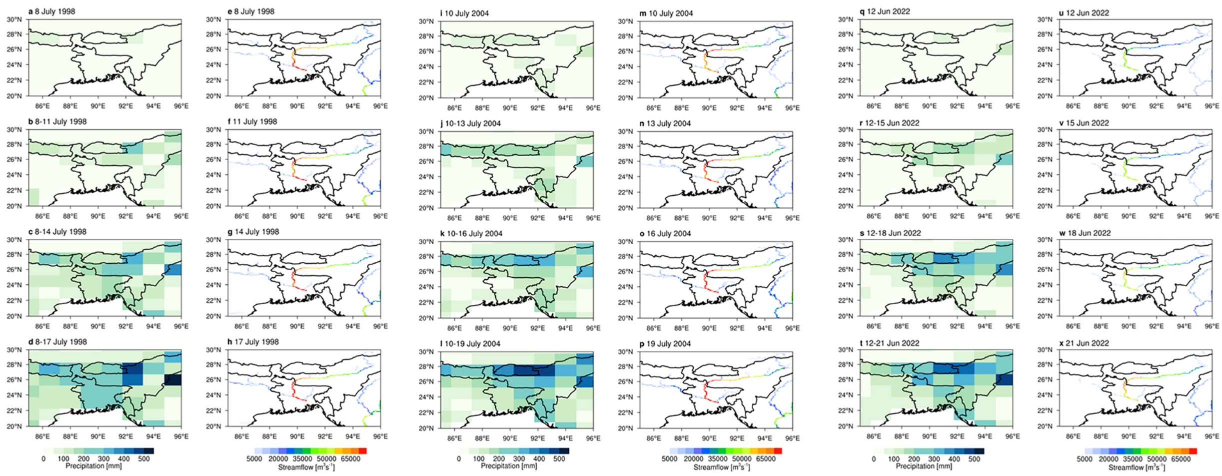

Fig. 4 Accumulated precipitation and simulated streamflow using VIC-RRM for the 1998, 2004, and 2022 flood cases in Bangladesh. (**a**–**d**), (**i**–**l**), (**q**–**t**) Accumulated precipitation (mm) pattern over 3, 6, and 9 days, starting from 8th July 1998; 10th July 2004, and 12th Jun 2022. (**e**–**h**), (**m**–**p**), (**u,x**) Distribution of simulated streamflow (m^3s^{-1}) at 3-day intervals from starting dates in 1998, 2004, and 2022 using VIC-RRM forced by runoff datasets of ERA5-L. Daily precipitation data obtained from the ERA5-L product.

cannot be used for water resource management and climate impact studies. Thus, there is bound to be a difference from the observed streamflow that includes human impacts, and the high bias and variability errors could imply that most areas in Bangladesh have their streamflow controlled by human activities. However, VIC-RRM can well captures the climatological seasonality of observed streamflow, which is controlled by that of precipitation. While human disturbance such as irrigation and reservoir can change the seasonality of streamflow, this study focuses on naturalized streamflow to better understand associations of streamflow with climate variability and change. In addition, the bias and variability errors can come from the errors of input data. In the Saint-Venant equations used in VIC-RRM, the streamflow depends on the amount of water[22]. Thus, this model can imply that the nonlinear model produces very inaccurate runoff due to an inappropriate amount of precipitation. Therefore, problems encountered in land or atmospheric models, such as vegetation parameters, soil parameters, and precipitation predictions can generate bias and variability errors between observations and simulations[21].

To assess the response of streamflow by precipitation in VIC-RRM, we select three major flood events in Bangladesh: July 1988, July 2004, and June 2022. Figure 4 shows the accumulated precipitation patterns on three, six, and nine days after the beginning date of each flood event (8 July 1998, 10 July 2004, and 12 Jun 2022), that is, it presents daily evolutions of the selected flood events at 3-day intervals. A common feature of the three flood events is a continuous increase in precipitation in the eastern part of Bhutan and increase in streamflow over the northern part of Bangladesh along the Brahmaputra River network. This is due to the geographical structure of Bhutan, which is at a higher elevation than Bangladesh. This result is consistent with the findings of previous studies that flood events occurred with the rapidly increased precipitation in the north-eastern part of Bangladesh[33,34,35]. Such extreme rainfall events in the upstream areas can cause floods in Bangladesh due to topographical effects, which can be examined from the long-term data of reconstructed daily streamflow from the VIC-RRM.

Usage Note

For VIC-RMM, river networks of Ganges, Brahmaputra, and Meghna Rivers were constructed using the 9-km DEM data to mimic a realistic river network for simulating daily streamflow with grid cells within three river basins. VIC-RRM is a simple linear transfer function model that assumes the water transport process is linear and time-invariant[21,22]. It incorporates natural topography derived from DEM data. The spatial resolution of the simulated streamflow is determined by the resolution of the DEM, which allows VIC-RRM to be applied to other river basins[24,36,] as high-resolution DEM data is now globally available.

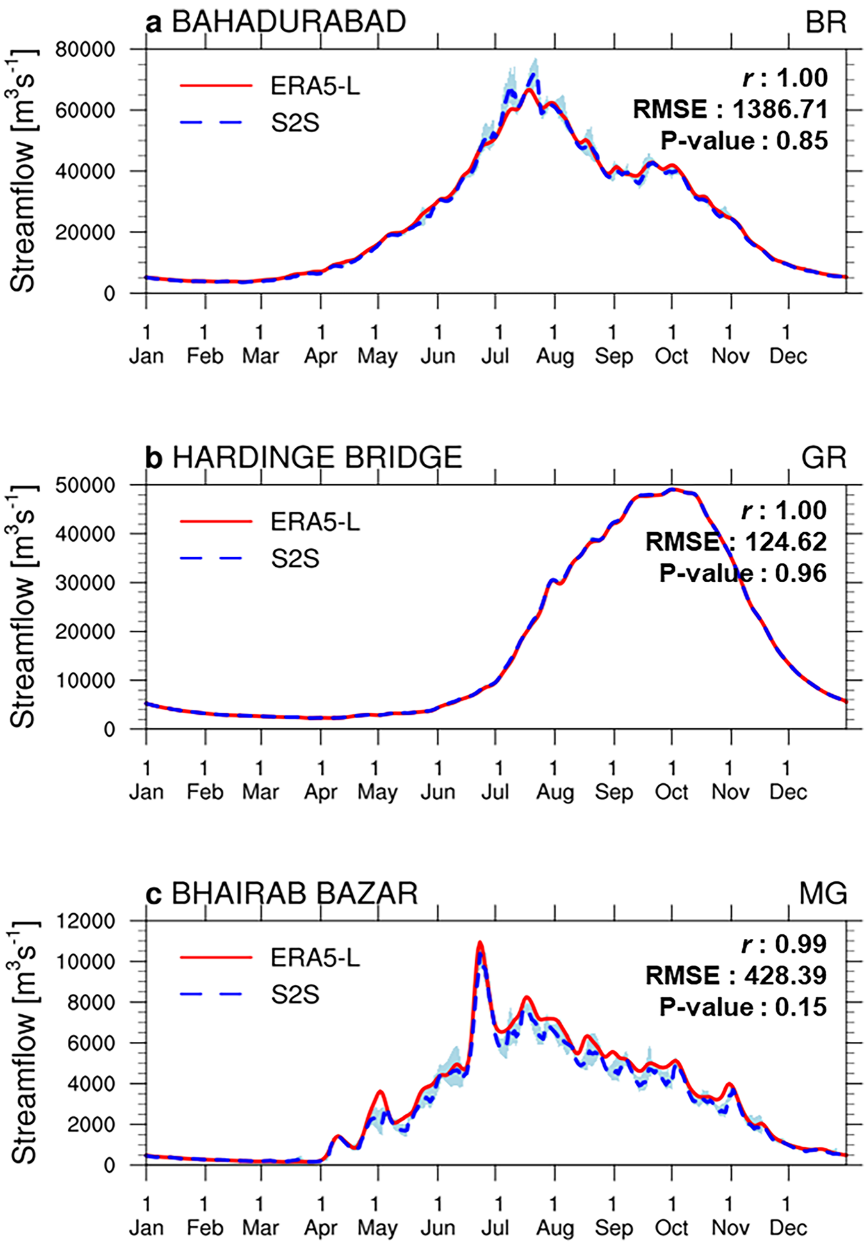

Fig. 5 Long-term averaged reconstructed (ERA5-L, red solid line) and forecasted (S2S, blue dashed line) streamflow over 2016–2023 at Bahadurabad, Hardinge Bridge, and Bhairab Bazar. The correlation coefficient (r) and RMSE between climatological observed and reconstructed streamflow are denoted on the top-right corner in each panel plot including the P-value for the difference of their means using the Student's t-test. Skyblue shading denotes the ensemble range for 50 ensemble members in the S2S forecast.

The simulated streamflow was evaluated against observed streamflow data at three gauge stations for three rivers in Bangladesh. The simulated streamflow by the VIC-RRM exhibits high performance in terms of temporal analysis related to the natural climatological seasonality while having large bias and variability errors at three gauge stations. These discrepancies between observed and simulated streamflow data can be attributed to the simplified hydrodynamic processes in VIC-RRM, the quality of observed streamflow records, or the impact of irrigation and reservoir storage, which need to be further investigated. The impact of local human disturbance was embedded in observed daily streamflow records over the Ganges-Brahmaputra-Meghna River basin where the streamflow records are available since the late 1990s after the construction of water-engineered facilities. In addition, most errors in simulated streamflow are likely caused by inaccurate runoff and precipitation data generated by nonlinear models such as land surface model and climate model[21]. Thus, bias and variability errors may arise from uncertainty in the ERA5-Land runoff data, and our streamflow data can provide an opportunity to reduce errors in streamflow through realistic parameterization of land surface models for the Ganges-Brahmaputra-Meghna River basin.

Station name	mKGE	r	β	γ	RMSE
Bahadurabad	0.97	1.00	0.99	1.03	1386.71
Hardinge Bridge	1.00	1.00	1.00	1.00	124.62
Bhairab Bazar	0.91	0.99	0.91	1.01	428.39

Table 3. Values of mKGE, r, β, γ, and RMSE between reconstructed streamflow and forecasted streamflow at three gauge stations during 2016–2023.

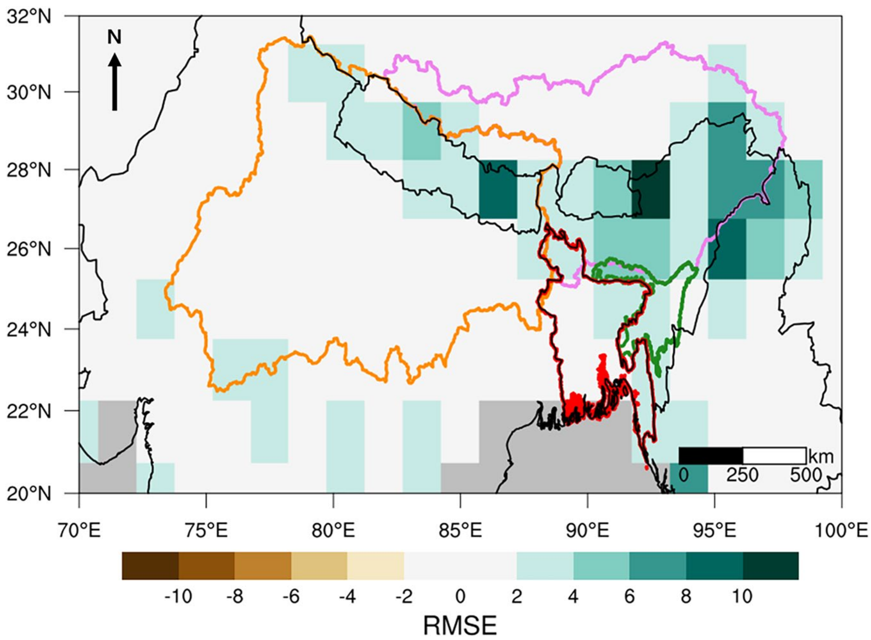

Fig. 6 RMSE for long-term averaged runoff between ERA5-L and S2S forecast during 2016–2023. Color lines (orange, pink, green, and red) denote the GR, BR, MR basins, and the Bangladesh national boundary.

On the other hand, VIC-RRM used in this study does not take into account the influence of man-made structures such as dams and reservoirs, so it has limitations that cannot perfectly match observational data. To simulate realistic streamflow while considering human influence, hydrological models using the modules related to irrigation[37] and reservoir[38] can be used. New modules of the VIC model have been developed, which can account for water withdrawal and consumption from different sectors, environmental flow requirements for water systems, and dam operations[39]. A recent study proposed the VIC model combined with a deep learning model[40] to predict reservoir inflow of a reservoir in China. The hybrid model showed significant improvement of model performance at six lead months.

Thus, the development of parameter data related to the local human disturbance is required for the anthropogenic streamflow simulation of the Ganges-Brahmaputra-Meghna Rivers, which will provide a unique opportunity to study the impact of local human disturbance on the regional hydrologic system.

Predicted water resource data helps water managers operate water infrastructure more efficiently and prepare for the impacts of floods and droughts. Using the forecasted runoff data as input for VIC-RRM, streamflow forecast data can be obtained. In this study, we used the surface and sub-surface runoff forecast data from the S2S forecast product as input from 2016 to 2023 to get the streamflow forecast data. Since the forecast data used in this study was generated at two-week intervals, VIC-RRM required the initial conditions of streamflow every two weeks. Thus, the regenerated streamflow data using the ERA5-L product was used for the initial state at two-week intervals.

Figure 5 shows the long-term averaged daily streamflow using ERA5-L product and S2S forecast product by VIC-RRM during 2016–2023 at the Bahadurabad, Hardinge Bridge, and Bhairab Bazar gauge stations. Bahadurabad and Hardinge Bridge have fairly similar long-term annual cycles of streamflow between ERA5-L and S2S, while the streamflow forecast is underestimated at the Bhairab Bazar gauge station. The performance of the forecasted streamflow data using the S2S forecast data was evaluated using the mKGE analysis based on the streamflow from ERA5-L. The mKGE values at the Bahadurabad, Hardinge Bridge, and Bhairab Bazar gauge stations are 0.97, 1.00, and 0.91. The three components (r, β, and γ) of mKGE are shown in Table 3. In terms of the naturalized streamflow, forecasted streamflows at three gauge stations are quite reliable and the ECMWF S2S forecast data is valuable for use in generating streamflow forecast data together with VIC-RRM in water resource management. In terms of the magnitude of errors, values of RMSE at three gauge stations are 1386.71, 124.62, 428.39. However, Student's t-test showed that the differences in mean values of simulated streamflow forced by the ERA5-L and S2S runoff data at three stations are insignificant with the p-values exceeding 0.05.

The error at Bahadurabad is relatively larger than other stations. It seems to come from the forecasted runoff errors of S2S in the BR basin region based on that of ERA5-L. Figure 6 shows the RMSE distribution for the total runoff, which is the sum of surface and sub-surface runoff, between ERA5-L and S2S over 2016–2023. Most errors are detected within the Brahmaputra basin region. The area averaged RMSE values for the BR, GR, MR basins are 3.38, 1.88, 2.75. As with the RMSEs for streamflow at gauge stations, the RMSE value for the basin averaged runoff at the Brahmaputra basin region is the largest. These results imply that inaccurate streamflow is generated due to the inappropriate prediction of runoff because the streamflow depends on the amount of water in VIC-RRM based on the Saint-Venant equations. Thus, the problems occurring in land or atmospheric prediction models such as vegetation parameters, soil parameters, and precipitation prediction are important[21]. In addition, in terms of naturalized streamflow, we confirmed that streamflow forecast data can be obtained by using reliable forecasted runoff data as input for VIC-RRM. We confirmed that the topographic data (i.e. DEM) and forecasted runoff datasets of the numerical forecast model can be used as input data in VIC-RRM to generate streamflow forecast data.

Code availability

Data processing and plotting were performed using the NCAR Command Language (NCL). VIC-RRM source codes used in this study are available from https://doi.org/10.7910/DVN/V2C6G2[32].

Received: 11 October 2024; Accepted: 15 April 2025;
Published online: 22 April 2025

References

1. Ficklin, D. L., Robeson, S. M. & Knouft, J. H. Impacts of recent climate change on trends in baseflow and stormflow in United States watersheds. *Geophys Res Lett* **43**, 5079–5088, https://doi.org/10.1002/2016gl069121 (2016).
2. Rumsey, C. A., Miller, M. P. & Sexstone, G. A. Relating hydroclimatic change to streamflow, baseflow, and hydrologic partitioning in the Upper Rio Grande Basin, 1980 to 2015. *J Hydrol* **584**, https://doi.org/10.1016/j.jhydrol.2020.124715 (2020).
3. Sadeghi, S. *et al.* Atlantic Ocean Sea Surface Temperatures and Southeast United States streamflow variability: Associations with the recent multi-decadal decline. *J Hydrol* **576**, 422–429, https://doi.org/10.1016/j.jhydrol.2019.06.051 (2019).
4. Eckstein, D., Künzel, V. & Schäfer, L. *The global climate risk index 2021*. (Bonn: Germanwatch, 2021).
5. Rahman, M. C. *et al.* Climate change and risk scenario in Bangladesh. *Asia-Pacific Journal of Regional Science* **7**, 381–404 (2023).
6. Bhuyan, M. A., Raihan, S. U., Barua, P. P., Hasan, M. & Sarker, P. K. Annual Flood Report 2021. 1–113 (Bangladesh Water Development Board, Pani Bhaban, 2021).
7. GRDC. Major River Basins of the World/Global Runoff Data Centre. *GRDC. 2nd, rev. ext. ed. Koblenz, Germany: Federal Institute of Hydrology (BfG)* (2020).
8. Sivapalan, M. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrol Process* **17**, 3163–3170, https://doi.org/10.1002/hyp.5155 (2003).
9. Wood, E. F. *et al.* Reply to comment by Keith J. Beven and Hannah L. Cloke on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water". *Water Resour Res* **48**, https://doi.org/10.1029/2011wr011202 (2012).
10. Allan, J., Bain, M., Pestegaard, K., Richter, B. & Sparks, R. The natural flow regime: A paradigm for river conservation (1997).
11. Wurbs, R. A. Methods for developing naturalized monthly flows at gaged and ungaged sites. *J Hydrol Eng* **11**, 55–64, https://doi.org/10.1061/(Asce)1084-0699(2006)11:1(55) (2006).
12. Kim, T. J. & Wurbs, R. A. Development of Monthly Naturalized Flow Using Water Rights Analysis Package (WRAP)-based Methods. *Ksce J Civ Eng* **15**, 1299–1307, https://doi.org/10.1007/s12205-011-1184-y (2011).
13. Dorchies, D. *et al.* Climate change impacts on multi-objective reservoir management: case study on the Seine River basin, France. *International Journal of River Basin Management* **12**, 265–283 (2014).
14. Laizé, C. L. R. *et al.* Projected Flow Alteration and Ecological Risk for Pan-European Rivers. *River Res Appl* **30**, 299–314, https://doi.org/10.1002/rra.2645 (2014).
15. Terrier, M. *et al.* Streamflow naturalization methods: a review. *Hydrolog Sci J* **66**, 12–36, https://doi.org/10.1080/02626667.2020.1839080 (2021).
16. Lin, P. R. *et al.* Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water Resour Res* **55**, 6499–6516, https://doi.org/10.1029/2019wr025287 (2019).
17. Kanishka, G. & Eldho, T. I. Streamflow estimation in ungauged basins using watershed classification and regionalization techniques. *J Earth Syst Sci* **129** https://doi.org/10.1007/s12040-020-01451-8 (2020).
18. Muñoz-Sabater, J. *et al.* 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 (2021).
19. Vitart, F. *et al.* The Subseasonal to Seasonal (S2s) Prediction Project Database. *B Am Meteorol Soc* **98**, 163-+, https://doi.org/10.1175/Bams-D-16-0017.1 (2017).
20. Jarvis, A., Reuter, H., Nelson, A. & Guevara, E. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture CIAT., available from http. *srtm. csi. cgiar. org* (2008).
21. Lohmann, D., Raschke, E., Nijssen, B. & Lettenmaier, D. P. Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. *Hydrolog Sci J* **43**, 131–141, https://doi.org/10.1080/02626669809492107 (1998).
22. Lohmann, D., NolteHolube, R. & Raschke, E. A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus A* **48**, 708–721, https://doi.org/10.1034/j.1600-0870.1996.t01-3-00009.x (1996).
23. Wu, H. *et al.* Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. *Water Resour Res* **50**, 2693–2717, https://doi.org/10.1002/2013wr014710 (2014).
24. Kim, B.-H., Kim, Y.-O. & Kam, J. Hyper-resolution naturalized streamflow data for Guem River in South Korea (1951-2020). *Sci. Data* **12**, 210, https://doi.org/10.1038/s41597-025-04486-y (2025)
25. Yamazaki, D., Kanae, S., Kim, H. & Oki, T. A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resour Res* **47**, https://doi.org/10.1029/2010wr009726 (2011).
26. Yamazaki, D., de Almeida, G. A. M. & Bates, P. D. Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map. *Water Resour Res* **49**, 7221–7235, https://doi.org/10.1002/wrcr.20552 (2013).
27. Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y. & Bates, P. D. Regional flood dynamics in a bifurcating mega delta simulated in a global river model. *Geophys Res Lett* **41**, 3127–3135, https://doi.org/10.1002/2014gl059744 (2014).

28. Kam, J. Diversity in the observed functionality of dams and reservoirs. *Environmental Research: Infrastructure and Sustainability* **1**, 031003 (2021).
29. Jones, P. W. First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon Weather Rev* **127**, 2204–2210, 10.1175/1520-0493(1999)127<2204:Fasocr>2.0.Co;2 (1999).
30. Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 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 (2009).
31. Kling, H., Fuchs, M. & Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J Hydrol* **424**, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011 (2012).
32. Kim, B.-H. & Kam, J. High-resolution gridded streamflow for Ganges-Brahmaputra-Meghna River Basins. *Harvard Dataverse*. https://doi.org/10.7910/DVN/V2C6G2 (2024).
33. Chowdhury, M. R. An assessment of flood forecasting in Bangladesh: The experience of the 1998 flood. *Nat Hazards* **22**, 139–163, https://doi.org/10.1023/A:1008151023157 (2000).
34. Salehin, M., Haque, A., Rahman, M., Khan, M. S. A. & Bala, S. K. Hydrological aspects of 2004 floods in Bangladesh. *Journal of Hydrology and Meteorology* **4**, 33–44 (2007).
35. Tuli, R. D., Rashid, K. J. & Akter, T. Impact analysis of the 2022 flood event in Sylhet and Sunamganj using Google Earth Engine. *Modern Cartography Series* **12**, 47-69, https://doi.org/10.1016/B978-0-443-23890-1.00003-7 (2024).
36. Ali, S., Kim, B.-H., Akhtar, T. & Kam, J. Past and future changes toward earlier timing of streamflow over Pakistan from bias-corrected regional climate projections (1962–2099), *J. Hydrol.* **617**, 128959, https://doi.org/10.1016/j.jhydrol.2022.128959 (2023)
37. Haddeland, I., Lettenmaier, D. P. & Skaugen, T. Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins. *J Hydrol* **324**, 210–223, https://doi.org/10.1016/j.jhydrol.2005.09.028 (2006).
38. Gao, H., Bohn, T. J., Podest, E., McDonald, K. C. & Lettenmaier, D. P. On the causes of the shrinking of Lake Chad. *Environ Res Lett* **6**, https://doi.org/10.1088/1748-9326/6/3/034021 (2011).
39. Droppers, B., Franssen, W. H. P., van Vliet, M. T. H., Nijssen, B. & Ludwig, F. Simulating human impacts on global water resources using VIC-5. *Geosci Model Dev* **13**, 5029–5052, https://doi.org/10.5194/gmd-13-5029-2020 (2020).
40. Xu, W. X. *et al*. Coupling Deep Learning and Physically Based Hydrological Models for Monthly Streamflow Predictions. *Water Resour Res* **60**, https://doi.org/10.1029/2023WR035618 (2024).

Acknowledgements

We thank ERA5-Land reanalysis data and ECMWF S2S forecast data for openly providing us. Reconstructed high-resolution (9-km) daily naturalized streamflow datasets for the Ganges-Brahmaputra-Meghna Rivers areas (70–100°E, 20–35°N) are freely available from Harvard Dataverse (https://doi.org/10.7910/DVN/V2C6G2)[32] including forecasted daily streamflow data using the S2S forecast product. We also acknowledge the availability of observed daily streamflow data at Yangcun, Bahadurabad, Hardinge Bridge, and Bhairab Bazar within the Ganges, Brahmaputra, and Meghna rivers from the Global Runoff Data Centre (GRDC) and the Bangladesh Water Development Board (BWDB) of Bangladesh. This study was supported by the National Research Foundation (NRF) of Korea (2021M3I6A1086808/RS-2021-NR057873) and RS-2023-00262044) and the Human Resource Program for Sustainable Environment in the 4th Industrial Revolution Society of the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) and NRF of Korea.

Author contributions

J.K. conceived and supervised the project. S.H. provided observed daily streamflow data at Bahadurabad, Hardinge Bridge, and Bhairab Bazar gauge stations in Bangladesh. B.-H.K. produced the streamflow data, devised and performed the analyses, and drafted the manuscript. J.K, S.H., and B.-H.K. provided input on the interpretation of the results and helped shape the analysis and the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-05014-8.

Correspondence and requests for materials should be addressed to J.K.