TEMPORAL DOWNSCALING: A COMPARISON BETWEEN ARTIFICIAL NEURAL NETWORK AND AUTOCORRELATION TECHNIQUES: PART 2-APPLICATION

Esse comunicado faz parte da síntese do artigo "TEMPORAL DOWNSCALING: A COMPARISON BETWEEN ARTIFICIAL NEURAL NETWORK AND AUTOCORRELATION TECHNIQUES OVER THE AMAZON BASIN IN PRESENT AND FUTURE CLIMATE CHANGE SCENARIOS" em fase final de publicação (DOI: 10.1007/s00704-009-0193-y) - Theoretical and Applied Climatology

David Mendes[1], José Antonio Marengo[1]

[1] Centro de Previsão de Tempo e Estudos Climáticos - CPTEC, Instituto Nacional de Pesquisas Espaciais - INPE, Cachoeira Paulista- SP, Brasil. david.mendes@cptec.inpe.br, jose.marengo@cptec.inpe.br.

ABSTRACT: Several studies have been devoted to dynamic and statistical downscaling for both climate variability and climate change. This paper introduces an application of temporal neural networks for downscaling global climate model output and autocorrelation functions. This method is proposed for downscaling daily precipitation time series for a region in the Amazon Basin. The downscaling models were developed and validated using IPCC AR4 model output and observed daily precipitation. In this paper five AOGCMs for the twentieth century (20C3M) (1970-1999) and three SRES scenarios (A2, A1B, and B1) were used. The performance in downscaling of the temporal neural network was compared to that of an autocorrelation statistical downscaling model with emphasis on its ability to reproduce the observed climate variability and tendency for the period 1970-1999. The model test results indicate that the neural network model significantly outperforms the statistical models for the downscaling of daily precipitation variability.

1. INTRODUCTION

Numerical models (general circulation models, or GCMs) representing physical processes in the atmosphere, ocean, cryosphere and land surface, are the most advanced numerical tools currently available for weather and climate forecasts, and for simulating the response of the global climate system to increasing greenhouse gas concentrations. A complete review of GCMs used in climate variability and change can be found in Meehl et al. (2007).

GCM simulations of local climate at individual grid points are often poor, especially in areas near mountains or coastlines. The notion that the increase of anthropogenic greenhouse gases will lead to significant global climate change over the next century is the accepted consensus of the scientific community. Human activities have been pointed out to have a significant contribution to the observed warming in the last 50 years, and in the projections of climate until the end of the Century XXI (Meehl et al. 2007). Human-related activities, as compared to natural climate variability (Zhang et al. 2007) are pointed out as the main cause of the observed warming in the 20th century, and the projected warming possible in the 21st century. In this context, an assessment of possible future changes of precipitation and temperature over the continents is highly relevant, considering the possible impacts of those changes and the issues of vulnerability that lead to consideration of adaptation measures.

For applications to impact studies, such as hydrological impacts of climate change, impact models are usually required to simulate sub-grid scale phenomenon and therefore require input data (such as precipitation and temperature) on a similar sub-grid scale. The methods used to convert GCM outputs into regional high-resolution meteorological fields required for reliable hydrological modeling are usually referred to as "downscaling" techniques (e.g. Hewitson and Crane, 1992). In recent years, a number of papers within the climatological community have adopted artificial neural networks as a tool for downscaling, principally in spatial resolution, from the large-scale atmospheric circulation to local or regional climate variables (Cavazos, 1999).

There are various downscaling techniques available to convert GCM outputs into daily meteorological variables appropriate for studies of hydrological impact and climate change variability (e.g. Dibike and Coulibaly, 2006). Widmann et al., (2003) developed a method to downscale precipitation, referred to as "local rescaling".

There are several different methods that can be used to derive the relationship between local and large-scale climates. There is statistical downscaling used for spatial downscaling; but mostly multiple linear regression, principle component analysis, and artificial neural networks are used. However, the procedure selected mainly depends on the objective of the study and its applications (Solman and Nuñez, 1999). Dynamical downscaling generates regional-scale information by developing and using regional climate models (RCMs) with the coarse GCM data used as boundary conditions. The RCMs represent an effective method of adding fine-scale detail to simulated patterns of climate variability and change, as they resolve better the local land-surface properties such as orography, coasts and vegetation and the internal regional climate variability through their better resolution of atmospheric dynamics and processes (Giorgi et al., 2004; Marengo et. al., 2009b).

Artificial Neural Networks (ANNs) denote a set of connectionist models inspired by the behavior of the human brain. In particular, the Multilayer Perceptron (MLP) is the most popular ANN architecture, where neurons are grouped in layers and only forward connections exist. This provides a powerful base learner, with advantages such as nonlinear mapping and noise tolerance, increasingly used in the Data Mining (DM) and Machine Learning (ML) fields due to its good behavior in terms of predictive knowledge (e.g. Rumelhart et al., 1995). The simplest form of ANN (e.g. Multilayer Perceptron) is reported to give results similar to those from multiple regression downscaling methods.

The objective of this study is to identify temporal empirical functions, using artificial neural networks (ANNs) and autocorrelation functions (ACs) that can capture the complex relationship between selected large-scale predictors and locally-observed meteorological variables for a given temporal scale (predictands).

The ANN method uses feed-forward which has temporal processing capability without resorting to complex and costly training methods. The emphasis of the feed-forward method is to evaluate and compare the optimal method with the most commonly used regression-based downscaling method and some of the IPCC AR4 models to which the downscaling technique is applied. The AOGCMs used in this paper derive from the IPCC project (CGCM3; CSIRO-MK3.5; ECHAM5-MPI; GFDL-CM2.1; and MIROC3.2-MEDRES) simulation for the twentieth century (present day – 20C3M) and SRES scenarios (A2, B1 and A1B). All five models represent the state-of-the-art AOGCMs (e.g. Boulanger et al., 2006, 2007).

2. **DATASETS AND PREDICTOR CHOICE**

3.1 Observed and model datasets.

a. The data used in these studies were from rain gauges located within the Brazilian Amazon Basin (Figure 1), which are part of the Brazilian national hydrometeorological network. They were provided by the National Water and Electric Energy Agency of Brazil (ANEEL), whose sources include the ANEEL network. Precipitation (P) is computed from rainfall observations in the Amazon Basin and is derived for the entire basin, using the records of 33 rainfall stations.

b. The AOGCM outputs are interpolated over the 2.5° x 2.5° grid defined for the observation. The period used for present conditions (20C3M run scenario) is 1970-1999, and the future is 2070-2099 as derived from five IPCC AR4 models. The five models (Table 1) represent state-of-the-art AOGCMs. In this paper we use the 20C3M run and SRES scenarios for futures A2, B1, and A1B, described in Nakienovic et al (2000). It is important to note, however, that the 20C3M simulation is intended to represent the same historical total-forcing scenarios, including both natural variability and the effect of human emissions on climate (e.g. Marengo et al., 2009b).

Figure 1 – Study area, with stations used.

Acronym	*Model*	*Resolutions*	*Source*
CGCM3	cccma_cgcm3_1_t63	T63L31	Canadian Centre for Climate Modeling and Analysis
CSIRO	csiro_mk3_0	T63L18	Australian Commonwealth Scientific Industrial and Research Organization
ECHAM	mpi_echam5	T42L19	Max-Planck-Institut für Meteorologie
GFDL2.1	gfdl_cm2_1	M45L24	Geophysical Fluid Dynamics Laboratory
MIROC-m	miroc3_2_medres	T42L20	Centre for Climate System Research, University of Tokyo; National Institute for Environmental Studies; Frontier Research Centre for Global Change

Table 1 - climate models with daily data for precipitation available from PCMDI. Column 1 is the acronym used in the text. Column 2 is the name of the model used in the PCMDI archive, columns 3 model resolution, and column 4 is the source of the model.

3. DOWNSCALING

3.1 Validation Results

The ANN were developed using various hidden nodes and layers. The final error after a given number of training cycles was observed. The number of intermediate nodes varied from three to eight and the number of iterations varied from 500 to 1 for converging to a desired mean square error and cross-evaluation on the validation set (Figure 2).

The synthesis of precipitation was carried out using statistical procedures for the purpose of comparison to the ANN results. The autocorrelation function for precipitation for the Amazon Basin is shown in Figure 4.

The time series plots in Figure 3 shows observed precipitation by day of season (JFM, JJA) and results of simulations using ANN from the AOGCMs. The use of ANN compared with autocorrelation results in a satisfactory performance, principally in daily variation (Figure 4).

The autocorrelation and partial autocorrelation analysis suggest the modeling of precipitation using multivariate autoregressive model (AR). An AR model was applied to generate the series and an inverse path of model fitting used to obtain the original variables. The residual series were tested for independence and normality. The test for normality indicated that the residual series followed a normal distribution.

Comparison of results obtained using the ANN compared with those obtained using an alternative statistical model indicates that the network is a potential competitive alternative tool for the analysis of multivariate time series.

Table 2 show the comparison of monthly means and standard deviation of the series generated using both the ANN and statistical model from the Amazon Basin. Table 3 gives the comparison of the monthly skewness of the generated series using the ANN and AR for the Amazon Basin.

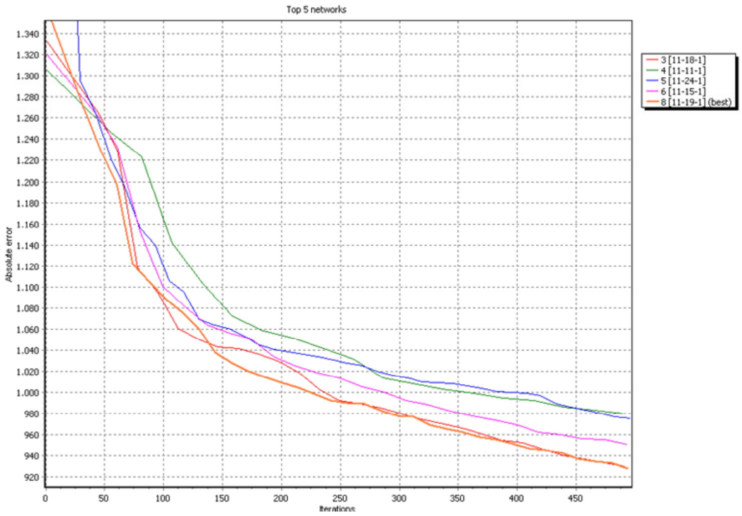

Figure 2 – The absolute error as a function of the number of iterations for various numbers of intermediate nodes.

The ANN preserved the mean skewness (skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable) of the generated series about as well as the statistical models did. Table 3 gives the comparison of the monthly values of skewness of the series generated by the use of the ANN and AR for the Amazon Basin. To conclude, each of these measures of performance in the Amazon Basin showed the overall high skill of the models (Section 3) in representing precipitation patterns and variability.

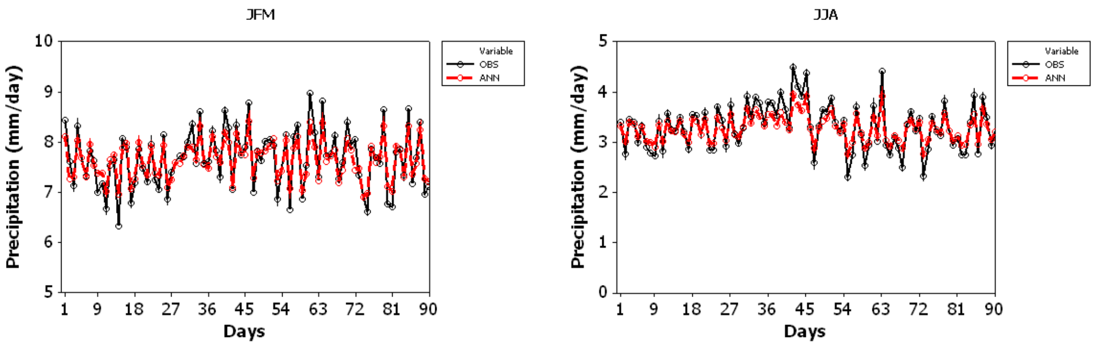

Figure 3 - Observed historical precipitation (black) by day of season (JFM, JJA), and results of simulations (line and open cicle) using Artificial Neural Network from AOGCMs, for Amazon Basin.

5.2 Downscaling scenarios

These data cover one period (2070-2099) and three scenarios (A2, A1B, B1). The ANN downscaling results in Figure 4 indicate a decrease of one-third both in the mean daily precipitation, with low difference between scenarios, principally between May and September and an increase between January and March (Figure 4; table 4).

Table 4a summarizes the downscaling results by presenting the simulated increase or decrease in monthly values of the difference 1970-1999 (present from 20C3M and observation data) and the future 2070-2099 from the 5 models, in mm day^{-1} and percentage (%), for each of the downscaling methods.

In the A2 scenario, ANN and autocorrelation method results predicted a small increase (+1,60 and 2,90 %, for ANN and AC, respectively). In the A1B scenario, ANN showed no increase/decrease, and AC gave a decrease of -2.50 %. For the B1 scenario, ANN gave a decrease (-0.90 %) and AC an increase (+0.68 %).

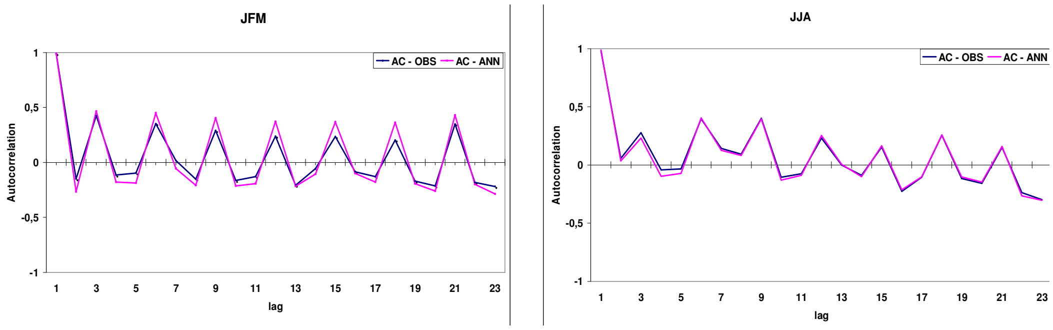

Figure 4 – Autocorrelation of historic precipitation in the Amazon Basin from observed data and ANN downscaling for the 1970-1999.

In summary, the result suggests a slight increase in the mean annual precipitation values in the study area about 1.78 % for the future years. Generally, there is good regional agreement between the signs of the precipitation changes in the AOGCM and the downscaled result in seasons (Table 4). The downscaling results from climate change scenarios (A2, A1B, and B1) presenting the increase or decrease in seasonal values of precipitation between the current (1970-1999) and (2070-2099) time period for each of the downscaling methods are as follows:

a) JFM (January-February-March) increase;

b) JJA (June-July-August) decrease.

	Mean			**Standard Deviation**		
	Observed	ANN	AR	Observed	ANN	AR
January	6,13	6,00	6,25	0,85	0,93	0,94
February	6,40	6,55	6,49	0,93	0,94	0,95
March	6,15	6,29	6,35	1,25	1,05	1,15
April	5,65	5,60	5,90	1,30	1,25	1,23
May	4,21	4,25	4,20	0,98	0,96	0,93
June	2,99	3,05	2,99	0,85	0,71	0,99
July	2,22	2,55	2,51	0,83	0,93	0,93
August	2,16	2,19	2,25	0,93	0,83	0,95
September	2,90	3,00	2,95	0,90	0,99	1,01
October	3,86	3,99	4,00	0,99	1,03	0,93
November	4,78	5,05	5,31	1,05	1,02	1,03
December	5,44	5,30	5,25	1,03	1,00	1,05

Table 2 – Comparison of mean and standard deviation of observed and generated precipitation series for Amazon Basin for1970-1999 (present conditions).

	MEAN		
	Observed	ANN	AR
January	0,05	0,03	0,03
February	0,01	0,01	0,02
March	0,03	0,05	0,06
April	0,01	0,04	0,03
May	0,05	0,09	0,05
June	0,06	0,03	0,04
July	0,03	0,05	0,03
August	0,01	0,03	0,07
September	0,02	0,02	0,08
October	0,04	0,09	0,05
November	0,01	0,04	0,03
December	0,04	0,03	0,04

Table 3 – Comparison of skewness of observed and generated precipitation series for Amazon Basin from 1970-1999 (Present day).

	Increase/decrease					
	A2		**A1B**		**B1**	
	ANN	AR	ANN	AR	ANN	AR
	*/+	*/+	*/+	*/+	*/+	*/+
JAN	0,86 / 14,0	0,76 / 12,40	0,68 / 11,1	0,87 / 9,4	0,26 / 4,2	-0,03 / -0,5
FEB	1,69 / 26,4	1,72 / 26,90	0,99 / 15,5	0,72 / 11,3	1,21 / 18,9	1,12 / 17,5
MAR	2,46 / 40,0	2,40 / 39,00	2,30 / 37,4	2,18 / 35,4	2,0 / 32,5	2,4 / 39,0
APR	0,78 / 13,8	1,24 / 21,90	0,64 / 11,3	0,91 / 16,1	0,55 / 9,7	0,60 / 10,6
MAY	-1,31 / -31,1	-1,36 / -32,3	-1,19 / -28,3	-1,09 / -25,9	-1,01 / -24,0	-0,65 / -15,4
JUN	-1,70 / -56,9	-1,43 / -47,8	-1,70 / -56,9	-1,44 / -48,2	-1,76 / -58,9	-1,78 / -59,5
JUL	-1,22 / -55,0	-1,20 / -54,1	-1,23 / -55,4	-0,97 / -43,7	-1,37 / -61,7	-1,22 / -55,0
AUG	-1,13 / -52,3	-1,17 / -54,2	-1,21 / -56,0	-1,05 / -48,6	-1,26 / -58,3	-1,17 / -54,2
SEP	-1,15 / -39,7	-1,11 / -38,3	-1,07 / -36,9	-1,15 / -39,7	-0,9 / -31,0	-1,01 / -34,8
OCT	-0,07 / -1,80	0,03 / 0,80	-0,04 / -1,0	-0,04 / -1,0	0,09 / 2,3	0,13 / 3,4
NOV	0,52 / 10,9	0,67 / 14,0	0,81 / 16,9	-1,06 / -22,2	0,77 / 16,1	1,22 / 25,5
DEC	1,09 / 20,0	1,0 / 18,40	1,05 / 19,3	0,88 / 16,2	0,95 / 17,5	0,81 / 14,9

Table 4 – Changes of monthly precipitation for the Amazon Basin interms of increase or decrease (mm day^{-1}) in comparison to the period 2070-2099 for different scenarios. Difference between 2070-2099 and 1971-1999 indicated by * and percentage difference (increase or decrease) indicated by $^{+}$.

4. CONCLUSIONS

This paper investigates the applicability of a temporal neural network as a downscaling method using an artificial neural network and an autocorrelation model for the generation of daily precipitation over the Amazon Basin (for the current years – 20C3M, and future scenarios). The ANN as well as the autocorrelation model both provided a very good fit to the data. This indicates that an ANN offers a viable alternative for multivariate modeling of precipitation time series.

The results obtained using the ANN model compared with those obtained using an alternative statistical model indicate that the network is a potentially competitive alternative tool for the analyses of multivariate time series. Comparison of the monthly values of skewness generated by the use of ANN with those generated by autocorrelation showed little difference between the two methods.

In relation to the three scenarios (A2, A1B, and B1), the ANN indicates a decrease by about a third both in the mean daily precipitation and very low difference between scenarios (May to September) and an increase between January and April (Figure 5). Performance of the ANN, principally for present-day conditions (1970-1999) for most seasons was better than that of the autocorrelation method.

However, one should also remember that all the downscaling in this study uses outputs from only one of various general circulation models. Previous studies showed that data taken from different GCMs could produce significantly different downscaling outputs.

In considering the method and results, it is important to note that our method is actually based on a hypothesis. The hypothesis is that the weight given to each of the various models when computing their differing estimates of twenty-first century climate conditions should depend on the skill of each in representing present climate conditions.

A major difficulty in using ANN for climate change lies in determining the network's capability to extrapolate. A comparison between ANN and a linear projection based on statistical downscaling allowed us to determine that the ANN penalizes climate change projections. The ratio between ANN and autocorrelation is sensitive to two factors: the bias and the divergence criteria. They represent respectively the error between the linear combination and present-day climate conditions and the variance between the models.

In conclusion, when applied to precipitation, the ANN approach makes it possible to compute the optimal set of weights for autocorrelation of the models (used in this paper), and a penalty function or probability that such a change occurred, based on the present-climate model biases and their projected dispersion.

The main advantages of this downscaling method (ANN) are its temporal processing ability and its ability to incorporate not only the concurrent, but also several preceding predictor values as input without any additional effort.

Acknowledgements

The present work was supported by the FAPESP grant 07/50145-4, and Jose A. Marengo was funded by the Brazilian Conselho Nacional de Desenvolvimento Cientifico e Tecnologico - CNPq. We are also grateful for the helpful comments made by anonymous reviewers.

5. REFERENCES

Boulanger, JP, Martinez F, Segura EC (2006) Projection of future climate change conditions using IPCC simulations, neural networks and bayesian statistics. Part 1. Temperature mean state and seasonal cycle in South America. *Climate Dynamics,* 27: 233-259.

Boulanger, JP, Martinez F, Segura EC (2007) Projection of future climate change conditions using IPCC simulations, neural networks and bayesian statistics. Part 2: Precipitation mean state and seasonal cycle in South America. Climate Dynamics, 2:, 255-271

Cavazos, T, (1999) Large-Scale Circulation Anomalies Conducive to Extreme Precipitation Events and Derivation of Daily Rainfall in Northeastern Mexico and Southeastern Texas. J. Climate, 12: 1506–1523.

Christensen, JH, Carter TR, Rummukainen M, Amanatidis G, (2007b) Evaluating the performance and utility of regional climate models: the PRUDENCE project *Climatic Change*, 81:1–6 DOI 10.1007/s10584-006-9211-6.

Dibike, YB, Coulibaly P, (2006) Temporal neural network for downscaling climate variability and extremes. Neural Network*,* 19: 135-144. DOI 10.1016/j.neunet.2006.01.003.

Gao, X., Pal JS, Giorgi F (2006) Projected changes in mean and extreme precipitation over the Mediterranean region from a high resolution double nested RCM simulation, Geophys. Res. Lett., 33, L03706, doi:10.1029/2005GL024954.

Giorgi, F, Bi X, Pal JS (2004) Mean interannual variability and trends in a regional climate change experiment over Europe. Part I: Present day climate (1961-1990). Climate Dynamics*,* 22: 733-756.

Giorgi F, Bi X, Pal JS, (2004) Mean, interannual variability and trends in a regional climate change experiment over Europe. I: present day climate (1961-1990). Climate Dynamics*.* 22: 7333-756.

Hewitson, BC, Crane RG, (1992) Large-scale atmospheric controls on local precipitation in tropical Mexico, Geophys. Res. Lett., 19: 1835–1838

Hewitson, BC, Crane RG, (1996) Climate downscaling : Techniques and application. Climate Res., 7**:** 85-95.

Hulme M, Jenkins GJ, Lu X, Turnpenny JR, Mitchell TD, Jones RG, Lowe J, Murphy JM, Hassell D, Boorman P, Macdonald R, Hill S (2002) Climate-Change Scenarios for the United Kingdom: The UKCIP02 Scientific Report. *Tyndall Centre for Climate Change Research. School of Environmental Sciences*. University of East Anglia, Norwich, UK, 120pp.

Marengo, J, Ambrizzi T (2006) Use of regional climate models in impacts assessments and adaptations studies from continental to regional and local scales: The CREAS (Regional Climate Change Scenarios for South America) initiative in South America. *Proceedings of 8 ICSHMO*, Foz do Iguaçu, Brazil, April 24-28, 2006, p. 291-296.

Marengo, JA, Jones R, Alves LM, Valverde M (2009a) Future change of temperature and precipitation extremes in South America as derived from the PRECIS regional climate modeling system. Int. J. Climatol., DOI: 10.1002/joc.1863

Marengo, JA, Ambrizzi T, Rocha RP, Alves LM, Cuadra SV, Valverde MC,. Ferraz SET, Torres RR, Santos DC (2009b) Future change of climate in South America in the late XXI Century: Intercomparison of scenarios from three regional climate models, Submitted to Climate Dynamics.

Mearns, L. (2004) NARCCAP North American Regional Climate Change Assessment Program A Multiple AOGCM and RCM Climate Scenario Project over North America. 12/17/2004. *AGU Fall Meeting*, San Francisco, USA.

Meehl, G, Covey, C, Delworth, T, Latif, M, McAvaney, B, Mitchell, J,, Stouffer FB,, R Taylor, (2007) The WCRP CMIP3 Multimodel data set. A new era in climate change research. Bull Am Met Soc., DOI:10.1175/BAMS-88-9-1383.

Rumelhart, DE, Durbin R, Golden R,, ChauvinY,(1995) Back-propagation: The basic theory. *Back Propagation: Theory, Architectures, and Applications.* Y. Chauvin and D. E. Rumelhart, Eds., Lawrence Earlbaum, 1–34.

Salathe, EP (2003) Comparison of Various Precipitation Downscaling Methods for the Simulation of Streamflow in a Rainshadow River Basin. Int. J. Climatol. 23: 887-90 1.

Solman, AS, Nuñez MN (1999) Local estimates of global climate changes: A statistical downscaling approach. Int. J. Climatol*.* 19: 835-861.

Weichert, A, Burger G, (1998) Linear versus nonlinear techniques in downscaling. Climate Research, 10: 83-93.

Widmann M.,., Bretherto CS, Salathe EP, (2003) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. J Climate, 16:799–816.

Zhang, X, Zwiers FW, Hegerl GC, Lambert FH, Gillett NP, Solomon S, Stott PA**,** Nozawa T, (2007) Detection of human influence on 20th century precipitation trends. Nature, 448**,** 461-466.