

**XV COBREAP – CONGRESSO BRASILEIRO DE ENGENHARIA DE
AVALIAÇÕES E PERÍCIAS – IBAPE-SP**

TRABALHO DE AVALIAÇÃO

**COMPARAÇÃO ENTRE REGRESSÃO LINEAR E
ESPACIAL NA AVALIAÇÃO DE
APARTAMENTOS NA CIDADE DE BELÉM**

RESUMO

Este trabalho vem estudar a regressão linear e a regressão espacial com o objetivo realizar uma comparação entre essas duas metodologias de cálculo utilizadas na avaliação de imóveis, em particular, apartamentos na cidade de Belém. Nas análises realizadas com base em informações do Censo demográfico (2000) e dados de imóveis coletados mediante pesquisa de mercado, verificou-se indícios de dependência espacial nos preços dos imóveis, e que os resultados mostraram-se confiáveis para os dois métodos analisando os índices econométricos.

Palavras chave: *Regressão linear, Regressão espacial, Análise espacial de dados, Imóveis em Belém.*

Comparação entre regressão linear e espacial na avaliação de apartamentos na cidade de Belém

Resumo

Este trabalho vem estudar a regressão linear e a regressão espacial com o objetivo realizar uma comparação entre essas duas metodologias de cálculo utilizadas na avaliação de imóveis, em particular, apartamentos na cidade de Belém.

Nas análises realizadas com base em informações do Censo demográfico (2000) e dados de imóveis coletados mediante pesquisa de mercado, verificou-se indícios de dependência espacial nos preços dos imóveis, e que os resultados mostraram-se confiáveis para os dois métodos analisando os índices econométricos.

Palavras chave: regressão linear, regressão espacial, análise espacial de dados, imóveis em Belém.

ABSTRACT

This work studies the linear and spacial regression analysis, with the objective of making a comparison between these methodologies used in real properties evaluation, particularly, apartments at Belém city.

The analysis was made with the information from the Demographic Census (2000), and real properties data colected based on market's research, one verified indications of spacial dependency in properties value, such as the results were in trust for the both methods, analysing the econométrics table of contents.

Keywords: *linear regression, spacial regression, spatial data analysis, real properties at Belém*

1 INTRODUÇÃO

De acordo com Moreira 2001, a Engenharia de Avaliações não é uma ciência exata, mas sim a arte de estimar os valores de imóveis com diversas especificidades em que o conhecimento profissional especializado de engenharia e o bom julgamento são condições essenciais.

O mercado imobiliário em Belém encontra-se aquecido nos últimos anos, impulsionado pelas linhas de financiamento imobiliário oferecidas pelo governo e administradas pela Caixa Econômica Federal, facilitando acesso ao crédito à todas as classes sociais, o que torna este setor um dos pontos chaves da economia Brasileira, sob o ponto de vista de geração de emprego e renda com grande impacto social. Entretanto, o lançamento de novos empreendimentos habitacionais não consegue acompanhar a crescente demanda pelo produto habitação. Este cenário, de forma macroeconômica pode induzir o avaliador a cometer equívocos, causados pela especulação imobiliária.

Compreender este mercado, e ainda, tratá-lo de forma matemática, utilizando-se das ferramentas estatísticas, tem-sido cada vez mais um desafio para a engenharia de avaliações. Com isso, na avaliação imobiliária, isto é, na determinação do valor de um imóvel, não se pode por em foco um único aspecto da

questão, convindo considerar simultaneamente, vários fatores, como por exemplo, a utilidade e o custo, por serem inseparáveis.

Por esse motivo, neste estudo, a análise do mercado habitacional foi realizada em caráter microeconômico com a inclusão no modelo da variável “Bairro” que representa o valor do rendimento nominal médio mensal dos responsáveis pelos domicílios particulares permanentes” publicada no censo demográfico-IBGE 2000, buscando a elaboração de um modelo inferencial com o objetivo de comparar os dados obtidos pelo modelo regressão linear e regressão espacial para a cidade de Belém, levando em conta a localização espacial do fenômeno de forma explícita, procurando identificar, para o mercado imobiliário de apartamentos, uma amostra que fosse capaz de distinguir, a valorização dos imóveis nos bairros pesquisados com a utilização da regressão linear e espacial.

2 METODOLOGIA

A metodologia utilizada nesta pesquisa, consistiu na elaboração de um modelo inferencial, que buscasse uma comparação dos resultados obtidos por regressão linear e por regressão espacial, identificando por meio da análise espacial de dados e com a utilização de dados socioeconômicos, as regiões onde há maior valorização imobiliária.

Com isso, a seqüência metodológica utilizada neste trabalho seguiu os seguintes passos:

- a. Pesquisa para identificar os imóveis em oferta no mercado imobiliário de Belém
- b. estimativa da equação de regressão de preços unitários de apartamentos
- c. utilização do valor do rendimento nominal médio mensal das pessoas com rendimento, responsáveis pelos domicílios particulares permanentes como variável sócio-econômica, publicada pelo IBGE no último censo demográfico no ano de 2000, afim de explicar a localização do imóvel

2.1 TRATAMENTO DOS DADOS

Os dados foram tratados por meio de regressão linear e análise espacial, utilizando-se os conceitos de econometria básica e espacial.

3 CRITÉRIOS DE CÁLCULO

A metodologia de cálculo vastamente utilizada e reconhecida na literatura é a modelagem por regressão, que resulta numa equação de regressão, que vem explicar o comportamento do mercado imobiliário, e é representada por:

$$Y_1 = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_k X_{ik} + e_i \quad i = 1, \dots, m \quad (1)$$

Na qual, Y_1 é a variável dependente representada pelo preço unitário do imóvel; X_{i1}, \dots, X_{ik} – são chamadas de variáveis independentes, correspondentes às características estruturais (área privativa, número de vagas de garagem, número de dormitórios, etc.) e de localização (bairro onde se situa o imóvel, posição no prédio, distância a pólos de influência); $b_1 X_1$ são os parâmetros do modelo e e_i são erros aleatórios.

Ao realizar a análise de regressão, têm-se o foco em dois objetivos: (1) encontrar um bom ajuste entre os preços observados e os valores estimados da

variável dependente; (2) descobrir quais das variáveis explicativas contribuem de forma significativa para este relacionamento linear. Com isso, a hipótese padrão é que as variáveis são independentes, ou seja, não são correlacionadas, e, conseqüentemente, que os resíduos e_i do modelo também são independentes e não correlacionados com a variável dependente, tem variância constante e apresentam distribuição normal com média zero.

Portanto, com a utilização do Método Comparativo de Dados de Mercado, preconizado na NBR-14653-2-2004, buscou-se um modelo que melhor se ajustasse aos dados e ao mesmo tempo tivesse um maior potencial de contribuir para a compreensão do fenômeno em estudo.

Contudo, no caso de dados espaciais, onde está presente a dependência espacial, é muito pouco provável que a hipótese padrão de observações não correlacionadas seja verdadeira. Comumente os resíduos continuam apresentando a autocorrelação espacial presente nos dados, que pode se manifestar por diferenças regionais sistemáticas, nas relações do modelo ou ainda por uma tendência espacial contínua.

Por esse motivo, no estudo da análise espacial de dados foi observada a autocorrelação espacial que trata de forma computacional a dependência espacial onde Tobler apud Câmara(2004) chama de primeira lei da geografia: “todas as coisas são parecidas, mais coisas mais próximas se parecem mais que coisas mais distantes”, ou ainda conforme Noel Cressie, “a dependência espacial está presente em todas as direções e fica mais fraca, à medida em que aumenta a dispersão na localização dos dados”. Desta forma, a autocorrelação espacial tem o objetivo de verificar como varia a dependência espacial, a partir da comparação entre os valores de uma amostra e de seus vizinhos, utilizando indicadores de uma estatística de produtos cruzados do tipo:

$$\Gamma(d) = \sum_{i=1}^n \sum_{j=1}^n w_{ij}(d) \mathbf{x}_{ij} \quad (2)$$

Onde, a matriz w_{ij} fornece uma medida de correlação entre as variáveis aleatórias que pode ser o produto destas variáveis, como no caso do índice de Moran para áreas.

Os principais testes para detectar a autocorrelação espacial são Moran I, LM Robusto(erro) e LM Robusto (defasagem). O índice global de Moran I é a expressão de autocorrelação considerando apenas o primeiro vizinho e consiste em um teste cuja a hipótese nula é de independência espacial, entretanto este teste não identifica o tipo de erro ou defasagem espacial, e por esse motivo, são realizados testes específicos como o LM robusto (erro) e o LM robusto (defasagem) para detectar os efeitos de auto-correlação e defasagem espacial na variável dependente respectivamente.

3.1 Teste LM Robusto (erro)

Segundo Dantas (2003), é um teste assintótico realizado a partir da estatística (3), que tem distribuição qui-quadrado com um grau de liberdade, sob a hipótese nula de não existência de auto-correlação espacial no termo do erro.

$$LM(erro) = \frac{[e^T W e / (s^2 / n)]^2}{tr(W^2 + W^T W)} \approx \chi^2_{(1)} \quad (3)$$

Onde e é o vetor de resíduos de mínimos quadrados; W a matriz de pesos espaciais; $s^2 = e'e/n$ a estimativa de máxima verossimilhança da variância do modelo; n o número de dados da amostra e tr o operador denominado traço da matriz.

Assim, se a estatística de teste for superior ao ponto crítico da distribuição qui-quadrado, com um grau de liberdade, rejeita-se a hipótese de não autocorrelação espacial nos resíduos do modelo clássico de regressão.

3.2 Teste LM Robusto (defasagem)

Assim como LM robusto (erro), este também é assintótico, e possui distribuição qui-quadrado com um grau de liberdade, sob a hipótese nula de não existência de defasagem espacial na variável dependente.

3.3 Modelo de defasagem espacial

A dependência espacial criada através da interação entre os preços dos imóveis causa o efeito chamado de defasagem espacial, conhecido como efeito de vizinhança que influi na determinação dos preços. Ou seja, quando há intenção entre um vendedor e um comprador em transacionar um imóvel, é levado em consideração, não apenas características locais, de padrão construtivo ou estruturais, mas também o preço dos imóveis vizinhos, que foram estudados, através de ocorrências identificadas como pontos localizados no espaço, denominados processos pontuais.

Foi adotado o modelo de defasagem espacial uma vez que Dantas (2006) apud Anselin, ressalta que o maior efeito espacial encontrado pela comparação entre o modelo de erro e o de defasagem espacial, está no que apresentar o maior valor absoluto.

Um aspecto fundamental da análise exploratória espacial é a caracterização da dependência espacial, mostrando como os valores estão correlacionados no espaço. Neste contexto, as funções utilizadas para estimar quanto o valor observado de um atributo numa região é dependente dos valores desta mesma variável nas localizações vizinhas, são a autocorrelação espacial e o variograma. Neste estudo, consideramos o índice global e Moran I , que é a expressão da autocorrelação considerando apenas o primeiro vizinho.

De maneira geral, o índice de Moran, consiste em um teste cuja hipótese nula é de independência espacial, obtendo-se o valor zero neste caso.

Valores positivos (entre 0 e +1) indicam para correlação direta, e negativos (entre 0 e -1) correlação inversa.

4 DESCRIÇÃO DO TRABALHO

Buscou-se, com base nos dados coletados, a determinação de um modelo de regressão, para que fosse realizada uma comparação entre os cálculos realizados por regressão linear e por regressão espacial. O modelo caracteriza-se por apresentar 62 dados de apartamentos, no qual a estimação foi realizada pelo método dos mínimos quadrados, sobre a variável ao nível socioeconômico, como o bairro onde foi demandado, número de quartos sociais, número de suítes, posição da unidade no prédio em relação a altitude, número de vagas de garagem, área privativa, as coordenadas geográficas longitude e latitude representadas por "coord X" e "coord Y".

Os dados foram pesquisados nos jornais de grande circulação e georeferenciados pela Companhia de Desenvolvimento da área Metropolitana de Belém (CODEM) que é detentora da base de dados georeferenciada da cidade. Posteriormente, buscou-se a identificação dos softwares que serviriam como ferramentas para a realização da pesquisa e obtenção dos resultados, então utilizou-se o SISPLAN V - Sistema de Planta de Valores, da Pelli Sistemas Engenharia, o ESRI® ArcMap™ 9.1, e o AutoCadMap 2000.

5 RESULTADOS E DISCUSSÃO

Para o modelo adotado, representado pela equação 4:

Valor unitário =

$$e^{(+6,851793483+0,1574122378*\text{dormitorios}+0,007969079395*\text{nivel/elevador}+0,3040538311*\text{suites}+0,1200427791*\text{vagasdegaragem}+5,840744856E-005*\text{bairro}-0,00731730436*\text{areaprivativa})}. \quad (4)$$

Onde:

- Valor unitário – Variável explicada (dependente) correspondente ao preço unitário do imóvel no período da pesquisa
- Dormitórios – variável quantitativa correspondente ao número de quartos sociais incluindo o número de suítes
- Suítes – variável quantitativa correspondente número de suítes encontradas no imóvel;
- Nível/elevador – variável proxy correspondente a posição do imóvel no prédio;
- Vagas de garagem – variável quantitativa correspondente ao número de garagens do apartamento;
- Área privativa – variável quantitativa correspondente a área privativa da unidade;
- Bairro – variável proxy de nível sócio econômico correspondente a localização do imóvel.

Os resultados obtidos encontram-se na Tabela 1. Onde se pode observar que os índices obtidos para as variáveis independentes estão coerentes com o comportamento mercado habitacional, pois, a medida que o número de dormitórios aumenta, o valor unitário tende a elevar-se. Também é possível concluir, que o mesmo acontece com as variáveis: nível/elevador (Figura 4) que reflete a posição do imóvel no prédio, ou seja, quanto mais alto a posição do apartamento no edifício (para prédios com elevador) em que se localiza a unidade, mais valorizada será; o número de suítes e o número de vagas de garagem.

Também, já era de se esperar que, quanto maior a renda média do chefe de família representada pela variável bairro (Figura 3) que caracteriza o índice sócio econômico utilizado, mais valorizados são os imóveis. Assim como, há uma expectativa de que o valor unitário decresça a conforme haja o aumento da área privativa.

Tabela 1 – cálculos obtidos por regressão linear

variáveis	equação	cresc	tcalc	sig RL
dormitorios	x	3,19	2,79	0,72
nivel/elevador	x	1,68	1,82	7,44
suites	x	12,93	6,84	0,01
vagas de garagem	x	2,42	1,93	5,91
bairro	x	1,7	1,98	5,29
area privativa	x	-12,85	-6,67	0,01
coord X				
coord Y				
Valor unitário	ln(x)			
correlação	0,7939			
determinação	0,6302			
R2 ajustado	0,5858			

Os ajustes indicaram um coeficiente de determinação (R^2 ajustado) de 58,58%, o que indica haver um bom poder de explicação do modelo.

Com isso, observa-se na Figura 2, uma forte aderência dos dados às curvas preço observado x valor estimado, calculadas pelos métodos de regressão linear e regressão espacial.

Figura 2 – preço observado x valor estimado

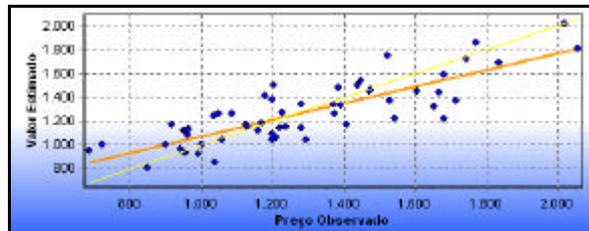


Figura 3 – bairro x valor unitário

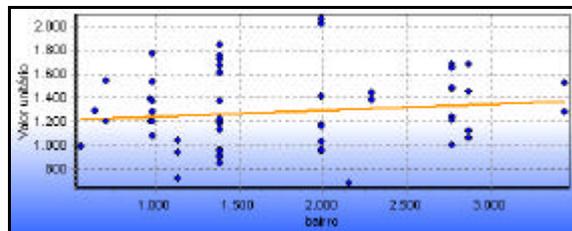
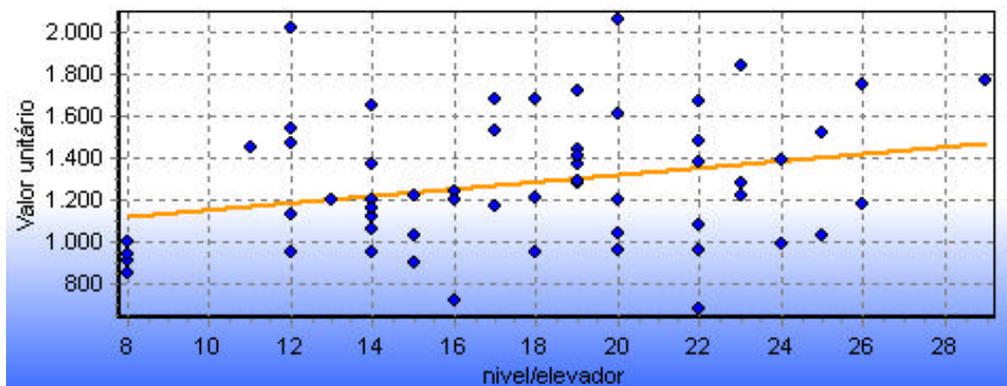


Figura 4 – nível/elevador x valor unitário



Entretanto, para que seja realizado o diagnóstico da presença de efeitos de dependência espacial foram realizados os testes LM Robusto(erro) e LM Robusto (defasagem) sobre os resíduos do método dos mínimos quadrados do modelo da Tabela 1. Indicando um um forte efeito de interação espacial nos preços dos imóveis pesquisados, uma vez que, o resultado obtido obteve uma significância inferior a 2%, conforme observa-se na Tabela 2.

Tabela 2 – resultados do teste LM Robusto

	valor	significância
LM Robusto (erro)	0,0638	0,8005
LM Robusto (defasaç	0,5520	0,4575

Diante disso, efetuaram-se os cálculos, admitindo a hipótese nula de não existência de defasagem espacial na variável dependente (sar1), pois LM Robusto (defasagem) apresentou um valor absoluto maior em relação ao LM Robusto (erro).

Os resultados obtidos pela metodologia da econometria espacial encontram-se na Tabela 03, na qual pode-se observar que os valores para a estatística t, refletem uma maior representatividade para as variáveis adotadas, em relação aos cálculos obtidos por regressão linear.

Pode-se observar ainda, efeitos espaciais significativos no valor calculado para o coeficiente de determinação (R2 ajustado) da ordem de 62,15%.

Tabela 3 – resultados obtidos por regressão espacial

variáveis	equação	cresc	tcalc	sig RE
dormitorios	x	3,19	3,1	0,18
nivel/elevador	x	1,68	1,88	5,98
suites	x	12,93	7,31	0
vaças de garagem	x	2,42	2,5	1,23
bairro	x	1,7	2,12	3,35
area privativa	x	-12,85	-7,27	0
coord X				
coord Y				
Valor unitário	ln(x)			
correlação	0,8095			
determinação	0,6556			
R2 ajustado	0,6215			
t.rho	0,6354			
sig rho(%)	52,52			

¹ É atribuída uma variável dependente Y a autocorrelação espacial ignorada. Esta abordagem é denominada como modelo espacial autoregressivo misto (“*Spacial Autoregressive – SAR*”), dado que se considera a dependência espacial através da adição ao modelo de regressão de um novo termo na forma de uma relação espacial para a variável dependente.

Com isso, a Figura 5, mostra a área de pesquisa, localizada na zona mais central da cidade de Belém, apresentando coloração alaranjada na zona mais valorizada, e desvalorização a medida que vai se afastando. Assim como, o histograma da distância entre os dados pesquisados, representado na Figura 6, destaca que a maior distância entre os imóveis pesquisados, está situada numa faixa inferior a 2.500 metros.

Figura 5 – área de pesquisa

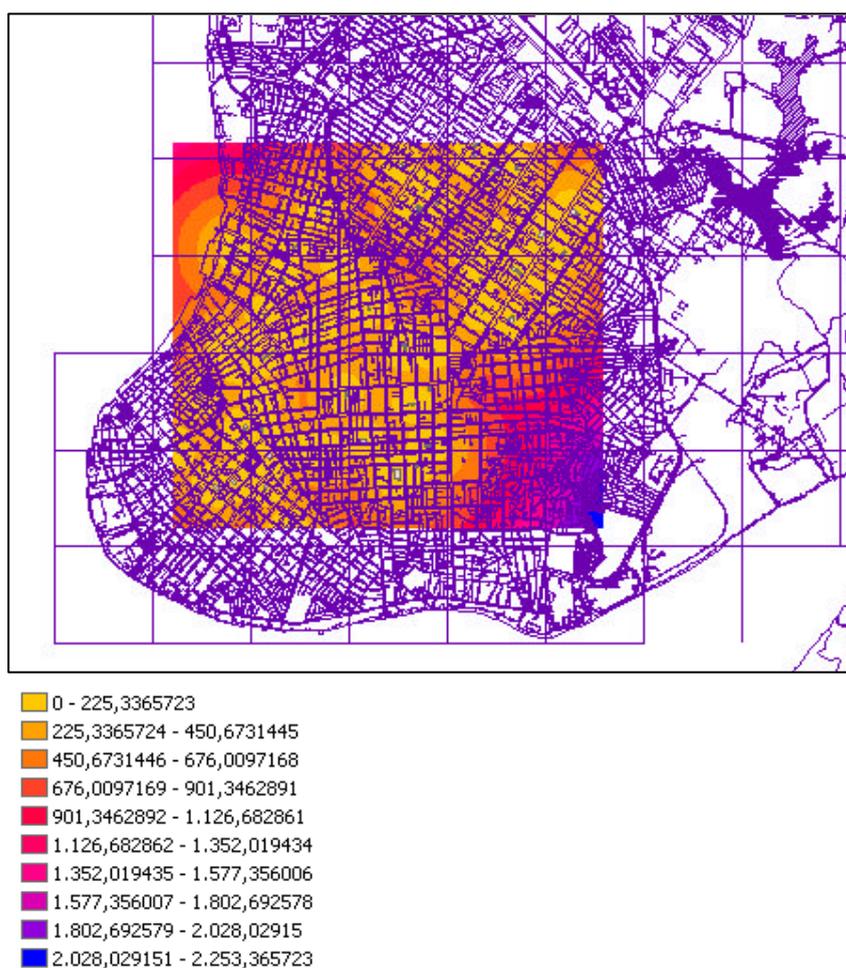
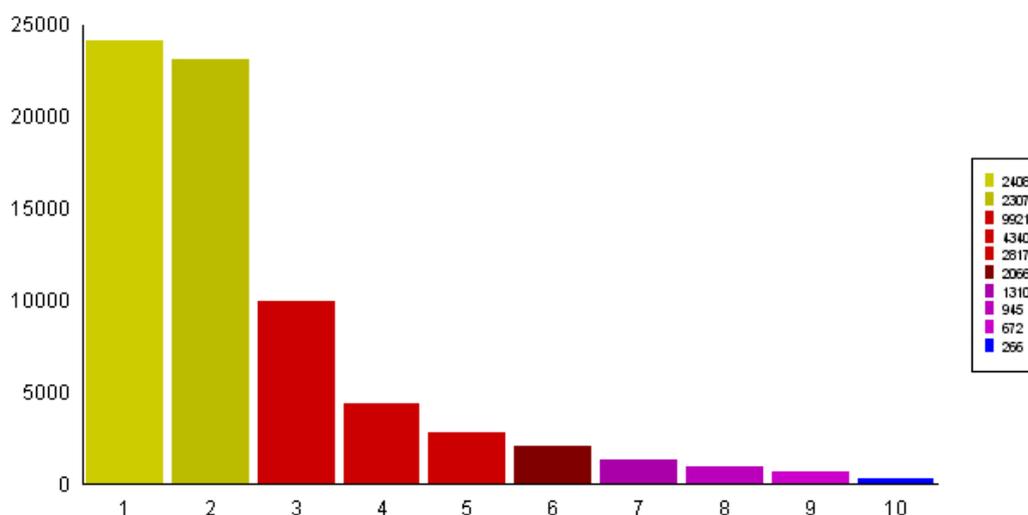


Figura 6 - Histograma da distância x dados pesquisados



6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Nos cálculos efetuados para a simulação do teste prático comparativo, onde foi realizada a projeção dos valores para avaliação de um dos imóveis constantes do banco de dados anexo a este trabalho, observou-se, analisando o modelo adotado que::

- - a vaga de garagem valoriza a unidade em 12%;
- - um apartamento com três dormitórios, é mais valorizado em 15% que outro de mesma área.
- - um apartamento com 02 suites, por exemplo, é em média 30% mais valorizado que outro de apenas uma;
- - para cada incremento de 1m² na área privativa o valor unitário decresce em 1%.
- - A altitude do apartamento no prédio representada pela variável “nível/elevador” representa uma valorização 0,7% de um andar para o outro imediatamente inferior;
- - A medida que aumenta a renda média do chefe de família, há uma valorização de 0,006% na localização do imóvel.

E ainda que, para o modelo utilizado a projeção do valor unitário para um determinado imóvel, realizado com o auxílio do software SisplanV resultou na diferença de 0,26% entre os cálculos efetuados por regressão linear e regressão espacial, projeção está, que foi considerada como um caso particular, sendo necessária em estudos futuros uma simulação com outra amostra para que se tenha uma discussão acerca da validação dos resultados.

Recomenda-se ainda, para a continuidade desta pesquisa a inclusão da variável “distancia ao pólo” a fim de verificar a área de influencia dos pólos valorizantes em relação a cada imóvel pesquisado.

Recomenda-se também que seja feito um estudo mais abrangente em parceria com a prefeitura, sobre a variável localização por meio da análise espacial de dados, com o intuito de atualizar a planta de valores genéricos do município de Belém.

7 REFERÊNCIAS BIBLIOGRÁFICAS

LESAGE, James - [Spatial Econometrics](#), Uma introdução à Econometria Espacial. *University of Toledo* (1999), disponível em <http://rri.wvu.edu/WebBook/LeSage/spatial/spatial.html> Acesso em 02/02/2007.

CÂMARA, Gilberto, *et al* - Análise Espacial de Dados Geográficos, EMBRAPA, Brasília, 2004, disponível em <http://www.dpi.inpe.br/gilberto/livro/analise/index.html>. Acesso em 02/02/2007

DANTAS, Rubens –Engenharia de Avaliações: uma introdução a metodologia científica – São Paulo: Pini, 1998

MAIA NETO, Francisco – Roteiro prático de avaliações e perícias judiciais. – 5. ed. ver atual – Belo Horizonte: Del Rey 2000

MOREIRA, Alberto Lélío – Princípios Engenharia e Avaliações. – 5. ed . – São Paulo: Pini, 2001.

ANSELIN, Luc; MORENO, Rosina – Properties of tests for spacial error components – University of Illinois, USA, october 22, 2002.

NETO, Antônio Pelli *et. Al.* – Um estudo de diagnóstico sobre a demanda habitacional e mercado imobiliário: o caso de Teresina – XXII CONGRESSO PAN AMERICANO DE AVALIAÇÕES e XIII COBREAP – CONGRESSO BRASILEIRO DE ENGENHARIA DE AVALIAÇÕES E PERÍCIAS – Fortaleza - CE, 31p., 2006.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo demográfico: dados distritais*. Belém, IBGE, 2000.

DANTAS, RUBENS; MAGALHÃES, André Matos. – Um modelo espacial de demanda habitacional para a cidade do Recife, XII COBREAP – Belo Horizonte - MG, 23p. 2003.

ABNT NBR 14.653-2 – Avaliação de Bens: Parte 2 – Imóveis urbanos, 1ª ed. 2004.

Sites Consultados

<http://pt.wikipedia.org/wiki/Bel%C3%A9m>, acesso em 02/02/2007.

<http://www.caminhos.ufms.br/matrizdados/pa/belem.html>, acesso em 02/02/2007