

---

# GWR e Geomarketing: Aventuras na Nova Zelândia

---

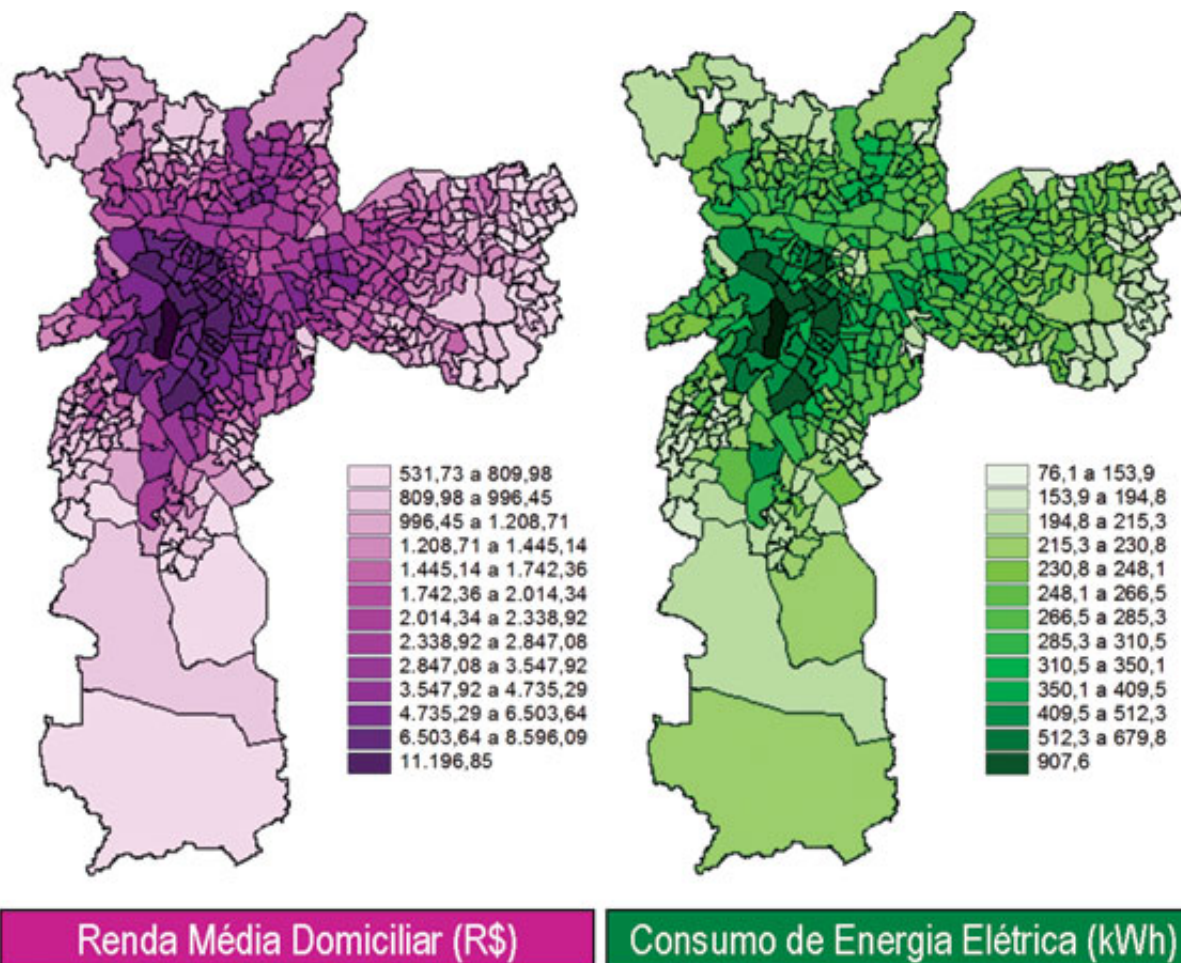
Por **MundoGEO** | 0100, 31 de Dezembro de 2007

Volto à InfoGEO já com saudades dos feedbacks dos leitores. Estou passando uma temporada como pesquisador visitante no Department of Information Science da Universidade de Otago, na Nova Zelândia. Oportunidade, no mínimo, fantástica. O aprendizado e a troca de experiências com pesquisadores como Peter Whigham e Tony Moore está sendo única.

Nesta edição volto a falar de estatística espacial (vejam a InfoGEO 46), desta vez apresentando aos leitores uma técnica nova, de simples implementação, mas com grande potencial, e que, de forma provocativa, fica como sugestão para um uso potencial para geomarketing.

Trata-se da GWR, ou Geographically Weighted Regression (Regressão Geograficamente Ponderada, estranha tradução literal para o português), apresentada pelos pesquisadores Fotheringham, Brunson e Charlton e consolidada em livro em 2002. É uma técnica espacial de análise exploratória que trabalha com indicadores locais ao invés dos modelos globais, mais tradicionais.

Vejam o exemplo a seguir, extraído da nossa conversa da edição 43 (GIS + Consumo de Energia = Indicadores Geográficos de Renda). Os mapas da Figura 1 apresentam a concentração de renda e de consumo de energia elétrica das 456 áreas de ponderação (conjuntos de setores censitários) do município de São Paulo. Vamos estudar o quanto o consumo de energia elétrica explica a renda familiar.



-> Figura 1

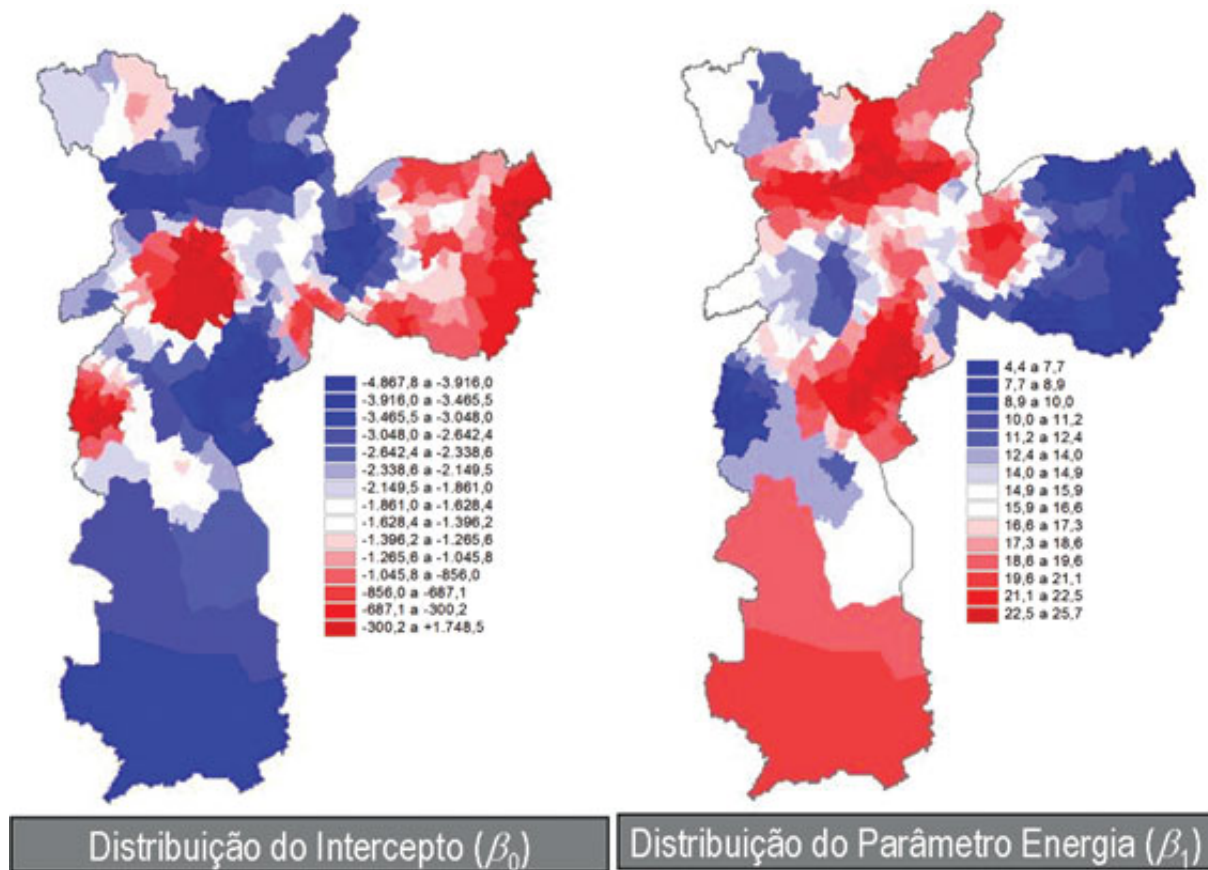
Um modelo global, tradicional, obtido através de regressão linear simples da renda, apresenta explicação de 86,8% da renda através do consumo de energia (coeficiente R<sup>2</sup>), o que é um valor bastante alto. Os parâmetros obtidos foram:

**Renda Estimada =  $\beta_0 + \beta_1 \cdot \text{Consumo de Energia}$**

**Renda Estimada (R\$) = -3034,71 + 19,55 . Consumo de Energia (kWh)**

O que significam esses parâmetros? Basicamente, se obtivermos o consumo de energia médio de uma determinada área estudada, multiplicarmos por 19,55 ( $\beta_1$ ) e subtrairmos 3.034,71 ( $\beta_0$ : intercepto) teremos uma boa estimativa da renda familiar média dessa área.

Mas será que essa fórmula é a melhor para todas as áreas de São Paulo? Para as áreas ricas (como Jardins e Higienópolis) e para as pobres (como Paraisópolis e Marsilac)? Aplicando a GWR (utilizando vizinhança de 9 áreas), obtivemos um coeficiente de explicação de diversos valores para esses parâmetros:  $\beta_0$  varia de -4.867 a +1.748 e  $\beta_1$  vai de 4,4 até 25,7. Os mapas da Figura 2 trazem essa variação (aliás, esse é o conceito de “mapeável” que está na tabela do artigo – os modelos globais não têm os parâmetros mapeáveis – eles são constantes no espaço). Além disso, o coeficiente de explicação do modelo subiu para mais de 97%!



-> Figura 2

Isso significa que as áreas de alta renda (centrais, pela Figura 1) têm baixa contribuição do consumo de energia elétrica (baixos valores de  $\beta_1$ ), mas partem de um patamar alto (altos valores de  $\beta_0$ ). Para as áreas de baixa renda, o fenômeno é exatamente o inverso. Assim, a estimativa de renda fica bem melhor e as diferenças locais do modelo se destacam.

Basicamente, se a variação espacial local aumenta, a confiabilidade das medidas globais como representações das condições locais diminui. Como estamos recheados de fenômenos de alta variabilidade espacial na vida real (socioeconômicos, ambientais, etc.), por que não experimentar a GWR?

Em técnicas globais (a tradicional regressão linear, por exemplo) utilizamos técnicas de minimização para obter a contribuição que cada variável explicativa traz para o fenômeno (variável resposta) que se deseja estudar. Para tal, normalmente todo o conjunto de observações é considerado, e as contribuições obtidas (os parâmetros da equação de regressão) são iguais em todo o espaço observado.

GWR e Geomarketing: Aventuras na Nova Zelândia

Indicadores sumarizados para a

região estudada como um todo

Desagregação local das estatísticas globais

Não "mapeáveis" Mapeáveis

Não indicados para visualização Adequados para visualização no GIS

Destaca SIMILARIDADES no espaço

Destaca DIFERENÇAS no espaço

Busca regularidades, ou “leis” Busca exceções, ou “hot spots”

Dados não espaciais ou espacialmente limitados

Dados espaciais

exemplo: exemplo:

Regressões clássicas (simples multivariadas)

Regressões espaciais com termo auto-regressivo (SAR)

•GWR – Geographically

Weighted Regression

•Regimes espaciais" hspace="5" src="http://mundogeo.com/wp-

content/uploads/2000/portugues/infogeo/50/pag37-2.gif" align="baseline" vspace="5" border="0" />

A GWR, por sua vez, ajusta um modelo de regressão a cada ponto observado, ponderando todas as demais observações como função da distância deste ponto. Em outras palavras, temos regressões diferentes para cada observação – assim, a contribuição (valor do parâmetro) de cada variável explicativa ao modelo é diferente para cada ponto. Vejam a tabela que compara as características dos modelos globais com os modelos locais (como a GWR).

Uma questão importante para os modelos locais é determinar qual o tamanho das regiões “locais”. Em outras palavras, qual vizinhança devo considerar no estudo local? Estatisticamente e computacionalmente, essa é a parte mais complicada da GWR. A técnica utiliza métodos adaptativos, como minimização do Akaike Information Criterion (AIC), ou cross-validation minimisation para sugerir ao usuário a melhor largura de banda da vizinhança a ser utilizada.

Parece complicado? É razoável? Sim! Quando tratamos de fenômenos em que o espaço é importante, e ocorre realmente variação espacial do fenômeno que buscamos estudar, é fundamental que tenhamos condições de tratar separadamente esse espaço e verificar como as variáveis estudadas se comportam diferentemente em cada local desse espaço.

Para dados socioeconômicos, os modelos de efeitos espaciais locais (como a GWR) tendem a apresentar melhores resultados que os da regressão simples ou das regressões espaciais com efeitos globais. Isso torna essa técnica bastante interessante para o geomarketing – os marqueteiros deveriam dar uma olhada!

Podem pensar o seguinte: técnica interessante, mas deve ser de difícil implementação e, por isso, cara. Pelo contrário, caros leitores. O software GWR3X, desenvolvido pelos próprios autores, em ambiente Windows, é barato e de fácil utilização. Para os adeptos do software livre, existe ainda uma extensão do poderoso pacote estatístico R especialmente desenhada para se trabalhar com a GWR (denominada SPGWR). Tanto o R quanto suas inúmeras extensões são totalmente gratuitos e disponíveis para download na internet.

Gostaram? Para maiores informações, teclem “GWR” (ou o nome por extenso) no Google e entrem no site oficial dos autores da técnica. Cuidado pois a “real” GWR que aparece no Google é a Great Western Railway, companhia ferroviária britânica.

**Eduardo de Rezende Francisco**

Mestre e doutorando em administração de empresas pela FGV-EAESP, bacharel em ciência da computação pelo IME-USP, atua em GIS e estratégias de marketing na AES Eletropaulo e é presidente da Gita Brasil ([www.gita.org.br](http://www.gita.org.br)). Atualmente é pesquisador visitante do Department of Information Science da Universidade de Otago, na Nova Zelândia  
**[eduardo.francisco@aes.com](mailto:eduardo.francisco@aes.com)**