

Application of artificial neural networks in modeling deforestation associated with new road infrastructure projects

Luisa Fernanda Gómez-Ossa & Verónica Botero-Fernández

Facultad de Minas, Universidad Nacional de Colombia, sede Medellín. lfomezos@unal.edu.co, vbotoero@unal.edu.co

Received: November 22th, 2015. Received in revised form: July 12th 2016. Accepted: December 12th, 2016

Abstract

Tropical deforestation is an ongoing process mainly caused by the construction of new roads, which, without proper environmental planning, contribute to biodiversity loss. Given that the artificial neural networks (ANNs) have the ability to capture nonlinear relationships, they were used to predict deforestation associated with new roads, such as the “Variante Porce” road and the “El Bagre-San Jacinto del Cauca” road in the department of Antioquia. ANN Training was carried out online using the back-propagation algorithm, part of the R software. The predictive capacity was evaluated using the area under the receiver operator characteristic curve (AUC). Also, a network that showed the best predictive capacity for the deforestation surface was generated for the baseline scenario and the simulated scenario incorporating the new roads. The comparison of scenarios suggested that new roads would increase the probability of deforestation for approximately 103.729 ha of forest.

Keywords: Artificial neural networks; prediction; deforestation; roads.

Aplicación de redes neuronales artificiales en la modelación de la deforestación asociada a nuevos proyectos de infraestructura vial

Resumen

La deforestación tropical es un proceso continuo causado principalmente por la construcción de nuevas vías, las cuales sin una planificación ambiental adecuada contribuyen a la pérdida de biodiversidad. Dado que las redes neuronales artificiales (RNAs) tienen la capacidad de capturar relaciones no lineales, se utilizaron para predecir la deforestación asociada a nuevas vías, como la Variante Porce y la vía El Bagre-San Jacinto del Cauca, en el departamento de Antioquia. El entrenamiento de las RNAs se realizó en modo on line con el algoritmo de retropropagación, en el software R. La capacidad de predicción se evaluó con el área bajo la curva ROC (AUC) y con la red que presentó mejor capacidad predictiva se generó la superficie de deforestación para el escenario base y el escenario simulado incorporando las nuevas vías. La comparación de escenarios indica que las nuevas vías incrementarían la probabilidad de deforestación de aproximadamente 103.729 ha de bosque.

Palabras clave: Redes neuronales artificiales; predicción; deforestación; vías.

1. Introduction

Due to biogeographic conditions, Colombia has a great amount of biodiversity, but it is unfortunately among the top ten countries that have shown important loss of forest habitat. It has an annual destruction rate of 0.5% [1]. It is calculated that by the year 2050 almost 80% of species extinction will be the result of deforestation [2]. This is an ongoing cycle in

Latin America that is brought about by the development of road infrastructure projects [3], which, without suitable environmental planning, contribute to the fragmentation of forests and biodiversity loss.

The main failures of environmental impact studies such as the descriptive nature of the assessments, the lack of measurable indicators and the absence of quantitative predictions [4] do not allow the potential impacts of new

How to cite: Gómez-Ossa, L.F. and Botero-Fernández, V., Aplicación de redes neuronales artificiales en la modelación de la deforestación asociada a nuevos proyectos de infraestructura vial. DYNA 84(201), pp. 68-73, 2017.

roads to be rigorously established, which could compromise the conservation of valuable resources in the long-term. These assessments should be complemented with spatial analysis models that help predict potential changes in the landscape caused by the construction of a road.

Statistical models have been widely used to model deforestation in the tropics, [5, 6] mainly in the Amazon [7,8]. Various pieces of research that included these models had identified the places most susceptible to deforestation and the main determinants. There has been no deep analysis or estimation of the condition of forests after the construction of new road infrastructure projects; this would help to improve the local management and conservation of forests.

Statistical models are attractive for decision makers due to their methodological stability and long application history. However, for these types of models, a priori assumptions should be made about the relationship between variables, which could not exist. The main advantage in the use of ANNs to make predictions is that this type of assumption is not required as they have the ability to capture nonlinear relationships. They can, therefore, be a more appropriate instrument when emphasis is placed on self prediction and not on the underlying relationships between independent and dependent variables [9].

ANNs have been widely applied in medicine and molecular biology. At the beginning of the 1990s the first studies were reported in the areas of ecology and environmental sciences [10]. Khoi and Murayama [11] used ANNs to predict the areas most vulnerable to deforestation, Larsen [12] used ANNs to evaluate the response of forest ecosystems in the presence of high amounts of carbon dioxide and ozone, and Pijanowski et al. [13] constructed an ANN model to predict changes in coverage and land use on a national scale.

The objective of this research is to construct neural network models, evaluate and choose the best model based on its performance (highest AUC value) to predict

deforestation from 1980-2000, and evaluate the change in the probability of deforestation for a scenario that simulates the construction of new roads.

2. Methodology

2.1. Study area

The study area includes the Bajo Cauca and Northeast sub-regions of the department of Antioquia (Fig. 1) which contain six and ten municipalities and have a total area of 8.585 km² and 8.645km² respectively. These sub-regions are mainly devoted to mining and cattle ranching and have only marginal agricultural production [14]. These regions still have forest areas with a great diversity of natural resources; however, their geographical locations has led to them being involved in the planning of new road infrastructure projects aimed to improve the department's connectivity with the northern part of the country. Therefore, it is an interesting zone to develop spatial deforestation models as a way of reducing the impacts caused by the construction of new roads.

Two roads were considered for this research: the “El Bagre - San Jacinto del Cauca” road, 72 Km in length; and the “Variante Porce” road that goes from the municipality of Yolombo up to the municipality of Zaragoza, approximately 150 km in length [15]. This road was originally designed in the study region to reduce the traveling time from Medellín to the coast; however, the feasibility studies were rejected by the environmental authorities in the 90s due to environmental impact on ecosystems. Their current stance is to rethink the assessments since this “Variante” would arrive in Zaragoza using 60 km less road than what was proposed in section 2 (T2) of the “Autopistas de la Prosperidad” (Fig. 1).

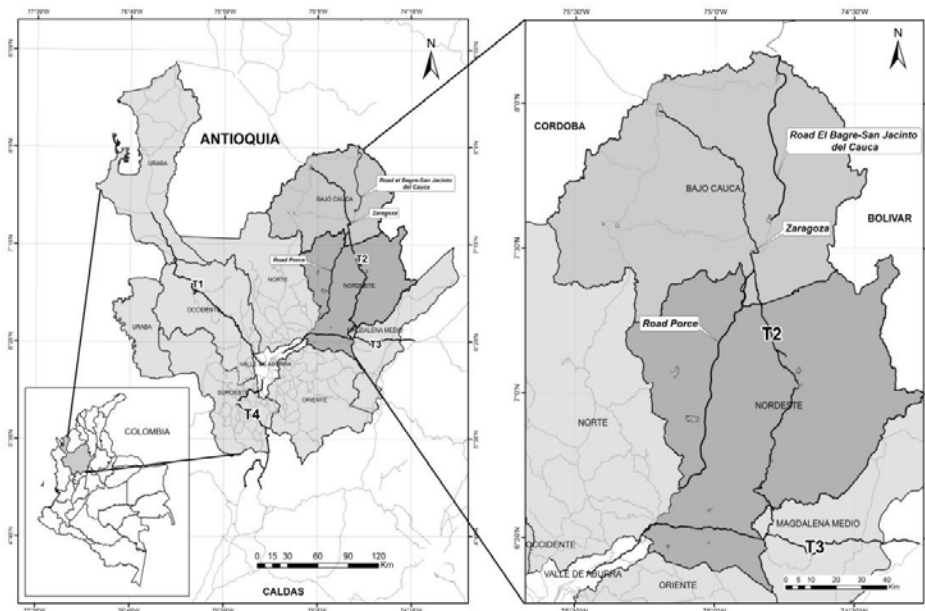


Figure 1. Location of the study area corresponding to the Northeast and Bajo Cauca sub-regions of the Department of Antioquia. Source: The authors.

2.2. Data and variables

Land coverage information for the years 1980 and 2000, using a scale of 1:25.000, was employed in the deforestation modeling. Land coverage information for 1980 was obtained after digitalizing more than 500 thematic plates [16] in analog format provided by the Ministry of Agriculture of Antioquia, and the land coverage for 2000 was the result of an inter-administrative agreement between the Ministry of Agriculture, the Administrative Department of National Planning and the “Instituto Geográfico Agustín Codazzi”. The information processing and homologation of coverage was carried out in a previous study using the Corine Land Cover methodology [16], and these were used as the basis to identify the changes from forest to no forest.

Six variables were built in raster format with a 30m pixel size. Orrego [16] indicates that it is best suited to represent deforestation and those variables that explain its occurrence.

Variables

- Euclidian distance to rivers.
- Euclidian distance to roads.
- Distance to forest border.
- Distance that represents the least cost to municipalities.
- Slope.
- Agricultural aptitude.

Although deforestation can be explained by a large number of variables, the underlying causes of deforestation in the study zone (poverty, poor forest management, insecurity in land tenure, etc.), which constitute the fundamental elements of the human-environmental relationship [17] are more difficult to identify than the proximate causes due to the availability of data. Therefore, biophysical and accessibility variables were used that provide a better spatial representation [18].

2.3. Neural network and validation training

The ANN training took place with the software R [19, 20] using the AMORE package, which, in turn, uses the back-propagation algorithm [21].

The proposed neural network has a unidirectional multilayer perceptron structure, with an input layer, a hidden layer and an output layer. Each explanatory variable is associated with a neuron in the input layer, and one neuron was assigned to the output that corresponded to the dependent variable. For this last layer, a logistic activation function was used because it is a categorical variable that takes values of zero and one (1= deforested, 0= not deforested)

According to the features of the package used, the weights start with random values within the range.

$$\pm \sqrt{\frac{3}{eo+os}} \quad (1)$$

Where eo is the number of connections between the input layer and the hidden layer, and os is the number of connections between the hidden layer and the output layer.

The training was carried out with a balanced sample of 250,000 points (3% of the data). That is to say, from this sample 50%, of the data corresponds to random deforested

points and the other 50% to random non-deforested points. The selection of this sample was made in order to improve the network performance and to be able to make better predictions since the learning network tends to be biased towards the class that has the majority in the training group [22,23] for a non-balanced sample.

The input data were standardized so that the values were in the range (0,1) since this speeds up the neural network training [24]. In addition, the standardization enables better initializations in the weights, which reduces the possibility of finding a local minimum. The data was transformed by using the maximum and minimum values in the original data set. By standardizing each variable, the ANNs assign equal importance to the data, so the problem of variables with different relevance due to their value range is avoided. The transformation was performed based on the following equation.

$$x'_i = (x_i - \text{mínimo})/(\text{máximo} - \text{mínimo}) \quad (2)$$

Using a standardized sample and changing the value of the training parameters, network models were trained. As Moreira and Fiesler [25] recommended, the values that were considered for the training rate were 0.01 and 0.05; both used a momentum of 0.5 and a rate of 0.1 with a momentum of 0.9. Similarly, the number of neurons in the hidden layer were adjusted and Hecht-Nielsen's suggestion [26] was taken as the basis. This author states that the use of $2n + 1$ hidden neurons (with n being the number of input neurons) can guarantee perfect adjustment of any ongoing function.

24 networks were constructed and trained. Due to the instability in training and poor error reduction after 5000 iterations, the cycle was interrupted for all networks, without significant accuracy losses. The logistic activation function was used for both the hidden layer and for the output layer, and the online learning method was used based on its ability to work with large training sets [27]. In addition, the network convergence is faster compared to the learning method in batch mode [28,29].

The quality of a model can be assessed in terms of its discriminatory capacity, that is to say the ability to predict the presence or absence of a feature [30]: in this case the presence or absence of deforestation. Therefore, the ROC curve (Receiver Operating Characteristic in English) and its associated area (AUC) were used. The ROC curve is independent of a threshold and provides a unifying criterion when evaluating the process of a model [31]. A detailed description of the ROC curve's application to evaluate the prediction of land use and land cover change can be found in [32].

To perform the validation, a random sample that was independent from the training set, and had a size of 250,000 points, was selected. Using the ROC package [33], the AUC values were estimated. This allowed some trends to be identified in the network's predictive capacity with respect to different sets of parameters (momentum, rate and amount of hidden neurons) as well as the best suited ANNs to predict deforestation.

The network that had best predictive capacity was chosen from 24-trained networks: that is to say, the one with the

highest AUC value. The deforestation surface for the baseline scenario and the simulated scenario were estimated with this network under the assumption that they both had similar deforestation patterns. For the simulated scenario estimation, the new roads were added and the variables distance to roads, distance to forest border and distance representing the least cost to the municipalities were updated. The original values for the other variables were maintained.

The pixels affected by the new roads were identified by overlapping the raster files of the deforestation probability surface in the baseline scenario with the estimated surface for the simulated scenario.

3. Results

The networks that showed high AUC values were those trained with a lower training rate and greater amount of neurons in the hidden layer; this can be observed in Fig. 2.

The ANN that showed best predictive capacity obtained an AUC value of 0.8212 (Table 1). Based on this, a deforestation probability surface was generated for the area covered by forests in 1980 as well as the deforestation probability surface for the simulated scenario.

The comparison of the scenarios suggests that 1,152,565 pixels, approximately 103,729 ha of forest (pixel area of 900 m²), would increase its probability of deforestation (Fig. 3); 75,622ha showed a low increase (0-0.11), 20.186ha an average increase (0.11-0.37), and 7.922ha a high increase (0.37-0.9). However, some pixels, approximately 7%, showed a reduction

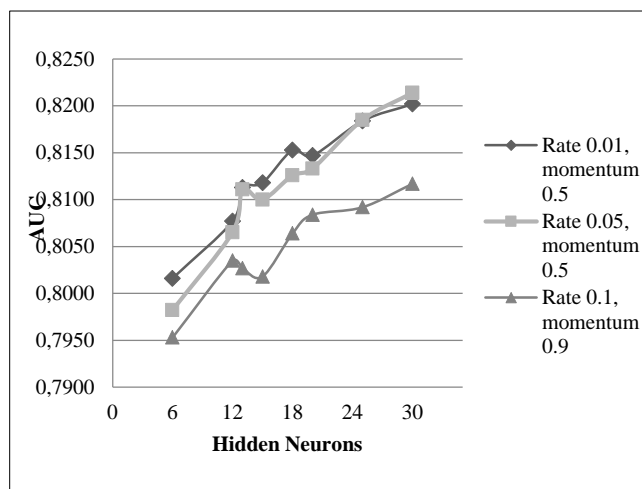


Figure 2. Area under the ROC curve (AUC), with respect to the training rates and the number of hidden neurons for a cycle of 5,000 iterations. Source: The Authors

Table 1. Parameters of the five artificial neural networks that showed best predictive capacity.

Momentum	Training rate	Hidden neurons	Error	AUC
0.5	0.01	20	0.173348	0.8147
0.5	0.01	25	0.170802	0.8184
0.5	0.05	25	0.172912	0.8185
0.5	0.01	30	0.170392	0.8202
0.5	0.05	30	0.171380	0.8214

Source: The Authors

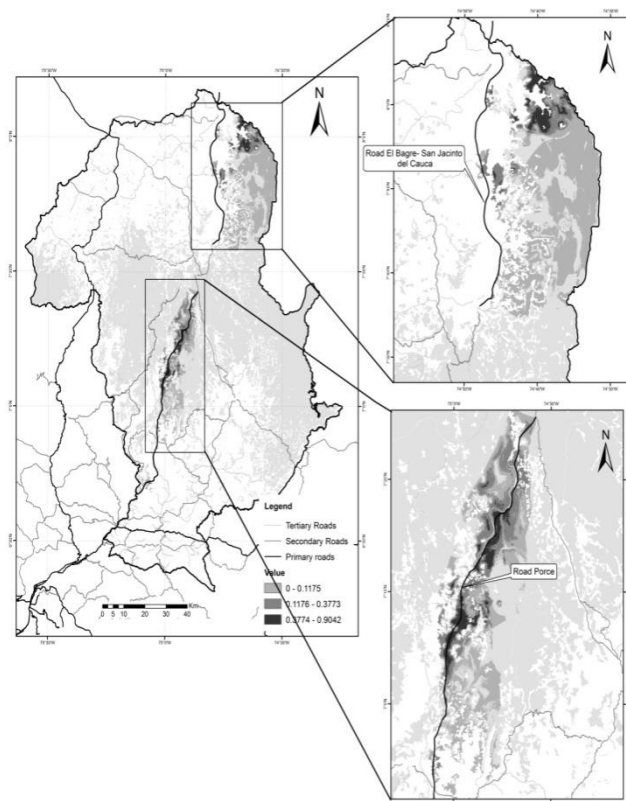


Figure 3. Location of the areas that had an increase in deforestation probability by the Construction of the “Variante Porce” and the road “El Bague-San Jacinto del Cauca”.

Source: The Authors

in the probability of deforestation for the simulated scenario, which goes against what was expected based on the literature, which indicates that the probability of deforestation increases with the proximity to roads [34, 35]. This may be attributable to the expected error in the simulation, taking into account that the predictive capacity for the best ANN was approximately 82%.

4. Discussion

As the training rate decreased, the AUC value increased, but we must bear in mind that even though the networks with low training rates show less error, compared with other rates, training is slower. Therefore, the value of the learning rate should be large enough to allow a quick learning process but small enough to ensure their effectiveness. For the current research, the training parameters that showed the best performance in terms of training times and predictive capacity was 0.05 for training rate and 0.5 for the momentum.

Another alternative for training would be the use of an adaptive network with a resilient back propagation algorithm in which the rate changes automatically during the training. This means that the problem of defining a global training rate that is appropriate for the whole learning process [36] is avoided.

The models generated in the training phase did not show

wide differences in terms of accuracy; the AUC values were between 0.79 and 0.82. This is consistent with Müller and Mburu [37] who also found similar results when modeling deforestation with different neural network architectures; and Okwuashi et al. [38] who simulated changes in land use in Lagos, Nigeria for three time periods, obtaining AUC values between 0.7574 and 0.7850.

The evaluation of the potential effect of the two roads proposed for the simulated scenario suggests an increase in the deforestation probability for 103.729ha of forest. Wilson et al. [39] in their study, carried out in the temperate zone forests in southern Chile, found an increase in deforestation probability of about 27,000ha after simulating the construction of a new road. In the central region of México, Nelson and Hellerstein [34] simulated the impact of new roads on deforestation and demonstrated that access to roads significantly influenced the change in land use.

The results of the research show an increase in the probability of deforestation within 6 km of roads; the probability diminishes as the distance increases. Mertens and Lambin [40] found a negative relationship between proximity to roads and deforestation at a distance beyond 7.5 km, and Barber et al. [36], in their study undertaken in the Amazon, found that 95% of the deforestation occurred within 5.5 km of roads. The results suggest that there is an agreement between different studies on the spatial pattern of environmental impacts resulting from the construction of roads. This offers the possibility to develop appropriate models for spatial scales that are larger than what is considered in this investigation.

Other than the immediate impact of road construction within or near to forest areas, roads are permanent elements of the landscape [41]. So, even when they do not directly influence deforestation, they represent a long-term determinant to forest conversion. This is the case with the simulation of the “El Bagre - San Jacinto del Cauca” road, which does not go through forest areas but generates an increase in the deforestation probability in the eastern part of the study area. This shows that the spatial configuration of the scenery is determined, to great extent, by the proximity of forests to roads.

Freitas et al. [41] indicate that although the variable of distance to roads is a strong predictor of forest dynamics in deforestation processes, the effect can only be detected when agricultural expansion has stabilized over a certain period. However, the effect of the construction of a new road such as the “variante Porce”, which would cross areas with low agricultural aptitude could be the increase in mining and excessive wood extraction and not the planting of agricultural crops. This is disturbing if we take into account that ecological studies suggest tropical forest based on poor soils have greater levels of biodiversity than those that have better quality soils [42].

The simulation shows the areas for which there is an increased deforestation probability are associated with the roads designed. Consequently, this estimation is a conservative one since many paths and temporal trails could be built from new roads facilitating access to forests. Therefore, the area that will experience an increase in deforestation could possibly be greater, which would put at risk the environmental services that house the forest ecosystems.

5. Conclusion

The simulation with the network that presented the best predictive capacity during the training process showed an increase in the deforestation probability for 103.729ha of forest. Although the results cannot be validated with real data since the proposed roads have not yet been built, the development of this type of analysis provides planners with an empirical orientation on the magnitude and location of damage. Therefore, after the completion of road projects strict surveillance by the environmental authorities is needed and the design of policies aimed at creating incentives for forest protection, which would enable the population to take part in forest conservation activities.

Although carrying out projections for change through empirical deforestation models generates uncertainty, the study shows how the ANN has the ability to capture changes and trends in terms of deforestation probability. Therefore, greater precision can be achieved by developing models for shorter time periods that help not to predict the exact location of future deforestation, but rather the places that have a high risk of conversion.

Furthermore, the ANNs could be important tools to select alternative routes of linear structures since they allow all the environmental information for the area in which it is intended for the civil work to be located to be integrated. This would help environmental science technicians to visualize landscape transformation while the location of the linear infrastructure is defined. This represents valuable information that is useful in order to formulate and implement compensation forestry projects under the mechanism for reducing emissions for deforestation and degradation (REDD).

Acknowledgements

We thank professor Sergio Alonso Orrego Suaza for providing land cover information and the Department of Geosciences and Environment at the Universidad Nacional de Colombia for providing the facilities to develop this research.

References

- [1] Savage, A., Guillen, R., Lamilla, I. and Soto, L., Developing an effective community conservation program for cotton-top tamarins (*Saguinus oedipus*) in Colombia. *American Journal of Primatology*, 72(5), pp.379-390, 2010. DOI: 10.1002/ajp.20770
- [2] Rodríguez, N., Armenteras, D. and Alumbroeros, J.R., Land use and land cover change in the Colombian Andes: Dynamics and future scenarios. *Journal of Land Use Science*, 8(2), pp.154-174, 2013. DOI: 10.1080/1747423X.2011.650228
- [3] Geist, H. and Lambin, E., What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence, Belgium: CIACO Iouva-in-la-Neuve. 116 P., 2001.
- [4] Mallard, F. and Francois, D., Effectiveness of the legal framework for natural areas protection relative to French road projects. *Land use policy*, 30(1), pp.582-591, 2012. DOI: 10.1016/j.landusepol.2012.05.006
- [5] Dávalos, L.M., Bejarano, A.C., Hall, M.A., Correa, H.L., Corthals, A. and Espejo, O.J., Forests and drugs: Coca-driven deforestation in tropical biodiversity hotspots. *Environmental Science and Technology*, 45, pp.1219-1227, 2011. DOI: 10.1021/es102373d
- [6] Armenteras, D., Cabrera, E., Rodríguez, N. and Retana, J., National and

- regional determinants of tropical deforestation in Colombia. *Regional Environmental Change*, 13(6), pp.1181-1193, 2013. DOI: 10.1007/s10113-013-0433-7
- [7] Soares-Filho, B., Coutinho, G. and Lopes, C., DINAMICA - A stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, 154(3), pp.217-235, 2002. DOI: 10.1016/S0304-3800(02)00059-5
- [8] Dalla-Nora, E.L., de Aguiar, A.P., Lapola, D.M. and Woltjer, G., Why have land use change models for the Amazon failed to capture the amount of deforestation over the last decade?. *Land Use Policy*, 39, pp.403-411, 2014. DOI: 10.1016/j.landusepol.2014.02.004
- [9] Brey, T., Jarre-Teichmann, A. and Borlich, O., Artificial neural network versus multiple linear regression Predicting P/B ratios from empirical data. *Marine Ecology Progress Series*, 140(1-3), pp. 251-256, 1996. DOI: 10.3354/meps140251
- [10] Lek-Ang, S., Deharveng, L. and Lek, S., Predictive models of collembolan diversity and abundance in a riparian habitat. *Ecological Modelling*, 120(2-3), pp.247-260, 1999. DOI: 10.1016/S0304-3800(99)00106-4
- [11] Khoi, D.D. and Murayama, Y., Forecasting areas vulnerable to forest conversion in the Tam Dao National Park Region, Vietnam. *Remote Sensing*, 2(5), pp. 1249-1272, 2010. DOI: 10.3390/rs2051249
- [12] Larsen, P., Cseke, L., Miller, R. and Collart, F., Modeling forest ecosystem responses to elevated carbon dioxide and ozone using artificial neural networks. *Journal of Theoretical Biology*, 359, pp.61-71, 2014. DOI: 10.1016/j.jtbi.2014.05.047
- [13] Pijanowski, B., Tayyebi, A., Doucette, J., Pekin, B., Braun, D. and Plourde J., A big data urban growth simulation at a national scale: Configuring the GIS and neural network based land transformation model to run in a High Performance Computing (HPC) environment, 51, pp. 250-258, 2014.
- [14] Gobernación de Antioquia, IDEA & IGAC., Antioquia características geográficas, Bogotá, Colombia: Imprenta Nacional de Colombia, 2007.
- [15] Secretaría de Infraestructura Física. Antioquia Mapa vial. República de Colombia, Departamento de Antioquia, 2009.
- [16] Orrego, S., Economic modeling of tropical deforestation in Antioquia (Colombia), 1980-2000: An analysis at a semi-fine scale with spatially explicit data. PhD Dissertation. Oregon State University, USA, 2009.
- [17] Geist, H. and Lambin, E., What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence, Belgium. CIACO Louvain-la-Neuve, 2001.
- [18] Thies, B., Meyer, H., Nauss, T. and Bendix, J., Projecting land-use and land-cover changes in a tropical mountain forest of Southern Ecuador. *Journal of Land Use Science*, 9(1), pp.1-33, 2012. DOI: 10.1080/1747423X.2012.718378
- [19] R Development Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, [online]. Vienna, Austria. Available at: URL <http://www.R-project.org/>, 2013.
- [20] R Studio, 2013. RStudio: Integrated development environment for R (Version 0.97.551) [Computer Software, online]. [Accessed 2013 Jan 5]. Boston, MA. Available at: <http://www.rstudio.org/>.
- [21] Castejón, M., Ordieres, J.B., Vergara, E.P., Martínez-de-Pisón, F.J., Pernía, A.V. and Alba, F., AMORE: AMORE flexible neural network package. R package version 0.2-12. [online]. Available at: URL <http://CRAN.R-project.org/package=AMORE>, 2010.
- [22] Nguyen, G.H., Bouzerdour, A. and Phung, S., Learning pattern classification tasks with imbalanced data sets. *Pattern recognition*. Vukovar, Croatia: Intech, pp. 193-208, 2009.
- [23] He, H. and Garcia, E., Learning from imbalanced data. *Knowledge and data engineering*, IEEE, 21(9), pp.1263-1284, 2009. DOI: 10.1109/TKDE.2008.239
- [24] Li, X. and Yeh, A.G.-O., Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), pp.323-343, 2002. DOI: 10.1080/13658810210137004
- [25] Moreira, M. and Fiesler, E., Neural networks with adaptive learning rate and momentum terms. IDIAP Technical Report, No 95-04, 1995.
- [26] Hecht-Nielsen, R., Kolmogorov's mapping neural network existence theorem. In *Proceedings of IEEE First Annual International Conference on Neural Networks*, pp. III-11, 1987.
- [27] Plagianakos, V.P., Magoulas, G.D. and Vrahatis, M.N., Learning rate adaptation in stochastic gradient descent. In N. Hadjisavvas and Pardalos, P., eds. *Advances in Convex analysis and global optimization*. nonconvex optimization and its applications, [online]. 2001. Springer US, pp. 433-444. Available at: DOI: 10.1007/978-1-4613-0279-7_27
- [28] Bengio, Y., *Neural networks for speech and sequence recognition*, London: International Thomson Computer Press, 1996.
- [29] Haykin, S., *Neural networks, A comprehensive foundation* 2a ed., Hamilton, Canada: Prentice- Hall International, Inc, 1999.
- [30] Freeman, E. and Moisen, G., A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217(1-2), pp.48-58, 2008. DOI: 10.1016/j.ecolmodel.2008.05.015
- [31] Fawcett, T., An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), pp.861-874, 2006. DOI: 10.1016/j.patrec.2005.10.010
- [32] Pontius Jr, P. and Batchu, K., Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India. *Transactions in GIS*, (4), pp. 467-484, 2003. DOI: 10.1111/1467-9671.00159
- [33] Sing, T., Sander, O., Beerenwinkel, N. and Lengauer, T., ROCr: Visualizing the performance of scoring classifiers. R package version 1.0-5, [online]. 2013. Available at: URL <http://rocr.bioinf.mpi-sb.mpg.de>.
- [34] Nelson, G. and Hellerstein, D., Do roads cause deforestation? Using satellite images in econometric analysis of land use. *American Journal of Agricultural Economics*, 79(1), pp.80-88, 1997. DOI: 10.2307/1243944
- [35] Barber, C., Cochrane, M.A., Souza Jr, C.M. and Laurance, W., Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biological Conservation*, 177, pp.203-209, 2014. DOI: 10.1016/j.biocon.2014.07.004
- [36] Günther, F. and Fritsch, S., Neuralnet: Training of neural networks. *The R Journal*, 2(1), pp.2073-4859, 2010.
- [37] Müller, D. and Mburu, J., Forecasting hotspots of forest clearing in Kakamega Forest, western Kenya. *Forest Ecology and Management*, 257, pp.968-977, 2009. DOI: 10.1016/j.foreco.2008.10.032
- [38] Okwuashi, O., Isong, M., Eyo, E., Eyoh, A., Nwanekezie, O., Olayinka, D., Udoudo, D. and Ofem, B., GIS cellular automata using artificial neural network for land use change simulation of Lagos, Nigeria. *Journal of Geography and Geology*, 4(2), pp.94-101, 2012. DOI: 10.5539/jgg.v4n2p94
- [39] Wilson, K., Newton, A., Echeverría, C., Wetson, C. and Burgman, M., A vulnerability analysis of the temperate forests of south central Chile. *Biological Conservation*, 122(1), pp.9-21, 2005. DOI: 10.1016/j.biocon.2004.06.015
- [40] Mertens, B. and Lambin, E., 1997. Spatial modelling of deforestation in southern Cameroon. *Applied Geography*, 17, pp.143-162, 1997. DOI: 10.1016/S0143-6228(97)00032-5
- [41] Freitas, S.R., Hawbaker, T.J. and Metzger, J.P., Effects of roads, topography, and land use on forest cover dynamics in the Brazilian Atlantic Forest. *Forest Ecology and Management*, 259(3), pp.410-417, 2010. DOI: 10.1016/j.foreco.2009.10.036
- [42] Huston, M., *The coexistence of species on changing landscapes*, Cambridge, U.K: Cambridge University Press, 1994.

L.F. Gómez-Ossa, is a BSc. in Forest Eng. who undertook her undergraduate degree at the Universidad Nacional de Colombia, in Medellín. She received her MSc. in Environment and Development in 2015 from the same university. Her scientific interest lies in the application of computational methods for environmental modeling; in particular, she works on understanding and exploring patterns of deforestation in the tropics. Currently, she works as a private consultant in the management of geographic information systems and environmental modeling areas. ORCID: 0000-0001-9971-4908.

V. Botero-Fernández, is BSc. in Civil Eng. who undertook her undergraduate degree at the Universidad Eafit, Medellín, Colombia. She received her MSc. in 1997 from the International Institute for GeoInformation Science and Earth Observation and her PhD from the Universidad de Utrecht, Netherlands in 2009. Her areas of interest are digital image processing and geo-informatics applied to environmental management. She is currently a full professor in the Geosciences and Environment Department, Faculty of Mines, Universidad Nacional de Colombia, in Medellín. ORCID: 0000-0002-6964-239X.