

1 **Machine learning and deep learning based streamflow**
2 **prediction in a hilly catchment for future scenarios using**
3 **CMIP6-GCMs data**

4 Dharmaveer Singh^{1&5*}, Manu Vardhan², Rakesh Sahu³, Debrupa Chatterjee¹, Pankaj
5 Chauhan⁴, Shiyin Liu⁵

6 1. Symbiosis Institute of Geo-informatics, Symbiosis International (Deemed University), Pune-411004
7 (India).

8 2. Computer Science and Engineering Department, National Institute of Technology Raipur, Raipur-492010
9 (India)

10 3. Computer Science and Engineering Department, Chandigarh University, Mohali- 140413 (India)

11 4. Wadia Institute of Himalayan Geology, Dehradun-248001 (India)

12 5. Institute of International Rivers and Eco-security, Yunnan University, Kunming - 650091 (China)

13 *Correspondence to: veermnit@gmail.com and shiyin.liu@ynu.edu.cn

14

15

16 **Abstract**

17 The alteration in river flow patterns, particularly those that originate in the Himalayas, has been caused by the
18 increased temperature and rainfall variability brought on by climate change. Due to the impending
19 intensification of extreme climate events, as predicted by the Intergovernmental Panel on Climate Change
20 (IPCC) in its sixth assessment report, it is more essential than ever to predict changes in streamflow for future
21 periods. Despite the fact that some research has utilised machine learning and deep learning based models to
22 predict streamflow patterns in response to climate change, very few studies have been undertaken for a
23 mountainous catchment, with the number of studies for the western Himalaya being minimal. This study
24 investigates the capability of five different machine learning (ML) models and one deep learning (DL) model,
25 namely the Gaussian Linear Regression Model (GLM), Gaussian Generalized Additive Model (GAM),
26 Multivariate Adaptive Regression Splines (MARS), Artificial Neural Network (ANN), Random Forest (RF),
27 and 1D-Convolutional Neural Network (1D-CNN), in streamflow prediction over the Sutlej River Basin in the
28 western Himalaya during the periods 2041-2070 (2050s) and 2071-2100 (2080s). Bias corrected data
29 downscaled at grid resolution of $0.25^{\circ} \times 0.25^{\circ}$ from six General Circulation Models (GCMs) of the Coupled
30 Model Intercomparison Project Phase 6-GCMs framework under two greenhouse gas trajectories (SSP245 and
31 SSP585) were used for this purpose. Four different rainfall scenarios (R_0 , R_1 , R_2 , and R_3) were applied to the
32 models trained with daily data (1979-2009) at Kasol (the outlet of the basin) in order to better understand how
33 catchment size and the geo-hydro-morphological aspects of the basin affect runoff. The predictive power of
34 each model was assessed using six statistical measures: the coefficient of determination (R^2), the ratio of the root
35 mean square error to the standard deviation of the measured data (RSR), the mean absolute error (MAE), the
36 Kling-Gupta efficiency (KGE), the Nash-Sutcliffe efficiency (NSE), and the percent bias (PBIAS). RF model
37 with rainfall scenario R_3 which outperformed other models during the training ($R^2=0.90$; $RSR=0.32$; $KGE=0.87$;
38 $NSE=0.87$; $PBIAS=0.03$) and testing ($R^2=0.78$; $RSR=0.47$; $KGE=0.82$; $NSE=0.71$; $PBIAS=-0.31$) period
39 therefore was chosen to simulate streamflow in the Sutlej River in the 2050s and 2080s under the SSP245 and
40 SSP585 scenarios. Bias correction was further applied to the projected daily streamflow in order to generate
41 reliable times series of the discharge. The mean ensemble of model results show that the mean annual
42 streamflow of the Sutlej River is expected to rise between 2050s and 2080s by 0.79 to 1.43% for SSP585 and by
43 0.87 to 1.10% for SSP245. In addition, streamflow will increase during the monsoon (9.70 to 11.41% and 11.64
44 to 12.70%) in the 2050s and 2080s under both emission scenarios, but it will decrease during the pre-monsoon (-
45 10.36 to -6.12% and -10.0 to -9.13%) and post-monsoon (-1.23 to -0.22% and -5.59 to -2.83%), as well as
46 during the winter (-21.87 to -21.52% and -21.87 to -21.11%). This variability in streamflow is highly correlated
47 with the pattern of precipitation and temperature predicted by CMIP6-GCMs for future emission scenarios, as
48 well as with physical processes operating within the catchment. Predicted declines in Sutlej River streamflow
49 over the pre-monsoon (April to June) and winter (December to March) seasons might have a significant impact
50 on agriculture downstream of the river, which is already having problems due to water restrictions at this time of
51 year. The present study will therefore assist in strategy planning for ensuring the sustainable use of water
52 resources downstream by acquiring a knowledge of the nature and causes of unpredictable streamflow patterns.

53

54 **Keywords:** Machine learning models; 1D-CNN; streamflow; climate change; CMIP6-GCMs; western Himalaya

55

56 1 Introduction

57 Human-induced global warming has altered patterns of the rainfall worldwide (Goswami et al., 2006; Trenberth,
58 2011), and also increased risks of extreme events such as the droughts and floods (Easterling et al., 2000;
59 Trenberth et al., 2015; Otto et al., 2017). It has impacted hydrology of many river basins globally, including
60 variation in streamflow (Gerten et al., 2008; Nepal and Shrestha, 2015; Singh et al. 2015a; Ali et al., 2018; Lutz
61 et al., 2019; Singh et al., 2022). A study of long-term (1948-2004) streamflow (discharge) data of 200 largest
62 rivers of the globe showed considerable change in their annual discharge, however, results were statistically
63 significant only for 64 rivers (Dai et al., 2009). Out of which 45 were marked with decreasing trends and the
64 remaining 19 showed increasing trends in their annual discharge. Similar decreasing and increasing trends in
65 discharge of the rivers were reported also at regional scale: Asia (Kundzewicz et al., 2009; Krysanova et al.,
66 2015), Europe (Stahl et al., 2010; Stahl and Tallaksen, 2012) and America (Pasquini and Depetris, 2007).
67 Moreover, it has been established that the effects of rainfall variation and extreme events on annual discharge
68 are likely strong compared with other drivers (Kundzewicz et al., 2009; Miller et al., 2012; Van der Wiel et al.,
69 2019). Zhao et al. (2021) examined how precipitation, evapotranspiration, and timing of snowmelt impacted
70 runoff in the Kaidu River Basin of China. They discovered that as global warming increased, the timing of
71 snowmelt became less significant while the influence of precipitation increased comparatively. A projected rise
72 of $\sim 2^{\circ}\text{C}$ to 5°C in mean annual global temperature by 2100 under higher greenhouse gas emission scenarios as
73 predicted from the General Circulation Models (GCMs) (Gao et al., 2017) will considerably affect the rainfall
74 pattern (intensity and amount) and may alter hydrological cycles (Okai and Kanae 2006; Haddeland et al.,
75 2014). This would subsequently impact availability of water resources and present challenges for their
76 management since a rise in the demand of water is also predicted (Lutz et al., 2019). Therefore, it is
77 indispensable to know the underlying hydrological dynamics occurring within a basin in context of climate
78 change for effective management and sustainable use of the water resources.

79
80 The underlying hydrological processes controlling rainfall-runoff generation in a basin can be understood with
81 the use of a hydrological model which is based on complex mathematical equations and theoretical laws
82 governing physical processes in the basin (Kirchner, 2006; Singh et al., 2019). It simulates/or predicts response
83 of the basin to climatological forcings such as the rainfall (Sood and Smakhtin, 2015) and generates synthetic
84 time series of hydrological data that can be used by water managers and scientists for varied applications
85 ranging from water budgeting and partitioning (Conan et al., 2003; Schreiner-McGraw and Ajami, 2020) to
86 inundation mapping and modelling (Mahto et al., 2022). A hydrological model is supposed not only to have a
87 good predictive power but also the ability of capturing relationships among the forcing factors and catchment
88 response so that an accurate estimate of rainfall-runoff could be made (Shortridge et al., 2016). However, until
89 now, there is no hydrological model that can simulate basin-behaviour universally well against all the
90 hydrological challenges inflicted from climate change and human-interventions (Yang et al., 2019). As a result,
91 many hydrological models have been devised considering functioning and robustness of models in explaining
92 underlying complexity in quantifying basin-scale response to small-scale spatial complexity of physical
93 processes (Shortridge et al., 2016; Herath et al., 2021). Broadly, these can be grouped into two categories:
94 physical or process-based models and empirical or data-driven models (Yang et al., 2019; Kabir et al., 2020).

95 The latter category of models uses a mathematical relationship established between runoff and affecting factors
96 in the basin for deriving the runoff (Adnan et al., 2019).

97

98 It is purported that the data-driven model despite of inherited limitations over physical interpretability of
99 processes has outperformed the physical models in terms of prediction accuracy in many hydrological
100 applications (Shortridge et al., 2016; Adnan et al., 2019; Kabir et al., 2020; Herath et al., 2021). Also, they are
101 preferred over the physical models for rainfall-runoff modelling/or streamflow prediction modelling due to
102 limited requirements of data as inputs, where data limitation is the major challenge (Beven, 2011). These models
103 in past were heavily criticised on the ground of being incompetent to model the non-linear behaviour of
104 streamflow (Yang et al., 2019). But recent developments in computational intelligence, in the areas of machine
105 learning (ML) and deep learning (DL) in particular, have greatly expanded the capabilities of empirical
106 modelling (Adnan et al., 2020; Fu et al., 2020; Rahimzad et al., 2021; Ghobadi and Kang, 2022). This resulted
107 in the development of many non-linear models such as the Artificial Neural Network (ANN), Random Forest
108 (RF), Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) models, which can capture and
109 model non-stationarity of the rainfall-runoff relationships (Yaseen et al., 2015; Shortridge et al., 2016; Adnan et
110 al., 2019; Yang et al., 2019; Xiang et al., 2020). Yang et al. (2019) applied three machine learning models
111 namely ANN, SVR, and RF to predict monthly streamflow over the Qingliu River basin in China under
112 changing environmental conditions between 1989 and 2010, and compared their results with the six process-
113 based hydrological models. They concluded that the ML model performed better than the process-based model
114 not just in terms of prediction accuracy, but also in terms of flexibility when it came to including other runoff
115 effect factors into the model. Similar outcomes for Lake Tana and the adjacent rivers in Ethiopia were also
116 reported by Shortridge et al. (2016), where ML models demonstrated noticeably lower streamflow prediction
117 errors than the physical models developed for the region. However, they inferred that linear machine learning
118 models, such as the Multivariate Adaptive Regression Splines (MARS) and Generalized Additive Model
119 (GAM), were sensitive to extreme climate events, so the degree of uncertainty in their predictions needed to be
120 carefully considered.

121

122 The limitations of such data-driven models can be overcome by adopting more advanced ML and DL models
123 (Xiang et al., 2020). Rasouli et al. (2012) compared the performance of the Multi-Linear Regression (MLR)
124 model with the Bayesian Neural Network (BNN), SVR, and Gaussian process (GP) in terms of daily streamflow
125 prediction for the Stave River, a mountainous basin, in British Columbia, and found that the BNN model
126 performed better than others. According to Hussain and Khan (2020), supervised learning model RF
127 outperformed Multilayer Perceptron (MLP) and SVR in terms of accuracy while predicting monthly streamflow
128 for the Hunza river in Pakistan by 33.6% and 17.85%, respectively. Recently, Deep Neural Network (DNN),
129 Convolutional Neural Network (CNN) and LSTM models, which are based on deep learning, have seen a surge
130 in the number of streamflow prediction applications due to their abilities to handle complex stochastic datasets
131 and abstracting the internal physical mechanism (Fu et al., 2020; Ghobadi and Kang, 2022). Based on statistical
132 performance evaluation criteria, Rahimzad et al. (2021) found that the LSTM outperformed the LR, SVM, and
133 Multilayer Perceptron (MLP) models in daily streamflow prediction over the Kentucky River basin in the USA.
134 However, Van et al. (2020) showed that CNN outperformed LSTM in streamflow modelling in the Vietnamese

135 Mekong Delta by a small margin. Comparing data-driven models to a given problem yield a range of results for
136 distinct geographical and climatic conditions (Hagen et al., 2021. Adnana et al. (2020) examined the predictive
137 accuracy of Optimally Pruned Extreme Learning Machine (OP-ELM), Least Square Support Vector Machine
138 (LSSVM), MARS, and Model Tree (M5Tree) models in order to estimate monthly streamflow in the Swat River
139 Basin (Hindukush Himalaya), Pakistan. They came to the conclusion that the LSSVM and MARS are the most
140 effective at forecasting streamflow. In contrast, Hussain et al. (2020) discovered that ELM outperformed 1-D-
141 CNN while forecasting streamflow on three time scales i.e., daily, weekly and monthly in the Gilgit River,
142 Pakistan. This suggests that it is challenging to find a data-driven model that is effective across all application
143 domains and scales (Yaseen et al., 2015; Fu et al., 2020).

144

145 The use of machine learning and deep learning based models for streamflow simulations within catchments is
146 generally limited to observable periods and resulting forecasts (Eng and Wolock, 2022). There are very limited
147 studies worldwide where these models were applied in predicting long-term streamflow for future periods in
148 context of climate change (Das and Nanduri, 2018; Thapa et al., 2021; Adib and Harun, 2022). This can be
149 attributed to the challenges associated with data assimilation brought on by the use of coarse resolution scenario
150 data obtained from General Circulation Models (GCMs), which limits their direct application in regional impact
151 assessment (Hagen et al., 2021; Adib and Harun, 2022). Das and Nanduri (2018) integrated Relevance Vector
152 Machine (RVM) and SVM models with Coupled Model Intercomparison Project Phase (CMIP5)-GCMs to
153 project monthly monsoon streamflow across the Wainganga basin (India) for monsoon season. Adib and Harun
154 (2022) studied variations in the monthly streamflow pattern of the Kurau River (Malaysia) from 2021 to 2080
155 by coupling ML models (RF and SVR) with Coupled Model Intercomparison Project Phase (CMIP6)-GCMs.
156 Despite of the significance potentials of the ML and DL models in streamflow prediction, relevant studies
157 assessing the application of these models for streamflow prediction under future scenarios over the mountainous
158 basins are limited due to non-availability of long-term data (Xenarios et al., 2019; Adnana et al., 2020). Thapa et
159 al. (2021) used a combination of the LSTM model and the CMIP5-GCMs scenarios to estimate streamflow
160 patterns in the Langtang basin of the Central Himalayas. Their analyses revealed a notable increase in
161 streamflow as a result of the predicted increase in precipitation. The projections from Coupled Model
162 Intercomparison Project Phase 3 (CMIP3)-GCMs and CMIP5-GCMs inherit limitations in simulating extreme
163 precipitation (Kim et al., 2020), which are the principal drivers for the runoff generation in the catchment. This
164 causes large uncertainty in streamflow predictions (Wang et al., 2021). Uncertainty in streamflow prediction can
165 be minimised by using scenarios from the CMIP6-GCMs which are likely to be more realistic than previous
166 generations, i.e., CMIP3-GCMs and CMIP5-GCMs, given their significant improvement in simulating rainfall
167 and temperature for historical records (Chen et al., 2020; Gusain et al., 2020; Kim et al., 2020). Therefore,
168 projected changes in streamflow patterns derived from CMIP6-GCMs scenarios would give a better
169 understanding of the catchment's future hydrological regime than previous ones. To the authors' knowledge, no
170 work has been published over a mountainous basin that integrates ML/DL models with CMIP6-GCMs scenarios
171 to predict changes in streamflow patterns for future periods. Hence, it is important to test whether machine
172 learning approaches can be effectively used over a mountainous river basin to predict streamflow using hydro-
173 meteorological variables and CMIP6-GCMs scenarios as the input data.

174

175 With a catchment area of 56874km² (up to Bhakara Dam), the Sutlej also pronounced as ‘Satluj’ is an important
176 river in the western Himalayas and runs through diverse climatic zones. The flow in the upper and middle
177 catchment is primarily impacted by glacier/snow melt induced by seasonal temperature shift and preceding
178 winter precipitation, while the lower section of the catchment area is mostly regulated by rainfall both in the
179 winter and during the monsoon season (Singh and Jain, 2002; Archer, 2003; Miller et al., 2012). Based on data
180 from the period 1986–1996, Singh and Jain (2002) estimated the mean yearly contribution of snow/glacier melt
181 and rainfall to the Sutlej River as being 59% and 41%, respectively. However, the discharge in the river peaks is
182 directly related to the peak in rainfall during the monsoon (Lutz et al., 2014). Recent studies on this basin has
183 raised concerns about the implications of climatic changes on streamflow since a warming climate has brought
184 changes in the amount and spatial-temporal distribution of precipitation (Singh et al., 2014; Singh et al., 2015b).
185 Previous research has only used process-based hydrological models and scenarios from CMIP3-GCMs and
186 CMIP5-GCMs to date when examining the effects of climate change (past and future) on streamflow patterns in
187 the region (Singh and Jain, 2002; Singh et al., 2015a; Ali et al., 2018; Shukla et al., 2021), which leaves a gap in
188 the use of machine and deep learning models and scenarios from the latest CMIP6-GCMs. This study very first
189 time examines the potential of five ML models and one deep learning model namely, Gaussian Linear
190 Regression Model (GLM), Gaussian Generalized Additive Model (GAM), MARS, ANN, RF and 1D-CNN in
191 streamflow prediction over the middle Sutlej River Basin (rainfall dominated zone) in western Himalaya using
192 different Shared Socio-economic Pathways (SSPs) scenarios from CMIP6-GCMs. The pattern of variations in
193 the Sutlej River's monthly, seasonal, and annual streamflow are assessed for the future periods 2041-2070
194 (2050s) and 2071-2100 (2080s) with respect to the reference period of 1979-2009 under SSP245 and SSP585. .
195 The findings of the study will help to develop a better plan for the operation of hydroelectric power projects and
196 water resources management in the catchment.

197 2 Study Area

198 The selected study area is a sub-catchment within the Satluj basin (Figure 1), with an area of 2457 km².
199 Topographically, it is very rugged (0-80°) and is dominated mostly by forests (56.20%), grassland (26.4%),
200 agricultural lands (17.1%), and glaciers and snow covers (0.3%) (Singh et al., 2015a). The presence of mountain
201 barriers in the sub-basin's north, large variation in altitudes (500–5000 m) and the aspect all contribute to the
202 region's diverse climate. It varies from hot and moist tropical climate in lower valleys to cool temperate climate
203 at about 2000 m, and tends towards alpine as the altitude increases beyond 2000 m. The mean annual discharge
204 (averaged over the period of 1979-2009) of the river gauged at Kasol was 12469.43 m³/s. There is large inter-
205 diurnal and monthly variation in pattern of the river discharge. The minimum and maximum daily discharge
206 recorded at Kasol was 64.30 m³/s and 2891m³/s, respectively. The early months of year, i.e., starting from
207 January up to March are characterised by low stream flow. After this a continuous and rapid rise in flow occurs,
208 being the maximum in the month of July (~22-23%). Then, it again starts decreasing and flow becomes the
209 minimum in the month of December (2-3%). The details of the sub-catchment are summarised in Table 1.

210 Figure 1: The location of the sub-catchment within Sutlej River Basin. The three hydro-meteorological stations
211 (Kasol, Sunni and Rampur) from which this study employed observed data for the years 1979 to 2009 are also
212 shown.

213 The sub-basin is bestowed with the large hydropower potential. There are three major hydroelectric power
214 projects: Sunni Dam Project of 1080 MW, Rampur Hydroelectric Power Project (RHEP) of 412 MW, and
215 Nathpa Jhakari Hydro-electric Power Project (NJHEP) of 1500 MW. The sub-basin is climatologically sensitive
216 and, at present, facing the challenges created due to climate change and human's interventions (Singh et al.,
217 2015b and 2015c). Change in future climate will alter patterns of flow in river and further could affect water
218 resources and hydroelectric power production (Singh et al., 2014).

219 Table 1: Characteristics of the study catchment over the evaluation period of 1979–2009.

220 **3 Description of the Data and Methods**

221 The methodology involved in predicting streamflow for the period 2041-2100 in the Sutlej River include: 3.1)
222 collection of hydro-meteorological data, 3.2) selection of machine and deep learning models, 3.3) performance
223 evaluation of the developed models, and 3.4) bias correction in streamflow projection. These are described in
224 details under following sub-headings:

225 **3.1 Hydro-meteorological data**

226 The daily rainfall, temperature (T_{\max} and T_{\min}), relative humidity, solar radiation, wind speed and discharge data
227 used to study performance of the different machine and deep learning models on streamflow modelling were
228 collected for 31 years i.e. 1979-2009. Rainfall, temperature and discharge data were obtained from the Bhakara
229 Beas Management Board (BBMB), while relative humidity, solar radiation and wind data were extracted from
230 the Global Weather Data (<http://globalweather.tamu.edu/>). These data were collected for three hydro-
231 meteorological stations namely, Kasol, Sunni and Rampur (Fig.1).

232

233 The downscaled outputs from the CMIP6-GCMs, the latest generation of climate models, were used for
234 streamflow prediction in future (2050s and 2080s). This framework of CMIP6-GCMs was run to simulate future
235 climate under four Shared Socio-economic Pathways Scenarios (SSPs), which are designed to explain potential
236 future greenhouse gas emissions under various global socioeconomic shifts that would occur by 2100 (Riahi et
237 al., 2017; Karan et., 2022). Even by using downscaled outputs, however, regional climate change projections
238 inherit biases from the GCM boundary conditions (Jose and Dwarakish, 2022), which were corrected in the
239 dataset detailed in Mishra et al. (2020) for South Asia. They used Empirical Quantile Mapping (EQM) method
240 for removing bias in the downscaled data. This dataset provides bias-corrected downscaled climate change
241 projections for 13 CMIP6-GCMs and four GHG emission scenarios (SSP126, SSP245, SSP370, and SSP585),
242 the latter are briefly summarised in Riahi et al. (2017). Climate projections from CMIP6-GCMs that have been
243 generated under the SSP245 and SSP585 scenarios were used in this study. SSP245, a medium scenario
244 represents the average pathway of future greenhouse gas emissions with radiative forcing of 4.5 W/m^2 by the
245 year 2100, while SSP585 is the upper limit of the range of scenarios scenario with radiative forcing of 8.5 W/m^2
246 by the end of this century (O'Neill et al., 2016). The data are available at a daily time-scale and horizontal spatial
247 resolution of $0.25^\circ \times 0.25^\circ$. Seven grids of the downscaled CMIP6-GCMs data cover the study area. The
248 temperature (T_{\max} and T_{\min}) data were adjusted for topographical bias by separating the study area into a number
249 of homogenous elevation bands spaced by at an interval of 1000m, and applying a temperature laps rate of

250 6.5°C/1000m within each grid. A Digital Elevation Model (DEM) of 30 m spatial resolution derived from
251 CartoSat-1 stereo data (www.bhuvan.nrsc.gov.in) was used for this purpose. The values of rainfall and
252 temperature at each grid were then averaged over the catchment using the Thiessen polygon method in order to
253 provide daily rainfall data integrated at the catchment scale for assessing changes in the future climate with
254 respect to the observed period i.e., 1979-2009.

255

256 Further, ranking of CMIP6-GCMs was done to find out the most appropriate models that can generate most
257 likely plausible scenarios of future climate in the catchment and ultimately being employed in streamflow
258 projection. Taylor diagram (Taylor, 2001), a robust graphical plot, is widely used to rank GCMs due to its
259 effectiveness in determining the relative strengths of the competing models and in evaluating overall
260 performance as a model evolves (Abbasian et al., 2019; Ghimire et al., 2021). It integrates three statistical
261 metrics, degree of correlation (r), centered root-mean-square error (CRMSE) and ratio of spatial standard
262 deviation (SD). Combining these metrics allows determining the degree of pattern correspondence and
263 explaining how exactly a model represents the observed climate (Taylor, 2001). Therefore, performance of 13
264 CMIP6-GCMs in modelling climatic variables (rainfall, T_{\max} and T_{\min}) in the Sutlej sub-basin was compared to
265 the observed data (1979-2009) using Taylor diagram (Fig. 2a-c). The models were then ranked as a result of this
266 comparison. High positive correlation ($r=0.84$ to 0.96) and low CRMSE ($<3^{\circ}\text{C}$) error were found in all 13
267 CMIP6-GCMs for temperature (T_{\max} and T_{\min}) (Fig. 2b-c). Additionally, it was found that models' standard
268 deviations, which ranged from 5.60 to 6.03°C for T_{\max} and 6.34 to 6.63°C for T_{\min} , were close to the SD of the
269 observed data (6.01°C and 6.07°C). These results imply that all CMIP6-GCMs may be able to predict most
270 likely future temperature over the catchment.

271 Figure 2: Taylor diagram showing comparative skills of 13CMIP6-GCMs in simulating climatic variables
272 (rainfall, T_{\max} and T_{\min}) over the Sutlej sub-basin during reference period (1979-2009). The degree of correlation
273 coefficient (r) between observed and CMIP6-GCMs, centered root-mean-square error (CRMSE) and departure
274 of the models' standard deviation (SD) from the observed data (dashed black arc line) are shown in Fig. 2a for
275 rainfall, Fig. 2b for T_{\max} and Fig. 2c for T_{\min} . The units of SD for rainfall and temperature is in cm and $^{\circ}\text{C}$,
276 respectively.

277 However, not all CMIP6-GCMs showed the high degree of similarity in predicting rainfall; in fact, two
278 (CanESM5 and NorESM2-LR) of the 13 models revealed a negative correlation (Fig. 2a). In the pool of 13
279 CMIP6-GCMs, only six models showed relatively higher correlation ($r\geq 0.56$), smaller CRMSE (<12 cm) errors,
280 and a high similarity to the standard deviation of the observed data (13.2 cm). They were: 1) Earth Consortium-
281 Earth 3 Veg Model (EC-Earth-Veg) , 2) Russian Institute for Numerical Mathematics Climate Model Version
282 4.8 (INM-CM4-8), 3) Russian Institute for Numerical Mathematics Climate Model Version 5.0 (INM-CM5-0),
283 4) Max Planck Institute for Meteorology Earth System Model version 1.2 with higher resolution (MPI-ESM1-2-
284 HR) , 5) Max Planck Institute for Meteorology Earth System Model version 1.2 with lower resolution (MPI-
285 ESM1-2-LR) and 6) Norwegian Earth System Model Version 2 with Medium Resolution (NorESM2-MR).
286 Further, within these models, the highest and lowest correlations between observed and simulated rainfall were
287 found for the INM-CM4-8 ($r=0.69$) and NorESM2-MR ($r=0.56$), respectively. These six CMIP6-GCMs were
288 finally selected to examine future patterns in streamflow for the periods 2050s and 2080s in the Sutlej River
289 Basin as they had also shown high performance in simulating temperatures ($r=0.90$ to 0.96).

290 3.2 Selection of machine and deep learning models for streamflow modelling

291 In this study, five machine and one deep learning models namely GLM, GAM, MARS, ANN, RF and one
292 dimensional Convolution Neural Network (1D-CNN) were selected and their performances in predicting
293 streamflow in Sutlej River were compared. These are regression based models which capture relationship
294 between the predictors (dependent variables) and predictand (independent variables) and provide value of the
295 output variables (Adnan et al., 2019; Kabir et al., 2020). The models were trained with daily observed data
296 recorded during 1979-2009 at Kasol (the gauging site) as well as simulated historical projections of CMIP6-
297 GCMs. The climatic projections of the grid corresponding to Kasol station were taken into consideration as the
298 input from the CMIP6-GCMs. However, prior to building the models, all of the data were normalized using
299 standard normalization techniques to get features on a common scale. Further, the entire data set was split into
300 training and testing datasets since a cross-validation method was adopted in this study. The training dataset
301 (80%) was used for fitting the models whereas testing dataset was used for checking model accuracy (20%).
302 Under the cross-validation method, the process was repeated until every part of the allocated data was used in
303 testing (Kabir et al., 2020). Six different program codes were written in python language for ANN, GAM, GLM,
304 MARS, RF and 1D-CNN simulations. Out of these six selected models, GLM, GAM and MARS are linear
305 models whereas other three i.e. ANN, RF and 1D-CNN are non-linear in nature (Shortridge et al., 2016; Yang et
306 al., 2019; Herath et al., 2021). Additionally, excluding GLM all of the remaining models are based on non-
307 parametric regression approach where functional relationship between predictor and predictand are not
308 predetermined but can be adjusted to capture unusual or unexpected features of the data (Shortridge et al.,
309 2016). A detailed description of these models can be found elsewhere (Shortridge et al., 2016; Adnan, 2019;
310 Yang et al., 2019; Kabir et al., 2020; Ghimire et al., 2021; Herath et al., 2021; Shu et al.,2021).

311
312 Since the 1D-CNN model is based on weight sharing, it needs less training parameters than other models
313 (Kiranyaz et al., 2021). It has mainly three layer, convolution layer, pooling layer and fully connected layer. The
314 primary job of the convolution layer is to nonlinearly map input data into a set of feature maps, or series of
315 feature vectors. When working as a visual cortical perceptron, filter kernels are convoluted with the input data of
316 their receptive fields. The convolution results with biases are then passed on to the activation function to create
317 feature maps. The pooling layer, which comes after each convolution layer, primarily serves to reduce the
318 dimension of feature maps and maintain the invariance of characteristic scale. The fully connected layer uses a
319 completely connected single layer perceptron to combine the feature maps that were acquired by the prior
320 convolution and pooling layers in order to build a higher level feature (Kiranyaz et al., 2021). In this study, one
321 convolution layer with 64 filters, a kernel of size 2, and a ReLU activation function was being employed. This
322 was followed by max pooling layer with pool size =2, and the falterm layer. After that two fully connected
323 layer applied with ReLU activation function and linear activation function, respectively. However, for
324 optimization, the adaptive moment estimation (Adam) algorithm was applied (Ghimire et al., 2021; Shu et
325 al.,2021). Six variables namely rainfall, T_{max} , T_{min} , relative humidity, solar radiation and wind speed were used
326 as the inputs for developing the models. Additionally, these models were simulated under four rainfall scenarios:
327 rainfall on the same day (R_0), rainfall lagged by one day (R_1) and rainfall lagged by two days (R_2) and rainfall
328 lagged by three days (R_3) to understand control of catchment size and geo-hydro-morphological characteristics

329 of the basin in generating runoff. While, remaining meteorological parameters were held constant during the
 330 processes.

331 3.3 Model performance evaluation

332 It has been found that overfitting in a model may lead to large errors in out-of-sample predictions (Hastie et al.,
 333 2009). Therefore, it has been evaded by establishing model parameters for GLM, GAM, MARS, ANN and RF
 334 through automated hyperparameter tuning methods. 500 bootstrap resamples of the training data set were
 335 generated for each parameter value to be assessed. Table 2 presents the information on the specific parameters
 336 evaluated for each model.

337 Table 2: The information on hyper parameters used for estimating model parameters.

338 The accuracy with which the simulated flow matches the observed flow during the training (calibration) and
 339 testing (validation) phases determines whether a hydrological model is appropriate for a given application
 340 (Refsgaard, 1997). Several methods, including quantitative statistics and graphical methods, has been developed
 341 in the past for assessing the accuracy of model predictions (Legates and McCabe, 1999). Moriasi et al. (2007)
 342 grouped these methods into three categories namely, standard regression, dimensionless, and error index,
 343 depending on how well each method explains the relationship between observed and simulated values, compares
 344 the relative performance of models, and quantifies the deviation in the units of the data of interest. Moreover, it
 345 has been established from previous studies that a single metric is inadequate to evaluate a model's performance,
 346 hence multiple metrics should be used (Adnan et al., 2020). Therefore, in this study, prediction accuracy of
 347 different models was compared using six statistical measures out of which one was standard regression
 348 (coefficient of determination (R^2)), two of which were dimensionless (Kling-Gupta efficiency (KGE) and Nash-
 349 Sutcliffe efficiency (NSE)), and the remaining three were being error index (ratio of the root mean square error
 350 to the standard deviation of the measured data (RSR)), the mean absolute error (MAE) and the percent bias
 351 (PBIAS)). These metrics are defined below by the equations (2-7):

352

$$R^2 \text{ (Van Liew et al., 2003)} = \left(\frac{\sum_{i=1}^n (Q_i - \bar{Q})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2} \times \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right) \quad (\text{range: 0 to 1}) \quad (1)$$

353

$$KGE \text{ (Gupta et al., 2009)} = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_p}{\sigma_{ob}} - 1\right)^2 + \left(\frac{P_i}{Q_i} - 1\right)^2} \quad (\text{range: 0 to 1}) \quad (2)$$

354

$$NSE \text{ (Nash and Sutcliffe, 1970)} = 1 - \left[\frac{\sum_{i=1}^n (Q_i - P_i)^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2} \right] \quad (\text{range: } -\infty \text{ to } 1) \quad (3)$$

355

$$RSR \text{ (Singh et al., 2004)} = \frac{\sqrt{\sum_{i=1}^n (Q_i - P_i)^2}}{\sigma_{ob}} \quad (\text{range: 0 to } \infty) \quad (4)$$

356

$$MAE \text{ (Adnan et al., 2020)} = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad (\text{range: 0 to } \infty) \quad (5)$$

357

$$PBIAS \text{ (Gupta et al., 1999)} = \left[\frac{\sum_{i=1}^n (Q_i - P_i)}{\sum_{i=1}^n (Q_i)} \right] \times 100 \quad (\text{range: } -100 \text{ to } 100\%) \quad (6)$$

358

359 where P_i are the predicted values and Q_i are the observed values, n accounts for the number of samples, Q^-
 360 represents the mean of observed data, and P^- is the mean of predicted data. However, r is the Pearson's
 361 correlation coefficient whereas σ_{ob} and σ_p refers to the standard deviation of observed and predicted values,
 362 respectively.

363

364 R^2 evaluates the percentage of the variation in the measured data that can be explained by the model, whereas
 365 NSE estimates the relative size of the residual variance in relation to the variance in the measured data (Nash
 366 and Sutcliffe, 1970; Van Liew et al.,2003). According to Mazrooei et al. (2021), NSE is sensitive to extreme
 367 flows; as a result, KGE is also used to evaluate a model's performance while considering extreme flows into
 368 account (Adib and Harun, 2022). Other metrics, like RSR, MAE, and PBIAS, shed light on the overall
 369 inaccuracies in the projected flow relative to the observed. The value of R^2 , KGE and NSE should all be 1 in an
 370 ideal model, whereas RSR and MAE and PBIAS values should be 0 (Nash and Sutcliffe,1970; Van Liew et
 371 al.,2003; Gupta et al.,2009; Adnan et al., 2020). Moriasi et al. (2007) developed a guideline for interpreting the
 372 results of these metrics and ranking for the hydrological models based on a thorough review of the available
 373 literature. They found that a model can be classified as very good, good, satisfactory, or unsatisfactory if its
 374 NSE value is between 0.75 and 1, 0.65 to 0.75, 0.50 to 0.65, or less than 0.50, respectively. Similarly, R^2 values
 375 between 0.6 to 0.7 are considered satisfactory, 0.85 to 1 are very good and below 0.5 are unsatisfactory (Van
 376 Liew et al., 2003). However, for RSR, numbers above 0.7 are considered to be poor, whereas values between 0
 377 and 0.5 are considered to be in the very good range. Thus, the lower is the RSR value, the better is the model.
 378 This is also true for PBIAS and MAE where lower values are favourable. According to Moriasi et al. (2007),
 379 PBIAS values of less than $\pm 10\%$ are considered to be highly acceptable, whilst values of more than $\pm 25\%$ are
 380 considered to be unsatisfactory. The negative number indicates that the model has overestimated its bias,
 381 whereas the positive value indicates that the model has underestimated its bias (Gupta et al., 1999).

382 3.4 Bias correction

383 Uncertainty in streamflow prediction may be caused by the GCMs' shortcomings (e.g., coarse spatial resolution,
 384 simplified physics and thermodynamic processes, numerical methods, or poor knowledge of climate system
 385 dynamics) in accurately replicating natural climate variability (Sperna Weiland et al., 2010). As a result, its
 386 quantification and correction are critical for generating a future time series of streamflow that is reliable and
 387 recommended to devising water resource management plans in the catchment. This study used the bias
 388 correction method proposed in Hawkins et al. (2013) to correct uncertainty (bias) between observed and
 389 CMIP6-GCMs predicted streamflow. The mathematical expression for this formulae is given below:

$$390 \quad Q_{bc} = \bar{Q}_{ob} + \frac{\sigma_{ob}}{\sigma_p} (Q_{future} - \bar{Q}_p) \quad (7)$$

391 where, Q_{bc} and Q_{future} is the bias corrected and raw daily discharge for future simulation, respectively. \bar{Q}_{ob}
 392 and \bar{Q}_p is the mean discharge of observed and historical simulation for reference period (1979-2009),
 393 respectively. σ_o and σ_p is the standard deviation in observed and historical simulation for reference period,
 394 respectively. This method captures variability in both observation and GCMs simulations Hawkins et al. (2013),
 395 which is the interest of this study.

396 4 Results

397 4.1 Streamflow simulation and evaluation of model performance

398 The simulation (1979-2009) results generated under different rainfall scenarios (R_0 , R_1 , R_2 and R_3) on daily time
399 scale for all six models (GLM, GAM, MARS, ANN, RF and 1D-CNN) during training and testing is shown in
400 Fig. 3 and Fig. 4, respectively. The model performed slightly better during training than testing periods. R^2 , NSE
401 and KGE values across models ranged from 0.69 to 0.90, 0.52 to 0.87, 0.69 to 0.91 and from 0.69 to 0.81, 0.49
402 to 0.74 and 0.68 to 0.82 during training and testing, respectively. Likewise, it was found that RSR, MAE and
403 PBIAS varied from 0.31 to 0.55, from 71.95 to 123.25 m^3/s and -2.11 to +4.31% during training, as well as from
404 0.56 to 0.46, from 123.06 to 106.64 m^3/s and -3.74 to +2.21% during testing, respectively. Non-linear models
405 (ANN, 1D-CNN and RF) outperformed linear models (GAM and GLM) in runoff prediction under all rainfall
406 scenarios (R_0 , R_1 , R_2 , and R_3), with the exception of MARS, which produced results that were more or less
407 comparable to those of the ANN model. Figures 3–4 show that both models (RF and 1D-CNN) satisfy the
408 performance requirements outlined by Moriasi et al. (2007) as the best models, but RF slightly outperformed
409 CNN in terms of error index. R^2 , NSE, KGE, RSR, and MAE and PBIAS values for the RF model during the
410 training ranged from 0.88 to 0.90, 0.85 to 0.87, 0.86 to 0.87, 0.32 to 0.34, 71.95 to 77.49 m^3/s and +0.03 to
411 +0.13%, respectively. For the 1D-CNN, however, it varied from 0.87 to 0.89, 0.85 to 0.87, 0.90 to 0.91, 0.34 to
412 0.35, 80.29 to 83.14 m^3/s , and -1.25 to +0.13%. Similar pattern with slightly lower values were revealed during
413 testing for the both models. This implies that RF can effectively capture non-linear interactions and can provide
414 insights about actual watershed functions (Shortridge et al., 2016). On the other hand, GLM showed the poorest
415 results. R^2 , NSE, KGE, RSR, MAE, and PBIAS values for the GLM model during the training varied from 0.69
416 to 0.71, 0.52 to 0.56, 0.71 to 0.72, 0.54 to 0.55, 134.80 to 140.56 m^3/s , and +2.63 to +2.73%, respectively.
417 During testing, they varied between 0.69 and 0.71, 0.49 and 0.54, 0.68 and 0.70, 0.54 and 0.56, 134.35 and
418 141.26 m^3/s , +1 and +1.31%, respectively. Furthermore, it was observed that the models with rainfall scenario
419 R_3 had revealed reasonably better results in comparison to R_0 , R_1 and R_2 scenarios, indicating delayed
420 contribution of rainfall-runoff to the river.

421 Figure 3: Evaluation of the model (ANN, 1D-CNN, GAM, GLM, MARS and RF) performance in simulating
422 streamflow under rainfall scenarios R_0 (Fig.3a), R_1 (Fig. 3b), R_2 (Fig.3c) and (Fig. 3d) R_3 at Kasol during
423 training phase using six statistical metrics (R^2 , KGE, NSE, RSR, MAE and PBIAS).

424 Figure 4: Evaluation of the model (ANN, 1D-CNN, GAM, GLM, MARS and RF) performance in simulating
425 streamflow under rainfall scenarios R_0 (Fig.4a), R_1 (Fig.4b), R_2 (Fig.4c) and (Fig.4d) R_3 at Kasol during testing
426 phase using six statistical metrics (R^2 , KGE, NSE, RSR, MAE and PBIAS).

427 Figure 5, 6, 7 and 8 shows comparison of observed and simulated streamflow under rainfall scenarios of R_0 , R_1 ,
428 R_2 and R_3 for all the models at Kasol, the outlet of the basin. As observed from the Figures (5-8), RF was able to
429 follow the curve better compared to the other models. It is also deduced from the comparison of scatter plots
430 wherein a relatively smaller deviation in the observed and estimated discharge of streamflow was found for the
431 RF model. GLM performed the worst out of the six models with respect to the time variation graphs. A
432 limitation faced by all the six models was the simulation of peak values. The models slightly underperformed at
433 the prediction of higher values of streamflow. These findings led to the ultimate decision to use the RF model

434 with rainfall scenario R_3 to predict streamflow in the Sutlej River in the future (2050s and 2080s) under the
435 SSP245 and SSP585 scenarios.

436 Figure 5: Comparison of observed and simulated streamflow for all six models (ANN, 1D-CNN, GAM, GLM,
437 MARS and RF) under rainfall scenarios R_0

438 Figure 6: Comparison of observed and simulated streamflow for all five models (ANN, 1D-CNN, GAM, GLM,
439 MARS and RF) under rainfall scenarios R_1

440 Figure 7: Comparison of observed and simulated streamflow for all five models (ANN, 1D-CNN, GAM, GLM,
441 MARS and RF) under rainfall scenarios R_2

442 Figure 8: Comparison of observed and simulated streamflow for all five models (ANN, 1D-CNN, GAM, GLM,
443 MARS and RF) under rainfall scenarios R_3 .

444 **4.2 Comparison of streamflow simulated with observed and CMIP6-GCMs data**

445 The uncertainty between observed and CMIP6-GCMs predicted streamflow during the reference period (1979-
446 2009) was investigated by comparing the streamflow simulated by RF model with observed and CMIP6-GCMs
447 data. A large difference in streamflow patterns was seen in the box-plot of observed and CMIP6-GCMs
448 simulated discharge (Fig. 9) derived for various months of the year, particularly from June through September
449 (monsoon season), when a pattern of intense daily rainfall was observed over the catchment. Additionally, it was
450 discovered through the analysis of probability exceedance curves generated using 10% of the time series' highest
451 flows that, despite the streamflow's in the two data sets being comparable throughout the pre-monsoon season
452 (Fig. 10c), they differ noticeably for high flows during the annual (Fig.10a) and monsoon season (Fig.10c).
453 Similar trends were seen in the comparison of the probability exceedance curves for low flows during the
454 monsoon season, although there was strong agreement for annual (Fig.10b) and pre-monsoon measurements
455 (Fig.10d). This may be due to the fact that orography has a considerable impact on regional Indian Summer
456 Monsoon (ISM) climate, making it challenging for climate models to predict daily monsoonal rainfall
457 accurately across the Himalaya (Turner and Annamalai, 2012; Niu et al., 2015; Choudhary et al., 2017). The
458 Regional Climate Model (RCM) based on CMIP5-GCMs was used by Sanjay et al. (2017) to study pattern of
459 change in precipitation and temperature over the HKH region. As a condition of the model's inability to
460 accurately represent complicated feedback mechanisms, the results revealed large uncertainty in the summer and
461 winter precipitation over the northwest Himalaya. This is also supported by the study of Kadel et al. (2018).
462 They evaluated the performance of 38 CMIP5-GCMs in simulating rainfall over the central Himalaya and came
463 to the conclusion that the majority of the models' studied performed poorly when it comes to reproducing the
464 spatial distribution of monsoonal rainfall. Although the most recent study by Gusain et al. (2020) in India
465 reported that ISM simulation using CMIP6-GCMs over CMIP5-GCMS had significantly improved, there are
466 discrepancies between the models and indicated uncertainty in predictions. Lalande et al. (2021) examined the
467 abilities of 26 CMIP6-GCMs to simulate the rate of precipitation across the Himalayan region and concluded
468 that the models consistently overestimated the rate of precipitation by 31% to 281%. Additionally, cold-bias in
469 temperature estimation was also reported. Therefore, bias correction as described in Section 3.4 was applied to
470 the projected streamflow for the future periods (2050s and 2080s) under all scenarios and for all six models in
471 order to provide accurate times series of the discharge.

472 Figure 9: Box-plot comparing observed and CMIP6-GCMs (mean ensemble of models) simulated streamflow
473 for various months of the year, derived over the period of 1979–2009. The line inside the box denotes the
474 median values of streamflow, while the upper and lower whiskers indicate the highest and minimum values,
475 respectively.

476 Figure 10: Probability exceedance curves developed using 10% of the highest and lowest flows from the
477 observed and CMIP6-GCMs (mean ensemble of models) over the time span of 1979–2009 for annual and
478 seasonal (pre-monsoon and monsoon) flows.

479 **4.3 Projected change in rainfall and temperatures in 2050s and 2080s under SSP245 and SSP585**

480 Figure 11 shows how the catchment's mean monthly rainfall is expected to change under SSP245 and SSP585 in
481 the 2050s and 2080s compared to the reference period (1979-2009). Within months and for the CMIP6-GCMs,
482 a sizable shift in the rainfall pattern is seen. With the exception of March, June, and September, the mean
483 ensemble of the models generally predicts a rise in rainfall throughout the year in the 2050s and 2080s under all
484 scenarios. The models also show significant variation in the seasonal and yearly rainfall patterns expected for
485 the catchment in the 2050s and 2080s under various emission scenarios. However, based on the mean ensemble
486 of the models, it is predicted that seasonal (Fig. 12) and annual (Fig. 13a) rainfall will increase generally in the
487 2050s and 2080s under SSP245 and SSP585. Pre-monsoon, monsoon, post-monsoon, and winter rainfall in
488 2050s will increase by 8.75 to 8.85%, 10 to 20.80%, 85 to 91.91%, and 12.48 to 14.16%, respectively, under
489 SSP245 and SSP585. However, under SSP245 and SSP585 in the 2080s, it will rise by 7.69 to 17.50%, 21.52 to
490 41.43%, 56.16 to 89.66%, and 22.48 to 12.43%, respectively. Under both scenarios in the 2050s and 2080s, pre-
491 monsoon and post-monsoon will have the lowest and highest percentage increases in rainfall, respectively. The
492 monsoon season, however, is anticipated to have the greatest rise in terms of quantity (~40-167mm). The
493 predicted range for the increase in mean annual rainfall is 13.85 to 18.61% in the 2050s and 17.91% to 34.31%
494 in the 2080s. It is observed that the predicted pattern of change in rainfall across the sub-basin under various
495 SSPs is consistent in terms of the direction of change with other studies conducted over the Sutlej and Himalaya
496 region. Lalande et al. (2021) reported an overall increase in mean annual precipitation over the Himalayan
497 region based on 10 CMIP6-GCMs. According to their analysis, the mean ensemble of model precipitation is
498 predicted to increase by 8.6% to 25.4% in 2081-2100 under SSP245 and SSP585. The same study also showed
499 an increase in the region's winter (November to April) and ISM (June to September) rainfall. This contradicts
500 past studies that showed a trend toward declining ISM rainfall after the 1950s (Sabin et al., 2020). They
501 postulated that the region's higher winter rainfall would have been caused by the strengthening of the western
502 disturbances; however, the intensification of the ISM is responsible for the region's enhanced summer rainfall.
503 Figure 11: Projected change in mean monthly rainfall in the sub-basin using different CMIP6-GCMs under
504 SSP245 and SSP585 scenarios in the 2050s (Fig.11a and Fig.12b) and 2080s (Fig.12c and Fig.12d).

505 Figure 12: Projected change in mean seasonal rainfall in the sub-basin using different CMIP6-GCMs under
506 SSP245 and SSP585 scenarios in the 2050s (Fig.12a and Fig.12c) and 2080s (Fig.12b and Fig.12d).

507 Figure 13: Projected change in mean annual rainfall (Fig.13a), T_{\max} (Fig.13b) and T_{\min} (Fig.13c) in the sub-basin
508 using different CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s and 2080s.

509 The analysis of the CMIP6-GCM projections leads to the conclusion that for all months and seasons in the
510 2050s and 2080s, maximum (excluding April and pre-monsoon in 2050s under SSP245) and minimum
511 temperatures will rise under both scenarios (Fig. 14 (a-d) and Fig.15 (a-d)). Similarly, increase in mean annual
512 T_{min} and T_{max} are also predicted in 2050s and 2080s under all scenarios (Fig.13b and 13c). The increase will be
513 relatively higher for the T_{min} as compared to the T_{max} . This is also reported by Singh et al. (2015c). The increase
514 in rainfall and temperature is typically higher under SSP585 than SSP245 in both eras (2050s and 2080s), as
515 expected, due to a larger increase in radiative forcing brought on by increased greenhouse gas emissions.

516 Figure 14: Projected change in mean seasonal maximum temperature (T_{max}) in the sub-basin using different
517 CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s (Fig.14a and Fig.14 c) and 2080s (Fig.14b and
518 Fig.14d).

519 Figure 15: Projected change in mean seasonal minimum temperature (T_{min}) in the sub-basin using different
520 CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s (Fig.15a and Fig.15c) and 2080s (Fig.15b and
521 Fig.15d).

522 4.4 Assessment of change in streamflow in 2050s and 2080s under SSP245 and SSP585

523 The Sutlej River's mean monthly streamflow change as compared to the reference period's observed flow (1979-
524 2009) is shown in Fig. 16 under scenarios SSP245 and SSP585 for the future periods (2050s and 2080s).
525 According to both scenarios and all six models, the Sutlej River's streamflow will decrease between January (-
526 33.80 to -14.38%), February (-32.40 to -14.15%), March (-23.55 to -0.84%), November (-21.06 to -5.14%) and
527 December (-29.88 to -18.38%) in the 2050s and 2080s. Moreover, except for MPI-ESM-2HR and MPI-ESM1-
528 2-LR, which show an increase in streamflow in the 2080s under the higher emission scenario, all of the CMIP6-
529 GCMs indicate a decrease in the river's discharge in June (-20.24 to -0.57%) under SSP245 and SSP585 for both
530 the periods. Similarly, excluding EC-Earth-Veg (under SSP245 in 2050s) and INM-CM5-0 (under SSP245 in
531 250s and 2080s and under SSP585 in 2050s), all of the CMIP6-GCMs indicate a decrease in the river's
532 discharge in May (-25 to -2.85%) during the study period. In contrast, under SSP245 and SSP585 in the 2050s
533 and 2080s, all of the CMIP6-GCMs predict a rise in the river's discharge in April (20.24 to -0.57%; excluding
534 SSP585 in 2080s), August (16.84 to 5.28%), and September (55.27 to 4.35%). But no clear pattern of
535 streamflow change is seen for the remaining months (July and October) of the year, making results difficult to
536 generalise because projected decrease/or increase in streamflow over the months is inconsistent among models
537 under various emission scenarios in the 2050s and 2080s. The variations in climate variable projections caused
538 by differing spatial resolutions and parametrization levels in the climate models may be the cause of these
539 discrepancies in streamflow estimates (Sperna Weiland et al., 2010; Singh et al., 2015a). According to Murphy
540 et al. (2004), the average of an ensemble of GCMs cancels out the errors of each individual model, and as more
541 models are used, the ensemble uncertainty decreases. Therefore, in order to reduce uncertainty in projection of
542 streamflow related to individual CMIP6-GCMs, streamflow pattern of the Sutlej River was analysed also using
543 the mean ensemble of all six GCMs.

544 Figure 16: Predicted change in monthly streamflow pattern of the Sutlej River with respect to the reference
545 period (1979-2009) in 2050s (Fig.16a and Fig. 16b) and 2080s (Fig.16c and Fig. 16d) under SSP245 and
546 SSP585 for different CMIP6-GCMs.

547 The mean ensemble of the models predicts that the Sutlej River's mean monthly streamflow (excluding April)
548 will decrease under both scenarios from November (-18.45 to -17.17%) to June (-10.90 to -8.06%) between
549 2050s and 2080s (Fig. 17). The river flow, which would have been expected to increase in April under both
550 scenarios in 2050s, will also decline in 2080s for the higher emission scenarios (SPP585). The maximum and
551 minimum streamflow declines are predicted to occur in the 2050s under SSP245 for the months of December (-
552 24.25%) and May (-7.77%), respectively. In comparison to SPP245, the decline generally will be slightly higher
553 under SSP585 in 2050s and, for the 2080s, the projected decrease in streamflow will not show much difference
554 under both the scenarios. Opposite to this, the mean ensemble of the models predicts that the Sutlej River's flow
555 will increase from July (5.50 to 5.91%) to October (3.01 to 11.42%) in the 2050s and 2080s under both the
556 scenarios. The maximum and minimum streamflow increases are predicted to occur in the 2080s under SSP245
557 for the months of September (25.82%) and July (5.50%), respectively. In all scenarios, the increase will be
558 slightly greater in the 2080s than it will be in the 2050s. When compared to SPP245, it will be higher for
559 SSP585 in scenarios.

560 Figure17: Comparison of monthly observed (1979-2009) and projected discharge of the multi-model ensembles
561 for period 2050s and 2080s under SSP245 and SSP585 scenarios.

562 The projected change in seasonal streamflow of the Sutlej River in 2050s and 2080s is shown in the Fig. 18. The
563 2050s and 2080s would see an increase in streamflow during the monsoon (4.46 to 16.14%) and a decrease
564 during the pre-monsoon (-17.40 to -0.51%) and winter (-28.81 to -12.42%) for all six CMIP6-GCMs, with the
565 exception of INM-CM5-0 in the 2050s under SSP245 and MPI-ESM-2HR and MPI-ESM1-2-LR in the 2080s
566 under SPP585, which indicate an increase in streamflow during the pre-monsoon rather than a decrease. The
567 predicted streamflow for the post-monsoon season, however, does not show a consistent pattern of change
568 across time within the models under SSP245 and SSP585 scenarios. But there is high probability, based on the
569 mean ensembles of models projections, that streamflow will also decline during the post-monsoon in 2050s (-
570 1.23 to -0.22%) and 2080s (-5.59 to -2.83%) under all scenarios. Similarly, the predicted decline for pre-
571 monsoon and winter will be between -10.36 and -6.12% and -21.87 and -21.52% under SSP245, and between -
572 10.0 and -9.13% and -21.87 and -21.11% under SSP585, respectively. With the exception of winter, when there
573 are no significant differences in the projected streamflow, the decline will be slightly larger in the 2080s than it
574 would be in the 2050s in all scenarios. In addition, the results of the mean ensemble of the models indicate that
575 the Sutlej River's flow will increase during the monsoon under both scenarios, from 9.70 to 11.41% in the 2050s
576 and 11.64 to 12.70% in the 2080s.

577 Figure 18: Predicted change in seasonal streamflow pattern of the Sutlej River with respect to the reference
578 period (1979-2009) in 2050s (Fig. 18a and Fig. 18c) and 2080s (Fig. 18a and Fig. 18c) under SSP245 and
579 SSP585 for different GCMs.

580 Similarly, Fig. 19 lists the projected change in mean annual streamflow for the Sutlej River in 2050s and 2080s
581 with respect to the reference period (1979-2009) under different emission scenarios. Although the nature of the
582 direction of change within models vary, the mean ensemble of the models reveals a persistent increasing pattern
583 in streamflow for all scenarios in 2050s and 2080s. The Sutlej River's annual stream flow will rise between 2050
584 and 2080 by 0.79 to 1.43% for SSP585 and 0.87 to 1.10% for SSP245, according to the mean ensemble of the
585 models. The rise is expected to be higher in 2080s as compared to 2050s under SSP585.

586 Figure 19: Predicted change in mean annual streamflow of the Sutlej River with respect to the reference period
587 (1979-2009) in 2050s and 2080s under SSP245 and SSP585 for different GCMs.

588 **5 Discussion**

589 This study reveals an increase in the Sutlej River's mean annual and monsoonal streamflow in the 2050s and
590 2080s in contrast to earlier studies (Singh et al., 2014; Ali et al., 2018) that reported a reduction based on long-
591 term investigation of station data over historical era. The pattern of rainfall and temperature predicted by
592 CMIP6-GCMs for future periods under the SSP245 and SSP585 emission scenarios, as well as physical
593 processes occurring within the basin, have contributed to this increase in the Sutlej River's streamflow. For
594 instance, it is speculated that the projected increase in mean streamflow during the monsoon season under both
595 scenarios in the 2050s and 2080 for all models is related to the projected percentage increase in rainfall amount
596 over the catchment and the melting of glaciers brought on by the increased maximum and minimum
597 temperatures. This increase in river streamflow and its propensity to raise silt load may have an impact on both
598 the capacity of reservoirs and the hydropower potential of hydroelectric facilities situated in the sub-basin and
599 downstream of it. On the other hand, despite the predicted increase in rainfall throughout the pre-monsoon, post-
600 monsoon, and winter seasons, the anticipated decrease in streamflow of the Sutlej River during pre-monsoon,
601 post-monsoon, and winter may be explained by the projected rise in temperatures, which may have led to
602 increased evaporation from the surface. Similar conclusions were reached by Adib and Harun (2022) who
603 studied the Kurau River in Malaysia and predicted a drop in streamflow during the months of January, April,
604 and October despite receiving more rainfall. Moreover, during winter and post-monsoon, most of precipitation in
605 upper part of the catchment occurs in form of snowfall which have minimal effect over runoff generation in the
606 catchment. Additionally, the large increase in monsoonal streamflow predicted during study periods is what led
607 to the projected increase in the Sutlej River's mean annual flow. Predicted decreases in Sutlej River streamflow
608 over the pre-monsoon (April to June) and winter (December to March) seasons may have a significant impact on
609 agriculture and hydropower generation downstream of the river, which is already struggling due to water
610 shortages at this time of year. Ali et al. (2018) predicted that the hydroelectric production from the Nathpa
611 Jhakri and Bhakra Nangal hydropower projects will decline during May to June in the future due to projected
612 decline in the streamflow of the Sutlej River.

613
614 The projected streamflow patterns for the Sutlej River under SSP245 and SSP585 in 2050s and 2080s show
615 similar tendencies, but with differing magnitudes, that have been found by past researchers using process-based
616 hydrological models. For instance, Singh et al. (2015a) used the SWAT (Soil Water Assessment Tool) model, a
617 semi-distributed hydrological model, to simulate streamflow for future periods using two CMIP3-GCMs models
618 (CGCM3 and HadCM3), and they discovered that the Sutlej River's mean annual streamflow would increase in
619 the range of 0.6 to 7.8% for the future periods (2050s and 2080s). Similar to this, using the Variable Infiltration
620 Capacity (VIC) and SWAT models, respectively, Ali et al. (2018) and Shukla et al. (2021) estimated increases
621 in the Sutlej River's mean annual streamflow for the 2050s and 2080s under RCP4.5 and RCP8.5. The study of
622 Shukla et al. (2021) estimated that under RCP4.5 and RCP8.5, the mean streamflow of the river would increase
623 by 14 and 21% (at Rampur), respectively, in the 2080s. The previous studies' observed substantially higher
624 increase in projected streamflow may be attributable to the CMIP3-GCMs' and CMIP5-GCMs' overestimation

625 of monsoonal precipitation over the Himalayan region (Choudhary et al., 2017; Sanjay et al., 2017; Gusain et
626 al., 2020; Lalande et al., 2021). Similar to this, the results of Singh et al. (2015a), Ali et al. (2018), and Shukla et
627 al. (2021) corroborated the expected decrease in streamflow during pre-monsoon and winter as well as rise
628 during monsoon. This suggests that the RF model can accurately predict runoff and analyse the effects of
629 climate change while capturing the nonlinearity of a hilly catchment.

630 **6 Conclusion**

631 This study compared the performance of the five machine learning models (GLM, GAM, MARS, ANN, and
632 RF) and one deep learning model (1D-CNN) which were further divided into linear (MARS, ANN, and RF) and
633 non-linear (ANN, 1D-CNN, and RF) models, in simulating rainfall-runoff responses over the hilly Sutlej River
634 Basin in order to determine the best model for predicting streamflow response to future climate change in the
635 2050s and 2080s under SSP245 and SSP585 using CMIP6-GCMs data. The important findings of the study are
636 summarised below:

637 In general, non-linear models (ANN, 1D-CNN and RF) outperformed linear models (GAM, GLM and
638 MARS) in runoff prediction under all rainfall scenarios (R_0 , R_1 , R_2 , and R_3). Among all the models, RF
639 and 1D-CNN were identified as the best models as per the model evaluation criteria. However, RF
640 outperformed CNN in terms of error index (MAE and PBIAS), and as a result, it was used to
641 investigate impact of future climate change on the Sutlej River pattern in the 2050s and 2080s under
642 SSP245 and SSP585 emission scenarios.

643 • The developed RF model slightly underperformed at the prediction of higher values of streamflow
644 during training and testing. This implies that it is less effective in predicting flash floods that are caused
645 by intense rainfall in the catchment. However, it was determined that the results produced by RF were
646 comparable to process-based hydrological models for long-term change study in streamflow pattern.

647
648 • Significant variations in the streamflow pattern were observed throughout the periods of months,
649 seasons, years, and for the CMIP6-GCMs. The differences in spatial resolution and parametrisation
650 levels of CMIP6-GCMs, which caused a noticeable change in the projected amounts of temperature
651 and precipitation during the study periods, may serve as an illustration of these variances in streamflow
652 prediction. The Sutlej River's mean annual streamflow based on the mean ensemble of models is
653 predicted to rise between the years 2050 and 2080 by 0.79 to 1.43% for SSP585 and by 0.87 to 1.10%
654 for SSP245. Additionally, under both emission scenarios, streamflow will decrease during the pre- and
655 post-monsoon (-1.23 to -0.22% and -5.59 to -2.83%), as well as during the winter (-21.87 to -21.52%
656 and -21.87 to -21.11%), but increase during the monsoon (9.70 to 11.41% and 11.64 to 12.70%) in the
657 2050s and 2080s.

658
659 • The increase in the Sutlej River's streamflow (annual and monsoon) is due to both physical processes
660 that occur within the basin and rainfall and temperature patterns that are predicted by CMIP6-GCMs
661 for future time periods under the SSP245 and SSP585 emission scenarios. The projected rise in mean
662 streamflow during the monsoon season is associated to both the projected percentage increase in

663 rainfall over the catchment and the melting of glaciers brought on by the increasing maximum and
664 minimum temperatures. On the other hand, the predicted increase in temperatures, which may have led
665 to increased evaporation from the surface, may be used to explain the anticipated reduction in
666 streamflow of the Sutlej River during pre-monsoon, post-monsoon, and winter.

667

668 • Additionally, the projected changes in the mean annual and seasonal streamflow of the river are
669 consistent with earlier research done using process-based physical hydrological models. Thus, the
670 outcomes of the overall study indicate that the RF model is efficient for simulating streamflow in the
671 Himalayan catchment, and that water availability during monsoon will rise as a result of an increase in
672 catchment precipitation, which would eventually lead to an increased sediment load and affect
673 hydropower generation. However, predicted reduction in streamflow during pre-monsoon, post-
674 monsoon and winter will put stress on agriculture and hydropower generation downstream of the river,
675 which is already struggling due to water shortages at this time of year. The administrators of local
676 water resources and the government organizations in charge of maintaining reservoirs down river may
677 find these details on streamflow patterns to be of great use.

678

679 **Code Availability:** The codes developed for this study is made available to the readers on reasonable request.

680

681 **Data Availability:** The observed station data are confidential and authors do not have permission for sharing
682 the data.

683

684 **Author's Contribution**

685 DS and SL conceptualized the problems, supervised the entire research activity from its inception to the
686 completion, contributed in data collection, processing, interpretation and wrote the research paper. MV and RS
687 contributed in the development of model, generation of figures and analysis of data. PC and DC contributed in
688 the data analysis and interpretation.

689

690 **Statements and Declarations**

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

693

694 **Acknowledgement**

695 Authors acknowledge National Natural Science Foundation of China (NFSC Grant no. 42171129) for funding
696 this research work, and Bhakara Beas Management Board (BBM), India for the hydro-meteorological data used
697 in this study. We thank Ms. Pratibha Shrivastava, a fourth-year B.Tech computer science student from NIT
698 Raipur, for her help in building the model.

699

700

701

702 **References**

- 703 Abbasian, M., Moghim, S., and Abrishamchi, A.: Performance of the general circulation models in simulating
704 temperature and precipitation over Iran, *Theor. Appl. Climatol.*, 135, 3, 1465-1483,
705 <https://doi.org/10.1007/s00704-018-2456-y>, 2019.
- 706 Adib M N, M., and Harun, S.: Metalearning Approach Coupled with CMIP6 Multi-GCM for Future Monthly
707 Streamflow Forecasting, *J. Hydrol. Eng.*, 27, 6, 05022004, [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0002176](https://doi.org/10.1061/(ASCE)HE.1943-5584.0002176), 2022.
- 709 Adnan R, M., Liang, Z., Heddad, S., Zounemat-Kermani, M., Kisi, O., and Li, B.: Least square support vector
710 machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin
711 using hydro-meteorological data as inputs, *J. Hydrol.*, 586, 124371,
712 <https://doi.org/10.1016/j.jhydrol.2019.124371>, 2020.
- 713 Adnan. R.M., Yuan, X., Kisi, O., Yuan, Y., Tayyab, M., and Lei, X.: Application of soft computing models in
714 streamflow forecasting. In *Proceedings of the institution of civil engineers-water. Manage.*, Vol. 172,
715 No. 3, pp, 123-134, Thomas Telford Ltd, <https://doi.org/10.1680/jwama.16.00075>, 2019.
- 716 Ali, S. A., Aadhar, S., Shah, H. L., and Mishra, V.: Projected increase in hydropower production in India under
717 climate change, *Sci. Rep.*, 8,1, 1-12, <https://doi.org/10.1038/s41598-018-30489-4>, 2018.
- 718 Archer, D.: Contrasting hydrological regimes in the upper Indus Basin, *J. Hydrol.*, 274,1-4, 198-210,
719 [https://doi.org/10.1016/S0022-1694\(02\)00414-6](https://doi.org/10.1016/S0022-1694(02)00414-6), 2003.
- 720 Beven, K. J.: *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2011.
- 721 Chen, H., Sun, J., Lin, W., and Xu, H.: Comparison of CMIP6 and CMIP5 models in simulating climate
722 extremes, *Sci. Bull.*, 65, 17, 1415-1418, <https://doi.org/10.1016/j.scib.2020.05.015>, 2020.
- 723 Choudhury B, A., Rajesh P, V., Zahan, Y. and Goswami B, N.: Evolution of the Indian summer monsoon
724 rainfall simulations from CMIP3 to CMIP6 models, *Clim. Dyn.*, 58, 9, 2637-2662,
725 <https://doi.org/10.1007/s00382-021-06023-0>, 2022.
- 726 Conan, C., De Marsily, G., Bouraoui, F., and Bidoglio, G.: A long-term hydrological modelling of the Upper
727 Guadiana River basin (Spain), *Phys. Chem. Earth., Parts A/B/C*, 28, 4-5, 193-200,
728 [https://doi.org/10.1016/S1474-7065\(03\)00025-1](https://doi.org/10.1016/S1474-7065(03)00025-1) 2003.
- 729 Dai, A., Qian, T., Trenberth, K. E., and Milliman, J. D.: Changes in continental freshwater discharge from 1948
730 to 2004, *J. Clim.*, 22, 10, 2773-2792, <https://doi.org/10.1175/2008JCLI2592.1>, 2009.
- 731 Das, J., and Nanduri U, V.: Assessment and evaluation of potential climate change impact on monsoon flows
732 using machine learning technique over Wainganga River basin, India, *Hydrol. Sci. J.*, 63, 7, 1020-
733 1046, <https://doi.org/10.1080/02626667.2018.1469757>, 2018.
- 734 Easterling, D. R., Meehl G, A., Parmesan, C., Changnon S, A., Karl, T. R., and Mearns, L. O.: Climate
735 extremes: observations, modeling, and impacts. *Sci.*, 289, 5487, 2068-2074, DOI:
736 10.1126/science.289.5487.206, 2000.
- 737 Eng, K., and Wolock D. M.: Evaluation of machine learning approaches for predicting streamflow metrics
738 across the conterminous United States (No. 2022-5058), US Geological Survey, 2022.
- 739 Fu, M., Fan, T., Ding Z, A., Salih S, Q., Al-Ansari, N., and Yaseen Z. M.: Deep learning data-intelligence
740 model based on adjusted forecasting window scale: application in daily streamflow simulation, *IEEE*.
741 *Access.*, 8, 32632-32651, 10.1109/ACCESS.2020.2974406, 2020.

742 Gao, Y., Gao, X., and Zhang, X.: The 2 C global temperature target and the evolution of the long-term goal of
743 addressing climate change—from the United Nations framework convention on climate change to the
744 Paris agreement, *Eng.*, 3, 2, 272-278, <https://doi.org/10.1016/J.ENG.2017.01.022>, 2017.

745 Gerten, D., Rost, S., von Bloh, W., and Lucht, W.: Causes of change in 20th century global river discharge,
746 *Geophys. Res. Lett.*, 35, 20, <https://doi.org/10.1029/2008GL035258>, 2008.

747 Ghimire, S., Yaseen Z, M., Farooque A, A., Deo R, C., Zhang, J. and Tao, X.: Streamflow prediction using an
748 integrated methodology based on convolutional neural network and long short-term memory networks.
749 *Sci. Rep.*, 11, 1, pp.1-26, <https://doi.org/10.1038/s41598-021-96751-4>, 2021.

750 Ghobadi, F., and Kang, D.: Improving long-term streamflow prediction in a poorly gauged basin using geo-
751 spatiotemporal mesoscale data and attention-based deep learning: A comparative study, *J. Hydrology.*,
752 615, 128608, <https://doi.org/10.1016/j.jhydrol.2022.128608>, 2022.

753 Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., and Xavier, P. K.: Increasing trend of
754 extreme rain events over India in a warming environment, *Science.*, 314, 5804, 1442-1445, DOI:
755 10.1126/science.113202, 2006.

756 Gupta H, V., Sorooshian, S., and Yapo P, O.: Status of automatic calibration for hydrologic models:
757 Comparison with multilevel expert calibration, *J. Hydrol. Engineering.*, 4, 2, 135-143, 1999.

758 Gusain, A., Ghosh, S., and Karmakar, S.: Added value of CMIP6 over CMIP5 models in simulating Indian
759 summer monsoon rainfall, *Atmos. Res.*, 232, 104680, <https://doi.org/10.1016/j.atmosres.2019.104680>,
760 2020.

761 Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F.,
762 Masaki, Y., Schewe, J. and Stacke, T.: Global water resources affected by human interventions and
763 climate change, *Proc. Natl. Acad. Sci.*, 111, 9, 3251-3256, <https://doi.org/10.1073/pnas.1222475110>,
764 2014.

765 Hagen J, S., Leblois, E., Lawrence, D., Solomatine, D. and Sorteberg, A.: Identifying major drivers of daily
766 streamflow from large-scale atmospheric circulation with machine learning, *J. Hydrol.*, 596, 126086,
767 <https://doi.org/10.1016/j.jhydrol.2021.126086>, 2021.

768 Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H.: *The elements of statistical learning: data mining,*
769 *inference, and prediction*, Vol. 2, pp., 1-758, New York: springer, 2009.

770 Hawkins, E., Osborne T, M., Ho C, K. and Challinor A, J.: Calibration and bias correction of climate projections
771 for crop modelling: an idealised case study over Europe, *Agric. For. Meteorol.*, 170, pp.19-31,
772 <https://doi.org/10.1016/j.agrformet.2012.04.007>, 2013.

773 Herath, H. M. V. V., Chadawalawada, J., and Babovic, V.: Hydrologically informed machine learning for rainfall–
774 runoff modelling: towards distributed modelling, *Hydrol. Earth Syst. Sci.*, 25, 8, 4373-4401,
775 <https://doi.org/10.5194/hess-25-4373-2021>, 2021.

776 Hussain, D., and Khan A, A.: Machine learning techniques for monthly river flow forecasting of Hunza River,
777 Pakistan, *Earth. Sci. Inf.*, 13(3), pp.939-949, <https://doi.org/10.1007/s12145-020-00450-z>, 2020a.

778 Hussain, D., Hussain, T., Khan A, A., Naqvi S A, A., and Jamil, A.: A deep learning approach for hydrological
779 time-series prediction: A case study of Gilgit river basin. *Earth Science Informatics*, 13(3), pp.915-
780 927.,2020b.

781 Jose D, M., and Dwarakish G, S.: Bias Correction and trend analysis of temperature data by a high-resolution
782 CMIP6 Model over a Tropical River Basin, *Asia-Pac. J. Atmos. Sci.*, 58, 1, 97-115,
783 <https://doi.org/10.1007/s13143-021-00240-7>, 2022.

784 Kabir, S., Patidar, S., and Pender, G.: Investigating capabilities of machine learning techniques in forecasting
785 stream flow. In *Proceedings of the Institution of Civil Engineers-Water Manage.*, Vol. 173, No. 2, pp.,
786 69-86, Thomas Telford Ltd, <https://doi.org/10.1680/jwama.19.00001>, 2020.

787 Kadel, I., Yamazaki, T., Iwasaki, T. and Abdillah M, R.: Projection of future monsoon precipitation over the
788 central Himalayas by CMIP5 models under warming scenarios, *Clim. Res.*, 75, 1, 1-21,
789 <https://doi.org/10.3354/cr01497>, 2018.

790 Karan, K., Singh, D., Singh P, K., Bharati, B., Singh T, P., and Berndtsson, R.: Implications of future climate
791 change on crop and irrigation water requirements in a semi-arid river basin using CMIP6 GCMs, *J.*
792 *Arid. Land.*, 14, 11, 1234-1257, <https://doi.org/10.1007/s40333-022-0081-1>, 2022.

793 Kim Y, H., Min S, K., Zhang, X.: Evaluation of the CMIP6 multi-model ensemble for climate extreme indices,
794 *Weat. Clim. Extremes.*, 29, 100269, doi: 10.1016/j.wace.2020.100269, 2020.

795 Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to
796 advance the science of hydrology, *Water. Resour. Res.*, 42, 3, <https://doi.org/10.1029/2005WR004362>,
797 2006.

798 Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M. and Inman, D.J.: 1D convolutional neural
799 networks and applications: A survey. *Mechanical systems and signal processing*, 151, 107398,
800 <https://doi.org/10.1016/j.ymsp.2020.107398>, 2021.

801 Krysanova, V., Wortmann, M., Bolch, T., Merz, B., Duethmann, D., Walter, J., ... and Kundzewicz, Z. W.:
802 Analysis of current trends in climate parameters, river discharge and glaciers in the Aksu River basin
803 (Central Asia), *Hydrol. Sci. J.*, 60, 4, 566-590, <https://doi.org/10.1080/02626667.2014.925559>, 2015.

804 Kundzewicz, Z. W., Nohara, D., Tong, J., Oki, T., Buda, S., and Takeuchi, K.: Discharge of large Asian rivers–
805 Observations and projections, *Quat. Int.*, 208, 1-2, 4-10, <https://doi.org/10.1016/j.quaint.2009.01.011>,
806 2009.

807 Lalande, M., Ménégoz, M., Krinner, G., Naegeli, K. and Wunderle, S.: Climate change in the High Mountain
808 Asia in CMIP6, *Earth. Syst. Dyn.*, 12, 4, 1061-1098, <https://doi.org/10.5194/esd-12-1061-2021>, 2021.

809 Legates D, R., and McCabe-Jr G, J.: Evaluating the use of “goodness-of-fit” measures in hydrologic and
810 hydroclimatic model validation, *Water. Resour. Res.*, 35, 1, 233-241,
811 <https://doi.org/10.1029/1998WR900018>, 1999.

812 Lutz A, F., Immerzeel W, W., Shrestha A, B., and Bierkens M F, P.: Consistent increase in High Asia's runoff
813 due to increasing glacier melt and precipitation, *Nat. Clim. Change.*, 4, 7, 587-592,
814 <https://doi.org/10.1038/nclimate2237>, 2014.

815 Lutz, A. F., Ter Maat, H. W., Wijngaard, R. R., Biemans, H., Syed, A., Shrestha, A. B., ... and Immerzeel, W.
816 W.: South Asian River basins in a 1.5 C warmer world. *Reg. Environ. Change.*, 19, 3, 833-847,
817 <https://doi.org/10.1007/s10113-018-1433-4>, 2019.

818 Mahato P, K., Singh, D., Bharati, B., Gagnon A, S., Singh B, B. and Brema, J.: Assessing the impacts of
819 human interventions and climate change on fluvial flooding using CMIP6 data and GIS-based

820 hydrologic and hydraulic models, *Geocarto. Int.*, 1-26,
821 <https://doi.org/10.1080/10106049.2022.2060311>, 2022.

822 Mazrooei, A., Sankarasubramanian, A., and Wood, A.W.: Potential in improving monthly streamflow
823 forecasting through variational assimilation of observed streamflow, *J. Hydrol.*, 600, 126559,
824 <https://doi.org/10.1016/j.jhydrol.2021.126559>, 2021.

825 Miller, J. D., Immerzeel W, W., and Rees, G.: Climate change impacts on glacier hydrology and river discharge
826 in the Hindu Kush–Himalayas, *Mt. Res. Dev.*, 32, 4, 461-467, [https://doi.org/10.1659/MRD-](https://doi.org/10.1659/MRD-JOURNAL-D-12-00027.1)
827 [JOURNAL-D-12-00027.1](https://doi.org/10.1659/MRD-JOURNAL-D-12-00027.1), 2012.

828 Mishra, V., Bhatia, U., and Tiwari, A.D.: Bias-corrected climate projections for South Asia from Coupled
829 Model Intercomparison Project-6, *Sci. Data.*, 7, 338, <https://doi.org/10.1038/s41597-020-00681-1>,
830 2020.

831 Moriasi D, N., Arnold J, G., Van-Liew M, W., Bingner R, L., Harmel R, D. and Veith T, L.: Model evaluation
832 guidelines for systematic quantification of accuracy in watershed simulations, *Trans. ASABE.*, 50, 3,
833 885-900, doi: 10.13031/2013.23153, 2007.

834 Murphy J, M., Sexton D, M., Barnett D, N., Jones G, S., Webb M, J., Collins, M., and Stainforth D, A.:
835 Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*,
836 430, 7001, 768-772, <https://doi.org/10.1038/nature02771>, 2004.

837 Nash J, E., and Sutcliffe J, V.: River flow forecasting through conceptual models part I—A discussion of
838 principles, *J. Hydrol.*, 10, 3, 282-290, [https://doi.org/10.1016/0022-1694\(70\)90098-3](https://doi.org/10.1016/0022-1694(70)90098-3), 1970.

839 Nepal, S., and Shrestha, A. B.: Impact of climate change on the hydrological regime of the Indus, Ganges and
840 Brahmaputra River basins: a review of the literature, *Int. J. Water. Resou. Dev.*, 31, 2, 201-218,
841 <https://doi.org/10.1080/07900627.2015.1030494>, 2015.

842 Niu, X., Wang, S., Tang, J., Lee D, K., Gutowski, W., Dairaku, K., McGregor, J., Katzfey, J., Gao, X.,
843 Wu, J., and Hong, S.: Projection of Indian summer monsoon climate in 2041– 2060 by multiregional
844 and global climate models, *J. Geophys Res.: Atmos*, 120, 5, 1776-1793,
845 <https://doi.org/10.1002/2014JD022620>, 2015.

846 Oki, T., and Kanae, S.: Global hydrological cycles and world water resources, *science.*, 313, 5790, 1068-1072,
847 DOI: 10.1126/science.1128845, 2006.

848 O'Neill B, C., Tebaldi, C., Van Vuuren D, P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E.,
849 Lamarque J, F., Lowe, J., and Meehl G, A.: The scenario model intercomparison project (ScenarioMIP)
850 for CMIP6, *Geosci. Model. Dev.*, 9, 9, 3461-3482, <https://doi.org/10.5194/gmd-9-3461-2016>, 2016.

851 Otto, F. E., Skeie, R. B., Fuglestedt, J. S., Berntsen, T., and Allen, M. R.: Assigning historic responsibility for
852 extreme weather events. *Nat. Clim. Change.*, 7, 11, 757-759, <https://doi.org/10.1038/nclimate3419>,
853 2017.

854 Pasquini, A. I., and Depetris, P. J.: Discharge trends and flow dynamics of South American rivers draining the
855 southern Atlantic seaboard: An overview, *J. Hydrol.*, 333, 2-4, 385-399,
856 <https://doi.org/10.1016/j.jhydrol.2006.09.005>, 2007.

857 Rahimzad, M., Moghaddam Nia, A., Zolfonoon, H., Soltani, J., Danandeh Mehr, A., and Kwon H,H.:
858 Performance comparison of an lstm-based deep learning model versus conventional machine learning

859 algorithms for streamflow forecasting, *Water. Resour. Manage.*, 35, 12, 4167-4187,
860 <https://doi.org/10.1007/s11269-021-02937-w>, 2021.

861 Rasouli, K., Hsieh W, W., and Cannon A, J.: Daily streamflow forecasting by machine learning methods with
862 weather and climate inputs, *J. Hydrol*, 414, 284-293, <https://doi.org/10.1016/j.jhydrol.2011.10.039>,
863 2012.

864 Refsgaard J, C.: Parameterisation, calibration and validation of distributed hydrological models, *J. hydrol.*, 198,
865 1-4, 69-97, [https://doi.org/10.1016/S0022-1694\(96\)03329-X](https://doi.org/10.1016/S0022-1694(96)03329-X), 1997.

866 Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., ... and Tavoni, M.: The
867 shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications:
868 an overview, *Global. Environ. Change.*, 42, 153-168, <https://doi.org/10.1016/j.gloenvcha.2016.05.009>,
869 2017.

870 Sabin T, P., Krishnan, R., Vellore, R., Priya, P., Borgaonkar H, P., Singh B, B., and Sagar, A.: Climate change
871 over the Himalayas. In *Assessment of climate change over the Indian region*, 207-222, Springer,
872 Singapore, https://doi.org/10.1007/978-981-15-4327-2_11, 2020.

873 Sanjay, J., Krishnan, R., Shrestha A, B., Rajbhandari, R. and Ren G, Y.: Downscaled climate change projections
874 for the Hindu Kush Himalayan region using CORDEX South Asia regional climate models, *Adv.*
875 *Clim. Change. Res.*, 8, 3, 185-198, <https://doi.org/10.1016/j.accre.2017.08.003>, 2017.

876 Schreiner-McGraw, A. P., and Ajami, H.: Impact of uncertainty in precipitation forcing data sets on the
877 hydrologic budget of an integrated hydrologic model in mountainous terrain, *Water. Resour. Res.*, 56,
878 12, <https://doi.org/10.1029/2020WR027639>, 2020.

879 Shortridge, J. E., Guikema S, D., and Zaitchik, B. F.: Machine learning methods for empirical streamflow
880 simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds,
881 *Hydrol. Earth Syst. Sci.*, 20, 7, 2611-2628, <https://doi.org/10.5194/hess-20-2611-2016>, 2016.

882 Shu, X., Ding, W., Peng, Y., Wang, Z., Wu, J. and Li, M.: Monthly streamflow forecasting using convolutional
883 neural network. *Water Resour. Manag.*, 35, 15, 5089-5104. [https://doi.org/10.1007/s11269-021-02961-](https://doi.org/10.1007/s11269-021-02961-w)
884 [w](https://doi.org/10.1007/s11269-021-02961-w), 2021

885 Shukla, S., Jain S, K., and Kansal, M. L.: Hydrological modelling of a snow/glacier-fed western Himalayan
886 basin to simulate the current and future streamflows under changing climate scenarios, *Sci. Total.*
887 *Environ.*, 795, 148871, <https://doi.org/10.1016/j.scitotenv.2021.148871>, 2021.

888 Singh, D., Gupta R, D., and Jain, S. K.: Assessment of impact of climate change on water resources in a hilly
889 river basin, *Arabian. J. Geosci.*, 8, 12, 10625-10646, <https://doi.org/10.1007/s12517-015-1985-2>,
890 2015a.

891 Singh, D., Gupta R, D., and Jain, S. K.: Statistical analysis of long term spatial and temporal trends of
892 temperature parameters over Sutlej River basin, India, *J. Earth. Syst. Sci.*, 124, 1, 17-35,
893 <https://doi.org/10.1007/s12517-015-1985-2>, 2015b.

894 Singh, D., Gupta R, D., and Jain, S. K.: Study of long-term trend in river discharge of Sutlej River (NW
895 Himalayan region), *Geogr. Environ. Sustainability.*, 7, 3, 87-96, [https://doi.org/10.24057/2071-9388-](https://doi.org/10.24057/2071-9388-2014-7-3-50-57)
896 [2014-7-3-50-57](https://doi.org/10.24057/2071-9388-2014-7-3-50-57), 2014.

897 Singh, D., Jain S, K., and Gupta, R. D.: Statistical downscaling and projection of future temperature and
898 precipitation change in middle catchment of Sutlej River Basin, India, *J. Earth. Syst. Sci.*, 124, 4, 843-
899 860, <https://doi.org/10.1007/s12040-015-0575-8>, 2015c.

900 Singh, D., Rai S, P., and Rai, D.: Application of geospatial techniques in hydrological modelling, In: Shah, S.,
901 Venkatramanan, V., Prasad, R., (eds), *Sust. Green. Techno. Environ. Management.*, Springer,
902 Singapore. https://doi.org/10.1007/978-981-13-2772-8_8, 2019.

903 Singh, D., Zhu, Y., Liu, S., Srivastava P, K., Dharpure J, K., Chatterjee, D., ... and Gagnon, A. S.: Exploring the
904 links between variations in snow cover area and climatic variables in a Himalayan catchment using
905 earth observations and CMIP6 climate change scenarios. *J. Hydrol*, 608, 127648,
906 <https://doi.org/10.1016/j.jhydrol.2022.127648>, 2022a.

907 Singh, J., Knapp H, V., Arnold J, G., and Demissie, M.: Hydrological modeling of the Iroquois river watershed
908 using HSPF and SWAT 1, *JAWRA. J. Am. Water. Resour. Assoc.*, 41, 2, 343-360, 2005.

909 Singh, P., and Jain, S. K.: Snow, and glacier melt in the Satluj River at Bhakra Dam in the western Himalayan
910 region, *Hydrol. Sci. J.*, 47, 1, 93-106, <https://doi.org/10.1080/02626660209492910>, 2002.

911 Sood, A., and Smakhtin, V.: Global hydrological models: a review, *Hydrol. Sci. J.*, 60, 4, 549-565,
912 <https://doi.org/10.1080/02626667.2014.950580>, 2015.

913 Sperna-Weiland F, C., Van-Beek L P, H., Kwadijk-J C, J., and Bierkens-M F, P.: The ability of a GCM-forced
914 hydrological model to reproduce global discharge variability, *Hydrol. Earth. Syst. Sci.*, 14, 8, 1595-
915 1621, <https://doi.org/10.5194/hess-14-1595-2010>, 2010.

916 Stahl, K., Hisdal, H., Hannaford, J., Tallaksen L, M., Van Lanen, H. A. J., Sauquet, E., ... and Jódar, J.:
917 Streamflow trends in Europe: evidence from a dataset of near-natural catchments, *Hydrol. Earth. Syst.*
918 *Sci.*, 14, 12, 2367-2382, <https://doi.org/10.5194/hess-14-2367-2010>, 2010.

919 Stahl, K., Tallaksen, L. M., Hannaford, J., and Van Lanen, H. A. J.: Filling the white space on maps of
920 European runoff trends: estimates from a multi-model ensemble, *Hydrol. Earth. Syst. Sci.*, 16, 7, 2035-
921 2047, <https://doi.org/10.5194/hess-16-2035-2012>, 2012.

922 Taylor K, E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res:*
923 *Atmos.*, 106, D7, 7183-7192, <https://doi.org/10.1029/2000JD900719>, 2001.

924 Thapa, S., Li, H., Li, B., Fu, D., Shi, X., Yabo, S., Lu, L., Qi, H. and Zhang, W.: Impact of climate change on
925 snowmelt runoff in a Himalayan basin, Nepal. *Environmental Monitoring and Assessment*, 193, 7, 1-
926 17, <https://doi.org/10.1007/s10661-021-09197-6>, 2021

927 Trenberth K, E., Fasullo J. T., and Shepherd T, G. Attribution of climate extreme events, *Nat. Clim. Change.*, 5,
928 8, 725-730, <https://doi.org/10.1038/nclimate2657>, 2015.

929 Trenberth, K. E.: Changes in precipitation with climate change, *Clim. Res.*, 47, 1-2, 123-138,
930 <https://doi.org/10.3354/cr00953>, 2011.

931 Turner A, G., and Annamalai, H.: Climate change and the South Asian summer monsoon. *Nat. Clim. Change.*,
932 2, 8, 587-595, <https://doi.org/10.1038/nclimate1495>, 2012.

933 Van der Wiel, K., Wanders, N., Selten F, M., and Bierkens M F, P.: Added value of large ensemble simulations
934 for assessing extreme river discharge in a 2 C warmer world, *Geophys. Res. Lett.*, 46, 4, 2093-2102,
935 <https://doi.org/10.1029/2019GL081967>, 2019.

936 Van S, P., Le H, M., Thanh D, V., Dang T, D., Loc H, H., and Anh D, T.: Deep learning convolutional neural
937 network in rainfall–runoff modelling, *J. Hydroinf.*, 22, 3, 541-561,
938 <https://doi.org/10.2166/hydro.2020.095>, 2020.

939 Van-Liew M, W., Arnold J, G., and Garbrecht J, D.: Hydrologic simulation on agricultural watersheds:
940 Choosing between two models, *Trans. ASAE.*, 46, 6, 1539, doi: 10.13031/2013.15643, 2003.

941 Wang, T., Zhao, Y., Xu, C., Ciais, P., Liu, D., Yang, H., ... and Yao, T.: Atmospheric dynamic constraints on
942 Tibetan Plateau freshwater under Paris climate targets. *Nat. Clim. Change.*, 11, 3, 219-225,
943 <https://doi.org/10.1038/s41558-020-00974-8>, 2021.

944 Xenarios, S., Gafurov, A., Schmidt-Vogt, D., Sehring, J., Manandhar, S., Hergarten, C., ... and Foggin, M.:
945 Climate change and adaptation of mountain societies in Central Asia: uncertainties, knowledge gaps,
946 and data constraints, *Reg. Environ. Change.*, 19, 5, 1339-1352, [https://doi.org/10.1007/s10113-018-](https://doi.org/10.1007/s10113-018-1384-9)
947 [1384-9](https://doi.org/10.1007/s10113-018-1384-9), 2019.

948 Xiang, Z., Yan, J., and Demir, I.: A rainfall-runoff model with LSTM-based sequence-to-sequence learning.
949 *Water. Resour. Res.*, 56, 1, <https://doi.org/10.1029/2019WR025326>, 2020.

950 Yang, Q., Zhang, H., Wang, G., Luo, S., Chen, D., Peng, W., and Shao, J.: Dynamic runoff simulation in a
951 changing environment: A data stream approach, *Environ. Modell. Soft.*, 112, 157-165,
952 <https://doi.org/10.1016/j.envsoft.2018.11.007>, 2019.

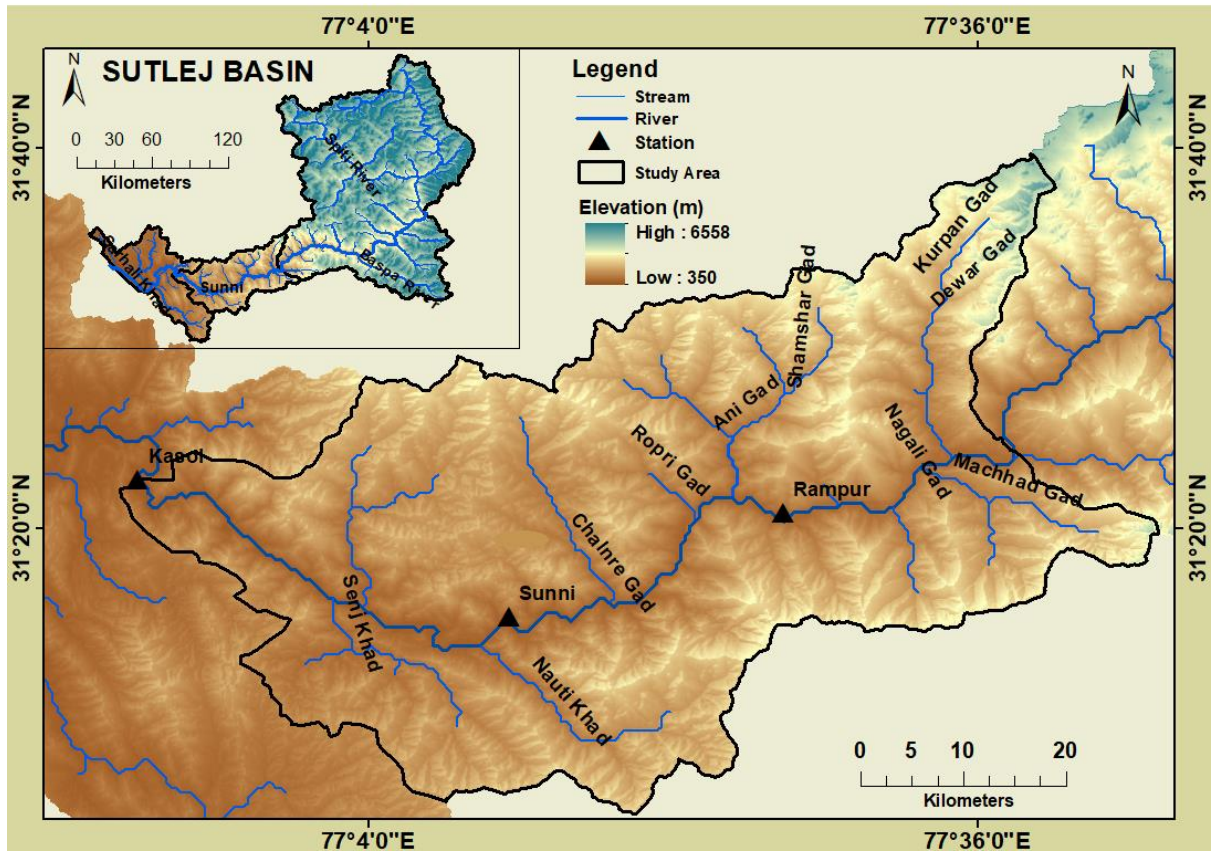
953 Yaseen Z, M., El-Shafie, A., Jaafar, O., Afan H, A., and Sayl K, N.: Artificial intelligence based models for
954 stream-flow forecasting: 2000–2015, *J. Hydrol.*, 530, 829-844,
955 <https://doi.org/10.1016/j.jhydrol.2015.10.038>, 2015.

956 Zhao, B., Sun, H., Yan, D., Wei, G., Tuo, Y., & Zhang, W.: Quantifying changes and drivers of runoff in the
957 Kaidu River Basin associated with plausible climate scenarios, *J. Hydrol.: Reg. Stud.*, 38, 100968,
958 <https://doi.org/10.1016/j.ejrh.2021.100968>, 2021.

959 Zhu, Y., Liu, S., Yi, Y., Xie, F., Grünwald, R., Miao, W., ... and Singh, D.: Overview of terrestrial water storage
960 changes over the Indus River Basin based on GRACE/GRACE-FO solutions, *Sci. Total. Environ.*, 799,
961 149366, <https://doi.org/10.1016/j.scitotenv.2021.149366>, 2021.

962

963



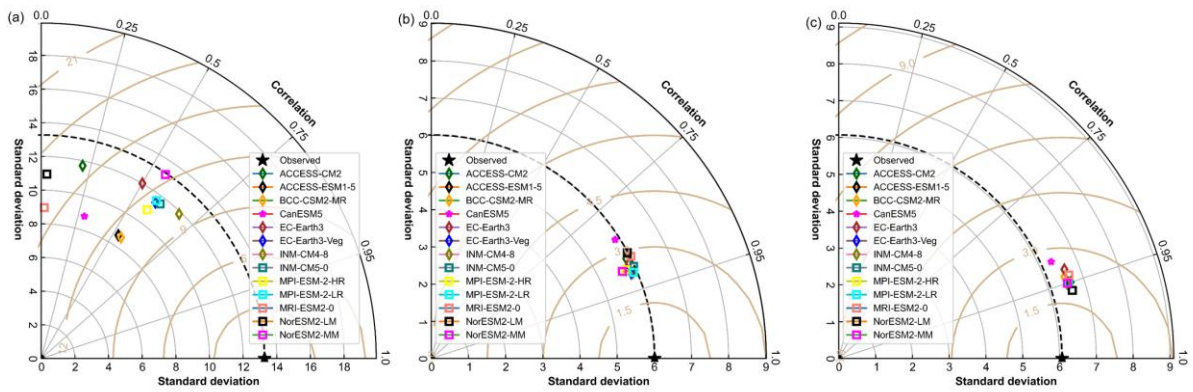
964

965

966

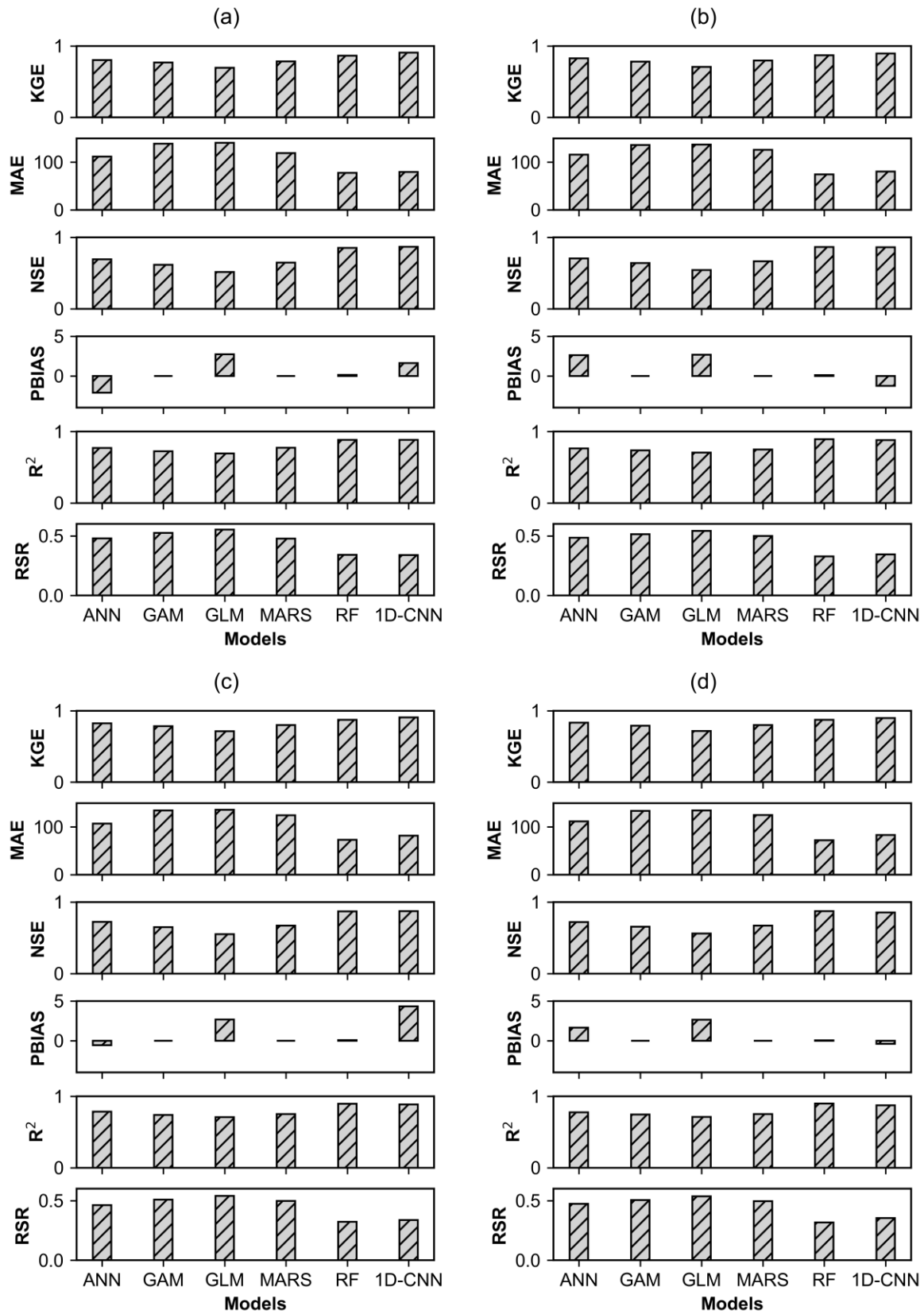
Figure 1: The location of the sub-catchment within Sutlej River Basin. The three hydro-meteorological stations (Kasol, Sunni and Rampur) from which this study employed observed data for the years 1979 to 2009 are also shown.

967



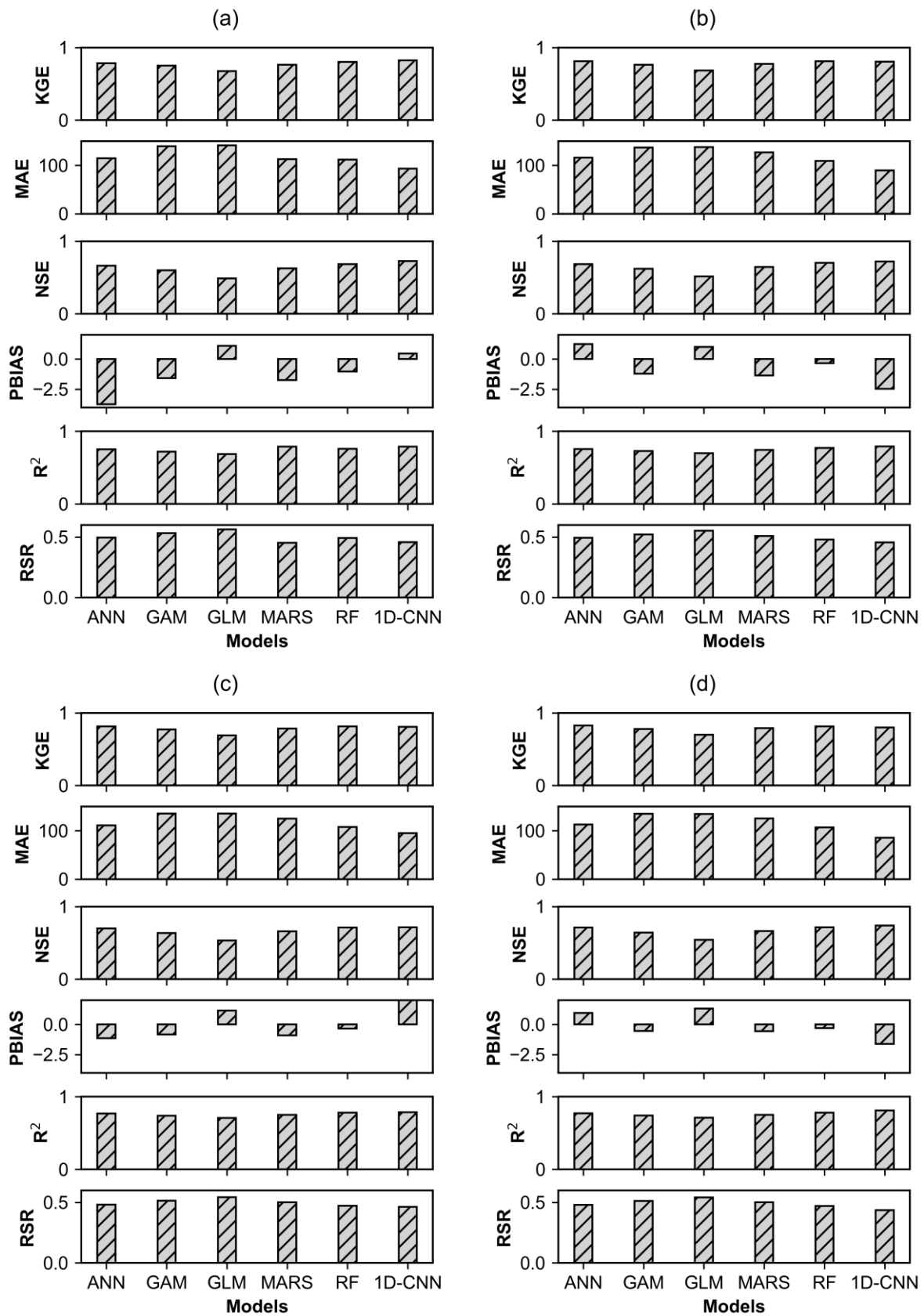
969

970 **Figure 2: Taylor diagram showing comparative skills of 13CMIP6-GCMs in simulating climatic variables (rainfall,**
 971 **T_{max} and T_{min}) over the Sutlej sub-basin during reference period (1979-2009). The degree of correlation coefficient (r)**
 972 **between observed and CMIP6-GCMs, centered root-mean-square error (CRMSE) and departure of the models'**
 973 **standard deviation (SD) from the observed data (dashed black arc line) are shown in Fig. 2a for rainfall, Fig. 2b for**
 974 **T_{max} and Fig. 2c for T_{min} . The units of SD for rainfall and temperature is in cm and $^{\circ}C$, respectively.**



975
 976
 977
 978

Figure 3: Evaluation of the models (ANN, GAM, GLM, MARS, RF and 1D-CNN) performance in simulating streamflow under rainfall scenarios R₀ (Fig.3a) , R₁ (Fig. 3b), R₂ (Fig.3c) and (Fig. 3d) R₃ at Kasol during training phase using six statistical metrics (R², KGE, NSE, RSR, MAE and PBIAS).



980

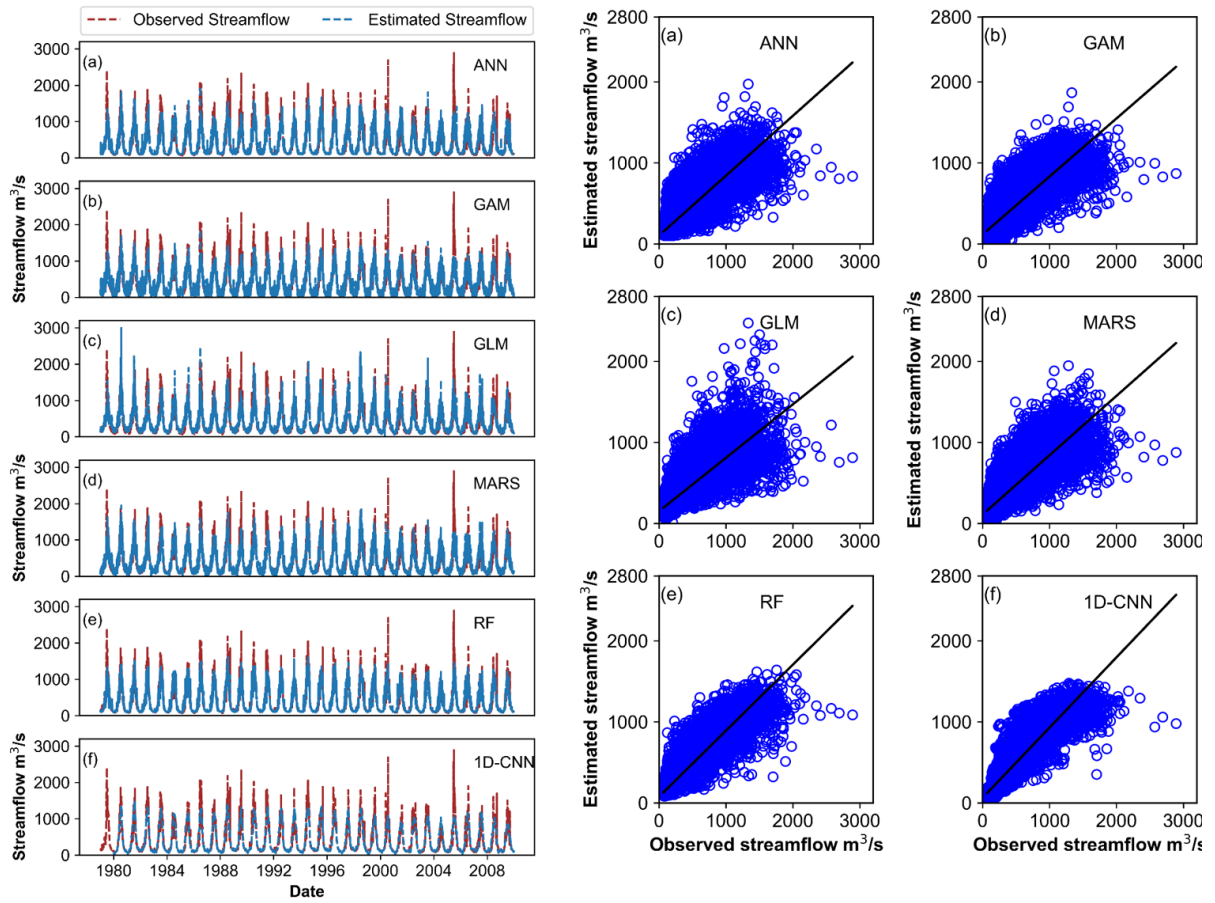
981

982

983

Figure 4: Evaluation of the models (ANN, GAM, GLM, MARS, RF and 1D-CNN) performance in simulating streamflow under rainfall scenarios R₀ (Fig.4a), R₁ (Fig. 4b), R₂ (Fig.4c) and (Fig. 4d) R₃ at Kasol during testing phase using six statistical metrics (R², KGE, NSE, RSR, MAE and PBIAS).

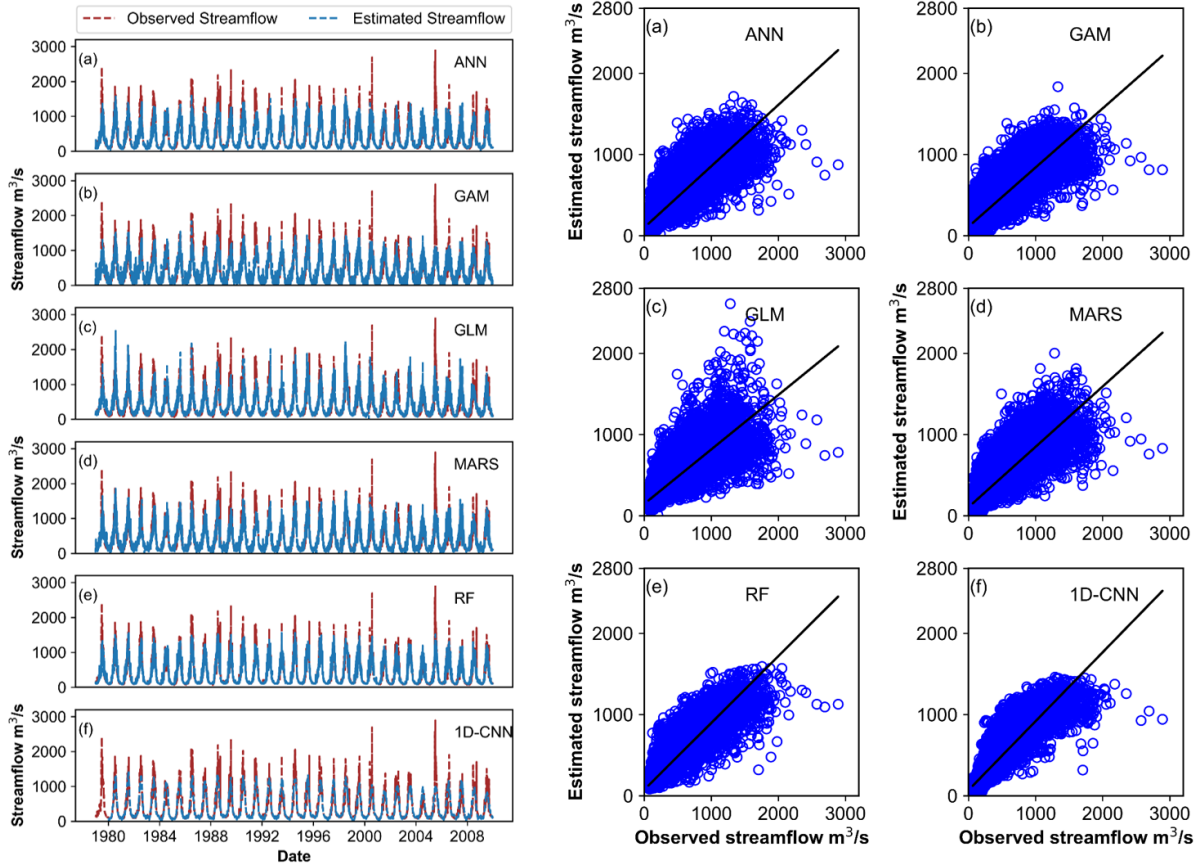
984



985

986 **Figure 5: Comparison of observed and simulated streamflow for all six models (ANN, GAM, GLM, MARS, RF and**
987 **1D-CNN) under rainfall scenarios R_0 .**

988



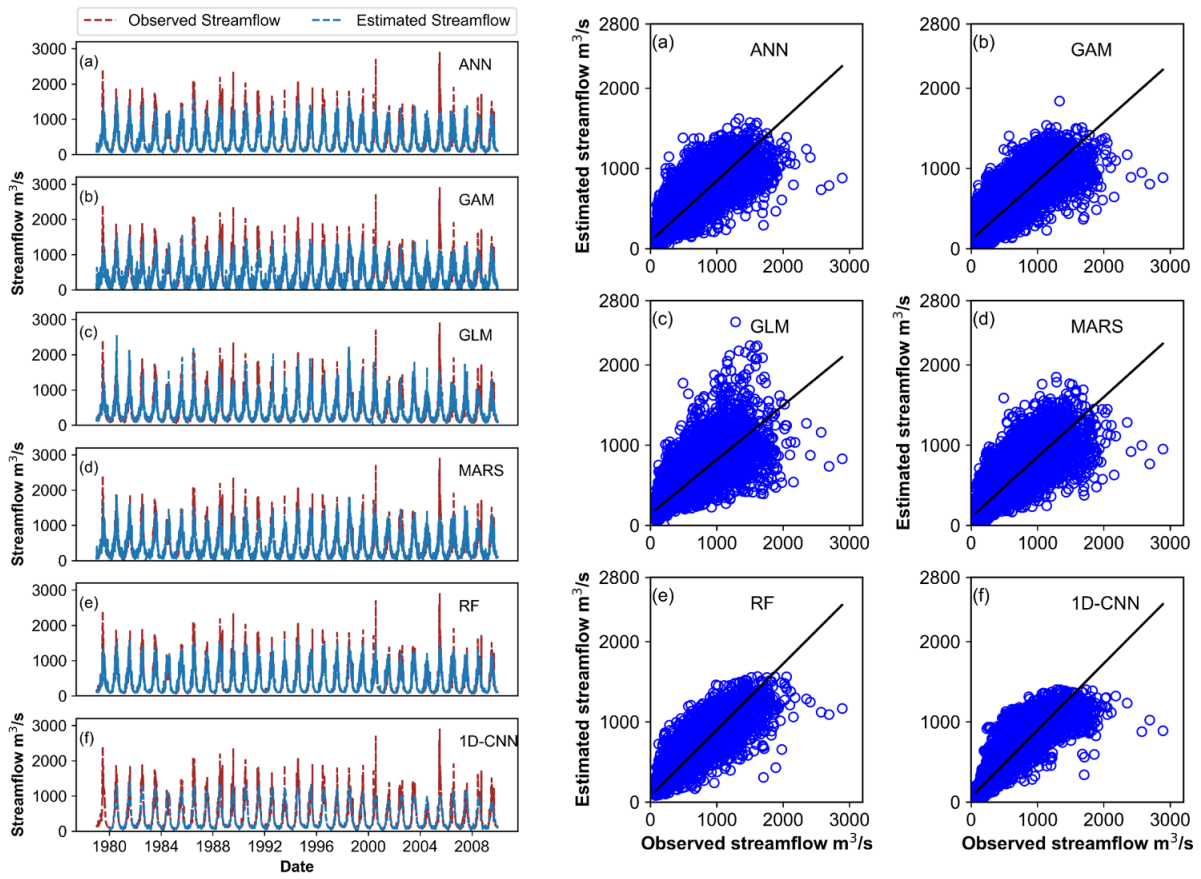
989

990 **Figure 6: Comparison of observed and simulated streamflow for all six models (ANN, GAM, GLM, MARS, RF and**
 991 **1D-CNN) under rainfall scenarios R_1 .**

992

993

994



995

996 **Figure 7: Comparison of observed and simulated streamflow for all six models (ANN, GAM, GLM, MARS, RF and**
 997 **1D-CNN) under rainfall scenarios R₂.**

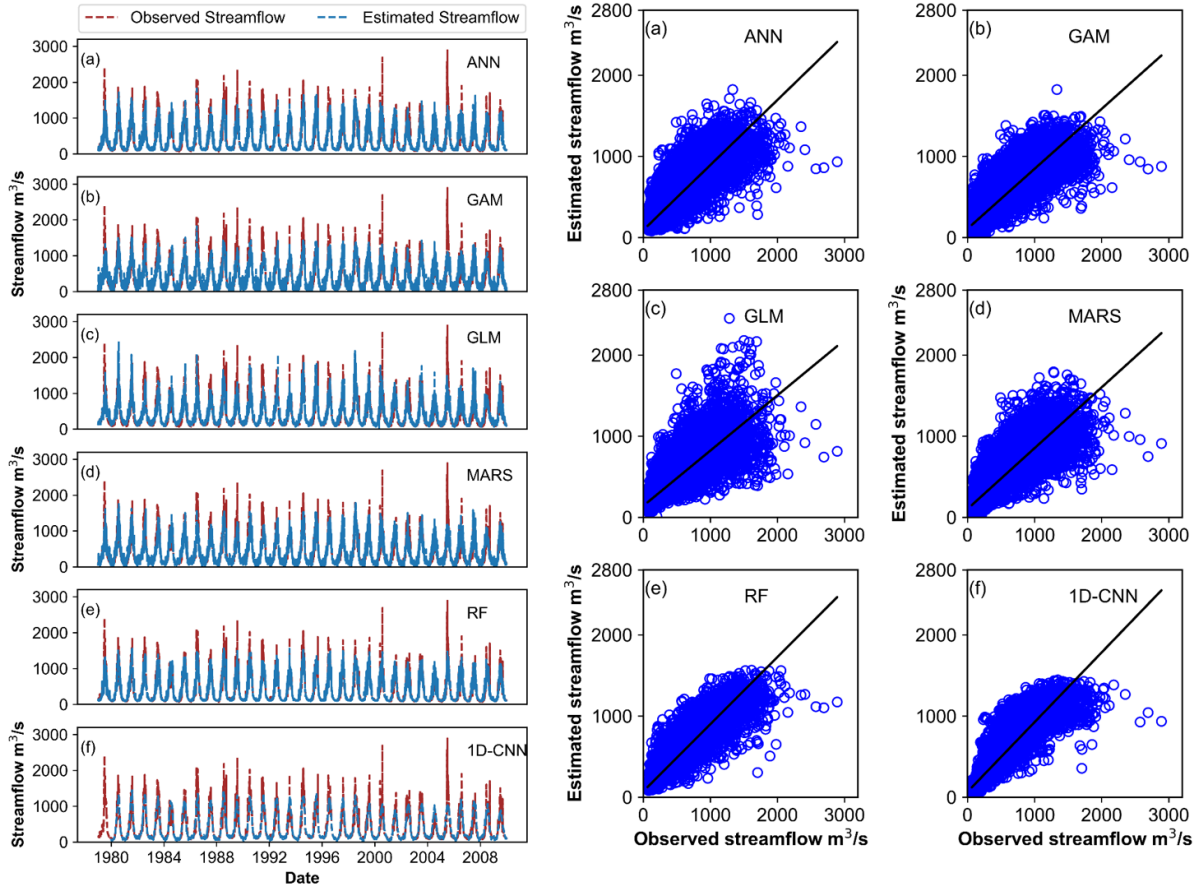
998

999

1000

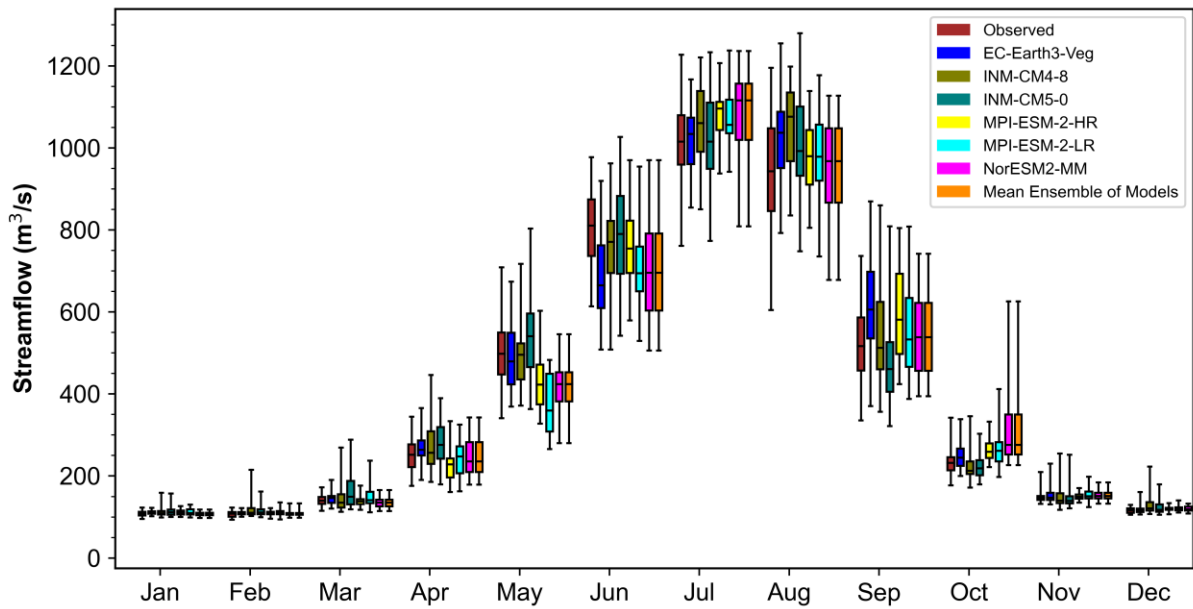
1001

1002



1003
 1004
 1005
 1006

Figure 8: Comparison of observed and simulated streamflow for all six models (ANN, GAM, GLM, MARS, RF and 1D-CNN) under rainfall scenarios R₃.



1007

1008 **Figure 9: Box-plot comparing observed and CMIP6-GCMs (mean ensemble of models) simulated streamflow for**
 1009 **various months of the year, derived over the period of 1979–2009. The line inside the box denotes the median values**
 1010 **of streamflow, while the upper and lower whiskers indicate the highest and minimum values, respectively.**

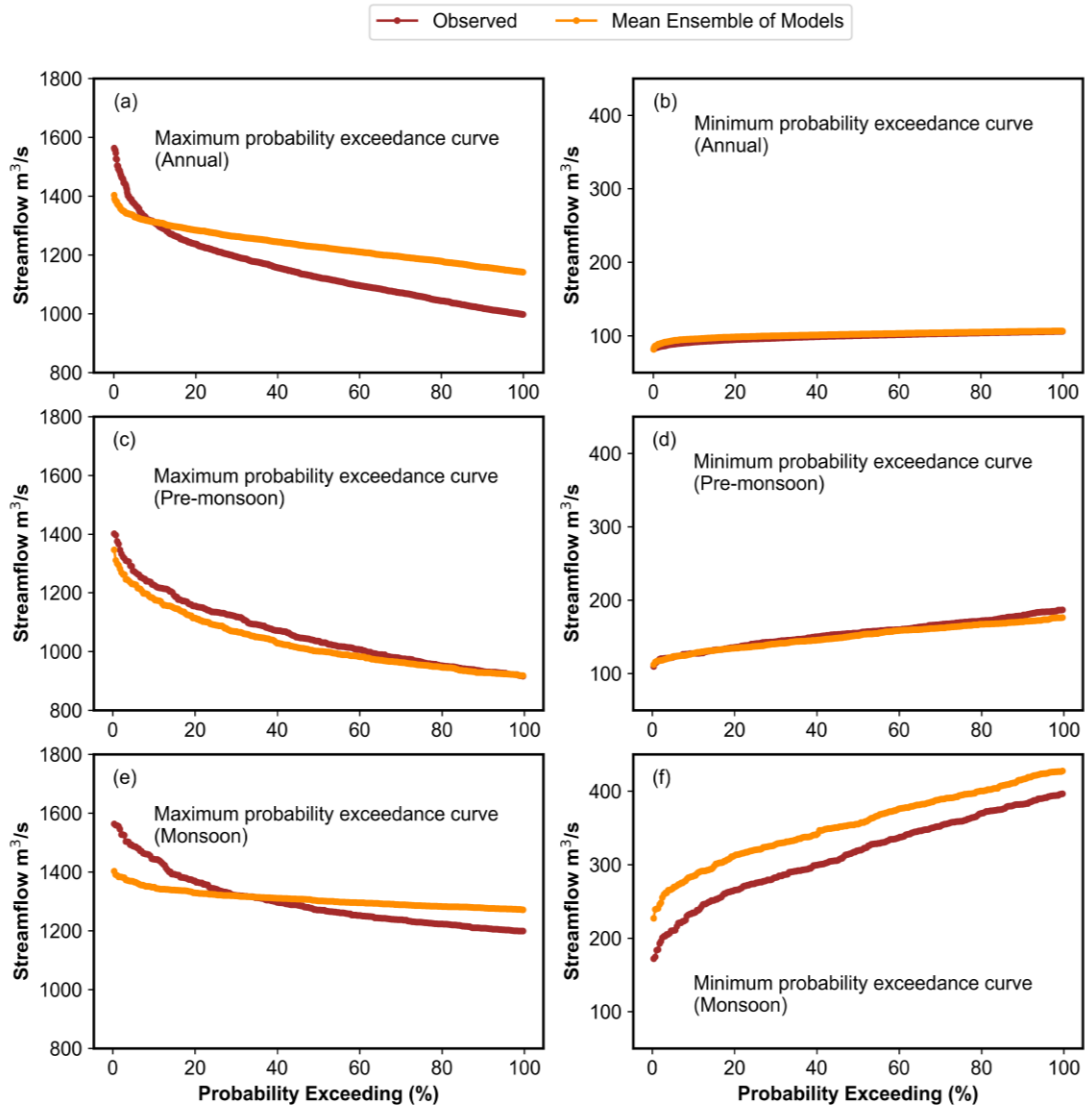
1011

1012

1013

1014

1015

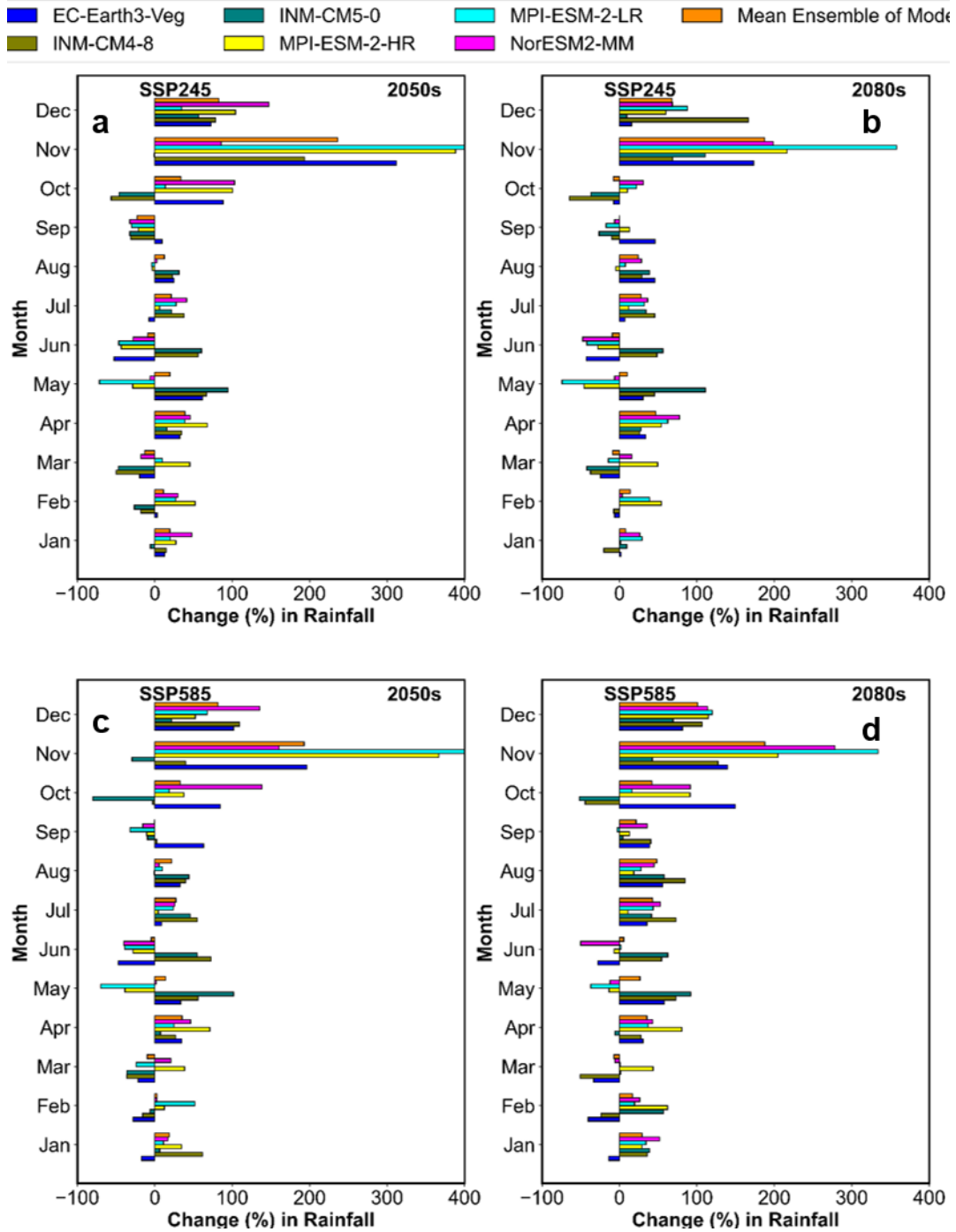


1016

1017 **Figure 10: Probability exceedance curves developed using 10% of the highest and lowest flows from the observed and**
 1018 **CMIP6-GCMs (mean ensemble of models) over the time span of 1979–2009 for annual and seasonal (pre-monsoon**
 1019 **and monsoon) flows.**

1020

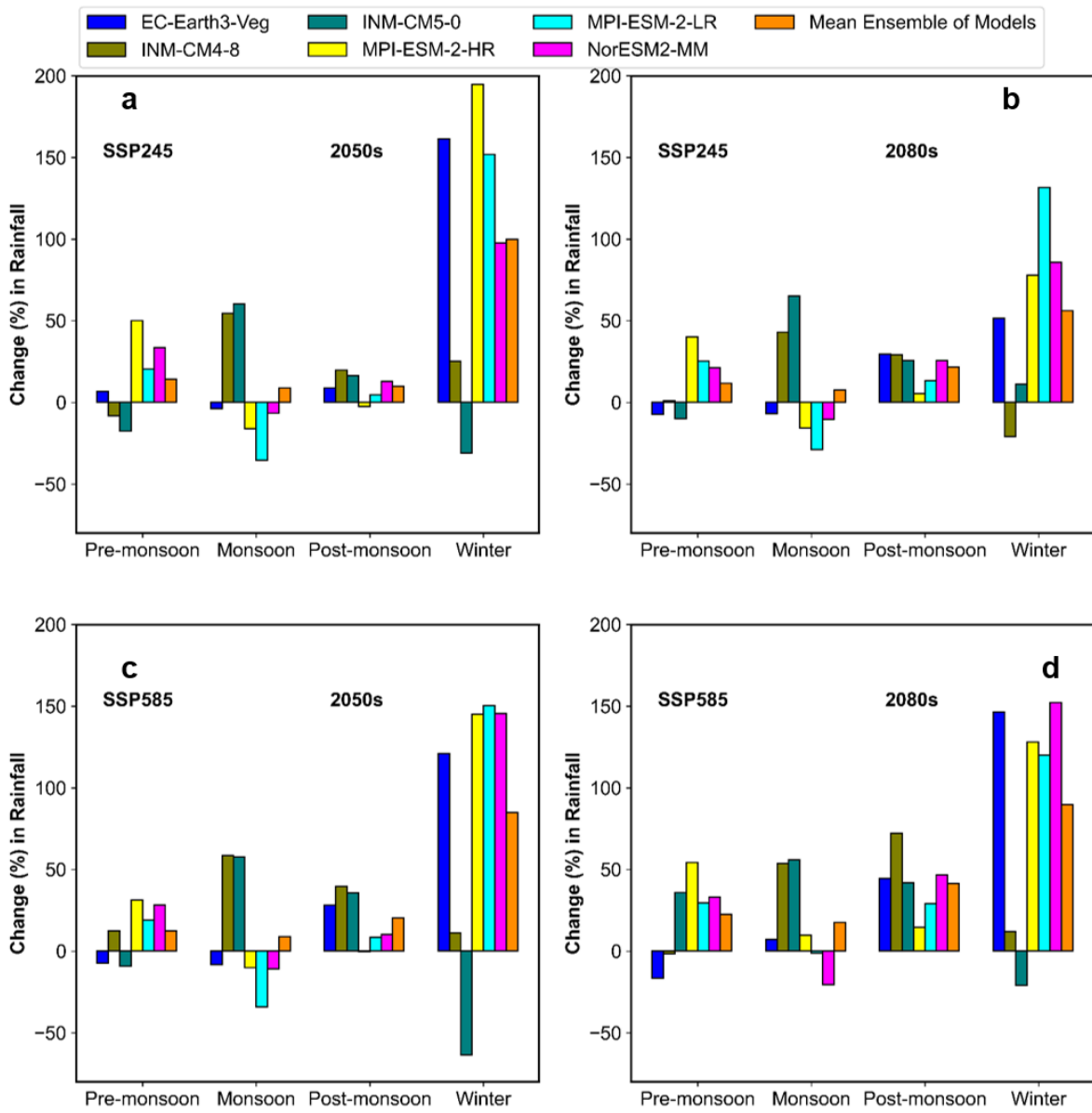
1021



1023

1024 Figure 11: Projected change in mean monthly rainfall in the sub-basin using different CMIP6-GCMs under SSP245
 1025 and SSP585 scenarios in the 2050s (Fig.11a and Fig.11b) and 2080s (Fig.11c and Fig.11d).

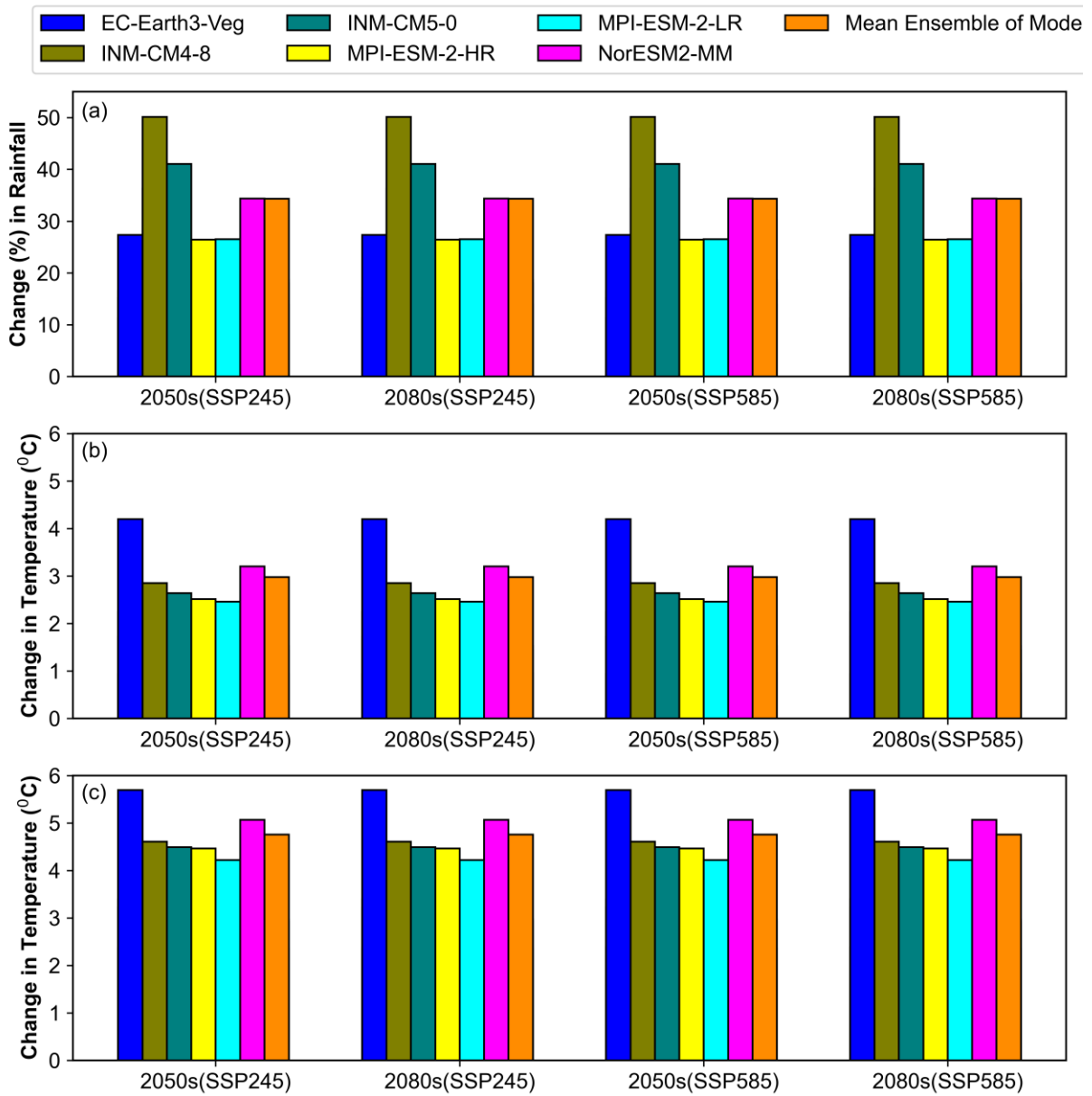
1026



1028

1029 **Figure 12: Projected change in mean seasonal rainfall in the sub-basin using different CMIP6-GCMs under SSP245**
 1030 **and SSP585 scenarios in the 2050s (Fig.12a and Fig.12c) and 2080s (Fig.12b and Fig.12d).**

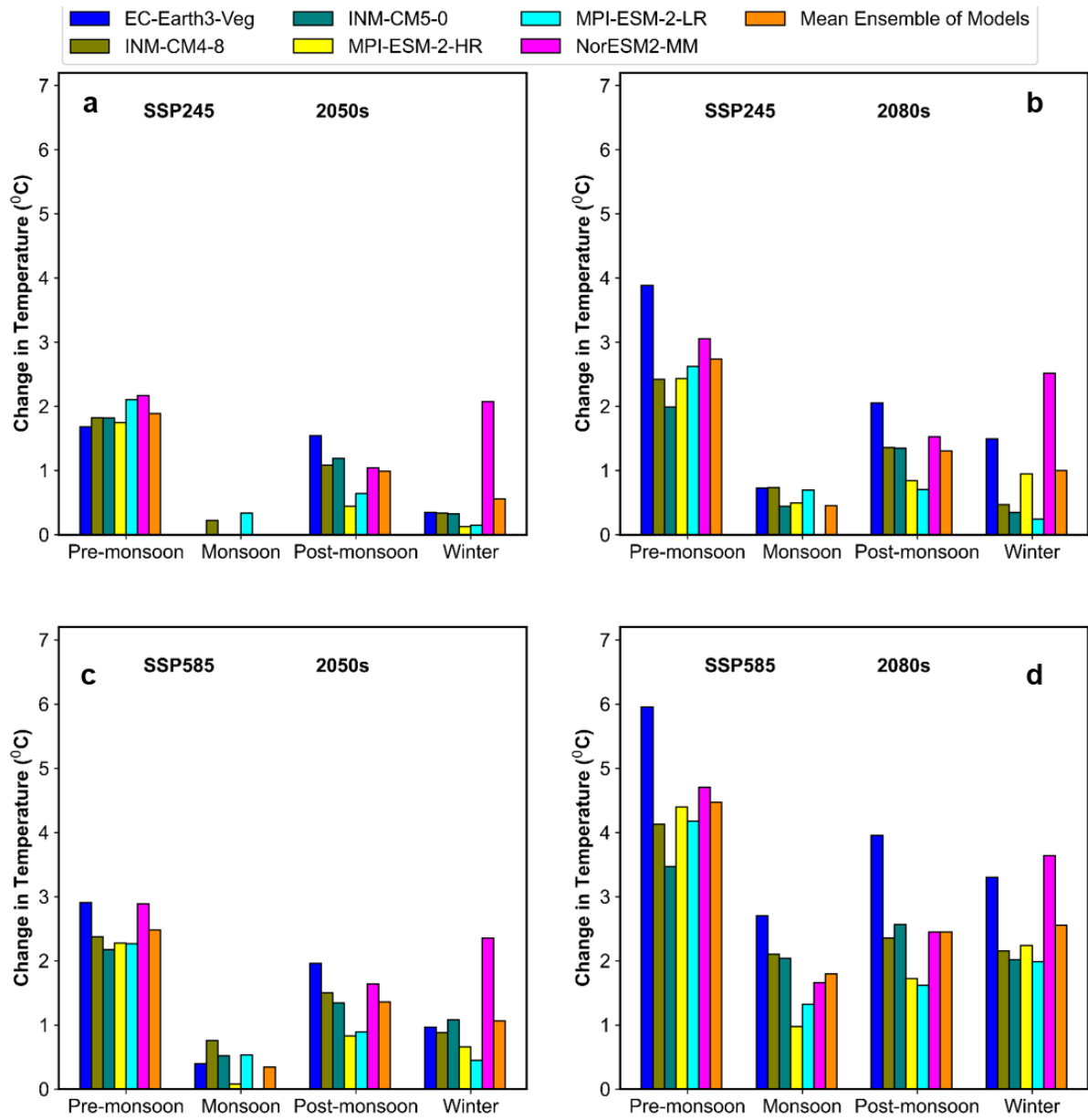
1031



1033

1034 **Figure 13: Projected changes in mean annual rainfall (Fig.13a), T_{max} (Fig.13b) and T_{min} (Fig.13c) in the sub-basin**
 1035 **using different CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s and 2080s.**

1036

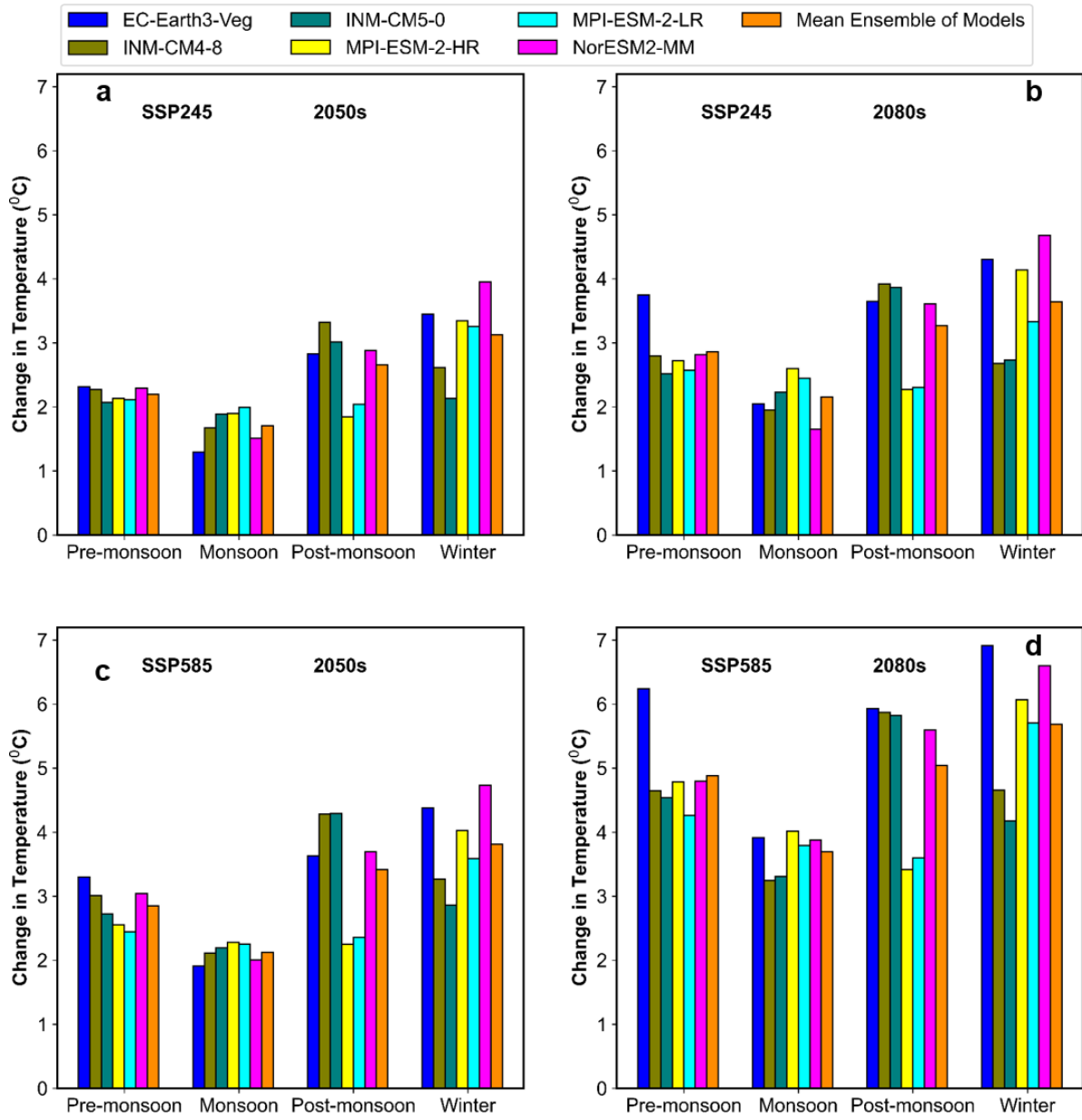


1037

1038

1039

Figure 14: Projected change in mean seasonal maximum temperature (T_{max}) in the sub-basin using different CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s (Fig.14a and Fig.14c) and 2080s (Fig.14b and Fig.14d).

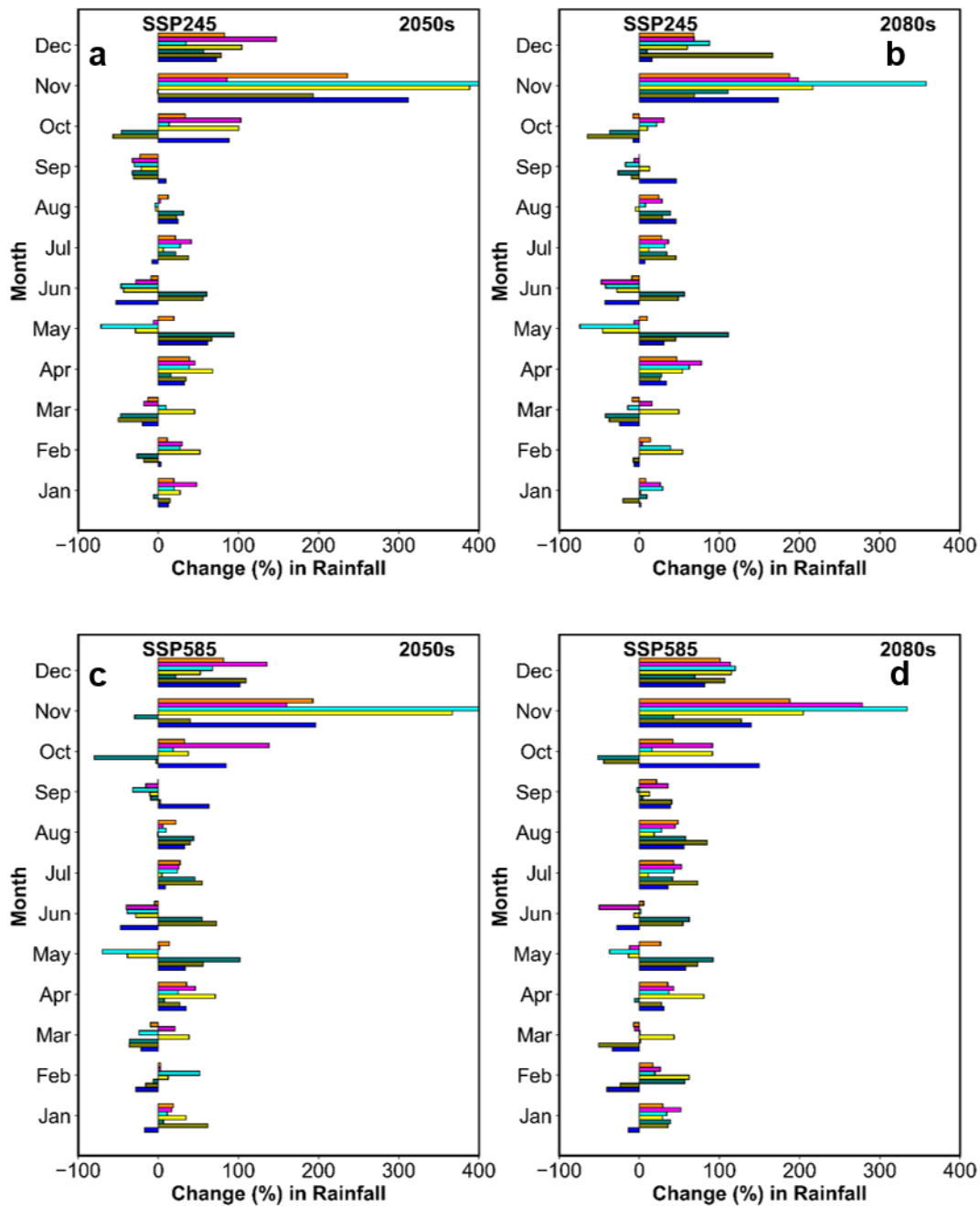


1040

1041

1042

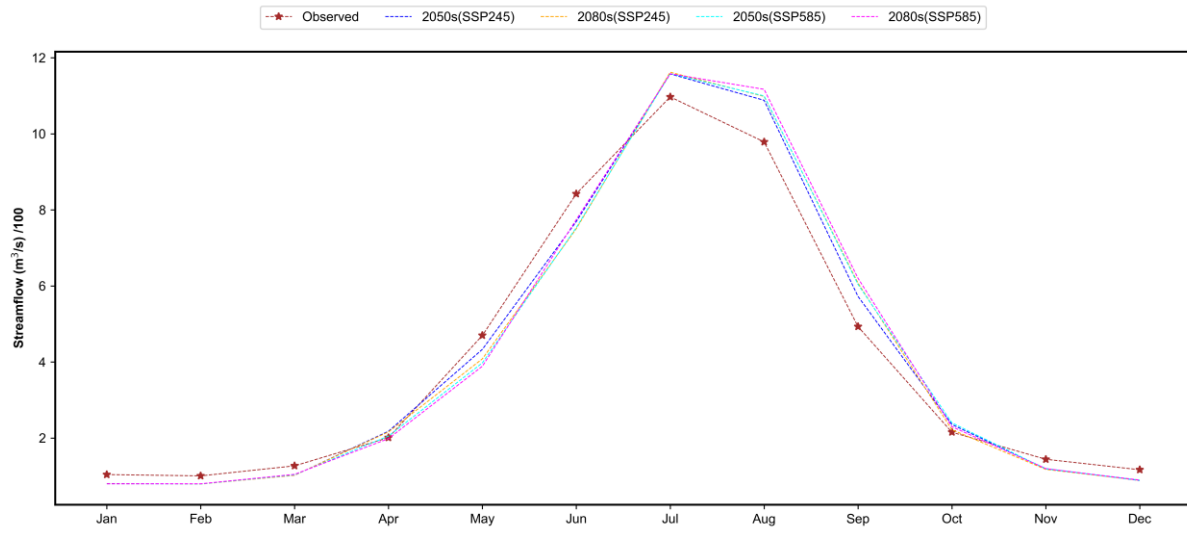
Figure 15: Projected changes in mean seasonal minimum temperature (T_{min}) in the sub-basin using different CMIP6-GCMs under SSP245 and SSP585 scenarios in the 2050s (Fig.15a and Fig.15c) and 2080s (Fig.15b and Fig.15d).



1043

1044 Figure 16: Predicted change in monthly streamflow pattern of the Sutlej River with respect to the reference period
 1045 (1979-2009) in 2050s (Fig. 16a and Fig. 16b) and 2080s (Fig. 16c and Fig.16d) under SSP245 and SSP585 for different
 1046 CMIP6-GCMs.

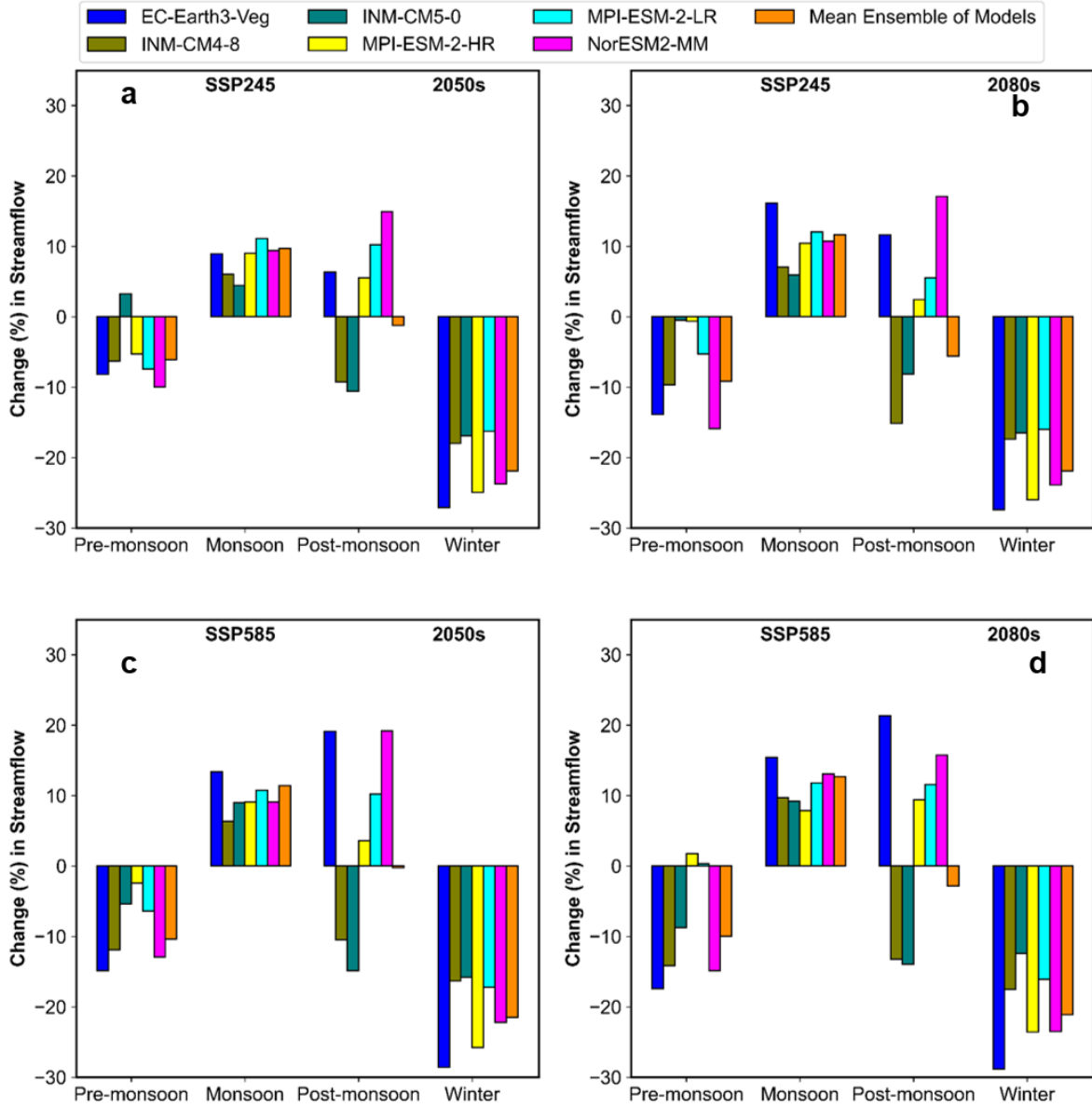
1047



1048

1049 **Figure17: Comparison of monthly observed (1979-2009) and projected discharge of the multi-model ensembles for**
 1050 **period 2050s and 2080s under SSP245 and SSP585 scenarios.**

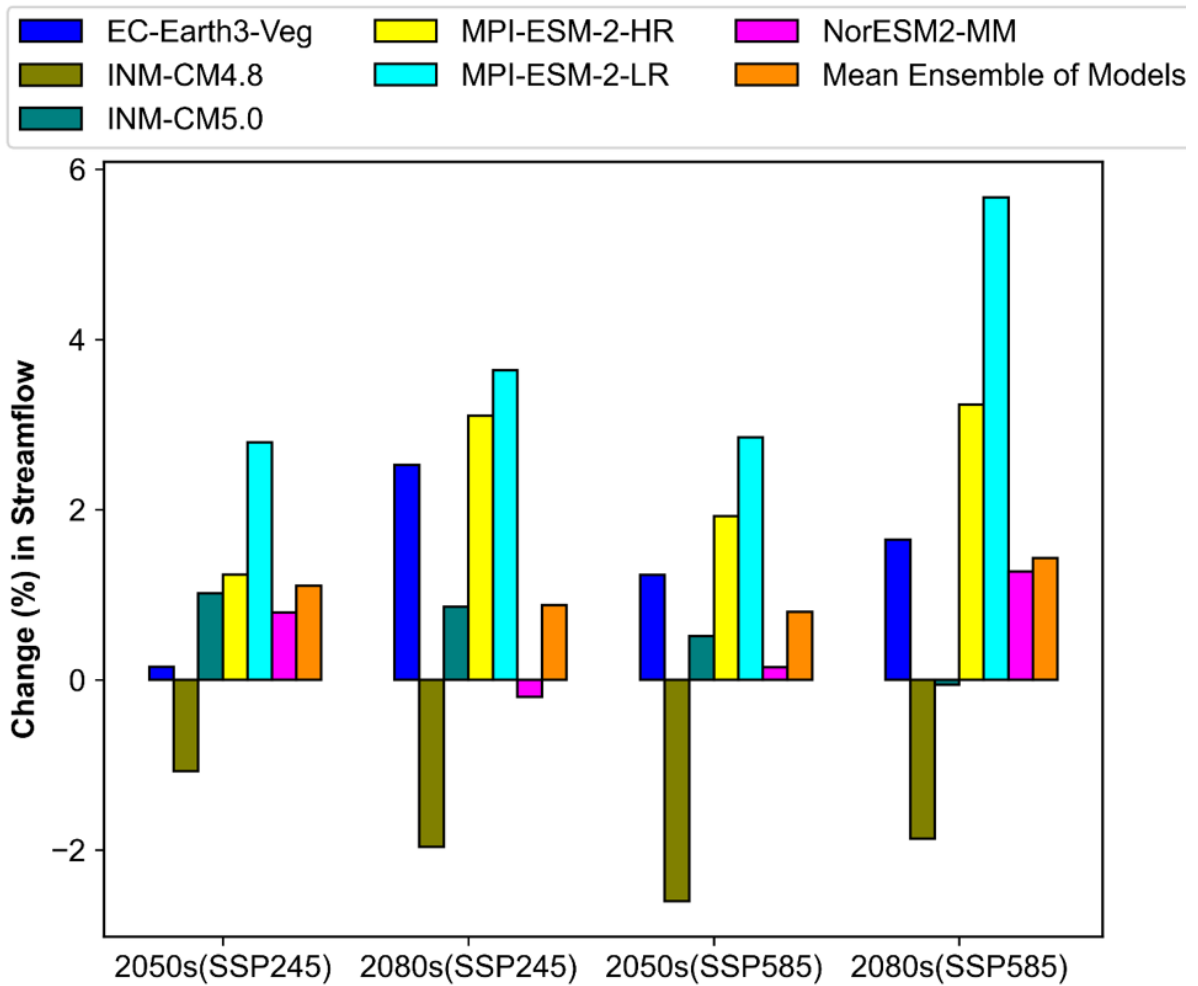
1051



1052

1053 **Figure 18: Predicted change in seasonal streamflow pattern of the Sutlej River with respect to the reference period**
 1054 **(1979-2009) in 2050s (Fig.18a and Fig.18c) and 2080s (Fig.18c and Fig.18d) under SSP245 and SSP585 for different**
 1055 **GCMs.**

1056



1057

1058 **Figure 19: Predicted change in mean annual streamflow of the Sutlej River with respect to the reference period**
 1059 **(1979-2009) in 2050s and 2080s under SSP245 and SSP585 for different GCMs.**

1060

1061 **Table 1:** Characteristics of the study catchment over the evaluation period of 1979–2009

| Parameters | Details |
|--|--|
| Details of the sub-catchment | |
| Drainage area of the sub-catchment (km ²) | 2457 km ² |
| Altitude | ~500-5000 m |
| Slope | 0-80° |
| Geology | Granite, Jutogh formation and Chail/Salkhala/Hemanta formation |
| Soil | Dystric cambisols, dystric regosols, and eutric fluvisols. |
| Streamflow measured at the outlet (Kasol) of the sub-catchment | |
| Average of annual streamflow | 411.2 m ³ /s |
| Minimum streamflow (daily) | 64.3m ³ /s |
| Maximum streamflow (daily) | 2891 m ³ /s |
| Standard deviation (SD) of annual streamflow | 1750.7m ³ /s |
| Coefficient of variation (CV) of annual streamflow | 0.1m ³ /s |
| Rainfall integrated over the sub-catchment | |
| Average of annual rainfall | 1001.3mm |
| Average of monsoon rainfall (July-September) | 403.0mm |
| Average of winter rainfall (December-March) | 277.3mm |
| Temperature integrated over the sub-catchment | |
| Average annual maximum temperature (T _{max}) | 28.3°C |
| Average annual minimum temperature (T _{min}) | 13.9°C |

1062

1063

Table 2: The information on hyper parameters used for estimating model parameters

| Model Name | Hyperparameter | Values |
|----------------------------------|--|---|
| Artificial Neural Network (ANN) | build_fn, warm_start, random_state, optimizer, loss, metrics, batch_size, validation_batch_size, verbose, callbacks, validation_split, shuffle, run_eagerly, epochs, | value = build_regressor value = False value = None value = rmsprop value = None value = None value = 64 value = None value = 1 value = None value = 0.0 value = True value = False value = 500 |
| Generalized Additive Model (GAM) | formula, family, data, weights, subset , na.action,offset, method, optimizer, control, scale, select, knots, sp, min.sp, H, gamma, fit, paraPen, G, drop.unused.levels, drop.intercept, discrete, | value = None value = gaussian() value = list() value = Null value = Null value = Null value = "GCV.Cp" value = c("outer","newton") value = list(), value = 0 value = False value = Null value = Null value = Null value = Null, value = 1 value = True value = Null value = Null value = True value = Null value = False |
| Generalized Linear Model (GLM) | endog, exog, | value = 1D value = 1D |

| | | |
|---|---|---|
| | family, offset, exposure, freq_weights, var_weights, missing, | value = sm.families.Gaussian(sm.families.links.log()) value = None value = None value = None value = None value = str |
| Multivariate Adaptive Regression Splines (MARS) | max_terms, max_degree , allow_missing, penalty, endspan_alpha, endspan, minspan_alpha, minspan, thresh , zero_tol, min_search_points, check_every, allow_linear, use_fast, fast_K, fast_h, smooth, enable_pruning, feature_importance_type, feature_importance_type, | value = 20 value = 3 value = False value = 3.0 value = 0.005 value = -1 value = 0.005 value = -1 value = 0.001 value = 1e-12 value = 100 value = -1 value = True value = False value = 5 value = 1 value = False value = True value = None value = 0 |
| Random Forest (RF) | n_estimators, criterion, max_depth, min_samples_split, min_samples_leaf, min_weight_fraction_leaf, max_features, max_leaf_nodes, min_impurity_decrease, | value=500 value="squared_error" value=None value = 2 value = 5 value = 0.0 value = auto value = None value = 0.0 |
| 1-Dimensional Convolution neural network | Conv1D_filter, | Value = 64 |

| | | |
|----------|--|---|
| (1D-CNN) | Conv1D_kernel_size, Conv1D_pool_size, Learning rate, Epic, Batch size, loss | Value = 2 Value =2 Value = 0.0001 Value = 30 Value = 280 Value = MSE |
|----------|--|---|

1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075