APPLYING A HIBRID SYSTEM WITH NEURAL NETWORKS TO REPRESENT THE HUMAN DIMENSION IN ENVIRONMENT PROJECTS: A CASE STUDY IN AN AMAZON COMMUNITY

Rômulo M. de Sousa, Antonio M. da Silveira, Roberto C. Limão, Nirvia Ravena

UFPA/Universidade Federal do Pará, Brazil
rmsousa@ufpa.br, morais@ufpa.br,limao@ufpa.br, ravena@ufpa.br

Abstract – This paper presents the results achieved by an application of Computational Intelligence technique in an Environmental and Social research activity in an innovative initiative, using a computational model formed by two Artificial Neural Networks (ANN) of distinct intentions to construct an environmental indicator and a characteristics map, to assist researchers and decision makers in the analysis of human intervention, through the measurement of the capacity to conserve the environment of a Amazonian community. The use of this computational model distinguishes itself in relation to the classic mathematical models, because it extracts the cognitive perceptions reached by researchers during the process of field research, avoiding distortions of the reality.

Keywords – Artificial Neural Network, Kohonen Neural Network, Self-Organization Map, Environment, Sustainable Development, Environment and Social Analysis.

1 Introduction

The objective of this paper is to present a solution to aid in the process of diagnoses and prediction in analyzing research projects related to human interventions in the environment, that will allow a representation that is closer to reality in observation, without discarding its component elements, adequately dealing with questions of complexity and uncertainty that are generally involved in information related to human behavior and its interaction with the environment.

Translating the intrinsic complexity of human interaction with the environment with the use of traditional models, may lead to generation of weakly formulated evidences, since the models based on classic mathematics are hard to prepare for situations of uncertainly, given that the variables involved, some tangible and others intangible, are normally not precise. Those variables may contain errors resulting from noises generated in obtaining the information and from imperfections in the system [1]. In these models, there is an attempt to describe the elements that are considered fundamental in the situation, deliberately ignoring elements recognized as secondary, with the good models being described as not excessively simplified, but as highlighting fundamental aspects of the situation [2]. One may perceive here the tendency mathematical models have of discarding elements of the reality that they represent, adjusting reality to the model and not the model to reality [3].

In observing reality it is not possible to know and project future conditions of environmental and human systems in a perfect manner, due to the intrinsic complexity of those systems. It is also know that experience can provide forecasting skills through experimentation [4]. Another point to observe is that, specifically in environmental studies, the models fail to provide understanding of situations related to value conflicts and uncertainty [5],[6].

In the NRC view [5] it is important to search for alternative methods that can be adjusted to problems in a natural manner, without the restrictions imposed by mathematical and statistical formalisms that most of the time discard information, or fail to consider nuances that are sometimes important for mapping and understanding complex environmental solutions.

As a solution for representing the relation entre humans and the environment, this paper proposes a neural computational model for generating an indicator that will measure sustainable behavior among the individuals inhabiting a given community. Based on indicators from the individuals analyzed, the proposed solution builds a graphic representation that demonstrates the sustainability capacity of individuals in the community.

The artificial neural networks (ANN) are an alternative application for solving the problems presented here, given their capacity for learning the perceptions of the world that are presented to them. Their fundamental difference in relation to processing of classical information is that in the latter, we proceed primarily by formulating a mathematical model of observations of the environment, validating it with real data, and then structuring the information system based on this model, whereas the ANN, directly utilize data from the real world, allowing the data to speak for themselves, providing an implicit model of the environment. They also carry out the function of processing information that is of interest [1].

The indicators are employed in this paper because they are a summarized form of expressing and representing the vast and complex environmental system. They have a series of purposes. They are very important for formulating, implementing and assessing environmental policies in a rational fashion [5]. However, the environmental indicators have been constructed in a restricted fashion. Initially, efforts at developing those indicators were undertaken by statisticians, resulting in sophisticated indicators with little repercussion, since, when rates combine different types of data they are rejected. In the view of many statisticians, this results from the fact than much detailed information is lost when the data are combined within rates [5]. Thus, building indicators by using traditional mathematical and statistical statistics to condense a large amount of information in a single rate has been the usual practice in research seeking to diagnose the environment. Such a procedure compromises the information gathered in the field, where the data are permeated by uncertainty, by their subjective character, which encompasses aspects of belief, culture, gender and other variants that are part of the process of constructing the perception of reality [7]. Nonetheless, it is exactly the subjectivity collected in field researchers that is the great challenge in constructing indicators. To apprehend it, it is not only necessary to simply process information, but fundamentally to tap into the cognitive capacity of environmental researchers, extracting their knowledge and their experience in interpreting and perceiving reality that one wishes to model, to amalgamate in constructing the indicator.

2 The Proposed Computational Model

The proposed solution model is founded on three fundamental elements: the specialist's knowledge, an ANN to capture this knowledge ant another ANN to map the results achieved. This structure may be observed in Figure 1.

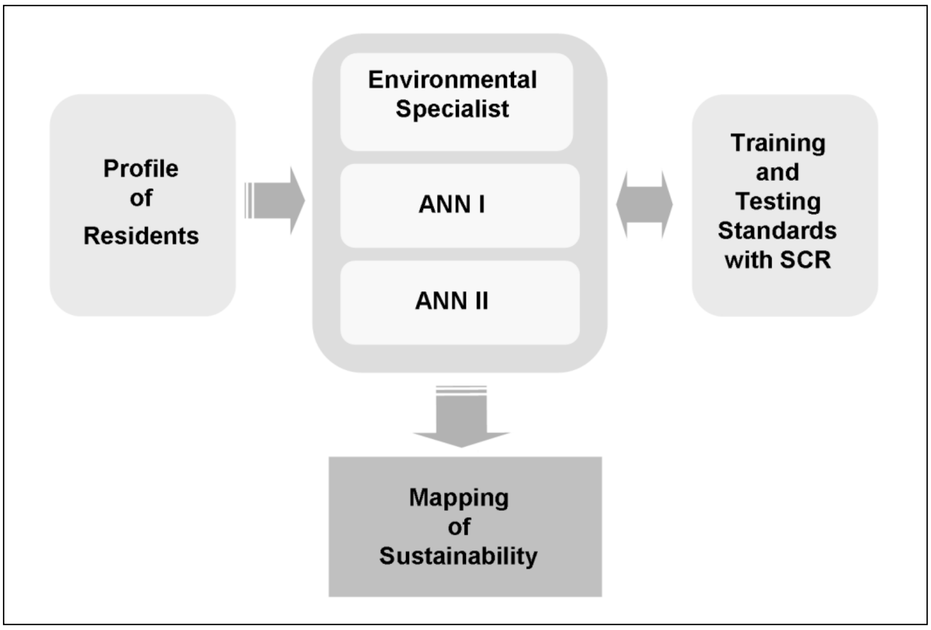

Figure 1 – Architecture of the neural model.

Inside the model, environmental specialist is inserted in the process of construction, with her or his knowledge and expertise employed in amalgamating the several variables that make up the profile of residents, through an analysis of the profile of residents, who were randomly selected to make up the entry standards for training in ANN I. As a result of that analysis, the specialist attributes a sustainable capacity rate, or SCR, for each profile selected.

Following the model description, we have the ANN I, a multilayered direct artificial neural network, which captures the specialist's knowledge through a process of supervised learning, using a backpropagation algorithm with the training entry standards. This network captures the criteria used by the specialist to attribute the SCR of selected residents used in training for the network. Once the learning process is concluded, the ANN I will provide an attribution of the other profiles that make up the research database, determining their respective SCR.

Finally, the ANN II, a Kohonen neural network for non-supervised learning, with a neural grid of size m x n, topographically maps its neurons, as the entry standards are presented to it. The ANN II is trained with the same entry standards utilized in the ANN I, and later on the same profiles are presented to the network, in order to identify the similarity of characteristics among the residents of the community, separating them by clusters.

As a result, we have the construction of maps of characteristics for analysis. These make it possible to assess the quantity of individuals distributed among the various groups, as well as the distribution of the SCR among individuals with the same similarity in socioeconomic and educational characteristics.

Based on the SCR and the maps of characteristics, respectively defined by the ANN I and ANN II, it is possible to carry out diagnoses and predictions for the community related to environmental questions.

The proposed solution is mainly directed towards researchers and decision makers in the environmental area, belonging to private, governmental and non-governmental institutions, who have an interest in expanding their analysis of human behavior in certain communities, at the most diverse scales (local, national, regional, global) [8], to diagnose and predict the effects of human intervention on the environment in which a given community is inserted.

3 Results Achieved

To generate results, the database was used from an environmental research project that studies an Amazonian community and the human intervention impact caused on the environment where this community is located [9]. The database takes into account the residents' profile. It is made up of socioeconomic and educational data, and has 1,174 items, made up of the following variables: age, sex, degree of kinship, state of origin, region of origin, marital status, school situation, school profile, type of school transport, school neighborhood, school scholarship funding, professional profile, type of transport to work, work neighborhood, income.

To make up the entry standards 250 resident profiles were selected at random. These profiles were analyzed by the environmental researchers of the project, who attributed to the SCR of each a value in the interval [1 a 10], based on their knowledge and expertise, with regard to environmental issues and the specific reality of the community. A value of "0" represents sustainable behavior with severe restrictions and a value of "10" represents optimal sustainable behavior.

After all the stages involved in the solution were concluded (section 2.2), all of the resident profiles have their respective SCR and are grouped by their characteristics of similarity. With this information, it is possible to construct the Sustainability Map, foreseen as a result of the proposed computational solution.

The Sustainability Map, represented in Figure 2, is a map of characteristics made up of 225 positions, distributed in a 15x15 matrix. That characteristic is due to the fact that the ANN II architecture calls for a neural grid with a size of 15x15. This dimension of the neural grid was defined after performance of several experiments to demonstrate the dimension that best fits the purposes of this paper. Thus, the map has 225 positions, to represent the 225 neurons of the neural grid, sequenced, from top to bottom and from left to right. Thus, in position (1, 1) we have the result of neuron 1 of the neural grid, and in position (15,1) the result of neuron 15, and so on successively.

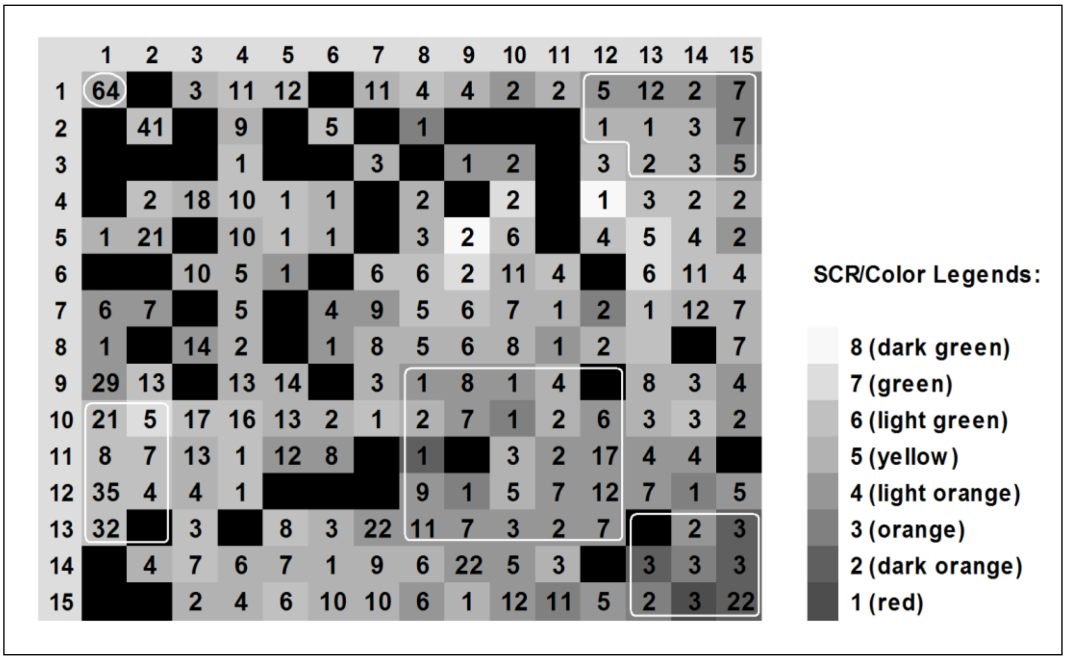

Figure 2 – Map of Sustainability Characteristics SCR.

The numerals in each position represent the quantity of residents who belong to the same cluster; in other words, they have a very close degree of similar characteristics. The colors in each position represent the distribution of the average SCR of each of the groupings, which occurs in the following manner: the positions colored red, orange and yellow represent groups that have a critical condition in sustainable behavior, that is, rates from [1 to 5], and the positions colored green in different tones represent groups with more favorable conditions of sustainability, with rates of [6 to 10]. In fact, value 8 was the highest rate found within the community.

One may observe in the Sustainability Map that the lower right corner contains the highest concentration of population with the worst rates and in the upper left corner we find the most favorable sustainability condition.

Through analyzing this map it is possible to observe the degree of sustainability not only of an isolated group, but also of regions made up of several groups. Such a visualization at a higher amplitude increases the specialist's capacity to prepare diagnoses and predictions, informing construction of strategies and actions that focus on the investigation and/or development of groups endowed with given degrees of sustainable behavior, as we shall see below:

- In position (1, 1), which is marked by a circle, one observe a group of 64 people with similar characteristics and who have a SCR of 5 (color yellow), which represents the last degree of SCR for exchanging a critical situation [1 a 5] for a more favorable situation [6 to 10]. That is the ideal situation, since it allows more specific actions to be applied in this group, to achieve an interesting quantity of individuals. Thus, it does require, at the outset, so much effort, considering that such a group has a degree of sustainability close to that of the situation of interest. If the same resources were applied in a random fashion, one could not be certain which results might be achieved and how many persons would really benefit from those actions;
- The positions which are located in the lower left, marked by a line, make up a region of persons with a high SCR, if one considers the reality of the community. It is important to understand more profoundly what characteristics lead the group to have a sustainability capacity quite superior to that of the other individuals in the community, which is feasible, since each grouping knows each one of its component individuals. Another relevant aspect is the fact that it is possible to direct actions directed to the group, in terms of maintaining the SCR of those individuals;
- The same thing may occur in the other regions demarcated in the Sustainability Characteristics Map, where groups such as the one positioned in the lower right corner are in the community's most critical region, and one which certainly needs immediate strategic actions, since they may be worsening the environmental conditions surrounding the community, thus affecting the other inhabitants. Thus, immediate recuperation of the SCR of those individuals may bring immediate benefits to the entire community.

The Sustainability Map provides an analysis of various community dimensions in a single graphic representation, at the same time in a succinct form, considering the condensation of various variables that represent the individuals, and which is broad in regard to the scope of its component scales, allowing the specialist to analyze the situation of a single individual, or of a group represented by a position on the map, or islands represented by a set of groups near each other, or even an overall view of the entire community, which is done through a spatial representation of the entire Sustainability Map.

The Sustainability Map provides the researcher with several possibilities for diagnoses and predictions based on its topography. It allows identification of most critical groups, or at least those most critical in relation to their capacity for sustainability, or those than can achieve the best conditions of sustainability with few investments, or even those that require priority for such actions. The research investigation process or extension actions become more focused, enabling one to accelerate towards the goals sought by the research project.

4 Conclusions

Utilization of the techniques of Computational Intelligence, more specifically the ANN, to aid in the process of diagnostics and forecasts with analysis of research projects regarding human intervention impacts on the environment, has made it possible to have a representation of reality as it appears, considering both the aspect that all the of variables surveyed have been employed, with none having been discarded, as well as the capacity for capturing the knowledge and expertise of the environmental specialist, through a cognitive process. This allows one to assess the profile of community residents and apply that knowledge acquired in the assessment process.

The possibility of visualizing that reality, though mapping similarities in individual characteristics, enables the environmental specialist to identify and analyze the regions made up of groupings of residents, in order to direct the process of research investigation, or to direct the application of extension actions in the community, more effectively directing available resources to improve the sustainability of certain groups of residents, or even to maintain the levels of sustainability that have already been achieved.

5 References

[1] S. Haykin, Neural Networks: A Comprehensive Foundation, 2th edition, **Prentice Hall**, (1998).
[2] J. F. Matos, Modelação Matemática, **Oficinas Gráficas Minerva do Comércio**, (1995).
[3] A. M. Silveira, M. L. C. Silva, e R. C. Limão, Evaluation of the Competitiveness of Organizations: An Essay Using Artificial Neural Networks (ANN), Frontiers in Artificial Intelligence and its Applications, **IOS Press**, (2003).
[4] S. Pfirman, Complex Environment Systems: Synthesis for Earth, Life, and Society in the 21st Century, **A report summarizing a 10-year outlook in environment research and education for the National Science Foundation**, (2003), 68.
[5] NRC, Decision Making for the Environment: Social and Behavioral Science Research Priorities. **Panel on Social and Behavioral Science Research Priorities for Environmental Decision Making**. G.D. Brewer and P.C. Stern, editors. Committee on the Human Dimensions of Global Change, Division of Behavioral and Social Sciences and Education. Washington, DC, **National Academy Press**, (2005).
[6] NAE, Measures of Environmental Performance and Ecosystem Condition, **National Academy Press**, (1999).
[7] N. B.Maia, H. L. Martos, W. Barrella, Indicadores ambientais: conceitos e aplicações. **Editora da PUC-SP**, (2001).

[8] N. Sizer, K. Miller, Criteria and Indicators for Forest ecosystem sustainability in amazônia: the international policy context and lessons from the temperate and boreal zone. In: **Propuesta de Tarapoto sobre critérios e indicadores de sostenibilidad Del bosque amazônico**, (1995), 67-81.
[9] V. Ravena-Cañete, A Descrição do Possível: a experiência de intervenção no Igarapé Mata-Fome e o levantamento de dados sócio-econômicos, **Editora Universidade da Amazônia**, (2006).
[10] J. Han, M. Kamber, Data mining: concepts and techniques, Academic Press, (2001).
[11] Matworks, **Neural Network Toolbox User's Guide: for use with MATLAB**, (2006).
[12] S. O. Rezende, Sistemas Inteligentes: Fundamentos e Aplicações, Manole, (2003).