TEMPORAL DOWNSCALING: A COMPARISON BETWEEN ARTIFICIAL NEURAL NETWORK AND AUTOCORRELATION TECHNIQUES: PART 1- METHODS

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).

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).

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).

2. METHOD

a. Artificial Neural Network

An ANN is a system based on the operation of a biological neural network, in other words, it is an emulation of biological neural system.

Advantages of the artificial neural network:

1) An ANN can perform tasks that a linear program cannot;
2) When an element of the ANN fails, it can continue without any problem, due to its parallel nature;
3) An ANN learns and does not have to be reprogrammed;
4) It can be implemented in any application;

Disadvantage of an ANN:

1) Large amounts of observational data may be required to establish statistical relationships for the current climate;
2) Specialized knowledge is required to apply the techniques correctly;
3) Relationships are only valid within the range of the data used for calibration; projection for some variables may lie outside this range;
4) It might not be possible to derive significant relationships for some variables.

ANN is among the newest signal-processing technologies in the engineer's toolbox. The field is highly interdisciplinary, but our approach will be restricted to the engineering perspective. Definitions and style of computation in an ANN are of an adaptive nature and often nonlinear systems learn to perform a function from data (input/output).

An input is presented to the ANN along with a corresponding desired, or target, response set for the output (when this is the case, the training is called supervised). An error field is constructed from the difference between the desired response and the system output. The error information is used as feedback to the system and adjusts the system parameters in a systematic fashion. The process is repeated until the performance is acceptable. It is clear from this description that the performance hinges heavily on the data.

The network diagram shown (Figure 1) is a full-connected two-layer, feed-forward, perceptron ANN. Full-connected means that the output from each input and hidden neuron is distributed to all of the neurons in the following layer. Feed-forward means that the values only move from the input layer to hidden layers and, then to the output layer, with no values fed back to earlier layers.

The goal of the training process is to find the set of weight values that will cause the output from the ANN to match the actual target values as closely as possible. There are several issues involved in designing and training a multilayer perceptron network:

1) Selecting how many hidden layer to use in the network;
2) Deciding how many neurons to use on each hidden layer;
3) Finding a globally optimal solution that avoids local minima;
4) Converging to on optimal solution in a reasonable period of time;
5) Validating the neural network to test for overfitting.

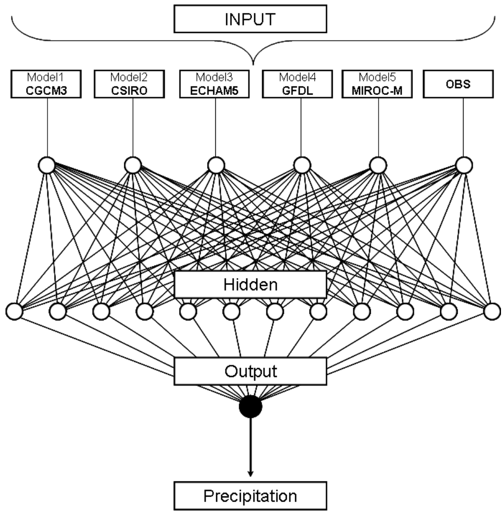

Figure 1 - Structure of the artificial neural network.

b. The statistical modeling (Autocorrelation)
Autocorrelation is the expected value of the product of a random variable or signal realization with a time-shifted version of itself obtained from a simple calculation and analysis of the autocorrelation function. We can discover a few important characteristics about our random process: These include:

1) How quickly our random signal or processes changes with respect to the time function;
2) Whether our process has a periodic component and what the expected frequency might be.

Since the autocorrelation functions are simply the expected value of a product, let us assume that we have a pair of random variables from the same process,

$X_1 = x(t_1)$ and $X_2 = x(t_2)$,

then the autocorrelation is often written as

$$R_{XXX}(t_1, t_2) = E[X_1, X_2]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) dx_2 dx_1 \quad (1)$$

The above equation is valid for stationary and non-stationary random processes. For stationary process, we can generalize this expression a little further. Given a wide-sense stationary process, it can be proven that the expected values from our random process will be independent of the origin of our time function. Therefore, we can say that our autocorrelation function will depend on the time difference and not some absolute time. For this discussion, we will $\tau = t_2 - t_1$, and thus we generalize our autocorrelation expression as

$$R_{XX}(t, t+\tau) = R_{XX}(\tau)$$

$$= [X(t), X(t+\tau)] \quad (2)$$

for the continuous-time case.
Below we will look at several properties of the autocorrelation function that hold for a stationary random process.
Autocorrelation is an even function for τ ;

$$R_{XX}(\tau) = R_{XX}(-\tau)$$

The mean-square value can be found by evaluating the autocorrelation where $\tau = 0$, which gives us;

$$R_{XX}(0) = \overline{X_2}$$

The autocorrelation function will have its largest value when $\tau = 0$. This value can appear again - for example in a periodic function at the values of the equivalent periodic points - but will never be exceeded:

$$R_{XX}(0) \geq |R_{XX}(\tau)|$$

If we take the autocorrelation of a periodic function, then $R_{XX}(\tau)$ will also be periodic with the same frequency.

3. PROCEDURE FOR TRAINING THE NETWORK

Training of the ANN is accomplished by providing inputs to the model, computing the output, and adjusting the interconnection weight until the desired output is reached. The error back-propagation algorithm is used to train the network, using the mean square error (MSE) over the training samples as the objective function. One part is used for training, the second is used for cross-validation and the third part is used for testing.
The architecture of the ANN in the present study consisted of an input layer, a hidden layer and an output layer. The number of intermediate units was obtained through a trial-and-error procedure. The error between the value predicted by the ANN and the value actually observed was then measured and propagated backwards along the feed-forward connection. The final error, after a given number of training cycles, was noted. The number of intermediate units that gave the minimum system error was accepted. During training, the performance of the ANN was also evaluated on the validation set.
The ANN and statistical procedures presented above were applied to modeling the daily precipitation data from five models (Table 1 – Part 2) derived from IPCC AR4, representing the current climate (i.e. 1970-1999), as well as daily observed precipitation measured during the concurrent period. The different parameters of each model are adjusted during calibration to get the best statistical agreement between observed and simulated meteorological variables.
The downscaling experiment was conducted with the one statistical method (autocorrelation) and the ANN methods (back-propagation) presented in section 3.The ANN training needs six predictors (five output models plus observation data) as input to the network, and the best-performing network is selected. A hyperbolic tangent activation function is used at both the hidden and output layers of the ANN and the networks are trained using a variation of feed-forward back-propagation algorithms.
A sensitivity analysis is done to determine the most relevant predictors, which need to be selected for further retraining. Sensitivity analysis provides a measure of the relative importance among the predictors (input of the ANN) by calculating how the model output varies in response to variation of an input.

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.

4. REFERENCES

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.

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.

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.

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

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.