

Revista Brasileira de Cartografia (2016), Nº 68/4, Edição Especial Geoinformação e Análise Espacial: 733-743
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto
ISSN: 1808-0936

ESTUDO COMPARATIVO DO USO DE KRIGAGEM PARA ESTIMAÇÃO DA ESCOLHA DO MODO PARTICULAR MOTORIZADO

A Comparative Study of Kriging Application to Estimate the Private Motorized Travel Mode

**Viviani Antunes Gomes¹, Cira Souza Pitombo¹, Samille Santos Rocha¹
& Ana Rita Salgueiro²**

¹Universidade de São Paulo – USP

Escola de Engenharia de São Carlos - Departamento de Engenharia de Transportes
Av. Trab. São-Carlense, 400 - Parque Arnold Schimidt, CEP: 13.566-590, São Carlos - SP, Brasil
viviani_antunes@yahoo.com.br; cira@sc.usp.br; millyblase@gmail.com,

²Universidade Federal do Ceará - UFC

Departamento de Geologia
Campus do Pici - Bloco 912, CEP: 60.440-900, Fortaleza, CE, Brasil
geo.ritasalgueiro@gmail.com

Recebido em 28 de Junho, 2015/ Aceito em 4 de Fevereiro, 2016
Received on June 28, 2015/ Accepted on February 4, 2016

RESUMO

Esse trabalho tem como objetivo analisar e comparar os resultados de duas técnicas de Krigagem (Ordinária e Indicativa) que foram aplicadas para estimar a escolha do modo de transporte particular motorizado (automóvel ou motocicleta) em coordenadas de valores conhecidos e desconhecidos. Os dados utilizados foram da Pesquisa Origem/Destino e da Pesquisa de Opinião sobre avaliação do transporte público, realizadas em 2007/2008 em São Carlos (SP). As técnicas foram aplicadas em uma região da cidade com 110 pontos amostrados. Inicialmente, aplicou-se a técnica da Árvore de Decisão para estimar a probabilidade de escolha do modo particular motorizado em domicílios pesquisados, determinando assim a variável numérica a ser interpolada posteriormente através da Krigagem Ordinária. Para a aplicação da Krigagem Indicativa utilizou-se a variável “modo principal” em sua forma discreta, onde “1” significava o uso do modo particular motorizado e “0” significava o não uso desses modos. Os resultados obtidos pelas duas técnicas de interpolação espacial foram similares (mapas interpolados e validação cruzada). A Krigagem Indicativa (KI), no entanto, obteve maior número de acertos. Além disso, com a KI foi utilizada a variável objeto de estudo em sua forma original, evitando possível propagação de erros.

Palavras chaves: Geoestatística, Krigagem, Escolha Modal, Interpolação Espacial.

ABSTRACT

This paper aims to analyse and compare the results of two techniques of kriging (Ordinary Kriging and Indicator Kriging) that were applied to estimate the private motorized travel mode choice (car or motorcycle) in known and unknown geographic coordinates. The data used was from the Origin/Destination and Public Transportation Opinion Survey, carried out in 2007/2008 at São Carlos (SP, Brazil). The techniques were applied in the region with 110 sample points. Initially, Decision Tree was applied to estimate the probability of mode choice in surveyed households, thus determining the numeric variable to be used in Ordinary Kriging. For application of Indicator Kriging it was used the variable “main travel mode” in a discrete manner, where “1” represents the use of motorized particular travel mode

and “0” characterizes others travel modes. The results obtained by the two spatial interpolation techniques were similar (Kriging maps and cross-validation procedure). However, the Indicator Kriging (IK) obtained the highest number of hit rates. In addition, with the IK it is possible to use the variable in its original form, avoiding error propagation.

Keywords: Geostatistics, Kriging, Travel Mode choice, Spatial Interpolation.

1. INTRODUÇÃO

A escolha modal é um aspecto altamente relevante a ser pesquisado com mais profundidade no âmbito da demanda por transportes. Esses estudos podem ajudar a entender o comportamento da população e, conseqüentemente, a elaborar políticas públicas. As razões de escolha de um modo de transporte em detrimento dos outros são complexas. Ortúzar e Willumsen (2011) apontam alguns fatores que podem interferir nessas escolhas, tais como: (1) as características do indivíduo (posse do carro, renda, densidade residencial, etc.), (2) as características da viagem (motivo de viagem, período do dia, etc.), e (3) as características do modo de viagem (tempo de viagem, custos monetários, conforto, conveniência, etc.).

Alguns trabalhos não só relacionam a escolha modal às características dos usuários, da viagem e do modo de transporte, mas também a um conjunto de variáveis relacionadas ao espaço urbano. Loureiro *et al.* (2006) afirmam que estudos relacionados a viagens urbanas são influenciados por atributos eminentemente geográficos, tais como: adensamentos residenciais e de atividades socioeconômicas, proximidade entre zonas, cobertura espacial da rede de transporte, impedância de viagem na malha rodoviária, etc. Pitombo *et al.* (2015) e Gomes *et al.* (2016) apontam a necessidade do estudo da relação da escolha modal e localizações espaciais dos domicílios, considerando a forte influência do espaço urbano nas decisões de utilização de determinado modo de transporte. Rajamani *et al.* (2003) investigaram e identificaram uma série de resultados sobre o impacto da forma de ocupação urbana na escolha modal.

A geoestatística, conjunto de técnicas de análise espacial em superfícies, permite entender a aparente aleatoriedade dos dados, além da sua possível estrutura espacial. Possibilita ainda desenvolver estudos sobre fenômenos,

cujas variáveis têm distribuição de valores associada à sua posição no espaço e estimar o valor da variável em locais não amostrados, levando em conta o comportamento espacial do fenômeno (YAMAMOTO & LANDIM, 2013; WACKERNAGEL, 2010; ISAACS & SRIVASTAVA, 1989).

Acredita-se que a inserção do atributo espacial no estudo das viagens no ambiente urbano possibilita uma melhor representação do padrão de comportamento. Esse trabalho tem como objetivo analisar e comparar os resultados obtidos com a aplicação de duas técnicas de Krigagem (Ordinária e Indicativa) para estimar a escolha por automóvel em coordenadas de valores conhecidos e desconhecidos.

Este artigo está estruturado em três seções, além desta introdução. A seção 2 apresenta a fundamentação teórica relativa às técnicas de análise espacial de dados utilizadas. A seção 3 descreve o método e resultados obtidos. Finalmente, a seção 4 traz as principais conclusões e discussões.

2. FUNDAMENTAÇÃO TEÓRICA

A estatística espacial, mescla conceitos de análise espacial com cálculos estatísticos tradicionais e permite estudar a correlação espacial dos aspectos que envolvem as análises de demanda por transportes como escolha modal, por exemplo. Entre as técnicas de estatística espacial, destaca-se a geoestatística que, de acordo com Camargo *et al.* (2004), permite descrever a distribuição das variáveis de estudo, identificar observações atípicas, não só em relação ao tipo de distribuição, mas também em relação aos vizinhos, e buscar a existência de padrões na distribuição espacial.

A geoestatística está fundamentada na Teoria das Variáveis Regionalizadas, desenvolvida por Georges Matheron e Daniel G. Krige (JOURNAL & HUIJBREGTS, 1978). A teoria diz que alguma medida pode ser vista como uma realização de uma função irregular,

na qual se tem um aspecto aleatório, cuja irregularidade impossibilita prever as variações de um ponto a outro, e um aspecto estruturado, que reflete as características estruturais do fenômeno regionalizado. Esta teoria é a base da geoestatística. Camargo *et al.* (2004) afirmam ainda que a variável regionalizada é distribuída no espaço e acrescentam que seus valores podem ser expressos pela soma de três componentes: a) uma componente estrutural, associada a um valor médio constante ou a uma tendência constante; b) uma componente aleatória, espacialmente correlacionada; e c) um ruído aleatório ou erro residual.

Conceitos básicos importantes da geoestatística como variograma, validação cruzada e krigagem são sumariamente descritos nas subseções seguintes.

2.1. Variograma

O variograma é uma ferramenta que permite através de uma representação gráfica apresentar quantitativamente a variação de um fenômeno regionalizado no espaço. O variograma serve para encontrar os pesos ótimos a serem associados às observações que irão estimar um ponto, uma área ou um bloco (CARR, 1990; YAMAMOTO & LANDIM, 2013).

Avalia-se o fenômeno através da mensuração das variâncias entre pares de pontos $Z(x)$ separadas por uma distância h , ou seja, a função variograma corresponde à média das variâncias entre os pontos (Equação 1).

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n(h)} [Z(x) - Z(x+h)]^2 \quad (1)$$

sendo “n” o número total de observações da amostra em cada *lag* (h).

Para a obtenção do variograma experimental, Lindner (2015), apresenta as definições dos parâmetros utilizados e relaciona os passos:

- Distância de corte (d_{corte}): limite superior para cálculo do variograma;

- *Lag*: segmentos delimitados a fim de auxiliar o cálculo das variâncias entre pares de pontos, cujos parâmetros estão indicados abaixo:

- Tamanho do *lag* (h): distância média entre amostras vizinhas;
- Número de *lags* (n_{lags}): valor escolhido de maneira que $h \times n_{\text{lags}} < d_{\text{corte}}$ possibilitando dessa forma a representação completa das amostras;
- Tolerância de *lag* (Δh): valores dentro dos limites ($h + \Delta h$) e ($h - \Delta h$) são considerados como pertencentes a h .

- Direção angular (θ): direção adotada para elaboração do variograma que melhor representa a distribuição espacial dos dados. Deste parâmetro, definem-se:

- Tolerância angular ($\Delta\theta$): valores dentro dos limites ($\theta + \Delta\theta$) e ($\theta - \Delta\theta$) são considerados como pertencentes a θ .
- Direção ortogonal ($\theta + 90^\circ$): direção ortogonal à direção adotada para elaboração do variograma

1) O processo inicia-se em um primeiro ponto n_1 da amostra. Para toda observação distante $h \pm \Delta h$ e com ângulo entre $\theta \pm \Delta\theta$ deste ponto, calcula-se a diferença entre os valores dos pares (Figura 1a);

2) Refaz-se o primeiro passo para todos os pontos da amostra (Figura 1b);

3) Em seguida obtém-se a variância γ para o *lag* de tamanho h . Este valor estará representado por um dos pontos γ no variograma experimental, onde o eixo das ordenadas é representado pela variância γ e o eixo abscissas, pela distância h (Figura 1c);

4) Repete-se os passos indicados acima para os *lags* $2h, 3h, \dots, n_{\text{lags}}*h$, afim de se obter $\gamma(2h), \gamma(3h), \dots, \gamma(n_{\text{lags}}*h)$ (Figura 1d), desde que os valores de distância não excedam a distância de corte (d_{corte}) (Figura 1e).

O ajuste do variograma experimental é feito a partir de diferentes modelos teóricos. Embora existam vários modelos de variogramas teóricos, os modelos esférico, exponencial e gaussiano conseguem explicar a grande maioria dos fenômenos espaciais (YAMAMOTO & LANDIM, 2013). A Tabela 1 apresenta as equações dos três modelos e a Figura 2 traz a suas representações:

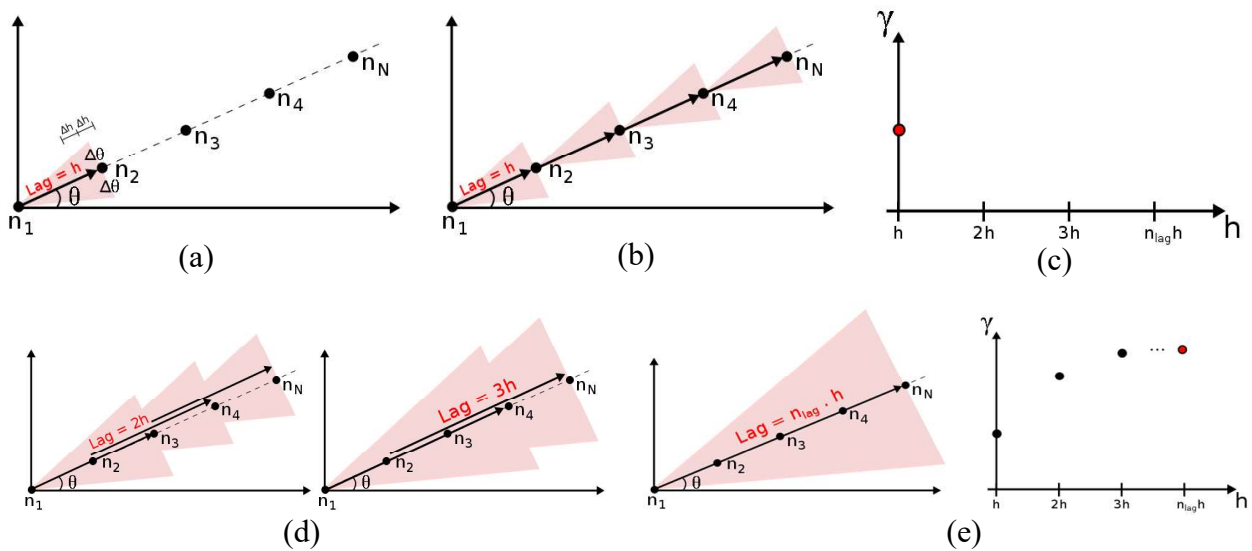


Fig. 1- Processo de obtenção do variograma experimental (LINDNER, 2015).

Tabela 1: Modelos teóricos de variogramas

Modelo	Equação
Esférico	$\gamma(h) = C_0 + C \left[1,5 \frac{h}{a} - 0,5 \left(\frac{h}{a} \right)^3 \right] \text{ para } h < a$ $\gamma(h) = C_0 + C \text{ para } h \geq a$
Exponencial	$\gamma(h) = C_0 + C \left[1 - \exp\left(-\frac{h}{a}\right) \right]$
Gaussiano	$\gamma(h) = C_0 + C \left[1 - \exp\left(-\left(\frac{h}{a}\right)^2\right) \right]$

A Figura 3 ilustra alguns parâmetros do variograma teórico. Representa o que, se espera dos dados de campo, isto é, que as diferenças $\{Z(x_i) - Z(x_i + h)\}$ decresçam à medida que h , a distância que os separa, decresce. É esperado que observações mais próximas geograficamente tenham um comportamento mais semelhante entre si do que aquelas separadas por maiores distâncias. Os principais parâmetros identificados na Figura 3 são sucintamente descritos em seguida.

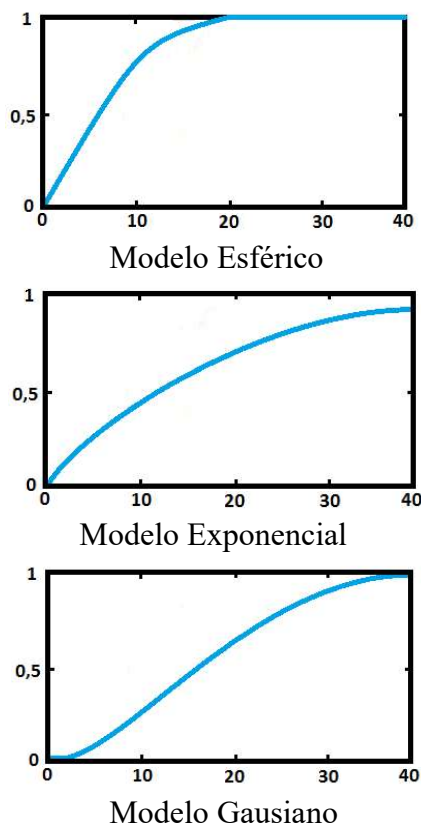


Fig. 2 – Modelos teóricos de variogramas.

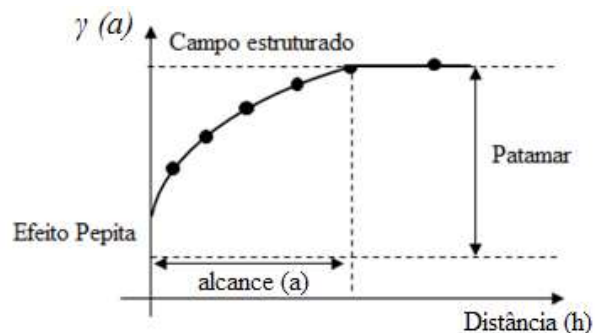


Fig. 3 - Parâmetros do variograma. Adaptado de Shinohara et al., 2006.

- (1) Alcance (a): distância dentro da qual as amostras apresentam-se correlacionadas espacialmente;
- (2) Efeito Pepita (C_0): é o ponto inicial da curva, é onde ela toca o eixo γ . O efeito pepita traduz a maior ou menor correlação espacial entre os pontos amostrados. Um valor alto deste índice indica que se encontram grandes variações em curtas distâncias;

(3) Patamar (Co+O):   o valor de γ (vari ncia) m ximo da curva correspondente   Amplitude.

2.2. Validao cruzada

Para validar os par metros dos modelos te ricos utiliza-se o teste do ponto fict cio, que consiste em remover um dado do conjunto de dados amostrais e, usando-se um estimador e funo ponderada relacionada com a dist ncia (variograma modelado), estima-se o valor retirado (JOURNEL, 1983). Tem-se, ent o, dois valores para o mesmo ponto, o observado e o estimado. O erro da estimaco pode ser calculado pela diferena entre o valor observado e o estimado para cada local amostrado.

2.3. Krigagem

A krigagem   uma t cnica de geoestat stica que possibilita fazer previs es do valor pontual de uma vari vel regionalizada em um determinado local dentro da  rea estudada. A t cnica assume que os dados recolhidos de uma determinada populao se encontram correlacionados no espao. A krigagem difere-se de outros interpoladores, pois parte do pressuposto de que pontos pr ximos no espao tendem a ter valores mais parecidos do que pontos afastados. Esse pressuposto   garantido por uma ponderao intr nseca ao m todo (SOARES, 2006).

O sistema de krigagem   constitu do por ponderadores com o objetivo de conduzir estimativas a erros nulos e minimizao da vari ncia. A definio dos ponderadores   feita a partir das curvas ajustadas dos variogramas do espao amostral e sua funo   ditar a influ ncia dos dados amostrais na estimaco de novos valores (OLEA, 1999).

A krigagem re ne um conjunto de m todos de estimaco, como: Krigagem Simples, Ordin ria, Universal, Indicativa, Disjunta, Cokrigagem e com Deriva Externa. Como j  foi citado antes, neste trabalho ser o avaliadas as aplicaes das t cnicas de Krigagem Ordin ria e Indicativa para a interpolao dos valores coletados.

A Equao 2 apresenta a formulao para o estimador de Krigagem Ordin ria:

$$z_o^*(u) = \sum_{\alpha=0}^n w_{o\alpha}(u)z(u) \quad (2)$$

sendo:

$z_o^*(u)$ - estimativa do valor na posio u ;

$w_{o\alpha}$ - ponderadores;

$z(u)$ - valor da vari vel na posio u .

Lindner (2015) aponta que as vari ncias de krigagem n o s o normalmente medidas de acur cia de estimativa local e que, para satisfazer esta necessidade, uma das solues   a utilizao da Krigagem Indicativa. No caso da Krigagem Indicativa, o objetivo passa a ser o de definir  reas com maior ou menor probabilidade de ocorr ncia de um determinado evento.

A Krigagem Indicativa requer que os valores do conjunto de dados amostrais sejam modificados segundo uma transformao n o linear $f(x)=0$ ou $f(x)=1$. Consiste em transformar uma vari vel aleat ria cont nua em uma vari vel bin ria definindo o n vel de corte, ou seja, transformar valores que est o abaixo de um determinado valor de corte em “0” e os que est o acima em “1” (YAMAMOTO & LANDIM, 2013).

A Krigagem Indicativa   dada pela Equao 3 abaixo:

$$\hat{F}(v_c) = \sum_{j=1}^n w_j i_j v_c \quad (3)$$

onde:

w_j - s o os pesos

i_j - s o os indicadores

v_c - n vel de corte

De acordo com Lima *et al.* (2008) a partir das frequ ncias acumuladas, da proporo de valores e dos variogramas experimentais, calcula-se as probabilidades para determinados locais da  rea amostral.

Neste trabalho n o se utilizou uma vari vel cont nua transformada a partir de um ponto de corte para a aplicao da KI, e sim, uma vari vel discreta “modo particular motorizado”, sendo “0” n o usa do autom vel ou moto e “1”, o uso do autom vel ou moto.

3. M TODO E RESULTADOS

O fluxograma da Figura 4 mostra a seq ncia dos procedimentos metodol gicos utilizados nesse trabalho. As sees seguintes trazem a descrio de cada etapa, bem como os resultados obtidos.



Fig. 4 - Fluxograma do método seguido.

3.1. Tratamento dos dados

Os dados utilizados são provenientes da entrevista domiciliar e da pesquisa de opinião a respeito da avaliação do transporte urbano, aplicadas em conjunto na Pesquisa Origem/Destino (O/D) de 2007/2008 na cidade de São Carlos, São Paulo (RODRIGUES DA SILVA, 2008). A cidade, de porte médio, está localizada no centro-oeste do estado de São Paulo, a 213 km de distância rodoviária da capital. Possui uma área total de 1.137,30 km², sendo a área urbana de aproximadamente 105 km². Sua população é composta de 221.936 habitantes, sendo que 96% reside na área urbana (IBGE, 2010).

Para a efetivação da entrevista domiciliar e da pesquisa de opinião foram selecionados aleatoriamente 5% dos domicílios contidos na base de informações cadastral cedida pelo Serviço Autônomo de Água e Esgoto (SAAE) do município. A referida base é formada por pontos cujas coordenadas em UTM - *Universal Transversa de Mercator*, são referentes aos hidrômetros dos domicílios e são apresentadas em metros.

A pesquisa de opinião contém dados qualitativos sobre o sistema de transporte de São Carlos e foi aplicada a um morador de cada domicílio pesquisado, gerando um banco com 2.791 casos. Este banco de dados, associado à entrevista domiciliar, foi utilizado como base para tratamento de dados e obtenção da amostra final. Para determinação da amostra, foram eliminados

os dados incoerentes ou ausentes, aqueles referentes às pessoas que não realizaram viagens, os domicílios com coordenadas geográficas repetidas e os indivíduos que utilizaram outros modos de transporte, diferentes do público, particular motorizado e não motorizado. Vale destacar que o modo não motorizado é referente apenas aos dados do modo a pé. Logo, o banco de dados final contém 1.216 casos tendo, 22 variáveis categóricas e 4 variáveis numéricas. A Tabela 2 traz as principais variáveis que compõem a amostra e a sua descrição. Na tabela também são descritas as variáveis utilizados em cada técnica.

Tabela 2: Principais variáveis da amostra

Variável	Descrição	Técnica
Principal Problema - modo a Pé	1 - risco de Atropelamento; 2 - roubos e assaltos; 3 - péssimas condições de calçadas; 4 - falta de arborização.	AD
Lotação dos ônibus	1- vazios; 2- com lotação adequada; 3- cheios; 4- superlotados.	AD
Frota	1- muito pequeno; 2- pequeno; 3-suficiente; 4-ótimo.	AD
Principal Problema - modo Ônibus	1 - rapidez, 2 - segurança, 3 - conforto, 4 - itinerário, 5 - horário.	AD
Principal Problema - modo Automóvel	1 - congestionamento; 2 - falta de estacionamento; 3 - custo elevado.	AD
Possuir CNH	1-sim 2-não	AD
Sexo do chefe da família	1 - masculino 2 - feminino	AD
Grau de instrução	1- segundo grau completo ou superior; 2 - segundo grau incompleto ou nível inferior; 3-não alfabetizado.	AD
Condição de Renda	1- 0,0 a 2 SM; 2- 2,1 a 8 SM; 3 - 8,1 a 20 SM; 4 - não respondeu.	AD
Modo principal	1-público; 2-particular motorizado; 3-não motorizado.	AD
Quantidade de Motocicletas	Quantidade de motocicletas por domicílio	AD

Quantidade de Automóveis	Quantidade de automóveis por domicílio	AD
Idade	Idade do chefe da família	AD
Total de Viagens Realizadas	Número total viagem realizada pelo chefe de família	AD
Modo particular motorizado	0: Não usa; 1: usa	KI
Probabilidades de escolha do modo particular motorizado	Variável contínua proveniente dos resultados da AD (intervalo entre 0 e 1)	KO

AD- Árvore de Decisão; KI – Krigagem Indicativa; KO – Krigagem Ordinária.

3.2. Aplicação da AD

Inicialmente, a técnica da Árvore de Decisão (AD) foi aplicada por Pitombo *et al.* (2015) com o objetivo de identificar as variáveis que contribuem para o entendimento da escolha modal e determinar uma variável contínua, para a aplicação da Krigagem Ordinária.

Identificou-se que, para o uso do modo de transporte particular motorizado, as variáveis com maior influência foram: possuir CNH (Carteira Nacional de Habilitação), possuir automóveis ou motocicletas no domicílio, ter idade inferior a 59 anos, ser do sexo masculino e considerar como principais problemas de utilização desse modo a falta de estacionamento e o custo elevado.

O modelo de AD gerou então, a partir das variáveis explicativas descritas na Tabela 2, probabilidades para a escolha modal. A probabilidade do uso do modo particular motorizado foi então a variável objeto de estudo. A Krigagem Indicativa, aplicada posteriormente, foi realizada com a variável *dummy* que caracteriza realmente o uso do modo particular motorizado (0: Não usa automóvel/moto – 1: Usa automóvel/moto).

Através da análise dos mapas exploratórios das probabilidades obtidas com a AD, percebeu-se que as variáveis não apresentavam um padrão espacial. Na tentativa de encontrar padrões espaciais das variáveis em uma menor dimensão, segregou-se a área de estudo em seis regiões considerando o critério da renda (Fig. 5). A ideia seria selecionar regiões com predominância de domicílios de menor renda, considerando a hipótese de que nestes locais, poderia haver padrões espaciais aparentes no uso do automóvel.

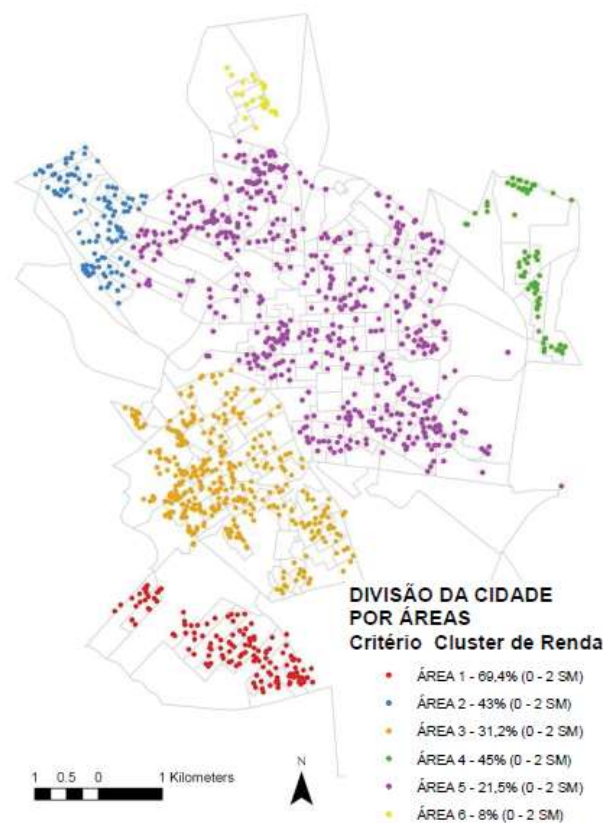


Fig. 5 - Seis regiões segundo renda e coordenadas geográficas. Fonte: Pitombo *et al.* (2015)

3.3. Seleção da região a ser interpolada

Pitombo *et al.* (2015) perceberam que a área 2 foi a que apresentou um padrão espacial, variando do centro para a periferia da região (Figura 6). A área 2 é constituída tanto por bairros considerados periféricos, como Loteamento Habitacional São Carlos I (Núcleo residencial Ivo Morganti), Dário Rodrigues (São Carlos II), São Carlos III e IV (Lourival Maricondi), Bairros Arnon de Melo (São Carlos V) e Conj. Habitacional Sta. Angelina, quanto por bairros de população de maior renda, como Parque Santa Felícia, jardim Ipanema, Montreal e condomínio Quebec, e engloba também, o campus universitário 2 da Universidade de São Paulo. Vale destacar que neste trabalho o foco é somente os resultados obtidos para o modo particular motorizado (automóvel e motocicleta).

A área 2 é formada por 110 pontos. Para a Krigagem Ordinária foram utilizadas as coordenadas em UTM para os 110 pontos, bem como as probabilidades de escolha do automóvel ou motocicleta. Já para a KI também foram utilizadas coordenadas geográficas e

variável *dummy* que representava o uso do auto/motocicleta.

Os dados foram trabalhados através do software GeoMS (*Geostatistical Modelling Software*), desenvolvido pelo Instituto Superior Técnico/Universidade Técnica de Lisboa e os mapas finais foram editados através do *Arcgis 10.1*.

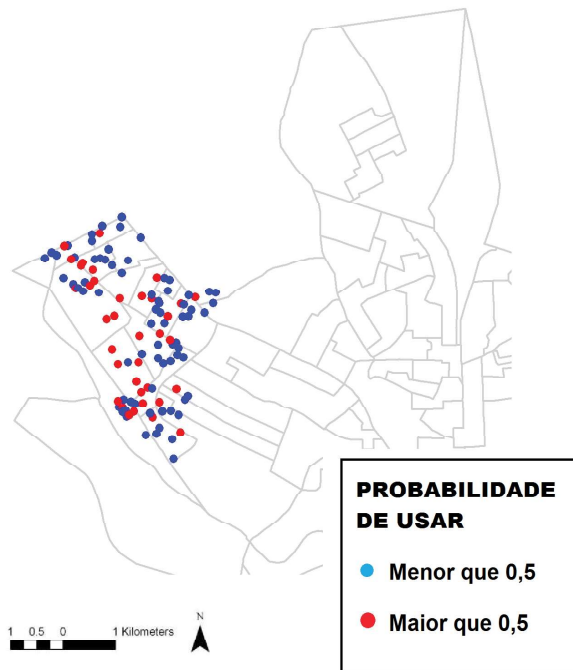


Fig. 6 - Distribuição da Probabilidade do modo particular motorizado – área 2

3.4. Modelagem de variogramas

Na modelagem dos variogramas utilizou-se duas variáveis descritas a seguir: uma contínua e outra *dummy*.

A variável contínua foi a probabilidade de escolha do modo particular motorizado, que depois foi aplicada na Krigagem Ordinária.

Transformou-se a variável “modo principal” em uma variável *dummy*, “1” para o uso do modo particular motorizado e “0” para os casos do não uso do modo particular motorizado. Posteriormente, essa variável foi utilizada na Krigagem Indicativa.

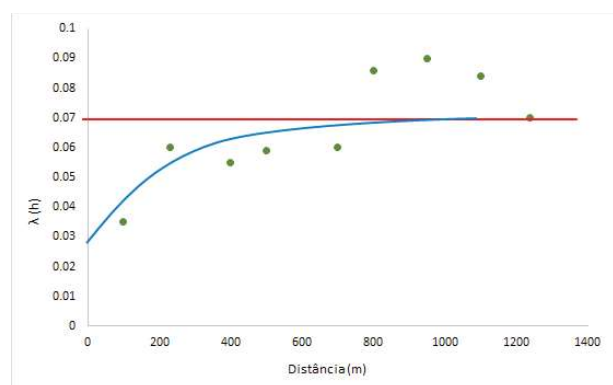
Após vários testes, identificou-se a direção principal (30°- maior dependência espacial dos dados) e a ortogonal (- 60°). A distância entre amostras (*h – Lag distance*) foi considerada na faixa de 130 metros (próxima a uma quadra – distância média entre as coordenadas

geográficas) e a distância máxima para o cálculo do variograma foi de 100.000 metros (*Cut distance*). Em seguida, foi feita a modelagem dos variogramas, determinando outros parâmetros, como o efeito pepita (*Co*), amplitude (*a*) e modelo teórico ajustado.

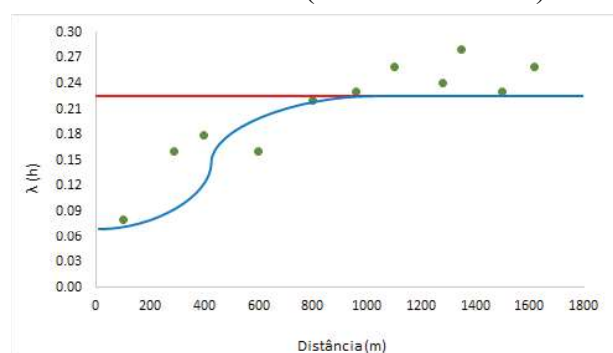
A Tabela 3 traz um resumo dos parâmetros de modelagem dos variogramas para os dois tipos de variáveis e a Figura 7 ilustra os variogramas modelados na direção 30°.

Tabela 3: Resumo dos parâmetros de modelagem dos variogramas – modo particular motorizado

Área 2	Variável contínua		Variável <i>dummy</i>	
	30°	-60°	30°	-60°
Direção	30°	-60°	30°	-60°
h (m)	131	132	131	132
Cut distance (m)	100.000	100.000		100.000
Efeito Pepita (<i>Co</i>)	0,029		0,070	
Alcance (a)	833 m	363 m	577 m	326 m
Patamar (<i>Co+C1</i>)	0,072		0,210	
Modelo	Esférico		Gaussiano	



Modelo Esférico (variável contínua)



Modelo Gaussiano (variável *dummy*)

Fig. 7 - Variogramas modelados na direção 30°.

3.5. Validação cruzada

A validação cruzada permite seleccionar o modelo matemático que melhor descreve a dependência espacial das observações das variáveis em função das distâncias. Para cada ponto existirá o valor real (amostrado) e o valor estimado, podendo-se assim obter o erro de estimação. Neste trabalho foram calculados o Erro Médio (EM); Erro Médio Normalizado (EMN) e Variância dos Erros (VAR) com o intuito de aferir a qualidade dos variogramas teóricos (Tabela 4). As equações para cálculo de tais medidas de desempenho também estão apresentadas a seguir.

$$EM = \frac{1}{N} \sum_{i=1}^N (x_i - y_i) \quad (4)$$

EM = Erro médio; X_i = valor estimado; y_i = valor observado.

$$EMN = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - y_i)}{y_i} \quad (5)$$

EMN = Erro médio normalizado

$$var(x - y) = E((x - y) - \mu)^2 \quad (6)$$

μ = Média dos erros; $x - y$ = erros

Tabela 4: Resultados da validação cruzada

Área 2	Krigagem ordinária	Krigagem indicativa
Erro médio	-0,003	-0,001
Erro médio Normalizado	0,6	0,3
Variância dos erros	0,079	0,104
% de acertos	49%	55%

Para o cálculo da % de acertos transformou-se os valores estimados (probabilidades do uso do automóvel ou motocicleta) em uma variável *dummy*, sendo que os pontos cujos valores estimados estavam acima da mediana (0,6919) foram considerados como “1” e aqueles com valores inferiores, como “0”. Esta transformação ocorreu apenas para o caso dos valores estimados na Krigagem Indicativa.

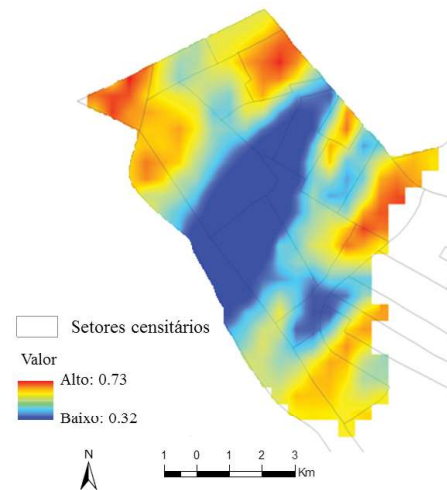
Observando as medidas de desempenho apresentadas na Tabela 4, verifica-se que as duas técnicas de krigagem são bem similares. A Krigagem Indicativa, no entanto, apresenta maior

percentual de acertos, menor erro médio e menor erro médio normalizado.

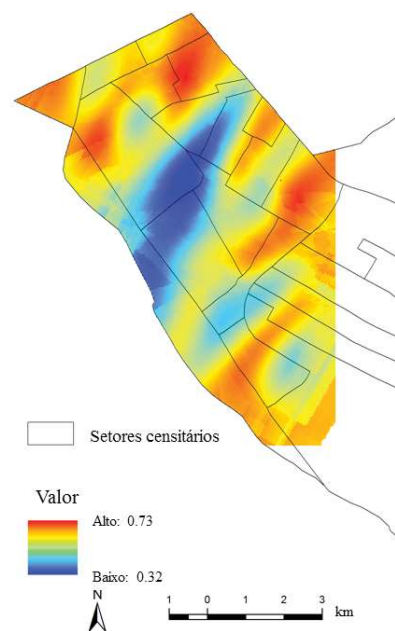
3.6. Krigagem

Os parâmetros da Tabela 3 foram utilizados para a composição dos mapas através dos processos de Krigagem Ordinária e Indicativa (Figura 8), a fim de estimar em locais não amostrados os valores das variáveis relativas à utilização do modo particular motorizado (automóvel e motocicleta).

O processo de Krigagem Ordinária foi aplicado por Pitombo *et al.* (2015) que estimaram a probabilidade de escolha modal em coordenadas conhecidas e desconhecidas. Vale ressaltar que, ao contrário da Krigagem Indicativa, a Krigagem Ordinária permite o uso apenas de variáveis contínuas.



Krigagem Ordinária



Krigagem Indicativa

Fig. 8 - Valores interpolados de probabilidades de escolha modal.

Observa-se que os mapas gerados pelas duas técnicas são bastante parecidos e indicam uma maior probabilidade da utilização do modo particular motorizado na área mais periférica da região. Essa probabilidade diminui à medida que se aproxima do centro da área, onde se localizam os conjuntos habitacionais com população com menor poder aquisitivo, o que sugere um quadro condizente com a realidade encontrada na região.

Considerando que no caso da aplicação da krigagem indicativa a variável aplicada não é contínua, a Figura 9 traz o mapa interpolado da variável *dummy* onde se tem a área azul (valor 1) representando a utilização do modo particular motorizado e a cor vermelha (valor 0) indicando a utilização de outro modo de transporte. Utilizou-se a mediana dos valores estimados para obter a transformação desses valores em 0 e 1 (valores maiores que a mediana foram classificados como 1).

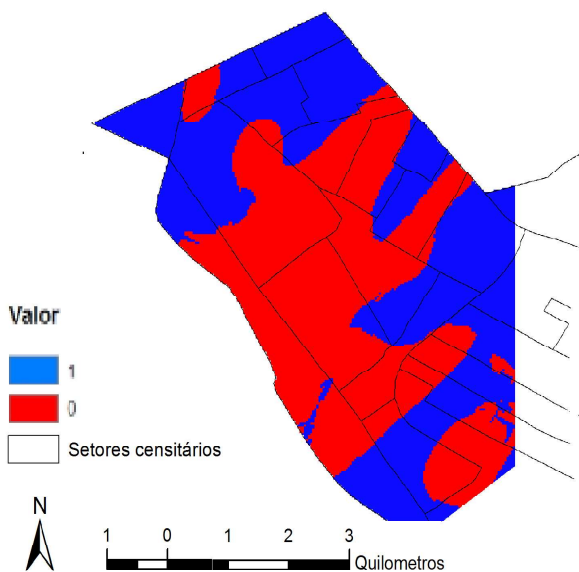


Fig. 9 - Mapa interpolado para a variável *dummy*.

Mais uma vez, observou-se uma predominância da utilização do modo particular motorizado nas regiões periféricas da região.

4. CONCLUSÕES

Este trabalho teve como objetivo analisar a aplicação de duas técnicas de Krigagem (Ordinária e Indicativa) para estimar a escolha por automóvel/motocicleta em coordenadas de valores conhecidos e desconhecidos, comparando os resultados alcançados a fim de identificar a resposta mais adequada ao caso em questão. Essa

verificação levou em consideração a análise dos variogramas, dos mapas estimados e as medidas de desempenho calculadas na validação cruzada.

Os resultados obtidos pela utilização das técnicas demonstram sua viabilidade na aplicação na área de demanda por transportes, especialmente no caso da escolha modal, sendo que entre as duas técnicas aplicadas nesse estudo (Krigagem Ordinária e Krigagem Indicativa), observou-se um incremento para o caso de utilização da Krigagem Indicativa.

A validação cruzada mostrou que na aplicação da Krigagem Indicativa obteve-se 55% de acertos, enquanto que na Krigagem Ordinária obteve-se 49% de acertos para as estimações de probabilidade de uso do modo particular motorizado. Com a KI também houve um menor erro médio e menor erro médio normalizado.

Vale destacar que na aplicação da Krigagem Ordinária, a variável interpolada “probabilidade de escolha do modo particular motorizado”, é resultante de um modelo gerado a partir da aplicação da Árvore de Decisão (AD), ou seja, indiretamente, outras variáveis influenciaram na caracterização da escolha do modo e, conseqüentemente, nos cálculos das probabilidades, propiciando o uso do procedimento metodológico para projeções futuras. Por outro lado, na aplicação da Krigagem Indicativa, considerou-se apenas a localização espacial da variável e o valor real dela, evitando assim possíveis propagações de erros. No entanto, o procedimento que consiste na aplicação e KI é adequado apenas para análises na data presente.

Ressalta-se ainda a necessidade de adaptação de dados de demanda por transportes para aplicação na modelagem geoestatística, já que o fenômeno estudado geralmente não apresenta continuidade espacial. As superfícies obtidas, no entanto, mostram manchas ou tendências de escolhas modais, de forma confirmatória, capazes de auxiliar no processo de decisão e implementação de políticas públicas.

REFERÊNCIAS BIBLIOGRÁFICAS

CAMARGO, E.; CAMARGO, C. G.; DRUCK, S.; CÂMARA, G. Análise Espacial de Superfícies. In: Druck, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.M.V. (Eds.). **Análise Espacial de Dados Geográficos**, Brasília, EMBRAPA. p.46-82, 2004.

- CARR, J.R. Application of Spatial Filter Theory to Kriging. **Mathematical Geology**. Vol. 22, No. 8, p.55-61, 1990 .
- GOMES, V., PITOMBO, C. , ROCHA, S. ; SALGUEIRO, A. Kriging Geostatistical Methods for Travel Mode Choice: A Spatial Data Analysis to Travel Demand Forecasting. **Open Journal of Statistics**, 6, p.514-527. doi: 10.4236/ojs.2016.63044, 2016.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. (2010) **Censo Demogrfico Brasileiro**. Disponvel em <http://www.ibge.gov.br>. Acesso em 30 de maro de 2015.
- ISAAKS, E.H.; SRVASTAVA, R. M. **An introduction to applied Geostatistics**. Oxford University Press, p.483-485, 1989.
- JOURNAL, A.G. ; HUIJBREGTS, JC.H. **Mininig geostatistics**. Nova York. Academic Press, 1978, 600 p., 1978.
- JOURNAL , A.G. Non-parametric estimation of spatial distribution. **Mathematical Geology**, v.15, p.445-468, 1983.
- MA, M. L. A.; MELO, H. R.; CHAVES, C. A. P.; Simo, F. B. Krigagem Indicativa aplicada  pesquisa de nitrato nas guas subterrneas da rea urbana de Porto Velho – RO. **XV Congresso Brasileiro de guas Subterrneas**. Natal. 11 a 14 de novembro de 2008.
- LINDNER, A. **Anlise desagregada de dados de demanda por transportes atravs de modelagem geostatstica e tradicional**. Dissertao (Mestrado). Escola de Engenharia de So Carlos, Universidade de So Paulo, So Carlos, p.44-81, 2015.
- LOUREIRO, C.F.G.; SILVA, H.N.; CARVALHO, L.E.X. Metodologia de anlise de regresso geograficamente ponderada aplicada ao fenmeno das viagens intermunicipais. **Panorama Nacional de Pesquisa em Transportes, XX ANPET**, v.1, p.479-491, 2006,.
- PITOMBO, C. S.; SALGUEIRO, A. R.; COSTA, A. S. G.; ISLER, C. A. A twostep method for mode choice estimation with socioeconomic and spatial information. **Spatial Statistics**. v . 11 p.45-64. DOI: 10.1016/j.spasta.2014.12.002, 2015.
- RAJAMANI, J., BHAT, C., HANDY, S., KNAAP, G., SONG, Y. Assessing the impact of urban form in nonwork trip mode choice after controlling for demographic and level-of-service effects. **Annual Meeting TRB**, p158-165, 2003.
- SHINOHARA, E. J.; RODRIGUES FILHO, O. S.; BERNUCCI, L. L. B.; QUINTANILHA, J. A. Uso da ferramenta de anlise geostatstica para estudo de atrito em pista de aeroporto. In: XX ANPET - Congresso de Pesquisa e Ensino em Transportes, 2006, Braslia. **Anais do XX ANPET - Congresso de Pesquisa e Ensino em Transportes**, 2006. v. 2. p. 1377-1386.
- RODRIGUES DA SILVA, A. N. **Elaborao de um Banco de Dados de Viagem para Auxlio ao Desenvolvimento de Pesquisas na rea de Planejamento dos Transportes**. Relatrio FAPESP, Processo n 04/15843-4. Escola de Engenharia de So Carlos, Universidade de So Paulo, So Carlos, p.180-203, 2008.
- SOARES, A. **Geostatstica para as Cincias da Terra e do Ambiente**. 2ª ed. Lisboa: IST Press, 214p, 2006.
- OLEA, R.A. **Geostatistics for Engineers and Earth scientists**. Boston: Kluwer Academic Publishers, p444-455, 1999.
- ORTZAR, J.D., WILLUMSEN, L.G., **Modelling Transport**. Londres: Wiley. 4ª ed., 586p, 2011.
- WACKERNAGEL, H. (2010) **Multivariate Geostatistics: An introduction with applications**. Springer, Third Edition, 381p.
- YAMAMOTO, J. K., LANDIM, P. M, B. **Geostatstica: Conceitos e Aplicaes**. So Paulo: Oficina de Textos. 215p, 2013.