CrossMark

Ensembling Downscaling Techniques and Multiple GCMs to Improve Climate Change Predictions in Cryosphere Scarcely-Gauged Catchment

Muhammad Azmat[1] • **Muhammad Uzair Qamar**[2] •
Shakil Ahmed[1] • **Muhammad Adnan Shahid**[3] •
Ejaz Hussain[1] • **Sajjad Ahmad**[4] •
Rao Arsalan Khushnood[1]

Abstract Future projections of climate variables are the key for the development of mitigation and adaptation strategy to changing climate. However, such projections are often subjected to large uncertainties which make implementation of climate change strategies on water resources system a challenging job. Major uncertainty sources are General Circulation models (GCMs), post-processing and climate heterogeneity based on catchment characteristics (e.g. scares data

✉ Muhammad Azmat
azmat@igis.nust.edu.pk

Muhammad Uzair Qamar
muhammad.uzair@uaf.edu.pk

Shakil Ahmed
shakilahmad@nice.nust.edu.pk

Muhammad Adnan Shahid
muhammad.shahid@uaf.edu.pk

Ejaz Hussain
ejaz@igis.nust.edu.pk

Sajjad Ahmad
pc277@yahoo.com

Rao Arsalan Khushnood
arsalan.khushnood@nice.nust.edu.pk

[1] School of Civil and Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), H-12, Islamabad, Pakistan

and high-altitude). Here we presents the comparisons between different GCMs, statistical downscaling and bias correction approaches and finally climate projections, with the integration of gridded and converted (monthly to daily) data for a high-altitude, scarcely-gauged Jhelum River basin, Pakistan. Current study relies on climate projections obtained from factorial combination of 5-GCMs, 2 statistical downscaling and 2 bias correction methods. In addition, we applied bias corrected APHRODITE, converted daily data using MODAWEC model and observed data. Further, five GCMs (CGCM3, HadCM3, CCSM3, ECHAM5 and CSIRO-MK3.5) were tested to scrutinize two suitable GCMs integrated with Statistical Downscaling Model (SDSM) and Smooth Support Vector Machine (SSVM). Results illustrate that the CGCM3 and HadCM3 were suitable GCMs for selected study basin. Both downscaling techniques are able to simulate precipitation, however, SSVM performed slightly better than SDSM. We found that the integration of CGCM3 with SSVM (SSVM-CGCM3) generates precipitation and temperature better than the CGCM3 (SDSM-CGCM3) and HadCM3 (SDSM-HadCM3) with SDSM. Furthermore, the low elevation stations were influenced by monsoon, significantly prone to rise in precipitation and temperature, while high-altitude stations were influenced by westerlies circulations, less prone to climate change. The projections indicated rise in basin-wide annual precipitation by 25.51, 36.76 and 45.52 mm and temperature by 0.64, 1.47 and 2.79 °C, during 2030s, 2060s and 2090s, respectively. The methods and results of this study can be adopted to evaluate climate change implications in the catchments of characteristics similar to Jhelum River basin.

1 Introduction

The substantial evidences justify existence of global warming, causing changes in the hydrological cycle, both regionally and globally, therefore, it is essential to identify any potential implications of climatic variations on hydrological process including extreme events e.g. floods, droughts and even expedition of snow- and glacier-melt (Azmat et al. 2016; Chen et al. 2012). To identify possible fluctuations in hydrological regime subjected to climate changes, several steps are needed to be taken for the climate projections and interconnectivity of steps called as "model chain". Previous studies explored that each step of model chain influences simulation, thereby contributes to uncertainty (Addor et al. 2014). Therefore, for each model chain step, research community developed different models and approaches with limitations. Steps of model chain comprise of emission scenarios e.g. GCMs, downscaling approaches/models, bias corrections and eventual projections.

GCMs are the principal source of projections (Dibike and Coulibaly 2005), however, direct use of GCMs is inhibited by the coarse spatial resolution which represents

---

[2]   Department of Irrigation and Drainage, Faculty of Agricultural Engineering & Technology, University of Agriculture, Faisalabad, Pakistan

[3]   Water Management Research Center, University of Agriculture Faisalabad, Faisalabad, Pakistan

[4]   Department of Civil Engineering, Mirpur University of Science and Technology, New Mırpur, Pakistan

significant obstacle in projection of climate variables at basin level. Conversion from coarse to fine is, therefore, essential for climate change studies to reduce uncertainties (Meenu et al. 2013; Wang et al. 2013). The downscaling of climate variables, particularly precipitation, was reported as highly uncertain by several researchers which makes the selection of appropriate downscaling techniques and GCMs, tremendously essential (Hagemann et al. 2013; Wang et al. 2013). Taking this into consideration, several statistical downscaling approaches/models have been developed to overcome uncertainty problems (Chen et al. 2012). Lin et al. (2017) employed k-nearest neighbor method (KNN), a spatiotemporal statistical downscaling technique with integration of Bergen Climate Model (BCM2.0) and third generation Coupled Global Climate Model (CGCM3.1) to assess change in hourly precipitation and stated that this technique was efficient tool to downscale hourly precipitation data over Dahan River in northern Taiwan. Chen et al. (2012) evaluated different statistical downscaling techniques and GCMs for the downscaling of climate variables, subsequently, utilized to examine the impact of climate change on runoff of the upper Hanjiang basin, China. However, different downscaling techniques with integration of GCMs are always challenging in high-altitude and scarcely-gauged catchments. Typically, high-altitude catchments confront two kinds of data issues, first, unavailability of daily climate data at appropriate length (Liu et al. 2009); second, the quality of available precipitation data (Azmat et al. 2016). Such issues make downscaling from global into local scale climate variables very challenging and misrepresentation of ground data. Therefore, in this study we examine different GCMs, downscaling and bias correction approaches to rely on suitable combination of model chain for future projections. To best our knowledge, most of the studies were conducted in low-elevation regions where data issues were not extensive. Even very few studies were carried out by considering uncertainty issues following comparative analysis. The specific objective of this study was to examine the efficiency of five different GCMs, two well-known statistical downscaling models, a multiple regression hybrid statistical downscaling model (SDSM) and a supervised learning based smooth support vector machine (SSVM) and finally adaptation of two bias correction techniques, distribution mapping and linear scaling, for the projected data correction. Subsequently, we examined potential changes in precipitation and temperature under future scenarios by adopting suitable method/approach/model in model chain to overcome uncertainty propagation, in Jhelum River basin, Pakistan.

2 Study Area

This study was conducted in the Jhelum River basin, a trans-boundary river catchment shared by Pakistan and India, located in Western Himalayas and draining the southern slopes of the Himalayan range. The catchment area upstream of the Mangla dam (only storage structure) is 33,867 km^2 (Fig. 1) and the estimated mean elevation of the study catchment is approximately 2094 m a.s.l. This catchment is influenced by two different precipitation patterns; the westerlies circulations and monsoon. Therefore, twofold precipitation peak occurs as recorded at most of the stations (i.e. Narran, Murree and Kotli) in the Jhelum River basin. The mean annual precipitation ranges from 485 mm at Astore to 1178 mm at Narran, respectively, while the mean annual temperature varies from 6.04 °C at Narran and 25.39 °C at Gujjar Khan station. Some key characteristics of the
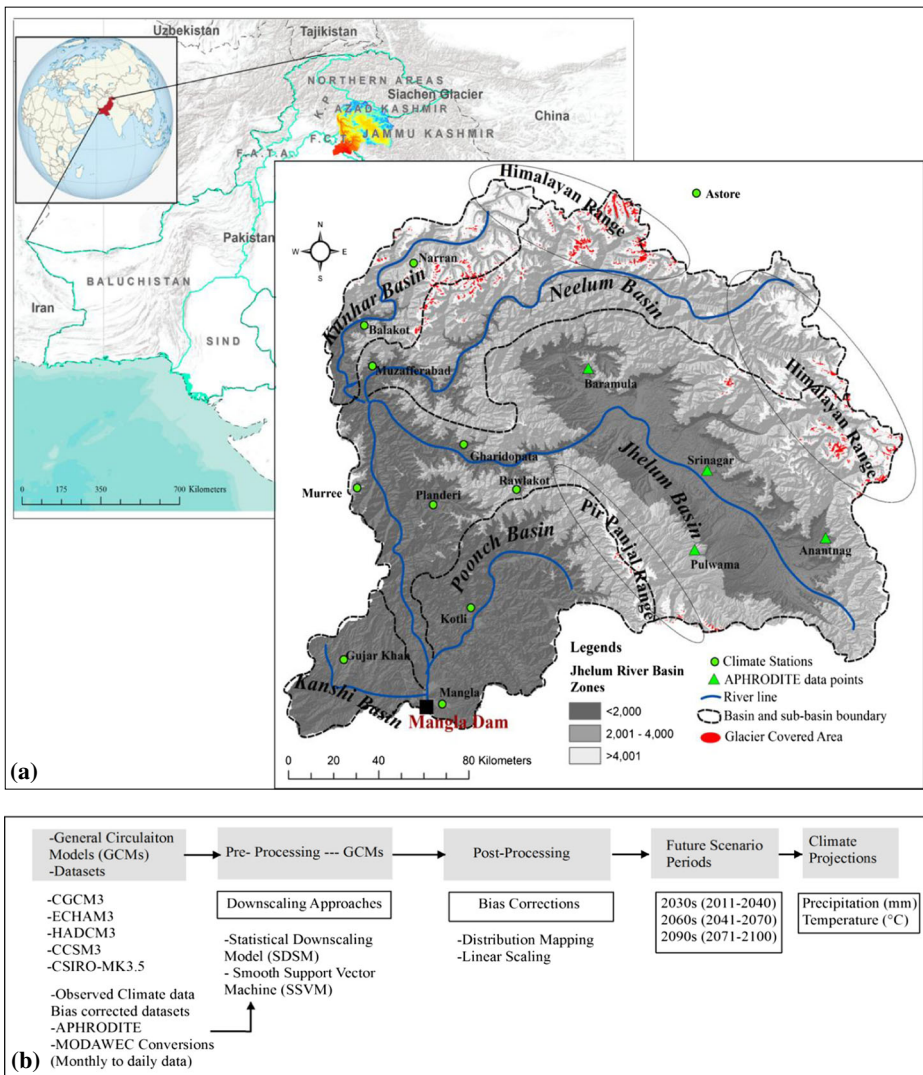
Fig. 1 (a) Geographical location of Pakistan including study area along with climate stations and data points of the APHRODITE over un-gauged part of the catchment, (b) Flowchart of methodology. The shaded boxes depict the major elements of model chain and white boxes show particulates of the element. Several methods were adopted for each element as listed below the box of corresponding element and process

climate stations and data points for the un-gauged part of the basin and altitude zones of the Jhelum River basin are given in Table 1 and shown Fig. 1a.

3 Datasets

Measured daily and monthly precipitation, T_{min} (minimum temperature) and T_{max} (maximum temperature) of 11 stations provided by Water and Power Development Authority

Table 1 Some key characteristics of different climate stations located in Jhelum River basin

Climate stations	Elevation (m)	Mean precipitation (mm)	T_{mean} (°C)	Climate variables temporal range	
				Daily	Monthly
Mangla	282	869	23.48	1970–2013	1950–2013
Planderi	1402	1472	17.11	1962–2006	1960–2013
Rawlakot	1683	1426	16.18	1960–2009	1960–2013
Muzaffarabad	731	1431	20.51	1961–2006	1955–2013
Astore	2394	485	9.72	1953–2013	1953–2013
Gharidopata	1223	1499	19.36	1970–2009	1955–2013
Kotli	1373	1227	22.13	1970–2007	1952–2013
Narran	2363	1278	6.04	1961–2013	1961–2013
Balakot	1014	1742	18.57	1970–2010	1960–2013
Gujar Khan	501	838	25.39	1970–2013	1961–2013
Murree	2203	1774	12.51	1970–2013	1960–2013
Climate data points for ungauged part of the catchment					**Data source**
Baramula	1842	987	12.71	1954–2007	APHRODITE
Srinagar	1658	778	13.78	1954–2007	APHRODITE
Pulwama	2721	1171	6.94	1954–2007	APHRODITE
Anantnag	1728	1086	12.89	1954–2007	APHRODITE

(WAPDA) and Pakistan Meteorological Department (PMD) over different available control periods as mentioned in Table 1. Since, the Jhelum River is a trans-boundary river catchment, the catchment located in India, we used daily precipitation and mean temperature (T_{mean}) provided by Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) project with 0.25° × 0.25° resolution (Yatagai et al. 2012). This data have already been successfully used by several researchers (Kumar et al. 2016; Lutz et al. 2014) in high Asia mountains (e.g. Himalayan, Hindu-Kush and Karakoram Range). The daily precipitation (APHRO_MA_V1003R1) and mean temperature (APHRO_MA_T_{AVE}_V1204R1) data products were downloaded and corrected (http://chikyu.ac.jp/precip) for 1961–2004 and 1961–2007, respectively.

The National Center for Environmental Protection (NCEP/NCAR) provided daily re-analysis atmospheric predictors for period of 1961–2000, used for the calibration and validation of downscaling models. Furthermore, the output of five different GCMs, Hadley Center's (HadCM3), Coupled Global Climate Model (CGCM3), Community Climate System Model version 3 (NCAR-CCSM3), MPI-ECHAM5 and CSIRO-MK3.5 were used to scrutinize the suitable GCMs for the Jhelum River basin.

3.1 Conversion of Monthly into Daily Data

The availability of long term daily climate data are key for downscaling of climate variables, however, the selected study area is subjected to twofold scarcely-gauged issues, first, trans-boundary catchment, second, limited daily data over a feasible temporal scale and range. While daily data at most of the stations were collected after 1970 and the only monthly records were available from 1960 or earlier (Table 1). To deal with this issue, monthly observed climate data was converted into daily data by using MODAWEC model. The MODAWEC model is a Fortran language based parametric weather generator. This model has been applied successfully by Liu et al. (2009) to convert monthly T_{min}, T_{max} (°C) and precipitation (mm) into daily time scale. It is a parametric weather generator, which uses rainfall as a trigging variable to generate rainfall amount and event occurrence independently adopting stochastic relationship, while, temperature variables generated based on the generated rainfall in first step. The MODAWEC model utilizes first-order Markov chain to differentiate dry- or wet-days based on transitional probability matrices. For the derivation of daily precipitation, the first-order Markov chain follows two states (wet or dry) with transitional probabilities of "wet day follow a dry-day" and "dry-day follows previous wet-day". The transition from wet to dry or dry to wet is governed by two transitional probabilities; first, in case of wet-day, the modified exponential distribution provide approximation of daily precipitation and second, the monthly given precipitation corrects final daily precipitation. The probability of wetness on corresponding day is based on dryness or wetness of previous day. The detailed discussion of MODAWEC model are provided by Liu et al. (2009). Further, the methodology flow chart of aforementioned equations is presented in Fig. S1 (supplement material). For temperature (minimum and maximum) generation, the MODAWEC employs algorithm designed by Richardson (1981) to approximate initial temperature values based on correlation with precipitation. The multivariate normal distribution generates the residual of daily T_{min} and T_{max}. Final corrected daily T_{min} and T_{max} were obtained by the correcting initial temperature values with help of monthly mean temperature.

MODAWEC utilizes standard deviations (SD) of T_{min} and T_{max} as input, however, in case of absence of SD the long term monthly extreme may substitute SD.

In this study, the monthly precipitation and temperature variables (T_{min} and T_{max}) are available at large temporal range (i.e. 1960–2013) for most climate stations located in the study basin, the monthly extremes of T_{min} and T_{max} were extracted from the observed data. The number of wet days for each year was obtained from Climatic Research Unit (CRU) available at 50×50 km spatial resolution for 1901–2000. The MODAWEC model is freely available on http://www.eawag.ch/en/department/siam/software/ with complete program version and user help manual. Further description of MODAWEC is also available in supplement material.

4 Methods

We used two different downscaling methods; Statistical Downscaling Model (SDSM) and Smooth Support Vector Machine (SSVM), to scrutinize two most suitable GCMs out of those five (CGCM3, HadCM3, CCSM3, ECHAM5 and CSIRO-MK3.5), for the projection of reliable climate data. Downscaling was performed using 40 years' baseline (1961–2000) observed data (T_{min} and T_{max} and precipitation). We calibrated and validated both downscaling techniques (SDSM and SSVM) for 30 years (1961 to 1990) and 10 years (1991 to 2000), respectively, using NCEP\NCAR reanalysis data. Subsequently, suitable GCMs for Jhelum River basin were utilized with downscaling of climate variables up to the end of the current century (2099). The generated future climate data were further bias corrected using distribution mapping and linear scaling methods with help of observed/control period of 1961–2000 (baseline) data. The change in future climate variables was analyzed for three time slices of 2011–2040 (2030s), 2041–2070 (2060s) and 2071–2099 (2090s). Further, detailed flow chart for the methodology adopted in this study is presented in Fig. 1b.

Since, precipitation data are very sensitive to hydrological predictions and are more important for the efficiency of downscaling techniques for the projection of precipitation data. Therefore, only precipitation data were utilized for the evaluation of downscaling techniques and analysis of GCMs using sophisticated statistical descriptors such as mean (mm/d), SD (mm/d), percentile (95%), percent wet days (%), maximum wet spell length (days), dry spell length (days) and 5-day maximum rainfall. In this study, 1 mm was selected as wet-threshold.

4.1 Statistical Downscaling Methods

4.1.1 Statistical Downscaling Model (SDSM)

Statistical Downscaling Model; a multiple regression hybrid downscaling approach, was developed by Wilby et al. (2002) and mathematical relationships are explained by Wilby and Dawson (2013). Model develops an empirical relationship between GCM predictors and local scale predictands to downscale climate variables. In this study, two major steps were taken to achieve the downscaling of climate variables. First, empirical relationship was established between predictors and predictands through processes such as quality control, screening, transformation and predictors' selection, calibration, and the weather generator by using observed predictands. Second, the future climate data (temperature

and precipitation) series were simulated using GCMs and selected NCEP\NCAR predictors in the first step. The most critical step in downscaling is to screen the best suitable predicators. Out of a total of twenty six, the most suitable large scale predictors were selected for each station during the "Screen Variables" process in the SDSM model (see Table S1). This model has been tested by several researchers for the downscaling of climate variables in different catchments (Chen et al. 2012; Najafi and Kermani 2017) to examine climate change implications. The SDSM model is freely available at http://co-public.lboro.ac.uk/cocwd/SDSM/sdsmmain.html. Further, detailed description along with mathematical relationships used in SDSM, is provided in supplement material.

4.1.2 Smooth Support Vector Machine (SSVM)

Support Vector Machine (SVM) is a supervised learning technique based on Vapnik-Chervonenkis (VC) and structural risk minimization (SRM) proposed by Vapnik (1998). It is an efficient approach to establish a workable relationship between model learning ability and complexity, even with inadequate data. Moreover, it has the capacity to predict and address several practical issues such as small sample size, high dimensionality and nonlinearity of data. It has been applied by several researchers (Meenu et al. 2013). Additionally, the learning algorithm of SVMs automatically chooses the model structure (number of hidden units). An advanced version of SVM, called the Smooth Support Vector Machine (SSVM) proposed by Lee and Mangasarian (2001), attempts to simplify the SVM training and minimize computational complexity which is very efficient for large data with linear kernel function. In SSVM with linear kernel function, it is modified by solving fast Newton algorithm, the smoothing approach utilized to convert a quadratic constrained optimization problem into a convex quadratic unconstrained problem. These properties converges a unique solution of unique problem of large datasets by considering smoothing parameter as infinity. The SSVM approach predicts climate data by establishing the statistical relationship between large scale GCM variables and local climate variables as generated by Chen et al. (2012). The selection of large scale predictors was performed SSVM models on the basis of correlation between observed predictands and large scale predictors by considering a threshold of 0.3. The complete description and derivation of the relationship can be found from supplement material and Lee and Mangasarian (2001). The detailed description of features, properties and algorithm of the SSVM is provided by Lee and Mangasarian (2001). The MATLAB written program of SSVM is provided in supplement material and also available on http://research.cs.wisc.edu/dmi/svm/ssvm/.

4.2 Bias Correction Approaches

We adopted distribution mapping and linear scaling approaches for the bias correction of APHRODITE gridded dataset and daily temperature and precipitation generated by the integration of SDSM and SSVM with selected GCMs. The idea behind bias correction was to cater the possible biases (uncertainties) between observed and other datasets (APHRODITE and downscaled climate variables). It is assumed that the bias correction algorithm and parameterization are applicable for both current (1961–2000) and future climate change scenarios (2030s, 2060s and 2090s). The distribution mapping used transfer function of precipitation to settle the occurrence distribution of temperature and precipitation. For the precipitation, the shape (α) and scale (β) parameter of Gamma distribution are the most

suitable for the distribution of precipitation event occurrence. The shape parameter controls the distribution profile, while, scale parameter is responsible to estimate dispersion of Gamma distribution (Teutschbein and Seibert 2012). In current study, a cumulative distribution function (CDF) was developed for observed and GCMs simulated data for a certain month on daily basis. Thus, GCMs simulated value of temperature or precipitation of corresponding day within the month sought on empirical CDF of GCMs simulation corresponding to the cumulative probability. Subsequently, the value of temperature or precipitation of that cumulative probability was located on CDF of observed data.

Linear scaling method needs monthly correction factor by taking difference between observed and GCMs simulated values of the present day. Theoretically, the corrected monthly values should be perfectly similar to the observed. The correction of precipitation data is based on the factor estimated using long term monthly mean observed and GCMs simulation. On the other hand, temperature data correction is based on additive factor estimated by taking difference between long term mean monthly observed and GCMs simulated temperature. The scaling factors (SFs) for the temperature variables (T_{min} and T_{max}) is estimated by an additive, established on the basis of difference between long-term observed mean and control run of GCM-simulations. The complete description and derivation of the relationship can be found from supplement material and Teutschbein and Seibert (2012).

5 Results and Discussion

5.1 Comparison of GCMs and Downscaling Models

The results for calibration period (1961–1990) using NCEP\NCAR reanalysis data with integration of SDSM and SSVM are given in Table 2a. Results show better performance of both models for T_{mean} and precipitation in comparison with observed data and are comparable with the results of several previous studies (Huang et al. 2011; Wilby et al. 2002). However, the simulated precipitation using SSVM was found slightly better than SDSM as confirmed by Meenu et al. (2013). The spatial heterogeneous nature of precipitation could be the responsible for low efficiency in comparison to temperature data (T_{mean}), therefore, only precipitation data was considered for the screening of two suitable GCMs for Jhelum River basin.

A comprehensive performance evaluation and comparison during calibration and validation of SDSM and SSVM were carried out by taking an average of daily observed and NCEP\NCAR simulated precipitation and temperature values of 15 stations (Table 2b). The bold and italic values indicated less biases between observed and simulated values and better efficiency in comparison with normal font values. SSVM performed better in reproducing precipitation for most of the statistics than SDSM. During the calibration period, most statistics indicated better performance of SSVM than SDSM, a few statistics such as maximum dry and wet spell length indicated better skill of SDSM than SSVM for all seasons (Table 2b). During the validation period, most statistics for SSVM was found to be higher than SDSM, therefore, SSVM was considered to be better than SDSM.

Comparisons between five GCMs were carried out to analyze suitability of SDSM and SSVM (Table 2c) using precipitation data generated for five different GCMs to determine two suitable GCMs for the Jhelum River basin. Table 2c indicates most of the bold values belong to CGCM3 while most italic values belong to HadCM3. However, other than CGCM3, the total number (4

Table 2 (a) Explained variance (E) and standard errors (SE) for calibration period (1961–1990), and statistical descriptors, (b) comparison between SDSM and SSVM simulated precipitation under NCEP\NCAR reanalysis with mean observed during calibration (1961–1990) and validation period (1991–2000) and, (c) comparison between five (5) different GCMs using mean observed and simulated precipitation of all climate stations by using SSVM and SDSM during validation period (1991–2000)

(a) SDSM

Parameters	E (%) Max	Min	SE (mm/day or °C) Max	Min
Precipitation	42.84	27.31	0.33	0.24
Tmean	83.37	65.91	0.57	0.49
SSVM				
Precipitation	48.12	29.67	0.35	0.26
Tmean	83.08	67.51	0.57	0.51

(b)

Statistics\Seasons	Models	Mean (mm/ day)	SD (mm/day)	Percentile (95%)	5-day max rainfall (mm)	%Wet	Maxi dry spell length	Maxi wet spell length
Calibration period (1961–1990)								
Annual (J-D)	Obs	4.47	11.03	21.6	173.82	0.52	71	18
	SDSM	4.12	**10.64**	24.78	146.37	0.67	**73**	**15**
	SSVM	**4.32**	10.26	**22.53**	**166.31**	**0.47**	66	12
Winter (O-M)	Obs	2.27	8.53	4.20	123.28	0.38	89	16
	SDSM	2.70	**8.28**	5.33	**118.8**	0.31	**86**	**16**
	SSVM	**2.39**	9.23	**4.81**	139.7	**0.41**	71	12
Pre-monsoon (A-J)	Obs	4.91	10.59	18.84	152.3	0.47	29	12
	SDSM	**4.76**	10.15	22.84	118.5	0.58	**31**	**13**
	SSVM	4.59	**10.34**	**20.88**	**131.9**	**0.52**	39	9
Monsoon (J-S)	Obs	6.14	13.22	26.39	203.12	0.67	19	18
	SDSM	**6.10**	13.91	28.42	219.30	0.53	**24**	**15**
	SSVM	6.31	**13.56**	**25.31**	**196.68**	**0.72**	31	21
Validation period (1991–2000)								
Annual (J-D)	Obs	4.26	11.15	26.30	113.29	0.47	76	21
	SDSM	4.85	10.31	29.09	**130.11**	0.43	65	**16**
	SSVM	**4.72**	**10.52**	**27.51**	141.45	**0.45**	**69**	13
Winter (O-M)	Obs	3.17	9.71	6.80	91.23	0.33	62	13
	SDSM	**3.75**	9.28	**8.30**	116.08	0.38	**59**	**14**
	SSVM	3.54	**9.85**	10.39	**108.3**	**0.31**	51	16
Pre-monsoon (A-J)	Obs	4.09	12.73	20.77	156.69	0.44	65	21
	SDSM	3.69	11.42	22.63	**149.22**	**0.42**	**62**	15
	SSVM	**3.78**	**11.83**	**21.61**	173.16	0.49	70	**19**
Monsoon (J-S)	Obs	5.76	13.61	29.64	233.37	0.71	44	15

Table 2 (continued)

|							
SDSM	6.37	12.89	33.51	248.69	**0.75**	**39**	**11**
SSVM	*5.98*	*14.12*	*31.90*	*238.54*	0.64	34	9
(c) GCMs							
Seasons		Validation Period (1991–2000)					
Observed							
Annual (J-D)	4.47	11.03	21.6	173.8	0.52	71	18
Winter (O-M)	2.27	8.53	4.20	123.3	0.38	89	16
Pre-monsoon (A-J)	4.91	10.59	18.84	152.3	0.47	29	12
Monsoon (J-S)	6.14	13.22	26.39	203.12	0.67	19	18
CGCM3							
Annual (J-D)	**4.42**	**8.93**	16.37	**147.3**	**0.56**	**96**	25
Winter (O-M)	**2.88**	**8.39**	5.22	*106.7*	0.19	*104*	22
Pre-monsoon (A-J)	4.09	**10.7**	**21.31**	*115.9*	**0.59**	**21**	*17*
Monsoon (J-S)	**6.40**	15.31	25.12	**149.6**	**0.72**	39	25
CCSM3							
Annual (J-D)	4.83	**11.38**	16.02	119.7	0.33	112	22
Winter (O-M)	3.24	9.83	**4.12**	98.5	0.51	**99**	25
Pre-monsoon (A-J)	4.17	10.17	23.57	101.2	*0.63*	42	19
Monsoon (J-S)	7.12	12.34	**25.73**	133.6	*0.78*	36	23
HadCM3							
Annual (J-D)	4.19	*10.11*	16.83	139.2	0.79	*107*	**16**
Winter (O-M)	*3.12*	8.71	6.97	101.6	0.08	112	21
Pre-monsoon (A-J)	4.22	10.28	22.19	109.8	0.22	38	18
Monsoon (J-S)	6.98	**13.29**	28.61	123.7	*0.52*	**31**	25
ECHAM5							
Annual (J-D)	5.03	12.23	**22.31**	113.5	0.69	113	23
Winter (O-M)	4.33	9.67	7.42	**112.2**	0.11	108	**20**
Pre-monsoon (A-J)	4.61	12.19	23.51	98.1	0.87	69	**14**
Monsoon (J-S)	7.51	*14.16*	27.35	127.4	0.93	51	29
CSIRO-MK3.5							
Annual (J-D)	*4.07*	13.41	22.54	*138.2*	*0.63*	117	*21*
Winter (O-M)	3.90	10.30	7.83	101.7	0.29	111	22
Pre-monsoon (A-J)	**4.58**	12.90	22.87	**121.0**	**0.29**	53	19
Monsoon (J-S)	6.84	14.98	28.59	*147.8*	0.95	44	**21**

Bold and Italic showed first and second ranked values, respectively, according to biasness in comparison with observed values

values for each) of bold shaded values belonging to CCSM3 and number of ECHAM5 and CSIRO-MK3.5 are greater than that of HadCM3 (3 values) suggesting that all GCMs performed efficiently. However, the maximum number of values (bold and italic) generated with CGCM3 and HadCM3 were found consistent with observed values. In addition, the downscaled values with CGCM3 were found slightly accurate than those of HadCM3. As a result, CGCM3 and HadCM3 were determined as the first and second appropriate models for Jhelum River basin.
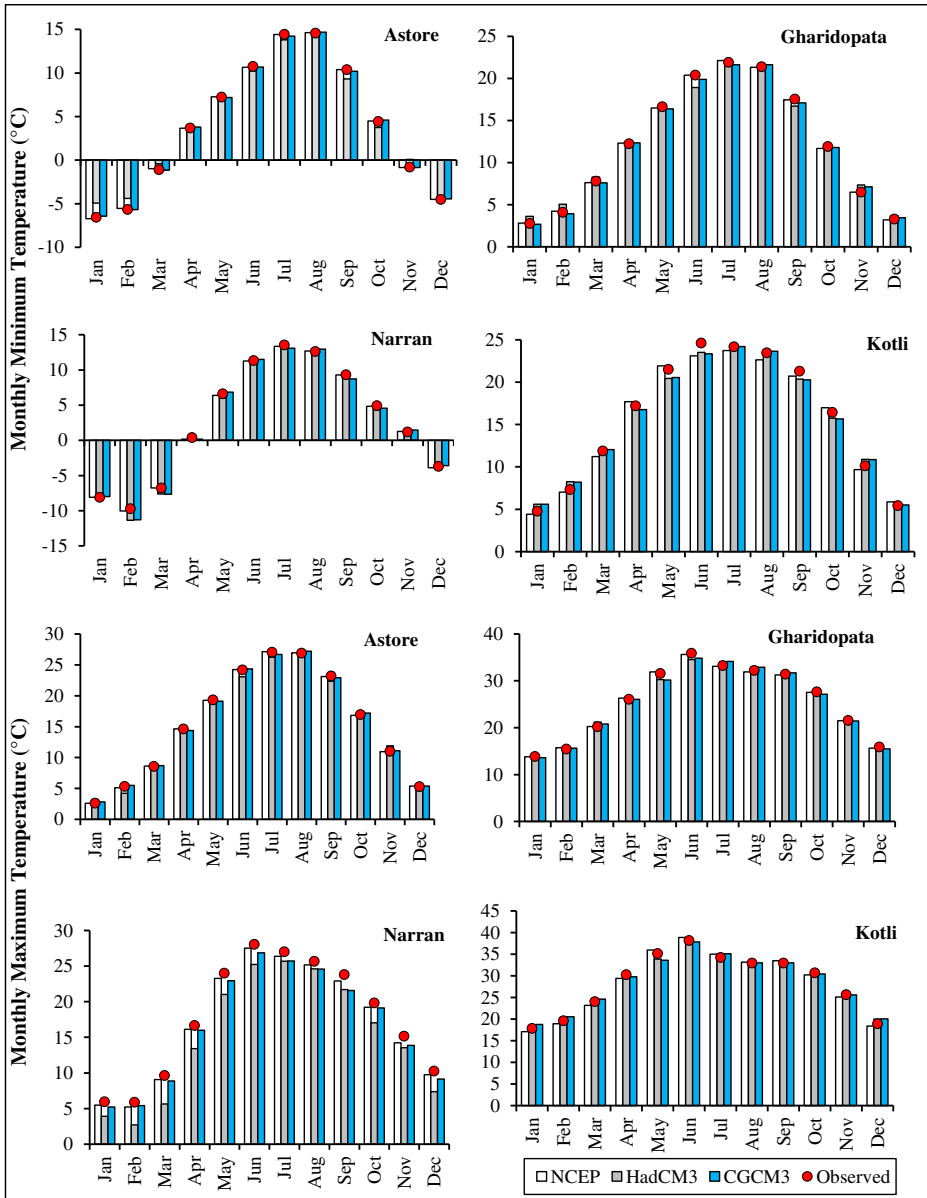
Fig. 2 Comparison between observed and downscaled maximum temperature using SSVM with NCEP\NCAR, CGCM3 and HadCM3 scenarios at four selected stations in Jhelum River basin, for calibration (1961–1990)

These models were selected for future climate projection at climate stations. Further, Figs. 2 and 3 provides comparisons between observed and simulated T_{min}, T_{max} and precipitation by employing NCEP\NCAR reanalysis, HadCM3 and CGCM3 predictors, during calibration period (1961–1990). A close agreement was found between observed (baseline) and simulated temperature variables (T_{min} and T_{max}) at four selected stations (i.e. Narran, Gharidopata, Kotli and Astore) for NCEP\NCAR, HadCM3 and CGCM3 scenarios using SSVM approach. However, simulated precipitation was overestimated particularly during extreme rainfall season (Fig. 3). The simulated
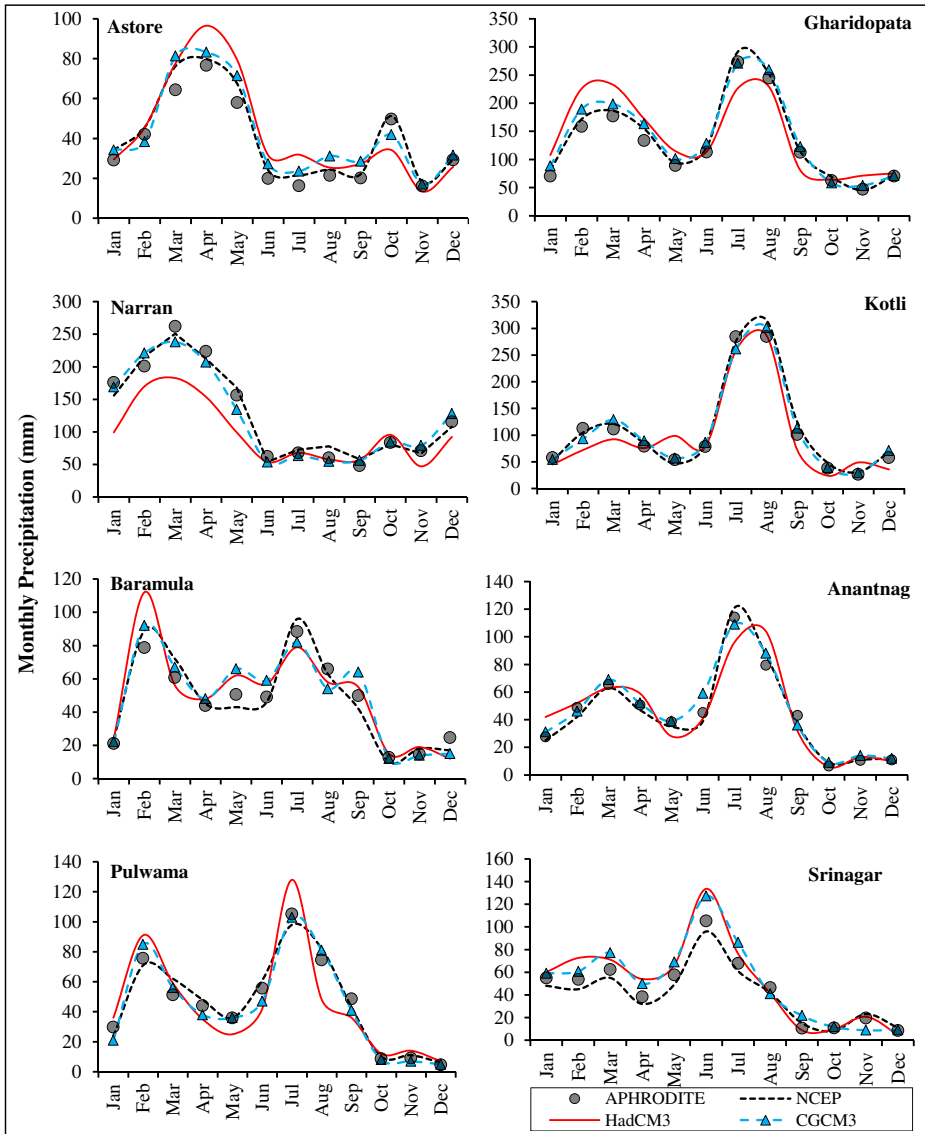
Fig. 3 Comparison between observed and downscaled precipitation with NCEP\NCAR, CGCM3 and HadCM3 scenario using SSVM at selected gauged (Astore, Gharidopata, Kotli and Narran) and un-gauged points (Baramula, Anantnag, Pulwama and Srinagar) in Jhelum River basin, for 1961 to 1990

precipitation under NCEP\NCAR reanalysis and CGCM3 were slightly better than those of HadCM3. Moreover, a slight inefficiency was found for the generation of precipitation at Baramula, Anantang, Pulwama and Srinagar with bias corrected APHRODITE precipitation. This may be associated with the existence of uncertainties in APHRODITE precipitation dataset which further leads to the low performance of downscaled precipitation.

Since large uncertainties were observed in projected datasets using SSVM and SDSM under the analyzed GCMs (CGCM3 and HadCM3), therefore, two widely used bias correction approaches were adopted to cater these uncertainties. Results before and after bias correction for the validation period (1991–2000) are presented in Table 3. The average values of performance indicators such as NS coefficient (%), root mean square error (RMSE) and R^2 for the downscaled climate variables (precipitation and T_{mean}) at stations during validation (1991–2000), are given in Table 3. The datasets for the climate variables with NCEP\NCAR generated good results in comparison to both selected GCMs. In the case of GCMs comparison, CGCM3 provides a better agreement with observed datasets as compared to HadCM3, particularly for precipitation. Moreover, the statistical descriptors for the precipitation showed that the SSVM with the integration of NCEP\NCAR (NCEP-SSVM) and GCMs (CGCM3-SSVM and HadCM3-SSVM) provided better results than SDSM (NCEP-SDSM, CGCM3-SDSM and HadCM3-SDSM) (Table 3). Results for the generation of mean temperature (T_{mean}) using aforementioned sub-models had similar conclusions. These results (SSVM performing better than SDSM for precipitation simulation) are consistent with Chen et al. (2012) and Meenu et al. (2013).

A significant improvement was realized in the climate datasets (precipitation and temperature) after adopting bias correction methods. However, it was observed that the distribution mapping method performed slightly better than the linear scaling approach. The values of R^2, NS coefficient and RMSE (mm) of the sub-models without bias correction of precipitation data varied between 0.49–0.69, 0.47–0.65 and 35.12–43.29 (mm), respectively. After bias correction of precipitation data, the values of R^2, NS and RMSE (mm) were varied between 0.72–0.81, 0.71–0.79 and 19.28–24.87, respectively (for linear scaling method) and 0.72–0.84, 0.69–0.83 and 15.34–19.24 (for distribution mapping approach), respectively. A significant improvement was realized in temperature because both GCMs and downscaling methods generated raw temperature data efficiently (as discussed in previous section). In general results show that distribution mapping performed efficiently than the linear scaling method for bias correction of climate variables. Further, Fig. S2 depicted a significant improvement in basin-wide simulated precipitation after bias correction, particularly, for peak rainfall season. Therefore, the bias corrected datasets using distribution mapping method were adopted to analyze the change in climate variables under future scenarios.
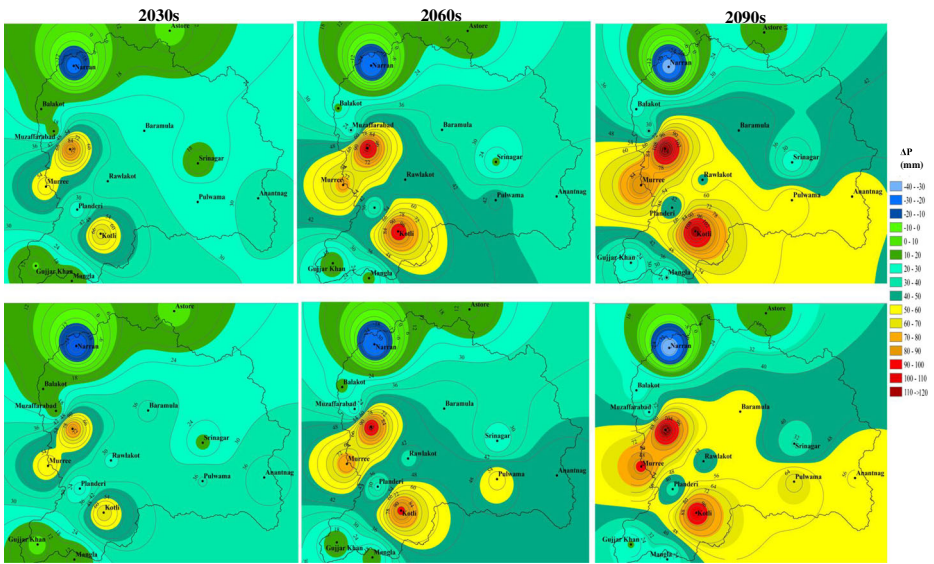
5.2 Projection of Climate Variables

The projected climate data (precipitation and temperature) using CGCM3 and HadCM3 by the integration of SSVM at different climate stations and the un-gauged part of the study area over mean annual and seasonal basis, is given in Table S2 and S3 and Figs. 4 and 5. The maximum average annual increment in precipitation (mm) was found at stations mainly influenced by the monsoon precipitation pattern such as Gharidopata, Kotli, Murree, Planderi, Rawalakot, Baramula, Anantnag and Pulwama, for 2030s in comparison with baseline period (1971–2000) (Fig. 4 and Table S2). An increase in precipitation ranges between 8.08 (Astore) and 89.50 mm (Gharidopata) using CGCM3. Interestingly, at Narran station, a decrease in precipitation by −23.61 and −26.94 mm was observed for 2030s using CGCM3 and HadCM3, respectively.

Table 3 Statistical comparison of downscaling techniques (SSVM and SDSM), NCEP/NCAR, scrutinized GCMs (CGCM3 and HadCM3) and bias correction techniques (mapping distribution and linear scaling approach) for the downscaling (with and without bias correction) and observed precipitation and T_{mean}, during validation period (1990–2000)

| | Without bias correction | | | With bias correction | | | | | |
| | | | | Linear scaling | | | Distribution mapping | | |
Precipitation	R^2	NS	RMSE (mm)	R^2	NS	RMSE (mm)	R^2	NS	RMSE (mm)
NCEP-SSVM	0.69	0.65	36.59	0.81	0.79	19.28	0.84	0.83	15.34
CGCM3-SSVM	0.53	0.51	41.23	0.75	0.74	21.19	0.79	0.77	17.91
HadCM3-SSVM	0.49	0.47	43.29	0.72	0.71	23.93	0.76	0.75	19.24
NCEP-SDSM	0.66	0.63	35.12	0.79	0.77	24.87	0.82	0.79	17.83
CGCM3-SDSM	0.50	0.47	40.09	0.76	0.75	21.66	0.77	0.73	16.98
HadCM3-SDSM	0.54	0.49	41.16	0.74	0.71	20.69	0.72	0.69	19.23
Tmean	R^2	NS	RMSE (°C)	R^2	NS	RMSE (°C)	R^2	NS	RMSE (°C)
NCEP-SSVM	0.96	0.90	1.51	0.99	0.98	0.19	0.99	0.98	0.14
CGCM3-SSVM	0.93	0.88	1.76	0.98	0.96	0.45	0.99	0.97	0.07
HadCM3-SSVM	0.91	0.86	1.93	0.98	0.97	0.49	0.98	0.96	0.18
NCEP-SDSM	0.94	0.89	1.69	0.98	0.97	0.47	0.99	0.98	0.19
CGCM3-SDSM	0.93	0.87	1.72	0.97	0.96	0.35	0.98	0.96	0.15
HadCM3-SDSM	0.90	0.85	2.09	0.98	0.97	0.33	0.98	0.97	0.17

Fig. 4 Spatial distribution of change in precipitation (mm) including contour lines, for 2030s, 2060s and 2090s, in the Jhelum River basin with CGCM3 (top) and HadCM3 (bottom) using SSVM

Similarly, for 2060s at Gharidopata station, the highest precipitation is expected to increase by 108.6 mm (CGCM3) and 103.9 mm (HadCM3). Meanwhile, minimum precipitation at Astore station was found to be increased by 12.48 mm (CGCM3) and 10.83 mm (HadCM3). Slight decrease of −29.54 mm and − 34.96 mm were found at the Narran station, for 2060s and 2090s, respectively, when using CGCM3. Table S2 and Fig. 4 shows that the maximum increase in precipitation is expected in 2090s at Gharidopata (125.5 mm) stations using CGCM3. Overall, a continuous increasing trend of precipitation was observed for future selected time windows (2030s, 2060s and 2090s) in comparison to the baseline. In all future scenarios, the precipitation values downscaled using HadCM3 were found slightly less than CGCM3. It was noticed that the precipitation increases at Astore (8.08–13.43 mm), Gujjar Khan (9.58–21.44 mm) and Mangla (13.12–23.72 mm) stations are very low than other stations. Moreover, decreasing trend in precipitation was only found at the Narran station. This may be associated with fact that the selected study area is mainly influenced by monsoon season (summer precipitation); however, Astore and Narran stations are mainly influenced by westerlies circulation (winter precipitation pattern). The decrease in precipitation at Narran station may be associated with the continuous decrease in temperature, particularly during winter months (Fig. 5). This decrease in temperature could be the cause of precipitation occurrence in the form of snow (solid) rather than liquid. In case of APHRODITE data, the precipitation is expected to increase in the future at all data points located in the un-gauged part of the basin, with maximum precipitation increment of 59.32 and 66.49 mm, with CGCM3 and HadCM3, respectively, at Pulwama for 2090s. Overall, the increase in precipitation at APHRODITE data points (Srinagar, Baramula, Anantnag and Pulwama) is not as high as that at Gharidopata and Kotli. This may be due to the fact that these stations, like Narran and in contrast to Gharidopata and Kotli, are mainly influenced by the westerlies circulation precipitation pattern, as confirmed by Azmat et al. (2017).

The projection of temperature variables for CGCM3 and HadCM3 with SSVM, is shown in Table S3 and Fig. 5. The minimum and maximum temperature at all stations are expected
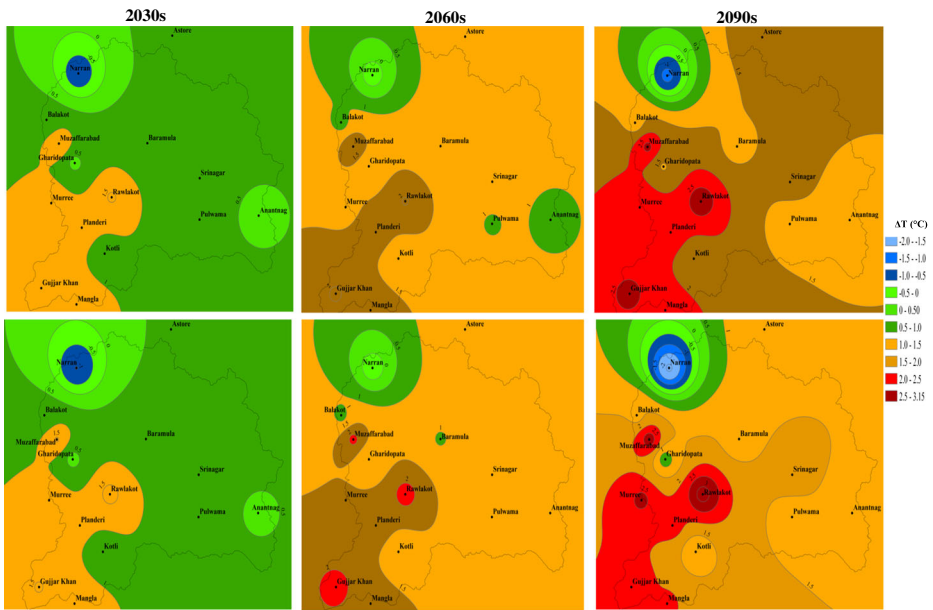
Fig. 5 Spatial distribution of change in mean temperature (°C) including contour lines for 2030s, 2060s and 2090s, in the Jhelum River basin with CGCM3 (top) and HadCM3 (bottom) using SSVM

to increase in the future except for Narran, where the minimum and maximum temperature [−0.16, −0.19 and − 0.30; −1.52, −1.95 and − 2.21 for 2030s, 2060s and 2090s, respectively] are expected to decrease, for CGCM3. The maximum rise in mean temperature is expected to occur at Srinagar for 2090s, by 1.67 or 1.79 °C, when using CGCM3 and HadCM3, respectively. The maximum increase in mean temperature is observed at Srinagar and Baramula. The maximum precipitation is expected to increase at stations located at relatively low elevation and mainly influenced by monsoon precipitation pattern (i.e. Gharidopata, Murree and Kotli) and this increase continuously happens under future scenarios (2030s, 2060s and 2090s). Conversely, the precipitation at high-altitude stations (e.g. Narran and Astore) mainly influenced by the westerlies circulation pattern is expected to decrease or a very slight increment will occur at the end of twenty-first century for CGCM3 and HadCM3 with SSVM (Fig. 4). Further, it was observed that the T_{mean} is expected to significantly increase at low elevation stations (i.e. Gharidopata, Kotli, Gujjar Khan, Planderi, Muzaffarabad and Mangla), while a slight increase (i.e. Astore, Pulwama and Anantnag) and decrease at Narran station is expected under future scenarios (Fig. 5). It seems that the slight increase in precipitation is may be due to the less increase in temperature at high-altitude stations and vice versa for the low elevation stations. Most of precipitation at high-altitude stations occurs in form of snow and the precipitation at Narran station will possibly occur in form of solid snow due to decrease in temperature. At low elevation stations, on the other hand, the rise in temperature is expected to contribute significantly in precipitation of liquid form even during the winter season also observed by Azmat et al. (2017).

Figure 6a indicates that the maximum increase in basin-wide precipitation was expected during monsoon or extreme rainfall season (July–Sep), while least increment was found during
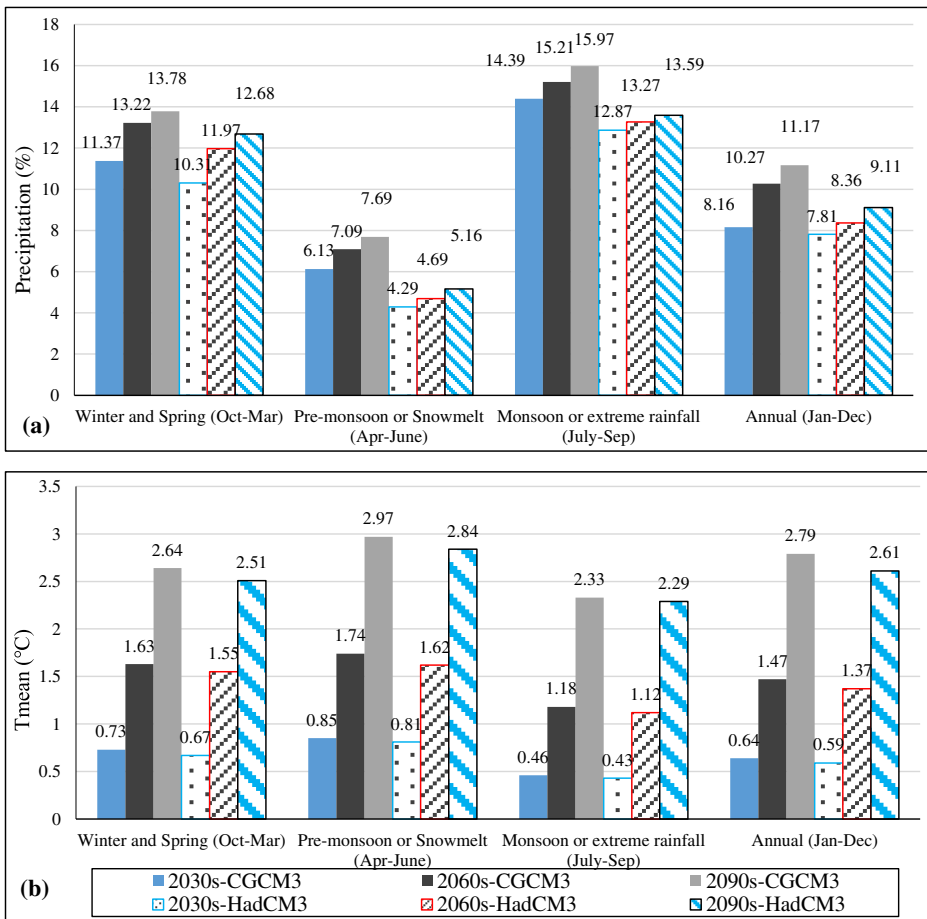
Fig. 6 Basin-wide seasonal change in **(a)** percent (%) change precipitation and **(b)** Tmean (°C), in Jhelum River basin for 2030s, 2060s and 2090s, with the scenarios of CGCM3 and HadCM3 using SSVM

pre-monsoon season (April–June), generated for CGCM3. Since, the major proportion of precipitation in study basin occurs during late winter (Feb and March) and monsoon season, this could be a reason for maximum increase in precipitation during the aforementioned seasons. In contrast to the precipitation prediction results, Fig. 6b shows that highest mean temperature rise is expected during pre-monsoon season (April–June). Overall, the mean basin-wide temperature is expected to increase by 0.759, 1.37 and 2.017 during 2030s, 2060s and 2090s, respectively, when using CGCM3. The increasing and decreasing trend of climate variables at different stations in Jhelum River basin are confirmed and stated by Azmat et al. (2017).

6 Conclusion

The current study aimed to provide a methodology to improve the predictions of climate variables by comparing GCMs (CGCM3, HadCM3, CCSM3, ECHAM5 and CSIRO-

MK3.5), downscaling techniques (SSVM and SDSM) and bias correction approaches in high-altitude, scarcely-gauged, Jhelum River basin.

- Comparative analysis of downscaling techniques showed that the SSVM performed slightly better than SDSM, particularly for downscaling of precipitation data, in high-altitude Jhelum River basin. Similar results were found for the downscaling of temperature data.
- All selected GCMs performed well, however, CGCM3 and HadCM3 performed better than CCSM3, ECHAM5 and CSIRO-MK3.5. Furthermore, CGCM3 provided better results in comparison to HadCM3. CGCMs and HadCM3 were chosen for the investigation of change in climate variables and their potential impact on hydrological behavior of high-altitude regions under future scenarios. In addition, the climate change analysis could be conducted using several GCMs and emission scenarios with the integration of different downscaling techniques to optimize robust and reliable conclusions.
- The availability of APHRODITE (gridded temperature and precipitation) datasets at long temporal range (1954–2007) could be an alternative option (after careful bias correction) for the climate change impact studies in high-altitude, scarcely gauged catchments. The long range precipitation record of APHRODITE data is more suitable for climate changes studies with integration of GCMs. However, the bias correction is mandatory to make APHRODITE datasets suitable for the downscaling of climate variables in scarcely-gauged, snow-fed mountainous catchments such as Jhelum River basin.
- The distribution mapping approach corrected temperature and precipitation fairly well in comparison with the linear scaling approach particularly for the precipitation. Distribution mapping convincingly showed better performance to overcome uncertainty in GCMs simulated data series for all statistical descriptors in comparison with linear scaling. Therefore, distribution mapping approach is a good choice for the bias correction of gridded datasets and GCM simulations.
- The MODAWEC model could be a good option for the catchments where long term monthly data is available, however the daily data are lacking or not up to the mark for the application.
- The results for the future climate change scenarios showed that a significant increase in precipitation (basin-wide) by 25.51, 36.76 and 45.52 mm and temperature (basin-wide) by 0.64, 1.47 and 2.79 °C during 2030s, 2060s and 2090s, respectively, are expected. Both climate variables (precipitation and temperature) are expected to increase at all stations except for Narran where both climate variables are expected to decrease.

Overall, current study relies on combination of modeling experimental framework, enabling the examining of downscaling uncertainties in a systematic way. This modeling framework provides novel insights into the future climate change in a high-altitude Jhelum River basin, because it permits in identifying robust changes.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Addor N, Rössler O, Köplin N, Huss M, Weingartner R, Seibert J (2014) Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. Water Resour Res 50:7541–7562

Azmat M, Choi M, Kim T-W, Liaqat UW (2016) Hydrological modeling to simulate streamflow under changing climate in a scarcely gauged cryosphere catchment. Environ Earth Sci 75:1–16. https://doi.org/10.1007/s12665-015-5059-2

Azmat M, Liaqat UW, Qamar MU, Awan UK (2017) Impacts of changing climate and snow cover on the flow regime of Jhelum River, western Himalayas. Reg Environ Chang 17:1–13. https://doi.org/10.1007/s10113-016-1072-6

Chen H, Xu C-Y, Guo S (2012) Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. J Hydrol 434:36–45. https://doi.org/10.1016/j.jhydrol.2012.02.040

Dibike YB, Coulibaly P (2005) Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. J Hydrol 307:145–163. https://doi.org/10.1016/j.jhydrol.2004.10.012

Hagemann S, Chen C, Clark DB, Folwell S, Gosling SN, Haddeland I, Hanasaki N, Heinke J, Ludwig F, Voss F, Wiltshire AJ (2013) Climate change impact on available water resources obtained using multiple global climate and hydrology models. Earth Syst Dynam 4:129–144. https://doi.org/10.5194/esd-4-129-2013

Huang J, Zhang J, Zhang Z, Xu C, Wang B, Yao J (2011) Estimation of future precipitation change in the Yangtze River basin by using statistical downscaling method. Stoch Env Res Risk A 25:781–792

Kumar R, Singh S, Kumar R, Singh A, Bhardwaj A, Sam L, Randhawa SS, Gupta A (2016) Development of a Glacio-hydrological model for discharge and mass balance reconstruction. Water Resour Manag 30:3475–3492

Lee Y-J, Mangasarian OL (2001) SSVM: a smooth support vector machine for classification. Comput Optim Appl 20:5–22. https://doi.org/10.1023/A:1011215321374

Lin G-F, Chang M-J, Wang C-F (2017) A novel spatiotemporal statistical downscaling method for hourly rainfall Water Resour Manag 31:1–25. https://doi.org/10.1007/s11269-017-1679-5

Liu J, Williams JR, Wang X, Yang H (2009) Using MODAWEC to generate daily weather data for the EPIC model. Environ Model Softw 24:655–664

Lutz A, Immerzeel W, Kraaijenbrink P (2014) Griddedmeteorological datasets and hydrological modelling in the Upper Indus Basin (Costerweg 1V, 6702 AA Wageningen, The Netherlands)

Meenu R, Rehana S, Mujumdar P (2013) Assessment of hydrologic impacts of climate change in Tunga–Bhadra river basin, India with HEC-HMS and SDSM. Hydrol Process 27:1572–1589. https://doi.org/10.1002/hyp.9220

Najafi R, Kermani MRH (2017) Uncertainty modeling of statistical downscaling to assess climate change impacts on temperature and precipitation. Water Resour Manag 31:1843–1858

Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17:182–190

Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J Hydrol 456:12–29

Vapnik VN (1998) Statistical learning theory vol 1. Wiley, New York

Wang X, Huang G, Lin Q, Nie X, Cheng G, Fan Y, Li Z, Yao Y, Suo M (2013) A stepwise cluster analysis approach for downscaled climate projection–a Canadian case study. Environ Model Softw 49:141–151

Wilby RL, Dawson CW (2013) The statistical downscaling model: insights from one decade of application. Int J Climatol 33:1707–1719

Wilby RL, Dawson CW, Barrow EM (2002) SDSM—a decision support tool for the assessment of regional climate change impacts. Environ Model Softw 17:145–157. https://doi.org/10.1016/S1364-8152(01)00060-3

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A (2012) APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull Am Meteorol Soc 93:1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1