



Artigo submetido ao Curso de Engenharia Civil da UNESC -
como requisito parcial para obtenção do Título de Engenheiro Civil



UTILIZAÇÃO DE MODELO DE REGRESSÃO ESPACIAL PARA AVALIAÇÃO DE TERRENOS NA CIDADE DE CRICIÚMA, SC.

Leonel de Moura Brizola Valduga (1), Evelise Chemale Zancan (2).

UNESC – Universidade do Extremo Sul Catarinense
(1)leonel_valduga@hotmail.com, (2)ecz@unesc.net

RESUMO

Este trabalho apresenta o estudo da utilização de modelagem por regressão espacial aplicada ao mercado de terrenos da cidade de Criciúma – SC. O modelo foi obtido por uma amostra de 460 dados, aferidos e interpretados por 5 variáveis independentes formadoras de valor: Área total, Distância a via principal mais próxima, Número máximo de pavimentos, Renda e Avenida Centenário. Criou-se uma matriz de vizinhança correspondente ao inverso do quadrado da distância entre os pontos e realizou-se um modelo de regressão linear múltipla. Os testes LM do erro e LM da defasagem espacial, apresentaram-se significantes ao nível de 1%, confirmando assim a autocorrelação espacial no modelo. Visto isso, realizou-se o modelo do erro espacial e o modelo da defasagem espacial. Todos os modelos enquadraram-se nos critérios da NBR 14.653-2 (2011). Os modelos foram comparados por meio dos critérios de Akaike e Schwarz, que representam a qualidade do ajuste do modelo. O modelo do erro espacial apresentou menores valores nos critérios em relação aos outros modelos, sendo então o mais indicado a ser utilizado para a avaliação de terrenos na cidade de Criciúma – SC.

Palavras-Chave: Regressão Espacial, Variável, Avaliação, Terrenos.

1. INTRODUÇÃO

Avaliar um bem é estimar o seu valor de mercado em data determinada. Segundo THOFERN (2008, p.11) o valor de mercado é o preço justo pago por um imóvel por um comprador desejoso de comprar para um vendedor desejoso de vender, ambos com pleno conhecimento do seu aproveitamento eficiente. Aproveitamento eficiente é aquele recomendável e tecnicamente possível para o local, numa data de referência, observada a tendência mercadológica nas circunvizinhanças entre os diversos usos permitidos pela legislação pertinente.

O respectivo valor de mercado de um bem resulta de um método matemático e estatístico de tratamento de dados, os quais, são extraídos por meio de coleta de

informações de preços de imóveis com características assemelhadas ao avaliando. Segundo COELHO (2011, p. 3) um importante processo no entendimento da avaliação é o conhecimento das variáveis, medidas que assumem valores diferentes, em diferentes pontos de observação, e podem ter uma relação de dependência ou independência sobre algum aspecto. Variáveis dependentes são aquelas que representam o valor de um imóvel, como o valor total ou o valor unitário. Já as independentes são as formadoras do valor do imóvel. Segundo a NBR 14653-2 (2011, p.14) estas variáveis referem-se às características físicas (área, frente, infraestrutura), de localização (bairro, distância a polo de valorização) ou econômicas (a vista ou a prazo, oferta ou venda). Ainda a mesma Norma define como variável quantitativa todas aquelas que podem ser mensuradas e como qualitativa todas aquelas variáveis que não podem ser medidas e nem contadas, apenas hierarquizadas, como por exemplo, topografia do terreno, pedologia e situação na quadra, que podem ser codificadas por meio de:

- 1) Variáveis *Proxy* – de acordo com a NBR 14653-2 (2011, p. 15) essa variável é utilizada para substituir outra de difícil mensuração e que se presume guardar com ela relação de pertinência. É obtida por meio de indicadores publicados ou inferidos em outros estudos de mercado.
- 2) Variáveis *Dummy* ou dicotômicas – são variáveis em que serão atribuídas apenas duas situações: sim ou não; atribuindo-se o valor 0 (zero) quando não possui a característica e 1 (um) caso contrário (Zancan, 1996, p. 70).
- 3) Códigos alocados – a NBR 14653-2 (2011, p. 15) define esta variável como uma escala lógica ordenada para diferenciar as características qualitativas dos imóveis. Por exemplo, o padrão do imóvel, pode ser codificado por: (1) Padrão baixo (2) Padrão médio (3) Padrão alto.

São diversas as metodologias utilizadas para avaliar, ficando a critério do profissional de engenharia de avaliações a escolha de qual método irá suprir de forma mais eficiente as necessidades exigidas pelo serviço de avaliação. Este trabalho, contará com a utilização de dois métodos de avaliação, sendo eles: Comparativo de dados de mercado por regressão linear múltipla (RLM) e regressão espacial (RE), ambos utilizam o tratamento científico através de inferência estatística.

O método comparativo direto de dados de mercado é o mais usualmente empregado, é a comparação direta do imóvel avaliado com outros imóveis de características semelhantes que estão inseridos dentro do mercado imobiliário em questão. Contudo este método só será aplicável se houver um conjunto de dados que seja considerado estatisticamente como uma amostra de mercado (DELFINO, 2013, p.2).

O modelo de regressão linear múltipla tem como objetivo estimar uma função que explique a variação de uma variável dependente Y , em relação a outras variáveis independentes, que são responsáveis pelo seu valor (FERMO, 2010, p.25).

A equação de regressão linear múltipla, com n variáveis é dada pela Equação 1:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad \text{Eq. [1]}$$

α : Intercepto;

β : Parâmetros de regressão das variáveis independentes;

X : Variáveis independentes;

ε : Erros aleatórios.

Uma das dificuldades encontradas na utilização de métodos comparativos de dados de mercado por regressão linear múltipla, na busca de modelos de avaliação, está em considerar a variável localização que pode agregar uma valorização ou desvalorização aos imóveis em questão. A exclusão da localização da modelagem pode causar sérios problemas de predição, pois dados localizados espacialmente (que é o caso dos imóveis), em geral, apresentam autocorrelação ou covariância espacial. Preditores que não levam isso em consideração estão muito longe da realidade (CRESSIE, 1993 apud HORNBERG, 2016, p.134).

“A autocorrelação ou dependência espacial pode afetar o termo de erro, a variável dependente ou ambos. A não consideração deste efeito, pode gerar sérios problemas de estimação, pois, na presença de autocorrelação espacial nos resíduos, os parâmetros “ β ” do modelo de regressão linear múltipla são ineficientes e, neste caso, testes de significância como t de student e F de Snedecor, bem como os intervalos de confiança inferidos, não são mais válidos e as decisões tomadas com base neles são enganosas. Ocorrido isso, o modelo de regressão linear múltipla deverá ser substituído pelos modelos espaciais, estimados por Econometria espacial.” (DANTAS, 2003, p.33).

Segundo HORNBERG (2016, p. 135) em análise de regressão se busca encontrar um bom ajuste ao modelo de dados, de forma a reduzir as diferenças entre o valor gerado pelo modelo e o valor obtido por meio de pesquisa de dados do mercado. Também procura-se compreender quais as variáveis independentes que apresentam considerável significância para a explicação do modelo. Uma das hipóteses, é que no modelo clássico as observações não sejam correlacionadas, sendo assim os resíduos do modelo são independentes e não correlacionados entre si, além de apresentar distribuição normal com média zero e variância constante. No caso onde há dependência espacial, é pouco provável que esta hipótese de observações não correlacionadas seja verdadeira.

Segundo (DANTAS, 2003, p.34), o primeiro passo para a utilização de modelos espaciais é o diagnóstico da presença de efeitos espaciais. A dependência espacial tem sido diagnosticada na literatura de duas formas distintas: pela análise gráfica do variograma ou utilizando testes estatísticos específicos, que utilizam matrizes de vizinhança.

De acordo com (ANSELIN, 1999) para diagnosticar a presença de efeitos de dependência espacial, é necessário definir previamente uma matriz de pesos espaciais (Matriz de vizinhança), conhecido como ω . Essa é uma matriz sintética em que cada elemento é igual a um se i e j forem vizinhos e igual a zero caso i e j não sejam vizinhos. Por convenção, os elementos diagonais da matriz ω são iguais a zero, ou seja, $\omega_{ii} = 0$.

Outras matrizes, consideram a importância dos vizinhos por meio de uma ponderação correspondente ao inverso da distância ou ao inverso do quadrado da distância entre eles. Em geral, a matriz de ω , é padronizado por linha, assumindo a nomenclatura ω^s . Neste caso, cada elemento de ω^s , representado por ω^s_{ij} é obtido dividindo-se ω_{ij} pela soma dos elementos da linha i a que pertence, ou seja $\omega^s_{ij} = \frac{\omega_{ij}}{\sum_j \omega_{ij}}$. Nesta matriz, os elementos das linhas somam 1. (DANTAS, 2003, p.34).

Os principais testes estatísticos utilizados para captar a autocorrelação espacial, são os testes de Moran I, Multiplicador de Lagrange (LM) Robusto do Erro e Multiplicador de Lagrange (LM) Robusto da Defasagem (DANTAS, 2003, p.35). O teste de Moran I, não nos fornece o tipo de efeito ocorrido no modelo (Erro ou defasagem), por isso ele será descartado no presente trabalho, sendo utilizados os testes mais específicos:

LM Robusto (Erro) para detectar efeitos de autocorrelação espacial no termo erro e LM Robusto (Defasagem) para verificar efeitos de defasagem espacial na variável dependente.

O teste LM Robusto (Erro) é realizado a partir da estatística representada pela Equação 2 que tem distribuição qui-quadrado com um grau de liberdade, sob a hipótese nula de não existência de autocorrelação espacial no termo do erro (ANSELIN, 1999).

$$LM(erro) = \frac{\left[\frac{e'W_e}{s^2}\right]^2}{\frac{n}{tr(W^2+W'W)}} \approx X^2 \quad Eq. [2]$$

e: Vetor de resíduos de mínimos quadrados;

W: Matriz de pesos espaciais;

$s^2=e'e/n$: Estimativa da máxima verossimilhança;

n: Número de dados da amostra;

tr: Operador denominado traço da matriz.

O teste LM Robusto (defasagem) é um teste assintótico, realizado a partir da estatística representada pela Equação 3, que tem distribuição qui-quadrado com um grau de liberdade, sob a hipótese nula de não existência de defasagem espacial na variável dependente (DANTAS, 2003, p.36)

$$LM(defasagem) = \frac{\left[\frac{e'Wy}{s^2}\right]^2}{\left[(WXB)' \frac{MWXb}{s^2} + tr(W'W+W^2)\right]} \approx X^2 \quad Eq. [3]$$

e: Vetor de resíduos de mínimos quadrados;

W: Matriz de pesos espaciais;

y: Vetor de observação da variável dependente;

$s^2=e'e/n$: Estimativa da máxima verossimilhança da variância do modelo;

X: Matriz de variáveis independentes;

b: Parâmetros estimados via mínimos quadrados ordinários;

tr: Operador denominado traço da matriz.

Segundo (TRIVELONI, 2005, p.41), a dependência espacial pode ser incorporada nos modelos clássicos de regressão de duas formas: em modelos de defasagem espacial, como um regressor adicional na forma de uma variável dependente espacialmente defasada (WY), ou em modelos de erro espacial como uma estrutura espacialmente defasada no erro da regressão ($W\varepsilon$).

No modelo de defasagem espacial a autocorrelação espacial ignorada é atribuída à variável dependente, conforme expressa a Equação 4.

$$Y = \rho WY + X\beta + \varepsilon \quad \text{Eq. [4]}$$

Y : variável dependente;

X : variáveis independentes;

β : coeficientes de regressão;

ε : erros aleatórios com média zero e variância σ^2 constante;

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

ρ : coeficiente espacial autoregressivo.

O modelo do erro espacial pode ser expressado formalmente conforme a Equação 5:

$$Y = X\beta + \varepsilon \quad \varepsilon = \lambda W\varepsilon + \xi \quad \text{Eq. [5]}$$

$W\varepsilon$ = erros com efeito espacial;

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

λ = coeficiente autoregressivo.

No método comparativo de dados de mercado por Regressão Linear Múltipla A NBR 14653-2: 2011, estabelece critérios estatísticos a serem atendidos, para aceitação do modelo de regressão.

O coeficiente de correlação indica a força da relação de causa e efeito entre a variação da variável dependente (Y) e a variável independente (X).

O coeficiente de determinação, define o percentual da variação total dos valores dos dados de amostra em torno da sua média aritmética originado nas diferenças analisadas pela equação da regressão (BIASI, 2017, p.4). O coeficiente de determinação é o quadrado do coeficiente de correlação e ele nos fornece o quanto a

variação da variável dependente (Y) está sendo explicada pelas variáveis independentes (X).

Para verificar a significância global do modelo, faz-se necessário a estatística Análise de Variância (ANOVA). Esta estatística é calculada conforme tabela 01, apresentada a seguir, e permite interpretar a aceitação do modelo de regressão gerado. Compara-se o F de Snedecor calculado (F_{calc}) com o F tabelado (F_{tab}) para o nível de significância desejado. Caso $F_{calc} > F_{tab}$, rejeita-se a hipótese de não existência de regressão e aceita-se o modelo em determinado grau de significância.

Tabela 01: Análise de variância

	<i>Graus de Liberdade</i>	<i>de</i>	Soma dos Quadrados	Quadrado Médio	F de Snedecor
Regressão	k		$\sum (y_i - \bar{y})^2$	$\sum (y_i - \bar{y})^2 / k$	$[\sum (y_i - \bar{y})^2 / k] / [\sum (\hat{y}_i - \bar{y})^2 / (n - k - 1)]$
Resíduo	n-k-1		$\sum (\hat{y}_i - \bar{y})^2$	$\sum (\hat{y}_i - \bar{y})^2 / (n - k - 1)$	
Total	n-1		$\sum (y_i - \hat{y}_i)^2$		

Fonte: Fermo (2006, p. 32)

k: Número de variáveis independentes testadas no modelo;

n: Número de dados utilizados;

y_i : Valor observado na pesquisa;

\hat{y}_i : Valor obtido na equação estimada;

\bar{y} : Média aritmética do valor observado.

Cabe salientar que o nível de significância máximo admitido para a rejeição da hipótese nula no modelo para o grau de fundamentação III é 1%, para grau II é 2% e para o grau I é 5% de acordo com a NBR 14653-2: 2011.

Na distribuição *t de student* testa-se o efeito de cada variável independente sobre a variável dependente, verificando a existência de significância estatística.

Quanto maior o valor em módulo da *estatística t (t obs)*, maior a significância estatística.

Conforme a NBR 14653-2: 2011 o nível de significância máximo para cada regressor para o grau de fundamentação III: 10%; II: 20% e grau I: 30%.

Para a aceitação de uma equação, é necessária a realização de uma análise da normalidade dos resíduos bem como o diagnóstico da existência ou não de heterocedasticidade. Para tanto, utilizam-se testes como o de Jarque-Bera e de Breusch-Pagan respectivamente (FERMO, 2010, p.27)

Segundo (TRIVELLONI, 2005, p.96), a *estatística z* fornecida pelos modelos espaciais corresponde ao equivalente da estatística *t* de *student* para o modelo de RLM. Ou seja, as probabilidades indicam o grau de significância de cada variável.

Quando os dados são tratados por regressão espacial, a NBR 14653-2:2011, indica, para realização da comparação entre os diferentes modelos gerados, a utilização dos critérios de Akaike e Schwarz. Estes critérios de informações e termos gerais, assumem a forma segundo a Equação 6.

$$IC = -2L + f(k, n) \quad \text{Eq. [6]}$$

IC: Critério de informação;

L: Máxima Log Verossimilhança;

$f(k, n)$: Função do número de variáveis independentes (k) e (n) é o número de observações.

Para Akaike $f(k, n) = 2k$; Para Schwarz $f(k, n) = k \cdot \ln(n)$.

Os critérios de Akaike e Schwarz representam a qualidade de ajuste do modelo e durante o processo de escolha entre os modelos, o melhor será aquele com menores valores encontrados.

2. MATERIAIS E MÉTODOS

Criciúma é um município brasileiro, localizado no extremo sul do estado de Santa Catarina com uma Área de 235,6 km². Segundo as estatísticas de 2014 do IBGE, conta com 204.667 habitantes.

A cidade é polo industrial em diversos setores, entre eles confecção, embalagens, cerâmico, plástico, extração do carvão mineral, construção civil e material gráfico.

Conhecida por ser a Capital Brasileira do Carvão, o seu subsolo abriga uma das maiores reservas minerais do país.

Colonizada por italianos, a cidade também recebeu poloneses, alemães, portugueses e árabes em diversas fases do seu desenvolvimento.

Conhecido o contexto geográfico e econômico da cidade, realizou-se a coleta de dados de terrenos da cidade de Criciúma – SC, por meio de imobiliárias locais, base de dados cedidas por Empresas do Ramo de Avaliações Imobiliária e Caixa Econômica Federal.

Ao todo, obteve-se uma amostra com 460 dados de ofertas de terrenos. Para facilitar a organização dos dados, foi elaborada uma planilha de Excel, contendo informações referentes ao endereço de cada imóvel (Rua e Bairro), características físicas que foram consideradas relevantes para a formação do valor do terreno, sendo: situação na quadra, topografia, área total, testada, número máximo de pavimentos, renda, distância a via principal mais próxima, Avenida Centenário, valor total, valor unitário e as coordenadas planas, Universal Transversa de Mercator (UTM).

Optou-se por trabalhar com as coordenadas em UTM, pois os softwares Quantum Gis Versão 2.18.11 (QGIS) e Sistema de Avaliação de Bens Versão 1.0 beta (SAB) utilizados no presente trabalho, possuem uma interface mais facilitada para se trabalhar em UTM.

Abaixo, segue um breve descrição das variáveis independentes pesquisadas:

Situação na quadra: Variável dicotômica. Adota-se o valor de 1 quando o terreno está localizado na esquina e 0 caso contrário.

Topografia: Variável dicotômica. Adota-se o valor de 1 quando o terreno é plano e 0 quando o é aclive ou declive.

Área total: Variável quantitativa. Trate-se da área total do terreno, mensurada em m².

Testada: Variável quantitativa. Trate-se da medida em metros da frente do terreno.

Número máximo de pavimentos: Variável quantitativa. Informação do plano diretor do município de Criciúma, que fornece o número máximo de pavimentos que pode ser construído sobre o terreno.

Renda: Variável proxy. Procura explicar o padrão da circunvizinhança onde está localizado o terreno a partir da renda média do chefe de família dessa mesma circunvizinhança, sendo esta pesquisa, realizada pelo IBGE no censo de 2010.

Distância a via principal mais próxima: Variável quantitativa. É a menor distância do terreno, até uma das vias principais da cidade de Criciúma/SC (conforme Figura 02).

Avenida Centenário: Variável dicotômica. Adota-se o valor de 1 quando o terreno está localizado na Avenida Centenário e 0 caso contrário.

Todos os dados provenientes de bases de dados da Caixa Econômica Federal ou de Empresas de Avaliação Imobiliária, possuíam informações completas referentes a endereço, bairro, topografia, situação na quadra, área do terreno, testada do terreno, valor do terreno e coordenadas em UTM.

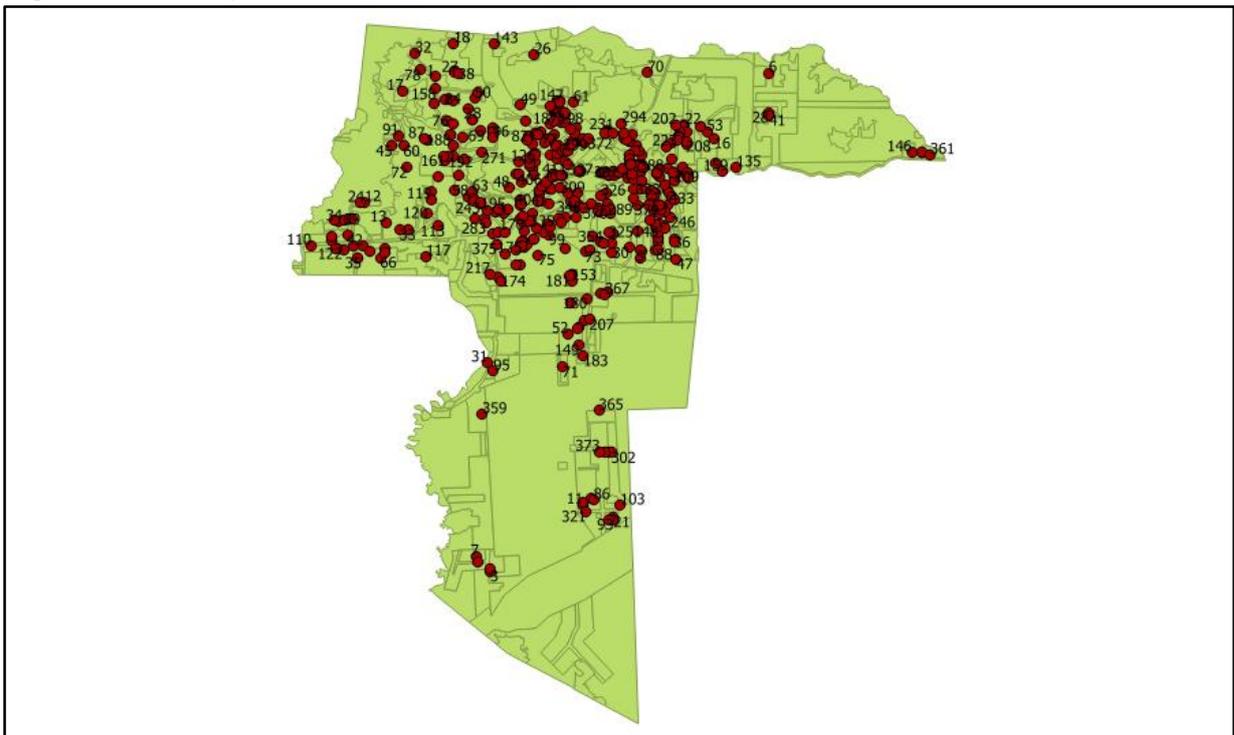
As informações do cadastro imobiliário de Criciúma (número máx de pavimentos) foram obtidas utilizando-se camadas de informações espaciais (*shapefile*) da cidade de Criciúma SC.

Arquivo *shapefile* é um formato popular de arquivo contendo dados espaciais em forma de vetor usado por sistemas de informação geográfica também conhecidos como SIG. Através do *shapefile*, é possível criar mapas de superfícies onde possam ser subtraídas informações.

O software utilizado para trabalhar com *shapefiles* foi o QGIS que é disponibilizado gratuitamente para *download* na internet.

Tendo em vista as informações anteriores, todos os dados da amostra foram plotados no software QGIS por meio de suas respectivas coordenadas UTM, assim como o *shapefile* do cadastro imobiliário, como mostrado na Figura 01. Em verde claro, está o *shapefile* contendo as informações do cadastro imobiliário, e em vermelho estão os dados da amostra rotulados pela sua respectiva numeração de controle.

Figura 01 – *Shapefile* Cadastro Imobiliário de Criciúma



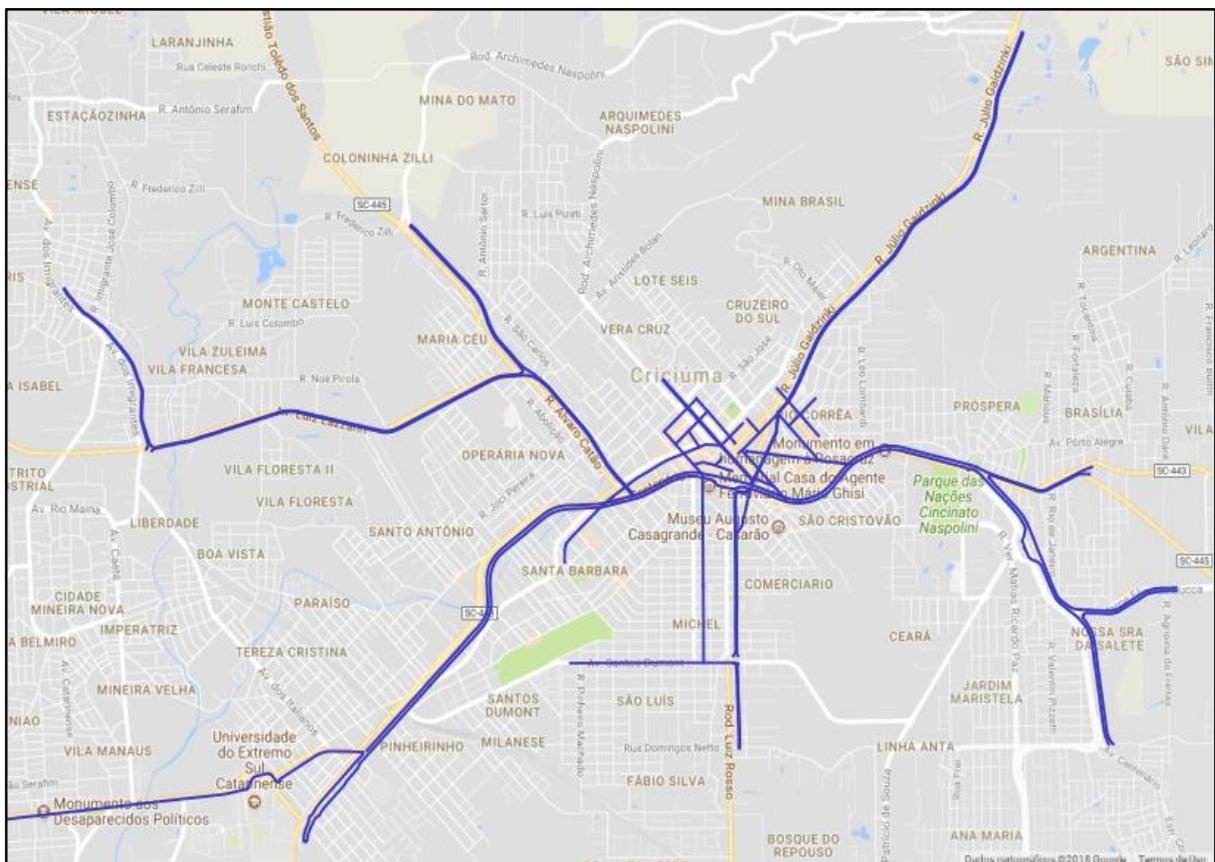
Fonte: O Autor

As informações foram repassadas do *shapefile* para todos os dados através da ferramenta “Unir atributos pela posição”, essa, permite a troca de informações entre os pontos (dados) e a camada (*Shapefile* do Cadastro Imobiliário de Criciúma) através da sobreposição.

A variável Renda, foi obtida empregando-se o mesmo procedimento utilizado para a variável número máximo de pavimentos. Porém não foi utilizado o *shapefile* do cadastro imobiliário, e sim o *shapefile* do mapa de renda do chefe de família, disponibilizado pelo IBGE (censo 2010).

A variável Distância as vias principais, também foi obtida pelo QGIS. Criou-se um *shapefile* das vias principais da cidade de Criciúma/SC. As vias principais estão em azul, conforme Figura 02.

Figura 02 – *Shapefile* das vias principais



Fonte: O Autor

Posteriormente, utilizou-se a ferramenta “Distância ao centro mais próximo” para calcular a distância de cada dado até a via principal mais próxima.

Com o mesmo software, plotou-se todos os imóveis espacialmente, possibilitando ter uma visão geral da representatividade da amostra de dados sobre a cidade de Criciúma conforme Figura 03.

Para melhor entendimento, o processo de modelagem foi separado em 03 principais etapas.

Etapa 01: Encontra-se um modelo de regressão linear múltipla que apresente maior adequação estatística conforme a NBR 14653-2: 2011.

Etapa 02: Cria-se a matriz de vizinhança. Posteriormente realiza-se a constatação de autocorrelação espacial através dos testes de LM Robusto Erro e LM Robusto Defasagem e confirma-se os mesmos apresentam significância.

Confirmada a existência de autocorrelação espacial, prossegue-se para a Etapa 03. Em caso de não existência de autocorrelação espacial, adota-se o modelo de regressão linear múltipla.

Etapa 03: Estando a matriz de vizinhança completa e reconhecida a autocorrelação espacial, inicia-se o processo de modelagem espacial através da Regressão Espacial. Após a Regressão Espacial ser realizada, compara-se os modelos da defasagem espacial, do erro espacial e da regressão linear múltipla através dos critérios de Akaike e Schwarz. O que apresentar menor valor nos critérios de Akaike e Schwarz deverá ser o adotado.

O Software utilizado para Modelagem, foi o SAB (Sistema de Avaliação de Bens - Versão 1.0 beta), desenvolvido e comercializado pela Dantas Engenharia.

Antes mesmo de ser iniciado o processo de modelagem, foram separados 06 dados para compor a amostra de controle e fazer a validação dos modelos encontrados.

3. RESULTADOS E DISCUSSÕES

Diversas foram as simulações, até obter-se o modelo de regressão linear múltipla que apresentou melhor resultado estatístico conforme critérios da NBR 14653-2: 2011.

Dos 460 dados iniciais da amostra, 431 foram os utilizados no modelo final.

A variável dependente adotada, foi o valor total. Entre as 08 variáveis independentes testadas, 05 apresentaram-se conjuntamente significativas para a formação do valor total, são elas: Área Total, Distância a via principal mais próxima, Número de Pavimentos Máximo, Renda e Avenida Centenário. A equação obtida está a seguir representada:

Equação 07 – Equação do modelo de Regressão Linear Múltipla.

$$\text{Valor total} = e^{(+6.615111456 + 0.9531798411 * \ln(\text{Área total}) - 4.671785602E-005 * \text{Dist Vias} + 0.3572263312 * \ln(\text{Nº Pav}) - 1344.114184 / \text{Renda} + 0.8261478387 * \text{Avenida Centenário Dicotômica})}$$

Fonte: O Autor

Na Tabela 02, apresenta-se os resultados estatísticos do modelo de regressão linear múltipla.

Tabela 02: Resultados estatísticos do RLM

<i>Estatística de regressão</i>	
Coeficiente de Correlação	0,9364
Coeficiente de Determinação	0,8768

Fonte: O Autor

O coeficiente de correlação obtido foi de 0,9364, demonstrando-nos uma correlação fortíssima da variável dependente com as variáveis independentes.

O coeficiente de determinação obtido foi de 0,8768, demonstrando-nos que aproximadamente 88% da variação do valor total dos terrenos na cidade de Criciúma, estão sendo explicados pelas 05 variáveis independentes utilizadas no modelo.

A análise de significância global do modelo, foi realizada por meio da análise de variância (ANOVA), apresenta-se os resultados conforme Tabela 03:

Tabela 03: Análise de Variância (ANOVA)

	<i>Graus de Liberdade</i>	Soma dos Quadrados	Quadrado Médio	F de Snedecor
Regressão	5	335,457	67,091	604,861
Resíduo	425	47,141	0,111	
Total	430	382,598		

Fonte: O Autor.

O *F de Snedecor* calculado é de 604,861; O *F tabelado* para a significância de 1% é de 3,32. Como $F_{calc} > F_{tab}$, rejeitou-se a hipótese de não existência de regressão, e aceitou-se o modelo de regressão para o nível de significância de 1%, enquadrando-se nesse quesito no grau III de fundamentação da NBR 14653-2: 2011.

Os resultados da significância individual das variáveis por meio do teste *t de student*, estão apresentados conforme Tabela 04.

Tabela 04 – Teste *t* de Student

<i>Variáveis</i>	<i>Transf.</i>	<i>t Obs.</i>	<i>Sig.(%)</i>
Área total	ln(x)	28,88	0.01
Distância a via principal	x	-5,55	0.01
Nº Pav	ln(x)	10,41	0.01
Renda	1/x	-16,38	0.01
Avenida Centenário	x	12,32	0.01
Valor total	ln(y)	30,46	0.01

Fonte: O Autor.

Observou-se que todos os valores da significância estão abaixo de 10% e enquadram-se no grau de fundamentação III segundo a NBR 14653-2: 2011.

Verificou-se também que entre as variáveis independentes, duas apresentaram-se mais influentes; Primeiramente, Área total, com o valor de *t obs* de 28,88 e posteriormente, Renda, com o valor de *t obs* de -16,38.

Pelos resultados que se encontram na Tabela 05, verifica-se que as hipóteses básicas de normalidade e homocedasticidade, foram aceitas a um nível de significância de 1%, quando realizados os testes de Jarque-Bera e Breusch-Pagan, respectivamente.

Tabela 05: Resultados estatísticos do RLM

<i>Teste</i>	<i>Valor</i>	<i>Probabilidade</i>
Jarque-Bera	2,59768	0,27285
Breusch-Pagan	25,32060	0,99988

Fonte: O Autor

Com o modelo de RLM atendendo aos critérios estabelecidos anteriormente, iniciou-se o processo de verificação da autocorrelação espacial utilizando os testes estatísticos de LM Robusto (Erro) e LM Robusto (Defasagem).

Para realização desses testes, necessitou-se previamente uma matriz de distância (Matriz de Vizinhança). Após diversas tentativas, observou-se que a matriz do inverso da distância ao quadrado apresentou-se a mais adequada para modelagem.

Na Tabela 06, apresenta-se os resultados do LM Robusto (Erro) e LM Robusto (Defasagem).

Tabela 06 – Teste LM Robusto (Erro e Defasagem)

<i>Teste</i>	<i>Valor</i>	<i>Prob</i>
LM Robusto (Defasagem)	3729,47313	0,00001
LM Robusto (Erro)	187,0282	0,00001

Fonte: O Autor.

Verificou-se que tanto o Teste LM Robusto (Defasagem) quanto o Teste LM Robusto (Erro) apresentaram-se significativos ao nível de 1%. Portanto, foi realizada a modelagem da defasagem espacial e a modelagem do erro espacial, para posterior comparação através dos critérios de Akaike e Schwarz

A equação obtida do Modelo da Defasagem Espacial está abaixo representada:

Equação 08 – Equação do Modelo de Defasagem Espacial.

$$\text{Valor total} = e^{(+2,2121+0,8639 * \ln (\text{Área total})-0,00001 * \text{Dist Vias}+0,2398 * \ln (\text{N}^\circ \text{ Pav})-855,2020 / \text{Renda}+0,5206 * \text{Avenida Centenário Dicotômica} +0,390963*W^2Y)}$$

Fonte: O Autor

Na tabela 07 apresenta-se os resultados estatísticos do Modelo da Defasagem Espacial.

Tabela 07: Resultados estatísticos do Modelo da Defasagem Espacial

<i>Estatística de regressão</i>	
Coeficiente de Correlação	0,9309
Coeficiente de Determinação	0,8666

Fonte: O Autor

O Coeficiente de Correlação de 0,9309, indicou uma correlação fortíssima da variável dependente com as variáveis independentes utilizadas no modelo.

O Coeficiente de Determinação de 0,8666, indicou que 86,66% da variação da variável dependente (Valor total) foi explicada pelas variáveis independentes utilizadas.

Os resultados da significância local por meio da *estatística z*, estão apresentados conforme Tabela 08.

Tabela 08 – Estatística z

<i>Variáveis</i>	<i>Transf.</i>	<i>Valor de z</i>	<i>Sig.(%)</i>
Área total	ln(x)	13,8037	0.01
Distância a via principal	x	-1,4745	0.1411
Nº Pav	ln(x)	3,2344	0.01
Renda	1/x	-5,5767	0.01
Avenida Centenário	x	3,90	0.01
ρ (coeficiente autoregressivo)		168,8979	0.01

Fonte: O Autor.

Observou-se que todas as variáveis independentes apresentaram-se significativas ao nível de 14,11% enquadrando-se no grau II de fundamentação conforme NBR 14.653-2:2011.

Destaca-se, que o ρ (Coeficiente Autoregressivo) é positivo e com um valor da estatística z de 168,8979 apresentou-se significativa a menos de 1%, confirmando novamente a existência de autocorrelação espacial na variável dependente do modelo.

A equação obtida do Modelo do Erro Espacial está abaixo representada:

Equação 09 – Equação do Modelo do Erro Espacial.

$$\text{Valor total} = e^{(+6,4676+0,9494*\ln(\text{Área total})-0,0001*\text{Dist Vias}+0,2972*\ln(\text{N}^\circ\text{Pav})-1045,6611/\text{Renda}+0,6838*\text{Avenida Centenário Dicotômica}+0,721906*W^2E)}$$

Fonte: O Autor

Na Tabela 09 apresenta-se os resultados estatísticos do Modelo do Erro Espacial.

Tabela 09: Resultados estatísticos do Modelo do Erro Espacial

<i>Estatística de regressão</i>	
Coeficiente de Correlação	0,9230
Coeficiente de Determinação	0,8520

Fonte: O Autor

O Coeficiente de Correlação de 0,9230, indicou uma correlação fortíssima da variável dependente com as variáveis independentes utilizadas no modelo.

O Coeficiente de Determinação de 0,8520, indicou que 85,20% da variação da variável dependente (Valor total) foi explicada pelas variáveis independentes utilizadas.

Os resultados da significância local por meio da *estatística z*, estão apresentados conforme Tabela 10.

Tabela 10 – Estatística z

<i>Variáveis</i>	<i>Transf.</i>	<i>Valor de z</i>	<i>Sig.(%)</i>
Área total	ln(x)	23,7662	0.01
Distância a via principal	x	-3,1862	0.01
Nº Pav	ln(x)	6,1173	0.01
Renda	1/x	-7,0193	0.01
Avenida Centenário	x	6,3220	0.01
λ (Coeficiente Autoregressivo)		23,9529	0.01

Fonte: O Autor.

Observou-se que todas as variáveis independentes apresentaram-se significativas ao nível de 1%, enquadrando-se no grau III de fundamentação, de acordo com a NBR 14.653-2:2011.

Destaca-se, que o λ (Coeficiente Autoregressivo) é positivo e com um valor da estatística z de 23,9529 apresentou-se significativa a menos de 1%, confirmando novamente a existência de autocorrelação espacial nos resíduos do modelo.

A Comparação entre os Modelos de Regressão Espacial e Regressão Linear Múltipla, foi realizada pelos critérios de Akaike e Schwarz, conforme Tabela 11.

Tabela 11: Comparação por critérios de Akaike e Schwarz

<i>Modelos</i>	<i>Akaike</i>	<i>Schwarz</i>
Regressão Linear Múltipla	281,3810	305,7776
Defasagem Espacial	459,1564	479,4870
Erro Espacial	180,5057	200,8362

Fonte: O Autor

Observou-se que quando comparado os três modelos, o Modelo do Erro Espacial apresentou os menores valores de Akaike e Schwarz. Como visto anteriormente, quanto menor os critérios de Akaike e Schwarz, melhor o modelo. Sendo assim, o Modelo do Erro Espacial, foi o que melhor explicou o comportamento do mercado de

terrenos e portanto foi o mais adequado para ser utilizado para a avaliação de terrenos na cidade de Criciúma/SC.

Também pode ser observado, que o Modelo de Regressão Linear Múltipla, apresentou menores valores de Akaike e Schwarz, se comparado ao Modelo de Defasagem Espacial. Conclui-se disto, que mesmo com significância na hipótese de autocorrelação espacial na variável dependente, não vale a pena utilizar o Modelo de Defasagem Espacial, visto que o Modelo de Regressão Linear Múltipla explica melhor o comportamento do mercado de terrenos de Criciúma/SC.

Com o intuito de validação dos 03 modelos em questão, utilizou-se uma amostra de controle com 06 dados para realizar a avaliação dos mesmos, conforme Tabela 12.

Tabela 12 – Amostra de controle

<i>Dado</i>	<i>Área Total</i>	<i>Dist vias</i>	<i>Nº Pav</i>	<i>Renda</i>	<i>Av Cent</i>	<i>X (UTM)</i>	<i>Y (UTM)</i>	<i>Valor Observado</i>
1	760	136,66	10	4.001,08	0	658.941,43	6.825.475,50	630.000,00
2	471	89,77	10	3.024,87	0	659.146,64	6.825.273,32	340.000,00
3	463,47	70,33	16	4.086,92	1	660.065,68	6.826.340,63	1.390.410,00
4	447,95	585,22	2	1.541,31	0	657.069,11	6.824.314,33	135.000,00
5	538,67	881,13	4	1.003,62	0	653.123,81	6.828.264,32	112.000,00
6	378,13	3.957,94	2	1.551,40	0	658.761,29	6.820.498,74	85.000,00

Fonte: O Autor

Segue os valores avaliados para cada modelo, juntamente com a variação em % do valor estimado em relação ao valor observado, conforme Tabela 13.

Tabela 13 – Validação dos Modelos.

<i>Valor Observado</i>	<i>RLM</i>	<i>Modelo Defasagem Espacial</i>	<i>Modelo Erro Espacial</i>
630.000,00	672.043,00 (7%)	571.613,00 (-9%)	642.402,00 (2%)
340.000,00	383.001,00 (13%)	353.657,00 (4%)	334.117,00 (-2%)
1.390.410,00	1.144.934,00 (-18%)	1.224.101,00 (-12%)	1.139.806,00 (-18%)
135.000,00	130.899,00 (-3%)	127.231,00 (-6%)	126.198,00 (-7%)
112.000,00	123.568,00 (10%)	114.222,00 (2%)	113.186,00 (1%)
85.000,00	95.681,00 (13%)	91.030,00 (7%)	86.799,00 (2%)

Fonte: O Autor

Verificou-se que em algumas ocasiões alguns modelos apresentavam estimações mais próximas ao valor observado, em relação aos outros modelos, mas de maneira geral, o modelo do Erro Espacial, apresentou os melhores resultados, como pode ser observado através da variação do valor estimado em relação ao valor observado. Confirmando novamente, que o modelo com menor critério de Akaike e Schwarz, é o mais adequado a ser utilizado para estudos do comportamento do mercado de terrenos da cidade de Criciúma/SC e também para realizar estimações de valores.

4. CONCLUSÃO

Como visto anteriormente, uma das dificuldades dos modelos de regressão linear múltipla está em considerar a variável localização e dados localizados espacialmente, como é o caso dos terrenos, em geral, apresentam autocorrelação espacial, e como consequência, gera-se problemas nas estimações do modelo.

Para verificação da autocorrelação espacial, criou-se uma Matriz de distância (Matriz de vizinhança) com o inverso do quadrado da distância dos dados. Posteriormente, utilizou-se os testes LM Robusto (Defasagem) e LM Robusto (Erro) e ambos apresentaram significância a nível de 1%.

Sendo ambos os testes significativos, criou-se tanto o modelo da defasagem espacial, quanto o modelo do erro espacial.

Os três modelos (Regressão linear múltipla, defasagem espacial e erro espacial) adequaram-se estatisticamente, todos eles possuindo um coeficiente de determinação acima de 0,80.

Os modelos de regressão linear múltipla e modelo do erro espacial apresentaram variáveis independentes significativas a um nível de 1%. Já o modelo de defasagem espacial, comportou-se pior nesse quesito, apresentando variáveis independentes significativas a um nível de 14,11%.

O fato do modelo de regressão linear múltipla apresentar autocorrelação espacial, não justifica isoladamente a utilização de modelos espaciais, visto que, os testes que indicam qual o melhor modelo a ser utilizado, são os critérios de Akaike e Schwarz.

Visto isso, realizou-se os testes de Akaike e Schwarz, e conforme tabela 12, concluiu-se que o modelo do erro espacial é o modelo que melhor explica o mercado de terrenos da cidade de Criciúma/SC, sendo o mais adequado a ser utilizado para

estimação de valores, seguido pelo modelo de regressão linear múltipla e posteriormente o modelo de defasagem espacial.

Utilizou-se uma amostra de controle com 06 dados escolhidos aleatoriamente para validação dos modelos. Os 03 modelos utilizados apresentaram resultados satisfatórios, conforme tabela 13.

Conclui-se que a modelagem espacial apresentou-se eficiente e levando em consideração a localização dos dados da amostra, possibilitou um melhor ajuste ao modelo de regressão linear múltipla, validando assim, a sua utilização para a avaliação de terrenos da cidade de Criciúma/SC.

A partir dos resultados e conclusões alcançadas nesta pesquisa, recomenda-se para futuros trabalhos a realização de modelos espaciais para outras tipologias de imóveis, como apartamentos, casas e glebas.

Outra recomendação seria realizar a comparação dos resultados obtidos nesse trabalho com resultados obtidos por métodos geoestatísticos.

ANSELIN, L. **Exploring Spatial Data with GeoDaTM: A Workbook**. 2005.

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

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14653/ 2011. Avaliações de Bens Parte 2: Imóveis Urbanos**. Rio de Janeiro, 2004, 37p.

BIASI, Luís Henrique. **Atualização do Modelo de Regressão Linear Múltipla Para Avaliação de Imóveis tipo Apartamento na Cidade de Criciúma - SC** – Engenharia Civil – Universidade do extremo Sul Catarinense – Unesc, Novembro 2017.

CASTANHETI, Carolina Brunel. **Avaliação dos terrenos em condomínios horizontais do município de Criciúma-SC** – Engenharia de Agrimensura – Universidade do extremo Sul Catarinense – Unesc, Março 2018.

COELHO, Jocilon. **Modelo de Regressão Linear Múltipla para Avaliação de Aluguéis de Salas Comerciais na Cidade de Araranguá-SC** – Engenharia Civil – Universidade do extremo Sul Catarinense – Unesc, Novembro 2011.

CRESSIE, N. A. C. **Statistics for Spatial Data**. Revised Edition, John Wiley & Sons, Inc., 1993

DANTAS, R. A. **Modelos Espaciais aplicados ao mercado habitacional – um estudo de caso para a cidade de Recife**. Recife, 2003 (Tese – Universidade Federal de Pernambuco – UFPE).

DELFINO, Vanessa. **Modelo de Regressão Linear Múltipla para Avaliação de Apartamentos na Cidade de Torres-RS** – Engenharia Civil – Universidade do extremo Sul Catarinense – Unesc, Novembro 2013.

FERMO, Graziela Olivo. **Influência da Ferramenta de Georreferenciamento na Avaliação de Apartamentos usando Regressão Espacial**. Florianópolis, 2010 (Dissertação de Mestrado – Universidade Federal de Santa Catarina – UFSC)

FERMO, Graziela Olivo. **Modelo de Regressão Linear Múltipla para Avaliação de Apartamentos na Cidade de Criciúma, SC** – Engenharia Civil – Universidade do Extremo Sul Catarinense – UNESC, junho 2006.

GUISSO, Dhionata. **Aplicação do Geoprocessamento para Análise do Comportamento de Algumas Variáveis Formadoras de Valores Utilizados na Avaliação de Imóveis Urbanos – Estudo de Caso de Alguns Dados da Cidade de Criciúma**. Engenharia de Agrimensura – Universidade do Extremo Sul Catarinense – UNESC, 2014.

GONZÁLEZ, Marco Aurélio Stumpf. **A Engenharia de Avaliações na Visão Inferencial**. 2ª ed. São Leopoldo: UNISINOS, 2000. 141 p.

HORNBURG, André. **Avaliação em Massa de Imóveis usando Regressão Espacial e Krigagem Bayesiana (um estudo de caso em Balneário Camboriú/SC)**. Revista eletrônica de Engenharia Civil, <https://revistas.ufg.emnuvens.com.br/reec/index>, volume 13, nº 1, Jan/2017.

SILVA, Everton da. **Cadastro Técnico Multifinalitário: Base Fundamental para Avaliações em Massa de Imóveis**. Florianópolis, 2006. (Tese – Universidade Federal de Santa Catarina - UFSC.)

THOFEHRN, Ragnar. **Avaliação de Terrenos Urbanos por Formulas Matemáticas**. 1ª ed. São Paulo: PINI, 2008. 235 p.

TRIVELLONI, Carlos Alberto Peruzzo. **Método para Determinação do Valor da Localização com Uso de Técnicas Inferências e Geoestatísticas na Avaliação em Massa de Imóveis**. Florianópolis, 2005 (Tese de Doutorado – Universidade Federal de Santa Catarina – UFSC)

ZANCAN, Evelise Chemale. **Avaliações de Imóveis em Massa para Efeitos de Tributos Municipais**. Florianópolis: Rocha, 1996.