Research Article

Downscaling Statistical Model Techniques for Climate Change Analysis Applied to the Amazon Region

David Mendes,1 José Antonio Marengo,2 Sidney Rodrigues,3 and Magaly Oliveira3

1 Climate Science Program, Federal University of Rio Grande do Norte, 59082-200 Natal, RN, Brazil
2 Instituto Nacional de Pesquisas Espaciais (INPE), Avenida dos Astronautas, 1.758 Jardim da Granja, 12227-010 São José dos Campos, SP, Brazil
3 World Wildlife Fund Brazil (WWF), SHISEQ L6/8 Conjunto, 71620-430 Brasilia, DF, Brazil

Correspondence should be addressed to David Mendes; davidmendes@ect.ufrn.br

Received 8 November 2013; Revised 25 January 2014; Accepted 27 January 2014; Published 29 May 2014

Copyright © 2014 David Mendes et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Amazon is an area covered predominantly by dense tropical rainforest with relatively small inclusions of several other types of vegetation. In the last decades, scientific research has suggested a strong link between the health of the Amazon and the integrity of the global climate: tropical forests and woodlands (e.g., savannas) exchange vast amounts of water and energy with the atmosphere and are thought to be important in controlling local and regional climates [1]. The same authors showed that the assessment of future cerrado land use scenarios is also necessary to understand the future climate and ecosystem health of the Amazon. The Amazon Biome is an area covered predominantly by dense tropical rainforest with relatively small inclusions of several other types of vegetation. It encompasses 6.7 million km² (twice the size of India) and is shared by eight countries (Brazil, Bolivia, Peru, Ecuador, Colombia, Venezuela, Guyana, and Suriname), as well as the overseas territory of French Guiana.

The Amazon Biome houses at least 11% of the world’s known biodiversity, including endemic and endangered flora and fauna, and its river accounts for 15-16% of the world’s total river discharge into the oceans. The Amazon River flows for more than 6,600 km and with its hundreds of tributaries and streams contains the largest number of freshwater fish species in the world. Equally impressive are the unfathomable numbers of mammals, birds, amphibians, and reptiles found across the biome.

1. Introduction

The Amazon is an area covered predominantly by dense tropical rainforest with relatively small inclusions of several other types of vegetation (e.g., Bromeliad, Heliconia, Orchids, water lily, and others). In the last decades, scientific research has suggested a strong link between the health of the Amazon and the integrity of the global climate: tropical forests and woodlands (e.g., Cerrado) exchange vast amounts of water and energy with the atmosphere and are thought to be important in controlling local and regional climates [1]. The same authors showed that the assessment of future cerrado land use scenarios is also necessary to understand the future climate and ecosystem health of the Amazon. The Amazon Biome is an area covered predominantly by dense tropical rainforest with relatively small inclusions of several other types of vegetation. It encompasses 6.7 million km² (twice the size of India) and is shared by eight countries (Brazil, Bolivia, Peru, Ecuador, Colombia, Venezuela, Guyana, and Suriname), as well as the overseas territory of French Guiana.

The Amazon Biome houses at least 11% of the world’s known biodiversity, including endemic and endangered flora and fauna, and its river accounts for 15-16% of the world’s total river discharge into the oceans. The Amazon River flows for more than 6,600 km and with its hundreds of tributaries and streams contains the largest number of freshwater fish species in the world. Equally impressive are the unfathomable numbers of mammals, birds, amphibians, and reptiles found across the biome.

1.1. The Importance of the Amazon Rainforest for Local and Global Climate. Tropical forests and woodlands (e.g., savannas) exchange vast amounts of water and energy with the atmosphere and are thought to be important in controlling local and regional climates.
Water released by plants into the atmosphere through evapotranspiration (evaporation and plant transpiration) and to the ocean by the rivers influences world climate and the circulation of ocean currents. This process works as a feedback mechanism, as the process also sustains the regional climate on which it depends.

Marengo et al. [2] say that Amazonian rainforest plays a crucial role in the climate system, helping to drive atmospheric circulations in the tropics by absorbing energy and recycling about half of the rainfall that falls on it. Amazonia can be categorized as a region at great risk due to climate variability and change. The risk is not only due to projected climate change but also through synergistic interactions with existing threats not related to climate change, such as land clearance, forest fragmentation, and fire [2].

Several papers on deforestation quote the following.

(1) Baidya Roy and Avisar [3] simulated the impact of deforestation on Rondônia State. They found that vegetation breezes formed and converged above deforested areas, resulting in strong updrafts carrying the moisture that had transpired from the forest and generating shallow-convection clouds.

(2) Gandu et al. [4] showed and evaluated the impact of deforestation on the climate of the eastern portion of the Amazon Basin. This region is primarily an area of native tropical rainforest but also contains several other natural 86 ecosystems such as mangroves and savanna.

(3) Avisser et al. [5] have summarized possible changes in rainfall accumulation due to a progressive expansion of deforestation in three scenarios.

(a) Rainfall will first decrease rapidly then more slowly as the deforested area expands.

(b) Given the abovementioned finding, it could be anticipated that rainfall will first remain unaffected (or possibly even increase) because of initial deforestation, but it will then fall rapidly as the deforested area crosses some threshold estimated at 30–50% of deforestation.

(c) Rainfall will decrease linearly with deforested area.

Oyama and Nobre [6] findings performed with General Circulation Model (GCM) suggest that deforestation may establish a permanent savanna in Amazonia, mostly over the east, but Ramos da Silva et al. [7] the rainfall toward the west and with stronger impacts during extreme events (e.g., El Niño). Thus, the combination of the large-scale effects on the drying of the eastern part of the basin and local mesoscale systems cause a much stronger effect over region.

Another important factor is increase of temperature, wind speed, and latent heat flux, causing a decrease because of expanding deforestation in the Amazon forest.

1.2. Extreme Tendency and the Precipitation over Amazon Basin. The climate change over Amazon Basin detected an observed warming of +0.57°C/100 year until 1997 [8]. Marengo et al. [9, 10] cited magnitude and size of the records and uncertainty is high since studies have found trends that vary in direction when different length periods are used.

An analysis of extreme rainfall over Amazon Basin has been hampered by the degree of spatial coherence and uncertainty due to reduced number of station distribution in region [9–11]. Several studies in the Amazon Basin, based on time series of rainfall data (e.g., [12, 13]), have produced conflicting results, principally in relation to trend rainfall; Marengo [14] identifies a weak positive rainfall trend in the whole Amazon Basin, with negative rainfall over northern and positive rainfall over southern section of the basin. Valverde and Marengo [15] show the annual variability of extreme precipitation indices, such as, maximum lengths of dry spell which were sensitive to identification periods of droughts. Similarly, it verified that extreme rains (>50 mm/day) occur with or without events El-Niño. In 2005, severe drought in the past 40 years and one of the most intense of the last 100 years affected Amazon Basin. Marengo et al. [9, 10] suggest that drought was not caused by El Niño, but by (1) the anomalously warm tropical North Atlantic, (2) the reduced intensity in northeast trade wind moisture transport into southern Amazonia during the peak summertime season, and (3) the weakened upward motion over this section of Amazonia, resulting in reduced convective development and rainfall. Marengo et al. [9, 10] showed that drought of 2005 affected western Amazon Basin, unlike El Niño-induced drought, which affects central and eastern Amazon Basin. In October 2006, rainfall began again and alleviated the situation, and, by February 2006, rainfall was above the normal. This rapid change in rainfall patterns directly linked to the change in the pattern of SST in the North Atlantic.

Cox et al. [16] examine the possible links of the drought of Amazon Basin in 2005 to climate change, verifying that the reduction of dry season (July to October) rainfall in western Amazon Basin is well-correlated with an index of north-south SST gradient across the equatorial Atlantic related to the Atlantic Multidecadal Oscillation (AMO). More recent study linking statistical downscaling were used (e.g., [17]) to identify extremes of Amazon climate. Mendes and Marengo [17] used neural network for downscaling with IPCC-AR4 models, the authors find that model test results indicate that the neural network model significantly outperforms the statistical models for the downscaling of daily precipitation variability; based on experience of Mendes and Marengo [17], this paper uses neural network model to correct climate extremes over Amazon Basin. Nobre et al. [18] showed and indicated that both this fraction of the deforested area and the spatial continuity of the vegetated area might be important for modulation of global climate change, principally significant remote atmospheric responses to Amazon deforestation scenarios. Cox et al. [16] showed the climate change simulation over the Amazon Basin and identified a link between the Atlantic Ocean that exerts an influence on the climate of Amazonia, which presents strong tendency for the sea surface temperature condition associated with drought to become much more common owing to continuing
reduction in reflective aerosol pollution in the Northern Hemisphere.

1.3. Downscaling and Artificial Neural Network for Amazon Basin. There are few studies that make the ANN compared to the Amazon Basin that cited; Mendes and Marengo [17], showed that 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) and future, 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 test results indicate that the neural network model significantly outperforms the statistical models for the downscaling of daily precipitation variability.

Espinoza et al. [19] describes the main circulation patterns (CP) in the Amazonian Basin over the 1975–2002 period and their relationship with rainfall variability; using an approach combining artificial neural network (Self-Organizing Maps) and Hierarchical Ascendant Classification. These results demonstrate the potential of using ANN to better understand local rainfall variability in the Amazon Basin. Particularly, this work supplies important information about large-scale atmospheric circulation and extreme rainfall relationship in the Amazon.

Other studies relate future climate in South America; for example, Boulanger et al. [20] developed a statistical method based on neural networks and Bayesian statistics to evaluate the models’ skills in simulating late twentieth century temperature over continental areas. As the study demonstrates, the use of neural networks, optimized by Bayesian statistics, leads to two major results. First, the model weight indices can be interpreted as optimal weights for a linear combination of the climate models. Second, the comparison between the neural network projection of twenty-first century conditions and a linear combination of such conditions allows the identification of the regions, which will most probably change, according to model biases and model ensemble variance.

As there are few studies that use ANN to identify climatic extremes over the Amazon Basin, we believe that this study presented here can be an initial way related to extreme rain conditions and other variables using an artificial technique.

2. Neural Network Model (ANN)

An ANN is a system based on the operation of biological neural network and in other works it is an emulation of biological neural system. Advantages of ANN are the following.

(a) An ANN can perform tasks that a linear program cannot.

(b) When an element of the ANN fails, it can continue without any problem by their parallel nature.

(c) An ANN learns and does not require reprogramming.

(d) It can be implemented without any problem.

Disadvantages of ANN are the following.

(a) It is necessary for the ANN training to operate.

(b) The architecture of ANN is different from the architecture of microprocessors therefore it needs to be emulated.

(c) It requires high processing time for large ANN.

ANN is among the newest signal-processing technologies in the engineer's toolbox. The field is highly interdisciplinary, but our approach will restrict the view to the engineering perspective. Definitions and style of computation an ANN is on adaptive, most after nonlinear system that learns to perform a function from data (input and output). An input presented to the ANN and a corresponding desired or target response set at the output (in this case the training is called supervised).

The ability to identify a relationship from given patterns makes it possible for ANNs to solve large-scale complex problems such as pattern recognition, nonlinear modelling, classification, association, and control.

Before its application to any problem, the network first trained whereby the target output at each output neuron is compared with the network output and the difference or error is minimized by adjusting the weights and biases through some training algorithm. Training in ANNs consists of three elements: (1) weights between neurons that define the relative importance of the inputs, (2) a transfer function that controls the generation of the output from a neuron, and (3) learning laws that describe how the adjustments of the weights are made during training. During training, a neuron receives inputs from a previous layer, weights each input with a prearranged value, and combines these weighted inputs.

In this study sigmoid function was employed as an activation function in the training of the network. The sigmoid function is a bounded, monotonic, nondecreasing function that provides a graded, nonlinear response.

The ANN diagram (Figure 1) is a full-connected two layers, feed-forward, perception ANN. “Fully 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 input to hidden to output layer; no values are feedback to earlier layer.

The learning of ANNs was accomplished by a back-propagation algorithm. Back-propagation is the most commonly used supervised training algorithm in the multilayered feed-forward networks. In back-propagation networks, information is processed in the forward direction from the input layer to the hidden layer and then to the output layer (Figure 1).

We use in this work the same argument as Mendes and Marengo [17], where 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. deciding how many neurons to use on each hidden layer,
2. selecting how many hidden layers to use in the network,
3. finding a globally optimal solution that avoids local minima,
4. validating the neural network to test for overfitting,
5. converging to an optimal solution in a reasonable period.

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 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 backward along the feed-forward connection.

Another important factor in training was a sensitivity analysis 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.

3. Observed and Model Datasets

The data used in this paper were from rain gauges located within the Amazon Basin, 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. Stations were also used in other countries of South America, Instituto de Hidrología, Meteorología y Estudios ambientales (IDEAM) da Colombia; Instituto Nacional de Meteorología y Hidrología do Equador (INAMHI); Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI); Servicio Nacional de Meteorología e Hidrología de Bolivia (SENAMHI); Instituto Nacional de Meteorología y Hidrología de Venezuela (INAMEH).

Vincent et al. [21] cited that data quality assessment is an important requirement before the calculation of indices since
any erroneous outlier can have a serious impact on the trends in extremes.

The coupled atmosphere-ocean GCMs (AOGCMs) outputs are 208 interpolated over the $2.5^\circ \times 2.5^\circ$ grid defined for the observation. The period used for present conditions (20C3M run scenario) is 1970–1999, and the future is 2011–2040, 2041–2070, and 2071–2100 as derived from ten IPCC AR4 models. The ten models (Table 1) represent state-of-the-art AOGCMs. It is important to note, however, that the 20C3M simulation is intended to represent the same historical forcing scenarios, including both natural variability and the effect of human emissions on climate.

4. Application and Result

Initially, we apply the technique of validating the ANN for two towns in the Amazon Basin, Belém and Manaus (Figures 2 and 3). Figures 2 and 3 show an excellent correlation between observation and ANNs, both in Belém (Figure 2(b)) and Manaus (Figure 3(b)) showing the reduced uncertainty observed data.

Comparison of results obtained using the ANNs compared with observation indicates that the network is a potential competitive alternative tool for the analysis of multivariate time series (Figures 2 and 3).
For most indices, the regionally averaged series are expressed in the index units, but several of the precipitation indices have millimeter units. Thus at the regional-wide scale, there is a consistency between indices that average daily rainfall. In this paper, four indexes are used for the precipitation climate extremes as follows:

(a) PRCPTOT: annual total wet-day precipitation (precipitation $\geq 1$ mm),

(b) CDII: annual total precipitation divided by the number of wet days (precipitation $\geq 1$ mm),

(c) CDD: maximum number of consecutive days with daily rainfall $<1$ mm,

(d) R25mm: annual count of days when precipitation $\geq 25$ mm.

The results for all indices (Figure 4) are summarized over all stations over the Amazon Basin, including the fraction of station with trends that are significant at the 5% level, in the 6–9% range, and that are not significant (statistical) by Kendall’s tau.

We use a nonparametric trend statistic Kendall’s tau for monotonic trends, which makes no assumption about the distribution of the data or the linearity of any trend (e.g.,
Table 1: Climate models with daily data for precipitation available from PCMDI.

<table>
<thead>
<tr>
<th>Models</th>
<th>Models data</th>
<th>Font Resolution</th>
<th>Spatial resolution Km</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSR NCAR/USA</td>
<td>1.4° × 1.4°</td>
<td>154 × 154 km</td>
<td></td>
</tr>
<tr>
<td>CSIRO-MK3.0 CSIRO/Australia</td>
<td>1.9° × 1.9°</td>
<td>209 × 209 km</td>
<td></td>
</tr>
<tr>
<td>ECHAM5/mpi-OM MAX-PLACK/Germany</td>
<td>1.9° × 1.9°</td>
<td>209 × 209 km</td>
<td></td>
</tr>
<tr>
<td>GFDL-CM2.0 NOAA/GFCL/USA</td>
<td>2.0° × 2.5°</td>
<td>220 × 275 km</td>
<td></td>
</tr>
<tr>
<td>BCCR-BCM2.0 BCCR/Norway</td>
<td>1.9° × 1.9°</td>
<td>209 × 209 km</td>
<td></td>
</tr>
<tr>
<td>GISS-ER NASA/GISS/USA</td>
<td>4.0° × 5.0°</td>
<td>440 × 440 km</td>
<td></td>
</tr>
<tr>
<td>IPSL-CM4 IPSL/France</td>
<td>2.5° × 3.7°</td>
<td>275 × 413 km</td>
<td></td>
</tr>
<tr>
<td>MRI-CGCM2.3.2 MRI/Japan</td>
<td>2.8° × 2.8°</td>
<td>308 × 308 km</td>
<td></td>
</tr>
<tr>
<td>PCM NCAR/USA</td>
<td>2.8° × 2.8°</td>
<td>308 × 308 km</td>
<td></td>
</tr>
<tr>
<td>BCC-CM1 BCC/China</td>
<td>1.9° × 1.9°</td>
<td>209 × 209 km</td>
<td></td>
</tr>
<tr>
<td>HadCM3 Hadley Centre/UK</td>
<td>3.75° × 2.5°</td>
<td>370 × 250 km</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Trend analysis for 1979–2000 for regional time series 507 of anomalies from the Amazon Basin.

<table>
<thead>
<tr>
<th>Index</th>
<th>Units</th>
<th>Trend, units/period Observation</th>
<th>Trend, units/period ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRECPTOT</td>
<td>mm</td>
<td>4.69</td>
<td>4.66</td>
</tr>
<tr>
<td>SDII</td>
<td>mm</td>
<td>0.34</td>
<td>0.43</td>
</tr>
<tr>
<td>CDD</td>
<td>Number of days</td>
<td>0.77</td>
<td>0.72</td>
</tr>
<tr>
<td>R25mm</td>
<td>Number of days</td>
<td>1.20</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Figure 4: Trends of Kendall’s Tau for all indices.

Regional time series of several precipitation indices can be seen in Table 2. Most stations show no significant trends for the different indices, due to the relatively short period of data, both for observation and for ANNs.

Checked is a small difference between observation and ANNs; this small difference is related to the homogeneity of the time series of precipitation. This is due to the fact that data quality assessment is an important requirement before the calculation of indices since any erroneous outlier can have a serious impact on the trends in extremes. There is a small trend more positive in all four indices (Table 2).

In the present study, we aim at evaluating the evolution of large-scale patterns rather than regionally averaged indices. A solution to the present problem is the use of an artificial neural network interpolation approach which should lead to determine optimal combination of models.

Figure 5(a) shows the difference between models (Table 1) versus observation data. The difference is normalized. The great difference is located over north, west, and southwest Amazon Basin, with values in 1 mm/day (Figure 5(a)). After application of Ann’s correction, the results are very good, principally over the central and east Amazon Basin. In the north, west and southwest improvement occurred, reducing the value of difference.

Figure 6 shows the trend for the R25mm. In this figure, we have elected, for ease of interpretation, to show just the sign of the trend as well as flag interpolation with trends significant at $P < 0.05$, using the ANN as calibration models. This figure shows significant high value over northeast and northwest Amazon Basin, with proximally 52 days with rainfall >25 mm. These two areas have very high annual rainfall value, above 3000 mm [14]. These results show that the calibration gets the signal variability of rainfall over the Amazon Basin.
Figure 5: Difference between models and observation data (mm/day) (a) and correction from artificial neural network (mm/day) (b). Period between 1979 and 2000.

Figure 6: Sign of the linear trend in rainfall indices. Very heavy precipitation days (annual count of days when precipitation ≥25 mm).

Figure 7: Sign of the linear trend in rainfall indices. Very heavy precipitation days (annual count of days when precipitation ≥25 mm) for 2079–2100 period and A2 scenarios.

This result is relatively coherent with observation data. Such values are beyond the range of the dataset used for training. The artificial neural network makes its extrapolation skill not questionable. Figure 7 shows a significant decrease in extreme rainfall across the Amazon Basin, decreasing by almost 50% days with rainfall exceeding 25 mm. In the areas where the 20th century (Figure 6) had many days with rainfall exceeding 25 mm, especially in the northwest Amazon Basin, this reduction was decreased of up to 30%.

Finally, a comparison between the outputs of the models and ANNs was very good. This technique of using ANN in the identification and correction of meteorological data is new and effective.
5. Discussion

This study is an examination of the trends in indices of rainfall extreme over the Amazon Basin during 1979 to 2000 and future scenarios A2 (IPCC, 2007) for 2071–2100. Fundamentally, it is necessary to use new techniques to identify extreme events. In this paper, we use artificial neural network as a tool for identification and calibration output model.

Amazon basin can be categorized as a region at great risk due to climate variability and change. Marengo et al. [2] cited that the risk is not only due to projected climate change but also through synergistic interaction with existing threats not to climate change. The Amazon Basin is affected by large climate variability and extreme events, with the strongest increase in 2005.

Marengo et al. [2] shows that large section of the western Amazon Basin experienced the most severe drought in the 40 years and also one of the most intense of the last 100 years; the authors suggest that the drought was not caused by El-Niño but by (a) the anomalously warm tropical North Atlantic, (b) the reduced intensity in northeast trade wind moisture transport into southern Amazon Basin during the summertime season, and (c) the weakened upward motion over this section of the Amazon Basin, resulting in reduced convection development and rainfall.

We have described the results of an analysis of extreme in climate data for output models and output from ANN. Regionally average dry spell length, average rainfall intensity, and annual 1 day maximum rainfall all show statistically significant increasing trends.

A synthesis of the results is as follows.

(i) Eleven models (11) were very good at detecting extreme rainfall indices.

(ii) The use of ANNs showed good similarity with the observation (Figures 2 and 3), in the cities of Belém and Manaus, with correlations of approximately 88.9% and 91.3%, respectively.

(iii) The method of trends of Kendall’s tau of the ANNs showed a “shortening” in the variability in all indices calculated (Figure 4).

(iv) The spatial distribution, especially in the correction process (Figure 5), represents a decrease in the differences of up to 50%.

(v) The index of extreme R25mm can be spatially well represented (Figure 6).

(vi) In projection of the 21st century, a decrease of days with rainfall above 25 mm up to 35% for the A2 scenario (IPCC, 2007).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors wish to thank the World Wildlife Fund (WWF) for their financial support in the development of this study. They thank the IPCC for facilitating the access to the simulation output and thank the international modeling group for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI).

References


