

Revista Brasileira de Cartografia (2012) N<sup>o</sup> 64/5: 703-721  
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto  
ISSN: 1808-0936

## **ANÁLISE DA DISTRIBUIÇÃO DE FOGOS ATIVOS DETECTADOS PELO SENSOR ATSR 2 E SUA RELAÇÃO COM VARIÁVEIS DEMOGRÁFICAS, AMBIENTAIS E INFRA-ESTRUTURAIS: COMPARABILIDADE ENTRE MÉTODOS DE REGRESSÃO OLS E GWR**

*Analysis of the Distribution of Active Fires Detected by Sensor ATSR 2 and its  
Relationship with Demographic, Environmental and Infrastructure Variables:  
Comparison Between the Methods of OLS Regression and GWR*

**Adélia Maria Oliveira de Sousa<sup>1</sup>; José Miguel Oliveira Cardoso Pereira<sup>2</sup>;  
Bernardo Wildung Cantante Mota<sup>2</sup> & João Roberto dos Santos<sup>3</sup>**

**<sup>1</sup>Universidade de Évora**

Escola de Ciências e Tecnologia, Departamento de Engenharia Rural,  
Instituto de Ciências Agrárias e Ambientais Mediterrânicas  
Apartado 94, 7002-554 Évora, Portugal  
asousa@uevora.pt

**<sup>2</sup>Universidade Técnica de Lisboa**

Instituto Superior de Agronomia  
Tapada da Ajuda, 1394-017 Lisboa, Portugal  
jmcpereira@isa.utl.pt  
bwcmota@gmail.com

**<sup>3</sup>Instituto Nacional de Pesquisas Espaciais – INPE**

Divisão de Sensoriamento Remoto - DSR  
Av. dos Astronautas, 1758 - Jardim da Granja,  
São José dos Campos, SP - 12227-010, Brasil  
jroberto@dsr.inpe.br

*Recebido em 19 de dezembro, 2011 / Aceito em 11 de abril, 2012  
Received on december 19, 2011/ Accepted on april 11, 2012*

### **RESUMO**

Neste estudo foi analisado a relação entre o produto dos fogos activos ATSR com variáveis ambientais, demográficas e infra-estruturas para o território Brasileiro. Foram comparados os métodos de regressão por mínimos quadrados (OLS) e regressão local ponderada (GWR). O método GWR permite a análise de regressão considerando as variáveis não-estacionárias espacialmente. Foi testada a hipótese do método GWR representar uma melhoria sobre o método de regressão global. O melhor ajustamento dos fogos activos ATSR foi obtido com o método GWR quando foi considerado

fatores humanos, como as variáveis de presença de áreas agrícolas e de distância às estradas, com  $R^2$  de 80%. Este estudo permite um melhor entendimento da distribuição espacial dos fogos ativos ATSR e mostra que o método GWR é um complemento válido relativamente aos métodos de análise espacial global.

**Palavras chaves:** Brasil; Fogos Ativos ATSR; GWR; OLS; Análise Espacial Não-Estacionária; Monitoramento.

## ABSTRACT

In this study it was analysed the relationship between the active fire ATSR products and environmental, demographic and infrastructure variables for the Brazilian territory. It was compared the Ordinary Least Square regression (OLS) and the Geographic Weighted Regression (GWR) methods. The method GWR allows the analysis of regression considering spatially non-stationary variables. The hypothesis that GWR represents no improvement over the global model was tested. The best adjustment of the ATSR active fires was found with the GWR method when it was considered the human factors, like agricultural area classification and distance to roads variables, with  $R^2 = 80\%$ . This study provides an improved understanding of spatial ATSR active fire relationships and shows that GWR is a valuable complement to global spatial analysis methods.

**Keywords:** Brazil; ATSR Active Fires; GWR; OLS; Non-stationary Spatial Analysis; Monitoring.

## 1. INTRODUÇÃO

O fogo no Brasil tem sido objeto de preocupação atingindo diversos sistemas ecológicos gerando impactos ambientais em escalas globais, regionais e locais. A atividade do fogo tem sido relacionada com as condições climáticas (sêcas severas), o aumento da fronteira agrícola e/ou melhoria da palatabilidade das áreas de pastoreio (KAUFFMAN *et al.*, 1988) e também, com a construção de vias de acesso no processo de extração madeireira, ou mesmo, de implantação de assentamento fundiário (NEPSTAD *et al.* 1999, CARDOSO *et al.*, 2003; LAURENCE *et al.*, 2002).

O estudo aqui apresentado tem como objetivo principal avaliar a relação de diversos fatores com a ocorrência e distribuição geográfica do fogo no Brasil, fazendo-se uso de procedimentos de análise de regressão. Um método convencional de análise citado na literatura, de larga utilização, é o de regressão por mínimos quadrados, ou *Ordinary Least Square* (OLS). Este método é uma forma expedita para estimar alguns parâmetros globais que permitem a análise da relação entre duas ou mais variáveis, relação esta constante no espaço. Essa regressão global tem algumas limitações, podendo seu uso não ser apropriado para algumas situações, o que torna necessário considerar outras alternativas (FOODY, 2003; COHEN *et al.*, 2003). Recentemente, outras técnicas de regressão têm sido adotadas em estudos de análise espacial, procurando melhorar os resultados decorrentes do tradicional método de regressão global.

Diversos estudos têm sido desenvolvidos com o objetivo de avaliar a relação espacial entre observações decorrentes de trabalho de campo como àquelas por sensoriamento remoto. Com dados obtidos por sensoriamento remoto este tipo de dados, FOODY (2003 e 2004) desenvolveu estudos da variação da diversidade de espécies de aves no continente Africano, usando dados de temperatura, precipitação e do índice de vegetação NDVI. WANG *et al.* (2005) modelaram a produção primária líquida a partir de dados do índice de vegetação NDVI. SHI *et al.* (2006), estudaram a distribuição de uma espécie de veado usando dados de cobertura do solo e fatores climáticos. SÁ *et al.* (2011) estudaram a relação entre a incidência do fogo e alguns fatores ambientais para a África sub-Saariana. O método de regressão OLS empregado dentre outros, não foi o mais apropriado, visto que não considera a variação espacial das relações entre variáveis. Um dos métodos então propostos foi uma regressão local ponderada, designado por *Geographically Weighted Regression* (GWR). Este método permite estimar parâmetros estatísticos considerando as coordenadas locais dos dados em análise, ou seja tendo em conta a sua variação espacial (FOTHERINGHAM *et al.*, 2002). O método de regressão local surge como o mais apropriado para estudos em áreas de grandes dimensões, onde a variação espacial é mais acentuada. Esta técnica fornece informação ponderada associada com cada ponto geográfico e permite estimar parâmetros de regressão, levando em conta a sua variação no espaço, o que pode revelar variações nas relações

entre as variáveis que, em outros tipos de análise seriam ignoradas (WANG *et al.*, 2005).

Para atingir os objetivos propostos apresentamos uma comparação entre o método tradicional de regressão global (OLS) e o recente método de regressão local (GWR), considerando como variável dependente o número de fogos ativos obtidos pelo sensor *Along Track Scanning Radiometer* (ATSR) designados por *World Fires Atlas* (WFA) para o território brasileiro. Como variáveis independentes foram consideradas a densidade populacional, a produtividade primária líquida, um índice climático (*Chandler Burning Index*), a percentagem de área de floresta, cerrado e agricultura e a distância às estradas.

## 2. DADOS SENSORIADOS

### 2.1 Fogos ativos

Os dados referentes aos fogos ativos utilizados foram derivados do sensor ATSR-2, lançado pela *European Space Agency's* (ESA) a bordo do satélite *European Remote Sensing Satellite* (ERS-2), integrado no Programa de Observação da Terra. Este sensor está equipado com três bandas na região do espectro eletromagnético do visível, uma banda no infravermelho médio e três bandas no infravermelho térmico. O ATSR produz dados de sensoriamento remoto com uma resolução espacial de 1 km no nadir. Estes dados produzem informação da ocorrência de fogos no seu estado ativo pela aplicação de dois algoritmos, usando dois limiares distintos, um de 312 °K e outro de 308 °K na radiância da banda centrada em 3,7 mm (ARINO & ROSAZ, 1999). Esta banda é muito sensível à radiância emitida a temperaturas entre 500 °K e 1000 °K, permitindo a detecção de fogo ativo mesmo que apenas uma pequena parte do pixel seja afetada pelo fogo.

Os produtos de fogos ativos deste sensor designados pelo *World Fires Atlas* (WFA), são produzidos apenas a partir de observações recolhidas durante a noite. Este fato minimiza os falsos alarmes devido à irradiância solar, como a reflexão por nuvens e por superfícies brilhantes no solo. Porém, MOTA *et al.* (2006) verificaram que mesmo assim, cerca de 25% das observações do WFA eram falsos alarmes, mostrando ser necessária a aplicação de diversos filtros para minimizar estes possíveis falsos alarmes.

Para a eliminação de falsos alarmes e de fogos que não tenham origem em vegetação foram dois os procedimentos implementados. O primeiro, a aplicação de uma série de filtros geográficos. O segundo procedimento foi a eliminação de alguns falsos alarmes, por inspeção visual, para identificar observações que ainda não tivessem sido eliminadas no primeiro processo e que porventura, ainda levassem a uma classificação errônea.

Esses dados assim corrigidos de alguns erros passam a designar-se como produto WFA corrigido (WFAc). Uma descrição mais pormenorizada da eliminação dos falsos alarmes pode ser encontrada em MOTA *et al.* (2006).

Os dados de entrada para o presente estudo foram o produto WFAc, correspondente à média do número de fogos para os seis anos (1997 a 2002), agregados em células de 0,5°, correspondendo a 2750 observações.

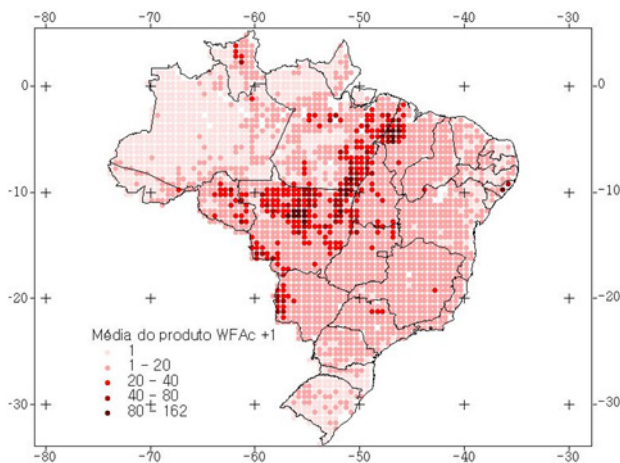
Devido à existência de células com número de fogos igual a zero, adicionou-se uma constante de valor 1 para todas as observações. Os pontos com o valor 1 indicam as células onde efetivamente não há registros de fogos para os seis anos (Fig. 1). A modificação da escala dos dados para o logaritmo da média deve-se à ocorrência de elevado número de fogos num baixo número de células, o que criava valores extremos muito elevados. Com esta modificação suavizaram-se estes valores extremos. Assim, os dados dos fogos ativos sobre os quais se desenvolveu o presente estudo correspondem à seguinte equação:

$$\text{Produto WFAc} = \text{Ln}(Y + 1) \quad (1)$$

onde *Ln* corresponde ao logaritmo de base natural e *Y* representa a contagem dos fogos ativos.

Aos dados apresentados na Fig. 1, já foram excluídos as áreas correspondentes às massas de água, identificados pelo mapa de ocupação do solo para a América do Sul GLC2000 (EVA *et al.*, 2002).

Também foram excluídos os dados referentes a valores de produtividade primária líquida nula, que corresponde a uma das variáveis usadas neste estudo descrita a seguir. Este procedimento foi efetuado através da aplicação de filtros geográficos aos dados das médias dos seis anos do produto WFAc.



**Fig. 1** - Média do produto WFAc, (1997 - 2002), numa grade com a resolução espacial de 0,5°, para o Brasil. Os pontos a branco correspondem a células cujos dados foram eliminados.

## 2.2 Dados demográficos, ambientais e infra-estruturais

### 2.2.1 Densidade populacional

A atividade humana está referida na bibliografia, como um fator impulsionador da ocorrência de fogos no Brasil. No entanto, este fator está sempre associado com outros aspectos que contribuem para a ocorrência de fogos, como a construção de vias de acesso (CARDOSO *et al.*, 2005).

A distribuição espacial da densidade populacional foi considerada um importante fator na tentativa de explicar a ocorrência de fogos neste país. Os dados desta variável foram obtidos através do *Gridded Population of the World (GPW)*, versão 3 (<http://sedac.ciesin.columbia.edu/plue/gpw/>), para o ano 2000, com uma resolução espacial de 30 segundos de arco (0,00833°), que corresponde aproximadamente a 1 km para a latitude média do Brasil.

### 2.2.2 Produtividade Primária Líquida

A produtividade primária líquida (PPL) representa a quantidade de biomassa vegetal acumulada por unidade de tempo e unidade de área, e depende do clima, do solo e do tipo de cobertura da superfície terrestre (ABER & MELILLO, 2001). A PPL apresenta valores elevados na floresta tropical, com vegetação densa e umidade no solo, decrescendo para os ecossistemas de Cerrado e pastos, onde a vegetação é menos densa, ocorrendo os valores mais baixos nas zonas áridas, no nordeste

do Brasil. É esperado que a maior incidência do fogo espera que se concentre em níveis intermédios de PPL. Segundo ABER & MELILLO (2001), a PPL média para a floresta tropical é de 900 gC/m<sup>2</sup>/ano, para o Cerrado é de 315 gC/m<sup>2</sup>/ano, para pastos de 225 gC/m<sup>2</sup>/ano e para áreas cultivadas é 290 gC/m<sup>2</sup>/ano.

Existem diversos modelos de estimativa da PPL à escala global, na sua maioria dependentes de observações climatéricas e de médias de biomassa. Neste estudo os dados de PPL para o Brasil foram obtidos do *Global Production Efficiency Model (GLOPEM)* (<http://glcf.umiacs.umd.edu/data/glopem/>) desenvolvido pelo Departamento de Geografia da Universidade de Maryland (USA). Este produto foi gerado com base na observação do estado da vegetação para um período de oito anos, através do *Normalized Difference Vegetation Index (NDVI)* e da temperatura de superfície (CHURKINA *et al.*, 1999), a partir de dados do sensor AVHRR, com resolução espacial de 8km (GOETZ *et al.*, 1999).

### 2.2.3 Índice climático

O índice climático *Chandler Burning Index (CBI)*, desenvolvido por CHANDLER *et al.* (1983) é utilizado para prever o risco de ocorrência de fogo e baseia-se no máximo da temperatura mensal e na temperatura do ponto de orvalho. Esta última é a temperatura a que o vapor de água contido na porção de ar de um determinado local sofre condensação. A ocorrência do fogo pode ser avaliada pela temperatura, umidade e pelo tipo de vegetação afetada. Nas regiões com temperaturas elevadas e valores baixos de umidade no solo, a vegetação tem baixo teor de água e está mais susceptível à ocorrência de fogo. Na situação oposta, com temperaturas baixas e umidade elevada no solo a vegetação está vigorosa, sendo mais difícil a ocorrência e propagação do fogo.

Os dados para calcular o CBI são puramente meteorológicos e resultam do *National Centers of Environmental Prediction / National Center for Atmospheric Research (NCEP / NCAR) Reanalysis 1* (<http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.html>). As variáveis usadas têm uma resolução temporal de seis horas: a temperatura na superfície e a umidade relativa são observadas quatro vezes por dia, entre 1 de Janeiro de 1997 e 31 de Dezembro de 2000 (período correspondente



à disponibilidade do produto WFA). A resolução espacial da grade para a qual existem os dados disponíveis é de 2,5° (aproximadamente 275 km). As quatro observações da temperatura de superfície permitem calcular a temperatura máxima diária,  $T_{max}$ . Considerando um processo de refrigeração com igual pressão atmosférica, a umidade relativa (HR) e a temperatura média mensal ( $T$ ), permitem o cálculo de temperatura do ponto de orvalho,  $T_{dew}$ , através da fórmula seguinte apresentada por IRIBARNE & GODSON (1981):

$$T_{dew} = T + 35(\log_{10}(0,01HR)) \quad (2)$$

O índice diário de umidade é calculado através da seguinte fórmula:

$$H = 100 \frac{[10^{0,0308T_{dew}}]}{10^{0,0308T_{max}}} \quad (3)$$

Considerando os efeitos de intensidade do fogo (I) e a sua velocidade de propagação, *Rate of Spread Index* (SI), o índice climático é determinado pela seguinte equação:

$$CBI = I \times SI \quad (4)$$

onde SI e I são dados por:

$$(5)$$

$$I = [1(SI = 124 \times 10^{-0,0142H - 0,054(T_{max})})] \quad (6)$$

Para cada um dos seis anos, calculou-se a média dos três valores mensais mais elevados do CBI. Na análise, usou-se apenas o mais elevado destes seis valores anuais. A selecção destes valores de cada ano deve-se à necessidade de garantir que extremos climáticos sejam representados, tal como os provocados pelo fenómeno *El Niño*. Quanto maior o valor deste índice maior é a probabilidade de ocorrerem fogos e de maiores dimensões tendo em conta as condições atmosféricas respectivas, ou seja, temperaturas elevadas e valores baixos de umidade.

Segundo CHANDLER *et al.* (1983), os valores obtidos para o CBI estão relacionados com o comportamento do fogo segundo a Tabela 1.

## 2.2.4 Uso e cobertura da terra

A ocorrência, severidade e extensão de área queimada está relacionada entre outros fatores, com o tipo e o estado da vegetação. As queimadas agrícolas são, na sua maior parte, controladas pelos agricultores, limitadas a parcelas definidas, na queima de resíduos de culturas anteriores, na utilização como fertilizantes em novas culturas, no desmatamento para obtenção de novas parcelas de terra ou ainda na renovação de pastagens (NEPSTAD *et al.*, 1999). No caso do desmatamento, as queimadas são realizadas sobre vegetação já cortada e seca. Contudo, há situações em que as queimadas fogem ao controle e se propagam para a vegetação circundante, são os chamados fogos de superfície, que tornam a floresta mais susceptível a futuras incidências. Caso a vegetação envolvente seja Cerrado, com porte arbóreo e/ou arbustivo e de elevada massa graminóide seca no substrato, a progressão da incidência do fogo pode ser facilitada para ocorrer em grandes dimensões, tanto pelo material senescente local como por direcções do vento.

Com o objetivo de avaliar a relação deste fenómeno com o tipo de ocupação do solo afetada, e com o mapa de vegetação GLC2000, procedeu-se ao cálculo da percentagem dos três tipos principais de ocupação do solo afetada por queimadas em cada célula de 0,5° de resolução espacial. Estes tipos de uso e cobertura da terra correspondem a Agricultura, Floresta e Cerrado que funcionam como três variáveis independentes.

## 2.2.5 Distância às estradas

A distância às estradas é um fator considerado de grande influência na ocorrência de queimadas no Brasil, sendo considerada a faixa de maior ocorrência até uma distância de 50 km às estradas pavimentadas e não pavimentadas (CARDOSO *et al.*, 2005; NEPSTAD *et al.*, 2001 e PERES, 2001). Segundo ALVES (2002), o desflorestamento ocorre numa faixa ao longo das principais vias de acesso, atingindo a distância máxima de 100 km às principais, que normalmente correspondem a estradas pavimentadas. O deslocamento da população para o interior do Brasil surge através dos acessos que são abertos na floresta, principalmente para permitir ligar cidades do interior entre si e destas às do litoral. A instalação das empresas madeireiras em áreas de floresta não

**Tabela 1** - Relação entre os valores de CBI e o comportamento do fogo. (Segundo CHANDLER *et al.*, 1983)

CBI	Comportamento do Fogo
1 – 19	Fogo rasteiro em vegetação arbustiva
20 – 39	Fogo de superfície, subcoberto
40 – 59	Fogo rápido de superfície com envolvimento ocasional das copas
60 – 79	Fogo rápido e intenso, de superfície, com envoltimentos frequente das copas
80	Fogo de copas provável

explorada também contribui para este fenômeno, onde o objetivo é o corte seletivo de árvores cuja madeira tem elevado valor comercial.

A abertura de clareiras de aproximadamente 50 ha ao longo destas vias é responsável por 75% do desmatamento total e 87% do desflorestamento observada entre 1991-1996.

Os dados referentes às estradas principais e secundárias do Brasil, foram obtidos do Instituto Brasileiro de Geografia e Estatística (IBGE, 1997). Com base no mapa digital das estradas, foi delimitada uma faixa de 50 km, correspondendo à zona de maior ocorrência do fogo relativamente às duas categorias de vias consideradas. Em seguida, procedeu-se ao cálculo da percentagem de células de 1 km que estão dentro da faixa de 50 km.

A Tabela 2 apresenta uma breve descrição das variáveis consideradas neste estudo, amostradas para uma grade com a resolução espacial de 0,5°.

### 2.3 Correlação entre as variáveis

Considerou-se pertinente a apresentação da correlação entre as variáveis independentes neste ponto, e não nos resultados por forma a facilitar a descrição da metodologia apresentada a seguir, que entra em conta com as variáveis definitivas utilizadas neste estudo.

Com o objetivo de eliminar as variáveis com elevada correlação entre si, procedeu-se ao cálculo da matriz de correlação entre as variáveis independentes. O método GWR é susceptível a problemas de multicolinearidade, sendo a baixa correlação muito importante para a estabilidade do modelo, que vai depender muito da distribuição do conjunto das variáveis independentes. Algumas consequências de multicolinearidade num modelo de regressão são as alterações na mudança do sinal do coeficiente de correlação quando se adiciona ou retira uma variável independente ao modelo ou quando se inclui ou exclui observações específicas na análise com erros padrão elevados (WHEELER & TIEFELSDORF, 2005).

Segundo a Tabela 3, destacou-se uma correlação elevada e negativa entre as variáveis de %Flor e %Agric (-0,81). Este fato pode ser explicado pelas práticas agropecuárias nas últimas décadas. A floresta têm sido cortada e queimadas, dando origem às parcelas de terras sujeitas aos usos agropecuários. No tempo, a floresta torna-se mais susceptível à ocorrência do fogo quando sujeita aos incêndios de superfície em anos anômalos como é o caso de anos sob a influência do fenômeno *El Niño*, que pode originar secas extremas e altas temperaturas. A variável %Flor, para além da

**Tabela 2** - Características das variáveis demográficas, ambientais e de infra-estrutura usadas.

Variáveis	Abreviatura	Unidades	Mínimo	Máximo	Fonte
Fogos Ativos	Produto WFAc	Nº médio	0	162	WFA
Densidade Populacional	Dens_pop	Nº habitantes/km <sup>2</sup>	0	6,868	GPW
Produtividade primária líquida	PPL	kgC/m <sup>2</sup> /ano	0,79	3,54	GLoPEM
<i>Chandler Burning Index</i>	CBI	s/ dimensão	0	68,29	NCEP/NCAR
% Agricultura	%Agric	Km <sup>2</sup> / 50 km	0	98,44	GLC2000
% Cerrado	%Cer	Km <sup>2</sup> / 50 km	0	93,18	GLC2000
%Distância às estradas (50km)	%Dist_est	Km <sup>2</sup> / 50 km	0	100	IBGE

**Tabela 3** - Matriz de correlação entre as variáveis independentes.

Variáveis	Dens_pop	PPL	CBI	%Agric	%Flor	%Cer	%Dist_est
Dens_pop	1						
PPL	0,01	1					
CBI	-0,02	-0,27	1				
%Agric	0,06	-0,11	0,36	1			
%Flor	-0,08	0,18	-0,42	-0,81	1		
%Cer	0,02	-0,23	0,43	0,11	-0,52	1	
%Dist_est	0,08	-0,12	0,3	0,61	-0,67	0,34	1

elevada correlação com a %Agric também apresenta correlação negativa significativa com a variável de %Cer (-0,52) e com a de Dist\_est (-0,67). O Cerrado domina em regiões mais a sul, sudeste e leste da Amazônia onde o clima é mais seco, com temperaturas elevadas e valores de umidade inferiores ao que acontece no domínio da floresta tropical úmida, correspondendo à elevada correlação inversa entre a %Flor e o %Cer.

A correlação negativa entre a variável Dist\_est e a %Flor (-0,67) deve-se ao fato de que nas zonas próximas às estradas, principais e secundárias, abertas na floresta, há dominância de extensas propriedades destinadas à agricultura e pastagens e/ou mesmo, de assentamentos rurais decorrentes dos projetos de colonização governamentais. As redes viárias, mesmo de péssima trafegabilidade em períodos de chuva, contribuem para a instalação rápida de população, que derrubam e queimam a vegetação, ocupando parcelas para uso exclusivo de atividades agropecuárias, o que justifica a correlação positiva entre as variáveis Dist\_est e %Agric (0,61).

Com esta análise, procedeu-se à eliminação da variável de %Flor, permitindo assim a criação de modelos de regressão sem variáveis de elevada correlação entre si.

Na Fig. 2 apresentam-se os mapas com a distribuição espacial de cada uma das variáveis.

### 3. MÉTODOS

O método GWR é uma recente melhoria na regressão linear tradicional, em que é tido em conta a não-estacionaridade da relação entre as variáveis espaciais, ou seja, em que os parâmetros a estimar das variáveis independentes sejam função da localização espacial. O método de estimação GWR permite estimar os parâmetros para cada ponto ( $\mu, \nu$ ) da área considerada. O programa que permite esta aplicação, designa-se pelo mesmo nome, e o

ajustamento é feito localmente através das coordenadas ( $\mu, \nu$ ), permitindo a variação ao longo do espaço dos parâmetros do modelo de regressão. Os dados que se encontram na vizinhança do ponto a estimar contribuem com um peso maior na estimativa dos parâmetros do ajustamento do que a informação mais distante (ZHANG *et al.*, 2004). Estes parâmetros estimados podem ser analisados, em forma de mapa, recorrendo a um sistema de informação geográfica permitindo uma análise das variações locais dos parâmetros estimados assim como dos respectivos erros (FOTHERINGHAM *et al.*, 2002; JETZ *et al.*, 2005).

O método dos mínimos quadrados (OLS), pode ser apresentado como:

$$y = \beta_0 + \beta_1 x_1 + \beta_n x_n + \varepsilon \quad (7)$$

onde  $y$  é a variável dependente,  $x_1$  a  $x_n$  são as variáveis independentes ou preditoras,  $\beta_0$  é a ordenada na origem,  $\beta_1$  a  $\beta_n$ , são os coeficientes estimados e  $\varepsilon$  é o erro padrão aleatório.

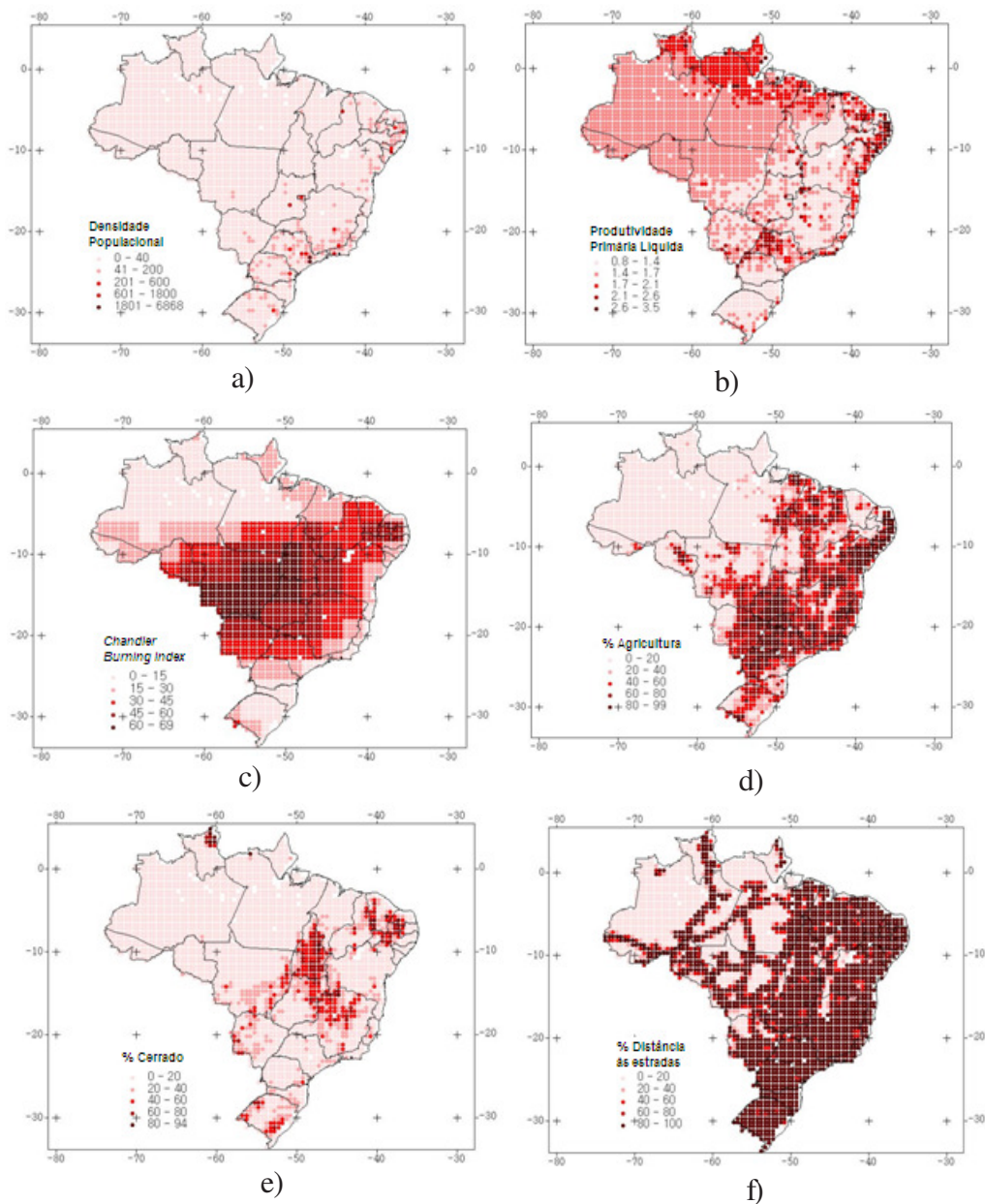
No método de estimação GWR o modelo de regressão é reescrito da seguinte forma:

$$y = b_0(m, n) + b_1(m, n)x_1 + \dots + b_n(m, n)x_n + e \quad (8)$$

onde a diferença para a Eq. 7 são as coordenadas ( $\mu, \nu$ ) dos dados.

O objetivo do método GWR é obter estimativas não paramétricas de cada uma das variáveis independentes, em cada ponto a estimar, considerando que a relação entre as variáveis não é espacialmente constante. Os parâmetros estimados dependem não só dos valores das variáveis independentes, mas também da função Gaussiana do tipo *kernel* escolhida, e do raio de influência ( $r$ ) desta função. No programa GWR, versão 2.0.3, pode-se escolher dois tipos de funções Gaussiana *kernel* (janela de valores de  $x$  colunas por  $x$  linhas): i) assumindo o raio da janela de influência constante para todos os pontos de observação, ou ii) raio variável, em função da densidade das observações. Nesta última opção, o raio é maior quando as observações estão mais espaçadas e menor se estão mais próximas. O método GWR pode ser usado considerando o diâmetro de um círculo de influência da mesma dimensão da área de estudo, ou seja, neste caso o modelo de regressão é global e





**Fig. 2** - Distribuição espacial das variáveis independentes, a) Densidade Populacional, b) Produtividade Primária Líquida, c) *Chandler Burning Index*, d) % Agricultura, e) % Cerrado e f) Distância às estradas (50km).

equivalente à regressão pelo método dos mínimos quadrados (SHI *et al.*, 2006).

Na estimativa dos parâmetros da regressão de cada ponto ( $\mu, v$ ), é necessário considerar a distância entre o ponto a estimar e as observações da sua vizinhança, ou seja, determinar quais as observações que são incluídas ou não na estimação dos parâmetros desse ponto. Numa primeira abordagem, é definido um círculo de raio ( $r$ ), o qual contém as observações que irão ser incluídas na regressão do ponto a estimar, sendo todas as restantes observações excluídas. No entanto, esta

decisão entre incluir ou excluir uma determinada observação parece muito drástica, tendo em conta que uma observação é incluída e a observação sua adjacente é excluída. Este problema é superado pela atribuição de pesos em função da distância ao ponto ( $\mu, v$ ) (BRUNSDON, *et al.*, 2001). O peso atribuído aos valores de cada variável decresce gradualmente com a distância ao ponto considerado e os parâmetros são estimados independentemente para cada ponto geográfico. A matriz dos parâmetros de regressão estimados é escrita da seguinte forma:



$$\hat{\beta}(\mu, \nu) = (X^T W(\mu, \nu) X)^{-1} X^T W(\mu, \nu) y \quad (9)$$

onde  $\hat{\beta}(\mu, \nu)$  é o vetor dos parâmetros de regressão estimados,  $W(\mu, \nu) = \text{diag}[W_1(\mu, \nu), \dots, W_n(\mu, \nu)]$  é a diagonal da matriz de pesos que é diferente para cada ponto,  $X$  é a matriz de variáveis independentes,  $X^T$  é a sua transposta e  $y$  é o vetor da variável dependente, com os valores estimados.

Para estimar os parâmetros pelo método da GWR, segundo a Eq. 9, é importante definir o critério de decisão de proximidade. Em análise espacial, é comum assumir-se que as observações mais próximas do ponto a estimar ( $i$ ), têm mais influência na estimação do que as mais distantes. Recorrendo à função inversa da distância,  $1/d_{ij}$ , define-se a matriz de pesos entre os pontos  $i$  e  $j$  ( $W_{ij}$ ). O raio da janela para cada ponto central  $i$  é uma constante ao longo da área de estudo, designado como *kernel* fixo. Quando  $i$  é igual a  $j$ , então  $W_{ij} = 1$  e com o aumento da distância  $d_{ij}$ ,  $W_{ij}$  decresce conforme a curva gaussiana.

Quando as observações não estão uniformemente distribuídas pela área de estudo utiliza-se um raio de influência que é variável para cada ponto a estimar. Neste caso a matriz dos pesos é calculada por:

$$\left\{ \begin{array}{l} w_{ij} = \exp \left[ -\frac{\left( \frac{d_{ij}}{b} \right)^2}{2} \right] \text{ para } d_{ij} < b \\ w_{ij} = 0 \text{ para } d_{ij} \geq b \end{array} \right. \quad (10)$$

O valor do peso é igual a 1 quando a distância entre os pontos  $i$  e  $j$  é nula, decrescendo para 0 quando a distância entre pontos é maior ou igual ao raio da janela ( $b$ ) selecionada (FOTHERINGHAM *et al.* 2002).

O raio de influência é um fator muito importante no cálculo da função *kernel*. Com o aumento da dimensão do raio, os parâmetros estimados aproximam-se dos valores resultantes de uma regressão global. O valor do raio de influência pode ser obtido por três formas: i) definido pelo utilizador; ii) validação cruzada (VC) e iii) Critério de Informação de Akaike (*Akaike Information Criterion* - AIC) (HURVICH *et al.*,

1998). O valor do raio pode ser estipulado pelo utilizador, limitando à partida o número de observações consideradas na estimação dos parâmetros em cada ponto.

Uma outra alternativa para obter o valor do raio é deixar que o programa determine esse valor apropriado através de medidas de avaliação da precisão dos modelos, pelo método de validação cruzada (VC) ou pelo AIC.

No presente trabalho utilizou-se o critério de seleção do AIC, para obtenção do raio de influência para a estimativa dos parâmetros empregando uma janela do tipo *kernel* fixa para toda a área de estudo. Este critério é uma medida que permite avaliar a qualidade do ajustamento de um modelo (HURVICH *et al.*, 1998), minimizando o valor de AIC (WINDLE *et al.*, 2009). O valor deste critério é calculado da seguinte forma:

$$AIC_c = 2n \log_e \left( \hat{\sigma} \right) + n \log_e (2\pi) + n \left\{ \frac{n + \text{tr}(S)}{n - 2 - \text{tr}(S)} \right\} \quad (11)$$

onde  $n$  é a dimensão da amostra,  $\hat{\sigma}$  é o desvio padrão estimado do erro padrão, e  $\text{tr}(S)$  é o traço da matriz-chapéu  $S$ .

Segundo FOTHERINGHAM *et al.* (2002) pequenas diferenças no valor do AIC (três unidades) para modelos diferentes podem dever-se a erros de amostragem e não a modelos com melhor ou pior ajustamento. Este critério permite comparar diferentes modelos e permite comparar o método de regressão OLS com o método de regressão GWR. Se o modelo pela regressão OLS tem um valor de AIC inferior ao do modelo ajustado pelo GWR, então não se justifica a aplicação deste método e sua complexidade de cálculos. Este critério ainda permite a comparação entre modelos obtidos variando o raio da função *kernel*, ou seja, usando o GWR, para dois raios diferentes obtendo-se dois modelos diferentes, onde o AIC permite avaliar qual dos modelos melhor se ajusta aos dados. Este fato torna-se possível porque o AIC entra em conta com os graus de liberdade dos diferentes modelos, permitindo assim uma comparação mais precisa entre modelos.

O programa GWR dispõe de duas formas de testar a variabilidade espacial da relação entre as variáveis, através do teste de significância de Monte Carlo (HOPE, 1968) ou do teste atribuído a LEUNG *et al.* (2000a). Ambos os testes de

significância permitem testar a estacionaridade dos parâmetros individuais, baseados na medição da sua variabilidade no espaço quando estimados pela GWR. A questão é se a variação espacial da relação entre as variáveis é suficiente para se rejeitar a hipótese de ser globalmente constante. O uso destes testes, principalmente o de Monte Carlo, tem como desvantagem os elevados cálculos que tornam o processo por vezes demorado, quando a quantidade de dados é elevada, caso em que, FOTHERINGHAM *et al.* (2002) sugerem que seja reduzido o número de iterações ou usada uma amostra menor.

No presente trabalho utilizou-se o teste de significância de Monte Carlo (HOPE, 1968). Este teste permite analisar a estacionaridade dos parâmetros individuais, baseados na medição da sua variabilidade no espaço quando estimados pela GWR.

Os conceitos do processo de estimação implementado pelo modelo de regressão geograficamente ponderada (GWR), estão descritos mais detalhadamente em FOTHERINGHAM *et al.* (2002) e sua aplicabilidade em questões ambientais pode ser observada nos trabalhos de FOODY (2003), WANG *et al.* (2005), ZANG *et al.* (2005), WINDLE *et al.* (2009), LIN & WEN (2011), SÁ *et al.* (2011).

Em cumprimento do procedimento metodológico adotado nesse estudo procedeu-se ao ajuste de modelos univariados e multivariados para avaliar a contribuição das diversas variáveis na explicação da distribuição espacial dos fogos ativos resultantes do produto WFAc (Tabela 4). A estratégia foi testar um modelo multivariado com as seis variáveis, seis modelos univariados e ainda combinações dessas variáveis. Realizaram-se três combinações de variáveis independentes, agrupando-as segundo fatores humanos (Modelo 2), fatores naturais (Modelo 3) e ainda daquelas cujos modelos univariados apresentaram melhores resultados segundo o critério de minimização do AIC. Esta última combinação representa, respectivamente, os três e os dois modelos com menor AIC (Modelo 10 e Modelo 11).

Com os modelos 2 e 3 pretende-se perceber qual a influência dos fatores humanos e naturais na ocorrência de fogos no Brasil. Os fatores humanos integram a distribuição e atividade da presença

**Tabela 4** - Variáveis independentes para cada modelo testado com o método de regressão OLS e GWR.

Modelos	Variáveis independentes
Modelo 1	Dens_pop, PPL, CBI, %Agric, %Cer e %Dist_est
Modelo 2	Dens_pop, %Agric e %Dist_est
Modelo 3	PPL, CBI e %Cer
Modelo 4	Dens_pop
Modelo 5	PPL
Modelo 6	CBI
Modelo 7	%Agric
Modelo 8	%Cer
Modelo 9	%Dist_est
Modelo 10	%Agric, %Cer e %Dist_est
Modelo 11	%Agric e %Dist_est

humana, que se refere a densidade populacional, a presença de agricultura e a construção de estradas. No caso dos fatores naturais, a produtividade primária líquida, o índice climático CBI e o tipo de cobertura vegetal afetada no caso específico, o Cerrado, vegetação típica do Brasil Central e com marcante efeito da sazonalidade, conforme citado por FRANÇA *et al.*, (2001), foram levados em consideração para verificar suas influências na incidência de queimadas.

## 4. RESULTADOS E DISCUSSÃO

### 4.1 Comparação entre os métodos de OLS e GWR

Foram comparados os resultados obtidos através da aplicação dos dois métodos de regressão OLS e GWR para os 11 modelos testados (Tabela 5).

Com base no método de regressão global o melhor modelo ajustado na explicação da variação espacial entre as variáveis independentes e os fogos ativos no Brasil foi o Modelo 1, que recorre às seis variáveis. Este modelo tem o menor valor de AIC (7617) e o valor mais elevado para o coeficiente de determinação ( $R^2 = 0,38$ ). Quando se compara o resultado obtido com o método GWR, verificamos que o mesmo modelo multivariado apresenta um melhor ajustamento dos dados relativamente ao método de regressão global, com valor de AIC inferior (4619) e um coeficiente de

**Tabela 5** - Valores de AIC, R<sup>2</sup> e o raio da janela kernel para os todos os modelos testados (OLS e GWR).

Modelos	OLS		GWR		
	R <sup>2</sup>	AIC	R <sup>2</sup>	AIC	Raio de janela (°)
Modelo 1	0,382	7617,06	0,783	4619	1,2194
Modelo 2	0,095	8660,56	0,785	4526	0,9521
Modelo 3	0,356	7723,61	0,782	4994	0,8545
Modelo 4	0,001	8929,85	0,771	4863	0,7057
Modelo 5	0,003	8923,38	0,768	4927	0,7342
Modelo 6	0,343	7776,38	0,778	4739	0,6636
Modelo 7	0,061	8758,36	0,793	4378	0,7523
Modelo 8	0,037	8826,84	0,804	4511	0,6845
Modelo 9	0,085	8686,35	0,791	4509	0,6972
Modelo 10	0,105	8629,91	0,801	4405	0,9198
Modelo 11	0,093	8666,23	0,8	4367	0,8316

determinação superior (R<sup>2</sup>=0,78). Estes resultados indicam que as variáveis se alteram espacialmente ao longo da área em estudo, tal como o produto WFAc. Esta melhoria de ajustamento verifica-se para os onze modelos testados, onde os valores de AIC obtidos pelo métodos GWR são sempre inferiores aos obtidos com o método OLS. No entanto, destaca-se o Modelo 11, com um valor de AIC mínimo de 4367 e um R<sup>2</sup>= 0,80, seguido do Modelo 7, com AIC = 4377,55 e do Modelo 10, com AIC = 4405.

Em outras aplicações ambientais desenvolvidas recentemente por LIN & WEN (2011), WINDLE *et al.* (2009) e SÁ *et al.* (2011), foram obtidas diferenças semelhantes entre os dois métodos, ou seja, o modelo GWR mostra um melhoramento significativo sobre o modelo OLS.

Segundo alguns autores (NEPSTAD *et al.*, 2001; LAURENCE *et al.* 2001 e 2002 e CARDOSO *et al.* 2003), a distância às estradas na Amazônia é o fator mais correlacionado com a expansão do desflorestamento e conseqüentemente com atividades onde o fogo está historicamente presente. Dois terços da conversão florestal da Amazônia ocorrem numa faixa até 50 km de distância às vias de acesso, ou seja, às estradas principais (NEPSTAD *et al.*, 2001). Associada a esta prática está a queima de biomassa acumulada após o desmatamento (NEGREIROS *et al.*, 1996). Com áreas desmatadas e limpas de resíduos desenvolvem-se atividades agrícolas e pecuárias, prática onde as queimadas continuam a ter

presença, quer na eliminação de resíduos resultantes da colheita de cultivos, como na renovação da palatabilidade das pastagens. Ainda se pode salientar a exploração seletiva de madeira, muitas vezes realizada sem cuidados de sustentabilidade florestal, o que deixa a floresta assim degradada mais vulnerável à ocorrência de fogo no período sêco. Estes fatores vão de encontro aos resultados obtidos neste estudo, cujo modelo com melhor ajustamento é baseado nestas duas variáveis, percentagem de agricultura e da distância às vias de acesso. Analisando os resultados para os modelos 2 e 3, verifica-se a contribuição dos fatores humanos para a ocorrência do fogo ativo, com um valor menor de AIC para o modelo 2.

Na Tabela 5, também se observa a variação do raio da função *kernel*, que mostra que quanto mais variáveis independentes tem o modelo maior é o raio da função *kernel* e, conseqüentemente, mais observações contribuem para a estimativa do produto WFAc. Relativamente aos modelos univariados (Modelos 4 a 9), o número de observações usadas em cada regressão local é menor, tornando o ajustamento mais local do que para os modelos multivariados.

O presente estudo foi desenvolvido sobre dados em coordenadas geográficas ( $\phi, \lambda$ ). Sendo assim, o raio fixo corresponde a uma distância planificada que varia ligeiramente com a latitude, dentro da área de estudo. Considerou-se esta variação insignificante para justificar a aplicação de um fator de correção.

O critério AIC não tem uma dependência com a dimensão do raio da função *kernel* porque considera os graus de liberdade do modelo. A dimensão média da janela para os modelos univariados (Modelos 4 a 9) com o método GWR é 0,70625° (aproximadamente 77 km para a latitude média do Brasil), quase metade da dimensão da função *kernel* escolhida pelo método GWR para o modelo multivariado, com todas as variáveis (aproximadamente 134 km).

A Tabela 6 apresenta a análise da variância para os três modelos que obtiveram melhores resultados considerando o valor de AIC e o R<sup>2</sup> (um modelo de cada grupo) para o método GWR e OLS.

Analisando os valores dos resíduos para os dois métodos confirma-se a melhoria significativa para o método GWR, apresentando um valor de



**Tabela 6** - Análise de variância dos métodos GWR e OLS, para os modelos 11, 7 e 2. Valor dos resíduos obtidos e a respectiva melhoria do método GWR sobre o OLS, valor da soma dos quadrados, SS (*Sum of squares*), da média dos quadrados, MS (*Mean squares*), dos graus de liberdade, DF (*Degree of Freedom*) e do teste F para  $p < 0,005$ .

Modelos	Resíduos	SS	DF	MS	F
Modelo 11	OLS	3751,5	3		
	GWR	437,7	2126,52	0,2058	25,9462
	Melhoria GWR	3313,8	620,48	5,3407	
Modelo 7	OLS	3882,2	2		
	GWR	454	2153,64	0,2108	27,3622
	Melhoria da GWR	3428,2	594,36	5,7679	
Modelo 2	OLS	3741,1	4		
	GWR	471,9	2140,67	0,2204	24,5013
	Melhoria da GWR	3269,2	605,33	5,4007	

resíduos inferior. Estes modelos apresentam valores semelhantes da soma dos quadrados dos resíduos (SS), sendo o Modelo 11 o que apresenta o menor valor pelo método GWR. Para o Modelo 7 a variável %Agric tem um ajustamento local mais adequado. O Modelo 2 apresenta o valor mais elevado de melhoria do método GWR entre os dois. Desta análise conclui-se que as variáveis de presença das áreas agrícolas e da distância às estradas (Modelo 11), ou seja, fatores humanos, explicam melhor a ocorrência do fogo no Brasil. A contribuição destes fatores foi descrita por NEPSTAD *et al.* (2001) e NEGREIROS *et al.* (1996).

Relativamente aos graus de liberdade de cada modelo verifica-se que no método GWR este valor é sempre superior, visto que correspondem ao número de parâmetros utilizados em cada regressão local, que varia consoante o número de variáveis que integram os modelos mas também com a dimensão da função *kernel*. No presente caso é o Modelo 7 que apresenta o número de graus de liberdade (DF) mais elevado.

O teste de Fisher (teste F) é geralmente utilizado para comparar variâncias e decidir se estas são ou não significativamente diferentes. Neste caso, pretende-se testar a hipótese nula, de que o método de regressão local representa melhor ajustamento dos dados do que o método de regressão global. Como resultado, rejeita-se a hipótese nula ( $P$ -value  $< 0,005$ ), comprovando-se que a relação entre as variáveis é não-estacionária e que o GWR tem melhor ajustamento com o produto WFAc. Na

Tabela 6, podemos verificar que todos os modelos têm valores elevados de teste F, indicando que as relações entre as variáveis independentes e a variável dependente não são constantes espacialmente.

Na Fig. 3 apresenta-se o mapa resultante da aplicação dos parâmetros estimados pelo Modelo 11 às variáveis independentes com o produto WFAc. A comparação entre a distribuição espacial do produto WFAc observado e estimado são similares (Fig. 1 e 3).

Os valores mais elevados do produto WFAc ocorrem na região do denominado “Arco do Desmatamento” (Estados do Mato Grosso, Pará e parte do Maranhão) contido na Amazônia, bem como em porções da costa nordestina (Estado de Alagoas) e na região do Pantanal (Estado de Mato Grosso do Sul). A contextualização na região do Arco do Desmatamento corresponde ao aumento da fronteira agrícola e pecuária, que vem ocorrendo na transição entre os biomas do Cerrado (típico do Brasil Central) e da Floresta Tropical Amazônica.

Ainda na Fig. 3, pode-se também observar que o modelo subestima ligeiramente as observações, com valores negativos para regiões onde não há ocorrências de fogos ativos e com valores máximos do produto WFAc de 4,4, sendo originalmente o valor máximo de 5. Os pontos com valores negativos do produto WFAc estimado pelo GWR com o Modelo 11, representam aproximadamente 2,8% do total das observações (Fig. 3). Estes pontos significam regiões onde foi estimado menor número de fogos ativos do que os observados, com base nas variáveis independentes, neste caso, as variáveis %Agric e %Dist\_est. Muitos



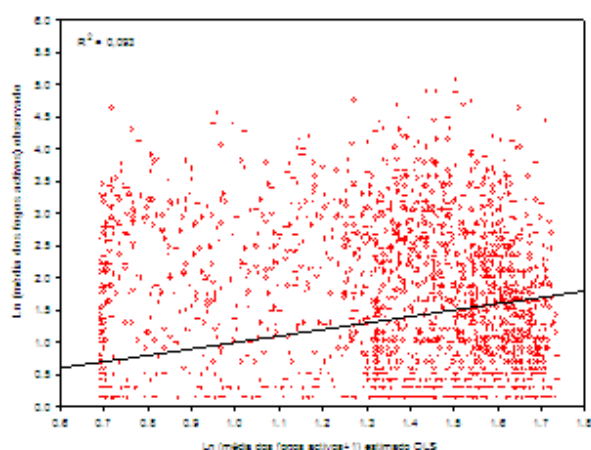
**Fig. 3** - Logaritmo natural da média do produto WFAc estimado pelo método GWR com o Modelo 11.



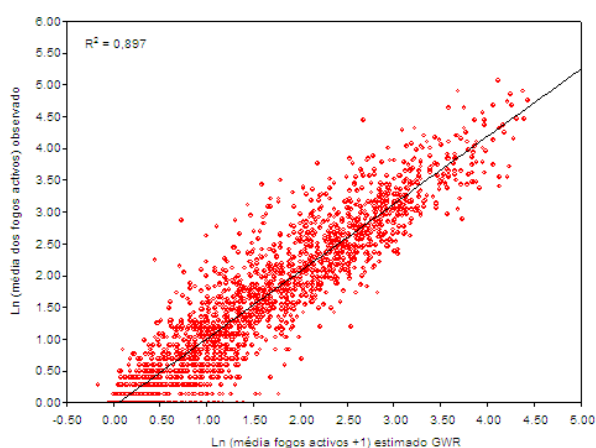
destes correspondem a valores observados nulos, ou seja, a células onde não são detectados fogos ativos ATSR. Localizam-se na sua maioria no Estado do Amazonas e em menor quantidade, nos Estados da Roraima, Pará e Amapá. Há ocorrências pontuais nas regiões sudeste e sul do Brasil, nos Estados de São Paulo e Santa Catarina. Este fato mostra alguma dificuldade do Modelo 11 se ajustar nas regiões onde não se observam fogos ativos.

A Fig. 4 mostra o diagrama de dispersão entre a média dos fogos ativos observados e estimados segundo o método de regressão OLS (Fig. 4 a) e GWR (Fig. 4 b), para o Modelo 11, e mostra a dificuldade do modelo na estimativa da variável resposta, que se traduz num valor do coeficiente de determinação muito baixo ( $R^2=9,3\%$ ).

A equação de regressão mostra uma boa relação com os dados WFAc e sua estimativa usando o método GWR, com o coeficiente de



a)



b)

**Fig.4** - Diagrama de dispersão entre logaritmo de base natural para a média do produto WFAc observado e estimado pelo método a) OLS e b) GWR (Modelo 11).

determinação elevado ( $R^2=89,7\%$ ). Como já se referiu, também neste diagrama podemos observar uma ligeira sub-estimativa do número de fogos ativos para o modelo em análise, usando o método de regressão GWR. O logaritmo da média do produto WFAc observado é 3338,88 e o estimado é 3294,18. Esta ligeira sub-estimação acontece principalmente para as áreas onde não foram observados fogos ativos. FOTHERINGHAM *et al.* (2002) e ZANG *et al.* (2004) sugerem a comparação entre os parâmetros estimados pelo GWR e os intervalos de confiança (IC) dos parâmetros estimados pelo OLS, para avaliar a não estacionaridade espacial da relação entre as variáveis. Outra forma de confirmar este fato é verificar se a distância interquartil (diferença entre o percentil 25 e 75) do método de estimativa local é maior do que  $\pm 1$  do erro padrão resultante da regressão global (FOTHERINGHAM *et al.*, 2002).

A Tabela 7 apresenta as estatísticas descritivas dos parâmetros estimados pelos dois métodos de regressão GWR e OLS.

Da análise da Tabela 7 pode-se verificar que o intervalo interquartil dos parâmetros estimados pelo GWR é sempre maior do que o intervalo  $\pm 1$  do erro padrão relativamente aos parâmetros equivalentes mas estimados pela OLS. Para o parâmetro de ordenada na origem  $\beta_0$ , a distância

**Tabela 7** - Estatística descritiva dos parâmetros estimados pelos métodos OLS e GWR (Modelo 11).

Estatísticas		$\hat{\beta}^0$	$\hat{\beta}^1$	$\hat{\beta}^2$
OLS	Estimado	0,6912	0,0042	0,0063
	Erro padrão	0,0387	0,0009	0,0006
	$\hat{\beta} - 1 SE$	0,6525	0,0033	0,0057
	$\hat{\beta} + 1 SE$	0,7299	0,0051	0,0069
	L. inferior IC 95%	0,6154	0,0024	0,0051
	L. superior IC 95%	0,7671	0,006	0,0075
GWR	Média	0,1059	0,0251	0,003
	Mínimo	138,1003	-0,761	-0,4324
	P. 25%	0,0197	-0,0052	-0,0029
	Mediana (50%)	0,8504	0,0025	0,0014
	P. 75%	2,3525	0,034	0,0068
	Máximo	43,9957	1,1723	0,3844

interquartil é 0,0197 a 2,3525 (2,3328) pelo método GWR, que é maior do que o intervalo de 0,6525 a 0,7299 (0,0774) que corresponde ao valor estimado de  $\pm 1$  erro padrão pelo método OLS. A mesma situação acontece para os parâmetros estimados  $\hat{\beta}_1$  e  $\hat{\beta}_2$  com distância inter-interquartil bastante maior do que os respectivos intervalos de  $\pm 1$  erro padrão para o valor estimado pelo método OLS.

Além desta comparação também podemos analisar a distribuição dos IC de 95% para os parâmetros estimados pela OLS relativamente às estatísticas obtidas pelo método GWR. O IC de 95% (0,6154; 0,7671) para o parâmetro de ordenada na origem estimado pela regressão OLS encontra-se entre o percentil 25 e a mediana, indicando que uma grande parte dos valores estimados para a ordenada na origem pelo método GWR são maiores do que os valores estimados pela regressão OLS. Para o  $\hat{\beta}_1$  o IC de 95% (0,0024; 0,0060) encontra-se entre o percentil 25 e o 75, que indica que os valores estimados para este parâmetro pelo GWR são aproximadamente da mesma ordem de grandeza dos valores estimados pela regressão OLS. O IC de 95% (0,6154; 0,7671) para o parâmetro de ordenada na origem estimado pela regressão OLS encontra-se entre o percentil 25 e a mediana, indicando que uma grande parte dos valores estimados para a ordenada na origem pelo método GWR são maiores do que os valores estimados pela regressão OLS. Para o  $\hat{\beta}_1$  o IC de 95% (0,0024; 0,0060) encontra-se entre o percentil 25 e o 75, que indica que os valores estimados para este parâmetro pelo GWR são aproximadamente da mesma ordem de grandeza dos valores estimados pela regressão OLS. O IC 95% para o  $\hat{\beta}_2$  (0,0051; 0,0075) encontra-se entre a mediana e o valor máximo, indicando que uma grande parte dos valores estimados para este valor pelo método GWR é menor do que os valores estimados pelo método OLS. Estes resultados comprovam que a relação entre os fogos ativos observados e variáveis independentes varia espacialmente (FOTHERINGHAM *et al.*, 2002; ZANG *et al.*, 2004).

O teste de significância de Monte Carlo para a hipótese nula de que as variáveis têm comportamento não estacionário espacialmente, mostra que todas elas têm o  $p$ -value  $< 0,000001$ , ou seja, não existe estacionaridade espacial.

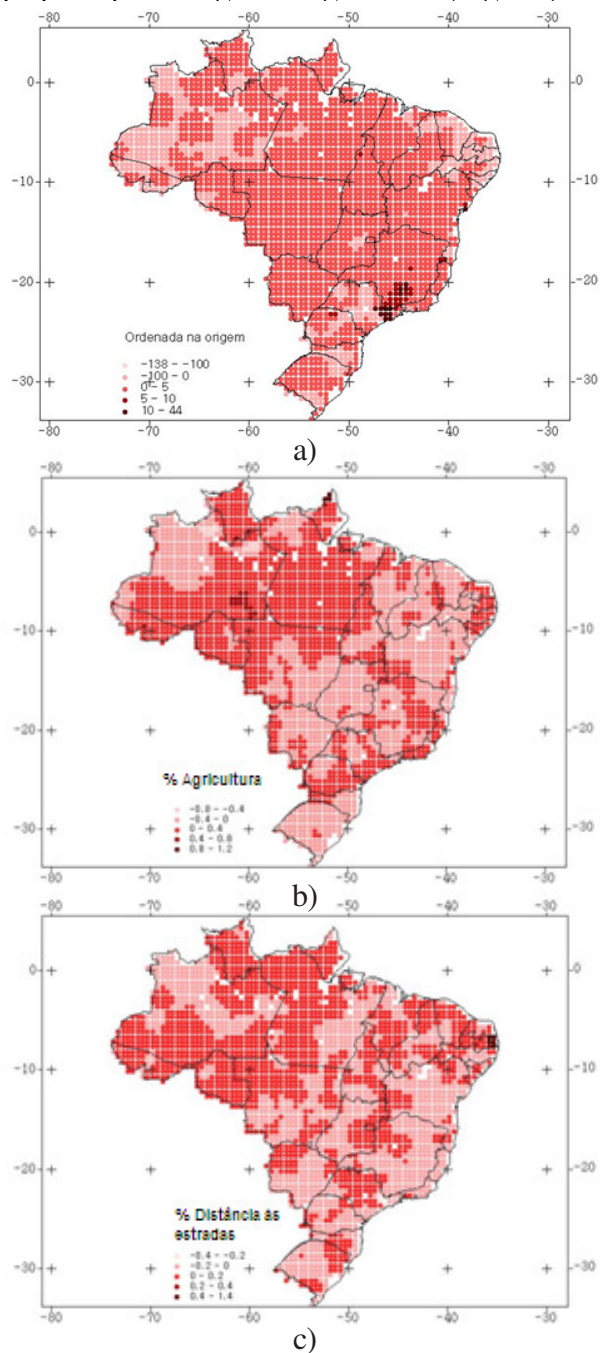
#### 4.2 Interpretação dos parâmetros estimados pelo GWR

Além de modelar as variáveis espacialmente, o método GWR possibilita a análise dos resultados através de mapas, o que permite avaliar de forma clara a sua distribuição espacial. Na Fig. 5 pode-se observar a distribuição espacial resultante da análise da distribuição geográfica dos valores da ordenada na origem, bem como dos parâmetros estimados das respectivas variáveis, constantes da aplicação do método de regressão GWR para o Modelo 11.

Estes parâmetros estimados representam a relação entre cada uma das variáveis independentes consideradas e o logaritmo do produto WFAc. A variação da ordenada na origem ( $\hat{\beta}_0$ ), tem uma mediana de 0,85 e um intervalo de variação entre -138,10 e 44,00 (Fig. 5 a). Estes valores representam a estimativa do produto WFAc, se assumirmos que todas as variáveis independentes apresentam valor nulo. A maior parte do território brasileiro tem valores positivos, ficando apenas 16% da área com valores negativos. As regiões com valores positivos mais elevados encontram-se no leste do Estado São Paulo e no Estado de Minas Gerais. As regiões com valores negativos encontram-se principalmente no Estado do Amazonas, na costa nordeste do Brasil, em São Paulo e nos Estados mais a sul do país, regiões onde a ocorrência de fogos é muito baixa ou nula, o que indica que o modelo tem alguma dificuldade de ajustamento para os valores baixos de fogos ativos. O peso dos valores negativos é maior, atingindo o valor -138 do que se registra com os valores positivos, contribuindo assim para a ligeira subestimação do número de fogos ativos.

O parâmetro de %Agric varia entre -0,761 e 1,172 (Fig. 5 b), com uma mediana de 0,002 e com 39% da área com valores negativos apesar de muito próximos de zero, o que indica a sua fraca relação com a média dos fogos ativos nestas regiões. Na realidade, quando aumenta uma unidade na variável %Agric numa célula de 0,5°, diminui ligeiramente a média do número de fogos, sendo o valor máximo

negativo de apenas -0,8. Na situação contrária quando se aumenta uma unidade desta variável, ocorre um aumento máximo de 0,8 a 1,2 na média do número de fogos. Este aumento máximo ocorre no norte do Estado de Amapá numa área de três células de  $0,5^\circ$ , local onde a % de agricultura em cada célula de  $0,5^\circ$  é muito baixa (0 a 20%). Destaca-se ainda numa região situada no canto sudeste do Estado do Amazonas, onde domina a floresta tropical úmida e há ocorrência de uma pequena percentagem de agricultura (Fig. 2 d).

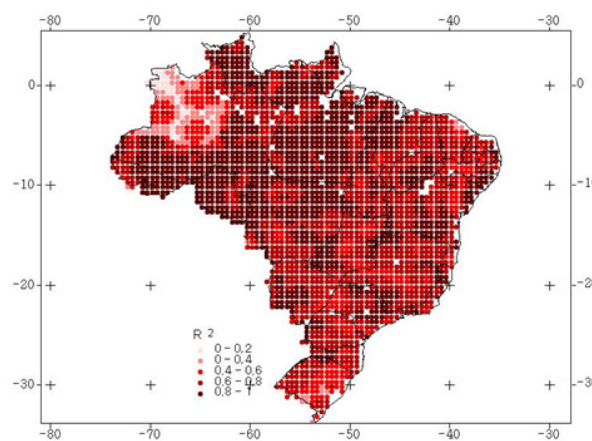


**Fig. 5** - Distribuição espacial dos parâmetros estimados pelo método GWR da a) constante, b) % Agricultura e c) % Distância às estradas (Modelo 11).

Para o parâmetro estimado referente à variável % Distância às estradas, o intervalo de variação é -0,432 a 1,384 (Fig. 5 c) e com uma mediana de 0,001. As regiões onde esta variável tem uma relação positiva com os fogos ativos encontram-se distribuídas pelo território brasileiro, destacando-se uma pequena região (cinco células) com valores mais elevados no Estado da Paraíba na região nordeste, junto à costa litorânea, onde a rede de estradas tem certo adensamento (facilitada pela relação com plantios específicos de cana de açúcar, cuja colheita é historicamente precedida por atividades de queima) e com valores intermédios do logaritmo da média do produto WFAc. Nesta região, a contribuição do aumento de 1% de área que está incluída na faixa de 50 km de distância às estradas corresponde ao aumento máximo de 1,4 no produto WFAc. As células com relação negativa com os fogos ativos estão distribuídas por todo o país, sendo sempre muito próximas de zero.

A variação espacial do  $R^2$ , resultante da aplicação do método de regressão local GWR para o Modelo 11, indica as regiões onde o modelo melhor explica a relação entre as variáveis independentes e os fogos ativos ATSR (Fig. 6).

Na distribuição espacial do coeficiente de variação, destacam-se zonas de valores elevados, essencialmente no norte do Brasil e baixos valores, sobretudo nos Estados do Amazonas e do Rio Grande do Sul. O modelo teve mais dificuldades no ajustamento, para as regiões onde os resultados de  $R^2$  são mais baixos, que correspondem a valores negativos para o número de fogos ativos ATSR estimados (Fig. 3). Regiões onde a presença dos fatores correspondentes às variáveis independentes no Modelo 11 tem menos presença.



**Fig. 6** - Distribuição espacial do coeficiente de determinação ( $R^2$ ) pelo método GWR (Modelo 11).



Os resíduos para cada método de regressão resultantes do Modelo 11, que representam a diferença entre os dados observados e os estimados pelo respectivo modelo, podem ser observados na Fig. 7.

Os resíduos obtidos pelo método de regressão OLS são maiores e têm grande variação espacial, enquanto pelo método GWR, são de menores dimensões e muito mais uniformes no espaço.

Para o método OLS, os resíduos têm uma distribuição semelhante à dos valores médios do produto WFac, onde valores mais elevados se distribuem ao longo do Arco do Desmatamento, na região da costa leste do país e no Pantanal, demonstrando alguma dificuldade no ajustamento do número elevado de fogos ativos com as respectivas variáveis, possivelmente necessitando de outras variáveis independentes para melhorar o seu resultado. Esta situação não acontece na distribuição

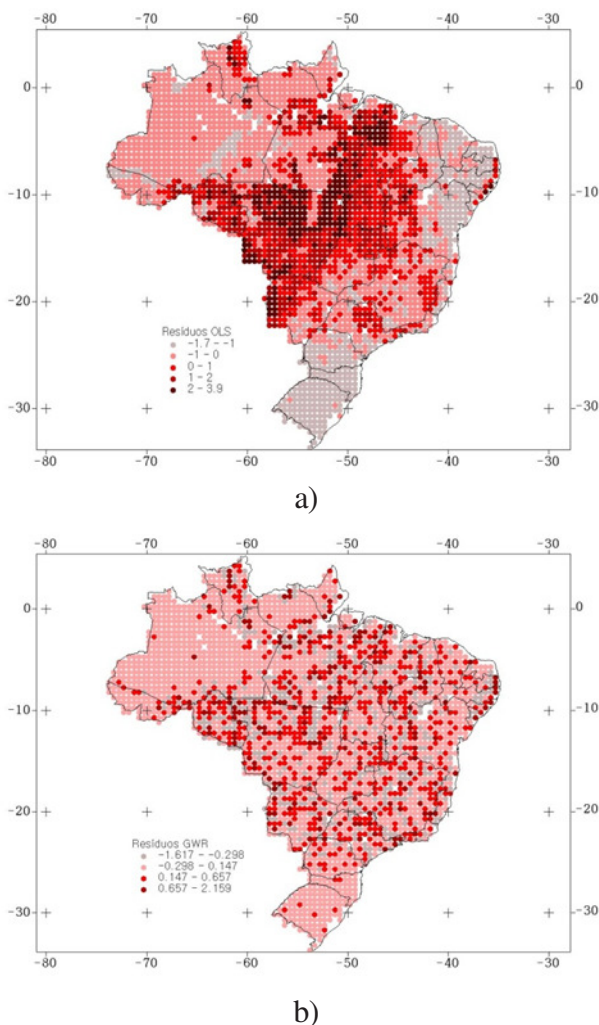
dos resíduos resultantes do método GWR, onde apenas 66 células de 0,5° de resolução espacial têm valores mais elevados (de 1 a 2,159), distribuídos pela área de estudo, não mostrando áreas de grandes dificuldades de ajustamento do modelo ao produto WFac.

## 5. CONCLUSÕES

Com este estudo podemos concluir a vantagem de utilizar o método de regressão GWR, em que para todos os modelos testados obteve-se uma melhoria significativa na estimação do número de fogos ativos comparativamente ao método de regressão OLS. Este resultado traduz a existência de variabilidade espacial entre as variáveis utilizadas. Da análise dos resultados obtidos para os diversos modelos testados, verificou-se que a inclusão de todas as variáveis consideradas não permite a melhor representação da realidade da ocorrência do fogo no Brasil, relativamente à utilização de apenas algumas das variáveis. Para os modelos univariados obteve-se um melhor ajustamento do número de fogos ativos, sendo superados pela combinação de apenas duas das variáveis, presença de áreas agrícolas e da distância às estradas. Este resultado permite concluir que a construção de acessos e a prática agrícola estão muito relacionadas com a ocorrência de fogo no Brasil.

Um fator importante a ter em conta neste tipo de estudo é a resolução espacial escolhida para a aplicação das técnicas de regressão, que foi uma resolução intermediária, considerando que algumas variáveis foram obtidas a 1 km (para a latitude média do Brasil, aproximadamente 0,00893°), mas para o cálculo do índice climático de *Chandler* foram usados dados climáticos que estavam disponíveis a 2,5°. Esta diferença de escalas pode distorcer os valores reais, devendo usar-se escalas próximas sempre que possível.

A análise realizada neste estudo seria interessante refazer considerando outros dados de fogos ativos, decorrentes do imageamento do sensor *Moderate-Resolution Imaging Spectroradiometer* (MODIS) Terra ou Aqua, bem como do sensor *Advanced Very High Resolution Radiometer* (AVHRR) da série *National Oceanic and Atmospheric Administration* (NOAA). Este tipo de análise pode ser útil no apoio ao desenvolvimento de módulos mais sofisticados, com base processual, que incorporam métodos de



**Fig. 7** - Distribuição espacial dos resíduos segundo o método a) OLS e b) GWR, (Modelo 11).



modelagem do fogo em modelos globais de dinâmica de vegetação, ao realçar como a importância de diferentes variáveis é heterogênea no espaço.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABER, J. D.; MELILLO, J. M. Terrestrial ecosystems. 2ª edição. Harcourt Science and Technology Company, **Harcourt Academic Press**. USA, 2001. 556p.
- ALVES, D. S. An analysis of the geographical patterns of deforestation in the Brazilian Amazonia in the period 1991-1996. In Deforestation and land use and forest change in the Amazon. Univ. da Florida, USA, p. 95, 2002.
- ARINO, O.; ROSAZ, J. M. 1997 and 1998 World ATSR fire atlas using ERS-2 ATSR-2 data. In Conferência e workshop Joint Fire Science. Boise Idaho, June 1999.
- BRUNSDON, C.; McCLATCHEY, J.; Unwin, D. J. Spatial variations in the average rainfall-altitude relationship in Great Britain: an approach using geographically weighted regression. **International Journal of Climatology**, v. 21, p. 455, 2001.
- CARDOSO, M.; HURTT, G.; MOORE III, B., NOBRE, C.; PRINS, E. Projecting the future fire activity in Amazonia. **Global Change Biology**, v. 9, p. 656, 2003.
- CARDOSO, M. F.; HURTT, G. C., MOORE III, B.; NOBRE, C. A.; BAIN, H. Field work and statistical analyses for enhanced interpretation of satellite fire data. **Remote Sensing of Environment**, v. 96, p. 212, 2005.
- CHANDLER, C.; CHENEY, P.; TRABAUD, L.; WILLIAMS, D. **Fire in Forestry. Forest Fire Behavior and Effects**. John Wiley & Sons, 1983. 450p.
- CHURKINA, G.; RUNNING, S. W.; SCHLOSS, A.L. Comparing global models of terrestrial net primary productivity (NPP): the importance of water availability. **Global Change Biology**, v. 5, p. 46. 1999.
- COHEN, W. B.; MAIERSPERGER, T. K.; FOWER, S. T.; TURNER, D. P. An improvement strategy for regression of biophysical variables and Landsat ETM+ data. **Remote Sensing of Environment**, v. 84, p. 561, 2003.
- EVA, H., MIRANDA, E. E., DI BELLA, C. M., GOND V., HUBER O., SGRENZAROLI M., JONES, S., COUTINHO, A., DORADO, A., GUIMARÃES, M., ELVIDGE, C., ACHARD, F., BELWARD, A.S., BARTHOLOMÉ, E., BARALDI, A., DE GRANDI, G., VOGT, P., FRITZ, S.; HARTLEY, A. A Vegetation Map of South America, EUR 20159 EN, European Commission Joint Research Centre, Luxemburg. <http://www.gvm.jrc.it/glc2000>, 2002.
- FOODY, G. M. Geographic weighting as a further refinement to regression modeling: An example focused on the NDVI-rainfall relationship. **Remote Sensing of Environment**, v. 88, p. 283, 2003.
- FOODY, G. M. Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. **Global Ecology and Biogeography**, v. 13, p. 315, 2004.
- FOTHERINGHAM, A. S.; BRUNSDON, C.; CHARLTON, M. **Geographically Weighted Regression: the analysis of spatially varying relationships**. John Wiley & Sons Inc. England, 2002. 282p.
- FRANÇA, H.; SETZER, A. AVHRR analysis of a savanna site through a fire season in Brazil. **International Journal of Remote Sensing**, v. 22, n. 13, p. 2449, 2001.
- GOETZ, S. J.; S. D. PRINCE; S. N. GOWARD; M. THAWLEY; J. SMALL. Satellite remote sensing of primary production: an improved production efficiency modeling approach. **Ecological Modeling**, v. 122, n. 3, p. 239, 1999.
- HOPE, A. C. A. A simplified Monte Carlo significance test procedure. **Journal Royal Statistical Society Series B**, v. 30, p. 582, 1968.
- HURVICH, C. M.; SIMONOFF, J. S.; TSAI C-L. Smoothing parameters selection in nonparametric regression using an improved Akaike information criterion. **Journal of the Royal Statistical Society Series B**, v. 60, p 271, 1998.
- IBGE. Paisagens do Brasil. Ministério do planejamento e coordenação geral. **Edição do Departamento de Documentação e Divulgação Geográfica e Cartográfica**, Instituto Brasileiro de Geografia e Estatística, 1997.

[http://www.ibge.gov.br/home/geociencias/default\\_prod.shtm#MAPAS](http://www.ibge.gov.br/home/geociencias/default_prod.shtm#MAPAS)

IRIBARNE, J. W.; GODSON, W. L. **Atmospheric Thermodynamics (Geophysics and Astrophysics Monographs)**. Springer, 1981. 278 p.

JETZ, W.; RAHBEK, C.; LICHSTEIN, J. W. Local and global approaches to spatial data analysis in ecology. **Global Ecology Biogeography**, v.14, p. 97, 2005.

KAUFMAN, Y. J.; TUCKER C. J.; FUNGI, I. Remote Sensing of biomass burning – method. In Proceedings of International Radiation Symposium, 18<sup>th</sup> 24<sup>th</sup> August 1988, Lille France (Deepak Publishing: Hampton VA), p. 488-493, 1988.

LAURENCE, W.; COCHRANE, M.A.; BERGEN, S., FEARNESIDE, P. M.; DELAMÔNICA, P.; BARBER, C.; D'ANGELO, S.; FERNANDES, T. The future of the Brazilian Amazon. **Science**, v. 291, p. 438, 2001.

LAURENCE, W.; ALBERNAZ, A. K.; SCHROTH, G., FEARNESIDE, P.; BERGEN, S.; VENTICINQUE, E.; COSTA, C. Predictors of deforestation in the Brazilian Amazon. **Journal of Biogeography**, v. 29, p. 737, 2002.

LEUNG, Y.; MEI, C-L, ZHANG, W-X. Statistical tests for spatial nonstationarity based on the geographically weighted regression model. **Environmental and Planning A**, v. 32, p. 9, 2000a.

LIN C.-H. & WEN, T.-H. Using Geographically weighted Regression (GWR) to Explore Spatial Varying Relationships of Immature Mosquitoes and Human Densities with the Incidence of Dengue. **International Journal of Environmental Research and Public Health**, v.8, p. 2798-2815, 2011.

MOTA, B. W.; PEREIRA, J. M.C., OOM, D., VASCONCELOS, M. J. P.; SCHULTZ, M. Screening the ESAATSR-2 World Fire Atlas (1997-2001). **Atmospheric Chemistry and Physics**, v. 6, p. 1409, 2006.

NEGREIROS, G. H.; STANBERG, D.; ALVARADO, E.; HINCKLEY, T.; NEPSTAD, D. C.; PEREIRA, M. Fire in the Transition between the Amazon Forest and Cerrado Ecosystems. In:

13th Conference on the Fire and Forest Meteorology, October 27-31. Lorne, Australia, 1996.

NEPSTAD, D. C.; MOREIRA, A. G.; ALENCAR, A. A. Flames in the rain forest: origins, impacts and alternatives to amazonian fires. Brasília, Brazil: **Pilot Program to Conserve the Brazilian Rain Forest, World Bank**, Brasilia, Brasil, 1999, 147p.

NEPSTAD, D., CARVALHO, G., BARROS, A. C., ALENCAR, A., CAPOBIANCO, J. P., BISHOP, J., MOUTINHO, P., LEFEBVRE, P., SILVA, U. L.; PRINS, E. Road paving, fire regime feedbacks, and the future of the Amazon forests, **Forest Ecology and Management**, v. 154, n. 3, p. 395, 2001.

PERES, C.A. Paving the way to the future of Amazonia. **Trends I Ecology & Evolution**, v. 16, n. 5, p. 217, 2001.

SÁ, A. C.; PEREIRA, J. M. C.; CHARLTON, M. E.; MORA, B.; BARBOSA, P.; FOTHERINGHAN, A. S. The pyrogeography of sub-Saharan Africa: a study of the spatial non-stationarity of fire-environment relationships using GWR. **Journal of Geographical Systems**, v. 13, p. 227, 2011.

SHI, H., LAURENT, E.J., LeBOUTON, J., RACEVSKIS, L., Hall, K. R., DONOVAN, M., DOEPKER, R. V., WALTERS, M.B., LUPI, F.; L., JIANGUO. Local spatial modelling of white-tailed deer distribution. **Ecological Modelling**, v. 190, p. 171, 2006.

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

WHEELER, D.; TIEFELSDORF, M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. **Journal of Geograph Systems**, v. 7, n. 2, p. 161, 2005.

WINDLE, M. J. S.; ROSE, G. A.; DEVILLERS, R.\_FORTIN, M.-J. Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. **International**

Council for the Exploration of the Sea. **Journal of Marine Science**. p.145, 2009.

ZHANG, L.; BI, H.; CHENG, P.; DAVIS, C. J. Modeling spatial variation in tree diameter-height relationships. **Forest Ecology and Management**, v. 189, p. 317, 2004.

ZHANG, L., GOVE, J. H.; HEATH, L. S. Spatial residual analysis of six modeling techniques. **Ecological Modeling**, v. 186, p. 154, 2005.