Leadership in Sustainable Infrastructure

Leadership en Infrastructures Durables

Vancouver, Canada

May 31 – June 3, 2017/ *Mai 31 – Juin 3, 2017*

A STOCHASTIC APPROACH TO DOWNSCALING OF MULTISITE DAILY TEMPERATURE SERIES IN THE CONTEXT OF CLIMATE CHANGE

Rahman, Mahzabeen[1,2], Nguyen, Van-Thanh-Van[1]

[1] McGill University, Canada

[2] mahzabeen.rahman@mail.mcgill.ca

Abstract: Downscaling techniques have been commonly used in many climate change impact assessment studies. In particular, statistical downscaling (SD) methods are widely employed because of their simplicity of implementation. However, most SD methods were developed for a single site without considering the observed spatial dependence of the hydrologic processes at different locations; this could significantly affect the accuracy of impact study results at the catchment scale. Therefore, in the present study an improved multivariate multisite SD approach was proposed for downscaling daily temperature series at many sites concurrently. This approach is based on multiple regression models for describing the linkages between large scale atmospheric predictors and local scale daily maximum and minimum temperatures, and on the use of the Singular Value Decomposition technique and multivariate autoregressive model for capturing the observed statistical properties of the stochastic components of these regression models. The feasibility of the proposed approach has been assessed using the available NCEP/NCAR re-analysis data and the daily extreme temperature data available in two regions with different climatic conditions: the southern Quebec-Ontario region in Canada and the Bangladesh region. Results of this illustrative application have indicated the ability of the proposed multivariate multisite SD to reproduce accurately the observed statistical properties of the daily extreme temperature series, and the spatial-temporal dependence of the underlying temperature processes at different locations.

1 INTRODUCTION

Climate change has been recognized as having a profound impact on the hydrologic cycle and therefore, many studies have been carried out to investigate this impact using outputs from Global Climate Models (GCMs). However, the GCM outputs are inadequate to represent local or regional scale features and variability due to their coarse spatial resolution (Nguyen et al. 2006). Hence, downscaling techniques have been proposed for assessing the climate change effects on hydrologic processes at the regional or local scale (Nguyen and Nguyen, 2008). Among two fundamental categories of downscaling techniques (dynamic and statistical), the statistical downscaling (SD) methods are widely used because of their simplicity of implementation and use (Wilby et al. 2002). However, most existing SD methods were developed for a single site without considering the observed spatial dependence of the hydrologic processes at different locations, which could significantly affect the accuracy of impact study results at the catchment scale. For example, Zareie et al. (2015) showed that the timing of streamflow during peak flow season was simulated more accurately by the use multisite downscaling as compared to the single site downscaling approach.

In view of the above issues, multi-site downscaling approaches have been a subject of intense interest in many recent studies (see, e.g., Khalili and Nguyen, 2016). In particular, for multisite downscaling of temperature series at different locations different methods have been proposed by Khalili et al. (2011), Khalili and Nguyen (2012), Rahman et al. (2013), Jeong et al. (2013), Srinivas et al. (2014). In the present study, a new multivariate multisite SD approach is proposed to provide further improvement in the accuracy of multisite downscaling of daily maximum and minimum temperature (Tmax and Tmin) series. The feasibility and accuracy of the proposed SD procedure was evaluated using daily temperature data available from two networks of weather stations located in two regions with completely different climatic conditions: cold-climate southern Quebec-Ontario region in Canada and tropical-climate region in Bangladesh.

2 METHODOLGY

The proposed multisite multivariate SD method is based on a combination of a multiple linear regression model to describe the direct linkages between global climate predictors and local Tmax and Tmin series, and a multivariate autoregressive-singular value decomposition (SVD-MAR) model to reproduce accurately the observed spatial-temporal statistical properties of the stochastic component of the regression model. Equation [1] represents the regression model for daily Tmax:

[1] $$\begin{pmatrix} Tmax_{i,m,1} \\ Tmax_{i,m,2} \\ \vdots \\ \vdots \\ Tmax_{i,m,s} \end{pmatrix} = \begin{pmatrix} \alpha_{Tmax_{0,m,1}} \\ \alpha_{Tmax_{0,m,2}} \\ \vdots \\ \vdots \\ \alpha_{Tmax_{0,m,s}} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{q} \alpha_{Tmax_{j,m,1}} \, p_{Tmax_{ij,m,1}} \\ \sum_{j=1}^{q} \alpha_{Tmax_{j,m,2}} \, p_{Tmax_{ij,m,2}} \\ \vdots \\ \vdots \\ \sum_{j=1}^{q} \alpha_{Tmax_{j,m,s}} \, p_{Tmax_{ij,m,s}} \end{pmatrix} + \begin{pmatrix} E_{Tmax_{i,m,1}} \\ E_{Tmax_{i,m,2}} \\ \vdots \\ \vdots \\ E_{Tmax_{i,m,s}} \end{pmatrix}$$

where, $Tmax_{i,m,k}$ is the maximum temperature on day i in month m and at station k from a network of s stations; $\alpha_{Tmax_{j,m,k}}$ is the j^{th} regression parameter for month m and at station k; $p_{Tmax_{ij,m,k}}$ is the value of the j^{th} predictor, from a total of q monthly significant predictors, on day i in month m and at station k; $E_{Tmax_{i,m,k}}$ is the residual value on day i in month m and at station k. In Eq. [1] the significant predictors were selected mainly based on partial correlation coefficient, although subjective selection has been made in many cases. The expression of Eq. [1] is only for *Tmax* but the similar representation is applicable for *Tmin* as well.

The residuals of Tmax and Tmin, which are the stochastic components of the regression models have been generated simultaneously by the SVD technique (Green, 1978). As the proposed methodology considers the simultaneous downscaling of Tmax and Tmin, the general matrix of the residual values E is obtained by the concatenation of the matrices E_{Tmax} and E_{Tmin} as described in Eq. [2], in which the elements of E_{Tmax} and E_{Tmin} are the residuals obtained for daily Tmax and Tmin, respectively.

[2] $$E = [E_{Tmax} \; E_{Tmin}]$$

The general idea behind using the SVD technique is that several series of the residual values need to be produced. The SVD technique is able to transform a matrix of several correlated variables into a matrix of a small number of principal components that are uncorrelated variables. These principal components are linear combinations of the standardized observed variables and are closer to the normal distribution than the observed variables. These principal components can therefore be generated independently, several times, following a standard normal distribution with zero mean and unit variance.

The SVD-MAR technique was used by Chaleeraktrakoon (1995), where a multivariate seasonal streamflow model was proposed to simulate streamflow at a single location as well as at many locations concurrently. The present study involves the combined use of SVD and MAR technique for generating simultaneously the residuals of daily Tmax and Tmin in the proposed SD model. Detailed description of SVD-MAR technique can be found in Chaleeraktrakoon (1995).

3 NUMERICAL APPLICATION

3.1 Data and Study Area

In the following, to assess the feasibility and accuracy of the proposed multivariate multisite SD procedure two sets of daily extreme temperature data representing two completely different climatic regions were selected. The first dataset is obtained from a network of seven weather stations located in the cold-climate southern Quebec-Ontario region in Canada, and the second dataset is from a network of four stations located in the tropical climate region in Bangladesh (Table1). In addition, significant climate predictors given by the NCEP reanalysis data (Kistler et al. 2001, Kalnay et al. 1996) for these two regions have been selected for this study using the same screening method proposed by Khalili et al. (2011).

Table 1: Name and coordinates of selected stations used in this study.

	Quebec-Ontario region			Bangladesh		
	Station Name	Latitude	Longitude	Station Name	Latitude	Longitude
Station 1	L'Assomption	45.81^0 N	73.43^0 W	Bogra	24.85^0 N	89.37^0 E
Station 2	Dorval	45.47^0 N	73.75^0 W	Jessore	$23.2\ ^0$ N	89.33^0 E
Station 3	Chenaux	45.58^0 N	76.68^0 W	Dhaka	$23.78\ ^0$ N	90.38^0 E
Station 4	Maniwaki UA	46.30^0 N	76^0 W	Satkhira	22.72^0 N	89.08^0 E
Station 5	Cornwall	45.02^0 N	74.75^0 W			
Station 6	Ottawa CDA	45.38^0 N	75.72^0 W			
Station 7	Morrisburg	44.92^0 N	75.19^0W			

3.2 Results

In this paper, both graphical and numerical criteria were used to demonstrate the accuracy of the proposed multisite SD method. More specifically, a total of 100 simulations of daily Tmax and Tmin were generated and the agreement between the observed and simulated temperature extremes was assessed using the box plots of means of Tmax and Tmin and based on the monthly average values of Mean Absolute Error (MAE) (Eq. [3]) and Root Mean Square Error (RMSE) (Eq. [4]) of daily Tmax and Tmin for each station. The mathematical expressions of MAE and RMSE are given as follows:

[3] $\mathrm{MAE} = \frac{1}{t}\sum_{t=1}^{t}\left|X_{Obs,i} - X_{Sim,i}\right|$

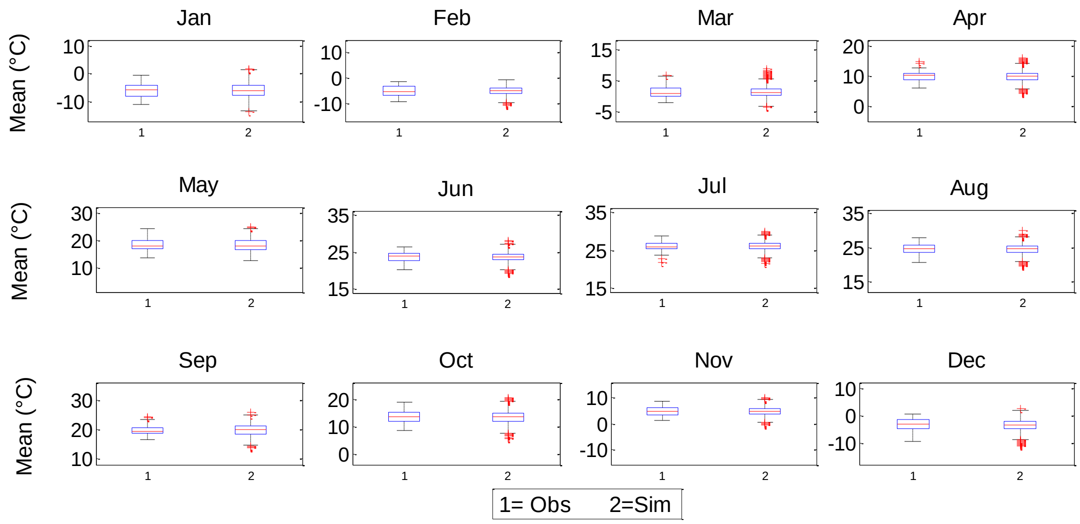
[4] $\mathrm{MSE} = \sqrt{\frac{1}{t}\sum_{i=1}^{t}\left(X_{Obs,i} - X_{Sim,i}\right)^2}$

where, t, $\overline{X}_{Obs}$, $X_{Obs,i}$ and $X_{Sim,i}$ represent the length of series, average of observed values, observed and simulated values for i^{th} observation respectively.

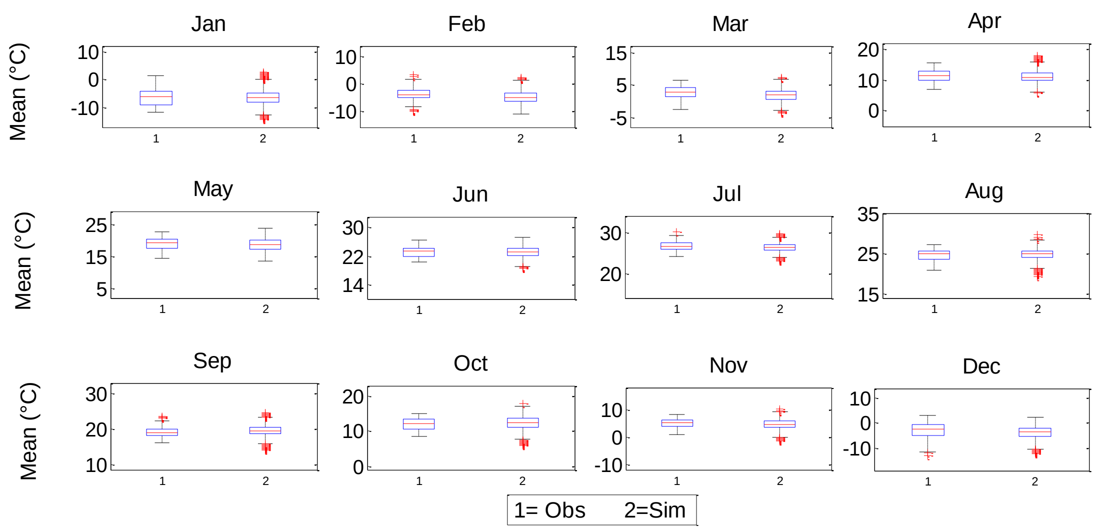
3.2.1 Results for Southern Quebec-Ontario Region

Figure 1 shows the boxplots of the mean of Tmax for each month. It can be seen that the observed means were accurately reproduced by the proposed SD procedure for almost all the months for both calibration (1961-1975) and validation (1976-1990) periods. Similar results were also found for Tmin. Table 2 presents the MAE and RMSE values for Tmax and Tmin for all stations. It is observed that very low values of MAE (RMSE) were produced ranging from 0.02~0.04 (0.03~0.06) and 0.05~0.10 (0.06~0.13) respectively for

Tmax and Tmin during calibration period. During validation period, those values were between 0.27~0.54 (0.34~0.69) and 0.29~0.59 (0.47~0.79) for Tmax and Tmin respectively.



(a)



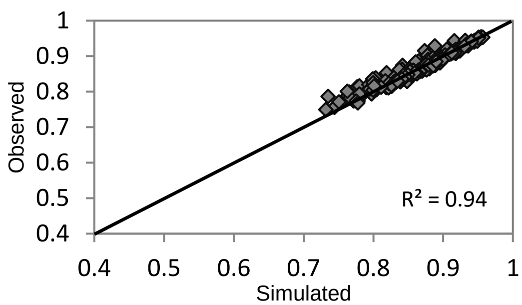
(b)

Figure 1: Boxplots of the mean of Tmax for (a) calibration and (b) validation periods for seven stations in southern Quebec-Ontario region. The band in the middle of each box denotes the median value, the boxes and whiskers denote the inter-quartile range (IQR) and 1.5×IQR respectively. Crosses beyond the whiskers are the outliers.

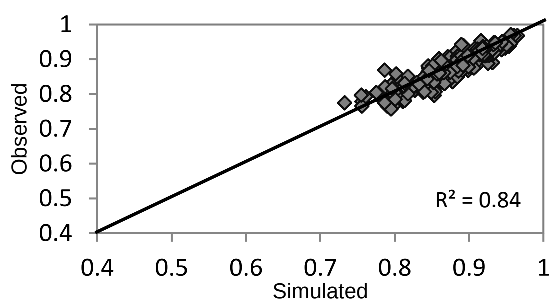
Table 2: MAE and RMSE values of Tmax and Tmin for seven stations in southern Quebec-Ontario region.

	Calibration Period				Validation Period			
	Tmax		Tmin		Tmax		Tmin	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Station 1	0.03	0.03	0.07	0.09	0.44	0.54	0.54	0.77
Station 2	0.02	0.03	0.05	0.06	0.41	0.50	0.35	0.55
Station 3	0.03	0.03	0.10	0.12	0.44	0.55	0.5	0.78
Station 4	0.02	0.03	0.05	0.06	0.27	0.34	0.43	0.73
Station 5	0.04	0.06	0.08	0.10	0.54	0.69	0.59	0.79
Station 6	0.03	0.03	0.08	0.10	0.38	0.54	0.50	0.73
Station 7	0.03	0.03	0.10	0.13	0.54	0.63	0.29	0.47

Figure 2 shows the accuracy of the proposed SD method in the description of the observed lag-0 inter-station correlation for Tmin as indicated by the high value of R^2 for calibration period ($R^2 = 0.94$) and for validation period ($R^2 = 0.84$). Similar results were found for Tmax as well (not presented here).



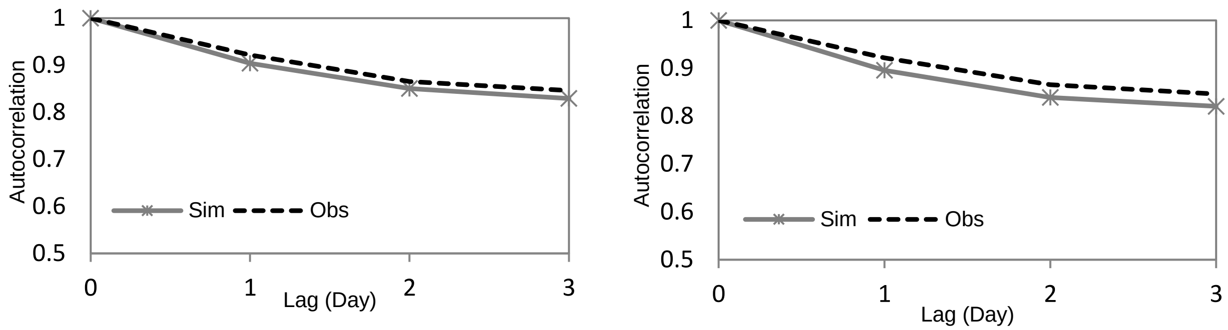
(a)



(b)

Figure 2: Lag-0 interstation correlations of daily Tmin during (a) calibration and (b) validation periods for the stations in southern Quebec-Ontario region.

The proposed SD procedure can also reproduce very well the observed autocorrelations for all stations for both Tmax and Tmin. For purposes of illustration, Figure 3 shows the very good agreement between observed and simulated values of the autocorrelations for lag one to lag three for Cornwall station for Tmax. This very good fit for Tmax and Tmin and for all stations has indicated the capability of the proposed SD method to reproduce accurately the temporal dependence of extreme temperature series at a given location of interest.

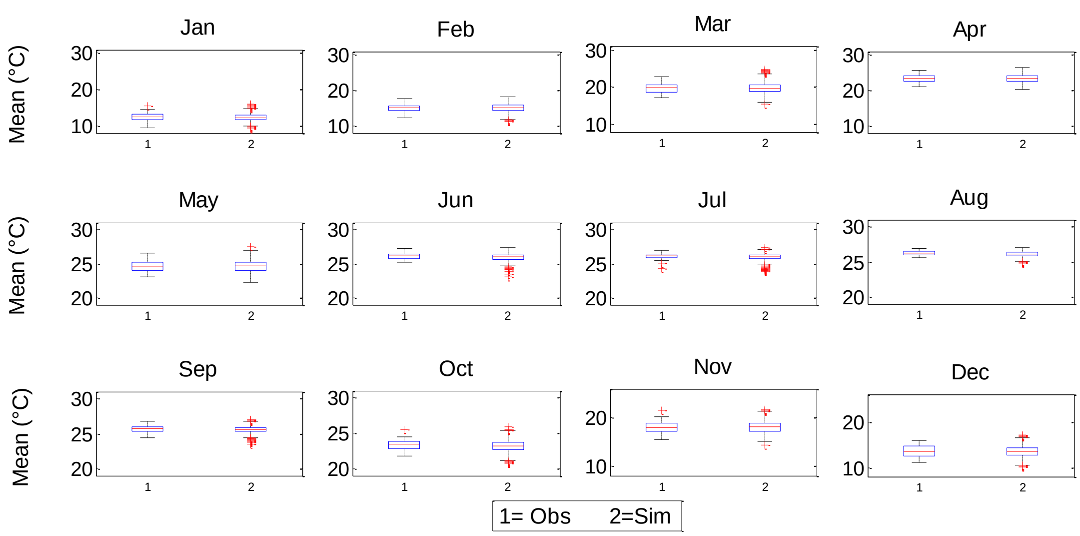


(a) (b)

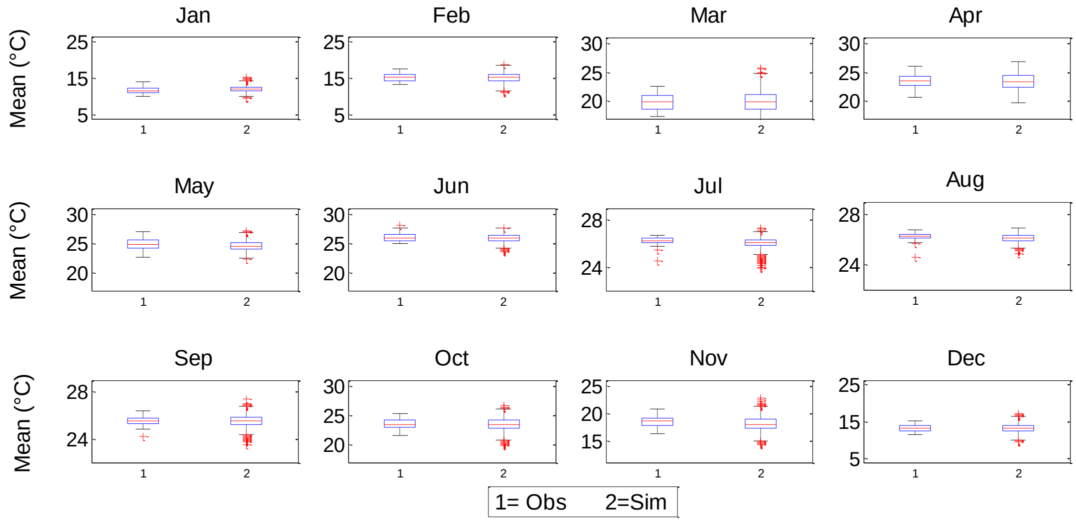
Figure 3: Observed and simulated autocorrelation coefficients of Tmax at Cornwall station during (a) calibration and (b) validation periods.

3.2.2 Results for Bangladesh

Figure 4 shows the very good agreement between observed and simulated values of monthly means of daily Tmin for all four stations in Bangladesh for both calibration (1981-1990) and validation (1991-2000) periods. The values of MAE (RMSE) for these stations ranged from 0.01 (0.01~0.02) and 0.04~0.07 (0.05~0.1) for Tmax and Tmin respectively (Table 3) during calibration period. For validation period those values are within 0.27~0.38 (0.35~0.46) and 0.16~0.38 (0.19~0.44).



(a)



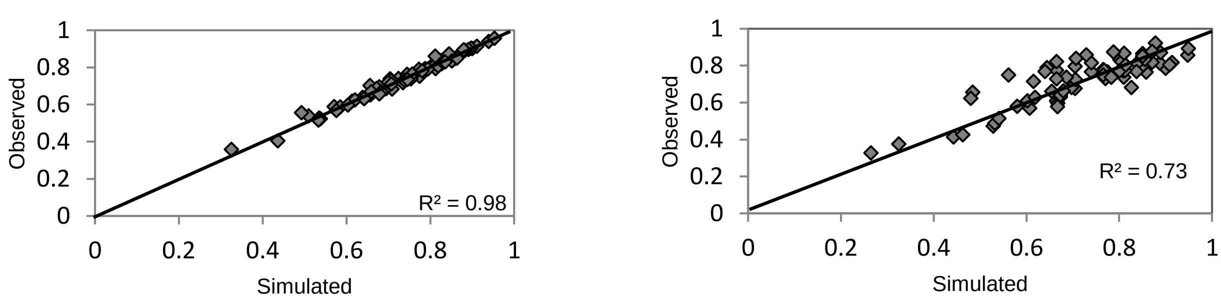
(b)

Figure 4: Same as Figure 1 but for Tmin for four stations in Bangladesh.

Table 3: Same as Table 2 but for four stations in Bangladesh.

	Calibration Period				Validation Period			
	Tmax		Tmin		Tmax		Tmin	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Station 1	0.01	0.02	0.05	0.07	0.38	0.46	0.38	0.44
Station 2	0.01	0.02	0.04	0.05	0.34	0.40	0.23	0.28
Station 3	0.01	0.02	0.07	0.10	0.27	0.35	0.16	0.19
Station 4	0.01	0.01	0.04	0.06	0.28	0.36	0.20	0.28

Figure 5 shows the accuracy in the reproduction of the lag-0 inter-station correlations for Tmax for calibration period ($R^2 = 0.98$) and for validation period ($R^2 = 0.73$). Similar results were obtained for Tmin, which are not given here. In addition to these results, the temporal dependence at a given location was reproduced adequately for all four stations in Bangladesh. For illustration purposes, the autocorrelations for Tmin during calibration and validation periods at Dhaka station is presented in Figure 6.



(a) (b)

Figure 5: Same as Figure 2 but for Tmax for four stations in Bangladesh.

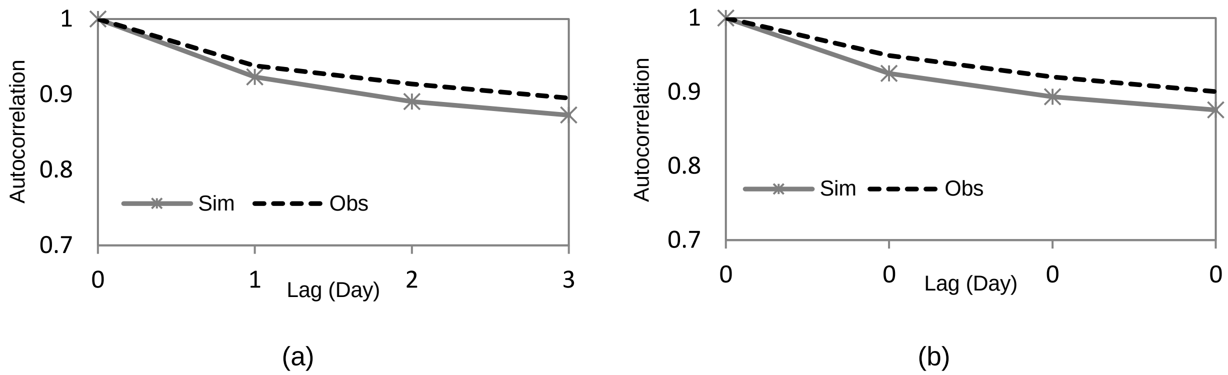


Figure 6: Same as Figure 3 but for Tmin at Dhaka station.

4 CONCLUSIONS

In this present study, an improved SD method has been proposed for downscaling daily Tmax and Tmin simultaneously at many sites concurrently. The proposed method was based on a combination of multiple linear regression models for describing the linkages between large scale atmospheric predictors and local scale daily maximum and minimum temperatures, and on the use of the SVD-MAR technique for capturing the observed statistical properties of the stochastic components of these regression models. The feasibility and accuracy of the proposed multivariate multisite SD approach have been assessed using two sets of daily temperature data available in two regions having different climatic conditions: the cold-climate southern Quebec-Ontario region in Canada and the tropical-climate region in Bangladesh. Results of these Illustrative applications have indicated the ability of the proposed SD to accurately reproduce the observed basic statistical properties of the historical temperature records, and the temporal and spatial dependences of daily temperature series at a given site as well as for different locations concurrently. The proposed multisite SD procedure could be used for improving the accuracy of climate change impact assessment studies.

Acknowledgements

The authors like to acknowledge the Data Access Integration (DAI) Team for providing the data. The DAI Portal (http://loki.qc.ec.gc.ca/DAI/) is made possible through collaboration among the Global Environmental and Climate Change Centre (GEC3), the Adaptation and Impacts Research Division (AIRD) of Environment Canada, and the Drought Research Initiative (DRI). The authors appreciate Institute of Water and Flood Management (IWFM) and Bangladesh Meteorological Department (BMD) for necessary information and data.

References

Chaleeraktrakoon, C. 1995. Stochastic Modelling and Simulation of Streamflow Processes, PhD. thesis, McGill University, QC,Canada.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C. 2007. Linking Climate Change Modelling to Impacts Studies: Recent Advances in Downscaling Techniques for Hydrological Modelling. *International Journal of Climatology*, **27**(12): 1547-1578.

Green, P. E. (1978), Analyzing multivariate data, The Dryden Press, Hinsdale, Illinois, USA.

Islam, A.S. 2009. Analyzing Changes of Temperature Over Bangladesh Due to Global Warming Using Historic Data. *Young scientists of Asia conclave*, *Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR)*,15-17.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G. and Woollen, J. 1996. The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, **77**: 437-471.

Khalili, M, and Nguyen, V.T.V. 2012. A Statistical Approach to Multi-Site Multivariate Downscaling of Daily Extreme Temperature Series for Climate-Related Impact Assessment Studies, *10th International Conference on Hydroinformatics - HIC 2012: Understanding Changing Climate and Environment and Finding Solutions*, Hamburg, Germany.

Khalili, M., Nguyen, V.T.V. and Gachon, P. 2013. A Statistical Approach to Multi-Site Multivariate Downscaling of Daily Extreme Temperature Series. *International Journal of Climatology,* **33**(1):15-32. Article first published online on 3 Nov 2011, DOI: 10.1002/joc.3402.

Khalili, M. and Nguyen V-T-V. 2016. An Efficient Statistical Approach to Multi-Site Downscaling of Daily Precipitation Series in the Context of Climate Change, *Climate Dynamics*, DOI 10.1007/s00382-016-3443-6, 18 pages.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., Dool, Hvd., Jenne, R., Fiorino, M. 2001. The NCEP/NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bulletin of the American Meteorological Society*, **82**:247-267.

Nguyen, V-T-V., Nguyen, T-D., and Gachon, P. 2006. On the linkage of large-scale climate variability with local characteristics of daily precipitation and temperature extremes: an evaluation of statistical downscaling methods, *Advances in Geosciences, Vol. 4: Hydrological Science*, N. Park et al. (Ed.), World Scientific Publishing Company, pp. 1-9.

Nguyen, V-T-V., and Nguyen, T-D. 2008. *Statistical Downscaling of Daily Precipitation Process for Climate-Related Impact Studies*, Chapter 16 in "Hydrology and Hydraulics", V.P. Singh (Ed.), Water Resources Publications, pp. 587-604.

Rahman, M., Khalili, M. and Nguyen, V.T.V. 2013. A Multivariate Statistical Approach to Downscaling of Daily Extreme Temperature Processes for Different Locations. *3rd Climate Change Technology Conference*. Concordia University, Montreal, QC.

Srinivas, V.V., Basu, B., Kumar, D. N. and Jain, S. K. 2014. Multi-Site Downscaling of Maximum and Minimum Daily Temperature Using Support Vector Machine. *International Journal of Climatology*, **34**: 1538-1560.

Wilby, R.L., Dawson, C.W. and Barrow, E.M. 2002. SDSM – A Decision Support Tool for the Assessment of Regional Climate Change Impacts. *Environmental Modelling & Software*, **17**: 147-159.

Zareie, A, Rahman, M. and Nguyen, V.T.V. 2015. Comparison of Multisite and Single-site Temperature Downscaling Effects on Streamflow and Run-off Simulation. *22nd Canadian Hydrotechnical Conference*, *Water for Sustainable Development: Coping with Climate and Environmental Changes (L'eau pour le développement durable: adaptation aux changements du climat et de l'environnement)*, Montreal, Quebec.