A Bayesian Hierarchical Network Model for Daily Streamflow Ensemble Forecasting

Alvaro Ossandon[1], Balaji Rajagopalan[1], Upmanu Lall[2], Nanditha J. S.[3], and Vimal Mishra[4]

[1]University of Colorado Boulder
[2]Columbia University
[3]Indian Institute of Technology Gandhinagar
[4]Indian Institute of Technology, Gandhinagar

November 23, 2022

Abstract

A novel Bayesian Hierarchical Network Model (BHNM) for ensemble forecasts of daily streamflow that uses the spatial dependence induced by the river network topology and hydrometeorological variables from the upstream contributing area between station gauges is presented. 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 four gauges in the Narmada basin in central India for the period 1978 – 2014 is presented. The covariates include daily streamflow from upstream gauges or from the gauge above of the upstream gauges depending on travel times and daily, 2-day, or 3-day precipitation from the area between two stations. The model validation indicates that the model is highly skillful relative to climatology and relative to a null-model of linear regression. We applied the BHNM out of sample to two high flooding years. High skill in both the timing and magnitude of the events is demonstrated.

A Bayesian Hierarchical Network Model for Daily Streamflow Ensemble Forecasting

Álvaro Ossandón[1,2], Balaji Rajagopalan[1,3], Upmanu Lall[4], Nanditha J. S.[5], and Vimal Mishra[5]

[1]Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO, USA
[2]Departamento de Obras Civiles, Universidad Técnica Santa María, Valparaíso, Chile
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[4]Department of Earth and Environmental Engineering, Columbia Water Center, The Earth Institute, Columbia University, New York, NY, USA
[5]Civil Engineering, Indian Institute of Technology, Gandhinagar, India

Key Points:

- We developed a Bayesian Hierarchical Network Model (BHNM) 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

Corresponding author: Álvaro Ossandón, alvaro.ossandon@colorado.edu

Abstract

A novel Bayesian Hierarchical Network Model (BHNM) for ensemble forecasts of daily streamflow that uses the spatial dependence induced by the river network topology and hydrometeorological variables from the upstream contributing area between station gauges is presented. 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 four gauges in the Narmada basin in central India for the period 1978 – 2014 is presented. The covariates include daily streamflow from upstream gauges or from the gauge above of the upstream gauges depending on travel times and daily, 2-day, or 3-day precipitation from the area between two stations. The model validation indicates that the model is highly skillful relative to climatology and relative to a null-model of linear regression. We applied the BHNM out of sample to two high flooding years. High skill in both the timing and magnitude of the events is demonstrated.

1 Introduction

Riverine floods are one of the major causes 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 1500 per year (The Data Centre of Central Water Commission), which results in associated damages worth 18 billion INR (CAG, 2017). The extreme rainfall events in the summer monsoon season result from synoptic-scale cyclonic depressions (Hunt et al., 2016; Hunt & Fletcher, 2019). Climate change is projected to enhance the frequency and intensity of extreme precipitation events (Ali & Mishra, 2018; Goswami et al., 2006; Papalexiou & 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, consequently, floods remain less so and vary widely across River Basins.

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, on the other hand, model the relationship between the current day's streamflow with input forcings such as precipitation, antecedent streamflow, and soil moisture, etc., statistically, from historical observations, thereby capturing the hydrologic processes implicitly. Here we consider a statistical model. A brief survey about statistical models 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; Kisi, 2008; Papacharalampous & Tyralis, 2018), autoregressive (AR; Kişi, 2004; Kisi, 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; Kisi, 2008), US (T. J. Chang et al., 1987; Papacharalampous & Tyralis, 2018), and China (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), US 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 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, and 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. Consequently, the main research question is whether can we model streamflow over the entire river network that captures the space-time dependence structure, non-stationarity, and robust estimation of uncertainties?

Motivated by this question, we develop a novel Bayesian Hierarchical Network Model (BHNM) inspired by the framework proposed in Ravindranath et al. (2019), who developed it for paleo-streamflow reconstruction in the Upper Missouri River Basin. We demonstrate this framework by its application to model and predict daily summer monsoon (July-August) streamflow at four 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.

2 Proposed Framework

The proposed Bayesian Hierarchical Network Model (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 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). In this, flow at a downstream gauge, i, at day t is dependent on: flow at one or two most immediate upstream feeder gauges at day $t-k$ with $k > 0$ (depending on travel time); precipitation and other hydrometeorological variables that represent inputs to the streamflow between the streamflow gauges. Thus, in a Bayesian framework, 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:

$$
\begin{aligned}
f\left(Q_t^{(1)}, \cdots, Q_t^{(n)} \middle| Q_{t-k}^{(2)}, \cdots, Q_{t-k}^{(n)}, \mathbf{X}_{t-k}^{(1)}, \cdots, \mathbf{X}_{t-k}^{(n)}\right) &= f\left(Q_t^{(1)} \middle| Q_{t-k}^{(j>1)}, \mathbf{X}_{t-k}^{(1)}\right) \\
&\cdot f\left(Q_t^{(2)} \middle| Q_{t-k}^{(j>2)}, \mathbf{X}_{t-k}^{(2)}\right) \cdot \ldots \cdot f\left(Q_t^{(i)} \middle| Q_{t-k}^{(j>i)}, \mathbf{X}_{t-k}^{(i)}\right) \\
&\cdot \ldots \cdot f\left(Q_t^{(n-1)} \middle| Q_{t-k}^{(n)}, \mathbf{X}_{t-k}^{(n-1)}\right) \cdot f\left(Q_t^{(n)} \middle| \mathbf{X}_{t-k}^{(n)}\right)
\end{aligned} \tag{1}
$$

Where $\mathbf{X}$ denotes the set of hydrometeorological covariates. 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 physical dependencies between streamflow gauges and their feeder gauges and hydrometeorological variables. 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 Gamma probability density function (other distributions as appropriate could be considered) 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:

$$
Q_t^{(i)} \sim Gamma\left(r_t^{(i)}, \lambda_t^{(i)}\right) \tag{2}
$$

where $r_t^{(i)} > 0$ 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 & Daniel, 2011) as follows:

$$
\lambda_t^{(i)} = \frac{\mu_t^{(i)}}{\left(\sigma_t^{(i)}\right)^2}; \qquad r_t^{(i)} = \frac{\left(\mu_t^{(i)}\right)^2}{\left(\sigma_t^{(i)}\right)^2}; \tag{3}
$$

Under the non-stationary 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 $t-k$:

$$\mu_t^{(i)} = \begin{cases} \beta_0^{(i)} + \mathbf{X}_{t-k}^{(i)}\boldsymbol{\beta}_x^{(i)} + \beta_Q^{(i)} Q_{t-k}^{(j>i)} & if\ feeder\ site\ applies \\ \beta_0^{(i)} + \mathbf{X}_{t-k}^{(i)}\boldsymbol{\beta}_x^{(i)} & if\ feeder\ site\ does\ not\ applies \end{cases} \tag{4}$$

$$\sigma_t^{(i)} = \begin{cases} \phi_0^{(i)} + \mathbf{X}_{t-k}^{(i)}\boldsymbol{\phi}_x^{(i)} + \phi_Q^{(i)} Q_{t-k}^{(j>i)} & if\ feeder\ site\ applies \\ \phi_0^{(i)} + \mathbf{X}_{t-k}^{(i)}\boldsymbol{\phi}_x^{(i)} & if\ feeder\ site\ does\ not\ applies \end{cases} \tag{5}$$

where $\beta_0^{(i)}$ and $\phi_0^{(i)}$ are the intercept terms for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\boldsymbol{\beta}_x^{(i)}$ and $\boldsymbol{\phi}_x^{(i)}$ are $m\times$ 1 vector of regression coefficients related to hydrometeorological variables for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\beta_Q^{(i)}$ and $\phi_Q^{(i)}$ are regression coefficients related to the feeder site for $\mu_t^{(i)}$ and $\sigma_t^{(i)}$; $\mathbf{X}_{t-k}^{(i)}$ is a $1 \times m$ vector of hydrometeorological variables on the day $t-k$; and $Q_{t-k}^{(j>i)}$ corresponds to the flow at the feeder site at the day $t-k$. All of the model covariates change with time to help capture nonstionarity.

2.2 Likelihood and Priors

The posterior distributions of the regression coefficients, $\boldsymbol{\theta} = [\boldsymbol{\beta}, \boldsymbol{\phi}]$, 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

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

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

$$p\left(\boldsymbol{\theta}^{(i)} \middle| Q_{t-k}^{(j>i)}, \mathbf{X}_{t-k}^{(i)}\right) = MVN\left(\ln\left(\boldsymbol{\beta}^{(i)}\right) \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right) \cdot MVN\left(\ln\left(\boldsymbol{\phi}^{(i)}\right) \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right) \cdot p\left(\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right) \cdot p\left(\boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right) \tag{7}$$

where $MVN\left(\ln\left(\boldsymbol{\beta}^{(i)}\right) \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right)$ and $MVN\left(\ln\left(\boldsymbol{\phi}^{(i)}\right) \middle| \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right)$ represent probability density of multivariate normal distributions with mean $\mathbf{0}$ and covariance matrix

$\mathbf{\Sigma}^{(i)}$ corresponding to the priors of the log of $\boldsymbol{\beta}^{(i)} = \left[\beta_0^{(i)}, \boldsymbol{\beta}_x^{(i)}, \beta_Q^{(i)}\right]$ and $\boldsymbol{\phi}^{(i)} = \left[\phi_0^{(i)}, \boldsymbol{\phi}_x^{(i)}, \phi_Q^{(i)}\right]$ at the gauge i, respectively; and $p\left(\mathbf{\Sigma}_{\boldsymbol{\beta}}^{(i)}\right)$ and $p\left(\mathbf{\Sigma}_{\boldsymbol{\phi}}^{(i)}\right)$ 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

$$\mathbf{\Sigma}_{\boldsymbol{\beta}}^{(i)}\ Inv\ wishart\,(\nu, A\mathbf{I})\,; \qquad \mathbf{\Sigma}_{\boldsymbol{\phi}}^{(i)}\ Inv\ wishart\,(\nu, B\mathbf{I})\,; \tag{8}$$

where ν corresponds to the degrees of freedom (m+1), $\mathbf{I}$ is an $(m+2)\times(m+2)$ identity matrix, and A and B are scalars properly set for $\mathbf{\Sigma}_{\boldsymbol{\beta}}^{(i)}$ and $\mathbf{\Sigma}_{\boldsymbol{\phi}}^{(i)}$, respectively. In equation 7 $\ln\left(\boldsymbol{\beta}^{(i)}\right)$ and $\ln\left(\boldsymbol{\phi}^{(i)}\right)$ were considered 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, selection of covariates, model structure for the Narmada basin, and the cross-validation procedures are described below.

3.1 The study Basin

The Narmada River basin (Figure 2), with 98,796 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 1120 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) for the period 1978 – 2014. Garudeshwar gauge station was not considered in this study since it had longer missing periods (summers of 1988, 1989, 1995, and 2004).

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 6995 meteorological stations across India (Pai et al., 2014) and is available at 0.25° spatial resolution from 1951-2018. 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 Covariates

We considered antecedent daily streamflow from upstream (feeder) gauges and spatial average precipitation from the area between the station gauges. The covariates are considered until the previous day (lag-1 day), i.e., 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. Due to the presence of dams for Mandleshwar and Sandiya, the spatial average precipitation was obtained from the area between the station gauge and the upstream dam. We obtained the best set of covariates for each station gauge based on the highest Pearson correlation coefficient. Besides, we checked their significance based on the posterior distribution of the model coefficients that they do not cross zero.

Figure 3 shows the best set of covariates for daily streamflow at each gauge. The best covariates we obtained based on the highest correlation for each gauge point. For Mandleshwar (Figure 3a, b), the best covariates were daily spatial average precipitation of $t-1$ day and streamflow at the upstream gauge of Hoshangabad of day $t-1$. For Handia (Figures 3c, d), two-day accumulated spatial average precipitation of days $t-$

1 and $t-2$ and day $t-1$ day streamflow at Hoshangabad were the best. For Hoshangabad gauge (Figures 3e, f), two-day spatial average precipitation of days $t-1$ and $t-2$ and day $t-1$ streamflow at the upstream gauge, Sandiya were the best covariates. For the headwater, Sandiya gauge (Figures 3g, h), three-day spatial average precipitation above the gauge of days $t-1$, $t-2$, and $t-3$, and the streamflow at day $t-1$ from the same gauge emerged as the best covariates. Since Sandiya does not have a gauge above it, we chose the flow at this gauge from day $t-1$ as a covariate. All the correlation coefficients (R) between daily streamflow covariates are ~ 0.8 and higher, while the correlations with precipitation are ~ 0.6. However, it is interesting to note that for Madleshwar gauge, the correlation of daily streamflow with precipitation is lowest ($\sim$ 0.5). This, we surmise, is due to a dam downstream near Handia gauge, which can control the flow at Mandleshwar gauge that is unrelated to the precipitation. Thus, we chose the streamflow at Hoshangabad gauge as a covariate for Mandleshwar. Also, the lagged time of the covariates is consistent with the travel time of the reaches. However, naturalized streamflow will likely have a stronger correlation between streamflow and precipitation.

3.4 Model Structure for the Narmada River Basin

The BHNM for the Narmada basin follows the generalized framework described in section 2.1. The schematic of the model for the basin is shown in Figure 4. The covariates identified and described in the previous section are incorporated in the model represented in the model equations below:

$$Q_t^{(i)} \sim Gamma\left(r_t^{(i)}, \lambda_t^{(i)}\right) \qquad i = 1,\ 2,\ 3,\ 4 \tag{9}$$

$$\lambda_t^{(i)} = \frac{\mu_t^{(i)}}{\left(\sigma_t^{(i)}\right)^2}; \qquad r_t^{(i)} = \frac{\left(\mu_t^{(i)}\right)^2}{\left(\sigma_t^{(i)}\right)^2}; \qquad i = 1,\ 2,\ 3,\ 4 \tag{10}$$

Mandleshwar:

$$\mu_t^{(1)} = \beta_0^{(1)} + \beta_1^{(1)} P_{1d,t-1}^{(1)} + \beta_Q^{(1)} Q_{t-1}^{(3)}$$

$$\sigma_t^{(1)} = \phi_0^{(1)} + \phi_1^{(1)} P_{1d,t-1}^{(1)} + \phi_Q^{(1)} Q_{t-1}^{(3)} \tag{11}$$

Handia:

$$\mu_t^{(2)} = \beta_0^{(2)} + \beta_1^{(2)} P_{2d,t-1}^{(2)} + \beta_Q^{(2)} Q_{t-1}^{(3)}$$

$$\sigma_t^{(1)} = \phi_0^{(1)} + \phi_1^{(2)} P_{2d,t-1}^{(2)} + \phi_Q^{(2)} Q_{t-1}^{(3)} \tag{12}$$

Hoshangabad:

$$\mu_t^{(3)} = \beta_0^{(3)} + \beta_1^{(3)} P_{2d,t-1}^{(3)} + \beta_Q^{(3)} Q_{t-1}^{(4)}$$

$$\sigma_t^{(3)} = \phi_0^{(3)} + \phi_1^{(3)} P_{2d,t-1}^{(3)} + \phi_Q^{(3)} Q_{t-1}^{(4)} \tag{13}$$

Sandiya:

$$\mu_t^{(4)} = \beta_0^{(4)} + \beta_1^{(4)} P_{3d,t-1}^{(4)} + \beta_2^{(4)} Q_{t-1}^{(4)}$$

$$\sigma_t^{(4)} = \phi_0^{(4)} + \phi_1^{(4)} P_{3d,t-1}^{(4)} + \phi_2^{(4)} Q_{t-1}^{(4)} \tag{14}$$

where $P_{xd,t-1}^{(i)}$ is the x-day spatial average precipitation for the gauge station i. The priors of $\boldsymbol{\beta}^{(i)}$ and $\boldsymbol{\phi}^{(i)}$ are for each streamflow gauge multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(i)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\phi}}^{(i)}$, respectively (equation 7). For the priors of the covariance matrix according to equation 8, for each station gauge, we consider weakly informative priors with $\nu = 4$, $A = 10$, and $B = 10$.

The model was implemented in R (R Core, 2017) using the program JAGS (Just Another Gibbs Sampler; Plummer, 2003) and the R package rjags (Plummer, 2019), which provides an interface from R to the JAGS 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 Gibbs sampling algorithm 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 100,000 iterations with a burn-in size value of 50,000 to ensure convergence. To reduce the sample dependence (autocorrelation), we chose a thinning factor of 50. 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.

3.5 Model Cross-Validation

Since this study's goal is to provide a daily streamflow forecast for risk-based flood mitigation, we chose to implement a leave-two-years-out cross-validation for years with high flow values. In this, two consecutive years from the common record (1978–2014), in which a high flow occurred, are chosen as validation years, and the BHNM built using the remaining observations, also known as the calibration years. The fitted model is applied to provide estimates for the two validation years. This cross-validation procedure was repeated four times, and the two consecutive years periods considered with high flows were: 1984-1985; 1990-1991; 1996-1997; 2013-2014. Figure 5 shows the time series of July-August daily streamflow for 1970-2014 at the Mandleshwar gauge station and black boxes that denote the four validation periods considered.

We included a comparison of the cross-validation results of the BHNM with those of a standard Multi-Linear Model (MLM). By "standard", we mean that a simple multi-linear regression model with the same covariates presented in section 3.3 fitted to daily streamflow at each gauge station via the Maximum Likelihood (ML) method. The parameters follow a multivariate normal distribution with mean and covariance matrix equal to the estimates and covariance matrix obtained from the ML (Bracken et al., 2018), thus, providing parameter ensembles and consequently, flow ensembles generated from the linear regression model for each parameter sample. Note that, unlike the BHNM, the multi-linear model is fitted at each gauge separately and thus does not contain correlation across the station gauges. Also, the uncertainty estimates from ML tend to be lower.

In this study, 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:

$$DI = \sum_{i=1}^{M+1} \left| \frac{count_i}{N} - \frac{1}{M+1} \right| 100 \tag{15}$$

where M is the number of ensemble 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 ith bin, and N is the sample size. Lower DI means that the ensemble better achieves the condition of reliability.

The continuous rank probability score (CRPS) evaluates the accuracy of the empirical/probabilistic forecasts 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

$$CRPS = \int_{-\infty}^{\infty} \left[F(Q) - H(Q - Q_o)\right]^2 dQ \tag{16}$$

$$H(Q - Q_0) = \begin{cases} 0 & Q < Q_o \\ 1 & Q \geq Q_o \end{cases} \tag{17}$$

where $F(Q)$ is the CDF associated with the forecast, Q_o is the observed streamflow, and $H(Q - Q_0)$ is the well-known Heaviside function. The continuous ranked probability skill score (CRPSS) is then defined accordingly:

$$CRPSS = 1 - \frac{CRPS_{forecast}}{CRPS_{reference}} \tag{18}$$

where $CRPS_{forecast}$ is the CRPS of the forecast model, $CRPS_{reference}$ is the 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.

The energy score (ES) assesses probabilistic forecasts of a multivariate quantity (Gneiting & Raftery, 2007; Gneiting et al., 2008):

$$ES = \frac{1}{M}\sum_{j=1}^{M}||\mathbf{Q}_j - \mathbf{Q}_o|| - \frac{1}{2M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}||\mathbf{Q}_i - \mathbf{Q}_j|| \tag{19}$$

where M is the size of the ensemble forecast, $\mathbf{Q}_j$ is the $n{\times}1$ vector of the jth 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 16), to which the energy score reduces in dimension $d = 1$. Then, the energy skill score (ESS) is defined as

$$ESS = 1 - \frac{ES_{forecast}}{ES_{reference}} \tag{20}$$

where $ES_{forecast}$ is the ES of the forecast model, $ES_{reference}$ is the ES of the reference forecast. As for the CRPSS, the ESS ranges from $-\infty$ to 1, and its values have the same meaning.

In this study, we considered climatology as the reference model.

4 Results

We present results from model calibration followed by results of leave two-year cross-validation and the forecast verification metrics.

4.1 Calibration

We calibrated the BHNM for the entire record (1978-2014). Figure 6 shows the predictive posterior distribution ensembles of July-August daily streamflow for Mandleshwar (the terminal gauge) for the whole time record (1978-2014) and the last two-year period 2013-2014 for a closer visualization of the timing of the high flows. The flow ensembles are generated from the Gamma distribution using the posterior samples of the model parameters. The flow ensembles are presented as time series of boxplots. The simulated median daily flows are generally lower than the observed flows; however, almost all the observed high flows are captured within the ensemble variability with few exceptions (e.g., high value in 1996, Figure 6a). The daily streamflow timing is captured very well by the posterior ensembles, as can be seen in Figure 6b for the two-year period 2013-2014. The scatter plots of daily observed streamflow vs. ensemble median of the posterior distribution and the related correlation coefficients (R) and relative bias for peak flows (computed only for dates where the observed flow exceeds the 90th quantile) are shown in the upper right corner. The R values were 0.83 and 0.92 for 1978-2014 and 2013-2014, respectively, while relative bias for peak flows were -0.19 and -0.21 of the observed mean high peak flows for 1978-2014 and 2013-2014, respectively.

The posterior flow ensembles for the Sandiya gauge (the headwater gauge) are shown in Figure 7 (same as Figure 6). In this case, there is a clear underestimation of the high streamflow values by the ensemble median, and the ensembles variability cannot fully capture most of them (Figure 7a). Also, there is a delay of one day in the peak streamflows' simulated timing (Figure 7b). The timing of the streamflows' peak is closely related to the travel time from the upstream gauge. Sandiya being the uppermost gauge in the basin, does not have this information available; thus, the model's simulated timing is off by a day. Additional covariates need to be explored for this gauge (uppermost gauge in general) from hydrometeorological variables that can capture the surface processes. The overall performance of the ensembles in terms of Perason correation, R values was good (0.83 and 0.9 for 1978-2014 and 2013-2014, respectively). As in Mandleshwar, for Sandiya, a negative relative bias was obtained (-0.32 and -0.29 of the observed mean high peak flows for 1978-2014 and 2013-2014, respectively).

4.2 Cross-Validation

The leave two-year out cross-validation, following the procedure described in section 3.5, was performed. The daily posterior ensemble flow forecasts for the four validation periods: 1984-1985, 1990-1991, 1996-1997, and 2013-2014 were obtained and shown in Figure 8. At each gauge, this consists of streamflow simulations for 496 days covering the four periods, and as in calibration, each day has 3,000 posterior predictive ensembles. Figure 8 shows the predictive posterior ensemble forecast of July-August daily streamflow presented as boxplot time series for the four validation periods at Mandleshwar and Sandiya gauge stations, along with the scatterplot of predictive median and observed flows – similar to Figure 6. As in Figure 6, in Figure 8a, it can be seen that for Mandleshwar gauge, the ensemble median slightly underestimates the observed for high streamflow values, but the ensembles variability can capture most of them except for a few events (July 30 of 1991 and July 29 of 1996). For Sandiya, as in the calibration, there is an evident underestimation of the high streamflow values by the ensemble median, and the ensemble variability cannot fully capture most of them (Figure 8b). While R values of the posterior and observed flows showed excellent performance (0.85 and 0.84 for Mandleshwar and Sandiya, respectively), negative relative bias values were obtained (-0.23 and -0.35 of the observed mean high peak flows for Mandleshwar and Sandiya, respectively). As in the calibration, better performance for intermediate station gauges was achieved (Figure S3 in the supplemental information).

4.3 Cross-Validation Skill Metrics

To assess the reliability of the ensembles forecast, rank histograms of the ensemble forecast of July-August daily streamflow during cross-validation periods for the BHNM and MLM models at Mandleshwar gauge station are presented in Figure 9. The shape of rank histograms and DI values demonstrate that a better spread is generated from the ensembles forecast of the BHNM since its rank histogram is almost uniform (Figure 9a) with a low DI value. For MLM, the U-shaped of the rank histogram and high DI value indicate a lack of variability in the ensemble. We obtained similar results for the other gauges (Figures S4-S6 in the supplementary information).

To assess the at site (marginal) probabilistic skill of the proposed model, we computed the CRPSS for two subsets: 496 days of the four validation periods, and for the days with high flows, which for each gauge are defined as the days when the observed streamflow exceeds their 75th percentile streamflow. The CRPSS calculated for each forecast day are shown as boxplots in Figure 10 for forecasts from BHNM (sky blue boxes) and MLM (gray boxes) for the two subsets. For the four validation periods combined (Figure 10a), most of the values remain above zero (over 75%) for both models. However, the variability of CRPSS from BHNM is lower than that of MLM. The median CRPSS from BHNM is higher than that of MLM except for Sandiya (median of the distribution is lower for BHNM, Figure 10a). BHNM exhibits better overall performance for high flow days than MLM in reduced variability and higher median values for all the gauges (Figure 10b). These findings are important for skillful forecasts of high flows that cause flooding in the basin since this forecast could be useful for flood early warning and mitigation. These results indicate a consistent and robust performance of the posterior ensembles from BHNM.

To assess the skill of the model ensembles in their ability to capture the joint dependence across gauges, we computed the ESS for 496 days of four validation periods. Figure 11 shows ESS distributions of BHNM (sky blue boxes) and MLM (gray boxes) for the four validation periods. The ESS for the BHNM remains entirely above zero compared to that from MLM, and the median ESS from BHNM is higher. These results indicate a higher skill of the BHNM ensembles to predict the joint flow distribution across gauges, which is especially crucial for flood mitigation across the river network and not just at single locations.

5 Summary and Discussion

We formulated and presented a Bayesian Hierarchical Network Model (BHNM) for daily streamflow. The model uses the spatial dependence induced by the river network topology and hydrometeorological variables from the upstream contributing area between the covariates' gauges. For the application presented, daily streamflow at each station is assumed to be distributed as a Gamma probability density function with spatial and temporal non-stationary in the distribution parameters. The distribution parameters for each day and at each gauge are modeled as linear functions of selected covariates. With suitable priors, the posterior distribution of the model parameters and consequently, the predictive posterior distribution ensembles of daily streamflow are obtained.

We applied this to forecast daily summer (July-August) streamflow at four gauges in the Narmada River Basin network in west-central India for the period 1978 – 2014, at one day lead time. The covariates included streamflow from upstream feeder gauges and spatial average precipitation from the area between stations, from previous 1, 2, or 3 days, that attempts to reflect the antecedent land conditions. The probabilistic skill and reliability of the ensemble forecasts individually and jointly were assessed by rank histograms and skill scores such as continuous cumulative rank probability skill score (CRPSS) and energy skill score (ESS). The model ensembles capture the daily streamflow and magnitudes quite well along with the flow peaks' timing, except for the headwater gauge. We computed the skill metrics in leave-two-year cross-validation for 496 days of the four validation periods and only days with high flows. We found the posterior ensemble forecast from BHNM to be highly skillful, consistent, and reliable compared to the traditional multi-linear model. The skill is higher and robust, especially for high flow days, raising the prospect that this BHNM can be used in real-time coordinated flood mitigation and early warning across the river basin.

The proposed Bayesian Hierarchical Network Model (BHNM) has benefits compared to stationary, at site Bayesian, and non-Bayesian models:

- Using the network structure in incorporating flow information from upstream gauges and precipitation from upstream contributing areas as covariates, communicates information through the network and captures the spatial correlation of flows simultaneously
- Compared to an at-site multi-linear model (MLM), BHNM shows better performance by capturing the river network's spatial dependence and the uncertainty at each station gauge.

The headwater gauges need special attention, for they do not have station gauges upstream to provide information about the basin hydrology. The modeling framework is general in that it can be adapted to model other space-time variables, admit other potential distributions such as Lognormal, Weibull, Generalized Extreme Value, etc. Further, this model can be combined with precipitation and basin hydrologic forecasts as covariates. Preliminary results combining hydrologic forecasts from the Variable Infiltration Capacity (VIC) model for the Narmada River basin with the BHNM showed good forecast skill with promising avenues for combining statistical and physical model forecasts.

This framework can be applied to basins with non-natural flow regimes by incorporating the right feeder gauge. Another alternative would be replacing the spatial average precipitation with another potential predictor more skillful such as the reservoir levels through its operation rule if known. These considerations allow that the model to replicate the effect of some human interventions such as dams.

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. Precipitation

from the India Meteorology Department (IMD) can be obtained from `https://imdpune.gov.in/Clim_Pred_LRF_New/lrf_Index.html`. Observed streamflow can be obtained from India Water Resources Information System (India-WRIS): `https://indiawris.gov.in/wris/#/`.

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. doi: 10.1007/s11269-017-1782-7

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. doi: 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*. doi: 10.1016/j.wace.2019.100212

Banerjee, R. (2009). Review of water governance in the Narmada river basin. , 40.

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. doi: 10.1002/2017WR020403

CAG. (2017). *Report of the Comptroller and Auditor General of India on Schemes for Flood Control and Flood Forecasting* (Tech. Rep.). New Delhi. Retrieved from `https://cag.gov.in/ag/manipur/en/audit-report/details/31347`

Can, , 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. doi: 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. doi: 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. doi: 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. doi: 10.1029/2005JD006917

Firat, M. (2008). *Comparison of Artificial Intelligence Techniques for river flow forecasting* (Vol. 12; Tech. Rep.). Retrieved from `www.hydrol-earth-syst-sci.net/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. doi: 10.5194/hess-7-693-2003

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511790942

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. doi: 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. doi: 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. doi: 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. doi: 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.

Govindaraju, R. S. (2000). Artificial Neural Networks in Hydrology. I: Preliminary Concepts. *Journal of Hydrologic Engineering*, *5*(2), 115–123. doi: 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. doi: 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. doi: 10.1175/1520-0493(2001)129⟨0550:IORHFV⟩2.0.CO;2

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

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. doi: 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. doi: 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. doi: 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.* Upper Saddle River: Prentice-Hall.

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian Networks and Decision Graphs.* New York, NY: Springer New York. doi: 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. doi: 10.2166/nh.2017.111

Kisi, Ö. (2008). Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes*, *22*(20), 4142–4152. doi: 10.1002/hyp.7014

Kişi, Ö. (2004). River Flow Modeling Using Artificial Neural Networks. *Journal of Hydrologic Engineering*, *9*(1), 60–63. doi: 10.1061/(asce)1084-0699(2004)9:1(60)

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. doi: 10.2166/hydro.2017.189

Londhe, S., & Gavraskar, S. S. (2015). Forecasting One Day Ahead Stream Flow Using Support Vector Regression. *Aquatic Procedia*, *4*, 900–907. doi: 10.1016/j.aqpro.2015.02.113

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. doi: 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. doi: 10.1016/j.jhydrol.2004.03.027

Pai, D., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution ($0.25° \times 0.25°$) 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. doi: 10.5194/adgeo-45-201-2018

Papalexiou, S. M., & Montanari, A. (2019). Global and Regional Increase of Precipitation Extremes Under Global Warming. *Water Resources Research*, *55*(6), 4901–4914. doi: 10.1029/2018WR024067

Pattanaik, D. R., Sahai, A. K., Mandal, R., Phani, M. K., Sahai, A. K., Phani Muralikrishna, R., … Mishra, V. (2019). *Evolution of operational extended range forecast system of IMD : Prospects of its applications in different sectors* (Vol. 70) (No. 2).

Plummer, M. (2003). Proceedings of the 3rd international workshop on distributed statistical computing. *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*, *124*(125.10), 1–10.

Plummer, M. (2019). rjags: Bayesian graphical models using MCMC. *R package version*, *4*(10), 19.

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. doi: 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. doi: 10.1029/2019WR024901

R Core, T. (2017). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Robert, C., & Casella, G. (2011). A short history of Markov Chain Monte Carlo: Subjective recollections from incomplete data. *Statistical Science*, *26*(1), 102–115. doi: 10.1214/10-STS351

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. doi: 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. doi: 10.1007/s11269-016-1235-8

Sivakumar, B. (2016). *Chaos in hydrology: Bridging determinism and stochasticity.* Springer Netherlands. doi: 10.1007/978-90-481-2552-4

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. doi: 10.1007/s00703-019-00672-x

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. doi: 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. doi: 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. doi: 10.1139/l2012-011

Wallemacq, P., & House, R. (2018). *UNISDR and CRED report: Economic losses,* *poverty & disasters: 1998-2017* (Tech. Rep.). Retrieved from https://www.undrr.org/publication/economic-losses-poverty-disasters-1998-2017

Wasko, C., & Sharma, A. (2017). Continuous rainfall generation for a warmer climate using observed temperature sensitivities. *Journal of Hydrology*, *544*, 575–590. doi: 10.1016/j.jhydrol.2016.12.002

Wilks, & Daniel. (2011). *Statistical Methods in the Atmospheric Sciences, Volume 100 - 3rd Edition.* Academic Press Inc.

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. doi: 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. doi: 10.1080/02626667.2018.1469756

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. doi: 10.1016/j.asoc.2007.07.011

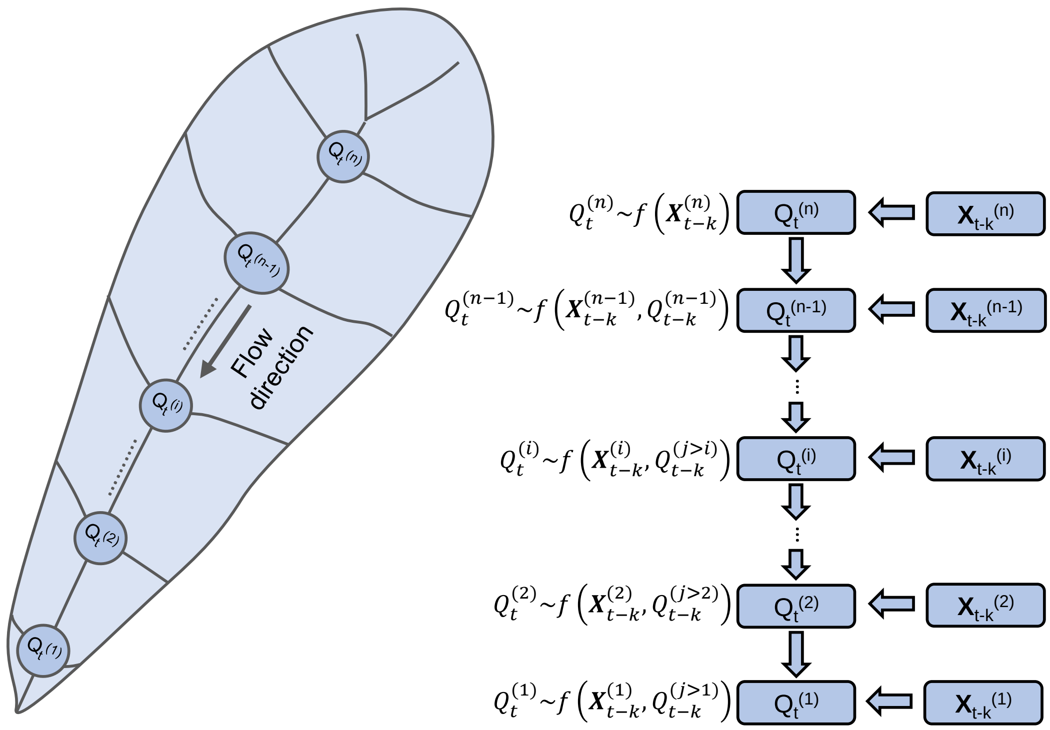

Figure 1. Conceptual sketch of Bayesian Hierarchical Network Model. n streamflow gauges and n hydrometeorological covariates vectors are shown in the graph for illustrating the concept of the graphical network model. Physically informed modeling structure using regional hydrometeorological covariates 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 all gauges (nodes) in the network simultaneously using a Bayesian estimation scheme.

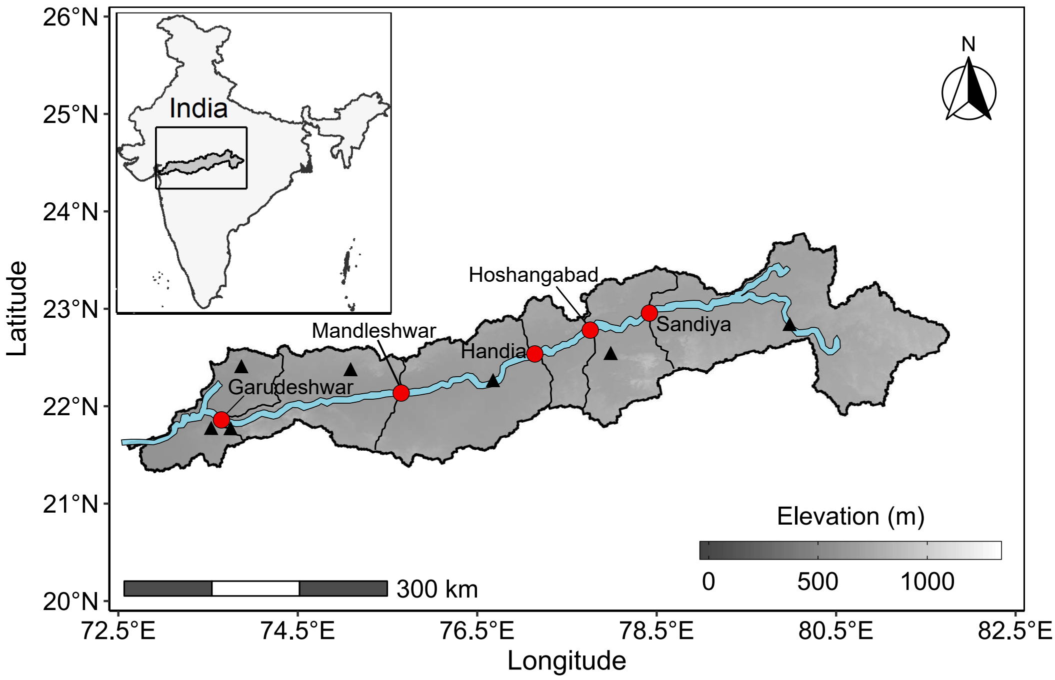

Figure 2. Map of the Narmada basin boundary in India showing the digital elevation model of the basin (SRTM DEM); the locations of five sub-basin outlets: Sandiya, Hoshangabad, Handia, Mandleshwar, and Garudeshwar; and some of the major dams in the basin are marked: Bargi, Tawa, Indirasagar, Jobat, and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

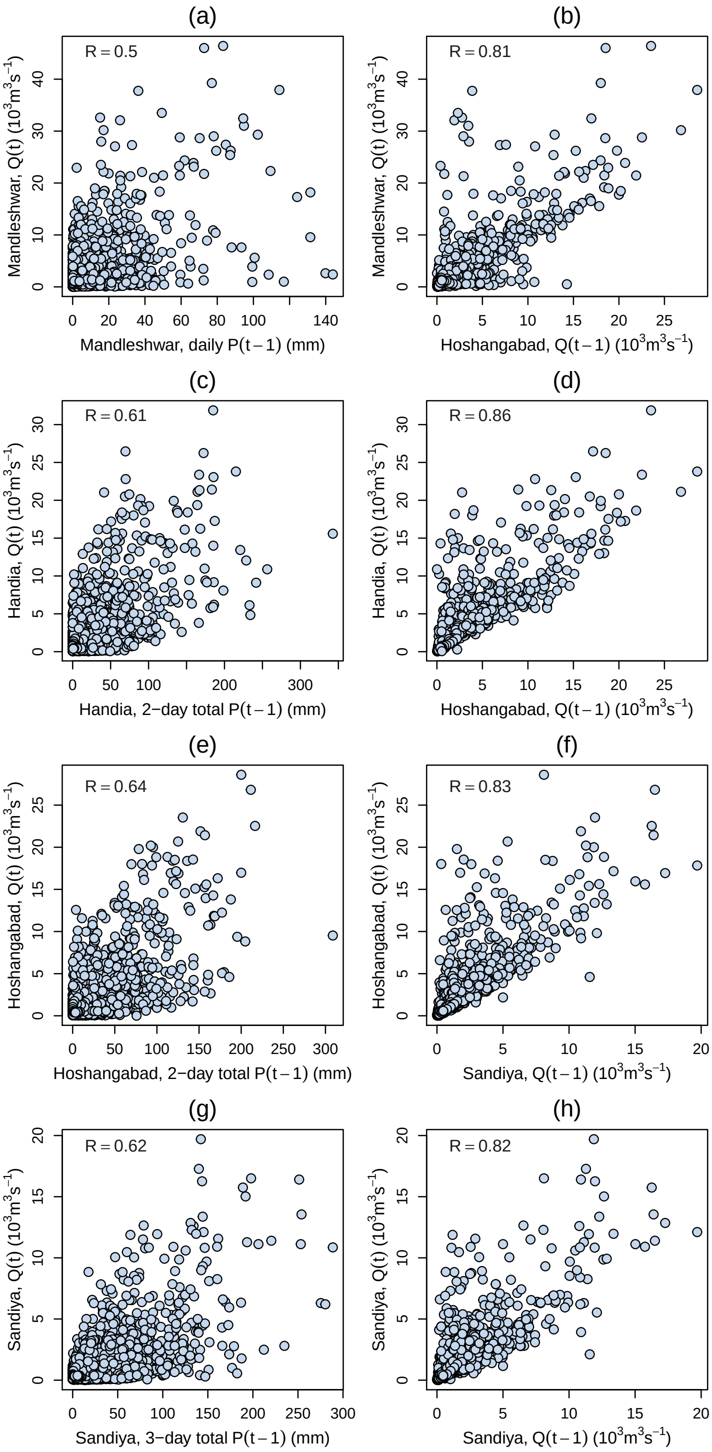

Figure 3. Scatter plots of daily streamflow on day t vs. best lag -1 day covariates selected for each station gauge: Mandleshwar streamflow vs. (a) daily spatial average precipitation, (b) and daily Hoshangabad streamflow; Handia streamflow vs. (c) 2-day spatial average precipitation, (d) and daily Hoshangabad streamflow; Hoshangabad streamflow vs. (e) 2-day spatial average precipitation, (f) and daily Sandiya streamflow; Sandiya streamflow vs. (g) 3-day spatial average precipitation, (h) and lag -1 day daily Sandiya streamflow. All Pearson correlation coefficients, R, are significant ($P - value < 0.1$).

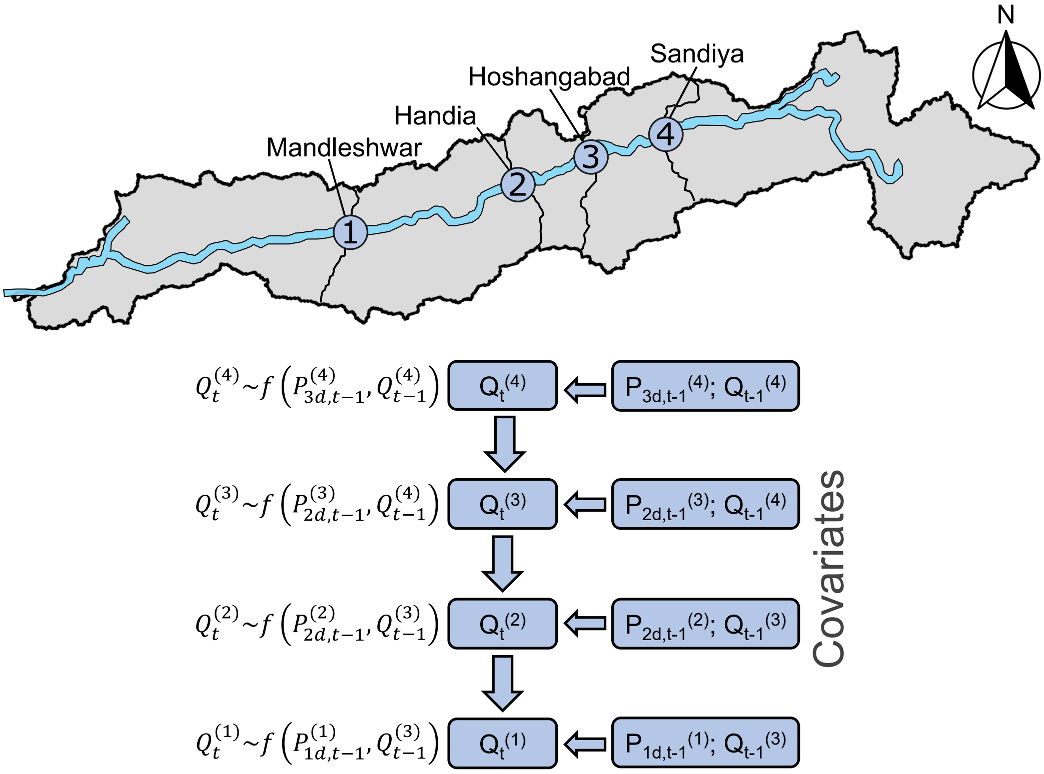

Figure 4. Schematic of the BHNM for the Narmada River basin.

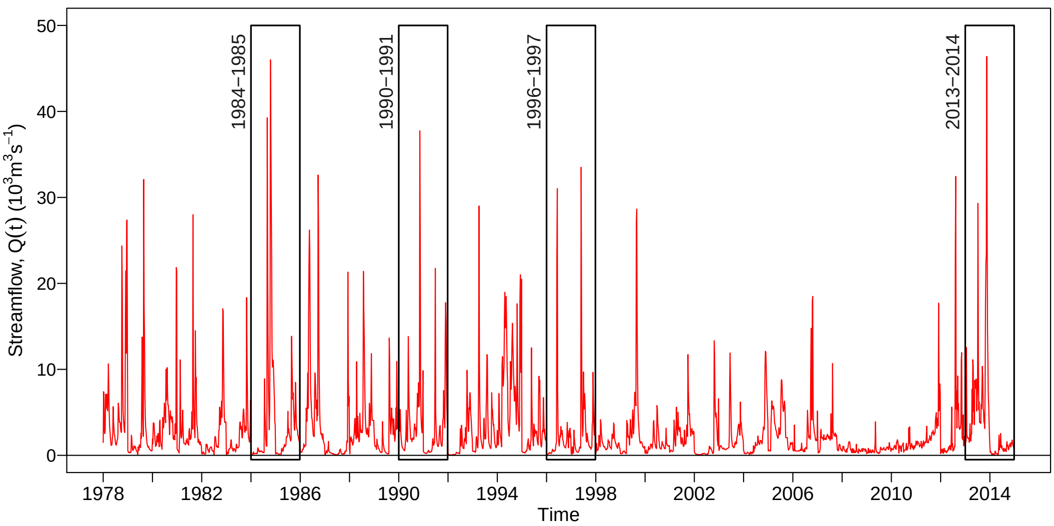

Figure 5. Time series of July-August daily streamflow for 1978-2014 at the Mandleshwar gauge station. Black boxes denote the four validation periods considered for the Cross-Validation.

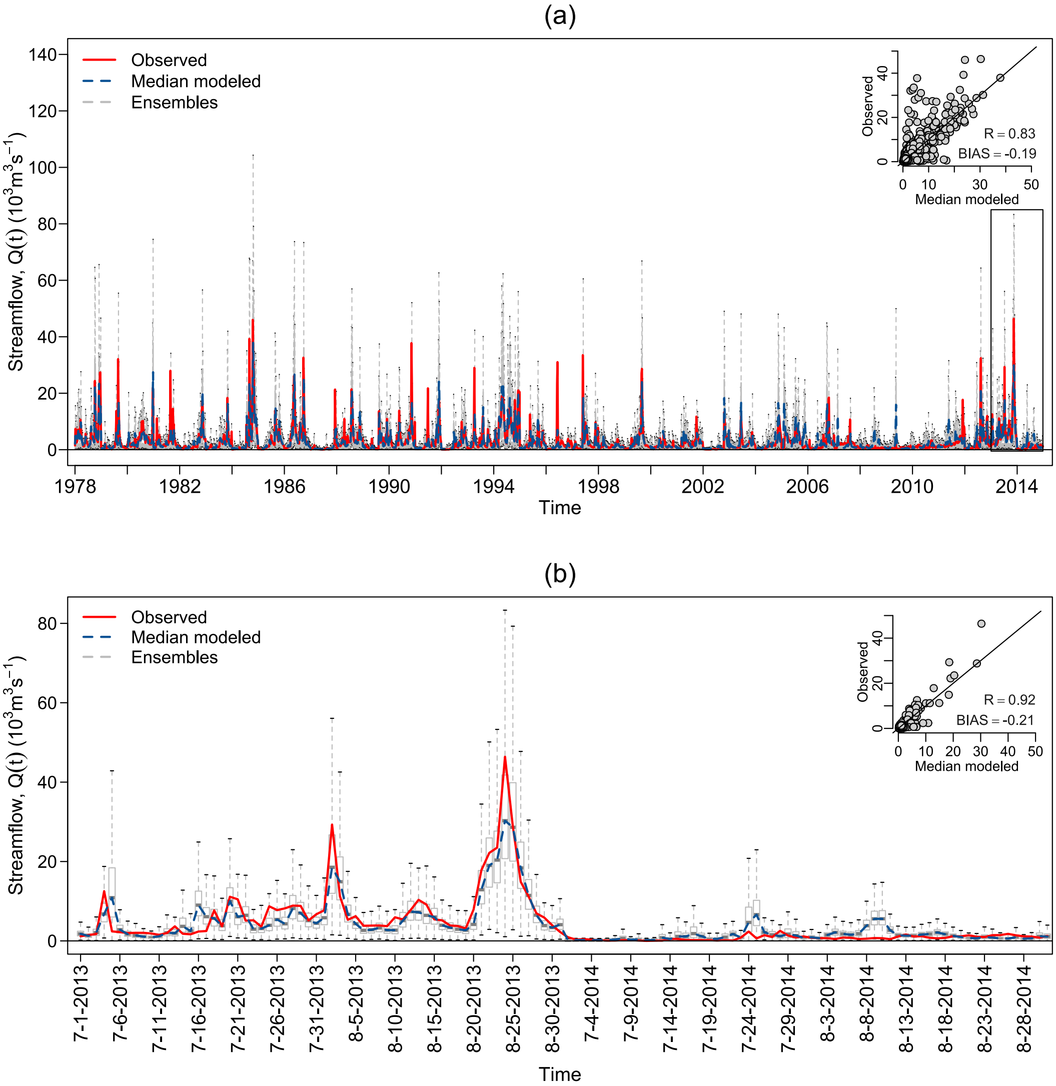

Figure 6. Predictive posterior distribution ensembles of simulated July-August daily streamflow for the Mandleshwar gauge station presented as boxplot time series for (a) entire record (1978-2014) and (b) 2013-2014. The boxplots represent the posterior distribution estimates of the daily streamflow. Outliers are not displayed. Red lines correspond to the observed daily streamflow and blue-dashed lines to the posterior median daily streamflow. Scatterplots of the posterior median and observed flows along with the 1:1 line and relative bias for peak flows (computed only for dates where the observed flow exceeds the 90th quantile) and R values are on the upper right of each panel. R values are significant ($P - value < 0.1$). The black box in panel a shows the temporal windows for time series in panel b.

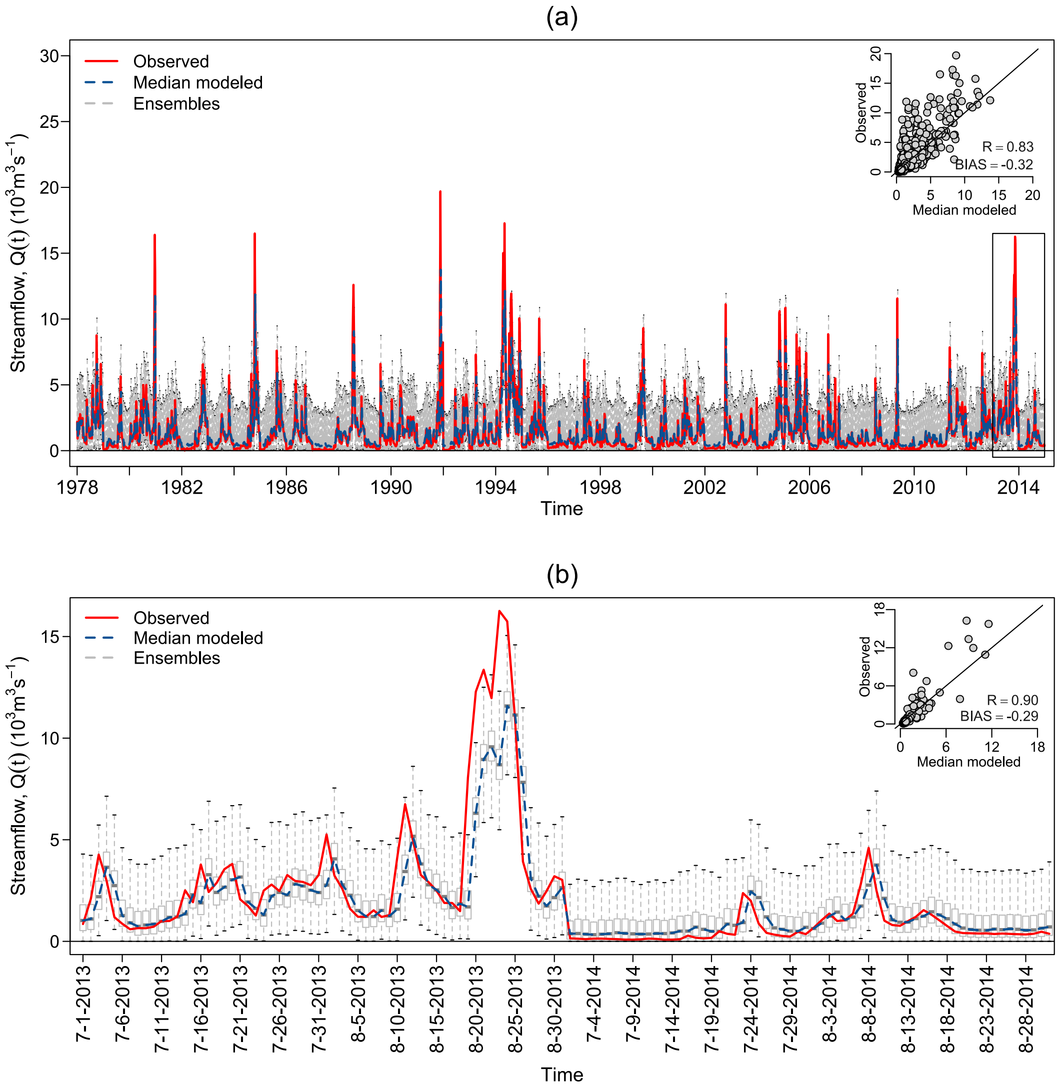

Figure 7. As in Figure 6, but for the Sandiya gauge station. R values are significant ($P - value < 0.1$).

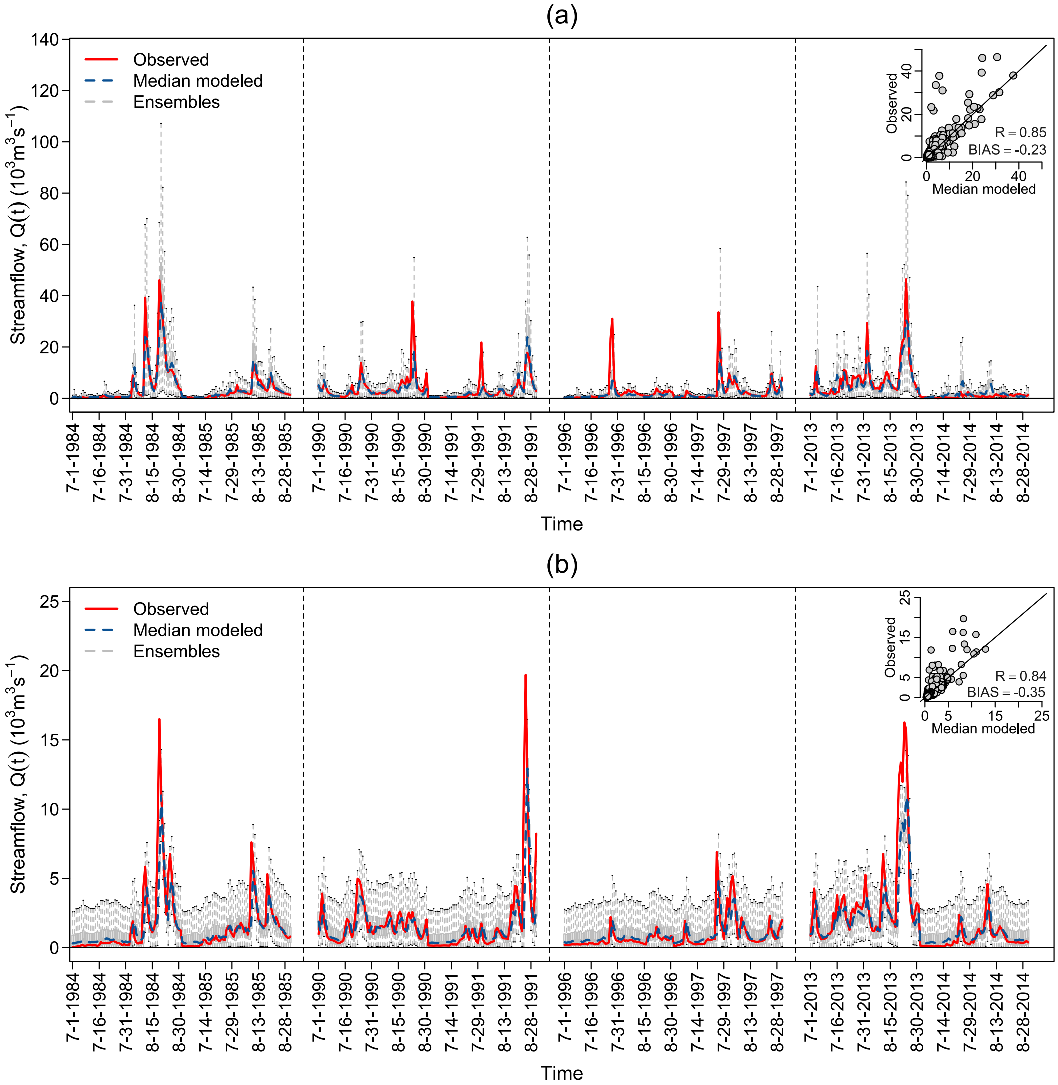

Figure 8. Same as Figure 6 but predictive posterior ensemble forecast of July-August daily streamflow presented as boxplot time series for the four validation periods (1984-1985, 1990-1991, 1996-1997, and 2013-2014) at (a) Mandleshwar and (b) Sandiya gauges. The boxplots represent the simulated posterior distribution of the daily streamflow. Outliers are not displayed. Red lines are the observed daily streamflow, and blue-dashed lines the posterior median daily streamflow. Scatter plots of the posterior median and observed flows along with the 1:1 line and bias for peak flows (computed only for dates where the observed flow exceeds the 90th quantile) and R values are on the upper right of each panel. R values are significant ($P - value < 0.1$). Black-dashed vertical lines indicate the division between validation periods.

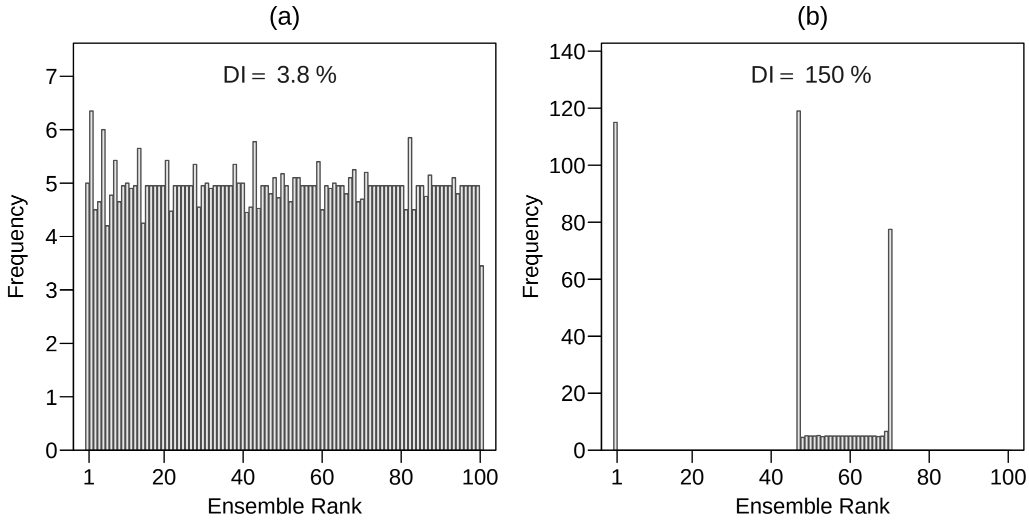

Figure 9. Rank histograms of the ensembles forecast of July-August daily streamflow during cross-validation periods. (a) the Bayesian Hierarchical Network Model and (b) the Multi-Linear Model at Mandleshwar gauge station. DI denotes the discrepancy index.

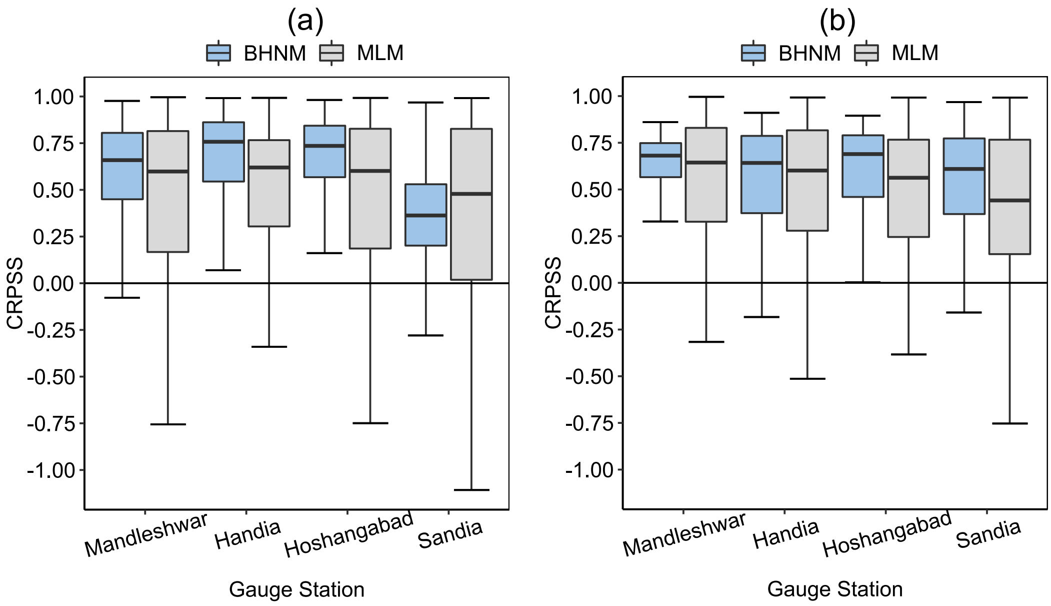

Figure 10. Boxplots of cumulative ranked probability skill score (CRPSS) statistic of streamflow ensembles from BHNM (sky blue boxes) and MLM (gray boxes) models for (a) 496 days of the four validation periods and (b) days with high flows. For boxplots, whiskers show the 95% credible intervals, boxes the interquartile range, and the horizontal lines inside the boxes, the median. Outliers are not displayed. Climatology was considered as the reference forecast model.

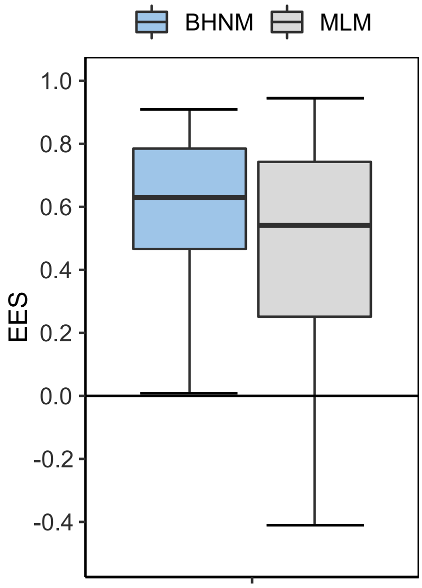

Figure 11. Boxplots of the energy skill score (RPSS) statistic of streamflow ensembles from BHNM (sky blue boxes) and MLM (gray boxes) for the four validation periods. For boxplots, whiskers show the 95% credible intervals, boxes the interquartile range, and the horizontal lines inside the boxes, the median. Outliers are not displayed. Climatology was considered as the reference forecast model.