Forest fire risk mapping in the Brazilian Amazon using MODIS images and artificial neural networks

Eduardo Eiji Maeda [1,2]
Antonio Roberto Formaggio [1]
Yosio Edemir Shimabukuro [1]
Gustavo Felipe Balué Arcoverde [1]
André Lima[1]

[1] Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil
{eduardo, formag, yosio, gustavo, andre}@dsr.inpe.br

[2] University of Helsinki
Gustaf Hällströmin katu 2, 00014, Helsinki, Finland
eduardo.maeda@helsinki.fi

Abstract. The present work describes a methodology based on Artificial Neural Networks (ANN) and multi-temporal images from the MODIS/Terra-Aqua sensors in order to detect areas with high risk of forest fire in the Brazilian Amazon. The hypothesis of this work is that, due to the characteristics of land use and land cover change dynamics in the Amazon forest, the temporal spectral profile of forest areas preparing to be burned can be separated from other areas. A study case was carried out in three municipalities in the north region of Mato Grosso State, Brazilian Amazon. Feedforward ANNs, with different architectures, were trained with a backpropagation algorithm, taking as inputs the NDVI values calculated from MODIS images acquired during five different periods preceding the forest fire season. Samples were extracted from areas where forest fires were detected in 2005, and also from forest and agricultural areas. These samples were divided to train, to validate and to test the ANN. The tests results achieved a mean squared error of around 0.07. When simulated in an entire municipality, the ANN model was efficient in showing the spatial distribution of the forest fire probability, which was coherent with the fire events observed in the following months.

Keywords: Forest fire, Artificial Neural Networks, Amazon forest, MODIS.

1. Introduction

The 2007 report of the United Nations (UN) Intergovernmental Panel on Climate Change (IPCC) highlights a number of consistent information about the human influence over the global warming (IPCC, 2007). This report clearly establishes the need to reduce immediately the emission of greenhouse gases and explicitly indicate the future consequences of not solving this problem.

In Brazil, the fire on the tropical forests is the principal source of greenhouse gas emission, representing about 75% of the total volume of CO_2 released in the country (MCT, 2004). Some works show that the deforestation in the Brazilian Amazon is responsible for over 200 million tons of CO_2-equivalent carbon per year, contributing significantly for the global amounts of greenhouse gas emission and, consequently, for the planet climate changes (Fearnside, 2007).

The Project for the Monitoring of Deforested Areas in the Amazon Region (PRODES) showed that the deforested area increased from 10 million hectares in the 1970’s to 67 million hectares in 2005. Consequently, the development of techniques to map and monitor burned areas in tropical forests offers a great deal in the supporting of public policies and in the measurement and control of the greenhouse gas emission (Almeida Filho and Shimabukuro, 2004; Aragão et al., 2007; Shimabukuro et al., in press).

However, the mapping of burned areas is still insufficient to support the authorities in the prevention of new fire events. In an effort to improve the forest fire prevention policies in the

Brazilian Amazon, Nepstad et al. (1998) developed a methodology that integrated the effects of drought and logging activities to assess forest fire susceptibility in the year of 1998. Among the data used for this methodology were soil water capacity maps, precipitation and evapotranspiration records, and maps with the spatial distribution of regional logging centers.

In other effort to estimate the fire probability in the Brazilian Amazon, Arima et al. (2007) added economical variables that reflected farm-gate prices of beef and soybean in a spatially explicit model. In this work the authors found a positive correlation between fire and the price of beef and soybean.

According to Jaiswal et al. (2002), the causes of the forest fires can be classified in three main categories: (I) natural causes, (II) deliberately caused by man and (III) accidentally caused by man. The fires in the Brazilian Amazon rainforests are typically associated with the implementation of new agricultural areas. Studying the eastern Amazon, Uhl and Kauffman (1990) described that the autogenic factors in primary forests create a microclimate that virtually eliminates the probability of fire. For this reason, a long process precedes the transition between rainforests and agricultural activities.

Usually, the selective logging is the first activity that takes place in a rainforest area. After the woods with commercial interest are over, the rest of the vegetation are cut down with the use of electric saws and tractors. As the moisture content is initially high, the vegetation is left to dry until the next dry season, when they are finally burned in order to clean the area. In many occasion this process is repeated until the area is clean enough to allow the use of agricultural machines (Cochrane, 2001; Cochrane and Laurance, 2002).

Therefore, the hypothesis of the present work was that, due to the characteristics of land use and land cover change dynamics in the Amazon forest, the temporal spectral profile of forest areas preparing to be burned can be separated from other areas. The MODIS sensor, launched in 1999 and 2002 on board of the Terra and Aqua satellites, respectively, offers almost daily images with a spatial resolution up to 250m, representing a great opportunity in the monitoring of the natural resources in large areas, as the Amazon region (Justice et al., 2002a).

This work describes a methodology based on Artificial Neural Networks (ANN) and multi-temporal images from the MODIS sensor in order to detect areas with high risk of forest fire in the Brazilian Amazon.

2. Material and Methods

2.1 Study Area

The study was carried out in three municipalities in the north part of the Brazilian State of Mato Grosso: Tabaporã, Porto dos Gauchos and Itanhangá (Figure 1).

The total area of the municipalities is approximately 18 thousands squared kilometers. The predominant vegetation types of the region are the Seasonal Forests and the Humid Tropical Forest. The agriculture is the main economical activity of the region. These activities started in the 1970’s when the Brazilian Government initiated public policies promoting the occupation and development of the north region of the country, resulting in the deforestation of large areas of the Amazon forest.

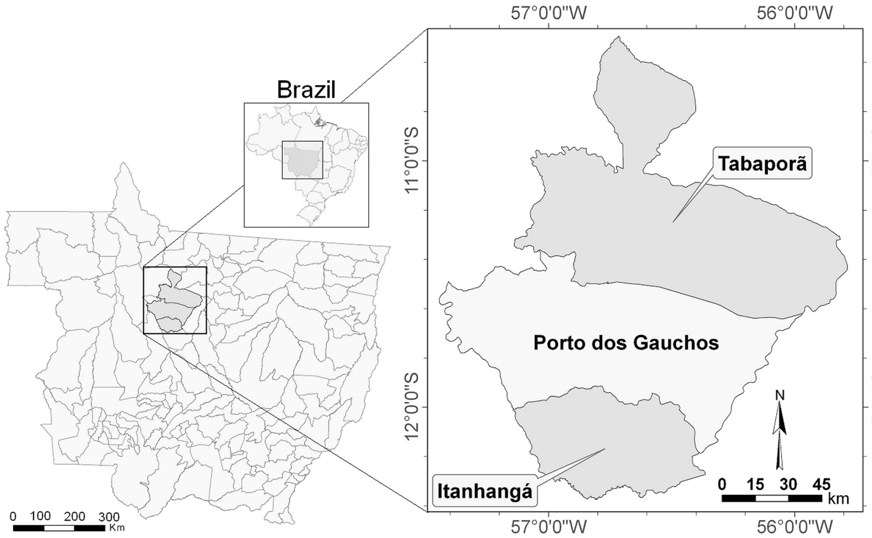

Figure 1. Study site location, in Mato Grosso State, Brazilian Amazon.

2.2 Methods

This work used multi-temporal values of the Normalized Difference Vegetation Index (NDVI) as inputs for feedforward ANNs. The NDVI values were obtained from the MOD13Q1 product (Justice et al., 2002a), which provides 16 days image compositions of vegetation indexes calculated from the MODIS/Terra-Aqua sensors images.

The time period selected to perform this study was the year of 2005. The dates of the images were chosen based on the results of previous works that show the peak of forest fire occurring approximately six months after the peak of the rainy season (Aragão et al., 2008).

Therefore, the months with higher number of fire occurrences in 2005 were August and September. Thus, in order to develop a methodology to predict the occurrence of fires, the analyzed images should be acquired before June. The periods chosen for the analysis were April and May, resulting in four 16 days image compositions. Along these periods, the forest areas that will be converted to agricultural areas are already being prepared to be burned. The previous months (January, February and March) did not make part of the analysis because of the high incidence of clouds in these periods, fact that could interfere in the operational usage of the method.

One additional image from late August/early September 2004 was also used to enhance the landu use changes occurred in the region, resulting, therefore, in five NDVI images. The dates and information of the images used are shown in Table 1.

Table 1. Date of the 16 days NDVI image compositions from the MODIS sensor used, and the respectively codes attributed to each one.

Code	Images date (dd/mm/yy)	Code	Images date (dd/mm/yy)
Aug04	28/08/ 04 - 12/09/04	**AprII**	23/04/05 – 08/05/05
AprI	07/04/05 – 22/04/05	**MayI**	09/05/05 – 24/05/ 05
		MayII	25/05/05 – 09/06/05

The samples used in the ANNs training process were selected within Tabaporã and Itanhangá municipalities based on a reference map. Each pixel area, with the five variations along the time, was considered as one sample. Although the objective of this work was only to indicate if a determined area has or not high probability of forest fire, the selected samples should include all the different situations found in the study area. The situations accounted were:

A- Agricultural areas in 2004 that were still agricultural areas in 2005.
B- Areas burned in 2004 and became agricultural areas in 2005.
C- Forest areas in 2004 that were burned in 2005.
D- Areas burned in 2004 and burned again in 2005.
E- Forest areas in 2004 that continued untouched areas in 2005.

Nine different ANNs architectures were tested, being the basic structure consisting of the input layer, one hidden layer and one output layer. The number of neurons tested in the hidden layer was 4, 6, 8, 10, 12, 14, 16, 18 and 20. The output layer was fixed in one neuron with a logarithmic-sigmoid transfer function, which means that it generates output values between 0 and 1, as the inputs can range from negative to positive infinity. Therefore, it was pre-defined that output values next to 1 would indicate areas with high probability of forest fire, while values next to 0, areas with low or none fire probability.

The ANNs were trained with the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963), which application to neural network training is described in Hagan and Menhaj (1994). One of the problems commonly faced on ANN training is the overfitting (German et al., 1992). The overfitting happens when the training error is set to a very small value, however, when new records are presented to the ANN the error becomes much higher, which means that the network memorized the training samples, but was not able to recognize unseen samples. To avoid the overfitting, the networks were trained using the early stopping technique (Prechelt, 1998). In this technique the total samples available were divided to train (50%), to validate (25%) and to test (25%) the ANNs. While the network is trained, the error of the validation set is monitored. If the network starts to overfit the training samples, the validation error starts to increase, as the train error continues to decrease. So when the validation error starts to increase successively, the train process is interrupted.

The test samples were not used during the training process, they were used to simulate the final networks and evaluate its performances. After defining the best architecture, the network that presented the best results was simulated in the Porto dos Gauchos municipality area. The result of the simulation was compared with the distribution of hot spots indicators of fire, captured by the MODIS sensors (Justice, 2002b), for the period from June to December 2005.

3. Results and Discussion

The NDVI temporal profiles of some sample segments taken from Tabaporã and Itanhangá municipalities are displayed in Figure 2.

It is observed that the NDVI temporal profiles from the forest samples are easily distinguished from the other areas, as its NDVI values are kept constantly high along the analyzed period. The profiles from areas that were just burned in 2005 (case C), are also easily differentiated, even though we can notice that in AprI the NDVI values of these burned areas are overlapped with other areas, showing the importance of a multi-temporal analysis.

Although the NDVI profiles in cases A, B and D seem initially hard to be distinguished, some particular characteristics of each case can be noticed. In case A, the NDVI values in Aug04 are very low, because in this period of the year the agricultural fields are starting to be prepared for the summer crop season, being predominantly cover with bare soil. In the pasture areas, although the soil is not revolved, the vegetation is very dry, because of the previously long dry season. The areas that were burned in 2004 and 2005 (D) had a more smooth transition along the analyzed period, giving that few or no vegetation grown happens in these areas in the rainy season, different from case A.

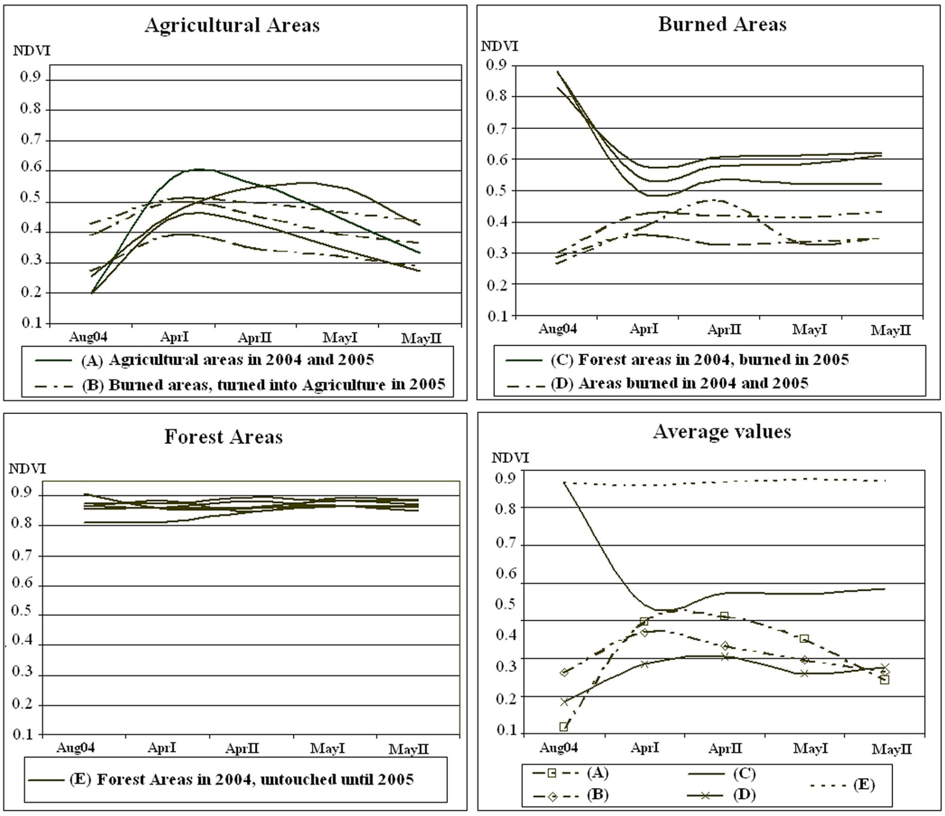

Figure 2. NDVI temporal profiles of some selected samples taken from Tabaporã and Itanhangá municipalities.

All the architectures tested achieved a mean squared error (MSE) oscillating close to 0.07. The increase in the architecture complexity, by adding neurons in the hidden layer, did not result in the improvement of the test performances. Hence, simple architectures, with just four neurons in the hidden layer are more appropriated for the solution of the studied problem, given that smaller networks are faster to train and have lower risk of overfitting (Fiszelew et al., 2007).

Slicing the output results, and considering values over or equal to 0.5 forest fire regions, and values lower than 0.5, regions with no fire risk, it was possible to evaluate the test results accuracy.

The global accuracy ranged around 90%, while the accuracy in the forest samples was close to 100%. As expected after analyzing the NDVI temporal profile of the samples, the main errors happened in distinguishing the agricultural areas from the areas that will be burned. Even so, the accuracy in identifying those areas was still satisfactory, reaching 85% for the areas with forest fire risk, and 90% for the agricultural areas. As seeing in the MSE analysis, no improvement in the accuracies occurred with the increase of neurons numbers in the hidden layer.

Hence, the network architecture chose to perform the simulation in Porto dos Gauchos municipality was [5 4 1], that is, five variables in the input layer, four neurons in the hidden layer and one neuron in the output layer.

The results of the simulation in Porto dos Gauchos municipality are displayed in Figure 3. The initial output values from the ANN model showed several noises, especially in areas that were not represented in the networks training samples, as boundaries of deforested areas, rivers and roads (Figure 3a and 3c). To get around this problem, a median filter (kernel size 5x5) was applied to smooth the result values (Jensen, 2005). The product of this procedure is exemplified in Figures 3b, and the entire fire risk map, as the final result of the simulation process is displayed in Figure 3d.

Although the filter application solved very properly the noise problems, in order to improve the results, future works should include in the training process samples from areas where this kind of errors were observed.

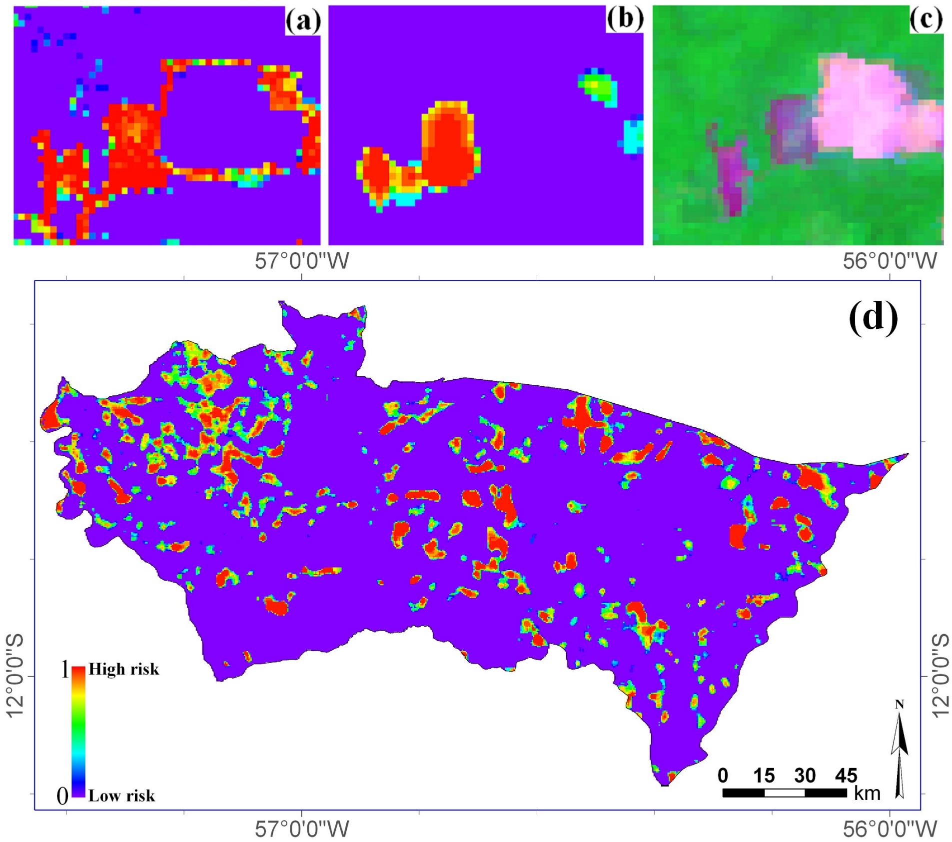

Figure 3. (a) raw simulation, with some noises observed specially in the boundaries of deforested areas; (b) simulation treated with a median filter (5x5); (c) MODIS image segment from September 2005; (d) final simulation result.

The approach carried out to assess the simulation results accuracy was to compare the fire risk map obtained from the ANN model with the hot spots captured by the MODIS sensor during the following months (June-December, 2005) (4a). The histogram of the hot spots illustrated in Figure 4b clearly shows that great part of the fire events observed after June got concentrated in the ANN model output values between 0.8 and 1, confirming the good agreement of the model.

A slight concentration of hot spots is also observed in the output values between 0 and 0.1. This fact was probably occasioned by the hot spots located in the border of burned areas, summed with eventual errors of the simulation and in the MODIS hot spots locations. Also, the filtering process may have decreased the areas of the segments where forest fire risk were detected, enhancing the probability of unmatched spots.

Some other mismatching cases were observed in areas were the ANN model indicated with high fire risk, but no hot spots were identified after June. In this cases, the main errors are attributed to areas that were burned before June and, therefore, these fire areas were simply classified, instead of predicted. The supplementary errors are attributed to agricultural areas that were incorrectly indicated with high fire risk.

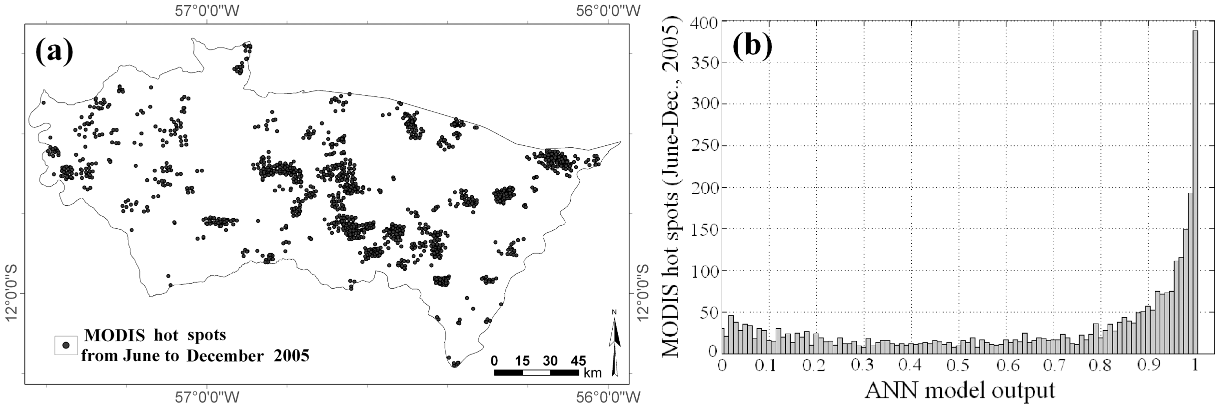

Figure 4. (a) Hot spots indicators of forest fire captured by the MODIS sensor from June to December 2005; (b) histogram showing the distribution of hot spots over the simulation results in Porto dos Gauchos municipality.

4. Conclusions

The ANNs training and tests achieved satisfactory results, reaching a MSE of around 0.07. Simple network architectures, with just four neurons in a single hidden layer are appropriated for the solution of the studied problem, providing a faster training and the same results of networks with more neurons. The highest errors were detected among the agriculture and forest fire risk areas, while the untouched forest areas were classified with almost 100% accuracy.

When simulated in Porto dos Gauchos municipality, the ANN model could indicate with relative precision the spatial distribution of areas where forest fire events took place in the year of 2005. The model result presented noises in areas that did not make part of the ANN training, as rivers, roads and deforested areas boundaries. Those noises were efficiently corrected by using a median filter with a kernel size of 5x5. However, it is recommended that future works should take those areas into account along the ANN training process.

The comparison of the simulation result with hot spots indicators of fire captured by the MODIS sensor from June to December 2005 presented good spatial agreement, being mostly the spots concentrated in output values from 0.8 to 1.

The ANN model presented in this work allowed a fast and relatively precise method to predict forest fire events in the studied area. Hence, it offers an excellent alternative for the support of forest fire prevention policies, and also to assist in the assessment of burned areas, reducing the uncertainty involved in the currents methods.

Nonetheless, further works are necessary in order to test and implement this method for the different vegetation types of the Brazilian Amazon and to develop an optimized operational methodology.

5 References

Almeida Filho, R.; Shimabukuro, Y. E. Monitoring biomass burning in the Brazilian Amazônia. **International Journal of Remote Sensing**, v. 25, n.24, p. 5537-5542, 2004.

Aragão, L. E. O. C. de; Malhi, Y.; Roman-Cuesta, R. M.; Saatchi, S.; Anderson, L. O.; Shimabukuro, Y. E. Spatial patterns and fire response of recent Amazonian droughts. **Geophysical Research Letters**, v.34, L07701-L07706, 2007.

Aragão, L. E. O. C. de; Malhi, Y.; Barbier, N.; Lima, A.; Shimabukuro, Y.; Anderson, L.; Saatchi, S. Interactions between rainfall, deforestation and fires during recent years in the Brazilian Amazonia. **Philosophical transactions of the royal society**, v.363, n.1498, p.1779–1785, 2008.

Arima, E. Y.; Simmons, C. S.; Walker, R. T.; Cochrane, M. A. Fire in the Brazilian Amazon: A spatially explicit model for policy impact analysis. **Journal of regional Science**, v.47, n.3, 541–567, 2007.

Cochrane, M. A. Synergistic interactions between habitat fragmentation and fire in evergreen tropical forests. **Conservation Biology**, v.15, n.6, p.1515-1521, 2001.

Cochrane, M. A.; Laurance, W. F. Fire as a large-scale edge effect in Amazonian forests. **Journal of Tropical Ecology**, v.18, n.3, p.311-325, 2002.

Fearnside, P. M. Uso da terra na Amazônia e as mudanças climáticas globais. Brazilian **Journal of Ecology**, v.10, n.2, p.83-100, 2007.

Fiszelew, A.; Britos, P.; Ochoa, A.; Merlino, H.; Fernández, E.; García-Martínez, R. Finding Optimal Neural Network Architecture Using Genetic Algorithms. **Advances in Computer Science and Engineering Research in Computing Science**, v.27, p.15-24, 2007.

Geman, S.; Bienenstock, E.; Doursat, R. Neural networks and the bias/variance dilemma. **Neural Computation**, v.4, n.1, p.1-58, 1992.

Hagan, M. T.; Menhaj, M. Training feedforward networks with the Marquardt algorithm. **IEEE Transactions on Neural Networks**, v.5, n.6, p.989-993, 1994.

Intergovernmental Panel on Climate Change (IPCC), **Summary for Policymakers: Synthesis Report**. Paris, 2007.

Jaiswal, R. K.; Mukherjee, S.; Raju, K. D.; Saxena, R. Forest fire risk zone mapping from satellite imagery and GIS. **International Journal of Applied Earth Observation and Geoinformation**, v.4, n.1, p.1-10, 2002.

Jensen, J. R. **Introductory Digital Image Processing: A Remote Sensing Perspective**, second ed. Upper Saddle River, NJ: Prentice Hall, 2005. 316p.

Justice, C. O.; Townshend, J. R. G.; Vermote, E. F.; Masuoka, E.; Wolfe, R. E.; Saleous, N.; Roy, D. P.; Morisette, J. T. An overview of MODIS Land data processing and product status. **Remote Sensing of Environment**, v.83, n.1-2, p.3-15, 2002a.

Justice, C.O.; Giglio, L.; Korontzi, S.; Owens, J.; Morisette, J. T.; Roy, D. P.; Descloitres, J.; Alleaume, S.; Petitcolin, F.; Kaufman, Y. The MODIS fire products. **Remote Sensing of Environment**, v.83, n.1-2, p.244-262, 2002b.

Levenberg, K. A Method for the Solution of Certain Non-linear Problems in Least Squares. **Quarterly of Applied Mathematics**, v.2, n.2, p.164–168, 1944.

Marquardt, D. W. An Algorithm for the Least-Squares Estimation of Nonlinear Parameters. SIAM **Journal of Applied Mathematics**, v.11, n.2, p.431-441, 1963.

MCT – Ministry of Science and Technology. **Comunicação inicial do Brasil à Convenção-Quadro das Nações Unidas sobre Mudanças do Clima**. Ministério da Ciência e Tecnologia. Brasília, 2004.

Prechelt, L. Automatic early stopping using cross validation: quantifying the criteria. **Neural Networks**, v.11, n.4, p.761-767, 1998.

Shimabukuro, Y. E.; Duarte, V.; Arai, E.; Freitas, R. M.; Lima, A.; Valeriano, D. M.; Brown, I. F.; Maldonado, M. L. R. Fraction images derived from Terra MODIS data for mapping burned areas in Brazilian Amazônia. **International Journal of Remote Sensing**, in press.

Uhl, C.; Kauffman, J. B. Deforestation, fire susceptibility and potential tree responses to fire in the eastern Amazon. **Ecology**, v.71, n.2, p.437–449, 1990.