UNIVERSIDADE FEDERAL DE LAVRAS

SUZANA RODRIGUES VIEIRA

INVESTIGATING THE DRIVING FORCES OF DEFORESTATION IN THE STATE OF MINAS GERAIS, BRAZIL

LAVRAS - MG

2011

SUZANA RODRIGUES VIEIRA

INVESTIGATING THE DRIVING FORCES OF DEFORESTATION IN THE STATE OF MINAS GERAIS, BRAZIL

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Ciências Florestais, para a obtenção do título de Mestre.

Orientador

Dr. Luis Marcelo Tavares de Carvalho

Coorientadora

Dra. Danielle Marceau

LAVRAS - MG

2011

Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da Biblioteca da UFLA

Vieira, Suzana Rodrigues.
Investigating the driving forces of deforestation in the state of Minas Gerais, Brazil / Suzana Rodrigues Vieira. – Lavras: UFLA, 2011.
90 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2011.
Orientador: Luis Marcelo Tavares de Carvalho.
Bibliografia.

1. Forças direcionadoras. 2. Desmatamento. 3. Mínimos quadrados ordinários. 4. Regressão geográfica ponderada. 5. Problema da unidade de área modificável. I. Universidade Federal de Lavras. II. Título.

CDD – 333.75

SUZANA RODRIGUES VIEIRA

INVESTIGATING THE DRIVING FORCES OF DEFORESTATION IN THE STATE OF MINAS GERAIS, BRAZIL

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Ciências Florestais, para a obtenção do título de Mestre.

APROVADA em 3 de junho de 2011.

Dra. Ana Paula Dutra de Aguiar	INPE
Dr. Natalino Calegario	UFLA

Dr. Luis Marcelo Tavares de Carvalho

Orientador

LAVRAS – MG

2011

Aos meus pais Márcio Vilela Vieira e Maria José Rodrigues Vieira, que me deram a vida, educação e a chance de construir meu futuro.

DEDICO

AGRADECIMENTOS

No meio a tanta dificuldade, encontrei a oportunidade e é por este motivo que gostaria primeiramente em agradecer à DEUS pelas bênçãos adquiridas e pela minha vida.

Aos meus pais, por todo incentivo nos momentos mais difíceis. Pelos bons exemplos de caráter, humildade e bondade. Por confiarem em mim desde o início.

Ao meu irmão pela amizade e à minha família por sempre torcerem por mim.

Ao meu orientador e amigo Passarinho pelo companheirismo e pelos conselhos.

À Dr. Danielle Marceau pela orientação durante os 6 meses na Universidade de Calgary.

À Universidade de Calgary em especial ao Departamento de Geomática pela bolsa de pesquisa concedida.

Aos professores Marco Aurélio, Faustinho, Lucas, Donizzeti, José Roberto, Zé Aldo, Zé Márcio Faria, Zé Márcio Rocha pelos ensinamentos durante o mestrado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo fornecimento da bolsa de pesquisa, fundamental para o desenvolvimento do trabalho.

Ao Programa de Pós-Graduação em Engenharia Florestal pela oportunidade de realização do mestrado.

Aos queridos amigos: Carolzete, Miroca, Marilda, Tati, Antônio, Boi, Pedro couve, Artur, Gazinho, Gustavo Salgado, Andressa, Bob, Clavícula, Tanga, Estopa, Tevico, Gordo, Calanguinho e Chapolim.

Aos demais colegas de pós-graduação que participaram dessa longa jornada comigo.

Às companheiras de república Patrika e Débora pelos incríveis momentos, pelas baladas, churrascos, viagens e também pelos incentivos nos momentos mais difíceis.

Muito obrigada!!

“O ser humano vivencia a si mesmo, seus pensamentos como algo separado do resto do universo - numa espécie de ilusão de ótica de sua consciência. E essa ilusão é uma espécie de prisão que nos restringe a nossos desejos pessoais, conceitos e ao afeto por pessoas mais próximas. Nossa principal tarefa é a de nos livrarmos dessa prisão, ampliando o nosso círculo de compaixão, para que ele abranja todos os seres vivos e toda a natureza em sua beleza. Ninguém conseguirá alcançar completamente esse objetivo, mas lutar pela sua realização já é por si só parte de nossa liberação e o alicerce de nossa segurança interior.”

Albert Einstein

ABSTRACT

In 2003, the government of Minas Gerais, Brazil, devised a vegetation monitoring system that provides important information to the government agencies such as land-cover maps, deforestation rates, volume estimates and carbon stocks. Nevertheless, the main factors behind deforestation in the region are still unidentified as well as in each scale these factors operate. The main goal of this study was to explore the spatial variability of the driving forces behind deforestation at three different scales using Ordinary Least Square (OLS) and Geographically Weighted Regression (GWR) techniques. Scales correspond to three aggregation levels: the municipalities (Aggregation level 1), the micro-regions (Aggregation level 2) and the watersheds (Aggregation level 3). The datasets was provided by the state's monitoring system and public agencies. The results reveal that the driving forces of deforestation are agricultural area, charcoal production, and monoculture forest area at Aggregation level 1 while charcoal production and monoculture forest area dominate at Aggregation levels 2 and 3. The GWR results present significant improvement compared to the OLS results. Additionally, the GWR technique provide useful insights for the government decision-makers about the spatial relationships between dominant driving forces and deforested areas at multiple scales. This study also confirms the influence of the modifiable areal unit problem (MAUP) in OLS and GWR models.

Keywords: Driving forces. Deforestation. Geographically weighted regression. Ordinary least square. Modifiable areal unit problem.

RESUMO

Em 2003, o governo de Minas Gerais desenvolveu um sistema de monitoramento da vegetação, fornecendo informações importantes para os órgãos governamentais, tais como mapas de uso do solo, taxas de desmatamento, estimativas de volume e estoque de carbono. Contudo, as principais causas do desmatamento na região ainda não foram identificadas assim como as escalas que essas causas atuam. O principal objetivo do estudo foi explorar a variabilidade espacial das forças direcionadoras do desmatamento em três diferentes escalas através das técnicas dos Mínimos Quadrados Ordinários (OLS) e da Regressão Geográfica Ponderada (GWR). No estudo, escala corresponde a três níveis de agregação: os municípios (Agregação de nível 1), as micro-regiões (Agregação de nível 2) e as bacias hidrográficas (Agregação de nível 3). A base de dados foi fornecida pelo sistema estadual de monitoramento e agências públicas. Os resultados revelam que as forças direcionadoras do desmatamento são a área agrícola, produção de carvão e reflorestamento na Agregação de nível 1, enquanto a produção de carvão vegetal e reflorestamento dominam os níveis de agregação 2 e 3. Os resultados GWR apresentam uma melhora significativa em comparação com os resultados OLS. Além disso, a técnica GWR fornece informações úteis para os órgãos governamentais sobre as relações espaciais entre as forças direcionadoras de desmatamento e as áreas desmatadas em várias escalas. Este estudo também confirma a influência do problema da unidade de área modificável (MAUP) em modelos OLS e GWR.

Palavras-chave: Forças direcionadoras. Desmatamento. Mínimos quadrados ordinários. Regressão geográfica ponderada. Problema da unidade de área modificável.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1 INTRODUCTION

Land-Use and Land-Cover Change (LUCC) studies have been providing information about environmental impacts caused by human activities as in the processes of agricultural expansion, urbanization and deforestation. LUCC is also the name given to an important initiative of the International Geosphere-Biosphere Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHDP) intended to propagate the emergence of the "land-change science" and to demonstrate its role within the Earth System (LAMBIN; GEIST; RINDFUSS, 2006). In the context of these programmes, land-cover is defined as the set of different features, such as natural vegetation, soils, crops and anthropogenic structures that cover the land surface, whereas land-use refers to activities carried out by humans while exploiting land-cover features (FRESCO, 1994).

Deforestation is an important component of LUCC, with rates averaging around 13 million hectares per year between 1990 and 2005 (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAO, 2008), being predominantly concentrated in tropical countries. In spite of this alarming pace, the process of deforestation is still poorly understood except for the well documented environmental impacts like reduced soil productivity, biodiversity loss, and hydrological degradation (FEARNSIDE, 2005), which are likely to deteriorate human life quality in the future. To reverse this tendency, policy makers and environmental managers must be supported by knowledge on the main driving forces behind deforestation and focus their decisions accordingly. Recently, such knowledge has been also considered fundamental to an emerging and promising global effort: the UN-REDD Programme (The United Nations Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries).

A number of studies have explored the relationship between land-use change and its driving forces (ASPINALL, 2004; BROW; PIJANOWSKY; DUH, 2000; MERTENS; LAMBIN, 1997; MINETOS; POLYZOS, 2010; SCHNEIDER; PONTIUS, 2001; TOLE, 1998) based on empirical approaches such as Ordinary Least Square (OLS) models, and Geographically Weighted Regression (GWR) models. Due to the complexity of structural mechanistic models and due to the difficulties in quantifying the involved factors, most studies use empirical models to analyze these relationships (SLUITER, 2005).

In empirical modeling, the global regression is a mathematical model that describes the relationships between dependent and independent variables through a linear function. For instance, OLS is a method for fitting linear models by minimizing the sum of predicted squared errors. In case of deforestation modeling, the dependent variable typically represents the deforested areas while the independent variables represent the driving forces of deforestation.

OLS has been widely used in studies related to deforestation process (GEOGHEGHAN et al., 2001; MENA; BILSBORROW; MCCLAIN, 2006) due to its facility of fit the model parameters. However, there is some evidence that the OLS models have some limitations and are not appropriate to deal with spatial data in reason of the non-stationarity problem and the spatial dependence (GAO; LI, 2011; WINDLE et al., 2009).

The non-stationarity problem occurs when a process or phenomenon is not constant over space and the spatial dependence states that near things are more related than distant things. Both issues are common when dealing with LUCC modeling. As deforestation varies in time and space, models must consider this variability to generate reliable results. Aiming at solving the problem of non-stationarity and spatial dependence, a new spatial regression called Geographically Weighted Regression (GWR) (BRUNDSON; FOTHERINGHAM; CHARLTON, 1998; FOTHERINGHAM; BRUNDSON;

CHARLTON, 1998) has been developed, improving the estimators from the OLS models.

GWR is a linear model that allows the analyst to assess the spatial variability of the data across a study area, adjusting the model for each location. Many researchers have reported the superiority of GWR models compared to OLS models (GAO; LI, 2011; WINDLE et al., 2009; ZHAO; YANG; ZHOU, 2010) in many purposes and study areas. The major advantage of GWR compared to OLS models is the capability of GWR to map the coefficients for each location, capturing spatially explicit relationships between variables.

In Brazil, investigations concerning the causes of deforestation are concentrated in the Amazon region due to the ecological importance of its biome. Several studies indicated that the main causes of deforestation are related to the expansion of infrastructure (LAURANCE, 2001; LAURANCE et al., 2002), expansion of cattle ranching (MARGULIS, 2004), proximity to markets (PFAFF, 1999), population density (LAURANCE et al., 2002), and climate conditions (CHOMITZ; THOMAS, 2003).

In 2003, the government of Minas Gerais, Brazil, devised a vegetation monitoring system that provides information to the government agencies such as land-cover maps, deforestation rates, volume estimates and carbon stocks. Nevertheless, the main factors behind deforestation in the region are still unidentified. Compared to other Brazilian states within the *Mata Atlântica* biome, Minas Gerais presents one of the highest rates of deforestation (FUNDAÇÃO SOS MATA ATLÂNTICA; INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE, 2010). Besides the *Mata Atlântica*, the *Cerrado* biome has been severely degraded in Minas Gerais, where deforestation rates are higher than in the Amazon biome. According to the state's monitoring system, the deforested area was 152,635 ha between 2003 and 2005, and 109,754 ha between 2005 and 2007 (CARVALHO; SCOLFORO, 2008;

SCOLFORO; CARVALHO, 2006). Thus, the conservation of remaining forests in Minas Gerais depends on government actions such as the creation of protected areas and effective surveillance of illegal logging. For this reason, accurate mapping, monitoring and LUCC modeling are important tools to guide government planning.

LUCC modeling is a useful tool to identify and quantify the sources of land-use and land-cover changes in a simple manner represented by a mathematical function, formalizing our knowledge on understanding land degradation and its consequences and specifically with the deforestation process, there are few studies that have used the GWR technique to explore the local variations of driving forces.

One important issue when working with empirical approaches on LUCC modeling is the scale of observation. Since the driving forces of deforestation operate at different scales, they cannot be captured at one single scale (MOREIRA et al., 2009) and vary from local to global scales.

Considering the data aggregation one component of scale, it allows an analyst to explore the results at different scales and is also used to reduce data volume and processing time. However, the aggregation can be a problem if loss of information occurs, compromising the final results. The aggregation problem and the scale problem are components of the modifiable areal unit problem (MAUP) (OPENSHAW; TAYLOR, 1979). Fotheringham and Wong (1991) studied the effects of MAUP in multivariate statistics. As a result, the parameters estimates were unpredictable at different scales and zoning systems. Moreover, the authors suggested that one way of assessing the MAUP is to report results at different levels of aggregations and with different zoning systems at the same scale.

To the best of our knowledge, there are no studies that had explored the driving forces behind deforestation in the state of Minas Gerais, nor attempts to analyze them at different scales. Thus, there is a lack of knowledge about the most important factors related to deforestation process, how these factors change from place to place, and how sensitive these factors are to different scales.

OBJECTIVES

The main goal of this study was to explore the driving forces behind deforestation in the state of Minas Gerais, in Brazil at three different scales corresponding to three aggregation levels: the municipalities (Aggregation level 1), the micro-regions (Aggregation level 2), and the watersheds (Aggregation level 3). The work was motivated by the following questions: (a) Among the set of variables considered in this study, what are the ones that most influence the deforestation in the state of Minas Gerais at each spatial scale? (b) What is the best technique to model the driving forces behind deforestation? (c) How does the relationship between deforested areas and the independent variables vary spatially within each scale?

The present study has contributed to LUCC modeling in the state of Minas Gerais by mapping the main driving forces behind deforestation. Furthermore, the study assessed at which scale each driving force operates.

The text is organized as follows. First, the study area and methodology used to explore the spatial variability of deforestation driving forces are described in the next section. Then, results and interpretations are presented with pertinent discussions. The final section brings conclusions and considerations for future studies.

2 BACKGROUND

2.1 Land-Use and Land-Cover Change (LUCC)

Currently, concerns about land-use and land-cover changes are often presented in the global research agenda on environmental issues (LAMBIN; GEIST; RINDFUSS, 2006) due to the influence of these processes on climate change. Mechanisms for the reduction of greenhouse gases emissions such as UN-REDD Programme (The United Nations Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries) create a financial value for carbon stored in forests, encouraging countries to reduce deforestation and consequently the national level of the emission of greenhouse gases.

Another effort to deal with climate change is called Land-Use and Land-Cover Change (LUCC). LUCC is a project of the International Geosphere-Biosphere Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHDP) attempting to propagate the emergence of the “land-change science” and demonstrating the roles of the land change within the Earth System (LAMBIN; GEIST; RINDFUSS, 2006). LUCC studies provide information about environmental impacts caused by humans as in the process of agricultural expansion, urbanization or deforestation.

Prior to considering the LUCC implementation it is necessary to recognize the differences between land-cover and land-use terms. In this context, land-cover is defined as the layer set of different types of features, such as natural vegetation, soils, crops and human structures that cover the land surface. Land-use refers to the purposes for which humans exploit the land cover (FRESCO, 1994).

According to FAO (2008), between 1990 and 2005 the rate of deforestation on the planet averaged about 13 million hectares a year, being predominantly concentrated in tropical countries. The causes of the deforestation process are still poorly understood while its consequences are severe for the environment, resulting in losses of soil productivity and biodiversity, hydrological degradation (FEARNSIDE, 2005) that also has a detrimental impact on human life.

LUCC modeling can be found in literature, presenting different objectives and techniques (MANSON, 2005; SOARES et al., 2002; VELDKAMP; FRESCO, 1996; VERBURG; VELDKAMP; ESPALDON, 2002). In general, the studies try to establish a connection between the main causes of deforestation and deforestation rates, describing the relationship (BROWN; PIJANOWSKY; DUH, 2000), mapping future areas of risk (MERTENS; LAMBIN, 1997), or projecting LUCC in conjunction with its consequences (LIU et al., 2009; TRISURAT; ALKEMADE; VERBURG, 2010).

Indeed, the first step towards understanding LUCC is to identify and quantify the sources of changes or the main causes that control them. The examination and understanding of the main causes or “driving forces” behind deforestation guide government programs in assisting deforestation inspection teams. Nevertheless, the examination of driving forces is not a straightforward task due to the complexity of the factors involved (GEIST; LAMBIN, 2002), to interactions in time and space (VELDKAMP; FRESCO, 1996) and to a variety of scenarios of land use and land cover change in a complex system.

Sluiter (2005) mentioned that three different approaches can be distinguished for the selection and quantification of the main causes of deforestation. The structural approach defines rules based on process information, theories and physical laws. The empirical approach uses statistical methods to define transition rules and finally, the expert knowledge approach is based on human experience. Of the three approaches, the empirical approach is most widely used to quantify the relationships between variables due to the complexity of structural mechanistic models and due to the difficulty of quantifying the factors involved. Empirical models of LUCC usually use remote sensing data and explanatory variables calculated in a GIS (Geographic Information System) (IRWIN; GEOGHEGAN, 2001). As stated by the authors, these models fit the spatial process and the land use change outcome reasonably well. Nonetheless, they are less successful at explaining the human behavior.

2.2 Driving forces of deforestation

According to Geist and Lambin (2002), the driving forces behind the deforestation process can be classified into proximate causes and underlying driving forces (Figure 1). The proximate causes can be divided into three main groups: a) infrastructure, b) agricultural expansion and c) wood extraction and the underlying driving forces can be divided into five large groups: a) demographic factors, b) economic factors, c) technological factors, d) policy and institutional factors and e) cultural factors.

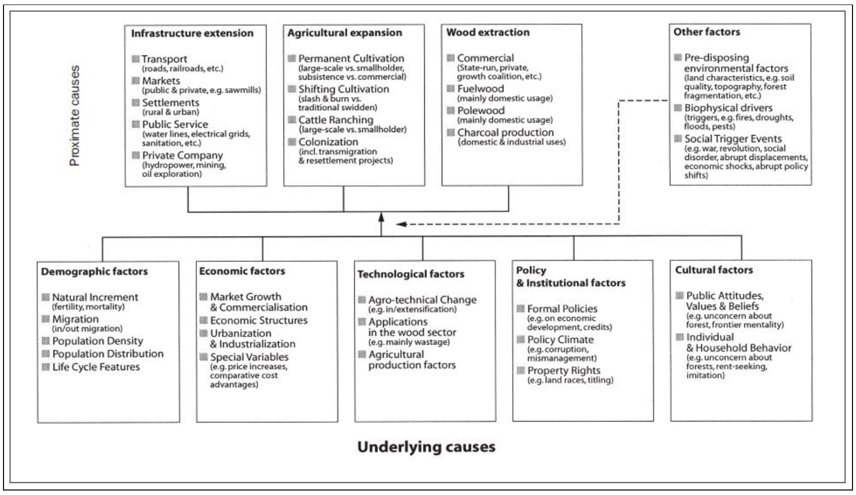


Figure 1 Proximate and Underlying driving forces
Source: Geist and Lambin (2002)

While some authors describe causes of deforestations as proximate and underlying driving forces (GEIST; LAMBIN, 2002; VERBURG et al., 2004) other authors classify them into direct and indirect causes (JAIMES et al., 2010). This study employs the terminology proposed by Geist and Lambin (2002).

According to Tole (1998), the key causes of deforestation in tropical countries include proximate and underlying driving forces such as expanding infrastructure, trade, debt, investment in human capital base, and resource-based economic expansion. Lambin et al. (2001) argued that neither population nor poverty alone constitutes the sole underlying causes of land-cover changes but that global forces are the main determinants of land-use changes, attenuating the local factors.

In Brazil, predominantly in the Amazon region, the proximate causes play a major role in LUCC. Infrastructure expansion such as building highways and roads has had a significant impact for decades and is still a problem for natural vegetation (KIRBY et al., 2006; LAURANCE et al., 2002). Exploring the intra-regional differences of land use in Brazilian Amazon, Aguiar, Câmara and Escada (2007) determined the main factors of deforestation, pasture, temporary and permanent agriculture using global regression and spatial regression models. The authors evaluated 50 explanatory variables divided into 7 categories such as access to markets, economy, agricultural structure, demography, technology, environmental and policy. The authors concluded that the human occupation in the Amazon region was heterogeneous in time and space and each region studied presents a specific spatial pattern of deforestation. Also, the authors revealed that the heterogeneous occupation patterns of the Amazon can only be explained when combining several factors such as favorable environmental conditions and access to local and national markets.

In general, various instances of research confirm that the main causes of deforestation are related to the expansion of infrastructure (LAURANCE, 2001; LAURANCE et al., 2002), and cattle ranching (MARGULIS, 2004), proximity to markets (PFAFF, 1999), population density (LAURANCE et al., 2002) and climate conditions (CHOMITS; THOMAS, 2003).

In the state of Minas Gerais, the deforestation process has started with mining for valuable minerals such as gold, iron and others. Compared to the Amazon region, the deforestation process in Minas Gerais is not accomplished by the expansion of infrastructure or climate conditions. The study developed by Carvalho, Scolforo and Cavalcanti (2009) showed the trend of land use conversion in the state. The authors investigated a large number of field samples of deforested areas from 2005 to 2007 and concluded that the main conversion of land use is a consequence of cattle ranching, agricultural activities, and monoculture forest areas. These results can be of value for governmental actions, but should not be viewed as reflecting historical aspects of conversion of land use in the state since the data collection only lasted for 2 years. Moreover, specific investigations in each region of the state will contribute to quantify the main driving forces behind deforestation.

2.3 Scale

It has been recognized that the scale of observation is a major concept in many sciences concerned with human activities and physical process occurring at the Earth's surface (MARCEAU, 1999). Scale is an important assumption that must be considered when developing LUCC models, especially when dealing with the driving forces of deforestation. The driving forces consisting of proximate and underlying causes interact in a complex system and cannot be captured in one single scale (MOREIRA et al., 2009) since driving forces vary significantly from local to global scales. Aiming at considering all driving forces in LUCC modeling, authors have been used multi-scale analysis (EVANS; KELLEY, 2004; VERBURG; CHEN, 2000). However, the relationships obtained at a certain scale of analysis may not be replicated when using other scales or in other study areas (VERBURG; CHEN, 2000).

Scale refers to spatial and temporal dimensions of an object or process, distinguished by a level of organization or hierarchy (TURNER; GARDNER; O'NEILL, 2001). It can also be defined as a continuum through which entities, patterns, and process can be observed and linked (MARCEAU, 1999). According to Kok (2001) scale is defined as a level of observation while level is defined as a level of organization. In this study, scale refers to a level of organization.

There are several reasons to employ scale into spatial analysis and into deforestation process. First, identifying the driving forces behind deforestation process requires an understanding of how people make land-use decisions and how the factors interact in specific contexts at the local, regional or global scale (VERBURG et al., 2006). For instance, proximate causes interact with the deforestation process in a local scale instead underlying causes which interact in a global scale. Second, many environmental problems, such as global warming, deforestation and water management cannot be handled at one single scale of observation (MARCEAU, 1999) due to the heterogeneity and complexity of surface (VERBURG et al., 1999).

In this context, several studies incorporated scale dependencies into LUCC models (JANTZ; GOETZ, 2005; KOK, 2001; MOREIRA et al., 2009; SOLER; ESCADA; VERBURG, 2009). In an urban land-use change model, Jantz and Goetz (2005) tested sensitivity analysis of a cellular automata model varying cells resolution. It was detected that the scale influences the measurement and quantitative description of land-use patterns and impacts on the behavior of model parameters that describe the land-use change processes.

Using logistic regression, Soler, Escada and Verburg (2009) revealed that the driving forces of deforestation in the Amazonian colonization frontier vary according to the territorial extension analyzed, confirming the hypothesis that different extents yield different relationships between LUCC patterns and theirs factors. Gao and Li (2011) detected spatially non-stationary and scale dependence relationships between landscape fragmentation and related factors through GWR modeling, indicating clear different patterns of parameters estimates in different scales. Koutsias, Martínez-Fernández and Allgower (2010) studied how factors related to wildfires vary from place to place. At the end, the authors highlighted the importance of investigation of cross-scale issues as well as further exploring how the relationships between wildfires and factors vary across different scales and data resolutions.

The choice of optimal scale is influenced by the type of data available and the main purpose of the research. In some cases, the census data provided by public agencies are the unique source of data for LUCC modeling even the advance of remote sensing techniques. However, census data can misrepresent some important information. Discrepancies between remote sensing data and census data for LUCC were highlighted by Pelorosso, Leone and Boccia (2009) in the Italian central Apennines.

Moreover, substantial variability in spatial data increases the difficulty of choosing the appropriate scale for analysis (MERTENS; LAMBIN, 1997; SERNEELS; LAMBIN, 2001) and can complicate the level of aggregation.

2.4 The Modifiable Areal Unit Problem (MAUP)

Data aggregation allows an analyst to explore the results at different scales and to reduce the time and volume of processing. However, the aggregation can be a problem if losses of information occur, damaging the final results.

The aggregation problem is one of the components of modifiable areal unit problem (MAUP) (OPENSHAW; TAYLOR, 1979). The MAUP occurs when the spatial zoning system used to analyze geographic data is modifiable, resulting in complications of some statistical analysis. Another component of MAUP is called scale problem and it is defined as the variation in results that may be obtained when the same areal data are combined into sets of increasingly larger areal units (OPENSHAW; TAYLOR, 1979).

Specifically, the aggregation problem refers to the variation in results produced by the use of alternative combinations of areal units at similar scales (OPENSHAW; TAYLOR, 1979). According to Openshaw and Taylor (1979), if the areal units are arbitrary and modifiable then the value of any work based upon them may not possess any validity independent of the units which are being studied.

Fotheringham and Wong (1991) studied the effects of MAUP in multivariate statistics. As a result, the parameters estimates were unpredictable at different scales and zoning systems. Moreover, the authors suggested that one way of assessing the MAUP is to report results at different levels of aggregations and with different zoning systems at the same scale.

Recently, with the development of remote sensing and GIS, MAUP may influence the results in the spatial analysis. Marceau, Howarth and Gratton (1994a, 1994b) were the first to recognize this relationship between the MAUP and remotely sensed imagery. The authors identified that changing the measurement scale and the aggregation levels affected the values of descriptive statistics were greatly affected. Hence, people should be aware of the MAUP issue in spatial analysis and attempt to address it when possible (DARK; BRAM, 2007).

Jelinksi and Wu (1996) focused on the importance of the MAUP in the results of landscape analysis based on NDVI (Normalized Difference Vegetation Index) images. The scale problem was explored through different aggregation of pixels in NDVI images and the aggregation problem was explored through two systematic procedures (i.e. considering the same number of pixels). All arrangements were evaluated using a spatial correlation index (Moran’s I and Geary’s c). One of the suggestions for dealing with MAUP is to conduct a sensitivity analysis which enables researches to assess which variables are sensitive to the variations in scale and aggregation. However, it should be noted that sensitivity analysis stipulates for a small number of variables, scales and levels of aggregation.

A multi-scale analysis of land-cover changes was done by Evans and Keley (2004). An agent-based modeling was developed to explore the scale dependence by changing the spatial resolutions of the input data. Different scales were obtained as a result of data aggregation, producing a series of datasets at seven spatial resolutions. The impact of the aggregation problem in the agent-based model performance resulted in the loss of agents due to errors of omission, homogenization of land-cover changes, reduction of precision and smoothing of topography. Thus, the results showed the scale-dependence of the model outcomes as well as some problems in aggregating data.

2.5 Vegetation monitoring system

Remote sensing techniques are important tools for obtaining land-cover and land-use change data from inaccessible areas on large scales in a short period of time. Currently, different data and techniques of remote sensing are available and allow several research teams to develop a vegetation monitoring system (BLAIR; RABINE; HOFTON, 1999; DEFREIES et al., 2007; MINCHELLA et al., 2009; SÁNCHEZ-AZOFEITA; HARRISS; SKOLE, 2001).

In Brazil there are three operating systems for change detection in the land cover - PRODES, DETER and DETEX - developed in areas located in the Amazon region by the National Institute for Space Research (INPE). These systems are complementary in their purposes.

Since 1998, the PRODES project - Estimate of Amazon gross deforestation – has been measuring the annual rate of deforestation (clearcutting) for increments greater than 6.25 hectares (CÂMARA; VALERIANO; SOARES, 2006). The program uses TM (Landsat), CCD (CBERS) and DMC (DMC) images, whose spatial resolution is approximately 30 meters, but whose temporal resolution is rather low. The results provided by PRODES consider only the areas that are in the final process of deforestation.

In contrast, the DETER project - Deforestation Detection in Real Time - provides monthly alerts of deforestation in areas larger than 25 hectares (INPE, 2008). The program uses MODIS (TERRA) and WFI (CBERS) sensors, whose spatial resolution is 250 meters. Finally, the DETEX project - Detecting Selective Logging - allows monitoring selected logging in the forest, through high-resolution images (20 meters) providing information at the first stages of deforestation.

Land-use and land-cover changes in the Amazon region are very often omitted from processing Landsat images due to the presence of clouds above the tree cover. Thus, the estimations of deforestation in some regions may result from a linear extrapolation, miscalculating important deforestation rates and neglecting critical areas.

In the state of Minas Gerais, the State Forestry Institute (IEF) in conjunction with the Secretary of Environment and Sustainable Development (SEMAD) and the Federal University of Lavras (UFLA) has been developing a vegetation monitoring system of the native flora and reforestation since 2003 (SCOLFORO; CARVALHO, 2006). The program aims to provide important information for the state government policies every two years. The information includes land-cover maps, deforestation rates, volume estimates, and carbon stocks.

The methodology of change detection was developed by Carvalho (2001) and it was based on NDVI (Normalized Difference Vegetation Index) image differences generated from Landsat images. The NDVI image difference with positive values indicated areas where vegetation had decreased while negative values indicate areas where vegetation had increased. An algorithm was applied to minimize possible noises and emphasize the change detections. According to the monitoring system, from 2003 to 2005 the deforestation rate was 152,635 ha and from 2005 to 2007 was 109,754 ha (CARVALHO; SCOLFORO, 2008; SCOLFORO; CARVALHO, 2006).

However, the information generated by the monitoring system in the state lacks detailed data on the patterns from land-use and land-cover changes as well as an overview of the critical areas and the driving forces behind deforestation. Thus, spatial modeling techniques are useful tools to analyze LUCC in the state.

A model is defined as a simplified mathematical representation of certain phenomena in the real world that involves independent and dependent variables. In case of LUCC modeling, the dependent variables represent typically the land-use and land-cover while the independent variables represent the factors that are involved in this process.

2.6 LUCC modeling

Many researchers have modeled the land-use and land-cover changes using different methods such as agent-based models (AMB) (EVANS; KELLEY, 2004; MANSON, 2005), cellular automata (CA) (ENTWISLE et al., 2008; MÉNARD; MARCEAU, 2007; SOARES et al., 2002), artificial neural networks (MAS et al., 2004), logistic models (ASPINALL, 2004; ECHEVERRIA et al., 2008; SCHNEIDER; PONTIUS, 2001; SERNEELS; LAMBIN, 2001), and econometric models (PFAFF, 1999).

An agent-based model (ABM) also called as multi-agent-system (MAS) is a class of computational system that aims to resolve complex problems through interaction of multi agents. The multi agents have the capability to learn about the land-use and land-cover process, move on the landscape and make decisions based on their preferences. Cellular automata (CA) in its turn are also a technique in a spatially explicit model, but it is static on landscape. The mechanism of CA is composed of five main components: a matrix space, a neighborhood configuration, a time step resolution, an ensemble of cell states and a set of transition rules. Basically, each state of cells depends on their previous state and on a set of transition rules according to configuration of neighborhoods. More details can be found in Mathey et al. (2008).

Logistic regression is one of generalized linear models (GLM) and it is used when the dependent variables is binary (0 or 1) and the independent variables are continuum or categorical. In LUCC modeling, the results provided by logistic regressions demonstrate the probability of as area's deforestation based on relationships between deforestation and its driving forces.

The optimal choice of the methods for LUCC modeling varies from case to case depending on the main goal of the study. However, available data may constrain and limit the application of the appropriate methodology. Thus, data availability impacts the choice of the suitable method for LUCC modeling.

In Brazil, the investigations of the deforestation process are concentrated in the Amazon region (AGUIAR; CÂMARA; ESCADA, 2007; KIRBY et al., 2006; LAURANCE et al., 2002; SOARES et al., 2002) and the methods vary from global regression to dynamic modeling.

In general, the techniques for LUCC modeling have two purposes: a) to explain the mechanism and process of change in use and land cover, and b) to project future changes (MILLINGTON; PERRY; ROMERO-CALCERRADA, 2007). For this study, the LUCC modeling will be focused only on explaining the mechanisms of deforestation, especially on identifying the main factors causing deforestation using two different regression techniques: Ordinary Least Square and Geographically Weighted Regression.

2.7 Ordinary Least Square

The linear models are mathematical functions that describe relationships between dependent variable and independent variables through a linear function. In the case of LUCC modeling, deforestation is the dependent variable and driving forces are independent variables. Ordinary Least Square (OLS) is a method of fitting linear models and it is obtained by minimizing the sum of squared errors predicted while taking into account such assumptions as the linearity of the relationship between variables, independence of the errors, homoscedasticity of the errors and normality of the error distribution. OLS can be represented as:

$$Y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

where:
Y: dependent variable observed;
x: independent variable;
β: parameters estimated;
ε: error.

It is evident that the parameters in the OLS equation are assumed to be spatially stationarity. Global estimation cannot capture local variations in relationships between deforested areas and the driving forces. However, it is known that the deforestation process varies in time and space and models must take into consideration this variability to generate reliable results.

Many researchers reported the spatial stationarity problem of OLS estimations (GAO; LI, 2011; WINDLE et al., 2009), that reduces the efficiency of the regression and misrepresents model results. Wang, Ni and Tenhunen (2005) reported the performance of OLS model in obtaining a net primary production (NPP) for different forest ecosystems in China. The authors compared the performance of the OLS model against the lag spatial and GWR models. The results indicated that the OLS technique was less successful to estimate NPP due to the stationarity of its parameters estimated.

To overcome this problem, geographically weighted regression (GWR) technique (BRUNDSON; FOTHERINGHAM; CHARLTON, 1998;

FOTHERINGHAM; BRUNDSON; CHARLTON, 1998) was developed to improve the global estimations, addressing spatial non-stationarity and spatial dependence.

2.8 Geographically Weighted Regression

Geographically Weighted Regression (GWR) is a linear model that allows the analyst to assess the spatial variability of the data across the study area. Specially, the GWR is used when spatial non-stationarity relationships between variables prevail (FOTHERINGHAM; BRUNDSON; CHARLTON, 2002). In contrast to the global regression or OLS model, GWR model can be expressed as follows:

$$Y(u,v) = \beta_{0t}(u,v) + \beta_{1t}(u,v)X_{1t} + \beta_{2t}(u,v)X_{2t} + \cdots + \beta_{nt}(u,v)X_{nt} + \varepsilon(u,v) \tag{2}$$

where:

Y: dependent variable observed;

(u,v): the coordinate location of the observation;

x: independent variable;

β: parameters estimated;

ε: error.

The component (u,v) indicates that the parameters will have a different influence on each location, based on geographical weightening. The estimator for the GWR model is conditioned on each local (u,v) and takes the form of:

$$\hat{\beta}(u,v) = (X^T W(u,v) X)^{-1} X^T W(u,v) y \tag{3}$$

$W(u,v)$ is a diagonal matrix of weights relative to the position of (u,v) in the study area; $X^T W(u,v) X$ is the geographically weighted variance-covariance matrix (the estimation requires its inverse to be obtained), and y is the vector of the values of the dependent variable.

The weights ($w_n(u,v)$) are given by the weight matrix and the observations which are spatially closer to the location where the local parameters are estimated will have greater influence than the others observations which are more distant. If all weights are equal to 1 then this corresponds to the global regression, because all variables in each location have the same weight in the regression.

$$W = \begin{bmatrix} w_1(u,v) & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & w_n(u,v) \end{bmatrix} \tag{4}$$

The weights themselves are computed from a fixed or adaptive Gaussian kernel function (Figure 2).

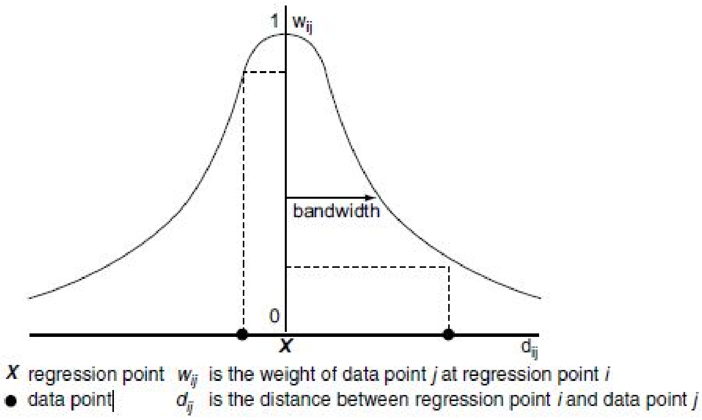


Figure 2 Kernel function
Source: Fotheringham, Brundson and Charlton (2002)

The selection of the kernel type controls the bandwidth and consequently it affects the GWR results. Figure 3 shows the difference between fixed and adaptive bandwidths in spatial kernels functions. Fixed kernel function can result in some standard errors in regions where exist few spatial data as opposed to the adaptive kernel function, which has larger bandwidths where data are sparse and smaller bandwidths where data are plentiful.

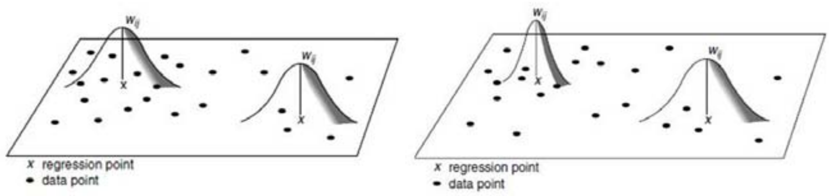


Figure 3 GWR with fixed and adaptive kernels respectively
Source: Fotheringham, Brundson and Charlton (2002)

The concern of GWR model calibration is how to select an appropriate bandwidth or decay function. The bandwidth selection can be done by means of a cross-validation (CV) approach or through an Akaike Information Criterion (AIC). The AIC approach is more flexibile when GWR is used in exploratory context and the CV is preferable when GWR is used for predicting (HARRIS; FOTHERINGHAM; JUGGINS, 2010). More details about both methods can be found in Fotheringham, Brundson and Charlton (2002).

GWR has been used in diverse ecological research projects such as vegetation distribution (AUSTIN, 2007; FOODY, 2003; GAO; LI, 2011; MILLER; FRANKLIN; ASPINALL, 2007; ZHAO; YANG; ZHOU, 2010), afforestation (CLEMENT et al., 2009), marine science (WINDLE et al., 2009), water quality (BIERMAN et al., 2011; TU; XIA, 2008), fire occurrence (KOUTSIAS; MARTÍNES-FERNÁNDEZ; ALLGOWER, 2010; TULBURE et al., 2010), deforestation (JAIMES et al., 2010; WITMER, 2005) as well as in social research (CAHILL; MULLIGAN, 2007; FARROW et al., 2005; GILBERT; CHAKRABORTY, 2011; OGNEVA-HIMMELBERGER; PEARSALL; RAKSHIT, 2009).

Clement et al. (2009) studied the factors such as social, geographical distance, physical and institutional variables in relation to afforestation in Northern Vietnam. The authors used remotely sensed and statistical data and the GWR technique to explore local variations in the relationship between the land afforested and its proximate factors. The results showed differences in the estimation of the afforestation between the remotely sensed and statistical data provided by the government. Moreover, the proximate causes of afforestation included the proximity of wood-processing industry, the distance from highways and land location from households. Zhao, Yang and Zhou (2010) applied the GWR technique to assess the spatial variability of the effect of climate conditions and site conditions on vegetation distribution. The results showed great improvement in knowledge about the vegetation distribution compared to those yielded by the OLS technique.

There are few studies of the deforestation process that have used the GWR technique to explore the local variations of driving forces. One of these studies was undertaken by Jaimes et al. (2010) who explored some potential variables associated with forest cover losses in the state of Mexico. When comparing the results of GWR to those of OLS models, the authors reported improvement on interpretation of the driving forces behind deforestation through the mapping of the parameter estimates obtained from GWR models. Another study led by Witmer (2005) characterized the relations between human activity and deforestation. The author explored the GWR results in conjunction with spatially population projections, speculating that in the coming decades, deforestation would be more intense in tropical regions of Africa.

The broad scope of the researches using GWR instead of OLS is due to the GWR's ability to handle with non-stationarity and spatial dependence of data (GAO; LI, 2011). Therefore, the GWR technique can be applied in studies that use spatial data and seek to map the relationships among variables. Miller, Franklin and Aspinall (2007) emphasized the importance of incorporating the spatial-dependence in predictive vegetation models. The authors analyzed four different statistical methods such as autoregressive models, geostatistics, geographically weighted regression, and parameter estimation models and at the end, they summarized the advantages and disadvantages of each technique.

2.9 Coefficient of determination and Sigma

One of the results provided by OLS and GWR models is the coefficient of determination (R^2), which expresses the proportion of variation in the dependent variables that explains the independent variable. Its values range from 0 to 1, with higher values being preferable. Unfortunately, when extra variables are added to the model the coefficient of determination increases significantly, giving a wrong impression of the fitted model.

Aiming to correct this problem, the adjusted coefficient of determination ($R^2 adj$) compensate the number of variables in the model and normalize the numerator and denominator on the basis of their degree of freedom, giving better parameter for models comparisons that it is also smaller than R^2 value. Model performance with higher R^2 *adj* presents the better goodness-of-fit.

Another result provided by OLS and GWR models is called Sigma and it represents the square root of the normalized residual sum of squares, where the residual sum of squares is divided by the effective degrees of freedom of the residual.

The R^2adj and Sigma will be used to compare the goodness-of-fit of OLS and GWR models.

2.10 Akaike Information Criteria (AIC)

Akaike Information Criterion (AIC) (AKAIKE, 1974) is a criterion that tries to balance the conflicting between accuracy (fit) and simplicity (small numbers of variables). The numerical value of AIC for a single model is not descriptive, but it is useful to rank different models on the basis of their twin criteria of fit and simplicity (CHATERJEE; HADI, 2006).

The AIC (corrected Akaike Information Critera) can be defined as:

$$AIC_c = 2n\,log_e(\hat{\sigma}) + n\,log_e(2\pi) + n\left\{\frac{n+tr(S)}{n-2-tr(S)}\right\} \quad (5)$$

Where:

n: sample size

$\hat{\sigma}$: estimated standard deviation of the error term

$tr(S)$: denotes the trace of the hat matrix, which is a function of the bandwidth.

For a more detailed overview of the AICc, the reader is referred to Akaike (1974) and Fotheringham, Brundson and Charlton (2002). As a general rule, the best model holds the lowest AICc value and then the model approximation is closer to reality.

Some authors have used AICc to compare OLS and GWR models (JAIMES et al., 2010; YU, 2006). All researchers revealed a better performance of GWR occasioned by the best estimation of parameters across space, spatial stationarity and spatial dependence. In this work, the AIC will be used to compare OLS and GWR performances.

2.11 Spatial Autocorrelation (Moran's I)

The "first law of geography" states that everything is related to everything else, but proximate things are more related than distant things. This law is a key concept of spatial data analysis, especially for spatial autocorrelation. Spatial autocorrelation measures how much the observed value of an attribute in a region is independent of the values of the same variable in relation to its neighbors, suggesting spatial dependence or spatial independence.

One of the OLS assumptions is that the errors terms be independent. However, this assumption is often violated due to spatial autocorrelation of the data, leading to a biased estimation of standard error parameters. Thus, autocorrelation should be taken into account in regression models because it may impair the ability to perform standard statistical hypothesis tests (LEGENDRE, 1993). The GWR technique can deal with the spatial autocorrelation problem, generating more reliable results. Zhang, Gove and Heath (2005) investigated spatial residuals from six different techniques such as generalized linear model (GLM), linear mixed model (LMM), classification and regression tree (CART), multivariate adaptive regression splines (MARS), artificial neural networks (ANN), and geographically weighted regression (GWR). All the models, except the GWR, yielded more residual clusters of similar values. These results confirm that the GWR technique has more desirable spatial distributions of errors terms.

One index very useful for measuring spatial autocorrelation is Moran's I (MORAN, 1950). This is a global index that tests the null hypothesis of autocorrelation in which values close to zero are expected among model residuals.

The Moran's I formula is given by formula:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{(\sum_{i=1}^{n}(y_i - \bar{y})^2) \sum \sum_{i=1}^{n} w_{ij}} \tag{6}$$

where:

n: total number of features

y_i: value of attribute data in region *i*

$\bar{y}$: average value of attribute data

w_{ij}: spatial weight between *i* and *j*

In general, Moran's I values range from -1 to 1, indicating negative or positive spatial autocorrelation, respectively.

4 METHODOLOGY

This section presents the implementation of the OLS and GWR techniques at three different scales. First, the potential explanatory variables datasets were collected based on the scientific literature and aggregated into three aggregation levels. Afterward, an exploratory analysis was conducted to select the most significant variables to insert into the models. Finally, the OLS and GWR models were applied to each aggregation level.

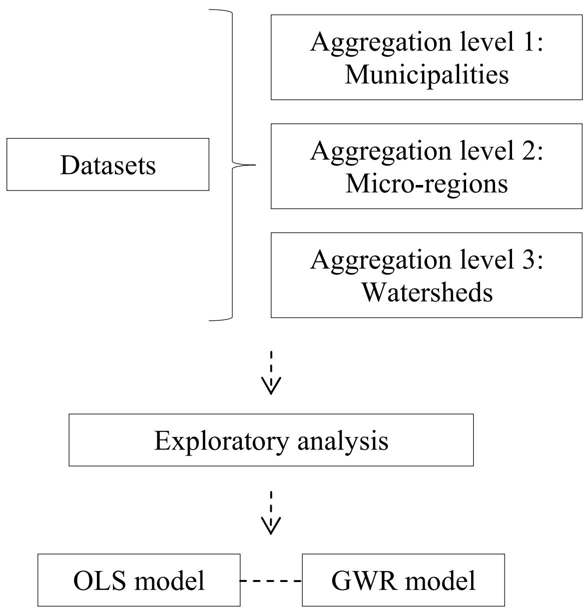


Figure 4 Schematic overview of the methodology

4.1 Study area

The state of Minas Gerais is located in south-eastern Brazil between latitudes 14º 03' 28" S and 23º 07' 02" S and longitudes 51º 07' 02" W and 39º 49' 58" W. It covers an area of approximately 590,000 km² distributed across

853 municipalities within three major Brazilian biomes: viz. *Cerrado*, *Mata Atlântica*, and *Caatinga* (Figure 5).

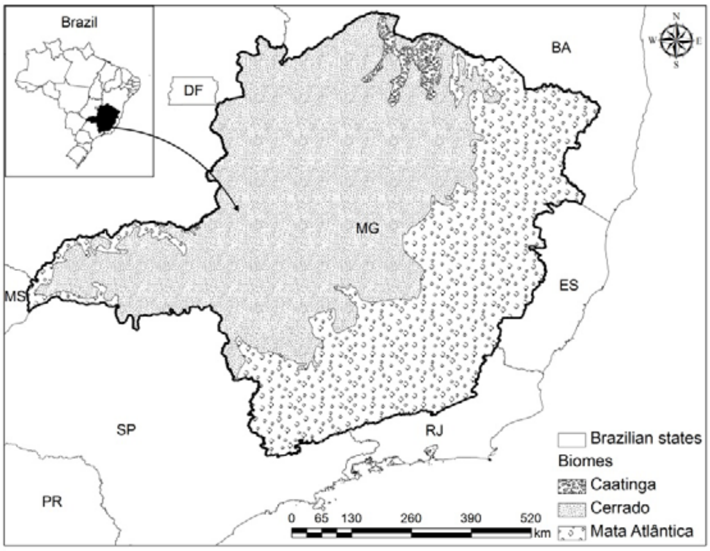


Figure 5 Location of the state of Minas Gerais, Brazil and its three major biomes

The *Cerrado* biome is the second largest Brazilian biome, exceeded only by the *Amazônia* biome. In the state of Minas Gerais, the *Cerrado* biome is the largest in area, followed by the *Mata Atlântica* and the *Caatinga* biomes. Also, the *Cerrado* biome is considered as an important biodiversity hotspot (MYERS et al., 2000) due to the highest levels of species richness and endemism. The *Mata Atlântica* biome is distributed along the Atlantic coast of the Brazil, reaching areas in Argentina and Paraguay. However, its remaining forests in the state of Minas Gerais are estimated in only 9.68% of the original area (FUNDAÇÃO SOS MATA ATLÂNTICA; INPE, 2010), being considered as one of the most endangered biomes. Finally, the *Caatinga* biome is distributed along the north-eastern states in Brazil as well as in the state of Minas Gerais. Nonetheless, compared to the *Cerrado* and *Mata Atlantica* biomes, the *Caatinga* is the least studied biome, requiring more researches about its diversity, ecology, and conservation.

Since 2003, the state of Minas Gerais in conjunction with the Secretary of Environment and Sustainable Development (SEMAD), the State Forestry Institute (IEF) and the Federal University of Lavras (UFLA) through the program entitled “Monitoring the Native Flora of Minas Gerais” (SCOLFORO; CARVALHO, 2006), has been providing important information to guide governmental strategies for sustainable forest management. This information is published every two years and it includes land-cover maps, deforestation rates, volume estimates, and carbon stocks.

According to the monitoring system (CARVALHO; SCOLFORO, 2008), the land cover is classified into 18 classes (Figure 6). Details regarding the classification process can be found in Scolforo and Carvalho (2006).

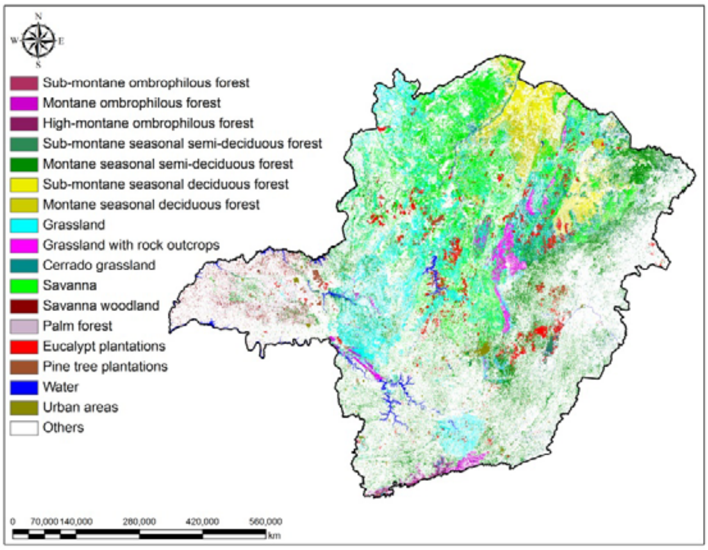


Figure 6 Land-cover map of the state of Minas Gerais for the year 2007

Besides the pronounced variability of vegetation types, the state is characterized by large socioeconomic and cultural differences across its macro-administrative regions (Figure 7). These regions were also used as benchmarking for a better comprehension and discussion of local results. Additionally, a deforestation map between the years 2003 to 2007 can be seen in the Figure 8.

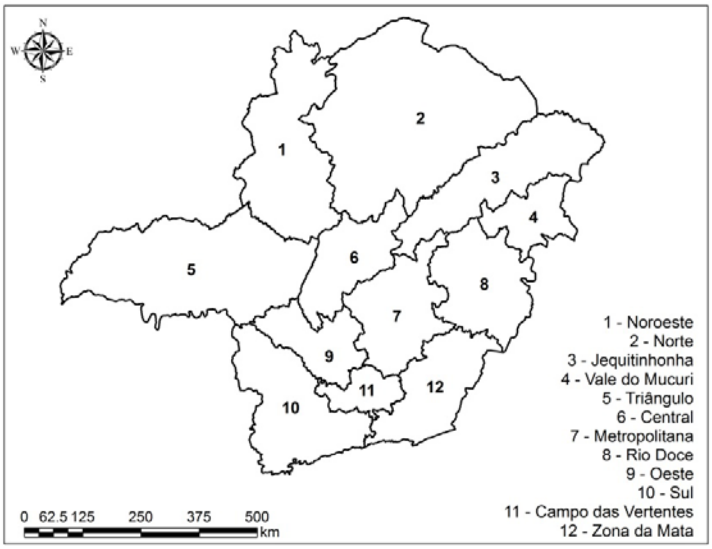


Figure 7 Macro-administrative regions in the state of Minas Gerais

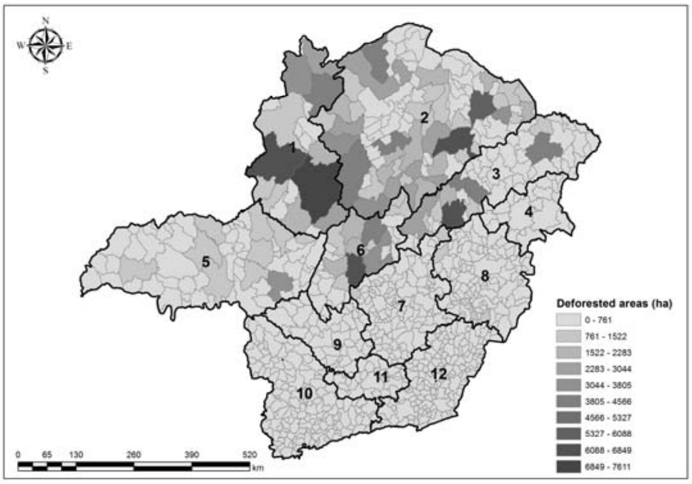


Figure 8 Deforested areas from 2003 to 2007 per municipalities

The state's population is about 19,595,309 inhabitants (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, 2010), concentrated in the Central region. The agrarian gross product index is higher in the Sul region, followed by Triângulo region (FUNDAÇÃO JOÃO PINHEIRO, 2002).

The state of Minas Gerais is also characterized by a number of steel industries that use charcoal as fuel during the production process. The charcoal comes from planted forest such as eucalyptus or from native forests located in the *Cerrado* biome. Due to the low occurrence of valuable species in the *Cerrado* biome and the industry pressure for steel mills, native forests are turned into charcoal and then, soils are prepared for the expansion of agricultural activities (CARVALHO; SCOLFORO; CAVALCANTI, 2009). According to Ughli, Goldemberg and Coelho (2008) the resources from native forests are increasingly scarce, especially in areas close to steel industries. As a result, the distance between the source of charcoal and steel industries are increasing. This situation may encourage steel industries to develop reforestation programs, aiming at supplying charcoal for the production process.

4.2 Datasets

The datasets used in the present study comprise deforestation rates registered between 2003 and 2007 (CARVALHO; SCOLFORO, 2008; SCOLFORO; CARVALHO, 2006). The deforestation rate is available per municipality and it was obtained through digital change detection applied to Landsat images. The NDVI image difference with positive values indicated areas where vegetation had decreased while negative values indicated areas where vegetation had increased. An algorithm was applied to minimize possible noises and emphasize change detections. Official figures show that 152,635 ha were deforested between 2003 and 2005, and 109,754 ha were deforested between 2005 and 2007 (CARVALHO; SCOLFORO, 2008; SCOLFORO; CARVALHO, 2006).

The independent or explanatory variables were collected based on previous research studies and on proximate and underlying driving forces proposed by Geist and Lambin (2002). In the present study, the independent variables are divided into two main groups: biophysical and socio-economic factors (Table 1).

Table 1 List of dependent variable and primary independent or explanatory variables

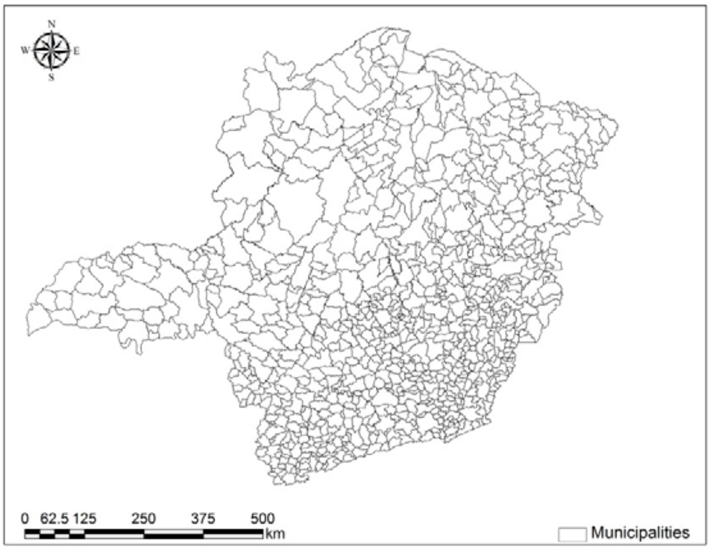
Dependent variable	**Code**	**Unit**
Deforested areas in the period of 2003 to 2007	DIF0307	ha
Independent variables		
Biophysical factors		
Monoculture forest area in 2007	R07	ha
Shortest distance to roads	SDR	m
Shortest distance to towns	SDC	m
Agricultural area in 2007	AP07	ha
Socio-economic factors		
Population in 2006	P06	no.
Agrarian gross domestic product in 2007	PI07	R$*2000
Charcoal production in 2007	PC07	ton
Charcoal price in 2007	PEC07	R$/ton
Cattle ranching in 2007	B07	no.

The biophysical factors such as proximity to roads, proximity to towns and monoculture forest areas were prepared in a Geographical Information System (GIS). The monoculture forest forest area comprises eucalyptus and pines plantations extracted from the state's land cover map. Proximity to roads and proximity to towns were calculated as the average of the shortest distance from each deforested area to the closest roads and towns, respectively. Agricultural area was obtained through the Applied Economic Research Institute (IPEA).

The selected socio-economic factors include agrarian gross domestic product, population as well as production and price of charcoal obtained through the Brazilian Institute of Geography and Statistics (IBGE) and IPEA. All variables are organized by municipalities.

4.3 The aggregation procedure

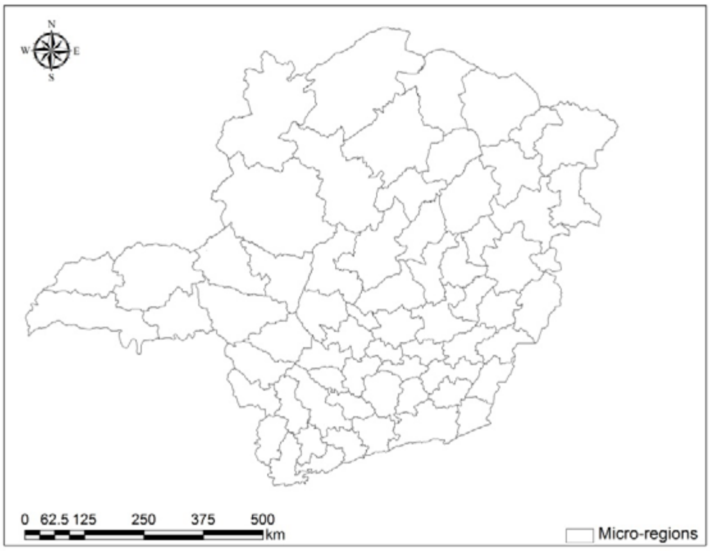
To evaluate the model results across scales, the dataset collected at the original level of municipalities was aggregated at two coarser spatial scales (Figure 8). The Aggregation level 1 (Figure 8a) refers to the original level of municipalities, containing 853 regions. These regions were aggregated into 66 regions corresponding to the micro-administrative regions to form Aggregation level 2 (Figure 8b). The micro-administrative regions perfectly coincide with the borders of the municipalities. Also, the 853 regions were further aggregated into 40 regions corresponding to the watersheds to form Aggregation level 3 (Figure 8c). In this case, the watersheds do not necessarily coincide with the municipal boundaries.

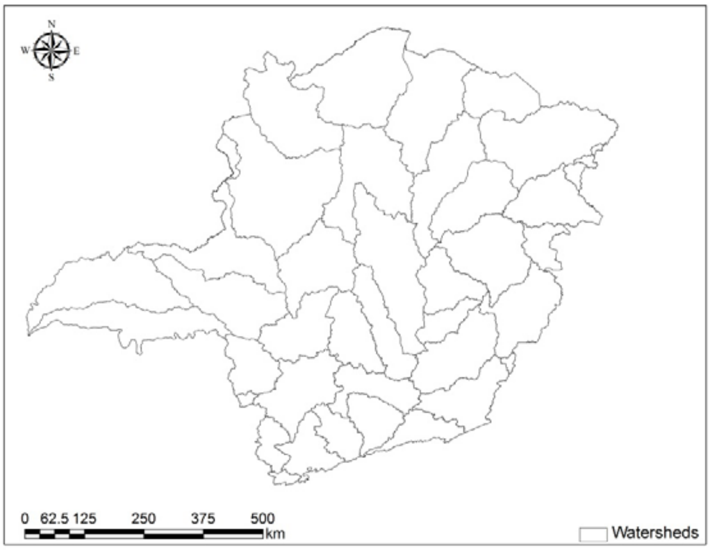


a)

Figure 7 Levels of analysis: a) Level 1; b) Level 2; c) Level 3

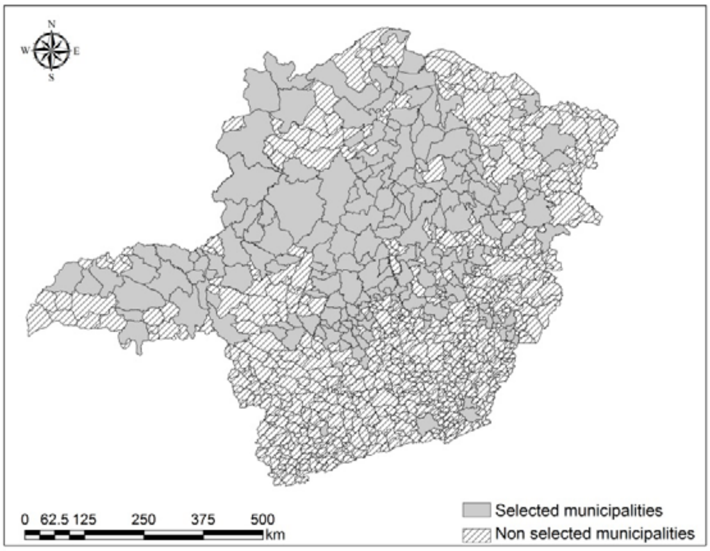
(...continue...)


N
W
E
S
0 62.5 125 250 375 500
km
Micro-regions
b)


N
W
E
S
0 62.5 125 250 375 500
km
Watersheds
c)

Data from Aggregation levels 2 and 3 were generated through a summation of values for all variables, except for the price of charcoal, which was calculated as the average of the charcoal price from all municipalities.

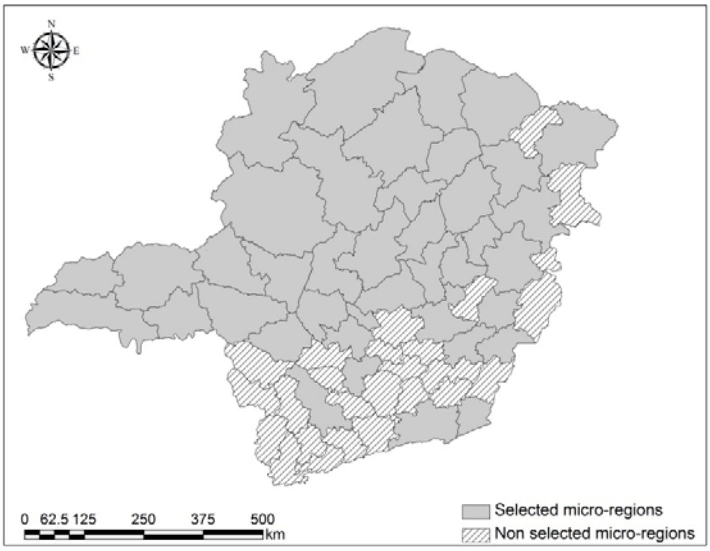
At Aggregation level 1, municipalities without the occurrence of monoculture forest forest, charcoal production, or agriculture, as well as municipalities with deforested area inferior to 10 hectares were considered as outliers and removed from further analysis, resulting in 192 municipalities. At Aggregation levels 2 and 3, the same procedure was used to remove outliers, resulting in 40 regions and 29 watersheds (Figure 9). According to Harris, Fotheringham and Juggins (2010) the outliers in the sample data have to be removed to ensure accurate results for GWR modeling.



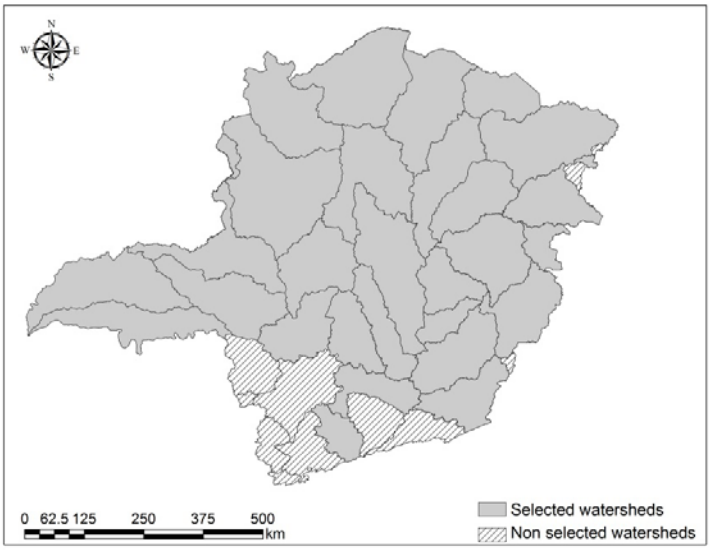
a)

Figure 8 a) Municipalities selected at Aggregation level 1; b) Micro regions selected at Aggregation level 2; c) Watersheds selected at Aggregation level 3

(...continue...)


N
W
E
S
0 62.5 125 250 375 500
km
Selected micro-regions
Non selected micro-regions

b)


N
W
E
S
0 62.5 125 250 375 500
km
Selected watersheds
Non selected watersheds

c)

4.4 Exploratory analysis

Since the normal distribution of variables is an assumption of the OLS and GWR models, the dependent variable and independent variables previously selected were tested against normality. When the variables were not normally distributed, a logarithmic transformation of the form Ln(Y) and Ln(X) was applied, improving the relationships between the dependent variable and independent variables. Moreover, a scatterplot matrix was used to display the associations among the dependent variable and independent variables. Results from multiple linear regressions are more reliable when independent variables are not strongly correlated. In this context, the strong correlation between variables is also referred as multicollinearity, which may invalidate the model. To avoid multicollinearity between the independent variables, a correlation matrix was calculated to measure the extent of association between them. Variables that presented a strong correlation (higher than 0.5) were excluded from further analysis.

Finally, seeking to refine the linear models, a forward stepwise method was carried to exclude non-significant independent variables.

4.5 Ordinary Least Square and Geographically Weighted Regression

Ordinary least squares regression (OLS) and geographically weighted regression (GWR) were used to explore the relationship between the dependent variable and the selected independent variables. OLS is a global estimation by linear models that assumes stationarity of the parameters across space and generates a single equation that shows the best fits. However, OLS can result in biased models, especially when dealing with spatial data. Thus, as the driving forces behind deforestation vary in time and space (VELDKAMP; FRESCO, 1996), the OLS method is not suitable for spatial analysis. Recent studies have demonstrated the poor performance of OLS compared to others models such as GWR (GAO; LI, 2011; ZHAO; YANG; ZHOU, 2010).

GWR appears as a new local spatial technique that addresses spatial data analysis (BRUNDSON; FOTHERINGHAM; CHARLTON, 1998; FOTHERINGHAM; BRUNDSON; CHARLTON, 2002). The technique is an improvement over global regression, providing more weight to observations spatially closer to locations being predicted. The weights assigned to each observation are given by a weight matrix scheme based on a distance kernel function.

The distance kernel function represents a Gaussian curve and can be fixed or adaptive. A fixed kernel function assumes that bandwidths are constant across the study area as opposed to adaptive kernel function that uses variable bandwidth. The adaptive kernel function produce spatial kernels able to adapt themselves to variations in the density of the data and avoiding large local variance estimation in areas where data are sparse (FOTHERINGHAM; BRUNDSON; CHARLTON, 2002). For this reason, we adopted an adaptive kernel function to implement the GWR model.

In this study, an algorithm that seeks to minimize the Akaike Information Criteria (AIC) was chosen to select the bandwidth for the adaptive Kernel function based on centroids of each analyzed entity. The AIC has the advantage of being more general in application than others statistics because it can be used for Poisson and logistic GWR, as well as for linear models (FOTHERINGHAM; BRUNDSON; CHARLTON, 2002).

In addition, a significance test-t with threshold of 95% of confidence was applied to the parameters of OLS and GWR masking out points where the relationship between variables were not significant. These values can be mapped, allowing the analyst to interpret the nature of the relationship between deforestation and driving forces.

To assess the goodness-of-fit for each model, we used AICc, R^2 adj and Sigma. The model with lower AICc values means a better fit to the observed data. AICc differences > 3 are assumed to represent significant difference between OLS and GWR models (GAO; LI, 2011; WINDLE et al., 2009). Models with higher R^2 adj and lower Sigma are also preferable.

Moran's I was calculated to evaluate the spatial pattern of OLS and GWR model residuals. Under the null hypothesis that no spatial autocorrelation exist among residuals, Moran's I has an expected value near zero. Values closer to 1 indicate positive autocorrelation and values closer to -1 indicate negative autocorrelation.

5 RESULTS AND DISCUSSION

This section summarizes our main modeling results and is organized as follows. The immediate section presents the selected variables, followed by a section about the OLS and GWR fitted models and another one about the visualization of GWR results. In these sections, results are compared among aggregation levels, highlighting the main differences found across the study area. Finally, the last section presents the spatial autocorrelation of model residuals for each aggregation level.

5.1 Selected variables

Figure 10 displays the scatterplot matrix among the dependent variable and explanatory variables. It is evident the improvement of the linear relationship at the transformed variables compared to the non-transformed variables.

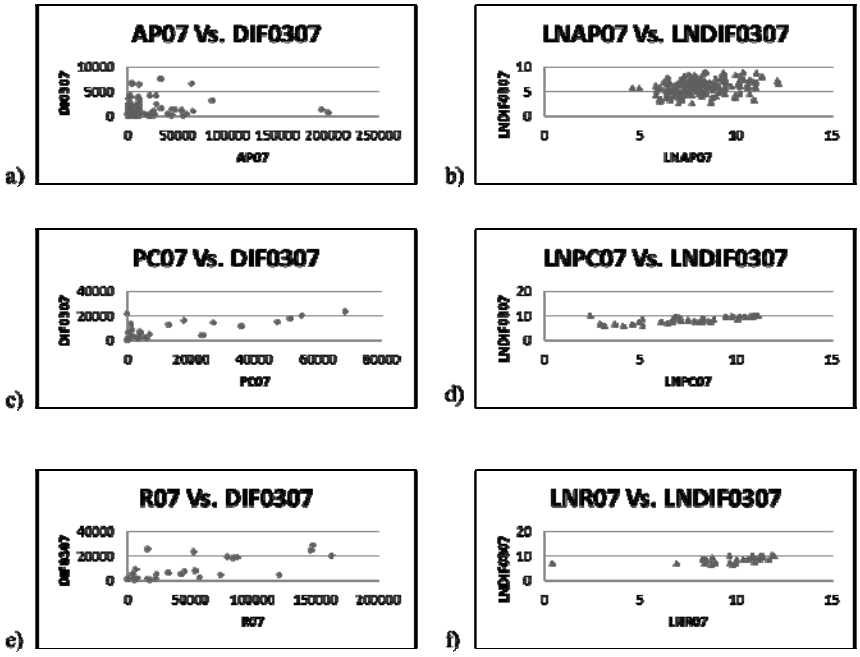


Figure 9 Scatterplot of transformed and non-transformed variables. a) non-transformed variables at Aggregation level 1; b) transformed variables at Aggregation level 1; c) non-transformed variables at Aggregation level 2; d) transformed variables at Aggregation level 2; e) non-transformed variables at Aggregation level 3

Table 2 shows summary statistics for the transformed and non-transformed dependent variables.

Table 2 Summary statistics for transformed and non-transformed dependent variables for each aggregation level

Levels	Aggregation level 1		Aggregation level 2		Aggregation level 3	
Statistics	DIF0307[1]	LNDIF0307[2]	DIF0307[1]	LNDIF0307[2]	DIF0307[1]	LNDIF0307[2]
Count	192	192	40	40	29	29
Average	931.33	5.96	6150.26	8.09	8785.93	8.43
Standard deviation	1357.30	1.42	6668.85	1.22	9106.78	1.26
Coeff. of variation	1.46	0.24	1.08	0.15	1.04	0.15
Minimum	12.87	2.55	273.06	5.61	506.88	6.23
Maximum	7611.03	8.94	23701.40	10.07	28681.80	10.26
Range	7598.16	6.38	23428.30	4.46	28174.90	4.04
Stnd. Skewness	15.26	-0.69	3.32	-0.23	2.19	-0.20
Stnd. Kurtosis	22.75	-1.56	0.63	-1.09	-0.57	-1.38

[1] Non-transformed deforested area during the period 2003 to 2007
[2] Transformed deforested area during the period 2003 to 2007

Measures of standard skewness and standard kurtosis demonstrate whether the data comes from a normal distribution. Values outside the range -2 to +2 indicate significant departures from normality. As observed in Table 3, non-transformed dependent variables (DIF0307) are not within the expected range of a normal distribution in all aggregation levels. Figure 11 shows the frequency histogram for the transformed and non-transformed dependent variables. All transformed variables are normally distributed.

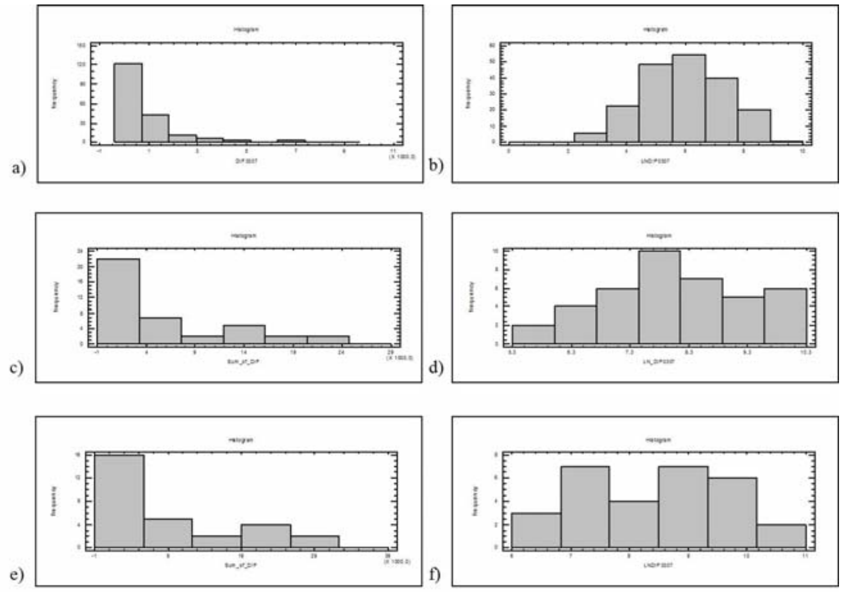


Figure 10 Frequency histogram for transformed and non-transformed variables. a) transformed variable at Aggregation level 1; b) non-transformed variable at Aggregation level 1; c) transformed variable at Aggregation level 2; d) non-transformed variable at Aggregation level 2; e) transformed variable at Aggregation level 3

At Aggregation level 1, multicollinearity was detected for the following variables: agrarian gross domestic product, cattle ranching, and shortest distance to towns (Table 3), whereas at Aggregation levels 2 and 3, multicollinearity was detected for agrarian gross domestic product, charcoal price, cattle ranching, and shortest distance to towns. These variables were removed from further analysis (Tables 4 and 5).

Table 3 Correlation matrix between independent variables at Aggregation level 1

Level 1	LNP06	LNAP07	LNPI07	LNPC07	LNPEC07	LNR07	LNB07	LNSDR	LNSDC
LNP06	1.00	0.47	0.56	0.19	0.08	0.14	0.56	0.08	0.34
LNAP07		1.00	0.74	0.26	0.05	0.14	0.62	0.37	0.51
LNPI07			1.00	0.29	-0.02	0.11	0.71	0.35	0.42
LNPC07				1.00	0.18	0.23	0.42	0.22	0.51
LNPEC07					1.00	-0.10	0.08	0.01	0.11
LNR07						1.00	0.01	0.05	0.20
LNB07							1.00	0.36	0.54
LNSDR								1.00	0.51
LNSDC									1.00

Table 4 Correlation matrix between independent variables at Aggregation level 2

Level 2	LNP06	LNAP07	LNPI07	LNPC07	LNPEC07	LNR07	LNB07	LNSDR	LNSDC
LNP06	1.00	0.30	0.57	0.00	-0.35	-0.03	0.48	0.00	-0.17
LNAP07		1.00	0.76	0.23	-0.01	-0.07	0.60	0.46	0.42
LNPI07			1.00	0.22	-0.13	-0.08	0.71	0.21	0.13
LNPC07				1.00	0.61	0.19	0.45	0.28	0.53
LNPEC07					1.00	0.10	0.05	0.15	0.48
LNR07						1.00	-0.14	-0.09	0.12
LNB07							1.00	0.37	0.39
LNSDR								1.00	0.60
LNSDC									1.00

Table 5 Correlation matrix between independent variables at Aggregation level 3

Level 3	LNP06	LNAP07	LNPI07	LNPC07	LNPEC07	LNR07	LNB07	LNSDR	LNSDC
LNP06	1.00	0.35	0.59	0.00	-0.31	0.41	0.51	-0.43	-0.16
LNAP07		1.00	0.90	0.22	0.03	0.40	0.70	0.13	0.28
LNPI07			1.00	0.23	-0.04	0.43	0.78	-0.10	0.13
LNPC07				1.00	0.80	0.36	0.44	0.33	0.74
LNPEC07					1.00	0.20	0.20	0.45	0.72
LNR07						1.00	0.32	-0.27	0.37
LNB07							1.00	0.10	0.35
LNSDR								1.00	0.47
LNSDC									1.00

The stepwise method generated the most significant variables for each aggregation level. As a result, from the nine initially selected variables, only three variables at Aggregation level 1 and two variables at Aggregation levels 2 and 3 were retained for OLS and GWR modeling (Table 6). At Aggregation level 1, the selected variables included charcoal production, monoculture and shortest distance to roads. At Aggregation levels 2 and 3, the selected variables included charcoal production and monoculture forest forest area.

Table 6 Explanatory or independent variables selected in each aggregation level for the OLS and GWR models

Aggregation level 1	Aggregation level 2	Aggregation level 3
Charcoal Production	Charcoal Production	Charcoal Production
Monoculture forest area	Monoculture forest area	Monoculture forest area
Shortest distance to roads		

The most significant driving forces detected confirm the hypothesis that the deforestation process in Minas Gerais is related to charcoal production and monoculture forest expansion.

5.2 OLS and GWR fitted models

OLS results are shown in Table 8. The OLS model explains 31 % of dependent variable variance at Aggregation level 1, 50% at Aggregation level 2 and 63 % at Aggregation level 3 (adjusted R square). Sigma values are lower at Aggregation level 3 (0.77) followed by Aggregation level 2 (0.86) and Aggregation level 1 (1.38). The AICc is greater at Aggregation level 1 (611.87) followed by Aggregation level 2 (104.42) and Aggregation level 3 (69.87). According to these parameters, the best performance is obtained at Aggregation level 3.

Table 6 OLS model results for each aggregation level

Aggregation levels	1	2	3
Sigma	1.38	0.86	0.77
AICc	611.87	104.42	69.87
R Square	0.32	0.53	0.65
R Square adjusted	0.31	0.5	0.63

Table 9 shows the results of GWR for each aggregation level. The GWR model explains 36 % of the dependent variable variance at Aggregation level 1, 73% at Aggregation level 2, and 68 % at Aggregation level 3 (adjusted R square). The Sigma value is the lowest at Aggregation level 2 (0.64), but reaches 0.72 and 1.13 at Aggregation level 3 and Aggregation level 1, respectively. These values confirm that the Aggregation level 2 presents the best GWR performance followed by Aggregation level 3 and Aggregation level 1.

Table 7 GWR model results for each aggregation level

Aggregation levels	1	2	3
Sigma	1.13	0.64	0.72
AICc	599.71	93.47	72.00
R Square	0.39	0.81	0.74
R Square adjusted	0.36	0.73	0.68

The lowest OLS and GWR performances at Aggregation level 1 implies that it is more difficult to fit a linear model to non-aggregated data composed of a large number of spatial units compared to aggregated data. The aggregation procedure uses average values from original datasets. As a result, the relationship between dependent and independent variables is more linear due to the decrease in the variance of the variable creating a smoothing effect (FOTHERINGHAM; WONG, 1991). Another view stated by Kok (2001) is referred as the aggregation error and occurs when non-linear relationships are translated at more aggregated scales. At Aggregation level 1, the adjusted R square value suggests that other influential variables should be considered into the model.

It is evident the improvement of GWR model performance analyzing AICc and adjusted R square criteria when comparing to the OLS model performance. The difference in AICc at Aggregation level 1 and 2 is greater than 3, indicating better performance in GWR results. The AICc value increase at Aggregation level 3 is not significant (<3).

5.3 OLS parameter estimates

The results of the OLS regression reveal that all parameter estimates are significant at 95% confidence level for all aggregation levels (Tables 11 and 12), except by the intercept parameter at Aggregation level 1 (Table 10). Moreover, all coefficients present positive correlation, where deforested areas and charcoal production has stronger influence followed by monoculture forest area at Aggregation levels 2 and 3 and followed by shortest distance to roads at Aggregation level 1.

Table 8 Parameter estimates for OLS regression model with significance level of 95% at Aggregation level 1

Parameter	Estimate	Standard error	*t*-value
Intercept	1,44	0,83	1,74
PC07	0,27	0,04	6,56
R07	0,09	0,03	3,23
SDR	0,26	0,09	2,82

Table 9 Parameter estimates for OLS regression model with significance level of 95% at Aggregation level 2

Parameter	Estimate	Standard error	*t*-value
Intercept	4,14	0,67	6,16
PC07	0,3	0,06	4,99
R07	0,19	0,06	3,01

Table 10 Parameter estimates for OLS regression model with significance level of 95% at Aggregation level 3

Parameter	Estimate	Standard error	*t*-value
Intercept	4,08	0,69	5,89
PC07	0,31	0,06	4,93
R07	0,2	0,07	2,86

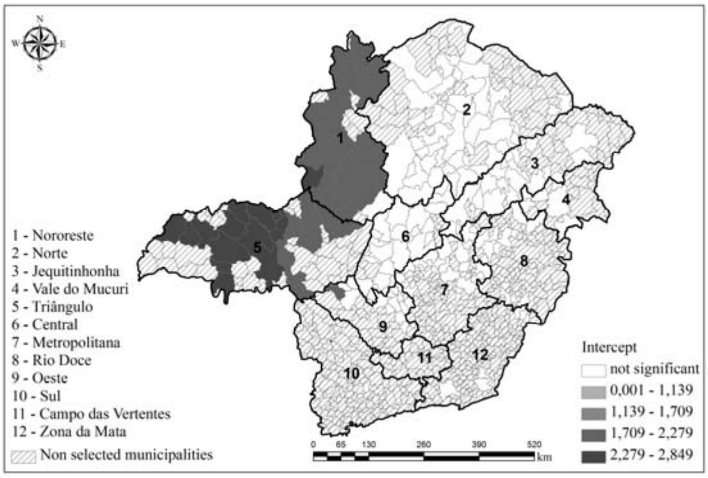
5.4 Visualizing GWR results

GWR models result in maps that highlight the spatial variability of the estimated parameters, assisting the interpretation of the factors related to the deforestation process in each local. The parameter estimates with significance threshold of 95% for each aggregation level are shown in Figure 12, 13 and 14.

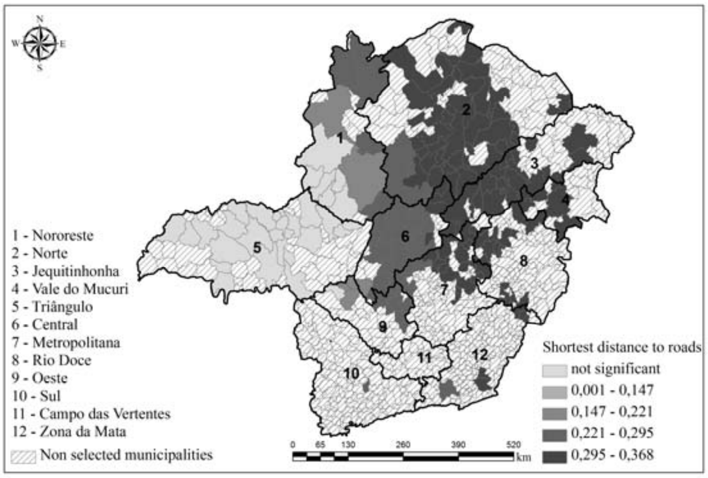
The regression coefficient β_0 represents the value of the dependent variable when all independent variables are zero. The coefficients β_n can be understood as the change in the dependent variable corresponding to a unit change in one independent variable when all other independent variables are constant.

At Aggregation level 1, all parameter estimates vary spatially according to both locations where they have a significant and positive association with deforested areas. The entities in white represent regions with the parameters were not significant with threshold of 95 %. The intercept parameter is not significant for municipalities located in the Norte, Jequitinhonha, Vale do Mucuri, Central, Metropolitana and Rio Doce regions. In contrast, the intercept parameter presents a stronger association in some municipalities located in the Triângulo region, suggesting that these municipalities may present deforested areas while other variables are zero (Figure 12a).

When analyzing the parameter estimates of the shortest distance to roads, it can be observed that the municipalities located in the Norte, Jequitinhonha, Vale do Mucuri, Metropolitana and Rio Doce regions present a strong relationship with deforested areas (Figure 12b). Charcoal production has a greater influence for most municipalities located in the Noroeste, Triângulo, Central, Metropolitana and Oeste regions (Figure 12c). Finally, monoculture forest area shows a greater influence in the Noroeste, Norte and Central regions (Figure 12d).



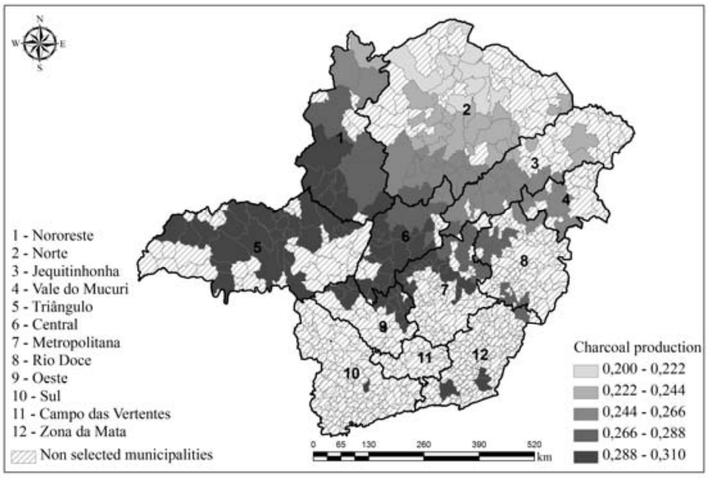
a)



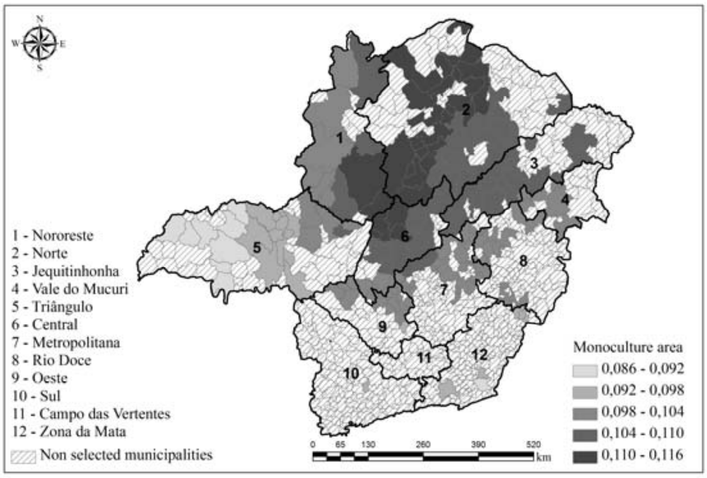
b)

Figure 11 Parameter estimates for GWR model at Aggregation level 1 with significance level. a) Intercept; b) Shortest distance to roads; c) Charcoal production; d) Monoculture forest area.

(…continue…)

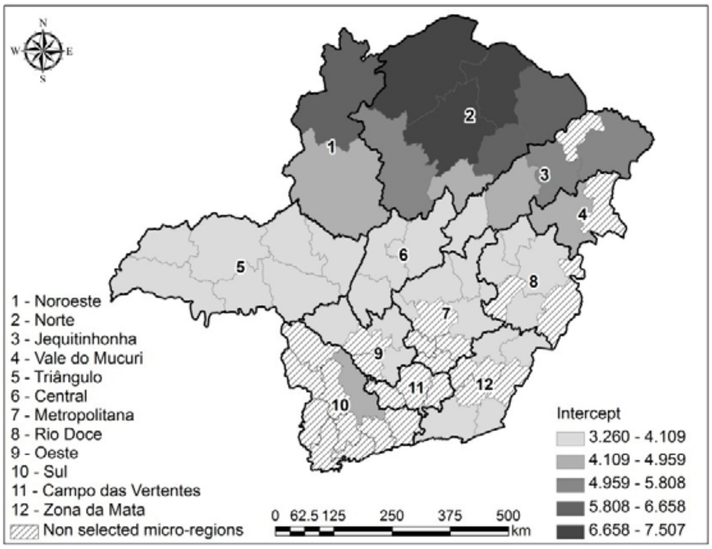


c)



d)

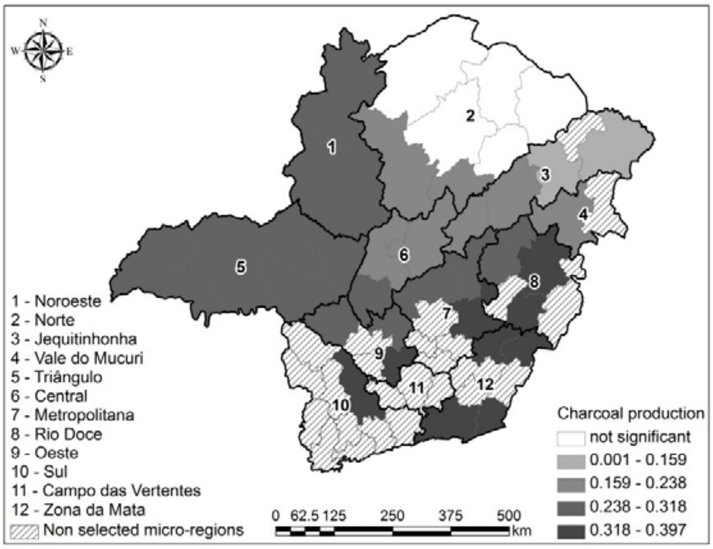
At Aggregation level 2, all estimated parameters vary spatially and present a positive relationship with deforested areas. The intercept parameter is significant for all micro administrative regions, but it is stronger in the Norte and Noroeste regions (Figure 13a). In contrast to Aggregation level 1, charcoal production has a greater influence at micro regions within the Metropolitana, Rio Doce, Oeste, Sul, and Zona da Mata regions (Figure 13b). The values for the monoculture forest area estimated parameters, presented in Figure 13c, indicate that the Jequitinhonha, Central and Rio Doce regions experience a greater deforestation process due to the expansion of monoculture forest plantations.



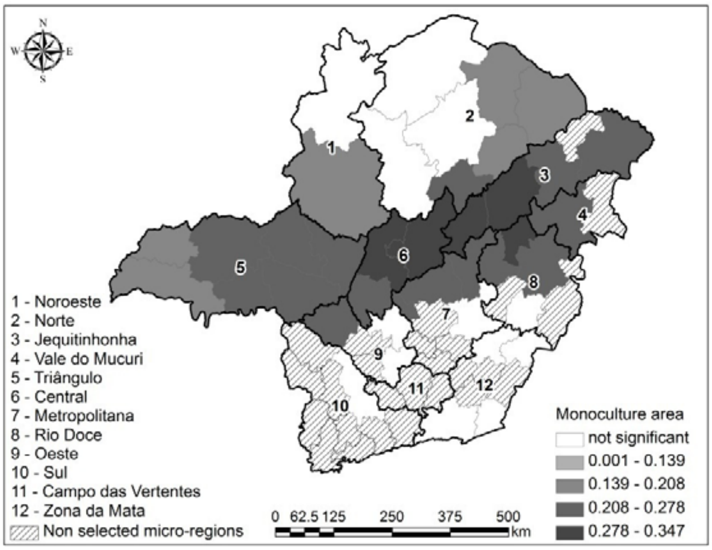
a)

Figure 12 Parameter estimates for GWR model at Aggregation level 2 with significance level. a) Intercept; b) Charcoal production; c) Monoculture forest area

(…continue…)



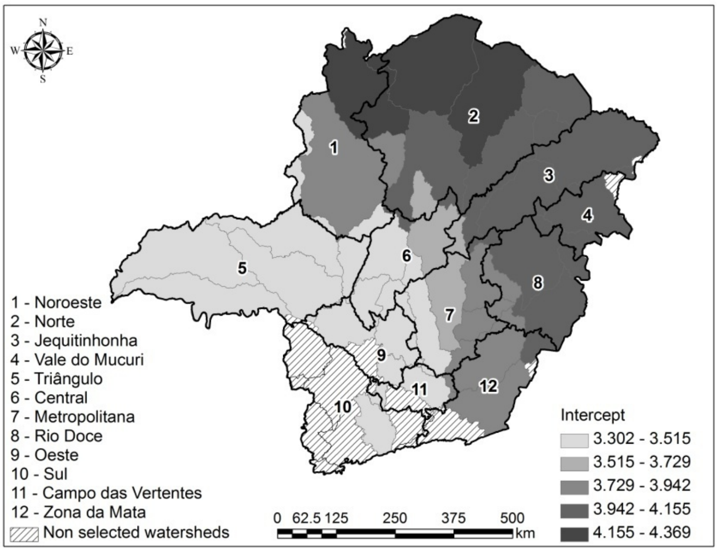
b)



c)

The estimated parameters at Aggregation level 3 reveal that the intercept is greater for watersheds located in the Noroeste and Norte regions (Figure 14a), while the charcoal production has a greater influence for watersheds located in the Norte, Jequitinhonha, Vale do Mucuri, and Rio Doce regions (Figure 14b).

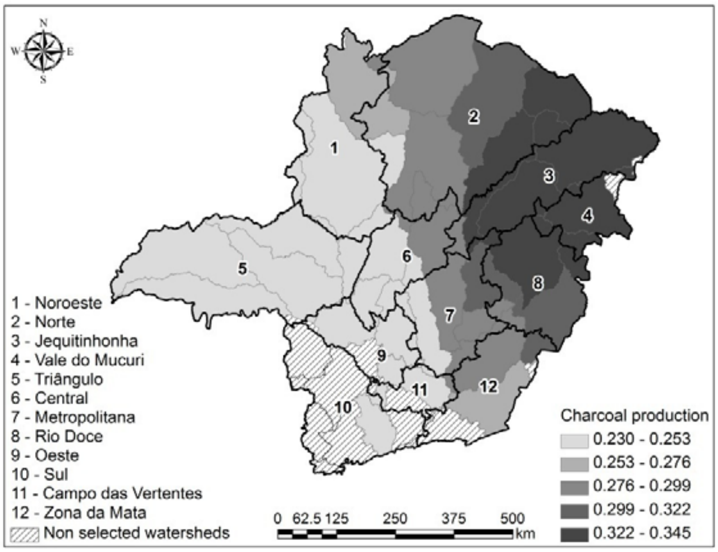
Monoculture forest area has a greater influence for some watersheds located in the Triângulo, Central, Metropolitana, Oeste, Sul and Campo das Vertentes regions (Figure 14c).



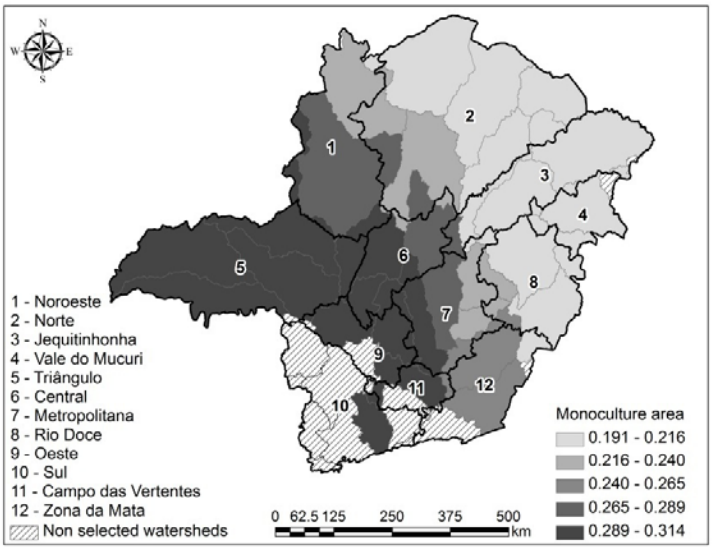
a)

Figure 13 Parameter estimates for GWR model at Aggregation level 3 with significance level. a) Intercept; b) Charcoal production; c) Monoculture forest area

(…continue…)



b)



c)

Spatial non-stationarity can be observed by analyzing all parameter estimates at all aggregation levels. Spatial non-stationarity detected by GWR models was previously reported in many studies (OGNEVA-

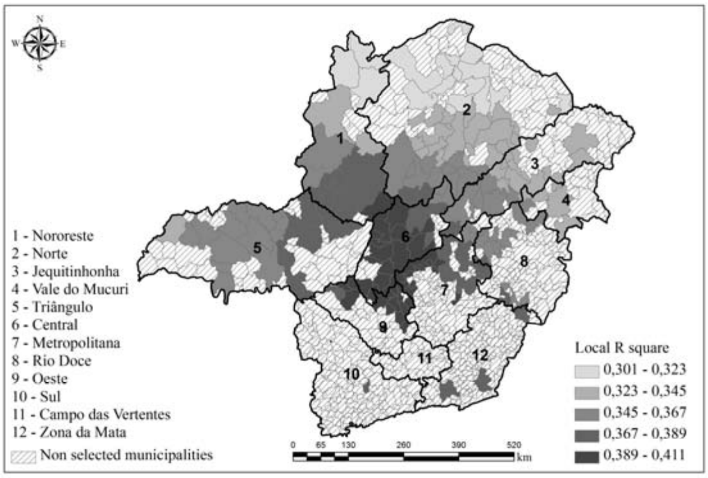
HIMMELBERGER; PEARSALL; RAKSHIT, 2009; WINDLE et al., 2009). In this study, it suggests that land-use and land-cover patterns vary across space as well as its driving forces. According to Aguiar, Câmara and Escada (2007) land-use data changes in one area tend to propagate to neighbor regions, being spatially variable. Also, the spatial non-stationarity indicates the spatial heterogeneity of the factors involved with deforestation process in Minas Gerais.

Differences were detected by comparing the parameter estimates at all aggregation levels. On the one hand, the strongest relationship between deforested areas and charcoal production at Level 3 is located in the Norte, Jequitinhonha, Vale do Mucuri, and Rio Doce regions. On the other hand, at Aggregation level 2, the charcoal production is strongly related to deforested areas in areas located in the Metropolitana, Rio Doce, Oeste, Sul and Zona da Mata regions. Another example is related to the monoculture forest area. At Aggregation level 3, the monoculture forest area shows a stronger relationship with deforested areas in administrative regions located in southwestern Minas Gerais, like in the Triângulo, Central, Metropolitana, Oeste, Sul and Campo das Vertentes regions, while at Aggregation level 2, the monoculture forest area shows a stronger relationship in regions located in the northeast of the state like in the Jequitinhonha, Central and Rio Doce regions.

These differences between aggregation levels confirm the hypothesis that the MAUP is present in this LUCC modeling and that it has a significant impact on GWR models. For each scale of analysis, the researcher is likely to find different relationships between the driving forces and deforested areas across the space.

The R square values show the proportion of variance explained by the model. In GWR models it is possible to map the R square for each local. Local R square values present variations from 0.30 to 0.41 at Aggregation level 1, 0.15 to 0.86 at Aggregation level 2 and 0.57 to 0.78 at Aggregation level 3 (Figure

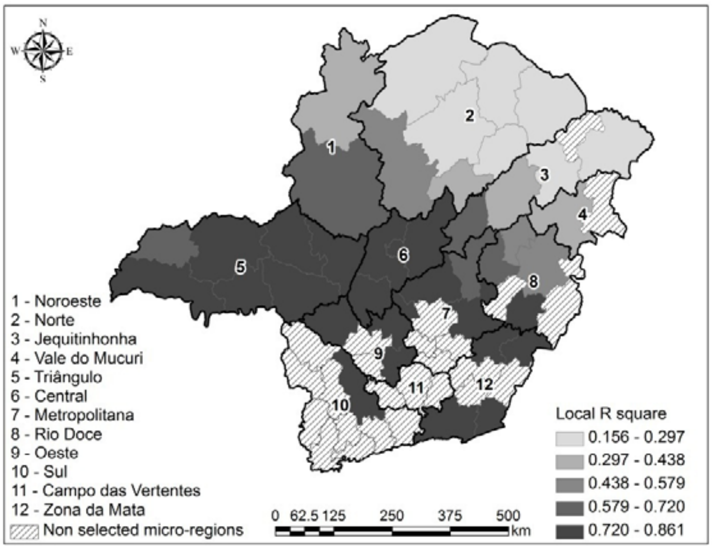
15). At Aggregation level 1, the higher R square values are located in the Central, Metropolitana and Oeste regions, suggesting that the GWR model generates better results for these regions. At Aggregation level 2, the higher R square values are located in the Triângulo, Central, Metropolitana, Rio Doce, Oeste, Sul and Zona da Mata regions, whereas at Aggregation level 3, the higher R square values are located in the Norte, Jequitinhonha, Vale do Mucuri, Central, Metropolitana, Rio Doce and Zona da Mata regions. The regions Central and Metropolitana present the best goodness-of-fit in all aggregation levels.

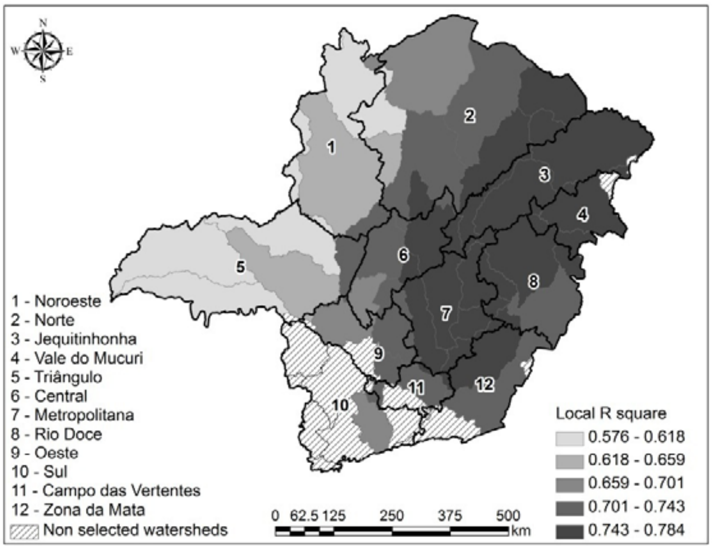


a)

Figure 14 Spatial distribution of local R square. a) Aggregation level 1; b) Aggregation level 2; c) Aggregation level 3

(…continue…)


N
W
E
S
1
2
3
4
5
6
7
8
9
10
11
12
1 - Noroeste
2 - Norte
3 - Jequitinhonha
4 - Vale do Mucuri
5 - Triângulo
6 - Central
7 - Metropolitana
8 - Rio Doce
9 - Oeste
10 - Sul
11 - Campo das Vertentes
12 - Zona da Mata
Non selected micro-regions
0 62.5 125 250 375 500
km
Local R square
0.156 - 0.297
0.297 - 0.438
0.438 - 0.579
0.579 - 0.720
0.720 - 0.861
b)


N
W
E
S
1
2
3
4
5
6
7
8
9
10
11
12
1 - Noroeste
2 - Norte
3 - Jequitinhonha
4 - Vale do Mucuri
5 - Triângulo
6 - Central
7 - Metropolitana
8 - Rio Doce
9 - Oeste
10 - Sul
11 - Campo das Vertentes
12 - Zona da Mata
Non selected watersheds
0 62.5 125 250 375 500
km
Local R square
0.576 - 0.618
0.618 - 0.659
0.659 - 0.701
0.701 - 0.743
0.743 - 0.784
c)

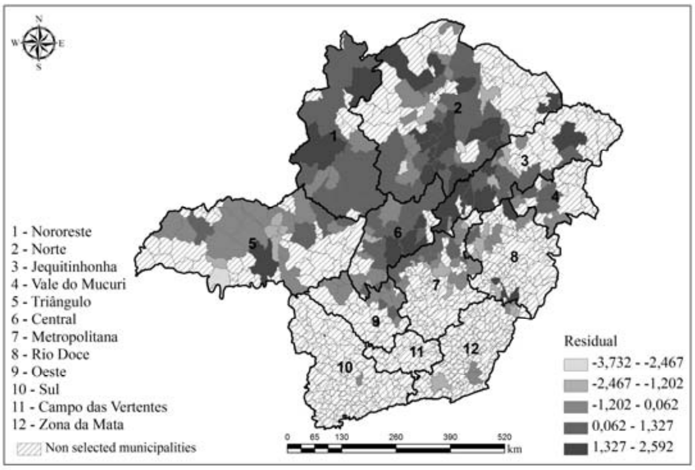
6 SPATIAL AUTOCORRELATION

Strong evidence of spatial autocorrelation is detected from OLS models residuals at Aggregation levels 1 and 2, differently from Aggregation level 3 (Table 13). The results were tested against the null hypothesis with 95% confidence level (p-value>0.05). The z score values measure standard deviation from the mean in a normal distribution. High values indicate that the pattern is not randomly distributed.

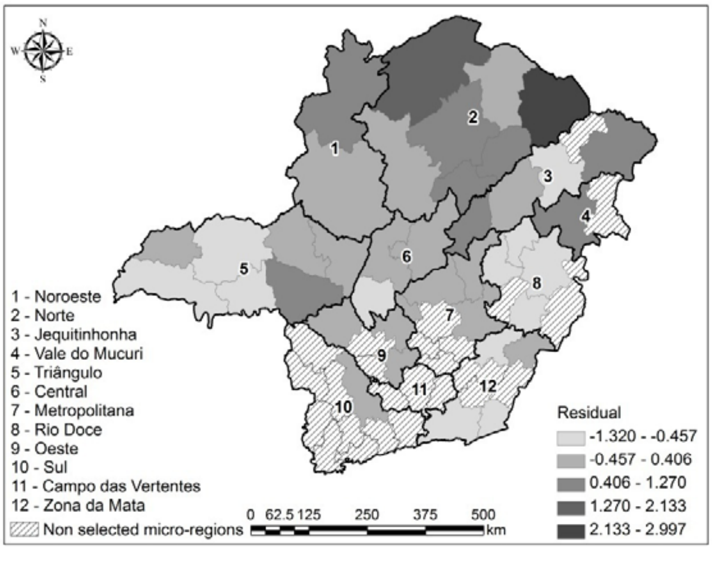
These results confirm the hypothesis that the conventional regression or OLS is not appropriated for spatial data, because the model does not consider the spatial non-stationarity neither spatial dependence, resulting in autocorrelated residuals (GAO; LI, 2011; TU; XIA, 2008; WINDLE et al., 2009). Figure 16 shows the OLS model residuals.

Table 11 Moran's I values for OLS model residuals for each aggregation level

Aggregation levels	Aggregation level 1	Aggregation level 2	Aggregation level 3
Moran's I	0.10	0.23	0.13
Z score (std. deviation)	5,20	3.36	1.15



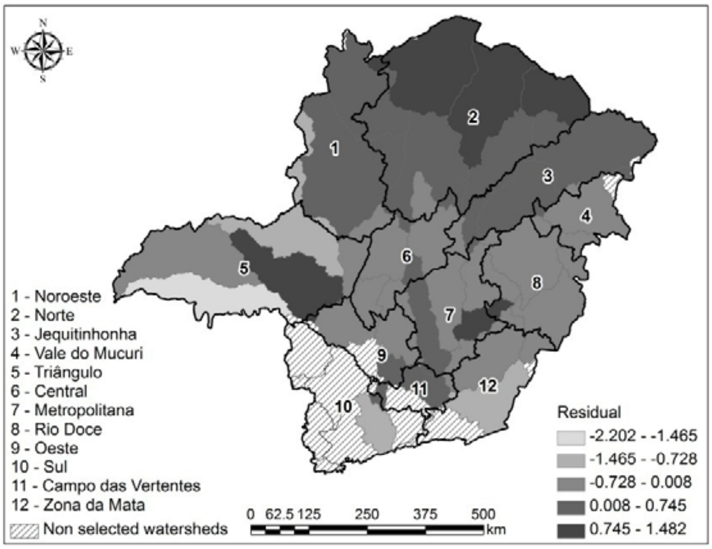
a)



b)

Figure 15 Spatial distribution of OLS residuals. a) Aggregation level 1; b) Aggregation level 2; c) Aggregation level 3

(…continue…)



c)

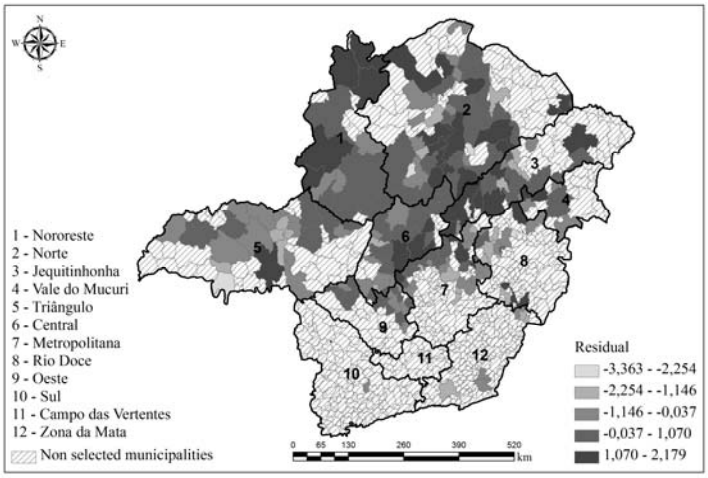
It is possible to identify some clusters in residuals located in the Norte region at all aggregation levels (Figures 16a, b, c).

Spatial autocorrelation is a requirement for the application of GWR (ZHAO; YANG; ZHOU, 2010), improving the performance compared to global regressions. A completely random spatial pattern of GWR model residuals was found at all aggregation levels (Table 14).

Additionally, Figure 17 shows the spatial distribution of GWR model residuals for the three aggregation levels. No clustering pattern of the GWR residuals can be identified across space for all the aggregation levels. The Moran's I analysis confirms the evidence that the GWR modeling is appropriate when dealing with spatial data, especially for LUCC modeling. These results are in accordance with the previous study conducted by Zhao, Yang and Zhou (2010) who applied GWR to estimate the effect of climate and site conditions on vegetation distribution in China. These authors also compared the results obtained from a GWR model and an OLS model using the Moran's I. They found spatial autocorrelation in OLS residuals and none in the GWR residuals.

Table 12 Moran's I values for GWR model residuals for each aggregation level

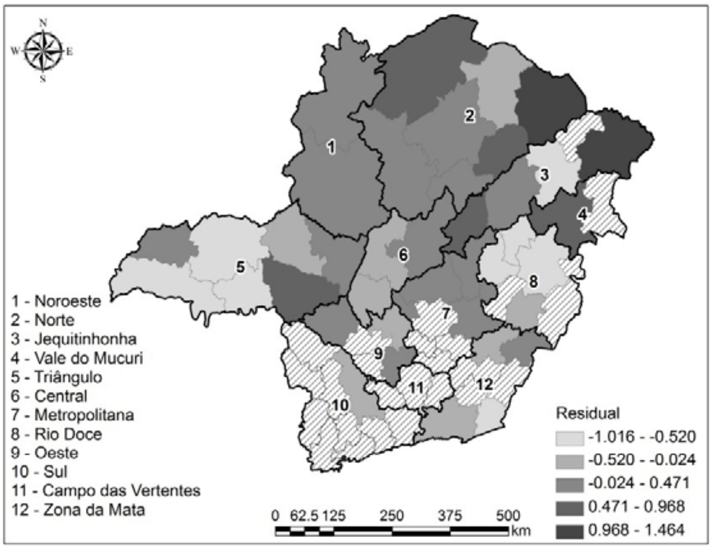
Aggregation levels	Aggregation level 1	Aggregation level 2	Aggregation level 3
Moran's I	0.06	0.02	0.08
Z score (std. deviation)	1.12	0.6	0.79



a)

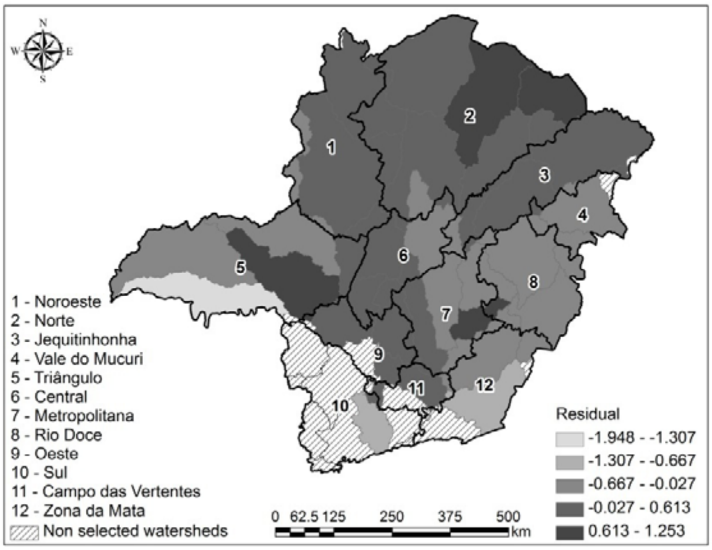
Figure 16 Spatial distribution of GWR residuals. a) Aggregation level 1; b) Aggregation level 2; c) Aggregation level 3

(…continue…)



1 - Noroeste
2 - Norte
3 - Jequitinhonha
4 - Vale do Mucuri
5 - Triângulo
6 - Central
7 - Metropolitana
8 - Rio Doce
9 - Oeste
10 - Sul
11 - Campo das Vertentes
12 - Zona da Mata
Residual
-1.016 - -0.520
-0.520 - -0.024
-0.024 - 0.471
0.471 - 0.968
0.968 - 1.464
0 62.5 125 250 375 500 km

b)



1 - Noroeste
2 - Norte
3 - Jequitinhonha
4 - Vale do Mucuri
5 - Triângulo
6 - Central
7 - Metropolitana
8 - Rio Doce
9 - Oeste
10 - Sul
11 - Campo das Vertentes
12 - Zona da Mata
Non selected watersheds
Residual
-1.948 - -1.307
-1.307 - -0.667
-0.667 - -0.027
-0.027 - 0.613
0.613 - 1.253
0 62.5 125 250 375 500 km

c)

7 CONCLUSIONS AND CONSIDERATIONS

This study examined the relationship between deforestation and its driving forces in the state of Minas Gerais, Brazil using three different aggregation levels. The use of GWR empirical models was motivated by the necessity of mapping the spatial distribution of the most important driving forces of deforestation across the state. The use of three aggregation levels was motivated by the importance of scale issues when performing LUCC modeling. To date, no study has examined these driving forces at different scales in Minas Gerais. Modeling the driving forces of deforestation in the state of Minas Gerais provided a valuable contribution to the LUCC science, as well as guidance to the decision-making process.

Comparing OLS and GWR techniques, the GWR performance presented improvements to model the driving forces behind deforestation. Additionally, both techniques showed different results at the three considered spatial scales.

Our exploratory analysis showed that the most significant driving forces of deforestation are shortest distance to roads, charcoal production and monoculture forest area at Aggregation level 1, and charcoal production and monoculture forest area at Aggregation levels 2 and 3. These forces are related to the current pressure on natural resources in Minas Gerais and should be investigated in more detail by governmental agencies. In this study, the deforestation database did not include information about the illegalness of deforestation activities, which might provide further insights on the causes of deforestation. Recent studies have shown that there still exists an increased use of wood from native forests of the *Cerrado* biome caused by the rising costs of wood from monoculture forest areas (REZENDE; SANTOS, 2010). According to the authors, charcoal is the most important product in the agro-forestry business in Minas Gerais and the state presents the largest monoculture forest area in Brazil. The exploratory analysis conducted in the present study confirms that charcoal production and monoculture forest area are important activities at all levels of aggregation.

The strong relationship with deforestation presented by charcoal production and monoculture forest area do not necessarily imply that they are the causative agents of deforestation, though these variables represent important economic activities for the state and are certainly related to other aspects not investigated in this study such as political and technological factors.

The identification of the most important variables related to deforestation may assist the choice of specific actions and optimal locations for surveillance of deforestation. However, the government decision-makers should be aware of the scale used in the analysis. For actions concerning local scales, Aggregation level 1 should be used with emphasis to municipalities showing the highest relationship between variables. Aggregated data might be considered for actions concerning the definition of more general statewide policy.

For a better evaluation of which factors are behind deforestation in the state of Minas Gerais, future studies should incorporate additional proximate and underlying variables such as technological, public policies, environmental and cultural factors. Models capable of projecting changes according to anticipated scenarios may be developed with the variables identified in the present study to support governamental strategic plans for the creation of protected areas and for the definition of effective deforestation monitoring systems.

Finally, it is important to highlight the non-stationary nature of the relationships between deforestation and its driving forces across the state of Minas Gerais. GWR analysis revealed that different factors determine forest loss in the state of Minas Gerais and that the adjusted GWR model is sensitive to variation in the aggregation levels. GWR is a helpful exploratory method to identify the driving forces of deforestation at different aggregation levels. The parameter estimates exhibit significant variation for all aggregation levels confirming the influence of MAUP in GWR models. Thus, the MAUP must be investigated into LUCC modeling as for each scale of analysis the researcher is likely to find different relationships between the driving forces and deforested areas across the space.

REFERENCES

AGUIAR, A. P. D.; CÂMARA, G.; ESCADA, M. I. S. Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: exploring intra-regional heterogeneity. **Ecological Modelling**, New York, v. 209, n. 2/4, p. 169-188, Apr. 2007.

AKAIKE, H. New look at statistical-model identification. **IEEE Transactions on Automatic Control**, New York, v. AC19, n. 6, p. 716-723, 1974.

ASPINALL, R. Modelling land use change with generalized linear models: a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. **Journal of Environmental Management**, New York, v. 72, n. 1/2, p. 91-103, Apr. 2004.

AUSTIN, M. Species distribution models and ecological theory: a critical assessment and some possible new approaches. **Ecological Modelling**, New York, v. 200, n. 1/2, p. 1-19, Feb. 2007.

BIERMAN, P. et al. A review of methods for analysing spatial and temporal patterns in coastal water quality. **Ecological Indicators**, London, v. 11, n. 1, p. 103-114, Jan. 2011.

BLAIR, J. B.; RABINE, D. L.; HOFTON, M. A. The laser vegetation imaging sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v. 54, n. 2/3, p. 115-122, 1999.

BROWN, D. G.; PIJANOWSKI, B. C.; DUH, J. D. Modeling the relationships between land use and land cover on private lands in the Upper Midwest, USA. **Journal of Environmental Management**, New York, v. 59, n. 4, p. 247-263, June 2000.

BRUNSDON, C.; FOTHERINGHAM, S.; CHARLTON, M. Geographically weighted regression: modelling spatial non-stationarity. **Journal of the Royal Statistical Society Series D-the Statistician**, London, v. 47, n. 3, p. 431-443, Aug. 1998.

CAHILL, M.; MULLIGAN, G. Using geographically weighted regression to explore local crime patterns. **Social Science Computer Review**, Durham, v. 25, n. 2, p. 174-193, 2007.

CÂMARA, G.; VALERIANO, D. M.; SOARES, J. V. **Metodologia para o cálculo da taxa anual de desmatamento na Amazônia Legal**. São José dos Campos: INPE, 2006. Available at: <http://www.obt.inpe.br/prodes/metodologia.pdf>. Access at: 26 Oct. 2009.

CARVALHO, L. M. T. **Mapping and monitoring Forest remmands:** a multiscale analysis of spatio-temporal data. 2001. 150 p. Thesis (Ph.D. in Production Ecology and Resources Conservation) - Wageningen University, Wageningen, 2001.

CARVALHO, L. M. T.; SCOLFORO, J. R. S. **Inventário florestal de Minas Gerais:** monitoramento da flora nativa 2005-2007. Lavras: UFLA, 2008. v. 1, 357 p.

CARVALHO, L. M. T.; SCOLFORO, J. R.; CAVALCANTI, H. C. Ocupação das áreas desflorestadas no Estado de Minas Gerais: atores e causas da modificação do uso do solo. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 14., 2009, Natal. **Anais...** Natal: SBSR/INPE, 2009. p. 5687-5694.

CHATERJEE, S.; HADI, A. S. **Regression analysis by example**. New York: Wiley Interscience, 2006. 375 p.

CHOMITZ, K. M.; THOMAS, T. S. Determinants of land use in Amazonia: a fine-scale spatial analysis. **American Journal of Agricultural Economics**, Saint Paul, v. 85, n. 4, p. 1016-1028, June 2003.

CLEMENT, F. et al. Drivers of afforestation in Northern Vietnam: assessing local variations using geographically weighted regression. **Applied Geography**, Oxford, v. 29, n. 4, p. 561-576, Apr. 2009.

DARK, S. J.; BRAM, D. The modifiable areal unit problem (MAUP) in physical geography. **Progress in Physical Geography**, London, v. 31, n. 5, p. 471-479, May 2007.

DEFRIES, R. et al. Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. **Environmental Science & Policy**, Exeter, v. 10, n. 4, p. 385-394, Aug. 2007.

ECHEVERRIA, C. et al. Spatially explicit models to analyze forest loss and fragmentation between 1976 and 2020 in southern Chile. **Ecological Modelling**, Amsterdam, v. 212, n. 3/4, p. 439-449, Apr. 2008.

ENTWISLE, B. et al. Population growth and its spatial distribution as factors in the deforestation of Nang Rong, Thailand. **Geoforum**, Oxford, v. 39, n. 2, p. 879-897, June 2008.

EVANS, T. P.; KELLEY, H. Multi-scale analysis of a household level agent-based model of landcover change. **Journal of Environmental Management**, New York, v. 72, n. 1/2, p. 57-72, Feb. 2004.

FARROW, A. et al. Exploring the spatial variation of food poverty in Ecuador. **Food Policy**, Guildford, v. 30, n. 5/6, p. 510-531, Dec. 2005.

FEARNSIDE, P. M. Deforestation in Brazilian Amazonia: history, rates, and consequences. **Conservation Biology**, Boston, v. 19, n. 3, p. 680-688, Sept. 2005.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **UN collaborative programme on reducing emissions from deforestation and forest degradation in developing countries (UN-REDD)**. Rome, 2008. Available at: <http://www.undp.org/mdtf/un-redd/docs/Annex-A-Framework-Document.pdf>. Access at: 20 Mar. 2009.

FOODY, G. M. Geographical weighting as a further refinement to regression modelling: an example focused on the NDVI-rainfall relationship. **Remote Sensing of Environment**, New York, v. 88, n. 3, p. 283-293, Mar. 2003.

FOTHERINGHAM, A. S.; BRUNDSON, C.; CHARLTON, M. E. **Geographically weighted regression:** analysis of spatially varying relationships. Chichester: Wiley, 2002. 266 p.

______. Geographically weighted regression: natural evolution of the expansion method for spatial data analysis. **Environment and Planning A**, London, v. 30, n. 11, p. 1905-1927, Nov. 1998.

FOTHERINGHAM, A. S.; WONG, D. W. S. The modifiable areal unit problem in multivariate statistical analysis. **Environment and Planning A**, London, v. 23, n. 6, p. 1025-1044, June 1991.

FRESCO, L. O. Imaginable futures, a contribution to thinking about land use planning. In: FRESCO, L. O. et al. (Ed.). **The future of the land:** mobilising and integrating knowledge for land use options. Chechester: J. Wiley, 1994. p. 137-156.

FUNDAÇÃO JOÃO PINHEIRO. **Produto interno bruto de Minas Gerais:** municípios e regiões. Belo Horizonte, 2002. Available at: <http://www.fjp.gov.br>. Access at: 20 Mar. 2010.

FUNDAÇÃO SOS MATA ATLÂNTICA; INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Atlas dos remanescentes florestais da Mata Atlântica período 2008-2010:** dados parciais dos estados avaliados até maio de 2010. São Paulo; São José dos Campos, 2010. 156 p.

GAO, J.; LI, S. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using geographically weighted regression. **Applied Geography**, Oxford, v. 31, n. 1, p. 292-302, Jan. 2011.

GEIST, H. J.; LAMBIN, E. F. Proximate causes and underlying driving forces of tropical deforestation. **Bioscience**, Washington, v. 52, n. 2, p. 143-150, Feb. 2002.

GEOGHEGAN, J. et al. Modeling tropical deforestation in the southern Yucatán peninsular region: comparing survey and satellite data. **Agriculture, Ecosystems and Environment**, Amsterdam, v. 85, n. 1/3, p. 25-46, Mar. 2001.

GILBERT, A.; CHAKRABORTY, J. Using geographically weighted regression for environmental justice analysis: cumulative cancer risks from air toxics in Florida. **Social Science Research**, Wallingford, v. 40, n. 1, p. 273-286, Jan. 2011.

HARRIS, P.; FOTHERINGHAM, S.; JUGGINS, S. Robust geographically weighted regression: a technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes. **Annals of the Association of American Geographers**, Washington, v. 100, n. 2, p. 286-306, June 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo demográfico 2010**. Rio de Janeiro, 2010. Available at: <http://www.ibge.gov.br>. Access at: 10 Dec. 2010.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Monitoramento da cobertura florestal da Amazônia por satélites**. São José dos Campos, 2008. Available at: <http://www.obt.inpe.br/deter/metodologia_v2.pdf>. Access at: 19 May 2010.

IRWIN, E. G.; GEOGHEGAN, J. Theory, data, methods: developing spatially explicit economic models of land use change. **Agriculture, Ecosystems & Environment**, Amsterdam, v. 85, n. 1/3, p. 7-23, Mar. 2001.

JAIMES, N. B. P. et al. Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. **Applied Geography**, Oxford, v. 30, n. 4, p. 576-591, Apr. 2010.

JANTZ, C. A.; GOETZ, S. J. Analysis of scale dependencies in an urban land-use-change model. **International Journal of Geographical Information Science**, London, v. 19, n. 2, p. 217-241, Apr. 2005.

JELINSKI, D. E.; WU, J. G. The modifiable areal unit problem and implications for landscape ecology. **Landscape Ecology**, Dordrecht, v. 11, n. 3, p. 129-140, 1996.

KIRBY, K. R. et al. The future of deforestation in the Brazilian Amazon. **Futures**, Guildford, v. 38, n. 4, p. 432-453, Apr. 2006.

KOK, K. **Scalling the land use system:** a modeling approach with case studies for Central America. 2001. 153 p. Thesis (Ph.D. in Land Dynamics) - Wageningen University, Wageningen, 2001.

KOUTSIAS, N.; MARTÍNEZ-FERNÁNDEZ, J.; ALLGOWER, B. Do factors causing wildfires vary in space?: evidence from geographically weighted regression 2010. **GIScience & Remote Sensing**, New York, v. 47, n. 2, p. 221-240, June 2010.

LAMBIN, E. F. et al. The causes of land-use and land-cover change: moving beyond the myths. **Global Environmental Change-Human and Policy Dimensions**, Oxford, v. 11, n. 4, p. 261-269, Dec. 2001.

LAMBIN, E. F.; GEIST, H.; RINDFUSS, R. R. Introduction: local processes with global impacts. In: LAMBIN, E. F.; GEIST, H. J. (Ed.). **Land-use and land-cover change:** local process and global impacts. Berlin: Springer, 2006. p. 1-8.

LAURANCE, W. F. The future of the Brazilian Amazon. **Science**, New York, v. 291, n. 5506, p. 988-988, 2001.

LAURANCE, W. F. et al. Predictors of deforestation in the Brazilian Amazon. **Journal of Biogeography**, Oxford, v. 29, n. 5/6, p. 737-748, Dec. 2002.

LEGENDRE, P. Spatial autocorrelation: trouble or new paradigm. **Ecology**, Durham, v. 74, n. 6, p. 1659-1673, Dec. 1993.

LIU, M. et al. Land use and land cover change analysis and prediction in the upper reaches of the Minjiang River, China. **Environmental Management**, New York, v. 43, n. 5, p. 899-907, Sept. 2009.

MANSON, S. M. Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico. **Agriculture, Ecosystems & Environment**, Amsterdam, v. 111, n. 1/4, p. 47-62, Apr. 2005.

MARCEAU, D. F.; HOWARTH, P. J.; GRATTON, D. J. Remote-sensing and the measurement of geographical entities in a forested environment: 1., the scale and spatial aggregation problem. **Remote Sensing of Environment**, New York, v. 49, n. 1, p. 93-104, Jan. 1994a.

______. Remote-sensing and the measurement of geographical entities in a forested environment: 2., the optimal spatial resolution. **Remote Sensing of Environment**, New York, v. 49, n. 1, p. 105-117, Jan. 1994b.

MARCEAU, D. J. The scale issue in the social and natural sciences. **Canadian Journal of Remote Sensing**, Ottawa, v. 25, n. 4, p. 347-356, Aug. 1999.

MARGULIS, S. **Causes of deforestation in Brazilian Amazon**. Washington: World Bank, 2004. 101 p. (World Bank Working Paper, 22).

MAS, J. F. et al. Modelling deforestation using GIS and artificial neural networks. **Environmental Modelling & Software**, Oxfordhire, v. 19, n. 5, p. 461-471, Sept. 2004.

MATHEY, A. H. et al. An object-oriented cellular automata model for forest planning problems. **Ecological Modelling**, Amsterdam, v. 212, n. 3/4, p. 359-371, Apr. 2008.

MENA, C. F.; BILSBORROW, R. E.; MCCLAIN, M. E. Socioeconomic drivers of deforestation in the Northern Ecuadorian Amazon. **Environmental Management**, New York, v. 37, n. 6, p. 802-815, Dec. 2006.

MENARD, A.; MARCEAU, D. J. Simulating the impact of forest management scenarios in an agricultural landscape of southern Quebec, Canada, using a geographic cellular automata. **Landscape and Urban Planning**, London, v. 79, n. 3/4, p. 253-265, Dec. 2007.

MERTENS, B.; LAMBIN, E. F. Spatial modelling of deforestation in southern Cameroon: spatial disaggregation of diverse deforestation processes. **Applied Geography**, Oxford, v. 17, n. 2, p. 143-162, Feb. 1997.

MILLER, J.; FRANKLIN, J.; ASPINALL, R. Incorporating spatial dependence in predictive vegetation models. **Ecological Modelling**, Amsterdam, v. 202, n. 3/4, p. 225-242, Oct. 2007.

MILLINGTON, J. D. A.; PERRY, G. L. W.; ROMERO-CALCERRADA, R. Regression techniques for examining land use/cover change: a case study of a mediterranean landscape. **Ecosystems**, New York, v. 10, n. 1, p. 562-578, Mar. 2007.

MINCHELLA, A. et al. Use of multitemporal SAR data for monitoring vegetation recovery of Mediterranean burned areas. **Remote Sensing of Environment**, New York, v. 113, n. 3, p. 588-597, Mar. 2009.

MINETOS, D.; POLYZOS, S. Deforestation processes in Greece: a spatial analysis by using an ordinal regression model. **Forest Policy and Economics**, London, v. 12, n. 6, p. 457-472, Dec. 2010.

MORAN, P. A. P. Notes on continuous stochastic phenomena. **Biometrika**, London, v. 37, n. 1/2, p. 17-23, 1950.

MOREIRA, E. et al. Dynamical coupling of multiscale land change models. **Landscape Ecology**, Dordrecht, v. 24, n. 9, p. 1183-1194, Sept. 2009.

MYERS, N. et al. Biodiversity hotspots for conservation priorities. **Nature**, London, v. 403, n. 6772, p. 853-858, 2000.

OGNEVA-HIMMELBERGER, Y.; PEARSALL, H.; RAKSHIT, R. Concrete evidence & geographically weighted regression: a regional analysis of wealth and the land cover in Massachusetts. **Applied Geography**, Oxford, v. 29, n. 4, p. 478-487, Apr. 2009.

OPENSHAW, S.; TAYLOR, P. J. A million or so correlation coefficients: three experiments on the modiafiable areal unit problem. In: WRIGLEY, N. (Ed.). **Statistical applications in the spatial science**. London: Routledge, 1979. p. 127-144.

PELOROSSO, R.; LEONE, A.; BOCCIA, L. Land cover and land use change in the Italian central Apennines: a comparison of assessment methods. **Applied Geography**, Oxford, v. 29, n. 1, p. 35-48, Jan. 2009.

PFAFF, A. S. P. What drives deforestation in the Brazilian Amazon?: evidence from satellite and socioeconomic data. **Journal of Environmental Economics and Management**, New York, v. 37, n. 1, p. 26-43, Feb. 1999.

REZENDE, J. B.; SANTOS, A. C. **A cadeia produtiva do carvão vegetal em Minas Gerais:** pontos críticos e potencialidades. Viçosa, MG: EPAMIG, 2010. 82 p. (Boletim Técnico, 95).

SANCHEZ-AZOFEIFA, G.; HARRISS, R. C.; SKOLE, D. L. Deforestation in Costa Rica: a quantitative analysis using remote sensing imagery. **Biotropica**, Washington, v. 33, n. 3, p. 378-384, Mar. 2001.

SCHNEIDER, L. C.; PONTIUS, R. G. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. **Agriculture, Ecosystems & Environment**, Amsterdam, v. 85, n. 1/3, p. 83-94, Mar. 2001.

SCOLFORO, J. R.; CARVALHO, L. M. T. **Mapeamento e inventário da flora nativa e dos reflorestamentos de Minas Gerais**. Lavras: UFLA, 2006. 288 p.

SERNEELS, S.; LAMBIN, E. F. Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. **Agriculture, Ecosystems & Environment**, Amsterdam, v. 85, n. 1/3, p. 65-81, Mar. 2001.

SLUITER, R. **Mediterranean land cover change:** modelling and monitoring natural vegetation using GIS and remote sensing. 2005. 145 p. Thesis (Ph.D. in Geoscience) - Utrecht University, Netherlands, 2005.

SOARES, B. S.; CERQUEIRA, G. C.; PENNACHIN, C. L. Dinamica: a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. **Ecological Modelling**, Amsterdam, v. 154, n. 3, p. 217-235, Sept. 2002.

SOLER, L. D.; ESCADA, M. I. S.; VERBURG, P. H. Quantifying deforestation and secondary forest determinants for different spatial extents in an Amazonian colonization frontier, Rondonia. **Applied Geography**, Oxford, v. 29, n. 2, p. 182-193, Feb. 2009.

TOLE, L. Sources of deforestation in tropical developing countries. **Environmental Management**, New York, v. 22, n. 1, p. 19-33, Mar. 1998.

TRISURAT, Y.; ALKEMADE, R.; VERBURG, P. H. Projecting land-use change and its consequences for biodiversity in Northern Thailand. **Environmental Management**, New York, v. 45, n. 3, p. 626-639, Mar. 2010.

TU, J.; XIA, Z. G. Examining spatially varying relationships between land use and water quality using geographically weighted regression I: model design and evaluation. **Science of the Total Environment**, Amsterdam, v. 407, n. 1, p. 358-378, Feb. 2008.

TULBURE, M. G. et al. Spatial and temporal heterogeneity of agricultural fires in the central United States in relation to land cover and land use. **Landscape Ecology**, Dordrecht, v. 26, n. 2, p. 211-224, Feb. 2010.

TURNER, M. G.; GARDNER, R. H.; O'NEILL, R. V. **Landscape ecology in theory and practice:** pattern and process. New York: Springer, 2001. 417 p.

UGHLI, A.; GOLDEMBERG, J.; COELHO, S. T. O uso de carvão vegetal na indústria siderúrgica brasileira e o impacto sobre as mudanças climáticas. **Revista Brasileira de Energia**, Rio de Janeiro, v. 14, n. 2, p. 67-85, mar./abr. 2008.

VELDKAMP, A.; FRESCO, L. O. CLUE: a conceptual model to study the conversion of land use and its effects. **Ecological Modelling**, Amsterdam, v. 85, n. 2/3, p. 253-270, Sept. 1996.

VERBURG, P. H.; CHEN, Y. Q. Multiscale characterization of land-use patterns in China. **Ecosystems**, New York, v. 3, n. 4, p. 369-385, Aug. 2000.

VERBURG, P. H. et al. Determinants of land-use change patterns in the Netherlands. **Environment and Planning B-Planning & Design**, London, v. 31, n. 1, p. 125-150, Mar. 2004.

______. Modeling land-use and land-cover change. In: LAMBIN, E. F.; GEIST, H. J. (Ed.). **Land-use and land-cover change:** local process and global impacts. Berlin: Springer, 2006. p. 117-135.

______. Spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. **Ecological Modelling**, Amsterdam, v. 116, n. 1, p. 45-61, Feb. 1999.

VERBURG, P. H.; VELDKAMP, W. S. A.; ESPALDON, R. L. V. Modeling the spatial dynamics of regional land use: the CLUE-S model. **Environmental Management**, New York, v. 30, n. 3, p. 391-405, Mar. 2002.

WANG, Q.; NI, J.; TENHUNEN, J. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. **Global Ecology and Biogeography**, Oxford, v. 14, n. 4, p. 379-393, Dec. 2005.

WINDLE, M. J. S. et al. Exploring the non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. **Ices Journal of Marine Science**, Dauphin, v. 67, n. 1, p. 145-154, 2009.

WITMER, F. **Simulating future global deforestation using geographically explicit models:** interim report IR 05-010. International Institute for Applied Systems Analysis, 2005. 31 p.

YU, D. L. Spatially varying development mechanisms in the Greater Beijing Area: a geographically weighted regression investigation. **Annals of Regional Science**, Berlin, v. 40, n. 1, p. 173-190, Mar. 2006.

ZHANG, L. J.; GOVE, J. H.; HEATH, L. S. Spatial residual analysis of six modeling techniques. **Ecological Modelling**, Amsterdam, v. 186, n. 2, p. 154-177, Feb. 2005.

ZHAO, N.; YANG, Y.; ZHOU, X. Application of geographically weighted regression in estimating the effect of climate and site conditions on vegetation distribution in Haihe Catchment, China. **Plant Ecology**, Dordrecht, v. 209, n. 2, p. 349-359, Feb. 2010.