

**O VALOR DE LOCALIZAÇÃO DOS IMÓVEIS:  
DETERMINAÇÃO POR MÉTODOS DE ANÁLISE ESPACIAL**

**Peruzzo Trivelloni, Carlos A.<sup>A</sup>, Hochheim, Norberto<sup>B</sup>**

<sup>A</sup>Eng<sup>o</sup>. Civil, Mestre em Eng. Civil, Doutor em Eng. Civil

Rodovia Admar Gonzaga 2025, bloco I-2, ap. 204. CEP 88034-001, Florianópolis - SC

Fone: (48) 3334 0351 e-mail: [carlos.peruzzo@gmail.com](mailto:carlos.peruzzo@gmail.com)

<sup>B</sup>Eng<sup>o</sup>. Civil, Mestre em Eng. de Produção e Doutor. CREA/SC – 14029-1. IBAPE-SC

Rua Heitor Luz, 44, Apto 44. CEP 88015-500 - Florianópolis - SC

Fone: (48) 3228 2846 Fax: (48) 3331 5167 e-mail: [hochheim@ecv.ufsc.br](mailto:hochheim@ecv.ufsc.br)

**Resumo.** Este trabalho apresenta um método para a determinação do valor da localização dos imóveis usando técnicas de estatística espacial. 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 é difícil de ser modelada adequadamente pelos métodos inferenciais tradicionais, provocando problemas de especificação nos modelos e autocorrelação espacial nos resíduos, comprometendo a confiabilidade da avaliação. Os modelos de estatística espacial, em especial os de regressão espacial e os métodos geoestatísticos (krigagem) são usados neste trabalho em forma combinada para estimar o valor da localização dos imóveis. Uma aplicação do método é realizada para uma amostra de mercado de dois bairros do Município de São José, Santa Catarina, Brasil. O método proposto permitiu encontrar um índice do valor da localização fortemente significativo no modelo de regressão para todos os tipos de imóveis da região e que modela de forma consistente o efeito do conjunto de pólos de valorização na região. O método elimina toda autocorrelação espacial nos resíduos do modelo de regressão melhorando o poder explicativo e a confiabilidade da avaliação.

**Palavras chave:** valor de localização, estatística espacial, regressão espacial, krigagem.

**Abstract.** This work presents a method for determination and estimation of location value on real estate data using spatial statistics techniques. 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 statistics models, especially spatial regression models and geostatistical models are used combined for estimation of location real estate 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 two districts of the municipality of São José, Estate of Santa Catarina, Brazil. The proposed method allowed estimating an index of location value for all real estate properties which seemed highly significant for all real estate types and modeled with consistency and coherency the effects of all valuation factors in the study area. The spatial structure of the location index eliminated all forms of spatial autocorrelation in the model, augmenting its explanation capacity and the reliability of the real estate valuation.

**Key words:** location value, spatial statistics, spatial regression, kriging.

## 1. INTRODUÇÃO

### 1.1. Contextualização

Entre todas as variáveis que influenciam no valor dos imóveis, as variáveis ou fatores referidos à localização são as mais complexas de analisar e modelar.

Os fatores de localização que participam da valorização de um imóvel são inúmeros: características gerais da cidade e do bairro, proximidade a centros comerciais, culturais, centros de educação, de saúde, fatores de segurança pública, características socioeconômicas da vizinhança, proximidade a fatores ambientais desejáveis ou indesejáveis, entre outros. A

lista completa de fatores que potencialmente e efetivamente interagem para valorizar uma localização e os imóveis ali construídos pode ser extremamente ampla.

A dinâmica urbana produz uma contínua mudança nos efeitos destes fatores sobre o valor dos imóveis: investimentos públicos, empreendimentos econômicos e sociais, atividades comerciais que abrem ou fecham, empreendimentos imobiliários, loteamentos, todos estes fatores produzem mudanças nas características de uma vizinhança e criam novos pólos de valorização imobiliária. Quase toda mudança na fisionomia urbana afeta direta ou indiretamente o valor dos imóveis localizados na sua área de influência, podendo transformar o valor de localização no fator mais dinâmico do valor dos imóveis.

## **1.2. Acessibilidade, vizinhança, autocorrelação espacial**

Os modelos de avaliação imobiliária tradicionalmente utilizados, baseados na inferência estatística, apresentam algumas dificuldades para lidar com esta complexidade de fatores.

Estes modelos procuram explicar o valor dos imóveis a partir da influência de variáveis explicativas, sendo estas geralmente classificadas como variáveis construtivas ou próprias do imóvel e variáveis de localização, também classificadas como de acessibilidade e vizinhança, que podem ser comuns a um conjunto de imóveis próximos entre si.

As variáveis de acessibilidade são geralmente consideradas como uma função da distância do imóvel a um determinado pólo de valorização ou desvalorização. Variáveis deste tipo podem não se mostrar estatisticamente significativas num modelo de regressão por vários motivos, desde alguns relacionados com a composição da amostra, até o grau de generalidade ou área de influência real que um pólo valorizador possa ter em relação ao conjunto de imóveis de uma região.

As variáveis de qualidade da vizinhança tentam medir os efeitos de localização desde um ponto de vista local, definindo áreas de influência onde a qualidade da vizinhança pode ser considerada semelhante para todos os imóveis. Estas áreas são também chamadas zonas homogêneas ou de valorização homogênea. As variáveis de vizinhança também podem apresentar problemas de significância estatística nos modelos de regressão relacionados com a definição e delimitação destas zonas.

Quando as variáveis de localização não são corretamente especificadas podem gerar outro tipo de problemas nos modelos de regressão. Além da perda de poder de explicação do modelo, o tratamento incompleto ou inadequado dos fatores de localização pode provocar a autocorrelação espacial dos resíduos do modelo de regressão. A presença de autocorrelação espacial nos resíduos invalida uma das hipóteses básicas da inferência estatística e questiona a validade dos testes de hipótese da regressão tornando ineficientes as estimativas e os resultados do modelo.

## **1.3. Modelagem da localização nos modelos hedônicos**

Na estimação do valor dos imóveis por modelos hedônicos ou modelos de inferência estatística as variáveis de localização são classificadas geralmente em duas categorias: variáveis de distância (para características de acessibilidade) e variáveis de zonas homogêneas (para as características qualitativas de vizinhança).

As variáveis de tipo distância são definidas para os possíveis centros de valorização de uma região. A principal dificuldade de usar este tipo de variáveis reside em identificar a área de influência real de cada pólo bem como a forma funcional de dependência do valor com a distância. A partir de uma certa distância a um pólo o efeito deste pode ser nulo ou

imperceptível. Definir a priori esta distância máxima é impossível sem uma análise estatística dos dados de mercado. Outra dificuldade está em definir e modelar a forma de variação funcional do valor com a distância. As mais utilizadas são a diminuição linear com a distância, a diminuição logarítmica, a inversa da distância e a inversa do quadrado da distância. A escolha entre estas formas requer de uma análise concreta dos dados de mercado, porém muitas vezes a insuficiência de informações não permite analisar eficientemente esta relação e uma forma funcional específica é adotada *a priori* pelo avaliador.

Estas dificuldades podem introduzir erros na especificação dos modelos de regressão levando à falta de significância estatística de variáveis de localização importantes. Outra limitante do uso de variáveis distância é a multicolinearidade de variáveis independentes, pois geralmente uma região tende a concentrar a presença de vários pólos de atratividade próximos entre si, conformando-se centros e subcentros de valorização; as variáveis distância podem apresentar forte correlação impedindo sua utilização como variáveis independentes e diminuindo o poder explicativo do modelo.

As variáveis de tipo zona homogênea definem-se para regiões teoricamente homogêneas em relação às características qualitativas de localização. Elas podem ser definidas em base a informações sociais, econômicas, construtivas, ambientais, e existem vários métodos propostos para identificar estas regiões com o uso de Sistemas de Informações Geográficas (Silva et al., 2004). As limitações que apresentam este tipo de variáveis referem à definição objetiva dos limites destas zonas, geralmente baseadas em setores definidos para outros objetivos; a validade da hipótese de homogeneidade ao interior destas regiões representa uma simplificação que pode ser excessiva, levando à falta de significância no modelo. Estudos realizados utilizando variáveis de zona homogênea em modelos hedônicos com amostras de grande tamanho têm mostrado resultados positivos, trabalhando com um grande número de setores homogêneos como variáveis dicotômicas. O uso de variáveis dicotômicas permite avaliar a média de valorização de cada região. A principal dificuldade para poder usar este tipo de modelagem está no tamanho da amostra necessária para obter um resultado adequado na inferência. Os trabalhos conhecidos utilizam amostras contendo milhares de informações de mercado, tamanho este disponível em cidades dos Estados Unidos de América ou Canadá onde este método é utilizado, mas de difícil aquisição em outras partes como no Brasil.

Também são utilizadas variáveis socioeconômicas como variáveis *proxy* do valor da localização: renda média das famílias ou *per capita*, nível de escolaridade, entre outras, são definidas para regiões consideradas homogêneas; este tipo de variáveis apresenta as mesmas limitações que as variáveis de zonas homogêneas: definição do suporte espacial e homogeneidade real.

Semelhantes às zonas homogêneas são usadas também variáveis de tipo zona fiscal, geralmente extraídas dos cadastros fiscais existentes, zonas com valor venal teoricamente homogêneo. Estas variáveis têm os erros e desatualização provenientes dos cadastros considerados, além da definição arbitrária de limites para estas zonas fiscais.

Outras variáveis utilizadas para representar a vizinhança são construídas a partir de Análise Fatorial, Análise de Componentes Principais ou Análise de Correspondências Múltiplas, representando uma série de características construtivas e socioeconômicas da vizinhança. Estas variáveis apresentam o mesmo problema da definição e delimitação do suporte geográfico além da disponibilidade real de informação completa e abrangente. A vantagem do uso de este tipo de variáveis é a redução da dimensionalidade das características de localização, redução importante para amostras de tamanho reduzido que serão analisadas por inferência estatística.

Visando contornar as dificuldades expostas acima, este trabalho apresenta um método para a determinação do valor da localização dos imóveis usando técnicas de estatística espacial. Modelos de estatística espacial, em especial os de regressão espacial e os métodos geoestatísticos (krigagem) serão usados de forma combinada para estimar o valor da localização dos imóveis.

## **2. FUNDAMENTAÇÃO TEÓRICA**

### **2.1. A importância da localização no valor de diferentes tipos de imóveis**

Segundo Chica Olmo (1994) os fatores de localização afetam no mesmo sentido todos os imóveis de uma região, sejam terrenos ou construções, porém o grau de influência para cada tipo de imóvel pode ser diferente.

Gloude-mans (2002) estuda a aplicação de modelos gerais de avaliação para imóveis construídos e terrenos baldios por inferência estatística de três cidades, concluindo que as variáveis de localização afetam todos os tipos de imóveis de forma semelhante sendo que a diferença relativa a cada tipo de imóvel pode ser modelada por inferência estatística utilizando-se fatores adequados para sua estimação.

No mesmo sentido Peruzzo Trivelloni (2005) conclui que os fatores de localização podem ser analisados para vários tipos de imóveis no mesmo modelo de avaliação desde que sejam considerados fatores de ajuste por tipo de imóvel que podem ser determinados por inferência estatística a partir dos dados de mercado.

### **2.2. Métodos de análise espacial**

A dependência espacial existente no valor dos imóveis, produzida principalmente pelos fatores de localização, tem mostrado que as técnicas tradicionais de inferência estatística podem levar a resultados ineficientes na presença de autocorrelação espacial dos resíduos dos modelos, levando a procurar e pesquisar o uso de técnicas de estatística espacial para o tratamento de dados do mercado imobiliário.

Na literatura de avaliação imobiliária existem duas grandes escolas para o tratamento espacial dos dados de mercado: alguns autores têm proposto o uso das técnicas de econometria espacial e seus modelos de regressão espacial, enquanto outros têm proposto o uso de técnicas geoestatísticas. A discussão sobre qual das duas escolas tem fundamentos teóricos e metodológicos mais adequados para sua aplicação na área da Engenharia de Avaliações tem levado a analisar com maior profundidade o tipo de dados espaciais que usam ambas técnicas de estatística espacial.

Anselin (2002) sugere fundamentar teoricamente a escolha de um ou outro método a partir da taxonomia proposta por Cressie (1993) para a análise espacial que distingue entre fenômenos de natureza pontual, modelos geoestatísticos ou de superfície para fenômenos contínuos e modelos de áreas para fenômenos discretos. Anselin (2002) incorpora na discussão os conceitos relacionados com a modelagem de dados espaciais e suas representações computacionais mais comumente utilizadas nos Sistemas de Informações Geográficas (SIG). Os modelos de dados distinguem dois tipos de dados: geobjetos e geocampos. Os geobjetos são entidades discretas no espaço geralmente representadas nos SIG como pontos, linhas ou polígonos enquanto os geocampos pertencem a uma distribuição espacial contínua representada geralmente como superfície. Na análise econômica os

geobjetos correspondem a agentes econômicos com localização bem definida tais como endereço, CEP, bairro e município. Os geocampos, por sua vez, podem ser representados como superfícies de preços ou de risco, por exemplo do valor da terra, qualidade do ar, entre outros.

Ainda segundo o mesmo autor, os tipos de dados que podem ser relacionados ao conceito de geobjetos espaciais são adequados aos tratamentos estatísticos por econometria espacial, enquanto que os geocampos deveriam ser analisados com modelos geoestatísticos ou de superfície. No caso da análise do mercado imobiliário o autor comenta que os dados podem apresentar simultaneamente comportamentos identificáveis como objetos ou como campos, abrindo a possibilidade de análise espacial por ambas metodologias.

Esta aparente ambivalência tem sua causa nos diversos fatores que influenciam no valor dos imóveis. Enquanto alguns fatores, como os construtivos, variam de imóvel para imóvel independentemente da sua localização espacial, dando aos dados a característica de objetos espaciais, outros fatores, os de localização, influenciam imóveis próximos de forma semelhante, produzindo uma variação do valor de forma contínua no espaço, dando ao valor dos imóveis o comportamento de um geocampo. Desta maneira diferentes fatores influenciadores dão aos imóveis e seus valores de mercado o comportamento ao mesmo tempo de objetos e de campos espaciais.

Dois grandes escolas metodológicas de estatística espacial têm sido propostas: a que propõe usar métodos de econometria espacial e a que propõe usar métodos geoestatísticos, tentando modelar estes efeitos.

Ambas metodologias, a econometria espacial e a geoestatística, têm sido aplicadas na análise de dados imobiliários apresentando vantagens e desvantagens, potencialidades e limitações.

O uso combinado de ambas metodologias (Peruzzo Trivelloni, 2005) permite superar algumas das deficiências e limitações destes métodos, potencializando ao mesmo tempo as suas vantagens e alcançando resultados satisfatórios.

A justificativa de usar as duas técnicas procura o tratamento mais adequado para os dois tipos de fatores influenciadores no valor dos imóveis: a análise por econometria espacial das características físicas e construtivas dos imóveis e a análise por geoestatística dos fatores de localização.

### **2.3. Econometria espacial**

A econometria espacial é a parte da econometria que lida com o tratamento das interações e estrutura espaciais de dados geograficamente distribuídos.

O foco na localização espacial dos dados e as conseqüências teóricas e metodológicas das suas interações espaciais têm obtido crescente atenção nos últimos anos, tanto do ponto de vista das aplicações técnicas como no embasamento teórico.

Segundo Anselin (1999a), o crescente interesse nos últimos anos na especificação e estimação de interações espaciais nos modelos econométricos tem sido provocado por diversos fatores.

Em primeiro lugar pela discussão teórica da existência de interação entre os agentes econômicos atuantes num determinado modelo. Estes agentes não devem ser considerados apenas como agentes atomizados, atuando isoladamente uns dos outros senão pelas suas interações, gerando efeitos de vizinhança a diferentes escalas de análise.

Em segundo lugar, a demanda de metodologias para o tratamento de fenômenos espaciais aumentou com a crescente disponibilidade de bases de dados georreferenciados e

utilização dos Sistemas de Informações Geográficas (SIG), que disponibilizam grande quantidade de informação onde podem ser explorados e modelados padrões espaciais dos dados. O uso conjunto dos SIG com análises estatísticas espaciais tem gerado um crescente interesse pela sua possível aplicação em várias áreas de conhecimento.

Finalmente, o reconhecimento de que as técnicas econométricas tradicionais falham na presença de autocorrelação espacial nos dados tem provocado crescente interesse no desenvolvimento teórico e metodológico da econometria espacial.

No contexto dos modelos de regressão os efeitos espaciais podem ser catalogados como efeitos de dependência espacial ou autocorrelação espacial e heterogeneidade espacial.

A autocorrelação espacial pode ser definida como a coincidência de valores similares em localizações similares.

A heterogeneidade espacial consiste na instabilidade estrutural na forma de variâncias não constantes dos erros (heterocedasticidade) ou dos coeficientes do modelo (coeficientes variáveis no espaço). Esta heterogeneidade deve ser considerada explicitamente desde que a estrutura por trás dessa instabilidade é espacial, ou seja, a localização das observações é determinante dessa instabilidade. Além disso, a heterogeneidade geralmente ocorre junto com autocorrelação espacial e as ferramentas econométricas tradicionais não são mais apropriadas neste caso. Em amostragens transversais (sem variação no tempo) os fenômenos de heterogeneidade espacial e de autocorrelação espacial podem ser considerados equivalentes.

Nas amostragens transversais não há suficiente informação para estimar a matriz de covariâncias diretamente dos dados.

Por isso em geral é necessário impor uma estrutura para essa covariância. Essa imposição se traduz na criação de um conceito de operador espacial de defasagem, que é uma média ponderada de uma variável aleatória em localizações vizinhas. Para isso deve ser definido o conceito de vizinhança de cada observação. O operador espacial de defasagem opera por meio de uma matriz de pesos espaciais  $W$  que define a estrutura de covariância dos dados para uma determinada vizinhança de cada observação. Por meio de pesos que representam a covariância entre dois elementos, o operador espacial calcula para cada observação a média ponderada da variável aleatória na vizinhança.

Para cada observação da variável aleatória “ $y$ ” num ponto “ $i$ ” o operador espacial  $Wy$  é obtido pela soma ponderada pelos pesos  $w_{ij}$  de todos os valores de “ $y$ ” nos pontos “ $j$ ” da vizinhança:

$$(Wy)_i = \sum (w_{ij} \cdot y_j)$$

os pesos  $w_{ij}$  não nulos são os que definem a vizinhança de cada observação.

Por facilidade de interpretação os pesos são padronizados por linha na matriz  $W$  de forma que a soma de pesos em cada linha é igual a 1. Conseqüentemente a defasagem espacial da variável “ $y$ ” no ponto “ $i$ ” pode ser interpretada como sua média ponderada na vizinhança de “ $i$ ”.

Os elementos que compõem a matriz  $W$  são não estocásticos e são externos ao modelo. Eles estão geralmente baseados nos conceitos de contigüidade dos dados ou na distância entre observações, baseados numa função decrescente com a distância entre observações. Esta estrutura de variação é definida pelo pesquisador diferentemente de outros métodos como os de estimação direta da covariância que calculam a covariância a partir dos próprios dados. Esta falta de obrigatoriedade de uma justificativa teórica na seleção dos pesos espaciais e o fato de serem exógenos ao modelo e sem ambigüidades explica a popularidade dos pesos derivados geograficamente, segundo Anselin (1999a).

## 2.4. Modelos de Regressão Espacial

A dependência espacial pode ser incorporada nos modelos clássicos de regressão de duas formas: como um regressor adicional na forma de uma variável dependente espacialmente defasada ( $Wy$ ), ou uma estrutura espacialmente defasada no erro da regressão ( $We$ ). O primeiro modelo é conhecido como modelo de Defasagem Espacial ou da variável dependente defasada e o segundo é o modelo do Erro Espacial ou do erro espacialmente correlacionado.

### 2.4.1. Modelo de Defasagem Espacial

Formalmente este modelo pode ser expresso da seguinte forma:

$$Y = \rho Wy + XB + \varepsilon$$

onde  $y$  é o vetor da variável dependente,  $\rho$  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  $\varepsilon$  é o vetor de resíduos do modelo.

### 2.4.2. Modelo do Erro Espacial

O modelo autoregressivo do erro ou modelo do erro espacial pode ser expressado formalmente da seguinte forma:

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

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

O modelo pode ser escrito também da seguinte forma:

$$y = XB + \lambda W\varepsilon + u$$

Segundo Anselin (1999a) os modelos de regressão espacial podem ser estimados utilizando vários métodos. O mais comumente usado é o de máxima verossimilhança, mas outros métodos também têm sido propostos, como os de variáveis instrumentais, mínimos quadrados espaciais, método dos momentos, método dos códigos, e métodos bayesianos entre outros.

Os primeiros estudos sobre a existência de autocorrelação espacial nos dados do mercado imobiliário foram propostos por Dubin (1988) e o primeiro trabalho de aplicação de modelos de regressão espacial foi proposto por Can (1990 e 1992).

A partir da discussão teórica sustentada por autores como Anselin e Dubin entre outros, novos estudos utilizando os modelos de regressão espacial foram desenvolvidos na área de avaliação imobiliária, como os de Pace e Gilley (1997), Pace et al. (1998), Dantas et al (2001), Besner (2002), Dantas (2003), entre outros.

Em geral todos estes trabalhos comprovam a existência de autocorrelação espacial nos resíduos dos modelos de mínimos quadrados e testam diferentes modelos de regressão espacial obtendo uma melhora importante nos resultados dos modelos e principalmente na adequação teórica e metodológica ao tratamento dos efeitos espaciais nos dados.



Porém, alguns elementos continuam em aberto na utilização dos modelos de regressão espacial: por um lado, a especificação da estrutura espacial dos dados e definição da matriz de pesos espaciais  $W$  depende da definição por parte do pesquisador, e a escolha desta matriz influencia fortemente os resultados alcançados. Por outro, os termos que incluem os efeitos espaciais na regressão, tanto o termo da variável dependente defasada como o termo do erro correlacionado, não apresentam uma interpretação teórica conclusiva como variáveis explicativas ou fatores de valorização dos imóveis, o que pode ser considerado uma dificuldade metodológica na hora de analisar e escolher o método de regressão mais apropriado. Desta forma, a escolha do modelo de regressão espacial tem sido baseada na comparação dos resultados estatísticos nos testes de significância de cada modelo, ao invés de critérios de escolha metodológicos.

## **2.5. Geoestatística. Teoria das Variáveis Regionalizadas**

Segundo Cardim (2000) a estatística clássica trabalha com a suposição de independência entre as amostras para representar um fenômeno, baseando-se na hipótese de que as variações de um local a outro são aleatórias. No entanto, quando uma determinada propriedade varia de um local para outro com uma certa continuidade, expressa através da dependência espacial entre as amostras, outra técnica, relativamente nova, a geoestatística, deve ser empregada para a realização das análises.

Para determinar qual das duas técnicas (estatística clássica ou geoestatística) deve-se utilizar, faz-se uso do semivariograma, que expressa a relação espacial entre as amostras. Havendo dependência espacial entra em ação a geoestatística, por outro lado, se não for detectada a dependência espacial, se utiliza a estatística clássica.

Verificada a dependência espacial podem-se estimar valores do fenômeno em estudo para os locais não amostrados, dentro da área em estudo, sem tendência e com variância mínima, por meio do método conhecido como interpolação por krigagem. A krigagem possibilita a construção de mapas de isolinhas, ficando definidas as classes dentro da área de estudo.

A krigagem compreende um conjunto de técnicas de estimação e predição de superfícies baseadas na modelagem da estrutura de correlação espacial. A hipótese implícita é que o processo estudado é estacionário.

Os passos num estudo empregando técnicas de krigagem incluem:

- a) Análise exploratória dos dados;
- b) Análise estrutural (modelagem da correlação espacial);
- c) Interpolação estatística da superfície.

O que diferencia a krigagem de outros métodos de interpolação é a estimação de uma matriz de covariância espacial que determina os pesos atribuídos aos diferentes dados, o tratamento da redundância destes, a vizinhança a ser considerada no procedimento inferencial e o erro associado ao valor estimado. A krigagem fornece estimadores não tendenciosos e eficientes.

Segundo Chica Olmo (1994) quando existe uma dependência entre uma variável aleatória e a sua localização espacial a variável pode ser chamada de regionalizada. A Teoria das Variáveis Regionalizadas (TVR) dedica-se à análise e resolução dos problemas de inferência espacial de variáveis que dependem da localização espacial onde são observadas. A TVR foi especificada por Matheron entre os anos 1962 e 1965 a partir dos estudos de Krige e outros. Inicialmente foi desenvolvida para ser aplicada na área de geologia e outras ciências

da Terra, sendo conhecida como Geoestatística. Posteriormente, devido ao seu caráter geral, a sua aplicação tem sido estendida para outras áreas do conhecimento.

A estrutura teórica da krigagem está baseada no conceito de variável regionalizada. Uma variável regionalizada é uma variável distribuída no espaço ou tempo cujos valores são considerados como realizações de uma função aleatória.

A variação espacial de uma variável regionalizada pode ser expressa pela soma de três componentes:

- a) uma componente estrutural, associada a um valor médio constante ou a uma tendência constante;
- b) uma componente aleatória, espacialmente correlacionada;
- c) um ruído aleatório ou erro residual.

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;

$\varepsilon''$  é um ruído aleatório não correlacionado, com distribuição normal com média 0 e variância  $\sigma^2$ .

A partir de uma amostra  $Z(x_1), \dots, Z(x_n)$ , deseja-se 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  $\lambda_i$  da seguinte maneira (Chica Olmo, 1994):

$$Z_k = \sum_i \lambda_i Z_i$$

Onde os ponderadores  $\lambda_i$  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.

Existem distintos métodos de krigagem em função das características da variável que deve ser estimada e dos objetivos da estimação. Se as estimativas estão referidas a suporte pontual e deseja-se realizar estimativas pontuais, utiliza-se a krigagem pontual. Se o objetivo é estimar o valor médio da variável sobre um suporte de tamanho maior utiliza-se a krigagem por blocos.

Segundo Chica Olmo (1994) as variáveis socioeconômicas geralmente não vem associadas a um suporte pontual, mas são definidas por unidade de superfície, por município, por bairro, por metro quadrado, etc.

Neste caso utiliza-se a estimação por krigagem por blocos. Quando os dados estão medidos sobre um suporte determinado e a estimação deseja-se realizar sobre um outro suporte de superfície, é necessário regularizar os dados de acordo com a nova superfície de estimação. É o caso, por exemplo, quando os dados são medidos para imóveis individuais e deseja-se obter uma estimação do valor para um conjunto próximo de imóveis, por exemplo, todos os imóveis de uma quadra.

Seja  $V$  a área do bloco sobre a qual deseja-se estimar a variável  $Z_k$ .

O novo valor regularizado, ou valor médio de  $Z(x)$  pode ser obtido da seguinte expressão:

$$Z_v(x) = 1/V \int_V Z(x)dx$$

Os primeiros estudos de aplicação dos métodos geoestatísticos na área de avaliação imobiliária foram apresentados por Dubin (1992). A autora propõe modelar o valor dos imóveis por regressão múltipla e modelar o erro do modelo por krigagem. Mas a aplicação do método geoestatístico é parcial pois não é realizado o estudo da estrutura espacial dos dados. Ao invés de realizar a análise da estrutura espacial por meio do semivariograma empírico, a estrutura espacial para estes resíduos é imposta, escolhendo a forma exponencial negativa e calculando apenas os parâmetros de ajuste por máxima verossimilhança.

Posteriormente e seguindo a mesma linha de pesquisa, Dubin (1998) e Basu e Thibodeau (1998) aplicam a mesma metodologia para novos estudos de valor de imóveis, realizando a krigagem dos resíduos dos modelos de regressão impondo uma determinada estrutura espacial de variação.

O primeiro estudo de aplicação que usa a metodologia geoestatística para a avaliação de imóveis de forma completa é proposto por Chica Olmo (1994). O autor argumenta que o valor dos imóveis pode ser considerado como uma variável regionalizada e portanto as técnicas da Teoria das Variáveis Regionalizadas (TVR) podem ser utilizadas para modelar a estrutura espacial deste valor. Para poder utilizar a TVR o estudo deve ser realizado sobre uma variável que seja homogênea na sua definição e composição, variando apenas em função da sua localização espacial. Dado que o mercado imobiliário consiste de objetos heterogêneos tanto na sua localização quanto nas suas características intrínsecas (características físicas e construtivas do imóvel) o estudo é realizado sobre uma amostra de dados de imóveis homogêneos nas suas características construtivas. Desta forma consegue-se estudar a estrutura espacial dos fenômenos que influenciam no valor dos imóveis, por meio do semivariograma, analisando os efeitos de anisotropia existentes e modelando por krigagem a variação do valor em toda a área de estudo.

Em definitivo, o autor modela a variação do valor de um tipo homogêneo de imóveis em função da sua localização, ou seja, está modelando o valor da localização, a menos de uma constante multiplicativa, mas somente para aquele tipo específico de imóveis homogêneos.

Posteriormente Cano Guervós (1999) tenta aplicar a mesma metodologia para uma amostra heterogênea de imóveis, onde propõe o uso da análise fatorial e da regressão múltipla para estimar o efeito das características construtivas dos imóveis e a modelagem por geoestatística do erro da regressão, realizando uma proposta similar ao estudo de Dubin (1992). Este autor também utiliza um modelo aditivo para o valor, considerando este como a soma de um valor de construção e um valor de localização. Desta forma propõe definir e estimar o valor da localização dos imóveis por meio do termo de krigagem dos resíduos e mais uma constante. Esta constante é calculada separando em duas partes o termo constante obtido na regressão, que é comum aos componentes de construção e localização, seguindo a proposta sugerida por Dubin (1992). Este método de divisão da constante apresenta algumas limitações pela própria definição, obtendo-se valores negativos para algumas localizações. Isto se deve, provavelmente, ao modelo aditivo e à metodologia de cálculo da constante, já que ela é obtida pela diferença entre duas constantes de modelos obtidos por métodos de estimação diferentes: o modelo de mínimos quadrados ordinários (que não se considera confiável por existir forte autocorrelação espacial nos resíduos) com o modelo de máxima verossimilhança. O próprio autor propõe estudar novas formas de poder dividir a constante do modelo por outro método para poder definir o valor da localização.

Gámez Martínez (1999) propõe modelar o valor dos imóveis usando geoestatística para modelar o efeito da localização e propõe um método de homogeneização das características

construtivas dos imóveis por regressão múltipla. Desta forma, são calculados fatores de homogeneização a partir dos elementos de mercado para as diferentes características construtivas dos imóveis. Depois de homogeneizadas as amostras, a metodologia geoestatística é usada para estimar a estrutura espacial do valor, por meio do estudo do semivariograma experimental. A proposta de homogeneizar os valores em função das características construtivas permite trabalhar com uma amostra com maior heterogeneidade mas não leva em consideração que a obtenção destes fatores por meio de uma regressão múltipla tradicional onde os erros estão fortemente autocorrelacionados mantém os problemas de ineficiência e confiabilidade dos testes de significância do modelo de mínimos quadrados, limitando a confiabilidade dos fatores de homogeneização encontrados.

## 2.6. Uso combinado de métodos espaciais

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 considerado como uma variável regionalizada e modelado por meio de técnicas geoestatísticas aplicadas sobre uma amostra de dados de mercado. Os modelos de regressão espacial permitem estimar o efeito das variáveis construtivas e características específicas dos imóveis, possibilitando sua homogeneização em relação a estas variáveis em uma amostra de dados de mercado.

A partir dos dados imobiliários homogeneizados por regressão espacial o valor da localização pode ser analisado e modelado utilizando as técnicas geoestatísticas.

## 3. MÉTODO

O método usado segue a proposta realizada por Peruzzo Trivelloni (2005).

O valor dos imóveis depende de um conjunto de fatores entre os quais se encontra a sua localização, as características físicas do imóvel, o momento da avaliação e as condições jurídicas e legais a que está condicionado o imóvel.

Quando considerada uma região geográfica determinada onde as condições fiscais são fixas e uma amostra de mercado em condições semelhantes de transação e coletada na mesma época, o valor dos imóveis poderá ser considerada 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.

O modelo utilizado é de tipo multiplicativo, permitindo examinar a interação entre os fatores de localização e físicos do imóvel, sendo também um modelo de caráter geral para todos os tipos de imóveis. A expressão do modelo geral 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 que depende do tipo de imóvel e os  $F_{CFi}$  representam fatores correspondentes às características físicas influenciadoras no valor para cada tipo de imóvel.

A variável VL representa o valor de localização. No valor VL estão influenciando simultaneamente todos os pólos de valorização, principais ou secundários, que afetam o valor dos imóveis em uma localização. Este conjunto de influencias atua de forma a gerar uma estrutura espacial do valor VL. Quanto mais próximo se encontra o imóvel dos pólos de valorização, maior será seu valor VL e este valor diminui paulatinamente na medida que aumenta a distancia aos pólos de valorização. A variável VL apresenta, portanto, características de dependência e autocorrelação espacial: dos imóveis próximos receberão efeito similar dos pólos de valorização e esta semelhança será menor quanto maior seja a distancia entre os imóveis.

O valor VL inclui também a influência da vizinhança imediata no valor de cada imóvel. As características de uma vizinhança como a estrutura urbana e os serviços públicos, a qualidade do ambiente natural e construído, as características socioeconômicas e culturais, entre outras, influenciam de forma geral a todos os imóveis nela localizados provocando uma determinada valorização média para todos os imóveis de uma região e estão presentes no valor VL. O efeito de adjacência que produz uma influência direta do valor de cada imóvel sobre os imóveis mais próximos gera um efeito de vizinhança que também está presente no valor VL com uma estrutura de autocorrelação espacial.

Os efeitos induzidos por vários pólos de valorização próximos entre si interagem de forma complexa afetando o valor VL.

Por estes motivos VL pode ser caracterizada como uma variável aleatória que apresenta uma estrutura espacial, ou seja, é uma variável regionalizada.

O modelo proposto permite compreender o comportamento do valor dos imóveis simultaneamente como objeto e como geocampo. De acordo com o modelo, VU depende da variável VL e de um conjunto de variáveis intrínsecas dependentes do tipo de imóvel.

Por ser VL uma variável regionalizada a modelagem mais apropriada do valor de localização é realizada usando métodos geoestatísticos. Para estimar VL é necessário homogeneizar os dados de mercado retirando o efeito no seu valor de todos os outros fatores ou variáveis.

Por se tratarem os imóveis de objetos espaciais a metodologia estatística aplicável para realizar esta homogeneização pertence a econometria espacial.

Para realizar esta homogeneização serão estimados os fatores utilizando o método de regressão espacial do erro. Este método é aplicado considerando como variáveis independentes apenas aquelas referidas às características tipológicas e construtivas dos imóveis:

$$L(VU) = L(F_{TIPO}) + L(F_{CF1}) + \dots + L(F_{CFn}) + e$$

onde:

$$e = \lambda We + u$$

ou seja, o erro encontra-se espacialmente correlacionado. Por não estarem incluídas no modelo as variáveis específicas de localização, os efeitos autocorrelacionados da localização encontram-se absorvidos pelo termo do erro da regressão espacial.

A utilização do modelo espacial do erro ao invés do modelo de regressão tradicional (não espacial) permite estimar os parâmetros da regressão de forma eficiente e confiável.

Estimados assim os parâmetros das variáveis construtivas e de tipo, os dados de mercado podem ser homogeneizados. Dividindo o valor VU pelos fatores correspondentes é obtido um valor homogeneizado VH que não depende do tipo de imóvel nem das suas

características construtivas mas apenas da localização do imóvel, sendo uma nova variável regionalizada que pode ser modelada por métodos geoestatísticos.

O método geoestatístico utilizado é a krigagem por blocos, que permite diminuir os efeitos de possíveis erros contidos no valor homogeneizado VH. A fonte destes possíveis erros pode estar nos próprios dados de mercado ou no processo de homogeneização definido se a especificação das variáveis independentes do modelo espacial do erro não for adequada.

A krigagem por blocos, ao tomar médias dos valores de VH para cada bloco está considerando o valor médio de cada localização sendo uma estimativa mais robusta que a krigagem pontual.

O resultado da krigagem é uma nova variável VL, definida para cada ponto do espaço, que representa o efeito de todos os fatores de localização sobre o valor dos imóveis.

## **4. APLICAÇÃO**

### **4.1. Caracterização da área e amostragem**

A área de estudo corresponde aos bairros Kobrasol e Campinas do município de São José, Estado de Santa Catarina (sul do Brasil).

Segundo Farias (2001) o município de São José, fundado por açorianos em 1750 está localizado na Grande Florianópolis, nas coordenadas geográficas 27°36'52" de latitude sul e 48°38'19" de longitude oeste, possuindo uma área geográfica de 116 km<sup>2</sup> e população em 1996 de 151.024 habitantes. Em 1998 São José se destacou como o sexto maior colégio eleitoral do estado com 94.480 eleitores. O município tem como limite a leste a Baía Sul da Ilha de Santa Catarina, a oeste São Pedro de Alcântara e Antônio Carlos, ao norte Biguaçu e Florianópolis e ao sul Palhoça. São José é um dos grandes pólos de desenvolvimento econômico de Santa Catarina. É o município de maior densidade demográfica de Santa Catarina, com 1.317 hab./km<sup>2</sup>. O Estado de Santa Catarina tem média de 50 hab./km<sup>2</sup> e Florianópolis 593 hab./km<sup>2</sup>.

Os bairros Kobrasol e Campinas, pertencentes ao Distrito Campinas de São José, constituem um dos centros mais dinâmicos do município, com forte movimento imobiliário, atividades comerciais variadas, centros de recreação, inúmeros restaurantes, centros educacionais do pré-escolar ao universitário e um número significativo de bancos, além da instalação de órgãos municipais importantes.

Em relação às características da construção imobiliária nestes bairros, até 1980 as construções limitavam-se praticamente a dois pavimentos sendo casas residenciais. A partir de 1980 começaram a ser construídos edifícios de 4 andares, posteriormente entre 1980 e 1990 começaram a ser construídos edifícios de 12 andares e atualmente é permitido até 14 pavimentos segundo o Plano Diretor do município.

Entre as obras públicas de recente execução, a construção da avenida Beira Mar de São José inaugurada no 2004 valorizou mais a região de Kobrasol e Campinas, especialmente a área mais próxima desta obra, provocando uma mudança nos valores imobiliários da região.

Foi realizada uma pesquisa do mercado imobiliário na área de estudo para levantamento dos dados do presente trabalho. A pesquisa incluiu a busca de imóveis a venda e foi realizada junto às imobiliárias que atuam na região, procurando anúncios de venda de imóveis publicados nos jornais, e diretamente em toda a área de estudo.

Foram incluídos na pesquisa todos os tipos de imóveis que estavam sendo comercializados na região, abarcando os seguintes tipos: terrenos, apartamentos, kitnetes, casas, salas e lojas comerciais. No total foram coletadas informações correspondentes a 249

imóveis a venda na região, correspondendo a 175 apartamentos, 21 kitinetes, 27 casas, 17 terrenos e 9 imóveis comerciais.

As características levantadas para todos os imóveis incluídos na amostra de mercado foram as seguintes:

- Tipo de imóvel. Cada imóvel foi classificado nos cinco tipos anteriormente mencionados: apartamento, kitinete, casa, terreno ou comercial. O tipo comercial inclui as salas e lojas comerciais.
- Localização do imóvel. A pesquisa incluiu a localização exata dos imóveis. Com as informações constantes no cadastro da Prefeitura Municipal foi definido para cada imóvel um geocódigo que define a localização do lote, sendo localizado cada imóvel da pesquisa de mercado no mapa do Município e registradas as coordenadas UTM de cada imóvel.

As outras variáveis levantadas na pesquisa para os apartamentos, kitinetes e imóveis comerciais foram as seguintes:

- A área total e privativa do imóvel medida em metros quadrados;
- O padrão construtivo do imóvel, medido de acordo com a classificação existente no atual cadastro da prefeitura, que classifica os padrões em quatro categorias: padrão luxo (padrão 1), alto (padrão 2), normal (padrão 3) ou baixo (padrão 4). O padrão construtivo foi levantado por vistoria externa dos imóveis e comparado com o padrão construtivo definido no cadastro imobiliário da Prefeitura. Como foi detectada uma desatualização nos dados do cadastro imobiliário foi mantido como valor da variável apenas o levantado na pesquisa;
- A existência ou não de garagem na unidade;
- A idade do imóvel, medida em anos, que foi obtida a partir dos dados constantes no cadastro da Prefeitura para os imóveis cadastrados e por meio de informações dos vendedores no caso dos imóveis recém construídos.
- Para as casas a pesquisa também incluiu a área construída e a área total do terreno, o padrão construtivo, e a idade.
- Para os terrenos a pesquisa incluiu a área total e as medidas da frente principal ou frentes do terreno e profundidade.

A localização dos dados da amostra na área de estudo é mostrada na Figura 1. Os imóveis da pesquisa de mercado encontram-se espalhados por toda a região, cobrindo a maior parte dela.

O fato dos imóveis da amostra estarem distribuídos em quase toda a área de estudo permite esperar um grau de confiabilidade importante na estimação do valor de localização em toda a região, pois existem imóveis próximos a praticamente todos os pólos de valorização.



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

#### 4.2. Tratamento dos dados

- Georreferenciamento dos imóveis identificando as coordenadas UTM para cada imóvel.
- Construção de variáveis dicotômicas por tipo de imóvel. A partir da variável TIPO foram construídas as seguintes variáveis dicotômicas: APTO, TERRENO, CASA, KIT, COMERCIA.
- Também as variáveis referentes ao padrão construtivo foram transformadas em variáveis dicotômicas. A partir da variável PADRAO da pesquisa foram construídas as variáveis dicotômicas denominadas Padrão 2, Padrão 3 e Padrão 4 para os imóveis de padrão alto, médio e baixo respectivamente (não há imóveis de padrão luxo na região).
- A variável idade mostrou baixa correlação com o valor unitário dos imóveis. Para medir o efeito da idade no valor unitário a variável IDADE foi subdividida por faixas de valor, definindo a primeira faixa com imóveis de até três anos de idade, a segunda faixa com



imóveis de até seis anos, a terceira com imóveis de até 10 anos, a quarta com imóveis de até 15 anos e a quinta com imóveis de mais de 15 anos. A análise mostrou que apenas os imóveis novos, com até 3 anos de idade, apresentam um efeito significativo de valorização. Por este motivo foi criada a variável dicotômica NOVO que vale 1 para os imóveis com menos de três anos de idade e 0 nos outros casos.

- A variável de área utilizada foi a área total construída para o caso de apartamentos, kitinetes, casas, comerciais e a área total dos terrenos.
- Como variável dependente para os modelos foi definida o Valor Unitário medido em reais por metro quadrado, definido pelo quociente entre o preço total à vista e a área total do imóvel.
- Interação das variáveis construtivas por tipo de imóvel. Como a amostra analisada apresenta 5 tipos de imóveis, e as variáveis construtivas podem apresentar um grau de influência diferente no valor segundo o tipo de imóvel, foi testada essa possível diferença na influência de cada variável para cada tipo de imóvel. Para isto, foi calculado o produto de cada variável construtiva pela variável dicotômica correspondente ao tipo de imóvel. Foram assim construídas novas variáveis a serem testadas como variáveis de homogeneização.

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 <sup>2</sup> , 0 se não é apto.
KIT_AT	contínua	Área total da kit em m <sup>2</sup> , 0 se não é kit
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, 0 em outro caso.
AP_P3	dicotômica	1 para aptos de padrão 3, 0 em outro caso.
KIT_P2	dicotômica	1 para kit de padrão 2, 0 em outro caso.
CASA_P2	dicotômica	1 para casas de padrão 2, 0 em outro caso.
CASA_P3	dicotômica	1 para casas de padrão 3, 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.

### 4.3. Aplicação do método

#### 4.3.1. Estimação da matriz de vizinhança

Para realizar a estimação da matriz de vizinhança são calculados fatores de homogeneização do valor unitário dos imóveis por mínimos quadrados.

A equação de regressão, considerando o logaritmo do valor unitário como variável dependente do modelo, é a seguinte:

$$L(VU) = 6.47401 + 0.00096*APT\_AT + 0.08687*APT\_GAR + 0.17361*AP\_P2 + 0.09797*AP\_P3 + 1.66178*KIT - 0.03908*KIT\_AT + 0.23986*KIT\_P2 + 0.46961*CASA - 0.00340*CASA\_AC + 0.00104*CA\_ATER + 0.53497*CASA\_P2 + 0.47212*CASA\_P3 - 0.24305*TERRENO + 0.00045*TER\_AT + 0.61893*COMERCIA + 0.12275*NOVO$$

Calculado o antilogaritmo, a equação de regressão também pode ser escrita da seguinte forma:

$$VU = 648.0767 * 1.0010^{APT\_AT} * 1.0908^{APT\_GAR} * 1.1896^{AP\_P2} * 1.1029^{AP\_P3} * 5.2687^{KIT} * 0.9617^{KIT\_AT} * 1.2711^{KIT\_P2} * 1.5994^{CASA} * 0.9966^{CASA\_AC} * 1.0010^{CA\_ATER} * 1.7074^{CASA\_P2} * 1.6034^{CASA\_P3} * 0.7842^{TERRENO} * 1.0005^{TER\_AT} * 1.8569^{COMERCIA} * 1.1306^{NOVO}$$

Uma vez obtidos os parâmetros de homogeneização é calculado para cada imóvel o valor unitário homogeneizado VH da seguinte forma:

$$VH = VU / [1.0010^{APT\_AT} * 1.0908^{APT\_GAR} * 1.1896^{AP\_P2} * 1.1029^{AP\_P3} * 5.2687^{KIT} * 0.9617^{KIT\_AT} * 1.2711^{KIT\_P2} * 1.5994^{CASA} * 0.9966^{CASA\_AC} * 1.0010^{CA\_ATER} * 1.7074^{CASA\_P2} * 1.6034^{CASA\_P3} * 0.7842^{TERRENO} * 1.0005^{TER\_AT} * 1.8569^{COMERCIA} * 1.1306^{NOVO}]$$

Com o valor unitário homogeneizado por mínimos quadrados foi calculado o semivariograma experimental para analisar a variação do valor em função da distância entre as observações. O semivariograma experimental isotrópico do valor unitário homogeneizado pode ser observado na Figura 2.

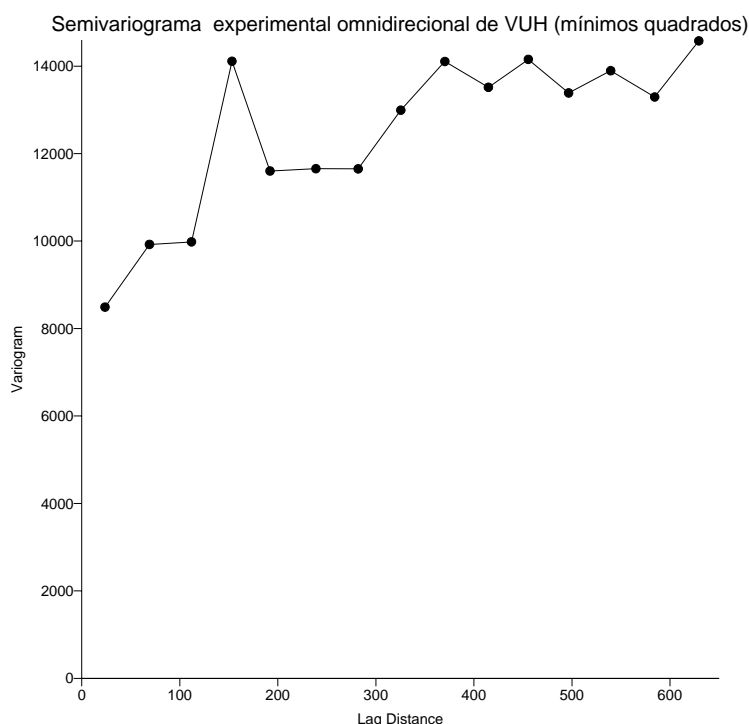


Figura 2. Semivariograma experimental isotrópico do valor unitário homogeneizado por mínimos quadrados.

A partir da análise do semivariograma experimental isotrópico da Figura 2 observa-se que os dados apresentam maior dependência espacial até uma distância um pouco superior aos 100 metros, e depois desta distância a dependência diminui rapidamente apresentando uma estrutura praticamente estacionária. Desta forma, as observações que se encontram muito próximas exercem uma forte influência entre si e esta influência diminui rapidamente com a distância, de forma que aquelas que se encontram separadas por distâncias maiores quase não apresentam influência entre si.

A matriz de pesos espaciais a ser estimada é uma matriz de adjacência, ou seja, uma matriz onde é definida uma distância máxima para dois imóveis serem considerados adjacentes ou vizinhos. Inicialmente os vizinhos têm peso igual a 1 e posteriormente estes pesos são padronizados por linha da matriz de forma que a soma dos pesos para cada imóvel seja igual a 1.

A estimação desta distância de vizinhança geralmente é determinada *a priori* nos modelos de regressão espacial, de forma externa aos dados. Neste caso a estimação foi realizada apoiada na análise da variância dos dados experimentais que oferece o semivariograma. Como a matriz de pesos espaciais não permite modelagem da anisotropia foi utilizado o semivariograma isotrópico.

Desta forma, tomando como base a análise da Figura 2, foi estimada a distância de vizinhança da matriz de pesos espaciais em 120 metros.

#### 4.3.2. Testes de autocorrelação espacial

Uma vez construída a matriz de pesos espaciais  $W$ , pode-se analisar a autocorrelação espacial nos resíduos do modelo de mínimos quadrados e estimar o modelo de regressão espacial do erro.

Com a matriz  $W$  definida foram calculadas as estatísticas correspondentes aos testes de autocorrelação espacial de Moran, os testes do Multiplicador de Lagrange para o modelo da variável dependente e do erro e suas estatísticas robustas. Os resultados, obtidos usando o software Geoda (Anselin, 2003) podem ser observados 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 uma forte autocorrelação espacial nos resíduos do modelo de regressão por mínimos quadrados. O teste de Moran e os dois testes do modelo espacial do erro mostram-se fortemente significativos.

Desta forma é comprovada a existência de uma estrutura espacial nos resíduos do modelo de mínimos quadrados. Este resultado era esperado, pois as variáveis explicativas da regressão representam apenas os fatores construtivos e de tipologia dos imóveis, não considerando nenhum fator de localização. A ausência dos fatores de localização no modelo faz com que o resíduo do modelo carregue o efeito destas variáveis não especificadas. O termo do erro, por conter o efeito dos fatores de localização, apresenta uma estrutura de dependência espacial e este resultado é comprovado com os testes de autocorrelação espacial.

Comprova-se também desta forma a limitação do método de mínimos quadrados para a estimação dos parâmetros obtidos por este modelo, e a carência de validade dos testes de significância nele calculados.

#### 4.3.3. Modelo espacial do erro

Com a matriz W definida foi estimado o modelo espacial do erro considerando as mesmas variáveis explicativas anteriormente descritas.

Os resultados obtidos para o modelo espacial do erro usando o software Geoda 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.

A estatística z corresponde ao equivalente para a regressão por máxima verossimilhança ao valor t de Student para o método de mínimos quadrados. As probabilidades indicam o grau de significância de cada variável de forma análoga que na regressão por mínimos quadrados.

Tabela 3. Modelo espacial do erro. Variáveis independentes.

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

Os parâmetros globais do modelo do erro e do modelo de mínimos quadrados podem ser observados na Tabela 4.

Da comparação dos resultados para os dois modelos surge a melhora substantiva do modelo espacial em todos os parâmetros de comparação. O coeficiente de determinação aumentou de 57.8% para 66.3%, o logaritmo da verossimilhança aumentou de 60.6 para 86.5 e os critérios de informação de Akaike e Schwarz, que são parâmetros indicativos da qualidade e poder explicativo da regressão também mostram que o modelo espacial foi superior ao de mínimos quadrados.

Tabela 4. Resultados do modelo do erro e de mínimos quadrados.

	MODELO DO ERRO	MÍNIMOS QUADRADOS
R <sup>2</sup>	0.6631	0.5780
Log Likelihood	86.5459	60.6014
Critério de Akaike	-139.0920	-87.2029
Critério de Schwarz	-79.2952	-27.4062
Erro padrão da regressão	0.1695	0.1965

#### 4.3.4. Cálculo dos parâmetros de homogeneização

Com os resultados obtidos pelo modelo espacial do erro podem ser calculados os fatores de homogeneização do valor dos imóveis.

Tomando antilogaritmos na equação de regressão pode ser expresso o valor unitário dos imóveis da seguinte forma:

$$\begin{aligned}
 VU = & 666.9446 * 1.0008^{APT\_AT} * 1.1357^{APT\_GAR} * 1.1137^{AP\_P2} * 1.0977^{AP\_P3} * 4.2832^{KIT} * \\
 & 0.9667^{KIT\_AT} * 1.1733^{KIT\_P2} * 2.9531^{CASA} * 0.9955^{CASA\_AC} * 1.0005^{CA\_ATER} * \\
 & 1.7415^{CASA\_P2} * 1.2517^{CASA\_P3} * 0.7828^{TERRENO} * 1.0004^{TER\_AT} * 1.6608^{COMERCIA} \\
 & * 1.1128^{NOVO} * 2.4048^{We}
 \end{aligned}$$

De acordo com estes resultados o valor unitário dos apartamentos aumenta 0.08% por cada metro quadrado a mais do apartamento, aumenta 13.57% se o imóvel tem garagem, aumenta 11.37% se é de Padrão 2 e 9.77% se é de Padrão 3. Também aumenta 11.28% se é um imóvel novo. As kitinetes diminuem seu valor unitário 3.33% por cada metro quadrado a mais que tem o imóvel na faixa de valores observados, e aumenta um 17.33% se é de Padrão 2 (as kitinetes da amostra não apresentam nenhum imóvel de Padrão 4, existindo somente os Padrões 2 e 3). Também no caso de imóvel novo aumenta 11.28%. O valor unitário das casas diminui com a área construída e aumenta com área do terreno e mostra maiores diferenças entre os imóveis dos diferentes padrões. O valor unitário dos terrenos aumenta com a área do terreno. Os fatores correspondentes às variáveis dicotômicas de tipo de imóvel servem para homogeneizar os valores unitários médios entre os imóveis de cada tipo para um valor homogêneo geral.

O valor unitário homogeneizado pode ser calculado com os novos fatores da mesma forma que foi calculado anteriormente. Os fatores considerados para realizar a homogeneização são os correspondentes às variáveis construtivas e de tipo de imóvel, como é mostrado a seguir:

$$\begin{aligned}
 VH = & VU / [1.0008^{APT\_AT} * 1.1357^{APT\_GAR} * 1.1137^{AP\_P2} * 1.0977^{AP\_P3} * 4.2832^{KIT} * \\
 & 0.9667^{KIT\_AT} * 1.1733^{KIT\_P2} * 2.9531^{CASA} * 0.9955^{CASA\_AC} * 1.0005^{CA\_ATER} * \\
 & 1.7415^{CASA\_P2} * 1.2517^{CASA\_P3} * 0.7828^{TERRENO} * 1.0004^{TER\_AT} * 1.6608^{COMERCIA} \\
 & * 1.1128^{NOVO}]
 \end{aligned}$$

O VH assim calculado carrega ainda os efeitos espaciais estimados no modelo espacial do erro, corrigindo apenas os efeitos das variáveis construtivas e de tipo. Os efeitos espaciais provenientes dos fatores de localização estão presentes na variabilidade de VH.

#### 4.3.5. Aplicação do método geoestatístico

Tendo calculado os valores unitários homogeneizados dos imóveis pode ser aplicado o método geoestatístico para modelar a estrutura espacial do valor da localização.

O valor homogeneizado VH representa a variável regionalizada que depende dos fatores de localização que influenciam no valor dos imóveis.

O primeiro passo na aplicação do método geoestatístico é a análise do semivariograma experimental de VH.

##### 4.3.5.1. Semivariograma experimental de VH

A partir da estimação do valor unitário homogeneizado dos imóveis foi calculado o novo semivariograma experimental.

O semivariograma isotrópico, ou seja, calculado sem considerar direções preferenciais, aparece na Figura 3. Nela pode ser comprovada a existência de uma estrutura de variação espacial que mostra menor variância para as observações mais próximas e maiores variâncias para as observações mais distantes.

O semivariograma apresenta um comportamento de efeito pepita, ou seja, a possibilidade de importantes variações para observações muito próximas, e um alcance e patamar definidos, onde o gradiente de variação se estabiliza aproximadamente a partir dos 400 metros de distância.

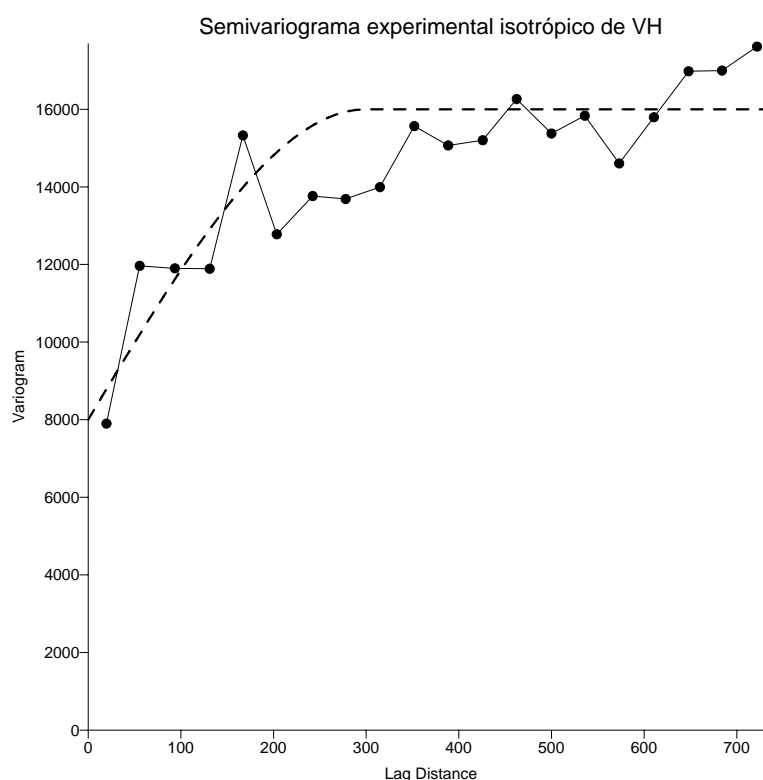


Figura 3. Semivariograma experimental isotrópico do valor unitário homogeneizado por regressão espacial.

Para analisar possíveis efeitos de anisotropia são calculados os semivariogramas direcionais para todas as possíveis direções.

A partir deste cálculo é comprovada a existência de anisotropia nos dados como pode ser observado nas Figuras 4 e 5.

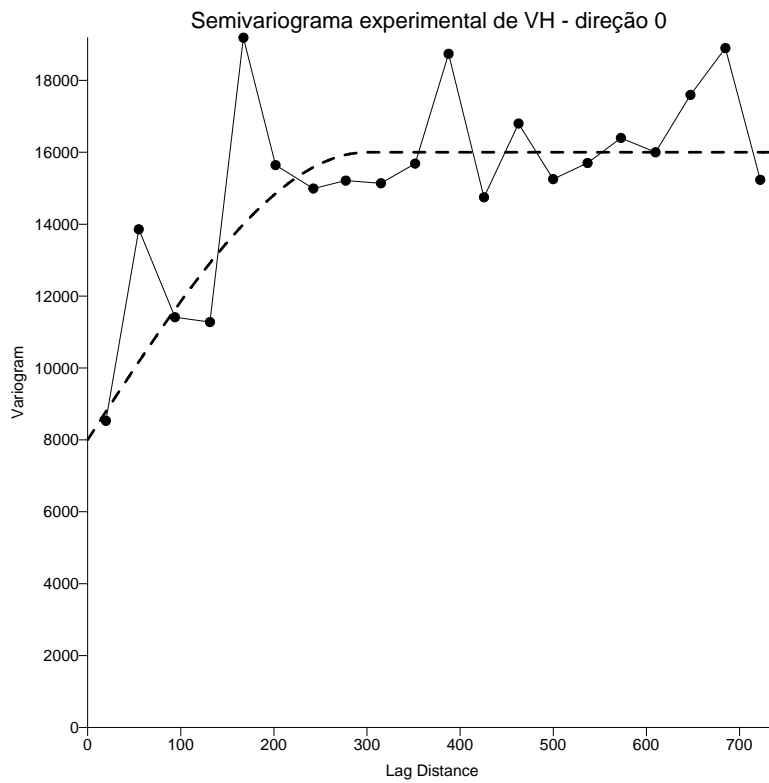


Figura 4. Semivariograma experimental de VH na direção 0° (LO).

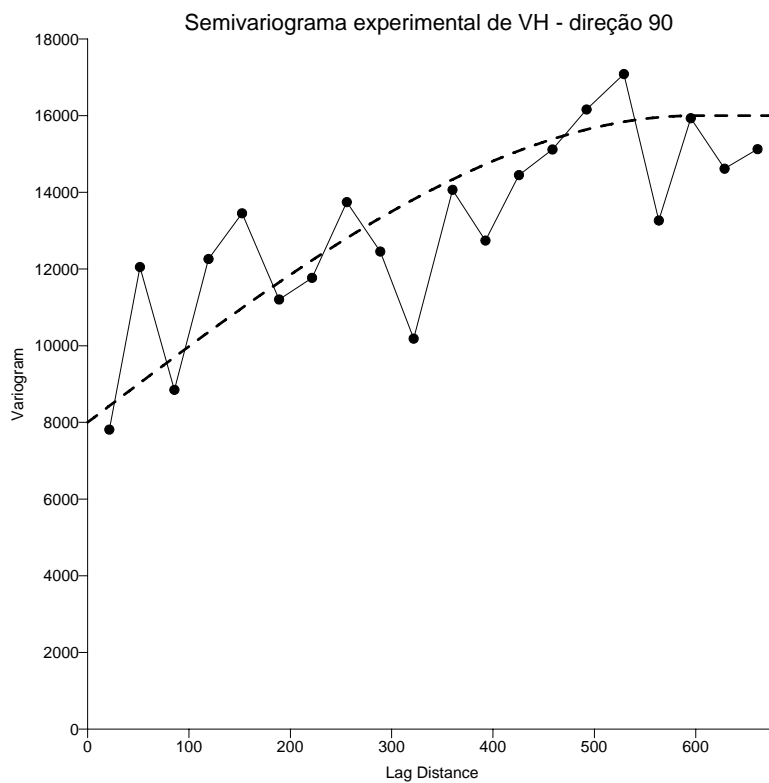


Figura 5. Semivariograma experimental de VH na direção 90° (NS).

As Figuras 4 e 5 mostram os semivariogramas direcionais para as direções 0 e 90 graus, ou seja, direções LO e NS respectivamente. Estas duas direções são as que apresentam semivariogramas com menor e maior alcances respectivamente sendo, portanto, as direções principais de anisotropia.

A partir da análise destas figuras pode ser comprovado que a influência entre observações na direção LO é mais reduzida que na direção NS. Apesar de terem alcances diferentes entre si, o patamar observado em ambos semivariogramas tem valor semelhante, caracterizando anisotropia geométrica da variável regionalizada.

#### 4.3.5.2. Modelagem da estrutura espacial de VH

A análise dos semivariogramas experimentais comprova a existência de um padrão de variação espacial que pode ser modelado pelo modelo teórico esférico de semivariograma, que permite uma variação mais intensa no início do semivariograma e uma estabilização em meseta depois de alcançar o valor da distância de estabilização. Os parâmetros do semivariograma esférico encontrado podem ser observados na Tabela 5.

Tabela 5. Parâmetros do semivariograma estimado.

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

Estes parâmetros mostram que os dados apresentam uma variabilidade devida à estrutura espacial que responde por 50% da variância total, sendo os outros 50% devidos a outras causas, como erros de medição ou a influência de outras variáveis que ainda afetam o valor dos imóveis.

O alcance máximo da dependência espacial é de 600 metros na direção NS enquanto que o fator de anisotropia igual a 2 indica que o alcance da influência entre observações na direção NS é igual ao dobro da distância de influência na direção LO.

Com os parâmetros do semivariograma assim definidos pode ser efetuada a krigagem dos valores.

#### 4.3.5.3. Krigagem por blocos dos valores homogeneizados VH

Utilizando os parâmetros do semivariograma e a definição de uma grade de 100 metros de lado foi realizada a krigagem por blocos da variável VH.

O resultado da krigagem de VH é a nova variável VL e pode ser observada graficamente através das curvas de isovalores de VL que mostra a Figura 6.

Nesta Figura as isolinhas foram desenhadas a cada 50 R\$/m<sup>2</sup>. A proximidade entre isolinhas permite identificar as zonas onde a variação do valor de localização é mais acentuada. Onde as isolinhas se encontram mais próximas existe um gradiente maior de variação. Já nas áreas onde as isolinhas se encontram mais espaçadas o valor de localização varia de forma menos acentuada no espaço.

A escolha do gradiente de cores permite também uma análise visual mais rápida e simples das áreas mais e menos valorizadas dentro da região, os pólos de valorização mais importantes e as áreas de menor valor.



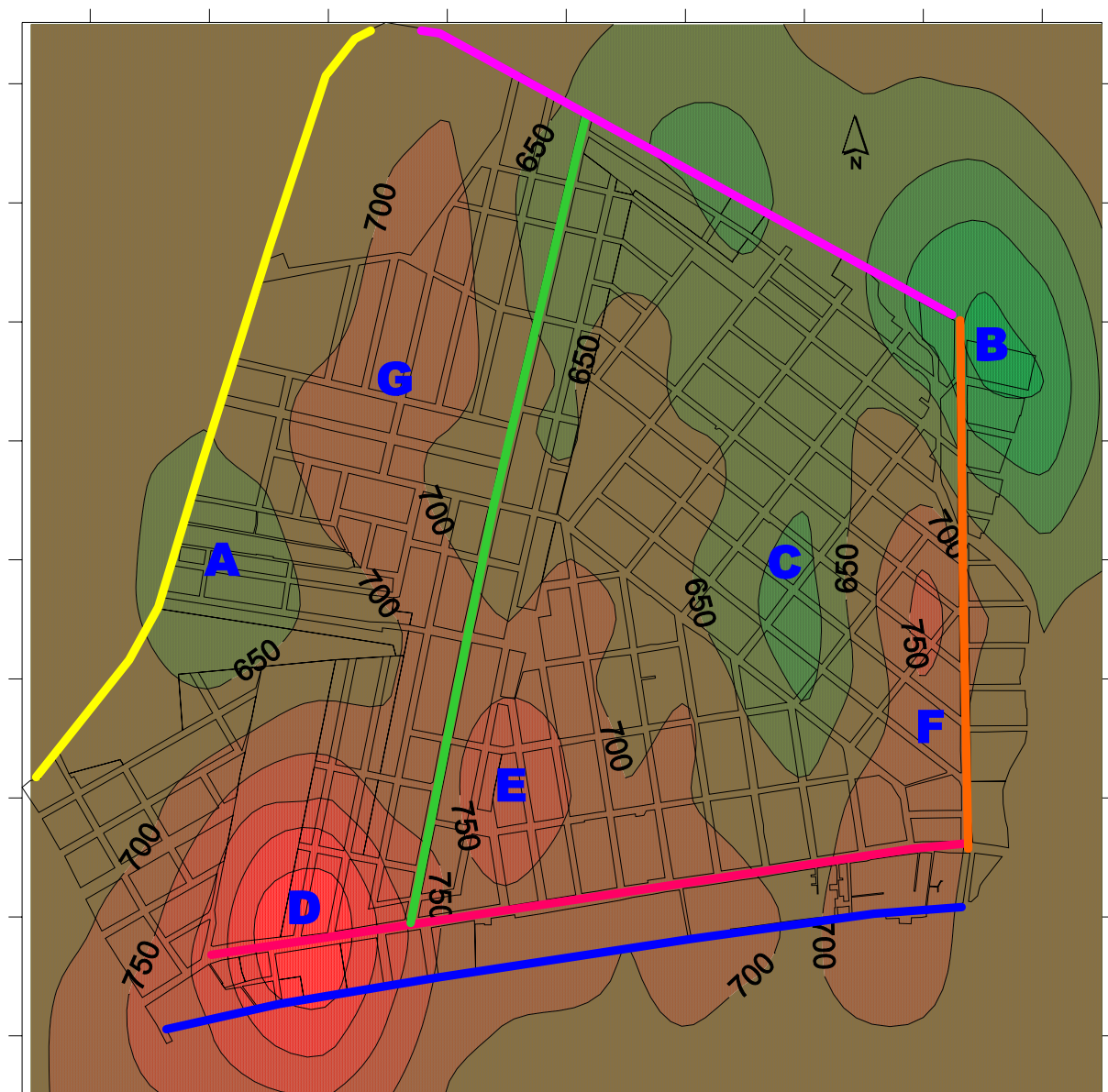


Figura 6. Curvas de isovalores da variável VL. Fonte: Peruzzo Trivelloni (2005).

A partir da análise da Figura 6 podem ser observados os principais centros de valorização ou desvalorização na área de estudo considerada. As isolinhas permitem determinar a localização e extensão de cada um destes centros ou regiões.

Como era esperado, a região mais valorizada da área de estudo se encontra nas proximidades da interseção da Avenida Presidente Kennedy (linha vermelha) e a Avenida Central do Kobrasol (linha verde). Se bem que quase toda a extensão da Avenida Kennedy está mostrando valores altos, os mais elevados se encontram nas proximidades da Avenida Central do Kobrasol. A krigagem mostra duas regiões, denominadas D e E na Figura 6, onde os valores são mais elevados. A região D é uma área de atual expansão comercial e residencial e por isso está sendo mais valorizada. A região E se encontra próxima ao Centro Comercial Campinas e a novos empreendimentos comerciais com prédios modernos de recente construção. Resulta importante salientar que a valorização da região mais próxima à Avenida Presidente Kennedy em toda a sua extensão também está influenciada pela recente

construção da avenida Beira Mar de São José (representada pela linha azul na Figura 6) que constitui um novo pólo de lazer que está contribuindo para valorizar a região próxima dela.

A Avenida Central do Kobrasol mostra maior valorização na sua parte sul, próxima à Avenida Presidente Kennedy, onde está localizado o Centro Comercial Kobrasol, onde existe maior concentração comercial e se encontram as principais agências bancárias na região. Na medida que a distância da Avenida Presidente Kennedy aumenta, os valores na Avenida Central do Kobrasol diminuem, enquanto a concentração da atividade comercial também diminui e a proporção de prédios residenciais mais antigos aumenta.

Existe outra região com boa valorização que é denominada G na Figura 6 e se caracteriza por ser uma área com várias escolas e centros educacionais, atividades de serviços e também como área residencial, com ruas tranquilas e amplas. Esta região tem boa acessibilidade aos principais centros comerciais e as vias de entrada e saída da área de estudo.

A região denominada F também se mostra como uma área valorizada, por estar próxima da rua Josué di Bernardi (linha laranja), com importante concentração comercial. Esta rua também mostra maior valorização na parte mais próxima da Avenida Presidente Kennedy, enquanto o valor diminui progressivamente com esta distância.

A Figura 6 mostra também as regiões menos valorizadas da região. Nela pode ser observado que o valor diminui próximo à rodovia BR 101 (linha amarela) e à BR 282 - Via Expressa (linha rosa). Entretanto, duas regiões específicas próximas a estas duas rodovias, as regiões denominadas A e B mostram maior desvalorização relativa.

No caso da região A, a desvalorização é devida às características da vizinhança, apresentando ruas muito estreitas, tendo difícil acessibilidade, e padrão construtivo inferior nos imóveis.

No caso da região B, a rua Josué di Bernardi passa por baixo de um viaduto da rodovia Via Expressa, comunicando a área de estudo com a região ao norte da rodovia. Existem apenas dois pontos de comunicação entre a área de estudo e a região ao norte da Via Expressa: uma neste ponto, pela Josué di Bernardi, e o outro num viaduto próximo da BR 101. No caso do viaduto localizado na região B, este comunica com uma região de favela, sendo este o motivo pelo qual o valor diminui abruptamente nas suas proximidades, tendo o menor valor dentro de toda a área de estudo.

Finalmente a região denominada C também apresenta menor valor. Esta área não se encontra nas proximidades dos principais centros comerciais nem de serviços tendo por este motivo menor atratividade que as outras.

Em resumo, o resultado obtido por meio das curvas de isovalores produto da krigagem permite analisar e interpretar consistentemente a complexa interação de fatores de valorização e desvalorização da região. Os resultados apresentam coerência com a realidade da área de estudo possibilitando uma fácil interpretação.

A variável VL obtida pela krigagem representa de forma consistente a valorização e desvalorização das diferentes localizações a partir das informações contidas na amostra de mercado.

O método da krigagem produz uma suavização nos dados, diminuindo os valores extremos devido às características da interpolação, aos parâmetros do semivariograma e também à homogeneização produzida pela krigagem por blocos.

A krigagem por blocos faz que possíveis valores atípicos presentes nos dados sejam suavizados. Estes valores extremos podem ser erros de medida ou refletir condições especiais de localização, sendo importantes neste último caso para realizar estimações adequadas da localização. Por este motivo foram mantidos todos os dados nas etapas de homogeneização e krigagem dos dados.

#### 4.3.6. Análise de significância da variável VL no modelo de avaliação em massa

Com o objetivo de verificar a significância estatística da variável VL obtida por krigagem e do modelo de valor proposto, foi calculado um novo modelo de regressão onde a variável VL é incluída como variável explicativa junto com as variáveis de tipo e construtivas dos imóveis anteriormente utilizadas.

Os resultados do novo modelo de mínimos quadrados mostram uma melhora importante no poder explicativo como pode ser observado na Tabela 6.

Da Tabela 6 pode ser concluído que a inclusão da variável VL produziu um forte aumento do poder explicativo da regressão em relação ao modelo espacial do erro: o coeficiente de determinação passa de 0.6632 para 0.7502 e 0.7330 como coeficiente ajustado.

Os resultados dos critérios de Akaike e Schwarz assim como todos os outros parâmetros de comparação mostram que a nova regressão apresenta melhores resultados que os modelos anteriormente calculados.

Tabela 6. Resultados da regressão por mínimos quadrados com a variável VL.

Variável dependente	L(VU)
$R^2$	0.7502
$R^2_a$	0.7330
F	43.54
Prob (F)	<0.0001
Log likelihood	125.878
Akaike info criterio	-217.757
Schwarz criterio	-157.96
Erro padrão da regressão	0.1512

Outro resultado muito importante é o referido aos testes de autocorrelação espacial dos resíduos da nova regressão.

Na Tabela 7 pode ser observado o resultado do diagnóstico de autocorrelação dos resíduos, onde todos os testes mostram que a nova regressão não apresenta autocorrelação espacial dos resíduos.

Tabela 7. Testes de autocorrelação espacial do novo modelo.

Teste	Probabilidade
Moran's I	0.7920
LM (lag)	0.5272
LM robusto (lag)	0.5744
LM (erro)	0.5430
LM robusto (erro)	0.5931

Os testes de Moran e do Multiplicador de Lagrange para os modelos do erro e da variável dependente mostram que não existe autocorrelação significativa nos resíduos da regressão sendo válidas, portanto, as hipóteses de mínimos quadrados e os testes de significância das variáveis explicativas e seus parâmetros.

Os resultados do novo modelo de regressão mostram que a variável VL conseguiu representar os fatores espacialmente correlacionados no valor unitário dos imóveis,

eliminando a autocorrelação espacial nos resíduos e melhorando o poder de explicação geral do modelo.

Da análise dos resíduos do modelo foi observada a existência de seis dados atípicos, com valores do resíduo padronizados superiores a 2.3 desvios padrões em valor absoluto.

Estes dados foram retirados da análise e os resultados para o novo modelo são apresentados na Tabela 8, mostrando um aumento do poder explicativo do modelo em todos os parâmetros.

Tabela 8. Resultados do modelo de regressão para 243 dados.

Variável dependente	L(VU)
R <sup>2</sup>	0.7661
R <sup>2</sup> <sub>a</sub>	0.7496
F	46.276
Prob (F)	<0.0001
Log likelihood	158.569
Critério Akaike	-283.137
Critério Schwarz	-223.755
Erro padrão	0.1308

Os resultados para os parâmetros das variáveis explicativas do modelo e suas significâncias podem ser observados na Tabela 9. Eles mostram que todas as variáveis explicativas são significativas ao nível de 20%, e todas exceto uma (AP\_P3) são significativas ao nível de 5%.

A variável explicativa VL, que representa o valor da localização dos imóveis, se mostra como uma das variáveis individuais com maior grau de significância. A incorporação desta variável mostrou uma melhora importante no poder explicativo deste modelo em relação ao modelo espacial do erro.

O modelo de avaliação em massa pode ser escrito da seguinte forma:

$$L(VU) = 5.1641 + 0.0020*VL + 0.1569*APT\_GAR + 0.1248*AP\_P2 + 0.0690*AP\_P3 + 1.6080*KIT - 0.0393*KIT\_AT + 0.1589*KIT\_P2 + 0.71*CASA - 0.0032*CASA\_AC + 0.0007*CA\_ATER + 0.4018*CASA\_P2 + 0.2700*CASA\_P3 - 0.2577*TERRENO + 0.0004*TER\_AT + 0.4554*COMERCIA + 0.1180*NOVO$$

Que também pode ser escrito da seguinte forma:

$$VU = e^{(5.1641 + 0.0020*VL + 0.1569*APT\_GAR + 0.1248*AP\_P2 + 0.0690*AP\_P3 + 1.6080*KIT - 0.0393*KIT\_AT + 0.1589*KIT\_P2 + 0.71*CASA - 0.0032*CASA\_AC + 0.0007*CA\_ATER + 0.4018*CASA\_P2 + 0.2700*CASA\_P3 - 0.2577*TERRENO + 0.0004*TER\_AT + 0.4554*COMERCIA + 0.1180*NOVO)}$$

Os testes de normalidade e homocedasticidade dos resíduos mostraram resultados satisfatórios. A Figura 7 mostra graficamente a normalidade dos resíduos do modelo, comparando o valor dos resíduos com seu valor normal esperado, obtendo-se um ajuste adequado.

Tabela 9. 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

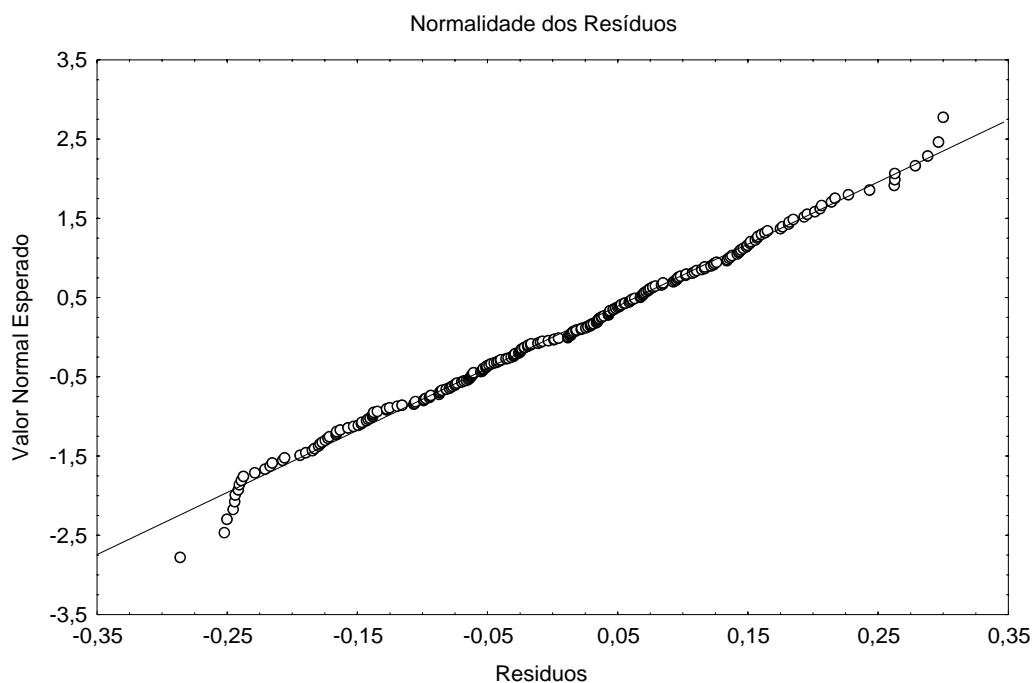


Figura 7. Gráfico de normalidade dos resíduos da regressão.

A distribuição dos resíduos padronizados do modelo de regressão em relação aos valores estimados pode ser observada na Figura 8, mostrando homocedasticidade dos resíduos e a inexistência de dados atípicos.

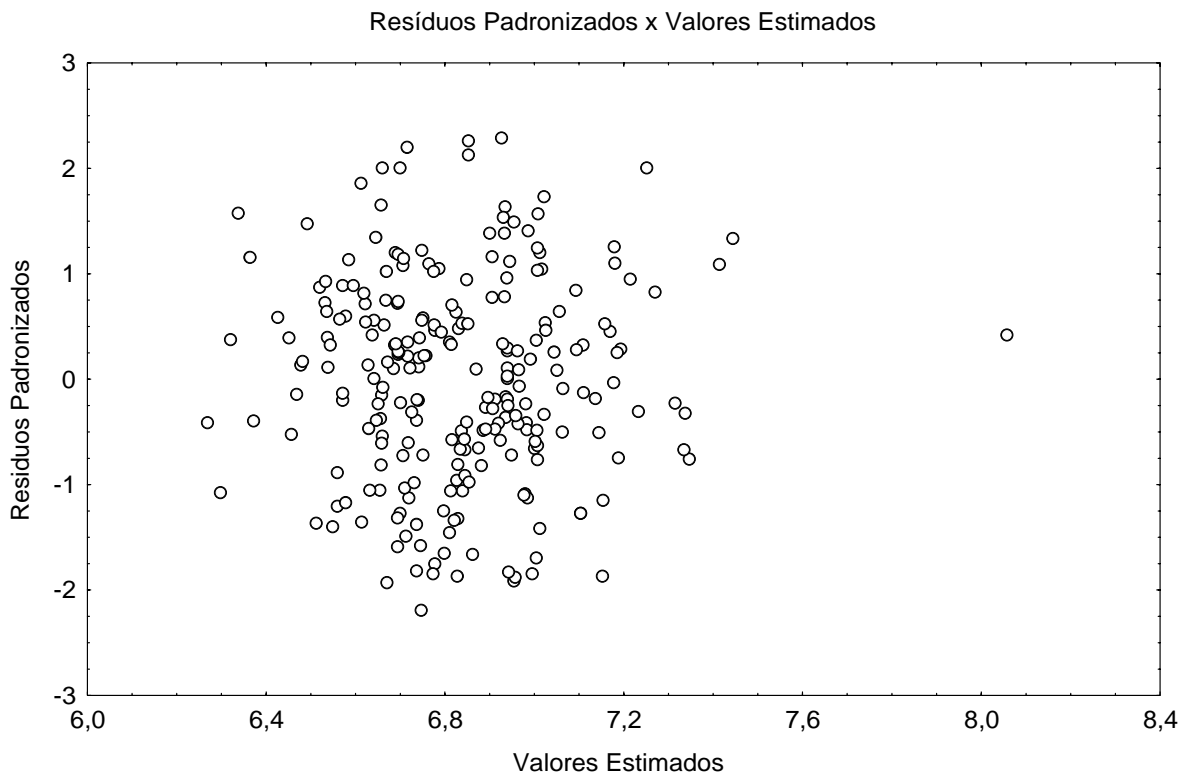


Figura 8. Gráfico dos resíduos padronizados e valores estimados da regressão.

Para analisar o desempenho do modelo para diferentes valores de VL, ou seja, para imóveis em regiões mais valorizadas e menos valorizadas, foi calculado o gráfico de dispersão entre os resíduos padronizados e a variável VL.

A Figura 9 mostra que a distribuição dos resíduos padronizados é uniforme para os diferentes valores de VL, mostrando que o modelo avaliou de forma adequada tanto os imóveis com maior valor de localização quanto os de menor valor de localização.

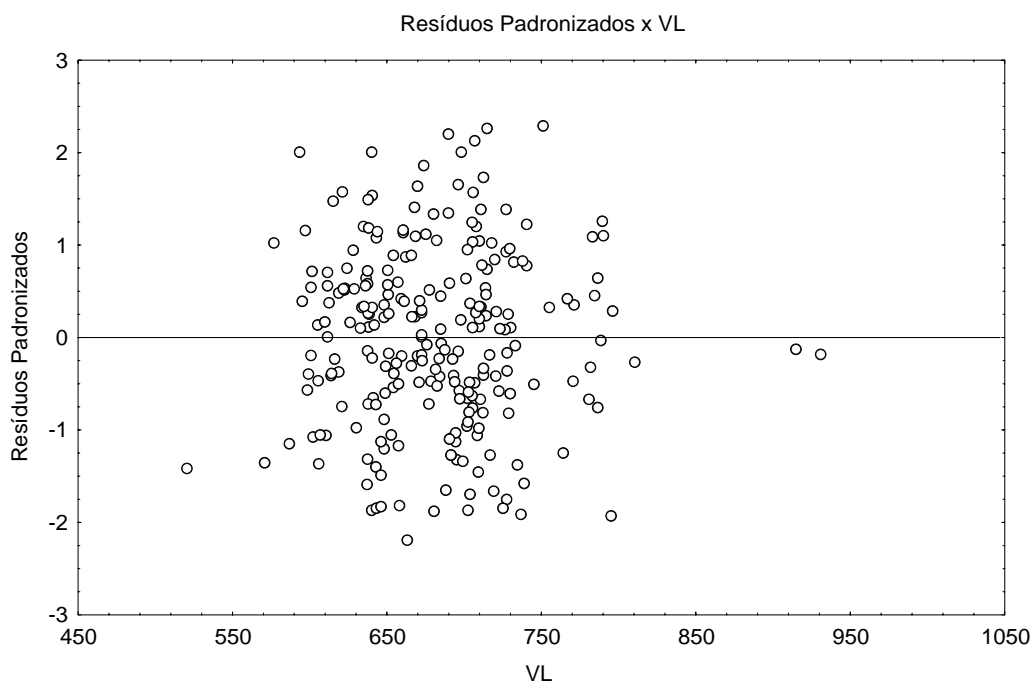


Figura 9. Gráfico de dispersão dos resíduos padronizados e a variável VL.

#### 4.3.7. Análise de performance da avaliação em massa

Os resultados obtidos pelo novo modelo de regressão mostram uma melhora importante na sua capacidade de predição em relação aos modelos anteriormente considerados. Para analisar o desempenho do modelo em termos de padrões de performance para uma avaliação em massa, foi considerada a norma da IAAO (*International Association of Assessing Officers*) de estudos de *ratios* para avaliações em massa.

Foram calculados os valores correspondentes da mediana de quocientes de avaliação (*ratios*), o coeficiente de dispersão (COD) e o diferencial relativo ao preço (PRD) que são os parâmetros de comparação recomendados por esta norma. Estes parâmetros foram estimados para o conjunto da amostra em geral, e por separado para os imóveis construídos e os terrenos, pois a norma estabelece valores de parâmetros diferentes para cada tipo de imóvel.

Segundo esta norma os parâmetros de performance da avaliação em massa devem ser calculados sem considerar os valores extremos do modelo.

Neste caso foram utilizados para o cálculo dos parâmetros todos os 243 dados do modelo de regressão final, ou seja, apenas os seis elementos considerados atípicos foram retirados. Os resultados obtidos podem ser observados na Tabela 10.

Tabela 10. Parâmetros de performance da avaliação em massa.

ESTATÍSTICA	AMOSTRA TOTAL	IMOVEIS CONSTRUÍDOS	TERRENOS
Mediana <i>ratios</i>	0.989	0.987	1.024
Média <i>ratios</i>	1.008	1.008	1.008
Média ponderada	0.991	0.991	0.997
Média desvios absolutos	0.104	0.104	0.093
COD	10.492	10.524	9.074
PRD	1.017	1.017	1.011

Os valores obtidos para a mediana dos *ratios*, o COD e o PRD devem ser comparados com os valores recomendados pela IAAO que se encontram na Tabela 11.

De acordo com estes valores, pode se concluir que os resultados obtidos na avaliação em massa são satisfatórios em relação aos padrões propostos pela norma da IAAO, tanto para os imóveis construídos quanto para os terrenos e para todos os parâmetros de avaliação.

Tabela 11. Valores padronizados para avaliação em massa propostos pela IAAO.

ESTATÍSTICA	VALORES RECOMENDADOS
Mediana dos <i>ratios</i> de avaliação	Entre 0.9 e 1.1
COD imóveis construídos	<15.0
COD terrenos	<20.0
PRD	Entre 0.980 e 1.030

#### 4.4. Aplicação para Planta de Valores Genéricos de Terrenos

O método proposto permite calcular o valor de localização e o valor dos imóveis dos tipos considerados na amostra de mercado, para toda a área de estudo.

Como resultado direto da krigagem pode ser obtido o mapa dos Valores de Localização em R\$/m<sup>2</sup> para o ponto médio de cada face de quadra da área de estudo como mostra a Figura 10. Desta forma, a partir do modelo proposto poderia ser calculada a Planta de Valores Genéricos dos imóveis da área de estudo, em função do tipo de imóvel, da localização e das características de cada imóvel.

Como aplicação prática do método proposto foi calculada a Planta de Valores Genéricos de Terrenos na área de estudo, ou seja, os valores médios por face de quadra de um terreno do tamanho padrão, no caso um terreno com área de 360 metros quadrados.

Aplicando o modelo para imóveis de tipo terreno com área de 360 metros quadrados foi obtida a Planta de Valores Genéricos de Terrenos, em R\$/m<sup>2</sup>, mostrada na Figura 11.

Os valores foram estimados a partir das seguintes equações:

$$L(VU) = 5.1641 - 0.2576 + 0.0004 * 360 + 0.0020 * VL$$

$$VU = 156.10 * 1.002^{VL}$$



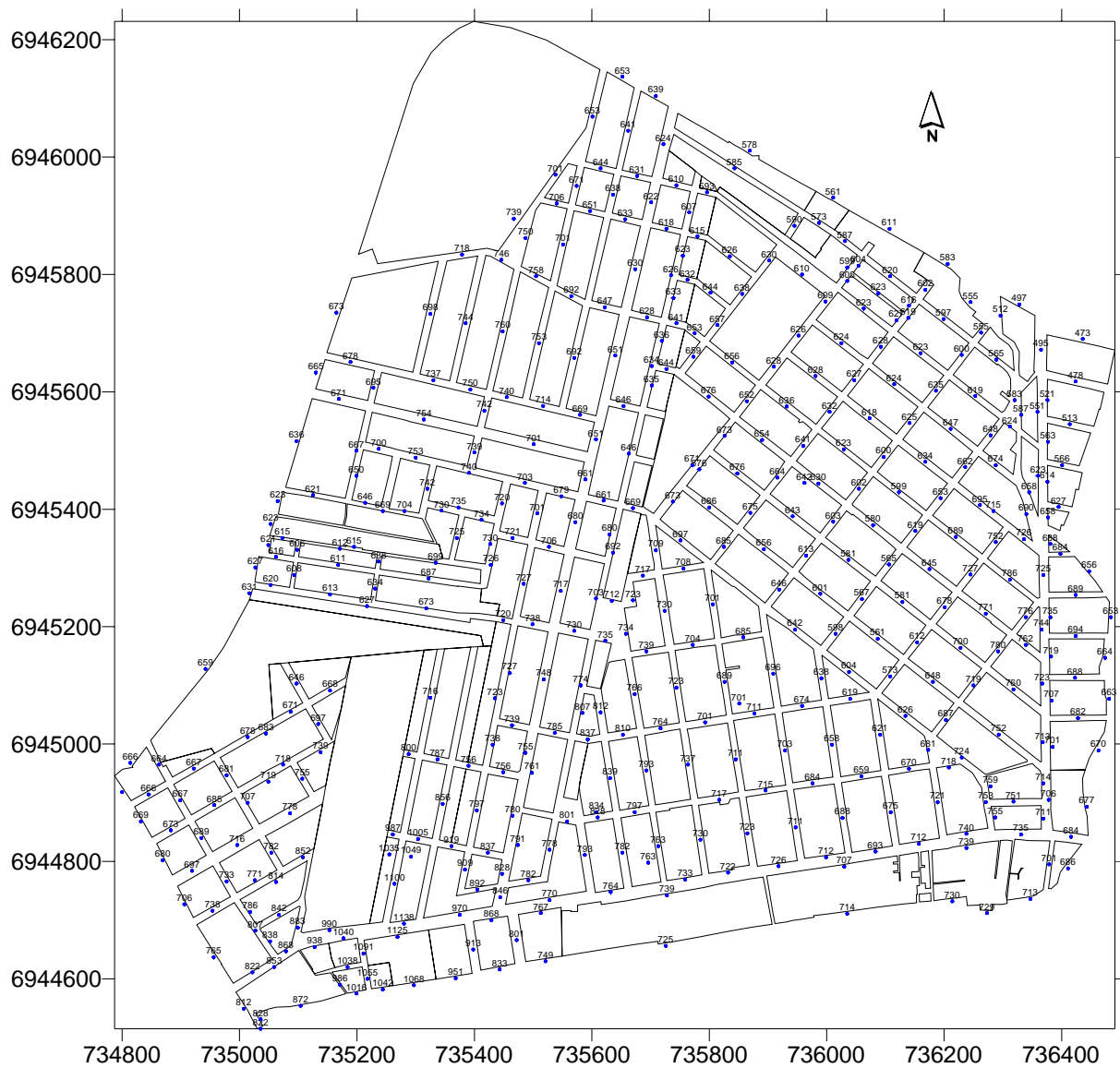


Figura 10. Mapa de Valor da Localização (R\$/m<sup>2</sup>) para cada face de quadra da área de estudo.



Figura 11. Planta de Valores Genéricos de Terrenos (R\$/m<sup>2</sup>) da área de estudo.

## **5. CONCLUSÕES**

### **5.1. Sobre o valor da localização**

1. O método proposto mostrou que o valor da localização de um imóvel, resultado da interação de diversos pólos e efeitos de valorização e desvalorização, pode ser considerado globalmente como uma variável regionalizada que representa todos estes efeitos e modelado por métodos geoestatísticos.
2. O método proposto permitiu estimar um índice do valor da localização que se mostrou fortemente significativo no modelo de regressão do valor unitário dos imóveis.
3. A variável de localização estimada VL refletiu de forma coerente e consistente as regiões de maior valorização da área de estudo em relação a todos os pólos de valorização.
4. A variável VL mostrou-se também adequada na modelagem das áreas menos valorizadas. Estas regiões mostraram-se relacionadas a condições de qualidade da vizinhança e acessibilidade pouco favoráveis. Problemas como a largura das ruas, vias de trânsito pesado, poluição sonora, maior distância aos principais centros comerciais ou a proximidade a uma favela foram refletidos adequadamente pelo índice de valorização VL.
5. O método permitiu estimar a variável regionalizada VL com um grau de confiabilidade adequada e uniforme em quase toda a área de estudo. Os desvios mais importantes de VL em relação aos dados de mercado foram produzidos naqueles pontos onde tomou valores extremos, os maiores ou os menores valores.
6. O estudo do semivariograma experimental dos valores homogeneizados permitiu analisar e modelar a estrutura espacial do valor dos imóveis, analisando o gradiente de variação em função da distância e da direção entre os dados de mercado, modelando a anisotropia existente.
7. A escolha da krigagem por blocos e a definição do tamanho dos blocos permitiu minimizar o efeito de possíveis erros aleatórios presentes nos dados de mercado, estimando médias consistentes do valor da localização em toda a área de estudo e sem produzir deformações importantes nos valores extremos.
8. A multiplicidade de pólos de valorização evidenciados pelos resultados, cada um caracterizado pela sua área de influência, gradiente e anisotropia específica, mostrou a adequação do método geoestatístico para modelar o valor da localização. Analisando a estrutura espacial dos dados, o método geoestatístico permitiu aproveitar de forma adequada a informação sobre a localização fornecida pelos dados de mercado.

### **5.2. Sobre o uso combinado de técnicas de Estatística Espacial.**

1. 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. Os resultados obtidos foram superiores aos conseguidos usando uma das técnicas de forma individual, possibilitando a análise de uma amostra heterogênea de dados e a determinação do valor da localização.
2. O uso combinado das duas técnicas estatísticas permitiu analisar de forma adequada as características mistas do valor dos imóveis como objetos e como variáveis regionalizadas. O modelo de regressão espacial do erro mostrou-se apropriado para a homogeneização dos valores em função das características construtivas na presença de uma estrutura espacial nos resíduos, enquanto que o método geoestatístico mostrou-se adequado para analisar a estrutura espacial dos valores homogeneizados e estimar o valor da localização dos imóveis.
3. A combinação de técnicas de regressão espacial e geoestatísticas permitiu resolver a heterocedasticidade induzida pelo modelo de regressão espacial do erro. A estimação de VL

pelo método geoestatístico não gerou heterocedasticidade no modelo final de regressão por mínimos quadrados.

### **5.3. Sobre a autocorrelação espacial.**

1. O método permitiu comprovar a existência de autocorrelação espacial nos resíduos do modelo de regressão por mínimos quadrados quando as variáveis de localização não são adequadamente representadas no modelo.
2. Foi comprovado que o método de homogeneização de valores por mínimos quadrados não é adequado quando se pretende homogeneizar dados em função das características construtivas e tipológicas dos imóveis, pois os resíduos do modelo apresentam autocorrelação espacial invalidando os resultados da estimação.
3. A inclusão da variável regionalizada VL no modelo de regressão do valor unitário dos imóveis conseguiu representar a dependência espacial nos dados, eliminando a autocorrelação espacial nos resíduos da regressão. Desta forma o método de mínimos quadrados pôde ser utilizado na estimação do modelo final, uma vez que a autocorrelação espacial foi corretamente modelada.

### **5.4. Sobre o modelo geral.**

1. O modelo final de regressão por mínimos quadrados incluindo a variável de localização estimada apresentou desempenho superior ao modelo espacial do erro melhorando o seu poder explicativo.
2. O modelo de avaliação em massa apresentou parâmetros de performance satisfatórios de homogeneidade e uniformidade, tanto para os imóveis construídos quanto para os terrenos.
3. O método adotado mostrou que é possível a utilização de dados de mercado de todos os tipos de imóveis para determinar o valor da localização na avaliação em massa de imóveis. Os fatores de localização mostraram um efeito similar para todos os tipos de imóveis e a estimação de valorização específica para cada tipo foi modelada por regressão.
4. O modelo multiplicativo proposto permitiu comprovar que existe interação entre os fatores de localização e construtivos, mostrando que estes não são independentes e que o valor dos imóveis não é corretamente formulado por meio da soma do valor do terreno e da construção como componentes independentes entre si.

### **5.5. Sobre o uso do modelo de regressão do erro no processo de homogeneização de valores.**

1. O modelo de regressão espacial do erro permitiu encontrar parâmetros de homogeneização dos fatores construtivos de forma eficiente e confiável, ajustando a estimação ao erro autocorrelacionado produzido pelos fatores de localização. Os valores obtidos por este processo resultaram efetivamente mais homogêneos, com medianas semelhantes e redução da variância para todos os tipos de imóveis. A homogeneização pelo modelo do erro manteve nos valores as diferenças devidas à localização dos imóveis, possibilitando sua posterior modelagem pelo método geoestatístico.
2. A significância do termo de autocorrelação no modelo do erro espacial mostrou a influência dos fatores de localização gerando uma estrutura espacial nos resíduos. O parâmetro de autocorrelação permitiu estimar a influência da vizinhança para cada dado pontual da amostra, mas não é suficiente para analisar a estrutura de variabilidade espacial dos resíduos. Desta forma, o modelo de regressão espacial por si só não permite a estimação do valor da localização nos outros pontos da área de estudo.

### **5.6. Sobre a aplicação do método para a elaboração de Plantas de Valores Genéricos.**

1. O método proposto permitiu a elaboração de uma Planta de Valores Genéricos na área de estudo para todos os tipos de imóveis incluídos na amostra de mercado.
2. O método proposto pode ser aplicado para estudar de forma sistemática e permanente o dinamismo de valorização e desvalorização imobiliária de uma região.
3. Desta forma o método proposto poderá ser de grande utilidade para os órgãos municipais responsáveis pelo cálculo e atualização das Plantas de Valores Genéricos dos municípios e permitirá também analisar a valorização imobiliária causada por obras públicas, avaliando a área efetiva e o índice de valorização num período determinado para o cálculo e cobrança de Contribuição de Melhoria.
4. O método proposto mostra a importância de contar com dados cadastrais georreferenciados e atualizados. A implementação de Sistemas de Informações Geográficas (SIG) nas Prefeituras Municipais pode facilitar a aplicação do método na elaboração de Plantas de Valores Genéricos atualizadas e acordes aos valores de mercado.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANSELIN, L. **Spatial Econometrics**. Discussion paper. Bruton Center, School of Social Sciences, University of Texas at Dallas, 1999a.
- ANSELIN, L. **The Future of Spatial Analysis in the Social Sciences**. *Geographics Information Sciences* 5, 1999b.
- ANSELIN, L. **Spatial Externalities, Spatial Multipliers and Spatial Econometrics**. Specialist Meeting on Spatial Externalities, Santa Barbara, California, 2001.
- 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.
- ANSELIN, L. **GeoDa 0.9 User's Guide**. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, 2003.
- BASU, S., THIBODEAU, T. **Analysis of Spatial Autocorrelation in House Prices**. *Journal of Real Estate and Economics*, Vol. 17, 1, pp. 61-85, 1998.
- BESNER, C. **A Spatial Autoregressive Specification with a Comparable Sales Weighting Scheme**. *Journal of Real Estate Research*, Vol. 24, 2, pp. 193-212, 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.
- CARDIM, M. **Fundamentos de Geoestatística e Aplicações**. Minicurso. IV Semana da Estatística da FCT/UNESP. São Paulo, 2000.
- CHICA OLMO, J. **Teoría de las Variables Regionalizadas. Aplicación en Economía Espacial y Valoración Inmobiliaria**. Ed. Universidad de Granada, 1994.
- CHICA OLMO, J. **Spatial Estimation of Housing Prices and Locational Rents**. *Urban Studies*, Vol. 32, N. 8, pag. 1331-1344, 1995.
- CRESSIE, N. A. C. **Statistics for Spatial Data**. Revised Edition, John Wiley & Sons, Inc., 1993.
- 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.
- DANTAS, R. A., MAGALHÃES, A. M., VERGOLINO, J. R. de O. **Modelos Espaciais aplicados ao Mercado de apartamentos do Recife**. XII COBREAP Congresso Brasileiro de Engenharia de Avaliações e Perícias, Anais. Belo Horizonte, 2003a.
- DANTAS, R. A., SÁ, L. A.C. M., PORTUGAL, J. L. **Elaboração de Plantas de Valores sob a ótica da inferência espacial**. XII COBREAP Congresso Brasileiro de Engenharia de Avaliações e Perícias, Anais. Belo Horizonte, 2003b.
- DAVIS, P. **The IAAO Standard on Ratio Studies: a Framework for Order and Progress in Assessment Performance Evaluation**. In: Seminário Internacional sobre Tributação Imobiliária, Anais. Porto Alegre, 2001.
- 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.
- DUBIN, R. **Spatial autocorrelation: a primer.** *Journal of Housing Economics* 7, pág. 304-327, 1998.
- DUBIN, R. **Predicting Prices Using Multiple Listings Data.** *Journal of Real Estate and Economics*, Vol. 17, 1, pp. 35-59, 1998.
- DUBIN, R. **Robustness of Spatial Autocorrelation Specifications: some Monte Carlo Evidence.** *Journal of Regional Science*, Vol. 43, 2, pp. 221-248, 2003.
- DUBIN, R. **Spatial Lags and Spatial Errors Revisited: some Monte Carlo Evidence.** *LSU Econometric Conference on Spatial and Spatiotemporal Econometric*, Baton Rouge, 2003.
- DUBIN, R., SUNG, C. H. **Specification of Hedonic Regressions: Non-nested Tests on Measures of Neighborhood Quality.** *Journal of Urban Economics*, vol. 27, pp. 97-110, 1990.
- FARIAS, V. F. **São José: 250 anos. Natureza, História e Cultura.** Ed. do autor, 2ª Edição, 2001.
- GLOUDEMANS, R. J. **An Empirical Analysis of the Incidence of Location on Land and Building Values.** *Lincoln Institute of Land Policy*, Working Paper, 2002.
- GLOUDEMANS, R. J., HANDEL, S., WARWA, M. **An Empirical Evaluation of Alternative Land Valuation Models.** *Lincoln Institute of Land Policy*, Working Paper, 2002.
- MCCLUSKEY, W. J., DEDDIS, W. G., LAMONT, I. **The Application of Spatially Derived Location Factors Whiting a GIS Environment.** *Proceedings of VIII Conference of Pacific Rim Real Estate Society (PPRES)*, 2002.
- 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. **Quick computation of Spatial Autoregressive Estimators.** *Geographical Analysis*, Vol. 29, 3, pp. 232-247, 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.
- SILVA, A. M. M., HEBERLE, A. L., HOCHHEIM, N. **Determinação de Zonas Homogêneas usando Sistema de Informações Geográficas: um Estudo de Caso em Florianópolis (SC).** I Simpósio de Ciências Geodésicas e Tecnologias da Geoinformação. Anais. Recife, 2004.

## **BREVE CURRÍCULO DOS AUTORES:**

### **<sup>A</sup>Carlos Alberto Peruzzo Trivelloni.**

- Engenheiro Civil graduado na Universidade da República (Uruguai, 1993).
- Mestre em Engenharia Civil pela Universidade Federal de Santa Catarina – UFSC, Área de Concentração Cadastro Técnico Multifinalitário (1998).
- Doutor em Engenharia Civil pela Universidade Federal de Santa Catarina – UFSC, Área de Concentração Cadastro Técnico Multifinalitário e Gestão Territorial (2005).
- No X COBREAP (Porto Alegre, 1999) foi agraciado com a medalha Eurico Ribeiro pelo melhor trabalho de avaliações, desenvolvido em co-autoria com Norberto Hochheim, intitulado: “Avaliação em massa de imóveis por inferência estatística e análise multivariada.”

### **<sup>B</sup>Norberto Hochheim**

- Engenheiro Civil graduado pela Universidade Federal de Santa Catarina – UFSC (1981)
- Mestre em Engenharia de Produção, área de Engenharia Econômica, pela Universidade Federal de Santa Catarina – UFSC (1986)
- Doutor pela Université de Nancy I, França (1991).
- Professor Titular da UFSC, lotado no Departamento de Engenharia Civil.
- Ministra na Graduação as disciplinas: Engenharia de Avaliações I, Engenharia de Avaliações II, Planejamento Econômico e Financeiro. No Programa de Pós-Graduação em Engenharia Civil (onde orienta trabalhos de mestrado e doutorado) ministra as disciplinas: Engenharia de Avaliações, Tópicos Avançados em Engenharia de Avaliações, Elaboração de Plantas de Valores Genéricos, Análise de Investimentos Imobiliários, Cadastro Técnico Urbano. Ministra também, pelo IBAPE/SC, cursos de Engenharia de Avaliações e Engenharia Econômica.
- No X COBREAP (Porto Alegre, 1999) foi agraciado com a medalha Eurico Ribeiro pelo melhor trabalho de avaliações, desenvolvido em co-autoria com Carlos Alberto Peruzzo Trivelloni, intitulado: “Avaliação em massa de imóveis por inferência estatística e análise multivariada.”