Journal of Agrometeorology 17 (1) : 29-35 (June 2015)

Evaluation of statistical corrective methods to minimize bias at different time scales in a regional climate model driven data

SAMANPREET KAUR[1], SK JALOTA[2], HARSIMRAN KAUR[2], BB VASHISHT[2], UR JALOTA[3] and PPS LUBANA[1]

[1]Department of Soil and Water Engineering [2]Department of Soil Science, PAU, Ludhiana

[3]Department of Mathematics, Arya College, Ludhiana – 141004 (Punjab), India

Corresponding author E-mail: samanpreet1974@gmail.com

ABSTRACT

The regional climate models provide sufficient information of the climate data, which can be used for simulating the impact of expected climate change on crop growth and hydrological processes. But future climate data derived from such models often suffers from bias and is not ready to use *per se* in crop growth/hydrological models, wherein reasonable and consistent meteorological daily input data is a crucial factor. The present study concerns the assessment and minimization of the bias in the PRECIS modeled data of maximum and minimum temperatures and rainfall for Ludhiana station, representing central Punjab of India. The correction functions for three corrective methods i.e. difference, modified difference and statistical bias correction at daily, monthly and annual time scales were developed and validated to minimize the bias. Amongst these, correction functions derived using modified difference method at daily time scale for rainfall and at monthly time scale for Tmax and Tmin were found to be the superseding.

Key words: Past climatic data, bias, bias correction methods, correction functions, time scale

Increasing observed temperature and rainfall over last few decades and related changes in the large-scale hydrological cycle due to anthropogenic interventions are posing an unprecedented challenge for crop production and hydrology (Parry *et al.* 2007; Bates *et al,* 2008). These challenges are likely to aggravate in future. Climate models are the main tools available for developing projections of climate change in the future. In climate change studies, general circular models (GCMs) and regional climate models (RCMs) are used to predict changing levels of CO_2, temperature and rainfall under different scenarios. The most commonly used GCMs are Hadley Centre Coupled Model version 3 (HadCM3), Commonwealth Scientific & Industrial Research Organization Mark 2 (CSIRO-Mk2) and Second version of Canadian Center for Climate Modeling and Analysis Coupled Global Climate Model (CCCMA-CGCM2). In recent years, the usage of RCMs has increased because of their improved ability to reproduce the present day climate (Xu *et al.,* 2005). Raw outputs of the climatic parameters from RCM models often suffer from systematic errors which may prevent their direct application for the analysis of the behavior of the climate system, its eventual changes and their local impacts. The errors in modeled daily rainfall and temperature may afflict the monthly or annual time trends and magnitude. Andreasson et al. (2004) pointed out that these biases are particularly pronounced for rainfall than temperature. Therefore, projected raw data must be made bias free using some corrections based on statistical corrective methods (Sharma *et al.,* 2007; Hansen *et al.,* 2006; Feddersen and Andersen, 2005). A number of statistical correction techniques from simple to advanced (Boberg *et al.,* 2007, Teutscbein and Seibert, 2010), to remove the bias in rainfall and temperature have been quoted in literature. The underlying assumptions is that the corrections methods and their parameters are valid for longer area and remain constant over time especially when moving from baseline to scenario simulations (Déqué, 2007; Hashino *et al,* 2007), but the improvements to the statistical properties of the data are limited to the specific time scale of the fluctuations and the site. Keeping this in view, the present study was undertaken with the objective to develop and validate correction functions using Statistical Bias correction (SBC) and other simple correction methods for matching the statistical parameters i.e. mean (μ), standard deviation (σ) and variance (σ^2) of the corrected modeled data with the observed.

MATERIALS AND METHODS

The study was carried out for Ludhiana (75° 52' E longitude and 30° 56' N latitude) with elevation of 230 m in the west and 273 m in the East. The data on rainfall (RF), minimum temperature (T_{min}) and maximum temperature (T_{max}) recorded at meteorological observatory for the period from1971 to 2010 was used.

Retrieval of climate data

Regional climate model data was obtained from the Indian Institute of Tropical Meteorology (IITM) - Pune, as the output of PRECIS (providing regional climate data for impact studies) at daily interval at a resolution of about 50 km. Simulated climate outputs from PRECIS for the present (1961–1990, Base Line) near term (2021–2050, Mid Century) and long term (2071–2100 End Century) for A1B, A2 and B2 scenario have been used. Bias of T_{max}, T_{min} was evaluated by analyzing trend at monthly time scale and of RF as periodic cumulative as well as their statistical parameters like μ, σ and σ^2.

Correction of modeled data reference to the observed

Following three methods were applied to bring the modeled data close to the observed in respect to time trends and magnitude.

Difference method : In this method, averaged daily difference of observed and modeled values (Δx) of a climate parameter (x) was taken for each Julian day (365 days) averaged from 15 years data (1971-1985). The (Δx) was considered as daily correction factor, which was added to the modeled uncorrected value (x $model_{uncor}$) to correct it (x $model_{cor}$) so that the values approach the observed ones

$$x\ model_{cor} = x\ model_{uncor} + (\Delta x).$$

Modified difference method : The modified difference method (dm) was similar to the difference method; however, some statistical parameters were added to improve the correction function. For example in temperature correction, μ and σ were added which aimed at shifting and scaling to adjust the μ and σ^2 (Leander and Buishand, 2007).

$$RF\ model_{cor} = (RF\ model_{unncor} + (\Delta x)) * (\sigma\ RF_{obs} / \sigma\ RF_{mod})$$

Statistical bias correction : Statistical bias correction (SBC) is a mathematical procedure (a functional) that maps the probability density function (PDF) of model data onto that of the observations. In climate generation studies, this is used to correct the cumulative distribution function (CDF) of the future modeled data in relation to the observed. Such corrections have already been applied separately for rainfall (Piani *et al,* 2009) and temperature data (Sennikovs and Bethers, 2009).

For all of the above discussed methods, transfer functions were also developed at different time scales *i.e.* daily (D), monthly (M) and annual (A) by cascading the original series of climate. This was done to take into account the little fluctuations of temperature/rainfall in some months of a year as a result of the systematic seasonal dependence of statistical expectation value within the month but rather due to the natural fluctuations from one day to the next (Haerter *et al,* 2010). The developed correction factors were validated on independent dataset of 5 years (from 1986 to 1990).

The correction capability of the correction functions was tested by coefficient of variation (CV) expressed as normalized percent root mean square error (NRMSE).

$$NRMSE = \frac{\left[\sum_{i=1}^{n}(P_i - O_i)^2 / n\right]^{0.5}}{\bar{O}} x100$$

Where Pi and Oi are the predicted and observed values, respectively, is the average of the observed data, and *i* is the number of observations ranging from 1 to n. The value equal to zero for a model shows perfect fit between the observed and predicted data.

RESULTS AND DISCUSSION

Correction of modeled data

Twenty years (1971-1990) monthly averages of the observed and PRECIS modeled T_{max}, T_{min} and RF for the Ludhiana location biases in time trends and statistical parameters (Jalota *et al.,* 2013). The statistical parameters i.e. μ, σ and σ^2 of T_{max} indicated that the means were comparable but standard deviation was 31% more in the modeled data. In case of T_{min}, μ and σ of modeled values were higher by 1°C (6%) and 2.3°C (25%) than that of the observed. The μ of modeled RF was 15% more than that of the observed and σ was 40% less. The annual mean wet days were 270 in the modeled and 325 in the observed.

Difference approach

Monthly and annual correction factors estimated with difference approach from the daily observed and modeled climate data are given in Table 1. However, daily are not given for sake brevity. Daily and monthly correction

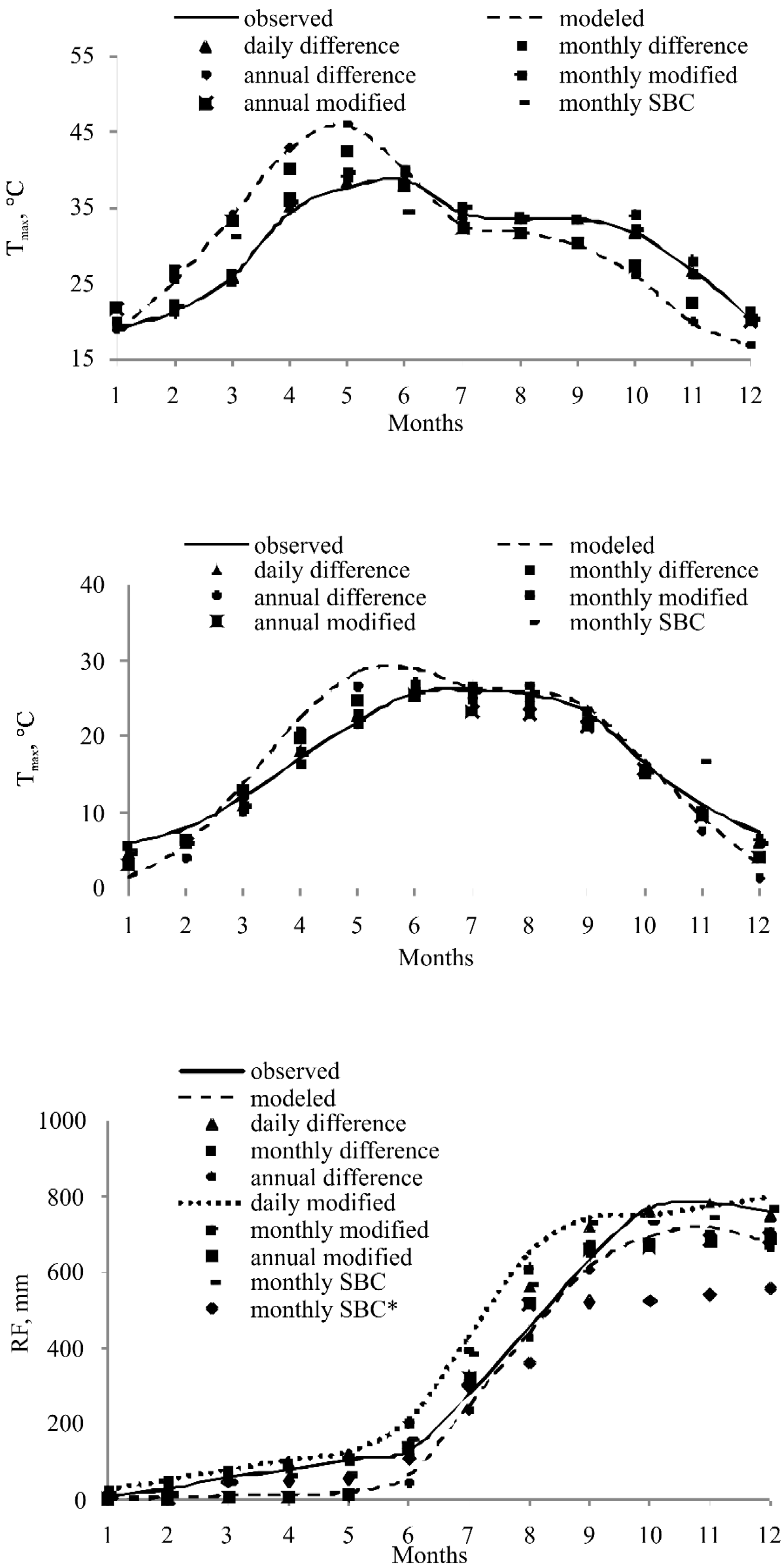

Fig. 1 : Monthly observed, modeled and model corrected T_{max}, T_{min} and RF by different statistical corrective methods

Table 1: Correction factor for difference approach (modeled minus observed) for different climate parameters at monthly and annual time scales

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
T_{max}, (°C)	1.0	5.8	7.7	6.5	4.0	-2.3	-2.2	-2.8	-4.7	-6.9	-5.7	-3.3	**3.3**
T_{min}, (°C)	-3.2	-0.2	2.9	4.3	5.8	3.2	0.7	0.7	2.0	1.7	-0.7	-3.2	**1.18**
RF, (mm/day)	-0.45	-0.54	-0.47	-.055	0.05	2.14	1.22	0.67	3.06	1.80	0.57	-0.10	**0.39**

Table 2 : Correction functions (modified difference) for correcting modeled daily temperature and rainfall data at monthly and annual time scale

Month	T_{min} (°C)	T_{max} (°C)	RF(mm)
January	T_{cor} = 8.23+0.74*(T_{mod}-5.03)	T_{cor} = 19.75+0.63*(T_{mod}-18.54)	RF_{cor} = (RF_{mod}-0.45)*1.09
February	T_{cor} = 6.76+0.73*(T_{mod}-6.55)	T_{cor} = 17.85+0.72*(T_{mod}-20.74)	RF_{cor} = (RF_{mod}-0.54)*1.09
March	T_{cor} = 7.90+0.74*(T_{mod}-10.84)	T_{cor} = 18.75+0.85*(T_{mod}-26.18)	RF_{cor} = (RF_{mod}-0.47)*1.09
April	T_{cor} = 12.35+0.74*(T_{mod}-16.69)	T_{cor} = 27.06+1.10*(T_{mod}-34.21)	RF_{cor} = (RF_{mod}-0.55)*1.09
May	T_{cor} = 15.88+0.88*(T_{mod}-21.7)	T_{cor} = 32.62+0.92*(T_{mod}-38.80)	RF_{cor} = (RF_{mod}+0.05)*1.09
June	T_{cor} = 21.89+1.01*(T_{mod}-25.1)	T_{cor} = 38.68+0.66*(T_{mod}-38.60)	RF_{cor} = (RF_{mod}+2.14)*1.09
July	T_{cor} = 24.60+1.87*(T_{mod}-25.3)	T_{cor} = 36.77+1.25*(T_{mod}-34.30)	RF_{cor} = (RF_{mod}+1.22)*1.09
August	T_{cor} = 24.04+2.15*(T_{mod}-24.8)	T_{cor} = 35.58+1.28*(T_{mod}-33.4)	RF_{cor} = (RF_{mod}+0.67)*1.09
September	T_{cor} = 19.69+1.53*(T_{mod}-21.7)	T_{cor} = 37.21+1.05*(T_{mod}-33.7)	RF_{cor} = (RF_{mod}+3.06)*1.09
October	T_{cor} = 13.86+0.84*(T_{mod}-15.6)	T_{cor} = 37.86+0.65*(T_{mod}-32.00)	RF_{cor} = (RF_{mod}+1.80)*1.09
November	T_{cor} = 10.34+0.83*(T_{mod}-9.7)	T_{cor} = 32.89+0.75*(T_{mod}-26.3)	RF_{cor} = (RF_{mod}+0.57)*1.09
December	T_{cor} = 9.02+0.86*(T_{mod}-5.8)	T_{cor} = 24.94+0.94*(T_{mod}-20.6)	RF_{cor} = (RF_{mod}-0.10)*1.09
Annual	T_{cor} = 14.59+ 0.81*(T_{mod}-15.77)	T_{cor} = 30.04 + 0.77*(T_{mod}- 29.8)	RF_{cor} = (RF_{mod}+0.39)*1.09

factors yielded corrected modeled values closer to the observed in terms of time trend and magnitude (Fig. 1). However, yearly correction factor yielded magnitude and time trends similar to that of the modeled and not of the observed. The statistical parameters of observed, modeled and corrected modeled T_{max} and T_{min}, indicated that σ and σ^2 were nearer to the observed with daily correction factor compared to that with correction factors at monthly and annual time scales. However, μ values remained higher because of originally higher modeled values than that of the observed. Similarly in case of RF, time trend of model corrected RF matched that of the observed using the daily correction factor, but total RF was 68 mm less than that of the observed (800 mm). Similar trends were observed with monthly correction factor but total precipitation was less by 153 mm than that of the observed. With annual correction factor, neither time trend nor total precipitation (633 mm) matched the observed one. The σ and σ^2 were least at monthly time scale. However, μ was 23 per cent less.

Modified difference :

Correction functions based on modified difference approach were explicitly developed for each of the calendar month and annually (Table 2). The use of these correction functions has matched the time trends and magnitude of the modeled corrected and observed temperatures (Fig. 1). In case of RF, time trends were similar to that with simple difference approach. The differences in μ, σ and σ^2 values were less in corrected modeled and observed T_{max} and T_{min} at monthly time scale compared to that of modeled and observed. However, variation in cumulative model corrected to that of observed rainfall was decreased to 2, 95 and 110 mm at daily, monthly and annual time scale, respectively.

Statistical bias correction :

The best fitted cumulative distribution function (CDF)

Table 3 : Monthly best fitted cumulative distribution functions and their scale parameters for modeled and observed data of rainfall

	Distribution	Modeled data						Observed data					
		α	β	γ	σ	μ	k	α	β	γ	σ	μ	k
Jan	Log normal			0	1.17	0.36				0	1.26	1.30	
Feb	Log pearson3	15.75	0.28	-4.13				37.68	-0.23	10.0			
Mar	Log normal			0.15	1.65	-0.70				0	1.23	1.09	
Apr	Log normal			0.29	1.27	-0.86				0.08	1.28	1.21	
May	Log normal			0	1.58	1.45				0	0.70	1.4	
Jun	Weibull	0.84	4.20	8.57				0.93	12.23	0			
Jul	Log pearson3	1.62	0.36	1.93				51.54	-0.20	12.4			
Aug	Log pearson3	1.86	0.27	1.82				85.80	-0.17	16.58			
Sep	Log normal			6.67	1.64	0.95				0.12	1.62	1.30	
Oct	Gen Ext Value				23.87	23.04	0.30				2.32	2.05	0.55
Nov	Gen Ext Value				0.98	2.59	0.87				2.68	2.23	0.58
Dec	Log normal			1.17	1.82	-0.39				0.16	1.39	1.36	

Table 4 : Statistical parameters of observed, modeled and model corrected T_{max}, T_{min} and RF by difference, modified difference, and statistical bias correction (SBC) methods

Parameter	Observed	Modeled	Difference			Modified difference			SBC
			Daily	Monthly	Yearly	Daily	Monthly	Yearly	Monthly
T_{max}									
Variance	45.6	87.4	46.8	50.1	87.4	—	49.6	51.8	44.6
NRMSE (%)	——	19	5	8	19	—	7	15	9
T_{min}									
Variance	58.0	101.4	66.5	67.3	101.4	—	68.3	66.5	63.9
NRMSE (%)	——	23	13	13	24	—	12	17	15
RF									
Variance	36.0	10.6	12.9	8.8	10.1	15.3	10.4	12.0	67.4
NRMSE (%)	——	16	12	23	28	11	15	20	40

along with their scale parameters to multi-year (1971-1985) data of observed and modeled RF at monthly time scale are given in Table 3. Similar functions were derived for T_{max} and T_{min}.

These functions transformed the modeled temperature data, close to the observed data in terms of time trends and magnitude. However, there was some deviation especially in T_{max} (Fig. 1). In case of RF, model corrected values matched closely to the observed if correction factor is applied to the individual day in multi- years and then averaged monthly, (Fig. 1). These results suggest that while correcting and projecting the multi-years modeled data using SBC correction factor for RF data, application of correction factor first and then averaging the modeled corrected data should be followed. However, cumulative precipitation from the annual developed correction function was under- estimated.

Best estimate

The coefficient of variation (%) as normalized root mean squared (NRMSE) for T_{max}, T_{min} and RF by different corrective statistical and other methods at different time scales are presented in Table 4. It is clear from this table that minimum coefficient of variation was observed with daily correction function of difference approach in both T_{max} and T_{min}. With monthly correction function of SBC involving difference method NRMSE was reduced by 13 per cent in T_{max} and 8 per cent in T_{min}. The NRMSE for modeled T_{max} was 19 per cent, which was reduced to minimum 5 per cent by difference method at daily time scale, 7 per cent by modified difference approach at monthly scale and 6 per cent by SBC method at monthly scale (Table 4). The corresponding value for modeled T_{min} was 23 per cent, which was reduced to 13 per cent by difference method at daily and monthly time scale, 12 per cent by modified difference approach at monthly scale and 15 per cent by SBC methods at monthly time scale. The NRMSE for the modeled cumulative rainfall was 16 per cent. It was reduced to 12 and 11 per cent by difference and modified difference approach at daily time scale. At monthly time scale, RMSE in modified difference approach was almost comparable to that of the modeled rainfall but was more (40%) in SBC. It may be ascribed to poor matching of corrected and observed rainfall magnitudes in the months of August, September and October. In these months of major portion of annual rainfall occur as heavy showers, which may not be in the domain of modeled data. Such lack of ability to correct temporal errors of major circulation system e.g. onset of monsoon is a limitation of this method and has also been pointed out by Piani *et al* (2010b).

CONCLUSIONS

This study indicates that biases exist in the data of T_{max}, T_{min} and RF derived from the PRECIS regional climate model for Ludhiana district. If correction is done, it adds significantly to minimize uncertainties in modeling climate change impacts. The best correction functions were obtained using modified difference approach at monthly time scale for T_{max} and T_{min}; and at daily time scale for rainfall. The effectiveness of a particular method was also found to be varied with the distribution of the data, time scale fluctuations.

ACKNOWLEDGMENT

The authors are thankful to Indian Council of Agricultural Research, New Delhi for funding this research under All India Co-ordinated Research Project on Groundwater Utilization and Emeritus Scientist Scheme. The authors also grateful to the Department for Environmental, Food and Rural Affairs (DEFRA), Government of United Kingdom, for sponsoring the joint Indo-UK program on Climate Change and the Ministry of Environment and Forest (MoEF), Government of India for coordinating its implementation. The thanks are due to the Hadley Centre for Climate Prediction and Research, UK Meteorological office for making available regional model-PRECIS. Support of the PRECIS simulation dataset is provided by Indian Institute of Tropical Meteorology, Pune.

REFERENCES

Andréasson, J., Bergström, S., Carlsson, B., Graham, L.P. and Lindström G. (2004). Hydrological change – Climate change impact simulation for Sweden. *Ambio*, 33(4–5):228–234.

Bates, B.C., Kundzewicz, Z.W., Wu, S. and Palutikof, J.P. (2008). Climate Change and Water. Technical paper. Geneva: Intergovernmental Panel on Climate Change, 210 pp.

Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Pettersson, A. and Rummukainen, M. (2001). Climate change impacts on runoff in Sweden – assessments by global climate models, dynamical downscaling and hydrological modeling. *Clim. Res*., 16:101–112.

Boberg, F., Berg, P., Thejll, P. and Christensen, J.H. (2007). Analysis of temporal changes in precipitation intensities using PRUDENCE data. Danish Climate Centre Report 07-03, 2007, Copenhagen.

Déqué, M. (2007). Frequency of precipitation and temperature extremes over France in an anthropogenic scenario; model results and statistical correction according to observed values. *Glob. Planet. Change,* 57:16-26.

Feddersen, H. and Andersen, U. (2005). A method for statistical downscaling of seasonal ensemble predictions. *Tellus*, 57A:398-408.

Gordon, C.C., Cooper, C., Senior, C.A., Banks, H., Gregory, J.M., Mitchell, J.F.B. and Wood, R.A. (2000) The simulation of SST, sea ice extents and ocean heat transport in a version of the Hadley centre coupled model without flux adjustment. *Clim. Dynam*., 16:147–168.

Haerter, J.O., Hagemann, S., Moseley, C. and Piani, C. (2010). Climate model bias correction and the role of timescales.

Hydrol. Earth Syst. Sci. Discuss., 7:7863-7898.

Hansen, J., Challinor, A., Ines, A., Wheeler, T. and Moronet, V. (2006). Translating forecasts into agricultural terms: Advances and Challenges. *Clim. Res*., 33:27–41.

Hashino, T., Bradlley, A.A. and Schwartz, S.S. (2007). Evaluation of bias correction methods ensemble stream flow water forecast. *Hydrol. Earth Syst. Sci. Discuss*., 11:939-950.

Jalota, S. K., Kaur, H., Kaur, S., and Vashisht, B. B. (2013). Impact of climate change scenario on yield, water and nitrogen –balance and use efficiency of rice-wheat cropping system. *Agric. Water Manage*., 116: 29-38.

Jones, R.G., Noguer, M., Hassell, D.C., Hudson, D., Wilson, S.S., Jenkins, G.J. and Mitchell, J.F.B. (2004). Generating High Resolution Climate Change Scenarios Using PRECIS. Met Office Hadley Centre: Exeter., UK, p. 40.

Leander, R. and Buishand, T. (2007). Resampling of regional climate model output for the simulation of extreme river flows. *J. Hydrol*., 332:487–496.

Parry, M., Canziani, O., Palutikof, J., Van der Linden, P. and Hanson, C. (2007). Climate Change: Impacts, Adaptation and Vulnerability. 1st edition, Cambridge University Press.

Piani, C., Haerter, J.O. and Coppola, E. (2010a). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol*., 99:187–192.

Piani, C., Weedon, G.P., Best, M., Gomes, S.M., Viterbo, P., Hagemann, S., Haerter, J.O. (2010b). Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol*., 395:199-215.

Sennikovs, J. and Bethers, U. (2009). Statistical downscaling method of regional climate model results for hydrological modelling. In: 18th World IMACS D MODSIM Congress, Cairns, Australia, pp 3962-3968.

Sharma, D., Das Gupta, A., Babel, M.S. (2007). Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand. Hydrol. *Earth Syst. Sci. Discuss*., 11:1373–1390.

Teuschbein, C. and Seibert, J. (2010). Regional climate models for hydrological impact studies at catchment scale: A review of recent modeling strategies. *Geography Compass*, 4/7:834-860.

Xu, C. Y., Widen, E. and Halldin, S. (2005). Modeling hydrological consequences of climate change – Progress and challenges. *Adv. Atmos. Sci*., 22 (6):789-797.

Received : April 2014 ; Accepted : December 2014