Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 13 No 11, 3934 doi: 10.2166/wcc.2022.251

Streamflow forecasting in a climate change perspective using E-FUSE

Rishith Kumar Vogeti[a], Sriman Pankaj Boindala[a], D. Nagesh Kumar[b] and K. Srinivasa Raju[a,*]

[a] Department of Civil Engineering, BITS Pilani Hyderabad Campus, Hyderabad, India
[b] Department of Civil Engineering, Indian Institute of Science, Bangalore, India
*Corresponding author. E-mail: ksraju@hyderabad.bits-pilani.ac.in

ABSTRACT

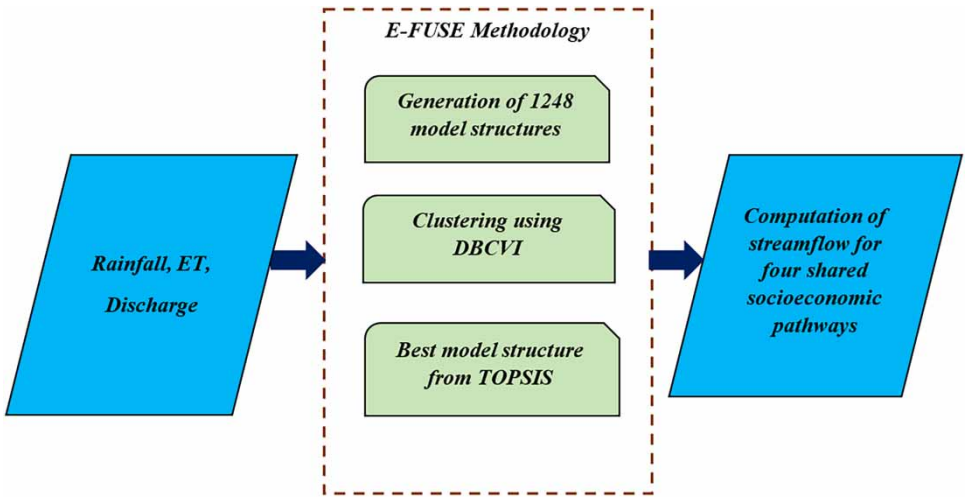
The present work aims to identify the best hydrological model structure suitable for the Lower Godavari River Basin, India, that forecasts streamflows. An extended version of the Framework for Understanding Structural Errors (FUSE), termed E-FUSE, is developed for this purpose. It consists of 1248 model structures. K means cluster analysis (KCA), and Davies Bouldin Cluster Validation Index (DBCVI) are used for identifying optimal clusters, whereas Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is employed for the best model structure. Correlation coefficient (r), normalized root mean square error (NRMSE), and mean bias error (MBE) are employed as evaluation criteria. The best model structure obtained exhibits ***r***, NRMSE and MBE of 0.734, 0.74 and -0.09 respectively during calibration and 0.69, 0.802 and -0.28 respectively during validation. The best model structure is then used to forecast discharges for a global climate model, EC-Earth3, and four Shared Socioeconomic Pathways, SSP126, SSP245, SSP370, and SSP585 scenarios. Analysis was made for three time horizons, namely, the near-future scenario (2021–2046), mid-future scenario (2047–2072), and far future scenario (2073–2099). It is observed that the July–September months contribute greatly to total runoff for four SSPs and three time horizons.

Key words: climate change, E-FUSE, KCA, runoff, SSP, TOPSIS

HIGHLIGHTS

- FUSE is extended by combining with K-means cluster analysis and multi-criteria decision-making technique, TOPSIS, to identify the best hydrologic model structure and applied to Lower Godavari River Basin, India.
- Runoff is forecasted using EC-Earth3 and four SSPs, namely SSP126, SSP245, SSP370, and SSP585 for near (2021–2046), mid (2047–2072), and far future (2073–2099) time horizons.

GRAPHICAL ABSTRACT



INTRODUCTION

Runoff response to meteorological variables in a catchment in a climate change scenario is always a research question and can be aided by effective hydrological modelling. However, reliable runoff assessment is governed by hydrologic model approximations, catchment data availability, and the quality of hydrologic measurements. Also, it requires holistic understanding of the architecture and flux parameterization of the individual hydrological model(s) to formulate one generalized framework that could perform effectively for any study area. Li *et al.* (2015) opined that suitable model structures with increased complexity are expected to yield realistic runoff simulations. This process enables policy-makers to decide on water security-related planning and management initiatives in data-rich and data-scarce catchments. Therefore, reliable runoff estimation is found to be essential.

Many studies utilized individual hydrological models, which worked effectively in a few catchments and showed no promise in others. As a result, a better framework must be built that can generate several models, increasing the likelihood of getting a suitable model within the framework. Such a structure would have the possibility of providing a better prediction across multiple catchments. In this context, Clark *et al.* (2008) initiated a Framework for Understanding Structural Errors (FUSE). It is the outcome of aggregating various components of four hydrologic models, namely, the National Weather Service Sacramento, Precipitation Runoff Modelling System (PRMS), TOPology MODEL (TOPMODEL), and Variable Infiltration Capacity (VIC). They formed 79 model structures based on the four mentioned hydrological models. These model structures were used for streamflow simulation of the French Broad and Guadalupe Rivers. They have concluded that FUSE simulated better for the French Broad River with a Nash Sutcliffe Efficiency (NSE) of 0.8 in comparison with Guadalupe River, with an NSE value ranging from 0.4 to 0.65.

Other researchers who explored FUSE are as follows: Staudinger *et al.* (2011) employed FUSE using 79 model structures to examine the influence of each model structure on low flows over the Narsjo catchment of the Norway region. They have concluded that lower soil architecture primarily influences low winter flows, whereas low summer flows are affected by a combination of model structures. The model structures were found satisfactory, having NSE greater than 0.8. Coxon *et al.* (2014) generated 78 modelling structures considering 10,000 parameter variations using uniform prior distribution over 24 UK catchments using FUSE. Performance metrics are the time step evaluation point, water balance signatures, and flow duration curve, and corresponding information is available from Littlewood & Croke (2008), Thoma *et al.* (2020), and Luan *et al.* (2021). They have concluded that the evaluation point metric is more dominant in evaluating the model performance than the other two, with an average precision of 60% and reliability of about 96%. Vitolo (2015) developed FUSE for generating 1248 model structures based on model building decisions, 10 state equations, 18 internal fluxes, and 24 model-specific and study-specific parameters over Plynlimon catchment, UK. FUSE performed satisfactorily in simulating low flow and medium flow events but lacked in capturing high flow events showing an NSE of 0.54. Lane *et al.* (2019) deployed FUSE for analyzing 1013 UK catchments. Model performance was assessed with NSE for daily flows and skill score for parameter uncertainty. It is concluded that over 80% of catchments have shown a satisfactory performance with an NSE value greater than 0.5 in simulating discharge. Newman *et al.* (2021) examined the model structure, model parameterization, and initial conditions for two USA watersheds having return periods ranging from 2 to 100,000 years. FUSE was employed to develop stochastic event-based hydrologic modelling. Kling Gupta Efficiency, NSE & Root Mean Square Error were employed to ascertain model performance. They have concluded that high rainfall significantly influences the variability of rare events, whereas initial conditions substantially impact the variance of more regular events. This study also revealed that Kling Gupta Efficiency is a better evaluation metric than the other two for the calibration of extreme events.

Quantifying the runoff under the influence of climate change is equally important as runoff dynamics are predominantly influenced by climate change (Herman *et al.* 2020) and is a challenging task. As a note, no studies on FUSE in Indian conditions or elsewhere are reported from a climate change perspective.

In this context, E-FUSE, an extended version of the FUSE in association with K-means cluster analysis (KCA), Davies-Bouldin cluster validation index (DBCVI) (Mahata *et al.* 2021), and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) (Farajpanah *et al.* 2020) is developed to determine the best hydrological model structure and related parameter variations for Lower Godavari River Basin (LGRB), India. LGRB is chosen for runoff simulation and forecasting purposes in the context of E-FUSE mainly due to the flexibility of diverse modelling options, which facilitates deriving realistic runoff simulations under drainage congestion, backwater effect, high evapotranspiration, and canopy losses (Amarnath & Thatikonda 2020). Accordingly, the objectives chosen are to:

i. Generate model structures in the E-FUSE environment
ii. Cluster the generated model structures using KCA and identify optimum cluster size using DBCVI
iii. Select suitable model structure (s) using TOPSIS.
iv. Forecast future runoff using a suitable model structure (obtained from iii) using Global Climate Model, EC-Earth3 in association with Shared Socioeconomic Pathways, SSP126, SSP245, SSP370, and SSP585, respectively.

A brief narrative of the case study is part of section 2. Section 3 describes FUSE, K-Means, DBCVI, TOPSIS, Evaluation metrics, EC-Earth3, and SSPs.

CASE STUDY

LGRB lies between the latitude of 16° 19′ to 22° 34′ N and the longitude of 73° 24′ to 83° 4′ E. The basin shares boundaries with Andhra Pradesh, Chhattisgarh, Odisha, Maharashtra, and Telangana. It has a basin length of 462 km and a total catchment area of 39,180 km^2. The catchment is dominated by red and black cotton soils. The majority of rainfall occurs during the southwest monsoon. Temperature ranges from 26° to 44° C. Demands that are expected to be met from LGRB are 664 Million Cubic Metres for drinking and industrial needs of Visakhapatnam, water diversion of approximately 2,265.35 Million Cubic Metres into Krishna river at Vijayawada. In addition, 226.54 Million Cubic Metres for stabilization of the Samarlakota branch canal, 141 Million Cubic Metres, and 42 Million Cubic Metres for water requirements for Odisha and Chhattisgarh states, respectively (Ministry of Water Resources GoI 2017). Table S1 (presented in the supplementary section) presents future rainfall and evapotranspiration (mm/day) under four SSPs.

These growing requirements necessitate accurate estimation of runoff which ensures water security. The upper cofferdam is 2.480 km long and 42.5 metres high. It is built to detain the natural watercourse to enable the construction of the Earth-cum Rockfill dam. A lower cofferdam, 1.616 km long and 30.50 m high, is built downstream of the spillway to ensure that the water does not flow back to the Earth-cum Rockfill dam site. The Ogee spillway with 48 vents is constructed to control the flow of water released during the flooding.

Perur is the inlet of the study area. Polavaram is considered an outlet instead of Dowleswaram, which is downstream of Polavaram. Dowleswaram was not considered an outlet as flow is influenced by left and right main canals, lower and upper cofferdams, and Earth-cum Rockfill dam. Figure 1 presents the map of LGRB describing the mentioned locations. Various studies have projected temperature rise over the study area, which intensifies the climate processes and has

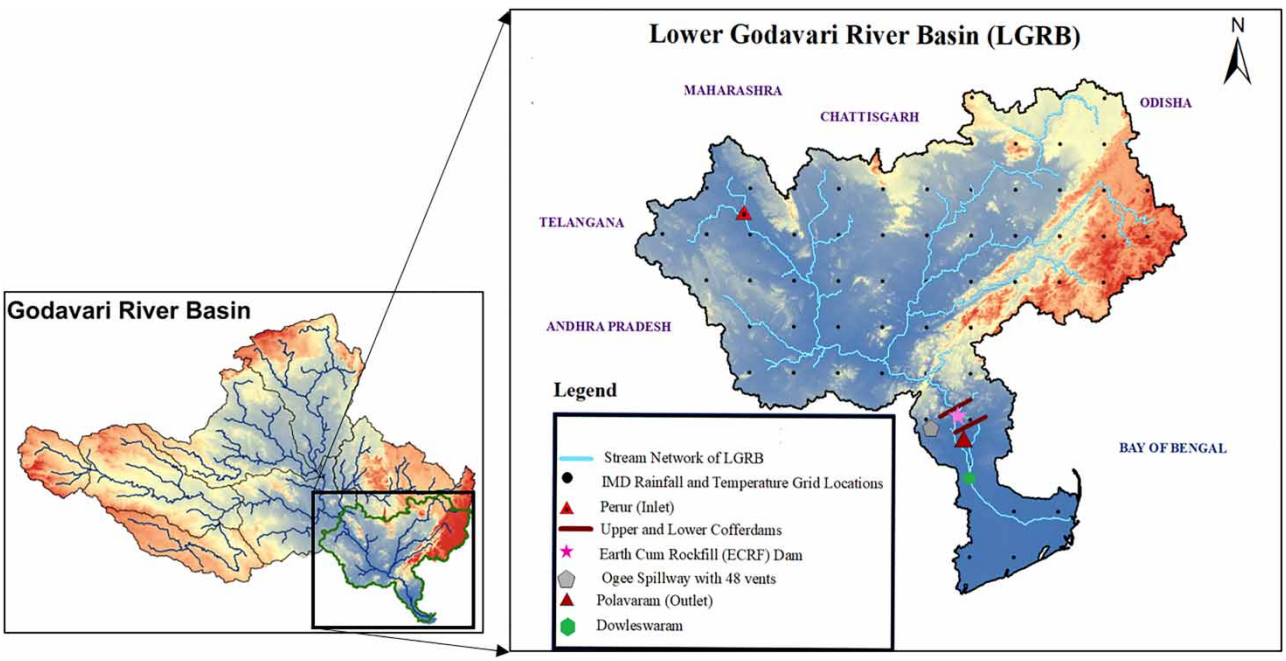


Figure 1 | Lower Godavari River Basin showing the location of cofferdams, spillway & Earth-cum Rockfill dam.

significant impacts on runoff of the catchment (Hengade *et al.* 2018). Therefore, it becomes essential to forecast the runoff from a climate change perspective and help policy-makers to formulate effective planning and management initiatives.

DESCRIPTION OF METHODS

FUSE is a modular modelling framework. The workflow of E-FUSE to identify the best hydrological model structure and parameter variation is presented in Figure 2. It is developed with PRMS (Leavesley *et al.* 1983), the National Weather

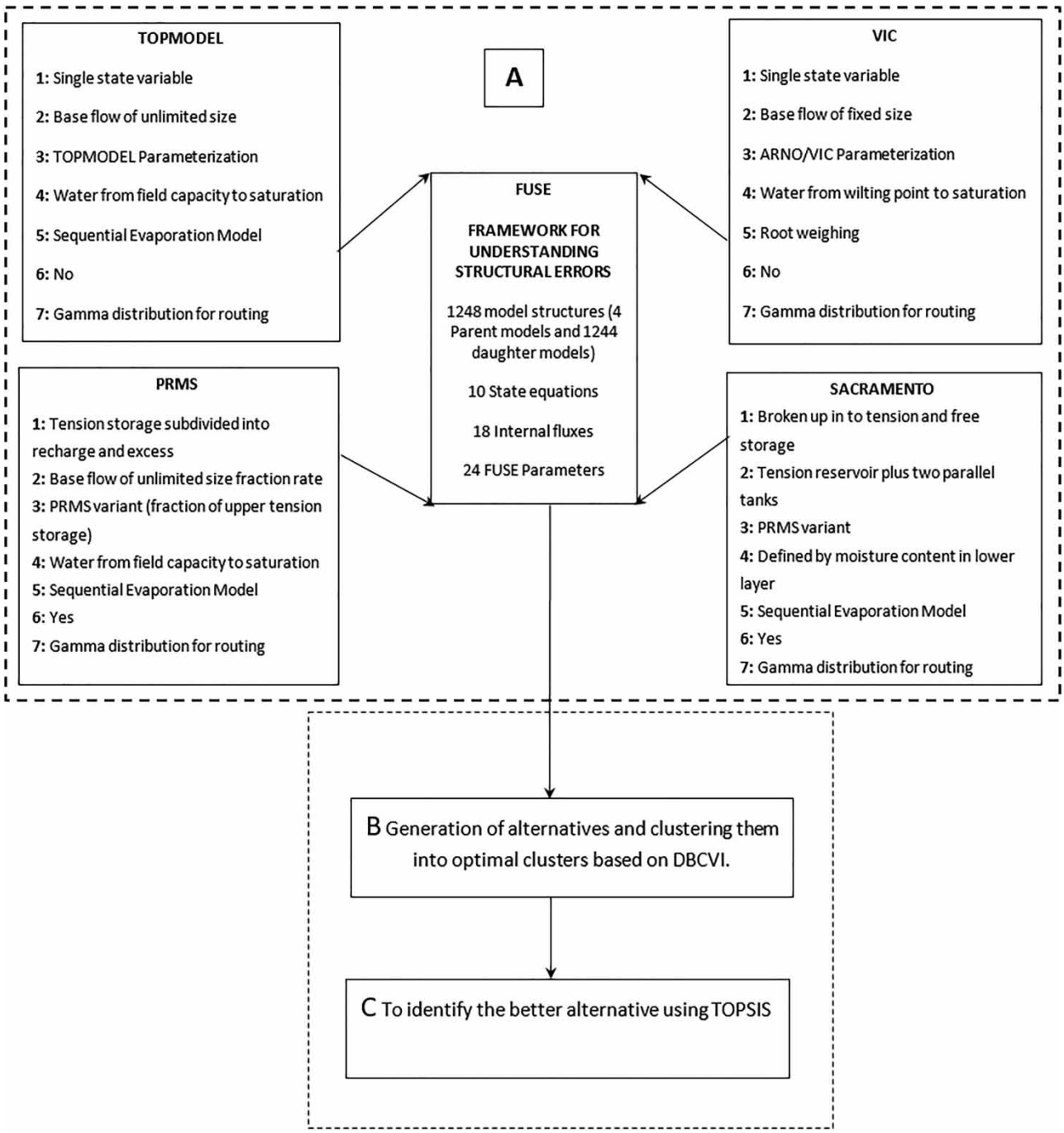


Figure 2 | Workflow of E-FUSE.

Service Sacramento (Burnash *et al.* 1973), VIC (Liang *et al.* 1994), and TOPMODEL (Beven & Freer 2001), terming them parent models. Each model differs in the design of the upper and lower soil layers and the parameterization of processes. FUSE could integrate components from various models to produce a variety of structures called daughter models (Clark *et al.* 2008).

The E-FUSE (refer to Figure 2) consists of three primary sections, namely, (A) generation of model structures/alternatives using FUSE, (B) application of the KCA on the alternatives and DBCVI to identify optimal clusters, and (C) TOPSIS to identify the best alternative. E-FUSE decision support system is developed in R-Platform. These are explained in brief as follows, along with evaluation metrics. Notations 1, 2, 3, 4, 5, 6, and 7 presented in Figure 2 represent Upper Layer, Lower Layer, Surface runoff, Percolation, Evaporation, Interflow, and Time delay in the runoff, respectively.

As part of **B**, KCA clusters data sets into relatively homogeneous sub-sets (Rao & Srinivas 2006). The mathematical background of KCA is presented in Figure S1 (provided in the supplementary material section). The number of clusters is an essential input for KCA. The present study uses DBCVI to determine the optimal number of clusters (Davies & Bouldin 1979). As part of **C**, TOPSIS is used for ranking alternatives (Opricovic & Tzeng 2004), and the mathematical background is presented in Table 1.

Performance metrics

Indicators are used simultaneously to judge the simulating ability of E-FUSE (Refer to Table 2). These are the correlation coefficient (***r***), the Normalized Root Mean Square Error (NRMSE), and the Mean Bias Error (MBE) (Moriasi *et al.* 2015), and are explained in Table 2.

Description of global climate model and SSPs

The present study uses EC-Earth3 (Mishra *et al.* 2020; Döscher *et al.* 2021). The change in the patterns of global society, economics, and demographics are being studied thoroughly by climate experts by exploring new pathways known as SSPs. The SSPs are based on socioeconomic trends and are intended to span the range of plausible futures (Riahi *et al.* 2017).

DATA COLLECTION AND PROCESSING

The application of E-FUSE requires considerable data and parameters. Daily rainfall, evapotranspiration, and discharge are the main inputs in E-FUSE, in addition to the parameters described in Figure 2. Daily rainfall of 0.25° grid resolution and temperature of 1° grid resolution are collected from India Meteorological Department for 1982–2020 over 52 grid locations resulting in 14,245 records. The nearest neighbourhood approach is used to interpolate temperature data to 0.25° grid. Temperature is used to compute evapotranspiration using the Penman-Monteith method. Daily discharge data over the Polavaram and Perur are obtained from Central Water Commission from 1982–2020. The Perur daily discharge deducted from Polavaram is considered as the observed discharge of LGRB for calibration purposes.

Future climate variables, namely, rainfall, maximum and minimum temperatures for 2021–2099, are obtained over LGRB (Mishra *et al.* 2020). These climate variables over 52 grid locations covering the LGRB are interpolated using the bilinear interpolation technique for four SSPs.

In addition to data collected in historical and future periods, the study also requires parameters and their corresponding ranges to address the effects of a hydrological system that influences the streamflow. Here, parameter refers to the values,

Table 1 | Mathematical background of TOPSIS

Mathematical expressions	Remarks
$k_j^*, k_j^{**}\ j = 1,2,\ldots..J$; J is the number of indicators	Ideal and anti-ideal solutions for each indicator j
$DS_a^+ = \sqrt{\sum_{j=1}^{J} (k_j(a) - k_j^*)^2}$	Distance measure of alternative a (DS_a^+) from ideal solution
$DS_a^- = \sqrt{\sum_{j=1}^{J} (k_j(a) - k_j^{**})^2}$	Distance measure (DS_a^-) of a from the anti-ideal solution
$CR_a = \frac{DS_a^-}{(DS_a^- + DS_a^+)}$	Relative closeness CR_a. Higher CR_a indicates a suitable alternative

Table 2 | Mathematical background of metrics

Mathematical expressions	Remarks
$r = \frac{\sum_{i=1}^{n}(O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(S_i - \bar{S})^2}}$	The higher r is preferred. Its range is -1 to $+1$. r of over 0.7 is considered satisfactory in the context of hydrological modelling O_i = Observed discharge (mm/day), $\bar{O}$ = Mean of the observed discharge (mm/day), σ = Standard deviation of the observed discharge (mm/day), S_i = Simulated discharge (mm/day), $\bar{S}$ = Mean of the simulated discharge (mm/day) & n = size of the sample considered
$NRMSE = \frac{1}{\sigma}\sqrt{\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{n}}$	The lower value of NRMSE is preferred. Its range is $-\infty$ to $+\infty$. NRMSE less than 0.7 is considered satisfactory in hydrological modelling
$MBE = \frac{\sum_{i=1}^{n}(O_i - S_i)}{\sum_{i=1}^{n}(O_i)}$	The lower value of MBE is preferred. Its range is -1 to $+1$. MBE range of -0.20 and 0.20 is considered satisfactory in hydrological modelling

which are model-dependent and study-specific, that influence the streamflow. There are 24 parameters suggested by Clark *et al.* (2008) whose ranges are required for this purpose (refer to Table 3). The drainage congestion is considered using parameters 12, 14, 18, 20, 21, and 24 and is amplified by higher base flow rates, percolation rates, interflow rates, and maximum saturated areas. Higher time delay values reflect more drainage congestion. The mean topographic index measures the surface flow accumulation at a given location, which will rise with an increase in area and a decrease in slope gradient. This will lead to drainage congestion and backwater effects. Thus, incorporating the parameters in runoff simulation enables a realistic amount of accounting for these challenges experienced in the study area. After identifying the study-specific and model-specific parameter thresholds, the next step is to generate numerous parameter variations. The higher parameter variations will provide more accurate results but will take a longer time for computation. Thus, depending on accuracy and time, the user can choose the number of parameter variations that can be selected. The Latin Hypercube Sampling (LHS) method (Devak & Dhanya 2017) generates the number of parameter variations. The total parameter threshold space is divided into equal (p) intervals with a probability of 1/p. Random parameter values are generated using stratified sampling only once for each interval for 24 hydrologic parameters. In this study, 20 parameter variations are considered as this would enable obtaining a behavioral threshold. The index of alternatives can be computed for the parameter variations from the generalized Equation (1).

The generalized equation formulated for computing the index of alternatives is as follows:

$$N_I = (N_{MR} - 1) \times (N_T) + (N_R) \tag{1}$$

Here, N_I, N_{MR} are the indices of alternative and model structure, N_T is the maximum number of parameter variations possible (20 in this case). N_R is the parameter variation at which index of alternative is required.

For example, if the given index model structure is 312, and the index of parameter variation is 6 (among 20 possible), then the resulting alternative index is (312-1) $\times$ 20 + 6 = 6,226. It means that index model structure 311 completed all 20 parameter variations, and index model structure 312 completed only six parameter variations as specified. If $N_T = N_R$, the equation would be modified as $N_I = N_{MR} \times N_T$.

After entering the number of runs in the decision support system, the user must decide which models can be used for runoff computation. Users can analyze four parent models, i.e., TOPMODEL, PRMS, VIC, SACRAMENTO, or resulting structures based on model-building decisions (refer to Figure 2).

Total model structures formed based on options available for model building decisions are rainfall error (2 levels), upper soil layer architecture (3 levels), lower soil layer architecture (4 levels), surface runoff (3 levels), percolation schemes (3 levels), evaporation (2 levels), interflow (2 levels) and routing schemes (2 levels) resulting in 1,728 combinations. Out of these, 480 model structures are incompatible. This incompatibility is caused by the interdependence of one model-building decision with the other (Clark *et al.* 2008). These are subtracted from the total of 1,728, yielding 1,248 model structures.

Table 3 | Model-specific and study-specific E-FUSE parameters

Parameter (1)	Functionality of Parameter (2)	Parameter abbreviation (3)	Units (4)	Ranges (5)	Optimum value (6)
1	Additive rainfall error	rferr_add	no units	(0,0)	0
2	Multiplicative rainfall error	rferr_mlt	no units	(1,1)	1
3	Fraction tension storage in the recharge zone	frchzne	no units	(0.05,0.95)	0.11
4	Fraction total storage in tension storage	fracten	no units	(0.05,0.95)	0.79
5	Depth of the upper soil layer	maxwatr_1	mm	(25,200)	166.69
6	Fraction of percolation to tension storage	percfrac	no units	(0.05,0.95)	0.16
7	Fraction storage in 1st base flow reservoir	fprimqb	no units	(0.05,0.95)	0.34
8	Base flow depletion rate 1st reservoir	qbrate_2a	day^{-1}	(0.001,0.25)	0.17
9	Base flow depletion rate 2nd reservoir	qbrate_2b	day^{-1}	(0.001,0.25)	0.11
10	Base flow depletion rate	qb_prms	day^{-1}	(0.001,0.25)	0.21
11	Depth of the lower soil layer	maxwatr_2	mm	(75,1200)	977.42
12	Base flow rate	baserte	mm/day	(0.865,865)	707.53
13	Fraction of roots in the upper layer	rtfrac1	no units	(0.05,0.95)	0.57
14	Percolation rate	percrte	mm/day	(185,8350)	3,507.23
15	Percolation exponent	percexp	no units	(1,20)	2.14
16	Sacramento model percolation multiplier for dry soil layer	sacpmlt	no units	(1,250)	11.43
17	Sacramento model percolation exponent for dry soil layer	sacpexp	no units	(1,5)	3.25
18	Interflow rate	iflwrte	mm/day	(0.865,865)	422.00
19	ARNO/VIC b exponent	axv_bexp	no units	(0.001,3)	2.39
20	Maximum saturated area	sareamax	no units	(0.05,0.95)	0.35
21	Mean value of the topographic index	loglamb	m	(5,10)	6.30
22	Shape parameter for the topo index Gamma distribution	tishape	no units	(2,5)	4.46
23	Base flow exponent	qb_powr	no units	(1,10)	4.55
24	Time delay in runoff	timedelay	days	(0.01,5)	4.93

Hence, these are used further for simulation and runoff forecasts. As a note, the historical and future climate variables are averaged over the grid locations covering LGRB for runoff forecasting.

RESULTS AND DISCUSSION

Calibration & validation

Among 14,245 records, 11,688 records, i.e., 1982–2013, were used for calibration, and the remaining 2,557 records, i.e., 2014–2020, for validation. The simulated outcomes obtained from each parameter variation (in this case 20) and corresponding model structure (in this case 1248), i.e., 24,960, are evaluated for ***r***, NRMSE, and MBE with the observed discharge. The weights of the evaluation metrics are assumed to be equal.

It is quite challenging to evaluate 24,960 alternatives to select the appropriate one that can be used for forecasting. KCA (refer to Figure S1 in the supplementary material section) is used to group these into a manageable set for further processing. KCA is analyzed with 5 to 20 clusters. Figure S2 presents the squared error value for each cluster. It is observed that square error is decreasing gradually up to cluster 12. After that, a small spike is observed, followed by fluctuations. In addition, the DBCVI is computed for the chosen cluster range and presented in Figure S3. It varies from 0.5607 to 0.3484 and found that

an optimum cluster size of 19 is preferred among the evaluated clusters. Thus, 24,960 alternatives are grouped into 19 clusters. Table 4 displays relevant information. The highest number of alternatives is observed in sub-cluster 18, whereas the lowest is in sub-cluster 16 (column 2). Columns 3 and 4 present model structure and parameter variation number. Columns 5, 6, and 7 present weighted evaluation metrics ***r***, NRMSE & MBE, respectively.

Based on the optimal clusters, the 19 representative alternatives are identified, describing the sub-clusters. TOPSIS is used to rank the representative alternatives based on the closeness measure and presented in Table 4. It is found that model number 605 with parameter variation 13, i.e., sub-cluster 7, is the best due to its higher closeness measure (columns 8–9). Table 3 (column 6) presents optimal values of parameters obtained corresponding to this best alternative. Upper soil layer depth occupying a value nearer to the higher threshold infers the high rainfall experienced in the basin. Lower soil layer depth occupying a value nearer to the higher threshold infers higher percolation, infiltration, and baseflow rates are contributing factors. Fraction of total storage in the tension zone occupies a value nearer to the lower threshold, inferring lesser total water storage found in the tension zone over the basin. Time delay in runoff occupying an optimal value closer to a higher threshold infers that the fluctuations in observed elevations in the catchment lead to an increase.

A scatter plot of observed discharge and simulated discharge obtained based on the best alternative is presented in Figure 3. Most of the discharge points are well within the 99% prediction interval for discharges less than 10 mm/day. The discharge points ranging from 10 mm/day to 40 mm/day were satisfactory, with few points falling outside the 99% prediction interval. A significant variation is observed in the discharge points having more than 40 mm/day, limiting the model's overall performance to ***r*** of 0.734, NRMSE of 0.74 & MBE of −0.09, respectively.

Figure 4 presents observed and simulated discharges. It is noted from Figure 4 that the best model structure is able to mimic the daily discharge records with observed records. Amplitude error is predominant over the phase and shape errors due to fluctuation in observed peak discharges. The model underestimated almost 70% of the amplitudes (peak discharges) in 1983, 1988, 1990, 1994, 1995, 1999, 2000, 2002, 2010, and 2013 respectively. These underestimations in the peak discharges are not due to large data usage. However, it may be attributed to the averaged inputs (Bárdossy & Anwar 2022). The model has

Table 4 | Clustering information and ranking of alternatives

Sub-Cluster number (1)	Number of Alternatives in each sub-cluster (2)	Model ID/ Model Structure (3)	Parameter Variation number (4)	Representative Alternative (5)	Weighted *r* (6)	Weighted NRMSE (7)	Weighted MBE (8)	C_a values (9)	Rank (10)
1	1,988	163	11	3,251	0.237	− 0.253	− 0.100	0.944	4
2	1,842	969	13	19,373	0.111	− 0.333	− 0.046	0.884	10
3	3,332	1,029	8	20,568	0.107	− 0.332	− 0.084	0.874	12
4	138	227	18	4,538	0.112	− 1.053	− 0.156	0.510	18
5	1,906	1,033	20	20,660	0.154	− 0.314	− 0.085	0.904	8
6	1,790	117	7	2,327	0.196	− 0.287	− 0.092	0.931	5
7	1,300	605	13	12,093	0.244	− 0.248	− 0.031	0.997	1
8	272	813	12	16,252	0.183	− 0.563	− 0.153	0.743	16
9	516	701	12	14,012	0.223	− 0.397	− 0.058	0.884	9
10	1,804	1,075	14	21,494	0.231	− 0.263	− 0.066	0.966	2
11	402	657	15	13,135	0.126	− 0.390	− 0.045	0.864	14
12	1,436	697	5	13,925	0.172	− 0.310	− 0.038	0.928	6
13	142	217	11	4,331	0.182	− 0.593	− 0.480	0.554	17
14	1,050	915	20	18,300	0.221	− 0.292	− 0.027	0.962	3
15	1,054	855	18	17,098	0.176	− 0.309	− 0.148	0.883	11
16	28	775	20	15,500	0.141	− 0.880	− 1.017	0.130	19
17	1,540	839	2	16,762	0.113	− 0.327	− 0.149	0.853	15
18	2,942	631	20	12,620	0.109	− 0.329	− 0.112	0.866	13
19	1,478	259	20	5,180	0.223	− 0.267	− 0.148	0.905	7

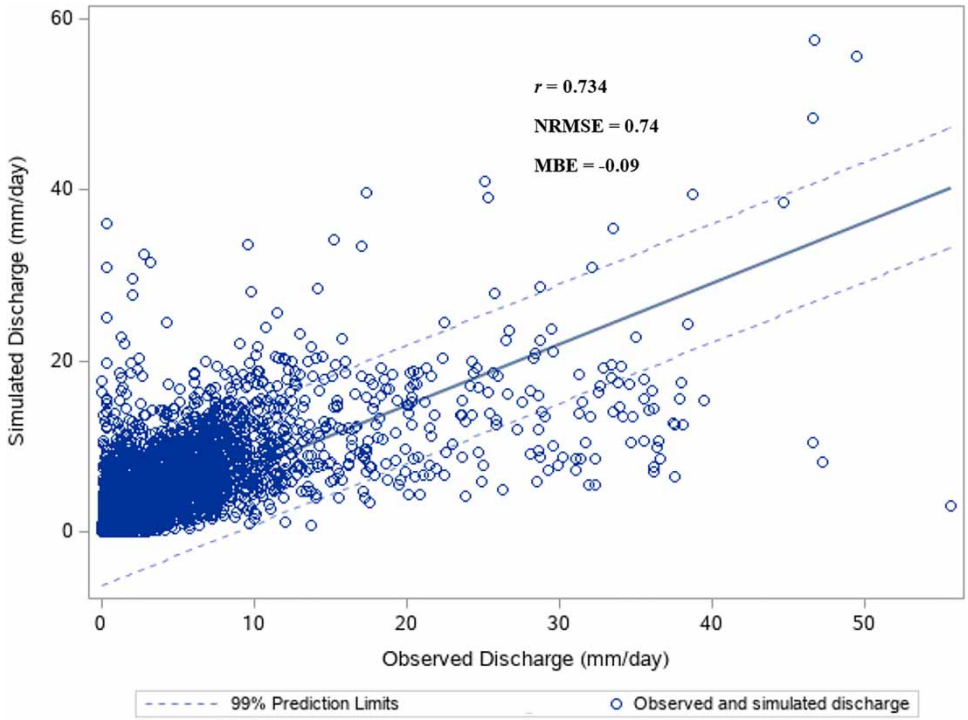


Figure 3 | Scatter plot during the calibration period (1982–2013).

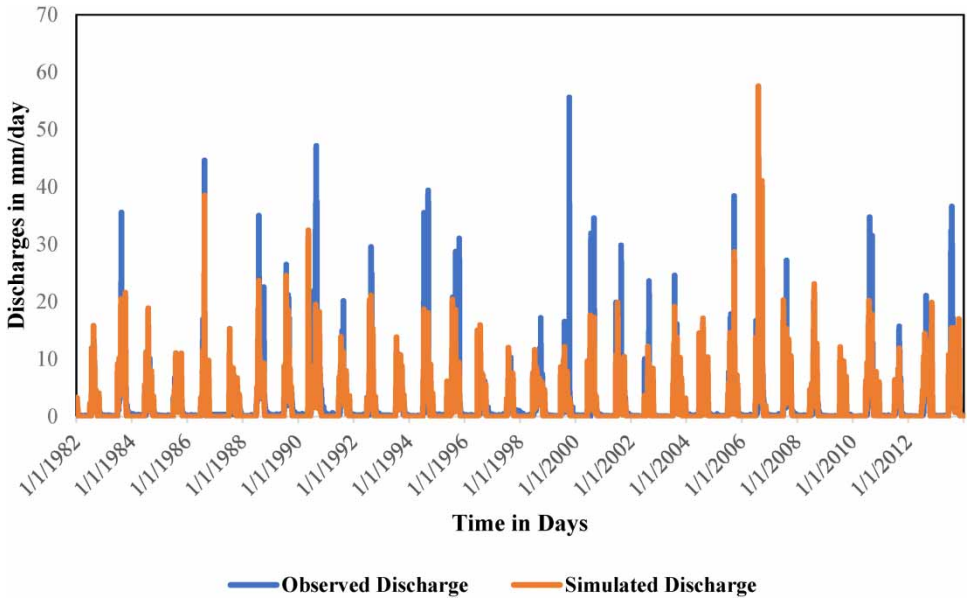


Figure 4 | Discharge plot for 11688 records in the calibration period (1982–2013).

shown the maximum deviation from the historical value in 1999 with a 78.2% decrease in amplitude. At the same time, the rest of the amplitudes are found to be overestimated during the years 1982, 1987, and 2006 respectively. The model has shown the maximum deviation from the historical value in 1987 with an 86.8% increase in amplitude.

Twenty-four calibrated parameters obtained from the best alternative (refer to column 6, Table 3; columns 9 and 10, Table 4) are used for validation purposes. The best model structure 605 and the corresponding best parameter variation

13 identified from E-FUSE for the calibration period is also used to validate the model. The daily records used for this are 2,557, i.e., 2014–2020. The relevant scatter plot is presented in Figure 5. It is found that most of the discharge points are well within the 99% prediction interval for discharges less than 10 mm/day. The discharge points ranging from 10 mm/day to 20 mm/day were satisfactory, with few points falling outside the 99% prediction interval. The overall model performance of the identified best model structure is limited to r of 0.69, NRMSE of 0.802, and MBE of −0.28, respectively, as the model has underestimated when flows exceed 40 mm/day.

The corresponding discharge plot for the validation period is presented in Figure 6. It is observed that simulated discharge is nearer to satisfactory performance when compared with the observed. It has been observed that the model has mostly underestimated the amplitude in the validation period. Significant deviations were found in 2014, 2016, 2018, and 2020 in peak values of about 65.6%, 58.25%, 41.36%, and 36.57%, respectively. However, in 2015, the model overestimated the peak value by 29.4%.

Effect of changing calibration-validation ratios (82%, 18%; 75%, 25%) and without/with discarding (20 mm/day; 40 mm/day) discharge points on model performance are studied. It is noticed that the effect of calibration-validation ratios & discarding discharge points greater than 20 mm/day and 40 mm/day have a minimal effect on model performance

It was found that the E-FUSE model has performed well, as observed from the metrics, despite forcing average data for a large area of over 39,180 km^2. Accordingly, E-FUSE is applied from a climate change perspective.

Forecasting runoff from a climate change perspective

Future runoff is based on rainfall and evapotranspiration related to EC-Earth 3-four SSPs and optimum parameters corresponding to best model structure 605 and parameter variation 13 (refer to Table 3). The forecasted runoff for four SSPs in three future time horizons is discussed under the following sub-sections.

Near future scenario (2021–2046)

Figures 7–10 represent the variation of future runoff for NFS for SSP126, SSP245, SSP370, and SSP585, and the relevant discussion is as follows:

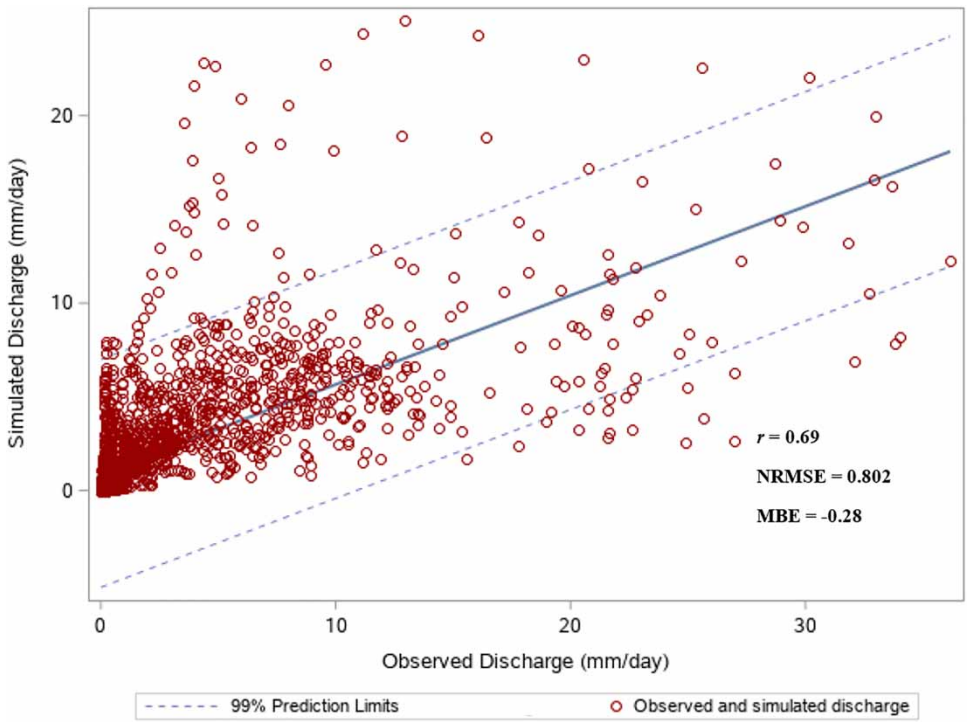


Figure 5 | Scatter plot during the validation period (2014–2020).

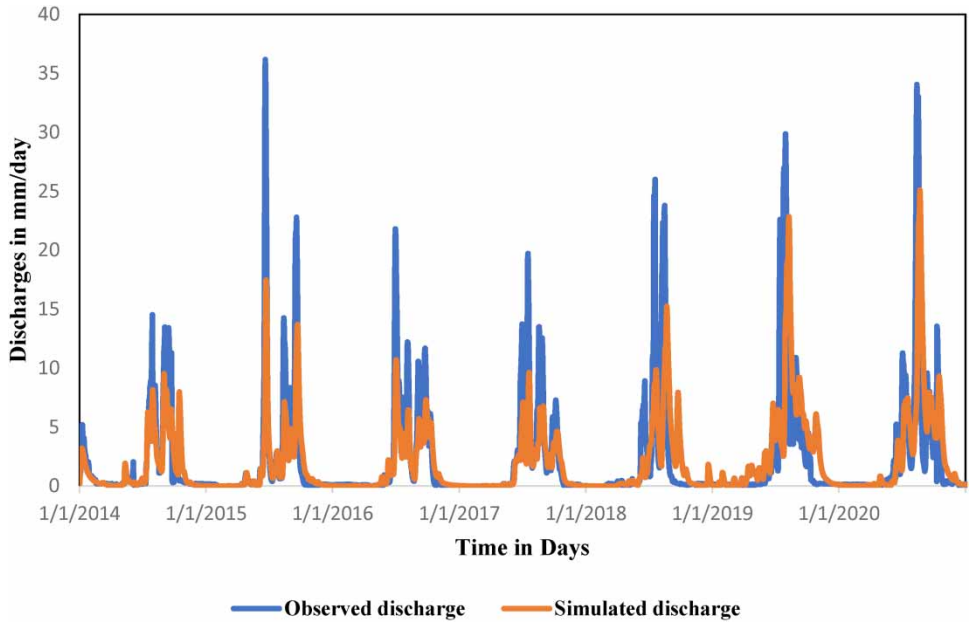


Figure 6 | Discharge plot during the validation period (2014–2020).

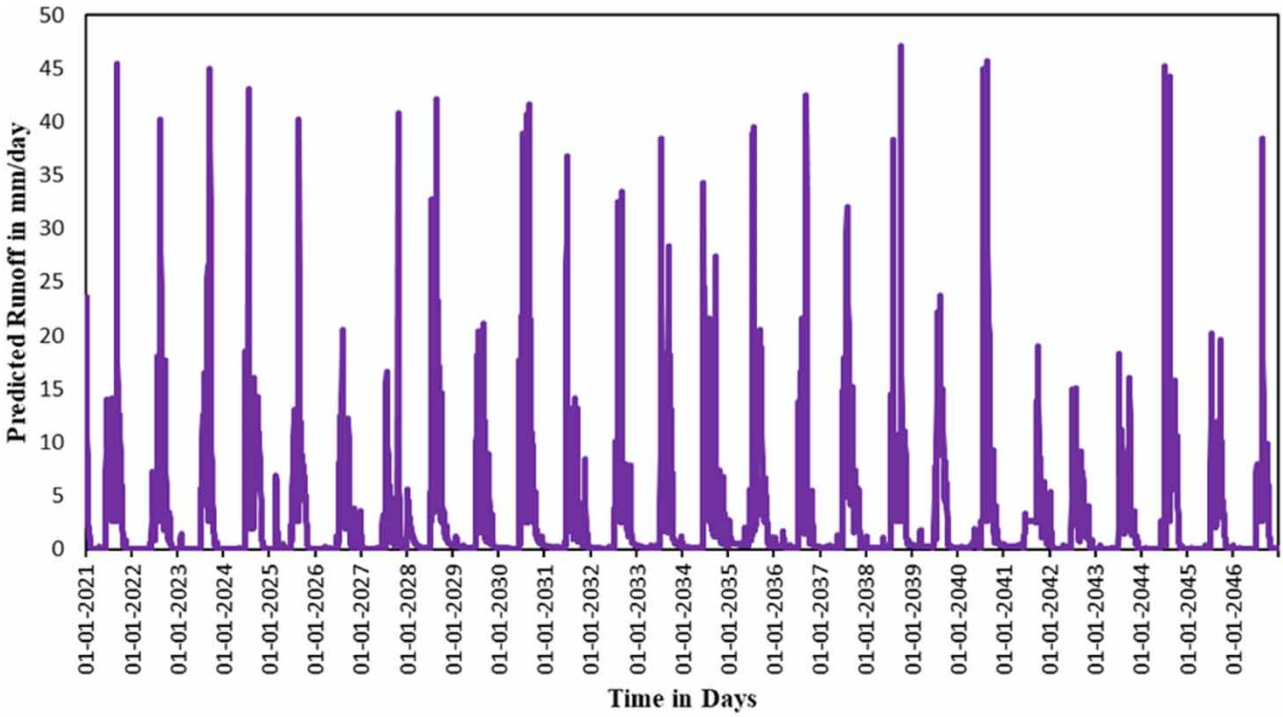


Figure 7 | Forecasted runoff in NFS (2021–2046) under SSP126.

- Maximum and minimum peak values of 47.15 mm/day and 16 mm/day may occur in 2038 and 2045 for SSP126; 39.74 mm/day and 11.03 mm/day may occur in 2042 and 2022 for SSP245; 46.47 mm/day and 14.37 mm/day may occur in 2046 and 2030 for SSP370; and 36.81 mm/day and 13.72 mm/day may occur in 2024 and 2035 for SSP585. Furthermore, SSP126, SSP245, SSP370, and SSP585 show an average increase of 27.78%, a decrease of 9.24%, a decrease of 4.12%, and an increase of 8.92% in peak annual runoffs in comparison to the baseline scenario.

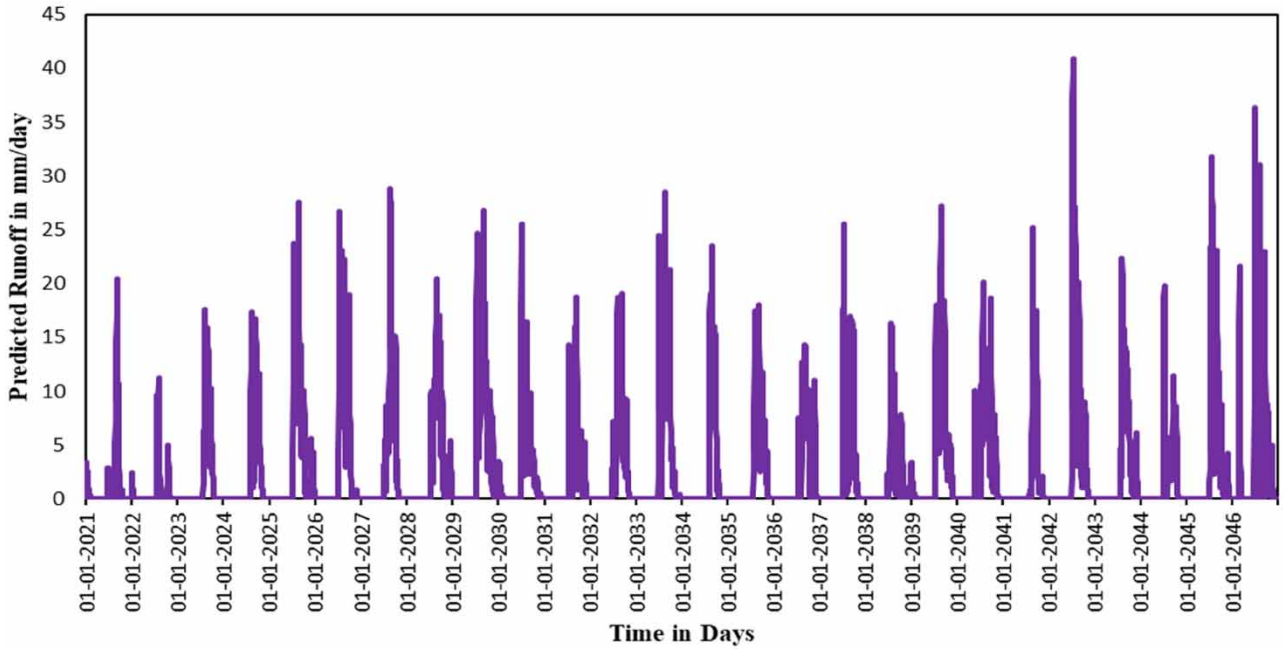


Figure 8 | Forecasted runoff in NFS (2021–2046) under SSP245.

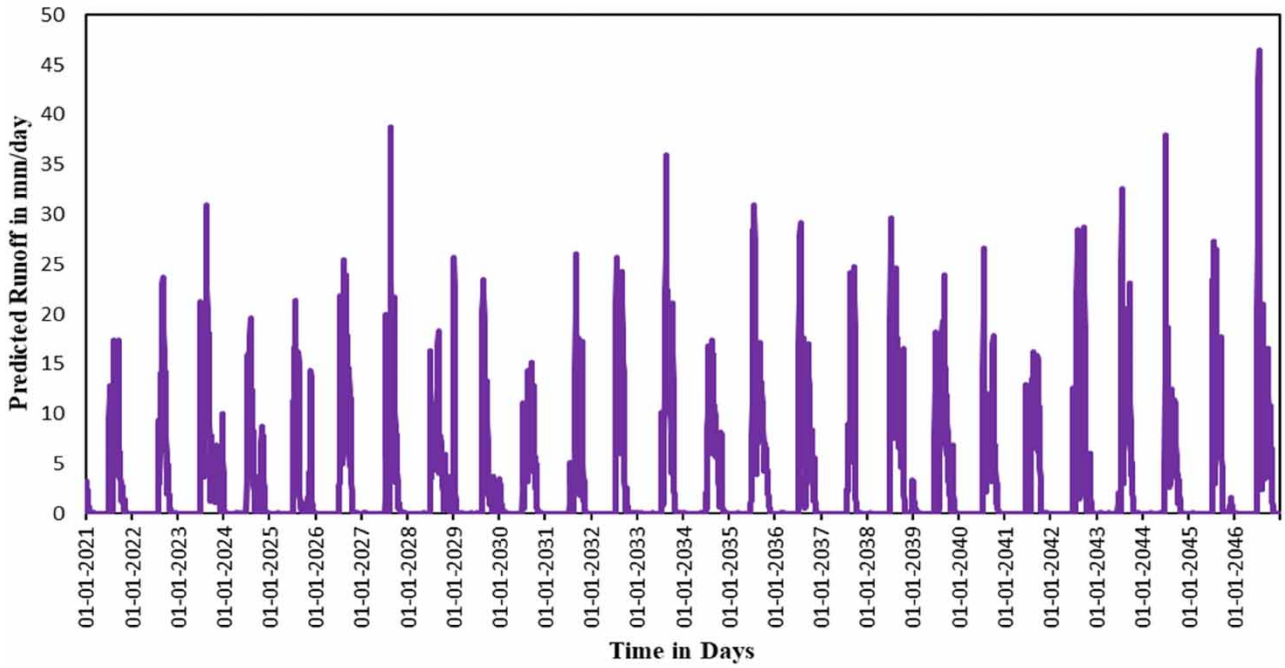


Figure 9 | Forecasted runoff in NFS (2021–2046) under SSP370.

- There is a significant increase in mean runoff in four SSPs compared to the mean runoff in the historical scenario. The maximum percentage increase in mean runoff is in the SSP370 scenario at 66.41%, and the minimum increase in mean runoff is in the SSP585 scenario at 40.01%.
- The highest and lowest monthly runoffs are forecasted for all four SSPs in August and April. These values are 226 mm/day and 2.65 mm/day for SSP126; 259.29 mm/day and 0.01 mm/day for SSP245; 277.58 mm/day and 0.02 mm/day for

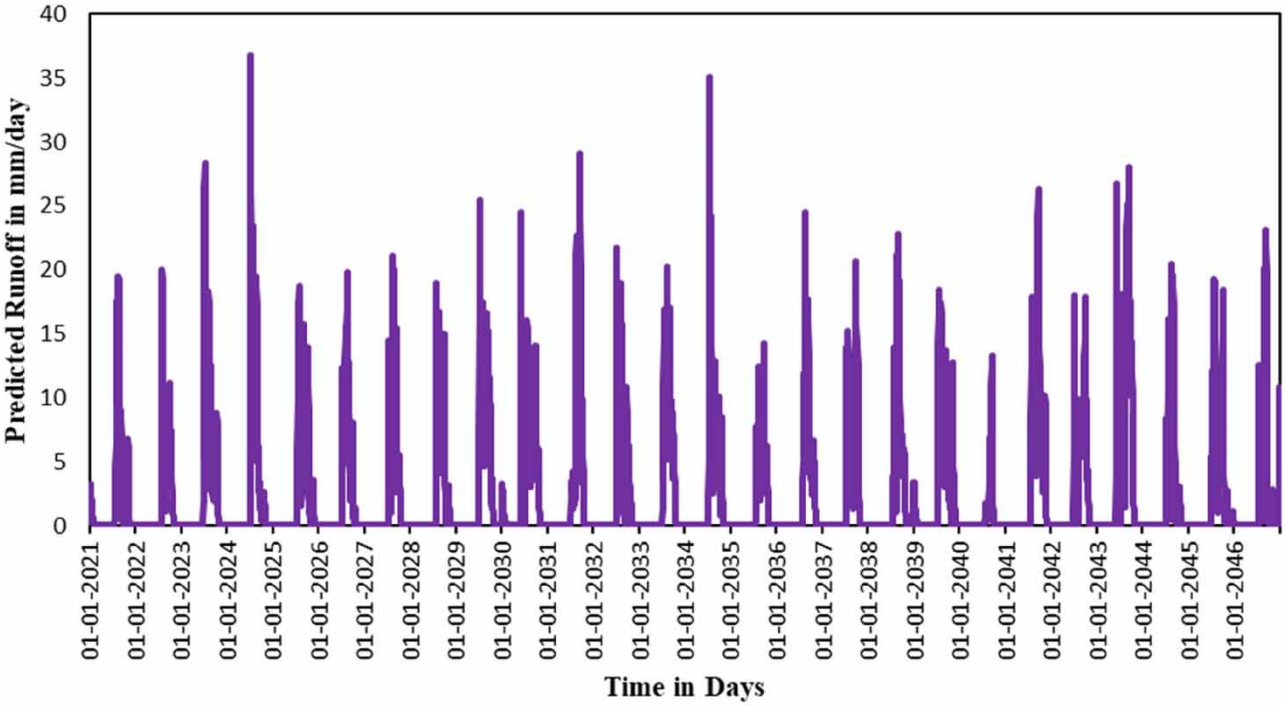


Figure 10 | Forecasted runoff in NFS (2021–2046) under SSP585.

SSP370; 247.16 mm/day and 0.01 mm/day for SSP585. In addition, July-September contributes 74.43%, 82.33%, 83.56%, and 83.11% of total runoff for SSP126, SSP245, SSP370, and SSP585. The forecasted monthly mean runoff is presented in Figure 11.

- Significant runoff occurrence events are decided based on the minimum value of the highest runoffs, 48.96 mm/day, 40.92 mm/day, 49.94 mm/day, and 71.63 mm/day observed in SSP126, SSP245, SSP370, and SSP585 respectively, i.e., 40.92 mm/day. Here, the runoffs exceeding 40.92 mm/day are considered significant runoff events. The number of significant runoffs, i.e., more than 40.92 mm/day for SSP126 and SSP370, are 11 and 3. In contrast, no significant runoff events are observed in SSP585. Thus, the total number of significant runoff events observed in NFS is 14.

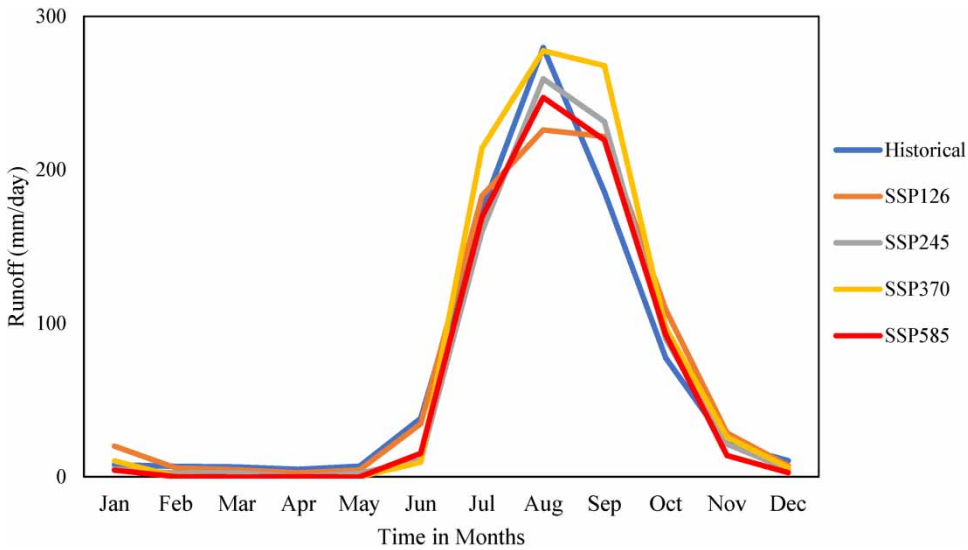


Figure 11 | Comparative analysis of mean runoff of four SSPs with the historical period in NFS (2021–2046).

Mid future scenario (2047–2072)

Variation of future runoff for MFS for four SSPs and the relevant discussion is as follows:

- Maximum and minimum peak values of 48.96 mm/day and 18.23 mm/day may occur in 2070 and 2067 for SSP126; 32.65 mm/day and 13.6 mm/day may occur in 2067 and 2069 for SSP245; 41.98 mm/day and 11.01 mm/day may occur in 2057 and 2054 for SSP370; and 43.54 mm/day and 14.75 mm/day may occur in 2072 and 2051 for SSP 585. Furthermore, SSP126, SSP245, SSP370, and SSP585 show an average increase of 32.94%, an increase of 1.42%, a decrease of 6.86%, and an increase of 9.62% in peak annual runoffs in comparison to the baseline scenario. An increase in mean runoff of 87.86% and 65.37% is observed in SSP585 and SSP370 scenarios (refer to Figures S4-S7 in the supplementary section).
- Figure 12 presents the highest and lowest monthly runoffs forecasted for various months. These are March and August in the case of SSP126 (288.11 mm/day and 0.68 mm/day); May and September in the case of SSP245 (292.13 mm/day and 0.01 mm/day); April and September in the case of SSP370 (301.32 mm/day and 0.01 mm/day); April and August in the case of SSP585 (296.47 mm/day and 0.01 mm/day). In addition, July-September contributes greatly 78.83%, 81.71%, 79.72%, and 78.10% of total runoff for SSP126, SSP245, SSP370, and SSP585.
- The number of significant runoff events (more than 40.92 mm/day) for SSP126, SSP370, and SSP585 are 12, 4, and 6, respectively. Thus, the total number of significant runoff events observed in MFS is 22.

Far future scenario (2073–2099)

Variation of future runoff for FFS for SSP126, SSP245, SSP370, and SSP585, and the relevant discussion is as follows:

- Maximum and minimum peak values of 47.6 mm/day and 24.86 mm/day may occur in 2096 and 2080 for SSP126; 40.24 mm/day and 17.12 mm/day may occur in 2074 and 2086 for SSP245; 49.94 mm/day and 15.52 mm/day may occur in 2094 and 2088 for SSP370; and 69.92 mm/day and 13.59 mm/day may occur in 2090 and 2080 for SSP585. SSP126, SSP245, SSP370, and SSP585 show an average increase of 35.6%, 8.39%, 21.45%, and 23.896%, respectively in peak annual runoffs in comparison to the baseline scenario. Percentage increases in the runoff for SSP585 and SSP126 are 111.56% and 55.22% (refer to Figures S8-S11 in the supplementary section).
- The plot for monthly mean runoff in FFS is presented in Figure 13. The highest and lowest monthly runoffs would occur in August and April in SSP126, SSP370, and SSP585. It is September and April in the case of SSP245. These values are 288.69 mm/day and 0.71 mm/day for SSP126; 314.20 mm/day and 0.01 mm/day for SSP245; 370.50 mm/day and 0.02 mm/day for SSP370; 383.77 mm/day and 0.02 mm/day for SSP585. In addition, July-September contributes greatly 81.13%, 79.25%, 76.05%, and 73.17% of total runoff for SSP126, SSP245, SSP370, and SSP585.

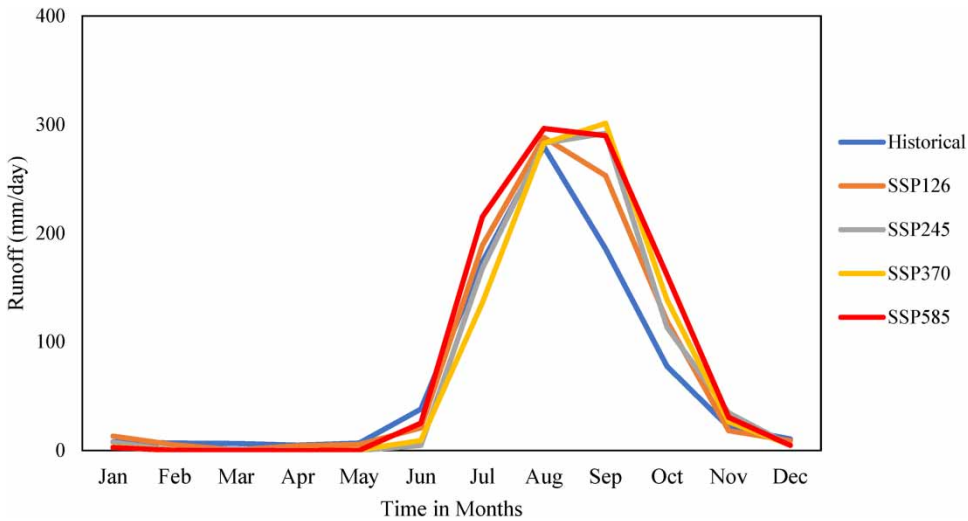


Figure 12 | Comparative analysis of mean runoff of four SSPs with the historical period in MFS (2047–2072).

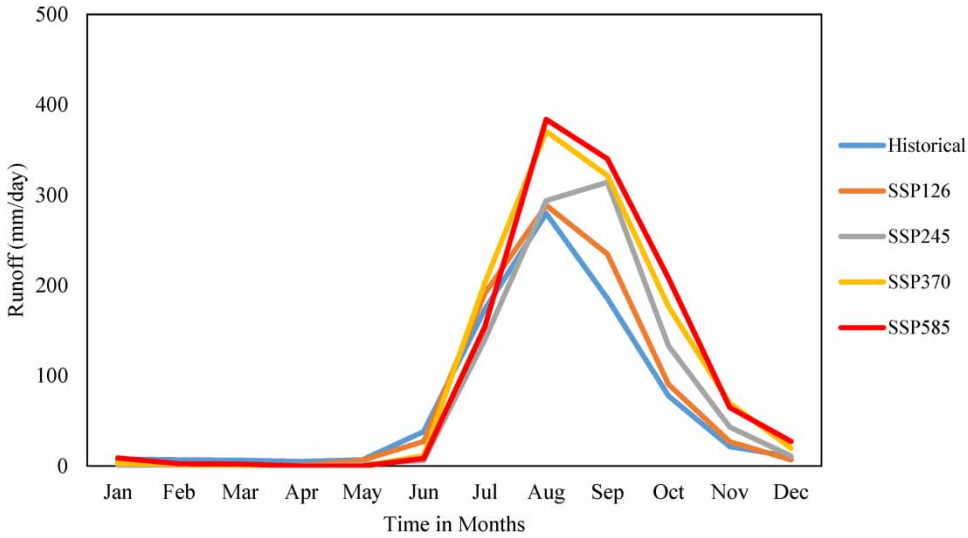


Figure 13 | Comparative analysis of mean runoff of four SSPs with the historical period in FFS (2073–2099).

- The number of significant runoffs, i.e., more than 40.92 mm/day for SSP126, SSP370, and SSP585, are 14, 15, and 16, respectively. Thus, the total number of significant runoff events observed in FFS is 45. It is also observed that more significant runoff events are found to be occurring in FFS as compared to NFS and MFS.

SUMMARY AND CONCLUSIONS

Streamflow forecasting in a given catchment area is a highly complex problem, especially under changing climate. In the past few decades, multitudes of hydrological models have been developed, and their performance has been tested in real-world case studies. However, no model is found to be suitable universally for all the catchments. Moreover, the model's accuracy largely depends on the availability of hydrological and meteorological data. This is in addition to the model structure and calibration mechanism. However, most catchments in developing countries are constrained by limited data availability. This restricts the hydrologists from focusing on hydrological models that use minimal data. They are expected to yield reasonably accurate outcomes. This promotes the need to obtain the best hydrological model that uses limited data and simultaneously forecasts the streamflow satisfactorily. In this context, this work proposes E-FUSE, the outcome of aggregating the components of four hydrologic models: the National Weather Service Sacramento, PRMS, TOPMODEL, and VIC.

E-FUSE was demonstrated in the LGRB. The 14,245 historical data sets comprise daily precipitation, evapotranspiration, and discharge. Out of which, 11,688 are used for calibration and the remaining for validation. Twelve hundred and forty-eight model structures based on the mentioned four hydrological models are formulated. Twenty-four parameters govern E-FUSE. Also, KCA and DBCVI are used for identifying optimal clusters, whereas TOPSIS is employed for the best model structure. This and the corresponding parameter set obtained from E-FUSE performed satisfactorily in both calibration and validation. The best model structure index is 605 with the 13^{th} parameter variation.

It is further used to forecast discharges from 2021 to 2099 for a global climate model, EC-Earth3, and four SSPs. A significant increase in mean runoff is observed from SSP126 to SSP585. Variations of runoff in NFS are from 40.01% to 66.41%; these are 65.37% to 87.86% in MFS and 55.22% to 111.56% in FFS. In MFS, an increase in mean peak runoff is observed in SSP126 and SSP585, and vice-versa in the other two SSPs. An increase in mean peak runoff is observed in all four SSPs in FFS. For three time horizons, the July-September months contribute greatly to total runoff for all four SSPs.

Increased runoff in NFS, MFS, and FFS may pose many challenges, including public and private property damage, aquatic habitat damage, sediment clogging at the left and right main canals at Polavaram, obstructing interlinking pathways to Penna and Krishna rivers, the possibility of eroding stream channels, and devastating inundation.

Slight deviations in the performance metrics in calibration and validation may be due to anthropogenic influence observed during the validation period and an insufficient number of parameter variations. Accordingly, further study can be extended to include additional conceptual parent model structure into E-FUSE, which may enhance flexibility in identifying better

model structure. Incorporating an efficient calibration strategy may also aid in finding suitable calibrated parameter values for the generated models. E-FUSE can be extended to multiple input formats instead of averaging to a single value, as employed in the present study.

ACKNOWLEDGEMENTS

This work is supported by the Council of Scientific and Industrial Research, New Delhi, through Project no. 22(0782)/19/EMR-II dated 24.7.19. The third author acknowledges the funding support by the Ministry of Earth Sciences, Govt. of India, through project# MoES/PAMC/H&C/41/2013-PC-II. Special acknowledgments to Ms C. Vitolo, Scientist, European Space Agency Centre for Earth Observations, for providing insights about FUSE; Dr RN Sankhua, Mr Shankar Rao, officials of Godavari River basin, for providing time for brainstorming discussions; & Prof Vimal Mishra, IIT Gandhinagar for making CMIP6 data available. Acknowledgments to India Meteorological Department and Central Water Commission for providing necessary data.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST STATEMENT

The authors declare there is no conflict.

REFERENCES

Amarnath, C. R. & Thatikonda, S. 2020 Study on backwater effect due to Polavaram Dam Project under different return periods. *Water* **12** (2), 576.

Bárdossy, A. & Anwar, F. 2022 Why our rainfall-runoff models keep underestimating the peak flows? *Hydrol. Earth Syst. Sci. Discuss* 1–30.

Beven, K. & Freer, J. 2001 A dynamic topmodel. *Hydrol. Process.* **15** (10), 1993–2011.

Burnash, R. J. C., Ferral, R. L. & McGuire, R. A. 1973 *A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers*. US Department of Commerce, National Weather Service, and State of California, Department of Water Resources, Sacramento, CA.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T. & Hay, L. E. 2008 Framework for understanding structural errors (FUSE): a modular framework to diagnose differences between hydrological models. *Water Resour. Res.* **44** (12), W00B02.

Coxon, G., Freer, J., Wagener, T., Odoni, N. A. & Clark, M. 2014 Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments. *Hydrol. Process.* **28** (25), 6135–6150.

Davies, D. L. & Bouldin, D. W. 1979 A cluster separation measure. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence ITPIDJ 0162-8828*. PAMI-1, 2, pp. 224–227.

Devak, M. & Dhanya, C. T. 2017 Sensitivity analysis of hydrological models: review and way forward. *J. Water Clim. Chang.* **8** (4), 557–575.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., Bernadello, R., Bousetta, S., Caron, L. P. & Carver, G. 2021 The EC-earth3 earth system model for the climate model intercomparison project 6. *Geosci. Model Dev. Discuss.* **1**, 2021.

Farajpanah, H., Lotfirad, M., Adib, A., Esmaeili-Gisavandani, H., Kisi, Ö., Riyahi, M. M. & Salehpoor, J. 2020 Ranking of hybrid wavelet-AI models by TOPSIS method for estimation of daily flow discharge. *Wat. Supp.* **20** (8), 3156–3171.

Hengade, N., Eldho, T. I. & Ghosh, S. 2018 Climate change impact assessment of a river basin using CMIP5 climate models and the VIC hydrological model. *Hydrol. Sci. J.* **63** (4), 596–614.

Herman, J. D., Quinn, J. D., Steinschneider, S., Giuliani, M. & Fletcher, S. 2020 Climate adaptation as a control problem: review and perspectives on dynamic water resources planning under uncertainty. *Water Resour. Res.* **56** (2), e24389.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., Greene, S., Macleod, C. J. & Reaney, S. M. 2019 Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. *Hydrol. Earth Syst. Sci.* **23** (10), 4011–4032.

Leavesley, G. H., Lichty, R. W., Troutman, B. M. & Saindon, L. G. 1983 Precipitation-runoff modeling system: user's manual. *Water Resour. Investig. Rep.* **83**, 4238.

Li, H., Xu, C. Y. & Beldring, S. 2015 How much can we gain with increasing model complexity with the same model concepts? *J. Hydrol.* **527**, 858–871.

Liang, X., Lettenmaier, D. P., Wood, E. F. & Burges, S. J. 1994 A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res. Atmos.* **99** (D7), 14415–14428.

Littlewood, I. G. & Croke, B. F. 2008 Data time-step dependency of conceptual rainfall – streamflow model parameters: an empirical study with implications for regionalization. *Hydrol. Sci. J.* **53** (4), 685–695.

Luan, J., Liu, D., Lin, M. & Huang, Q. 2021 The construction of the flow duration curve and the regionalization parameters analysis in the northwest of China. *J. Water Clim. Chang.* **12** (6), 2639–2653.

Mahata, K., Das, R., Das, S. & Sarkar, A. 2021 Land use land cover map segmentation using remote sensing: a case study of Ajoy river watershed, India. *Intell. Syst.* **30** (1), 273–286.

Ministry of Water Resources, GoI 2017 Polavaram Project Authority.

Mishra, V., Bhatia, U. & Tiwari, A. D. 2020 Bias-corrected climate projections for South Asia from coupled model intercomparison project-6. *Sci. Data.* **7** (1), 1–13.

Moriasi, D. N., Gitau, M. W., Pai, N. & Daggupati, P. 2015 Hydrologic and water quality models: performance measures and evaluation criteria. *Trans. ASABE.* **58** (6), 1763–1785.

Newman, A. J., Stone, A. G., Saharia, M., Holman, K. D., Addor, N. & Clark, M. P. 2021 Identifying sensitivities in flood frequency analyses using a stochastic hydrologic modeling system. *Hydrol. Earth Syst. Sci.* **25** (10), 5603–5621.

Opricovic, S. & Tzeng, G. H. 2004 Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS. *Eur. J. Oper. Res.* **156** (2), 445–455.

Rao, A. R. & Srinivas, V. V. 2006 Regionalization of watersheds by fuzzy cluster analysis. *J. Hydrol.* **318** (1–4), 57–79.

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O. & Lutz, W. 2017 The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change.* **42**, 153–168.

Staudinger, M., Stahl, K., Seibert, J., Clark, M. P. & Tallaksen, L. M. 2011 Comparison of hydrological model structures based on recession and low flow simulations. *Hydrol. Earth Syst. Sci.* **15** (11), 3447–3459.

Thoma, D. P., Tercek, M. T., Schweiger, E. W., Munson, S. M., Gross, J. E. & Olliff, S. T. 2020 Water balance as an indicator of natural resource condition: case studies from great sand dunes national park and preserve. *Glob. Ecol. Conserv.* **24**, e01300.

Vitolo, C. 2015 *Exploring Data Mining for Hydrological Modelling. PhD Thesis*, Imperial College, London, UK.

First received 6 July 2022; accepted in revised form 16 October 2022. Available online 26 October 2022