

---

# A INFLUÊNCIA DA AQUISIÇÃO DA INFORMAÇÃO POSICIONAL NA AVALIAÇÃO EM MASSA DE IMÓVEIS POR MEIO DE REGRESSÃO ESPACIAL

AUTORES:

ANDERSONN MAGALHÃES DE OLIVEIRA

Universidade Federal de Pernambuco - UFPE

Centro de Tecnologia e Geociências - CTG

Programa de Pós-graduação em Ciências Geodésicas e Tecnologias da Geoinformação, Recife, PE

andersonn.oliveira@ufpe.br

M.SC. ERISON ROSA DE OLIVEIRA BARROS

Universidade Federal de Pernambuco - UFPE

Centro de Tecnologia e Geociências - CTG

Departamento de Engenharia Cartográfica, Recife, PE

erison.barros@ufpe.br

---

**RESUMO**—Um problema enfrentado pelos avaliadores de imóveis e pelas prefeituras brasileiras é a falta de qualidade da informação posicional dos imóveis, pois estes dados são coletados muitas vezes sem critérios ou a partir de aquisições de produtos cartográficos sem especificações técnicas de qualidade satisfatória para essa atividade. Sabendo que é interesse do Governo Federal Brasileiro em desenvolver um índice de preço de imóveis para o Brasil através do Sistema Nacional de Gestão de Informações Territoriais – Sinter, decreto Nº 8.764, de 10 de maio de 2016, disposto no art. 41 da Lei nº 11.977, de 7 de julho de 2009, este trabalho traz subsídios e parâmetros metodológicos de avaliação geoespacial de imóveis para estas ferramentas de gestão territorial que estão sendo implementadas no país. Junto a isso, a utilização de modelos de regressão espaciais vem sendo cada vez mais utilizado na confecção de plantas de valores genéricos, conseqüentemente se faz necessária a utilização da informação posicional, visto que as influências de contágio espaciais entre os próprios dados justificam parte do erro até então não explicados por modelos clássicos de regressão. Neste Contexto, o presente trabalho trata de verificar as distorções provocadas pela utilização de diferentes modos de aquisição de dados espaciais, GPS/GNSS, Google Earth e ESIG da Prefeitura do Recife, nos modelos de regressão.

**ABSTRACT** - One problem faced by many Real Estate brokers and Brazilian City Halls is the lack of quality information of the properties, since these data are acquired many times without criteria, or from acquisitions of cartographical products without technical specifications of a certain quality for this activity. Since it is known that the Brazilian Federal Government has an interest to develop a Real Estate Price Index for Brazil, through the Territorial Information Management National System – Sinter, decree Nº 8.764, from May 10, 2016, article 41, Law nº 11.977, from July 7, 2009, this work contributes with resources and methodological parameters of geospatial evaluation of properties, for the use of these territorial management tools, which are being implemented in the country. Together with that, the utilization of space regression models is being used more each day in the making of generic value plants. Therefore, the use of the positional information is needed, since the spatial contagion information among the data itself justify part of the error, not explained until now by classic regression models. In this context, the present paper verifies the distortions caused by the use of different spatial data acquisition models, GPS/GNSS, Google Earth and ESIG from Recife's City Hall, in the models of regression.

---

## 1 INTRODUÇÃO

A grande maioria dos municípios brasileiros não dispõe de informação cartográfica de qualidade, dificultando, muitas vezes, até a localização do próprio bem avaliando, outras possuem apenas mapas ilustrativos das vias, que, porém, são elaborados sem escala. Essa escassez de informação dificulta a pesquisa de imóveis para a modelagem de avaliação de imóveis, diante disso, nos deparamos com ruas sem indicativo de nome, dificultando posterior marcação dos locais das mesmas em mapas analógicos e digitais. Uma Planta cadastral em formato digital permite a obtenção das coordenadas de forma precisa e rápida. Já as ferramentas como: o Google Maps, Google Earth, Bing Maps e entre

outros sistemas de informação Geográfica que possuem acesso gratuito, oferecem imagens de boa resolução para os principais municípios brasileiros, carente deste tipo de informação. O problema que se têm é que grande parte destas imagens não está disponível em escala adequada e não informa a qualidade posicional em seus metadados, quando se tem. Outra ferramenta disponível aos gestores públicos dos municípios brasileiros são os receptores GPS/GNSS de navegação que têm baixo custo, porém, sabe-se que sua precisão é baixa. Considerando que a precisão na obtenção das coordenadas é diferente em cada tipo de ferramentas, se pretende investigar, a partir da bibliografia especializada, como se comporta esta diferença na precisão no valor de mercado para os imóveis residenciais.

A questão é que a maioria dos avaliadores toma como verdade que as informações do imóvel foram verificadas nas documentações obtidas (matrícula e IPTU), que são, por premissa, consideradas boas e válidas, mas muitas das vezes essas informações não são fidedignas e precisam ser verificadas. A matrícula é responsabilidade do serviço de registro, a segunda é de responsabilidade do serviço de cadastro do município. Essas duas fontes de informação deveriam ter uma única resposta, mas que na prática isso não acontece devido à falta de intercâmbio de informações entre Cadastro e Registro de Imóveis. Diante desses questionamentos, se faz necessário um estudo detalhado de qual é a influência da qualidade da informação posicional para a determinação do valor de mercado de um imóvel com precisão.

Outra questão que norteia o desenvolvimento desta pesquisa é o interesse do Governo Federal Brasileiro em desenvolver um índice de preço de imóveis para o Brasil, comprovada com a aprovação do decreto que institui o Sistema Nacional de Gestão de Informações Territoriais – Sinter que foi recentemente regulamentado pelo disposto no art. 41 da Lei nº 11.977, de 7 de julho de 2009 e que implementa Índice de Preços de Imóveis a que se refere o Decreto nº 7.565, de 15 de setembro de 2011. Este trabalho poderá subsidiar parâmetros metodológicos, de avaliação geoespacial para estas ferramentas de gestão que estão sendo implementadas no país. Visto que o país ainda não possui uma base cartográfica cadastral confiável para todo o território nacional.

Diante desta problemática, pretende-se aplicar as ferramentas e métodos da engenharia de avaliações, em especial ao uso e aplicação de geoestatística inferencial em regressões múltiplas pelo método comparativo de dados do mercado, para ampliar e definir parâmetros para um tipo de avaliação cada vez mais comum nas cidades brasileiras, neste caso a cidade do Recife, bairro do espinheiro e adjacências, que consiste na avaliação em massa de apartamentos residenciais com a utilização dos métodos de geoestatística espacial. Sabe-se que nem sempre informações como a área, preço, localização e entre outras variáveis explicativas são suficientes para se firmar uma decisão de estimação de valor de determinado imóvel ofertado, especialmente quando se trata de imóveis de mesma região e localizações distintas, visto que o valor de determinado imóvel influencia nos valores de imóveis próximos, a certo raio de contágio espacial que explica parte dos resíduos ajustados (DANTAS et al, 2004). No caso, o trabalho se utiliza de ferramentas de regressão espacial e metodologia clássica de regressão, devido que ao se implementar a influência espacial se obtém modelos muito mais refinados que a metodologia clássica.

A estimação do valor de mercado por geoestatística e regressão espacial se deu através da definição do Método comparativo direto de dados de mercado, onde se identifica o valor de mercado do bem por meio de tratamento técnico dos atributos dos elementos comparáveis, constituintes da amostra. No presente estudo, onde através dos dados coletados se propiciou a modelagem da equação que define e mensura o comportamento do valor dos imóveis, a equação modelada contempla e condiciona a estimação do valor de mercado com todas as influências espaciais dos dados georreferenciados. Foi utilizado software de modelagem espacial como o Geoda e o SpaceStat para os testes de significância, resíduos e entre outros testes estatísticos, sob a ótica da interpretação da engenharia de avaliações. Para coleta, análise e visualização e representação dos dados geoespaciais, se usou os recursos do QuantunGis - QGIS. Usando como base cartográfica duas bases de informação: o Google Earth e o ESIG da Prefeitura do Recife-PE.

O objetivo deste trabalho é analisar a influência de erros de coleta de informações posicionais dos imóveis utilizados como amostra no desenvolvimento de modelos de regressão espacial, focalizando na avaliação da qualidade da informação posicional destas duas ferramentas de aquisição de dados posicionais: Google Earth e GPS/GNSS. Analisando a eficácia da usabilidade dessas formas de aquisição para avaliação de imóveis em massa. Os resultados indicados pela pesquisa bibliográfica apresentada, a partir de três tipos de aquisição de dados geométricos de posição, ESIG, GPS/GNSS e Google Earth, indicam que são ferramentas excelentes, mas que implicam em cuidados e critérios. Sugerindo que em futuras atualizações da norma de avaliação de bens NBR 14.653-2 contemplem normatizações para avaliações em massa, e que as formas de aquisição dos dados espaciais pesem no grau de fundamentação e precisão.

### 1.1 Desenvolvimento da representação de dados espaciais para avaliação em massa no Brasil.

Em um trabalho publicado em 1880, sobre “As teorias matemáticas e físicas da geodesia superior”, o cientista alemão F. R. Helmert definiu a Geodesia como: “A ciência da medição e representação da superfície da Terra”. (KAHMEN & FAIG, 1988).

As atividades geodésicas têm experimentado uma verdadeira revolução com o advento do Sistema de Posicionamento Global (GPS), hoje denominado de Sistema Global de Navegação por Satélite (GNSS). A capacidade

que este sistema possui de permitir a determinação de posições, estáticas ou cinemáticas, aliando rapidez e precisão muito superiores aos métodos clássicos de levantamento, provocou a necessidade de repensar metodologias de aquisição de dados espaciais, visto que isto proporcionou grandes avanços na qualidade e rapidez de geração de produtos cartográficos analógicos. Com os avanços das ciências da computação, motivaram que os mesmos passassem a ser digitais, e posteriormente propicia o surgimento dos mapas na web (Google Maps, Google Earth e OpenStreetMaps). Isto proporcionou uma grande demanda de aplicações, dentre elas, suporte a avaliação em massa de imóveis.

A dificuldade em cartografia é transferir tudo o que existe numa superfície curva, como a Terra, para uma superfície plana que é o mapa. Considerando-se que uma figura esférica não se desdobra em um plano, permanece na planificação deformações. A confecção de um mapa exige, antes de tudo, o estabelecimento de um método segundo o qual, a cada ponto da Terra corresponda a um ponto no mapa e vice-versa. Diante disso, vem a dificuldade de se especializar uma grande quantidade de dados amostrais em um país de dimensões continentais, como é o caso do Brasil, com intuito de desenvolver um Índice de Preços de Imóveis para todo território nacional, baseado na criação de um sistema de gestão territorial como é o caso do Sinter - Sistema Nacional de Gestão de Informações Territoriais, estabelecido pelo decreto executivo 8.764/2016 de 10 de maio de 2016.

## 1.2 Avaliação de imóveis e sistema de informação territorial

O conjunto de conhecimentos nas áreas de engenharia e arquitetura, bem como de outras ciências, a fim de se determinar tecnicamente o valor, frutos e custo de um bem, é o ramo da engenharia que se chama Engenharia de Avaliações (DANTAS, 1998). A avaliação de imóveis, sejam urbanos ou rural faz-se presente na maioria dos negócios, pendências entre pessoas, empresas ou ambos. Geralmente é necessário avaliar imóveis para a compra e venda dos mesmos, na determinação do preço real de empresas, em atendimento à legislação, na partilha de heranças, divórcios, no lançamento de impostos, nas hipotecas imobiliárias, na demarcação de terras, nas indenizações, enfim, em um número expressivo de ações inerentes aos relacionamentos humanos, onde o valor de um bem assume importância crucial. Diante disso são necessárias informações qualitativas e quantitativas sobre os imóveis que foram vendidos ou estão prestes a ser vendidos. A maioria dessas informações são retiradas dos bancos de dados cadastrais de imóveis que podem ser públicos ou particulares.

Segundo Barros (2011) os bancos de dados cadastrais apresentam necessidades especiais de rigor, por isso o controle de qualidade é uma questão fundamental na construção e atualização de sistemas de informações cadastrais. Estes Sistemas de Informações Cadastrais subsidiam na determinação do valor de mercado a partir de seu uso. Um problema enfrentado pelos profissionais da avaliação e as prefeituras brasileiras (usuários) é a falta de qualidade da informação posicional dos imóveis, pois estes dados são coletados muitas vezes sem critérios ou a partir da medição tirada de produtos cartográficos cadastrais, sem especificações técnicas de qualidade satisfatória para essa atividade, como plantas, cartas cadastrais, mapas, ortofotocartas, SIG's ou até croquis. Segundo Silva (2003), para que um produto cartográfico atenda aos objetivos do usuário é necessário que a metodologia para a confecção de tal documento atenda a certos padrões de qualidade, padrões estes que garantirão um produto final que supra de maneira eficaz as necessidades do usuário. De toda forma, faz-se necessário o estabelecimento de indicadores de qualidade, que propiciarão aos usuários o atendimento às suas necessidades. Os produtos cartográficos não oficiais sem padrão de qualidade como o Google Earth e entre outros dificultam a identificação da incerteza posicional das amostras.

## 1.3 Área de estudo

Na região do bairro do Espinheiro e adjacências, localizado no Grande Recife em Pernambuco, é a área que se considerou na modelagem. Região primordialmente residencial, porém, foram encontradas escolas, padarias, farmácias, restaurantes e igrejas. A região compõe um complexo de bairros da zona norte do Recife, de considerável especulação imobiliária, com poucos terrenos em oferta e boa oferta de serviços e infraestrutura urbana. (Ver Figura 1). Conforme vistoria de reconhecimento da região, se verificam alguns imóveis disponibilizados em oferta, com indicação de liquidez média.

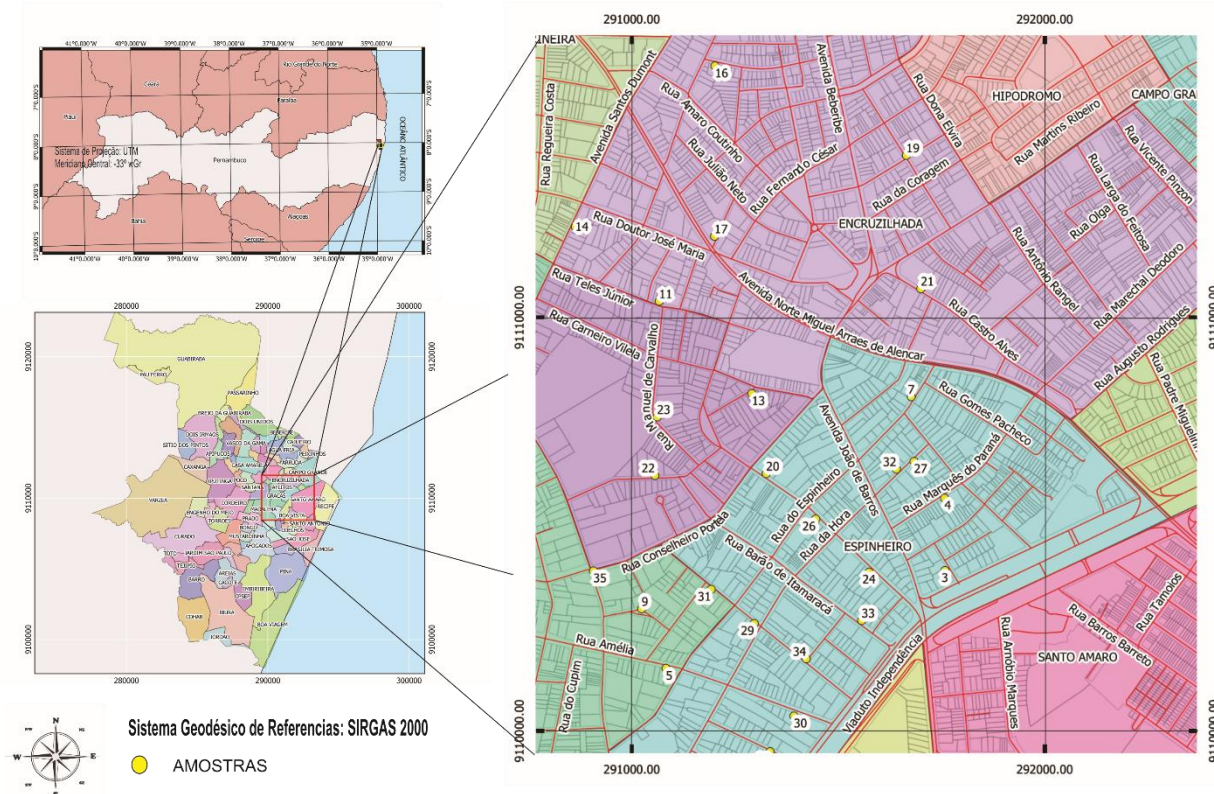


Figura 1—Área de Estudo

## 2 MODELAGEM ESPACIAL DE AVALIAÇÃO DE IMÓVEIS

No planejamento de uma pesquisa, o que se pretende é a composição de uma amostra representativa de dados de mercado de imóveis com características, tanto quanto possível, se usando toda a evidência disponível. Esta etapa, que envolve estrutura e estratégia da pesquisa, se deve iniciar pela caracterização e delimitação do mercado em análise, com o auxílio de teorias e conceitos existentes ou hipóteses da formação do valor.

A coleta de dados foi realizada com informações coletadas através de pesquisa de campo, corretor de imóveis, informações de proprietários, dados de setores censitários do IBGE, sites de imobiliárias e com o uso complementar de informações por telefone. Foram consideradas informações de coleta: o valor de venda, a área do imóvel, coordenadas projetivas UTM, quantidade de quartos, quantidade de suítes, padrão de acabamento e conservação, IDH dos bairros, renda média mensal dos chefes de família para o bairro (Fonte: IBGE), valor cobrado do condomínio e seu respectivo andar, sendo que nem todas foram consideradas na modelagem final, por não possuírem correlação significativa com a variável explicada e por questões de micronumerosidade. Sendo que no desenvolver do tratamento estatístico, pelas condicionantes dos estudos da regressão múltipla, foram evidenciados dados com resíduos bastante elevados, ou até com ofertas mais elevadas que o da prática do mercado, diminuindo assim o poder de explicação do modelo, portanto foram dados desconsiderados por serem *outliers*. A quantidade de dados obedeceu a função  $N = 7,5*(z+1)$ , onde,  $N$  é o número mínimo de dados e  $Z$  é o número de variáveis independentes. Visto que para análises espaciais, para um melhor ajustamento se faz necessário um vasto número de amostras para uma boa modelagem, portanto se utilizou da quantidade máxima de dados possíveis à realização do presente trabalho.

### 2.1 Regressão e Dependência Espacial

A preocupação estatística ao analisar dados, é a de criar modelos que explicitem estruturas do fenômeno em observação. O modelo de regressão é um dos métodos estatísticos mais usados para investigar a relação entre variáveis. A análise de regressão é a metodologia estatística que estuda (modela) a relação entre duas ou mais variáveis, no caso, a regressão envolve três ou mais variáveis, portanto, estimadores. Ou seja, ainda uma única variável dependente, porém duas ou mais variáveis independentes (explanatórias), (DANTAS, 1998). A finalidade das variáveis independentes adicionais é melhorar a capacidade de predição em confronto com a regressão linear simples. Isto é, reduzir o coeficiente do intercepto, o qual, em regressão, significa a parte da variável dependente explicada por outras variáveis,

que não a considerada no modelo. Mesmo quando estamos interessados no efeito de apenas uma das variáveis, é aconselhável incluir as outras capazes de afetar Y, efetuando uma análise de regressão múltipla, por duas razões: Para reduzir os resíduos estocásticos. Reduzindo-se a variância residual (Erro Padrão da Estimativa), aumentando a força dos testes de significância. E, por conseguinte, eliminar a tendenciosidade que poderia resultar se simplesmente ignorássemos uma variável que afeta Y substancialmente.

O ideal é obter o mais alto relacionamento explanatório com o mínimo de variáveis independentes, sobretudo em virtude do custo na obtenção de dados para muitas variáveis e também pela necessidade de observações adicionais para compensar a perda de graus de liberdade decorrente da introdução de mais variáveis independentes. Para o caso, ainda se faz necessário o confronto entre o ajustamento por Regressão Clássica e o ajustamento por Regressão Espacial, os testes são fundamentais para a análise do comportamento dos ajustamentos através de testes estatísticos específicos, a fim de identificar questões de significância global do modelo, significância individual dos parâmetros, coeficiente de ajustamento do modelo, análise de multicolinearidade pela matriz de correlações, teste de normalidade, de homoscedasticidade, teste de Moran I, teste dos multiplicadores de Lagrange e dentre outros como Akaike, Schwartz, Jarque-Bera e Breusch-Pagan,. Ainda com todos os testes citados, se faz indispensável as observações no que diz respeito a análise exploratória, fundamental em qualquer tratamento estatísticos dos dados.

O modelo pode ser representado por  $(y = \beta_0 + x_1\beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \varepsilon_i)$ , para n observações e k variáveis, que pela forma matricial se apresenta  $(Y = X\beta + \varepsilon)$ , sendo X a matriz das variáveis independentes e Y,  $\beta$  e  $\varepsilon$  os vetores de retorno, de parâmetro e de resíduos da regressão, respectivamente (GUJARATI, 2003). Para aprimorar o paradigma ao conjunto, se utiliza do Método dos Quadrados Mínimos com o propósito de minimizar os resíduos, onde o estimador minora a soma dos quadrados dos resíduos  $(\sum_{i=0}^n e_i^2)$  para ajustar ao máximo o modelo aos dados (DANTAS, 1998), sendo os parâmetros estimados por  $(b = (X'X)^{-1} X'y)$ .

Para eficiência desse método, se dispõem por premissa que os regressores sejam fixos e sem perturbações, que os parâmetros sejam constantes, que o modelo linear ocorra na forma de  $(y = X\beta + \varepsilon)$ , e que os erros apresentem distribuição normal, ausência de correlações entre as variáveis independentes, homoscedasticidade e aleatoriedade com esperança nula  $(E(\varepsilon) = 0)$ . Visto que a medida da qualidade do modelo em relação a estimação correta do vetor resposta será dada pelo coeficiente de determinação (2.1), atendendo também aos testes de significância dos coeficientes para estimadores com distribuição centrada em  $\beta$  e variância do erro  $(\sigma^2(X'X)^{-1})$ , dado pela estatística t para a variável j na forma de (2.2), onde  $a_{jj}$  é elemento de  $((X'X)^{-1})$ , e aplicando  $t_j$  na curva acumulada da distribuição t de Student com  $(n - k - 1)$  graus de liberdade, e para o teste de significância global do modelo ao utilizar o teste F através da estatística (2.3), onde testa a hipótese nula  $(H_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0)$  para nenhuma variável explicando o fenômeno e a hipótese alternativa  $(H_1$  ao menos  $\beta_j(j = 1, \dots, k) \neq 0)$  onde ao menos uma variável contribui na variabilidade de preços. No entanto, mesmo atendendo esses pressupostos, as variáveis escolhidas devem apresentar relevância e adequabilidade para garantia de um modelo linear, consistente e eficiente (DANTAS, 1998).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.1)$$

$$t_j = \frac{b_j}{s\sqrt{a_{jj}}} \sim t(n - k - 1) \quad (2.2)$$

$$F_c = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \hat{y}_i)^2} \times \frac{(n - k - 1)}{k} \quad (2.3)$$

Mesmo atendendo esses pressupostos, pela ausência dos efeitos espaciais e por questões da própria sensibilidade desse modelo a dados inconsistentes, a metodologia apresenta tendenciosidade e ineficiência (ANSELIN, 1988). Dada a moderna proposta de trabalho proposta por (DANTAS et al, 2004), chamada de Avaliação por Inferência Espacial, da qual se utiliza do processo do contágio espacial, (ANSELIN, 1988), os efeitos residuais espaciais são facilmente parametrizados e contemplados à parte explicada do modelo, se adequando a garantia de um modelo linear, consistente e eficiente (DANTAS, 1998). O diagnóstico da presença de efeitos espaciais pode se dá por métodos de testes estatístico específicos, como o de Moran I e pelo Multiplicador de Lagrange (LM), nesse caso elaborando e se utilizando de matrizes de Pesos Espaciais de vizinhança.

Nesse caso, os testes LM indicam qual metodologia a ser utilizada para a incorporação dos efeitos espaciais ao modelo, que pode ser o Modelo de Erro Espacial (ANSELIN, 1988) é dado matricialmente por  $(Y = X\beta + u)$ , onde  $(u = \lambda Wu + \varepsilon)$  e  $y \approx N(0, \sigma^2 I)$ , sendo W a matriz de pesos espaciais e  $\lambda$  o coeficiente de autocorrelação espacial do erro u. Ou ainda a incorporação dos efeitos espaciais através do Modelo de Defasagem Espacial (ANSELIN, 1988), matricialmente por  $(Y = \rho WY + X\beta + \varepsilon)$ , onde  $y \approx N(0, \sigma^2 I)$ , sendo WY a variável dependente espacialmente defasada e  $\rho$  o coeficiente de autocorrelação espacial de WY. Nesses modelos se estima os parâmetros via maximização

da função de Log verossimilhança. Feito isso, se entende que das amostras extraídas foi estimada a aproximação do verdadeiro comportamento da população, e por sua vez eficiente para o processo avaliatório.

## 2.2 Diagnóstico da dependência espacial

O início do diagnóstico de dependência espacial parte da análise dos testes estatísticos de Moran I (2.4), Multiplicador de Lagrange LM erro (2.5), e Multiplicador de Lagrange LM defasagem (2.6). No caso da estatística de Moran, se identifica que os resíduos, desde a modelagem clássica de regressão, apresentam correlação espacial, não identificando, porém, qual o tipo de efeito espacial ocorre no modelo, dada a condição da hipótese nula para estatística não significativa. Sendo assim, seguindo as mesmas condições de Moran, deve se dá a análise das estatísticas LM erro e LM defasagem, onde a distinção entre as significâncias dos testes se tornam evidentes, se tomando então o mais significativo para realização do modelo, (ANSELIN, 1988). Ainda sobre a comparação entre modelos, o uso dos testes de Akaike e Schwarz, fornecem uma medida de melhor ajustamento para cada modelo.

$$I = \left(\frac{n}{S}\right) \left[ \frac{e'We}{e'e} \right] \quad (2.4)$$

$$LM(erro) = \frac{\left[ \frac{e'We}{(s^2/n)} \right]^2}{[tr(W^2 + W'W)]} \quad (2.5)$$

$$LM(defasagem) = \frac{\left[ e \frac{Wy}{s^2} \right]^2}{(WXb)' \frac{MWXb}{s^2} + tr[W'W + W^2]} \approx X^2 \quad (2.6)$$

## 2.3 Modelos espaciais gerados

Para as modelagens utilizadas nesse estudo, foi utilizado como variável explicada o PU (Preço Unitário por metro quadrado), e como variáveis explicativas: Área (m<sup>2</sup>), Padrão de acabamento do apartamento (*dummy* por tricotomia: baixo, médio e alto) e Número de Quartos (suítes consideradas apenas como quarto), contudo, mais atributos foram considerados na coleta dos dados, embora a posteriori por questões de micronumerosidade fossem descartadas em função da significância individual dos parâmetros.

Em relação à informação da posição espacial, se utilizou três métodos de aquisição de dados, plataforma ESIG (Prefeitura do Recife) como banco de dados oficial, GPS com utilização de GPS/GNSS de navegação, e plataforma Google Earth, no caso, para cada dado do modelo foi realizada essas três formas de aquisição, cuja finalidade foi de se obter o melhor modelo de regressão para cada tipo de aquisição.

Tabela 1 – Resumo das saídas Obtidas das Regressões com o Uso do SpaceStat

	ESIG			Google			GPS		
	Clássico	LAG	ERRO	Clássico	LAG	ERRO	Clássico	LAG	ERRO
R <sup>2</sup>	0.738227	0.745727	0.713736	0.738227	0.743769	0.713104	0.738227	0.741015	0.762016
P value	0	0	0	0	0	0	0	0	0
AKAIKE	467.9215	468.5292	467.6391	467.9215	468.8574	467.2269	467.9215	469.4511	467.7215
SCHWARTZ	473.5262	475.5352	473.2439	473.5262	475.8634	472.8317	473.5262	476.4571	473.3263
Variable	Probability	Probability	Probability	Probability	Probability	Probability	Probability	Probability	Probability
Coef		0.154182	0.499759		0.246388	0.261226		0.452389	0.606827
Intercept	0.0	0.249758	0.0	0.0	0.254172	0.0	0.0	0.125582	0.0
Área	0.000006	0.0	0.0	0.000006	0.0	0.0	0.000006	0.0	0.0
Padrão	0.000003	0.0	0.0	0.000003	0.0	0.0	0.000003	0.0	0.0
Quartos	0.105523	0.067617	0.096054	0.105523	0.082611	0.106584	0.105523	0.062446	0.06085
Jarque-Bera	0.434493			0.434493			0.434493		
Breush-Pagan	0.668702	0.874737	0.779973	0.668702	0.840466	0.857738	0.668702	0.815607	0.561934

Diante dos resultados obtidos, para os modelos espaciais do ESIG e Google Earth se adotou o modelo da Defasagem Espacial devido aos testes de LM (Defasagem), poder de explicação do modelo e significância individual dos parâmetros apresentarem resultados determinantes, embora os testes de Akaike e Schwartz terem confirmado a hipótese do Erro Espacial. Para o caso do GPS, os testes do Multiplicador de Lagrange indicaram um alto grau de significância, minorando assim a hipótese da influência espacial, sendo a regressão clássica o modelo melhor explicado. Para todas as regressões os testes de normalidade e homoscedasticidade apresentaram aderência.

Tabela 2 - Resumo das Saídas dos Testes Estatísticos do Multiplicador de Lagrange

Test	ESIG			Google			GPS		
	MI/DF	Value	Probability	MI/DF	Value	Probability	MI/DF	Value	Probability
Moran's I (error)	0.037062	1128007	0.259317	0.058195	1511135	0.130754	-0.02946	0.069468	0.944617
Lagrange Multiplier (error)	1	0.213086	0.644359	1	0.541078	0.461986	1	0.134814	0.713492
Robust LM (error)	1	0.397759	0.528249	1	0.216284	0.641886	1	183329	0.175739
Lagrange Multiplier (lag)	1	1248609	0.263818	1	114739	0.284096	1	0.447402	0.503571
Robust LM (lag)	1	1433282	0.23123	1	0.822596	0.364422	1	2145878	0.142953
(SARMA)	2	1646368	0.439032	2	1363674	0.505687	2	2280692	0.319708

### 3 INFLUENCIA DA INCERTEZA POSICIONAL NA AQUISIÇÃO DOS DADOS ESPACIAIS

A precisão é o principal requisito das avaliações de imóveis e o único que pode ser mensurado objetivamente. As avaliações não podem ser comprovadas diretamente, pois os imóveis geralmente não são vendidos logo após as avaliações. Ademais, normalmente os imóveis não são transacionados necessariamente pelo valor de mercado. Assim, a precisão das avaliações é verificada indiretamente, pela comparação do valor estimado pelo modelo com os preços de venda para os imóveis da amostra.

Em princípio, as avaliações devem ter precisão, mas a imprecisão é relacionada com a parte não explicada do modelo. Os erros seriam devidos à falta de um controle ou registro central das transações, à heterogeneidade dos imóveis, à confidencialidade das informações e as oscilações normais do mercado. A diferença entre preços de venda e valores estimados pode ser decorrente, além de erros de modelagem, do próprio processo de formação dos preços (de acordo com o funcionamento do mercado imobiliário), pois os preços praticados incorporam as distorções do mercado, tais como as diferenças de informação entre os agentes. Outro fator que influencia na precisão da avaliação, mas que não se tem parâmetros de quanto influência no valor do bem avaliado, é a qualidade da informação métrica e posicional das amostras.

#### 3.1 Qualidade da informação Posicional.

As coordenadas de um ponto, georreferenciado ou não, geralmente são grandezas obtidas indiretamente a partir das medições de outras, por exemplo, a definição de uma área a partir das medições dos limites dos seus lados, a obtenção das coordenadas cartesianas de um ponto a partir das medições polares (ângulos e distâncias) – bi ou tridimensional – entre outras grandezas. Segundo Galdino (2006), se as grandezas não forem medidas diretamente, mas derivadas de outras medidas com desvios casuais derivados destas grandezas (por exemplo a área de um retângulo é produto dos seus lados), deve ser observado como as variâncias dos dados de saída se propagam sobre os valores medidos - lei de propagação dos erros.

Desta forma procede-se com a determinação da propagação dos erros, obtendo-se ao final dos procedimentos, a variância resultante deste acúmulo de erros propagados pela incerteza e natureza da medição. A avaliação de imóveis usa a teoria da estimação, que é o processo que usa os resultados extraídos da amostra para produzir inferência sobre a população da qual foi extraída aleatoriamente da amostra. A aplicação da propagação do erro, está em concordância com o método de aquisição de dados.

O uso, por exemplo, das imagens de alta resolução espacial obtidas por meio de sensores aerotransportados ou a bordo de plataformas orbitais, se apresenta com uma solução para estimativas, como fontes de informação posicional e geométrica de um imóvel, desde que esteja ortorretificadas e georreferenciadas. Além da utilização de imagens de satélite de alta resolução, faz parte do escopo do trabalho de alguns profissionais a utilização de um receptor GPS/GNSS de navegação como principal instrumento de aquisição de amostras, utilizando as observáveis GPS/GNSS coletadas a partir de métodos de posicionamentos absoluto.

Um correto posicionamento é fundamental na determinação precisa do valor do bem avaliando, enfim, elementos que devem ter sua influência verificada e que dependem essencialmente da perfeita localização dos dados coletados no

campo, em mapas, ou plantas. O uso de informações posicionais de base cartográfica cadastrais, imagens de satélite, Google Earth, levantamentos georreferenciados com o uso de receptores GNSS/GPS de navegação sem a adoção de critérios qualidade geométrica, podem mascarar o resultado de uma avaliação, uma vez que todo bem avaliado possui influência posicional agregado ao seu valor. Entretanto, não existem parâmetros legais que estabeleçam a precisão posicional de um imóvel urbano, ficando o mesmo limitado a informação coletada nas bases cartográficas cadastrais, quando existirem, e limitadas a escala do produto cartográfico.

### 3.2 Verificação da Incerteza Posicional

Para se analisar a incerteza posicional dentre as formas de aquisição dos dados espaciais GNSS/GPS e Google Earth, se tomou como duas referências aceitas como confiáveis para se realizar a análise de melhor concordância dentre os métodos. Sendo o primeiro, os dados do próprio atributo explicado utilizado para as modelagens, no caso, o Preço Unitário real e, por conseguinte os dados do Preço Unitário obtidos pela estimação através do modelo de defasagem espacial do ESIG, visto que dentre as formas de aquisição é o que apresenta confiabilidade superior, por se tratar de dados oficiais e de qualidade posicional superior aos demais.

Desse modo, por se tratar de dois métodos de classificação, se utilizou do índice Kappa (3.1) para se obter uma medida de concordância entre as duas formas de aquisição, baseado nas relações de aderência entre os resíduos do modelo superior, resíduos dos dados reais e resíduos do modelo avaliado, em função do desvio médio estimado dos modelos das aquisições como parâmetro do quão aderente uma forma de aquisição está em relação à outra.

Dessa forma, por elaboração de matrizes de confusão para aquisição tipo GPS e tipo Google Earth, o índice Kappa mensura o grau de concordância e consequentemente as hipóteses de concordância, de perfeita até a hipótese nula, (FLEISS, 1981). Na equação  $p_o$  é a probabilidade de concordância e  $p_e$  é a probabilidade hipotética de acaso.

Tabela 3 - Matrizes de Confusão Utilizadas na Obtenção do Índice Kappa.

GPS/ESIG	GPS/Real			Google/E SIG	Google/Real		
	Dentro	Fora	TOTAL		Dentro	Fora	TOTAL
Dentro	7,000	4,000	11,000	Dentro	14,000	1,000	15,000
Fora	9,000	10,000	19,000	Fora	1,000	14,000	15,000
TOTAL	16,000	14,000	30,000	TOTAL	15,000	15,000	30,000

$k = 0,1484716$ 
 $k = 0,8666667$

Visto isso, conforme (LANDIS & KOCH, 1977), a sugestão da interpretação dos índices obtidos é que a concordância de GPS, com índice de 14% é considerada “Ruim”, e no caso da concordância de Google Earth, com índice de 86% é de “Excelente” aderência. Dessa forma, os testes com utilização de matriz de confusão e índice Kappa indicam que os resultados da forma de aquisição de dados espaciais através do Google Earth são mais aderentes e correlacionados com a superfície real do comportamento dos preços dos imóveis nos bairros do estudo.

$$k = \frac{p_o - p_e}{1 - p_e} \quad (3.1)$$

Os recursos tecnológicos disponíveis à cartografia têm viabilizado avanços consideravelmente importantes, através, principalmente, dos softwares que permitem velocidade na obtenção de dados, na troca de informações espaciais e em sofisticadas análises dos fenômenos espaciais. Tradicionalmente, tanto produtores como usuários de dados espaciais reconhecem sua importância, mas dedicam poucos esforços para entender a qualidade dos dados espaciais e para a avaliação dos documentos cartográficos em meio analógico. A ausência de dados precisos implica no fornecimento de informações não confiáveis, que consequentemente irão gerar erros comprometedores à qualidade dos estudos.

As imagens disponíveis no Google Earth são atualizadas com frequência e não são disponíveis em tempo real. A resolução espacial das imagens do Google Earth para a maioria dos locais é de 15 metros (SILVA & NAZARENO, 2009).

O procedimento para a análise da qualidade posicional aconteceu em dois momentos:



- Na análise das discrepâncias entre as coordenadas dos pontos retirados do centroide dos lotes no Google Earth e as coordenadas dos pontos homólogos no ESIG;
- Na análise de discrepância entre as coordenadas dos pontos coletados com receptor GNSS de navegação na frente do lote e as coordenadas dos centroides dos lotes no ESIG.

A dificuldade encontrada foi a visualização e a identificação dos limites das propriedades que são usadas como amostras no Google Earth que geralmente não condiz com que se encontra *in loco*. Uma vez que, os limites físicos não coincidem com limite legal, assim como o limite visual no mapa. Na imagem do Google Earth só é possível identificar o limite físico e não o legal, tornando a descrição e identificação do imóvel usado para avaliação, incompleta e imprecisa.

A coleta dados com receptor GNSS de navegação teve como dificuldade a ocupação do centroide das dos dados da amostra. Diante disso, adotou-se como metodologia a coleta do centro médio da testada dos lotes das amostras, o que gerou uma grande diferença entre as coordenadas dos centroidese a coordenada do centro médio da testada do lote, associada incerteza posicional das coordenadas adquiridas por este método de levantamento, causado pelas interferênciasdo Sinal GNSS, possível caso de multicaminhamento, e o tipo de receptor utilizado.

#### 4 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

No geral, o trabalho conseguiu realizar a metodologia proposta, foi realizada a coleta de dados dos preços dos imóveis com os respectivos atributos explicativos das modelagens realizadas. Foi utilizado três tipos de aquisição de dados espaciais, ESIG, Google Earth e GPS/GNSS, que deram suporte para utilização de regressões espaciais através da utilização de matrizes de pesos espaciais. O primeiro, ESIG, por se tratar de base cartográfica de qualidade posicional confiável foi utilizado, assim como os próprios parâmetros da amostra, como parâmetro para a concordância posicional desses últimos, Google Earth e GPS/GNSS. Para cada forma de aquisição foi realizado um modelo de regressão clássica e por testes estatísticos, a análise da influência espacial sobre o modelo. Conforme os pressupostos deste tipo de estatística se adotou modelo espacial conforme a indicação dos referidos testes, tendo como hipótese nula a própria regressão clássica. Para todas as situações o poder de explicação do modelo, significância global do modelo, significância individual dos parâmetros, normalidade e homoscedasticidade foram satisfatórios. Para ESIG e Google Earth a modelagem por defasagem espacial foi a adotada, enquanto que GPS/GNSS não satisfaz as condições de espacialidade sendo adotado o próprio modelo clássico.

De posse do melhor modelo para cada forma de aquisição, se iniciou a comparação entre a aquisição GPS e Google Earth, a fim de se encontrar qual técnica a apresentar melhor aderência a superfície real. Metodologia pela qual se fez uso de matrizes de confusão, e tomou o índice Kappa como fator determinante para a análise de concordância posicional. Os índices Kappa mostraram que a aderência da forma de aquisição GPS/GNSS foi insatisfatória, enquanto que a forma de aquisição Google Earth apresentou concordância excelente, que pode ser justificado pela consideração da utilização da posição do centroide dos referidos lotes que diferem entre uma forma de aquisição e a outra.

É interessante ressaltar aos usuários de imagens fornecidas pelo Google Earth que se deve ter cautela acerca dos limites de precisão e das aplicações possíveis, uma vez que por trás de uma alta resolução espacial podem conter erros desconhecidos, que podem ser desprezados pelos usuários comuns, podendo trazer consequências nas decisões apoiadas sobre estas imagens. Finalmente, se recomenda que trabalhos desta natureza sejam realizados e divulgados de modo a alertar sobre a necessidade de se verificar qual a qualidade do produto cartográfico disponível e a necessidade de cuidados em relação aos erros grosseiros no ato da aquisição destas posições, principalmente em avaliações de imóveis.

#### REFERÊNCIAS

ANSELIN, L. (1988) – **Spatial Econometrics: Methods and Models**. Dordrecht: KluwerAcademic.

BARROS, E.R.O. **Uma Proposta para o Controle de Qualidade do Processo de Certificação de Imóveis Rurais**. Programa de Pós-graduação em Ciências Geodésicas e Tecnologias da Geoinformação (dissertação). Recife: UFPE. 2011. 170p

BONISSONE, P. P.; CHEETAM, W.; GOLIBERSUCH, D. C.; KHEDKAR, P. **Automated residential property valuation: An accurate and reliable approach based on soft computing**. In: RIBEIRO, R.; ZIMMERMANN, H.; YAGER, R. R.; KACPRZYK, J. (Eds.). *Soft computing in financial engineering*. Heidelberg: Physica-Verlag, 1998.

DANTAS, Rubens Alves. (1998) – **Engenharia de Avaliações: uma introdução à metodologia científica.** 1ª ed. São Paulo: PINI.

DANTAS, R., SÁ, L. e PORTUGAL, J. (2004) – **A importância da inferência espacial na elaboração de planta de Valores.** Anais do XXI Congresso Panamericano de Avaliações, Cartagena – Colômbia.

FLEISS, Joseph L. **Statistical methods for rates and proportions.** New York: John Wiley, 1981. p 212-236.

GALDINO, Carlos Alberto Pessoa Mello. **Cadastro de Parcelas Territoriais vinculadas ao Sistema de Referência Geocêntrico – SIRGAS 2000.** Dissertação de Doutorado Programa de Pós-Graduação em Engenharia Civil - PPGEC da Universidade Federal de Santa Catarina – UFSC, 2006.

GILBERTSON, B. G. **Appraisal or valuation: An art or a science? Real Estate Issues, v.26, n.3, p.86-89, Fall, 2001.**

KAHMEN, H.; FAIG, Wolfgang. **Surveying.** Berlim: De Gruyter, 1988. 577p. ISBN 3-11-008303-5.

LANDIS, J.R. e KOCH, G.G. **The measurement of observer agreement for categorical data.** Biometrics, v.33, n.1, p. 159-174, 1977.

MATHERON, G. (1965) – **Les Variables Regionalisées et Leur Estimation.** Masson, Paris.

SILVA, ArlyGautéama Rodrigues; SILVA, Daniel Luís Andrade. **A Avaliação da Qualidade de documentos Cartográficos Aspectos Geométricos, Temáticos e Temporais.** Instituto Militar de Engenharia IME, Rio de Janeiro 2003.

SILVA, L.A.; NAZARENO, N.R.X. **Análise do Padrão de Exatidão de Cartográfica da imagem do Google Earth tudo como área de estudo a imagem da cidade de Goiânia.** Anais XVI Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil, 25-30 abril 2009, INPE, p. 1723-1730.