

# Fogos activos do sensor ATSR 2 e sua relação com variáveis demográficas, ambientais e infra-estruturais usando os métodos de regressão global e geograficamente ponderada para o Brasil

Adélia Maria Oliveira de Sousa<sup>1</sup>  
José Miguel Cardoso Pereira<sup>2</sup>  
Bernardo Wildung Cantante Mota<sup>3</sup>

<sup>1</sup>Departamento de Engenharia Rural, Universidade de Évora  
Instituto das Ciências Agrárias Mediterrânicas  
Pólo da Mitra, 7002-554 Évora, Portugal  
asousa@uevora.pt

<sup>23</sup> Departamento de Engenharia Florestal, Instituto Superior de Agronomia,  
Universidade Técnica de Lisboa  
Tapada da Ajuda, 1394-017 Lisboa, Portugal  
<sup>2</sup>jmcperreira@isa.utl.pt, <sup>3</sup>bwcmota@gmail.com

**Abstract.** In this study it was analysed the correlation 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. This last method allows the analysis of regression considering spatially non-stationary variables. The best adjustment of the ATSR active fires was found with the GWR method when it was considered the agricultural area classification and distance to roads variables.

**Palavras chave:** Fogos activos, ATSR, Brasil, OLS, GWR, active fire, ATSR, Brazil, OLS, GWR.

## 1. Introdução

O fogo no Brasil tem sido objecto de preocupação e polémica, atingindo diversos sistemas ecológicos gerando impactos ambientais a escalas globais, regionais e locais. A actividade do fogo tem sido relacionada principalmente com o uso da terra e com as condições climáticas (Cardoso *et al.*, 2003), nomeadamente o aumento da desflorestação que ocorre em consequência do aumento da densidade populacional, da construção de vias de acesso e da ocorrência de estações secas severas (Laurence *et al.*, 2002).

O objectivo deste estudo é avaliar a relação de diversos factores com a ocorrência e distribuição geográfica do fogo no Brasil, que pode ser avaliada pelo método de análise de regressão entre o fogo e esses factores. O método mais utilizado é a análise de regressão convencional, pelo método dos 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 para este método. A regressão global tem algumas limitações, o seu uso pode não ser apropriado para algumas situações e devem ser consideradas outras alternativas (Foody, 2003). Recentemente têm sido aplicadas outras técnicas de regressão, em diversos estudos de análise espacial com o objectivo de melhorar os resultados relativamente ao método de regressão global. Um dos métodos propostos é a 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

(Foody, 2003; Fotheringham *et al.*, 2002). Este 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, tendo 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).

No presente trabalho apresentamos uma comparação entre o método tradicional de regressão global, OLS e o recente método de regressão local (GWR), na tentativa de modelar o padrão de distribuição espacial da ocorrência de fogo no território Brasileiro.

## 2. Dados

### Fogos activos

Os dados referentes aos fogos activos utilizados foram do sensor *Along Track Scanning Radiometer 2* (ATSR-2), lançado pela *European Space Agency's* (ESA) a bordo do satélite *European Remote Sensing Satellite* (ERS-2) a 21 de Abril de 1995, integrado no Programa de Observação da Terra. Este sensor está equipado com três bandas na região do espectro electromagnético do visível, uma banda na região do infravermelho médio e três bandas na região do infravermelho térmico. O ATSR produz dados de detecção remota com uma resolução espacial de 1 km no nadir. Estes dados produzem informação da ocorrência de fogos no seu estado activo 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 µm (Arino & Rosaz, 1999). Esta banda é muito sensível à radiância emitida a temperaturas entre 500 °K e 1000 °K, permitindo assim a detecção de fogo activo mesmo que apenas uma pequena parte do pixel seja afectada pelo fogo.

Os produtos de fogos activos deste sensor designados por *World Fires Atlas* (WFA), são produzidos apenas a partir de observações recolhidas durante a noite. Este facto 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 e aplicados diversos filtros de forma a eliminar estes possíveis falsos alarmes. Os dados corrigidos de alguns erros passam a designar-se como produto WFA corrigido (WFAc).

Os dados de partida para o 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 (Figura 1).

Devido à existência de células com número de fogos igual a zero, adicionou-se uma constante de valor 1 a todas as observações. A escala dos dados foi alterada para o logaritmo da média que se deve à ocorrência de elevado número de fogos num baixo número de células, o que criava valores extremos muito elevados. Com esta modificação suavizaram-se os valores extremos dos dados.

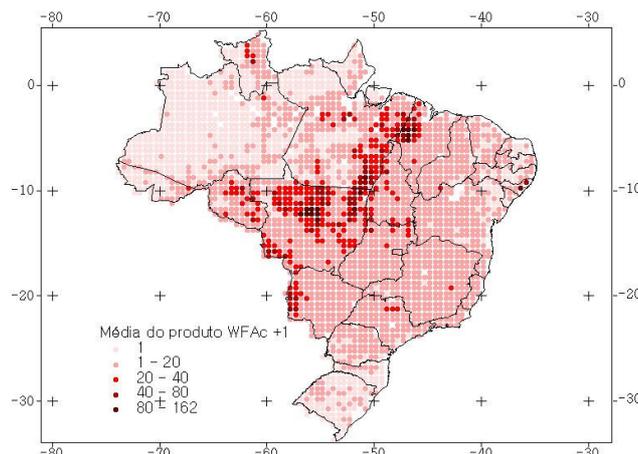


Figura 1. Média do produto WFAc, (1997 - 2002) adicionado de um valor, para uma grelha com resolução espacial de 0,5°. Os pontos a branco correspondem a células cujos dados foram eliminados.

### Dados demográficos, ambientais e infra-estruturais

A Tabela 1 apresenta uma breve descrição de todas as variáveis consideradas neste estudo. As variáveis foram amostradas para uma grelha com a resolução espacial de 0,5°.

A variável dependente é o número de fogos activos WFAc e as variáveis independentes são a densidade populacional, a produtividade primária líquida, um índice climático designado por *Chandler Burning Index* (Chandler *et al.*, 1983), a percentagem de área de cerrado, de agricultura e a distância às estradas para uma célula de 0,5°.

Tabela 1. Breve descrição das variáveis usadas neste estudo.

Variáveis	Abreviatura	Unidades	Mínimo	Máximo	Fonte
Fogos Activos	WFAc	Nº médio	0	162	WFA
Densidade Populacional	Dens_pop	Nºhabitantes/km <sup>2</sup>	0	6.868	GPW
Produtividade primária líquida	PPL	kgC/m <sup>2</sup> /ano	0,79	3,54	GLoPEM
Chandler Burning Index	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

Na Figura 2 apresentamos os mapas com a distribuição espacial de cada uma das variáveis independentes indicadas na Tabela 1.

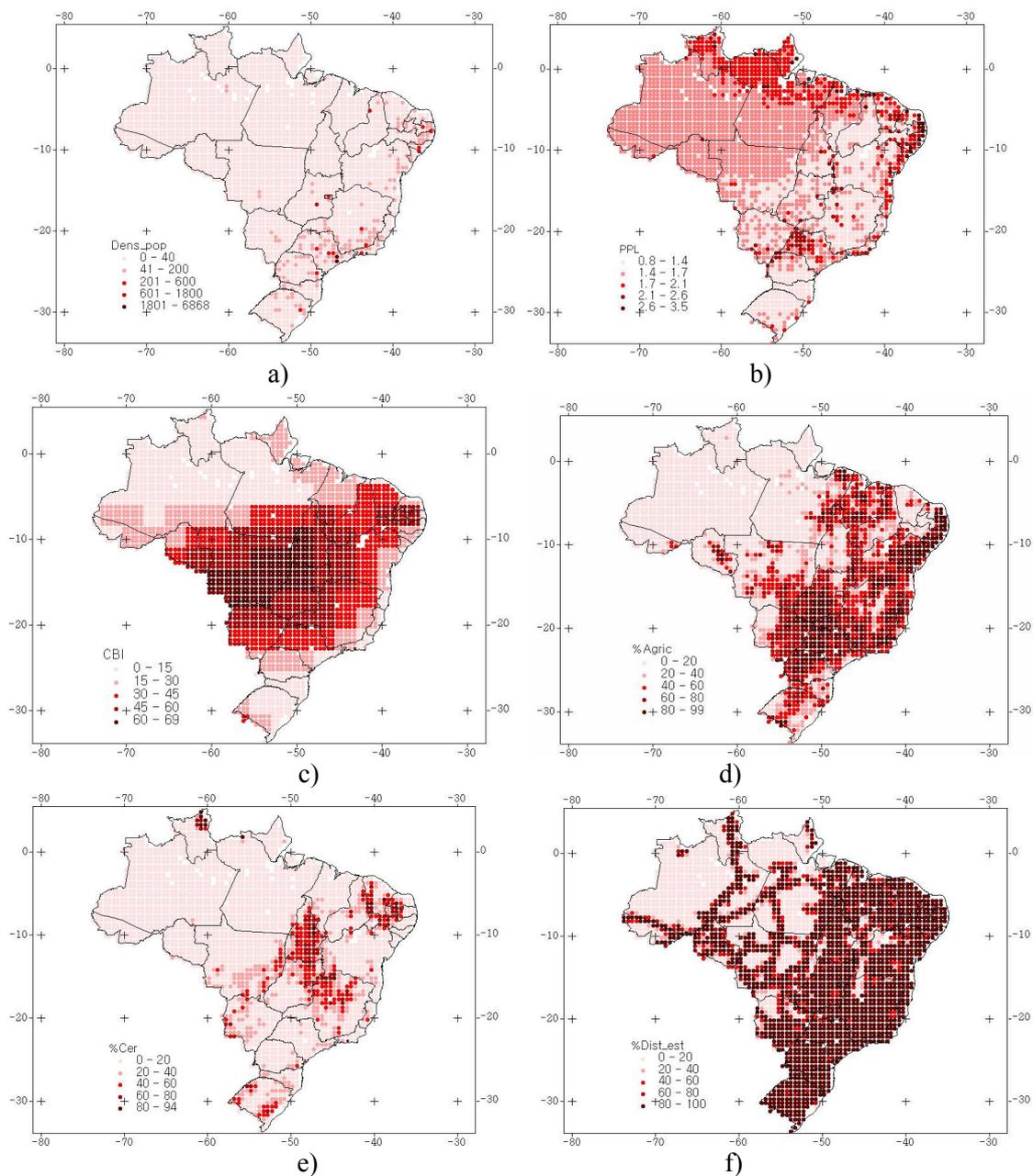


Figura 2. Distribuição espacial das variáveis independentes, a) Densidade populacional, b) Produtividade Primária Líquida, c) Chandler Burning Index, d) % de Agricultura, e) % de Cerrado e f) Distância às estradas (50km).

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

$$y = \beta_0(\mu, \nu) + \beta_1(\mu, \nu)x_1 + \dots + \beta_n(\mu, \nu)x_n + \varepsilon$$

O objectivo 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 constante espacialmente. 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 se 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 equivalente à regressão pelo método dos mínimos quadrados (Shi *et al.*, 2005).

Na estimativa dos parâmetros da regressão de cada ponto  $(\mu, \nu)$ , é necessário considerar a distância entre o ponto a estimar e as observações da sua vizinhança, ou seja, determinar quais as observações a incluir ou não na estimação dos parâmetros desse ponto. Utilizou-se o critério de selecção do *Akaike Information Criterion* (AIC) para obter o raio de influência para a estimativa dos parâmetros usando 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, comparar diferentes modelos e a regressão OLS com a regressão pelo método GWR. Para testar a variabilidade espacial da relação entre as variáveis, utilizou-se o teste de significância de Monte Carlo (Hope, 1968). Este teste 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.

Os conceitos do processo de estimação implementado pelo modelo de regressão geograficamente ponderada (GWR), podem ser obtidos mais detalhadamente em Fotheringham *et al.* (2002).

### **Modelos de regressão**

Procedeu-se ao ajuste de modelos univariados e multivariados para avaliar a contribuição das diversas variáveis na explicação da distribuição espacial dos fogos activos resultantes do produto WFAc. Foram testados diferentes modelos com diferentes combinações das variáveis independentes (Tabela 2), para ambos os métodos de regressão, OLS e GWR. A estratégia foi apresentar o modelo multivariado com as seis variáveis, os seis modelos univariados e ainda combinações dessas variáveis. Testaram-se três combinações de variáveis independentes, agrupando-as segundo factores humanos (Modelo 2), factores 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 factores humanos e naturais na ocorrência de fogos no Brasil. Os factores humanos integram a distribuição e actividade da presença humana, que se refere à densidade populacional, à presença de agricultura e à construção de estradas. No caso dos factores naturais, a produtividade primária líquida, o índice climático CBI e o tipo de coberto Cerrado, pretende-se verificar a sua influência na actividade do fogo.

#### 4. Resultados e conclusões

Na Tabela 2 apresentam-se os resultados obtidos para os onze modelos testados, pelos dois métodos de regressão global e local. Com base no método de regressão OLS o melhor modelo ajustado na explicação da variação espacial entre as variáveis independentes e os fogos activos 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 ( $R^2 = 0,38$ ). No entanto, este modelo não representa um bom ajustamento, relativamente aos resultados obtidos pela GWR.

Quando se compara o resultado obtido pelo método GWR, verificamos que o mesmo modelo multivariado apresenta melhor resultado, com valor de AIC = 4619 bastante inferior e  $R^2 = 0,78$ , bastante superior, indicando que permite um melhor ajustamento dos dados relativamente ao método de regressão global. Esta melhoria de ajustamento verifica-se para os onze modelos testados, com o método GWR sendo os valores de AIC sempre inferior aos obtidos com o método OLS. No entanto, o Modelo 11 é o que apresenta melhor ajustamento com o método GWR, com um valor de AIC = 4367 mínimo e um  $R^2 = 0,80$ , seguido do Modelo 7 (com a variável %Agric), com AIC = 4377.

Tabela 2. Variáveis independentes para cada modelo testado com o método de regressão OLS e GWR. Com os respectivos valores de AIC,  $R^2$  e raio da janela *kernel*. Um grau corresponde aproximadamente 110 km para a latitude média do Brasil.

Modelos	Variáveis independentes	OLS		GWR		
		$R^2$	AIC	$R^2$	AIC	Raio de janela (°)
Modelo 1	Dens_pop, PPL, CBI, %Agric, %Cer e %Dist_est	0,382	7617,06	0,783	4619	1,2194
Modelo 2	Dens_pop, %Agric e %Dist_est	0,095	8660,56	0,785	4526	0,9521
Modelo 3	PPL, CBI e %Cer	0,356	7723,61	0,782	4994	0,8545
Modelo 4	Dens_pop	0,001	8929,85	0,771	4863	0,7057
Modelo 5	PPL	0,003	8923,38	0,768	4927	0,7342
Modelo 6	CBI	0,343	7776,38	0,778	4739	0,6636
Modelo 7	%Agric	0,061	8758,36	0,793	<b>4378</b>	0,7523
Modelo 8	%Cer	0,037	8826,84	0,804	4511	0,6845
Modelo 9	%Dist_est	0,085	8686,35	0,791	4509	0,6972
Modelo 10	%Agric, %Cer e %Dist_est	0,105	8629,91	0,801	4405	0,9198
Modelo 11	%Agric e %Dist_est	0,093	8666,23	0,800	<b>4367</b>	0,8316

A variação do raio da função *kernel* mostra que quanto mais variáveis independentes tem o modelo maior é o raio da função *kernel* e, conseqüentemente, mais observações contribuem para a estimação de cada observação. 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 teste de Fisher (teste F) foi utilizado para testar a hipótese nula, de que o método de regressão local representa melhor ajustamento dos dados do que com o modelo de regressão global. Como resultado, rejeita-se a hipótese nula ( $P\text{-value} < 0,005$ ), comprovando-se que a relação entre as variáveis é não estacionária e que o GWR tem melhor ajustamento para a variável que representa o número de fogos activos.

Os mapas com a distribuição geográfica dos valores dos parâmetros estimados e dos resultados da aplicação do método GWR apresentados a seguir correspondem ao modelo 11 que mostrou melhor ajustamento dos dados. O mapa resultante da aplicação dos parâmetros estimados pelo Modelo 11 apresenta-se na Figura 3.

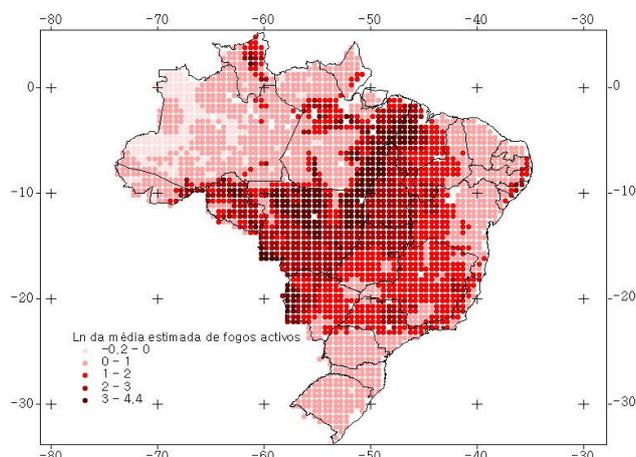


Figura 3. Média do produto WFAc estimado pelo método GWR (Modelo 11).

A Figura 4 mostra o diagrama de dispersão entre a média dos fogos activos observados e estimados segundo o método de regressão global ( $R^2 = 0,093$ ) e local ( $R^2=0,90$ ), para o Modelo 11.

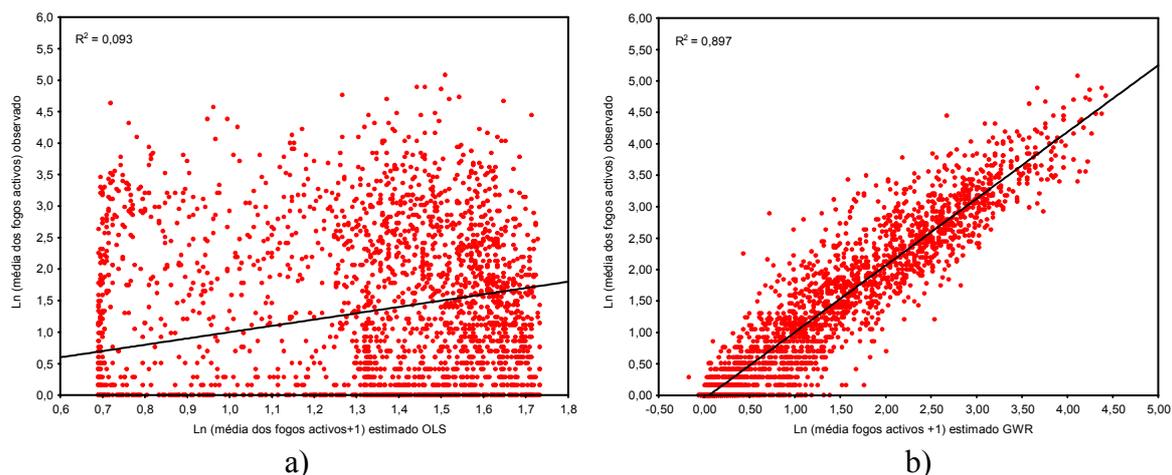


Figura 4. Diagrama de dispersão para a média do produto WFAc observado e estimado pelo método a) OLS e b) GWR (Modelo 11).

Na Figura 5, apresentam-se os resíduos obtidos pelo método de regressão OLS e GWR. O valor dos resíduos é maior e têm grande variação espacial pelo método global,

comparando com o método local onde obtemos valores menores e muito mais uniformes no espaço.

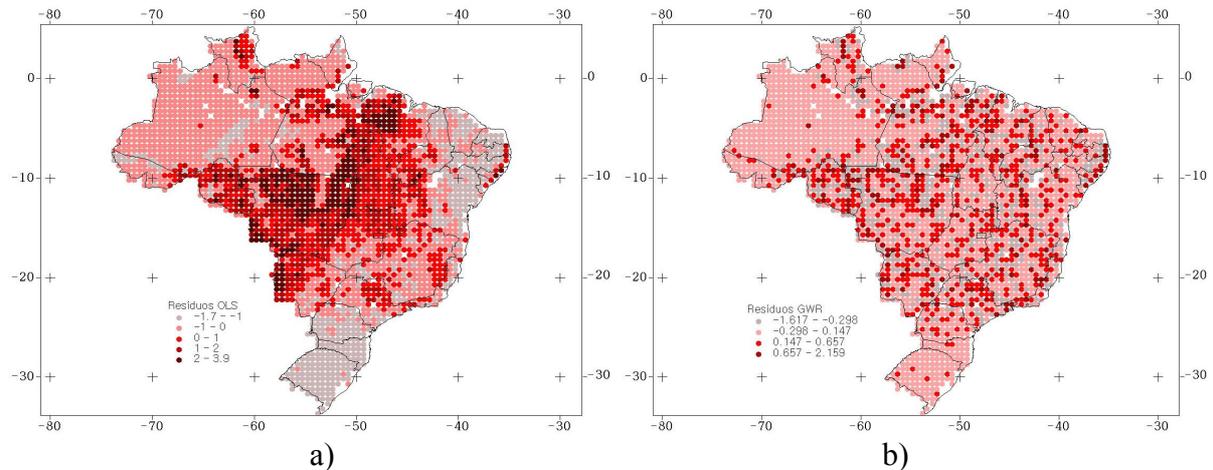


Figura 5. Distribuição espacial da resíduos segundo o método a) OLS e b) GWR (Modelo 11).

Para todos os modelos testados constatou-se uma melhoria significativa na estimação do número de fogos activos aplicando o método de regressão GWR comparativamente ao método OLS, o que se traduz pela existência de variabilidade espacial entre as variáveis utilizadas. Da análise dos resultados, verificou-se que a inclusão de todas as variáveis consideradas não permite a melhor representação da realidade da ocorrência do fogo no Brasil. Para os modelos univariados a junção da área agrícola e da distância às estradas superam todos os outros modelos, indicando a sua importância na ocorrência do fogo no Brasil.

## 5. Referências

- Arino, O. & Rosaz, J.M. 1997 and 1998 World ATSR fire atlas using ERS-2 ATSR-2 data. **Contributed papers for the Joint Fire Science Conference and workshop**. Boise Idaho, June 1999.
- Cardoso, M., Hurtt, G., Moore III, B., Nobre, C., & Prins, E. **Projecting the future fire activity in Amazonia. Global Change Biology**, v. 9, p. 656-669, 2003.
- Chandler, C., Cheney, P., Trabaud, L. e Williams, D. Fire in Forestry, Volume I, **Forest Fire Behavior and Effects**. Editora John Wiley & Sons, 1983.
- 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. 283-293, 2003.
- Fotheringham, A. S., Brunson, C. & Charlton, M. **Geographically Weighted Regression: the analysis of spatially varying relationships**. Editora John Wiley & Sons Inc. England, 2002.
- Hope, A. C. A. (1968). A simplified Monte Carlo significance test procedure. **Journal of the Royal Statistical Society Series, B** 30, p. 582-98.
- Laurence, W., Albernaz, A. K., Schroth, G., Fearnside, P., Bergen, S., Venticinque, E. & Costa, C. Predictors of deforestation in the Brazilian Amazon. **Journal of Biogeography**, v. 29, p. 737-748, 2002.
- Mota, B. W., Pereira, J. M.C., Oom, D., Vasconcelos, M. J. P., & Schultz, M. Screening the ESA ATSR-2 World Fire Atlas (1997-2001). **Atmospheric Chemistry and Physics**, v. 6, p. 1409-1424, 2006.
- Shi, H., Laurent, E.J., LeBouton, J., Racevskis, L., Hall, K. R., Donovan, M., Doepker, R. V., Walters, M.B., Lupi, F. & Liu, Jianguo. Local spatial modelling of white-tailed deer distribution. **Ecological Modelling**, v. 190, p. 171-189, 2006.
- Wang, Q., Ni, Jian & Tenhunen, J. Application of a geographically-weighted regression analysis net primary production of Chinese forest ecosystems. **Global Ecology and Biogeography**, DOI: 10.1111/j.1466-822x.2005.00153.x, 2005.
- Zhang, L., Bi, H., Cheng, P. & Davis, C. J. Modeling spatial variation in tree diameter-height relationships. **Forest Ecology and Management**, n.189, p. 317-329, 2004.