Spatial modeling of the Amazon deforestation

Ricardo Alexandrino Garcia

Professor of Economic Geography and Researcher at Center for Remote Sensing Federal University of Minas Gerais (UFMG) alexandrinogarcia@gmail.com

Britaldo Silveira Soares Filho

Professor of Cartography and Researcher at Center for Remote Sensing Federal University of Minas Gerais (UFMG) britaldo@csr.ufmg.br

Abstract. The greatest challenge in establishing a function of socioeconomic development for an environmental degradation potential, as in the specific case of deforestation, is in obtaining measures for a group of variables that give some indication of the dependent variable behavior in a the future. This paper aims to explore, in a spatial context, the variables that better explain the Amazon deforestation. We reach a parsimonious model, capable of showing the environmental degradation behavior of that area in the future. Based on the statistical exploration of the variables that are directly or indirectly involved in the process of Amazon deforestation, the main objective of this paper is to elaborate classic and spatial lineal models that establish functional relationships between the variables and the percentage of the deforested areas. The estimated models are intended for subsidizing the simulation of future scenerios, integrating socioeconomic variations and deforestation impact levels.

Keywords: deforestation; Amazon; spatial models.

I - Introduction

The greatest challenge in establishing a function of socioeconomic development for an environmental degradation potential, as in the specific case of the deforestation, consists in obtaining measures for a group of variables that give some indication of the dependent variable behavior in a next future. With this purpose, stocks of socioeconomic variables, in time t , are related with deforestation data in a time close to t . In this case, the aim is, besides to understand the past, to obtain an indication of the path of the analyzed phenomenon – the deforestation – as a function of the socioeconomic conjunctures. A way suggested for that consists in considering the time that a stock of a certain socioeconomic conjuncture takes to produce an extension of changes, and in using this relationship to calculate the elapsed time until a certain area will need adjustments, in terms of the relationship between pressure and impact.

In this sense, it's advisable to relate stocks densities – calculated, for instance, as a function of the municipal area – with the extension of the environmental degradation (in the specific case, the density of deforested areas). In terms of environmental degradation, the deforestation is still concentrated along an area known as "deforestation arch" (Maranhão, eastern Pará, Mato Grosso, Rondônia and eastern Acre) and along some areas of Amazonas river margin. There is also a spreading potential of that process in some areas of Pará, along BR-163 road (which will be asphalted), for all the remaining of Mato Grosso and part of Rondônia. Thus, this paper aims to explore, in a spatial context, the variables that better explain the Amazon deforestation, reaching a parsimonious model, capable to show the environmental degradation behavior of that area, in a near future.

Based on the statistical exploration of the variables that are directly or indirectly involved in the process of Amazon deforestation, the main objective of this paper is to elaborate classical and spatial lineal models that establish functional relationships between the variables and the percentage of the deforested areas. The estimated models are intended for subsidizing the simulation of future sceneries, integrating socioeconomic variations and deforestation impact levels.

Basically, the methodology consists in: (i) statistical exploration of economical and demographic variables, directly or indirectly involved in the process of Amazon deforestation; (ii) elaboration of mathematical transformations in the selected variables, in way to obtain the best adjustment for each selected variable and its respective dependent variables; (iii) estimation of classical linear models that establish functional relationships between the variables and the deforestation; (iv) analyze the residues of those models to find possible outliers and estimate new models (in case there is need of heterocedastic control); (v) application of auto-correlation spatial tests in the previous models; and (vi) adjusting spatial models, in the case of presence of spatial autocorrelation.

II - Methodology

II.1 – Linear Regression

The main objective of a linear regression is to estimate the function that best describes the relationship between a dependent variable and a set of other explicative variables, as expressed in Equation 1. Equation 1

$$
y = X\beta + \varepsilon ,
$$

in which y is the dependent variable (a N lines vector), X is a matrix that contains K co-variables (N lines and K columns), β is the K variables regression coefficients vector, and ε is an error term.

To estimate the linear regression coefficients we used the ordinary least square (OLS) procedure. In matrix notation, vector β can be estimated according to b $=$ $(X'X)^{-1}X'V$. There is also the possibility to estimate the linear regression coefficients by other methods. In both cases, one must presuppose a normal distribution of errors, zero mean and constant variance. Equation 2

$$
\varepsilon_i \sim N(0, \sigma^2)
$$

The maximum likelihood (ML) method is not really necessary when one needs to estimate the coefficients of a "classic" linear regression because the generated parameters would be similar to the OLS parameters. Nevertheless, when one wants to compare classical and spatial models, it is necessary to analyze measures that the OLS method does not allow, as the information criterion. The most common are Akaike (AIC) and Schwartz (SC).

Once errors have a normal distribution, zero mean and constant variance, the ML method uses a density probability function to estimate the parameters σ^2 and β of the regression equation. The point is to maximize the likelihood function (L) logarithm, to express it as shown in Equation 3: Equation 3

$$
\ln L = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i} (y_i - \hat{y}_i)^2
$$

where $\hat{y} = X\beta$.

The main factors that induce to regression model specification errors are the multicolinearity, errors no-normal distribution, heteroskedasticity and spatial dependency.

Although there is no specific multi co linearity test, there are some diagnoses that may indicate potential problems. Condition number is one of them and values higher than 20 or 30 may be considered as suspicious results (Anselin, 1992).

Heteroskedasticity occurs when errors variance are not constant, what affects the model specification and its measure adjust. Two examples are Lagrange test, developed by Breusch & Pagan (BP) and Koenker & Basset (KB) test.

When errors distribution is not normal, KB test is more consistent for small samples.

In general, residue analysis identifies outliers. The outliers can damage the dependent variable modeling, as they affect the linear model coefficients significance and the correlation indices $(R$ and R^2), indicators of model adjustment. The literature indicates a series of procedures and techniques to deal this kind of problem (Werkema & Aguiar, 1996). The most common procedures are: (i) to take outliers off the model; (ii) to introduce two specification variables (dummies) in the model, one for positive outliers and other for negative outliers. This exercise chose the second one.

II.2 – Spatial Regression

Spatial dependence occurs when the dependent variable or the error, in a locality, is correlated to the dependent variable or the errors of other localities. Several tests to spatial dependency diagnoses confirmation are: Moran's I (error); Lagrange Multiplier (error); Robust LM (error); Lagrange Multiplier (lag); Robust LM (lag); Lagrange Multiplier (error and lag) (Anselin, 1992). The lag spatial model is represented by Equation 4. Equation 4

$$
y = \rho W y + X \beta + \varepsilon,
$$

where ρ is the autoregressive coeficient; W is the spatial weight matrix; y is the dependent variable (N lines vector); X is co-variables information matrix; β is the regression coeficients and is ε a random error term.

W can be understood as the representation of a phenomenon spatial interaction. In a binary matrix, unit i is unit j 's neighbor if the spatial weight matrix cell, a_{ii} , is equal to 1. When the matrix is normalized, the sum in each line is equal to 1. The choice for the weight matrix depends on the investigated phenomenon spatial structure. In general, the chose matrix is the first order neighborhood when it is expected that the phenomenon spatial structure is circumscribed to a certain locality and its immediate neighbors. In spite of the Amazon geographic space heterogeneity and the peculiarity of the deforestation process, we believe that this kind of matrix is more adequate.

As it is represented, the equation that describes the Lag Spatial model works upon dependent variable y, but a more accurate analysis will reveal that it is not exactly like this. The expanded form of that Equation (8) shows that the model works upon the dependent variable as well as upon the random effects (residues) (Anselin, 2002).

Equation 5

y = ρ Wy+X β + ε

Equation 6

 y - ρ Wy = X β + ε

Equation 7

 $(I-\rho W)y = X\beta + \varepsilon$

Equation 8

$$
y = (I - \rho W)^{-1} X \beta + (I - \rho W)^{-1} \varepsilon
$$

The second form of spatial autocorrelation in a regression model is related with the errors terms. In this case, they are called spatial error. This spatial dependence can be expressed by spatial models for the error terms, if they are autoregressive or moving averages. An autoregressive model can be described as:

Equation 9

$$
y = X\beta + \varepsilon
$$

$$
\varepsilon = \lambda W \varepsilon + \xi
$$

where $W_{\mathcal{E}}$ is the spatial lag matrix of the error terms, λ is the autoregressive coefficient and ξ is the non-biased error term.

A movable average model for the error terms can be expressed as: Equation 10

$$
\varepsilon = \lambda W \xi + \xi
$$

The calculation of a model whose dependent variable presents strong spatial autocorrelation can be made in two different ways. The first uses instrumental variables, with the employment of the two steps ordinary minimum square method (2SLS). The second is based on the maximization of the likelihood function (ML), under the presupposition of normal distribution. The first, however, has the advantage of being flexible to the presupposition of normal residues distribution.

The 2SLS method, in spite of being flexible to the presupposition of residues normality, presents R2 as the only adjustment indication parameter of the space modeling statistics. This is also the only parameter that allow the comparison between the spatial and classic methods.

The spatial lag modeling obtained by the ML method presents a larger number of comparison parameters between the spatial models and the classic models., but is subjected to a larger series of factors than imply in estimate biases.

The option to apply both methods in the estimate of the spatial lag models of the dependent variables was an attempt to make the Amazon municipal districts deforestation adjustment models safer and to allow comparability among the several modeling used in this exercise. The literature points out that the 2SLS method is more robust than ML. Once it produces good results, there will be a tendency that the second will fit better. (Anselin, 1992). The lag spatial models obtained by 2SLS method will be, therefore, presented first.

III - Results

Given the temporary nature of the selected variables, we decided to work with three different models: the first has as dependent variable the percentage of deforested area (from the original forest area), by municipal district, in 1997; in the second, the dependent variable is the percentage of deforested area (from the original forest area), by municipal district, in 2001; and, in the third, the

dependent variable is difference between the percentage of deforested area in 2001 and 1997, for municipal district.

From the 792 Amazon municipal districts, only 630 present data about the deforested areas stocks, either because they don't possess expressive original forest, or because Prodes [\(http://www.obt.inpe.br/prodes/](http://www.obt.inpe.br/prodes/)) only has information about equatorial forests, leaving out the municipal districts of the savannah area. In the specific case of the deforested area growth model between 1997 and 2001, we opted to work only with those that growth was different from zero, what represented 399 municipal districts.

III.1 – Modeling Amazon deforested areas up to 1997

The 1997 deforest area model variables.

The first step to model the deforested area percentage in the Amazon municipal districts was to make an exploratory analysis of the available variables. The exploratory analysis consisted in the application of hierarchical groupings analysis methods, in the analysis of the main components and in the analysis of lineal regression (stepwise). Table 1 lists the variables that were explored to compose the 1997 model.

Table 1: BRAZILIAN AMAZON – 1997 AMAZON DISTRICTS DEFOREST AREA MODEL (1997): EXPLORED VARIABLES.

Sources: 1 and 2, INPE (2002); 3 e 4, IBGE - PPM, 1997; 5 to 8, IBGE - PAM, 1997; 9 to 12 and 16, IBGE – 1996 Populational Counting; 13 to 14, IBGE – 1995/1996 Censo Agropecuário and e 1996 Populational Counting; 17 to 22, Garcia et al, 2004; 23 to 26, Andrade & Serra, 1999; 27, INPE (2002).

After analysis of the variables related to the stock of the Amazon deforestation translated by the percentage of deforestation, in 1997 – we determined six explanatory variables. Table 2 describes some of the statistical parameters behavior.

Table 2: BRAZILIAN AMAZONIAN – 1997 AMAZON DISTRICTS DEFOREST AREA MODEL – SELECTED VARIABLES – DESCRIPTIVE STATISTICS.

Descriptive Statistics		Std			Correlação
		Mean Deviation Min Max			Index (R)
1997 Deforest area (%)	46.76	34.57 0.07		100	1.00000
Mean distance to asphalted road	61.6	98.9	0.7	702	-0.54535
Cattle by Km ² in 1997	22.3	27.1		241	0.54572
Government Index in 1996/2000	0.35	0.41			-0.00372
Populational Density in 1996	20.2	95.7		0.1 1782	0.09158
Tilled ground in 1997 (%)	3.32	4.9		50	0.35272
Agriculture Value by Km ² in 1997	1.8	3.1	n	24	0.30645

Mathematical adjustment of the model variables.

Before we started the classical modeling, we opted to work those variables mathematically, searching for the best adjustment between them and the dependent variables. For that, the package LAB Fit Curves Adjusts (Silva and Silva, 2003) was used. The procedure consisted in finding the best two parameters non-linear function that would be capable to produce the best adjustment between independent and dependent variables. In other words, we searched the best mathematical smoothing procedure for the independent variables. Table 3 summarizes the fitting information of the three proposed models, for each independent variable.

Mathematical adjustment	Function name	Algebric Expression	Parameters Correlation Correlação B A		New Index (R)	Previous Index (R)
1997 Deforest Area (%)						
Mean distance to asphalted road	Modified power	$Y = A*B^{**}X$	0.869	0.982	0.750	-0.545
Cattle by Km ² in 1997	Modified Log	$Y = A^*Ln(X+B)$	0.188	1.142	0.675	0.546
Government Index in 1996/2000	Linear				-0.004	-0.004
Populational Density in 1996	Hyperbole Inverse	$Y = X/(A+B^*X)$	4.524	1.109	0.690	0.092
Tilled ground in 1997 (%)	Hyperbole Inverse	$Y = X/(A+B^*X)$	0.011	1.231	0.675	0.353
Agriculture Value by Km ² in 1997	Hyperbole Inverse	$Y = X/(A+B^*X)$	0.424	1.341	0.541	0.306

Table 3: MATHEMATICAL ADJUSTMENT OF MODEL VARIABLES - 1997.

The values of the independent variables were corrected based on the non-linear relationship between them and their respective dependent variables. The linear correlation indexes got better in almost all the cases, as shown on Table 3. The

better results were due to the smoothing mathematical transformation in the original data. The next step was the deforestation models adjustment.

Deforested area "OLS" Model

Table 4 brings the information regarding the linear modeling, based on seven variables, selected to explain the fraction the original forest that was deforest up to 1997. The model was well identified, with linear correlation index among the predicted results (\hat{y}) and the observed values of the density of deforestation (y) of 0.84. The result suggests that the model explains, approximately, 72% of the random error, ε, total variance.

Table 4: DEFOREST AREAS "OLS" MODEL RESULTS - 1997.

 Regression Statistics R multiple R-Square R-square adjusted N 0.845 0.7144 0.7117 630

The residues analysis reveals, however, that 6% of the 630 municipal districts presented residues superior to two pattern deviations, and were characterized as outliers - differing residues, with values superior to two deviations pattern. The residues spatial distribution, generated by the model, can be observed in Map 1. A visual analyzes suggests the strong presence of spatial correlation, what will confirmed further.

Map 1. BRAZILIAN AMAZON – RESIDUES SPATIAL DISTRIBUTION – OLS MODEL, 1997. Source: IBGE: Municipalities digital layer - 2000.

The information regarding the Amazon municipal districts deforestation modeling, with the introduction of the outliers identification variables (heteroskedastic control) are described in Table 5. The outlier indicative variables addition improved results of R and R^2 in linear models. In this case, the R^2 value increased considerably, from 0,72 (Table 4) to 0,82 (Table 5). Besides that, one can observe the improvement of all other variables significance coefficients.

Table 5: OUTLIERS CONTROL "OLS" MODEL RESULTS: 1997.

Spatial Lag Model - 2SLS

Box 1 summarizes results of the tests that were applied to the percentage deforest areas "OLS" model, for the 630 Amazon municipal districts. Results suggest that both effects may be present: spatial dependence of the deforested area percentage and spatial dependence of the model residues. In this case, one decided for the spatial lag model, due to the external effects, commented in the methodological section.

Box 1: OUTLIERS CONTROL OLS MODEL SPATIAL DEPENDENCE TESTS: 1997.

Table 6 displays the information about the 1997 Amazon municipal districts deforestation lag spatial model, adjusted by the 2SLS method. The model was well identified and presented a high linear correlation index (0.92), what suggests that the model explains approximately 86% of the total variance of the random errors, ε. The comparison between those results and the classical modeling results reveal that there was an increase in the R^2 value (from 0,82 – Table $4 -$ to 0,85 – Table 5). The spatial model also presented improvements in the coefficients levels of significance of all other variables. The spatial correlation coefficient significance high level suggests the confirmation of the dependent variable autocorrelation.

Table 6: LAG SPATIAL MODEL (2SLS) RESULTS: 1997.

Regression Statistics

R multiple R-square R-adjuted Square N

0.922 0.851 - 630

Spatial Lag Model – Maximum likelihood

The information about to the 1997 Amazon municipal deforestation lag spatial model, adjusted by the ML method, are shown in the Table 7. The results show a high lineal correlation index (0,93), what suggests that the model explains, approximately, 87% of the random error total variance, ε. The comparison between the ML and the 2SLS models reveal that R2 increased from 0,85 (Table 6) to 0,87 (Table 7). The spatial correlation coefficient significance high level suggests the confirmation of the dependent variable autocorrelation.

Table 7: MAXIMUM LIKELIHOOD LAG SPATIAL MODEL RESULTS: 1997.

III.2 – Modeling Amazon deforested areas in 2000

The 2000 deforest areas model variables

Table 8 lists the variables that were explored to compose the 2001 model. The incorporation of the 2000 Brazilian Demographic Census allowed the exploitation of a larger range of variables. The number of selected variables was, however, almost the same of the previous model – 7 variables.

Table 8: BRAZILIAN AMAZON – 2001 AMAZON DISTRICTS DEFOREST AREA MODEL (1997): EXPLORED VARIABLES.

Sources: 1 and 2, INPE (2002); 3, IBGE - PPM, 2001; 4 and 5, IBGE - PAM, 2001; 8 – Agricultural Census, 1995/1996 and 2000 Demographic Census; 6, 7 and 9 to 23, 2000 Demographic Census; 24 and 25, INPE (2002); 26 TO 31, Garcia et al, 2004.

The comparison between the data of Tables 2 and 9 reveals that there was considerable increase of the deforested Amazon area from 1997 to 2001. In 1997, the average deforestation for each Km^2 was around 46,7% (Table 2) and in 2001 it was around 49,3%. Results indicate that the municipal average deforestation annual growth rate was around 1,32%.

Table 9: BRAZILIAN AMAZONIAN – 2001 AMAZON DISTRICTS DEFOREST AREA MODEL – SELECTED VARIABLES – DESCRIPTIVE STATISTICS.

The variables mathematical adjustment

With the intention to have the best adjustment between independent and dependent variables a mathematical adjustment was made. Table 10 summarizes the information of the adjustments made for each independent variable of the three proposed models.

Table 10: MATHEMATICAL ADJUSTMENT OF MODEL VARIABLES - 2001.

Deforested area "OLS " Model

In the lineal modeling of the Amazon municipal districts deforestation level in 2000 – measured in square kilometers –, we used the same strategy that was used in the previous modeling. In that sense, Table 11 brings the information on the variables selected for the deforestation model in 2000. By the results it is possible to predict, approximately, 73% of the total variance of the random error.

Table 11: DEFOREST AREAS "OLS" MODEL RESULTS – 2001. Regression Statistics R multiple R-Square R-square adjusted N 0.850 0.7217 0.7186 630

The analysis of the residues of the model also indicates the outliers' presence, however in a smaller degree, around 5%. The residues spatial distribution generated by the model can be observed in Map 2. As was foreseen, results suggest strong spatial correlation. It is interesting to notice that the 2000 model presented deviations pattern similar to the 1997 model (0,182), in spite of the decrease of the forest stock in the area.

Map 2. BRAZILIAN AMAZONIAN – RESIDUES SPATIAL DISTRIBUTION – OLS MODEL: 2001. Source: IBGE: Municipalities digital layer - 2000.

Table 12 displays the information regarding the Amazon municipal districts deforestation modeling, in 2000, with the introduction of the outlier identification variables. Just as in the previous modeling, the increase of the R^2 value was considerable (from 0.73 – Table 12 – to 0.81 – Table 13) and the results regarding the coefficient significance levels of all other variables were also more satisfactory.

Table 12: OUTLIERS CONTROL "OLS" MODEL RESULTS: 2000.

Regression Statistics

R-MultipleR-SquareR-adjusted square N

0.897 0.8052 0.8023 630

Lag Spatial Model - 2SLS

Box 2 summarizes results of the tests that were applied to the percentage deforest areas "OLS" model, up to 2001, for the 630 Amazon municipal districts. Results suggest that both effects may be present: spatial dependence of the deforested area percentage and spatial dependence of the model residues.

Box 2: OUTLIERS CONTROL OLS MODEL SPATIAL DEPENDENCE TEST: 2001.

Table 13 shows the Amazon municipal deforestation Lag Spatial Model information. The 2001 model was adjusted by the 2SLS method. Results show a high linear correlation index (0,93) which suggests that the model explains approximately 87% of the random error total variance, ε. The comparison of these results with those of the classic modeling reveals that there was an increasing in the R2 value (from $0.81 -$ Table $13 -$ to $0.86 -$ Table 17). The significance coefficient high level also suggests the confirmation of the dependent variable autocorrelation.

Table 13: LAG SPATIAL MODEL (2SLS) RESULTS: 2001. Regression Statistics R multiple R-square R-adjuted Square N

0.930 0.8646 - 630

Spatial Lag Model – Maximum Likelihood

Table 14 displays the 2001 Amazon municipal deforest lag spatial model information. The model was adjusted by the ML method and presented the highest linear correlation index (0,93), comparing to other models. The linear correlation index value suggests that the model explains 87% of the random error total variance, ε. The comparison between these results with those of the 2SLS modeling reveals that there was an increase of the R2 value (from 0,86 – Table $13 -$ to $0.87 -$ Table 14).

Table 14: MAXIMUM LIKELIHOOD LAG SPATIAL MODEL RESULTS: 2001.

Regression Statistics

Log. Max.R-Square AIC N

437.612 0.8716 -853.2230 630

III.3 Modeling the Amazon deforestation between 1997 and 2001

Variables of the model of deforestation between 1997 and 2001

Table 15 lists the variables that were explored to compose the 1997 to 2001 model. Some of the variables represent the difference between the observed stocks in 1997 and 2001.

Table 15: BRAZILIAN AMAZON – AMAZON DISTRICTS DEFORESTATION MODEL (1997/2001): EXPLORED VARIABLES

Sources: 3, IBGE - PPM, 1997 e 2001; 3 e 4, IBGE - PAM, 1997 and 2001; 4 and 5, IBGE – 1996 Populational Counting and 2000 Demographic Census; 6, IBGE – Censo Agropecuário de 1995/1996, 1996 Populational Counting and 2000 Demographic Census; 7 to 9, IBGE - 2000 Demographic Census.

In the case of the Amazon deforestation growth model, we opted to work with the municipal districts that presented deforested stocks variation between 1997 and 2001. With that, the number of municipal districts was reduced to 399, from the original 630.

Table 16 suggests that the average deforestation increase (4% on average) was accompanied by positive variations of the Amazon municipal districts migratory rate and volume, among those municipal districts that presented a decrease in the original forest stock and an increase in the agricultural activity.

Table 16: BRAZILIAN AMAZONIAN – 1997/2001 AMAZON DISTRICTS DEFORESTATION MODEL – SELECTED VARIABLES – DESCRIPTIVE STATISTICS.

Variables Mathematical Adjustment

The independent variables were mathematically adjusted, looking for the best adjustment between them and the dependent variables. Table 17 summarizes the adjustment information for each independent variable of the proposed model.

Table 17: 1997/2001 DEFORESTATION MODEL INDEPENDENT VARIABLES MATHEMATICAL ADJUSTMENT.

The independent variables values were corrected based on the non linear relationship between them and their respective dependent variables. This adjustment improved the linear correlation indexes in almost all the adjusted variables (Tables 7 to 9). Once the independent variables were selected and adjusted, the subsequent step was the adjustment of the deforestation models.

Deforestation "OLS" Model between 1997 and 2001

The Brazilian Amazon municipal deforestation growth model is summarized in Table 18. Results show that, comparing to the stock models, the adjustment was not so satisfactory. The linear correlation index between the deforestation density predicted and observed values was 0,56, what suggests that the model explains only 31% of the random error total variance.

Table 18: DEFORESTATION "OLS" MODEL RESULTS – 1997/2001.

Regression Statistics

R multiple R-Square R-square adjusted N

The residues analysis also indicates the presence of outliers. In the case of the growth model, about 5% of the municipal districts presented discrepancy in the residues. In the three proposed models, the outliers were related to residues superior to two pattern deviations. Map 3 shows the residues spatial distribution. As it was expected, results suggest strong spatial correlation. The difference of the geographical composition observed Map 3, in relation to Maps

1 and 2, is due to the absence of the places that, according to the data, did not present deforestation levels modification between 1997 and 2001.

Map 3. BRAZILIAN AMAZONIAN – RESIDUES SPATIAL DISTRIBUTION – OLS MODEL: 1997/2001.

Source: IBGE: Municipalities digital layer - 2000.

Regression Statistics

In spite of the introduction of outlier identification variables, the increase of the value of R^2 was not capable to reproduce the previous models levels. The comparison between Tables 14 and 15 shows that the parameter value increased from 0,31 to 0,60, approximately. The identification variables bring, however, a sensitive improvement in the coefficients significance levels, for all the other model variables. Table 19 displays the information regarding the Amazon municipal districts deforestation model, between 1997 and 2000, with heteroskedastic control.

Table 19: OUTLIERS CONTROL "OLS" MODEL RESULTS: 1997/2001.

Lag Spatial Model – 2SLS.

Box 3 summarizes the results of the tests that were applied to the 1997 to 2001 deforestation linear regression model, for the 399 Amazon municipal districts. The results suggest, as in previous cases, the possible presence of both deforested area percentage and model residues spatial dependence. The chosen procedure was the same.

The 1997 to 2001 Deforestation Growth Lag Spatial Model was adjusted by the 2SLS method. In spite of the high spatial correlation significance coefficient (ρ) level, the $R²$ increase was smaller than the increase observed in the previous lag spatial models, but was higher than the increase observed in the non-spatial adjustment. The comparison between Tables 19 and 20, shows that parameter increased from 0,60 to 0,67, approximately. The spatial modeling allowed, however, a sensitive improvement in the coefficients significance levels of all other model variables of the model, specially in the constant value.

Table 20: LAG SPATIAL MODEL (2SLS) RESULTS: 1997/2001.

Regression Statistics

R multiple R-square R-adjuted Square N

0,819 0,6713 - 399

Lag Spatial Model – Maximum Likelihood

Table 21 displays the 1997 to 2001 Amazon municipal deforest lag spatial model information. The model was adjusted by the ML method and presented a high linear correlation index (0,82), what suggests that the model explains approximately 67% of the random error total variance, ε. The comparison between these results with those of the 2SLS modeling reveals that there was not an increase of the R^2 value (it remained around 0,67 – Table 20). This model also presented an improvement in the coefficient significance levels for all other model variables, when compared with the previous model (2SLS).

Table 21: MAXIMUM LIKELIHOOD LAG SPATIAL MODEL RESULTS: 1997 TO 2001. Regression Statistics

Log. Max.R-Square AIC N

982.373 0.6709 -1946.750 399

IV – Final Remarks

The main objective of this paper was to elaborate classical and spatial lineal models that establish functional relationships between the variables and the percentage of the deforested areas between 1997 and 2001, based on the statistical exploitation of variables that are, direct or indirectly, involved in the process of Amazon deforestation.

After the outliers' control, the classic lineal models presented high adjustment indexes and high estimated parameters significance degree. The tests accused strong presence of dependent variables spatial autocorrelation, what indicated that the lag spatial modeling is more adequate. Table 22 summarizes the information on the adjustment obtained for each one of the estimated models.

Table 22: RESUMED DEFORESTATION MODEL OF THE AMAZON DISTRICTS: STOCK 1997– 2001 AND DEFORESTATION GROWTH.

	Heterocedastic			
Model	Control	R	R^2	Log- likelihood

Based on Table 22, it is possible to affirm that there was a progression in the statistics that indicate the estimated models adjustment level. The climax was the deforest variables spatial modeling, made by the maximum likelihood method.

REFERENCES

ANSELIN, L. SpaceStat TUTORIAL. Urbana-Champaign, University of Illinois, 1992.

ANSELIN, L. Spatial Externalities, Spatial Multipliers and Spatial Ecopnometrics. Urbana-Champaign, University of Illinois, 2002.

GARCIA, Ricardo Alexandrino; SOARES FILHO, Britaldo Silveira; SAWYER, Diana Oya. Dimensões sócio-econômicas e movimentos populacionais: uma regionalização da Amazônia brasileira. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, XIV, 2004, Caxambú. Anais do XIV ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS. 2004. I

NSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) Censo Agropecuário de 1995/1996. (CD-ROM). Rio de Janeiro, Br: IBGE. 1996.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) Censo Demográfico de 2000. (CD-ROM). Rio de Janeiro, Br: IBGE. 2000.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) Contagem de 1996. (CD-ROM). Rio de Janeiro, Br: IBGE. 1996.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA. Banco Municipal de Informações. 2a. Ed. Rio de Janeiro, IBGE, 1998. (Disponível em CD-ROM).

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA. Perfil dos Municípios Brasileiros. Rio de Janeiro, IBGE, 1999. (Disponível em CD-ROM).

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA. Pesquisa Agrícola Municipal. Rio de Janeiro, IBGE, 2001. (Disponível em CD-ROM).

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA. Pesquisa Pecuária Municipal. Rio de Janeiro, IBGE, 2001. (Disponível em CD-ROM).

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE) (2002) Monitoring of Brazilian Amazonian Forest by Satellite 2000 <http://www.grid.inpe.br/papers.html> (June 2002).

SILVA, Wilton P. e SILVA, Cleide M. D. P. S, LAB Fit Ajuste de Curvas (Regressão não-Linear e Tratamento de Dados) V 7.2.14 c (2003-2004), online, disponível em: <http://www.angelfire.com/rnb/labfit/index_p.htm>.