

**UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS
NÍVEL MESTRADO**

LUÍS FERNANDO TAVARES VIEIRA BRAGA

**VALORAÇÃO DE IMÓVEIS NO RIO GRANDE DO SUL:
UMA ANÁLISE A PARTIR DE REGRESSÃO ESPACIAL**

SÃO LEOPOLDO-RS

2010

LUÍS FERNANDO TAVARES VIEIRA BRAGA

**VALORAÇÃO DE IMÓVEIS NO RIO GRANDE DO SUL:
UMA ANÁLISE A PARTIR DE REGRESSÃO ESPACIAL**

Dissertação apresentada como requisito parcial para a obtenção ao título de Mestre em Economia pelo Programa de Pós-Graduação em Economia – Nível Mestrado, da Universidade do Vale do Rio dos Sinos – UNISINOS.

Orientador: Prof. Dr. Tiago Wickstrom Alves

**SÃO LEOPOLDO-RS
2010**

LUÍS FERNANDO TAVARES VIEIRA BRAGA

**VALORAÇÃO DE IMÓVEIS NO RIO GRANDE DO SUL:
UMA ANÁLISE A PARTIR DE REGRESSÃO ESPACIAL**

Dissertação apresentada como requisito parcial para a obtenção ao título de Mestre em Economia pelo Programa de Pós-Graduação em Economia – Nível Mestrado, da Universidade do Vale do Rio dos Sinos – UNISINOS.

Aprovado em ___/___/___

BANCA EXAMINADORA

Prof. Tiago Wickstrom Alves - UNISINOS - Presidente/Orientador

Prof. Rubens Alves Dantas – UFPE

Prof. Guilherme Vaccaro - UNISINOS

Prof. Igor A. Clemente de Moraes - UNISINOS

Prof. André Filipe Zago de Azevedo - UNISINOS - Suplente

Visto e permitida a impressão
São Leopoldo, ___/___/___.

Prof. Dr. André Filipe Zago de Azevedo
Coordenador Executivo PPG em Economia

A Deus,
a minha mãe Aura Regina
e a minha filha Camila

AGRADECIMENTOS

No momento em que concluo esta dissertação é que me dou conta o quão importantes determinadas pessoas foram no decorrer desta jornada.

À UNISINOS, ao pessoal da secretaria da pós-graduação, em especial ao Coordenador do Mestrado professor Dr. André Filipe Zago de Azevedo, por ter me proporcionado a flexibilização de alguns prazos, permitindo com isso a conclusão deste trabalho.

Ao professor e orientador Dr. Tiago Wickstrom Alves, grande mestre do saber, amigo de todas as horas, responsável pela permanente motivação, sempre atento, nunca permitindo a mudança do foco. Tiago, obrigado pelo conhecimento transmitido, pela paciência e por acreditar que, sem dúvida, eu obteria êxito.

À Caixa Econômica Federal, em especial a Arquiteta Silvia Mary Corteletti, Supervisora de Avaliações, Custos e Perícias da GIDUR de Porto Alegre, fundamental no auxílio da definição do tema desta pesquisa e por ter disponibilizado os dados aqui utilizados.

À UCPEL e à UFRGS, pelos conhecimentos adquiridos durante as graduações e pós-graduações que fiz.

Ao Professor Dr. Luiz Fernando Heineck, meu grande mestre da Engenharia Civil da UFRGS, responsável pela sugestão de cursar Ciências Econômicas.

Aos amigos e colegas de mestrado, André, Eliezer, Bertussi, Andressa, Marcelo, Manuela, Alexandre, Sergio, Carlos. Em especial a Cristiane, hoje colega e professora da UCPel, companheira de inúmeras viagens durante o curso.

Ao Vagner, fundamental na parte inicial do trabalho, com os softwares aqui utilizados. Transformou-se num grande amigo, companheiro e entusiasta pelo trabalho. Com certeza que sem ele o caminho teria sido muito mais árduo e demorado.

Ao professor e pesquisador Dr. Glauber Acunha Gonçalves da FURG, pelos conhecimentos iniciais transmitidos em georreferenciamento e pelo auxílio na obtenção de milhares de coordenadas geográficas utilizadas nos dados desta pesquisa.

A Elisane, essencial na parte mais árdua e braçal deste trabalho quando me auxiliou no georreferenciamento de milhares de dados da amostra aqui utilizada.

A Ana, sempre à disposição e presente nos momentos mais difíceis desta jornada.

A Roberta, por estar ao meu lado, principalmente nestes momentos finais.

As professoras Lisbela e Maria Regina, pelas correções ortográficas e adaptações normativas realizadas neste trabalho.

À minha Camila, não um agradecimento, mas um pedido de desculpas, pelo período de ausência em que não pude participar de brincadeiras, de conversas e de troca de afeto.

Embora ausente, meu coração esteve sempre contigo.

Meu agradecimento especial vai para meus pais, Carlos Alberto (in memoriam) e Aura, meus grandes alicerces. Ele ausente, mas, que soube passar para mim o gosto pelo saber e ela presente que, com amor, entendeu minhas necessidades e soube cobrar de mim atitudes necessárias para garantir a finalização deste trabalho.

Certamente, sem ela nada teria sido possível.

Não se deve ir atrás de objetivos fáceis.

É preciso buscar o que só pode ser
alcançado por meio dos maiores esforços.

Albert Einstein

RESUMO

Este trabalho visa determinar a influência que os fatores sócio-econômicos, em conjunto com as variáveis construtivas usuais, provocam na valoração de imóveis no Estado do Rio Grande do Sul. Sendo o imóvel um bem com características distintas das demais, muitos pesquisadores buscam elementos diferentes para explicar esse comportamento distinto. Neste contexto, sobressai um dos fatores com grande relevância, a vizinhança. Os métodos inferenciais tradicionais dificultam a modelagem adequada pela multiplicidade dos fatores que influenciam o valor dos imóveis de uma determinada região. Sendo assim, os modelos de regressão espacial foram utilizados na estimação do valor unitário dos mesmos (VU). Já o modelo de regressão LAG foi utilizado para uniformizar a amostra de dados dos imóveis que se mostrou heterogênea. Não obstante, a krigagem demonstrou a estimativa do valor de um imóvel para determinada região. A aplicação dos métodos foi realizada para uma base de dados obtida junto a Caixa Econômica Federal, contendo imóveis transacionados no Estado do Rio Grande do Sul, no período de 2006 a 2008. Ademais, os métodos de regressão aplicados confirmaram índices fortemente significativos nos modelos obtidos para todos os imóveis de uma região. Com isso, a estrutura espacial dos índices estimados minimizou a autocorrelação existente nos resíduos do modelo de regressão, melhorando a confiabilidade da avaliação.

Palavras-chave: Econometria, Avaliações de Imóveis, Regressão Espacial.

ABSTRACT

This work aims to determine the influence that the social economic factors together with the usual variable constructive cause in the valuation of property, in Rio Grande do Sul State. Property is a material good with distinct characteristics from the other goods. Many researchers seek different elements to explain this situation. In this context one factor stands out with great relevance, the neighborhood. The traditional inferential methods difficult the adequate modeling because of the multiplicity of the factors that influence the value of the properties in a given region. Spatial regression models were used to estimate their unit value (UV). The regression model LAG was used to standardize the data sample of the properties, it was heterogeneous. Kriging showed the estimated value of a property for a given region. The application of the methods was performed for a database obtained from Caixa Econômica Federal, containing properties transacted in Rio Grande do Sul State, from 2006 to 2008. The applied regression methods confirmed strongly significant indices on the obtained models for all the properties in the region. The spatial structure of the estimated indices minimized the autocorrelation existing in the residuals of the regression models, improving the reliability of assessment.

Key-Words: Econometric, Evaluations of flats, Spatial Regression

LISTA DE FIGURAS

Figura 1: Gráfico de dispersão de Moran.	29
Figura 2: Brasil Político.....	43
Figura 3: Localização dos municípios estudados.	45
Figura 4: Distribuição geográfica dos imóveis analisados.	47
Figura 5: Exemplo de classificação de dados.	52
Figura 6: Origem dos dados ponderados.	53
Figura 7: Origem de médias e percentuais.	53
Figura 8: Cálculo do valor ponderado e valor aritmético.	54
Figura 9: Curvas de isovalor para a variável dependente Valor Unitário (VU).	57
Figura 10: Curvas de isovalor para a variável explicativa: Área Construída.	59
Figura 11: Análise de resíduos da variável VU.	65
Figura 12: Análise dos resíduos em mapas condicionais.	66

LISTA DE TABELAS

Tabela 1: Variáveis independentes, simbologia utilizada e efeito esperado	40
Tabela 2: Composição da amostra da base por tipo de imóvel	46
Tabela 3: Variáveis independentes, simbologia utilizada e sem efeito esperado.....	48
Tabela 4: Variáveis indicadoras de criminalidade por município.	50
Tabela 5: Melhor ajuste do modelo de regressão por regressão de mínimos quadrados ordinários – modelo I.....	60
Tabela 6: Teste de autocorrelação espacial do modelo de regressão – modelo I.....	60
Tabela 7: Significância das variáveis explicativas de regressão – modelo I.....	61
Tabela 8: Resultados da regressão no modelo LAG – modelo II.....	62
Tabela 9: Significância das variáveis explicativas da regressão no modelo LAG – modelo II	62

SUMÁRIO

1. INTRODUÇÃO	14
1.1. Definição do problema e justificativa	15
1.2. Objetivo geral	20
1.3. Método e delimitação do estudo	20
2. REFERENCIAL TEÓRICO	22
2.1. Aspectos teóricos	22
2.2. Estudos empíricos	24
3. REGRESSÃO ESPACIAL	27
3.1. Índice Global de associação espacial: Índice de Moran: I	28
3.2. Modelos de regressão espacial	30
3.3. Modelo de defasagem espacial	31
3.4. Modelo de erro espacial	32
3.5. Critério de Akaike	32
3.6. Critério de Schwarz	34
3.7. Krigagem	35
4. PROCEDIMENTOS METODOLÓGICOS	39
4.1. Modelo de regressão	39
4.2. Testes de validação da regressão espacial e do modelo de transbordamento espacial	41
4.3. Amostra, fonte e tratamento dos dados	42
4.3.1. Região de análise	42
4.3.2. Fonte dos dados	46
4.3.3. Tratamento dos dados	51
4.4. Matriz de pesos espaciais	54
4.5. Critério de Seleção do Modelo	55
4.6. Análise	56
4.6.1. Número de variáveis construtivas relevantes	63
4.6.2. Número de variáveis socioeconômicas relevantes	63
4.6.3. Análise dos resíduos	64

5.	CONCLUSÕES.....	68
6.	INDICAÇÃO PARA TRABALHOS FUTUROS	69
7.	REFERÊNCIAS	70

1. INTRODUÇÃO

O valor de um imóvel depende de uma ampla gama de variáveis como: atributos construtivos, fatores estruturais da obra, qualidade dos materiais; fatores locacionais, pólos de atração, efeitos de vizinhança, acessibilidade, políticas públicas, crédito dirigido para determinadas faixas de renda, investimentos em infra-estrutura; taxaço de tributos, entre outros. Esses elementos combinados geram uma complexidade no processo de estabelecimento do preço de mercado de um determinado imóvel.

Essa complexidade tem gerado, nos últimos anos, incremento dos procedimentos matemáticos para determinação do valor dos imóveis, como, por exemplo, uso de regressão espacial e análise envoltória de dados.

No Brasil, os primeiros trabalhos publicados de engenharia de avaliações foram em revistas técnicas do ramo, em São Paulo, nos anos de 1918 e 1919. Posteriormente, em 1941, o Engenheiro Luiz Carlos Berrini lançou os primeiros livros sobre o assunto versando sobre avaliações de terrenos e de imóveis.

Embora já com a primeira norma brasileira da ABNT, surgida em 1952, foi durante a década de 70 que a engenharia de avaliações começou a despertar mais atenção dos profissionais interessados. Desde então, a avaliação de imóveis, além das técnicas expeditas, começou a utilizar fatores de homogeneização determinísticos e fórmulas empíricas. Mesmo assim, a falta de segurança permanecia com relação à geração do valor.

Surge, então, na década de 80, a metodologia de pesquisa científica aplicada à engenharia de avaliações, utilizando regressão clássica, consolidando-se na década de 90 com a publicação da norma pertinente pela ABNT.

Mais recentemente, diversos autores têm evidenciado melhorias nas estimações com uso de regressão espacial e Krigagem em função de existência de autocorrelação espacial.

No entanto, essas evoluções nos processos estatísticos de avaliação não têm levado em conta alguns índices econômico-sociais. A análise da valoração dos imóveis com regressão espacial e inclusão destas variáveis, na explicação do preço dos imóveis, é o tema desta dissertação.

1.1. Definição do problema e justificativa

Segundo Dantas et al. (2007), a avaliação de um imóvel deve levar em conta: i) atributos decorrentes dos seus aspectos físicos como: área, padrão construtivo, número de vagas na garagem, etc.; ii) localização, tais como: local do imóvel, distância a pólos de influência, etc.; iii) , aspectos econômicos como: condições de pagamento do imóvel, natureza do evento: em oferta ou efetivamente vendido, etc.

Ao se avaliar a argumentação dos autores pode-se verificar que, em resumo, os fatores construtivos evidentemente entram no valor do imóvel, bem como sua localização, mas na visão econômica, são ressaltados apenas aspectos microeconômicos, isso é aqueles relacionados com a restrição orçamentária do consumidor.

Para González e Formoso (2000), os imóveis possuem um comportamento diferenciado economicamente de outros bens, em função dos efeitos de seus atributos específicos, especialmente: custo elevado, a heterogeneidade, a imobilidade e a durabilidade. Por outro lado, o mercado é atomizado, contando com a participação simultânea de muitos agentes, não coordenados. A combinação destes elementos permite explicar grande parcela das variações de preços.

A questão da heterogeneidade decorre do fato de que cada imóvel possui quantidades diferentes de cada um dos atributos valorizados pelo mercado. Por isso, são chamados de “bens compostos”, e a comparação entre eles exige a ponderação dos vários atributos de interesse (BALCHIN; KIEVE, 1986; ROBINSON, 1979).

Segundo Gonzalez (1997), “avaliar” é buscar o valor e o valor de um bem é essencialmente determinado pelo segmento de mercado onde ele pode ser transacionado. Embora existam muitas interpretações, o valor de mercado de um imóvel pode ser definido

como o valor mais provável que um dado imóvel pode atingir, numa transação normal, em determinadas condições econômicas. Em última análise, o valor de mercado trata de um equilíbrio microeconômico momentâneo, decorrente das condições de oferta e demanda, específicas e gerais, para um dado imóvel e mercado. O valor de mercado representa este “equilíbrio potencial”.

Neste contexto, para Moreira (apud GONZÁLEZ; FORMOSO, 2000), não havendo informações sobre transações (de razoável semelhança e em quantidade mínima para a análise), o analista pode utilizar alguns métodos que permitem a obtenção do valor indiretamente. São exemplos destes os chamados métodos da renda e evolutivo.

Pela antiga NBR-5676/89 – Norma Brasileira de Avaliações de Imóveis Urbanos, o trabalho avaliatório poderia ser classificado de acordo com os níveis alcançados em: expedito, normal, rigoroso e rigoroso especial. Os níveis, rigoroso e rigoroso especial sempre necessitavam de processos de inferência estatística, o nível normal tinha algum grau de subjetividade e o tratamento dispensado aos elementos era através de estatística descritiva, homogeneizando os dados por fatores ou ponderações empíricas e o nível expedito definia o trabalho avaliatório onde preponderava a subjetividade, onde não era utilizado qualquer instrumento matemático de suporte, nem a comprovação expressa dos elementos e critérios que conduziram a convicção do valor.

Embora esta norma estabelecida em 1989 exigisse inferência estatística nos trabalhos classificados como “rigorosos”, os procedimentos mais utilizados até 2004 eram de forma expedita, ou seja, mesmo que existisse uma norma que sinalizava para a melhoria nos processos de avaliação, esses só eram utilizados em poucos trabalhos avaliatórios.

Dois aspectos são relevantes para essa situação. Um refere-se ao custo de criar e manter um banco de dados que permitisse o uso de regressão para determinação do valor. O segundo a falta de softwares que fossem amigáveis e facilitassem os processos de estimação.

Com o estabelecimento da NBR 14653-2 de 2004 - que passou a exigir modelos econométricos para a inferência do valor dos imóveis nas avaliações – generalizou-se o uso desse procedimento, embora os modelos, em sua expressiva maioria, se restringissem a regressão múltipla pelo método dos mínimos quadrados ordinários com dados em corte.

Segundo Hayashi (2000), a questão é que esses modelos têm como preposição a não existência de correlação entre as variáveis explicativas e o erro da regressão. No entanto, os imóveis são altamente correlacionados em termos de valores no espaço.

Por exemplo, imóveis idênticos, localizados em bairros diferentes, terão elevada probabilidade de apresentarem valores diversos e esses serem decorrentes dos valores dos imóveis existentes na vizinhança. Esse aspecto é que tem sido o motivador da busca de outros métodos de estimação do valor dos imóveis.

Para Belsley et al. (1980) e Daniel e Wood (1980), os demais aspectos referentes as hipóteses de regressão e as características do mercado de imóveis não seriam fatores restritivos para o emprego da análise de regressão clássica, pois no caso de surgimento de outliers, há tratamentos bem documentados e a solução é razoavelmente fácil. A multicolinearidade, causada por inter-relações das variáveis explicativas, pode ser eliminada pelo emprego de técnicas que transformam os dados, segundo González (1993), Harmann (1976) e Maddala (1988), utilizando a análise fatorial. A heterocedasticidade pode ser contornada, conforme Judge et al.,(1985) e Neter et al.,(1990), com o emprego de mínimos quadrados generalizados, desde que sejam obtidas estimativas para a matriz de ponderação.

Entretanto, segundo González e Formoso (2000), para a autocorrelação do termo erro permanecem algumas dificuldades, inclusive fazem restrições às análises de regressão pela inexistência de um modelo teórico que permita definir a forma funcional da função de regressão.

Em função destes elementos, têm ocorrido diversas tentativas de estimar os valores dos imóveis com o apoio de outras metodologias.

Por exemplo, Baptistella (2005) utilizou técnicas de Redes Neurais Artificiais de forma alternativa a Regressão Linear Múltipla, na estimação dos valores venais de imóveis urbanos da cidade de Guarapuava/PR, concluindo serem ambas eficientes com maior precisão para o modelo com redes neurais.

Já Novaes e Paiva (2003) aplicaram Análise Envoltória de Dados (DEA) com o objetivo não só de realizarem a avaliação de imóveis, mas também, com o de comparar os resultados com aqueles encontrados com o uso de regressão linear múltipla. Em seu estudo, concluíram que a utilização de DEA pode ser particularmente vantajosa quando o avaliador se deparar com uma população de dados de mercado ou quando o objetivo do trabalho for a determinação de intervalos para o valor final do imóvel e este apresentar atributos próximos aos extremos observados na amostra.

Não obstante aos diferentes instrumentos utilizados, o método que tem sido mais empregado, recentemente, é o de regressão espacial. Esse, que teve seu impulso dado

fortemente por Anselin (1988, 1999) permite incluir nos modelos de regressão a influência do espaço e com isso elimina a autocorrelação das variáveis explicativas com o erro da função de regressão. Sendo que, segundo Trivellone (2005) “entre todas as variáveis que influenciam no valor dos imóveis, as variáveis ou fatores referidos à localização dos imóveis são as mais complexas de analisar e modelar.”

Embora, não exista teoria referente ao processo de modelagem da influência do espaço no valor dos imóveis, esta variável é de suma importância no estabelecimento do valor de mercado, como destaca Trivelloni (2005):

Os fatores de localização que participam da valorização de um imóvel são inúmeros: características 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.

Ainda, segundo Meen e Andrew (2003) há uma dicotomia entre os pesquisadores que modelam a partir de uma perspectiva macroeconômica ou financeira e aqueles que aproximam por uma perspectiva de economia urbana. Os primeiros dão ênfase particular nos aspectos de investimento de moradia e usam o modelo de vida-ciclo. Estes levam muito pouco em consideração a dimensão de espaço. Já na literatura de economia urbana, os aspectos de espaço são centrais como nas teorias de localização residenciais. Sendo a influência do espaço no valor dos imóveis captada pela modelagem de regressão espacial, então, esses modelos não só permitem estimar o valor dos imóveis, como também detectar a influência do efeito marginal das variáveis explicativas e, entre elas, a localização.

Já Holly *et al.* (2006) realizaram uma investigação empírica de dependência regional nos Estados dos EUA, trabalhando com dados sobre os rendimentos e o crescimento real dos preços dos imóveis. Foram analisados dados de 49 Estados americanos, excluindo Alasca e Havaí, ao longo de 29 anos, entre 1975 e 2003, e constataram evidências de que os preços reais dos imóveis são conduzidos por fundamentos de rendimento.

Recentemente, têm ocorrido avanços significativos nos procedimentos estatísticos de estimação do modelo de inferência. Talvez o mais relevante tenha sido a inclusão da análise espacial. Essa, segundo Trivelloni (2006), evidenciou a insuficiência das técnicas tradicionais de inferência para avaliação em massa, levando ao desenvolvimento de técnicas de estatística

espacial para o tratamento de dados imobiliários, cujo processo de avaliação imobiliária pode ser realizado por duas metodologias para o tratamento espacial dos dados: a econometria espacial e seus modelos de regressão espacial, por um lado, e o uso da geoestatística, por outro.

A autocorrelação espacial refere-se à redundância de informação entre duas realizações de um fenômeno quando elas ocorrem próximas uma da outra. Sua presença distorce os resultados obtidos pela aplicação de modelos estatísticos tradicionais.

González e Formoso (2000) concluem que a correlação espacial é um dos principais problemas estatísticos nas análises econométricas realizadas sobre o mercado imobiliário, que ainda não foi solucionado, e a busca de soluções a serem aplicadas em modelos hedônicos ou de outros métodos que permitam diminuir a dificuldade de estimação é importante para o aperfeiçoamento da avaliação de imóveis.

Outro problema é o da formatação do modelo, que afeta os pressupostos de linearidade da equação e de que as variáveis importantes tenham sido incluídas. Quais variáveis incluir e em que formato é um problema estatístico não-trivial e as revisões da literatura indicam estas dificuldades através da falta de uniformidade nos textos sobre economia urbana (BALL, 1973; SMITH et al., 1988).

Para White (1992), freqüentemente, a suposição de que o formato real do modelo sob análise seja conhecido não é correta e, neste caso, as estimativas dos parâmetros são inconsistentes, provocando suspeita sobre as inferências realizadas com a equação.

A complexidade de fatores como a proximidade de centros comerciais, de educação, de saúde e segurança pública, bem como os padrões construtivos e os polos ambientais desejáveis ou indesejáveis, provocados pela dinâmica urbana, produzem uma contínua mudança nos efeitos destes fatores sobre o valor dos imóveis como: investimentos públicos na região, modificações nas atividades comerciais locais; empreendimentos imobiliários; entre outros.

Na realidade, quase toda mudança urbana afeta o valor dos imóveis próximos, transformando a localização no fator mais dinâmico da sua valorização. Surgem as variáveis de vizinhança, espaciais ou de localização. Muitos desses fatores são de fato alterações de demanda que se refletem no preço dos imóveis e esses não são considerados nos modelos atuais, pois, partem da análise de que a demanda está dada para período determinado (normalmente um ano).

Outro trabalho interessante foi o desenvolvido por Marques *et al.* (2009) com o objetivo de compreender os fatores determinantes nos preços dos imóveis da cidade de Aveiro, através de um modelo hedônico de preços, incidindo sobre o tratamento da heterogeneidade e a dependência espacial nos dados. Primeiramente, concluíram que existe uma forte heterogeneidade espacial nas unidades espaciais analisadas, em termos de características físicas e localização. Entretanto, segundo a análise de dependência espacial (*spatiallag* e *spatialerror*), e, com base em várias matrizes de pesos, concluíram que não existe autocorrelação espacial nem na variável dependente, nem no termo de erro, confirmando que nem sempre a regressão espacial é a solução para a melhoria dos modelos. Com relação aos atributos explicativos para o preço dos imóveis, embora a localização seja importante, concluem que estes atributos são determinantes para a formação dos preços.

Tendo em vista a complexidade do processo de avaliação e dos avanços dos métodos de inferência e que esses ainda desconsideram o efeito demanda e variáveis sócio-econômicas no modelo, questiona-se: quais são os fatores determinantes da formação dos preços dos imóveis?

Essa é a questão central no qual esse trabalho se propõe responder.

1.2. Objetivo geral

Estimar os preços dos imóveis no Rio Grande do Sul com o objetivo de ampliar o entendimento dos fatores determinantes da formação de seus preços.

1.3. Método e delimitação do estudo

Utilizou-se, nesta pesquisa, a base de dados da Gerência de Desenvolvimento Urbano – RS, cedida pela Caixa Econômica Federal, que continha 20698 observações de imóveis que foram financiados pela mesma, no Estado do Rio Grande do Sul, durante os anos de 2006 a 2008.

Esta amostra possui dados de imóveis por municípios do Rio Grande do Sul, exceto Porto Alegre, excluída da mesma devido a diferenças significativas dos demais municípios. O tratamento destes dados pode ser observado no Capítulo 4.

O método utilizado para realizar a inferência dos fatores determinantes do valor dos imóveis foi o de regressão espacial com dados em corte. Utilizou-se, também, krigagem como procedimento para avaliar a distribuição espacial do erro e a defasagem espacial do valor dos imóveis.

Logo, os resultados desta pesquisa referem-se apenas aos imóveis comercializados nos municípios do interior do Estado do Rio Grande do Sul.

2. REFERENCIAL TEÓRICO

A valoração de imóveis tem gerado diversos procedimentos de modelagem para determinação dos valores de mercado dada a complexidade decorrente das características dos mesmos. O objetivo aqui é descrever a evolução das técnicas mais utilizadas pelos pesquisadores, com foco na regressão espacial que será utilizada nesta pesquisa.

Desta forma, estrutura-se primeiramente uma revisão dos estudos que buscaram formar um quadro teórico a respeito do mercado de imóveis e da formação de seu valor e após, faz-se uma análise dos trabalhos empíricos de avaliação de imóveis.

2.1 Aspectos teóricos

Essa seção foi baseada em Dantas (1998), todas as demais citações serão identificadas no parágrafo correspondente.

Grande parte dos bens públicos, particulares e empresariais do mundo, consistem em bens imóveis. Neste contexto, existe uma especialidade da engenharia chamada Engenharia de Avaliações reunindo vários conhecimentos desta área, além de outras áreas das ciências sociais, exatas e da natureza, para poder determinar tecnicamente o valor de um bem, de seus direitos, frutos e custos de reprodução, subsidiando tomadas de decisões a respeito de valores, custos e alternativas de investimentos, envolvendo bens de qualquer natureza, tais como: imóveis, máquinas e equipamentos, automóveis, móveis e utensílios, etc.

Segundo Abunahman (2000), avaliação é, pois, uma aferição de um ou mais fatores econômicos, especificamente definidos, em relação a propriedades descritas com data determinada, tendo como suporte a análise de dados relevantes.

Os métodos e técnicas de Engenharia de Avaliações permitem responder, de forma segura, questões que envolvam quaisquer tipos de bens ou decisões sobre investimentos, com comprovação científica, fugindo de suposições, opiniões subjetivas e empíricas. Esta área engloba análise de investimentos, análise de balanços, estatística básica, estatística inferencial, tecnologia da amostragem, matemática aplicada, matemática financeira, micro e macroeconomia, engenharia econômica, economia urbana, planejamento urbano, sociologia urbana, pesquisa social, econometria, álgebra linear, direito imobiliário, marketing e mercado de capitais, tornando esta área multidisciplinar uma das mais complexas espacialidades.

Conforme Trivellone (2005), o mercado imobiliário é de um oligopólio, geralmente, dominado por uma grande parte de agentes, empreendedores e incorporadores que concorrem entre si, tornando os preços praticados distantes do valor de mercado em função do seu grau de concentração, sendo composto por três elementos fundamentais: os bens levados a mercado, as partes desejosas em vendê-los e as partes interessadas em adquiri-los. Esta confusão entre valor de mercado e preço do imóvel ocorre devido a diversos fatores particulares da negociação, o que conduz o preço para cima ou para baixo do valor de mercado, sendo que na média tendem a convergir.

A plurivalente, que assume valor correlacionado à finalidade para o qual é avaliado, podendo assumir diversos valores. E a univalente, que conceitua o valor único para um bem em determinado instante de tempo, independente do fim a que este se destina.

O mercado imobiliário não é, por sua natureza, de concorrência perfeita, logo, apenas alguns segmentos aproximam-se deste como o mercado de locações, e outros que se distanciam como o de vendas de imóveis novos. Na prática, o máximo que se consegue é estimar o preço médio de mercado, através das amostras de preços que trazem todas as imperfeições do mercado, fazendo com que numa transação imobiliária o que se paga é o preço e não o valor.

Para Gonzalez (2000), o mercado imobiliário possui um comportamento diferenciado dos mercados de outros bens, devido a possuir inúmeras fontes de divergências e desigualdades entre os imóveis como: grande vida útil, fixação espacial, singularidade, elevado prazo de maturação e alto valor agregado, impossibilitando, assim, a comparação

direta e permitindo explicar grande parcela das diferenças de valores entre os imóveis, em um dado momento.

O mais utilizado é método comparativo de dados de mercado que relaciona dados assemelhados quanto às características intrínsecas e extrínsecas, desde que a amostra seja representativa. Nesta pesquisa, os dados devem conter os preços dos imóveis praticados no mercado, além de suas características físicas, locacionais e econômicas.

Os dados podem conter preços de transação ou oferta, sendo que esses últimos são menos consistentes, em virtude da ausência da parte compradora, orientando para um limite superior de preço que se pagaria pelo bem no mercado.

Na Engenharia de Avaliações, considera-se, geralmente, como variável dependente o preço praticado no mercado (total ou unitário) e, como variáveis independentes, as respectivas características físicas (área, frente, número de dormitórios, etc.), locacionais (bairro, logradouro, etc.) e econômicas (oferta ou transação, época, etc.).

Adota-se o modelo de regressão linear múltipla quando mais de uma variável independente é necessária para explicar a variabilidade dos preços praticados no mercado. Como o mercado imobiliário enquadra-se neste contexto, usualmente é utilizado este modelo de regressão, onde a estimação dos parâmetros é feita pelo método dos mínimos quadrados ordinários (MMQ).

2.2 Estudos empíricos

Esta seção apresenta alguns estudos empíricos relevantes sobre avaliações de imóveis.

Bráulio (2005) pesquisou um procedimento de construção de um modelo estatístico para avaliação de imóveis em função de suas características, utilizando técnicas de Análise Multivariada, em apoio às técnicas tradicionais de estatística, aplicando a Análise de Agrupamentos (*Cluster Analysis*) nos dados de imóveis para obtenção de classes homogêneas. Construiu um modelo de Regressão Linear Múltipla de cada classe de itens para cada tipo de imóvel (44 apartamentos, 51 casas e 24 terrenos) da cidade de Campo Mourão – PR, concluindo que cada classe destes três tipos de imóveis apresentou um bom ajuste aos dados,

uma boa capacidade preditiva e que a metodologia multivariada é viável, trazendo grande precisão e tratamento correto das diferenças.

Uma pesquisa realizada por Morcelli (2006), onde observou dados relativos da cidade de Santa Maria, desenvolvendo equações de regressão que tentaram explicar o valor de lotes urbanos, em função de variáveis comuns como: área, testada¹, profundidade média, valor do imposto, zona, subzona, valor da taxa de lixo, etc. Estas equações geraram resultados para todas as orientações (SE, SO, NE e NO), trazendo utilidade para questões de gestão da massa de dados e de estudos de valores de impostos. Utilizou, inicialmente, 23 variáveis e, posteriormente, apenas 8 ficaram no modelo final, identificando as variáveis que devem ser levadas em conta, quando do desejo de aplicação de equações de regressão à planta de valores..

Resende e Cypriano (2010) pesquisaram a valorização dos lotes urbanos na cidade de Toledo-PR, no período de 1998 a 2008, utilizando um modelo econométrico através do método dos mínimos quadrados ordinários (MMQ), estimando parâmetros que demonstraram significância e as elasticidades referentes à influência das variáveis renda per capita do município, distância dos lotes urbanos até a região das faculdades/universidade, distância dos lotes urbanos ao centro da cidade, tamanho dos lotes em metros quadrados, e a variável tendência, sendo uma variável que representa o tempo, sobre o preço dos lotes urbanos. Os resultados obtidos indicaram que essas variáveis influenciam significativamente nos preços dos lotes.

Canan (2005) buscou, em seu trabalho, o desenvolvimento de fatores de homogeneização para padrão de acabamento e estado de conservação, calculados através de metodologia científica, para residências localizadas na região central da cidade de Cascavel – PR, utilizando a inferência estatística, a partir de um modelo matemático de regressão, permitindo a adoção de fatores de homogeneização com isenção de subjetividade, transmitindo à avaliação uma maior confiabilidade do rigor estatístico. As variáveis qualitativas (padrão de acabamento e estado de conservação) influenciaram de forma positiva no valor final do imóvel, através do método comparativo de dados de mercado, permitindo a determinação do valor, levando em consideração as diferentes tendências e flutuações do mercado imobiliário. O estudo mostrou que é possível encontrar quaisquer fatores através da inferência estatística, para homogeneizar dados, bastando que sejam criados modelos isolados de regressão com a variável independente que se pretende encontrar o fator e a variável

¹ Medida frontal do terreno em metros.

dependente coletada. Com a dinâmica da economia, o banco de dados deve ser atualizado, pois, os imóveis sofrem alteração em seus valores no decorrer do tempo, crescem ou decrescem dependendo da situação e, com a adição de novos dados, o modelo de regressão será alterado. Com isso, cada atualização exigirá uma nova análise e um novo ajustamento da regressão, podendo, portanto, mudar a relação entre as variáveis independentes com a variável dependente, fazendo com que, periodicamente, se consigam os fatores de homogeneização para residências, localizados no centro da cidade pesquisada.

3. REGRESSÃO ESPACIAL

São formadas pelas atualizações das atuais ferramentas utilizadas nas análises de regressão espacial. Anselin (1996) definiu como ESDA (Análise Espacial Exploratória de Dados) e pode ser considerada continuação da EDA (Análise Exploratória de Dados) e conceituamos em três elementos básicos:

- Matriz de proximidade espacial (W): matriz quadrada que define a variação espacial de dados de área, onde cada elemento W_{ij} equivale a uma medida de proximidade entre A_i e A_j , sendo A_i e A_j as zonas analisadas;
- Vetor de desvios (Z): descoberto utilizando a média geral (μ), onde cada elemento é calculado subtraindo-se o valor de μ do valor do atributo (y_i) para cada região;
- Vetor de médias ponderadas (Wz): é definido pela medida útil para a realização do cálculo da variação da tendência espacial. Também conhecido por média dos valores dos vizinhos, ou média móvel espacial.

Segundo Câmara et al. (2002) a estimativa da média espacial é obtida pela equação

$$\hat{\mu}_i = \frac{\sum_{j=1}^n W_{ij} \times y_j}{\sum_{j=1}^n W_{ij}} \quad (1)$$

Onde:

$\hat{\mu}_i$ = médias ponderadas;

$\sum_{j=1}^n W_{ij}$ = matriz de proximidade espacial.

Segundo Câmara et al. (2002), um aspecto essencial na análise exploratória espacial é a caracterização da dependência espacial, o qual evidencia como os valores estão correlacionados no espaço.

3.1 Índice Global de associação espacial: Índice de Moran: I

O Índice de Moran fornece uma medida comum da associação espacial entre conjunto de dados. Seu valor varia de -1 a 1 . Caso estes forem valores próximos de zero, indicam a inexistência de autocorrelação espacial significativa entre os valores dos objetos e seus vizinhos. Logo, valores positivos para o índice indicam autocorrelação espacial positiva, ou seja, os valores do atributo de um objeto comprovam uma disposição a serem similares aos valores dos seus vizinhos. Já os valores negativos para o índice, advertem autocorrelação negativa (ANSELIN, 1996).

Segundo Florax e Graaff (2004), o índice de Moran é dado por:

$$I = \frac{n}{S_o} \cdot \frac{\hat{\epsilon}'W\hat{\epsilon}}{\hat{\epsilon}'\hat{\epsilon}} \quad (2)$$

Onde:

n : é o número de observações

S_o : é a soma dos elementos da matriz de pesos espaciais W , se os valores não estão normalizados.

$\hat{\epsilon}$: é o vetor $n \times 1$ dos resíduos decorrentes do Método dos Mínimos Quadrados Ordinários (MMQ)

W : é a matriz de pesos espaciais

Segundo Monastério e Ávila (2004), o valor uniformizado de uma variável em cada uma das abscissas e , no eixo das ordenadas, assim como a média do valor padronizado da mesma variável com os vizinhos, está representado com o gráfico de dispersão de Moran.

Ressalta-se que o Índice de Moran não especifica a forma de contagem espacial, apenas permite verificar a existência ou não do transbordamento espacial.

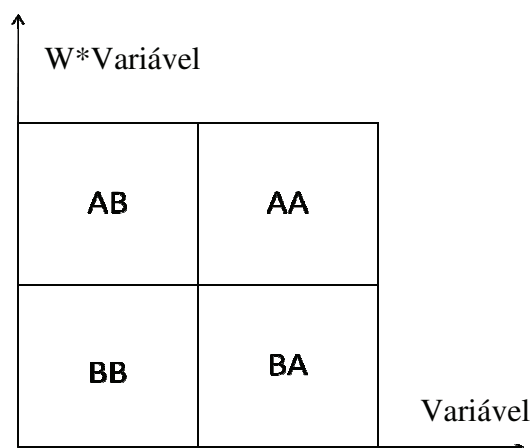


Figura 1: Gráfico de dispersão de Moran.

Fonte: elaborado pelo autor

Almeida (2004) ainda afirma que o gráfico está dividido em quatro quadrantes que correspondem a quatro diferentes padrões de associação local espacial, entre as regiões e seus vizinhos.

Primeiro quadrante (AA) Alto-Alto: Este quadrante apresenta regiões de altos valores para a variável em análise, assim como altos valores acima da média, para a mesma variável em questão.

Segundo quadrante (BA) Baixo-Alto: neste as regiões com valores baixos cercados por vizinhos que apresentam valores altos.

Terceiro quadrante (BB) Baixo-Baixo: ao contrário do primeiro, apresenta valores baixos para as variáveis em análise, e para as regiões que a cercam.

Quarto quadrante (AB) Alto-Baixo: é formado pelas regiões com altos valores para as variáveis em análise cercados por regiões de baixos valores.

Logo, as regiões AA e BB, respectivamente, primeiro e terceiro quadrante, demonstram autocorrelação espacial positiva, pois apresentam alto e baixo valores de variáveis em análise, sempre cercadas por alto e baixo valores.

Para indicar regiões onde a correlação local é demasiadamente diferente dos demais dados, é comum a utilização de mapas do tipo: LISA Map e Moran Map, pois tais regiões apresentam dinâmica espacial própria, isto é, áreas de não-estacionariedade, e necessitam de uma análise mais detalhada. Autocorrelação com significância inferior a 95% demonstra áreas com forte semelhança em relação a sua vizinhança.

O mapa bidimensional, do tipo diagrama de espalhamento, conhecido como Box Map, indica cada polígono, dentro do quadrante respectivo de espalhamento.

3.2 Modelos de regressão espacial

Normalmente, ao realizar uma análise de regressão, o objetivo é o de obter uma adequação entre os valores sugeridos pelo modelo e os valores observados da variável dependente. Também se deseja encontrar as variáveis explicativas que contribuem de forma expressiva para o relacionamento linear. A hipótese padrão é que as observações não são correlacionadas. Logo, os resíduos serão independentes e não correlacionados com a variável independente, também apresentando Distribuição Normal com média zero e variância constante.

Existem poucas possibilidades de que a hipótese padrão de observações não correlacionadas seja verdadeira em caso de dados, onde há presença de dependência espacial. Geralmente, os resíduos tendem a manter a autocorrelação espacial constatada nos dados, e pode ser verificada por diferenças regionais sistemáticas, ou também por uma tendência espacial contínua.

Assim, a análise dos resíduos da regressão, a procura de ocorrências de estrutura espacial, pode informar a necessidade da utilização de um modelo espacial. Ferramentas de análises gráficas e o mapeamento de resíduos podem fornecer a primeira indicação de que os valores observados estão mais correlacionados do que seria esperado sob uma condição de independência (FOTHERINGHAM et al., 2000). Além da análise gráfica, testes estatísticos podem auxiliar na verificação de existência da autocorrelação espacial nos resíduos da regressão, podendo ser utilizada a análise do índice I de Moran.

A análise de regressão em dados espaciais unifica, na modelagem, a dependência espacial entre os dados, auxiliando a escolha do melhor modelo a ser utilizado. Quando realizada a análise exploratória de dados, o objetivo é definir a estrutura de dependência nos dados, definindo a forma de incorporação dessa dependência ao modelo de regressão. Existem dois tipos básicos de modelagem que permitem incorporar o efeito espacial: as de forma

Global e as de forma Local (ANSELIN, 2002; CÂMARA et al., 2002 e FOTHERINGHAM et al., 2000).

Os modelos de forma Global, avaliados neste estudo, foram testados no software GeoDA, de domínio público e disponível em (<http://geodacenter.asu.edu>), sendo de dois tipos: Modelo de Defasagem Espacial e Modelo do Erro Espacial.

3.3 Modelo de defasagem espacial

Anselin (2002) define que no modelo de defasagem espacial, a autocorrelação espacial mascarada é atribuída à variável dependente Y. A dependência espacial é analisada através da adição, ao modelo de regressão, de um novo termo no formato de uma relação espacial com a variável dependente, e representada pela equação a seguir:

$$y = \rho WY + X\beta + \varepsilon \quad (2)$$

Onde:

y = variável dependente;

X = variáveis independentes;

β = coeficientes de regressão;

ε = erros aleatórios com média zero e variância s^2 ;

W = matriz de vizinhança espacial ou matriz de ponderação espacial;

ρ = coeficiente espacial autorregressivo.

A hipótese nula para a não existência de autocorrelação é que $\rho = 0$. O objetivo mínimo é incorporar a autocorrelação espacial como parte do modelo.

3.4 Modelo de erro espacial

Ainda o mesmo autor define o segundo tipo de modelo de regressão espacial com parâmetros globais, também referido como Modelo de Erro Espacial, e considera que os efeitos espaciais são um ruído, ou seja, necessitam ser excluídos. Assim, os efeitos da autocorrelação espacial são associados ao termo de erro $W\varepsilon$ e o modelo é representado pela equação.

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + \xi \end{aligned} \tag{3}$$

Onde:

$W\varepsilon$ = erros com efeito espacial;

ξ = erros aleatórios com média zero e variância s^2 ;

λ = coeficiente autorregressivo.

A hipótese nula para a não existência de autocorrelação é que $\lambda = 0$, logo, se não for possível rejeitar H_0 o termo de erro não é espacialmente correlacionado.

Câmara et al. (2002) comprovam que, na prática, a definição de utilização entre os dois tipos de modelos de regressão espacial com parâmetros globais é complexa, pois, apesar da diferença nas suas motivações, apresentam-se muito próximos em termos formais.

3.5 Critério de Akaike

Esta seção se baseia em Castelar *et al* (2003).

Sob o nome de "um critério de informação" (AIC), em 1971, e propôs em Akaike (1974), é uma medida da qualidade do ajuste de cerca de modelo estatístico. Ele se baseia no conceito de entropia, de fato oferecendo uma medida relativa da perda de informações quando um determinado modelo é usado para descrever a realidade e pode ser dito para descrever o

equilíbrio entre viés e variância na construção do modelo, ou vagamente falar que de precisão e complexidade do modelo.

A AIC não é um teste do modelo no sentido de testar a hipótese, mas um teste entre os modelos - uma ferramenta para seleção de modelos. Dado um conjunto de dados, vários modelos concorrentes podem ser classificados de acordo com suas AIC. A escolha fica a cargo do menor AIC, enquanto os de maiores serão automaticamente rejeitados.

No caso geral, a AIC é:

$$AIC = 2k - 2\ln(L) \quad (4)$$

Onde:

k é o número de parâmetros no modelo estatístico;

L é o valor maximizado do risco de função para o modelo estimado.

Durante o restante desta entrada, será assumido que os erros do modelo são normalmente e independentemente distribuídos. Deixe- n é o número de observações e a soma dos quadrados dos resíduos (RSS) é:

$$RSS = \sum_{i=1}^n \varepsilon_1^2 \quad (5)$$

Assumimos ainda que a variância dos erros do modelo seja desconhecida, mas igual para todos eles.

Maximizar a verossimilhança com respeito a essa variação torna-se o AIC

$$AIC = 2k + n[\ln\left(\frac{2\pi RSS}{n}\right) + 1] \quad (6)$$

Isso pode ser simplificado pela factoring o prazo $n * \ln(2\pi)$. Este é um termo constante adicionado ao valor do AIC de todos os modelos concorrentes. Portanto, não pode afetar a ordem em que classificá-las e nós podemos remover este prazo. Quando temos também o fator constante n , AIC simplifica para:

$$AIC = 2k + n[\ln(RSS)] \quad (7)$$

Aumentar o número de parâmetros livres a serem estimados melhora a qualidade do ajuste, independentemente do número de parâmetros livres no processo gerador dos dados. Daí a AIC não só recompensas bondade de ajuste, mas inclui também uma pena que é uma função crescente do número de parâmetros estimados. Esta penalidade

desencoraja overfitting. O modelo preferido é aquele com o menor valor de AIC. A metodologia da AIC tenta encontrar o modelo que melhor explica os dados com um mínimo de parâmetros livres. Em contrapartida, as abordagens tradicionais mais para começar a modelar a partir de uma hipótese nula. A AIC penaliza parâmetros livres menos fortemente do que o critério de Schwarz.

Logo o AIC um modelo de como fechar os seus valores tendem a ser equipados com os verdadeiros valores, em termos de um determinado valor esperado. Mas é importante perceber que o valor AIC atribuído a um modelo serve apenas para classificar os modelos concorrentes e dizer-lhe qual é o melhor entre as alternativas dadas. Os valores absolutos da AIC para modelos diferentes não têm significado, apenas diferenças em relação pode ser atribuído significado.

3.6 Critério de Schwarz

O Critério de Informação de Schwarz (BIC) é um resultado assintótico derivado sob a hipótese de que a distribuição dos dados está na família exponencial.

Onde:

x = os dados observados;

n = número de pontos de dados em x , o número de observações, ou equivalentemente, o tamanho da amostra;

k = número de parâmetros a serem estimados. Se o modelo estimado é uma regressão linear, k é o número de regressores, incluindo a interceptação;

$p(x|k)$ = a probabilidade dos dados observados, dado o número de parâmetros, ou, a probabilidade dos parâmetros dado o conjunto de dados;

L = o valor maximizado do log-likelihood para o modelo estimado.

A fórmula para o BIC é:

$$-2 \cdot \ln p\left(\frac{x}{k}\right) \approx BIC = -2 \cdot \ln L + k \ln(n) \quad (8)$$

Partindo do pressuposto de que os erros do modelo ou distúrbios são independentes e identicamente distribuídas segundo uma distribuição normal e que a condição de fronteira que a derivada da verossimilhança no que diz respeito à variação real é zero, isto torna-se (até uma constante aditiva, que depende apenas de n , e não o modelo):

(9)

Onde:

$\hat{\sigma}_e^2$ é a variância do erro.

A variância do erro neste caso é definida como:

$$- \quad \quad \quad 2 \quad \quad \quad (10)$$

Dados quaisquer dois modelos estimados, o modelo com o menor valor do BIC é o único a ser o preferido. O BIC é um aumento da função de $\hat{\sigma}_e^2$ e um aumento da função de k . Ou seja, variação não explicada na variável dependente e o número de variáveis explicativas aumentarem o valor do BIC. Assim, implica menor BIC quer menos variáveis explicativas, melhor ajuste, ou ambos. O BIC penaliza os parâmetros livres geralmente mais forte do que o critério de Akaike, mas isso depende do tamanho de n e a magnitude relativa de n e k .

É importante ter em mente que o BIC pode ser utilizado para comparar modelos estimados apenas quando os valores numéricos da variável dependente são idênticos para todas as estimativas estão sendo comparados. Os modelos que estão sendo comparados não precisam ser aninhados, ao contrário do caso quando os modelos estão sendo comparados com um F ou teste de verossimilhança.

3.7 Krigagem

Essa seção foi baseada em Landim (2003), todas as demais citações serão identificadas no parágrafo correspondente.

O processo conhecido como krigagem teve início na África do Sul, onde alguns pesquisadores, com destaque para o engenheiro de minas D. G. Krige e o estatístico H. S.

Sichel, que desenvolveram, de forma empírica, uma técnica de estimativa para o cálculo de reservas minerais. Esta metodologia recebeu o nome de *Geoestatística* para o estudo das chamadas *variáveis regionalizadas*, ou seja, variáveis com condicionamento espacial.

A análise geoestatística preocupa-se com a compreensão, dos fenômenos naturais que governam atuam sobre variáveis regionais, de sua estimativa, ou algumas de suas características espaciais, utilizando informações e relações, a partir de um conjunto discreto de amostras. Também existe uma preocupação com a avaliação dos erros de estimativa, de modo a estabelecer o grau de segurança em previsões e os padrões ótimos de amostragem, assegurando que um erro máximo de estimativa não seja excedido.

Inicialmente, sua aplicação era apenas para situações em geologia mineira na lavra e prospecção, mas, posteriormente, estendeu-se para outros campos, especialmente, nesses últimos anos, com aplicação em climatologia, econometria espacial, geologia ambiental, hidrogeologia, entre outros. Atualmente, quase todas as últimas versões de softwares para Confecção de Mapas ou Sistemas de Informações Georreferenciadas apresentam métodos geoestatísticos.

O termo geoestatística é um tópico especial da estatística aplicada que trata de problemas referentes às variáveis verdadeiramente aleatórios e àquelas totalmente determinísticas, por apresentarem uma aparente continuidade no espaço. Essas são representadas por funções numéricas ordinárias que assumem um valor definido a cada ponto no espaço e, matematicamente, descrevem um fenômeno natural.

A continuidade geográfica dada manifesta-se pela propriedade que a variável tem de apresentar valores muito próximos em dois pontos vizinhos e, paulatinamente, mais diferentes à medida que os pontos vão ficando mais distantes. Embora a variável regionalizada seja contínua no espaço, geralmente, não é possível conhecer os seus valores em todos os pontos, mas sim, apenas em alguns que foram obtidos por amostragem.

Quando se estuda o comportamento das variáveis regionalizadas é necessário lançar mão de duas ferramentas fundamentais dos métodos geoestatísticos: o *semivariograma* e a *krigagem*.

Os semivariogramas expressam o comportamento espacial da variável regionalizada ou de seus resíduos mostrando:

- o tamanho da zona de influência em torno de uma amostra;
- a anisotropia, quando os semivariogramas mostram diferentes comportamentos para diferentes direções de linhas de amostragem e de estudo da variável;
- a continuidade, pela forma do semivariograma, situação conhecida como efeito pepita, onde mede os erros de medição ou ao fato de os dados não terem sido coletados a intervalos suficientemente pequenos para mostrar o comportamento espacial subjacente do fenômeno em estudo.

A Krigagem é um processo de estimativa de valores de variáveis distribuídas no espaço, e/ou no tempo, a partir de valores adjacentes, enquanto considerados como interdependentes pelo semivariograma, sendo um método de estimativa por médias móveis. Com ela é viável encontrar a melhor estimativa possível para locais não amostrados, pela minimização da variância do erro, sem que se tenha total garantia que o mapa obtido pela krigagem tenha o mesmo semivariograma e a mesma variância que os dados originais, pois, trata-se, pela própria natureza do método, de um mapa com valores suavizados.

Segundo Cressie (1990), a krigagem pode ser usada para atender várias situações. O método fornece, além dos valores estimados, o erro associado a tal estimativa, distinguindo-a dos demais algoritmos à disposição. Esse é entendido como um estimador que se baseia numa série de técnicas de análise de regressão, sejam elas lineares ou transformações não-lineares, procurando minimizar a variância estimada, a partir de um modelo prévio que leva em conta a dependência estocástica entre os dados distribuídos no espaço. Com ela torna-se possível melhorar a estimativa para locais não amostrados, pela minimização da variância do erro.

Existem várias formas de krigagem, sendo as mais usuais: krigagem ordinária, krigagem universal, krigagem indicativa e cokrigagem.

Conforme Jakob (2002), os métodos tradicionais de análise de dados georreferenciados utilizam dados na forma vetorial com pontos, linhas e polígonos tentando representar as características físicas/geográficas da área de estudos. Quando se utiliza *softwares* como o Arcview na análise de dados tabulares, toda a unidade de análise assume o mesmo valor. Logo, se a análise for a nível municipal, cada um assumirá um determinado valor, e seu vizinho outro valor, podendo gerar enorme contraste entre estes, fazendo com que o mapa final possua grandes distorções. Para contornar este problema é melhor utilizar um método de análise baseado na interpolação, preservando os valores intermediários, gerando como resultado uma superfície contínua de dados mais suavizados e minimizando os

contrastes entre os polígonos. Neste caso, a Krigagem é considerada uma boa metodologia de interpolação de dados.

4. PROCEDIMENTOS METODOLÓGICOS

Este capítulo aborda os procedimentos metodológicos necessários para a obtenção dos resultados desejados. Descreve a obtenção dos dados analisados neste trabalho, desde o seu tratamento, seleção de variáveis significantes, cálculo do centróide dos municípios estudados, técnicas de krigagem e obtenção da matriz de peso espacial.

Foi utilizado o software OpenGeoDa para a realização das regressões espaciais testadas e obtenção da matriz de pesos espaciais. Para o mapeamento das curvas de isovalor e a geração da krigagem, foi empregado o software ArcGis 9.3.

4.1 Modelo de regressão

O modelo espacial, segundo Anselin (1999), pode ser expresso do seguinte modo:

$$VU = \rho W_{VU} + X\beta + \lambda W\varepsilon + \varepsilon \quad (11)$$

Onde:

VU é o vetor da variável dependente, definida como o quociente entre o valor total do imóvel pela sua área correspondente;

ρ é o coeficiente de autocorrelação espacial;

W a matriz de pesos espaciais, definida pelo método *Queen*;

X é a matriz das observações das variáveis independentes;

β é o vetor de parâmetros a serem estimados;

λ é o coeficiente de autocorrelação espacial;

u é o vetor de resíduos não correlacionados;

ε é o vetor de resíduos do modelo que se espera ser: $\varepsilon \sim N(0, \sigma^2 I)$.

A matriz X ficou composta pelas variáveis significantes, conforme tabela 1.

Destaca-se que diversas variáveis econômicas foram testadas para avaliar a capacidade de explicação do valor dos imóveis. Foi um processo exploratório no qual se compôs o modelo com um número amplo de variáveis e foram realizados diversos testes, redundante às que estavam no modelo, e variáveis omitidas, entre as que não estavam selecionadas, pelo teste da razão de máxima verossimilhança (LR). Ao mesmo tempo, foram avaliando a estabilidade dos parâmetros estimados pelo teste de Wald². As variáveis utilizadas nesse processo e que não compuseram o modelo final, por não serem significantes, constam na Tabela 3.

Tabela 1: Variáveis independentes, simbologia utilizada e efeito esperado

Simbologia	Variáveis	Efeito Esperado
ACAB_POND	Acabamento – refere-se ao padrão de acabamento conforme os materiais aplicados na construção.	$\partial(VU)/\partial(ACAB_POND) > 0$
ACONST_POND	Área Construída – é a quantidade medida em metros quadrados do total da área do imóvel.	$\partial(VU)/\partial(ACONST_POND) < 0$
AG_CAIXA	Caixa Econômica - é uma instituição financeira que tem a finalidade de estimular a poupança popular, aplicando os depósitos recebidos e os demais recursos em operações de crédito que visem à promoção social e ao bem estar da população.	$\partial(VU)/\partial(AG_CAIXA) > 0$
BANH_POND	Banheiros – número de banheiros existentes no imóvel.	$\partial(VU)/\partial(BANH_POND) > 0$
CONSERV_PO	Conservação – corresponde ao estado de conservação do imóvel obedecendo as condições ali presentes.	$\partial(VU)/\partial(CONSERV_POND) > 0$
DEN_DEM	Densidade demográfica – corresponde ao valor obtido pela relação entre a população e área do município. Medida em habitantes por quilometro quadrado.	$\partial(VU)/\partial(DEN_DEM) > 0$

²O teste R1 é dado por $LR = 2(l_{ir} - l_r) \sim \chi^2$ (mgl). Onde LR é o teste da razão de máxima verossimilhança; l é o log da função de verossimilhança restrita (r) e irrestrita (ir); LR segue uma distribuição qui-quadrado com o número de graus de liberdade igual ao número restrições impostas ao modelo (Gujarati, 2006). O teste de Wald (W_t), testa restrições sobre os parâmetros estimados e é calculado, supondo $y = X\beta + \varepsilon$, com restrições lineares $R\beta = r$, por: $W_t = \frac{(R\hat{\beta} - r)' [R(X'X)^{-1}R']^{-1} (R\hat{\beta} - r) / q}{(e'e)/(n-k)} \sim F(q, n-k)$. W_t segue uma distribuição F e a

definição dos graus de liberdade dada na equação dependem de q , o número de restrição, e $n-k$, número de observações menos o número de parâmetros estimados (SOARES; CASTELAR, 2003).

DOMIC	Domicílio – locais estruturalmente separados e independentes que se destinam a servir de habitação a uma ou mais pessoas, ou que esteja sendo utilizado como tal.	$\partial(VU)/\partial(DOMIC) < 0$
FROTA	Frota de Veículos registrados - diz respeito a todos os veículos automotores, elétricos, articulados, reboque ou semi-reboque registrados perante o órgão executivo de trânsito do Estado ou do Distrito Federal, no município de domicílio ou residência de seu proprietário, na forma da lei.	$\partial(VU)/\partial(FROTA) > 0$
HOMIC	Homicídios – é a quantidade de homicídios ocorridos e registrados.	$\partial(VU)/\partial(HOMIC) > 0$
HOSPITAIS	Hospital - é uma instituição prestadora de serviços médico-hospitalares-ambulatoriais, com leitos e instalações apropriadas ao desempenho de suas atividades, e que conta com, pelo menos, um médico e a equipe de enfermagem.	$\partial(VU)/\partial(HOSPITAIS) > 0$
PAV_POND	Pavimentação – corresponde ao tipo de pavimento utilizado no local do imóvel.	$\partial(VU)/\partial(PAV_POND) > 0$
PIB_PER	PIB per capita - é o Produto Interno Bruto a preço de mercado dividido pela população.	$\partial(VU)/\partial(PIB_PER) > 0$
TERR_POND	Terreno – é a quantidade medida em metros quadrados do total da área do terreno	$\partial(VU)/\partial(TERR_POND) > 0$

4.2 Testes de validação da regressão espacial e do modelo de transbordamento espacial

O teste utilizado para avaliar o uso de um modelo de regressão espacial foi o Índice de Moran (*I-Moran*). Como este teste não especifica o tipo de relação espacial, então os testes para verificar quais dos efeitos espaciais utilizados no modelo, se de defasagem ou do erro, ou ambos, foram os Multiplicadores de Lagrange (*LM*).

O teste de Moran confirma ou não a presença de autocorrelação espacial. Se p-valor do for significativo, ou seja, igual ou inferior a 5%, há efeito espacial envolvido, dado que o mesmo é um teste conduzido sob hipótese nula de inexistência de autocorrelação espacial, conforme Florax e Graaff (2004).

Já a definição em relação à escolha do modelo, isto é, defasagem espacial ou erro espacial, é feita pelos testes dos Multiplicadores de Lagrange (*LM*), que são: Multiplicador de Lagrange de Defasagem Espacial (*LM-lag*) e Multiplicador de Lagrange do Erro Espacial (*LM-error*). Estes possuem duas estatísticas que são o LM e LM-robusto. Desta forma, quando LM-lag (simples ou robusto) for significativo, ou seja, inferior a 5%, há um indicativo a favor do modelo de defasagem espacial (ρWVU) e se o LM-error (simples ou robusto) também for significativo, ou seja, inferior a 5%, há um indicativo a favor do modelo do erro espacial ($\lambda W\varepsilon$), (ANSELIN, 2005).

Como os testes LM foram todos significantes (conforme tabela 6), então optou-se por aquele com maior significância que foi o de defasagem espacial.

4.3 Amostra, fonte e tratamento dos dados

Nesta seção, descreve-se a amostra utilizada na pesquisa e o tratamento de dados realizado. Desta forma, tem-se, inicialmente, a descrição da região de análise, e sua composição em termos de observações. Após, tem-se as variáveis testadas no modelo e que não foram significativas, bem como suas fontes de dados. Por fim, tem-se a estruturação dos dados para utilização no modelo descrito na seção 4.1.

4.3.1 Região de análise

A região de análise foi o Estado do Rio Grande do Sul, que possui 496 municípios (IBGE, 2000) e está localizado no extremo sul do Brasil, conforme pode ser observado na Figura 2.

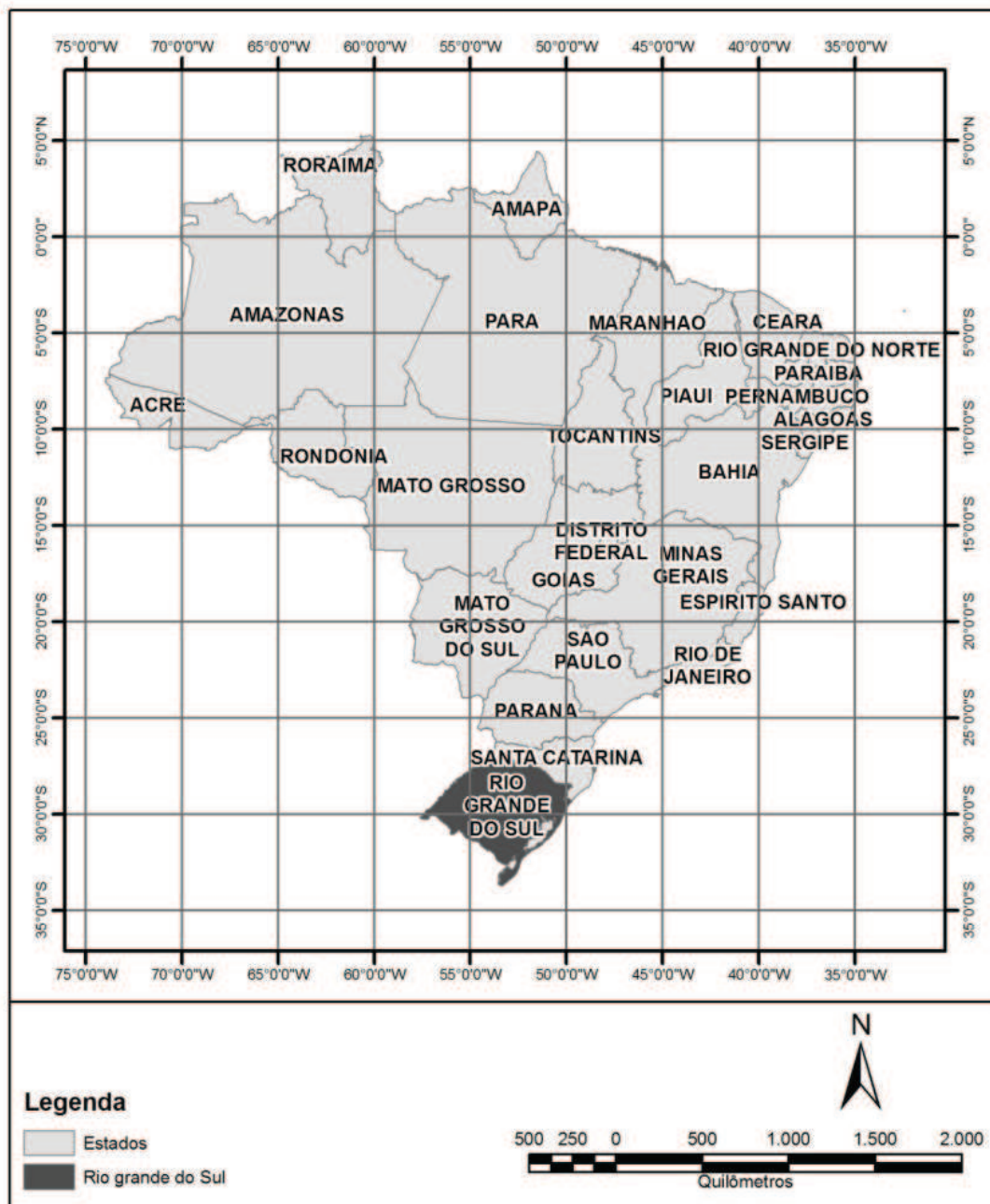


Figura 2: Brasil Político
Fonte: elaborado pelo autor

A partir dos 496 municípios do Estado do Rio Grande do Sul, retirou-se Porto Alegre da amostra, pois em decorrência de sua dimensão, os imóveis possuem um grau de variabilidade dos valores significativamente elevado, cuja variabilidade é decorrência da localização interna dos imóveis neste município. Evidentemente que os imóveis de outras cidades também têm seus valores influenciados pelo bairro em que se localizam, mas espera-se que, em termos médios, esses valores se distingam de outros municípios em função de

variáveis macroeconômicas e da localização do mesmo. Considerar esse valor médio para Porto Alegre crê-se que se estaria gerando uma distorção de tal magnitude que ela refletiria valores irrealistas. Essa percepção ocorre na Caixa Econômica Federal (CEF), através da Gerência de Desenvolvimento Urbano (GIDUR/CEF), pois, quando a mesma analisa dados de imóveis e valores de financiamento, trabalha com dois bancos de dados. Um só para Porto Alegre e outro para os demais municípios do Estado. Isso já é um indicativo da existência de diferenças de valores nas variáveis analisadas.

Cabe destacar que a amostra fornecida pela CEF não continha dados de todos os municípios do Rio Grande do Sul, basicamente, em função da ausência de transação financiada pela CEF no mesmo. Foram 206 municípios em que não houve informação. Desta forma, dos 496 municípios do Estado, restaram 290. Ainda, após a análise dos primeiros resultados, decidiu-se por retirar da amostra os municípios de Canoas e de Triunfo, devido à existência, no primeiro, de uma refinaria de petróleo e, no segundo, de um pólo petroquímico. Estas atividades nestes municípios geram um valor adicionado muito superior aos demais e são valores que não permanecem de forma significativa no município, gerando valores irrealistas para o PIB per capita. Em função destas atividades, presentes nestes municípios, o valor adicionado é muito superior aos dos demais. Estes valores acabam não permanecendo de forma significativa no município, gerando valores irrealistas para o PIB per capita. Em virtude desta distorção é que se decidiu pela retirada dos mesmos.

Em resumo, ficaram 288 municípios analisados. Sua distribuição espacial pode ser visualizada na Figura 3.

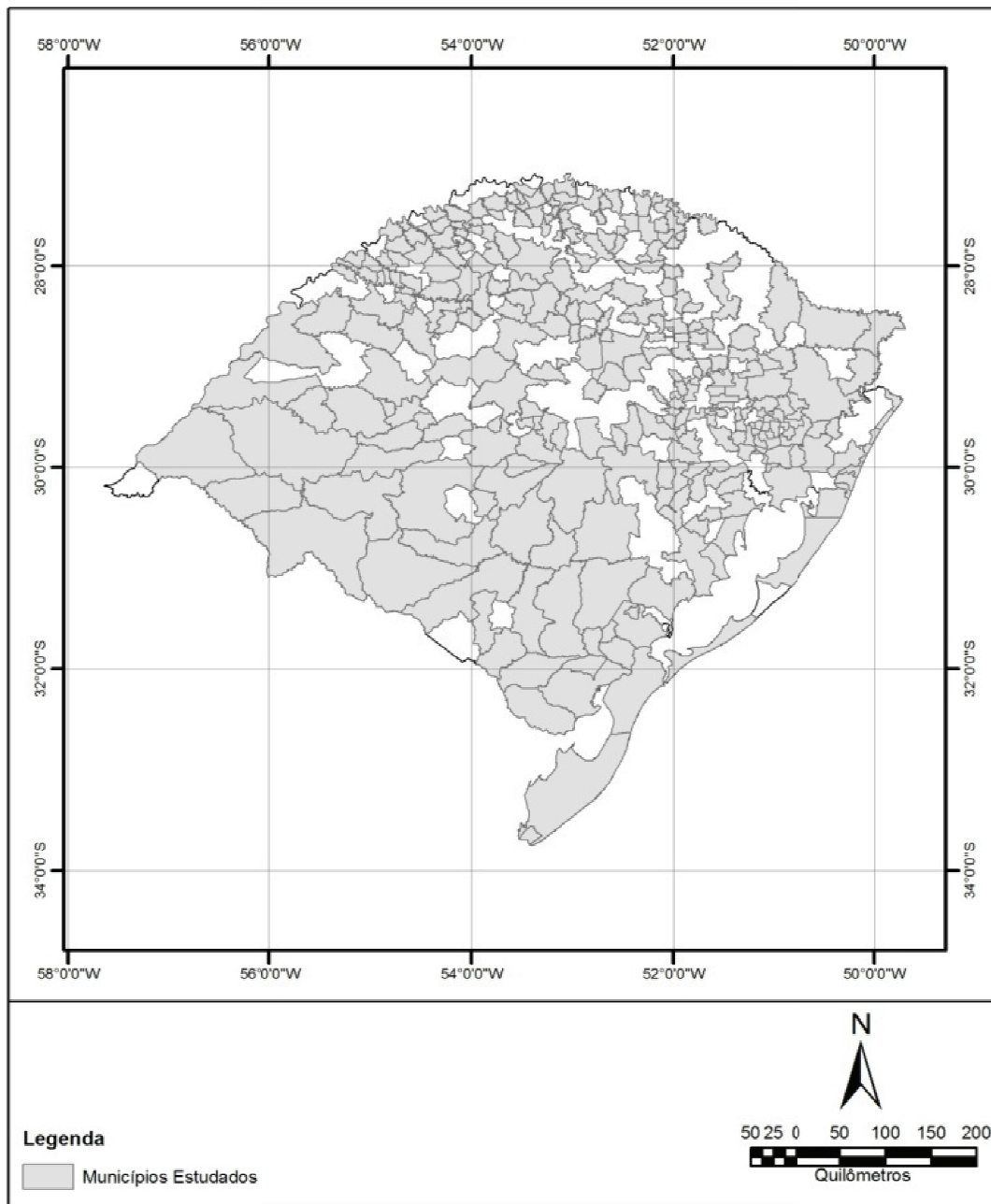


Figura 3: Localização dos municípios estudados.
Fonte: elaborado pelo autor

Ao se observar a Figura 3, pode-se verificar que a amostra utilizada nesta pesquisa contém uma cobertura espacial significativamente ampla do Rio Grande do Sul e não está concentrada em uma região específica.

4.3.2 Fonte dos dados

A pesquisa utilizou-se de dados de imóveis financiados pela Caixa Econômica Federal, cedidos pela Gerência de Desenvolvimento Urbano do Rio Grande do Sul – GIDUR, no período compreendido entre os anos de 2006 e 2008, em um total de 20.698 imóveis. O Anexo A contém os municípios e o número de observações em cada um deles, bem como o valor médio dos imóveis em cada.

Posteriormente, foram catalogados os dados e fez-se uma análise lógica da coerência dos mesmos e se verificaram diversas inconsistências como: áreas exorbitantes em relação ao número de dormitórios e ao valor total, valores exorbitantes em relação à área do imóvel, ausência de dormitórios em alguns dados, etc.

Assim, decidiu-se por retirar essas observações aberrantes de forma que dos 20.698 imóveis restaram 19.203, compostos por 7.457 apartamentos e 11.746 casas, como pode ser visto na **Tabela 2**.

A mostra a distribuição percentual dos imóveis da amostra por tipo.

Tabela 2: Composição da amostra da base por tipo de imóvel

Tipo	Ocorrências	% sobre o total
Apartamento	7.457	38,83
Casa	11.746	61,17
Total	19.203	100%

Fonte: elaborada pelo autor com base nos dados da pesquisa.

Às informações dos imóveis acrescentaram-se dados de latitude e longitude, a fim de permitir a vinculação das informações socioeconômicas e construtivas a sua região geográfica de origem. Figura 4 apresenta a distribuição dos imóveis em termos espaciais no Estado do Rio Grande do Sul.

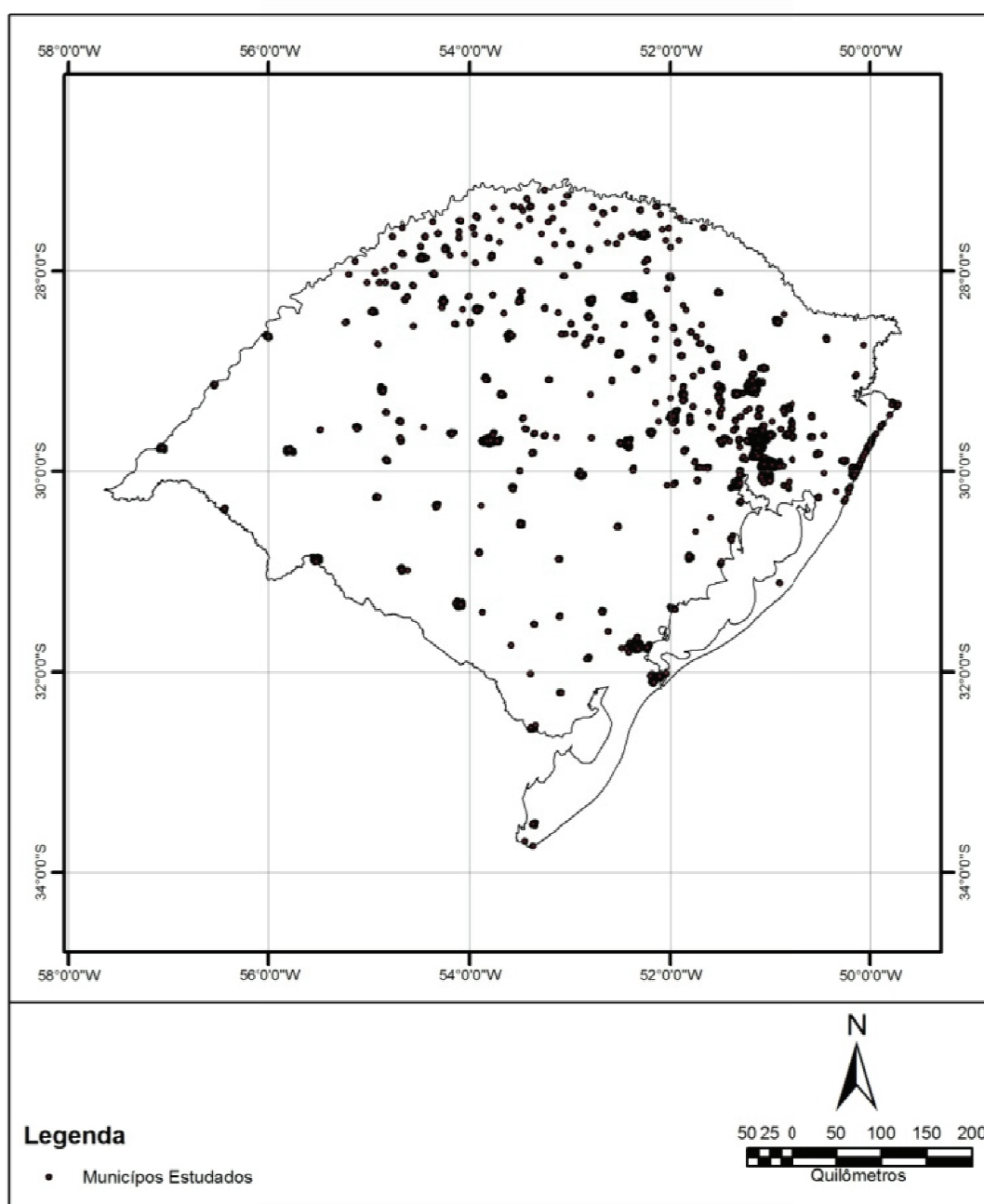


Figura 4: Distribuição geográfica dos imóveis urbanos analisados.
Fonte: Elaborado pelo autor

Conforme mencionado anteriormente (na apresentação das variáveis explicativas dadas na Tabela 1), as variáveis socioeconômicas, avaliadas como possíveis fatores de determinação do valor dos imóveis e que não se apresentaram significativas encontram-se na Tabela 3. Essas foram testadas no modelo de regressão, compondo a matriz X , e sua exclusão do modelo final deu-se, como já descrito, pelo teste de máxima verossimilhança e pelo teste de Wald. Na Tabela 3, tem-se, além da variável, a denominação utilizada e o sinal que se

esperava quando de sua inclusão no modelo. Como essas variáveis não foram relevantes, suprimiu-se do modelo final, considerando um nível de significância de 5%. A fonte dos dados encontra-se no final da tabela.

Tabela 3: Variáveis independentes, simbologia utilizada e sem efeito esperado.

Variáveis	Descrição	Sinal esperado
AGEN_BANC	Banco comercial - é uma instituição financeira bancária, privada ou pública, constituída sob a forma de sociedade anônima e especializada em operações de curto e médio prazo.	$\partial(VU)/\partial(AGEN_BANC) > 0$
CASAM	Casamento - é o processo pelo qual se constitui uma relação legal entre o homem e a mulher, podendo ser civil ou religioso com efeito civil. Os dados referem-se ao lugar de registro.	$\partial(VU)/\partial(CASAM) > 0$
COMBUS	Combustíveis – valor referente a quantidade em litros dos total de combustíveis automotores de qualquer espécie.	$\partial(VU)/\partial(COMBUS) > 0$
DIVORC	Divórcio - é a dissolução legal do casamento e que confere às partes o direito a novo casamento civil ou religioso. Os dados referem-se ao lugar da ação do processo.	$\partial(VU)/\partial(DIVORC) < 0$
DORM_POND	Dormitórios – é o número de dormitórios existentes no imóvel.	$\partial(VU)/\partial(DORM_POND) < 0$
ED_SUP	Ensino superior – é uma instituição de ensino superior, privada ou pública, especializada em graduação ou pós-graduação.	$\partial(VU)/\partial(ED_SUP) > 0$
ELEITORES	Eleitores - são os brasileiros maiores de 18 anos de idade, de ambos os sexos, que estão obrigados ao exercício do voto, e os maiores de 16 anos, que podem optar pelo exercício do voto.	$\partial(VU)/\partial(ELEITORES) > 0$
EXP_VIDA	Expectativa de Vida ao Nascer - número médio de anos que se espera que as pessoas vivam a partir do nascimento.	$\partial(VU)/\partial(EXP_VIDA) > 0$
EXPOR_R\$	Exportações - refere-se ao valor de venda das mercadorias exportadas, acrescido de todas as despesas assumidas pelo exportador até colocá-las a bordo, incluindo as taxas portuárias, de Previdência, da Comissão de Marinha Mercante e outras que incidem sobre o valor do frete.	$\partial(VU)/\partial(EXPOR_R\$) > 0$
ICMS	Imposto sobre circulação de mercadorias e sobre prestação de serviços de transporte e de comunicações - é um tributo estadual, incidente sobre a entrada ou a saída de mercadorias, bem como sobre o início dos serviços de transporte interestadual ou intermunicipal, as comunicações e o fornecimento de alimentação, bebidas e outras mercadorias, incluídos os serviços.	$\partial(VU)/\partial(ICMS) > 0$
IDESE_GERAL	Índice de desenvolvimento socioeconômico (IDESE) - é um índice sintético que abrange um conjunto amplo de indicadores sociais e econômicos com o objetivo de mensurar	$\partial(VU)/\partial(IDESE_GERAL) > 0$

	o grau de desenvolvimento dos municípios do Estado. O IDESE é resultado da agregação, com a mesma ponderação (0,25), de quatro blocos de indicadores: Domicílio e Saneamento, Educação, Saúde e Renda. Cada um dos blocos, por sua vez, resulta da agregação de diferentes variáveis.	
INDUST	Indústria – valor referente ao nível de Produção Industrial no município.	$\partial(VU)/\partial(INDUST) > 0$
IPI	Imposto sobre Produtos Industrializados - é o imposto que incide sobre produtos industrializados, nacionais e estrangeiros, obedecidas as especificações constantes da Tabela de Incidência do Imposto sobre Produtos Industrializados - TIPI (Lei n.º 4.502, de 30 de novembro de 1964, art. 1º, e Decreto-lei n.º 34, de 18 de novembro de 1966, art. 1º).	$\partial(VU)/\partial(IPI) > 0$
IPTU	Imposto predial e territorial urbano - é um tributo municipal, incidente sobre a propriedade, o domínio útil ou a posse de bem imóvel localizado na zona urbana do município.	$\partial(VU)/\partial(IPTU) < 0$
IPVA	Imposto sobre propriedade de veículos automotores - é um tributo estadual, tendo como fato gerador a propriedade de veículos automotores	$\partial(VU)/\partial(IPVA) > 0$
IRPJ	Imposto de Renda Pessoa Jurídica - é a arrecadação classificada conforme o município do domicílio fiscal da empresa. Abrange todas as formas de tributação da Pessoa Jurídica: lucro real, lucro presumido e lucro arbitrado.	$\partial(VU)/\partial(IRPJ) > 0$
ISSQN	Imposto sobre serviços de qualquer natureza - é um tributo municipal, incidente exclusivamente sobre serviços relacionados ao trabalho e a atividades legalmente consideradas como de prestação de serviços por empresas ou profissionais autônomos.	$\partial(VU)/\partial(ISSQN) > 0$
ITBI	Imposto sobre a transmissão de bens imóveis - é um tributo estadual, incidente sobre a transmissão de bens imóveis e de direitos a ele relativos. É gerado por contrato de compra e venda, doação e ato de arrematação ou transmissão por óbito. A Constituição de 1988 extinguiu este imposto, criando o ITCO no âmbito estadual e o ITBI no municipal. A presença do ITBI estadual após 1988 é explicada pela existência de valores referentes a fatos geradores anteriores à referida alteração.	$\partial(VU)/\partial(ITBI) > 0$
ITIBI	Imposto sobre transmissão "inter vivos" de bens imóveis e de direitos reais sobre imóveis - é um tributo municipal, incidente sobre a transmissão a qualquer título, por ato oneroso, de bens imóveis e de direitos reais sobre os mesmos, exceto os de garantia, bem como sobre a cessão de direito.	$\partial(VU)/\partial(ITIBI) > 0$
LEITOS	Leito - refere-se ao número de internações máximas que o hospital consegue atender, dispondo de acomodações básicas para o paciente.	$\partial(VU)/\partial(LEITOS) > 0$
OBITOS	Óbito - diz respeito ao desaparecimento definitivo dos sinais de vida em qualquer momento posterior ao nascimento. O registro refere-se ao lugar de residência do falecido.	$\partial(VU)/\partial(OBITOS) < 0$

PIB	Produto Interno Bruto a preço de mercado - é igual ao Valor Adicionado a preço básico, subtraído dos serviços de intermediação financeira, indiretamente medidos e adicionado dos impostos sobre produtos, líquidos de subsídios.	$\partial(VU)/\partial(PIB) > 0$
POP_RURAL	População rural - abrange pessoas moradoras fora dos limites das áreas urbanas, inclusive nos aglomerados rurais (povoados, núcleos, etc.).	$\partial(VU)/\partial(POP_RURAL) > 0$
POP_URBANA	População urbana - corresponde às pessoas moradoras nas cidades, vilas ou áreas urbanas isoladas, cujos limites são definidos por lei municipal.	$\partial(VU)/\partial(POP_URBANA) > 0$
POPUL_TOTAL	População total – corresponde a abrangência da população do município, sendo resultado da soma da população urbana e a população rural.	$\partial(VU)/\partial(POPUL_TOTAL) > 0$
SEP_JUD	Separação judicial - é a dissolução legal da sociedade conjugal, sem permitir às partes o direito a novo casamento civil ou religioso. Os dados referem-se ao lugar da ação do processo.	$\partial(VU)/\partial(SEP_JUD) < 0$

Fonte: FEE – Fundação de Economia e Estatística

As variáveis criminais relacionadas aos municípios que foram adicionadas à base de dados seguem na Tabela 4, a seguir:

Tabela 4: Variáveis indicadoras de criminalidade por município.

Variáveis	Descrição	Sinal Esperado
ARMA_MUNI	Armamento e Munição	$\partial(VU)/\partial(ARMA_MUNI) < 0$
CORRUPC	Corrupção	$\partial(VU)/\partial(CORRUPC) < 0$
ENTOR_POSSE	Entorpecente – Posse	$\partial(VU)/\partial(ENTOR_POSSE) < 0$
ENTOR_TRAF	Entorpecente – Tráfico	$\partial(VU)/\partial(ENTOR_TRAF) < 0$
ESTELIO	Estelionato	$\partial(VU)/\partial(ESTELIO) < 0$
EXT_M_SEQ	Extorsão Mediante Seqüestro	$\partial(VU)/\partial(EXT_M_SEQ) < 0$
EXTOR	Extorsão	$\partial(VU)/\partial(EXTOR) < 0$
FUR_VEIC	Furtos de Veículos	$\partial(VU)/\partial(FUR_VEIC) < 0$
FURTOS	Furtos	$\partial(VU)/\partial(FURTOS) < 0$
LATROC	Latrocínios	$\partial(VU)/\partial(LATROC) < 0$
ROUB_VEIC	Roubo de Veículos	$\partial(VU)/\partial(ROUB_VEIC) < 0$
ROUBOS	Roubos	$\partial(VU)/\partial(ROUBOS) < 0$
TT_DELITOS	Total de Delitos	$\partial(VU)/\partial(TT_DELITOS) < 0$

4.3.3 Tratamento dos dados

O software Open Geoda, apresenta certas restrições devido ao volume de dados. Este tem capacidade de processamento de apenas 999 observações, utilizando matriz de pesos espaciais. Diante dessa limitação, decidiu-se por construir uma informação para cada município com base nos valores médios ponderados dos imóveis comercializados naquele município. Essa informação denominou-se de imóvel representativo. Com esse procedimento, transformou-se a amostra em 288 observações. Se esse procedimento sofre a limitação de conter valores médios por região, perdendo graus de liberdade, por outro lado tem como vantagem o fato de eliminar valores distorcidos.

A partir do imóvel representativo, então, tem-se em cada município analisado, os dados que representam os valores únicos, de um imóvel central em um município, ou seja, um imóvel que apresente dados ponderados de todas as variáveis e possa representar a média ponderada destes imóveis em um único ponto georeferenciado.

Da Figura 5 até a Figura 8, tem-se a apresentação de um exemplo dos cálculos realizados para a ponderação do valor dos imóveis. Essas informações referem-se à ponderação feita para o município de Agudo, onde aparece a tipologia e o valor de cada dado. Utilizando recursos do Microsoft Excel, pode-se ver a lógica do cálculo de um valor ponderado simples, seguidos pelas exigências do software, como descrito a seguir no exemplo da variável *Valor_do_Imóvel*.

A seqüência adotada foi de:

- Classificação dos dados simplificados por ordem crescente, utilizando como primeira diretriz o campo Municípios, seguido por Tipo e a seguir por Valor do Imóvel na ordem Decrescente, exemplificada na **Erro! Fonte de referência não encontrada.**;

Município	Tipo	Valor do Imóvel
Agudo	apartamento	R\$ 105.000,00
Agudo	apartamento	R\$ 52.000,00
Agudo	apartamento	R\$ 45.000,00
Agudo	apartamento	R\$ 38.500,00
Agudo	casa	R\$ 130.000,00
Agudo	casa	R\$ 105.000,00
Agudo	casa	R\$ 90.000,00
Agudo	casa	R\$ 85.000,00
Agudo	casa	R\$ 80.000,00
Agudo	casa	R\$ 65.000,00
Agudo	casa	R\$ 65.000,00
Agudo	casa	R\$ 65.000,00

Figura 5: Exemplo de classificação de dados.

Fonte: elaborado pelo autor

- Calcular o total por **Tipo** de todos os imóveis considerando o **Tipo**. Este total contém dois valores, um valor total para apartamento (Soma_Apartamento) e outro para valor total para casa (Soma_Casa) no mesmo município (passo 1);
- Contar a quantidade por **Tipo** de todos os imóveis considerando o **Tipo**. Este total contém dois valores, um valor total para apartamento (Apartamento_Contar) e outro para casa (Casa_Contar) no mesmo município (passo 2);
- Obter o **Total_Geral** através da soma dos valores do total por tipo, somando Soma_Apartamento com Soma_Casa (passo 3);
- Do mesmo modo, obter o **Contar_Geral**, somando as quantidades por tipo, somando Apartamento_Contar com Casa_Contar (passo 4);
- Dividindo os valores encontrados em **Total_Geral** pelo **Contar_Geral**, obtemos o valor **Media_por_Tipo**, conforme demonstrado na **Erro! Fonte de referência não encontrada.** (passo 5);

Município	Município por Tipo	Tipo	Valor do Imóvel	Média por Tipo	Soma do Valor Apto e Valor Casa	Percentual Apto e Casa	Valor Ponderado Casa / Apto	Valor Ponderado Final por Município	Valor Aritimético Final por Município
Agudo	Agudo apartamento	apartamento	R\$ 38.500,00						
Agudo	Agudo apartamento	apartamento	R\$ 45.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 52.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 105.000,00						
	Agudo apartamento Contar	4	R\$ 240.500,00	R\$ 60.125,00		41,25%	R\$ 24.802,85		
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 80.000,00						
Agudo	Agudo casa	casa	R\$ 85.000,00						
Agudo	Agudo casa	casa	R\$ 90.000,00						
Agudo	Agudo casa	casa	R\$ 105.000,00						
Agudo	Agudo casa	casa	R\$ 130.000,00						
	Agudo casa Contar	8	R\$ 685.000,00	R\$ 85.625,00	R\$ 145.750,00	58,75%	R\$ 50.302,85	R\$ 75.105,70	R\$ 72.875,00

Figura 6: Origem dos dados ponderados.

Fonte: elaborado pelo autor

- A *Média por tipo* foi calculada para apartamentos e para casas de um mesmo município. A soma das duas médias mostrará a *Soma do Valor Apto e Valor Casa*, representado na Tabela 6 (passo 6);
- O campo *Percentual Apto e Casa* é o resultado da divisão entre a *Soma do Valor Apto e Valor Casa* pela *Media por Tipo* de apartamentos e a *Soma do Valor Apto e Valor Casa* pela *Media por Tipo* de casas, também representado na **Erro! Fonte de referência não encontrada.** (passo 7);
- O campo *Percentual Apto e Casa* é o resultado da divisão entre a *Soma do Valor Apto e Valor Casa* pela *Media por Tipo* de apartamentos e a *Soma do Valor Apto e Valor Casa* pela *Media por Tipo* de casas, da mesma forma representado na **Erro! Fonte de referência não encontrada.** (passo 8);

Município	Município por Tipo	Tipo	Valor do Imóvel	Média por Tipo	Soma do Valor Apto e Valor Casa	Percentual Apto e Casa	Valor Ponderado Casa / Apto	Valor Ponderado Final por Município	Valor Aritimético Final por Município
Agudo	Agudo apartamento	apartamento	R\$ 38.500,00						
Agudo	Agudo apartamento	apartamento	R\$ 45.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 52.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 105.000,00						
	Agudo apartamento Contar	4	R\$ 240.500,00	R\$ 60.125,00		41,25%	R\$ 24.802,85		
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 80.000,00						
Agudo	Agudo casa	casa	R\$ 85.000,00						
Agudo	Agudo casa	casa	R\$ 90.000,00						
Agudo	Agudo casa	casa	R\$ 105.000,00						
Agudo	Agudo casa	casa	R\$ 130.000,00						
	Agudo casa Contar	8	R\$ 685.000,00	R\$ 85.625,00	R\$ 145.750,00	58,75%	R\$ 50.302,85	R\$ 75.105,70	R\$ 72.875,00

Figura 7: Origem de médias e percentuais.

Fonte: elaborado pelo autor

- O *Valor_Ponderado_Casa/Apto* obtido pelo produto do *Percentual_Apto_e_Casa* pela *Média_por_Tipo* deve ser calculado para casas e apartamentos;
- Finalizando o *Valor_Ponderado_Final_por_Município*, é a soma dos dois valores encontrados no campo *Valor_Ponderado_Casa/Apto*;
- O *Valor_Aritmético_Final_por_Município* é a média aritmética entre os valores da *Média_por_Tipo*.

Município	Município por Tipo	Tipo	Valor do Imóvel	Média por Tipo	Soma do Valor Apto e Valor Casa	Percentual Apto e Casa	Valor Ponderado Casa / Apto	Valor Ponderado Final por Município	Valor Aritmético Final por Município
Agudo	Agudo apartamento	apartamento	R\$ 38.500,00						
Agudo	Agudo apartamento	apartamento	R\$ 45.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 52.000,00						
Agudo	Agudo apartamento	apartamento	R\$ 105.000,00						
	Agudo apartamento Contar	4	R\$ 240.500,00	R\$ 60.125,00		41,25%	R\$ 24.802,85		
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 65.000,00						
Agudo	Agudo casa	casa	R\$ 80.000,00						
Agudo	Agudo casa	casa	R\$ 85.000,00						
Agudo	Agudo casa	casa	R\$ 90.000,00						
Agudo	Agudo casa	casa	R\$ 105.000,00						
Agudo	Agudo casa	casa	R\$ 130.000,00						
	Agudo casa Contar	8	R\$ 685.000,00	R\$ 85.625,00	R\$ 145.750,00	58,75%	R\$ 50.302,85	R\$ 75.105,70	R\$ 72.875,00

Figura 8: Cálculo do valor ponderado e valor aritmético.

Fonte: elaborado pelo autor

Aplicar o mesmo procedimento para as variáveis construtivas, qualitativas ou quantitativas, do imóvel, como: Área Construída, Terreno, CUn, CUBs, VUn, Dormitórios, Banheiros, , Conservação, Acabamento, Pavimentação e Esgoto.

4.4 Matriz de pesos espaciais

A matriz de pesos espaciais é um dos aspectos determinantes no estudo do efeito espacial do mercado da habitação.

Segundo Anselin (2005), esta matriz é usualmente baseada na contiguidade espacial (*rook* e *queen*) ou em distâncias métricas (*distancedecay*). Estas matrizes possuem, em cada célula da diagonal principal, valores iguais a zero e cada célula W_{ij} igual a zero se i não for vizinho de j ; de outra forma assumem o valor 1.

Conforme Anselin (1999), o operador de defasagem espacial é considerado como uma média ponderada de variáveis aleatórias relacionadas a unidades geográficas vizinhas, sendo fundamental a definição do conjunto de vizinhança para cada unidade geográfica. Este autor recomenda que os elementos que compõem a matriz W sejam não estocásticos e exógenos (gerados fora do modelo), evitando problemas de identificação no modelo e auxiliando na obtenção de estimadores consistentes e assintoticamente normais.

Os w_{ij} 's são elaborados baseados na relação de contigüidade existente entre as unidades geográficas, referindo-se as seguintes definições: “*rookcontiguity*”, “*bishopcontiguity*”, “*double linear*”, “*doublerook*” e “*queencontiguity*”. Também existem outras alternativas de pesos espaciais baseadas no inverso da distância ou no inverso do quadrado da distância.

A matriz de peso espacial encontrada neste estudo foi gerada pelo método “*queen*” a partir dos valores unitários ponderados a fim de escolher o modelo que melhor resultados demonstrariam pela vizinhança contigua dos municípios.

Sabe-se que os municípios possuem áreas muito diferentes e os municípios vizinhos podem estar com fronteiras próximas ou afastadas, prejudicando a utilização dos métodos baseados em distância. Já o método “*queen*” possibilita detectar todos os contados que possuem borda com a fronteira existente em todas as direções.

4.5 Critério de Seleção do Modelo

Uma vez determinado o modelo de contágio espacial, que foi de defasagem espacial, a composição entre os modelos analisados para determinação do modelo final, deu-se através de três estatísticas.

A primeira estatística foi o coeficiente de determinação (R^2). Este coeficiente é dado pela seguinte fórmula:

$$R^2 = \frac{(\sum(x_i - \bar{x})(y_i - \bar{y}))^2}{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2} \quad (11)$$

O coeficiente efetivamente utilizado foi o coeficiente de determinação ajustado (R^2 ajustado).

A justificativa de sua adoção foi decorrente de que sempre quando incluímos inúmeras variáveis, mesmo que tenham muito pouco poder explicativo sobre a variável resposta, aumentarão o valor de R^2 , o que é ruim pois incentiva a inclusão indiscriminada de variáveis, indo contra o princípio da parcimônia.

Este coeficiente ajustado, conforme fórmula (12), evita a penalização quando da inclusão de variáveis pouco explicativas, mantendo o mesmo critério do coeficiente de determinação originário onde quanto maior o seu valor melhor.

$$\bar{R}^2 = 1 - \frac{n-1}{n-(k+1)}(1 - R^2) \quad (12)$$

Onde:

$(k - 1)$ representa o número de variáveis explicativas mais a constante.

Note que a inclusão de mais variáveis com pouco poder explicativo prejudica o valor do R^2 ajustado porque aumenta k em uma unidade e não traz muito incremento em R^2 .

A segunda e a terceira estatísticas foram, respectivamente, pelos critérios de Akaike e de Schwartz. Conforme Castelar *et al* (2003), quanto menores os valores encontrados, melhor a qualidade do modelo.

4.6 Análise

Analisando, em termos espaciais, a distribuição dos valores unitários dos imóveis (VU), que é a variável dependente do modelo a ser estimado, percebe-se que existe concentração de valores em determinadas regiões e que esses se reduzem a partir destes pontos de concentração. Essa distribuição, realizada a partir da krigagem, onde se estabeleceu as curvas de isovalor, pode ser observada na **Erro! Fonte de referência não encontrada.**

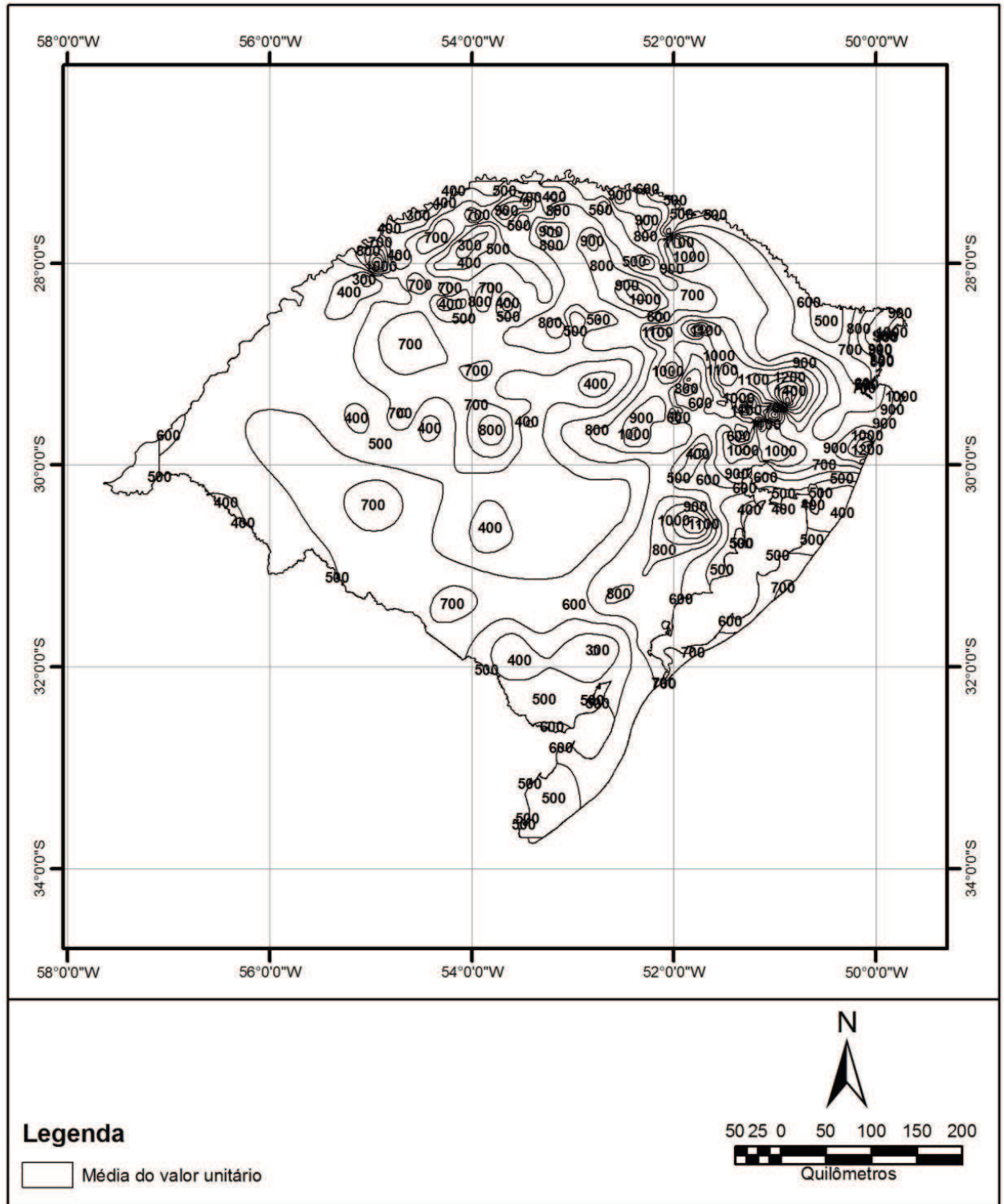


Figura 9: Curvas de isovalor para a variável dependente Valor Unitário (VU).

Fonte: elaborado pelo autor

Os valores mais altos encontram-se na região da Serra, confirmando o esperado, uma vez que essa é uma região turística conhecida em nível nacional, como Gramado e Canela, e fortemente industrializada como o caso de Caxias do Sul e Bento Gonçalves. A partir destes focos de elevado valor, há uma redução do VU gerando contornos bem definidos e não

uniformes em que os VUs estão se reduzindo, quando se movimenta em direção ao sul do Estado até as regiões fronteiriças.

Áreas com VUs de 700 e 500 são bastante amplas na metade sul e se estreitam, quando se observa sua incidência para o centro e norte do Estado do Rio Grande do Sul. Ainda, em áreas contidas ou contíguas a essas, percebem-se focos de depressão de valores como os VUs de 400.

A análise dos formatos das curvas de isovalor para o VU permite antecipar que a hipótese de existência de autocorrelação espacial deva ser confirmada quando do teste de **I-Moran**. Porém, essa hipótese pode ser irreal, se as variáveis construtivas mantiverem o mesmo padrão de distribuição espacial. Neste sentido, um dos elementos mais significativos, em termos de custos e de valoração do imóvel é sua área construída. Assim, realizou-se a construção das curvas de isovalor para essa variável, como mostra a Figura 10.

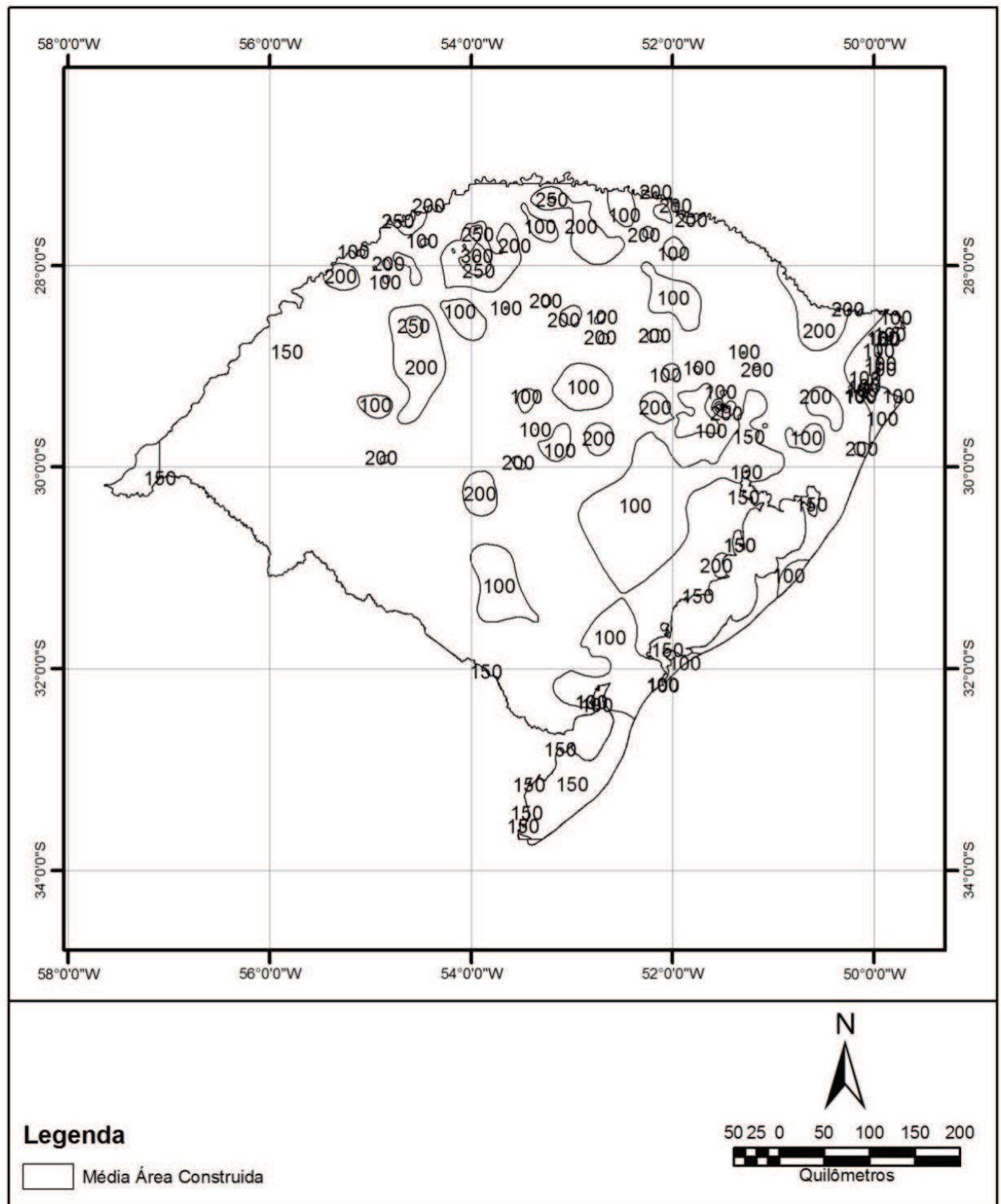


Figura 10: Curvas de isovalor para a variável explicativa: Área Construída.

Fonte: elaborado pelo autor

Nessa percebe-se que os valores mais altos encontram-se na região Noroeste do Estado. A região da serra gaúcha, não apresentou o destaque nas maiores médias de área construída. Essa informação corrobora para o fortalecimento da hipótese de autocorrelação

espacial, uma vez que os valores de área construída apresentam um formato no espaço significativamente diferente dos valores unitários. Logo, o VU deverá ter influência da região de localização mantido o controle das demais variáveis explicativas.

Para confirmar a existência ou não de autocorrelação espacial e com isso validar a regressão espacial (Anexo B), fizeram-se diversas regressões com a inclusão de variáveis construtivas e socioeconômicas, conforme descrito no capítulo 3. Para cada uma delas, mesmo contendo variáveis redundantes, calculava-se o I-Moran. Nesse processo de depuração do modelo, obteve-se o resultado dado na Tabela 5, 6 e 7.

Tabela 5: Melhor ajuste do modelo de regressão por regressão de mínimos quadrados ordinários – modelo I

VARIÁVEL DEPENDENTE	VU
R ²	0.453701
R ² A	0.427782
F	17.5044
PROB (F)	<0.0001
LOG VEROSSIMILHANÇA	-1863.75
CRITÉRIO AKAIKE	3755.5
CRITÉRIO SCHWARZ	3806.78
ERRO PADRÃO DA REGRESSÃO	156.399

Tabela 6: Teste de autocorrelação espacial do modelo de regressão – modelo I

TESTE	PROBABILIDADE
MORAN'S I (ERROR)	0.0000539
LM (LAG)	0.0000533
LM ROBUSTO (LAG)	0.0643163
LM (ERRO)	0.0001729
LM ROBUSTO (ERRO)	0.2734961

Tabela 7: Resultados da Regressão Clássica para o modelo I

VARIÁVEL	COEFICIENTE	ERRO PADRÃO	ESTATÍSTICA-T	PROBABILIDADE
CONSTANTE	-117.7098	95.22571	-1.236113	0.2174757
ACAB_POND	171.8044	32.50796	5.284994	0.0000003
ACONST_PON	-1.34712	0.3949554	-3.410816	0.0007453
AG_CAIXA	61.73802	25.57798	2.413717	0.0164468
BANH_POND	81.41429	32.73306	2.487219	0.0134699
CONSER_PO	128.4538	30.31026	4.237966	0.0000309
DEN_DEM	0.03477071	0.03192862	1.089014	0.2771035
DOMIC	-0.008704205	0.003300537	-2.637209	0.0088365
FROTA	0.006476919	0.002085742	3.10533	0.0021000
HOMIC	3.348134	2.55062	1.31267	0.1903916
HOSPITAIS	7.69578	25.177	-0.3056671	0.7600898
PAV_POND	44.18212	31.03985	1.4234	0.1557576
PIB_PER	0.005387097	0.002251402	2.392774	0.0173952
TERR_POND	0.180552	0.0575395	3.13788	0.0018874

Analisando os resultados da regressão, dois elementos são relevantes. O primeiro refere-se ao valor do I-Moran. Seu p-valor é altamente significativo, inferior a 1%. Logo, a regressão clássica gerará resultados viesados, uma vez que o erro não terá distribuição normal em função de uma variável relevante ter sido omitida do modelo. Assim, a regressão espacial poderá contribuir ao inserir o efeito de vizinhança no mesmo. Logo, não se analisarão os resultados da regressão para esse modelo, mas sim para o realizado com regressão espacial.

O segundo resultado refere-se aos testes dos multiplicadores de Lagrange. Por esses, ambos os modelos são significativos, com um p-valor mais expressivo para modelo de defasagem espacial. Assim, tem-se um indicativo de que ambos os modelos são viáveis, mas com uma pequena indicação a favor do modelo de defasagem espacial.

Desta forma, reestimou-se a regressão anterior, utilizando como matriz de pesos espaciais a relação de vizinhança dos valores unitários, isso é, W_{VU} . Observou-se, então, o p-valor associado ao parâmetro da variável e os indicadores de Akaike e Schwarz. O modelo que resultou em melhor estimativa, contendo apenas variáveis significantes, estimado por regressão espacial, encontra-se na Tabela 9, extraída do anexo C.

Tabela 8: Estatística da regressão espacial por defasagem espacial – modelo II

VARIÁVEL DEPENDENTE	VU
R ²	0.495306
LOG VEROSSIMILHANÇA	-1853.73
CRITÉRIO AKAIKE	3737.46
CRITÉRIO SCHWARZ	3792.41
ERRO PADRÃO DA REGRESSÃO	150.325

Tabela 9: Resultados das variáveis explicativas da regressão espacial por defasagem espacial – modelo II

VARIÁVEL	COEFICIENTE	ERRO PADRÃO	VALOR	PROBABILIDADE
W_VUN_POND	-1.216643	0.005993549	-202.992	0.0000000
CONSTANTE	487.2929	153.8812	3.166682	0.0015420
ACAB_POND	237.4804	52.53094	4.520771	0.0000062
ACONST_PON	-1.910317	0.6386188	-2.991327	0.0027778
AG_CAIXA	139.0533	41.35365	3.362541	0.0007724
BANH_POND	87.2119	52.92876	1.647722	0.0994096
CONSER_PO	138.0597	48.98461	2.818431	0.0048260
DEN_DEM	0.1317794	0.05159465	2.554129	0.0106454
DOMIC	-0.0211343	0.005333374	-3.962651	0.0000742
FROTA	0.01159433	0.003370483	3.43996	0.0005819
HOMIC	16.13492	4.121771	3.914561	0.0000906
HOSPITAIS	77.32497	40.68569	-1.900545	0.0573616
PAV_POND	84.57635	50.15796	1.6862	0.0917573
PIB_PER	0.01108821	0.003638073	3.047825	0.0023052
TERR_POND	0.1917788	0.09297919	2.0626	0.0391506

O R² melhorou na comparação do modelo clássico com o espacial, passando de 0,45 para 0,50. Registre-se aqui que o R² ajustado do modelo clássico era de 0,43 e não foi possível compará-lo com o do modelo espacial em virtude do software utilizado não apresentá-lo, como mostra o anexo C.

Da mesma forma, os critérios de Akaike e Schwarz se reduziram, evidenciando um modelo melhor para os dados com regressão espacial, como demonstra a Tabela 8, com origem no anexo C.

4.6.1 Número de variáveis construtivas relevantes

Das “7” variáveis construtivas analisadas, as significativas foram: padrão de acabamento, área construída, número de banheiros, estado de conservação, pavimentação urbana e área do terreno.

Destas, apenas a área construída teve um comportamento inversamente proporcional ao valor unitário, confirmando a correlação esperada para as mesmas. As demais tiveram um comportamento diretamente proporcional ao valor unitário, como também já era esperado.

A área construída relaciona-se inversamente com o valor unitário, diferentemente da área total que possui uma relação direta.

As outras variáveis construtivas têm a relação direta, pois melhor acabamento, mais banheiros, melhor conservação, mais área de terreno e pavimentação, valorizam o imóvel.

4.6.2 Número de variáveis socioeconômicas relevantes

Das “45” variáveis socioeconômicas analisadas, as significativas foram: número de agências da CEF, densidade demográfica, número de domicílios, frota urbana, número de homicídios, número de hospitais e PIB per capita.

Destas, apenas o número de domicílios teve um comportamento inversamente proporcional ao valor unitário, confirmando que uma maior oferta de imóveis no mercado provoca uma redução no valor unitário dos mesmos. As demais tiveram um comportamento diretamente proporcional ao valor unitário.

O PIB per capita era esperado, uma vez que uma renda maior deveria levar a maiores gastos com residência e a avaliação marginal dos valores gastos com imóveis mais valorizados deve ser maior em relação àqueles de menor renda.

Uma maior frota urbana é decorrente do desenvolvimento do município e, por conseqüência, da valorização imobiliária refletida no VU., o mesmo ocorrendo com a densidade demográfica.

A presença de hospitais nas cidades funciona como polo de atração urbana, resultando num aumento da demanda por imóveis e provocando uma relação direta com o VU.

O número de homicídios revelou uma relação direta com o VU, possivelmente devido a elevação da insegurança pública em alguns locais, provocando um aumento na demanda por imóveis em regiões mais seguras, valorizando-os.

A existência de agências da Caixa Econômica Federal relaciona-se, diretamente, com o VU, crê-se que isso ocorra, basicamente, em função do aumento da disponibilidade e facilidade ao crédito imobiliário, sendo essa uma informação relevante para avaliação das políticas públicas de Habitação realizadas pelo Governo Federal

4.6.3 Análise dos resíduos

No gráfico representativo das isolinhas da variável VU_Pond, confrontado com a análise de resíduos, pode-se comprovar que os valores esperados para cada região sofrem variações proporcionais à variação do próprio VU, sendo possível identificar os locais onde ocorrem maiores discrepâncias entre os valores esperados e os valores observados.

Assim sendo, observa-se que os imóveis situados nas regiões onde não ocorrem resíduos, as isolinhas representam proximidade entre o valor esperado do VU e o seu valor observado.

Ao contrário, é possível localizar várias ilhas com resíduos positivos e negativos, onde os valores observados e estimados dos VUs estão distantes. Conclui-se que, os imóveis nestas regiões foram comercializados com valores acima ou abaixo do esperado, como exemplo o ocorrido na região da Serra Gaúcha.

Note-se que as isolinhas representam o valor observado dos imóveis transacionados naquela região e a escala de cores representa a análise de resíduos, conforme

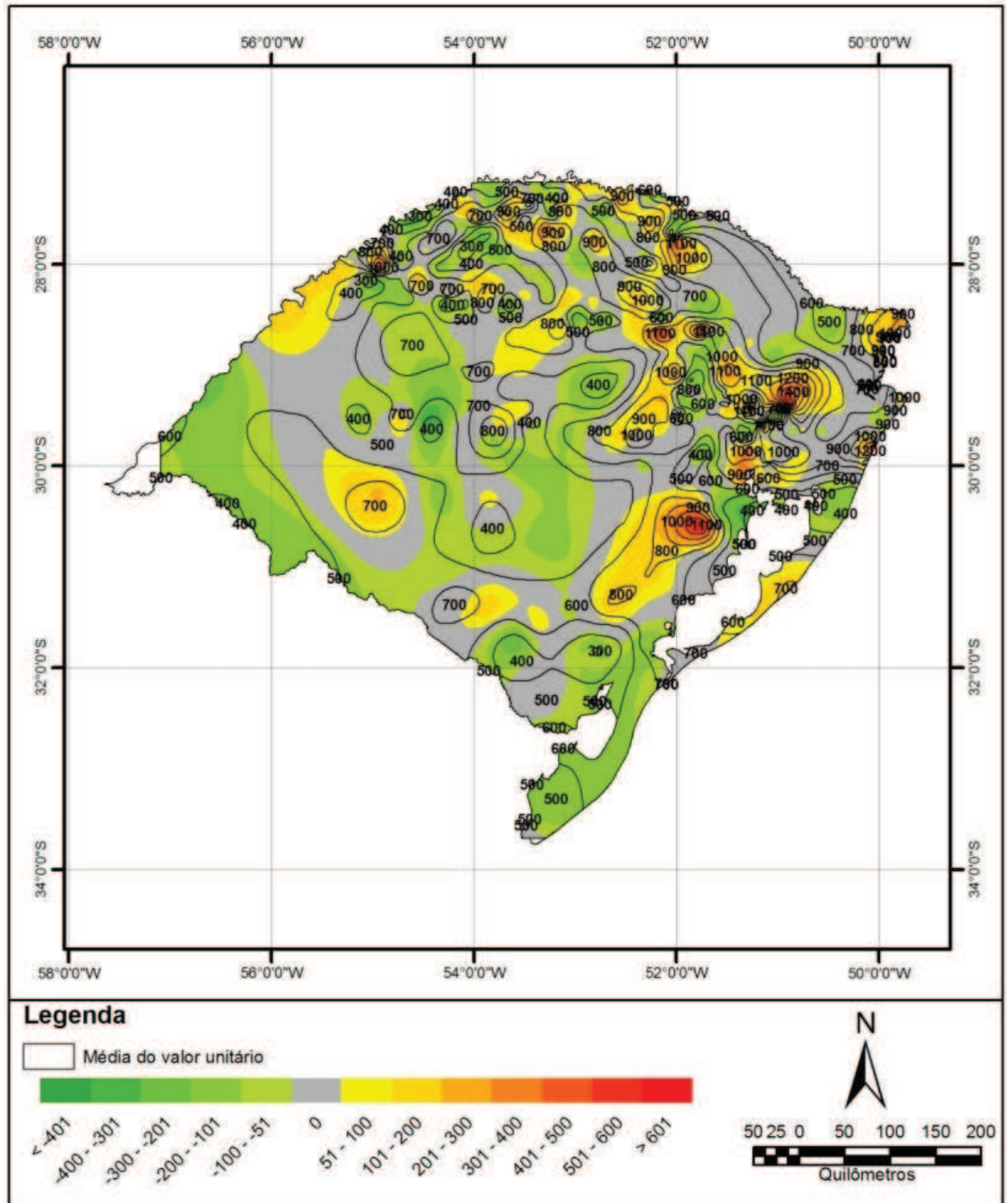


Figura 11: Análise de resíduos da variável VU.

Fonte: elaborado pelo autor

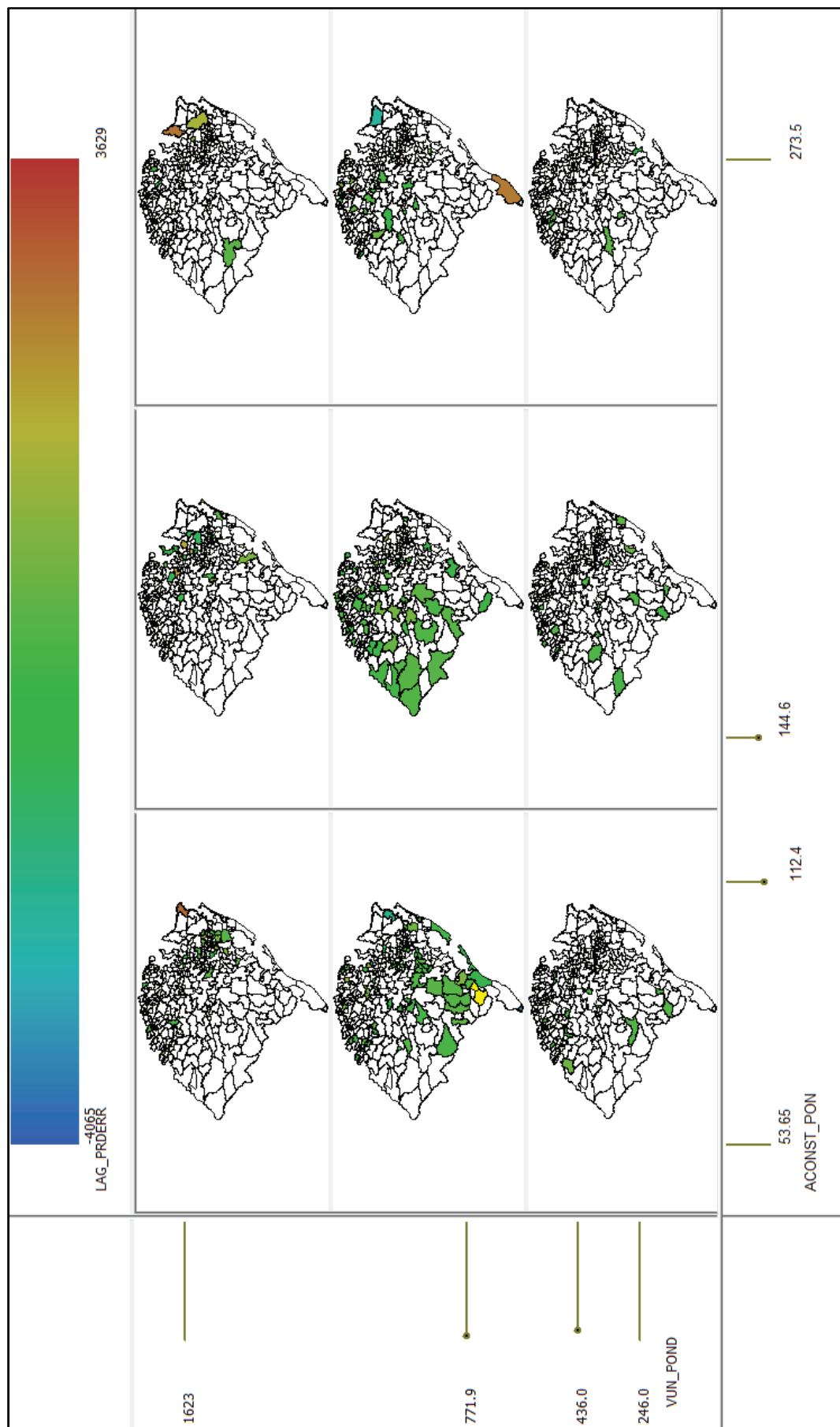


Figura 12: Análise dos resíduos em mapas condicionais.

Fonte: elaborado pelo autor

Figura 12 refere-se à análise dos resíduos em comparação com o valor unitário e a área construída. O eixo dos Y apresenta três faixas de valores unitários, sendo que seus extremos foram extraídos da base de dados, enquanto os valores centrais foram configurados para melhor compreensão dos gráficos condicionados. O mesmo ocorre no eixo dos X, apenas alterando a variável para área construída.

Os municípios destacados nas áreas de análise demonstram o valor dos resíduos para o cruzamento comparativo entre VU e área construída. Os transbordamentos ocorridos são indicados pelos extremos da escala de cores, representados nos respectivos mapas.

Observa-se que os municípios do extremo sul do Estado apresentaram transbordamento positivo para a faixa superior da área construída e a faixa intermediária dos valores unitários.

Também ocorreu transbordamento negativo, na Região Noroeste, na mesma faixa de valores..

5 CONCLUSÕES

Este trabalho buscou mensurar a melhoria dos modelos tradicionais de avaliações de imóveis com a introdução de variáveis sócio-econômicas.

Esta melhoria dá maior confiabilidade às avaliações destes imóveis no mercado imobiliário, pois vai além dos pressupostos básicos de regressão múltipla e avança na metodologia espacial.

Conclui-se, baseado nos resultados dos modelos econométricos, que as variáveis que mostraram uma correlação estatisticamente significativa com o Valor Unitário dos imóveis foram: área construída, área do terreno, pavimentação urbana, padrão de acabamento, estado de conservação, PIB per capita, frota urbana, número de agências da Caixa Econômica Federal, número de domicílios, número de banheiros, número de hospitais, densidade demográfica e número de homicídios. Percebe-se que a variável área construída se associou negativamente com o VU, enquanto as demais se associaram positivamente com o mesmo.

Observou-se, também, que a utilização de regressão espacial foi favorável pela melhoria dos índices estatísticos do modelo.

Sabe-se que a avaliação dos imóveis, baseada em dados que apresentem autocorrelação espacial, quando é realizada pelos modelos de regressão múltipla, pode oferecer resultados discrepantes ou tendenciosos, sem nenhuma consistência.

Para dar maior confiança nos trabalhos avaliatórios, deve-se utilizar a Metodologia da Regressão Espacial, pois evita a violação das principais características dos estimadores. Logo, sempre deverá ser testada a possibilidade de utilização de regressão espacial.

Houveram muitas limitações deste trabalho, sendo a fundamental a impossibilidade de realizar a regressão espacial com dados georeferenciados para os 19.203 imóveis, compostos por 7.457 apartamentos e 11.746 casas.

Assim utilizou-se um agrupamento de 288 informações para unidades ponderadas, para cada município, o que gera perda de variabilidade e do grau de explicação do modelo.

6 INDICAÇÃO PARA TRABALHOS FUTUROS

Dando seqüência ao trabalho apresentado, a sugestão para a melhora dos modelos de regressão apresentados são as seguintes:

- Utilização de variáveis dicotômicas representando imóveis do tipo apartamento e casa, novos e usados, bairro e centro;
- Testar a presença ou não de instituições comerciais de financiamento no município;
- Fundir variáveis econômicas em uma única representativa para cada município;
- Apresentar o variograma;
- Analisar a influência da distância do imóvel, ao centro do município;
- Testar a taxa de ocupação, ou seja, a variável número de habitantes por número de domicílios;
- Comparar as variáveis utilizadas, com os resultados de autores que também estudaram as variáveis citadas;
- Escalonar cidades conforme número de amostragem de transações imobiliárias: pequenas, médias e grandes;
- Utilizar as variáveis homicídios, frota, etc., ponderadas por 100 mil habitantes;
- Analisar o valor ponderado do Desvio Padrão da Média de Preços no teste de variáveis;
- Testar se número de terrenos vagos influenciam na valoração dos imóveis;
- Comparar metodologias já utilizadas com a apresentada;
- Averiguar bolha de preços, regiões metropolitanas, áreas comparáveis, precificação e cenário de estresse, entre outros.

A sugestão de análise da base completa, explorando sua totalidade, demonstra necessidade de utilização de ferramentas computacionais com algoritmos que permitam o teste com um maior número de variáveis e observações.

7 REFERÊNCIAS

- ABUNAHMAN, S.A. *Engenharia Legal e de Avaliações*. São Paulo: PINI, 2000.
- ALMEIDA, E. S. *Curso de Econometria Espacial Aplicada*. Piracicaba: s.ed., 2004.
- ANSELIN, L. *Spatial Econometrics: Methods and models*. DORDRECHT: Kluwer Academic, 1988.
- ANSELIN, L. *Spatial Externalities, Spatial Multipliers and Spatial Econometrics*. Specialist Meeting on Spatial Externalities, Santa Barbara, Califórnia, 2001.
- ANSELIN, L.. *Exploring Spatial data with GeoDaTM: A Workbook*. Center for Spatially Integrated Social Science, University of Illinois, Urbana-Champaign, Illinois, 2005.
- ANSELIN, L.. *Spatial econometrics*. Bruton Center, School of Social Sciences, University of Texas, Dallas, Richardson, 1999.
- ANSELIN, L.; BERA, A.; *Spatial dependence in linear regression models with an introduction to spatial econometrics*. Orgs.: ULLAH, A. *Handbook of Applied Economic Statistics*, Giles: Marcel Dekker, 1997.
- ANSELIN, L. The Moran scatterplot as ESDA tool to assess local instability in spatial association. In: Fisher, M.; Scholten, H. J.; Unwin, D. *Spatial Analytical Perspectives on GIS*. London: Taylor & Francis, 1996, p. 111-126.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 5676. Avaliação de Imóveis Urbanos*. Rio de Janeiro, 1989.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 14653-1. Avaliação de Bens. Parte 1: Procedimentos gerais*. Rio de Janeiro, 2001.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 14653-2. Avaliação de Bens. Parte 2: Imóveis urbanos*. Rio de Janeiro, 2004.
- BALCHIN, P.N.; KIEVE, J.L. *Urban land economics*. 3. ed. London: McMillan, 1986.
- BALL, M.J. *Recent empirical work on the determinants of relative house prices*. Urban Studies, v.10, p.213-233, 1973.
- BAPTISTELLA, Marisa. *O uso de redes neurais e regressão linear múltipla na engenharia de avaliações: determinação dos valores venais de imóveis urbanos*. (130.pdf)

BELSLEY, D.A.; KUH, E.; WELSCH, R.E. *Regression diagnostics: Identifying influential data and sources of co linearity*. New York: John Wiley, 1980.

BRAULIO, S. N. *Proposta de uma Metodologia para Avaliação de Imóveis Urbanos baseado em Métodos Estatísticos Multivariados*. Curitiba: Universidade Federal do Paraná, 2005. Dissertação de Mestrado em Métodos Numéricos em Engenharia.

BROWN, G.R. *Property Investment and Performance Measurement: A Reply*. Journal of Valuation 4: 33-44. 1985.

CÂMARA, Gilberto et al. *Análise espacial de áreas*. Disponível em <http://www.dpi.inpe.br/gilberto/livro/analise/cap.5-areas.pdf>. Acesso em 14/08/10.

CANAN, K. A. *Determinação Científica de Fatores de Homogeneização para Padrão de Acabamento e Estado de Conservação analisando Residências Localizadas no Centro da Cidade de Cascavel- PR.* Cascavel: Faculdade Assis Gurgacz - FAG, 2005. Especialização em Perícias Judiciais e Avaliações em Engenharia.

CRESSIE, N. *The Origins of kriging*. Math. Geology, v.22, p.239-52, 1990.

DANIEL, C.; WOOD, F.S. *Fitting equations to data*. 2. ed. New York: John Wiley, 1980.

DANTAS, R. A. *Engenharia de Avaliações: Uma Introdução à Metodologia Científica*. São Paulo: PINI, 1998.

FOTHERINGHAM, A.S. et al. *Quantitative Geography*. London: Sage, 2000.

FLORAX, R.J.G.M.; GRAAFF, T. *The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies*. Free University Amsterdam, Orgs.: ANSELIN, L.; FLORAX, R.J.G.M.; REY, S.J. *Advances in Spatial Econometrics: Methodology, Tools and Applications*. New York: Springer, 29-65,2004.

GONZALEZ, M.A.S. *A formação do valor de aluguéis de apartamentos residenciais na cidade de Porto Alegre*. Porto Alegre: UFRGS, 1993. Dissertação de Mestrado em Engenharia.

GONZALEZ, M.A.S.. *A Engenharia de Avaliações na visão inferencial*. São Leopoldo: Editora da UNISINOS, 1997.

GONZALEZ, M.A.S.; FORMOSO, C.T. *Análise conceitual das dificuldades na determinação de modelos de formação de preços através de análise de regressão*. Engenharia Civil – UM – n°8, UFRGS, 2000.

GUJARATI, D. *Econometria Básica*. 4. ed. Rio de Janeiro: Elsevier Campus, 2006.

HARMANN, H.H. *Modern factor analysis*. 3. ed. Chicago: University of Chicago, 1976.

HAYASHI, F. *Econometrics*. New Jersey: Princenton University Press, 2000.

HOLLY, S.; PESARAN, M.H.; YAMAGATA. T., *A Spatio - Temporal Model of House Prices in the US*, Cambridge *Working Papers in Economics* 0654, Faculty of Economics, University of Cambridge, 2006.

JAKOB, A.A.E. *Krigagem como Método de Análise de Dados Demográficos*. UNICAMP/NEPO. XIII Encontro da Associação Brasileira de Estudos Populacionais, Ouro Preto, Minas Gerais, Brasil, 2002.

JUDGE, et al. *Introduction to the theory and practice of econometrics*. 2. ed. New York: John Wiley, 1985.

LANDIM, P.M.B. *Análise Estatística de Dados Geológicos*. 2 ed. São Paulo: UNESP 2003.

LIZIERI, C.; VENMORE-ROWLAND, P. Valuation Accuracy: A Contribution to the Debate. *Journal of Property Research* 8-2: 115-122, 1991.

MADDALA, G.S. *Introduction to econometrics*. New York: McMillan, 1988.

MARQUES, J.L.; CASTRO, E.A.; BHATTACHARJEE, A. *A localização urbana na valorização residencial: Modelos de autocorrelação espacial*. Congresso de Desenvolvimento Regional de Cabo Verde, 2009.

MATYSIAK, G.; WANG, P. Commercial Property Market Prices and Valuations: Analysing the Correspondence. *Journal of Property Research* 12: 181-202, 1995.

McALLISTER, P. *Uncertainties in Property Performance Measurement: The Problem of Valuation Error*. International Conference on Financial Management of Property and Construction. Newcastle. Northern Ireland: 185-199, 1995.

MEEN, Geoffrey; ANDREW, Mark. *Modelling Regional House Prices: A Review of the Literature*. (out1.doc)

MONASTERIO, Leonardo M.; ÁVILA, Rodrigo Peres de, 2004. Uma Análise Espacial Do Crescimento Econômico Do Rio Grande Do Sul (1939-2001), *Anais do XXXII Encontro Nacional de Economia* [Proceedings of the 32th Brazilian Economics Meeting] 113, ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia [Brazilian Association of Graduate Programs in Economics], 2004.

MORCELLI, A. T. G. *Construção de Modelos de Regressão para estimar o Valor dos Lotes Urbanos do Setor 11 de Santa Maria-RS*. Santa Maria: Universidade Federal de Santa Maria, 2006. Dissertação de Mestrado em Engenharia da produção.

MOREIRA, A.L. *Princípios de Engenharia de Avaliações*. São Paulo: Pini / Escola Nacional de Habitação e Poupança, 1984.

NETER, J.; WASSERMANN, W.; KUTNER, M.H. *Applied linear statistical models*. 3. ed. Burr Ridge: Richard D. Irwin, 1990.

NOVAES, L.F.; PAIVA, S.A. *Situações onde a aplicação da Análise Envoltória de dados (DEA) apresenta vantagens sobre o uso da regressão linear múltipla*. (avaliação de imóveis). XII COBREAP – Belo Horizonte/MG, 2003.

PINDYCK, R.S.; RUBINFELD, D.L. *Econometria. Modelos & Previsões*. 4. ed. Rio de Janeiro: Elsevier Campus, 2004.

RESENDE, T. G; CYPRIANO, L.A. Valorização dos lotes urbanos na cidade de Toledo: uma análise econométrica no período de 1998-2008. *Revista Brasileira de Gestão Urbana*, v.2, n.1, p. 107-116, jan/jun.2010.

ROBINSON, R. *Housing economics and public policy*. London: McMillan, 1979.

SMITH, L.B.; ROSEN, K.T.; FALLIS, G. Recent development in economic models of housing markets. *Journal of Economic Literature*, v.26, p.29-64, Mar. 1988.

SOARES, Ilton G. e CASTELAR, Ivan. *Econometria Aplicada com uso do Eviews*. Fortaleza: UFC/CAEN, 2003. 276 p.

TRIVELLONI, C.A.P. *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.

TRIVELLONI, C.A.P. *Utilização de Métodos de Análise Espacial na Avaliação em Massa de Imóveis*. Congresso Brasileiro de Cadastro Técnico Multifinalitário. UFSC, Outubro 2006.

WHITE, H. *Artificial neural networks. Approximation and learning theory*. Cambridge (USA): Blackwell, 1992.

WOOLDRIDGE, J.M. *Introdução à Econometria. Uma Abordagem Moderna*. 2. ed. São Paulo: Pioneira Thomson Learning, 2006.

ANEXOS

Anexo A

Tabela de Municípios Analisados

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
84	AGUA SANTA	-28,177	-52,034	5
122	AGUDO	-29,645	-53,24	12
23	AJURICABA	-28,239	-53,771	6
125	ALECRIM	-27,655	-54,764	6
130	ALEGRETE	-29,783	-55,792	208
133	ALEGRIA	-27,827	-54,057	2
136	ALPESTRE	-27,249	-53,035	4
144	ALVORADA	-29,99	-51,084	268
150	AMETISTA DO SUL	-27,361	-53,182	3
158	ANTONIO PRADO	-28,858	-51,283	23
161	ARAMBARE	-30,915	-51,498	5
164	ARATIBA	-27,394	-52,3	7
165	ARROIO DO MEIO	-29,401	-51,945	6
50	ARROIO DO SAL	-29,551	-49,889	6
170	ARROIO DOS RATOS	-30,077	-51,729	11
110	ARROIO GRANDE	-32,238	-53,087	25
155	ARVOREZINHA	-28,872	-52,175	7
14	AUGUSTO PESTANA	-28,517	-53,992	3
181	AUREA	-27,698	-52,049	1
3	BAGE	-31,331	-54,107	332
185	BALNEARIO PINHAL	-30,247	-50,233	6
186	BARAO	-29,377	-51,496	3
70	BARAO DE COTEGIPE	-27,621	-52,38	3
115	BARRA DO RIBEIRO	-30,291	-51,301	8
203	BARROS CASSAL	-29,093	-52,583	2
209	BENTO GONCALVES	-29,171	-51,519	179
132	BOA VISTA DO BURICA	-27,669	-54,11	6
218	BOM JESUS	-28,668	-50,417	8
222	BOM PRINCIPIO	-29,489	-51,353	13
228	BOM RETIRO DO SUL	-29,609	-51,943	1
31	BOSSOROCA	-28,73	-54,9	3
239	BROCHIER	-29,551	-51,593	5
241	BUTIA	-30,12	-51,962	9
184	CACAPAVA DO SUL	-30,512	-53,491	82
248	CACEQUI	-29,884	-54,825	22
245	CACHOEIRA DO SUL	-30,039	-52,894	96
145	CACHOEIRINHA	-29,951	-51,094	889
83	CAIBATE	-28,288	-54,638	14
103	CAICARA	-27,274	-53,432	4

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
193	CAMAQUA	-30,851	-51,812	268
105	CAMBARA DO SUL	-29,048	-50,145	5
263	CAMPINA DAS MISSOES	-27,989	-54,839	2
265	CAMPINAS DO SUL	-27,716	-52,628	4
267	CAMPO BOM	-29,679	-51,053	62
226	CAMPO NOVO	-27,675	-53,803	7
270	CANDELARIA	-29,669	-52,789	2
264	CANDIDO GODOI	-27,952	-54,752	3
273	CANELA	-29,366	-50,816	83
276	CANGUCU	-31,395	-52,676	95
51	CAPAO DA CANOA	-29,761	-50,03	80
176	CAPAO DO LEAO	-31,763	-52,484	14
288	CAPELA DE SANTANA	-29,7	-51,325	3
166	CAPITAO	-29,269	-51,989	2
293	CARAZINHO	-28,284	-52,786	188
141	CARLOS BARBOSA	-29,298	-51,504	23
299	CARLOS GOMES	-27,718	-51,914	2
302	CASCA	-28,561	-51,978	14
305	CATUIPE	-28,25	-54,012	3
275	CAXIAS DO SUL	-29,168	-51,179	1407
301	CENTENARIO	-27,761	-51,999	2
280	CERRITO	-31,856	-52,813	3
192	CERRO GRANDE DO SUL	-30,59	-51,739	1
60	CERRO LARGO	-28,149	-54,738	16
9	CHAPADA	-28,055	-53,068	5
171	CHARQUEADAS	-29,955	-51,625	16
124	CHIAPETA	-27,923	-53,941	3
317	CHUI	-33,691	-53,457	1
286	CIDREIRA	-30,161	-50,234	9
319	CIRIACO	-28,344	-51,876	1
292	COLORADO	-28,524	-52,994	6
321	CONDOR	-28,208	-53,487	9
65	CONSTANTINA	-27,735	-52,992	11
180	CORONEL BARROS	-28,383	-54,066	1
234	CORONEL BICACO	-27,716	-53,701	1
210	COTIPORA	-28,994	-51,696	7
328	CRISSIUMAL	-27,5	-54,101	23
21	CRUZ ALTA	-28,639	-53,606	159
231	CRUZEIRO DO SUL	-29,513	-51,985	1
333	DAVID CANABARRO	-28,388	-51,848	1
266	DOIS IRMAOS	-29,58	-51,085	72
339	DOM PEDRITO	-30,983	-54,673	99

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
341	DONA FRANCISCA	-29,622	-53,357	10
330	DOUTOR MAURICIO CARDOSO	-27,506	-54,361	3
344	DOUTOR RICARDO	-29,086	-51,992	2
172	ELDORADO DO SUL	-30,084	-51,616	101
291	ENCANTADO	-29,236	-51,87	97
146	ENCRUZILHADA DO SUL	-30,544	-52,522	3
43	ENTRE RIOS DO SUL	-27,528	-52,733	1
324	ENTRE-IJUIS	-28,36	-54,268	1
67	ERECHIM	-27,634	-52,274	227
207	ERVAL GRANDE	-27,391	-52,571	1
213	ERVAL SECO	-27,549	-53,504	15
54	ESPUMOSO	-28,725	-52,85	34
360	ESTACAO	-27,911	-52,26	4
418	ESTANCIA VELHA	-29,648	-51,174	26
254	ESTEIO	-29,861	-51,179	692
361	ESTRELA	-29,502	-51,966	17
346	EUGENIO DE CASTRO	-28,525	-54,149	4
306	FARROUPILHA	-29,225	-51,348	158
367	FAXINAL DO SOTURNO	-29,575	-53,445	4
206	FAXINALZINHO	-27,424	-52,673	3
369	FELIZ	-29,451	-51,306	8
307	FLORES DA CUNHA	-29,029	-51,182	18
373	FONTOURA XAVIER	-28,983	-52,346	9
375	FORMIGUEIRO	-30	-53,499	5
149	FREDERICO WESTPHALEN	-27,359	-53,394	22
41	GARIBALDI	-29,256	-51,534	24
182	GAURAMA	-27,584	-52,094	4
243	GENERAL CAMARA	-29,905	-51,76	2
349	GETULIO VARGAS	-27,89	-52,228	34
382	GIRUA	-28,028	-54,35	22
384	GLORINHA	-29,881	-50,767	7
386	GRAMADO	-29,379	-50,874	10
383	GRAVATAI	-29,944	-50,992	1179
391	GUABIJU	-28,541	-51,69	2
200	GUAIBA	-30,114	-51,325	225
156	GUAPORE	-28,846	-51,89	16
62	GUARANI DAS MISSOES	-28,141	-54,558	5
353	HERVAL	-32,024	-53,396	13
398	HORIZONTINA	-27,626	-54,308	73
4	HULHA NEGRA	-31,404	-53,869	4
215	HUMAITA	-27,563	-53,974	7

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
404	IBIRUBA	-28,628	-53,09	80
405	IGREJINHA	-29,574	-50,79	61
15	IJUI	-28,388	-53,915	175
407	IMBE	-29,975	-50,128	93
304	INDEPENDENCIA	-27,833	-54,188	1
157	IPE	-28,82	-51,279	7
138	IRAI	-27,194	-53,251	7
409	ITAARA	-29,61	-53,765	13
412	ITAQUI	-29,125	-56,553	12
416	IVOTI	-29,591	-51,161	15
309	JABOTICABA	-27,631	-53,277	1
79	JACUTINGA	-27,729	-52,535	8
420	JAGUARAO	-32,566	-53,376	78
422	JAGUARI	-29,497	-54,69	21
19	JULIO DE CASTILHOS	-29,227	-53,682	38
29	LAGOA VERMELHA	-28,209	-51,526	28
390	LAGOAO	-29,235	-52,796	3
167	LAJEADO	-29,467	-51,961	154
89	LAVRAS DO SUL	-30,813	-53,895	37
432	LIBERATO SALZANO	-27,6	-53,073	1
419	LINDOLFO COLLOR	-29,597	-51,209	7
436	MACHADINHO	-27,567	-51,668	4
413	MANOEL VIANA	-29,589	-55,483	3
401	MARAU	-28,449	-52,2	113
437	MARCELINO RAMOS	-27,462	-51,906	1
439	MARIANO MORO	-27,354	-52,147	8
46	MARQUES DE SOUZA	-29,328	-52,093	1
424	MATA	-29,566	-54,46	2
440	MATO LEITAO	-29,524	-52,129	1
61	MATO QUEIMADO	-28,256	-54,616	4
240	MINAS DO LEO	-30,127	-52,048	2
236	MIRAGUAI	-27,494	-53,686	3
289	MONTENEGRO	-29,689	-51,461	172
444	MORMACO	-28,692	-52,692	3
279	MORRO REDONDO	-31,107	-50,921	3
417	MORRO REUTER	-29,538	-51,081	2
113	MOSTARDAS	-31,107	-50,921	2
446	MUCUM	-29,165	-51,868	5
450	NAO-ME-TOQUE	-28,459	-52,821	77
403	NICOLAU VERGUEIRO	-28,536	-52,464	2
204	NONOAI	-27,362	-52,771	8
410	NOVA ALVORADA	-28,678	-52,166	2

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
392	NOVA ARACA	-28,659	-51,745	3
452	NOVA BASSANO	-28,724	-51,705	10
329	NOVA CANDELARIA	-27,607	-54,107	3
427	NOVA ESPERANCA DO SUL	-29,41	-54,829	7
406	NOVA HARTZ	-29,583	-50,902	4
121	NOVA PALMA	-29,472	-53,469	6
388	NOVA PETROPOLIS	-29,376	-51,114	27
365	NOVA PRATA	-28,784	-51,61	31
281	NOVA SANTA RITA	-29,857	-51,274	9
455	NOVO HAMBURGO	-29,678	-51,131	955
285	OSORIO	-29,887	-50,27	98
458	PALMARES DO SUL	-30,258	-50,51	7
313	PALMEIRA DAS MISSOES	-27,899	-53,314	15
355	PALMITINHO	-27,355	-53,555	6
460	PANAMBI	-28,293	-53,502	16
393	PARAI	-28,594	-51,786	9
463	PARAISO DO SUL	-29,669	-53,149	1
162	PAROBE	-29,629	-50,835	54
326	PASSO FUNDO	-28,263	-52,407	653
1	PEDRAS ALTAS	-31,733	-53,584	8
175	PEDRO OSORIO	-31,864	-52,823	15
17	PEJUCARA	-28,423	-53,656	4
10	PELOTAS	-31,772	-52,343	1594
453	PICADA CAFE	-29,444	-51,136	11
457	PINHAL	-27,511	-53,215	7
72	PINHEIRO MACHADO	-31,578	-53,381	40
468	PIRAPO	-28,045	-55,199	2
352	PIRATINI	-31,448	-53,104	3
137	PLANALTO	-27,329	-53,059	5
188	POCO DAS ANTAS	-29,45	-51,671	1
69	PONTE PRETA	-27,654	-52,488	2
287	PORTAO	-29,702	-51,242	16
126	PORTO MAUA	-27,575	-54,668	1
469	PORTO XAVIER	-27,906	-55,138	13
471	QUARAI	-30,388	-56,451	15
252	RESTINGA SECA	-29,813	-53,375	19
111	RIO GRANDE	-32,035	-52,099	578
345	RIO PARDO	-29,99	-52,378	5
298	RIOZINHO	-29,641	-50,453	2
39	ROCA SALES	-29,284	-51,868	52
332	RODEIO BONITO	-27,471	-53,169	6
475	ROLANTE	-29,651	-50,576	57

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
78	RONDA ALTA	-27,767	-52,802	7
334	ROQUE GONZALES	-28,131	-55,026	1
129	ROSARIO DO SUL	-30,258	-54,914	24
477	SAGRADA FAMILIA	-27,707	-53,136	2
294	SALDANHA MARINHO	-28,393	-53,095	3
56	SALTO DO JACUI	-29,088	-53,213	15
82	SALVADOR DAS MISSOES	-28,126	-54,835	4
94	SALVADOR DO SUL	-29,438	-51,511	2
295	SANTA BARBARA DO SUL	-28,358	-53,247	10
464	SANTA CRUZ DO SUL	-29,718	-52,426	337
337	SANTA MARIA	-29,684	-53,807	815
387	SANTA MARIA DO HERVAL	-29,498	-50,993	1
478	SANTA ROSA	-27,871	-54,481	127
112	SANTA VITORIA DO PALMAR	-33,519	-53,368	28
246	SANTANA DA BOA VISTA	-30,872	-53,115	6
340	SANTANA DO LIVRAMENTO	-30,891	-55,533	113
421	SANTIAGO	-29,192	-54,867	160
396	SANTO ANGELO	-28,299	-54,263	158
297	SANTO ANTONIO DA PATRULHA	-29,818	-50,52	10
379	SANTO ANTONIO DAS MISSOES	-28,511	-55,228	8
454	SANTO AUGUSTO	-27,851	-53,777	47
128	SANTO CRISTO	-27,824	-54,663	32
380	SÃO BORJA	-28,661	-56,004	28
426	SÃO FRANCISCO DE ASSIS	-29,55	-55,131	27
107	SÃO FRANCISCO DE PAULA	-29,448	-50,584	9
250	SÃO GABRIEL	-30,336	-54,32	14
169	SÃO JERONIMO	-29,959	-51,722	14
482	SÃO JOSE DO NORTE	-32,015	-52,042	5
106	SÃO JOSE DOS AUSENTES	-28,748	-50,066	1
456	SÃO LEOPOLDO	-29,76	-51,147	1073
11	SÃO LOURENCO DO SUL	-31,365	-51,978	100
503	SÃO LOURENCO DO SUL	-31,365	-51,978	100
336	SÃO LUIZ GONZAGA	-28,408	-54,961	69
372	SÃO MARCOS	-28,971	-51,068	14
32	SÃO MIGUEL DAS MISSOES	-28,563	-54,554	2
476	SÃO PAULO DAS MISSOES	-28,021	-54,936	43
80	SÃO PEDRO DO BUTIA	-28,124	-54,887	5
247	SÃO PEDRO DO SUL	-29,621	-54,179	39
223	SÃO SEBASTIAO DO CAI	-29,587	-51,376	16
251	SÃO SEPE	-30,161	-53,565	59

ID DO MUNICÍPIO	MUNICÍPIO	LATITUDE	LONGITUDE	NÚMERO DE IMÓVEIS
448	SAO VALENTIM DO SUL	-27,558	-52,524	16
425	SAO VICENTE DO SUL	-29,692	-54,679	17
268	SAPIRANGA	-29,638	-51,007	22
483	SAPUCAIA DO SUL	-29,839	-51,144	399
7	SARANDI	-27,944	-52,923	25
214	SEBERI	-27,478	-53,403	9
269	SEDE NOVA	-27,635	-53,946	4
139	SELBACH	-28,629	-52,953	11
487	SERAFINA CORREA	-28,712	-51,935	29
489	SERTAO	-27,98	-52,26	1
191	SERTAO SANTANA	-30,46	-51,603	1
438	SEVERIANO DE ALMEIDA	-27,433	-52,116	2
177	SOLEDADE	-28,818	-52,51	97
86	TAPEJARA	-28,068	-52,014	78
359	TAPERA	-28,626	-52,87	27
198	TAPES	-30,673	-51,396	13
385	TAQUARA	-29,651	-50,781	121
229	TAQUARI	-29,8	-51,864	10
484	TAQUARUCU DO SUL	-27,4	-53,467	4
354	TENENTE PORTELA	-27,371	-53,758	2
217	TEUTONIA	-29,448	-51,806	7
108	TORRES	-29,335	-49,727	42
318	TRAMANDAI	-29,985	-50,134	189
274	TRES COROAS	-29,517	-50,778	19
135	TRES DE MAIO	-27,773	-54,24	43
224	TRES PASSOS	-27,456	-53,932	10
492	TUCUNDUVA	-27,657	-54,44	18
18	TUPANCIRETA	-29,081	-53,836	23
127	TUPARENDI	-29,081	-53,836	13
131	URUGUAIANA	-29,755	-57,088	123
221	VACARIA	-28,512	-50,934	131
370	VALE REAL	-29,398	-51,254	1
230	VENANCIO AIRES	-29,606	-52,192	86
494	VERA CRUZ	-29,715	-52,506	9
159	VERANOPOLIS	-28,936	-51,549	49
183	VIADUTOS	-27,569	-52,02	2
117	VIAMAO	-30,081	-51,023	346
99	VICTOR GRAEFF	-28,56	-52,748	5
259	VILA MARIA	-28,535	-52,154	5
91	VILA NOVA DO SUL	-30,344	-53,883	2
495	VISTA ALEGRE	-27,367	-53,49	5
283	XANGRI-LA	-29,801	-50,044	15

Anexo B

Regression				
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION				
Data set	:	CENT_UNI_MUNICIPIOS_DADOS		
Dependent variable	:	VUN_POND	Number of observations:	288
Mean dependent var	:	651.381	Number of variables	14
S.D. dependent var	:	211.601	Degrees of Freedom	274
R-squared	:	0.453701	F-statistic	17.5044
Adjusted R-squared	:	0.427782	Prob(F-statistic)	3.23539e-029
Sum squared residual	:	7.04464e+006	Log likelihood	-1863.75
Sigma-square	:	25710.4	Akaike info criterion	3755.5
S.E. of regression	:	160.344	Schwarz criterion	3806.78
Sigma-square ML	:	24460.5		
S.E. of regression ML	:	156.399		
Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	-117.7098	95.22571	-1.236113	0.2174757
ACONST_PON	-1.34712	0.3949554	-3.410816	0.0007453
TERR_POND	0.180552	0.0575395	3.13788	0.0018874
PAV_POND	44.18212	31.03985	1.4234	0.1557576
ACAB_POND	171.8044	32.50796	5.284994	0.0000003
CONSERV_PO	128.4538	30.31026	4.237966	0.0000309
PIB_PER	0.005387097	0.002251402	2.392774	0.0173952
FROTA	0.006476919	0.002085742	3.10533	0.0021000
AG_CAIXA	61.73802	25.57798	2.413717	0.0164468
DOMIC	-0.008704205	0.003300537	-2.637209	0.0088365
BANH_POND	81.41429	32.73306	2.487219	0.0134699
HOSPITAIS	7.69578	25.177	-0.3056671	0.7600898
DEN_DEM	0.03477071	0.03192862	1.089014	0.2771035
HOMIC	3.348134	2.55062	1.312675	0.1903916
REGRESSION DIAGNOSTICS				
MULTICOLLINEARITY CONDITION NUMBER		36.798228		
TEST ON NORMALITY OF ERRORS				
TEST	DF	VALUE	PROB	
Jarque-Bera	2	320.8949	0.0000000	
DIAGNOSTICS FOR HETEROSKEDASTICITY				
RANDOM COEFFICIENTS				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	13	35.93628	0.0006070	
Koenker-Bassett test	13	11.14429	0.5987311	
SPECIFICATION ROBUST TEST				
TEST	DF	VALUE	PROB	
white	104	N/A	N/A	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
FOR WEIGHT MATRIX : Matriz_Peso_Espacial_2.GAL (row-standardized weights)				
TEST	MI/DF	VALUE	PROB	
Moran's I (error)	0.169797	4.0381660	0.0000539	
Lagrange Multiplier (lag)	1	16.3281671	0.0000533	
Robust LM (lag)	1	3.4224293	0.0643163	
Lagrange Multiplier (error)	1	14.1048655	0.0001729	
Robust LM (error)	1	1.1991276	0.2734961	
Lagrange Multiplier (SARMA)	2	17.5272948	0.0001563	
===== END OF REPORT =====				

Anexo C

Regression				
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION				
Data set	: CENT_UNI_MUNICIPIOS_DADOS			
Spatial weight	: Matriz_Peso_Espacial_2.GAL			
Dependent variable	: VUN_POND	Number of observations	: 288	
Mean dependent var	: 651.381	Number of variables	: 15	
S.D. dependent var	: 211.601	Degrees of Freedom	: 273	
Lag coeff. (Rho)	: -1.21664			
R-squared	: 0.495306	Log likelihood	: -1853.73	
Sq. Correlation	: -	Akaike info criterion	: 3737.46	
Sigma-square	: 22597.7	Schwarz criterion	: 3792.41	
S.E of regression	: 150.325			
variable	Coefficient	std.Error	z-value	Probability
W_VUN_POND	-1.216643	0.005993549	-202.992	0.0000000
CONSTANT	487.2929	153.8812	3.166682	0.0015420
ACONST_PON	-1.910317	0.6386188	-2.991327	0.0027778
TERR_POND	0.1917788	0.09297919	2.0626	0.0391506
PAV_POND	84.57635	50.15796	1.6862	0.0917573
ACAB_POND	237.4804	52.53094	4.520771	0.0000062
CONSERV_PO	138.0597	48.98461	2.818431	0.0048260
PIB_PER	0.01108821	0.003638073	3.047825	0.0023052
FROTA	0.01159433	0.003370483	3.43996	0.0005819
AG_CAIXA	139.0533	41.35365	3.362541	0.0007724
DOMIC	-0.0211343	0.005333374	-3.962651	0.0000742
BANH_POND	87.2119	52.92876	1.647722	0.0994096
HOSPITAIS	77.32497	40.68569	-1.900545	0.0573616
DEN_DEM	0.1317794	0.05159465	2.554129	0.0106454
HOMIC	16.13492	4.121771	3.914561	0.0000906
REGRESSION DIAGNOSTICS				
DIAGNOSTICS FOR HETEROSKEDASTICITY				
RANDOM COEFFICIENTS				
TEST		DF	VALUE	PROB
Breusch-Pagan test		13	22.62451	0.0464092
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : Matriz_Peso_Espacial_2.GAL				
TEST		DF	VALUE	PROB
Likelihood Ratio Test		1	3727.496	0.0000000
===== END OF REPORT =====				