

# MODELAGEM GEOESTATÍSTICA APLICADA À AMOSTRA SISTEMÁTICA DESAGREGADA: ESTIMAÇÃO DE GERAÇÃO DE VIAGENS

**Samille Santos Rocha**

**Anabele Lindner**

**Cira Souza Pitombo**

Universidade de São Paulo

Escola de Engenharia de São Carlos/Departamento de Engenharia de Transportes

## RESUMO

O objetivo do presente trabalho é avaliar melhorias nas estimativas de modelagem geoestatística para geração de viagens por transporte coletivo (TCO), na hora pico, a partir da proposta de uma amostra sistemática desagregada. Os dados utilizados são oriundos da Pesquisa O/D, realizada na Região Metropolitana de Salvador (RMS), em 1995. O método possui duas etapas principais relacionadas à aplicação de Krigagem Ordinária (KO): (1) a partir dos centroides de macrozonas associados aos valores padronizados de produção e atração de viagens por TCO; (2) a partir de amostra sistemática desagregada proposta. Os mapas de Krigagem, para ambas as variáveis e amostragens obedeceram à distribuição espacial similar dos valores observados. Além disso, a validação dos valores estimados com a KO apresentou bons resultados para as duas amostragens (erro médio e variância de resíduos baixos) e melhor correlação para a amostragem sistemática artificial aqui proposta.

## ABSTRACT

The aim of this paper is to evaluate improvements in the geostatistical modeling estimations for trip generation by transit, proposing a systematic disaggregated sample. The data used is from the Origin-Destination Survey, carried out in 1995 in the Salvador Metropolitan Area. The method has two main steps related to the application of Ordinary Kriging (OK): (1) from the centroids of macrozones associated to the standardized values of production and attraction of trips by transit; (2) from the systematic disaggregated sample proposed. The kriging maps for both variables and samples had a spatial distribution similar to the observed values. Moreover, the accuracy of KO application showed good results for the two samples (low average of error and variance of errors) and better correlation for the systematic sampling here proposed.

## 1. INTRODUÇÃO

A Geoestatística emergiu a partir da década de 1950, quando se observou que a estatística básica não corresponderia às necessidades de análises de fenômenos espaciais. A Geoestatística estuda o comportamento espacial de variáveis e permite estimar seu valor em áreas geograficamente amostradas ou não amostradas. A técnica é usada para entendimento e modelagem da variabilidade espacial de atributos.

Esta técnica, recentemente utilizada no Planejamento de Transportes, pode auxiliar na estimação e interpolação da demanda por transportes. A dependência espacial das características relacionadas ao transporte urbano permite supor que a Geoestatística seja uma ferramenta conveniente de análise. Contudo, percebe-se que ainda é pouco aplicada na área de transportes. Tanto na literatura nacional quanto na literatura internacional, trabalhos envolvendo análise de demanda por viagens e Geoestatística são absolutamente raros.

Uma referência da Geoestatística aplicada à análise de demanda por transportes pode ser encontrada no estudo de Costa *et al.* (2013). O trabalho estima a escolha modal por meio de dados desagregados de Pesquisa Origem/Destino realizada na cidade de São Carlos (SP), através de Krigagem Ordinária (KO). Pitombo e Sousa (2009) realizaram um estudo que utilizou Krigagem Ordinária e uma base de dados de valores agregados de geração de viagens na Região Metropolitana de São Paulo (RMSP). Destacou-se que a produção e atração de viagens aumentavam de valor nas regiões da periferia para o centro. Pitombo *et al.* (2010) utilizaram técnicas de análise espacial de dados, a Geoestatística multivariada (Krigagem com Deriva Externa) e a Regressão Geograficamente Ponderada, para estimar a produção e atração

de viagens urbanas por modo e motivo de viagem, na RMSP. Já Rocha (2014) estimou geração de viagens por transporte coletivo na hora pico, através de Krigagem Ordinária e Krigagem com Deriva Externa, com dados da Região Metropolitana de Salvador.

Teixeira (2003) aplicou a Krigagem Ordinária para relacionar aspectos socioeconômicos à geração de viagens. O autor apresentou uma metodologia para definir o zoneamento por áreas homogêneas relacionadas ao planejamento de transportes, considerando aspectos econômicos e de viagens da população através dos setores censitários (IBGE). Como referência internacional, Miura (2009) descreve a abordagem da Krigagem Universal na estimação de tempos de viagem num espaço imaginário. Foi verificado que a Krigagem pode auxiliar na estimação de tempos de viagem entre dois pontos arbitrados, podendo auxiliar também na determinação de congestionamentos.

O objetivo do presente trabalho é avaliar melhorias nas estimativas de modelagem Geoestatística para geração de viagens por transporte coletivo, a partir de uma proposta de desagregação dos dados associados a Macrozonas de tráfego da Região Metropolitana de Salvador.

## 2. GEOESTATÍSTICA

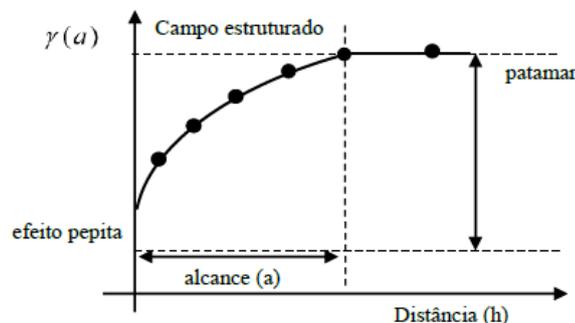
A revisão deste trabalho está focada em descrever e conceituar a ferramenta utilizada para estimar geração de viagens. Desta maneira, são analisados nos próximos itens três processos fundamentais da Geoestatística: o entendimento das variáveis regionalizadas, a modelagem dos variogramas experimentais e a Krigagem.

### 2.1 Variáveis Regionalizadas

As Variáveis Regionalizadas (VR) representam um conjunto de dados distribuídos espacialmente de forma homogênea dentro de certa distância e direção. As VR apresentam dois aspectos contraditórios: um aspecto aleatório e outro estrutural. O caráter aleatório corresponde à variação inesperada dos pontos no espaço e o estrutural mostra as correlações existentes entre os diversos pontos onde ocorre o fenômeno (Sousa e Muge, 1990).

### 2.2 Variograma

O variograma representa a variação de um fenômeno regionalizado no espaço por meio de uma descrição matemática da relação entre a variância de pares de observações e a distância ( $h$ ) que as separa (Câmara *et al.*, 2002). Os parâmetros importantes do variograma são o efeito pepita, patamar e a amplitude, apontados na Figura 1.



**Figura 1:** Representação geral de um variograma e de seus componentes

O efeito pepita  $\gamma(0)$  ou  $C_0$  representa a alta variabilidade da variável em distâncias pequenas (Soares, 2006). O patamar  $\gamma$  mede o quanto da variância é explicada pelo modelo. O valor do

patamar inclui o efeito pepita. A partir do patamar, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras torna-se aproximadamente constante (Camargo *et al.*, 2004). O alcance, ou amplitude, é a distância dentro da qual os dados são espacialmente correlacionados e corresponde à distância horizontal associada ao patamar (Camargo *et al.*, 2004).

A construção do variograma experimental é um ponto fundamental na modelagem geoestatística, pois sintetiza a informação relativa à estrutura espacial do fenômeno, influenciando fortemente os resultados da Krigagem. O variograma deve ser ajustado de acordo com modelos teóricos, que podem ser do tipo esférico, gaussiano, exponencial, dentre outros. Dos modelos com patamar, um dos mais utilizados na Geoestatística é o modelo esférico, podendo ser comparado com a distribuição normal na estatística clássica (Landim, 2010). Este modelo apresenta comportamento linear junto à origem e ocorre em função de dois parâmetros: o patamar e a amplitude.

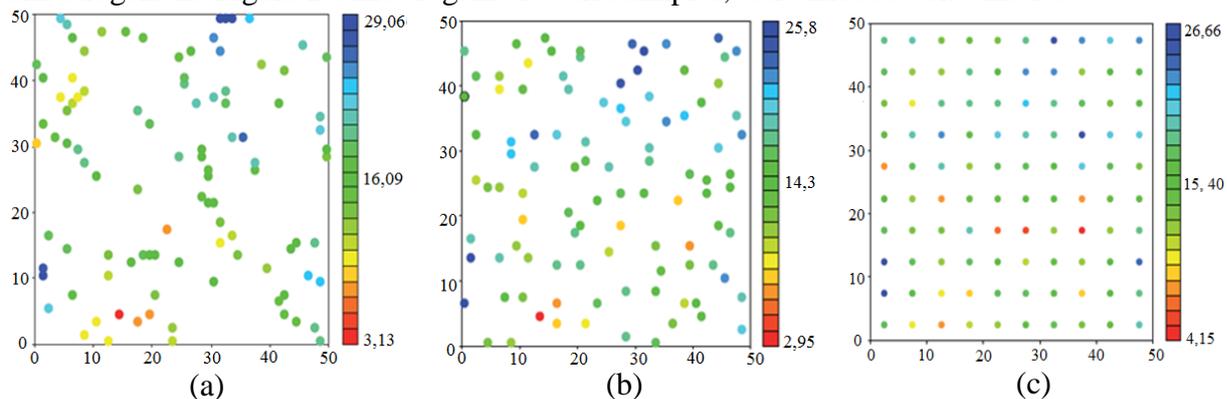
### 2.3 Krigagem

A utilização da Krigagem possibilita a inferência de um valor médio do atributo e sua posição espacial, a partir de uma quantidade  $n$  de amostras vizinhas e do variograma ajustado (Felgueiras, 1999).

Um dos tipos de Krigagem mais utilizado é a Krigagem Ordinária (KO). A KO é utilizada quando se admite as médias dos atributos constantes e desconhecidas. A estimativa em um ponto não amostrado resulta da combinação linear dos valores encontrados na vizinhança próxima (Yamamoto e Landim, 2013).

### 2.4 Amostragens geoestatísticas

Uma amostra é um conjunto de valores do fenômeno espacial que deve garantir uma boa representação da realidade. Cabe à Geoestatística e suas técnicas mensurar o grau de incertezas advindas das estimativas baseadas em amostras. São apresentados três tipos de amostragens na Figura 2: amostragem aleatória simples, estratificada e sistemática.



**Figura 2:** Amostragens aleatórias: (a) amostragem simples; (b) amostragem estratificada; (c) amostragem sistemática. Fonte: Yamamoto e Landim (2013).

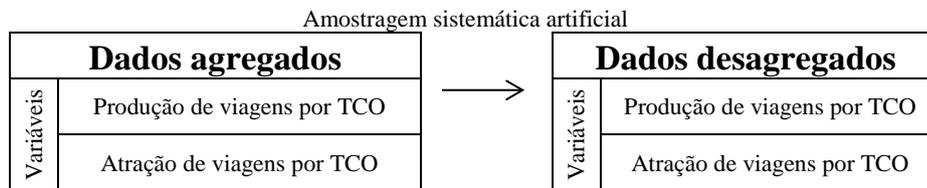
Na Geoestatística, a amostragem aleatória simples pode ser obtida através de um sorteio, onde a componente aleatória são as coordenadas geográficas selecionadas ao acaso e todos os elementos têm a mesma probabilidade de ocorrência (Yamamoto e Landim, 2013). Na amostragem aleatória estratificada, divide-se a população em subgrupos e coletam-se as amostras aleatoriamente dentro de cada estrato (Barbetta, 2005).

Na amostragem aleatória sistemática, os intervalos são regulares e definidos antecipadamente. Segundo Yamamoto e Landim (2013), na Geoestatística, a amostragem sistemática é a melhor dentre as três, porque recobre toda a área de estudo, evitando área com vazios ou com agrupamento de amostras.

Neste trabalho, gerou-se uma amostragem sistemática artificial, no intuito de melhorar as estimativas geoestatísticas. Os resultados obtidos através das coordenadas dos centroides das macrozonas da RMS foram comparados aos resultados obtidos através da amostra gerada artificialmente.

### 3. MÉTODO

O presente trabalho descreve a estimação de geração de viagens por Transporte Coletivo (TCO) em coordenadas cujos valores da variável de interesses são conhecidos, através da técnica de interpolação espacial - Krigagem Ordinária (KO). O trabalho utilizou amostras de pontos (dados da Pesquisa Origem/Destino da RMS de 1995). Além dos dados originais da Pesquisa O/D, é proposta uma amostra sistemática artificial para geração de dados desagregados. Resultados provenientes de modelagem geoestatística aplicada às duas amostras são apresentados. A Figura 3 esquematiza o método proposto.



**Figura 3:** Etapas do processo metodológico aplicado ao estudo

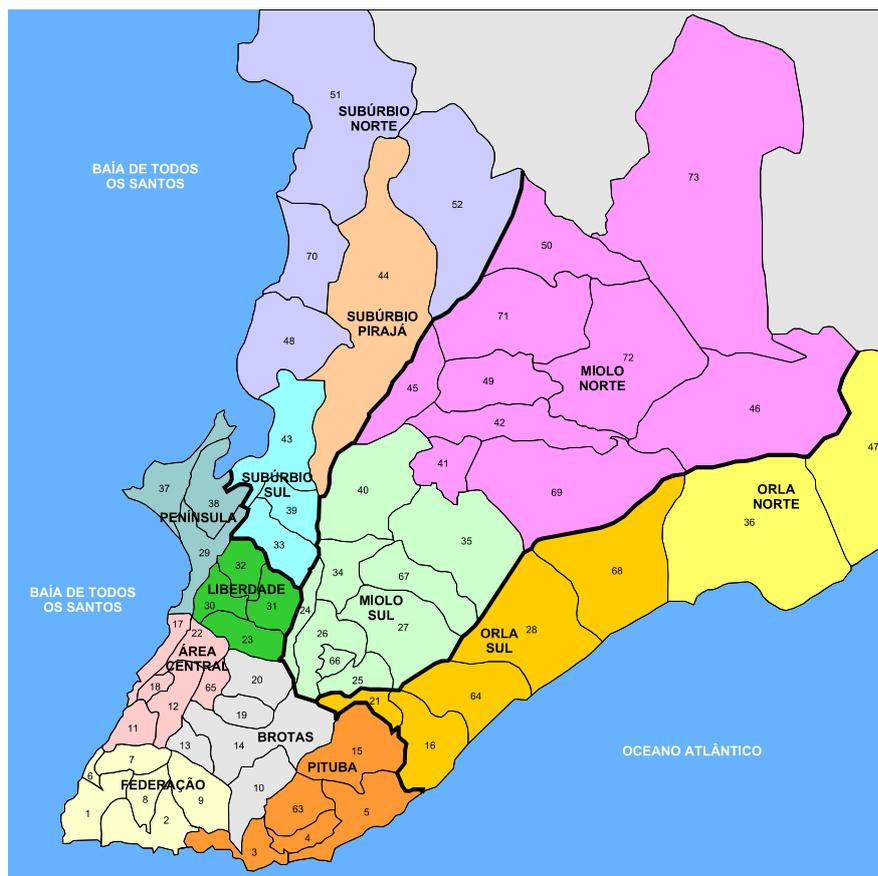
### 3.1 Dados

#### 3.1.1 Dados agregados

A área de estudo, a Região Metropolitana de Salvador (RMS) é composta por 13 municípios compondo uma área de 4.375 km<sup>2</sup> de extensão. A RMS possui uma população de cerca de 3.600.000 habitantes, e deste total, cerca de 2.700.000 residem na capital (IBGE, 2010). Salvador é a terceira capital mais populosa do país, e sua região metropolitana é a sétima em população (Ribeiro, 2012).

O banco de dados aqui utilizado é referente à Pesquisa Origem-Destino (OD) de 1995 e compreende o município de Salvador, além de 10 municípios. A grande maioria dos deslocamentos da RMS neste período concentrava-se em Salvador e o ônibus era o principal modo de transporte motorizado utilizado. Por isso a importância de analisar a geração de viagens pelo modo de transporte coletivo por ônibus (TCO) em Salvador (SETPS, 1995).

A unidade de área da pesquisa OD foi Macrozona de tráfego (MZ). Estas macrozonas (63 no total) foram classificadas em sub-regiões e em quatro grandes regiões: Área Urbana Consolidada (AUC), Orla, Miolo e Subúrbio (Figura 4). O Subúrbio Ferroviário e o Miolo de Salvador são as áreas mais populosas de Salvador e com menor renda. Portanto, existe uma provável tendência que a maior parte das viagens nestas áreas seja realizada por transporte coletivo. Em 1995, segundo a Pesquisa OD, a região com população economicamente mais rica em Salvador era a Orla, seguida da AUC.



**Figura 4:** Sub-regiões e macrozonas da RMS. Fonte: STEPS, 1995.

Para a análise geoestatística, as coordenadas geográficas dos centróides das macrozonas foram associadas às variáveis de produção de viagens por TCO e atração de viagens por TCO. Foram utilizados valores padronizados das duas variáveis em questão obtidos pela Equação 1.

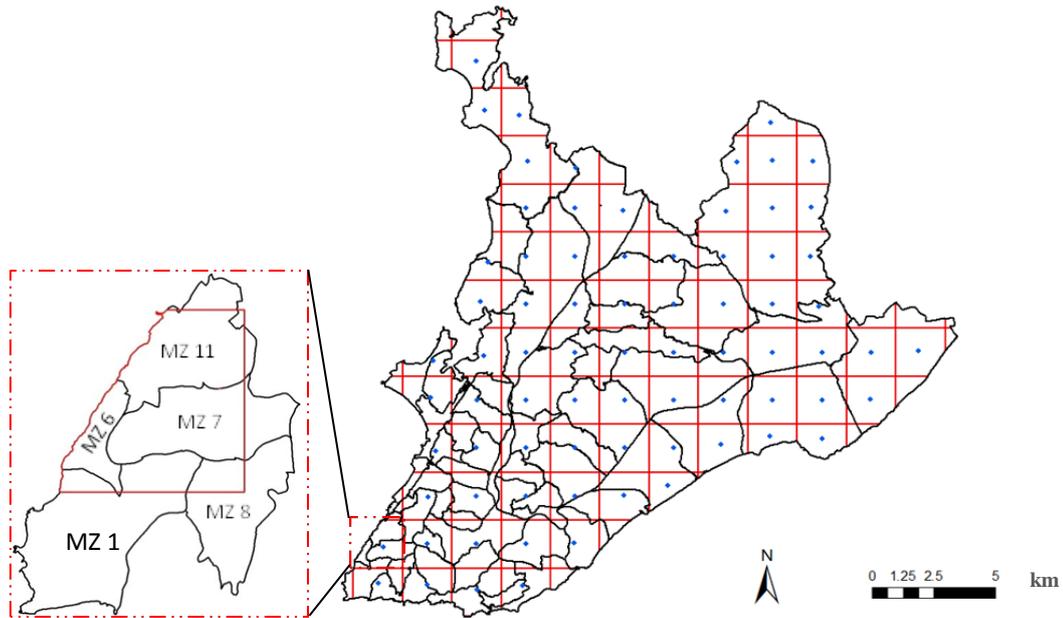
$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

em que z: valor padronizado da variável; x: valor absoluto da variável (produção/atração de viagens por TCO);  $\mu$ : média e  $\sigma$ : desvio padrão.

### 3.1.2 Dados desagregados

A fim de 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 e Atração de viagens por TCO associados às coordenadas dos centroides das células.

Após alguns testes, a determinação da malha se deu pela geração de células de 2000 x 2000 m. Em seguida, foram gerados centroides para cada célula obtida. As células quadradas obtidas e os seus centroides associados estão ilustrados na Figura 5.



**Figura 5:** Malha gerada para amostragem sistemática artificial

Notou-se que cada célula era composta por diferentes macrozonas. Tornou-se necessário, então, mensurar a ocupação, em porcentagem, das macrozonas em cada célula. A célula 2 (ou quadrícula 2), por exemplo, é composta pela MZ 1, MZ 6, MZ 7, MZ 8 e MZ 11, conforme ilustrado na Figura 5. A MZ 7, por exemplo, ocupa 32,8% da célula 2. A fim de determinar o valor das variáveis de geração de viagens para cada quadrícula, a partir dos valores conhecidos das macrozonas, atribuiu-se um peso para cada MZ ocupada nas quadrículas. O exemplo da MZ 7 na quadrícula 2, teve um peso de 0,328, a MZ 11 teve um peso de 0,335, e assim por diante (vide Tabela 1). Multiplicou-se o peso de cada MZ pelo seu respectivo valor de geração de viagem. O somatório dessas multiplicações resultou no valor da quadrícula e foi associado à respectiva coordenada geográfica do centróide.

**Tabela 1:** Exemplo de ponderação da área e determinação da variável para amostragem sistemática artificial (célula ou quadrícula 2)

Área da quadrícula (km <sup>2</sup> )	ID (MZ), Equivalente a 1 célula da malha	Área ocupada pela MZ em cada célula (km <sup>2</sup> )	Peso (%)	Produção de viagens por MZ	Valor da variável após compatibilização	Média ponderada por célula
2,975	1	0,424	0,142	0,18	0,03	0,18
	6	0,366	0,123	0,18	0,02	
	7	0,974	0,328	0,15	0,05	
	8	0,191	0,064	0,22	0,01	
	11	0,997	0,335	0,20	0,07	

### 3.2 Etapas para modelagem geostatística

#### 3.2.1 Análise estrutural e modelagem de variogramas

A primeira etapa da modelagem geostatística tratou primeiramente do entendimento do comportamento das variáveis, bem como visualização de possíveis padrões espaciais. Verificou-se a existência de anisotropia ou isotropia e a direção com maior variabilidade dos dados (direção principal), bem como seu ângulo ortogonal. Na análise estrutural (modelagem

da estrutura espacial da variável regionalizada) foram descritos os atributos que determinam a função do variograma: efeito pepita, alcance ou amplitude, patamar (*sill*) e o modelo teórico de ajuste do variograma. O resultado final desta etapa se deu pela construção de variogramas experimentais e ajuste de variogramas teóricos.

### 3.2.2 Verificação do modelo

Esta etapa trata da verificação do modelo por meio da validação cruzada (teste do ponto fictício ou *cross validation*). A validação cruzada compreende a mensuração dos erros. A validação cruzada considera os valores observados e estimados, possibilitando o conhecimento da qualidade do modelo do variograma (Pitombo e Sousa, 2009; Soares, 2006).

### 3.2.3 Interpolação Geoestatística

Conhecendo-se a qualidade dos parâmetros e do modelo de variograma adotado, a última etapa refere-se à interpolação espacial através da Krigagem. Foram escolhidos parâmetros como as coordenadas, o número de blocos, espaçamento entre blocos, discretização, número de amostras, tipo de busca das amostras e ângulos do elipsóide.

## 3.3 Softwares

Os procedimentos de modelagem geoestatística foram realizados no GeoMs (*Geostatistical Modelling Software*), desenvolvido pelo Instituto Superior Técnico/Universidade Técnica de Lisboa. Os mapas finais foram editados através do Arcgis 9.3.

## 4. RESULTADOS E DISCUSSÕES

### 4.1.1 Análise estrutural e modelagem de variogramas

A partir da análise do mapa de pontos visualizados no GeoMS, verificou-se a direção que apresenta maior variabilidade dos dados para a variável regionalizada de produção de viagens por TCO. Percebeu-se, por meio da distribuição espacial dos dados, as áreas que apresentam maiores valores da variável regionalizada, sendo estas localizadas na região do Subúrbio Ferroviário de Salvador. Esta área é composta pelos bairros mais populosos e com renda mais baixa da cidade. Já a tendência contrária é observada na região central.

Comparando os resultados da análise das variáveis de produção de viagens da amostra agregada com a desagregada, percebeu-se que a distribuição espacial das variáveis continuou similar aos pontos analisados por centroides das macrozonas, apresentando, porém, uma quantidade maior de dados e diminuição da distância entre estes.

Após visualização espacial e realização de alguns testes, optou-se por modelar os variogramas experimentais para a direção 50° NE e o seu respectivo ângulo ortogonal 320° SE para os dados de produção de viagens por TCO por centroides das macrozonas. Este ângulo representou melhor a continuidade espacial das amostras.

A variável de atração de viagens por TCO possui comportamento isotrópico, e mesmo adaptada à malha regular de 2000 por 2000 m, não demonstrou nenhuma continuidade ou variação aparente dos dados em alguma direção. O variograma foi modelado com o ângulo máximo de abertura (180° - omnidirecional).

Os dados desagregados não possuem uma direção principal de variabilidade. As variáveis foram também modeladas com o ângulo máximo de abertura (variograma omnidirecional). Verificou-se que a amostragem artificial gerou uma melhora significativa no ajustamento dos

variogramas experimentais, em relação à amostragem pelos centroides das macrozonas devido à maior quantidade de pontos localizados na área de estudo.

Os parâmetros de azimute  $\theta$ , tolerância angular ( $R$ ), *lag distance* ( $h$ ), *cut distance* foram utilizados para a modelagem dos variogramas experimentais das duas variáveis analisadas. No total, foram gerados três variogramas para a amostragem original e dois variogramas para a amostragem obtida pela malha regular (amostra sistemática).

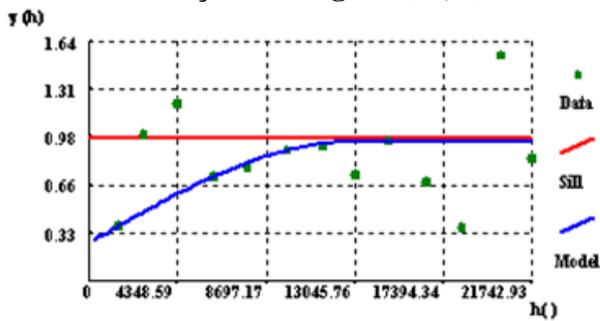
Na etapa de modelagem dos variogramas, foram determinados os seguintes parâmetros: efeito pepita, alcance ou amplitude, patamar e o modelo teórico de ajuste do variograma, apresentados na Tabela 2. Os variogramas ajustados estão ilustrados na Figura 6.

**Tabela 2:** Parâmetros utilizados na modelagem dos variogramas

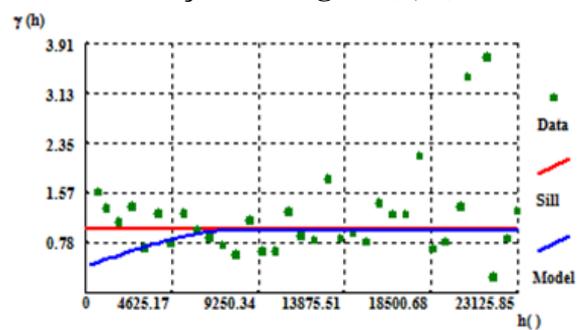
	Variáveis e direções	Efeito Pepita (C0)	(C1)	Alcance	Patamar (C0-C1)	Modelo Teórico
Amostra agregada	Produção de viagens (50°)	0,263	0,69	12.882,68	0,95	Esférico
	Produção de viagens (-40°)	0,263	0,69	9.832,59	0,95	Esférico
	Atração de viagens (omnidirecional)	0,488	0,49	7.400,27	0,98	Esférico
Amostra desagregada	Produção de viagens (omnidirecional)	0,182	0,668	10.845,03	0,85	Esférico
	Atração de viagens (omnidirecional)	0,068	0,139	11.336,76	0,207	Esférico

**Amostra agregada (a)**

**Produção de viagens (50; 0)**

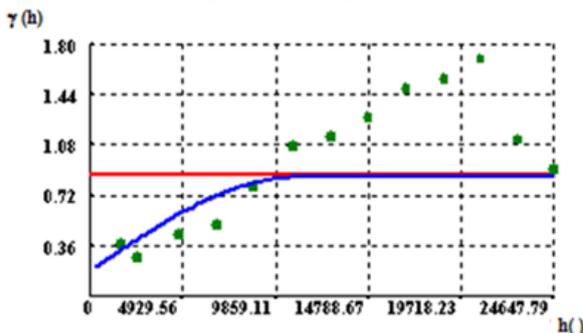


**Atração de viagens (0; 0)**

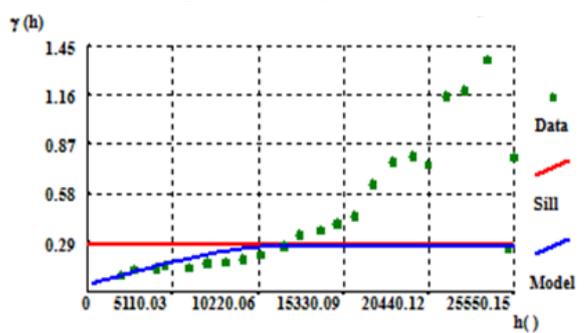


**Amostra desagregada (b)**

**Produção de viagens (0; 0)**



**Atração de viagens (0; 0)**



**Figura 6:** Modelagem dos semivariogramas

Observa-se claramente a melhoria na estruturação espacial da variável regionalizada Atração de viagens por TCO, através da proposta da amostragem sistemática artificial. O variograma correspondente à Atração de viagens para a amostra original não apresenta nenhuma estrutura espacial (efeito pepita puro - Figura 5 (a)). Enquanto que o variograma correspondente à variável Atração de viagens por TCO para amostra desagregada apresenta uma estrutura que se assemelha ao modelo esférico (Figura 5 (b)), com valor de efeito pepita reduzido (0,068).

#### 4.1.2 Verificação do modelo

A análise dos resultados da validação cruzada para as duas variáveis, em ambas as amostras de pontos, apresentou, no geral, bons resultados (Tabela 3). A média de erros e a variância dos resíduos apresentaram valores próximos de zero.

No entanto, o coeficiente de correlação entre valores observados e estimados das variáveis para as amostras agregadas apresentou baixos valores, devido ao modelo teórico do variograma ter baixa variância espacial e apresentar alto efeito pepita. Considerando a amostra desagregada, o coeficiente de correlação apresentou significativo aumento tanto para produção de viagens, quanto para atração de viagens por TCO. Os demais resultados são satisfatórios, com média de erros e variância de resíduos próximos a zero.

**Tabela 3:** Dados obtidos pela validação cruzada da estimação das variáveis

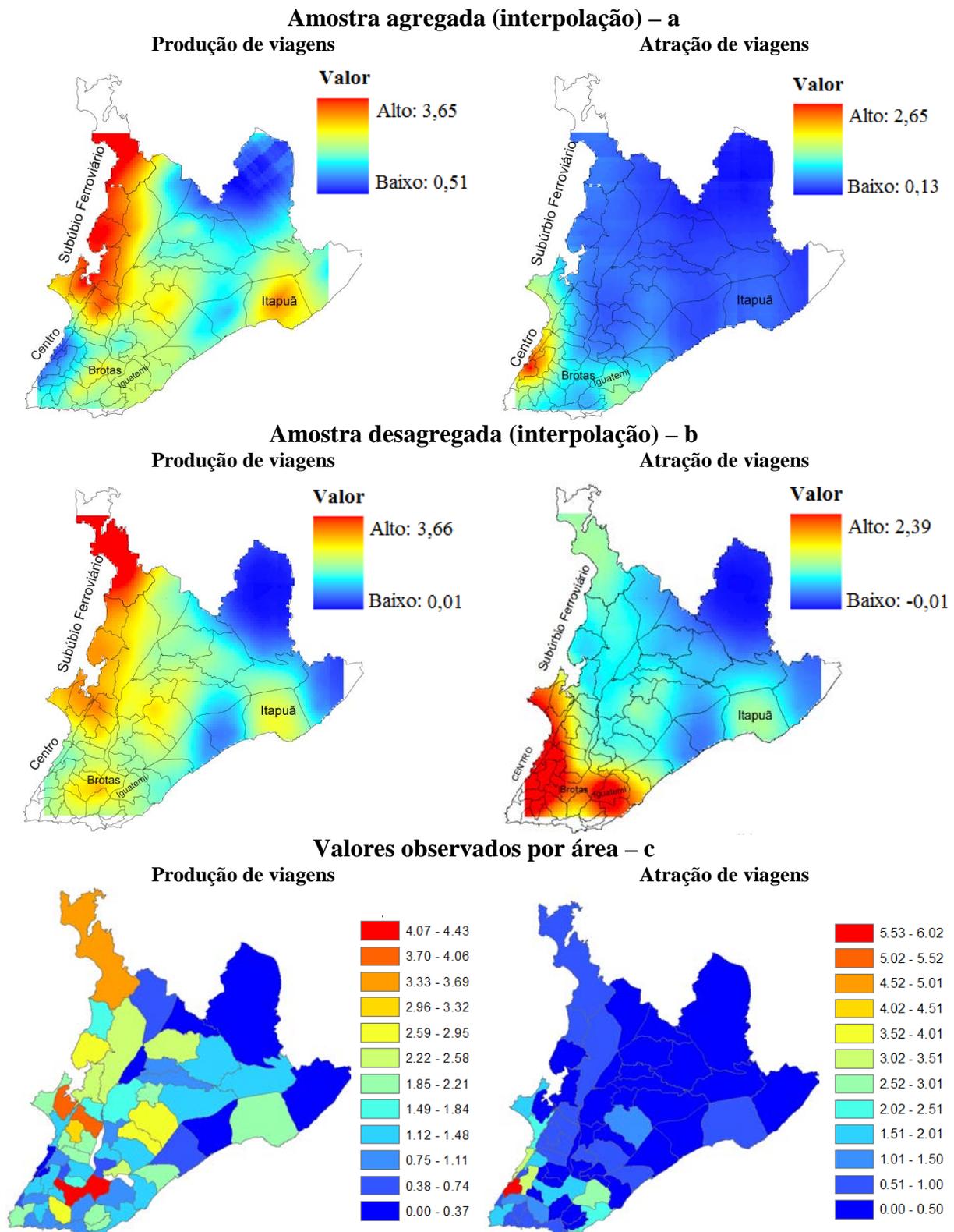
Variável	Coefficiente de correlação	Média dos erros	Variância dos erros
Produção de viagens por TCO	0,20	0,0005	0,476
Atração de viagens por TCO	0,40	-0,0030	0,747
Produção de viagens por TCO	0,84	-0,0041	0,386
Atração de viagens por TCO	0,77	-0,0208	0,409

Alguns autores defendem que a validação cruzada pode ser utilizada como método para a escolha do melhor modelo a ser adotado na estimação. Esta técnica retira o valor de todos os dados amostrados e obtém as estimativas feitas pela Krigagem, através dos parâmetros estabelecidos pelo ajuste do modelo teórico do variograma. Desta maneira, haverá para cada ponto de mesma coordenada o valor conhecido e o valor estimado. A partir destes resultados são calculados outros atributos estatísticos, como a média e variância dos erros, além do coeficiente de correlação de pearson

#### 4.1.3 Interpolação espacial

Os mapas de interpolação, obtidos através da Krigagem Ordinária, para o caso da amostra agregada e amostra sistemática artificial são mostrados na Figura 7 (a,b). A mesma figura ainda apresenta mapas temáticos de valores observados das variáveis produção e atração de viagens por TCO para facilitar comparação visual de valores obtidos através da interpolação e valores reais (Figura 7c).

Fazendo-se apenas uma análise visual entre as interpolações (Figura 7- a,b) e os valores observados por área (Figura 7 – c), observa-se que, para produção de viagens por TCO, a interpolação proveniente da amostra desagregada assemelha-se mais aos dados observados, enquanto que para Atração de viagens por TCO, a interpolação proveniente da amostra agregada aproximasse melhor dos valores observados.



**Figura 6:** Interpolação espacial para as variáveis Produção e Atração de viagens por TCO: (a) amostra agregada; (b) amostra desagregada; (c) valores observados por área.

## 5. CONCLUSÕES E RESSALVAS SOBRE PROCEDIMENTO METODOLÓGICO

A proposta deste trabalho alcançou resultados satisfatórios para a aplicação da Geoestatística na análise de demanda por transportes. Contudo, algumas dificuldades foram encontradas, sobretudo, na modelagem dos variogramas, principalmente para a atração de viagens por TCO. Verificou-se que a alta agregação de dados amostrais prejudica a modelagem dos variogramas, devido ao alto espaçamento entre amostras, além da quantidade reduzida de observações. Diante destas dificuldades e das várias tentativas de modelação dos variogramas em diferentes direções, foi proposta a desagregação dos dados analisados.

Na obtenção de dados para a amostra sistemática artificial, atentou-se para algumas considerações:

- A desagregação de dados desconsidera mudanças abruptas nos valores de variáveis socioeconômicas e de viagens, que em geral, não apresentam continuidade espacial equivalente a amostras de solos ou de água, por exemplo, comumente utilizadas em modelagem geoestatística;
- Adotou-se a premissa da continuidade espacial dos valores das variáveis através da utilização da técnica de geração de amostras desagregadas. No entanto, sabe-se que as amostras agregadas possuem descontinuidades mais evidentes;
- Os dados devem ser preferencialmente gerados através de técnicas de simulação e não necessariamente, média ponderada de valores das macrozonas.

A metodologia empregada neste estudo é uma tentativa inicial de melhorar modelagem Geoestatística a partir de dados muito agregados, sabendo que muitas vezes, apenas estes estão disponibilizados. Portanto, o método proposto para desagregação amostral tem o intuito de suavizar os efeitos da falta de continuidade espacial das variáveis em estudo, evidenciada no caso das amostras agregadas por macrozonas de tráfego.

A técnica aqui estudada demonstra potencial de estimação de demanda por viagens, além de obtenção de dados de viagens em várias coordenadas. A Tabela 4 compara os aspectos favoráveis da utilização da amostra original com a amostra sistemática artificial.

**Tabela 4:** Quadro comparativo entre pontos favoráveis da utilização das amostras.

Atributos	Amostra original	Amostra gerada
Confiabilidade	x	
Similaridade visual de mapas interpolados e observados	x	
Número de amostras		x
Estruturação espacial das variáveis		x
Poder preditivo		x
Qualidade da validação cruzada		x
Agregação		x
Distância entre amostras		x

A amostra original pode ser considerada como mais confiável por se tratar de dados pesquisados, obtidos em campo. A amostra sistemática artificial é produzida a partir da original, podendo acumular erros o processo.

Com a utilização da amostra sistemática artificial, o número de amostras aumentou, desagregou-se a amostra original e, conseqüentemente, diminuiu-se a distância entre amostras, facilitando no processo de estimação geoestatística.

Outro aspecto favorável da desagregação dos dados foi verificado pelo aprimoramento da estrutura espacial das variáveis regionalizadas. Considerando os resultados de validação cruzada, observou-se significativa melhora das estimativas e interpolações por Krigagem Ordinária com a amostragem sistemática artificial.

Tal método pode ser testado em várias cidades, sobretudo considerando a ausência de pesquisas domiciliares (amostras desagregadas) na maior parte das regiões brasileiras.

#### **AGRADECIMENTOS**

O presente trabalho foi realizado com apoio do CNPq e do Instituto de Transporte e Logística (ITL) da CNT.

#### **REFERÊNCIAS BIBLIOGRÁFICAS**

- Barbetta, P. A. (2005) *Estatística Aplicada às Ciências Sociais*. 5 ed. Florianópolis: UFSC, 340 p.
- Câmara, G.; Carvalho, M.S.; Cruz, O.G.; Correia, V. (2002) Análise Espacial de Áreas. *Análise Espacial de Dados Geográficos*, Fuks, S.D.; Carvalho, M.S.; Câmara, G.; Druck, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.V.N (eds.). Divisão de Processamento de Imagens – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- Camargo, E.; Camargo, C. G.; Druck, S.; Câmara, G. (2004) Análise Espacial de Superfícies. *Análise Espacial de Dados Geográficos*. Druck, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.M.V. (eds.). Brasília, EMBRAPA, 2004. p.46-82.
- Costa, A. S. G.; Pitombo, C. S.; Salgueiro, A. R. (2013) Proposta de um Método para Estimção de Escolha Modal através da Geoestatística. *Anais do XXVII Congresso de Pesquisa e Ensino em Transportes*, ANPET, Belém.
- Felgueiras, C. A. (1999) Modelagem Ambiental com Tratamento de Incertezas em Sistemas de Informação Geográfica: O Paradigma Geoestatístico por Indicação. Tese (Doutorado). São José dos Campos: INPE.
- IBGE (2010) Censo Demográfico Brasileiro 2010. Disponível em: < [http://downloads.ibge.gov.br/downloads\\_top.php](http://downloads.ibge.gov.br/downloads_top.php)>. Acesso em 04 de Novembro 2013.
- Landim, P.M.B. (2010) Introdução à análise estatística de dados geológicos multivariados. DGA, IGCE, UNESP/Rio Claro, Texto Didático 15, 229 p. Disponível em: < <http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/MULTIVAR.pdf> > Acesso em 12 de Setembro de 2014.
- Miura, H. (2009) A Study of Travel Time Prediction using Universal Kriging. *Sociedad de Estadística e Investigación Operativa*, p. 257-270.
- Pitombo, C. S.; Sousa A. J.; Birkin, M. (2010) Comparing Different Spatial Data Analysis to Forecast Trip Generation. *12th World Conference on Transport Research Society, WCTR*, p. 25
- Pitombo, C. S.; Sousa, A. J. (2009) Aplicação de Conceitos Geoestatísticos para Análise de Geração de Viagens Urbanas. Em: Congresso de Pesquisa e Ensino em Transportes. *Anais do XXIII Congresso de Pesquisa e Ensino em Transportes*, ANPET, Vitória.
- 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.
- Rocha, S. S. (2014) Análise de Geração de Viagens Urbanas por Transporte Coletivo através de Técnicas de Geoestatística. Dissertação (Mestrado). Universidade Federal da Bahia. Mestrado em Engenharia Ambiental Urbana, Salvador.
- SETPS (1995) Pesquisa Domiciliar Origem/Destino. Salvador.
- Soares, A. (2006) *Geoestatística para as Ciências da Terra e do Ambiente*. 2 ed. Lisboa: IST Press, 214p.
- Sousa, A. J.; Muge, F. (1990) *Elementos de Geoestatística*. Lisboa: Laboratório de Mineralogia e Planejamento Mineiro.
- Teixeira, G. L. (2003) Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes Utilizando Estatística Espacial. 156 f. Dissertação (Mestrado). Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental, Brasília.
- Yamamoto, J. K.; Landim, P. M. B. (2013) *Geoestatística: Conceitos e Aplicações*. São Paulo: Oficina de Textos. 215p.