Downscaling technique uncertainty in assessing hydrological impact of climate change in the Upper Beles River Basin, Ethiopia

Girma Yimer Ebrahim, Andreja Jonoski, Ann van Griensven and Giuliano Di Baldassarre

ABSTRACT

We investigate the uncertainty associated with downscaling techniques in climate impact studies, using the Upper Beles River Basin (Upper Blue Nile) in Ethiopia as an example. The main aim of the study is to estimate the two sources of uncertainty in downscaling models: (1) epistemic uncertainty and (2) stochastic uncertainty due to inherent variability. The first aim was achieved by driving a Hydrologic Engineering Centre-Hydrological Modelling System (HEC-HMS) model with downscaled daily precipitation and temperature using three downscaling models: Statistical Downscaling Model (SDSM), the Long Ashton Research Station Weather Generator (LARS-WG) and an artificial neural network (ANN). The second objective was achieved by driving the hydrological model with individual downscaled daily precipitation and temperature ensemble members, generated by using the stochastic component of the SDSM. Results of the study showed that the downscaled precipitation and temperature time series are sensitive to the downscaling techniques. More specifically, the percentage change in mean annual flow ranges from 5% reduction to 18% increase. By analyzing the uncertainty of the SDSM model ensembles, it was found that the percentage change in mean annual flow ranges from 6% increase to 8% decrease. This study demonstrates the need for extreme caution in interpreting and using the output of a single downscaling model.

Key words | downscaling, hydrologic model, stream flow, uncertainty

Girma Yimer Ebrahim (corresponding author)
Andreja Jonoski
Ann van Griensven
Giuliano Di Baldassarre
UNESCO-IHE Institute for Water Education,
Westvest 7,
2611AX, Delft,
The Netherlands
E-mail: *g.ebrahim@unesco-ihe.org*

Ann van Griensven
Vrije Universitiet Brussel,
Pleinlaan 2,
1050 Brussels,
Belgium

INTRODUCTION

The impacts of climate change may significantly alter the natural environment on our planet and consequently create less favourable conditions for overall socio-economic development. Since nearly all natural and socio-economic systems critically depend on water, the impacts of climate change on the hydrological cycle are of high significance. The alteration of water resources availability and changing patterns of hydrological disasters such as floods and droughts may be direct impacts of climate change, with many indirect effects on agriculture, food and energy production and overall water infrastructure.

On a global scale, climate change and resulting changes in the hydrological cycle are commonly predicted from global circulation models (GCMs) which are driven by various possible scenarios with greenhouse gas emissions as inputs. However, outputs from GCMs (precipitation, temperature and other climatic variables) are available only on large scales, usually ranging over several hundreds of kilometres (Solomon *et al.* 2007), which are inadequate for analysis and prediction of climate change impacts on hydrological systems at a local or river basin scale. To overcome this problem, different downscaling methods have been developed and proposed. These all aim to provide adequate hydrological variables at a local scale from GCM outputs. These variables are then used as inputs to basin-scale hydrological models for predicting climate-induced changes in

doi: 10.2166/nh.2012.037

flow patterns. This article focuses on the downscaling methods and uncertainties associated with their implementation.

As described by Elshamy *et al.* (2008), the specific steps in assessing climate change impacts on local-scale hydrological systems are as follows:

1. Choosing one or more emission scenarios of the International Panel on Climate Change (IPCC) special report (Nakicenovic *et al.* 2000), which depend on the future economy and energy use policies.
2. Choosing one or more GCM.
3. Downscaling of the GCM output to the specific river-basin scale.
4. Using the downscaled GCM outputs as inputs for a hydrological model.
5. Analysis of hydrological model results by comparing them to the corresponding results related to current climate or different possible future climates.

This approach has become very popular as it potentially allows the quantification of changes in floods, flow duration curves or other parts of the hydrological cycle which may be of interest in particular studies (Blöschl & Montanari 2010).

However, the entire modelling chain is affected by the range of uncertainties which need to be taken into account in decision-making processes (Di Baldassarre *et al.* 2011). The first steps of the modelling chain have uncertainties in estimates of emission scenarios and climate modelling, which are commonly assessed in studies of global socio-economic development and climate modelling studies, respectively (beyond the scope of this study). The uncertainty of downscaling methods, which are in the middle of the modelling chain, are the main focus of this study. At the end of the modelling chain are the hydrological modelling and local impact assessment. Hydrological modelling is affected by significant uncertainty as many sources of error propagate through the models and affect its output (Montanari 2007). These sources can be classified (Götzinger & Bárdossy 2008) as: (1) observation uncertainty (Daren Harmel & Smith 2007), which is the approximation in the observed hydrological variables used as input or calibration data (e.g. rainfall, temperature and river discharge); (2) parameter uncertainty (Beven & Binley 1992), which is induced by imperfect model parameterization; and (3) model structural uncertainty (Butts *et al.* 2004; Refsgaard *et al.* 2006), which originates from the inability of models to perfectly schematize the (often incompletely understood) physical processes involved. While hydrological model uncertainty is fully recognized, it is not considered in this study. Local impact uncertainty, which depends on actual target of the analysis (availability of water resources for various purposes, analysis of floods or droughts, etc.), is also not considered in this study. We focus instead on the uncertainty in climate change studies associated with downscaling techniques.

As mentioned, downscaling is necessary because even the most fully developed GCMs are currently only able to calculate atmospheric processes at a spatial resolution of several hundred kilometres. This is much too coarse to be used as input data for climate impact hydrological models (Busuioc *et al.* 2001; Giorgi *et al.* 2001; Wood *et al.* 2004; Dibike & Coulibaly 2005). This scale mismatch comes from the different nature of the models. Hydrological models deal with small- or subcatchment-scale processes whereas GCMs simulate planetary-scale and parameterize many regional and smaller-scale processes. Downscaling techniques have therefore emerged as a means of relating the scale mismatch between what GCMs are able to provide and the increasingly small scales required for impact studies (Wilby & Wigley 2000; Dibike & Coulibaly 2005; Maurer & Hidalgo 2008). Downscaling is a term given to the process of deriving finer-resolution data (e.g. for a particular site) from coarser-resolution GCM data.

There are two main approaches to downscaling: dynamical and statistical. Dynamic downscaling involves extraction of local-scale information by developing and using regional climate models (RCMs) with the coarse GCM data used as boundary (Wilby & Wigley 2000; Haylock *et al.* 2006). Statistical downscaling involves the extraction of local-scale information by statistically relating large-scale climate features to fine-scale climate for the region of interest. The latter approach is mostly performed on station data (observations in appoint) and it can provide point-scale climatic variables from GCM-scale output. An excellent reference source for climate models and climate science more generally is the Fourth Assessment report of the IPCC (Solomon *et al.* 2007).

Since RCM with higher resolution was not available for the current study, the dynamic downscaling method was not considered and the focus here is on statistical downscaling methods. Statistical downscaling techniques are straightforward techniques for constructing finer-resolution climate scenarios from GCM outputs at the local level. Although there are various statistical downscaling techniques, it is not yet clear which method provides the most reliable estimate of daily rainfall and temperature series for hydrological impact studies (Dibike & Coulibaly 2005). The uncertainty introduced by the application of downscaling techniques tends to increase as scale is reduced and for more extreme hydrological events (Blöschl & Montanari 2010).

A few studies have examined the uncertainty of different downscaling models. However, as far as the authors are aware, there have not been any such studies in the Blue Nile where long-term and well-distributed climatic information is not available and the topography is complex.

Khan *et al.* (2006) carried out uncertainty analysis of three downscaling models, namely: Statistical Downscaling Model (SDSM), Long Ashton Research Station Weather Generator (LARS-WG) and artificial neural networks (ANNs) in the Chute-du-Diable sub-basin located in the Saguenay–Lac-Saint-Jean watershed in northern Quebec, Canada. Uncertainty assessment was carried out by comparing monthly means and variance of downscaled and observed daily precipitation and daily minimum and maximum temperature. Khan *et al.* (2006) applied both parametric and non-parametric tests for downscaled data, and demonstrated that SDSM is the most efficient in reproducing various statistical characteristics of observed data in its downscaled result with a 95% confidence level. The ANN is the least capable and LARS-WG is between SDSM and ANN.

Dibike *et al.* (2008) carried out similar uncertainty analysis for the SDSM model. They used two complementary methods, namely: Wilcoxon signed-rank testing and the bootstrap confidence interval estimation technique. Khan & Coulibaly (2009) also addressed the issue of SDSM uncertainty for hydrological impact assessment. In their study they used the SDSM model to undertake uncertainty analysis of a SDSM in simulating daily river and reservoir flow in the Serpent and the Chute-du-Diable watershed, respectively, located in northeast Canada. This study found that the uncertainty bands of the mean ensemble flow (i.e. flow simulated using the mean of the ensemble members of the downscaled meteorological variables) were able to encompass all other flows simulated with various individual downscaled meteorological ensemble members. The study claimed that, when dealing with uncertainty with regard to hydrological impact of climate change, there is no need to use individual downscaled members; rather, it would be sufficient to use the mean of the ensemble flow and the uncertainty band.

The main objective of the study presented here is to examine the uncertainty of three downscaling techniques – SDSM, ANN and LARS-WG – for hydrological impact assessment. We first obtain downscaling results from these three techniques in terms of precipitation and temperature, the most important variables for hydrological impacts. Secondly we examine runoff uncertainty by using a Hydrological Engineering Centre-Hydrological Modelling System (HEC-HMS) rainfall–runoff model using the downscaled precipitation and temperature data as inputs. We aim to estimate two types of uncertainty: (1) epistemic uncertainty (associated with the lack of knowledge, imprecision in data and observations), estimated by making comparative analysis of results from the three downscaling methods; and (2) stochastic uncertainty (associated with the randomness observed in nature, regarded as irreducible) as defined by Walker *et al.* (2003), which is estimated using precipitation and temperature ensemble members generated from the stochastic component of the SDSM. To this end, the Upper Beles River Basin, Upper Blue Nile, Ethiopia is used as application example.

STUDY AREA AND DATA USED

The study area selected for this study is the Beles sub-basin of the Upper Blue Nile in Ethiopia. The main stem of the Beles River originates on the face of the escarpment across the divide to the west of the south-western portion of Lake Tana (Figure 1). It then flows on in a westerly direction and enters the Blue Nile just before it crosses the Ethiopia–Sudan frontier. It is the only major right-bank

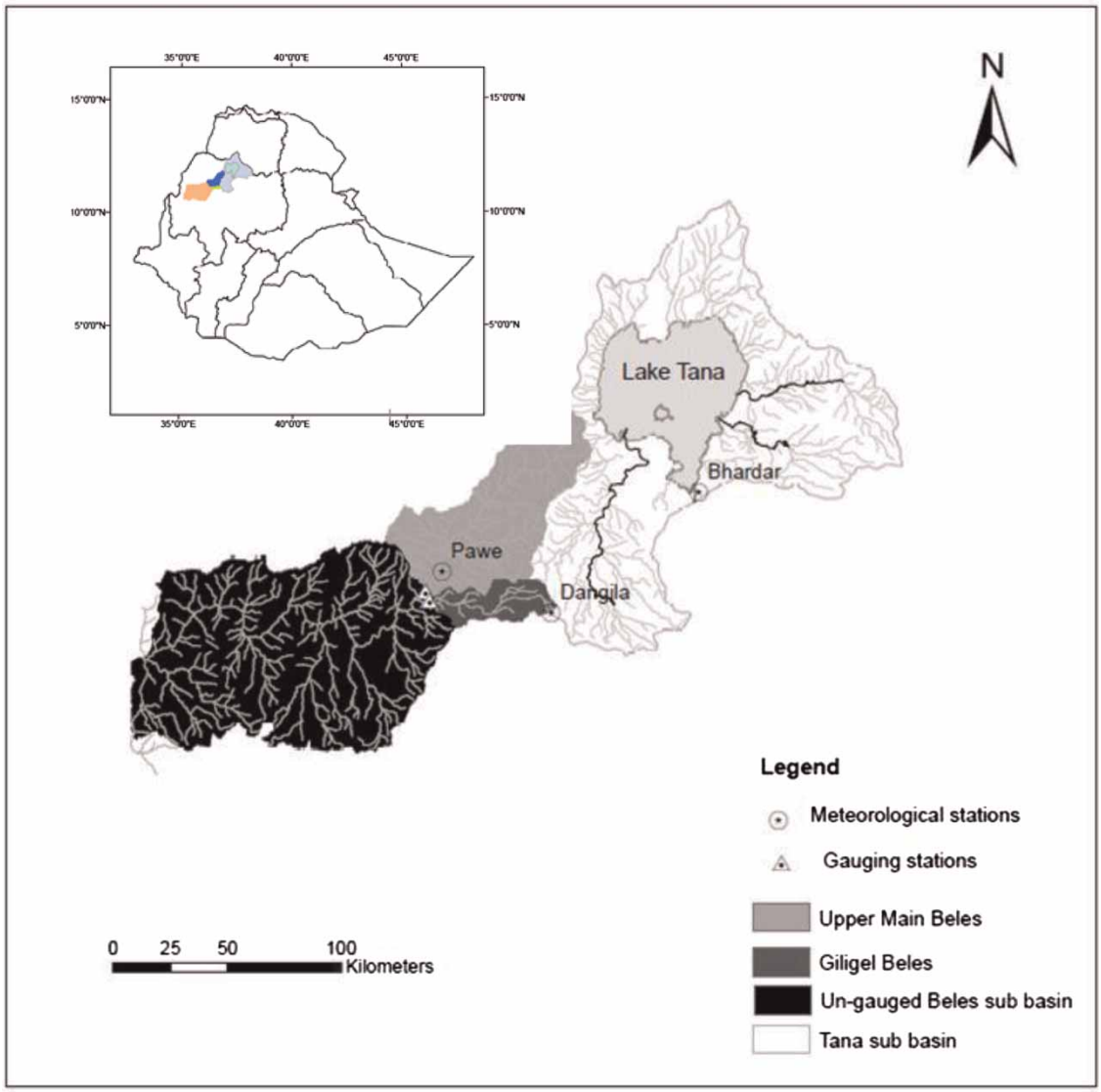

Figure 1 | Tana and Beles basin map.

tributary of the Blue Nile. The Beles basin covers an area of about 14,000 km^2 and geographically it extends from 10°56′ to 12°N latitude and 35°12′ to 37°E longitude. The basin has two gauged subcatchments, namely Upper Main Beles and Giligile Beles which have an area of 3,474 km^2 and 675 km^2, respectively. The focus of this study is the Upper Main Beles sub-basin. The elevation in the Upper Main Beles basin ranges from 2,718 m above mean sea level (a.m.s.l) in the north at the watershed divide between the Tana and Beles basins to 1,015 m (a.m.s.l) at the gauging station.

The climate in the study area is warm and subtropical. The annual mean minimum and maximum temperature ranges over 16.5–32.5 °C (Pawe station). Precipitation is moderately abundant (about 1,000 mm yr^{-1}), even when adjacent areas are very severely affected by drought. Rainfall in the study area increases with elevation. Annual potential evapotranspiration (ET) is about 1,500 mm. Climatic seasons are mainly controlled by the annual migration of the Inter-tropical Convergence Zone (ITCZ) and the associated atmospheric circulation, which are modulated by the complex topography of the region.

Three meteorological stations in and around the Upper Beles sub-basin, namely Pawe, Dangila, and Bhardar, are used in the downscaling experiments (geographical coordinates are listed in Table 1 and the stations are shown in Figure 1). Precipitation, temperature, wind speed, radiation and relative humidity data were gathered from the National Meteorological Agency of Ethiopia. For Bhardar station, 30 years of data (1961–1990) were used for the downscaling experiment. Since the other two stations have shorter records of only 15 years (1987–2001), daily precipitation and daily maximum and minimum temperature records have been used as predictands. In this study, only the downscaling results from Bhardar station are discussed for two reasons: (1) it has a longer time series that can show a wide range of temporal variability; and (2) the longer time series enables comparison of downscaling models. Downscaling results for the other two stations are also summarized in tables.

Table 1 | Locations of meteorological stations used in the downscaling experiment

	Locations in degree Latitude (N)	Longitude (E)	Altitude (m a.s.l)
Pawe	11.28	36.39	1,050
Bhardar	11.6	37.42	1,770
Dangila	11.28	36.91	2,180

Atmospheric predictor variables used to calibrate the downscaling models were obtained from the National Centre for Environmental Prediction (NCEP) reanalysis (Kalnay *et al.* 1996). Future climate change scenarios from the Hadley Centre's coupled ocean/atmosphere climate model (HadCM3) (Gordon *et al.* 2000) were used. HadCM3 was selected due to the availability of both the outputs for the considered emission scenario and the primary downscaling variables archived daily. Pre-processed predictor variables for NCEP reanalysis and HadCM3 were obtained from the Canadian Institute for Climate Studies (CICS). The archive of NCEP and GCM output contains 26 daily predictor variables (describing atmospheric circulation, thickness and moisture content at the surface, 850 and 500 hPa levels), listed in Table 2. Predictor variables are provided on a grid-box-by-grid-box basis of size 2.5° latitude and 3.75° longitude. In the present study, predictors from a grid-box corresponding to 12.5° latitude and 37.5° longitude were employed for A2 (medium–high emissions) scenario of the IPCC Special Report on Emission Scenarios (SRES) (Nakicenovic *et al.* 2000). Other climate variables corresponding to future climate scenarios such as monthly wind speed, radiation and relative humidity were obtained from the Hadley Centre, UK.

Table 2 | Large-scale daily atmospheric variables from the NCEP reanalysis and HadCM3 simulation output that are used as potential inputs to the SDSM and ANN. All predictors were normalized with respect to the 1961–1990 mean and standard deviation, except the wind direction (http://www.cccsn.ec.gc.ca/?page=pred-help)

No.	Predictors	Description	No.	Predictors	Description
1	mslpaf	Men sea level pressure	14	p5_ zhaf	500 hpa divergence
2	p_faf	Men sea level pressure	15	p8_faf	850 hpa airflow strenght
3	p_uaf	Surface air flow strength	16	p8_uaf	850 hpa zonal velocity
4	p_vaf	Surface zonal velocity	17	p8_vaf	850_hpa meriodinal velocity
5	p_zaf	Surface meriodinal velocity	18	p8_zaf	850 hpa vorticity
6	p_thaf	Surface vorticity	19	p850af	850 hpa geopotential height
7	p_zhaf	Surface wind direction	20	p8thaf	850 hpa wind direction
8	p5_faf	Surface divergence	21	p8zhaf	850 hpa divergence
9	p5_uaf	500 hpa air flow strength	22	r500af	Relative humidity at 500 hpa
10	p5_vaf	500 hpa zonal velocity	23	r850af	Relative humidity at 850 hpa
11	p5_zaf	500 hpa meriodinal velocity	24	Rhumaf	Near surface relative humidity
12	p500af	500 hpa voritcity	25	shumaf	Surface specific humidity
13	p5thaf	500 hpa geopotential height	26	tempaf	Mean temperature at 2 m

METHODS

In this study we adopted a top-down approach where outputs from GCMs were downscaled to obtain local climate signals to be used as input to a hydrological model to assess the potential impact of climate change. The following three-step methodology was followed. First, local precipitation and temperature values were obtained using statistical downscaling approaches for the baseline (1961–1990) and future (2040–2069) climates. Second, a hydrological model was applied to simulate daily river flows. Third, the simulated future river flows were compared with those simulated with the baseline. The following sections describe the downscaling models, downscaling experiments and the hydrological modelling.

DOWNSCALING MODELS

Statistical downscaling model

The SDSM (Wilby *et al.* 2002) is a hybrid of the multiple linear regression and stochastic downscaling models. It is a freely available decision-support tool. In SDSM downscaling, a multiple linear regression model is developed between a few selected large-scale predictor variables and local-scale predictands such as temperature and precipitation. The stochastic component of SDSM enables the generation of multiple simulations with slightly different time series attributes, but the same overall statistical properties. Precipitation in SDSM is modelled as a conditional process in which local precipitation amount is correlated with the occurrence of wet days. Minimum and maximum temperatures are modelled as unconditional processes, where a direct link is assumed between the large-scale predictors and the local-scale predictand. The SDSM model reduces the task of downscaling into a series of discrete processes such as quality control and data transformation, screening of predictor variables, model calibration and weather and scenario generation (for a more detailed description, see Wilby *et al.* (2002)).

LARS-WG stochastic weather generator

LARS-WG is a public-domain stochastic weather generator model developed by Mikhail Semenov of Rothamsted Research, UK (Semenov & Barrow 1997). It generates a suite of climate variables, namely precipitation, maximum and minimum temperature and solar radiation, by utilizing semi-empirical distributions for the lengths of wet and dry day series. Precipitation in this case is considered as the primary variable, and the other three variables on a given day are conditioned on whether the day is wet or dry. The simulation of precipitation occurrence is therefore based on distributions of the length of continuous sequences of wet and dry days, whereas daily minimum and maximum temperatures are considered as a stochastic process with daily means and daily standard deviations conditioned on the wet or dry status of the day. Downscaling with LARS-WG is based on relative monthly change in mean daily precipitation amount, daily wet and dry series, mean daily temperature and standard deviations between the current and future GCM outputs.

Artificial neural networks

ANNs are simplified representations of the structure and function of the brain system that perform parallel and distributed information processing, mimicking the basic structure and operations of the neuron system. The neurons and corresponding synapses are represented as nodes and several connecting links, respectively. The strength of the connections in terms of weighting factors of output from previous nodes characterizes the synaptic efficiency of signal transmission. A bias term provides the threshold required to commence the firing of signals to the next nearby nodes. The summation of all inputs into the neural node at a specific time is filtered through the activation functions in order to control the amplitude of nodal outputs. Haykin (1994) has provided theoretical and analytical description of the neural networks in greater detail.

The architecture of ANNs is versatile enough to describe a wide range of relationships between geophysical variables. In particular, an appropriate topology of the neural network can satisfactorily represent a non-linear association that could not be explained by explicit mathematical

formulations. Among the various types of topologies, a multilayer feed-forward neural network has received considerable attention due to its ability and efficiency in approximating physical relationships. The layer of neural networks consists of nodes that have a similar order of connecting links before and after the processing node. The computations across layers proceed in one direction and are not reversible. Although feed-forward neural networks are popular in many application areas, they lack time delay and/or feedback connections necessary to provide a dynamic model for temporal sequences processing. Time-lagged feed-forward networks (TLFN) and recurrent neural networks (RNN) are the two major groups of dynamic neural networks used in time series analysis. However, the latter require complex training algorithms and are hence computationally expensive (Dibike & Coulibaly 2006). The advantage of dynamic neural networks is that they have a memory structure for holding past samples of the input signals. The application of TLFN in downscaling daily precipitation and temperatures can be found in the work of Dibike & Coulibaly (2006). In this study, multi-layer perceptron (MLP) was used for downscaling daily precipitation amount and minimum and maximum temperatures.

Table 3 | Large-scale atmospheric predictor variables for SDSM used to downscale daily temperature and precipitation at Bhardar station. The partial correlation coefficient r shows the explanatory power that is specific to each predictor. All are significant at $p = 0.05$

Predictand	Predictors (NCEP reanalysis)	Partial r
Precipitation	Surface meridional velocity	−0.053
	Relative humidity at 500 hpa	0.07
	Mean temperature at 2 m	−0.039
Maximum temperature	500 hpa zonal velocity	0.121
	500 hpa geopotential height	0.078
	Relative humidity at 500 hpa	−0.078
	Relative humidity at 850 hpa	−0.069
	Mean temperature at 2 m	0.49
Minimum temperature	500 hpa geopotential height	0.291
	Relative humidity at 500 hpa	0.191
	Surface specific humidity	0.379
	Mean temperature at 2 m	0.248

Table 4 | Large-scale climate predictors selected for computing surface meteorological variables at different stations with SDSM model (variables corresponding to each predictor no. as for Table 2)

	Predictor no. Precipitation	Maximum temperature	Minimum temperature
Pawe	5,9,23	2,17,26	3,5,12
Dangila	9,22,24	1,2,4,12,22,26	3,5,12,26

DOWNSCALING EXPERIMENTS

Statistical downscaling model

In SDSM downscaling, large-scale predictor variables representing the current climate conditions (1961–1990) derived from the NCEP reanalysis datasets were used to investigate the percentage of variance explained by each predictand–predictor pair. Predictor variables were selected based on correlation analysis, partial correlation analysis and visual inspection of scatter plots. Table 3 shows the partial correlation coefficients for each predictor–predictand relationship for Bhardar station, while Table 4 shows selected predictor variables for the other two stations used in the downscaling experiments. After potential predictor variables were identified, the coefficients of the multiple regression equation that relate potential large-scale atmospheric predictor variables derived from NCEP and local-scale surface variables were computed in the calibration process.

The SDSM model was calibrated and validated separately for daily precipitation, daily maximum and minimum temperature. Fifteen years of data (1961–1975) for Bhardar station and 10 years of data (1987–1996) for Pawe and Dangila stations were used for calibration. During calibration, some of the SDSM setup parameters such as event threshold, bias correction and variance inflation were adjusted until a good statistical agreement between observed and simulated climate variables was obtained. To assess the performance of the model for further downscaling experiments, the SDSM model was validated using the identified atmospheric predictor variables derived from the NCEP reanalysis dataset with 15 years of data (1976–1990) for Bhardar station and 5 years of data (1997–2001) for Pawe and Dangila stations.

The calibrated and validated SDSM model using observed data and large-scale predictors from NCEP reanalysis dataset was used to downscale the future climate scenario by forcing the empirical relationship between observed station data and NCEP predictors with predictor variables supplied by the HadCM3 model. One hundred synthetic daily time series were generated for A2a SRES emission scenarios for a period of 139 years (1961–2099). The generated time series for the period covering the 2050s (2040–2069) was used for hydrological impact studies. The same period (2040–2069) was also used for analysis of hydrological impact with LARS-WG and ANN downscaling, explained in the following sections.

Table 5 | Quarterly probability distributions for the length of wet and dry series of precipitation compared using the chi-squared goodness-of-fit test. The probabilities of each season are the probability that the observed and synthetic wet and dry lengths come from the same probability distribution

	Degree of freedom	Chi-squared	Probability
Dec–Feb	2	0.79	0.672
	6	60.67	0
Mar–May	5	0.73	0.981
	8	1.64	0.99
Jun–Aug	8	3.26	0.917
	6	0.21	1
Sept–Nov	9	1.03	0.999
	9	1.91	0.993

LARS-WG stochastic weather generator

Unlike SDSM and ANN, LARS-WG does not use large-scale atmospheric variables directly in the model; rather, it uses direct GCM outputs of daily precipitation and maximum and minimum temperature. Its calibration involves calculating the relevant statistical parameters for each meteorological variable from observed historical data. These calculated parameters or those modified based on future climate scenario were then used to stochastically generate realistic climate data corresponding to the present or future climate scenario, respectively. In order to ensure that the simulated data probability distributions are close to the true long-term observed distribution, the statistical characteristics of the observed data were used to generate 300 years of synthetic data for each climate variable considered.

Statistical characteristics of the observed and synthetic weather data were also analyzed using the *t*-test, *f*-test and chi-squared test to determine if there were any statistically significant differences. The chi-squared test was used to compare observed and synthetic data histograms. As can be seen in Table 5, except for a particularly dry winter spell, the chi-squared goodness-of-fit test for the seasonal mean wet and dry spell lengths resulted in confidence values of above 95%. The significant difference during the winter season is likely to be due to LARS-WG smoothing the observed data.

Tables 6 and 7 show *t*-test and *f*-test values respectively, which were used to test the equality of means and variances between the observed and simulated maximum and minimum temperature values. Each of these tests calculates *p*-values, which are used to accept or reject the hypothesis that there is no difference between the observed and simulated climate variables. As shown in Table 6, the difference between observed and simulated monthly minimum and maximum temperature values are insignificant at a significance level of 0.01. However, the *p*-value computed for the *f*-test (Table 7) shows a significant difference at significance level of 0.1 for some of the months. Again, the differences are likely to be due to the departures of the observed values from the smooth curves which LARS-WG fits to average daily mean values of minimum and maximum temperature values. (For a detailed description of possible reasons of significant difference between observed and simulated data in LARS-WG, see Semenov & Barrow (2002)). In most cases, the stochastic weather generator is able to generate synthetic data with statistical properties quite similar to the observed regional climate variables, namely daily precipitation as well as daily minimum and maximum temperature values.

To generate daily precipitation and temperature data corresponding to future scenarios using LARS-WG, site analyses were carried out using the daily GCM data (HadCM3) for both the baseline (current) and future periods. To incorporate the change in climate, the relative change in monthly mean precipitation and monthly mean wet and dry series lengths were calculated from the HadCM3 output of the baseline and future scenarios. Similarly, for each month relative change in mean temperature and standard deviation were calculated from the parameter files generated during

Table 6 | Maximum and minimum temperature per month: observed and WG mean, *t*-values and *p*-values

	Maximum temperature				Minimum temperature			
	Obs mean	WG mean	*t*-value	*p*-value	Obs mean	WG mean	*t*-value	*p*-value
J	26.4	26.6	−1.66	0.10	6.6	6.8	−0.76	0.45
F	27.7	27.8	−0.71	0.48	8.2	8.4	−0.74	0.46
M	29.3	29	2.39	0.02	11.4	11	1.37	0.17
A	29.6	29.7	−0.8	0.43	12.2	12.5	−1	0.32
M	28.7	28.9	−1.49	0.14	13.7	13.3	1.71	0.09
J	26.5	26.3	1.65	0.10	13.5	13.6	−0.63	0.53
J	23.8	24	−1.89	0.06	13.3	13.4	−0.71	0.48
A	23.6	23.6	0.00	1.00	13.1	12.8	1.84	0.07
S	24.8	24.8	0.00	1.00	12.3	12.5	−1.15	0.25
O	26.0	26.0	0.00	1.00	11.7	11.8	−0.56	0.57
N	26.1	26.0	1.09	0.28	9.7	9.6	0.42	0.67
D	26.0	25.8	1.99	0.05	6.9	7.7	−3.36	0.01

Table 7 | Maximum and minimum temperature per day: observed and WG standard deviation, *f*-values and *p*-values

	Maximum temperature				Minimum temperature			
	Obs stdv	WG stdv	*f*-value	*p*-value	Obs stdv	WG stdv	*f*-value	*p*-value
J	1.66	1.7	1.05	0.339	2.82	2.74	1.06	0.237
F	1.84	1.8	1.04	0.384	2.89	2.94	1.03	0.531
M	1.77	1.75	1.02	0.634	3.36	3.31	1.03	0.545
A	1.88	1.74	1.17	0.001	3.47	3.19	1.18	0.001
M	1.88	1.98	1.11	0.039	2.33	2.55	1.2	0.001
J	1.98	1.84	1.16	0.002	1.88	1.79	1.1	0.048
J	1.43	1.52	1.13	0.015	1.62	1.48	1.2	0.001
A	1.43	1.31	1.19	0.002	1.53	1.69	1.22	0.002
S	1.23	1.26	1.05	0.34	1.87	1.81	1.07	0.179
O	1.12	1.22	1.19	0.001	2.01	2.03	1.02	0.695
N	1.27	1.22	1.08	0.1	2.64	2.62	1.02	0.749
D	1.37	1.44	1.1	0.044	2.68	2.88	1.15	0.004

the site analysis, corresponding to climate variables (HadCM3 outputs) for the baseline and future scenarios. This information was then used to construct the climate scenario file which LARS-WG uses to determine how the weather generator parameter values (obtained from observed precipitation and temperature data) should be perturbed to generate regional climate scenarios.

Artificial neural networks

The neural network model for this study was developed using Neuro-Solutions neural network development environment. The 26 predictor variables obtained from the NCEP reanalysis datasets were used as input to the neural network model. Daily precipitation amount and maximum and minimum temperatures were specified as desired output and modelled separately. For Bhardar station, from the 30 years of data representing the current climate condition, the first 15 years of data (1961–1975) were used for training and the remaining 15 (1976–1990) were used for testing. Similarly for Pawe and Dangila stations, of the 15 years of data representing the current climate condition, the first 10 years of data (1987–1996) were used for training and the remaining 5 years (1997–2001) for testing.

The neural network design and development were carried out as follows. With the 26 predictor variables as input, the model was trained and sensitivity analysis was carried out to determine the most relevant predictors which need to be retained for further analysis or retraining. The neural network was retrained with a few selected predictor variables until acceptable validation performance was achieved. Table 8 presents the selected predictor variables based on sensitivity analysis. Hyperbolic tangent activation function was used at both the hidden and

Table 8 | Large-scale daily climate predictors selected for computing surface meteorological variables at different stations with ANN model (variables corresponding to each predictor no. as for Table 2)

	Predictor no. Precipitation	Maximum temperature	Minimum temperature
Pawe	1,5,17,19,21,22,24	1,4,7,16,17,19,21,23,24,26	1,5,16,17,23,24,25,26
Bhardar	1,5,9,16,17,23,25,26	1,2,5,12,17,19,26	1,4,5,7,21,24
Dangila	1,2,4,12,22,26	1,4,9,12,16,17,19,23	1,3,5,16,24

output layers of the neural networks and the neural network was trained using a learning rule of momentum. The numbers of hidden nodes were optimized during training. The trained ANN using NCEP predictor variables was used to downscale a future climate scenario using predictor variables from HadCM3 outputs.

HYDROLOGICAL MODELLING

Various hydrological models, ranging from simple conceptual models to comprehensive physically based models, have been used to study the impact of climate change scenarios including: Hydrologiska Byråns Vattenbalansavdelning (HBV; Dibike & Coulibaly 2005); Soil and Water Assessment Tool (SWAT; Jha *et al.* 2006); Cascade Consulting Hydrological (CAS-Hydro; Reaney & Fowler 2008); Precipitation–Runoff Modelling System (PRMS; Najafi *et al.* 2011); HEC-HMS (Simonovic *et al.* 2007); Catchment Hydrological model (CATCHMOD; Wilby 2005); and Topography-based Hydrological model (TOPMODEL; Bastola *et al.* 2011).

There are numerous criteria which can be used when selecting the appropriate hydrological model. The four common fundamental criteria are: (1) required model outputs; (2) hydrological processes that need to be modelled to estimate the desired outputs; (3) availability of input data; and (4) cost. Cunderlik (2003) carried out a comparison study of existing hydrological models for choosing an appropriate hydrological model for the assessment of water resources risk and vulnerability to changing climatic conditions. In that study, lumped semi-distributed and fully distributed models were compared. Results of the study showed that semi-distributed model HEC-HMS is one of the first ranked hydrological models. Semi-distributed models were found to be a more attractive choice than lumped or distributed models as they are a good compromise between the high simplification of the governing processes used in lumped models and extensive data requirements of the distributed models. The HEC-HMS model was selected to investigate the potential impact of climate change in this study due to the model's physical basis, limited data requirement, easy availability (public domain) and wider use.

HEC-HMS is a comprehensive hydrologic model developed by the Hydrologic Engineering Centre (HEC) of United States Army Corps of Engineers (USACE). It is designed to simulate the precipitation–runoff processes of watershed systems in a wide range of geographic areas, such as large river basins and small urban or natural watersheds (Feldman 2000). The current version of HEC-HMS is a highly flexible package. It includes different methods to simulate infiltration losses, transform excess precipitation, base flow estimation and channel routing. For this particular study, a continuous Soil Moisture Accounting (SMA) model was used. The SMA method allows for long-term continuous simulation of hydrologic processes that occur and change over time in a given watershed. This can be achieved by simulating the movement of precipitation through storage volumes that represent canopy interception, surface depressions, the soil profile and two groundwater layers. Computational components of this algorithm also include ET, surface runoff and groundwater flow. As well as precipitation, the only other input to the SMA algorithm is potential ET rate. Catchment monthly potential ET rate was determined using the Penman–Monteith method (Allen *et al.* 1998) using Ethiopia Meteorological Office monthly climate data (radiation, minimum and maximum temperature, air humidity and wind speed) as input. The spatial and temporal distribution of precipitation was determined using the gauge weight method. The Thiessen

polygon technique was used to determine the spatial gauge weights.

HEC-HMS has four main model components: basin model, meteorological model, control specifications and input data (time series, paired data and gridded data). The basin model, for instance, contains information relevant to the physical attributes of the model such as basin areas and river reach connectivity. Likewise, the meteorological model holds rainfall data. The control specifications section contains information pertinent to the timing of the model run. The input data component stores parameters or boundary conditions for basin and meteorological models. The Army Corps of Engineers Geospatial Hydrologic Modelling Extension (HEC-GeoHMS) (Doan 2000), an ArcView extension, was used for creating the basin model, background map file and stream network from the digital elevation model (DEM) which in turn were used as input to HEC-HMS model. A DEM was downloaded from the Consortium for Spatial Information (CGIAR_CSI) which provides the NASA Shuttle Radar Topographic Mission (SRTM) 90 × 90 m resolution digital elevation data for the entire world. The sub-basin physical characteristics such as longest flow lengths, centroidal flow length and slopes derived from the DEM were used for estimating hydrologic parameters. The longest flow path information, for example, was used for estimating time of concentration. Background maps can be loaded in HEC-HMS that provide spatial context for the hydrologic elements composing the basin model. The maps are not actually used in the computational process, but they can be very helpful in showing the spatial relationship between the elements. They are commonly used for showing the boundaries of a watershed or the location of streams.

Observed daily river discharges obtained from the Ministry of Water Resources of Ethiopia were used to calibrate the hydrologic model. Model parameters were determined by minimizing the sum of squares of differences between observed and simulated flows. Both automated and manual calibration methods were used to determine a practical range of parameter values, preserving the hydrograph shape and minimum error in volumes. The emphasis was on simulating low and average flows accurately rather than peak flows. The period 1994–1998 was used for model calibration and model fit was validated for 1999–2001. The maximum infiltration rate and the maximum soil depth, as well as the percolation rates and groundwater components, were found to be more sensitive to simulated flows.

The coefficient of determination (R^2) and Nash–Sutcliffe simulation efficiency (E) (Nash & Sutcliffe 1970) were used as evaluation criteria. The coefficient of determination (R^2) during calibration and validation in a daily time step were found to be 0.7 and 0.62, respectively, and Nash–Sutcliffe simulation efficiency (E) was found to be 0.66 during calibration and 0.6 during validation. Figure 2 shows observed and simulated daily river flow for the validation period. In general, the validation result shows that the model performance is reasonably good in simulating low flows for periods outside the calibration period. Table 9 summarizes calibrated model parameters. The obtained calibrated parameters, however, may still not represent one unique solution. It is always possible to construct an equally well-calibrated model with different sets of parameters (the known 'equifinality' problem).

RESULTS

For this study we considered only temperature and precipitation downscaling, which are the essential parameters for simulating the effect of climate change on river flows. For both parameters, downscaling results were presented in the subsequent sections for the present-day and future climate simulations.

Comparison of downscaling model validation statistics

The model biases (the difference between the observed and downscaled result using relevant NCEP predictor) for Bhardar station during the validation period (1976–1990) for the three downscaling models are shown in Figures 3–5. The model biases for Pawe and Dangila stations for all three meteorological variables in the validation period (1997–2001) are provided in Tables 10 and 11, respectively. In order to ensure that the derived scenarios are a result of projected future climate, the downscaling algorithms were trained with NCEP and observed data and validated with the baseline climate of NCEP and HadCM3 derived predictors. As shown in Table 12, the amount of explained

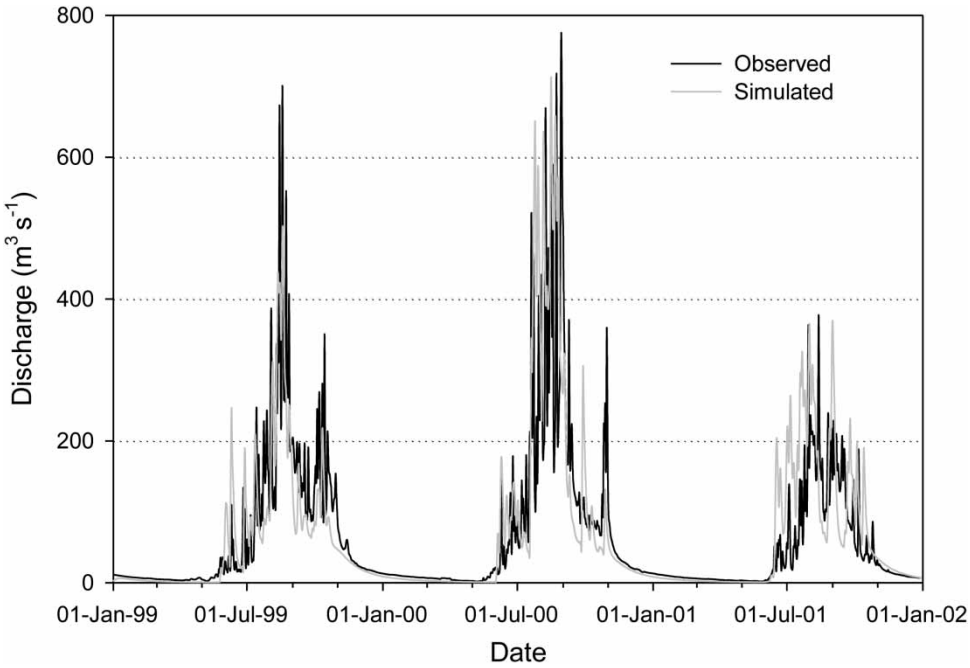

Figure 2 | Observed and simulated daily flows in the Beles River for the validation period.

Table 9 | Calibrated model parameters

Parameter	Units	Values	Parameter	Units	Values
Muskingum k	hr	6	Groundwater 1 storage	mm	50
Muskingum x	hr	0.1	Groundwater 2 percolation	mm hr^{-1}	0.6
Max infiltration	mm hr^{-1}	1.9	Groundwater 1 coefficient	hr	600
Soil storage	mm	200	Groundwater 2 storage	mm	80
Tension storage	mm	150	Groundwater 2 percolation	mm hr^{-1}	0.6
Soil percolation	mm hr^{-1}	0.9	Groundwater 2 coefficient	hr	1,800
Storage coefficient	hr	60			
Time of concentration upper sub basin	hr	24	Time of concentration lower sub-basin	hr	12

variance (R^2) obtained for the validation period (using HadCM3 daily predictors) varies between 0.17 and 0.21 for precipitation, 0.39 and 0.52 for maximum temperature and 0.33 and 0.37 for minimum temperature. The low explained variance for precipitation underlines the more stochastic nature of precipitation and the difficulty in capturing its characteristics and variability in the downscaling. Wilby *et al.* (2002) reported that, for heterogeneous variables such as daily precipitation occurrence or amounts, explained variance of less than 40% is more likely. Since precipitation depends on other intermediate processes such as humidity, cloud cover, etc., it is identified by many researchers as one of the most problematic variables in downscaling (Zhang *et al.* 2012).

The comparison of means of observed and downscaled daily precipitation in each month (Figure 3) shows that SDSM and LARS-WG models were able to reproduce the observed mean daily rainfall better than ANN in all months. One reason could be that SDSM and LARS-WG consider precipitation downscaling as a conditional process

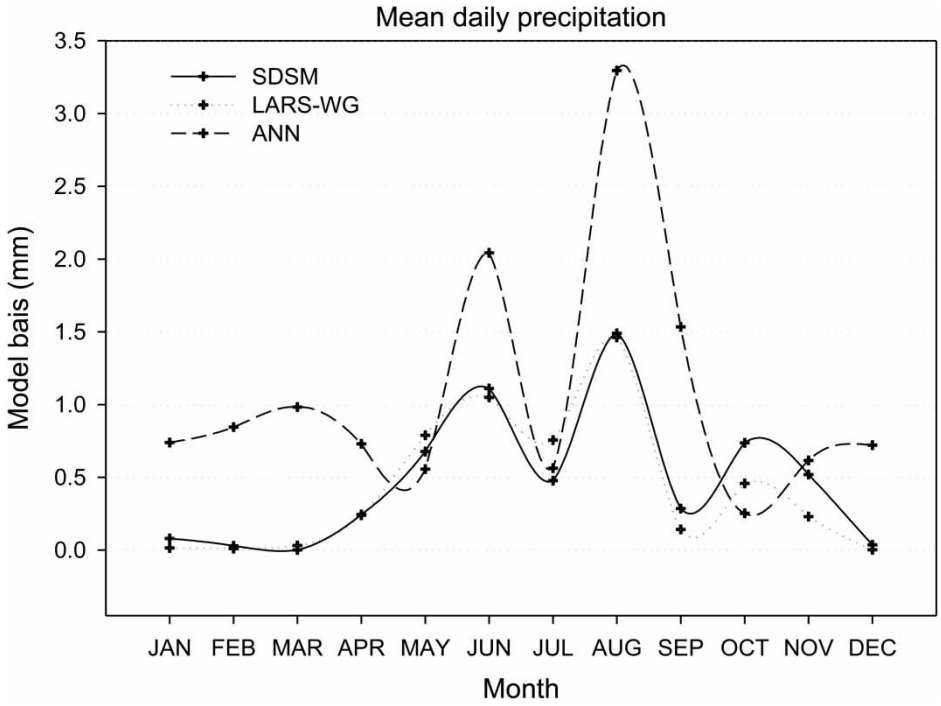

Figure 3 | Model biases in absolute values (the difference between the observed and downscaled mean daily precipitation using relevant NCEP predictors) for Bhardar station during validation period (1976–1990).

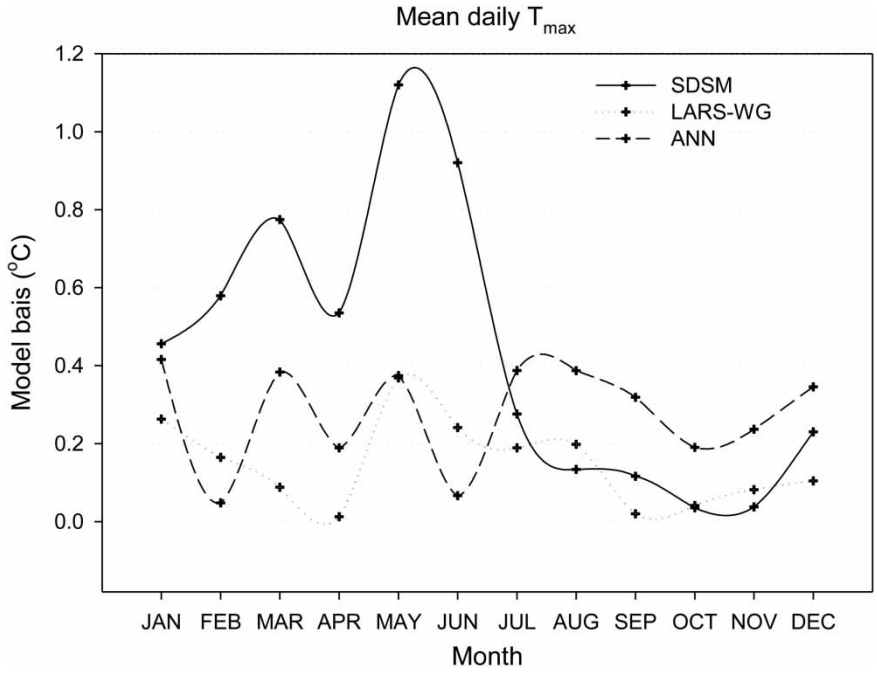

Figure 4 | Model biases in absolute values (the difference between the observed and downscaled mean daily maximum temperature using relevant NCEP predictors) for Bhardar station during validation period (1976–1990).

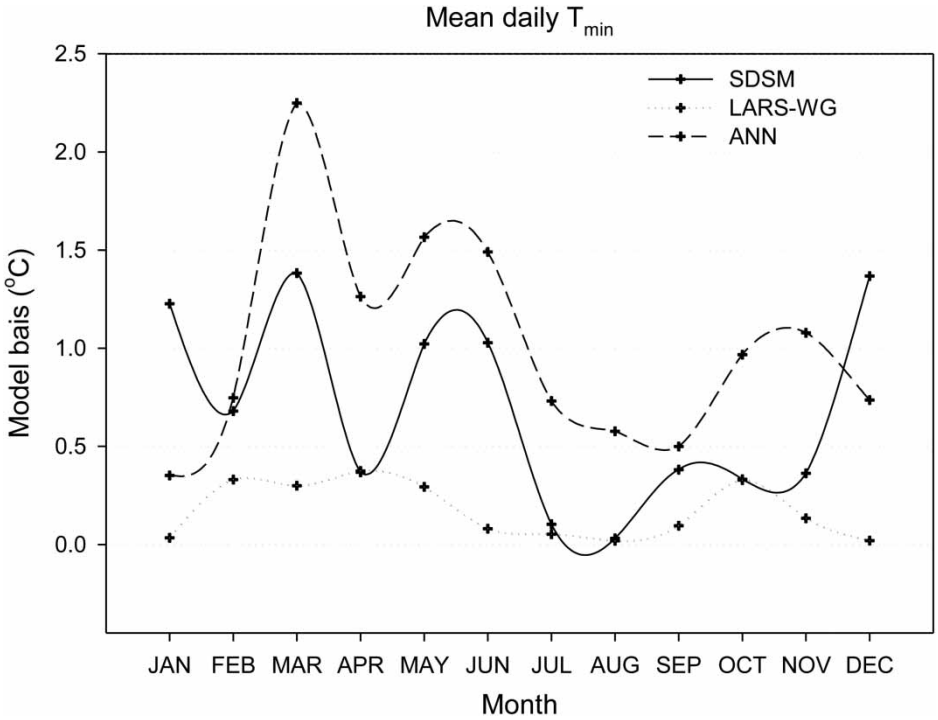

Figure 5 | Model biases in absolute values (the difference between the observed and downscaled mean daily minimum temperature using relevant NCEP predictors) for Bhardar station during validation period (1976–1990).

Table 10 | Model bias between observed and simulated values for the validation period for Pawe station

	Mean daily precipitation (mm)			Mean daily T_{max} (°C)			Mean daily T_{min} (°C)		
	SDSM	LARS	ANN	SDSM	LARS	ANN	SDSM	LARS	ANN
J	0.02	0.01	0.22	0.24	0.31	0.42	−0.40	0.74	1.24
F	0.03	0.11	0.57	0.23	−0.17	−0.23	0.46	0.80	0.62
M	0.08	0.14	0.60	0.45	−0.17	−0.25	−1.79	−0.82	−1.97
A	0.32	0.31	0.98	0.67	0.32	−0.40	−0.27	−0.02	−1.22
M	−2.08	0.18	−0.61	1.00	0.62	0.52	0.27	0.09	−0.24
J	−0.71	−1.69	−1.61	0.93	0.18	1.22	0.43	0.75	0.34
J	1.39	1.39	1.39	0.16	0.17	0.47	0.29	−0.17	0.27
A	−0.84	2.66	−1.96	0.08	−0.03	0.36	0.37	−0.05	0.50
S	1.74	1.63	0.83	−0.38	−0.62	0.79	0.27	0.36	0.67
O	−0.50	−0.88	−1.39	1.01	0.30	2.44	−0.40	−0.95	−0.26
N	0.80	0.30	0.35	0.83	0.41	1.32	−0.38	0.07	0.45
D	−0.06	0.04	0.22	0.54	0.08	0.48	−0.83	−0.76	0.23

in which local precipitation is correlated with the occurrence of wet days, whereas the ANN model does not consider precipitation as a conditional process. It only establishes a direct non-linear relationship between large-scale predictors and local-scale predictands without considering the precipitation occurring process. Indeed, in another study, the ANN model was also found to be poor in simulating daily precipitation, particularly for wet-day occurrence,

Table 11 | Model bias between observed and simulated values for the validation period for Dangila station

	Mean daily precipitation (mm)			Mean daily T_{max} (°C)			Mean daily T_{min} (°C)		
	SDSM	LARS	ANN	SDSM	LARS	ANN	SDSM	LARS	ANN
J	0.03	0.12	0.22	0.72	0.24	0.44	0.54	0.56	1.99
F	0.02	−0.02	0.54	−0.06	−0.56	−0.47	0.47	−0.46	0.79
M	0.87	–0.36	0.49	0.20	−0.41	−0.05	−0.49	−0.55	−0.02
A	0.46	0.58	1.18	0.70	−0.13	−0.11	0.36	−0.22	−0.27
M	−1.92	−0.70	−1.56	0.61	0.45	0.99	−0.05	0.79	–0.04
J	−1.05	−0.54	−2.43	0.59	−0.37	0.95	0.74	0.59	0.57
J	0.71	0.39	−0.35	0.90	0.55	1.05	−0.04	−0.32	0.13
A	1.65	0.33	−0.08	0.59	−0.03	1.16	0.39	−0.34	0.21
S	0.28	1.79	−0.27	0.09	−0.66	0.29	0.29	0.61	1.09
O	−1.46	−0.13	−2.06	0.56	0.33	1.04	0.13	−0.07	−0.13
N	−0.42	−0.14	−0.42	0.06	0.12	0.79	1.13	1.23	1.22
D	0.54	0.26	0.19	0.00	−0.56	0.14	0.31	−0.14	1.16

Table 12 | Correlation coefficients and explained variances between observed daily precipitations, maximum and minimum temperature and corresponding relevant HadCM3 derived predictor variables for Bhardar station during validation. For the case of LARS-WG, large-scale predictor variables are not used directly in the model; rather it is based on direct GCM output

	Precipitation		Maximum temperature		Minimum temperature	
	Correlation	R^2	Correlation	R^2	Correlation	R^2
SDSM	0.46	0.21	0.71	0.51	0.61	0.37
ANN	0.45	0.2	0.62	0.39	0.57	0.33

due to the simplistic treatment of days with zero amounts (Khan *et al.* 2006). In the case of LARS-WG, as shown in Table 6, model biases are found to be insignificant for both maximum and minimum temperature downscaling.

Comparative downscaling results

The monthly statistics of downscaling results for daily precipitation as well as daily maximum and minimum temperature for the three downscaling models are summarized and plotted in Figures 6–8. Figure 6 shows a decreasing trend in summer precipitation by both the SDSM and ANN models. However, the overall result of the LARS-WG seems to suggest the possibility of an increase in mean daily precipitation in the region for future scenarios considered in this study. As shown in Table 13, there is a large divergence regarding the change in summer precipitation which ranges from +5.6% (LARS-WG) to –27.3% (SDSM) at Bhardar station, –4.5% (LARS-WG) to +57.6% (ANN) at Pawe station and +14.5% (LARS-WG) to –31.1% (ANN) at Dangila station by the 2050s. Figures 7 and 8 show a consistently increasing trend both in maximum and minimum temperature values for all months of the year. However, large differences emerge in the rates of annual warming downscaled with the three models (Table 13). The rate of annual maximum temperature change between the 1961–1990 and 2050s range from: +1.5 °C (SDSM) to +3 °C (ANN) at Bhardar station, +0.7 °C (SDSM) to +2.5 °C (LARS-WG) at Pawe station and +2 °C (ANN) to +4 °C (SDSM) at Dangila station. The annual rate of change of minimum temperature at Bhardar station also ranges from +1.2 °C (SDSM) to 3 °C (ANN).

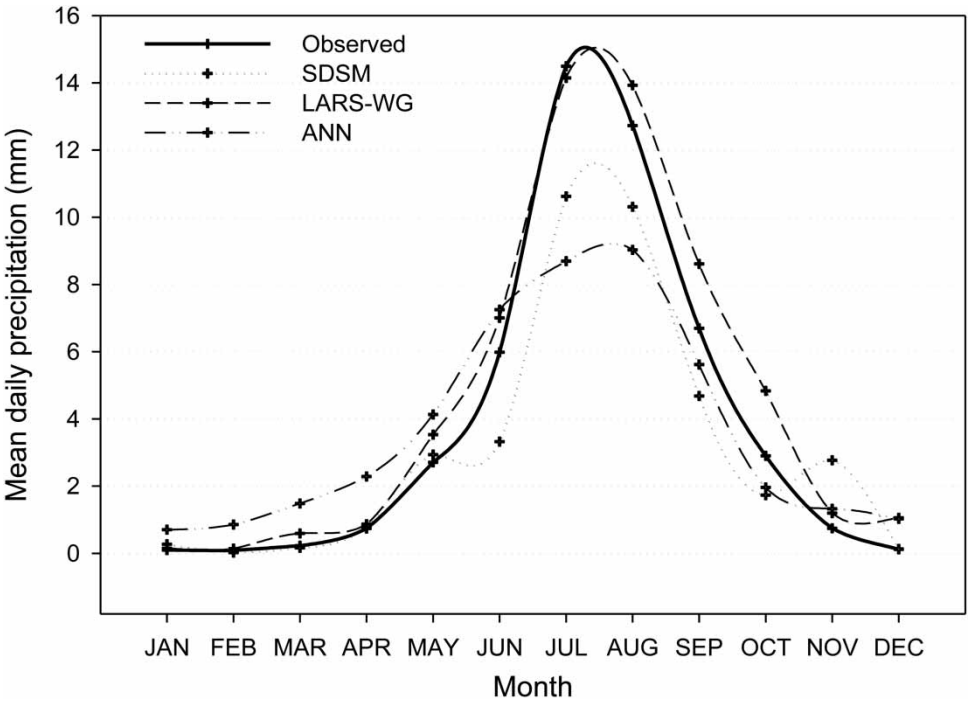

Figure 6 | General trend in mean daily precipitation corresponding to climate change scenario (2040–2069) and observed mean daily precipitation (1961–1990) at Bhardar station.

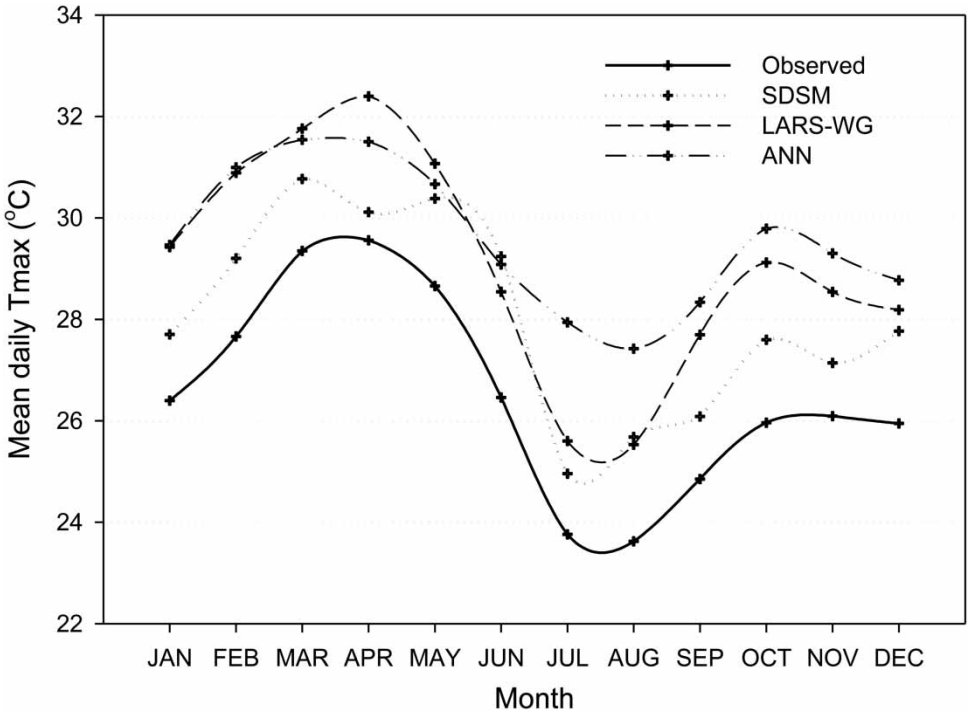

Figure 7 | General trend in mean daily maximum temperature corresponding to climate change scenario (2040–2069) and observed mean daily maximum temperature (1961–1990) at Bhardar station.

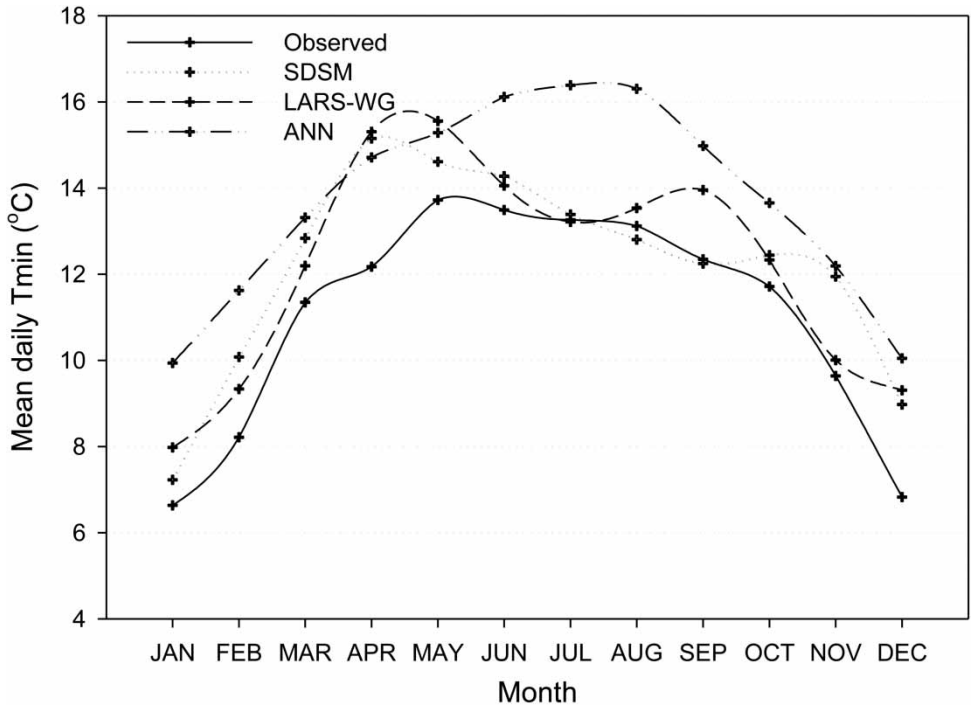

Figure 8 | General trend in mean daily minimum temperature corresponding to climate change scenario (2040–2069) and observed mean daily minimum temperature (1961–1990) at Bhardar station.

Table 13 | Changes in downscaled temperature and precipitation (2050s) with respect to the 1961–1990 average

		Bhardar			Pawe			Dangila		
		SDSM	LARS	ANN	SDSM	LARS	ANN	SDSM	LARS	ANN
Maximum temperature (°C)	Summer	2.0	1.9	3.5	1.5	4.4	1.0	5.5	3.7	2.6
	Annual	1.5	2.5	3.0	0.7	2.5	1.1	4.0	2.5	2.0
Minimum temperature (°C)	Summer	0.4	0.3	2.5	0.4	0.1	1.8	1.7	− 0.3	− 0.2
	Annual	1.4	1.2	3.0	1.5	1.1	1.1	1.6	1.1	0.0
Precipitation (%)	Summer	− 27.3	5.6	− 25.0	4.1	− 4.5	57.6	− 0.1	14.5	− 31.1
	Annual	− 21.3	17.9	− 7.0	2.7	22.6	− 5.4	2.3	12.1	20.3

DISCUSSION

The ultimate goal of downscaling is to generate an estimate of meteorological variables corresponding to a given scenario of future climate, so that these meteorological variables can be used as a basis for hydrological impact assessments. After calibrating and validating the hydrological model with the historical data, we have used this model to simulate flows corresponding to future climate conditions corresponding to the downscaled precipitation and temperature data. This will help to identify any specific trend in the mean flows in the Beles River, corresponding to the future time horizon considered in this study. The future simulation (2050s) was carried out with the downscaled precipitation and temperature data from the three downscaling models, namely SDSM, ANN and LARS-WG. The generated river flows were then used for assessing the epistemic and stochastic uncertainty.

The epistemic uncertainty (i.e. the uncertainty due to imperfect knowledge) was first assessed by comparing the impact of projected climate change on simulated river flows for the baseline (1961–1990) and future climate conditions (2040–2069), corresponding to downscaled precipitation and temperature time series using three conceptually different downscaling methods (SDSM, LARS-WG and ANN). Subsequently, the stochastic uncertainty (i.e. uncertainty due to inherent variability, e.g. climate variability) was assessed by simulating river flows for the same time period using different ensemble members of downscaled precipitation and temperature time series.

The latter was performed using only SDSM, because the random element in SDSM introduces variability in each of the downscaled time series and hence has been used to derive 100 daily precipitation as well as maximum and minimum temperature values. However, LARS-WG and ANN are limited to producing only one time series of the downscaled variables; it was therefore not possible to assess this uncertainty component using these models. In this study, the SDSM was used in two ways: (1) the mean of the 100 downscaled ensemble members was used for the hydrological model to simulate river flow; and (2) among the 100 time series (ensembles), 11 individual time series (1st, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th and 100th) were used as input for hydrological simulations.

The impact of projected climate change in mean annual river flow ranges from –3.3% (SDSM) to +17.5% (LARS-WG) between the baseline and 2050s (Table 13). Large differences in flow simulation also emerge in mean river flow changes in summer (–4.4% for SDSM to +12.3% for ANN) and in winter (+18.4% for SDSM to +55% for ANN) between the baseline and 2050s, as shown in Table 14. The percentage change in mean annual flow change between the current and future scenario considered in this study using different ensemble members downscaled with SDSM model ranges between 6% increase and 8% reduction (Table 15).

The epistemic uncertainty seems to be larger than the stochastic uncertainty. Figure 9 shows the comparison of simulated change in the average monthly mean flows of Beles River, corresponding to the downscaled precipitation and temperature data of the current (1961–1990) and future (2040–2069) climate using the three downscaling models. As can be seen in Figure 9, the major difference in simulated flow occurs in the month of June. In this month, LARS-WG showed significant reduction in flow. Although the magnitudes are different, SDSM and ANN showed a possible reduction in flows from June to September and a possible increase from October to December. While ANN showed an increase in mean monthly flows in all months, SDSM showed a decrease in mean monthly flows for almost all months. In contrast, in the case of LARS-WG, the number of months where river flow increases is less than the number of months where river flow decreases. Figure 10 also shows the average mean flow change in river flow simulated with the downscaled

Table 14 | Simulated change in mean Beles flow between current (1961–1990) and future (2040–2069) climates corresponding to precipitation and temperature values downscaled with the three downscaling models

	Mean annual flow increase or decrease (%)	Mean winter flow increase or decrease (%)	Mean summer flow increase or decrease (%)
SDSM	−3.3	18.4	−4.4
LARS-WG	−4.8	29.5	−14.6
ANN	17.5	55	12.3

Table 15 | Simulated change in mean Beles flow between current (1961–1990) and future (2040–2069) climates corresponding to precipitation and temperature ensemble members of SDSM downscaled values

Ensemble member no.	Average increase or decrease (%)	Ensemble member no.	Average increase or decrease (%)	Ensemble member no.	Average increase or decrease (%)
1	2.2	40	−0.72	80	5.92
10	−0.7	50	5.04	90	−4.29
20	−5.5	60	−8.20	100	−6.36
30	0.48	70	−4.07		

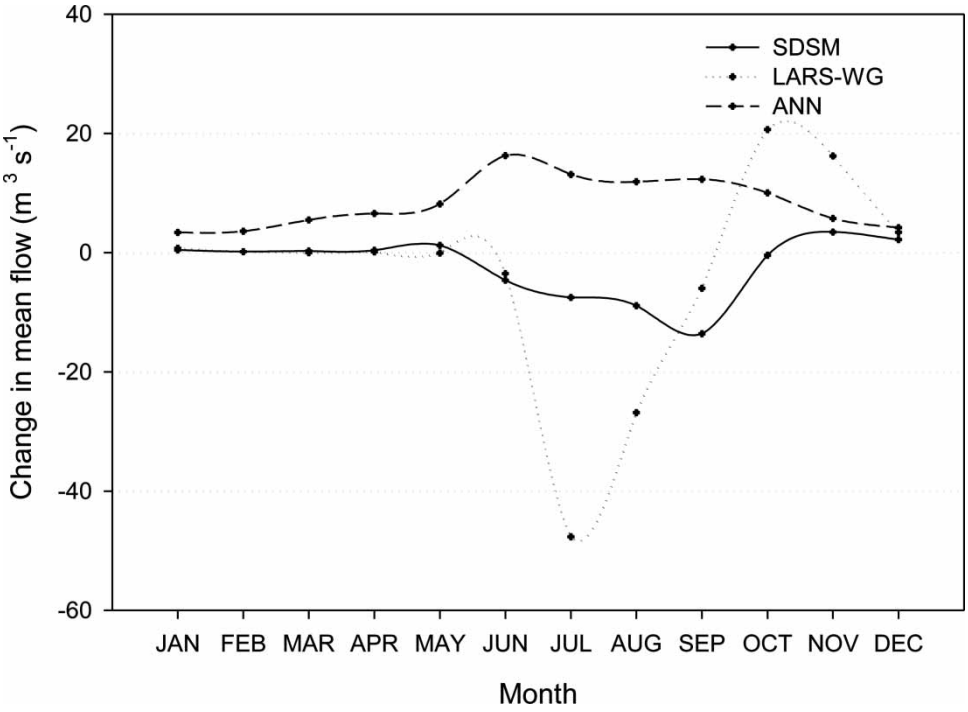

Figure 9 | Simulated change in the average monthly mean flows in Beles River corresponding to the downscaled precipitation and temperature data of the current (1961–1990) and future (2040–2069) 2050s climate with the three downscaling models.

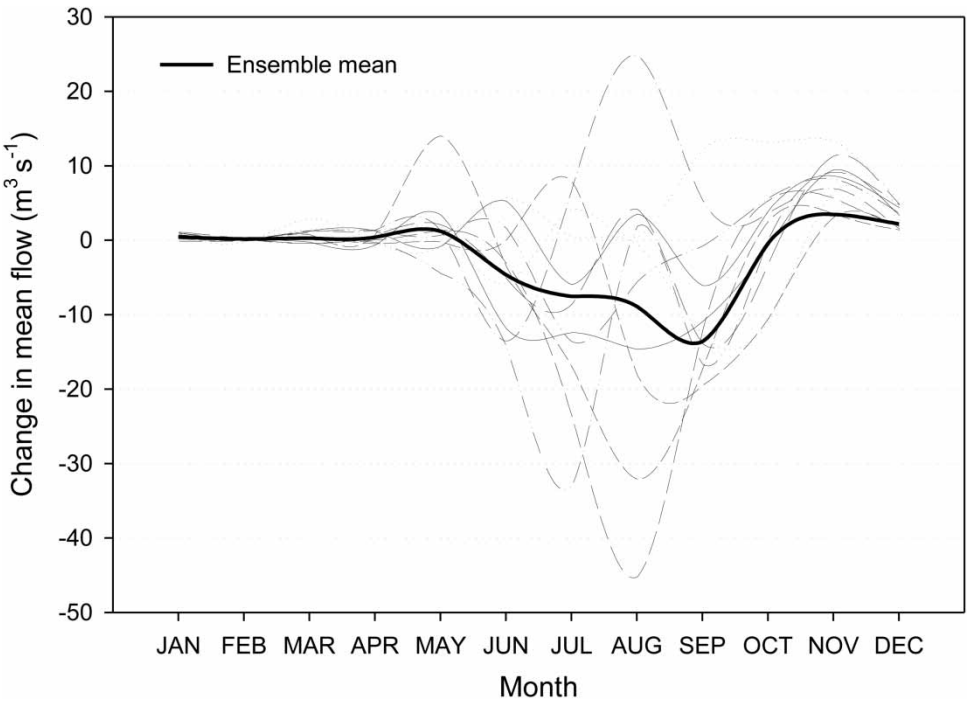

Figure 10 | Simulated change in the average monthly mean flows in Beles River corresponding to the downscaled precipitation and temperature data of the current (1961–1990) and future (2040–2069) 2050s climate using SDSM. The heavy solid line represents the simulated river flow using mean of the downscaled variables and the thin lines represent the simulated flow using individual downscaled ensemble members.

climate variables using the SDSM model for different ensemble members (1st, 10th, 20th … 100th).

CONCLUSIONS

A comparison of three downscaling models (SDSM, ANN and LARS-WG) for daily maximum and minimum surface temperature and precipitation values for A2 SRES emission scenario were presented. The comparative downscaling results show that both maximum and minimum temperatures have an increasing trend throughout the year for all the stations in a future climate. For the case of precipitation, this trend is however not clear; moreover, there is a remarkable variability between stations and downscaling methods. More specifically, the precipitation downscaling at Bhardar and Dangila stations by both SDSM and ANN downscaling models shows a decreasing trend in summer precipitation (rainy season of the region) while LARS-WG shows an increase in summer precipitation. In contrast, while SDSM and ANN show an increase at Pawe station, LARS-WG shows a decrease in summer precipitation (Table 13).

The result of projected impact on river flow caused by climate change is subject to various sources of cascading uncertainties. This study investigated the impact of climate change on river flows using an HEC-HMS rainfall–runoff model driven by downscaled daily precipitation and temperature values. The ability of downscaling methods to represent the baseline climate is a necessary condition (not sufficient) to have reasonable confidence on the reliability of the future climate computed from the scenario runs. This study specifically addressed uncertainties associated with downscaling methods.

The epistemic uncertainty of downscaling methods was investigated by comparing three conceptually different SDSMs; the stochastic uncertainty was investigated using stochastic components in one of the downscaling models (SDSM). Regarding epistemic uncertainty, depending on the statistical downscaling methods used, we obtained a highly variable projection of future stream flow changes. The results of the three downscaling models showed that the percentage change in mean annual flow ranges from a 4.8% reduction to a 17.5% increase. The results also show a possible decrease (15%) and an increase (12%) in mean monthly river flow during the summer. Regarding stochastic uncertainty, the percentage change in mean annual flow simulated with different ensemble members of the SDSM model showed mean annual flow change ranging from a 6% increase to an 8% decrease. These uncertainties are particularly associated with the lack of ability of models to downscale future precipitation values which have a major impact on the stream flow simulations. An important consequence of this finding is that uncertainty across the downscaling models seems to be larger than the stochastic uncertainty. The uncertainty shown here clearly indicates how the outcome of a hydrologic impact study can be affected by the choice of a downscaling technique.

It should be noted that the standard error of a downscaling estimate is inversely proportional to the sample size (record length). There is therefore a bias associated with the short record length of Pawe and Dangila considered in the downscaling experiments. Relatively shorter records may not reveal the full extent of natural variability and also may not capture some of the less frequent climate events (e.g. droughts); long-term records can however place recent trends and extremes in a broader context. Only changes in the climatic inputs to the catchment system were considered here; it was assumed that catchment land use remained constant and that there would be no changes in catchment physical properties. Finally, downscaling in this study is based on a single GCM model output (HadCM3). However, previous studies showed that data taken from different GCMs could differ significantly. Analyzing the uncertainties from climate forcings through multi-model ensembles and constraining the hydrological model with ensemble spread of information therefore remains the recommended approach.

ACKNOWLEDGEMENTS

This study was made possible through a scholarship grant from the Water Resources Planning and Management Project, Nile Basin Initiative. The authors would like to acknowledge the Ethiopian Meteorological Agency and Ministry of Water Resources for providing station data, the Canadian Climate Research Data Distribution Centre for providing GCM data and the Hadley Centre in the UK for

access to daily values of the HadCM3 outputs. The authors are very grateful to Yonas Birhan Dibike and Solomon Seyoum Demissie for their continuous support during this research. The authors also thank Wendy Sturrock for improving the English. The research has been partly funded by the UNESCO-IHE Partnership Research Fund (UPaRF) within the ACCION project (Adaptation to Climate Change Impact on the Nile river basin) and EU/FP7 AQUAREHAB Project (Grant Agreement No. 226565).

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998 Crop evapotranspiration. Guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. *FAO, Rome* **300**, 6541.

Bastola, S., Murphy, C. & Sweeney, J. 2011 The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Advances in Water Resources* **34** (5), 562–576.

Beven, K. & Binley, A. 1992 The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes* **6** (3), 279–298.

Blöschl, G. & Montanari, A. 2010 Climate change impacts – throwing the dice? *Hydrological Processes* **24** (3), 374–381.

Busuioc, A., Chen, D. & Hellström, C. 2001 Performance of statistical downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation. *International Journal of Climatology* **21** (5), 557–578.

Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H. 2004 An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology* **298** (1–4), 242–266.

Cunderlik, J. 2003 Hydrologic Model Selection for the CFCAS Project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Condition. *Water Resources Research Report*, University of Western Ontario, Department of Civil and Environmental Engineering. Available from: http://ir.lib.uwo.ca/wrrr/9/.

Daren Harmel, R. & Smith, P. K. 2007 Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *Journal of Hydrology* **337** (3–4), 326–336.

Di Baldassarre, G., Elshamy, M., van Griensven, A., Soliman, E., Kigobe, M., Ndomba, P., Mutemi, J., Mutua, F., Moges, S. & Xuan, Y. 2011 Future hydrology and climate in the River Nile basin: a review. *Hydrological Sciences Journal* **56** (2), 199–211.

Dibike, Y. B. & Coulibaly, P. 2005 Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *Journal of Hydrology* **307** (1/4), 145–163.

Dibike, Y. B. & Coulibaly, P. 2006 Temporal neural networks for downscaling climate variability and extremes. *Neural Networks* **19** (2), 135–144.

Dibike, Y. B., Gachon, P., St-Hilaire, A., Ouarda, T. & Nguyen, V. T. V. 2008 Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theoretical and Applied Climatology* **91** (1), 149–170.

Doan, J. 2000 *Geospatial Hydrologic Modeling Extension HEC-GeoHMS-User's Manual-Version 1.0*. Davis, California.

Elshamy, M., Seierstad, I. & Sorteberg, A. 2008 Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios. *Hydrology and Earth System Sciences Discussions* **5** (3), 1407–1439.

Feldman, A. 2000 *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*. US Army Corps of Engineers, Hydrologic Engineering Center. Available from: http://rivers.snre.umich.edu/639rivmod/hms_technical.pdf.

Giorgi, F., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L., Fu, C., Arritt, R., Bates, B., Benestad, R., Boer, G., Buishand, A., Castro, M., Chen, D., Cramer, W., Crane, R., Crossly, J., Dehn, M., Dethloff, K., Dippner, J., Emori, S., Francisco, R., Fyfe, J., Gerstengarbe, F., Gutowski, W., Gyalistras, D., Hanssen-Bauer, I., Hantel, M., Hassell, D., Heimann, D., Jack, C., Jacobeit, J., Kato, H., Katz, R., Kauker, F., Knutson, T., Lal, M., Landsea, C., Laprise, R., Leung, L., Lynch, A., May, W., McGregor, J., Miller, N., Murphy, J., Ribalaygua, J., Rinke, A., Rummukainen, M., Semazzi, F., Walsh, K., Werner, P., Widmann, M., Wilby, R., Wild, M. & Xue, Y. 2001 Regional Climate Information – Evaluation and Projections. In: *Climate Change 2001: The Scientific Basis*. Contribution of Working Group to the Third Assessment Report of the Intergovernmental Panel on Climate Change (J. T. Houghton *et al.* (eds)). Cambridge University Press, Cambridge, 881 pp. Available from: http://www.grida.no/climate/ipcc_tar/wg1/pdf/TAR-10.PDF.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. & Wood, R. A. 2000 The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16** (2), 147–168.

Götzinger, J. & Bárdossy, A. 2008 Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research* **44** (12), W00B07.

Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation*. Macmillan, New York, NY.

Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L. & Goodess, C. M. 2006 Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology* **26** (10), 1397–1415.

Jha, M., Arnold, J. G., Gassman, P. W., Giorgi, F. & Gu, R. R. 2006 Climate change sensitivity assessment on Upper Mississippi River Basin stream flow using SWAT. *JAWRA Journal of the American Water Resources Association* **42** (4), 997–1015.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R. & Joseph, D. 1996 The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77** (3), 437–471.

Khan, M. S. & Coulibaly, P. 2009 Assessing hydrologic impact of climate change with uncertainty estimates: Bayesian neural network approach. *Journal of Hydrometeorology* **11** (2), 482–495.

Khan, M., Coulibaly, P. & Dibike, Y. 2006 Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology* **319** (1–4), 357–382.

Maurer, E. P. & Hidalgo, H. G. 2008 Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences* **12** (2), 551–563.

Montanari, A. 2007 What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes* **21** (6), 841–845.

Najafi, M. R., Moradkhani, H. & Jung, I. W. 2011 Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrological Processes* **25** (18), 2814–2826.

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T. Y. & Kram, T. 2000 Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change. Pacific Northwest National Laboratory, Richland, WA (USA), Environmental Molecular Sciences Laboratory (USA).

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I: a discussion of principles. *Journal of Hydrology* **10** (3), 282–290.

Reaney, S. M. & Fowler, H. J. 2008 Uncertainty estimation of climate change impacts on river flow incorporating stochastic downscaling and hydrological model parameterisation error sources. BHS 10th National Hydrology Symposium, Exeter.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J. & van der Keur, P. 2006 A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* **29** (11), 1586–1597.

Semenov, M. A. & Barrow, E. M. 1997 Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change* **35** (4), 397–414.

Semenov, M. A. & Barrow, E. M. 2002 LARS-WG: A Stochastic Weather Generator for use in Climate Impact Studies. Version 3.0. User Manual. Available from: http://www.rothamsted.ac.uk/mas-models/larswg.php.

Simonovic, S., Prodanovic, P. & Helsten, M. 2007 Inverse modelling of climatic change impacts in the Upper Thames River Basin. Challenges for Water Resources Engineering in a Changing World, 18th Canadian Hydrotechnical Conference, Winnipeg, Manitoba.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M. & Miller, H. 2007 IPCC 2007: Climate Change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York.

Walker, W., Harremoës, P., Rotmans, J., Van der Sluijs, J., Van Asselt, M., Janssen, P. & Von Krauss, M. 2003 Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment* **4** (1), 5–17.

Wilby, R. L. 2005 Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrological Processes* **19** (16), 3201–3219.

Wilby, R. L. & Wigley, T. M. L. 2000 Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology* **20** (6), 641–661.

Wilby, R. L., Dawson, C. W. & Barrow, E. M. 2002 SDSM – a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling and Software* **17** (2), 145–157.

Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. 2004 Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change* **62** (1), 189–216.

Zhang, L., Lu, W., An, Y., Li, D. & Gong, L. 2012 Response of non-point source pollutant loads to climate change in the Shitoukoumen reservoir catchment. *Environmental Monitoring and Assessment* **184** (1), 581–594.

First received 4 February 2011; accepted in revised form 22 December 2011. Available online 30 July 2012