

Mapping the Risk of Burning in the Brazilian Amazon with the Use of Logistic Regression and Fuzzy Inference

Camil Wadiah Salame · Joaquim Carlos
Barbosa Queiroz · Gilberto de Miranda Rocha ·
Mario Miguel Amin

Received: 10 April 2011 / Accepted: 10 January 2012 / Published online: 9 February 2012
© International Association for Mathematical Geosciences 2012

Abstract Through the PRODESDIGITAL Project, the National Institute for Space Research (INPE) has been mapping vegetal coverage in the Brazilian Legal Amazon using Landsat satellite images. INPE not only identifies deforested areas but also releases a daily map of burning areas. Burning is frequently used as the cheapest way to deforest, to clear areas for farming and to increase the soil's fertility in a short period of time. In many instances, these fires get out of control and end up accidentally invading areas of forest exploited by the lumber industry, agricultural plantations and pastures. However, the deforestation and burning area maps alone are insufficient for monitoring and control on a regional scale. It is necessary to know in an analytical way how actions (deforestation and burnings) that characterize human occupation are occurring in a region. In this research, two methods are used that allow the inclusion of co-variation in the estimation of variables of interest: logistic regression, which is a statistical method that considers a categorical or discrete response variable and discrete and/or continuous co-variation; and fuzzy logic, which makes use of artificial intelligence methods to incorporate into computational models information based on the knowledge or experience of a specialist. In both methods, the response variable may be related to the probability of occurrence of the environmental variable under study.

Keywords Legal Amazon · Deforestation · Burning · Fuzzy logic

C.W. Salame (✉) · J.C.B. Queiroz
Department of Statistics (ICEN), Federal University of Para, Brazil, 66075-110, Belem, Para, Brazil
e-mail: camilsalame@yahoo.com

G. de Miranda Rocha
Environment Nucleus (NUMA), Federal University of Para, Brazil, 66075-110, Belem, Para, Brazil

M.M. Amin
Nucleus for Amazon Research (NAEA), Federal University of Para, Brazil, 66075-110, Belem, Para,
Brazil

1 Introduction

The monitoring of the forest area of the Legal Amazon (encompassing the States of Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, Tocantins, Mato Grosso and the western part of Maranhão) performed by the National Institute for Space Research (INPE) revealed deforestation rates that varied between 1 and 3×10^6 ha/year in the period 1991 to 1999 and the loss of about 6×10^7 ha (more than half a million km²) of forest until 2000. This deforestation results from an Amazon occupation process occurring from the second half of the 20th century and is associated with the expansion of the agriculture margin, the construction of highway transportation systems and development poles. Environmental damage caused by this deforestation has led to significant social problems, such as foundry concentration, the instability of farmers in the agricultural field, precarious urbanization, and social conflicts followed by varying degrees of violence. The agricultural margin expansion and the deforestation in the Legal Amazon are closely connected to the context of the reorganization of Brazilian agriculture, followed by the accelerated industrialization starting in the 1950s and, more recently, by attempts to adapt Brazil to the globalizing economy. In this context, many factors may have contributed to the high rates of deforestation, among which are the public and private funds available, the population dynamics, the organization of the production systems and various physical conditions.

Due to public politics, the deforestation analysis accomplished by the PRODES Project (analog or digital) is not sufficient to provide a basis for governmental action because the results are obtained yearly and, many times, are derived from informative action, that is, after actions have already taken place. Since the data are grouped by state and published a year after a deforestation event occurs, the federal and state governments cannot anticipate dynamic changes in the use of Amazonian soil. Likewise, it is necessary to supplement the PRODES data with other initiatives that allow the Brazilian State to develop preventive actions to combat illegal deforestation activities.

Among previous research, Arima et al. (2007) implemented an explicit space model to estimate the probability of burning in the Brazilian Amazon by using logistic regression. This model was improved regarding its use of predictive variables, such as the price of beef and soy bean oil, in a conceptual model addressing pasture care and the use of fire to clear it and, by doing so, simulating the impact of paved highways, live cattle exportation and fire prevention work. The analysis showed that fire is positively related to the price of beef and soy bean oil and that the creation of new preservation areas makes up for the negative impacts of prior devastating acts. Sousa and Duarte (2006) developed and compared two mathematical models, one based on logistic regression, and the other on fuzzy group theory, to define the indications for scintigraphy test performance as a laboratory exam. There were 194 patients identified who had their calcium and parathyroid hormone levels measured at baseline and underwent scintigraphy for parasitoids performed in diagnostic laboratories in São Paulo during the period from January 2000 to December 2004. The models' performance was compared by the use of receiver operating characteristic (ROC) curves. The results showed that models had statistically significant differences in their development ($p = 0.026$). The fuzzy model was particularly useful because, unlike the logistic model, it had the ability to use the parathormone information in

an interval in which the calcium values showed little discrimination. Therefore, the mathematical model based on fuzzy group theory seemed to be more adequate than the one based on logistic regression for deciding whether to perform scintigraphy for parasitoids. Since it was the result of a methodological exercise, inferences about objective behavior may be inappropriate, as there are no representative population data.

These two models, logistic regression and fuzzy group theory, are used in the current study involving the mapping of burning risk in the state of Pará, Brazil, during the period from 1994 to 2004. The efficiency of each method is evaluated, and comparisons are made in relation to the adjustment quality of the data between the two models. The influence of variables such as vegetation, deforestation, weather and distance along the road over registered fire heat areas (response variables) in the considered period is evaluated.

2 Study Area

The Legal Amazon in Brazil is in the states of Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima, Tocantins and part of Maranhão, corresponding to an area of approximately 5 million km². In this area, the part that has a forest physiognomy occupies about 4 million km². The state of Pará (Fig. 1) occupies an area of 1,247,703 km² with a population of 6.2 million inhabitants. It is divided into 6 mesoregions identified as Marajó, Metropolis of Belém, Low Amazon, South Western Pará, North Eastern Pará and South Eastern Pará. Due to the forest, the state of Pará has very humid weather, and it is hot due to its proximity to the equator. The main highways in the state of Pará are BR010, which extends from the capital Belém to Brasília, PA150, which divides the southeast and northeast areas of Pará, PA256 and BR230, which cross the state from east to west, and BR163, which connects Santarem City in the west of Pará to Cuiabá, the capital of Mato Grosso State (Fig. 1).

3 Methodology

3.1 Logistic Regression

Logistic regression is a regression technique available for models that describes the relation among different independent variables and a binary dependent variable (Kleinbaum 1994). McCullagh and Nelder (1989) presented four functions used in the modeling of data whose variables are binary: logistic functions, probit functions, log-log complementary functions, and log-log functions. In many areas, the logistic function has been used frequently not only because its theoretical properties are simpler but also because of its simple interpretation as the odds ratio logarithm.

In the logistic regression model, $P(y_i = 1) = p_i$ and $P(y_i = 0) = 1 - p_i$ are considered the success and failure probabilities, respectively. In reality, success represents the event's occurrence and failure, its non-occurrence. When using a logistic regression model, the interest may be in the effect of a specific risk factor or in

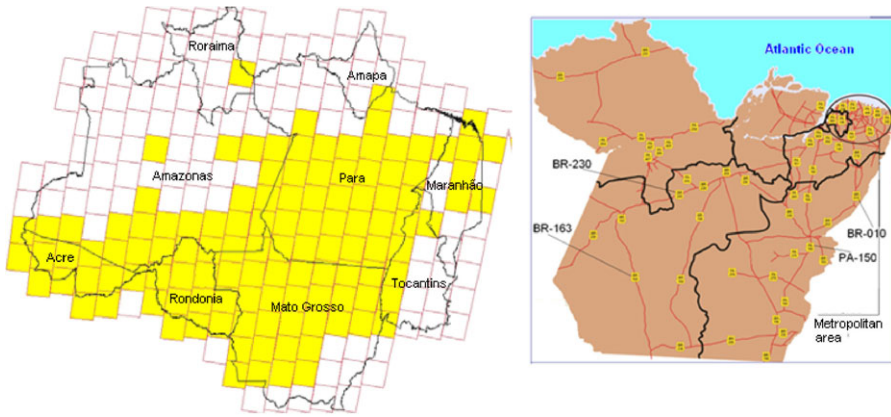


Fig. 1 Brazilian states which are part of the Legal Amazon (*left*). Images in *yellow* were taken by LANDSAT. Main highways (*right*) in the State of Pará

the identification of several factors associated with the response variable. In this research, the variables that are considered risk factors for the occurrence of burning (the response variables) are: deforestation, weather, vegetation and the distance to roads. Linear regression models are not appropriate for this type of analysis because in these models the variable value expected, i.e., the response, given a group of explanatory variables ($E(Y/X) = X'b$), assumes values of $(-\infty < E(Y/X) < +\infty)$. Unless restrictions are imposed, this violates the probability laws once $0 \leq E(Y/X) = p \leq 1$.

A simple solution to this problem is obtained by the use of a function $g(p_i)$ that traces the values in the real number field over a $[0,1]$ interval. The function $g(\pi_i)$ may be defined as

$$\ln\left(\frac{\tilde{\pi}_i}{1 - \pi_i}\right) = \alpha + \beta\mathbf{X}, \tag{1}$$

$$g(\pi_i) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \sum_{r=0}^{p-1} x_i \beta^r, \quad i = 1, \dots, n, \tag{2}$$

where

$$\pi_i = \frac{1}{1 + e^{(-\alpha - \beta\mathbf{X})}}. \tag{3}$$

Therefore,

$$y_i = \pi(x_i) + e_i, \tag{4}$$

where the error term e_i assumes only two values. If $y_i = 1$, $e_i = 1 - p(x_i)$ with a probability equal to $p(x_i)$. On the other hand, if $y_i = 0$, $e_i = p(x_i)$ with a probability equal to $1 - p(x_i)$. Thus, e_i has an average of zero and variance equal to $p(x_i)[1 - p(x_i)]$, and the distribution of the variable response is binomial with a probability $p(x_i)$. The

estimated parameters are obtained from the probability function, which is expressed by

$$L(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}. \tag{5}$$

Therefore, it is easier to maximize the $L(\beta)$ logarithm defined as

$$L(\beta) = \ln(\beta) = \sum_{i=1}^n y_i \ln \left[\frac{\pi(x_i)}{1 - \pi(x_i)} \right] + \sum_{i=1}^n \ln [1 - \pi(x_i)]. \tag{6}$$

In contrast to the linear model, the equations obtained in the maximization process require iterative methods. According to Hosmer and Lemeshow (1989), the most frequently used method is the Newton–Raphson.

In general, the likelihood ratio statistic, like an F statistic in classical multiple linear regression, requires the identification of two models to be compared, one of which is a special case of the other. The larger model is sometimes called the full model, and the smaller model is sometimes called the reduced model; that is, the reduced model is obtained by setting certain parameters in the full model equal to zero. The likelihood ratio (LR) statistic is given by (Kleinbaum 1994)

$$LR = -2 \ln \left(\frac{L_1}{L_2} \right) = -2 \ln \left(\frac{\text{Reduced model}}{\text{Full model}} \right). \tag{7}$$

3.2 Fuzzy Groups and Fuzzy Logic

The theory of fuzzy groups is, in great part, an extension of the traditional theory of groups. With the development of fuzzy logic, it became possible to apply the partial truth concept, making it possible to present any value between the extremes of completely true or completely false. Thus, a group may be formed by elements that belong there partially. In this way, for a universal group U (classic), the fuzzy subgroup F in U takes the form of a dominant pertinence function

$$\mu_F : U \rightarrow [0, 1], \tag{8}$$

where $\mu_F(x)$ represents the degree to which the element x from U belongs in the fuzzy subgroup F . This function maps the pertinence degree of element x in subgroup F , representing, for example, $\mu_F(x) = 0$ and $\mu_F(x) = 1$, the non-pertinence and the total pertinence of x in F , respectively. The fuzzy subgroup definition is formalized by enlarging the image of the characteristic function that was the group $\{0,1\}$ to the interval $[0,1]$, causing the classic group to become a particular case of the fuzzy group (Zadeh 1965; Barros and Bassanezi 2001).

The main idea of fuzzy groups is the pertinence degree (the value that indicates the degree to which an element belongs to a group). The groups are represented in a qualitative way, and elements of each group are characterized by varying the degree of attribution. For example, consider that A and B are fuzzy subgroups in the universal group U . Operations on the fuzzy groups such as union, interception and complementation also result in fuzzy groups. Following the standard operations for union,

interception and complementation for fuzzy groups defined by Zadeh, the following pertinence functions are obtained

$$\begin{aligned}\mu_{A \cup B}(x) &= \text{máximo}[\mu_A(x), \mu_B(x)], & x \in U; \\ \mu_{A \cap B}(x) &= \text{mínimo}[\mu_A(x), \mu_B(x)], & x \in U; \\ \mu_{A'}(x) &= [1 - \mu_A(x)], & x \in U.\end{aligned}\tag{9}$$

These are standard operations of a possible operator group known as t -norm for interception operators and t -conorm for union operators. If A and B are classical groups, the pertinence functions described above satisfy these equalities, showing the coherence of these definitions (Barros and Bassanezi 2001).

Using a linguistic approach, fuzzy logic becomes closer to a human's reasoning, as the results are expressed using the pertinence degree concept, and it defers from classical logic as the fuzzy propositions reach values other than true and false (Reznik 1997). Fuzzy linguistic models represent a significant portion of the applications of fuzzy group theory. The linguistic models may be understood as specialist systems that linguistically describe a complex object to be analyzed. The basis of these models is a group of IF-THEN-type rules with vague predicates. The fuzzy groups are parameters of the model that associate with the rules from its structure (Ortega 2001).

The model's basic structure includes four components: (i) an entrance fuzzy group, which is an initial stage where the entrances to the system are built by fuzzy groups and the pertinence functions are formulated; (ii) a base of fuzzy rules, considered the fuzzy linguistic model's nucleon in which knowledge is expressed through a group of rules (IF-THEN) composed of fuzzy propositions described in a linguistic manner; (iii) fuzzy inference, in which each prior fuzzy component proposition is translated mathematically by fuzzy logic techniques (in this stage, the t -norms, t -conorms and inference rules are defined and used to obtain the fuzzy relations that the model is based on); and (iv) defuzzification, which is a final process that allows a fuzzy group to be represented by a whole number. Among inference methods, the most widely used are those of Mamdani and of Takagi–Sugeno–Kang (TSK). Overall, the latter method characterizes it as function of the entrance value of each rule. If the consequent is a fuzzy group and the interference is of the min-max type, then the model is of the Mamdani-type (Ortega 2001; Barros and Bassanezi 2001).

4 Analysis and Discussion of Results

The behavior of the environmental variables describing burning occurrences and deforestation and the effect of distance from the main roads on these variables are presented in Figs. 2 and 3. The southeastern region of Pará was the mesoregion that presented the highest burning rate, with 98,965 points of fires accounting for 69.3% of the state's points, followed by the southwestern mesoregion with 23,603 points (or 16.5%). On the other end of the spectrum are the Marajo and Belém metropolitan mesoregions with burning rates equivalent to 1,249 and 84 points, respectively, both of which are less than 1% (Fig. 2, left). The highest deforestation rate, equivalent to an area of 118.62 km² or 53.34% of the state's total deforestation, also occurred in

Table 1 Distribution of deforestation in 2003 by mesoregion in the state of Pará

Mesoregion	Area (km ²)	Brute deforestation area (km ²)	Brute deforestation area (%)	Accumulated deforestation area (km ²)	Accumulated deforestation area (%)
Marajó	104,606.90	3,547.46	1.60	3,547.46	1.60
Belém Metropolitan	6,906.30	4,951.66	2.23	8,499.12	3.82
Low Amazon	341,968.30	20,262.39	9.11	28,761.50	12.93
Southwest	417,571.20	25,780.65	11.59	54,542.16	24.53
Northeast	83,553.80	49,220.97	22.13	103,763.13	46.66
Southeast	298,558.00	118,618.00	53.34	222,381.23	100.00
Pará	1,253,164.50	222,381.23	100.00		

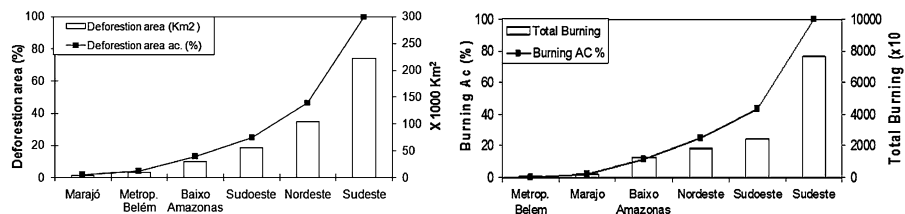


Fig. 2 Distribution of deforestation (*left*) and burning (*right*) in 2003 with respect to distance from roads in the state of Pará (Rodney 2005)

the southeast of Pará, followed by the northeast of Pará with 49.22 km² (or 22.14%). The Marajó and Belém metropolitan mesoregions presented the lowest deforestation rates of approximately 3.55 km² and 4.95 km², respectively, both of which are less than 3% of the total deforestation area (Table 1 and Fig. 3 on the right).

Figure 3 shows the distribution of deforestation and burning in the state of Pará prior to 2003 along the main highways (BRs and PAs), and indicates in a nonlinear manner that the shorter the distance to a road, the higher the quantities of burning and deforestation. Regarding deforestation, the radius of influence from roads is a maximum of 110 km; however, this value is about 150 km for the burning parameter.

The relative distribution of deforestation in Pará may be evaluated by Lorenz’s curve, which shows the deforested area distribution in relation to the total state area (sampled area) distribution (Fig. 4). The deforestation intensity distribution may be evaluated by estimating the concentration area (Fig. 4), which is given by the spatial concentration rate, defined as $I = (C - 550)/450$, where C represents the sum of percentages related to the deforested area in relation to the total area. The smaller this area is, the more uniform is the state’s deforestation distribution, and the larger this area is, the more concentrated is the deforestation in the studied area (Pará State).

In the current study, the spatial concentration rate presents a value of 0.572; therefore, the deforestation concentration is about 57.2% in the state of Pará. Figure 5 shows how the distribution of deforestation in Pará State is very irregular, with a higher concentration existing next to the BR-010 and PA-I50 highways located in

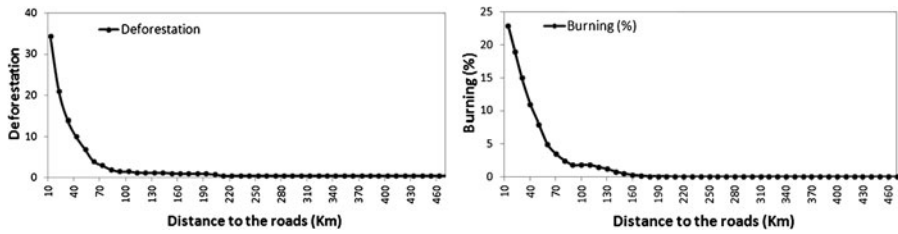


Fig. 3 Distribution of deforestation (*left*) and fire (*right*) with respect to distance from main highways (BRs and PAs) in the state of Pará from 1996 to 2003 (Rodney 2005)

Fig. 4 Deforestation distribution using Lorenz’s curve (Rodney 2005)

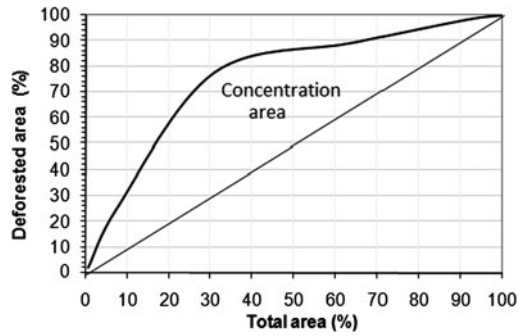
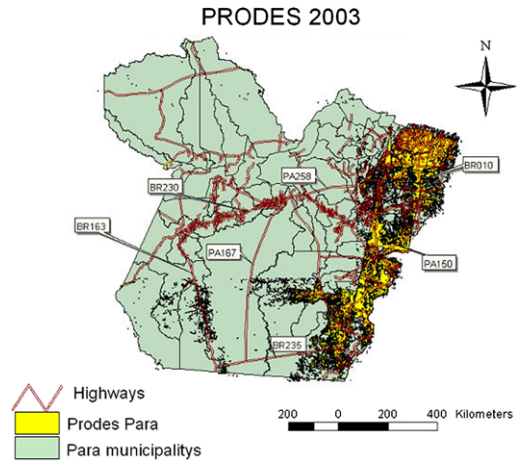


Fig. 5 PRODES Project images of deforestation in 2003



the mesoregions (southeast, northeast and Belém Metropolitan) with the most infrastructure in the state.

4.1 Mapping with the Use of Logistic Regression

The variables in the logistic model are burning occurrences (response variable) due to deforestation, distance from roads, weather and vegetation. The results presented are

Table 2 Variable codification

Value	Burning	Deforestation	Vegetation	Weather
0	Non-burnt area	Non-deforested area	Valued	Dry
1	Burnt area	Deforested area	Not valued	Transition
2				Humid

Table 3 Categorization of the variable Weather

Weather	C1	C2
0 (Dry)	1	0
1 (Humid)	0	1
2 (Transition)	0	0

Table 4 Logistic regression results

Variable	Coef.	Z	p-value	Odds ratio	95% CI	
					Min	Max
Constant	1	1.26	0.21			
C1	1.07	8.23	0.00	2.91	2.26	3.76
C2	0.32	13.05	0.00	1.38	1.32	1.45
Distance	-0.02	-65.74	0.00	0.99	0.98	0.99
Vegetation	0.19	5.13	0.00	1.20	1.12	1.29
Deforest	2.08	59.66	0.00	8.01	7.48	8.57

for fire sites that occurred in the period June 1, 1998, to December 31, 2003. Tables 2 and 3 show the codification used in the logistic model for the categorical variables.

The statistical program *MINITAB* is used to adjust the logistic regression model (Table 4). The significance of the variables in the model is evaluated at the descriptive level (p -values). All parameters are significant (p -value = 0.00). The odds ratio allows us to conclude that

- (i) For the variable *DEFOREST*, the chance of an area burning is about 8 times higher if the area is deforested (in relation to a non-deforested area).
- (ii) For the variable *VEGETATION*, the chance of an area burning is about 1.2 times higher if the vegetation does not have lumber value (in relation to the vegetation with lumber value).
- (iii) For the variable *WEATHER*, the chance of an area burning is about 2.91 times higher if the weather is dry (in relation to not being dry) and about 1.38 times lower if the weather is humid (in relation to not being humid).

As expected, the statistics $LR = 19284.413$ and p -value = 0.00 indicate that there is enough evidence that at least one of the coefficients is different from zero. Thus, considering what is revealed above, the logistic model adjusted to the data is

Table 5 Rate of burning presence probability

Variable	Value				
Weather	0	0	1	1	2
Veget	1	0	1	0	0
Deforest	1	0	1	1	1
Logit	33.198	31.347	25.737	23.886	20.654
<i>P</i>	0.9651	0.9583	0.9292	0.9160	0.8875

$$\text{Prob}(\text{Burning} = 1) = \frac{e^{(2.08 * \text{Deforest} - 0.051 * \text{Dist} + 1.069 * C1 + 0.032 * C2 + 0.185 * \text{Veget})}}{1 + e^{(2.08 * \text{Deforest} - 0.051 * \text{Dist} + 1.069 * C1 + 0.032 * C2 + 0.185 * \text{Veget})}} \quad (10)$$

From the adjusted model, the probabilities of burning due to the presence of deforestation, vegetation and weather may be determined as

$$\begin{aligned} \text{Logit } P_i &= 2.08 * \text{Deforest} - 0.015 * \text{Dist} + 1.069 * C1 \\ &\quad + 0.032 * C2 + 0.185 * \text{Veg}, \quad (11) \\ P_i &= \frac{1}{1 + e^{-\text{Logit } P_i}}. \end{aligned}$$

In the analysis of the probability of fire occurrence in the state of Pará, the combinations that most stand out may be observed in Table 5. A higher burning probability of 96.51% occurs when a region is characterized by deforestation, dry weather and vegetation with lumber value. The second highest probability of 92.92% represents a region with transitional weather, vegetation with lumber value and deforestation. For a deforested region with dry weather and vegetation without lumber value, the probability is 95.83%. A probability of 91.60% represents an area with no transitional weather, vegetation with lumber value and deforestation. Finally, a region with a chance of burning of 88.75% is characterized by deforestation, humid weather and vegetation with no lumber value.

The results of the logistic regression related to the probability of burning occurrence were mapped and imported to a Geographic Information System program (ARCVIEW) in which a map of burning risk in the state of Pará was assembled (Fig. 6).

An interval of 0.25 is considered to present the results. Thus, four levels of burning risk are defined, as shown in Table 6. It is revealed that about 20% (250,633 km²) of Pará is at a low risk of being burnt due to a lack of infrastructure (roads). The areas that have regular road infrastructure, BR-230 (Transamazon) and BR-163 (Cuiabá-Santarém) account for 26% (326,000 km²) of the state and are evaluated as having a medium risk of burning. The areas that have good road infrastructure (BR-010, BR-316 and PA-150) are at a high (29%) and very high (25%) risk of being burnt. These areas add up to 54%, which corresponds to 677,000 km² of the state of Pará.

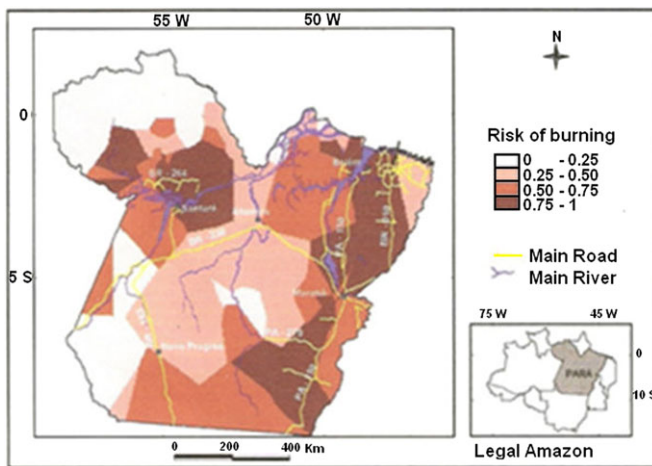


Fig. 6 Risk of burning in the state of Pará from 1998 to 2003 as determined by logistic regression (Rodney 2005)

Table 6 Probability of burning in the state of Pará

Risk Class	Interval	Area (km ²)	%
Low	0–0.25	250,632.90	20
Medium	0.25–0.50	325,822.80	26
High	0.50–0.75	363,417.70	29
Very High	0.75–1	313,291.10	25
Total		1,253,164.50	100

4.2 Mapping with the Use of Fuzzy Logic

Considering that variables such as weather, vegetation and deforestation fail to form a well-defined geographical border, fuzzy logic was chosen for the numerical treatment of these variables based on the knowledge acquired with the logistic regression method. Subsequently, the response variable (burning) is mapped based on the identified fuzzy model. The model uses a fuzzy inference method (Takagi and Sugeno 1985) involving a group of IF-THEN-type rules, with the additional characteristic of allowing the use of previous knowledge of the rules. Constants (singletons) are used as consequences of each rule. The consequence (then) of each rule is obtained by the fuzzy intersection (*t*-norm, using the minimum) of the entrance variables, and its value represents the degree of the rule's activation. The final exit to the model is a whole number, a weighted-average of the rules according to the activation's degree (with values between 0 and 1), and the pertinence degree of the consequent situation derived from the established conditions.

The fuzzy groups for both the entrance variables (weather, vegetation, deforestation) and the exit variables (burning) are linguistic expressions. In this case, the linguistic categories established for the entrance variables were: dry, transition and hu-

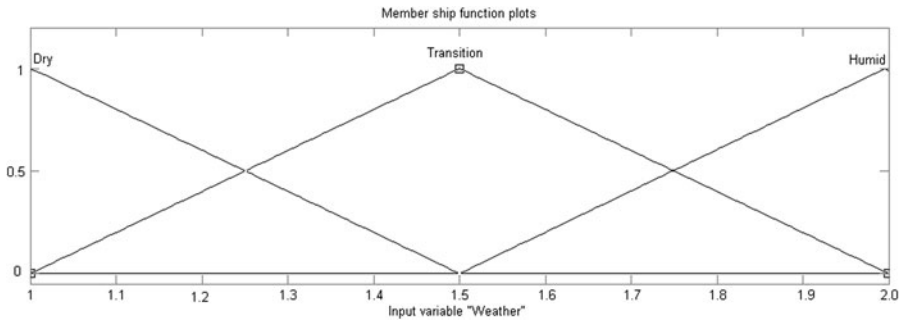


Fig. 7 Categories used for the entrance variable Weather

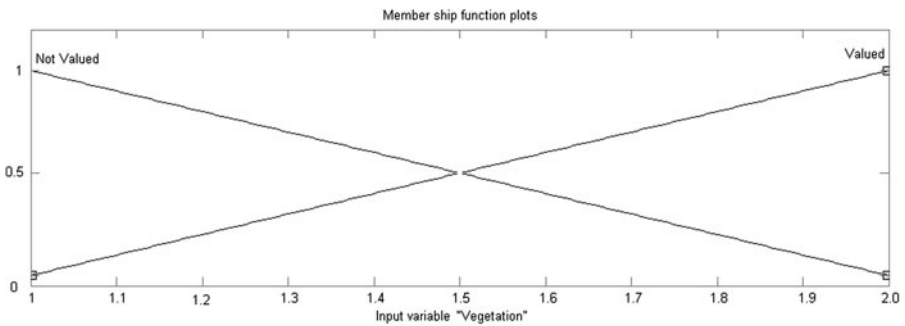


Fig. 8 Categories used for the entrance variable Vegetation

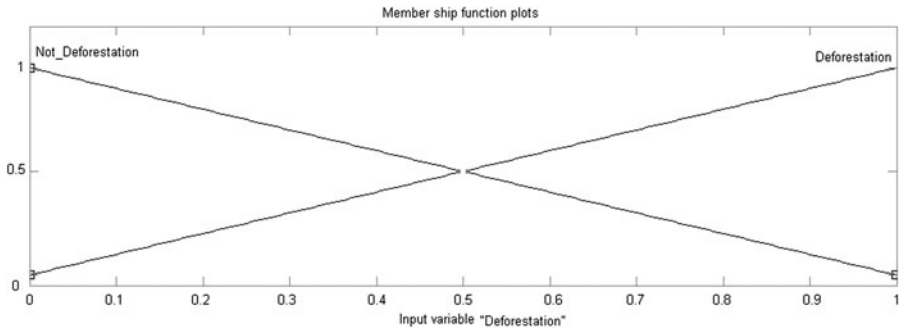


Fig. 9 Categories used for the entrance variable Deforestation

mid for the weather variable; with wood value and without wood value for the vegetation variable; and deforested and not deforested for the deforestation variable. For the exit variable (burning), five linguistic burning risk categories were defined: very low, medium low, high and very high, to which were attributed singular pertinence degree values (singletons) of 0.0, 0.2, 0.4, 0.6, 0.8 and 1.0, respectively. These singular values are the translation of what is taken to be numerically interpreted by a

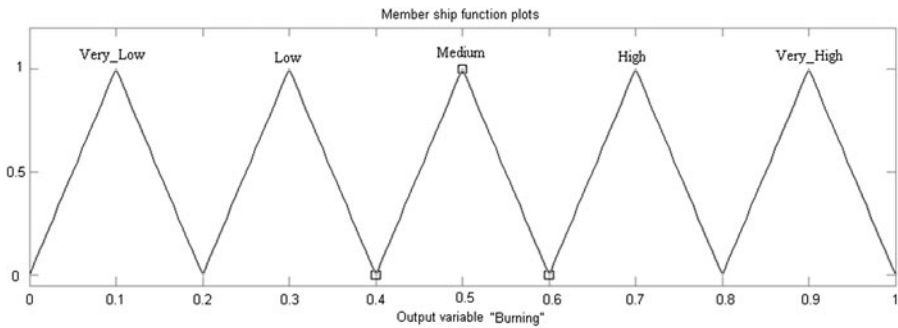


Fig. 10 Categories used for the response variable Burning

linguistic expression. Figures 7, 8, 9, 10 show the categories and pertinent functions (triangular functions are used) for each variable.

A group of IF-THEN-type rules derived from entering established empirical knowledge from consulting projects about the burning situation into the entrance and exit variables are given below: (i) If (weather is dry) and (vegetation is valued) and (deforest is deforest) then (burning is very high); (ii) If (weather is dry) and (vegetation is valued) and (deforest is without deforestation) then (burning is high); (iii) If (weather is dry) and (vegetation is not valued) and (deforest is deforest) then (burning is very high); (iv) If (weather is dry) and (vegetation is not valued) and (deforest is without deforestation) then (burning is high); (v) If (weather is humid) and (vegetation is valued) and (deforest is deforest) then (burning is very high); (vi) If (weather is humid) and (vegetation is valued) and (deforest is without deforestation) then (burning is high); (vii) If (weather is humid) and (vegetation is not valued) and (deforest is deforest) then (burning is very high); (viii) If (weather is humid) and (vegetation is not valued) and (deforest is without deforestation) then (burning is medium); (ix) If (weather is transition) and (vegetation is valued) and (deforest is deforest) then (burning is very high); (x) If (weather is transition) and (vegetation is valued) and (deforest is without deforestation) then (burning is medium); (xi) If (weather is transition) and (vegetation is not valued) and (deforest is deforest) then (burning is very high); (xii) If (weather is transition) and (vegetation is not valued) and (deforest is without deforestation) then (burning is low).

Figure 11 shows schematically the functioning of the fuzzy model. In this case, it is shown by means of values suggested for the case of dry weather (with the value of 1.04), vegetation (with a high lumber value of 1.99) and a deforested area of about 50%. The final value obtained for burning risk is 0.752. This value, which is the average weighted by the activation degree of the exits in each rule, is indicative of the category HIGH, suggesting a high risk of burning in this situation.

With the values acquired for burning risk from the many values of the entrance variables, risk maps can be created as long as the variables are geo-referred. The resulting map is presented in Fig. 12. There is a degree of similarity to the map obtained with the use of logistic regression. However, the map based on fuzzy logic seems to better define the areas of burning risk. On this map, the highest burning risk follows the main roads such as PA-150, BR-320 and BR-010. In Fig. 13, the main

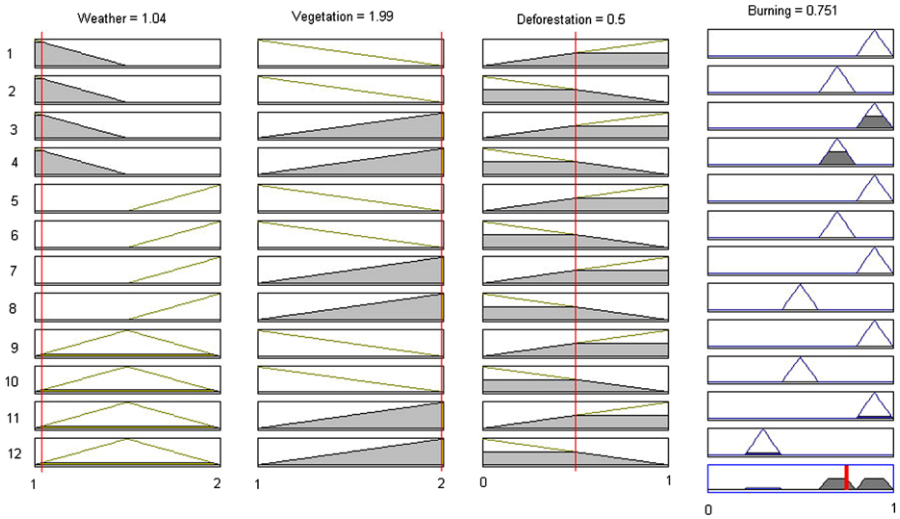


Fig. 11 Rule viewer that simulates the entire fuzzy inference process for specific values of the input variables

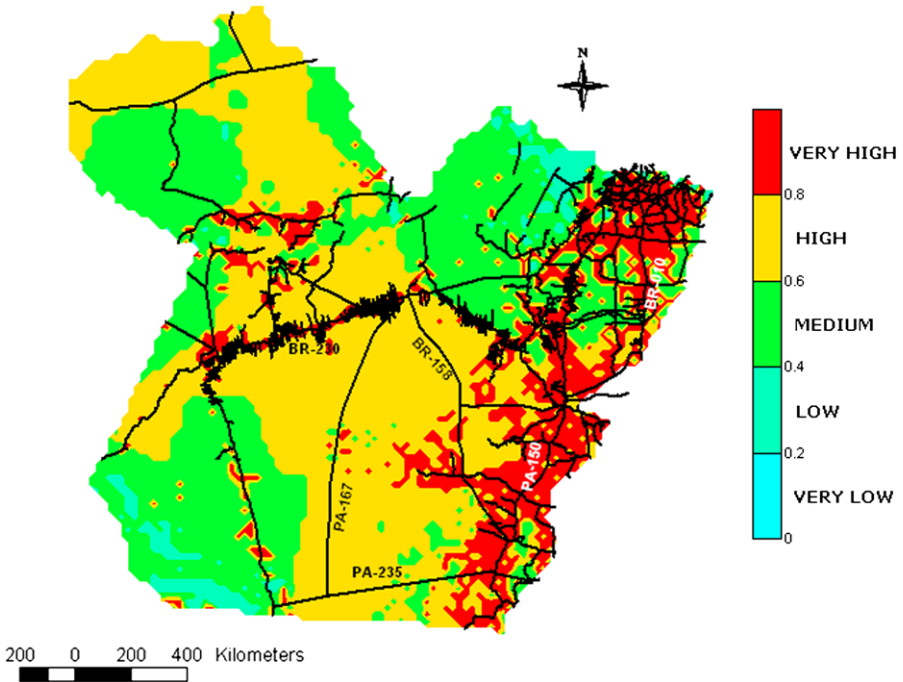


Fig. 12 Map of burning risk in the state of Pará from 1998 to 2003 based on fuzzy logic

lumber areas in the state of Pará are shown. A comparison of the maps in Figs. 12 and 13 reveals that the areas with a greater burning risk are located very close to the

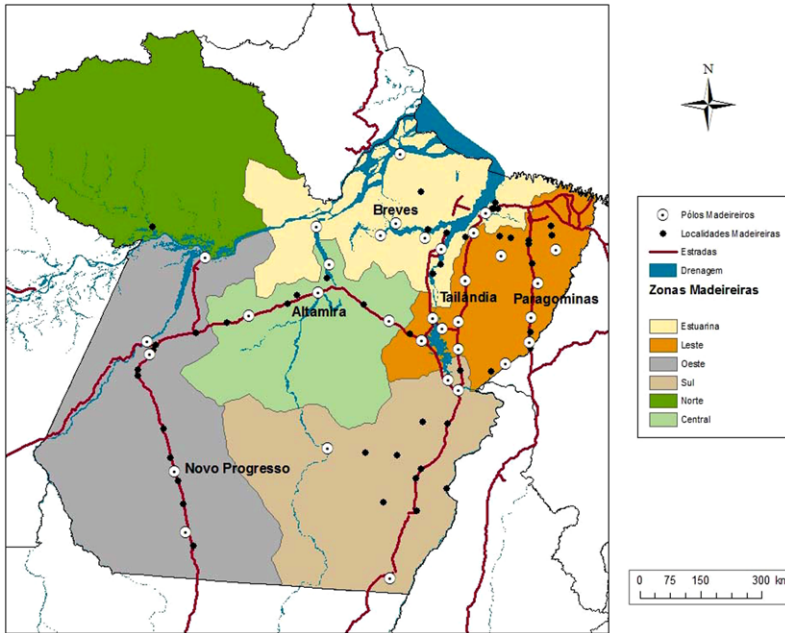


Fig. 13 Lumber areas in the state of Pará in 2007 (INPE)

main lumber zones in southern and northeast Pará. The regions with a lower burning risk are located, in general, in preserved areas, Indian areas and military areas shown in Fig. 14.

5 Conclusion

The methodology used in this assignment shows that burning occurrences are influenced by factors such as weather, vegetation and deforestation. The two methods used to map burning risk show similar results. Specifically, the map built with the use of fuzzy logic yields somewhat better results than the map created by logistic regression. These results are important for initiating public actions. First, the areas with high burning risk should be inspected to reduce the number of potential fires. The lowest rates of burning are associated with a low density of roads. In other cases, they are associated with areas in which the use of fire is forbidden, that is, in protected areas (Indian reservations, integral protection and sustainable use) and low-access infrastructure (roads). Second, the substitution of fire for fertilizer and machinery, that is, agriculture intensification, has the potential of inducing fires. The current deforested areas match the mesoregions that have good infrastructural access. These regions must be technological investment targets for farming.

The construction of maps using these methods significantly helps revealing possible spatial relations between burning or deforestation sites and Indian territory and/or lumbering zones, for example. Maps acquired through different techniques show that the results are consistent with reality.



Fig. 14 Indian territory in the state of Pará in 2003 (IBGE)

Acknowledgements The authors thank FAPESPA (Fundação de Pesquisa do Estado do Pará, Brazil) for funding this research.

References

- Arima EY, Simmons CS, Walker RT et al (2007) Fire in the Brazilian Amazon: A spatially explicit model for policy impact analysis. *J Reg Sci* 47:541–567
- Barros LC, Bassanezi RC (2001) Introdução à teoria fuzzy—aplicações em biomatemática. IMECC-UNICAMP, Campinas, São Paulo
- Hosmer DW, Lemeshow S (1989) Applied logistic regression. Wiley series in probability and statistics. Wiley, New York
- Kleinbaum DG (1994) Logistic regression: A self-learning text. Springer, New York
- McCullagh P, Nelder JA (1989) Generalized linear models. Chapman and Hall, London
- Ortega NRS (2001) Aplicação da teoria de lógica fuzzy a problemas de biomedicina. Tese de Doutorado, Programa de Física, USP, São Paulo
- Reznik (1997) Fuzzy controllers: How to design them, how they work
- Rodney SR (2005) Avaliação da distribuição do desmatamento em queimadas no estado do Pará. Monografia aperfeiçoamento/especialização em estatística. Universidade Federal do Pará, Belém
- Sousa CA, Duarte PS (2006) Lógica fuzzy e regressão logística na decisão para prática de cintilografia das paratiróides. *Rev Saúde Públ* 40:898–906
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modelling and control. *IEEE Trans Syst Man Cybern* 15:116–132
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8:338–353