

Water Resources Research®

RESEARCH ARTICLE

10.1029/2021WR029920

Key Points:

- We developed a Bayesian Hierarchical Network Model for ensemble forecasts of daily streamflow with the attendant uncertainties
- The model provides ensemble forecasts at all the locations on a river network simultaneously, capturing the spatial and temporal correlation
- The framework can be applied to any river network and with appropriate covariates

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Ossandón,
alvaro.ossandon@colorado.edu

Citation:

Ossandón, Á., Rajagopalan, B., Lall, U., Nanditha, J. S., & Mishra, V. (2021). A Bayesian hierarchical network model for daily streamflow ensemble forecasting. *Water Resources Research*, 57, e2021WR029920. <https://doi.org/10.1029/2021WR029920>

Received 5 MAR 2021

Accepted 8 AUG 2021

© 2021. American Geophysical Union.
All Rights Reserved.

A Bayesian Hierarchical Network Model for Daily Streamflow Ensemble Forecasting

Álvaro Ossandón^{1,2} , Balaji Rajagopalan^{1,3} , Upmanu Lall⁴ , J. S. Nanditha⁵ , and Vimal Mishra⁵ 

¹Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO, USA,

²Departamento de Obras Civiles, Universidad Técnica Federico Santa María, Valparaíso, Chile, ³Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, ⁴Department of Earth and Environmental Engineering, Columbia Water Center, The Earth Institute, Columbia University, New York, NY, USA,

⁵Civil Engineering, Indian Institute of Technology, Gandhinagar, India

Abstract A novel Bayesian Hierarchical Network Model (BHNM) for ensemble forecasts of daily streamflow is presented that uses the spatial dependence induced by the river network topology and hydrometeorological variables from the upstream contributing area between station gauges. Model parameters are allowed to vary with time as functions of selected covariates for each day. Using the network structure to incorporate flow information from upstream gauges and precipitation from the immediate contributing area as covariates allows one to model the spatial correlation of flows simultaneously and parsimoniously. An application to daily monsoon period (July–August) streamflow at three gauges in the Narmada basin in central India for the period 1978–2014 is presented. The best set of covariates include daily streamflow from upstream gauges or from the gauge above the upstream gauges depending on travel times and daily precipitation from the area between two stations. The model validation indicates that the model is highly skillful relative to a null-model of generalized linear regression, which represents the analogous non-Bayesian forecast. The ensemble spread of BHNM accounts for the forecast uncertainty leading to reliable and skillful streamflow predictions.

1. Introduction

Riverine floods are a major cause of destruction of property and loss of life each year across the world (Tanoue et al., 2016; Wallemacq & House, 2018). This is the case in India, where floods occur mostly during the summer monsoon season of June–September, when the country receives more than 80% of annual rainfall. Extensive damages to life and property occur annually during the monsoon season floods in India. The deaths caused by flood events substantially increased in the 21st century (EM-DAT) with an average death toll of 1,500 per year (The Data Centre of Central Water Commission), and associated damages worth 18 billion INR (CAG, 2017). The extreme rainfall events in the summer monsoon season result from synoptic-scale cyclonic depressions (Hunt & Fletcher, 2019; Hunt et al., 2016). Climate change is projected to enhance the frequency and intensity of extreme precipitation events (Ali & Mishra, 2018; Goswami et al., 2006; Papalexioi & Montanari, 2019; Wasko & Sharma, 2017) and damages caused by floods will further increase. This highlights the importance of accurate flood forecasting. India has achieved significant progress in predicting extreme precipitation events using Numerical Weather Prediction models and Ensemble Prediction systems (Pattanaik et al., 2019; Sridevi et al., 2020). While precipitation forecasts are increasingly becoming skillful, forecasts of streamflow and specifically of floods remain less so with wide variation in skill across River Basins and events.

For daily streamflow forecasting, physically based and statistical models are two broad categories of widely used approaches (e.g., Yuan et al., 2015; Zhang et al., 2018). Physically based models consider different hydrological processes and their interactions and model them with deterministic equations. Statistical models focus on an identification of the empirical relationship between the current day's streamflow with input forcings such as precipitation, antecedent streamflow, and soil moisture, and so on, using historical observations, that index key hydrologic processes. Here we consider a statistical model. A brief survey of statistical models used in this context is provided below.

Typical statistical models used for rainfall-runoff are largely regression based using linear, non-linear, and machine learning techniques. Multiple linear regression (MLR; Gaume & Gosset, 2003; Kişi, 2008; Papacharalampous & Tyralis, 2018), autoregressive (AR; Kişi, 2004, 2008; Sivakumar, 2016), and autoregressive moving average models (ARMA; Can et al., 2012; T. J. Chang et al., 1987; Sivakumar, 2016) are reported. The streamflow on a day is modeled as a function of streamflow and precipitation from preceding days. Precipitation from the current day is included to incorporate daily precipitation forecasts when available. Such models have been applied for daily streamflow forecasting in Europe (Can et al., 2012; Gaume & Gosset, 2003; Kişi, 2004, 2008), United States (T. J. Chang et al., 1987; Papacharalampous & Tyralis, 2018), and China (Y. Sun et al., 2019). To address non-linearity, machine learning techniques such as artificial neural network (ANN; Abdollahi et al., 2017; Govindaraju, 2000; Isik et al., 2013), adaptive neuro-fuzzy inference system (ANFIS; F. J. Chang & Chen, 2001; Jang et al., 1997; Li et al., 2018; Zounemat-Kermani & Teshnehlab, 2008) and, support vector machines (SVM; Ghorbani et al., 2016; Karimi et al., 2018; Londhe & Gavraskar, 2015), are gaining prominence. These models have been applied in Europe (Firat, 2008; Gaume & Gosset, 2003; Hadi & Tombul, 2018), Asia (F. J. Chang & Chen, 2001; Pramanik & Panda, 2009; Shiau & Hsu, 2016), United States and Canada (Isik et al., 2013; Moradkhani et al., 2004; Vafakhah, 2012). Studies have also found machine learning models to be more skillful than linear models (Firat, 2008; Hadi & Tombul, 2018; Vafakhah, 2012). However, they have been found to be uninterpretable ("black box"), prone to overfitting, and usually do not quantify uncertainty in the parameters and model estimates.

Traditional statistical models as mentioned above assume stationarity of the daily streamflow process and are typically implemented at single sites individually. However, to capture spatial correlation such as daily streamflows on a river network, multivariate versions are needed, which are not easy to develop in the traditional approaches. Lastly, the uncertainties in parameters and model estimates are not formally modeled, consequently, underestimating of extremes is common. In order to model and mitigate flooding on a river network, forecasts are required at all the sites simultaneously capturing their space-time correlation structure along with all the attendant uncertainties.

Bayesian hierarchical modeling frameworks are emerging as attractive alternatives for nonstationary analysis of hydroclimate fields in space and time. We call out few relevant applications here. Spatial models for precipitation extremes to obtain stationary return levels have been developed and applied to western United States and other parts of the world (Bracken et al., 2016; Dyrddal et al., 2015; Renard, 2011; Yan & Moradkhani, 2015). Extensions to nonstationary extremes of spatial rainfall and streamflow extremes have been developed in other studies (Bracken et al., 2018; Hanel et al., 2009; Ossandón et al., 2021; X. Sun et al., 2014). The Bayesian hierarchical framework has been used in stochastic generation of precipitation and temperature (e.g., Lima & Lall, 2009; Verdin et al., 2019). Applications to streamflow extremes to obtain return levels for flood mitigation have been the early motivation of these methods (e.g., Kwon et al., 2008; Lima & Lall, 2010; Lima et al., 2016; Luke et al., 2017; Sampaio & Costa, 2021). However, multi-site streamflow modeling on a river network is limited and more so for modeling daily streamflow. Ravindranath et al. (2019) offer one of the early applications of Bayesian modeling framework for multi-site flow on a river basin exploiting the network structure to capture spatial dependence. In this, streamflow processes are modeled as spatial Markov process, thus, the dependence structure is captured by the covariates considered on the network. They applied this for paleo-reconstruction of annual flows in the Upper Missouri River Basin.

Our research in this paper is motivated by the need for a daily streamflow modeling framework on a river network for potential use in real-time forecasting, that can capture the space-time dependence structure, and robust estimation of uncertainties. To this end, we develop a novel Bayesian Hierarchical Network Model (BHNM) inspired by the framework proposed in Ravindranath et al. (2019). We demonstrate this framework by its application to model and predict daily summer monsoon (July–August) streamflow at three gauges in the Narmada River Basin network in central India. The manuscript is organized as follows. In Section 2, the framework, in general, is described. The application set up for the Narmada basin network is then described, followed by the specific form of the model structure and model cross-validation procedure in Section 3. The results are described in Section 4, and Section 5 presents a summary and discussion of the results.

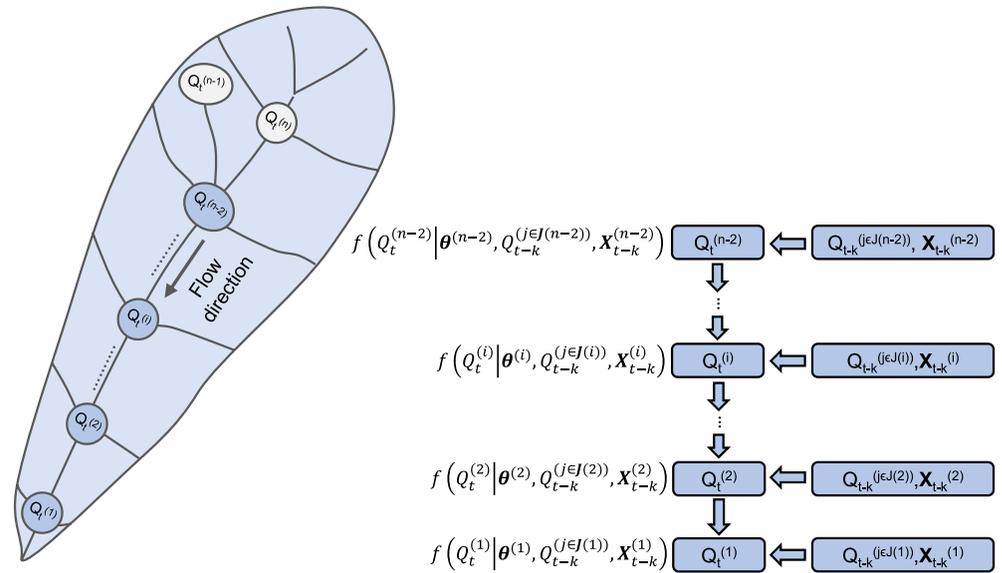


Figure 1. Conceptual sketch of Bayesian Hierarchical Network Model. n streamflow gauges and $n-2$ hydrometeorological covariates vectors are shown in the graph for illustrating the concept of the graphical network model. Physically informed modeling structure using regional hydrometeorological covariates and feeder streamflow gauges is explored using factorization into lower-dimensional conditional probability distributions as shown in the directed graph. The conditional distributions generated at each stage of the chart serve as statistical interpretations of the modeling structure and provide the basis for converting the graphical model into a set of equations for estimating the parameters of the streamflow network's likelihood function for $n-2$ gauges (nodes, blue circles) in the network simultaneously using a Bayesian estimation scheme. $\theta_t^{(i)} = [\beta_t^{(i)}, \phi_t^{(i)}]$ denotes the vector of distribution parameters at the gauge i and day t .

2. Proposed Framework

The proposed BHNM for daily streamflow has two components: the general model structure and calculation of the likelihood function and specification of priors.

2.1. General Model Structure

In order to model the daily streamflow at $n-2$ locations simultaneously, the model structure takes advantage of the feature of the river network by treating the streamflow processes as a spatial Markov process (Ravindranath et al., 2019). Flow at a downstream gauge, i , on day t is dependent on: flow at the most immediate ($i+1$) or second most immediate ($i+2$) upstream feeder gauge at day $t-k$ with $k > 0$ (k represents the lead time of the forecast); precipitation and other hydrometeorological variables that represent local inputs to the streamflow between the streamflow gauges. Note that the approach does not model the headwater gauge since there are no feeder gauges available as predictors for it. The ability to capture the timing of high flow events in advance resides in feeder gauges. Thus, in a Bayesian framework, assuming conditional distributions at each location as independent, the joint conditional probability density of streamflow at the gages on the network on day t , conditioned on the suite of covariates (flow and hydrometeorological variables from upstream) as:

$$\begin{aligned}
 & f(Q_t^{(1)}, \dots, Q_t^{(n-2)} | \theta^{(1)}, \dots, \theta^{(n-2)}, Q_{t-k}^{(2)}, \dots, Q_{t-k}^{(n)}, \mathbf{X}_{t-k}^{(1)}, \dots, \mathbf{X}_{t-k}^{(n-2)}) \\
 &= f(Q_t^{(1)} | \theta^{(1)}, Q_{t-k}^{(j \in J(1))}, \mathbf{X}_{t-k}^{(1)}) \cdot f(Q_t^{(2)} | \theta^{(2)}, Q_{t-k}^{(j \in J(2))}, \mathbf{X}_{t-k}^{(2)}) \cdot \dots \\
 & f(Q_t^{(i)} | \theta^{(i)}, Q_{t-k}^{(j \in J(i))}, \mathbf{X}_{t-k}^{(i)}) \cdot \dots \cdot f(Q_t^{(n-2)} | \theta^{(n-2)}, Q_{t-k}^{(j \in J(n-2))}, \mathbf{X}_{t-k}^{(n-2)})
 \end{aligned} \tag{1}$$

Where \mathbf{X} denotes the set of hydrometeorological covariates, $\theta^{(i)}$ corresponds to the vector of regression coefficients of the distribution parameters at the gauge i , and $\mathbf{J}(i)$ to an index vector of length n with only zeros except for the position of the feeders which contain ones (e.g., for $\mathbf{J}(n-2)$, only its n and $n-1$ elements

are non-zero values, see Figure 1). The right-hand side of Equation 1 is the mathematical factorization of the joint conditional density as a product of individual conditional densities using the fundamental Bayes rule (Jensen & Nielsen, 2007). This factorization is consistent with the assumption of independent marginal distributions and the physical dependencies between streamflow gauges and their feeder gauges and hydro-meteorological variables. We hypothesize that the spatial dependence of the flows will be captured by the covariates. A conceptual sketch of the BHNM for daily streamflow is shown in Figure 1.

The daily streamflow at each gauge is conditionally assumed to follow a probability density function f with parameters that can vary with time through a multi-level specification in terms of other predictors. Thus, daily streamflow at each gauge i at the day t is expressed as:

$$f\left(Q_t^{(i)} \mid \alpha_t^{(i)}, \lambda_t^{(i)}, Q_{t-k}^{(j \in \mathbf{J}(i))}, \mathbf{X}_{t-k}^{(i)}\right), i = 1, 2, \dots, n - 2 \quad (2)$$

where $\alpha_t^{(i)}$ and $\lambda_t^{(i)}$ are the distribution parameters at the gauge i and day t . These parameters can be expressed in terms of the expected value, $\mu_t^{(i)}$, and variance, $(\sigma_t^{(i)})^2$, of $Q_t^{(i)}$ as follows:

$$\lambda_t^{(i)} = g_1\left(\mu_t^{(i)}, (\sigma_t^{(i)})^2\right), \quad \lambda_t^{(i)} = g_2\left(\mu_t^{(i)}, (\sigma_t^{(i)})^2\right) \quad (3)$$

Under the nonstationary assumption, $\mu_t^{(i)}$, and $\sigma_t^{(i)}$ are modeled as linear functions of the flow at the upstream feeder gauges depending on travel time, and m hydrometeorological variables at day $t-k$. Since precipitation, $P_{t-k}^{(i)}$, can be considered as a hydrometeorological covariate, for avoiding fitting issues due to there are days with zero rainfall (e.g., the log transformation of the covariates is required for a lognormal distribution), we considered a step function for the mean and standard deviation which yields the following expressions

$$\mu_t^{(i)} = \begin{cases} \beta_1^{(i)} + \beta_2^{(i)} Q_{t-k}^{(j \in \mathbf{J}(i))} & P_{t-k}^{(i)} = 0 \\ \beta_3^{(i)} + \beta_4^{(i)} Q_{t-k}^{(j \in \mathbf{J}(i))} + \beta_x^{(i)} \mathbf{X}_{t-k}^{(i)} & P_{t-k}^{(i)} > 0 \end{cases} \quad (4)$$

$$\sigma_t^{(i)} = \begin{cases} \phi_1^{(i)} + \phi_2^{(i)} Q_{t-k}^{(j \in \mathbf{J}(i))} & P_{t-k}^{(i)} = 0 \\ \phi_3^{(i)} + \phi_4^{(i)} Q_{t-k}^{(j \in \mathbf{J}(i))} + \phi_x^{(i)} \mathbf{X}_{t-k}^{(i)} & P_{t-k}^{(i)} > 0 \end{cases} \quad (5)$$

where $\beta_1^{(i)}$, $\beta_2^{(i)}$ and $\phi_1^{(i)}$, $\phi_2^{(i)}$ are the intercept and slope regression coefficients for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$ when there is not rainfall on the day $t-k$. For days when rainfall on the day $t-k$ is not zero, $\beta_3^{(i)}$ and $\phi_3^{(i)}$ are the intercept terms for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\beta_4^{(i)}$ and $\phi_4^{(i)}$ are regression coefficients related to the feeder site (vector with more than one component when there are more than one feeder) for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\beta_x^{(i)}$ and $\phi_x^{(i)}$ are $1 \times m$ vector of regression coefficients related to hydrometeorological variables for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\mathbf{X}_{t-k}^{(i)}$ is a $m \times 1$ vector of hydrometeorological variables on the day $t-k$; and $Q_{t-k}^{(j \in \mathbf{J}(i))}$ corresponds to the flow at the feeder site at the day $t-k$ (vector with more than one component if the index vector $\mathbf{J}(i)$ has more than one non-zero component). All of the model covariates change with time to help capture nonstationarity. For days with zero rainfall, we only consider streamflow at the feeder site as covariate. Also, modeling the standard deviation, $\sigma_t^{(i)}$, as nonstationary allow addressing the heteroscedasticity—that is, forecast uncertainty typically increases with the magnitude of the predictand—usually present in forecast models (Scheuerer & Hamill, 2015; Zhao et al., 2015).

2.2. Likelihood and Priors

The posterior distributions of the regression coefficients, $\theta = [\beta, \phi]$, given the data (observed daily streamflow at each gauge and values of hydrometeorological variables) and considering a record length of T days by Bayes' rule, is

$$f(\boldsymbol{\theta}|data) \propto \prod_{t>k} \prod_{i=1}^n \prod_{j \in \mathbf{J}^{(i)}} f\left(Q_t^{(i)} \middle| \boldsymbol{\theta}^{(i)}, Q_{t-k}^{(j \in \mathbf{J}^{(i)})}, \mathbf{X}_{t-k}^{(i)}\right) \cdot f\left(\boldsymbol{\theta}^{(i)} \middle| Q_{t-k}^{(j \in \mathbf{J}^{(i)})}, \mathbf{X}_{t-k}^{(i)}\right) \quad (6)$$

where the term $f\left(Q_t^{(i)} \middle| \boldsymbol{\theta}^{(i)}, Q_{t-k}^{(j \in \mathbf{J}^{(i)})}, \mathbf{X}_{t-k}^{(i)}\right)$ corresponds to Equation 2, and $f\left(\boldsymbol{\theta}^{(i)} \middle| Q_{t-k}^{(j \in \mathbf{J}^{(i)})}, \mathbf{X}_{t-k}^{(i)}\right)$ can be rewritten as

$$f\left(\boldsymbol{\theta}^{(i)} \middle| Q_{t-k}^{(j \in \mathbf{J}^{(i)})}, \mathbf{X}_{t-k}^{(i)}\right) = MVN\left(\boldsymbol{\beta}^{(i)} \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right) \cdot MVN\left(\boldsymbol{\phi}^{(i)} \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right) \cdot f\left(\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right) \cdot f\left(\boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right) \quad (7)$$

where $MVN\left(\boldsymbol{\beta}^{(i)} \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right)$ and $MVN\left(\boldsymbol{\phi}^{(i)} \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right)$ represent probability density of multivariate normal distributions with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}^{(i)}$ corresponding to the priors of $\boldsymbol{\beta}^{(i)} = \left[\beta_1^{(i)}, \beta_2^{(i)}, \beta_3^{(i)}, \beta_4^{(i)}, \beta_x^{(i)}\right]$ and $\boldsymbol{\phi}^{(i)} = \left[\phi_1^{(i)}, \phi_2^{(i)}, \phi_3^{(i)}, \phi_4^{(i)}, \phi_x^{(i)}\right]$ at the gauge i , respectively; and $f\left(\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right)$ and $f\left(\boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right)$ are the priors of the covariance matrix of $\boldsymbol{\beta}^{(i)}$ and $\boldsymbol{\phi}^{(i)}$, which based on Gelman and Hill (2006) are assumed to follow an inverse-Wishart distribution to ensure a positive definite covariance matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)} \text{ Inv wishart}\left(\nu, \mathbf{A}\mathbf{I}\right); \quad \boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)} \text{ Inv wishart}\left(\nu, \mathbf{B}\mathbf{I}\right); \quad (8)$$

where ν corresponds to the degrees of freedom ($m+5$), \mathbf{I} is an $(m+4) \times (m+4)$ identity matrix, and A and B are scalars properly set for $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}$, respectively. $\boldsymbol{\beta}^{(i)}$ and $\boldsymbol{\phi}^{(i)}$ were restricted to be greater or equal than zero in order to ensure positive shape and rate parameters. The model parameters, as can be seen, are modeled jointly to capture their inter-correlations.

3. Application to Narmada River Basin, India

We demonstrate the BHNM with application to Narmada River Basin in west-central India. The study basin, data, potential covariates, implementation and model fitting, and the cross-validation procedure are described below.

3.1. The Study Basin

The Narmada River basin (Figure 2), with 97,882 km² (Narmada basin organization, 2019), originates in the Amarkantak hills of central India and is the largest river that drains into the Arabian Sea in the West. It is a narrow and elongated basin that stretches in the East-West direction (Figure 2). It is an important source of water resources for the populous States of Madhya Pradesh and Gujarat. The basin receives an average rainfall of 1,120 mm, with most of it arriving during the summer monsoon season of June–September. The upper parts of the basin at higher elevations receive higher precipitation relative to the lower basin (Banerjee, 2009). The flooding in the basin mostly occurs during July–August, the focus of our application. The basin and the key streamflow gauges are shown in Figure 2.

3.2. Data

Observed daily streamflow during the peak monsoon season (July–August) at four gauge stations in the Narmada basin: Sandiya, Handia, Hoshangabad, and Mandleshwar were obtained from India Water Resource Information System (IWRIS) (Figure 2 and Table 1) for the period 1978–2014. Garudeshwar gauge station was not considered in this study since it had long missing periods (summers of 1988, 1989, and 1995). The four gauges have drainage areas between 32,494.8 and 71,738.6 km², elevations between 141 and 301 m, mean streamflow, mean seasonal (July–August) streamflow, and max seasonal streamflow seasonal (July–August) streamflow range from 481 to 997 m³s⁻¹, from 1,410 to 2,938 m³s⁻¹, and from 19,700 to 46,398 m³s⁻¹, respectively. Reservoirs inflow and release were not considered since they are not available or have long periods of missing values.

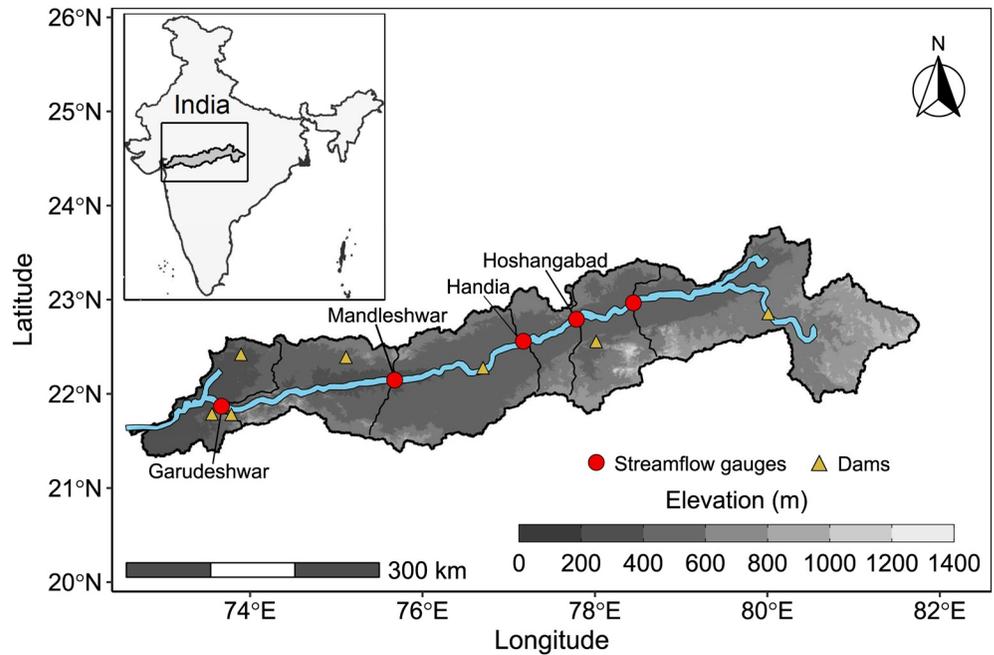


Figure 2. Map of the Narmada basin boundary in India showing the digital elevation model of the basin (SRTM DEM); the locations of five sub-basin outlets: Sandiya, Hoshangabad, Handia, Mandleshwar, and Garudeshwar; and some of the major dams in the basin are marked: Bargi, Tawa, Indirasagar, Jobat, and Sardar Sarovar (from upstream to downstream direction).

For the hydrometeorological variable, we used daily gridded precipitation data from the India Meteorology Department (IMD) for 1978–2014. The gridded precipitation data was prepared using the inverse distance weighted scheme based on observations from 6,995 meteorological stations across India (Pai et al., 2014) and is available at 0.25° spatial resolution from 1951 to 2018. The gridded daily precipitation captures the key features such as high seasonal rainfall over the core monsoon region and orographic rainfall in the Western Ghats and foothills of Himalaya (Pai et al., 2014). Previous studies have widely used the IMD precipitation for hydrometeorological studies (Ali et al., 2019; Shah & Mishra, 2016).

3.3. Potential Covariates

As potential covariates for the monsoon season (July–August), we considered antecedent daily streamflow from an upstream (feeder) gauge j , $Q_{t-1}^{(j \in J(i))}$ ($J(i)$ only have one non-zero component), and s -days accumulated spatial average precipitation from the area between the station gauges i and $i+1$, $P_{sd,t-1}^{(i)}$. The covariates are considered until the previous day (lag-1 day), that is, we have a 1-day lead time for the streamflow forecast. The antecedent streamflow and precipitation capture the hydrologic basin characteristics and forcing input before the streamflow signal on any given day. We did not model the headwater station (Sandiya) since there is not any feeder gauge for it. For the spatial average precipitation, we considered durations of 1, 2, and

Table 1
Basic Data Corresponding to the Streamflow Gauges in the Narmada River Basin Considered in This Study

Gauge	Area (km ²)	Elevation (m)	Mean Streamflow (m ³ s ⁻¹)	Mean seasonal Streamflow (m ³ s ⁻¹)	Max seasonal Streamflow (m ³ s ⁻¹)
Mandleshwar	71,739	141	997	2938	46,398
Handia	51,115	260	785	2487	31,880
Hoshangabad	44,487	292	676	2092	28,600
Sandiya	32,495	301	481	1,410	19,700

3-days, for example, the 2-days accumulated spatial average precipitation corresponds to the cumulative spatial average precipitation of days $t-1$ and $t-2$. Due to the presence of dams for Mandleshwar, the spatial average precipitation was obtained from the area between the station gauge and the upstream dam. Also, for Mandleshwar, we considered the second most immediate gauge (Hoshangabad) as its feeder since the Handia gauges cannot record the peaks of high flow events at least 1 day in advance to Mandleshwar. This, we surmise is due to the operations of the reservoir downstream of Handia.

As initial exploratory analysis, we assessed the strength of the relationship between the covariates and daily streamflow at each gauge by computing the Spearman's rank correlations (see Figure S1). It can be seen that all the covariates showed significant correlation with the predictands.

At all gauges for days when $P_{sd,t-1}^{(i)} = 0$, we considered $Q_{t-1}^{(j \in \mathbf{J}(i))}$ as the covariate. The high lag 1 correlation at each gauge was verified (see Figure S2).

3.4. Implementation and Model Fitting

We fitted various candidate BHNMs, each one with different combination of covariates presented in the previous section (streamflow at feeder gauges and different l -days accumulated spatial average precipitation from the previous day) and potential distribution choice for the marginal—Gamma and Lognormal. For the Gamma distribution, rate and shape parameters were parametrized in terms of the mean and variance (Wilks, 2011). For Lognormal, the log transformation was applied to the predictands and covariates, and the standard deviation was set stationary because of its positive restriction ($\sigma > 0$). We considered the deviance information criterion (DIC; Spiegelhalter et al., 2002) as metric for selecting the best model. The DIC corresponds to a hierarchical modeling generalization of the Akaike information criterion (AIC; Akaike, 2011) and facilitates Bayesian model selection. The DIC was computed for each candidate BNM and the candidate with the minimum DIC was selected as the best model.

Candidates BNM were implemented in R (R Core, 2017) using the program STAN (Stan Development Team, 2014) and the R package RStan (Stan Development Team, 2020) which provides an interface from R to the STAN library for Bayesian data analysis. Posterior distributions of the parameters and predictive posterior distributions of the streamflows (ensembles) for all days were estimated using the No-U-Turn Sampler (Hoffman & Gelman, 2014) for the Markov Chain Monte Carlo method (Gelman & Hill, 2006; Robert & Casella, 2011) based on the priors assigned. We ran three parallel chains with different initial values, and each simulation was performed for 8,000 iterations with a burn-in size value of 4,000 to ensure convergence. To reduce the sample dependence (autocorrelation), we chose a thinning factor of 4. The scale reduction factor \hat{R} (Gelman & Rubin, 1992) was used to check the model convergence in that \hat{R} values less than the critical value of 1.1 suggests good convergence of the model. In all of our runs the \hat{R} values were less than 1.1 at 3,000 samples, indicating model convergence. Consequently, the posterior distributions of the parameters and the predictive posterior distribution of daily streamflows consists of 3,000 ensembles for each candidate BNM. Also, for each candidate we checked that its regression coefficients were significant (posterior PDFs do not contain zero in the 95% credible interval). For the priors of the covariance matrix according to Equation 8, for each gauge, we considered weakly informative priors with $\nu = 6$, $A = 10$, and $B = 10$.

3.5. Model Cross-Validation

To assess the out-of-sample predictability of the best model, we performed leave-1-year-out cross-validation for the period 1978–2014 (37 years). There is a total of $37 \times 62 = 2294$ days of data. The observations in a year are selected as validation data (62 days), the data from the remaining 36 years (2232 days), as training data. The trained BNM is applied to provide estimates for the one validation year. This cross-validation procedure was repeated 37 times.

We used a generalized linear model (GLM; Dobson & Barnett, 1992; Dunn & Smyth, 2018) with a Gamma distribution and the covariates selected from the DIC above, as the equivalent non-Bayesian alternative considered as the reference forecast. Since in Markov model the prediction at any gauge depends only on the

immediately gauge, a comparison with a GLM is valid. The GLM model was fitted at each station separately via the Maximum Likelihood (ML) method.

To evaluate the forecast accuracy we considered traditional deterministic metrics such as the correlation (R), the Nash-Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970), and the relative bias (BIAS) in percent, which are computed for the ensembles mean. In addition, to evaluate the probabilistic skill and reliability of the forecast, three verification metrics were computed: rank histograms, the continuous ranked probability skill score (CRPSS), and the energy skill score (ESS).

Rank histograms indicate the level of uniform distribution of observations throughout the ensemble forecast and, thus, its reliability. A rank histogram is computed from the rank or position of the observed value relative to the ensemble members over a number of cases (the length of the validation records; Hamill, 2001; Mendoza et al., 2015). If the ensemble at a given point is reliable, the resulting rank histogram should be uniform (flat rank histogram). Overpopulation of the lowest or highest ranks is a sign of positive or negative biases in the ensemble forecast. A lack of variability in the ensemble will show up as a U-shaped, or concave, rank population. Overpopulation of the middle ranks means an excess of dispersion (overdispersion). It should be noted that a flat rank histogram is a necessary but not sufficient condition for determining that the ensemble is reliable (Hamill, 2001).

Along with the rank histogram, we also consider a discrepancy index (DI) to quantify the departure of the histogram from uniformity (Delle Monache et al., 2006; Mendoza et al., 2015). It is computed as follows:

$$DI = \sum_{i=1}^{M+1} \left| \frac{count_i}{T} - \frac{1}{M+1} \right| 100 \quad (9)$$

where M is the number of ensemble forecast members (so $M + 1$ is the number of bins in the rank histogram), $count_i$ is the number of times the observed event falls into the i th bin, and T is the sample size. Lower DI means that the ensemble better achieves the condition of reliability.

The continuous rank probability score (CRPS) provides an integrated evaluation of the forecast accuracy and reliability of the forecast ensemble spread by estimating the area between the cumulative distribution functions of the forecasted streamflow and the observed streamflow (Gneiting & Raftery, 2007; Hersbach, 2000). For a station gauge on a specific day, it is defined as

$$CRPS = \int_{-\infty}^{\infty} [F(Q) - H(Q - Q_o)]^2 dQ \quad (10)$$

$$H(Q - Q_o) = \begin{cases} 0 & Q < Q_o \\ 1 & Q \geq Q_o \end{cases} \quad (11)$$

where $F(Q)$ is the CDF associated with the forecast, Q_o is the observed streamflow, and $H(Q - Q_o)$ is the well-known Heaviside function. The CRPSS is then defined accordingly:

$$CRPSS = 1 - \frac{\overline{CRPS}_{fcast}}{\overline{CRPS}_{ref}} \quad (12)$$

where \overline{CRPS}_{fcast} is the average CRPS of the forecast model, \overline{CRPS}_{ref} is the average CRPS of the reference forecast. The CRPSS ranges from $-\infty$ to 1. $CRPSS < 0$ indicates that the reference forecast has higher skill than the forecast model, $CRPSS = 0$ implies equal skill, and $CRPSS > 0$ implies that the forecast model has a higher skill, with $CRPSS = 1$ being a perfect score. For a deterministic forecast, such as GLM, CRPS corresponds to the mean absolute error (MAE; Zhao et al., 2015).

The energy score (ES) assesses the multivariate (jointly across the gauges) accuracy and reliability of the forecasts ensemble spread (Gneiting & Raftery, 2007; Gneiting et al., 2008):

$$ES = \frac{1}{M} \sum_{j=1}^M \|\mathbf{Q}_j - \mathbf{Q}_o\| - \frac{1}{2M^2} \sum_{i=1}^M \sum_{j=1}^M \|\mathbf{Q}_i - \mathbf{Q}_j\| \quad (13)$$

where M is the number of ensemble forecast members, \mathbf{Q}_j is the $n \times 1$ vector of the j th ensemble forecast at day t , \mathbf{Q}_o is the $n \times 1$ vector of observed streamflow at day t , and $\|\cdot\|$ denotes the Euclidean norm. This is a direct generalization of the continuous ranked probability score (Equation 10), to which the energy score reduces in dimension $d = 1$. Then, the energy skill score (ESS) is defined as

$$\text{ESS} = 1 - \frac{\overline{\text{ES}}_{\text{fcst}}}{\overline{\text{ES}}_{\text{ref}}} \quad (14)$$

where $\overline{\text{ES}}_{\text{fcst}}$ is the average ES of the forecast model, $\overline{\text{ES}}_{\text{ref}}$ is the average ES of the reference forecast. As for the CRPSS, the ESS ranges from $-\infty$ to 1, and its values have the same meaning. For GLM, ES corresponds to the Euclidean norm of the mean absolute error (MAE) from the gauges modeled.

We also evaluated the BHNm performance for different types of events by considering sample stratification for the probabilistic metrics (Bellier et al., 2017) by defining three event categories (strata). The three categories defined are “low,” “normal,” and “high” forecast events, defined as those days when the ensemble forecast mean ($\overline{Q}(t)$) is below the median observed ($Q_{50\text{th}}$), between 50th and 80th observed quantiles ($Q_{50\text{th}}$ and $Q_{80\text{th}}$) and above the 80th observed quantile ($Q_{80\text{th}}$), respectively.

To evaluate the reliability and accuracy of the joint forecast, we computed the basin average specific streamflow (i.e., streamflow per unit area). Then, the days in each strata are obtained as described above. Table 2 shows the threshold considered for each gauge and the basin average specific streamflow. Time series of observed daily streamflow at three gauges for 1978–2014 showed that above the upper threshold ($Q_{80\text{th}}$) are most of the high flow events while between 0 and the lower threshold ($Q_{50\text{th}}$), only low flow events are contained (see Figure S3).

The uncertainty of CRPSS and ESS metrics was generated by applying bootstrap to the time position vector of the observed streamflow.

4. Results

4.1. Best Model Selection

Candidates BHNm were fitted for the entire period (1978–2104) and their DIC values are shown in Table 3. It is important to mention that for all candidate BHMs, significance of the regression coefficients (β and ϕ) were checked (posterior PDFs do not contain zero in the 95% credible interval). As mentioned in Section 3.4, the standard deviation (σ) was considered stationary for the lognormal distribution. The best BHNm was model 1, which considered as covariates the streamflow at the feeder site ($Q_{t-1}^{(j \in J(i))}$) and the 1-day accumulated spatial average precipitation ($P_{1d,t-1}^{(i)}$) at the day $t-1$ for the gauge i . Gamma distribution is seen to perform better than Lognormal—which is consistent with other studies (e.g., Razack & Lasm, 2006), in that Lognormal exhibits higher variability especially for high flows from back transformation.

The schematic of the best BHM model for the basin is shown in Figure 3 and the model is presented below. The daily streamflow at each gauge i , $Q_t^{(i)}$, follow a Gamma distribution

Table 2 Thresholds Considered for Stratification Analysis		
Gauge	Streamflow threshold ($\text{m}^3 \text{s}^{-1}$)	
	$Q_{50\text{th}}$	$Q_{80\text{th}}$
Mandleshwar	1,460	3,919
Handia	1,324	3,520
Hoshangabad	1,000	2,914
Specific streamflow threshold (md^{-1})		
Basin average	4.9	20.6

Table 3
DIC Values for Candidates BHM Considered in the Best Model Selection

$f(\alpha_t^{(i)}, \lambda_t^{(i)})$	Model	Covariates			DIC
		1: Mandleshwar	2: Handia	3: Hoshangabad	
Gamma	1	$Q_{t-1}^{(3)}, P_{1d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{1d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{1d,t-1}^{(3)}$	11,148
	2	$Q_{t-1}^{(3)}, P_{2d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{2d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{2d,t-1}^{(3)}$	11,763
	3	$Q_{t-1}^{(3)}, P_{3d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{3d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{3d,t-1}^{(3)}$	12,355
Lognormal ^a	4	$Q_{t-1}^{(3)}, P_{1d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{1d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{1d,t-1}^{(3)}$	12,104
	5	$Q_{t-1}^{(3)}, P_{2d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{2d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{2d,t-1}^{(3)}$	12,334
	6	$Q_{t-1}^{(3)}, P_{3d,t-1}^{(1)}$	$Q_{t-1}^{(3)}, P_{3d,t-1}^{(2)}$	$Q_{t-1}^{(4)}, P_{3d,t-1}^{(3)}$	12,470

Note. The best model corresponds to model 1, which has the lowest DIC.

DIC, deviance information criterion.

^a $\sigma_t^{(i)} = \phi_t^{(i)}$ for models 4, 5, and 6 (see Equation 5).

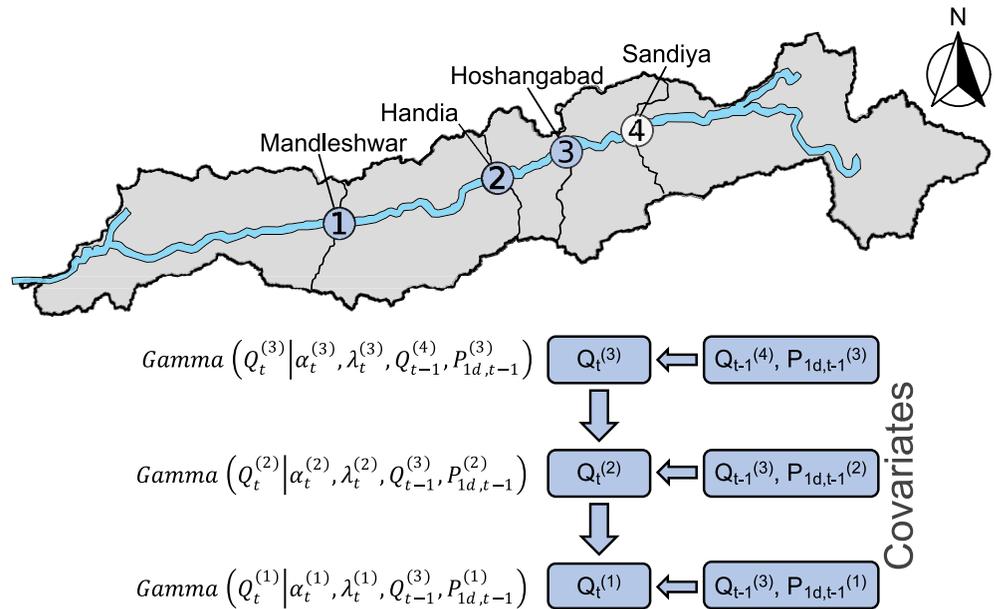


Figure 3. Schematic of the Bayesian Hierarchical Network Model for the Narmada River basin. $Q_t^{(i)}$ corresponds to the observed streamflow at gauge i and day t (predictand), $P_{1d,t-1}^{(i)}$ to 1-day accumulated spatial average precipitation at gauge i and day t , and $Q_{t-1}^{(j \in \mathcal{J}(i))}$ to the daily streamflow at gauge j and day t . For each gauge $Q_{t-1}^{(j \in \mathcal{J}(i))}$ was considered as the only covariate when $P_{1d,t-1}^{(i)} = 0$.

$$f\left(Q_t^{(i)} \mid \alpha_t^{(i)}, \lambda_t^{(i)}, Q_{t-1}^{(j \in J^{(i)})}, P_{1d,t-1}^{(i)}\right), \quad i = 1, 2, 3 \quad (15)$$

where $\alpha_t^{(i)} > 0$ is the shape parameter and $\lambda_t^{(i)} > 0$ is the rate parameter at the gauge i and day t . These parameters can be expressed in terms of the expected value, $\mu_t^{(i)}$, and variance, $\left(\sigma_t^{(i)}\right)^2$, of $Q_t^{(i)}$ (Wilks, 2011) as follows:

$$\alpha_t^{(i)} = \frac{\left(\mu_t^{(i)}\right)^2}{\left(\sigma_t^{(i)}\right)^2}, \quad \lambda_t^{(i)} = \frac{\mu_t^{(i)}}{\left(\sigma_t^{(i)}\right)^2} \quad (16)$$

Expressions for $\mu_t^{(i)}$ and $\left(\sigma_t^{(i)}\right)^2$ at each station are the following:

Mandleshwar ($i = 1$):

$$\mu_t^{(1)} = \begin{cases} \beta_1^{(1)} + \beta_2^{(1)} Q_{t-1}^{(3)} & P_{1d,t-1}^{(1)} = 0 \\ \beta_3^{(1)} + \beta_4^{(1)} Q_{t-1}^{(3)} + \beta_5^{(1)} P_{1d,t-1}^{(1)} & P_{1d,t-1}^{(1)} > 0 \end{cases} \quad (17)$$

$$\sigma_t^{(1)} = \begin{cases} \phi_1^{(1)} + \phi_2^{(1)} Q_{t-1}^{(3)} & P_{1d,t-1}^{(1)} = 0 \\ \phi_3^{(1)} + \phi_4^{(1)} Q_{t-1}^{(3)} + \phi_5^{(1)} P_{1d,t-1}^{(1)} & P_{1d,t-1}^{(1)} > 0 \end{cases} \quad (18)$$

Handia and Hoshangabad ($i = 2, 3$):

$$\mu_t^{(i)} = \begin{cases} \beta_1^{(i)} + \beta_2^{(i)} Q_{t-1}^{(i+1)} & P_{1d,t-1}^{(i)} = 0 \\ \beta_3^{(i)} + \beta_4^{(i)} Q_{t-1}^{(i+1)} + \beta_5^{(i)} P_{1d,t-1}^{(i)} & P_{1d,t-1}^{(i)} > 0 \end{cases} \quad (19)$$

$$\sigma_t^{(i)} = \begin{cases} \phi_1^{(i)} + \phi_2^{(i)} Q_{t-1}^{(i+1)} & P_{1d,t-1}^{(i)} = 0 \\ \phi_3^{(i)} + \phi_4^{(i)} Q_{t-1}^{(i+1)} + \phi_5^{(i)} P_{1d,t-1}^{(i)} & P_{1d,t-1}^{(i)} > 0 \end{cases} \quad (20)$$

where $P_{1d,t-1}^{(i)}$ is the 1-day spatial average precipitation for the gauge station i . Posterior distributions of the model coefficients ($\beta^{(i)}$ and $\phi^{(i)}$) at each gauge showed that BHNM captures the correlation between the coefficients (see Figures S4–S9).

Figure 4 shows the predictive posterior distribution ensembles of streamflows from the best BHNM for July–August daily streamflow for Mandleshwar (the terminal gauge) for the whole record (1978–2014, Figure 4a) and the year 2013 (which is a high flow year) (Figure 4c) for a closer visualization of the timing of the flows. The flow ensembles are generated from the Gamma distribution using the posterior samples of the model coefficients (Equations 17–20). The flow ensembles are presented as time series of boxplots. Also, Figure 4b displays the scatter plot of daily observed streamflow versus ensemble mean streamflow. The ensembles mean daily flows are generally slightly lower than the observed flows (Figures 4a and 4c); however, almost all the observed high flows are captured within the ensemble spread with few exceptions (e.g., high value in 1996, Figure 4a). The daily streamflow timing is captured very well by the posterior ensembles, as can be seen in Figure 4c for the high flow year 2013. Correlation coefficients (R), NSE, and BIAS for 2018–2014 and 2013 are displayed in Figures 4b and 4c, respectively. The R values are 0.83 and 0.92 for 1978–2014 and 2013, respectively, while BIAS and NSE values are 1.8% and -8.1% , and 0.69 and 0.82 for 1978–2014 and 2013, respectively.

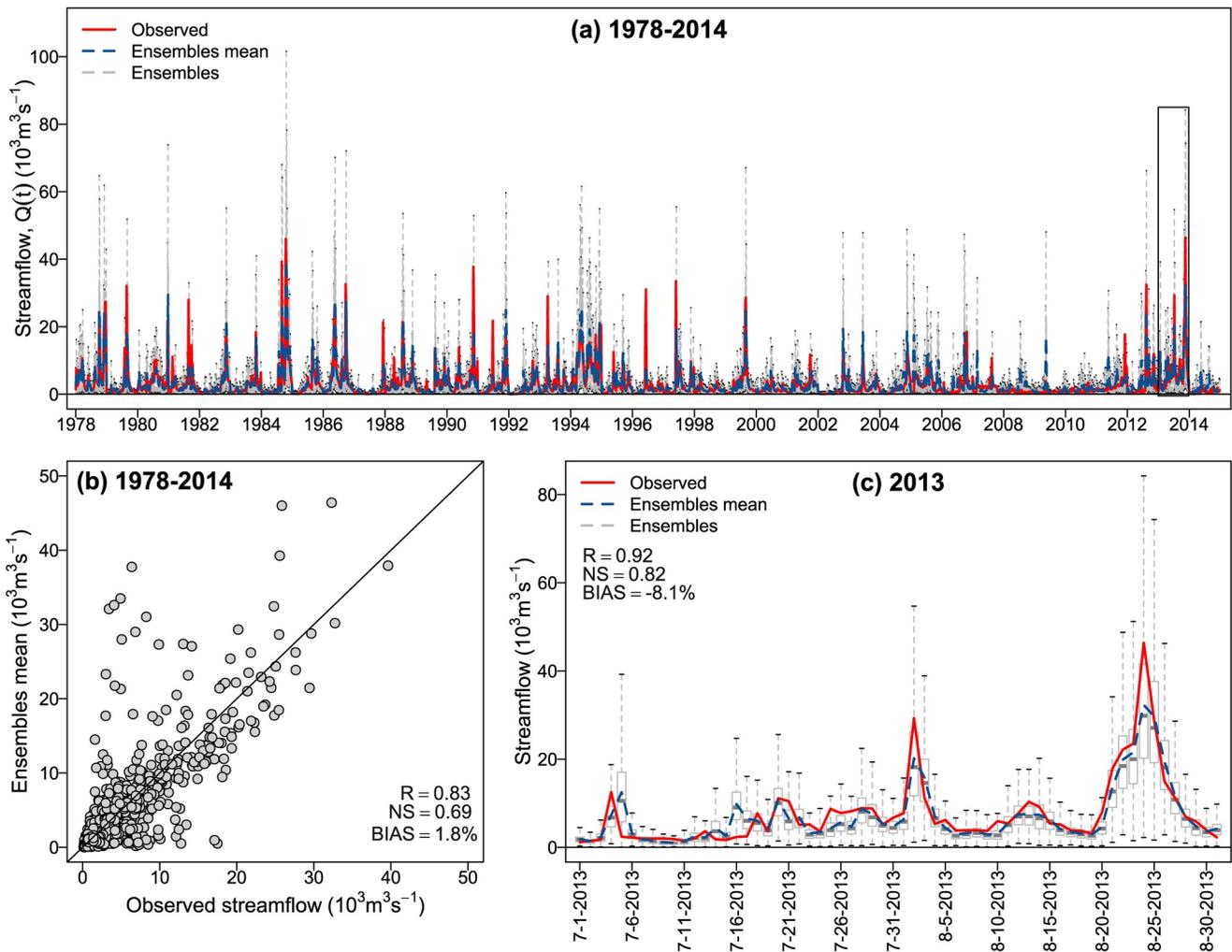


Figure 4. (a) Predictive posterior distribution ensembles of July–August daily streamflow, and (b) scatter plot of daily observed streamflow versus predictive posterior ensembles mean streamflow for the entire record (1978–2014), and (c) predictive posterior distribution ensembles of July–August daily streamflow for 2013 (high flow year) at Mandleshwar gauge. The boxplots represent the posterior distribution estimates of the daily streamflow. Outliers are not displayed. Red lines correspond to the observed daily streamflow and blue-dashed lines to the posterior ensembles mean daily streamflow. R, Nash-Sutcliffe efficiency (NSE), and relative bias (BIAS) values for 1978–2014 and 2013 are displayed in panels (b) and (c). R values are significant (p - value < 0.05).

Table 4
R, NSE, and BIAS Values for the Entire Record (1978–2014) and the Year 2013 (High Flow Year)

Gauge	1978–2014			2013		
	R	NSE	BIAS (%)	R	NSE	BIAS (%)
Mandleshwar	0.83	0.69	1.8	0.92	0.82	–8.1
Handia	0.89	0.77	1.1	0.92	0.85	0.4
Hoshangabad	0.87	0.75	–9.6	0.92	0.84	–6.9

BIAS, relative bias; NSE, Nash-Sutcliffe efficiency.

We achieved even better performance for the other two gauges (Figures S10 and S11) as can be seen in Table 4 which displays a summary of the accuracy metrics.

In addition, Figure 5 shows the spatial dependence of July–August daily observed streamflow (Figure 5a) and ensemble mean streamflow from BHNM for the three gauges. The spatial dependence was assessed using Pearson correlation coefficient for each pair of gauges. It is seen that the spatial dependence structure is strong for the observed data (R ranging from 0.77 to 0.93), and ensemble mean streamflow from the BHNM captures them very well. Note that the GLM can also capture this spatial dependence structure (see Figure S12) as this is mainly from the skillful covariates considered.

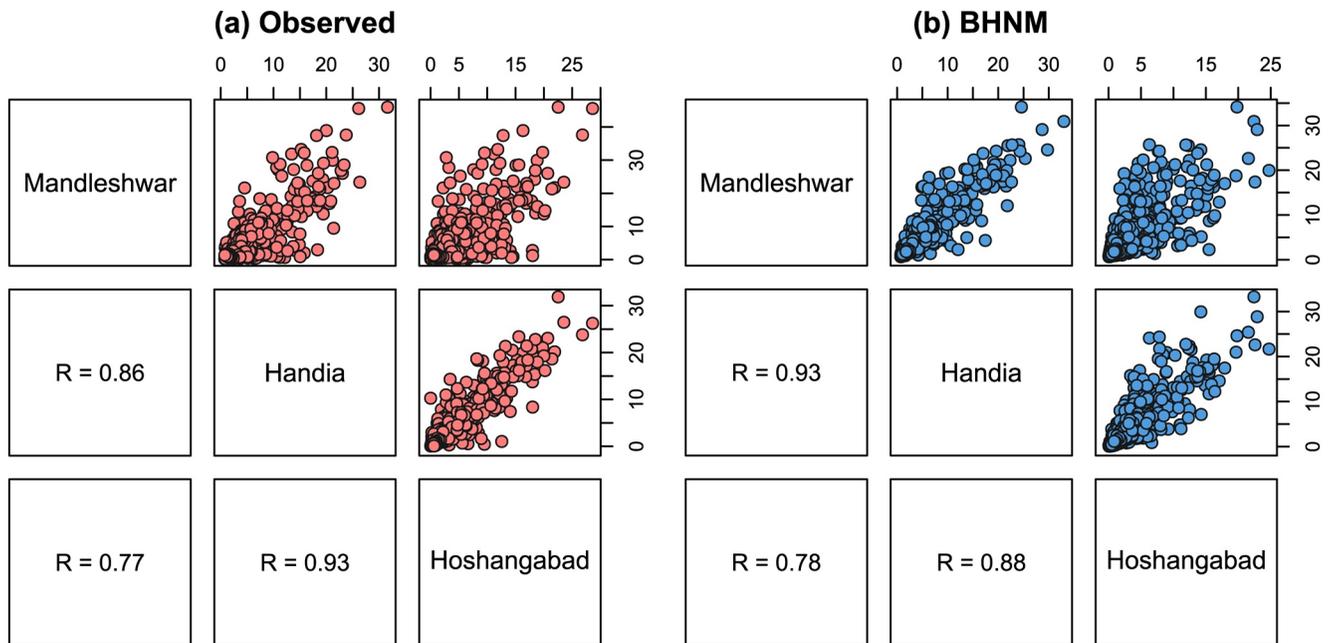


Figure 5. Pairwise scatter plots (upper triangular matrix) and Pearson correlation coefficients (R, lower triangular matrix) of July–August daily (a) observed streamflow and (b) ensembles mean forecast streamflow from Bayesian Hierarchical Network Model for the three gauges analyzed in this study. Streamflow values are in ($103\text{m}^3\text{s}^{-1}$). All Pearson correlation coefficients are significant (p - value < 0.05).

4.2. Cross-Validation

The leave 2-year out cross-validation, following the procedure described in Section 3.5, was performed for both BHNM and GLM models. Figure 6 shows the predictive posterior ensemble forecast of July–August daily streamflow at Mandleshwar for the year 2013. As in Figure 4c, Figure 6a shows that for Mandleshwar, the ensemble mean slightly underestimates the observed for high streamflow values, but the ensembles spread can capture most of them except for a few events (July 30, 1991 and July 29, 1996). While accuracy metrics (R, NSE, and BIAS) present similar values as for BHNM calibration, which an indication of good out-of-sample predictability for the BHNM for Mandleshwar and the other two gauges (see Figures S13 and S14). Table 5 summarizes performance accuracy metrics for both models BHNM and GLM. It can be seen that the BHNM provides higher performance compared to GLM.

4.3. Probabilistic Performance Metrics at Site

To assess the at site (marginal) probabilistic skill, we computed the CRPSS for BHNM (GLM was considered as the reference forecast) and rank histograms at each gauge and for the entire period (1978–2014) and the three strata (low, normal, and high forecasts).

Figure 7 shows CRPSS distribution of BHNM for the entire period and three strata (see the color legend). At all the gauges the skills are very good (95% of the distributions above ~ 0.4) for the entire period (i.e., total) with a clear increase for Hoshangabad (95% of the distributions above 0.65). Performance for low flow forecasts is similar to that for the entire period, except at Handia, where the CRPSS distribution is high (above 0.5). For normal flow forecast there is a decrease of the CRPSS compared to total, this reduction is more evident for Handia. For high flow forecast events, there is a clear increase in skill compared

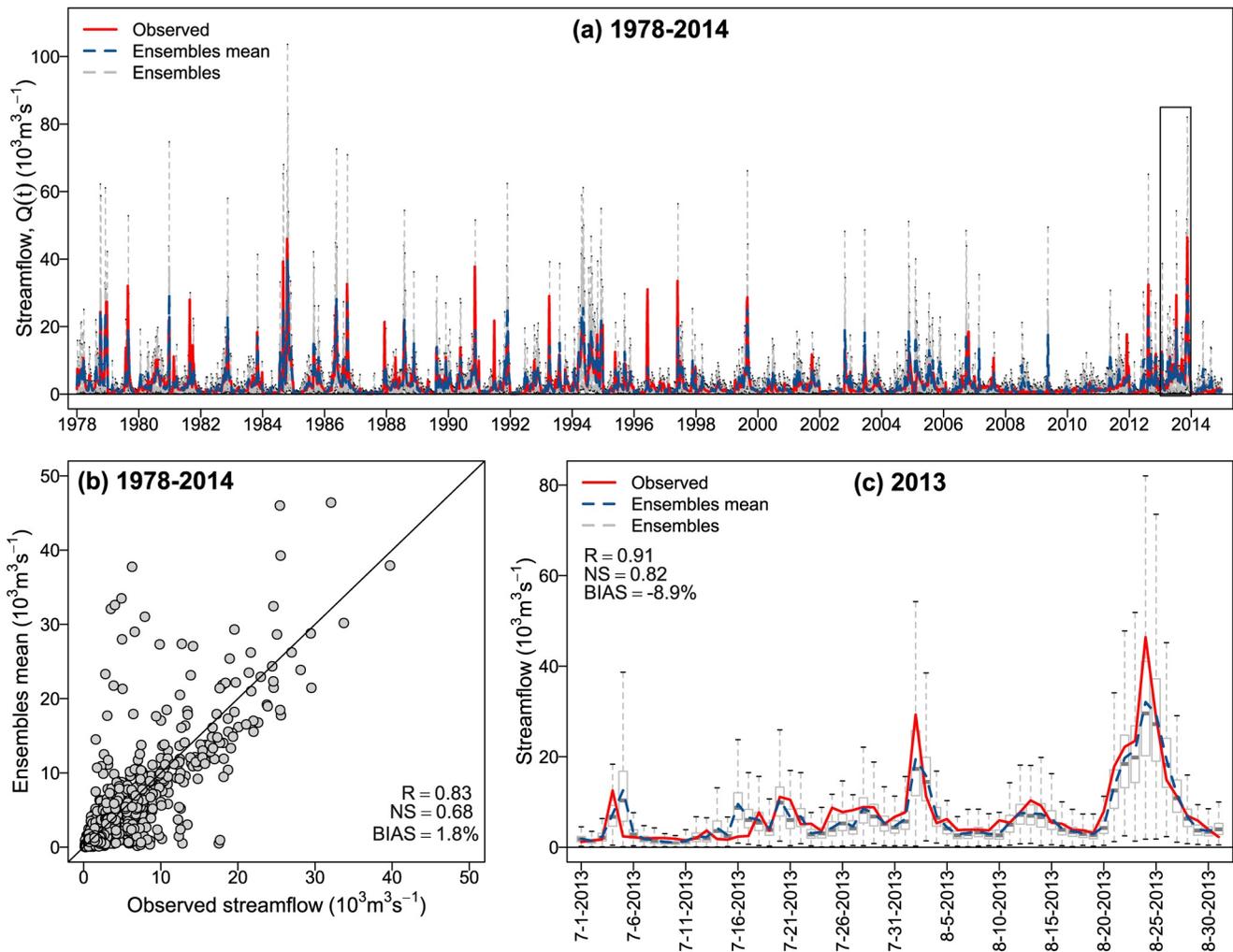


Figure 6. As in Figure 4 but predictive posterior ensemble forecast of July–August daily at Mandleshwar for the cross-validation.

Table 5
Summary of Accuracy Metrics for BHNM and GLM for the Period 1978–2014

Model	Metric	Mandleshwar	Handia	Hoshangabad
BHNM	R	0.83	0.89	0.86
	NSE	0.68	0.77	0.74
	BIAS (%)	1.8	1.1	-4.6
GLM	R	0.75	0.87	0.81
	NSE	0.33	0.59	-0.19
	BIAS (%)	25.2	25.8	71.3

Note. For BHNM, the metrics were computed considering the ensembles mean.

BHNM, Bayesian Hierarchical Network Model; GLM, generalized linear regression.

to the total. CRPSS values for the entire period and strata shows that the BHNM is highly skillful relative to the GLM.

Figure 8 displays rank histograms of the BHNM ensembles streamflow along with discrepancy indexes (DI) obtained from the cross-validation for each station (columns) and the entire period and the three strata (rows). Ensembles for the entire period and the first two strata (low and normal flow forecast events) at the three gauges show over-dispersive (underconfident) rank histograms but close to a uniform shape (red line corresponds to the median value) that indicates an adequate reliability overall. This feature is also confirmed by not so high DI values (most of them below 40%). For High flow forecast events (last row), rank histograms show a slightly over-forecasting bias for Mandleshwar and Handia gauges with DI values of 50.5% and 35.2%, respectively. For Hoshangabad, rank histogram shows a slightly under-dispersion (overconfident) with a DI value of 35.6%.

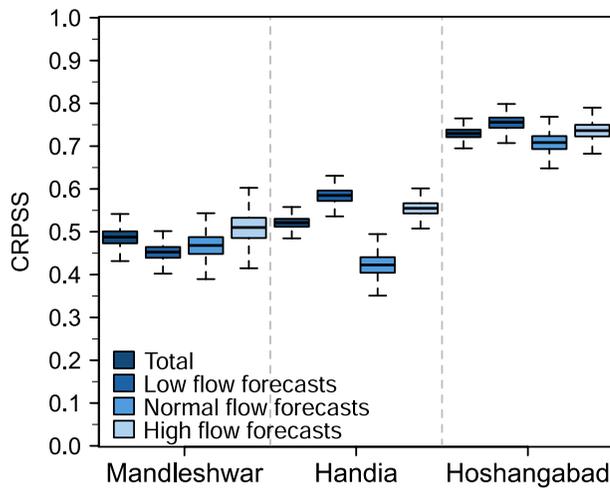


Figure 7. Boxplots of continuous ranked probability skill score at the three gauges for the total period; low flow forecasts strata ($\hat{Q}(t) < Q_{50th}$); normal flow forecasts strata ($Q_{50th} \leq \hat{Q}(t) \leq Q_{80th}$); and high flow forecasts strata ($\hat{Q}(t) > Q_{80th}$). The whiskers show the 95% credible intervals, boxes the interquartile range, and horizontal lines inside the boxes, the median. Outliers are not displayed. Generalized linear regression was considered as the reference forecast model.

4.4. Probabilistic Multivariate Performance Metrics

To assess the joint skill and reliability of the BHNM ensembles, we computed the rank histograms for the total period and each strata using the basin average specific streamflow as was described in Section 3.5. For computing the ESS for the three strata, we considered the events (days) classification obtained for the Basin average specific streamflow.

Figure 9 shows ESS distributions of BHNM for the entire period and three strata. ESS distribution for the total period and three strata show that BHNM provides a better joint accuracy and reliability compared to GLM (95% of the distributions above 0.46). Also, compared to the marginal skill (CRPSS), ESS shows a better joint performance (ESS values higher than CRPSS), except for Hoshangabad.

Figure 10 displays rank histograms of the BNHM ensembles basin average specific streamflow along with discrepancy indexes (DI) obtained from the cross-validation for the entire period and the three strata. Although rank histograms for the entire period (Figure 10a) and low flow forecast (Figure 10b) show an over-forecasting bias, they are closer to the uniform shape (DI values below 25%) than at site ranks histograms (Figure 8). For normal flow forecast strata, shape of rank histogram is close to the uniform shape, except for the highest ensemble percentile bin, and also present a low DI value (20%). Finally, rank histogram for high flow forecast presents an under-dispersion (overconfident) shape, but a lower DI value (~31%) than the three gauges (last row in Figure 8). Thus, rank histograms also show a better reliability than at site rank histograms (Figure 8) overall.

The better joint performance is also checked for the predictive posterior ensemble forecast of July-August daily basin average specific streamflow (see Figure S15).

5. Summary and Discussion

In this study, we formulated and presented a BHNM for daily streamflow modeling and forecasting. The model uses the spatial dependence induced by the river network topology and hydrometeorological variables from the upstream contributing area between the covariates' gauges. The daily streamflow at each gauge is assumed to be conditionally independent and distributed as general probability density function with time-varying distribution parameters. The distribution parameters for each day and at each gauge are modeled as linear functions of potential covariates.

We applied this to forecast daily summer (July–August) streamflow at three gauges in the Narmada River Basin network in west-central India for the period 1978–2014, at 1 day lead time. The best BHNM was selected based on the lowest DIC, considered a Gamma distribution as the suitable distribution for each gauge, and included daily streamflow from upstream feeder gauges and 1-day accumulated spatial average precipitation from the area between two consecutive gauges from the previous day as covariates.

BHNM ensembles forecast showed to be more skillful than GLM for both at the site and jointly, that is, for both cases, skill scores (CRPSS or ESS) for the total period and each forecast strata were above zero. In general, BHNM provides reliable marginal (each gauge) and joint forecast distribution, but with a slight over-dispersion for some strata.

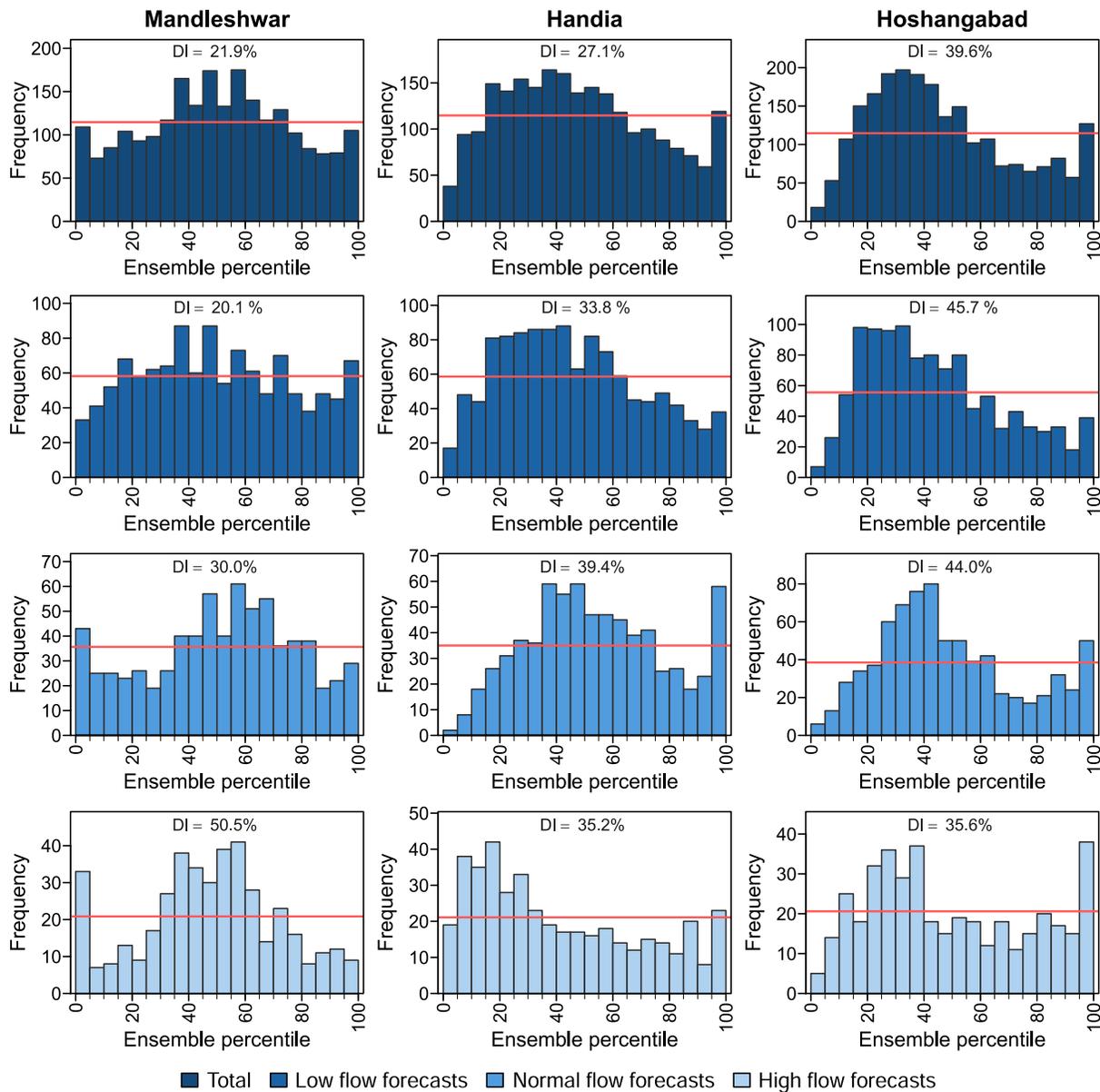


Figure 8. Rank histograms of the Bayesian Hierarchical Network Model ensembles forecast of July–August daily streamflow obtained from the cross-validation for the total period; low flow forecasts strata ($\hat{Q}(t) < Q_{50th}$); normal flow forecasts strata ($Q_{50th} \leq \hat{Q}(t) \leq Q_{80th}$); and high flow forecasts strata ($\hat{Q}(t) > Q_{80th}$). Horizontal red lines correspond to the median frequency. DI denotes the discrepancy index.

Overall, the BHNM framework has the following benefits:

1. Using the network structure in incorporating flow information from upstream gauges and precipitation contributing areas between gauges as covariates, communicates information through the network and captures the spatial correlation of flows simultaneously.
2. Allows capturing the timing of the high flows through via feeder gauges that capture high flows peak in advance as covariates.
3. Can capture non-Gaussian features.
4. Allows capturing the heteroscedasticity of the data by considering nonstationary variance.
5. Compared to standard non-Bayesian alternative GLM, BHNM provides robust estimates of uncertainties.

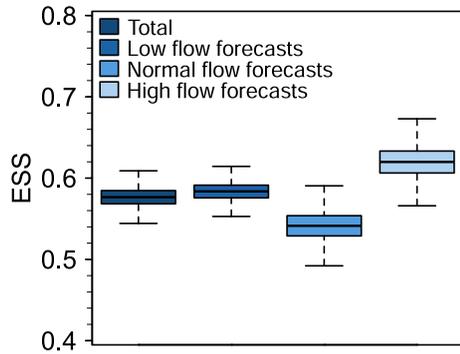


Figure 9. Boxplots of the energy skill score (ESS) of daily streamflow ensembles from Bayesian Hierarchical Network Model for the total period; low flow forecasts strata ($\hat{Q}(t) < Q_{50th}$); normal flow forecasts strata ($Q_{50th} \leq \hat{Q}(t) \leq Q_{80th}$); and high flow forecasts strata ($\hat{Q}(t) > Q_{80th}$). The whiskers show the 95% credible intervals, boxes the interquartile range, and horizontal lines inside the boxes, the median. Outliers are not displayed. Generalized linear regression was considered as the reference forecast model.

Although the BHNM was applied to a river network with only four gauges (headwater gauge was only used as a predictor), the framework can be conceptually easily extended to a bigger and more complex basin, as was shown by (Ravindranath et al., 2019). The model's success lies in the availability of skillful covariates and correct identification of the feeder gauges. In that sense, the ability to capture the high flow timing depends on identifying feeder gauges that have a lag of at least 1 day for the peak flow relative to the predictand. This requirement could be a problem in river networks with multiple junctions where some feeder gauges cannot capture the peak flow at least 1 day in advance. This could be alleviated by considering all possible combinations of covariates which will increase computational time.

Potential extensions of this framework include combining this with deterministic physical model forecasts to provide multi-model ensemble forecasts, use of other covariates such as reservoir levels, releases and so on to capture human affects and also in snow-melt basins.

All of these findings indicate that this BHNM can be used in real-time coordinated flood mitigation and early warning across the river network basin at all locations.

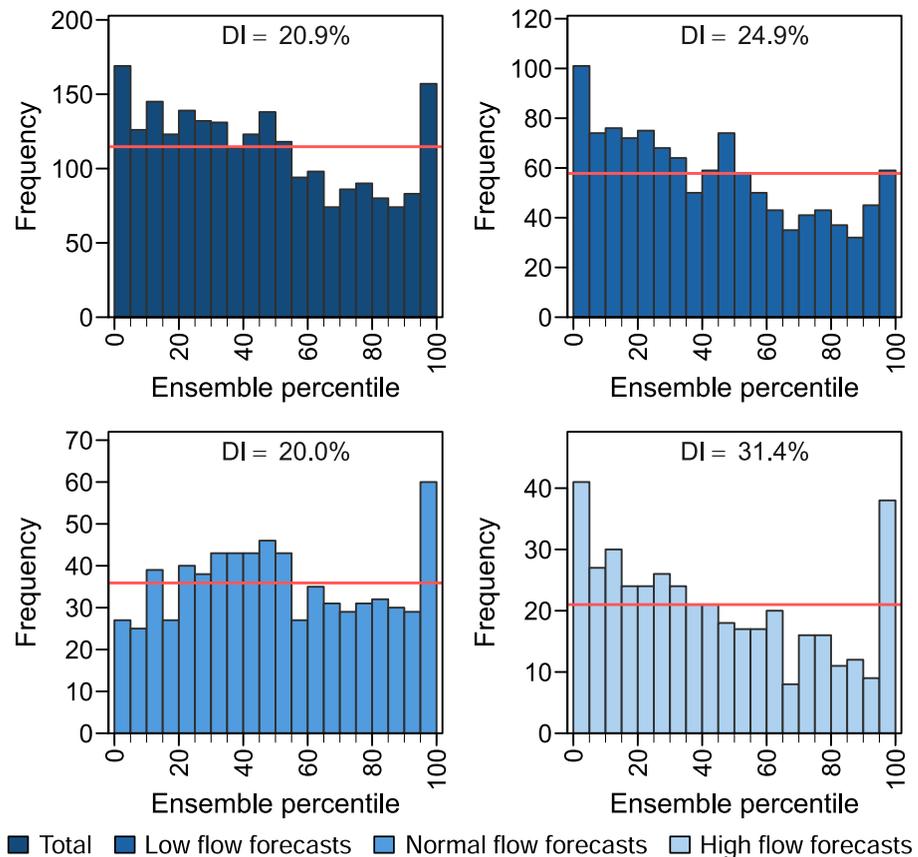


Figure 10. Rank histograms of the Bayesian Hierarchical Network Model ensembles forecast of July–August daily basin average specific streamflow obtained from the cross-validation for the total period; low flow forecasts strata ($\hat{Q}(t) < Q_{50th}$); normal flow forecasts strata ($Q_{50th} \leq \hat{Q}(t) \leq Q_{80th}$); and high flow forecasts strata ($\hat{Q}(t) > Q_{80th}$). Horizontal red lines correspond to the median frequency. DI denotes the discrepancy index.

Data Availability Statement

The data set used in this study, which consists of time series of potential covariates and daily monsoon period (July–August) streamflow for four station gauges in the Narmada River basin, can be downloaded from <https://www.hydroshare.org/resource/34e0b581418442068e796762aa4fc4d8/>.

Acknowledgments

This project was funded by the Monsoon Mission project of the Ministry of Earth Sciences, India. We also acknowledge the support from the Fulbright Foreign Student Program and the National Agency for Research and Development (ANID) Scholarship Program/DOCTORADO BECAS CHILE/2015–56150013 to the first author.

References

Abdollahi, S., Raeisi, J., Khalilianpour, M., Ahmadi, F., & Kisi, O. (2017). Daily mean streamflow prediction in perennial and non-perennial rivers using four data driven techniques. *Water Resources Management*, 31(15), 4855–4874. <https://doi.org/10.1007/s11269-017-1782-7>

Akaike, H. (2011). Akaike's information criterion. In M. Lovric (Ed.), *International encyclopedia of statistical science* (p. 25). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_110

Ali, H., & Mishra, V. (2018). Increase in subdaily precipitation extremes in India under 1.5 and 2.0 C warming worlds. *Geophysical Research Letters*, 45(14), 6972–6982. <https://doi.org/10.1029/2018GL078689>

Ali, H., Modi, P., & Mishra, V. (2019). Increased flood risk in Indian sub-continent under the warming climate. *Weather and Climate Extremes*, 25, 100212. <https://doi.org/10.1016/j.wace.2019.100212>

Banerjee, R. (2009). *Review of water governance in the Narmada river basin* (Tech. Rep.). Society for Promotion of Wastelands Development.

Bellier, J., Zin, I., & Bontron, G. (2017). Sample stratification in verification of ensemble forecasts of continuous scalar variables: Potential benefits and pitfalls. *Monthly Weather Review*, 145(9), 3529–3544. <https://doi.org/10.1175/MWR-D-16-0487.1>

Bracken, C., Holman, K. D., Rajagopalan, B., & Moradkhani, H. (2018). A Bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis. *Water Resources Research*, 54(1), 243–255. <https://doi.org/10.1002/2017WR020403>

Bracken, C., Rajagopalan, B., Cheng, L., Kleiber, W., & Gangopadhyay, S. (2016). Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain. *Water Resources Research*, 52(8), 6643–6655. <https://doi.org/10.1002/2016WR018768>

CAG. (2017). *Report of the comptroller and auditor general of India on schemes for flood control and flood forecasting* (Tech. Rep.). Ministry of Water Resources.

Can, I., Tosunoğlu, F., & Kahya, E. (2012). Daily streamflow modelling using autoregressive moving average and artificial neural networks models: Case study of Çoruh basin. Turkey. *Water and Environment Journal*, 26(4), 567–576. <https://doi.org/10.1111/j.1747-6593.2012.00337.x>

Chang, F. J., & Chen, Y. C. (2001). A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. *Journal of Hydrology*, 245(1–4), 153–164. [https://doi.org/10.1016/S0022-1694\(01\)00350-X](https://doi.org/10.1016/S0022-1694(01)00350-X)

Chang, T. J., Delleur, J. W., & Kavvas, M. L. (1987). Application of Discrete Autoregressive Moving Average models for estimation of daily runoff. *Journal of Hydrology*, 91(1–2), 119–135. [https://doi.org/10.1016/0022-1694\(87\)90132-6](https://doi.org/10.1016/0022-1694(87)90132-6)

Delle Monache, L., Hacker, J. P., Zhou, Y., Deng, X., & Stull, R. B. (2006). Probabilistic aspects of meteorological and ozone regional ensemble forecasts. *Journal of Geophysical Research*, 111(D24), D24307. <https://doi.org/10.1029/2005JD006917>

Dobson, A. J., & Barnett, A. G. (1992). *An introduction to generalized linear models* (Vol. 34). Chapman and Hall. <https://doi.org/10.2307/1269239>

Dunn, P. K., & Smyth, G. K. (2018). *Generalized linear models with examples in R*. Springer New York. <https://doi.org/10.1007/978-1-4419-0118-7>

Dyrddal, A. V., Lenkoski, A., Thorarinsdottir, T. L., & Stordal, F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics*, 26(2), 89–106. <https://doi.org/10.1002/env.2301>

Firat, M. (2008). Comparison of Artificial Intelligence Techniques for river flow forecasting. *Hydrology and Earth System Sciences*, 12, 123–139. <https://doi.org/10.5194/hess-12-123-2008>

Gaume, E., & Gosset, R. (2003). Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology? *Hydrology and Earth System Sciences*, 7(5), 693–706. <https://doi.org/10.5194/hess-7-693-2003>

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472. <https://doi.org/10.1214/ss/1177011136>

Ghorbani, M. A., Zadeh, H. A., Isazadeh, M., & Terzi, O. (2016). A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. *Environmental Earth Sciences*, 75(6), 1–14. <https://doi.org/10.1007/s12665-015-5096-x>

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>

Gneiting, T., Stanberry, L. I., Grimit, E. P., Held, L., & Johnson, N. A. (2008). Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. *Test*, 17(2), 211–235. <https://doi.org/10.1007/s11749-008-0114-x>

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., & Xavier, P. K. (2006). Increasing trend of extreme rain events over India in a warming environment. *Science*, 314(5804), 1442–1445. <https://doi.org/10.1126/science.1132027>

Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering*, 5(2), 115–123. [https://doi.org/10.1061/\(asce\)1084-0699\(2000\)5:2\(115\)](https://doi.org/10.1061/(asce)1084-0699(2000)5:2(115))

Hadi, S. J., & Tombul, M. (2018). Forecasting daily streamflow for basins with different physical characteristics through data-driven methods. *Water Resources Management*, 32(10), 3405–3422. <https://doi.org/10.1007/s11269-018-1998-1>

Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3), 550–560. [https://doi.org/10.1175/1520-0493\(2001\)129<0550:IORHFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2)

Hanel, M., Buishand, T. A., & Ferro, C. A. T. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *Journal of Geophysical Research*, 114(D15), D15107. <https://doi.org/10.1029/2009JD011712>

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5), 559–570. [https://doi.org/10.1175/1520-0434\(2000\)015<0559:DOTCRP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2)

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(47), 1593–1623.

Hunt, K. M. R., & Fletcher, J. K. (2019). The relationship between Indian monsoon rainfall and low-pressure systems. *Climate Dynamics*, 53(3–4), 1859–1871. <https://doi.org/10.1007/s00382-019-04744-x>

- Hunt, K. M. R., Turner, A., & Parker, D. E. (2016). The spatiotemporal structure of precipitation in Indian monsoon depressions. *Quarterly Journal of the Royal Meteorological Society*, 142(701), 3195–3210. <https://doi.org/10.1002/qj.2901>
- Isik, S., Kalin, L., Schoonover, J. E., Srivastava, P., & Graeme Lockaby, B. (2013). Modeling effects of changing land use/cover on daily streamflow: An Artificial Neural Network and curve number based hybrid approach. *Journal of Hydrology*, 485, 103–112. <https://doi.org/10.1016/j.jhydrol.2012.08.032>
- Jang, J., Sun, C., & Mizutani, E. (1997). *Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence*. Prentice-Hall.
- Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs*. Springer New York. <https://doi.org/10.1007/978-0-387-68282-2>
- Karimi, S., Shiri, J., Kisi, O., & Xu, T. (2018). Forecasting daily streamflow values: Assessing heuristic models. *Hydrology Research*, 49(3), 658–669. <https://doi.org/10.2166/nh.2017.111>
- Kisi, Ö. (2004). River flow modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 9(1), 60–63. [https://doi.org/10.1061/\(asce\)1084-0699\(2004\)9:1\(60\)](https://doi.org/10.1061/(asce)1084-0699(2004)9:1(60))
- Kisi, Ö. (2008). Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes*, 22(20), 4142–4152. <https://doi.org/10.1002/hyp.7014>
- Kwon, H.-H., Brown, C., & Lall, U. (2008). Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. *Geophysical Research Letters*, 35(5), L05404. <https://doi.org/10.1029/2007GL032220>
- Li, X., Sha, J., Li, Y. M., & Wang, Z. L. (2018). Comparison of hybrid models for daily streamflow prediction in a forested basin. *Journal of Hydroinformatics*, 20(1), 206–220. <https://doi.org/10.2166/hydro.2017.189>
- Lima, C. H., & Lall, U. (2009). Hierarchical Bayesian modeling of multisite daily rainfall occurrence: Rainy season onset, peak, and end. *Water Resources Research*, 45(7). <https://doi.org/10.1029/2008WR007485>
- Lima, C. H., & Lall, U. (2010). Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. *Journal of Hydrology*, 380(3–4), 438–449. <https://doi.org/10.1016/j.jhydrol.2009.11.016>
- Lima, C. H., Lall, U., Troy, T., & Devineni, N. (2016). A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. *Journal of Hydrology*, 541, 816–823. <https://doi.org/10.1016/j.jhydrol.2016.07.042>
- Londhe, S., & Gavaskar, S. S. (2015). Forecasting one day ahead stream flow using support vector regression. *Aquatic Procedia*, 4, 900–907. <https://doi.org/10.1016/j.aqpro.2015.02.113>
- Luke, A., Vrugt, J. A., AghaKouchak, A., Matthew, R., & Sanders, B. F. (2017). Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States. *Water Resources Research*, 53(7), 5469–5494. <https://doi.org/10.1002/2016WR019676>
- Mendoza, P. A., Rajagopalan, B., Clark, M. P., Ikeda, K., & Rasmussen, R. M. (2015). Statistical postprocessing of high-resolution regional climate model output. *Monthly Weather Review*, 143(5), 1533–1553. <https://doi.org/10.1175/MWR-D-14-00159.1>
- Moradkhani, H., Hsu, K. L., Gupta, H. V., & Sorooshian, S. (2004). Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *Journal of Hydrology*, 295(1–4), 246–262. <https://doi.org/10.1016/j.jhydrol.2004.03.027>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Ossandón, Á., Rajagopalan, B., & Kleiber, W. (2021). Spatial-temporal multivariate semi-Bayesian hierarchical framework for extreme precipitation frequency analysis. *Journal of Hydrology*, 600, 126499. <https://doi.org/10.1016/j.jhydrol.2021.126499>
- Pai, D., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution (0.25 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *MAUSAM*, 65(1), 18.
- Papacharalampous, G. A., & Tyralis, H. (2018). Evaluation of random forests and Prophet for daily streamflow forecasting. *Advances in Geosciences*, 45, 201–208. <https://doi.org/10.5194/adgeo-45-201-2018>
- Papalexioy, S. M., & Montanari, A. (2019). Global and regional increase of precipitation extremes under global warming. *Water Resources Research*, 55(6), 4901–4914. <https://doi.org/10.1029/2018WR024067>
- Pattanaik, D. R., Sahai, A. K., Mandal, R., Phani, M. K., Sahai, A. K., Phani Muralikrishna, R., et al. (2019). *Evolution of operational extended range forecast system of IMD: Prospects of its applications in different sectors* (Vol. 70).
- Pramanik, N., & Panda, R. K. (2009). Application of neural network and adaptive neuro-fuzzy inference systems for river flow prediction. *Hydrological Sciences Journal*, 54(2), 247–260. <https://doi.org/10.1623/hysj.54.2.247>
- Ravindranath, A., Devineni, N., Lall, U., Cook, E. R., Pederson, G., Martin, J., & Woodhouse, C. (2019). Streamflow reconstruction in the upper Missouri River Basin using a novel Bayesian Network Model. *Water Resources Research*, 55(9), 7694–7716. <https://doi.org/10.1029/2019WR024901>
- Razack, M., & Lasm, T. (2006). Geostatistical estimation of the transmissivity in a highly fractured metamorphic and crystalline aquifer (Man-Danane Region, Western Ivory Coast). *Journal of Hydrology*, 325(1–4), 164–178. <https://doi.org/10.1016/j.jhydrol.2005.10.014>
- R Core, T. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Renard, B. (2011). A Bayesian hierarchical approach to regional frequency analysis. *Water Resources Research*, 47(11). <https://doi.org/10.1029/2010WR010089>
- Robert, C., & Casella, G. (2011). A short history of Markov Chain Monte Carlo: Subjective recollections from incomplete data. *Statistical Science*, 26(1), 102–115. <https://doi.org/10.1214/10-STS351>
- Sampaio, J., & Costa, V. (2021). Bayesian regional flood frequency analysis with GEV hierarchical models under a new spatial dependency structures. *Hydrological Sciences Journal*, 66(3), 422–433. <https://doi.org/10.1080/02626667.2021.1873997>
- Scheuerer, M., & Hamill, T. M. (2015). Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. *Monthly Weather Review*, 143(11), 4578–4596. <https://doi.org/10.1175/MWR-D-15-0061.1>
- Shah, R. D., & Mishra, V. (2016). Utility of global ensemble forecast system (GEFS) reforecast for medium-range drought prediction in India. *Journal of Hydrometeorology*, 17(6), 1781–1800. <https://doi.org/10.1175/JHM-D-15-0050.1>
- Shiau, J. T., & Hsu, H. T. (2016). Suitability of ANN-based daily streamflow extension models: A case study of Gaoping River Basin, Taiwan. *Water Resources Management*, 30(4), 1499–1513. <https://doi.org/10.1007/s11269-016-1235-8>
- Sivakumar, B. (2016). *Chaos in hydrology: Bridging determinism and stochasticity*. Springer Netherlands. <https://doi.org/10.1007/978-90-481-2552-4>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, 64(4), 583–639. <https://doi.org/10.1111/1467-9868.00353>
- Sridevi, C., Singh, K. K., Suneetha, P., Durai, V. R., & Kumar, A. (2020). Rainfall forecasting skill of GFS model at T1534 and T574 resolution over India during the monsoon season. *Meteorology and Atmospheric Physics*, 132(1), 35–52. <https://doi.org/10.1007/s00703-019-00672-x>
- Stan Development Team. (2014). *Stan modeling language user's guide and reference manual*.

- Stan Development Team. (2020). *RStan: the R interface to Stan*.
- Sun, X., Thyer, M., Renard, B., & Lang, M. (2014). A general regional frequency analysis framework for quantifying local-scale climate effects: A case study of ENSO effects on Southeast Queensland rainfall. *Journal of Hydrology*, *512*, 53–68. <https://doi.org/10.1016/j.jhydrol.2014.02.025>
- Sun, Y., Niu, J., & Sivakumar, B. (2019). A comparative study of models for short-term streamflow forecasting with emphasis on wavelet-based approach. *Stochastic Environmental Research and Risk Assessment*, *33*(10), 1875–1891. <https://doi.org/10.1007/s00477-019-01734-7>
- Tanoue, M., Hirabayashi, Y., & Ikeuchi, H. (2016). Global-scale river flood vulnerability in the last 50 years. *Scientific Reports*, *6*(1), 1–9. <https://doi.org/10.1038/srep36021>
- Vafakhah, M. (2012). Application of artificial neural networks and adaptive neuro-fuzzy inference system models to short-term streamflow forecasting. *Canadian Journal of Civil Engineering*, *39*(4), 402–414. <https://doi.org/10.1139/l2012-011>
- Verdin, A., Rajagopalan, B., Kleiber, W., Podestá, G., & Bert, F. (2019). BayGEN: A Bayesian space-time stochastic weather generator. *Water Resources Research*, *55*(4), 2900–2915. <https://doi.org/10.1029/2017WR022473>
- Wallemacq, P., & House, R. (2018). *Economic losses, poverty and disasters: 1998-2017* (Tech. Rep.). UNISDR and CRED.
- Wasko, C., & Sharma, A. (2017). Continuous rainfall generation for a warmer climate using observed temperature sensitivities. *Journal of Hydrology*, *544*, 575–590. <https://doi.org/10.1016/j.jhydrol.2016.12.002>
- Wilks, D. S. (2011). *Statistical methods in the atmospheric Sciences, volume 100*. (3rd ed.). Academic Press Inc.
- Yan, H., & Moradkhani, H. (2015). A regional Bayesian hierarchical model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment*, *29*(3), 1019–1036. <https://doi.org/10.1007/s00477-014-0975-3>
- Yuan, X., Wood, E. F., & Ma, Z. (2015). A review on climate-model-based seasonal hydrologic forecasting: Physical understanding and system development. *Wiley Interdisciplinary Reviews: Water*, *2*(5), 523–536. <https://doi.org/10.1002/wat2.1088>
- Zhang, Z., Zhang, Q., & Singh, V. P. (2018). Univariate streamflow forecasting using commonly used data-driven models: Literature review and case study. *Hydrological Sciences Journal*, *63*(7), 1091–1111. <https://doi.org/10.1080/02626667.2018.1469756>
- Zhao, T., Wang, Q. J., Bennett, J. C., Robertson, D. E., Shao, Q., & Zhao, J. (2015). Quantifying predictive uncertainty of streamflow forecasts based on a Bayesian joint probability model. *Journal of Hydrology*, *528*, 329–340. <https://doi.org/10.1016/j.jhydrol.2015.06.043>
- Zounemat-Kermani, M., & Teshnehlab, M. (2008). Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Applied Soft Computing Journal*, *8*(2), 928–936. <https://doi.org/10.1016/j.asoc.2007.07.011>