

Utilização de Métodos de Análise Espacial na Avaliação em Massa de Imóveis

Dr. Eng. Carlos Alberto Peruzzo Trivelloni
carlos.peruzzo@gmail.com

Resumo. Este trabalho apresenta uma proposta de utilização de métodos de análise espacial para avaliação em massa de imóveis. Entre todas as variáveis que influenciam no valor dos imóveis, as de localização são as mais complexas de analisar e modelar. A multiplicidade de fatores ambientais, sociais e econômicos que influenciam no valor dos imóveis apresenta dificuldades aos métodos inferenciais tradicionais, provocando falhas de especificação, autocorrelação espacial nos resíduos, e comprometendo a confiabilidade da avaliação. Os métodos de regressão espacial e geoestatísticos (krigagem) são usados neste trabalho em forma combinada para estimar o valor da localização dos imóveis. Uma aplicação é realizada para uma amostra de mercado do Município de São José, Santa Catarina. O método permitiu encontrar um índice do valor da localização fortemente significativo no modelo de regressão em massa para todos os imóveis, eliminando a autocorrelação espacial nos resíduos e aumentando o poder explicativo e a confiabilidade da avaliação.

Palavras chave: avaliação em massa, valor de localização, análise espacial, regressão espacial, krigagem.

Abstract. This work presents a method for mass appraisal estimation of real estate using spatial analysis. Between all variables influencing on real estate value, location factors are the most difficult of analyzing and modeling. The multiplicity of environmental, social and economics factors are difficult of being modeled with traditional inferential techniques, introducing specification problems and spatial autocorrelation on model errors. Spatial models, especially spatial regression and geostatistical models are used for estimation of location value. The spatial error model and block kriging allow estimate a regionalized variable, which represents the location value of real estate. An application of this method is presented with a sample market of the municipality of São José, Estate of Santa Catarina. The proposed method allowed estimating an index of location value highly significant and modeled with consistency and coherency the effects of all valuation factors. The location index eliminated residual spatial autocorrelation in the model, augmenting its explanation capacity and the reliability of the real estate valuation.

Keywords: mass appraisal, location value, spatial statistics, spatial regression, kriging.

1. Introdução

Entre todas as variáveis que influenciam no valor dos imóveis, aquelas referidas à localização são as mais complexas de analisar e modelar, principalmente nos modelos de avaliação em massa.

Os modelos tradicionalmente utilizados na avaliação em massa, baseados na inferência estatística por mínimos quadrados, geralmente apresentam dificuldades para lidar com a complexidade de fatores de localização.

Os fatores de localização que valorizam os imóveis são inúmeros: proximidade a centros comerciais, centros de educação, de saúde, segurança pública, padrão construtivo, pólos ambientais desejáveis ou indesejáveis, entre outros. A lista completa destes fatores pode ser extremamente ampla.

Ao mesmo tempo, a dinâmica urbana produz uma contínua mudança nos efeitos destes fatores sobre o valor dos imóveis: novos investimentos públicos ou privados, novos empreendimentos econômicos, atividades comerciais, empreendimentos imobiliários, loteamentos, entre outros, produzem mudanças significativas nas características de cada vizinhança e nos valores imobiliários. Quase toda mudança urbana afeta o valor dos imóveis próximos, transformando a localização no fator mais dinâmico da sua valorização.

Os métodos de avaliação em massa devem considerar estas características espaciais de forma adequada. A estatística espacial constitui uma ferramenta indispensável nesse sentido.

2. Modelos tradicionais de avaliação em massa

Os modelos tradicionais de avaliação imobiliária procuram explicar o valor a partir da estimação por inferência estatística das variáveis influenciadoras, sendo estas classificadas em características construtivas próprias do imóvel e características de localização, de acessibilidade e vizinhança.

As variáveis de acessibilidade, definidas em função da distância do imóvel a um pólo de valorização ou desvalorização, podem não ser estatisticamente significativas nos modelos de inferência por vários motivos, como a composição insuficiente da amostra ou a especificação equivocada da função distância no modelo.

As variáveis de vizinhança, que definem a localização a partir de áreas onde este efeito pode ser considerado semelhante para os imóveis, podem resultar não significativas por insuficiência na amostragem ou por erros na delimitação destas zonas.

Quando as variáveis relacionadas com a localização não são corretamente especificadas, outros problemas podem surgir. Além da perda de poder de explicação, a presença de autocorrelação espacial nos resíduos invalida uma das hipóteses básicas da inferência, tornando ineficientes as estimativas e inválidos os testes de hipótese.

3. Métodos de análise espacial

A dependência espacial do valor dos imóveis, produzida pela sua localização, tem mostrado a insuficiência das técnicas tradicionais de inferência para avaliação em massa, levando ao desenvolvimento de técnicas de estatística espacial para o tratamento de dados imobiliários.

Na literatura de avaliação imobiliária existem duas metodologias para o tratamento espacial dos dados: a econometria espacial e seus modelos de regressão espacial, por um lado, e o uso da geoestatística, por outro.

Anselin (2002) fundamenta a escolha por um ou outro modelo em função do tipo de dados espaciais considerados. Os modelos espaciais trabalham com dois tipos de dados: os geobjetos e os geocampos. Os geobjetos são entidades discretas, geralmente representadas nos Sistemas de Informações Geográficas como pontos, linhas ou polígonos, enquanto os geocampos pertencem a uma distribuição espacial contínua representada geralmente como superfícies. Segundo este autor, os geobjetos devem ser tratados por econometria espacial, enquanto que os geocampos devem ser analisados com modelos geoestatísticos. No caso do mercado imobiliário os dados apresentam simultaneamente comportamentos identificáveis como geobjetos e geocampos, devido a natureza dos fatores influenciadores no seu valor, abrindo a possibilidade de tratamento por ambas metodologias.

O uso combinado de ambas técnicas (Peruzzo Trivelloni, 2005) permite superar algumas das deficiências e limitações dos métodos espaciais, potencializando as suas vantagens com resultados satisfatórios. O uso combinado permite o tratamento adequado para os dois tipos de variáveis influenciadoras no valor: a análise por econometria espacial das características construtivas e a análise por geoestatística dos fatores de localização.

3.1 Econometria espacial

Segundo Anselin (1999), a econometria espacial é a parte da econometria que lida com o tratamento da estrutura espacial e interações de dados geograficamente distribuídos. A dependência espacial pode ser

incorporada nos modelos de regressão de duas formas: no modelo de Defasagem Espacial e no modelo do Erro Espacial. Formalmente, o modelo de Defasagem Espacial é expresso da seguinte forma:

$$Y = \rho W y + X B + \varepsilon$$

onde Y é o vetor da variável dependente, ρ é o coeficiente de autocorrelação espacial, W a matriz de pesos espaciais, X é a matriz das observações nas variáveis independentes dos dados, B é o vetor de parâmetros e ε é o vetor de resíduos do modelo.

Por outro lado, o modelo do Erro Espacial é expresso formalmente da seguinte forma:

$$\begin{aligned} y &= X B + \varepsilon, \\ \varepsilon &= \lambda W \varepsilon + u \end{aligned}$$

onde λ é o coeficiente de autocorrelação espacial, u é o vetor de resíduos não correlacionados e os outros termos são os mesmo do modelo anterior.

Estudos sobre autocorrelação espacial nos dados do mercado imobiliário e aplicação dos modelos de regressão espacial podem ser encontrados em Dubin (1988), Can (1990 e 1992), Pace e Gilley (1997), Pace et al. (1998), Dantas (2003), Dantas et al (2002), Peruzzo Trivelloni (2005), entre outros.

3.2 Geoestatística. Teoria das Variáveis Regionalizadas

Segundo Chica Olmo (1994) quando existe dependência entre uma variável aleatória e a sua localização espacial a variável é chamada regionalizada. A Teoria das Variáveis Regionalizadas (TVR) dedica-se à análise e inferência espaciais de variáveis que dependem da localização espacial onde são observadas. Segundo este autor, a variação espacial de uma variável regionalizada pode ser expressa pela soma de três componentes. Se o vetor x representa uma posição em uma, duas ou três dimensões, então o valor da função aleatória Z , em x , é dada por:

$$Z(x) = \mu(x) + \varepsilon'(x) + \varepsilon''$$

Onde:

$\mu(x)$ é uma função determinística que descreve a componente estrutural de Z em x ;

$\varepsilon'(x)$ é um termo estocástico correlacionado que varia localmente;

ε'' é um ruído aleatório não correlacionado, com distribuição normal com média 0 e variância σ^2 .

Considerando uma amostra $Z(x_1), \dots, Z(x_n)$, e sendo necessário estimar o valor $Z(x_0)$ num ponto x_0 , o estimador de krigagem Z_k define-se como uma combinação linear dos $Z(x_i)$ e os ponderadores λ_i da seguinte maneira (Chica Olmo, 1994):

$$Z_k = \sum_i \lambda_i Z_i$$

Onde os ponderadores λ_i são obtidos de forma que o estimador Z_k seja ótimo, ou seja, a variância do erro seja mínima. Os ponderadores são calculados a partir da estrutura geométrica dos dados, atribuindo valores menores aos pontos mais afastados e valores maiores aos pontos mais próximos.

Estudos de aplicação dos métodos geoestatísticos na área de avaliação imobiliária podem ser encontrados em Dubin (1992), Chica Olmo (1994), Cano Guervós (1999), Gámez Martínez (1999), Peruzzo Trivelloni (2005), entre outros.

4. Metodologia

Quando analisado o mercado imobiliário em uma região geográfica específica por meio de uma amostra de mercado em condições semelhantes de transação e coletada na mesma época, o valor dos imóveis poderá ser considerado pelo modelo geral a seguir:

$$V = f(L, T, CF_T)$$

Onde **V** representa o valor do imóvel, **L** representa as variáveis de localização, **T** representa o tipo de imóvel e **CF_T** representa as características físicas do imóvel em função do tipo.

Segundo Peruzzo Trivelloni (2005) o valor de localização dos imóveis, considerado como a interação ou efeito combinado de todos os fatores influenciadores na vizinhança do imóvel, pode ser analisado como uma variável regionalizada e modelado por técnicas geoestatísticas.

A expressão geral do modelo utilizado é a seguinte:

$$VU = VL * F_{TIPO} * F_{CF1} * F_{CF2} * \dots * F_{CFn}$$

Onde **VU** é o valor unitário do imóvel, **VL** representa o valor da localização, **F_{TIPO}** é um fator de correção dependente do tipo de imóvel e os **F_{CFi}** representam fatores correspondentes às características construtivas para cada tipo de imóvel.

Para estimar **VL** é necessário homogeneizar os dados de mercado em relação às características construtivas. Sendo os imóveis objetos espaciais a metodologia estatística aplicável para homogeneização pertence a econometria espacial.

A área de estudo corresponde aos bairros Kobrasol e Campinas do município de São José, Santa Catarina. Os bairros Kobrasol e Campinas constituem um dos centros mais dinâmicos do município. Foi realizada uma pesquisa de mercado na área de estudo para levantamento dos valores praticados para diferentes tipos de imóveis. Foram coletadas informações correspondentes a 249 imóveis a venda na região, correspondendo a 175 apartamentos, 21 kitnetes, 27 casas, 17 terrenos e 9 imóveis comerciais. A localização dos dados da amostra na área de estudo aparece na Figura 1. As características levantadas para todos os imóveis foram as seguintes: tipo de imóvel, localização, área total e privativa, padrão construtivo, existência ou não de garagem na unidade, idade do imóvel. Para as casas a pesquisa incluiu a área construída e área total do terreno. As variáveis construtivas para os diferentes tipos de imóveis foram definidas como aparece na Tabela 1.

Tabela 1. Resumo de variáveis de tipo e construtivas dos imóveis da amostra.

NOME	TIPO	DESCRIÇÃO
APT	dicotômica	1 se é apartamento, 0 se não.
KIT	dicotômica	1 se é kitinete, 0 se não.
CASA	dicotômica	1 se é casa, 0 se não.
COMERCIA	dicotômica	1 se é comercial, 0 se não.
TERRENO	dicotômica	1 se é terreno, 0 se não.
APT_AT	contínua	Área total do apto em m ² , 0 se não é apto.
KIT_AT	contínua	Área total da kitinete em m ² , 0 se não é kitinete.
CASA_AC	contínua	Área total construída da casa, 0 se não é casa.
CA_ATER	contínua	Área total do terreno para casas, 0 se não é casa.
TER_AT	contínua	Área total do terreno, 0 se não é terreno.
AP_P2	dicotômica	1 para aptos de padrão 2 (alto), 0 em outro caso.
AP_P3	dicotômica	1 para aptos de padrão 3 (médio), 0 em outro caso.
KIT_P2	dicotômica	1 para kit de padrão 2 (alto), 0 em outro caso.
CASA_P2	dicotômica	1 para casas de padrão 2 (alto), 0 em outro caso.
CASA_P3	dicotômica	1 para casas de padrão 3 (médio), 0 em outro caso.
APT_GAR	dicotômica	1 para apto com garagem, 0 em outro caso.
NOVO	dicotômica	1 para imóvel com idade até 3 anos, 0 em outro caso.

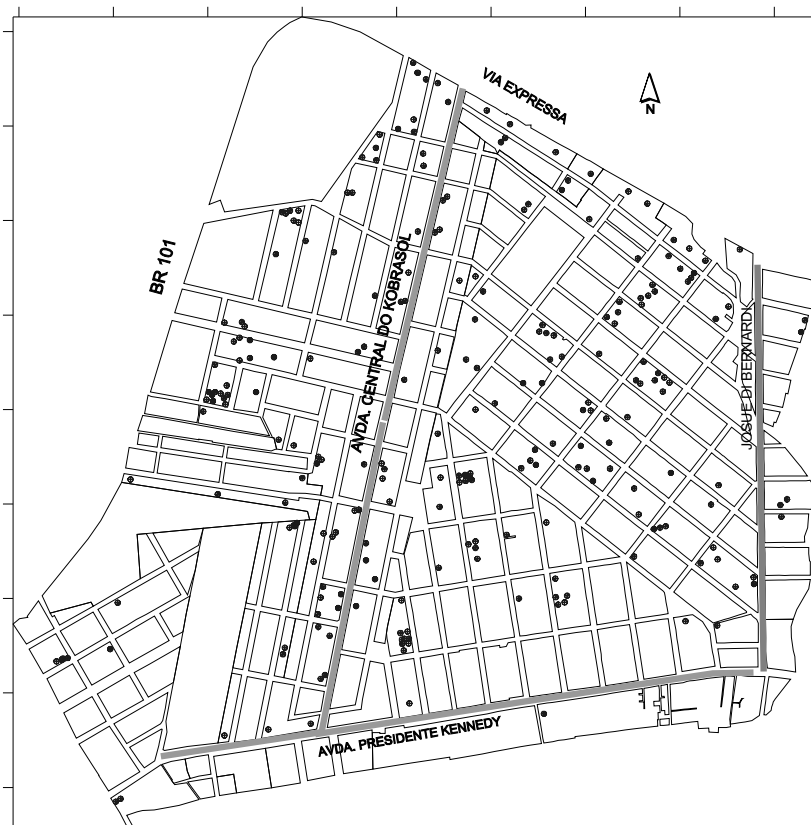


Figura 1. Localização dos dados da pesquisa de mercado. Fonte: Peruzzo Trivelloni (2005).

4.1 Regressão espacial.

A análise de autocorrelação espacial dos resíduos do modelo de mínimos quadrados pode ser observada na Tabela 2.

Tabela 2. Diagnóstico de autocorrelação espacial.

TESTE	VALOR	PROBABILIDADE
Moran's I	6.2724	0.0000
LM (lag)	1.4601	0.2269
LM robusto (lag)	0.1139	0.7357
LM (err)	30.3870	0.0000
LM robusto (err)	29.0408	0.0000

Os resultados mostram a existência de forte autocorrelação espacial: o teste de Moran e os dois testes do modelo espacial do erro mostram-se fortemente significativos.

Os resultados para o modelo de regressão espacial do erro são apresentados na Tabela 3. A variável LAMBDA representa o termo de autocorrelação espacial do erro e se mostra fortemente significativa, indicando a existência de forte autocorrelação espacial.

4.2 Aplicação do método geoestatístico.

Tomando antilogaritmos na equação de regressão, o valor unitário homogeneizado pode ser calculado a partir da expressão anterior como é mostrado a seguir:

$$VH = VU / [1.0008^{APT_AT} * 1.1357^{APT_GAR} * 1.1137^{AP_P2} * 1.0977^{AP_P3} * 4.2832^{KIT} * 0.9667^{KIT_AT} * 1.1733^{KIT_P2} * 2.9531^{CASA} * 0.9955^{CASA_AC} * 1.0005^{CA_ATER} * 1.7415^{CASA_P2} * 1.2517^{CASA_P3} * 0.7828^{TERRENO} * 1.0004^{TER_AT} * 1.6608^{COMERCIA} * 1.1128^{NOVO}]$$

Tabela 3. Modelo espacial do erro.

VARIÁVEL	COEFICIENTE	ERRO PADRÃO	VALOR Z	PROBABILIDADE
CONSTANTE	6.50271	0.08896	73.09392	0.00000
APT AT	0.00075	0.00053	1.41539	0.15695
APT GAR	0.12721	0.03580	3.55383	0.00038
AP P2	0.10769	0.07750	1.38960	0.16465
AP P3	0.09323	0.07173	1.29975	0.19369
KIT	1.45470	0.37314	3.89854	0.00010
KIT AT	-0.03390	0.00990	-3.42323	0.00062
KIT P2	0.15984	0.10076	1.58632	0.11267
CASA	1.08286	0.21638	5.00436	0.00000
CASA AC	-0.00448	0.00065	-6.88557	0.00000
CA ATER	0.00054	0.00017	3.19073	0.00142
CASA P2	0.55473	0.11194	4.95574	0.00000
CASA P3	0.22451	0.10541	2.12985	0.03318
TERRENO	-0.24482	0.10069	-2.43156	0.01503
TER AT	0.00036	0.00009	3.82849	0.00013
COMERCIA	0.50730	0.10238	4.95519	0.00000
NOVO	0.10692	0.04494	2.37892	0.01736
LAMBDA	0.87748	0.02278	38.52470	0.00000

O semivariograma experimental de VH aparece na Figura 2, onde pode ser analisada a estrutura de variação espacial, mostrando menor variância para as observações mais próximas e maiores variâncias para as observações mais distantes.

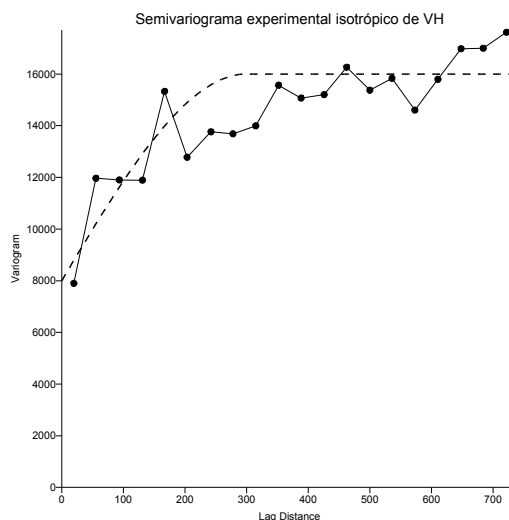


Figura 2. Semivariograma experimental isotrópico do valor unitário homogeneizado por regressão espacial.

A análise dos semivariogramas experimentais para todas as direções comprova a existência de um padrão de variação espacial esférico de semivariograma. Os parâmetros deste semivariograma podem ser observados na Tabela 4.

Tabela 4. Parâmetros do semivariograma estimado.

PARÂMETRO	NOTAÇÃO	VALOR
Efeito Pepita	Co	8000
Patamar	C	8000
Alcance	A	600
Ângulo de Anisotropia	α	90°
Fator de Anisotropia	F	2.0

Utilizando os parâmetros do semivariograma e uma grade de krigagem de 100 metros de lado foi realizada a krigagem da variável VH, obtendo a nova variável VL. As curvas de isovalores de VL aparecem na Figura 3 desenhadas a cada 50 R\$/m². Nesta Figura podem ser observados os principais centros de valorização ou desvalorização na área de estudo. As isolinhas permitem analisar a localização e raio de influência de cada um destes pólos.

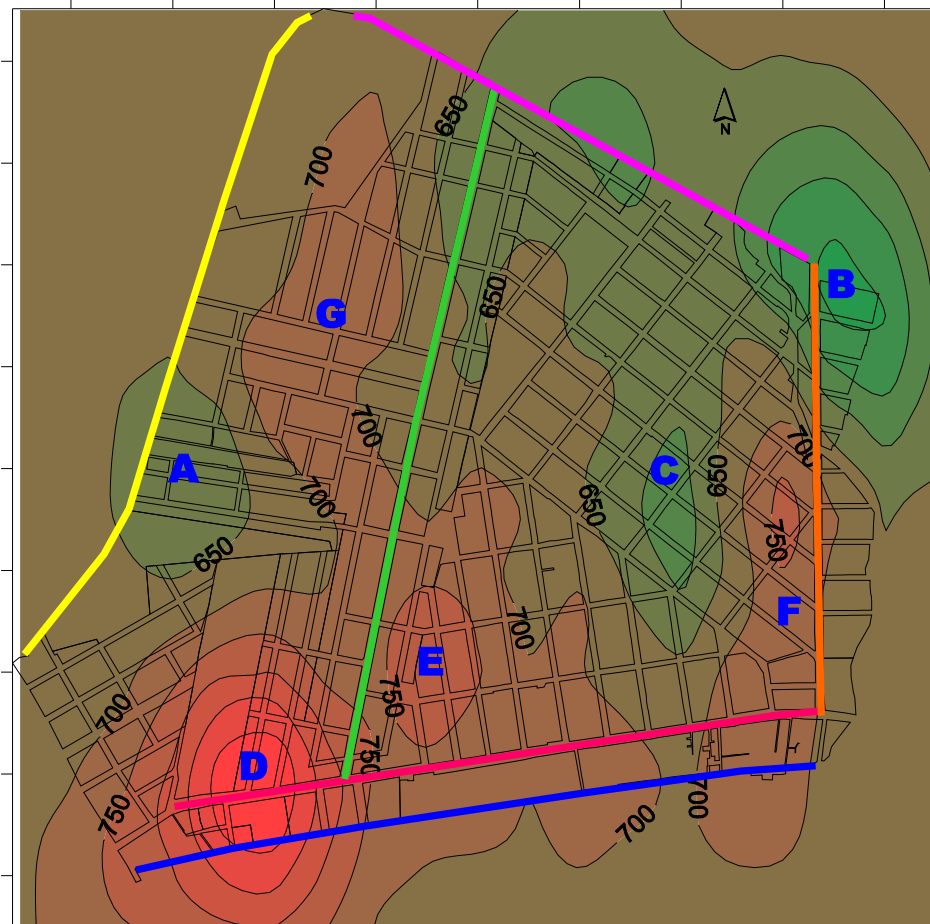


Figura 3. Curvas de isovalores da variável VL. Fonte: Peruzzo Trivelloni (2005).

A região mais valorizada da área de estudo se encontra nas proximidades da Avenida Kennedy (linha vermelha) e a Avenida Central do Kobrasol (linha verde). A Figura mostra duas regiões, denominadas D e E onde os valores são maiores, em áreas de atual expansão comercial e residencial, próximas também ao Centro Comercial Campinas. A valorização da região mais próxima à Avenida Kennedy é resultado também da Beira Mar de São José (linha azul na Figura 6). Outra região valorizada, denominada G na Figura 6, é caracterizada pela presença de escolas e centros educacionais, atividades de serviços e área residencial, com ruas tranquilas e amplas, tendo boa acessibilidade aos principais centros comerciais e as vias de entrada e saída da área de estudo. A Figura 6 permite analisar também as regiões menos valorizadas. Os valores diminuem próximos à BR 101 (linha amarela) e à BR 282 (linha rosa). As regiões A e B mostram a menor valorização. No caso da região A, a desvalorização responde a presença ruas estreitas de difícil acessibilidade e padrão construtivo inferior nos imóveis. No caso da região B, a presença próxima de uma região de favela diminui sensivelmente o valor dos imóveis tendo sendo os menores de toda a área de estudo. Finalmente a região denominada C, onde o valor também diminui, se encontra mais afastada dos principais centros comerciais da região, tendo por este motivo menor atratividade.

Como observado, a krigagem permite analisar e interpretar consistentemente a complexa interação de fatores de valorização e desvalorização da região.

4.3 MODELO DE AVALIAÇÃO EM MASSA. APLICAÇÃO PARA PLANTA DE VALORES GENÉRICOS DE TERRENOS

Para verificar a significância estatística da variável VL foi calculado novo modelo de regressão por mínimos quadrados onde VL é incluída como variável explicativa. Os resultados do novo modelo de regressão, observados na Tabela 5, mostram a forte significância da variável VL, eliminando a autocorrelação espacial nos resíduos e o poder de explicação geral do modelo.

Tabela 5. Significância das variáveis explicativas da regressão.

VARIÁVEL	COEFICIENTE	ERRO PADRÃO	VALOR t	PROBABILIDADE
Constante	5.1641	0.1250	41.3113	0.0000
APT GAR	0.1569	0.0250	6.2891	0.0000
AP P2	0.1248	0.0578	2.1577	0.0320
AP P3	0.0690	0.0531	1.3001	0.1949
KIT	1.6080	0.2742	5.8636	0.0000
KIT AT	-0.0393	0.0075	-5.2234	0.0000
KIT P2	0.1589	0.0762	2.0865	0.0381
CASA	0.7146	0.1934	3.6956	0.0003
CASA AC	-0.0032	0.0007	-4.8872	0.0000
CA ATER	0.0007	0.0002	4.3959	0.0000
CASA P2	0.4018	0.1123	3.5770	0.0004
CASA P3	0.2700	0.0989	2.7299	0.0068
TERRENO	-0.2577	0.0782	-3.2958	0.0011
TER AT	0.0004	0.0001	5.3034	0.0000
COMERCIA	0.4554	0.0673	6.7691	0.0000
NOVO	0.1180	0.0297	3.9768	0.0001
VL	0.0020	0.0002	11.9759	0.0000
Variável dependente	L(VU)			
R ²	0.7661			
R ² _a	0.7496			
F	46.276			
Prob (F)	<0.0001			
Erro padrão	0.1308			

O método proposto permite calcular o valor para todos os tipos de imóveis incluídos na amostra e em toda a área considerada. Aplicando o modelo anterior foi obtido o valor unitário para o ponto médio de cada face de quadra da área de estudo.

Como aplicação do método foi calculada a Planta de Valores Genéricos de Terrenos na área de estudo, estimando os valores médios por face de quadra de um terreno do tamanho padrão de 360 metros quadrados.

Aplicando o modelo foi obtida a Planta de Valores Genéricos de Terrenos, em R\$/m², mostrada na Figura 4.

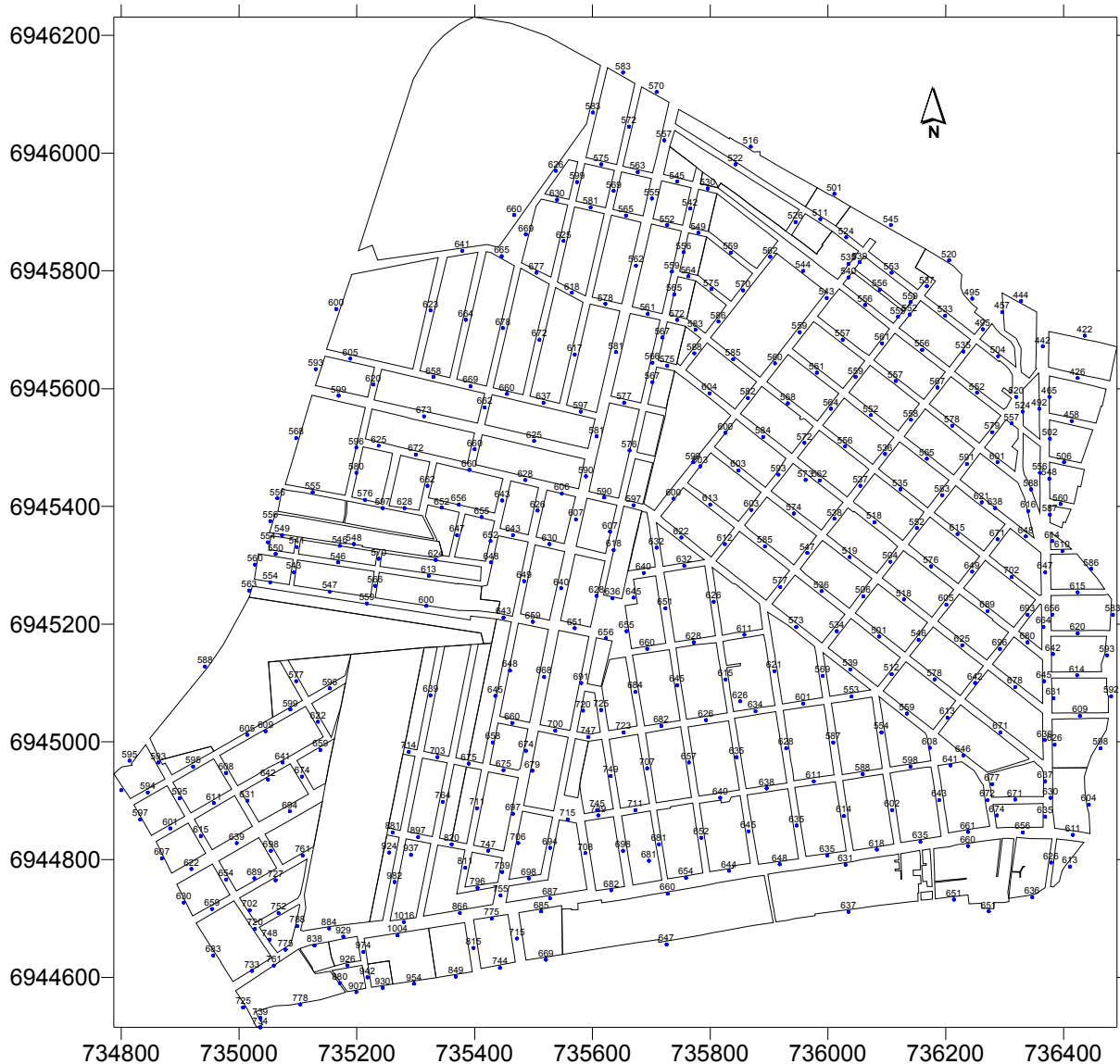


Figura 4. Planta de Valores Genéricos de Terrenos (R\$/m²) da área de estudo (coordenadas UTM). Fonte: Peruzzo Trivelloni (2005).

5. Conclusões.

- O método utilizado mostrou que o valor da localização de um imóvel é o resultado da interação de diversos pólos e efeitos de valorização e desvalorização. Este valor pode ser considerado como uma variável regionalizada e modelado por geoestatística. A multiplicidade de pólos de valorização ficou evidenciada, cada um deles caracterizado por uma área de influência, gradiente e anisotropias de variação específicas.
- O método permitiu estimar um índice do valor da localização fortemente significativo no modelo de regressão.
- O resultado comprovou a existência de autocorrelação espacial nos resíduos da regressão quando as variáveis de localização não são adequadamente consideradas, invalidando neste caso o uso de inferência por mínimos quadrados.

•As técnicas de econometria espacial e de geoestatística, tradicionalmente utilizadas de forma separada e excludente, podem ser usadas em forma combinada na avaliação em massa de imóveis, permitindo analisar de forma adequada as diferentes características dos imóveis.

6. Referências bibliográficas.

Anselin, L. *Spatial Econometrics. Discussion paper*. Bruton Center, School of Social Sciences, University of Texas at Dallas, 1999.

Anselin, L. *Under the Hood. Issues in the Specification and Interpretations of Spatial Regression Models*. Agricultural Economics, vol.27, Issue 3, pp 247-267, 2002.

Can, A. *The Measurement of Neighborhood Dynamics in Urban House Prices*. Economic Geography, vol. 66, pp. 254-272, 1990.

Can, A. *Specification and estimation of hedonic housing price models*. Regional Science and Urban Economics, n. 22, pag. 453-474, 1992.

Cano Guervós, R. *Aproximación al Valor de la Vivienda. Aplicación a la Ciudad de Granada*. Ed. Universidad de Granada, 1999.

Chica Olmo, J. *Teoría de las Variables Regionalizadas. Aplicación en Economía Espacial y Valoración Inmobiliaria*. Ed. Universidad de Granada, 1994.

Dantas, R. A. *Modelos Espaciais aplicados ao Mercado Habitacional. Um Estudo de Caso para a Cidade do Recife*. Tese de Doutorado em Economia. Universidade Federal de Pernambuco. Recife, 2003.

Dantas, R. A., Magalhães, A. M., Rocha, F. J. S. *La importancia de la regresión espacial en la tasación inmobiliaria*. I Congreso Internacional en Tasación y Valoración. Valencia, 2002.

Dubin, R. *Estimation of Regression Coefficients in the Presence os Spatially Autocorrelated Error Terms*. The Review of Economics and Statistics, Vol. 70, 3, pp. 466-474, 1988.

Dubin, R. *Spatial autocorrelation and neighborhood quality*. Regional Science and Urban Economics, n. 22, pag. 433-452, 1992.

Pace, R. K., Gilley, O. W. *Using the Spatial Configuration of the Data to Improve Estimation*. Journal of Real Estate and Economics, vol. 14, número 3, pp. 333-340, 1997.

Pace, R. K., Barry, R., Sirmans, C. F. *Spatial Statistics and Real Estate*. Journal of Real Estate Finance and Economics, Vol. 17, N. 1, 1998.

Peruzzo Trivelloni, C. A. *Método para Determinação do Valor da Localização com Uso de Técnicas Inferenciais e Geoestatísticas na Avaliação em Massa de Imóveis*. Tese de Doutorado, Universidade Federal de Santa Catarina, 2005.