

Revista Brasileira de Cartografia (2016), N^o 68/4, Edição Especial Geoinformação e Análise Espacial: 705-715
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto
ISSN: 1808-0936

INTERPOLAÇÃO ESPACIAL DE VIAGENS URBANAS POR TRANSPORTE COLETIVO ATRAVÉS DE AMOSTRA SISTEMÁTICA DESAGREGADA ARTIFICIAL

*Spatial Interpolation of Transit Urban Trips through an
Artificial Systematic Disaggregated Sample*

Samille Santos Rocha¹, Cira Souza Pitombo¹ & Ana Rita Salgueiro²

¹Universidade de São Paulo – USP

Escola de Engenharia de São Carlos

Avenida Trabalhador Sancarlense, 400, Departamento de Transportes, CEP: 13566-590, São Carlos, SP, Brasil
samille@usp.br, cirapitombo@gmail.com

²Universidade Federal do Ceará – UFC

Departamento de Geologia

Campus do Pici - Bloco 912, CEP: 60440-900, Fortaleza, CE, Brasil
geo.ritasalgueiro@gmail.com

Recebido em 29 de Novembro, 2014/ Aceito em 23 de Junho, 2015

Received on November 29, 2014/ Accepted on June 23, 2015

RESUMO

O principal objetivo deste trabalho é estimar a produção de viagens por Transporte Coletivo, na hora pico, em pontos não amostrados espacialmente, através da técnica de krigagem com deriva externa. Desta forma, as interpolações espaciais são realizadas considerando o par de variáveis POPULAÇÃO (variável secundária) X PRODUÇÃO DE VIAGENS (variável primária). Os dados utilizados são oriundos da Pesquisa Origem-Destino, realizada na Região Metropolitana de Salvador, Bahia, em 1995, e dados do censo 2010. Com isso, é realizada a estimação de produção de viagens para o ano de 2010 por macrozona de tráfego (MZ), a partir de dados censitários, e interpolados dados de produção de viagens através da variável população. Vale ressaltar que a realização da interpolação foi feita com uma amostra sistemática desagregada artificial, proveniente de dados agregados. Este trabalho possui essencialmente duas contribuições: estima dados de viagens urbanas em diversos pontos desconhecidos; e possibilita a obtenção de dados de viagens urbanas a partir de variáveis como população, facilmente disponíveis.

Palavras-chave: Produção de Viagens por TCO, Geoestatística, Amostra Desagregada.

ABSTRACT

The main aim of this work is to estimate transit trip production in known and unknown geographic coordinates through kriging with external drift. Thus, the spatial interpolation is performed considering the pair of variable: POPULATION (secondary variable) X TRIP PRODUCTION (primary variable). The data used are from Origin-Destination Survey, performed in Salvador Metropolitan Area, Bahia (Brazil), in 1995, and census data 2010. Thus, trip production estimation is carried out to the year of 2010, based on census data. It is noteworthy that the interpolation was performed with a systematic disaggregated sample from aggregate data.. This paper presents mainly two contributions: estimates

urban trips data in several unknowns' points; and allows the achievement of urban travel data from variables such as population, easily available.

Keywords: Transit Trip Production, Geostatistics, Systematic Disaggregated Sample.

1. INTRODUÇÃO: SÍNTESE DA LITERATURA

Estudos de demanda por transportes, fundamentais para o planejamento urbano, podem ser gerados a partir de dados espacialmente dependentes. No entanto, técnicas estatísticas e modelos tradicionais de previsão de demanda por transportes assumem que as observações não estão correlacionadas no espaço (LOPES e RODRIGUES da SILVA, 2005).

No planejamento de transportes alguns conceitos são bastante utilizados, e alguns deles são objeto de estudo deste trabalho. Segundo Ortúzar e Willumsen (2011), *viagem* pode ser definida como o movimento de único sentido a partir de um ponto de origem para um ponto de destino; *produção de viagens* (enfoque deste estudo) é a quantidade de viagens originadas numa determinada zona de tráfego ou região; *atração de viagens* é quantidade de viagens atraídas em uma zona ou região; *geração de viagens* é o número total de viagens produzidas e atraídas por zona de tráfego ou região.

Entende-se que o comportamento relativo a viagens urbanas está relacionado com características dos indivíduos e domicílios, bem como localização espacial dos domicílios, dos destinos e a distribuição de atividades no meio urbano. Assim, a incorporação de variáveis relacionadas ao espaço e de coordenadas geográficas aos estudos de demanda por transportes torna-se importante para aperfeiçoamento das estimativas (PITOMBO et al., 2015).

A influência de variáveis espaciais e a disponibilidade cada vez maior de informação georreferenciada viabilizaram a incorporação de dados espaciais, mais efetivamente, em estudos de demanda por transportes e comportamento subjacente a viagens. Tais trabalhos têm sido identificados como áreas emergentes de pesquisa (Páez et al., 2013).

Recentemente, pesquisadores têm obtido sinais de interdependência espacial no comportamento de viajantes. Bhat e Zhao (2002) identificaram fatores espaciais que devem ser

incorporados na modelagem de demanda por transportes. Os autores formularam um *modelo logit de múltiplo nível espacial* para estimação de geração de viagens e atividades na Região Metropolitana de Boston. Adjemian et al. (2010) mostraram que a escolha de posse de diferentes tipos de veículos é espacialmente dependente. Peer et al. (2013) aplicaram Regressão Geograficamente Ponderada para estimação de tempo de viagens.

Dentre as técnicas de estatística espacial destaca-se a geoestatística, que possibilita o desenvolvimento de estudos sobre fenômenos, cujas variáveis têm distribuição de valores associada à sua posição no espaço, permitindo principalmente estimar o valor de uma variável numa localização não amostrada. Contudo, percebe-se que esta técnica ainda é pouco aplicada na área de demanda por transportes. Tanto na literatura nacional quanto na literatura internacional, trabalhos envolvendo tais análises e geoestatística são absolutamente raros.

Pitombo e Sousa (2009) estimaram taxas de geração de viagens, através de procedimentos geoestatísticos e dados agregados. Já Pitombo et al. (2010) comparam duas diferentes técnicas de análise espacial de dados para a previsão de geração de viagens urbanas. As técnicas utilizadas foram: geoestatística multivariada e a Regressão Geograficamente Ponderada. Em seguida, Costa et al. (2013) estimaram a probabilidade de escolha modal, através de dados desagregados e geoestatística.

Teixeira (2003) aplica a krigagem para relacionar aspectos socioeconômicos à geração de viagens. Este último autor apresenta um método para definir áreas homogêneas relacionadas ao planejamento de transporte, considerando aspectos econômicos e de viagens da população através dos setores censitários. Este zoneamento foi definido a partir de dados censitários de rendimento do responsável do domicílio (IBGE), utilizando a krigagem ordinária.

Outro estudo faz uma análise da previsão de tempo de viagem por automóvel utilizando a krigagem universal. A previsão baseia-se na característica que pontos próximos (no espaço)

terão tempos de viagem parecidos. Tal estimativa baseia-se no fato de que a distância é proporcional ao tempo de rotas de viagem que possuem origens e destinos próximos. (MIURA, 2010).

Este trabalho possui dois objetivos importantes na adequação de dados de viagens para modelagem geoestatística: (1) propor uma amostragem sistemática desagregada através de dados agregados de Pesquisa Origem/Destino e censo demográfico (IBGE, 2010) e (2) estimar e realizar uma interpolação espacial de produção de viagens por Transporte Coletivo (TCO) através de técnicas de krigagem com deriva externa.

A krigagem com deriva externa é utilizada quando uma variável principal (ou primária) apresenta relação de dependência com a variável auxiliar externa (secundária). Neste trabalho é feita estimação da produção de viagens por TCO a partir da variável secundária população. A Figura 1 descreve as principais etapas metodológicas e apresenta a estrutura do presente trabalho.

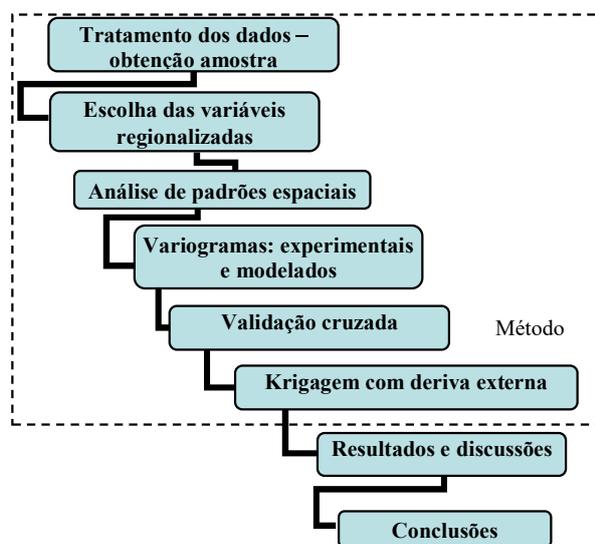


Fig. 1 - Procedimentos metodológicos e estrutura do trabalho.

2. MÉTODO

Esta seção apresenta algumas características da região de estudo – Região Metropolitana de Salvador (RMS), bem com o tratamento dos dados da Pesquisa OD de 1995 e do censo 2010. Além disso, as principais etapas da análise geoestatística serão discutidas.

2.1 Tratamento dos dados

Os dados deste trabalho são provenientes da Pesquisa Origem/Destino, realizada na Região Metropolitana de Salvador (RMS) em 1995 e do censo 2010. Foram usados dados agregados e posteriormente desagregados padronizados: população e produção de viagens por TCO. A padronização dos dados foi feita a partir da Equação 1. Em seguida foi realizada uma normalização, considerando que o menor dado (maior valor em módulo dos valores negativos) tomasse valor zero e os demais valores se tornassem relativos ao novo valor nulo.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

onde,

z = Valor padronizado de ;

μ = Média;

σ = Desvio padrão.

Os dados agregados foram desagregados a partir de uma malha sistemática regular, com espaçamento de 2000x2000 metros. Os mesmos foram associados às coordenadas geográficas dos centroides das quadrículas da malha. Este tipo de amostragem foi uma tentativa de adaptação dos dados originais da pesquisa OD da RMS (1995), para aplicação da geoestatística, já que um dos pressupostos desta técnica é a continuidade espacial do fenômeno estudado. Dados agregados por unidade de área e associados aos seus centroides são espacialmente discretos.

Em uma análise final, foi possível fazer a estimação espacial da produção de viagens por TCO, para o ano de 2010, através de técnicas de geoestatística e de dados censitários (IBGE, 2010).

Neste trabalho foi realizado tratamento de duas bases de dados, nas duas etapas ilustradas no esquema da Figura 2.

2.1.1 Pesquisa O/D de 1995

A Região Metropolitana de Salvador (RMS) tem 4.375 km² de extensão e possui uma população de cerca de 3.600.000 habitantes. Deste total, cerca de 2.700.000 reside na capital (IBGE, 2010). Em 2010, Salvador ainda detinha 75% da população da RMS. Além de ser centro financeiro, polo educacional e de saúde,

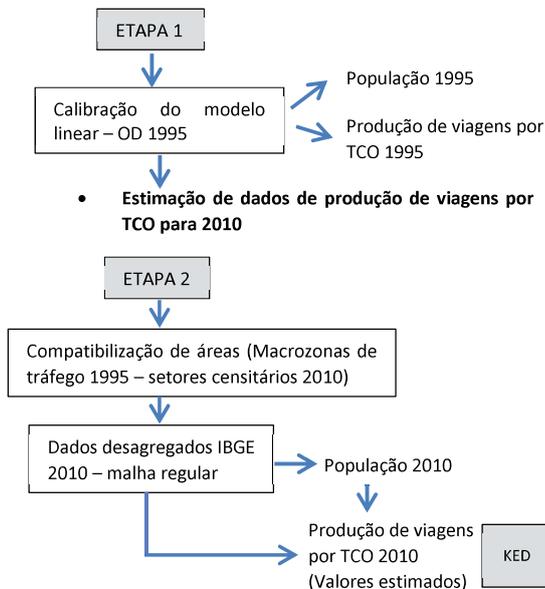


Fig. 2 - Esquema de tratamento de dados do trabalho.

Salvador tem a sua economia voltada ao setor de serviços, a exemplo do comércio, transportes, telecomunicações, turismo e atividades culturais (RIBEIRO, 2012).

A Região Metropolitana de Salvador possuía 13 municípios, no entanto, a área de estudo considerada neste trabalho, compreende apenas o município de Salvador. A grande maioria dos deslocamentos da RMS neste período concentrava-se na capital, e o ônibus era o principal modo de transporte utilizado (SETPS, 1995).

Esta pesquisa revelou que eram realizadas 3.691.889 viagens diárias na área de estudo, sendo que 22% (815.723) concentrava-se no período de pico da manhã, entre 6 e 8 horas. Cerca de 55% das viagens eram realizadas por ônibus.

A unidade de área administrativa da Pesquisa OD de 1995 é a macrozona de tráfego (MZ). Estas macrozonas foram classificadas em sub-regiões e em quatro grandes regiões: Área Urbana Consolidada (AUC), Orla, Miolo e Subúrbio (Figura 3).

O Subúrbio Ferroviário e o Miolo são as áreas mais populosas de Salvador (Figura 4), correspondendo juntas a 46% do total da população do município.

2. 2 Calibração do modelo linear – Etapa 1

A técnica de Regressão Linear Múltipla foi utilizada para prever a produção de viagens

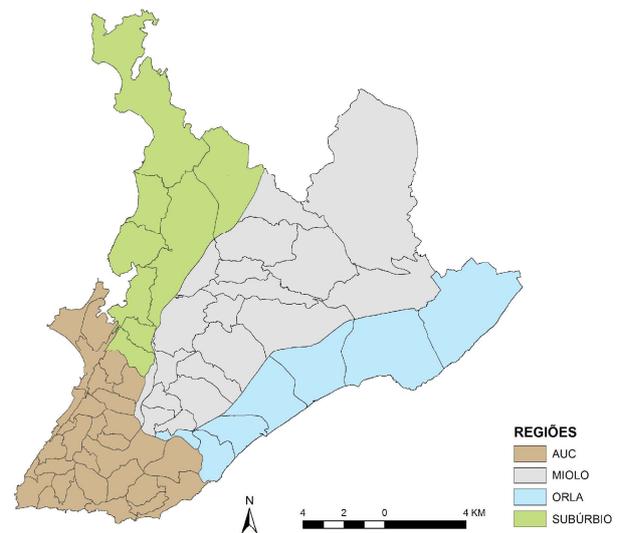


Fig. 3 - Região Metropolitana de Salvador.

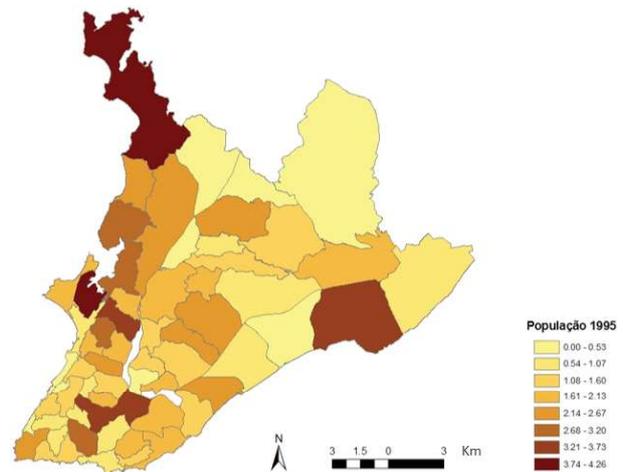


Fig. 4 - População por MZs de tráfego. Fonte dos dados: SETPS (1995)

por Transporte Coletivo para 2010, através das variáveis população e renda média do responsável por domicílio. O modelo foi inicialmente calibrado com a amostra de dados correspondente à pesquisa OD de 1995 e as mesmas variáveis. A equação calibrada com os dados de 1995 corresponde à Equação 2. Com a utilização deste modelo, foi possível estimar dados de produção de viagens por TCO para 2010.

$$Prod_{TCO} = 0.17 \times Pop - 0.218 \times Ren \quad (2)$$

para,
 $Prod_{TCO}$ = Produção de viagens por TCO;
 Pop = População da macrozona equivalente;
 Ren = Renda média do responsável do domicílio.

A Tabela 1 traz os principais resultados do modelo linear obtido com os dados de 1995. Esta tabela mostra que modelo de regressão linear múltipla obteve um elevado coeficiente de determinação, além disso, os parâmetros estimados fazem sentido na explicação do fenômeno produção de viagens. É possível verificar que a variável população foi a mais significativa para estimar a variável analisada.

Tabela 1: Resultados do modelo calibrado para estimação de produção de viagens

Modelo		Variáveis independentes		
R ²	0.964	coef		t
Soma dos quadrados				
Regressão	3.75E+09	Constante	0	
Resíduos	1.38E+08	Renda	-2.18E-01	-3.706
Total	3.88E+09	População	0.177	30.39

Vale ressaltar que para aplicação da krigagem com deriva externa, são necessários valores de variáveis primárias e secundárias. Por esta razão, primeiramente foram estimados dados de população (variável secundária) em 2010, para posteriormente, com base nos valores estimados sobre população, se conhecer valores de produção de viagens para 2010 (variável primária) em diversas coordenadas.

2.3 Tratamentos dos dados do censo de 2010: compatibilização de áreas – Etapa 2

Na segunda etapa do tratamento dos dados foram utilizados dados disponíveis pelo IBGE (2010). Inicialmente, foram compatibilizados os setores censitários, referentes ao último censo, com as MZs de tráfego da Pesquisa OD de 1995 (dados originais).

Segundo Ortúzar e Willumsen (2011) o zoneamento do tráfego deve ser compatível com outras divisões administrativas, em particular com os setores censitários. A zona deve ser mais homogênea possível em relação ao uso do solo e característica da população.

A análise das duas áreas e sua compatibilização foi feita através do *software* Arcgis 9.3. Neste SIG foi feita a sobreposição dos setores censitários do censo 2010 com MZs da pesquisa OD 1995, para extrair dados atuais de população para as MZs. Selecionados os setores pertencentes a cada MZ, foi utilizada a ferramenta PROC-V do Excel para localizar

os setores dentro de uma zona, através do seu código. Após a compatibilização foi extraído o total da população e a renda média do responsável por domicílio por MZ em 2010.

Os dados de renda foram corrigidos em relação ao percentual da diferença do salário mínimo em relação a 1995. Para a correção da renda em relação ao valor do Salário Mínimo (SM) de 1995, foi encontrado o percentual de reajuste do valor do SM por ano. Foi possível verificar os valores dos SM, e obter a taxa de crescimento de 1995 a 2010. Encontrou-se um valor de aproximadamente 170% de aumento do SM entre os dois períodos.

2.4 Amostra desagregada

Para desagregar o banco de dados, adotou-se a utilização da amostragem sistemática artificial, em que foram seguidos os seguintes passos: (1) Determinação da malha; (2) Geração de centroides por cada célula que compõe a malha; (3) Determinação de valores das variáveis padronizadas Produção de viagens por TCO e população associados as coordenadas dos centroides das células. As células obtidas e os seus centroides associados estão ilustrados na Figura 5.

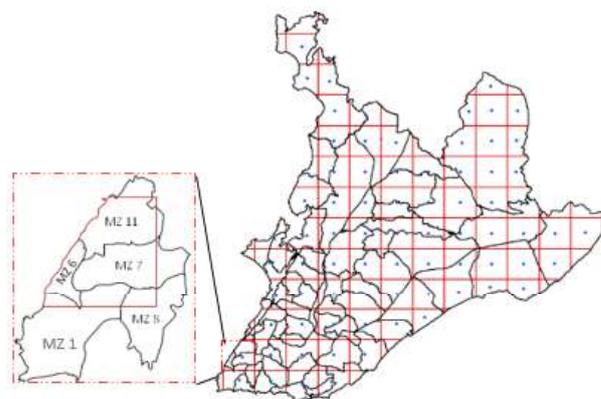


Fig. 5 - Malha gerada para amostragem sistemática artificial

Cada célula era composta por diferentes macrozonas. Tornou-se necessário, então, mensurar a ocupação, em porcentagem, das macrozonas em cada célula, a fim de determinar o valor da variável produção de viagens para cada quadrícula. A partir dos valores conhecidos das macrozonas, atribuiu-se um peso para cada MZ, considerando sua área inserida em cada quadrícula. Por fim, multiplicou-se o peso de

cada MZ pelo seu respectivo valor de produção de viagem e população. O somatório dessas multiplicações resultou no valor atribuído as quadrículas, que foram representados pela coordenada geográfica de seus centroides.

O diagrama de dispersão das variáveis analisadas para estimação de viagens é apresentado na Figura 6.

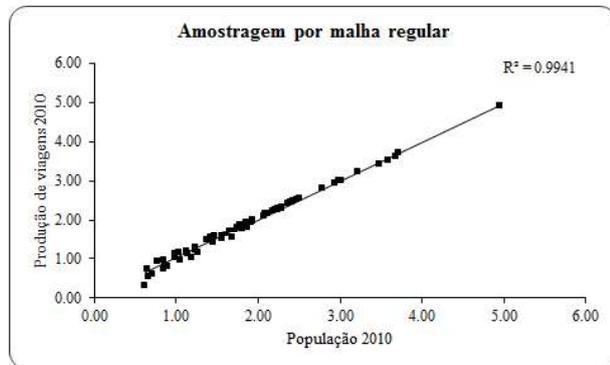


Fig. 6 - Diagrama de dispersão de valores observados – Produção de viagens x População.

Este diagrama é referente as variáveis produção de viagens por TCO e população por centroide das quadrículas, par de variáveis altamente correlacionadas entre si.

2.5 Escolha das Variáveis Regionalizadas (VR)

Camargo *et al.* (2004) destacam que as variáveis regionalizadas variam no espaço e no tempo, não se distribuindo aleatoriamente no espaço e sim, dentro de uma certa distância e direção. A variação espacial de uma variável regionalizada pode ser expressa 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.

Neste trabalho, as variáveis regionalizadas escolhidas para tratamento geoestatístico foram: População 2010 (variável secundária); Produção de viagens por TCO 2010 (variável primária).

2.6 Análise de padrões espaciais

Através da visualização da dispersão de cada uma das variáveis no espaço pode-se ter uma ideia de como estas se comportam, quais as direções que provavelmente devem ser observadas e enfim adotar parâmetros para a próxima etapa dos procedimentos geoestatísticos,

que é o cálculo de variogramas experimentais.

O objetivo desta análise consiste em verificar a forma como a variável em estudo se distribui no espaço, de modo a investigar se as variáveis em análise são regionalizadas. Foram gerados mapas da distribuição dos pontos das variáveis: população e produção de viagens por TCO (2010). A Figura 7 mostra a dispersão dos valores das variáveis Produção de viagens por TCO.

Observa-se que valores de produção de viagens por TCO têm uma tendência de crescimento conforme proximidade do subúrbio ferroviário (região com menor renda per capita e mais populosa).

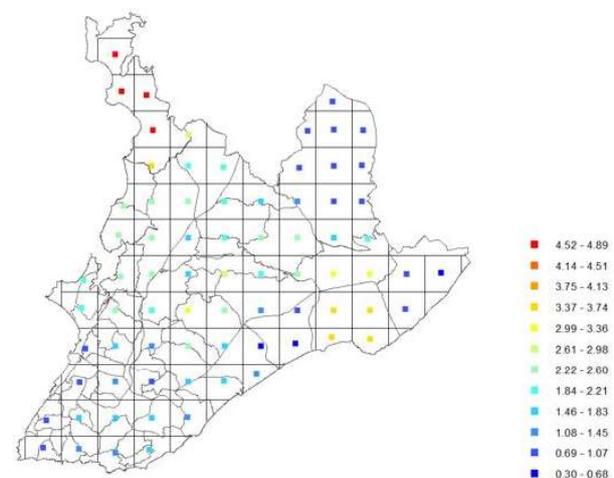


Fig. 7 - Distribuição dos valores padronizados na RMS: Produção de viagens por TCO 2010.

2.7 Obtenções de variogramas experimentais e ajustados

O variograma permite representar quantitativamente a variação de um fenômeno regionalizado no espaço. O variograma é a descrição matemática do relacionamento entre a variância de pares de observações (pontos) e a distância separando estas observações (h).

Na determinação experimental do variograma, para cada valor de h , distância entre as observações, consideram-se todos os pares de observações n , conforme Equação 3.

$$y(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_j)]^2 \quad (3)$$

onde,

$y(h)$ é o estimador do variograma (também

conhecido por semivariograma); $N(h)$ é o número de pares de valores medidos para cada valor de h ; $Z(x+h)$ e $Z(x)$ são pares de observações separados pelo vetor h .

Há alguns parâmetros do variograma, ilustrados na Figura 8, que devem ser observados para aplicação da Krigagem.

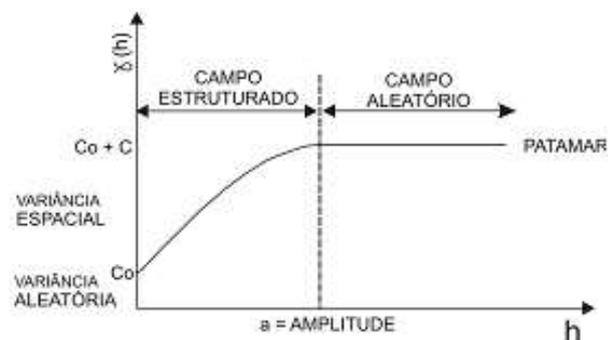


Fig. 8 – Variograma teórico representando o comportamento espacial da variável.

(1) Amplitude (a): distância dentro da qual as amostras apresentam-se correlacionadas espacialmente; (2) Patamar (C_0+C): é o valor de γ máximo da curva correspondente à Amplitude, onde, na maioria dos casos, o patamar é equivalente a variância total dos dados; (3) Efeito Pepita (C_0): é o ponto inicial da curva, é onde ela toca o eixo γ . O efeito pepita traduz o quanto pequenas distâncias são parecidas ou diferentes. Um valor alto deste índice indica que se encontram grandes variações em curtas distâncias.

Após a determinação do variograma experimental, e análise da continuidade espacial dos dados, obtém-se um variograma teórico, através de um modelo que o ajusta em função da variância dos dados ($\gamma(h)$), determinada pela distância entre os pontos observados.

Existem alguns tipos de variogramas, dentre eles estão o cúbico, linear, esférico, exponencial e o gaussiano. Estes três últimos são os mais usuais, pois explicam a variabilidade espacial de grande parte dos fenômenos espaciais (Yamamoto e Landim, 2013). As Equações 4a, 4b e 4c apresentam matematicamente os principais modelos teóricos de variogramas.

Esférico

$$y(h) = C_0 + C \left[1,5 \frac{h}{a} - 0,5 \left(\frac{h}{a} \right)^3 \right], \text{ para } h < a$$

$$y(h) = C_0 + C, \text{ para } h \geq a \tag{4a}$$

Exponencial

$$y(h) = C_0 + C \left[1 - \exp \left(-\frac{h}{a} \right) \right] \tag{4b}$$

Gaussiano

$$y(h) = C_0 + C \left[1 - \exp \left(-\left(\frac{h}{a} \right)^2 \right) \right] \tag{4c}$$

Foram estabelecidos os parâmetros para construção dos variogramas experimentais, referentes aos valores de produção de viagens por TCO e população. Em seguida, foi feita a modelagem dos variogramas. Através do modelo teórico ajustado foram determinados alguns parâmetros, tais como efeito pepita (C_0) e Amplitude (a). Neste trabalho não foi considerado o patamar do variogramas, pois observou-se que a variabilidade espacial das variáveis aumenta de acordo com a distância, sem se tornar constante a partir de determinado ponto.

A Tabela 2 apresenta o resumo dos parâmetros de modelagem de cada um dos variogramas e a Figura 9 apresenta a modelagem dos mesmos. O modelo de ajuste escolhido para os modelar os variogramas foi o modelo teórico esférico, pois, dentre os modelos testados, foi o que melhor se ajustou à distribuição espacial dos dados.

Tabela 2: Resumo dos parâmetros de modelagem dos variogramas

Variáveis (por MZ) e Direções	Efeito Pepita (C_0)	(C1)	Alcance	C total
População 2010 (omnidirecional)	0,159	2,44	28799,25	2,60
Produção de viagens 2010 (omnidirecional)	0,225	2,178	26591,50	2,40

2.8 Validação cruzada

A Validação Cruzada, também conhecida como teste do ponto fictício, é a técnica que obtém a estimativa dos pontos amostrados (coordenadas conhecidas), usando os valores dos pontos vizinhos e o modelo teórico do variograma ajustado. Portanto, para cada ponto existirá o valor verdadeiro (amostrado) e o valor estimado, podendo-se assim obter o erro de estimação. Com o objetivo de aferir a qualidade dos variogramas ajustados ou teóricos, foram calculados parâmetros como coeficiente de

correlação, erro médio e variância média dos erros que serão descritos na próxima seção. Neste trabalho, foi considerada uma amostra com 60% dos pontos amostrados para ajuste dos variogramas e 40% dos pontos para validação da modelagem geoestatística.

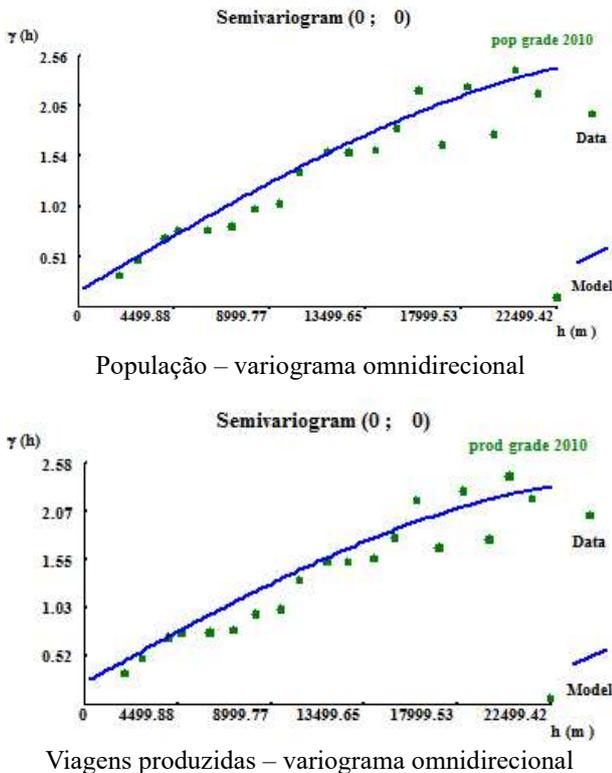


Fig. 9 - Variogramas modelados.

2.9 Krigagem com Deriva Externa

Yamamoto e Landim (2013) descrevem Krigagem como um processo de estimativa de valores de variáveis distribuídas no espaço a partir de valores adjacentes, considerados como interdependentes pelo variograma. A krigagem é um estimador utilizado principalmente para previsão do valor pontual de uma variável regionalizada em um determinado local dentro do campo geométrico.

Muitos problemas têm natureza multi-variada e a interdependência entre os dados deve ser incorporada nas análises. Em muitas situações, uma variável é obtida de forma mais prática, ou mais fácil ou mais barata ou mais rápida do que outras e, dada essa interdependência, é possível se obter valores de uma através de outra variável.

A krigagem com deriva externa permite a estimação de uma variável primária a partir de informações de uma variável secundária

amplamente amostrada. Estas precisam ser multicolocalizadas (ROCHA *et al.*, 2009). Neste caso, além dos pontos observados da variável primária, o valor da variável secundária é conhecido por toda a malha a ser estimada (OLEA, 1999).

Considere-se $Z(x)$ como a variável primária e $Y(x)$ a variável secundária, ambas apresentando dependência linear (Equação 5), conforme escreve Wackernagel (2010). O estimador da krigagem com deriva externa pode ser escrito conforme Equação 6.

$$E[Z(x)] = a_0 + b_1 y(x) \quad (5)$$

sendo:

a_0 e b_1 os coeficientes implicitamente estimados junto com a variável a ser estimada $Z(x)$; e $y(x)$ valor da variável secundária.

$$y(x_0) = \sum_{i=1}^n \lambda_i y(x_i) \quad (6)$$

Assim, neste trabalho são considerados os parâmetros dos modelos teóricos dos variogramas das variáveis primárias e secundárias e os parâmetros da malha de pontos a estimar (2000 x 2000 metros, neste trabalho).

3. RESULTADOS E DISCUSSÕES

Para validação da modelagem geoestatística foram calculadas algumas medidas estatísticas de ajuste, descritas na Tabela 3.

Observa-se que a validação cruzada apresentou valores pequenos de erros (médio e quadrático). Além disso, os valores de coeficiente de *Pearson* entre valores observados e estimados são elevados.

Tabela 3 - Resultados da validação cruzada.

Variável	Coefficiente de correlação	Média de erros	Erro quadrático médio
População (2010)	0,880	0,004	0.279
Produção de viagens por TCO (2010)	0,998	0,000	0,002

Pode-se concluir que o método de desagregação de dados necessita ainda de melhor adaptação e aprimoramento, como a utilização

de técnicas de simulação, em detrimento da obtenção de médias por peso. No entanto, os resultados indicam que a técnica de Krigagem com deriva externa pode ser adequada para as aplicações descritas neste trabalho, sobretudo considerando variáveis com distribuição espacial similar como população e produção de viagens, por exemplo.

3.1 Mapas interpolados

Finalmente, estimou-se dados de viagens através da krigagem com deriva externa em pontos não amostrados, a partir de dados de população.

Foram gerados mapas de interpolação para população, que auxiliou na estimação por krigagem com deriva externa para a produção de viagens em 2010 (Figuras 10 e 11).

O mapa interpolado de produção de viagens por TCO para 2010 é bastante similar ao mapa de população, devido à alta relação existente entre população e produção de viagens.

Observa-se que há maior concentração de produção de viagens na região do Subúrbio Ferroviário. Essa tendência que vai diminuindo à medida que se aproxima da AUC, região mais central de Salvador.

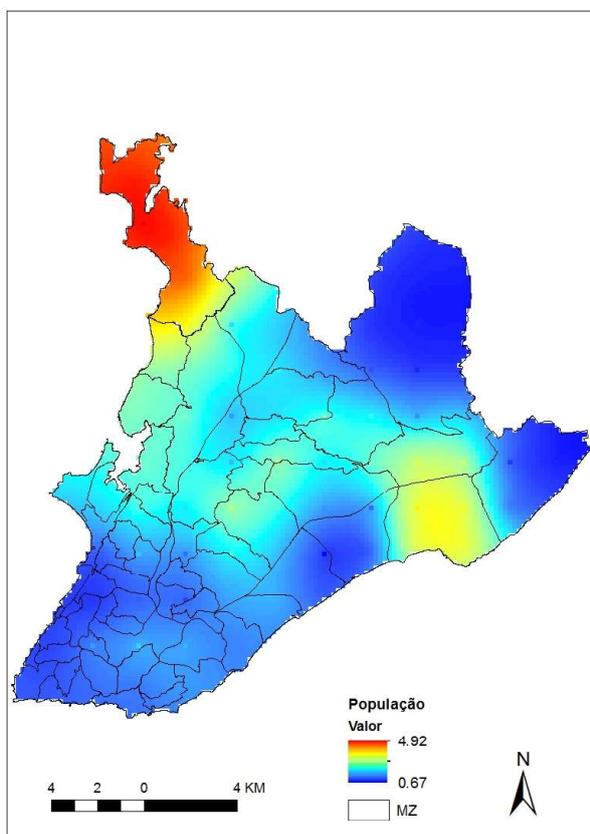


Fig. 10 - Valores interpolados de população 2010

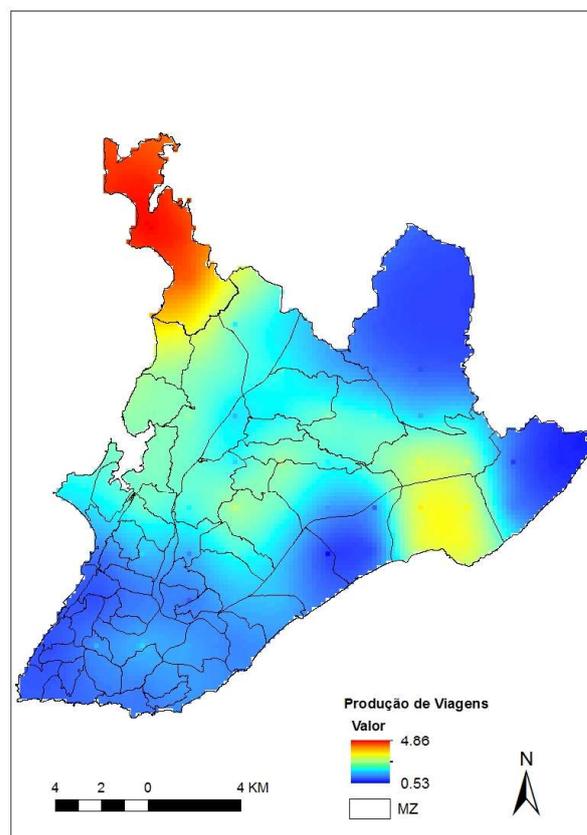


Fig. 11 - Valores interpolados de Produção de viagens por TCO para 2010

4. CONCLUSÕES

Os resultados obtidos neste trabalho permitiram determinar valores de produção de viagens por TCO em coordenadas geográficas amostradas e não amostradas através da utilização de Krigagem com Deriva Externa para dados de produção de viagens para 2010, obtidos a partir de dados populacionais.

Para a aplicação da krigagem, a amostra original (por macrozona de tráfego) foi adaptada, utilizando-se uma malha regular para obter dados menos agregados. Com a amostragem sistemática da Pesquisa O/D da RMS de 1995, os valores foram padronizados. Desta forma, coordenadas de centroides da malha obtida para a área estudada foram associadas a valores de viagens produzidas e população residente por macrozona de tráfego.

Em outra etapa, dados populacionais do censo demográfico brasileiro de 2010 foram compatibilizados com as macrozonas de tráfego, possibilitando a aplicação da mesma técnica de amostragem sistemática, realizada para os dados de 1995. Desta forma, foi possível fazer uma

interpolação espacial considerando estimativas de produção de viagens por TCO para o ano de 2010 (variável primária) a partir da variável população de 2010 (variável secundária).

Para construção e modelação dos variogramas, foram considerados modelos esféricos e omnidirecionais sem patamar. Vale ressaltar que foi possível a construção de mapas de krigagem para estimação da produção de viagens por TCO em aproximadamente 10.900 novas células, com espaçamento de 200x200 metros.

Conclui-se que o resultado da krigagem na RMS é condizente com a realidade do local. Observou-se que o aumento da quantidade de viagens produzidas por TCO se dá do centro para o subúrbio ferroviário na região. Locais com menor renda e maior população possuem maior tendência a produzir mais viagens por TCO.

A validação cruzada mostrou bons resultados considerando erros médios de estimação, erro quadrático médio e ótimos valores de coeficiente de correlação entre viagens estimadas (krigagem) e observadas (próximos a 1, o que indica alta relação linear positiva). Foi observada, além disso, alta correlação tanto de valores estimados quanto de valores observados entre os pares de variáveis Produção x População.

A partir da interpolação espacial proposta neste trabalho, foi possível obtenção de estimativas de produção de viagens em qualquer ponto que compõe a superfície interpolada. Esta é uma contribuição significativa, sobretudo considerando a carência de dados relativos a viagens urbanas na região de estudo.

Uma questão importante a ser levada em consideração é que as variáveis estimadas não são naturais, foram levantadas a partir da pesquisa O/D e dados censitários. Deve-se levar em consideração, no entanto, o caráter inédito e incipiente do estudo, no que diz respeito à utilização de técnicas pouco comuns em estudos de análise de demanda por transportes.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

ADJEMIAN, M. K.; LIN, C.; WILLIAMS, J. (2010) Estimating spatial interdependence in automobile type choice with survey data. **Transportation Research Part A**, vol 44, p. 661-675. DOI: 10.1016/j.tra.2010.06.001.

BHAT, C., ZHAO, H. (2002) The spatial analysis of activity stop generation. **Transportation Research Part B** 36 (2002) p. 557–575. DOI:10.1016/S0191-2615(01)00019-4

CAMARGO, E.; CAMARGO, C. G.; DRUCK, S.; CÂMARA, G. (2004). 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.

COSTA, A. S; PITOMBO, C. S.; SALGUEIRO, A. R. (2013). Estimação de escolha modal através da geoestatística. Anais do **XXVII Congresso Congresso do Ensino e Pesquisa e Transportes (ANPET)**, Belém, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Censo Demográfico Brasileiro 2010**. Disponível em < <http://www.ibge.gov.br>>. Acesso em 20 de agosto de 2012.

LOPES, S. B; RODRIGUES DA SILVA, A. N. (2005) Modelos de previsão de demanda por transportes empregando análise de dependência espacial. Anais do **XIX Congresso do Ensino e Pesquisa e Transportes (ANPET)**, Recife, 2005.

MIURA, H. (2010) A study of travel time prediction using universal kriging. **Sociedad de Estadística e Investigación Operativa. TOP** 18(1): p. 257–270. DOI: 10.1007/s11750-009-0103-6

OLEA, R. A. 1999. Geostatistics for Engineers and Earth Scientists. **Kluwer Academic Publishers**, Norwell, Massachusetts, 303p.

PÁEZ, A.; LÓPEZ, F. A.; RUIZ, M.; MORENCY, C. (2013) Development of an indicator to assess the spatial fit of discrete choice models. **Transportation Research Part B**, vol 56, p. 217-233. DOI:10.1016/j.trb.2013.08.009

- PEER, S.; KNOCKAERT, J.; KOSTER, P. TSENG, Y. Y.; VERHOEF, E.T. (2013) Door-to-door travel times in RP departure time choice models: An approximation method using GPS data. **Transportation Research Part B**, vol 58, p. 134-150. DOI: 10.1016/j.trb.2013.10.006.
- PITOMBO, C. S.; SOUSA, A. J. (2009) Aplicação de Conceitos Geoestatísticos para Análise de Geração de Viagens Urbanas. Anais do **XXIII Congresso de Pesquisa e Ensino em Transportes (ANPET)**, Vitória, 2009.
- PITOMBO, C. S.; SOUSA, A. J.; BIRKIN, M. (2010) Comparing different spatial data analysis to forecast trip generation. In: **World Conference on Transport Research Society, 2010, Lisboa**. Proceedings of the 12th WCTR. Lisboa, 2010, p. 1-23.
- PITOMBO, C.S.; A. R. SALGUEIRO; A.S.G. DA COSTA; C. A. ISLER (2015) A two-step method for mode choice estimation with socioeconomic and spatial information. **Spatial Statistics**, vol 11, p. 45-64.
- RIBEIRO, L. C. (2012) Os Estados e as Regiões Metropolitanas constitutivas do Observatório das Metrôpoles no Censo 2010. **Instituto Nacional De Ciência E Tecnologia CNPq/FAPERJ/CAPES**. Observatório das Metrôpoles: Rio de Janeiro, 2012. 75p.
- ROCHA, M. M.; YAMAMOTO, J. K.; FONTELES, H. R. N. (2009) Cokrigagem Ordinária versus Krigagem com Deriva Externa: Aplicações para a Avaliação do Nível Potenciométrico em um Aquífero Livre. São Paulo: **Revista do Instituto de Geociências – USP**. Série Científica, v.9, n.1, p. 73-84.
- TEIXEIRA, G. L. (2003) Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes utilizando Estatística Espacial. Dissertação de Mestrado, 156p. – **Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental**. Brasília, 2003.
- SINDICATO DAS EMPRESAS DE TRANSPORTE DE PASSAGEIROS DE SALVADOR - SETPS (1995) **Pesquisa Domiciliar Origem/Destino**. Salvador.
- WACKERNAGEL, H. (2010) Multivariate Geostatistics: An introduction with applications. **Springer**, Third Edition, 387p.
- YAMAMOTO, J. K.; LANDIM, P. M. B. (2013) Geoestatística: conceitos e aplicações. São Paulo: **Oficina de Textos**, 215p.