UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
MESTRADO EM ENGENHARIA AMBIENTAL URBANA

SAMILLE SANTOS ROCHA

ANÁLISE DE GERAÇÃO DE VIAGENS URBANAS POR TRANSPORTE COLETIVO ATRAVÉS DE TÉCNICAS DE GEOESTATÍSTICA

Salvador
2014

SAMILLE SANTOS ROCHA

ANÁLISE DE GERAÇÃO DE VIAGENS URBANAS POR TRANSPORTE COLETIVO ATRAVÉS DE TÉCNICAS DE GEOESTATÍSTICA

Dissertação apresentada à Escola Politécnica da Universidade Federal da Bahia, como requisito para obtenção do título de Mestre em Engenharia Ambiental Urbana.

Orientadora: Prof. Drª Cira Souza Pitombo
Co-orientadora: Prof. Drª Ana Rita Salgueiro (Universidade de Aveiro)

Salvador
2014

R672 Rocha, Samille Santos

Análise de geração de viagens urbanas por transporte coletivo através de técnicas de geoestatística/ Samille Santos Rocha. – Salvador 2014.

154 f. : il. color.

Orientadora: Profª Drª. Cira Souza Pitombo
Co-orientadora: Profª. Drª Ana Rita Salgueiro.

Dissertação (mestrado) – Universidade Federal da Bahia. Escola Politécnica, 2014.

1. Transporte coletivo. 2. Geoestatística. 3. Viagens. I. Pitombo, Cira Souza. II. Salgueiro, Ana Rita. III. Universidade Federal da Bahia. IV. Título.

CDD: 388.4

UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA

Mestrado em Engenharia Ambiental Urbana

ATA DA SESSÃO PÚBLICA
DE
DEFESA DE DISSERTAÇÃO DE MESTRADO

Aos 04 dias do mês de junho de 2014, às 16:00 horas, na Escola Politécnica da Universidade Federal da Bahia, presente a Comissão Julgadora, integrada pelos Professores CIRA SOUZA PITOMBO (Orientadora da Candidata), Ilcel Marília Dantas Pinto (Examinador Interno), Juan Pedro Moreno Delgado (Examinador Interno) e Antonio Nelson Rodrigues da Silva (Examinador Externo) iniciou-se a Defesa de Dissertação de Mestrado da candidata SAMILLE SANTOS ROCHA, intitulada **"ANÁLISE DE GERAÇÃO DE VIAGENS URBANAS POR TRANSPORTE COLETIVO ATRAVÉS DE TÉCNICA DE GEOESTATÍSCA".**

Concluída a apresentação, argüição e defesa oral da dissertação, conforme disposto no Regimento Interno do MEAU, procedeu-se ao julgamento na forma regulamentar, tendo a Comissão Julgadora considerado o candidato:

Aprovado .....(Avaliação da Profª. Cira Souza Pitombo);
APROVADA .....(Avaliação do Profª Ilce Marília Dantas Pinto);
Aprovado .....(Avaliação do Profº Juan Pedro Moreno Delgado);
APROVADA ..... .(Avaliação da Profº Antônio Nélson Rodrigues da Silva).

Encerrada a sessão, foi lavrada a presente ata que vai acompanhada dos pareceres escritos de cada examinador e assinada pela Comissão Julgadora.

Salvador, 04 de junho de 2014

A Comissão Julgadora:

Cira Pitombo
Profª. Drª. Cira Souza Pitombo *(Orientadora)*

Profª. Drª. Ilce Marília Dantas Pinto (*Examinador Interno*)

Profº. Drº. Juan Pedro Moreno Delgado (*Examinadora Interno*)

Profº. Drº. Antônio Nélson Rodrigues da Silva (*Examinador Externo*)

AGRADECIMENTOS

A Deus, pelo dom da vida, e por todas as conquistas realizadas até aqui.

À minha família, pelo constante incentivo e apoio, especialmente à minha mãe Arlete, minha irmã Larissa e meu avô José Rocha (*in memorian*) pelo amor e dedicação incondicionais.

À Prof. Cira, por este desafio, pela confiança, amizade, dedicação e paciência na realização deste trabalho. Também, pelo exemplo de profissionalismo, e pelas brilhantes ideias, que só enriqueceram nosso trabalho.

À Prof. Ana Rita, co-orientadora deste trabalho, por sempre poder contar com sua ajuda e pelas imprescindíveis contribuições (mesmo do outro lado do Atlântico).

Aos professores do MEAU, Ilce e Juan, por estarem sempre presentes ao longo desses dois anos de pesquisa, e, principalmente, pelas indispensáveis dicas no andamento do trabalho.

Ao Profº Clímaco, Instituto de Geociências da UFBA, pela ajuda na análise da dinâmica socioeconômica e espacial de Salvador nos períodos analisados.

Às colegas do MEAU, Ana e Ingrid, pela amizade e cumplicidade. A Daiane pela disponibilidade e ajuda para obtenção de dados importantes. A Angélica pela disponibilidade em ajudar sempre nos momentos de dúvidas.

Às colegas de apartamento, Adryelle e Thaíse, pela amizade e companheirismo, desde a época da Residência Universitária. A Oscar e a Nala, pelo amor, carinho e companhia.

Às amigas Fernanda, Mayara e Renata, pela amizade de uma vida inteira.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), pelo auxílio financeiro.

ROCHA, Samille Santos. Análise de Geração de Viagens Urbanas por Transporte Coletivo Através de Técnicas e Geoestatística. 154f. il. 2014. Dissertação (Mestrado) – Escola Politécnica, Universidade Federal da Bahia, Salvador, 2014.

RESUMO

O objetivo principal deste trabalho é estimar geração de viagens por transporte coletivo (TCO) em coordenadas conhecidas e desconhecidas. A ferramenta de análise espacial de dados utilizada é a Geoestatística (Krigagem Ordinária e Krigagem com Deriva externa). Os dados utilizados foram: a Pesquisa Origem-Destino (OD/1995), realizada na Região Metropolitana de Salvador (RMS) e dados de população e renda do último censo demográfico (IBGE, 2010), compatibilizados com as Macrozonas de tráfego (MZs). Foram utilizados dois diferentes tipos de amostragem para cada banco de dados (1995 e 2010): dados agregados, representados pelos centróides das MZs e dados obtidos através de uma malha regular, também associados aos centróides das células da malha. Assim, foram estimadas, em coordenadas conhecidas e desconhecidas, a geração de viagens por TCO, População e Empregos. Foram obtidos mapas de interpolação a partir das diferentes análises para estas variáveis. Observou-se que tanto população quanto a produção de viagens por TCO apresentam um padrão espacial crescente em direção ao Subúrbio Ferroviário da cidade. Já a atração de viagens e empregos não possuem continuidade espacial aparente, apresentando valores mais elevados apenas na região central de Salvador. Para a amostragem por centróides das MZs a validação dos resultados iniciais foi satisfatória, apresentando baixos valores de média de erros e variância dos erros. No entanto, houve baixa correlação entre os dados observados e estimados para ambas variáveis. Já para os dados analisados através da amostragem sistemática artificial, a correlação entre estes dados apresentou melhores resultados que a amostragem agregada por MZs. A análise através da krigagem com deriva externa gerou ainda melhores resultados. Além disso, foi possível prever variáveis de produção de viagens por TCO através de dados censitários populacionais. Este trabalho é uma etapa incipiente que busca melhorias de estimativas provenientes de modelagem geoestatística aplicada à análise de demanda por transportes.

Palavras-chave: geração de viagens por TCO, geoestatística, malha regular.

ROCHA, Samille Santos. Analysis of urban trip generation by public transportation through Geostatistical techniques. 154f. il. 2014. Dissertation (Master) – Polytechnic School, Federal University of Bahia, Salvador, 2014.

ABSTRACT

The main objective of this work is to estimate trips generation by public transportation in known and unknown coordinates. The spatial data analysis tool used is Geostatistics (Ordinary Kriging and Kriging with external drift). The data used were: the Origin-Destination Survey (OD/1995) realized in the Salvador Metropolitan Area (SMA) and population and income data from the last census (IBGE, 2010), matched with traffic Macrozones (ZMs). Two different types of sampling were used for each database (1995 and 2010): aggregated data, represented by the centroid of ZMs and data obtained through a regular grid, also associated with the centroid of the cells. So, the trip generation by public transportation, population, and employment were estimated to known and unknown coordinates. Interpolated maps were obtained from the different analyses for these variables. It was observed that both population and trip production by public transportation present an increased spatial pattern towards the railroad suburb of the city. Already the trip attraction and employment do not have an apparent spatial continuity, with higher values only in the central region of Salvador. For the sample of centroid of ZMs the initial validation of results was satisfactory, presenting low values of average of errors and variance of errors. However, there was a low correlation between the observed and estimated data for both variables. As for the data analyzed through artificial systematic sampling, the correlation between these data demonstrated better results than the sample of aggregate data by ZMs. The analysis through kriging with external drift generated still better results. Moreover, it was possible to predict variables of trip production by public transportation through population census data. This work is an incipient stage that seeks improvements of estimates by geostatistical modeling applied for transportation demand analysis.

Keywords: trip generation by public transportation, geostatistics, regular grid.

LISTAS DE FIGURAS

Figura 1.1: Estrutura da dissertação 17

Figura 3.1: Parâmetros do variograma.....41

Figura 3.2: Modelos usuais de variogramas.....42

Figura 3.3: Amostragens aleatórias: (a) amostragem simples. (b) amostragem estratificada e (c) amostragem sistemática 48

Figura 4.1: Região Metropolitana de Salvador. 54

Figura 4.2: População por MZ de tráfego (1995).....55

Figura 4.3: Renda média do responsável por domicílio nas MZs de tráfego (2000) .56

Figura 4.4: Percentual de domicílios que possuem 1 ou mais veículos (1995).....57

Figura 4.5: População por MZ de tráfego (2010).....57

Figura 4.6: Renda média do responsável por domicílio (em salários mínimos) nas MZs de tráfego (2010)......58

Figura 4.7: Sub-regiões e Macrozonas 60

Figura 4.8: Exemplo da obtenção da malha regular e amostra sistemática artificial da área de estudo 65

Figura 5.1: Mapa de pontos da variável população (1995) 69

Figura 5.2: Mapa de pontos da variável produção de viagens por TCO (1995) 70

Figura 5.3: Mapa de pontos da variável empregos (1995). 71

Figura 5.4: Mapa de pontos da variável atração de viagens por TCO (1995). 72

Figura 5.5: Mapa de pontos da variável população (1995) - amostra sistemática artificial 73

Figura 5.6: Mapa de pontos da variável produção de viagens por TCO (1995) - amostra sistemática artificial.....74

Figura 5.7: Mapa de pontos da variável empregos (1995) - malha regular 75

Figura 5.8: Mapa de pontos da variável atração de viagens por TCO (1995) - malha regular 75

Figura 5.9: Mapa de pontos da variável população (2010). 77

Figura 5.10: Mapa de pontos da variável produção de viagens (2010).....77

Figura 5.11: Mapa de pontos da variável população (2010) - amostra sistemática artificial. 78

Figura 5.12: Mapa de pontos da variável produção de viagens por TCO (2010) - amostra sistemática artificial..... 79

Figura 5.13: Modelagem dos variogramas. (A e B) população; (C e D) produção de viagens por TCO; (E) empregos; (F) atração de viagens por TCO. 84

Figura 5.14: Modelagem dos variogramas. (A) população; (B) produção de viagens; (C) empregos; (D) atração de viagens (amostra sistemática artificial). 86

Figura 5.15: Modelagem dos variogramas das variáveis população (A e C) e produção de viagens (B e D) 2010 (amostragem por centróide das MZs e por malha regular).....88

Figura 6.1: População estimada por MZ (1995) - amostra por centróide da MZ - KO 97

Figura 6.2: Produção de viagens por TCO estimada por MZ (1995) - amostra por centróide da MZ - KO 98

Figura 6.3: Produção de viagens por TCO estimada por MZ (1995) - amostra por centróide das MZs - KDE 99

Figura 6.4: Empregos estimados por MZ (1995) - amostragem por centróide das MZs - KO.....102

Figura 6.5: Atração de viagens por MZ - amostragem por centróide das MZs - KO.. 103

Figura 6.6: Atração de viagens (1995) - amostragem por centróide das MZs - KDE ... .104

Figura 6.7: Mapa de interpolação população (1995) - amostragem sistemática artificial - KO.....107

Figura 6.8: Mapa de interpolação da variável produção de viagens por TCO (1995) - amostragem sistemática artificial - KO 108

Figura 6.9: Mapa de interpolação da variável produção de viagens por TCO (1995) - amostragem sistemática artificial - KDE.....109

Figura 6.10: Mapa de interpolação da variável empregos (1995) - amostragem sistemática artificial - KO 112

Figura 6.11: Mapa de interpolação da variável atração de viagens por TCO (1995) - amostragem sistemática artificial - KO 113

Figura 6.12: Mapa de interpolação da variável atração de viagens por TCO (1995) - amostragem sistemática artificial - KDE 114

Figura 6.13: Mapa de interpolação da variável população (2010) - KO 117

Figura 6.14: Mapa de interpolação da variável produção de viagens por TCO (2010) - KO 118

Figura 6.15: Mapa de interpolação da variável produção de viagens por TCO (2010) - KDE 119

Figura 6.16: Mapa de interpolação da variável população (2010) - amostragem sistemática artificial - KO 122

Figura 6.17: Mapa de interpolação da variável produção de viagens por TCO (2010) amostragem sistemática artificial – KO 123

Figura 6.18: Mapa de interpolação da variável produção de viagens por TCO (2010) amostragem sistemática artificial – KDE 124

Figura 6.19: Mapas de interpolação krigados por KO: amostras por centróide x amostras por malha regular ..... 125 e 126

Figura 6.20: Mapas de interpolação krigados por KDE: amostras por centróide x amostras por malha regular 127

Figura 7.1: Resumo dos dados e resultados 130

LISTA DE TABELAS

Tabela 2.1: Variação inter-zonal de produção de viagens através de dois diferentes sistema de zoneamento 31

Tabela 2.2: Alguns trabalhos que abordam análise espacial e transportes 35

Tabela 3.1: Alguns trabalhos que envolvem geoestatística e transporte 52

Tabela 4.1: Exemplos de dados socioeconômicos agregados por MZ 59

Tabela 4.2: Exemplo do banco de dados de entrada.....61

Tabela 4.3: Resultados do modelo de regressão linear 62

Tabela 4.4: Evolução do SM no período de 1995 à 2010 62

Tabela 4.5: Exemplo do ajuste da renda para 2010.....63

Tabela 5.1: Estatística descritiva das variáveis população e produção de viagens por TCO (1995) 70

Tabela 5.2: Estatística descritiva das variáveis empregos e atração de viagens por TCO (1995) 72

Tabela 5.3: Estatística descritiva das variáveis população e produção de viagens por TCO 1995 (malha regular).....74

Tabela 5.4: Estatística descritiva das variáveis empregos e atração de viagens por TCO 1995 (malha regular).....76

Tabela 5.5: Estatística descritiva das variáveis população e produção de viagens por TCO (2010) 76

Tabela 5.6: Estatística descritiva das variáveis população e produção de viagens por TCO 2010 (malha regular).....79

Tabela 5.7: Parâmetros utilizados para a elaboração dos variogramas das variáveis empregos, população, produção e atração de viagens por TCO (1995).....82

Tabela 5.8 Parâmetros utilizados para a elaboração dos variogramas das variáveis população e produção de viagens por TCO (2010).....82

Tabela 5.9: Parâmetros utilizados na modelagem dos variogramas das variáveis população, empregos e geração de viagens por TCO (1995).....83

Tabela 5.10: Parâmetros utilizados na modelagem dos variogramas das variáveis população e produção de viagens por TCO (1995) - malha regular.....85

Tabela 5.11: Parâmetros utilizados na modelagem dos variogramas (dados compatibilizados, IBGE 2010) amostra por centróides das MZs. 87

Tabela 5.12: Parâmetros utilizados nas Krigagens 89

Tabela 6.1: Exemplo de dados obtidos na etapa da validação cruzada.....92

Tabela 6.2: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (1995) 93

Tabela 6.3: População observada e estimada 95

Tabela 6.4: Dados obtidos através da Validação Cruzada da estimação das variáveis empregos e atração de viagens (1995).....100

Tabela 6.5: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (1995) – amostragem sistemática artificial 105

Tabela 6.6: Dados obtidos através da Validação Cruzada da estimação das variáveis empregos e atração de viagens (1995) – amostragem sistemática artificial.....110

Tabela 6.7: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (2010) 115

Tabela 6.8: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (2010) – amostragem sistemática artificial 120

Tabela 7.1: Quadro resumo da Validação Cruzada 131

SUMÁRIO

CAPÍTULO 1: INTRODUÇÃO.....14
1.1 CONSIDERAÇÕES INICIAIS.....14
1.1. OBJETIVOS.....15
1.1.1 OBJETIVO GERAL.....15
1.1.2 OBJETIVOS ESPECÍFICOS.....16
1.2. JUSTIFICATIVA.....16
1.3. ESTRUTURA DO TRABALHO.....17

CAPÍTULO 2: MODELOS DE GERAÇÃO DE VIAGENS.....20
2.1. MODELOS TRADICIONAIS: CONCEITOS.....20
2.2. MODELO QUATRO ETAPAS.....22
2.3 MODELOS DE GERAÇÃO DE VIAGENS: CONSIDERAÇÕES TEÓRICAS.....26
2.3.1. Principais modelos de geração de viagens.....29
2.4 ANÁLISE ESPACIAL E DEMANDA POR TRANSPORTES: REVISÃO DA LITERATURA.....32

CAPÍTULO 3: GEOESTATÍSTICA.....36
3.1. CONTEXTUALIZAÇÃO.....36
3.2 VARIÁVEL REGIONALIZADA (VR).....38
3.3 VARIOGRAMA.....39
3.4 KRIGAGEM.....43
3.4.1 KRIGAGEM ORDINÁRIA.....45
3.4.2 KRIGAGEM COM DERIVA EXTERNA.....46
3.5 MÉTODOS DE AMOSTRAGEM.....48
3.6. TRANSPORTES E GEOESTATÍSTICA: TRABALHOS ANTERIORES.....50

CAPÍTULO 4: CARACTERIZAÇÃO DA ÁREA DE ESTUDO E DADOS.....53
4.1 ÁREA DE ESTUDO.....53
4.2 TRATAMENTO DOS DADOS E OBTENÇÃO DAS AMOSTRAS.....58
4.2.1 Pesquisa OD 1995: Dados agregados (Etapa 1).....58
4.2.2 Tratamento dos dados do censo de 2010: compatibilização de áreas – Etapa 2.....61
4.2.3 Obtenção da amostragem sistemática artificial – Etapa 3 e Etapa 4.....64

CAPÍTULO 5: APLICAÇÃO DAS TÉCNICAS DE GEOESTATÍSTICA 67

5.1 ETAPAS PARA MODELAGEM GEOESTATÍSTICA 67

5.2 ANÁLISE ESPACIAL EXPLORATÓRIA – GEODATA 68

5.2.1 População e Produção de viagens por TCO 1995 (amostragem original por centróides das MZs).....68

5.2.2 Empregos e Atração de viagens por TCO 1995 (amostragem original por centróides das MZs).....71

5.2.3 População e Produção de viagens por TCO 1995 (amostragem sistemática artificial).....73

5.2.4 Empregos e Atração de viagens por TCO 1995 (amostragem sistemática artificial).....74

5.2.5 População e Produção de viagens por TCO 2010 (amostra por MZ).....76

5.2.6 População e Produção de viagens por TCO 2010 (amostra sistemática artificial).....78

5.3 DETERMINAÇÃO DE PARÂMETROS E MODELAGEM DOS VARIOGRAMAS EXPERIMENTAIS 80

5.4 MODELAGEM DOS VARIOGRAMAS.....83

5.4.1 Geração de Viagens - Pesquisa OD 1995 (amostragem por centróides das MZs).....83

5.5 ESCOLHA DOS CRITÉRIOS PARA A KRIGAGEM - GeoKrig e GeoKed 88

CAPÍTULO 6: RESULTADOS E DISCUSSÕES 91

6.1 VALIDAÇÃO CRUZADA.....91

6.1.1 População e Produção de viagens por TCO 1995 (amostragem original por centróide das MZs).....93

6.1.2 Interpolação de mapas: População e Produção de viagens (1995) amostragem original por centróides das MZs 94

6.1.3 Empregos e Atração de viagens por TCO 1995 (amostragem original por centróides das MZs).....100

6.1.4 Interpolação de mapas: Empregos e Atração de viagens (1995) amostragem original por centróides das MZs 100

6.1.5 População e Produção de viagens por TCO 1995 (amostragem sistemática artificial).....105

6.1.6 Interpolação de mapas de krigagem: População e Produção de viagens (1995) amostragem sistemática artificial 106

6.1.7 Empregos e Atração de viagens por TCO 1995 (amostragem sistemática artificial).....110

6.1.8 Interpolação de mapas de krigagem: Empregos e Atração de viagens (1995) amostragem sistemática artificial 110

6.1.9 População e Produção de viagens por TCO 2010 (amostra por MZs) 115

6.1.10 Interpolação de mapas: População e Produção de viagens (2010) amostragem original por centróides das MZs 116

6.1.11 População e Produção de viagens por TCO 2010 (amostra sistemática artificial) 120

6.1.12 Interpolação de mapas: População e Produção de viagens (2010) amostragem sistemática artificial 120

CAPÍTULO 7: CONCLUSÕES E RECOMENDAÇÕES 129

7.1 SÍNTESE DO TRABALHO 129

7.2 PRINCIPAIS CONCLUSÕES E CONTRIBUIÇÕES 132

7.3 RECOMENDAÇÕES PARA TRABALHOS FUTUROS 133

REFERÊNCIAS 135

Apêndice A 142

Apêndice B 143

Apêndice C 144

Apêndice D 145

Apêndice E 146

Apêndice F 147

Apêndice G 148

Apêndice H 149

Apêndice I 150

Apêndice J 151

Apêndice K 152

Apêndice L. 153

Apêndice M 154

CAPÍTULO 1: INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

Para um planejamento de transporte urbano eficaz é importante definir uma boa oferta e estabelecer adequados padrões de utilização do transporte. Nesse sentido, é imprescindível quantificar a demanda por transportes, caracterizando a sua distribuição na área de estudo. Esta avaliação pode ser feita utilizando modelos de planejamento (CAMPOS, 2007).

A partir da década de 1970 começam a surgir as primeiras técnicas de modelagem de transporte. A modelagem tem o papel de apoiar o planejamento de transporte e a tomada de decisão, através de técnicas que permitem quantificar e estimar a demanda por transportes, estabelecendo padrões de viagens por zona de tráfego, por exemplo (ORTÚZAR e WILLUMSEN, 2011). Um dos modelos mais tradicionais no planejamento de transporte é o **Quatro Etapas**. Conhecido também como modelo sequencial, este modelo é um dos mais utilizados para caracterizar o planejamento de transportes, e tem como objetivo prever a demanda futura de transportes a fim de apresentar melhores dimensionamentos da oferta, utilizando taxas médias de viagens (LOPES, 2005). Este modelo clássico é representado como a sequência de quatro etapas: geração de viagens, distribuição de viagens, divisão modal e alocação de tráfego.

A primeira etapa deste modelo, a geração de viagens, está condicionada ao padrão de uso do solo e à dinâmica socioeconômica de uma região. A produção e a atração de viagens dependem de características socioeconômicas, como população, nível de renda, matrículas escolares, quantidade de empregos, como também da estrutura física, rede viária e do sistema de transporte de uma cidade. Segundo Lopes (2005), dentre as variáveis utilizadas na análise da geração de viagens, destacam-se renda, posse de veículos, tamanho da família, densidade residencial e acessibilidade. O total da população é um elemento importante para se estimar a produção de viagens (BRUTON, 1979). No caso da atração de viagens, é importante se considerar outros fatores, como áreas de atividades industriais, de comércio e serviços, e, consequentemente, a quantidade de empregos. No entanto,

os modelos tradicionais não incorporam o fator espacial nas suas análises. A análise espacial tem como objetivo avaliar propriedades e relacionamentos, considerando a localização espacial do fenômeno estudado (LOPES, 2005). Para Câmara *et al.* (2002) os processos associados espacialmente são denominados fenômenos geográficos, que podem ser explorados e modelados levando-se em consideração a sua distribuição espacial. Segundo esses autores, estudos considerando a análise espacial vêm sendo aprimorados graças à crescente utilização dos Sistemas de Informação Geográfica (SIG). Os SIG permitem a transformação e manipulação dos dados geográficos no apoio à tomada de decisão por demonstrar padrões e transformar dados brutos em informação espacialmente localizada.

A análise de demanda por transportes relaciona-se a 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 previsão de demanda torna-se importante para aperfeiçoamento das estimativas. A existência de autocorrelação espacial de variáveis relacionadas ao comportamento relativo a viagens amplia as possibilidades de análises a partir da utilização de técnicas que permitam incorporar a localização espacial (Costa *et al*., 2013).

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 desconhecida.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

O objetivo principal deste estudo é estimar a geração de viagens por Transporte Coletivo (TCO) em coordenadas conhecidas e desconhecidas através de técnicas de geoestatística: krigagem ordinária e krigagem com deriva externa.

1.1.2 OBJETIVOS ESPECÍFICOS

- Estimar dados de população, empregos e geração de viagens em coordenadas desconhecidas, através de krigagem ordinária;

- Estimar dados de atração e produção de viagens por TCO a partir das variáveis empregos e população, respectivamente, através da Krigagem com Deriva Externa;

- Comparar a estimação e interpolação espacial da geração de viagens por TCO, considerando dois tipos de dados (amostra sistemática artificial – grade regular - ou por centróide das macrozonas de tráfego (MZs);

- Propor estimativas futuras de geração de viagens, inclusive em coordenadas desconhecidas, baseadas em dados censitários.

1.2 JUSTIFICATIVA

Considerando-se a importância da localização espacial de fenômenos geográficos, sobretudo na geração de viagens, o presente trabalho visa estimar a produção e atração de viagens por transporte coletivo, inclusive em coordenadas desconhecidas, no horário de pico da manhã (6 às 8 horas) na Região Metropolitana de Salvador (RMS).

Optou-se por uma técnica pouco utilizada no planejamento de transportes, já que a estatística tradicional não aborda a análise espacial. Este fator dificulta as análises, pois a estatística defende que todas as amostras são aleatórias, negligenciando a localização destas no espaço (COSTA, 2013).

A Geoestatística é uma técnica recentemente utilizada no Planejamento de Transportes. A dependência espacial das características relacionadas ao espaço urbano permite sua adequada utilização no estudo da geração de viagens. Com isso, busca-se analisar as características socioeconômicas e de viagens da população, através do uso de ferramentas de análise espacial, como a krigagem.

Pretende-se realizar um estudo a fim de aprimorar a estimação de viagens, estendendo os dados conhecidos a coordenadas geográficas desconhecidas, colaborando, desta forma, para corroborar a viabilidade da aplicação da Geoestatística como ferramenta de análise espacial urbana.

Outro aspecto importante deste estudo é a proposta de um procedimento para estimação da geração de viagens em coordenadas desconhecidas a partir de dados atuais, como o censo demográfico de 2010 (IBGE, 2010).

1.3 ESTRUTURA DO TRABALHO

Este estudo propõe a estimação da geração de viagens por transporte coletivo, a partir de técnicas da geoestatística. A Figura 1.1 representa o esquema metodológico, que buscou atender aos objetivos deste estudo. Cada etapa, sumariada a seguir, está detalhada nos próximos capítulos.

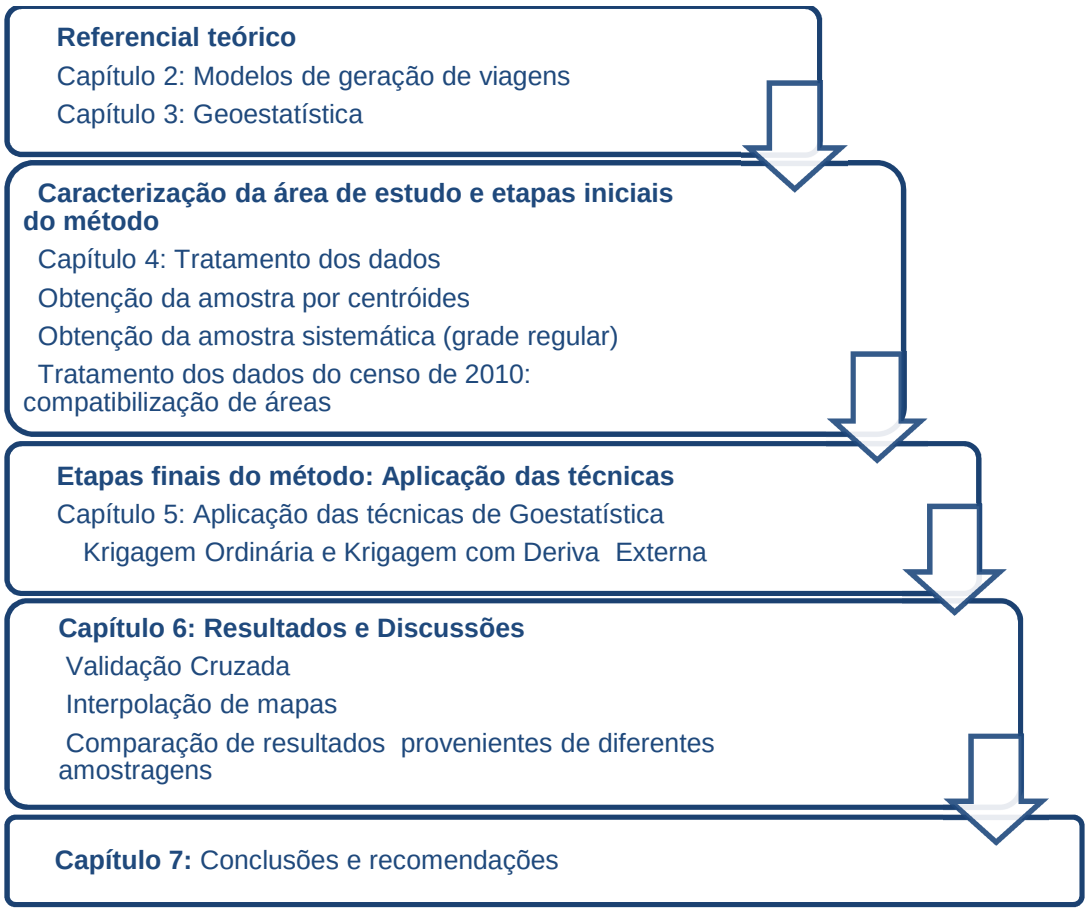


Figura 1.1: Estrutura da dissertação

Etapa 1: Referencial teórico

A primeira fase deste estudo consiste na revisão bibliográfica sobre modelos de geração viagens, além dos principais conceitos da geoestatística. Desta forma, os capítulos 2 e 3 do presente trabalho descrevem alguns aspectos referentes aos modelos de geração de viagens, suas aplicações e a abordagem de análise espacial em transportes, além de conter os principais conceitos da geoestatística, respectivamente.

Etapa 2: Tratamento dos dados

Nesta etapa foi feito o tratamento dos dados da Pesquisa O/D de 1995, realizada na Região Metropolitana de Salvador (RMS), e selecionadas as variáveis mais usuais nos modelos de geração de viagens. Também foi feita a compatibilização das áreas das Macrozonas da O/D de 1995 e setores censitários de 2010, para obtenção de dados atuais estimados de produção de viagens na RMS.

Portanto, esta etapa do método compreende tratamento de dados da RMS (1995) e do censo demográfico de 2010, além de geração de amostra sistemática artificial dos pontos observados.

Etapa 3: Aplicação das técnicas

A aplicação da Krigagem ordinária e Krigagem com deriva externa na estimação da geração de viagens está descrita no capítulo 5 deste trabalho. Após a realização da etapa anterior, tratamento dos dados da área de estudo, foi possível realizar a estimação da produção e atração de viagens com auxílio do *software* geoMS (*Geostatistical Modelling Software*), desenvolvido em 1999, por uma equipe de investigadores do atual CERENA do Instituto Superior Técnico (Lisboa).

Etapa 4 : Resultados e discussões

Após a etapa de krigagem é necessário avaliar os resultados das estimações. Antes da execução dos mapas da krigagem é necessário avaliar os erros, ou validar a estimação para verificar sua confiabilidade. A validação cruzada (*cross validation*) ou

teste do ponto fictício é a etapa de mensuração dos erros da estimação. Desta maneira, haverá, para cada ponto, o valor conhecido e o valor estimado. Assim, é possível verificar o erro de estimação.

A interpolação espacial da geração de viagens irá gerar mapas de superfície, representando a estimação da produção e atração de viagens para 1995 e de produção de viagens para 2010.

CAPÍTULO 2: MODELOS DE GERAÇÃO DE VIAGENS

Este capítulo traz os principais conceitos sobre modelos de geração de viagens, Primeiramente, é feita uma abordagem sobre os modelos de demanda tradicionais, como os modelos de demanda diretos e sequenciais, com enfoque para o modelo Quatro Etapas. Em seguida traz uma revisão da literatura sobre modelos de demanda por transportes. Na última seção deste capítulo encontra-se uma revisão sobre análise espacial e sua aplicação em estudos de demanda por transportes.

2.1 MODELOS TRADICIONAIS: CONCEITOS

O planejamento de transporte urbano visa a previsão de demandas futuras, através do manejo de informações importantes, como dados socioeconômicos e demográficos da população e características de uso e ocupação do solo.

A previsão da demanda é uma etapa fundamental no planejamento de transportes, pois ela mostra o padrão de deslocamento de pessoas ou mercadorias, ajudando a estabelecer, de forma prioritária, a oferta de transportes. Assim, o objetivo da análise da demanda por transportes é compreender os seus determinantes, as interações que os envolvem e como eles influenciam o volume do tráfego, a fim de obter uma projeção futura da demanda por transportes (KAWAMOTO, 2010). A partir da metade de século passado, começam a surgir várias técnicas para modelagem de transporte.

Novaes (1986) observa três níveis de análise utilizados em estudos na previsão de demanda por transporte:

- Previsões a curto prazo: abordam apenas situações atuais, onde as análises são feitas através de pesquisas de medições, levantamentos etc. Nesta etapa não são feitas previsões de variáveis socioeconômicas, como população, empregos, uso do solo, entre outras, considerando que a distribuição espacial destas variáveis permanecerá a mesma.

- Previsões a médio e longo prazo sem considerar efeitos nas atividades socioeconômicas: nesta etapa há projeções mais detalhadas das variáveis

socioeconômicas por zona, sem se estabelecer um relacionamento direto entre a demanda prevista de transportes e o seu efeito no uso do solo e suas atividades.

- Previsões de longo prazo com avaliação dos efeitos nas atividades socioeconômicas e uso do solo: é um misto de planejamento de transportes e urbano regional. São feitas previsões detalhadas envolvendo o uso do solo e as atividades existentes no espaço geográfico.

Um modelo de demanda pode ser entendido como uma função matemática que serve para quantificar a demanda por transportes, baseado em variáveis que condicionam tal demanda entre a origem e destino das viagens (ou pares de zonas), o motivo, o modo de transporte e o caminho escolhido pelos indivíduos (KAWAMOTO, 2010).

A execução de um modelo de demanda e sua representação da realidade irá depender das variáveis adotadas e da relação entre elas, como também do grau de importância que cada uma irá exercer sobre o modelo.

Kawamoto (2010) classifica dois tipos de modelos aplicados à previsão de demanda por transportes: modelos sequenciais e modelos diretos ou simultâneos. A diferença entre os dois modelos é o processo de modelagem de cada um.

A análise da demanda por transportes pode ser realizada para investigar novas estratégias gerenciais ou operacionais através de modelos diretos. Os modelos diretos são utilizados para modelar simultaneamente a geração, distribuição de viagens e divisão modal, podendo evitar a propagação de erros durante a modelagem (TEIXEIRA, 2003).

Assim, através de uma função matemática (Equação 2.1), representa-se o volume de viagens que ocorre entre uma origem e um destino, por um determinado modo e motivo de viagem, em função de diversas características.

$$Q_{ijm}^{n} = \alpha . X_1^{\beta} . X_2^{\gamma} . X_3^{\delta} \ldots \qquad \text{(eq 2.1)}$$

sendo,

Q_{ijm}^{n} a quantidade de usuários que vão da zona i para j, pelo motivo n, através do modo m;

X_i representam as variáveis do modelo e,

$\alpha, \beta, \gamma, \delta$ são os parâmetros que definem o peso de cada variável.

No entanto, a partir do momento em que aumenta a quantidade de dados analisados, como o número de pares de zonas, os motivos e modos utilizados, as análises se tornam mais complexas, e o emprego dos modelos diretos se mostra limitado.

Os modelos sequenciais subdividem o estudo de previsão da demanda de acordo com a sequência do processo das decisões dos indivíduos, como por exemplo, a viagem da zona ‘a’ para a zona ‘b’, por determinado modo, motivo e determinada rota (KAWAMOTO, 1994). O modelo mais tradicional na previsão de demanda por transportes é o modelo Quatro Etapas. As principais características desse modelo estão descritas a seguir.

2.2 MODELO QUATRO ETAPAS

O objetivo do modelo Quatro Etapas é estimar a demanda atual e futura por transportes (TEIXEIRA, 2003). O modelo é dividido em quatro fases distintas, mas interligadas ou sequenciais:

I) Geração de viagens;
II) Distribuição de viagens;
III) Divisão modal;
IV) Alocação do tráfego.

I) Geração de viagens:

A primeira etapa do modelo investiga o tipo de viagens de pessoas e veículos, e tem como objetivo estimar a produção e atração de viagens em cada zona de tráfego

através do uso de taxas médias de viagens (LOPES, 2005). Esta fase é essencial no planejamento de transportes, pois é nela que se analisam os principais fatores que determinam a geração de viagens em uma zona.

A geração de viagens está condicionada aos tipos de usos do solo urbano, características das residências, como tamanho da família, propriedade de veículos, tipos de domicílios, tipo de ocupação dos residentes, e fatores fundamentais, como a renda familiar, população e quantidade de empregos.

A faixa etária da população também pode ser considerada um fator relevante na taxa de produção de viagens de uma zona de tráfego, pois idades diferentes geram demandas distintas por viagens. Outro fator analisado por Ortúzar e Willumsen (2011) é a questão da acessibilidade. Segundo estes autores, zonas com mais infraestrutura viária e acessibilidade tendem a atrair mais viagens. Assim, o primeiro procedimento para análise da geração de viagens é determinar as variáveis explicativas do volume do tráfego. As principais características dessa etapa estão melhor descritas na seção 2.3 deste capítulo.

II) Distribuição de viagens:

Definido o volume de tráfego produzido e atraído nas zonas de tráfego, é necessário conhecer a distribuição das viagens, identificando o volume de tráfego entre a zona de origem e as demais zonas (NOVAES, 1986). Desta forma, pode ser gerada a matriz de viagens, conhecida como matriz origem-destino (O-D). Também, é possível analisar viagens por períodos do dia (hora de pico), modo de transporte e motivo de viagem (LOPES, 2003).

Existem alguns modelos que analisam a distribuição de viagens, como os que utilizam o fator de crescimento. Este modelo possui algumas restrições à sua utilização, principalmente em países em desenvolvimento e em áreas que terão modificações na infraestrutura do sistema de transportes, pois essas mudanças não são detectadas pelo modelo (MELLO, 1975). Dos modelos que utilizam o fator de crescimento, o mais representativo é o modelo Fratar, que apesar de antigo é muito utilizado. Neste modelo, a previsão de viagens futuras é feita a partir da multiplicação do volume do tráfego atual pelo produto dos fatores de crescimento

para duas zonas. É feito o ajuste da atratividade relativa para as outras zonas e o cálculo do valor de viagens distribuídas a partir de uma determinada zona, obtendo-se os fluxos dos pares de zonas (KAWAMOTO, 1994).

Outros modelos utilizados na previsão da distribuição de viagens são os modelos gravitacionais.

Uma das vantagens deste modelo é que se considera o efeito da distância espacial ou facilidade de iteração entre as regiões definidas pela função de impedância (CAMPOS, 2013). Além disso, a taxa de geração de viagens em uma zona é proporcional à sua “massa”, nesse caso podendo ser representada por variáveis como população, emprego, etc. (NOVAES, 1986). Deve-se observar que a distância pode ser substituída por outros fatores, como tempo e custo da viagem, ou por diversos fatores compostos, denominados de impedância (resistência ao deslocamento) ou custo generalizado (atribuído a todos os fatores que exercem influência na impedância, com base na percepção do usuário) (NOVAES, 1986).

Estes modelos passaram a ser utilizados na área de transportes a partir da década de 50 do século passado, e ao longo do tempo foram feitas algumas mudanças na escolha das variáveis utilizadas no modelo, como por exemplo, tempo, custo, viagens produzidas na zona de origem e atraídas nas zonas de destino. Diferentemente da lei de gravitação universal, que utiliza apenas uma constante, a adaptação desse modelo para o estudo de transportes utiliza uma constante de proporcionalidade para cada zona de origem. A calibração do modelo pode ser feita através de regressão linear, conhecendo-se o número de viagens produzidas e atraídas para cada zona (KAWAMOTO, 2010). Podem-se citar ainda outros modelos de distribuição de viagens, tais como: entrópico, modelos sintéticos e modelos de oportunidades intervenientes.

III) Divisão modal:

Conhecendo-se a distribuição das viagens em relação às zonas de tráfego, em seguida é feita a divisão modal, ou seja, a análise da escolha do modo de transporte do total de viagens realizadas. Diversos fatores podem influenciar na escolha modal, como por exemplo, velocidade, conforto e custo (BRUTON, 1979). As variáveis

utilizadas na divisão modal são essencialmente características socioeconômicas, das viagens e do sistema de transporte (CAMPOS, 2013).

A etapa de escolha modal é realizada a partir de modelos comportamentais, com o objetivo de refletir o comportamento do usuário em relação às escolhas de deslocamentos, sendo o Modelo Logit o mais utilizado (KAWAMOTO, 2010).

Este modelo é baseado em três hipóteses básicas, nas quais o ser humano é capaz de *perceber*, *comparar* e *escolher* o objeto de maior utilidade, e no caso do estudo da divisão modal, o objeto é representado por um determinado modo de transporte.

Desta forma, o usuário pode avaliar a utilidade desse modo através da sua quantificação em função dos atributos do modo e de variáveis socioeconômicas do usuário. Outra hipótese admitida pelo Modelo Logit é a existência de um erro de percepção da utilidade do objeto. A utilidade percebida pode ser descrita como a utilidade verdadeira mais um erro qualquer de percepção:

$$U = V + \varepsilon \qquad \text{(eq 2.2)}$$

para,

U = a utilidade percebida;

V =a utilidade verdadeira;

ε é o erro de percepção, que segue a distribuição acumulada de Weibull.

Conhecendo-se a utilidade associada a cada modo, calcula-se a probabilidade de escolha de determinado modo pelo usuário. Então, é possível conhecer os valores estimados de viagens através do produto do total de viagens entre duas zonas e a probabilidade de escolha de cada modo (NOVAES, 1986).

IV) Alocação de tráfego:

Por fim, a última etapa do modelo, a alocação de tráfego, investiga qual o caminho percorrido por estas viagens e a sua distribuição pela rede de transportes. A preocupação com a alocação do tráfego gerado surgiu com o aumento da utilização do transporte individual, o que ocasionou o surgimento de congestionamentos

constantes nas grandes cidades. Os modelos que analisam a alocação do tráfego procuram mensurar as condições do sistema viário de absorver o fluxo de viagens gerado pelos modos de transporte, principalmente o transporte individual (CAMPOS, 2013).

Os modelos de alocação investigam dois parâmetros: o tempo de deslocamento e a capacidade de cada via investigada, ou até mesmo o custo de deslocamento (NOVAES, 1986), além de investigar o menor caminho, para diminuir o tempo de viagem (CAMPOS, 2013).

A primeira técnica empregada para analisar a alocação do tráfego é o modelo “Tudo ou Nada”. Esta técnica admite que, definida a melhor rota entre um par OD, todo fluxo passaria por esta via, sem considerar sua capacidade. Outros modelos surgiram posteriormente com o intuito de aprimorar esta técnica, tais como alocação incremental e métodos estocásticos baseados em simulações. Este modelo possibilita uma análise microscópica, no qual se analisam inclusive os problemas de congestionamento nas interseções.

2.3 MODELOS DE GERAÇÃO DE VIAGENS: CONSIDERAÇÕES TEÓRICAS

O entendimento das viagens urbanas tem como pressuposto considerar alguns tópicos essenciais. Para Ortúzar e Willumsen (2011), viagens urbanas podem ser definidas como de base domiciliar e não domiciliar. Viagens de base domiciliar têm o domicílio como origem ou destino da viagem. Já em viagens de base não domiciliar, tanto a produção quanto a atração de viagens não têm relação com o domicílio do indivíduo. Os autores definem outros termos bastante utilizados no planejamento de transportes e estudados neste trabalho.

Viagem: é o movimento de único sentido a partir de um ponto de origem para um ponto de destino. Deslocamentos a pé são considerados a partir de uma distância de 300 metros.

Geração de viagens: pode ser definida como o número total de viagens produzidas e atraídas por zona de tráfego, sejam elas de base domiciliar ou não.

Produção de viagens: são as viagens originadas numa determinada zona.

Atração de viagens: quantidade de viagens atraídas em uma zona, sejam elas de base domiciliar o não.

Um conjunto de características define o padrão de viagens em uma zona. Verificou-se que o estudo da geração de viagens pode ser melhor compreendido se classificados os motivos das viagens. A subdivisão da geração de viagens por motivo facilita a análise do modelo (ORTÚZAR e WILLUMSEN, 2011). Segundo esses autores, os principais motivos identificados na geração de viagens são:

- trabalho;
- educação;
- compras;
- sociais ou de lazer;
- outros motivos.

Os dois primeiros motivos caracterizam viagens rotineiras ou obrigatórias, enquanto os demais podem ser classificadas como viagens mais esporádicas ou opcionais. Por isso a importância da classificação de viagens por motivo, já que são as viagens rotineiras que representam maior demanda por transportes. A última categoria abrange todas as viagens menos frequentes, como ir ao médico ou para negócios pessoais.

A quantidade de viagens também pode oscilar dependendo do período do dia, geralmente concentrando grande fluxo nos horários de pico. Neste período, a maioria das viagens é do tipo obrigatória, ou seja, viagens por motivo de trabalho e educação. Muitas vezes o pico de viagens varia de acordo com aspectos do uso do solo em cada região, podendo alterar em diferentes áreas da cidade. No horário de pico se concentra cerca 10 a 20% do total das viagens diárias (CAMPOS, 2013). Neste trabalho optou-se por analisar a produção e atração de viagens por transporte coletivo no horário de pico da manhã.

Viagens produzidas e atraídas acontecem em função de um conjunto de características. Na etapa de produção de viagens são tratadas, principalmente, características relacionadas à população ou a indivíduos. Ortúzar e Willumsen (2011) abordam alguns desses fatores:

* renda;
* posse do carro;
* tamanho da família;
* estrutura familiar;
* valor da terra;
* densidade residencial;
* acessibilidade.

Os fatores relacionados a características individuais ou familiares (renda, posse de carro, estrutura familiar e tamanho da família) são considerados nos estudos de geração de viagens domésticas. Já fatores como o valor da terra e densidade residencial são utilizados para estudar zonas de tráfego. A acessibilidade é definida como a facilidade de acesso de pessoas a bens ou equipamentos, e vem sendo incorporada aos conceitos principais do planejamento de transportes, já que a facilidade de acesso ou locomoção dos indivíduos pode melhor representar alguns aspectos da etapa de geração e viagens.

Na etapa de atração de viagens, os fatores mais utilizados têm sido características de serviços industriais, comerciais, educacionais, empregos, dentre outros (ORTÚZAR e WILLUMSEN, 2011).

Os mesmos autores ainda destacam dois aspectos relacionados à geração de viagens: a estabilidade temporal e a estabilidade geográfica. A estabilidade temporal significa que algumas características permanecerão constantes durante um longo período de tempo. Por exemplo, as viagens pelo modo ônibus algumas vezes aparecem estáveis no decorrer do tempo, para determinados tipos de domicílios. Por outro lado, viagens por transporte individual, como o carro, são dependentes de outros fatores, como por exemplo, valor dos combustíveis, ou mesmo facilidade de aquisição de um automóvel, como vem ocorrendo nos últimos anos no Brasil. A estabilidade geográfica é outro fator importante na análise de geração de viagens, e que pode auxiliar na redução dos custos de pesquisas. Viagens são relacionadas a grupos homogêneos de pessoas ou famílias, logo pode-se considerar que regiões com características semelhantes terão dados relacionados a viagens similares.

2.3.1. Principais modelos de geração de viagens

A partir do início da década de 1950 começam a surgir modelos para representar a geração de viagens. A maioria dos métodos procura prever o número de viagens produzidas ou atraídas por família ou zona de tráfego, em função das características já mencionadas.

a) Fator de crescimento

Uma técnica bastante utilizada para prever a geração de viagens é o fator de crescimento. Esta técnica prevê o número de viagens futuras por zona de tráfego em função de variáveis que têm influência na geração de viagens (CAMPOS, 2013). Com este método, a previsão do número de viagens é determinada através da aplicação de um fator de crescimento aos dados atuais. Dessa forma, o resultado da estimação será igual aos dados atuais, multiplicado pelo fator de crescimento. O fator de crescimento pode ser determinado através de qualquer variável que explique a geração de viagens (MELLO, 1975), onde:

$$Ti = Fi \, .ti \qquad \text{(eq 2.3)}$$

Ti = número de viagens futuras produzidas na zona de tráfego i;

Fi = fator de crescimento;

ti = número de viagens do ano base.

O fator de crescimento é utilizado para estimar o aumento de viagens em determinada zona de tráfego, associando dados de tráfego com dados de uso do solo. Segundo Ortúzar e Willumsen (2011) o problema deste método é como conhecer o fator de crescimento Fi, mas ele pode ser relacionado a variáveis como população, renda, posse de automóveis, etc.

Este método tem caráter simplista e não analisa fatores essenciais na geração de viagens, como características socioeconômicas, e é muito pouco utilizado no planejamento de transportes, por ser considerado rudimentar (BRUTON, 1979). Quando utilizado, deve ser aplicado para planos a curto prazo. Outro problema relacionado a essa técnica é que, ao se assumir um valor de crescimento constante, erros podem estar inseridos no modelo. Isso pode comprometer todo processo de

modelagem, já que a geração de viagens é a primeira etapa do planejamento de transportes (LOPES, 2003).

b) Regressão linear simples e mútipla

A utilização deste método visa construir uma relação linear entre o número de viagens existentes (variável dependente) e os vários fatores que influenciam as viagens (variáveis independentes). Este modelo relaciona a demanda por transportes com as variáveis explicativas, como população, renda, produção industrial, empregos, dentre outras que podem ser utilizadas para explicar o funcionamento do sistema de transporte de uma área (MELLO, 1975). Esta técnica é bastante disseminada no estudo de transportes para estimar a geração futura de viagens através de duas ou mais variáveis relacionadas, onde são atribuídos pesos a cada uma das variáveis (BRUTON, 1979).

Ortúzar e Willumsen (2011) consideram alguns fatores importantes sobre a utilização da regressão linear na modelagem da geração de viagens. A aplicação de modelos de regressão para a previsão da geração de viagens, a partir de análises de zonas, só pode explicar a variação no comportamento de viagens entre zonas. Assim, esse modelo apenas prever adequadamente a produção e atração de viagens inter-zonais.

Podem acontecer casos onde algumas zonas não contenham informações sobre certas variáveis dependentes. Por exemplo, provavelmente não pode ocorrer produção de viagens de base domiciliar em zonas não residenciais. Estes autores chamam essas zonas de zonas nulas. Segundo Ortúzar e Willumsen (2011), elas devem ser excluídas das análises. Zonas que não fornecem dados úteis tendem a produzir coeficientes que superestimam a precisão dos dados estimados pela regressão.

Quando os estudos são feitos para zonas, ou seja, agregando-se as variáveis, o tamanho da zona pode interferir na intensidade dos erros. Vale ressaltar que mesmo utilizando taxas de viagens, a variabilidade interzonal diminui com o aumento do tamanho da zona. Os mesmos autores ilustram um exemplo desse efeito na previsão da produção de viagens entre dois diferentes tipos de zoneamento.

Tabela 2.1: Variação interzonal de produção de viagens através de dois diferentes sistema de zoneamento

Sistema de zoneamento	Valor médio de viagens por zona	Variância
75 zonas pequenas	8.13	5.85
23 zonas maiores	7.96	1.11

Fonte: Ortúzar e Willumsen, 2011.

Como pode ser notado, o tamanho da zona influencia na variância da estimação do valor médio de viagens. Os autores apontam como redução para variações intra-zonais a diminuição do tamanho das zonas, ou desagregação dos dados, utilizando dados de indivíduos ou domicílios.

Além disso, é importante mencionar que existem variáveis qualitativas na análise de geração de viagens (gênero, ocupação, etc.), que geralmente apresentam um comportamento não linear. Uma alternativa para este problema é a utilização de variáveis *dummy*, que tornam o modelo de regressão linear uma ferramenta flexível, característica importante para tratar muitos problemas encontrados em estudos empíricos (MISSIO e JACOB, 2007).

c) Classificação cruzada ou análise de categorias

Essa técnica utiliza dados desagregados por categorias de residência (CAMPOS, 2007). As viagens são agrupadas e relacionadas à estrutura familiar e às condições econômicas de cada família. Este modelo considera que as taxas de geração de viagens, independente da categoria da família, são constantes no futuro (ORTÚZAR e WILLUMSEN, 2011).

A vantagem desta técnica é que os agrupamentos de classificação cruzada não têm relação com o zoneamento da área de estudo. Não são necessárias suposições prévias sobre a forma da relação entre os dados, e o relacionamento entre eles pode ser diferente no tipo de análise. No entanto, este método apresenta algumas desvantagens, como a necessidade de grande quantidade de amostras. Outra desvantagem é a ausência de um método eficaz para a escolha das variáveis de

classificação, ou para escolher melhores agrupamentos de uma determinada variável (ORTÚZAR E WILLUMSEN, 2011).

2.4 ANÁLISE ESPACIAL E DEMANDA POR TRANSPORTES: REVISÃO DA LITERATURA

A Primeira Lei da Geografia afirma que todas as coisas no espaço estão relacionadas, no entanto, eventos mais próximos no espaço são mais relacionados entre si (TOBLER, 1970). Isto indica que existe uma relação de dependência entre fenômenos, sejam eles naturais ou sociais, porque são, sobretudo, fenômenos inseridos no espaço geográfico. Assim, os resultados a partir de análise espacial dependem da localização dos objetos analisados.

Os efeitos locacionais se manifestam de duas formas: através da dependência e da heterogeneidade espacial. O primeiro caso, muitas vezes se refere ao problema de autocorrelação espacial. Já a heterogeneidade espacial ocorre uma vez que cada local tem um grau de singularidade intrínseca no que diz respeito ao resto do sistema espacial. Não considerar a dependência e heterogeneidade espacial pode tornar os resultados de análise de dados espaciais inválidos (MILLER, 1999).

Os modelos tradicionais de demanda por transportes não correspondem à atual dinâmica sócio-espacial das cidades. O problema de muitos modelos tradicionais de demanda por transportes é que os mesmos não analisam um componente básico na estrutura urbana e de transportes: o fator espacial. Sabe-se que a geração de viagens está relacionada a diversos fatores socioeconômicos, e isso possibilita análises mais confiáveis e próximas do real, incorporando sob esta ótica a localização espacial dos fenômenos. Padrões de viagens regulares são definidos a partir da análise da relação espacial de dados referentes a estudos de transportes. Autores como Doubleday (1976), Bolduc *et al*. (1988), Said e Young (1989) e Yao e Morikawa (2005) destacaram que a relação entres variáveis como posse de veículos e estrutura familiar, explicam as viagens produzida, e auxiliam na análise de demanda por transportes.

Técnicas que consideram a dependência espacial de dados passaram a ser utilizadas na análise de demanda e planejamento de transportes. Nas últimas

décadas, muitos trabalhos, como Teixeira (2003), Lopes (2005 e 2010) e Pitombo *et al.* (2010) mostram que a inserção da análise espacial nos estudos de transportes aprimora as estimações, principalmente após os avanços incorporados aos Sistemas de Informações Geográficas (SIG).

Outros autores destacam a importância da dependência espacial na análise de dados geográficos. Wang *et al.* (2011) discutem como técnicas de regressão aplicadas a análises espaciais na área de transportes são limitadas, apesar de muito utilizadas na previsão de demanda por viagens e acidentes de trânsito. Os autores discutem estas limitações, abordando entre elas a existência do problema da unidade de área modificável, falácia ecológica e a dependência espacial. Os dois primeiros problemas são relacionados, e afirmam que os resultados da modelagem estatística podem variar tanto devido ao nível de agregação dos dados, quanto através da modificação da escala ou o tamanho de agregação das zonas.

Lopes (2005) confronta a análise tradicional e modelos de análise espacial em um estudo relacionado ao planejamento de transportes, mostrando os efeitos e avanços que a inserção da dependência espacial atribui aos modelos tradicionais. A análise de demanda apresentada mostra que através da análise espacial foi possível produzir estimativas mais precisas em relação aos modelos tradicionais, inclusive apresentando menos erros na estimação. Além disso, os modelos tradicionais apresentaram maiores coeficientes de correlação entre as variáveis independentes, podendo gerar problemas de multicolinearidade com inclusão de mais de uma delas no modelo.

Em outro estudo, Lopes (2010) analisa os efeitos espaciais e a relação entre transporte e uso do solo no estudo de mobilidade, avaliando a aplicação de técnicas que considerem os efeitos espaciais na modelagem urbana. Considerando a dependência espacial e a relação entre uso do solo e transportes, a autora aplicou um modelo desenvolvido na Áustria, o *Metropolitan Activity Relocation Simulator* (MARS), com o qual foi possível estabelecer oito cenários distintos de planejamento urbano em um período de 30 anos para Porto Alegre.

Uma área urbana tem diferentes características de emprego, moradia e áreas de lazer, por exemplo. Os indivíduos participam de diferentes atividades em diferentes

momentos. Sendo assim, há uma complexa estrutura dinâmica regular que não pode ser adequadamente percebida por uma representação estática da realidade do espaço urbano (GOODCHILD, KLINKENBERG e JANELLE, 1993[1]; JANELLE, KLINKENBERG e GOODCHILD, 1998[2] *apud* MILLER, 1999). Isso indica as complexas e dinâmicas inter-relações entre os fenômenos espaciais e suas localizações. Diferentes análises sobre viagens podem ser mais relevantes no estudo do comportamento espacial que um tipo de visão estática, frequentemente utilizada nos modelos tradicionais de análise de demanda por transportes.

Tsutsumi e Seya (2008) fizeram um estudo para avaliar o impacto de projetos de transportes em um contexto regional. Houve significativo progresso em pesquisas empíricas sobre os mercados imobiliários que consideram a dependência espacial. Este estudo pretendia mensurar este impacto sobre os preços da terra, utilizando modelos estatísticos espaciais. Vários tipos de modelos de preços de terra foram construídos empregando os métodos existentes de econometria espacial e geoestatística. Usando esta segunda técnica foi possível interpolar o preço da terra e sua dinâmica.

Adjemian *et al.* (2010) discutem as razões da existência de efeitos espaciais na posse de veículos. Os resultados apontaram para a importância de estudar as relações espaciais em pesquisa de transporte, além de destacar as desvantagens de ignorar o seu papel. Esta pesquisa aponta que os fatores espaciais devem ser considerados a fim de estimar corretamente modelos de escolha modal, evitando inferências equivocadas.

Assim, considera-se que o efeito da análise espacial no planejamento de transportes pode contribuir para análises mais precisas, já que as variáveis utilizadas apresentam-se espacialmente distribuídas no espaço, podendo ser analisadas considerando-se o seu comportamento e a relação entre elas.

A Tabela 2.2 reúne alguns estudos que mostram a relação da dependência espacial e estudos de transportes. Todos os trabalhos apresentados na tabela indicam que a

---

[1] GOODCHILD, M. F.; KLINKENBERG, B.; JANELLE, D. G. (1993). *A Factorial Model of Aggregate Spatio-Temporal Behavior: Application to the Diurnal Cycle.* Geographical Analysis 25, 277-294.

[2] JANELLE, D. G.; KLINKENBERG, B.; GOODCHILD, M. F. (1998). *The Temporal Ordering of Urban Space and Daily Activity Patterns for Population Role Groups*. Geographical Systems 5, 117-137.

inclusão dos efeitos espaciais e técnicas alternativas de planejamento de transportes melhoram os resultados.

Tabela 2.2: Alguns trabalhos que abordam análise espacial e transportes

Tipo de análise \ Autores	Miller (1999)	Bhat e Zhao (2002)	Lopes (2005)	Lopes; Brondino; Rodrigues da Silva (2006)	Eom *et al.* (2006)	Souza *et al.* (2008)	Tsutsumi e Seya (2008)	Adjemian *et al.* (2010)	Lopes (2010)	Lopes; Rodrigues da Silva; Pfaffenbichler (2011)	Wang *et al.* (2011)
Dependência espacial	x	x	x	x	x	x	x	x	x		x
Heterogeneidade espacial	x										
Problema da unidade de área modificável (MAUP)	x								x		x
Análise espacial e transportes		x	x	x	x		x		x		x
Análise do uso do solo e espaço urbano							x		x	x	
Escolha modal								x			
Regressão espacial			x	x					x		
Acidentes de trânsito						x					

CAPÍTULO 3: GEOESTATÍSTICA

Este capítulo apresenta os principais conceitos de Geoestatística, como as Variáveis Regionalizadas e os Variogramas, além dos principais tipos de Krigagem. Serão discutidas também técnicas de amostragem, com destaque para umas das técnicas aplicadas neste trabalho, a amostragem sistemática, obtida através de uma malha regular. Por fim, serão descritas aplicações da Geoestatística em estudos de Transportes encontrados na literatura.

3.1 CONTEXTUALIZAÇÃO

Através da geoestatística, foi desenvolvida uma teoria para explicar fenômenos nos quais os valores das variáveis estão associados a coordenadas geográficas. A mesma faz parte de uma categoria de análise espacial de dados, na qual são estimadas superfícies contínuas a partir de um conjunto de amostras de campo que podem estar regularmente ou irregularmente distribuídas no espaço.

Seu principal objetivo é caracterizar a dispersão espacial e/ou espaço-temporal de um fenômeno, avaliar medidas de incerteza, considerando a sua variabilidade espacial. De acordo com Landim *et al.* (2002c), a geoestatística analisa o comportamento das Variáveis Regionalizadas (VR), a serem descritas detalhadamente na seção 3.1, e tem como base três pressupostos:

Ergodicidade: a esperança relacionada à média de todas as estimações de uma variável é igual a média de uma única estimação em certo domínio;

Estacionariedade: o fenômeno é homogêneo na área a ser estimada;

Hipótese íntrisica: as diferenças entre valores são localmente estacionárias, ou seja, a variância entre dois pontos, separados pela distância h, é válida para toda área de estudo, analisados sob a mesma distância e direção dos pontos (ROCHA, 2004).

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. Esta técnica foi utilizada primeiramente em estudos geológicos, e sua evolução começa a partir do momento que o estudo dos recursos naturais começa a ter mais visibilidade. Nesse contexto, a estatística tradicional se mostrou limitada, por não representar adequadamente fenômenos espaciais, já que não considera a localização espacial das amostras.

Os modelos geoestatísticos resultam da associação entre a componente naturalista (geologia, hidrologia, etc.) com os fundamentos teóricos da matemática e da estatística (SOARES, 2006). Pode-se considerar que as áreas que mais utilizam a geoestatística são as Ciências Naturais, todavia, recentemente vem sendo aplicada com êxito nas Ciências Sociais.

A geoestatística analisa o comportamento espacial de variáveis e permite estimar seu valor em uma área geograficamente amostrada, tornando possível estimar dados e se conhecer a localização espacial das observações. Com isso, a geoestatística é utilizada para obter estimações e interpolações de dados para uma área onde o valor da variável apenas é conhecido pontualmente (COSTA, 2013). As técnicas da geoestatística são usadas para entendimento e modelagem da variabilidade espacial de atributos. Portanto, a localização do fenômeno é um fator importante na análise espacial, pois a referência da localização é fundamental no processo de tomada de decisão.

Modelos geoestatísticos são oriundos de fenômenos aleatórios, mas que obedecem a uma tendência espacial (processos estocásticos). Quanto mais próximos de um ponto amostrado, valores semelhantes ocorrerão, e caso contrário, à medida que a distância aumentar, os valores encontrados serão menos parecidos com o valor da amostra analisada (hipótese da estacionariedade). Com isso, para se realizar uma inferência geoestatística é preciso ter como pressuposto “a concepção de um processo aleatório que melhor caracterize um conjunto de dados (...)”. Para isso é indispensável a escolha de um conjunto de parâmetros que permita a inferência espacial dos dados (SOARES, 2006).

O arcabouço teórico-metodológico da geoestatística consiste em técnicas que quantifiquem a continuidade espacial do fenômeno estudado, baseado na variabilidade estrutural, e a incerteza relacionada aos dados analisados (SOARES,

2006). A geoestatística compreende três processos fundamentais: o entendimento das variáveis regionalizadas, a modelagem dos variogramas experimentais e a etapa de interpolação ou estimação de dados não amostrados, a krigagem.

A vantagem da geoestatística em relação aos métodos tradicionais de interpolação é o conhecimento do comportamento de variáveis em determinada direção (ROCHA, 2004). Ela considera a variabilidade espacial da variável assim como as suas características intrínsecas.

Embora as variáveis que caracterizem a análise de fenômenos naturais e as variáveis relacionadas à análise de demanda por transportes sejam bastante diferentes, pode-se afirmar que existe um fator de associação entre tais estudos: a dependência espacial das variáveis envolvidas. Assim, considerando que a análise das viagens urbanas é fortemente influenciada por aspectos geográficos, conclui-se que é possível considerar as características socioeconômicas e proximidade espacial de atributos. Por isso, estudos relacionados às viagens podem ser realizados utilizando técnicas de geoestatística, fazendo-se adaptações adequadas (COSTA, 2013).

3.2 VARIÁVEL REGIONALIZADA (VR)

As Variáveis Regionalizadas (VR) representam um conjunto de dados distribuídos espacialmente, de forma a se perceber uma determinada tendência espacial dos dados em análise. De acordo com Matheron (1970), uma VR é um termo puramente descritivo, anterior a toda interpretação probabilística.

As VR apresentam dois aspectos contraditórios: um aspecto aleatório e outro estrutural. Os valores assumidos pela VR têm uma componente aleatória, intrínseca ao fenômeno estudado, e outra espacial, segundo a qual esses valores não se distribuem espacialmente de forma aleatória, mas de acordo com determinada distância e direção. O caráter aleatório corresponde à variação inesperada dos pontos no espaço, e o caráter estrutural mostra as correlações existentes entre os diversos pontos onde ocorre o fenômeno (SOUSA e MUGE, 1990). Portanto, a teoria das VR possui dois objetivos importantes: teoricamente, representar as

características estruturais de uma forma matematicamente adequada e, na prática, estimar uma VR a partir de uma pequena amostragem (MATHERON, 1970).

As VR apresentam um aspecto em comum: a forte dependência da localização espacial, por isso são denominadas de variáveis regionalizadas (VR). Qualquer variável dependente do espaço em que, além do caráter aleatório, apresenta um caráter estrutural, pode ser tratada como VR e ser analisada e estimada por técnicas de Geoestatística.

As VR possuem, ainda, outras caraterísticas, entre as quais estão: (1) continuidade e (2) anisotropia (SOUSA e MUGE, 1990).

- A continuidade indica como VR se dispersa espacialmente, ou seja, a estrutura espacial do fenômeno. A continuidade de uma VR pode ser analisada através do comportamento do variograma junto à origem.
- Em alguns casos, a VR pode apresentar um comportamento anisótropico, ou seja, demonstrando uma continuidade espacial distinta em diferentes direções.

3.3 VARIOGRAMA

O variograma é a função que representa quantitativamente a variação de um fenômeno regionalizado no espaço (CÂMARA *et al*., 2002). A continuidade espacial da VR pode ser medida por uma quantidade positiva que traduz as diferenças dos valores da variável medidos em pontos distanciados de uma distância qualquer (h) (PITOMBO e SOUSA, 2009). Portanto, para a construção do variograma, considera-se que a VR tenha um comportamento fracamente estacionário, e que os valores esperados sejam os mesmos para uma determinada área (LANDIM *et al*., 2002c).

O variograma pode ser entendido como uma descrição matemática da relação entre a variância de pares de observações e a distância (*h*) que as separa (JAKOB e YOUNG, 2006).

O variograma é uma ferramenta estrutural que permite descrever os traços marcantes da VR (SOUSA e MUGE, 1990). Portanto, seu ajuste é feito através de

uma curva média em função do conjunto de parâmetros, que melhor represente a continuidade espacial.

O estimador da função variograma da variável regionalizada Z entre dois pontos, xi e xj, separados de uma distância h é dado por:

$$y(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(xi + h) - Z(xj)]^2 \qquad \text{(eq 3.1)}$$

$y(h)$ é a semivariância para o intervalo h (distância);

$N(h)$ é o número de pares de dados que distam h entre eles e alinhados segundo a direção do vetor h;

$[Z(xi + h) - Z(xi)]$ é o valor da diferença de um par de dados que distam h entre eles.

Geralmente, dados amostrais estão alojados no espaço de forma irregular. Para obter cálculo dos variogramas experimentais são escolhidas classes de distância e de ângulos que irão melhor representar os dados (SOUSA e MUGE, 1990). Para cada direção θ escolhe-se um ângulo de abertura δ, conhecido como *tolerância*, que irá conseguir um maior número de pares de pontos para a direção θ e distância *h.* O aumento da tolerância angular, e também de h, inclusão um maior número de pares de pontos na análise, no entanto, este procedimento é indicado apenas para áreas com pouca densidade amostral (áreas escassamente amostradas).

Quando o conjunto de atributos possui comportamento anisotrópico, o ajuste do modelo teórico do variograma (Figura 3.1) deve ser baseado em duas direções angulares ortogonais: a direção de máxima e mínima continuidade espacial (FELGUEIRAS, 1999).

Nos casos isotrópicos é feito o cálculo do variograma para todas as direções, ou seja, com tolerância angular máxima de 180º, chamado de variograma omnidirecional. No entanto, segundo Soares (2006), na ocorrência de fenômenos anisotrópicos o aumento da tolerância pode sobrestimar as amplitudes das direções

de menor continuidade e, o inverso, subestimar as amplitudes de direções com maior continuidade espacial.

Shinohara *et al*. (2006) afirmam que a modelagem do conjunto de variogramas construídos permite obter um conjunto de parâmetros essenciais para a determinação da estrutura espacial da variável, e subsequente estimação da mesma. Alguns parâmetros importantes do variograma são ilustrados na Figura 3.1 e descritos em seguida.

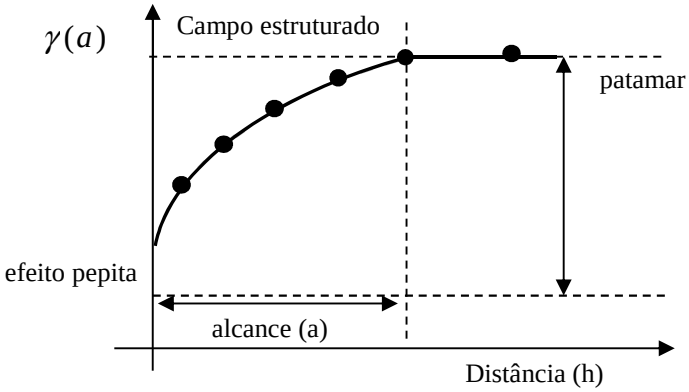


Figura 3.1: Parâmetros do variograma. Adaptado de Shinohara *et al*., 2006.

- efeito pepita C_0 – indica o nível mínimo que o procedimento amostral adotado consegue detectar nas distâncias (h) consideradas. O efeito pepita representa a incerteza à pequena escala, como também a alta variabilidade da variável em distâncias pequenas (SOARES, 2006). Esta descontinuidade junto á origem pode ocorrer devido a erros de medição, ou resultante de uma variável bastante irregular (SOUSA e MUGE, 1990). Segundo estes autores, o efeito pepita está condicionado a dois fatores: i) possível existência de erros de amostragem e ii) microregionalizações não detectáveis à escala de análise. No entanto, é impossível saber se a ocorrência do efeito pepita provém dos erros de medição ou da variabilidade de pequena escala, não captada pelos dados amostrais (CAMARGO *et al.*, 2004).

- patamar – o patamar (*sill*) do modelo 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 entre as amostras, porque a variância da diferença entre pares de amostras torna-se aproximadamente constante (CAMARGO *et al*., 2002).

- alcance ou amplitude (a) – é a distância dentro da qual os dados são espacialmente correlacionados. Corresponde à distância a partir da qual o patamar é atingido (CAMARGO *et al.*, 2002).

- anisotropia (ou efeito de direção) – de modo contínuo, a variabilidade dos dados acontece em uma dada direção particular (SOARES, 2006). Os variogramas são calculados em diferentes classes de ângulos, de maneira a testar a existência desta componente direcional, denominada anisotropia. Nesse caso, escolhe-se a direção que corresponde à maior variabilidade espacial dos dados. Em casos onde há anisotropia a função do variograma varia de acordo a direção (YAMAMOTO e LANDIM, 2013), sendo que o único parâmetro que pode variar entre variogramas é a amplitude, todos os outros se mantêm iguais.

A construção do variograma experimental é um ponto fundamental da variografia, pois esta função matemática deve sintetizar informação relativa à estrutura espacial do fenômeno, influenciando fortemente os resultados da krigagem.

Sousa e Muge (1990) desaconselham ajustes automáticos (do tipo mínimos quadrados), pois desta forma não é possível incorporar o conhecimento prévio do fenômeno.

Após definidos os parâmetros dos variogramas (efeito pepita, patamar e alcance), é necessário ajustá-los através de uma função que represente a continuidade espacial e as características estruturais do fenômeno (SOARES, 2006). Antes da krigagem, os variogramas experimentais são modelados ou ajustados por uma curva (Figura 3.2).

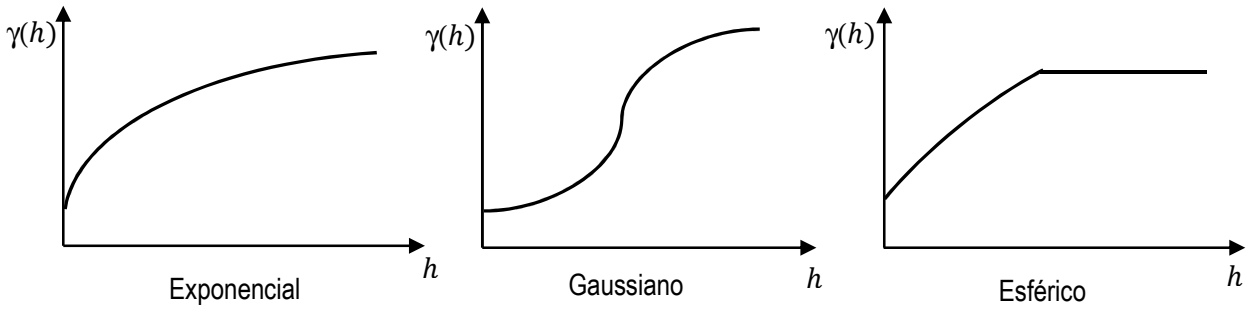


Figura 3.2: Modelos usuais de variogramas

Na geoestatística existem dois tipos de variogramas: com e sem patamar. Em modelos com patamar, o alcance corresponde à distância equivalente a 95% do patamar. Os modelos sem patamar continuam aumentando de acordo com a distância. Estes modelos são utilizados para representar fenômenos que possuem capacidade infinita de dispersão (CAMARGO *et al.*, 2004).

Dos modelos com patamar, um dos mais utilizados na geoestatística é o modelo esférico. Este modelo apresenta comportamento linear junto a origem e ocorre em função de dois parâmetros: o patamar e a amplitude. Ele é o tipo mais utilizado, podendo ser comparado com a distribuição normal na estatística clássica (LANDIM, 2010).

3.4 KRIGAGEM

Na krigagem são encontrados os pesos ótimos que serão relacionados às amostras para estimar um ponto, uma área ou um bloco, com base em informações obtidas a partir do variograma ajustado. De acordo com a posição geográfica dos pontos, diferentes pesos serão associados a eles, já que a função estabelecida no variograma leva em consideração a distância entre locais de amostragem (LANDIM, 2006).

A krigagem leva em consideração tanto a distância entre amostras como o seu agrupamento. Deve-se considerar, no entanto, que para as áreas não amostradas, a krigagem não se faz interpolação de valores tão próximos ou exatos à realidade, mas ela propõe-se a indicar as áreas com maiores possibilidades de ocorrência de certo fenômeno (CAMARGO *et al.*, 2004). No local exato onde há pontos amostrais, o resultados da estimação deve ser os mesmo valor da amostra. Utilizando a krigagem, é possível inferir um valor médio do atributo e sua posição espacial, a partir de uma quantidade *n* de amostras vizinhas (FELGUEIRAS, 1999).

Krigagem é uma técnica da geoestatística de estimação linear que obedece dois critérios em relação ao seu erro: o não-enviesamento e variância mínima da estimação. A primeira condição significa que a média dos valores estimados seja igual à média dos dados observados, de forma que a média de erros seja igual a

zero. Mesmo obedecendo ao primeiro critério, a variância dos valores estimados também determina a qualidade da interpolação espacial (SOARES, 2006).

Para Felgueiras (1999), a geoestatística oferece técnicas de interpolação através da representação dos atributos espaciais, possibilitando também a modelagem da incerteza em relação ao valor estimado de uma variável aleatória conhecida.

Os estimadores clássicos de inferência espacial são métodos deterministas que não possuem medidas de incerteza associadas e não consideram medidas de continuidade estrutural (SOARES, 2006). Para Camargo *et al.* (2004) a krigagem se difere de outros métodos de interpolação devido a maneira como os pesos são atribuídos às amostras distintas. Por exemplo, na interpolação baseada no inverso do quadrado das distâncias, os pesos são definidos como o inverso do quadrado da distância que separa o valor interpolado dos valores observados. Na krigagem, uma análise espacial determina os pesos a partir da análise do variograma ajustado.

Segundo Sousa e Muge (1990), as técnicas de estimação tradicionais não contêm um método que determine a qualidade da estimação. Neste sentido, a geoestatística ou estimadores lineares de krigagem são adequados, já que levam em consideração o conjunto de atributos relacionados à qualidade da estimação local.

Os tipos de krigagem mais utilizados são a Simples e Ordinária. A Krigagem Ordinária (KO) é utilizada quando se admite as médias dos atributos constantes e desconhecidas.

Há alguns fenômenos onde os valores dos atributos não possuem comportamento homogêneo. Para estas situações existem alguns tipos de krigagem que consideram certas características de deriva, ou seja, variações nas médias locais. A krigagem Simples é uma delas, e é utilizada quando o formalismo teórico do modelo atribui o conhecimento da média da função aleatória, ou quando se conhece a tendência (deriva) dos dados, com a possibilidade de assumir estes valores como a média dos valores amostrados e não amostrados (SOARES, 2006).

Este é um tipo de krigagem mais geral, no qual o valor das médias do conjunto de dados aleatórios apresenta caráter não-estacionário, ou seja, não ocorre dependência ou homogeneidade espacial dos dados. A krigagem simples é utilizada

quando se pretende estimar a média global dos valores, já que ela não reproduz de maneira eficaz variações locais.

Outro tipo de krigagem que não trabalha com o fator estacionariedade, ou constância das médias, é a krigagem com deriva externa (KDE). Neste tipo de krigagem, uma variável conhecida pode ser utilizada, através de uma função linear, a fim de estimar outra variável a ela linearmente correlacionada (SOARES, 2006).

Neste trabalho serão utilizadas a Krigagem Ordinária e Krigagem com Deriva Externa como técnicas para estimação e interpolação da geração e viagens por transporte público. Tais técnicas estão descritas nas subseções 3.4.1 e 3.4.2.

3.4.1 KRIGAGEM ORDINÁRIA

A Krigagem Ordinária é o tipo mais usual de krigagem. A KO é uma variante da Krigagem Simples, no entanto, nela assume-se que a média m(x) é constante, porém desconhecida. Pode ser analisada a partir de dados com certa tendência, pois possui técnicas específicas da análise de tendência e anisotropia (JAKOB e YOUNG, 2006). Portanto, pode-se aplicar o algoritmo da Krigagem Ordinária na análise de dados correlacionados espacialmente quando os valores da variável regionalizada apresentam média constante, e possibilitam minimizar a variância do erro de estimação.

O valor desconhecido da variável Z(x) pode ser estimado a partir de uma combinação linear dos *n* valores observados, (JOURNEL, 1988).

$$Z^*_{\mathbf{x}_0} = \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i) \qquad \text{(eq 3.2)}$$

Onde $Z(\mathbf{x}_i)$ representa a quantidade de valores amostrados e λ_i o peso associado a cada um destes valores.

Segundo o mesmo autor, o erro deve ser ajustado para zero, qualquer que seja o valor desconhecido da média, o que significa que as médias sejam iguais:

$$E[Z_{\mathbf{x}_o} - Z^*_{\mathbf{x}_o}] = 0 \qquad \text{(eq 3.3)}$$

A Krigagem Ordinária é um interpolador confiável, no sentido de que os valores interpolados irão coincidir com os valores dos pontos amostrais. Desta forma, a variância da Krigagem Ordinária, garante a confiabilidade dos valores estimados (CAMARGO *et al.*, 2004). Neste tipo de krigagem, a substituição de uma única média estacionária por médias ou tendências locais justifica a vantagem de sua utilização (DEUTSCH e JOURNEL, 1998[3], *apud* FELGUEIRAS, 1999).

3.4.2 KRIGAGEM COM DERIVA EXTERNA

Em alguns casos, quando duas variáveis possuem alta correlação entre si, admite-se que uma variável supostamente conhecida (secundária) pode ser utilizada, através de uma função linear, a fim de estimar outra variável a ela correlacionada (primária) (SOARES, 2006). De acordo com o mesmo autor, a função linear que prevê a deriva entre a variável principal e a secundária pode ser expressa por:

$$m\,(x) \;=\; a_0 \;+\; a_1\, Y(x) \qquad \text{(eq 3.4)}$$

onde,

$m\,(x)$ é o valor da deriva;

$a0$ e $a1$ são os coeficientes implicitamente estimados junto com a variável a ser estimada $Z(x)$;

$Y(x)$ valor da variável secundária.

Assim, a krigagem com deriva externa pode ser aplicada em casos onde a variável primária possui relação de dependência espacial com a variável correlacionada, a secundária (FERNANDES, 2009).

Nesse caso, a variável primária pode ser conhecida em qualquer ponto através do valor da variável secundária (SOARES, 2006). Ou seja, necessariamente, o valor da variável secundária deverá ser conhecido em todos os pontos que se deseja realizar a estimação da variável primária. As duas variáveis, a primária e a secundária, e o

[3] Deutsch, C. V.; Journel, A. G. *GSLIB Geostatistical software library and user's guide.* New York: Oxford University Press, 1998. 369p.

modelo de deriva externa têm que possuir alto grau de dependência linear (ROCHA *et al*., 2009).

Considere-se $Z*(x)$ como a variável primária e $Y(x)$ a variável secundária, ambas apresentando dependência linear, conforme escreve Wackernagel (2010):

$$\mathrm{E}[\mathrm{Z}(x)] = a_0 + b_1\, y(x) \qquad \text{(eq 3.5)}$$

Ainda segundo Wackernagel (2010), o estimador da krigagem com deriva externa, pode ser escrito como:

$$Z^*_{KED}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad \text{(eq 3.6)}$$

com a primeira condição de não enviesamento sendo:

$$\sum_{i=1}^{n} \lambda_i = 1 \qquad \text{(eq 3.7)}$$

e a segunda condição de não enviesamento como:

$$y(x_0) = \sum_{i=1}^{n} \lambda_i\; y(x_i) \qquad \text{(eq 3.8)}$$

Os valores dos ponderadores são obtidos pela resolução de um sistema de equações lineares, conforme:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i\; C_0^r(x_j - x_i) - \mu_1 - \mu_2\, y(x_i) = C^r(x_i - x_0) \text{ para } i = 1, \ldots, n \\ \sum_{i=1}^{n} \lambda_i = 1 \\ \sum_{i=1}^{n} \lambda_i\, y(x_i) = y(x_0) \end{cases} \qquad \text{(eq 3.9)}$$

onde μ1 e μ2 são os multiplicadores de Lagrange e $C^r(x_i - x_0)$ é a covariância centrada dos resíduos, entre a variável primária e a secundária, da j-ésima amostra em relação à i-ésima.

3.5 MÉTODOS DE AMOSTRAGEM

Os métodos de amostragens devem garantir uma boa representação da realidade (GOES *et al*., 1991). Estimativas baseadas em amostras estão sujeitas a um determinado grau de incerteza. Neste aspecto a geoestatística e suas técnicas se destacam, pois, além de realizar a estimativa, é possível avaliar também esta incerteza (YAMAMOTO e LANDIM, 2013). Estes autores apresentam esquematicamente três tipos de amostragem, apresentadas na Figura 3.3: amostragem aleatória simples, estratificada e sistemática. Estes tipos de amostragem serão discutidos a seguir.

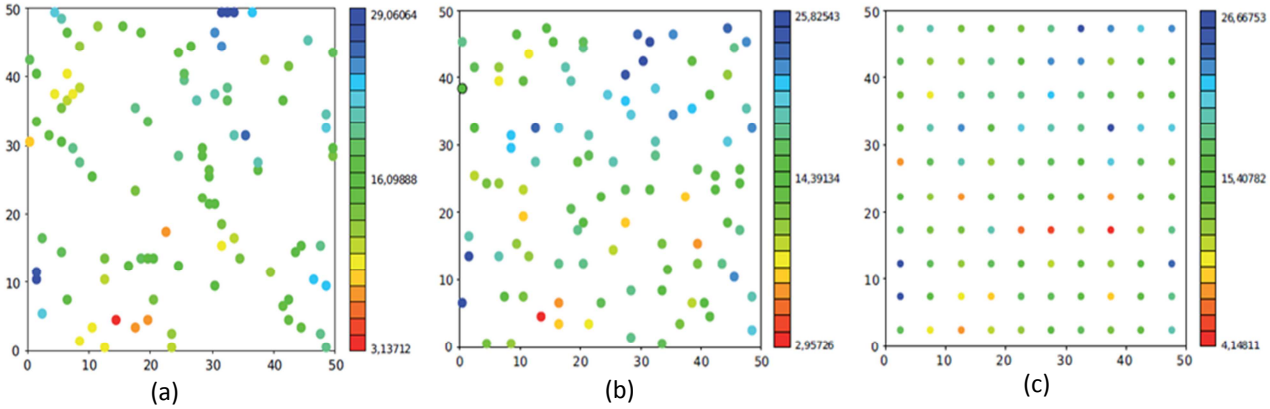


Figura 3.3: Amostragens aleatórias: (a) amostragem simples. (b) amostragem estratificada e (c) amostragem sistemática. Fonte: YAMAMOTO e LANDIM (2013).

Amostragem aleatória simples

Pode ser obtida através de um sorteio, onde todos os elementos do plano de amostragem têm a mesma probabilidade de ocorrência. Por exemplo, em uma população de N elementos, a amostra é obtida por n elementos sorteados sem reposição, sendo n < N. Em geoestatística a componente aleatória traduz-se igualmente pelas coordenadas geográficas selecionadas ao acaso (YAMAMOTO e LANDIM, 2013).

Amostragem aleatória estratificada

Neste tipo de amostragem divide-se a população em subgrupos ou estratos. Feito isto, deve-se proceder como na amostragem aleatória simples, onde as amostras são coletadas aleatoriamente dentro de cada estrato (ROCHA, 2004). Segundo Barbetta (2005), é preciso obter estratos com grupos homogêneos internamente em relação ao conjunto total da população.

A diferença da amostragem estratificada para a amostragem aleatória simples é que se divide a área em subdomínios, e a seleção das amostras pode ser extraída proporcionalmente à sua representatividade em cada domínio (ROCHA, 2004). Esse tipo de amostragem é conhecido como amostragem estratificada proporcional. Nela a proporção de cada estrato da população é considerada na amostragem. Assim, se um estrato corresponder a 10% da população a sua representatividade na amostragem também será de 10% (BARBETTA, 2005).

Amostragem aleatória sistemática

Na amostra sistemática, os intervalos são regulares e definidos antecipadamente (GOES *et al.* 1991). O intervalo de seleção é dado por N/n, onde se escolhe um primeiro número aleatoriamente, para garantir a aleatoriedade da amostragem. As demais amostras são coletadas obedecendo ao intervalo de seleção.

A amostragem sistemática é feita nos nós da malha pré-estabelecida (YAMAMOTO e LANDIM, 2013). Se os dados possuírem em um arranjo aleatório, o resultado da amostragem será igual ao da amostragem aleatória simples. Caso os dados sejam regionalizados, como ocorre em estudos de geoestatística, os resultados serão melhores que a amostragem anterior (ROCHA, 2004).

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 será gerada uma amostragem sistemática artificial, com o intuito de melhorar as estimativas geoestatísticas. Devem-se comparar os resultados obtidos através dos centróides das Macrozonas da Região Metropolitana de Salvador com os resultados obtidos através da amostra gerada artificialmente. O processo de geração da amostra artificial será melhor descrito no Capítulo 4, posteriormente.

Foram utilizados os centróides gerados em uma malha regular ao invés de utilização dos nós da malha.

3.6 TRANSPORTES E GEOESTATÍSTICA: TRABALHOS ANTERIORES

Há alguns anos a geoestatística deixou de ser uma técnica utilizada apenas pelas ciências naturais, como a geologia, e vem se destacando como ferramenta na análise de fenômenos sociais. Os componentes que descrevem o padrão de viagens urbanas podem ser entendidos como parte das ciências sociais, pois no estudo e planejamento de transportes podem ser consideradas características individuais e sociais, como fatores socioeconômicos e variáveis que definem características comportamentais. Assim, os modelos de previsão de viagens urbanas podem ser considerados como parte das ciências sociais (PITOMBO *et al.*, 2010).

Dentre as pesquisas que possuem este enfoque, destacam-se trabalhos desenvolvidos no Mestrado de Engenharia Ambiental Urbana da Universidade Federal da Bahia, inseridos na linha de pesquisa Gestão do Território e Sistemas de Transporte. Os trabalhos fazem parte do projeto de pesquisa "*Aplicação de geoestatística multivariada para análise de geração de viagens: dados desagregados georreferenciados*", no qual este trabalho foi desenvolvido, e tem como participantes outras universidades nacionais e portuguesas. O projeto teve duas etapas definidas, a primeira utilizando dados desagregados (COSTA, 2013), e a segunda etapa utilizando dados agregados da Região Metropolitana de Salvador (RMS) de 1995, onde foi possível também estimar dados agregados de produção de viagens por macrozonas para o ano de 2010, com base em dados censitários de população e renda.

Costa *et al.* (2013) utilizaram a geoestatística para estimar a probabilidade de escolha modal no município de São Carlos (SP). O trabalho utilizou inicialmente a técnica de Árvore de Decisão para conhecer a probabilidade de escolha modal e determinar a variável numérica para a krigagem. Foram estimadas probabilidades de escolha modal em 5.048 mil novas células desconhecidas e observou-se uma tendência espacial de escolha do automóvel, ônibus ou modo a pé condizente com a

realidade socioeconômica de uma região da cidade (bairro Santa Felícia e conjuntos habitacionais).

Pitombo e Sousa (2009) utilizaram a geoestatística para analisar a geração de viagens na Região Metropolitana de São Paulo (RMSP), estimando taxas de produção e atração de viagens por modo de transporte e motivo de viagem. Com essa técnica foi possível conhecer os valores de viagens em 40.000 células desconhecidas, como também mensurar a qualidade da estimação, através da validação cruzada e obtenção de taxas médias de erros e correlação entre valores observados e estimados.

Pitombo *et al.* (2010) utilizaram técnicas de análise de dados espaciais, 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 mesma região. Foi proposta uma metodologia de aplicação conjunta, a Análise de Componentes Principais, para extrair as variáveis independentes e impedir problemas de multicolinearidade e as técnicas de análise espacial de dados. Os resultados foram positivos para as duas técnicas, que apresentaram altos coeficientes de correlação e erros baixos entre valores observados e estimados. Observou-se que a variável emprego está mais relacionada à atração de viagens, bem como outras variáveis como renda alta, influenciam a geração de viagens por transporte individual. Verificou-se que a aplicação da Regressão Geograficamente Ponderada apresenta bons resultados em relação à visualização dos padrões de deslocamentos espaciais de cada uma das variáveis envolvidas. Já com a estimação a partir da krigagem com deriva externa, foi possível estimar os valores das variáveis primárias (produção e atração de viagens) em 40.000 pontos, com base nos valores das variáveis secundárias.

Teixeira (2003) aplica a krigagem para relacionar aspectos socioeconômicos à geração de viagens. Este último autor apresenta uma metodologia para definir áreas homogêneas relacionadas ao planejamento de transportes, 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 de carro utilizando a krigagem universal. A previsão baseia-se na característica que pontos próximos (no espaço de quatro dimensões) terão tempos de viagem parecidos. Tal estimativa baseia-se no fato que a “distância” é proporcional ao tempo e rotas de viagem que possuem origens e destinos próximos. A partir de 200 observações, foi possível verificar que resultados indicam que 95% do intervalo de predição está entre 10 e 30 minutos de viagem entre dois pontos quaisquer. Com a validação dos resultados, verificou-se um coeficiente de correlação de 0.90 entre valores observados e estimados (MIURA, 2010).

Souza (2013) utiliza a geoestatística como técnica de análise espacial para testar a existência de autocorrelação espacial nas análises de geração de viagens, a partir de dados do Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro, publicado em 2004. A partir destas informações foi possível estimar a produção e atração de viagens em qualquer coordenada, considerando os valores de taxas médias das variáveis nos centróides das zonas de tráfego. A Tabela 3.1sintetiza estudos de transportes a partir da aplicação da Geoestatística.

Tabela 3.1: Alguns trabalhos que envolvem Geoestatística e Transporte

Tipo de análise \ Autores	Teixeira (2003)	Silva (2006)	Pitombo e Sousa (2009)	Miura (2010)	Pitombo; Sousa; Birkin (2010)	Costa (2013)	Ferreira e Raia Jr. (2010)	Wang; Zhuang; Liu (2012)	Souza (2013)
Escolha modal			x		x	x			
Estimação de viagens			x		x				x
Tempo de viagem				x					
Zoneamento	x								
Acidentes de trânsito							x		
Impactos ambientais causados pelo tráfego veicular urbano		x							
Variação espacial da velocidade veicular								x	

CAPÍTULO 4: CARACTERIZAÇÃO DA ÁREA DE ESTUDO E DADOS

Este capítulo apresenta as principais características da região de estudo – Região Metropolitana de Salvador (RMS). Também serão descritas as variáveis utilizadas para análise geoestatística, bem como o tratamento e obtenção das amostras finais.

4.1 ÁREA DE ESTUDO

No Brasil, o processo de formação de regiões metropolitanas teve início a partir da década de 1970, como parte da política de desenvolvimento urbano, vinculado à expansão industrial. As regiões metropolitanas brasileiras passaram a ser regulamentadas através de legislação federal e incluíram os principais centros urbanos do país. Estas áreas são regiões de planejamento, formadas por mais de um município, com um núcleo que as nomeia. Elas são alvo de programas especiais, com normas e recursos federais, no entanto, muitas vezes não atendem aos problemas regionais, restringindo-se a aspectos setoriais (SANTOS, 2009).

Sob este contexto, a Região Metropolitana de Salvador (RMS) foi criada pela Lei Complementar número 14, de 8 de junho de 1973. Em 2000, a RMS possuía 10 municípios: Camaçari, Candeias, Dias d'Ávila, Itaparica, Lauro de Freitas, Madre de Deus, Salvador, São Francisco do Conde, Simões Filho e Vera Cruz. Em 2008 foram anexados os municípios Mata de São João e São Sebastião do Passé. No ano seguinte foi incluído o município de Pojuca (RIBEIRO, 2012). Atualmente, a RMS é composta por 13 municípios (Figura 4.1).

Salvador é a terceira capital mais populosa do país, e sua região metropolitana é a sétima em população (RIBEIRO, 2012). A 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, mesmo depois da inclusão de mais três municípios à RMS, Salvador ainda detinha 75% da população da RMS. Além de ser centro financeiro, polo educacional e de saúde, Salvador tem a sua economia voltada ao setor de serviços, a exemplo do comércio, transportes, telecomunicações, turismo e atividades culturais (RIBEIRO, 2012).

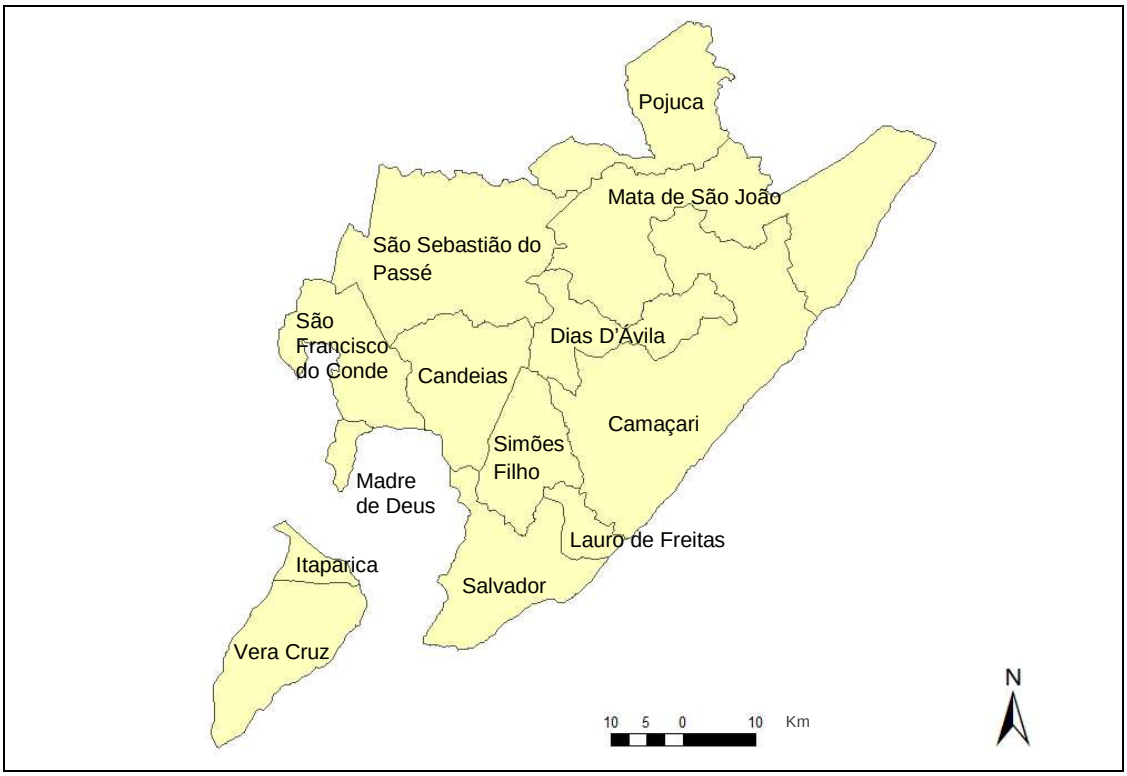


Figura 4.1: Região Metropolitana de Salvador. Fonte: IPEA (2014)

A área de estudo da Pesquisa Origem-Destino (OD) de 1995, banco de dados utilizado no presente trabalho, 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 utilizado. Por isso a importância de analisar a geração de viagens pelo modo de transporte coletivo por ônibus em Salvador (SETPS, 1995).

Esta pesquisa revelou que eram realizadas 3.691.889 viagens diárias em Salvador, sendo que 22% (815.723) concentrava-se no período de pico da manhã, entre 6 e 8 horas. Cerca de 55% destas viagens eram realizadas por ônibus, sendo trabalho e estudo os principais motivos.

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 o mais homogênea possível em relação ao uso do solo e característica da população. A unidade de classificação da pesquisa OD de 1995, são as macrozonas de tráfego (MZs). Estas resultam do agrupamento de zonas menores, que por sua vez são provenientes do agrupamento de setores censitários.

Estas macrozonas foram classificadas em sub-regiões e em quatro grandes regiões: (1) Área Urbana Consolidada (AUC); (2) Orla; (3) Miolo e (4) Subúrbio. O Subúrbio Ferroviário e o Miolo são as áreas mais populosas de Salvador (Figura 4.2), correspondendo juntas a 46% do total da população do município.

No Subúrbio ferroviário também está a população com menor renda (Figura 4.3). Nesta região, a renda mensal dos responsáveis pelos domicílios varia 1,87 a 5,4 salários mínimos (IBGE, 2000).

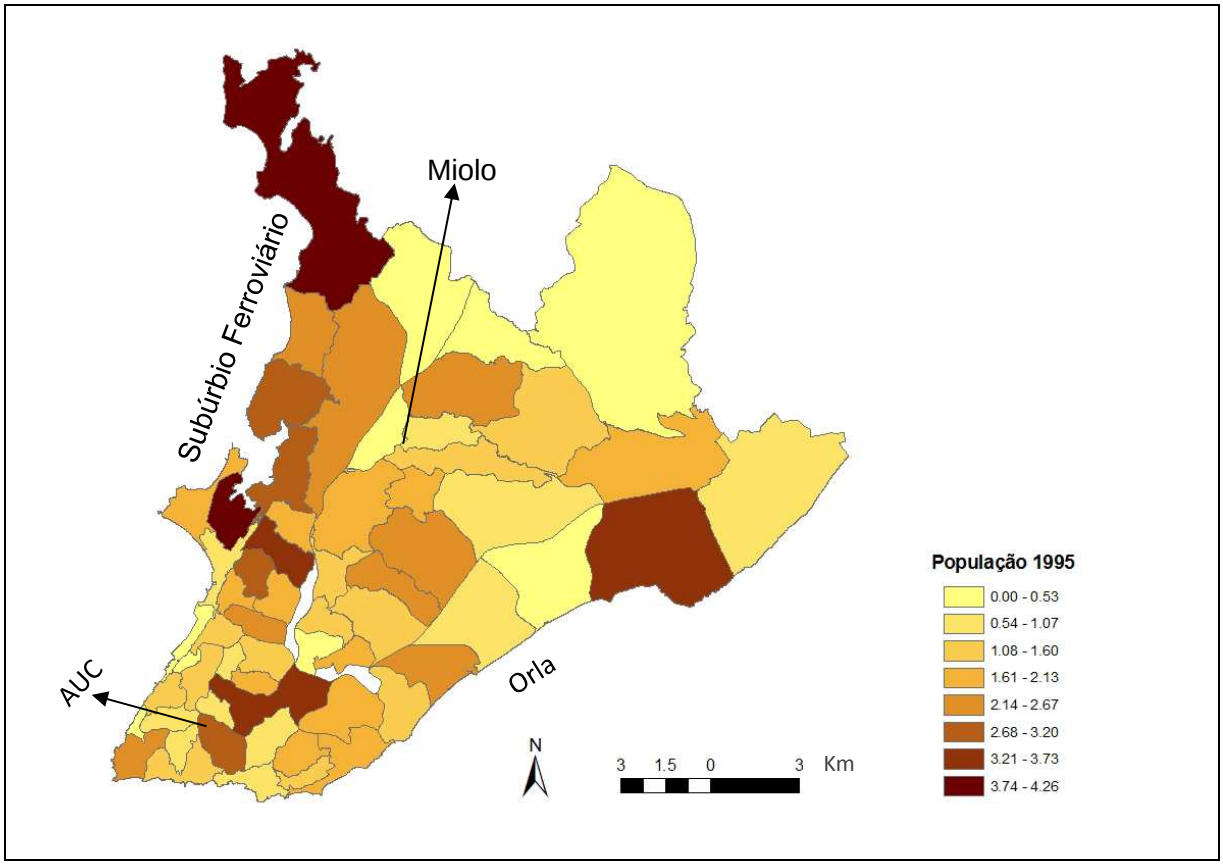


Figura 4.2: População por MZs de tráfego. Fonte dos dados: SETPS (1995)

O nível de renda da população brasileira é um dos fatores principais na escolha do modo de transporte. Portanto, existe uma provável tendência que a maior parte das viagens motorizadas, em áreas que possuem população com renda mais elevada, seja realizada por transporte coletivo.

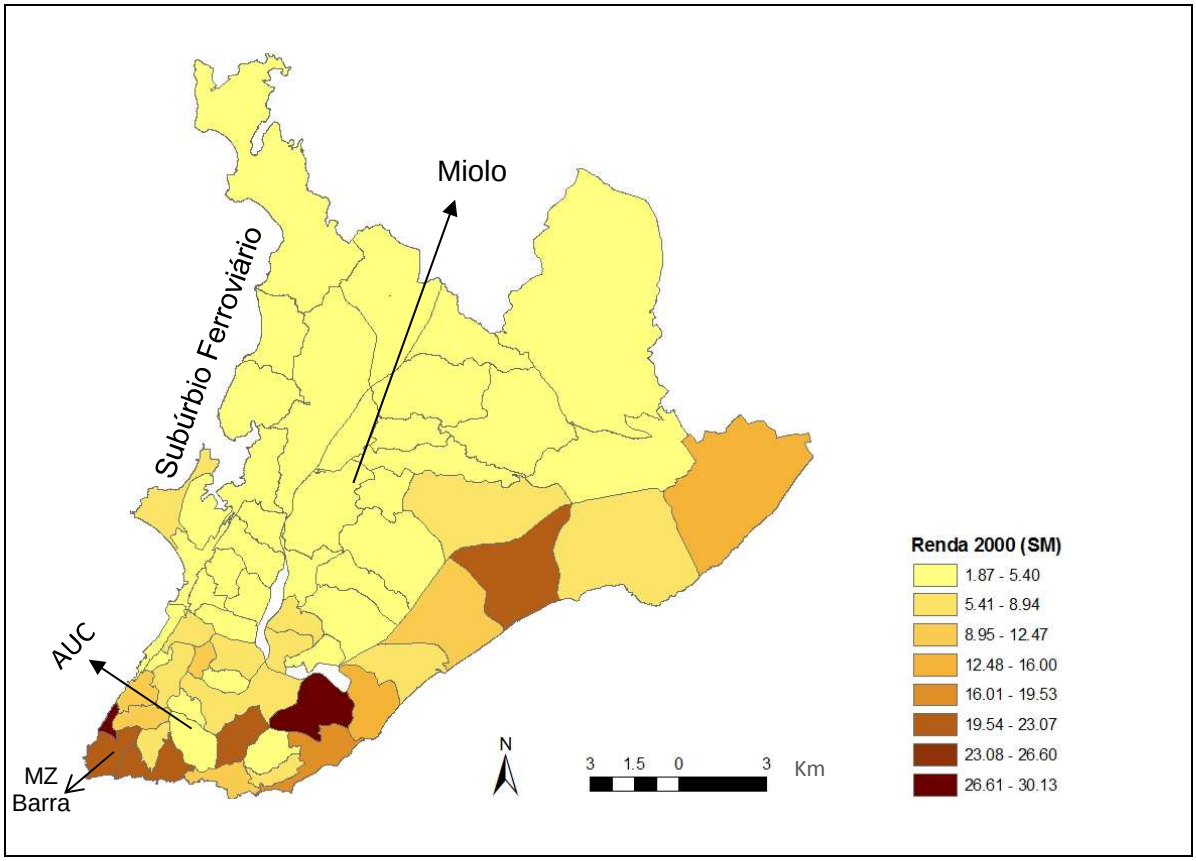


Figura 4.3: Renda média do responsável por domicílio por MZs de tráfego. Fonte dos dados: IBGE (2000)

Em 1995, segundo a Pesquisa OD, a região com população economicamente mais rica em Salvador é a Orla e a AUC, que possuem responsáveis de domicílios com rendimento de até 30 salários mínimos. Na orla e em bairros do centro, a renda média dos responsáveis por domicílio são as mais elevadas de Salvador (Figura 4.3). Nestas áreas localizam-se bairros dotados de maior infraestrutura urbana (como as MZs Barra e Pituba).

As MZs da Orla Atlântica possuíam maior quantidade de veículos por domicílio (Figura 4.4). Estas macrozonas, segundo a pesquisa OD (1995) produziam mais viagens por transporte particular. Juntas, as regiões da Barrra e Pituba possuíam uma taxa de 83% das viagens produzidas por transporte particular (SETPS, 1995). Mesmo assim, no período analisado, o automóvel possuía pouca participação na repartição modal das viagens de Salvador.

A segunda etapa deste trabalho, a estimação da geração de viagens, foi feita a partir de dados de população e renda do censo de 2010. Estes dados mostraram que o padrão espacial da população nas MZs continua similar aos dados de 1995, com

maior quantidade da população na região do Subúrbio Ferroviário, e uma taxa populacional menos expressiva nos bairros do miolo de Salvador (Figura 4.5).

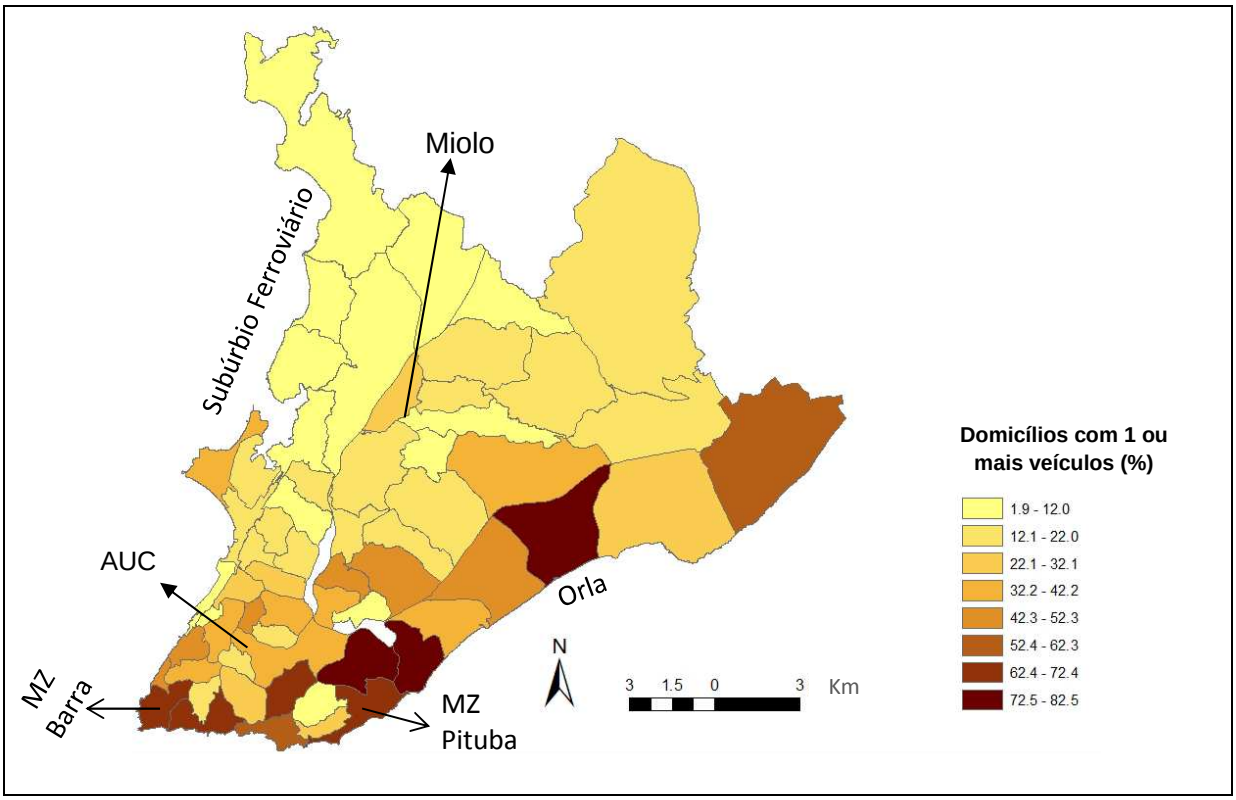


Figura 4.4: Percentual de domicílios que possuem um ou mais veículos. Fonte dos dados: SETPS (1995)

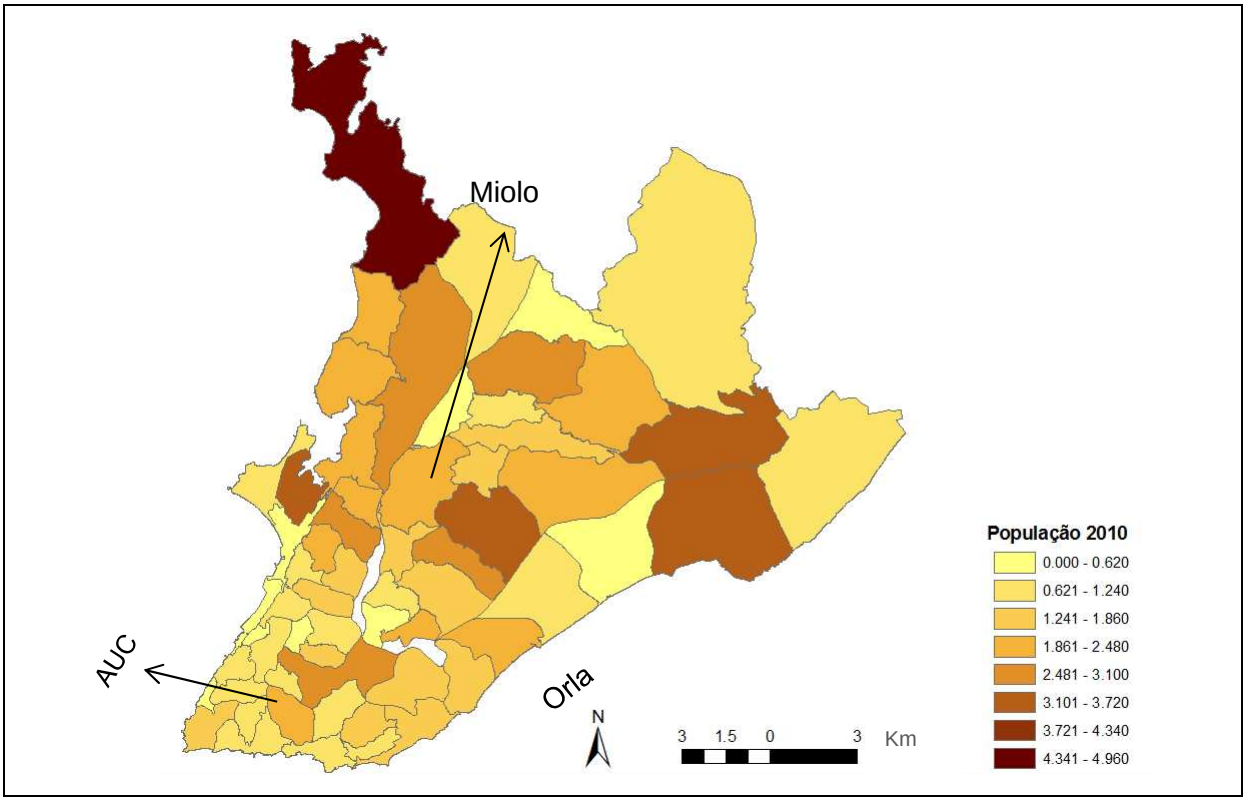


Figura 4.5: População por MZ de tráfego. Fonte dos dados: IBGE (2010)

A segunda etapa deste trabalho, a estimação da geração de viagens, foi feita a partir de dados de população e renda do censo de 2010. Estes dados mostraram que o padrão espacial da população nas MZs continua similar aos dados de 1995, com maior quantidade da população na região do Subúrbio Ferroviário, e uma taxa populacional menos expressiva nos bairros do miolo de Salvador (Figura 4.5). A Figura 4.6 apresenta a distribuição da renda em Salvador em 2010. A partir desta figura é possível verificar que a orla, especialmente a MZ de Patameres, possui a maior concentração de renda da cidade. O inverso se observa na região Miolo e Subúrbio ferroviário.

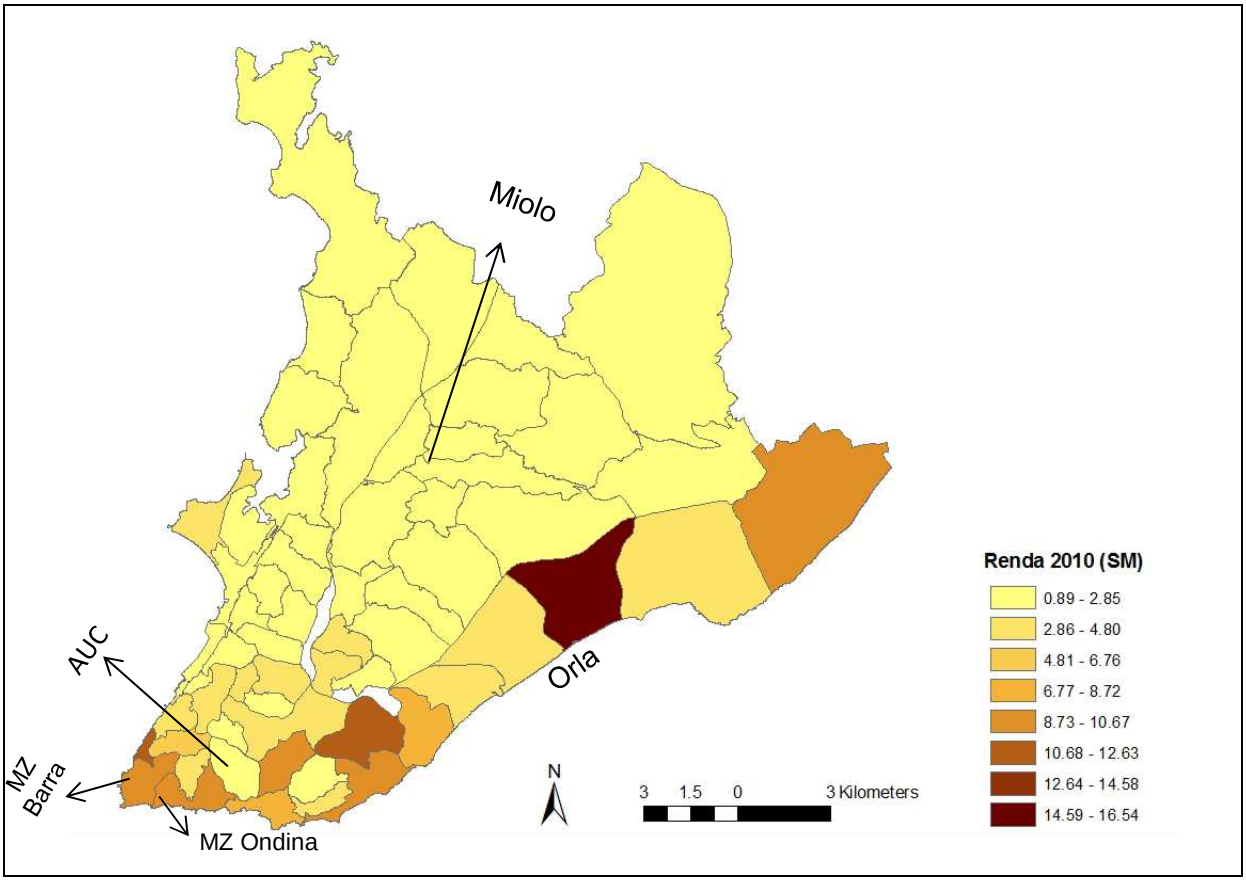


Figura 4.6: Renda média do responsável por domicílio (em salários mínimos) nas MZ de tráfego. Fonte dos dados: IBGE (2010).

4.2 TRATAMENTO DOS DADOS E OBTENÇÃO DAS AMOSTRAS

4.2.1 Pesquisa OD 1995: Dados agregados (Etapa 1)

Os dados da pesquisa OD da RMS foram agrupados em 63 macrozonas de tráfego. Estas macrozonas são resultado do agrupamento de 173 zonas. Este processo foi

orientado pela equipe técnica do Sindicato das Empresas de Transporte de Passageiros de Salvador (SETPS), que dividiu o espaço urbano de Salvador de acordo com suas características gerais de uso e ocupação do solo, renda, além dos fatores geográficos e viários. Estas macrozonas se agrupam somando 13 sub-regiões maiores. A sub-região Miolo Norte é que possui maior quantidade de MZs, um total de 10. Já a sub-região Subúrbio Pirajá é constituída por apenas uma macrozona. As demais sub-regiões são formadas em média por cinco MZs. A Tabela 4.1 descreve as variáveis agregadas utilizadas neste estudo. Para a análise geoestatística, as coordenadas geográficas dos centroides das macrozonas foram associadas às variáveis socioeconômicas citadas na Tabela 4.1. A Figura 4.7, em seguida, mostra a subdivisão das MZs de tráfego da RMS em 1995.

Tabela 4.1: Exemplos de dados socioeconômicos agregados por MZ. (Fonte: SETPS, 1995)

Variáveis	Descrição
População Residente por MZ	População total residente por MZ em 1995
Empregos por MZ	Número total de empregos por MZ em 1995
Produção de viagens por TCO por MZ	Apresenta o total de viagens produzidas por Transporte Coletivo por Ônibus por MZ
Atração de viagens por TCO por MZ	Apresenta o total de viagens atraídas por Transporte Coletivo por Ônibus por MZ

As variáveis descritas acima foram padronizadas. Optou-se pela por este método devido à alta variabilidade dos valores absolutos (Equação 4.1).

$$z = \frac{\chi - \mu}{\sigma} \qquad \text{(eq. 4.1)}$$

onde,

z é o valor padronizado da variável;

χ é o valor absoluto da variável;

μ é a média aritmética dos dados observados;

σ é o desvio padrão dos dados observados.

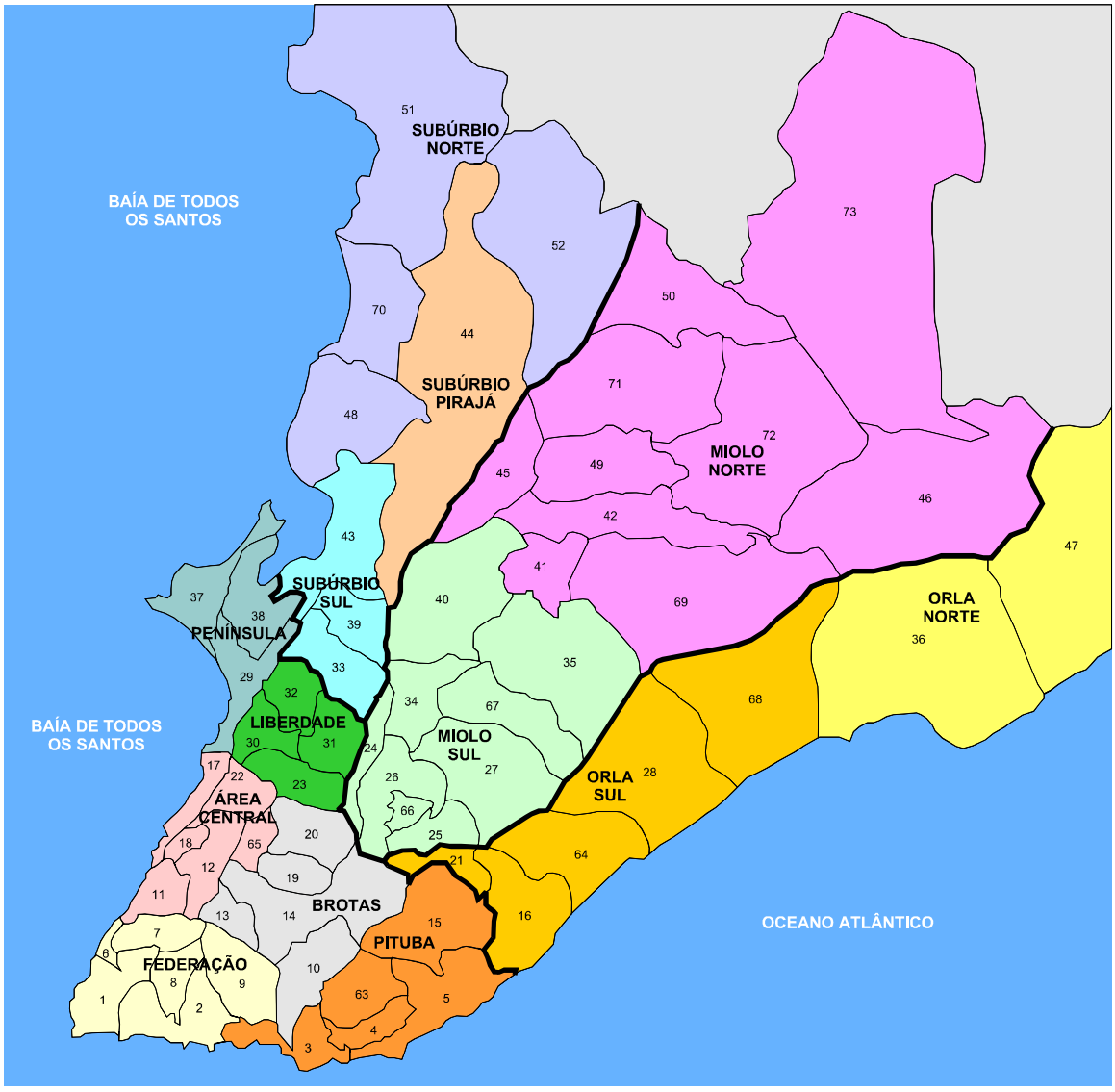


Figura 4.7: Sub-regiões e Macrozonas (Fonte: STEPS, 1995)

O banco de dados utilizados para estimação geoestatística é formado por 63 macrozonas, as coordenadas geográficas (UTM) correspondentes aos centróides das respectivas MZs e pelos dados padronizados referentes às variáveis produção de viagens por TCO e população, atração de viagens por TCO e empregos, variáveis altamente correlacionadas (coeficiente de *Pearson* equivalente a 0,93 e 0,89, respectivamente).

A Tabela 4.2 mostra exemplo de alguns dados de entrada e as variáveis utilizadas em números absolutos e padronizados. Além das variáveis analisadas, o banco de dados e composto pelas coordenadas UTM dos centróides das MZs (IDs).

Tabela 4.2: Exemplo do banco de dados de entrada

X	Y	ID	População	Produçao TCO	População Z	Produçao TCO Z
551453	8562590	1	50699	9347	2,139	2,177
552992	8562160	2	27551	4539	1,138	1,042
555434	8561660	3	22015	4010	0,899	0,917
557172	8561870	4	37921	8174	1,587	1,900
558542	8562530	5	46196	6013	1,944	1,390
551469	8563580	6	10977	1989	0,422	0,440
552551	8563750	7	22753	3482	0,931	0,793
552916	8562860	8	21648	4830	0,883	1,111
554321	8562900	9	65752	12327	2,790	2,880

Após a conversão dos dados em números padronizados, o procedimento de análise espacial dos mesmos foi realizado no GeoMs (*Geostatistical Modelling Software*). Os mapas finais foram editados através do Arcgis 9.3. A aplicação da Krigagem com auxílio do GeoMs e os resultados obtidos são descritos no Capítulo 5 – Aplicação das técnicas de Geoestatística.

4.2.2 Tratamento dos dados do censo de 2010: compatibilização de áreas – Etapa 2

Na segunda etapa do trabalho utilizaram-se dados de população e renda do responsável da família, divulgadas pelo IBGE (2010). Inicialmente, foram compatibilizados os setores censitários, referentes ao último censo, com as MZs de tráfego da Pesquisa OD de 1995. A análise das duas áreas e sua compatibilização foi feita através do *software* Arcgis 9.3. 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 a partir do percentual da diferença do salário mínimo de 2010 em relação a 1995. A técnica de Regressão Linear Múltipla foi utilizada para prever a produção de viagens por Transporte Coletivo para 2010, através das variáveis população e renda média do responsável por domicílio. O modelo foi calibrado com a amostra de dados correspondente à Pesquisa OD de 1995, sendo utilizadas as variáveis população e renda média por domicílio para as MZs de tráfego. A equação calibrada com os dados de 1995 corresponde à equação 4.2.

$$Y = 0.177 \times \textit{População} - 0.218 \times \textit{Renda} \qquad \text{(eq 4.2)}$$

A Tabela 4.3 apresenta os parâmetros estatísticos resultantes do modelo de regressão linear para a estimação da produção de viagens. O coeficiente de determinação do modelo é bastante elevado, e a variável população possui maior influência no resultado do valor estimado da produção de viagens por TCO.

Tabela 4.3: Resultados do modelo de regressão linear

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

Com os valores das variáveis independentes para o ano de 2010, com renda devidamente corrigida, foram estimados os valores de produção de viagens por TCO para 2010. Para a correção da renda em relação ao valor do 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 (Tabela 4.4).

Tabela 4.4: Evolução do SM no período de 1995 à 2010

ANO	VALOR DO SM	TAXA DE AUMENTO (%)
1995	R$ 100,00	0
1996	R$ 112,00	12,00
1997	R$ 120,00	7,14
1998	R$ 130,00	8,33
1999	R$ 136,00	4,62
2000	R$ 151,00	11,03
2001	R$ 180,00	19,21
2002	R$ 200,00	11,11
2003	R$ 240,00	20,00
2004	R$ 260,00	8,33
2005	R$ 300,00	15,38
2006	R$ 350,00	16,67
2007	R$ 380,00	8,57

2008	R$ 415,00	9,21
2009	R$ 465,00	12,05
2010	R$ 510,00	9,68
	Total de aumento	173,33

Fonte: Tribunal Regional do Trabalho – TRT

Assim, o valor da renda de cada MZ de 2010 foi dividido pela taxa de ajuste do SM (1,73). Desta forma foi possível encontrar o valor da renda corrigida para as MZs de tráfego em 2010, de forma a igualar o valor da moeda (ou o poder de compra do SM) para os dois períodos analisados. A Tabela 4.5 mostra uma síntese da correção da renda em 2010.

Tabela 4.5: Exemplo do ajuste da renda para 2010

MZ	Renda do responsável (2010)	Renda ajustada (renda de 2010/ 1.7)
29	985.93	579.96
33	689.08	405.33
37	1738.41	1022.59
38	774.83	455.78
39	795.67	468.04
43	548.40	322.58
48	691.14	406.55
51	593.82	349.30
70	730.88	429.92

Os valores estimados para o ano de 2010 não puderam ser comparados com valores observados porque o censo de 2010 não divulga dados de geração de viagens por TCO. Observou-se que a variável renda é inversamente proporcional à variável que se estimou (produção de viagens por TCO). Com isso, percebeu-se que onde há dados de renda elevada o número de viagens por TCO é mais baixo, e vice-versa. Esta etapa possibilitou estimar a produção de viagens nos centróides das MZs para o ano de 2010.

Assim como os dados da pesquisa OD de 1995, os valores obtidos através da Regressão Linear foram padronizados. A partir disso foram aplicadas técnicas de Geoestatística, como a Krigagem com Deriva Externa, onde foi possível prever a produção de viagens (variável primária) através de população (variável secundária), para o ano de 2010, além da estimação de viagens e população através da Krigagem Ordinária.

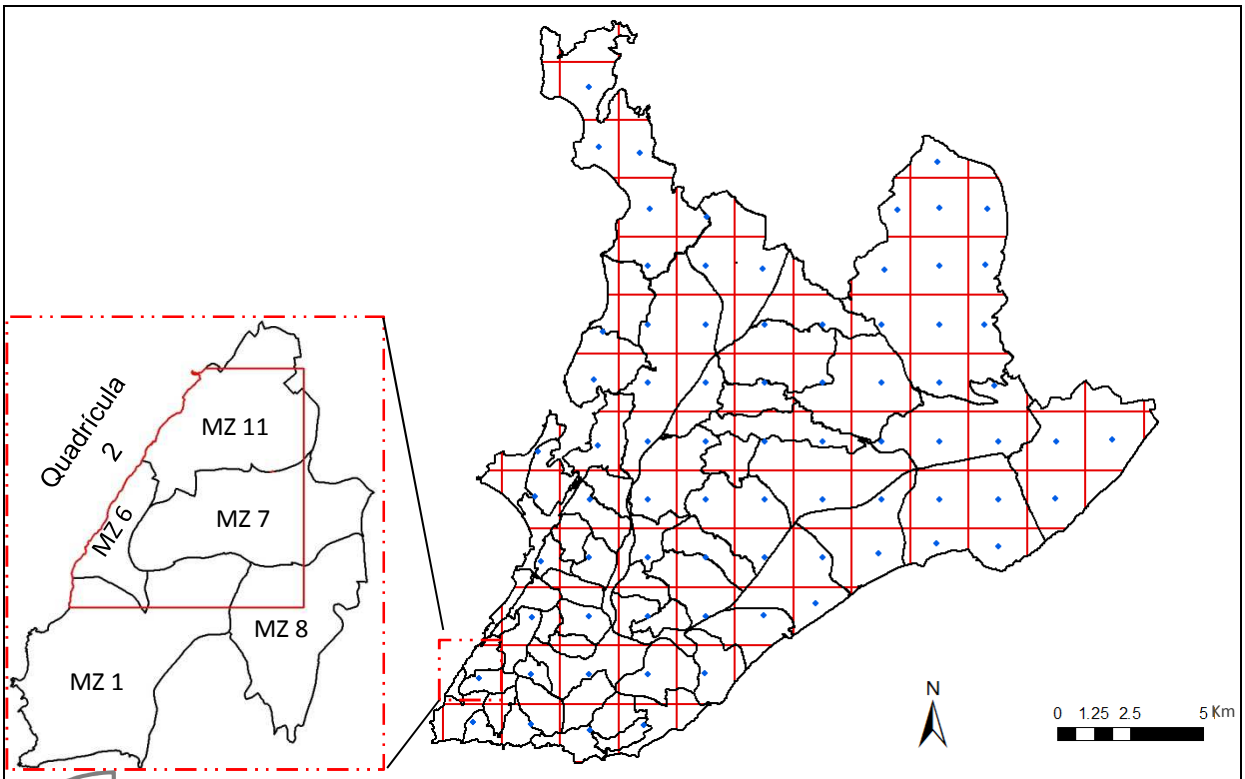
Houve apenas a estimação e interpolação espacial da produção de viagens para o ano de 2010, já que dados de população são facilmente encontrados nos censos demográficos. Além disso, pesquisas deste tipo não oferecem dados sobre empregos (variável altamente correlacionada com a atração de viagens), dificultando sua estimação.

4.2.3 Obtenção da amostragem sistemática artificial – Etapa 3 e Etapa 4

As análises anteriores mostraram que a estimação por krigagem a partir dados analisados através de centróides das MZs gera baixos valores de correlação entre valores observados e estimados. Então a amostra optou-se por utilizar um tipo de amostragem que gerasse um banco de dados mais desagregado. Com isso optou-se por gerar uma amostragem sistemática artificialmente, obtida através de uma malhar regular.

Esta malha regular foi espaçada de 2000 x 2000m (Figura 4.9). Em seguida, foram gerados centróides provenientes das células originadas pela malha. O valor das variáveis (produção, atração, empregos e população) em cada célula da foi obtido através da porcentagem (peso) que a MZ ocupa dentro de cada célula da malha. Estabeleceu-se, então, uma média ponderada dos valores das variáveis, considerando a ocupação de cada célula em diferentes macrozonas (Figura 4.8), e estabeleceu-se o novo valor para cada uma das variáveis, associados aos centróides da malha.

Após esta etapa, foi possível aplicar as técnicas Geoestatísticas, krigagem ordinária e krigagem com deriva externa, para uma amostragem mais densa (dados desagregados), em relação aos centróides das MZs.



Área da quadrícula (Km²)	ID (MZ), Equivalente a 1 célula da malha	Área ocupada pela MZ em cada célula (Km²)	Peso (%)	POP Z por MZ	Valor da variável após compatibilização	Média ponderada por célula
2,975	1	0,424	0,142	2,14	0,305	1,15
	6	0,366	0,123	0,42	0,052	
	7	0,974	0,328	0,93	0,305	
	8	0,191	0,064	0,88	0,057	
	11	0,997	0,335	1,31	0,439	

Figura 4.8: Exemplo da obtenção da malha regular e amostra sistemática artificial da área de estudo.

Este artifício com o intuito de gerar artificialmente dados desagregados tem como objetivo principal aprimorar as estimativas geoestatísticas. No entanto, algumas ressalvas devem ser feitas.

- Variáveis de transportes muitas vezes são discretas, com mudanças de valores abruptas. Neste trabalho os valores estão associados a centroides artificiais, desprezando-se a continuidade espacial de valores de variáveis;

- Esta técnica de geração de dados desagregados considera que haja continuidade nos valores das variáveis, o que não é exatamente o caso para geração de viagens, população e empregos;
- Os dados devem ser preferencialmente gerados através de técnicas de simulação e não necessariamente, média ponderada de valores das macrozonas;
- Salienta-se que se trata de uma tentativa inicial de melhorar a modelagem geoestatística a partir de dados muito agregados, como no caso deste trabalho.

Estas ressalvas são extremamente importantes, no entanto, os pontos amostrais iniciais já são, de certa forma, criações artificiais, assim como dados agregados pelos resultados dos censos e por MZs, aos quais foram atribuídas coordenadas aos centróides das macrozonas. Com isso, conclui-se que amostragem inicial já provém de uma amostra também artificial. Portanto, estas ressalvas não se aplicam apenas à amostragem sistemática. Assim, os mesmos problemas mencionados podem ocorrer em ambas as amostragens. Outra importante consideração é que, uma vez padronizados os dados originais, tem-se para a amostragem sistemática uma segunda padronização dos dados.

CAPÍTULO 5: APLICAÇÃO DAS TÉCNICAS DE GEOESTATÍSTICA

Este capítulo aborda a aplicação das ferramentas de interpolação espacial, a Krigagem Ordinária e a Krigagem com Deriva externa. A aplicação das duas técnicas permite a estimação de dados em coordenadas desconhecidas, a partir de dados amostrados nos centróides das áreas analisadas.

5.1 ETAPAS PARA MODELAGEM GEOESTATÍSTICA

Segundo Câmara *et al.* (2002) são necessários os seguintes passos para análises geoestatísticas que empregam técnicas de krigagem:

- análise exploratória dos dados;
- análise estrutural (modelagem da estrutura espacial da variável regionalizada);
- interpolação estatística da superfície (krigagem).

A seção seguinte trata da primeira etapa para modelagem geoestatística, a análise exploratória de dados. A análise estrutural (modelação dos variogramas) será descrita em seguida. Também serão descritos os parâmetros principais adotados para interpolação espacial. Os principais resultados estão apresentados no Capítulo 6 deste trabalho.

Neste trabalho estima-se a geração de viagens por transporte coletivo por centróide de MZ de tráfego e através de uma amostra de pontos obtida através de uma grade regular – amostra sistemática artificial. Utilizou-se o Geostatistical Modelling Software – GeoMs – para a estimação geoestatística para as diferentes análises. Este *software* foi desenvolvido em 1999 pelo CERENA, Centro de Recursos e Ambiente, Departamento de Minas e Georrecursos do Instituto Superior Técnico (Universidade Técnica de Lisboa - atual Universidade de Lisboa). O GeoMs é composto por diferentes módulos: neste trabalho foram utilizados o geoDATA, geoVAR, geoMOD, geoKRIG, GeoKed, geoVIEW e geoTRDATA.

5.2 ANÁLISE ESPACIAL EXPLORATÓRIA – GEODATA

O entendimento do comportamento das variáveis, bem como a visualização de possíveis padrões espaciais são etapas realizadas no módulo geoData do *software* geoMS. Nesta etapa, conhecendo o valor padronizado (Z) da geração de viagens por TCO, população e empregos por MZ, considerou-se como referência o centróide georreferenciado de cada MZ, além dos centróides da grade regular, também com coordenadas conhecidas, para estimar valores para áreas não amostradas.

Com isso, foi possível visualizar a distribuição espacial das amostras através dos mapas de pontos, descritos a seguir. Vale ressaltar que além do *software* GeoMs, foi utilizado o SIG Arcgis 9.3 para gerar os mapas de distribuição espacial dos pontos.

Outra ferramenta utilizada no módulo geoData foi a análise da estatística descritiva univariada e bivariada das amostras, onde foi possível analisar a distribuição de frequência dos dados, através de histogramas e *box-plot*, sendo possível analisar também a dispersão e assimetria dos dados. Alguns parâmetros estatísticos que podem ser analisados são: média, variância, Coeficiente de *Pearson* e desvio padrão.

A seguir, se encontram descritos os principais resultados da análise exploratória de dados para geração de viagens por TCO, quantidade de empregos e população.

5.2.1 População e Produção de viagens por TCO 1995 (amostragem original por centróides das MZs)

A população é um fator que impulsiona fortemente a produção de viagens (Bruton, 1979). Assim, observou-se que as variáveis população e produção de viagens por TCO (pico da manhã) apresentavam o mesmo padrão espacial e alta correlação.

A partir da análise do mapa de pontos no geoData verificou-se a direção que apresenta maior variabilidade dos dados para as variáveis população e produção de viagens de por TCO (Figura 5.1 e Figura 5.2, respectivamente). Através destas figuras, que representam os centróides das macrozonas das variáveis, foi possível

identificar maiores e menores valores de população e produção de viagens em determinadas regiões.

A partir da análise da distribuição espacial dos dados foi possível verificar as áreas que apresentam maiores valores da variável regionalizada (população e produção de viagens) e perceber que as mesmas localizam-se 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.

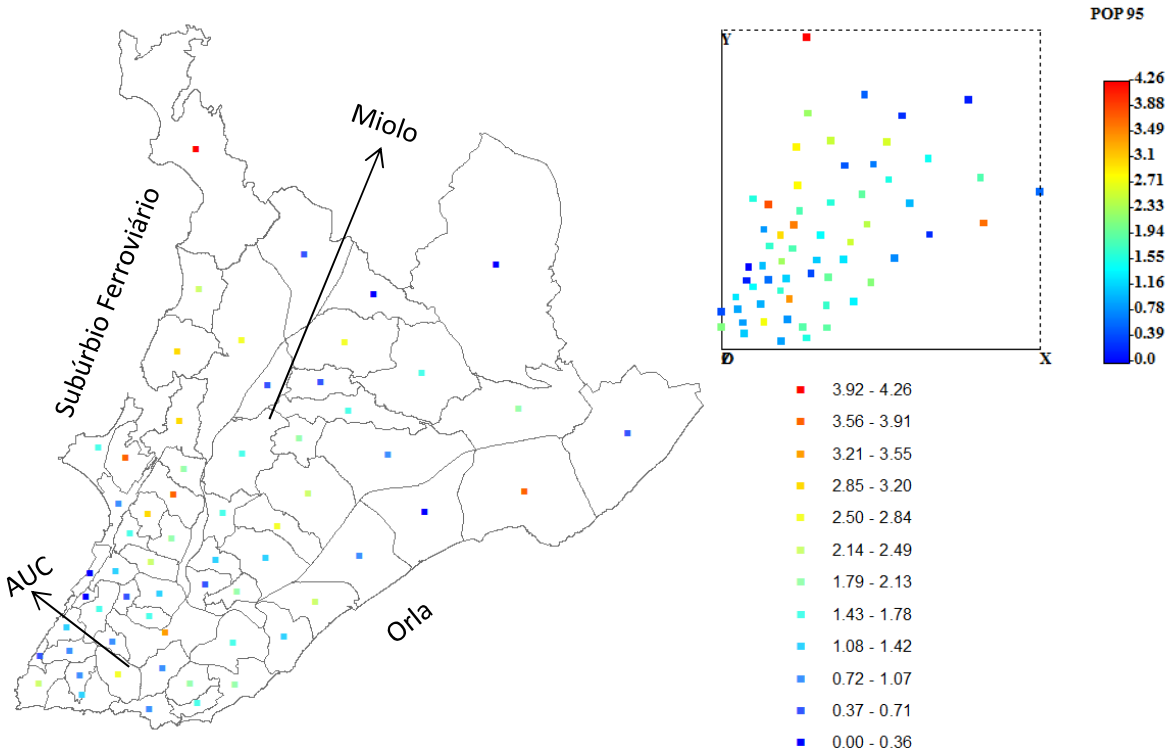


Figura 5.1: Mapa de pontos da variável população (1995)

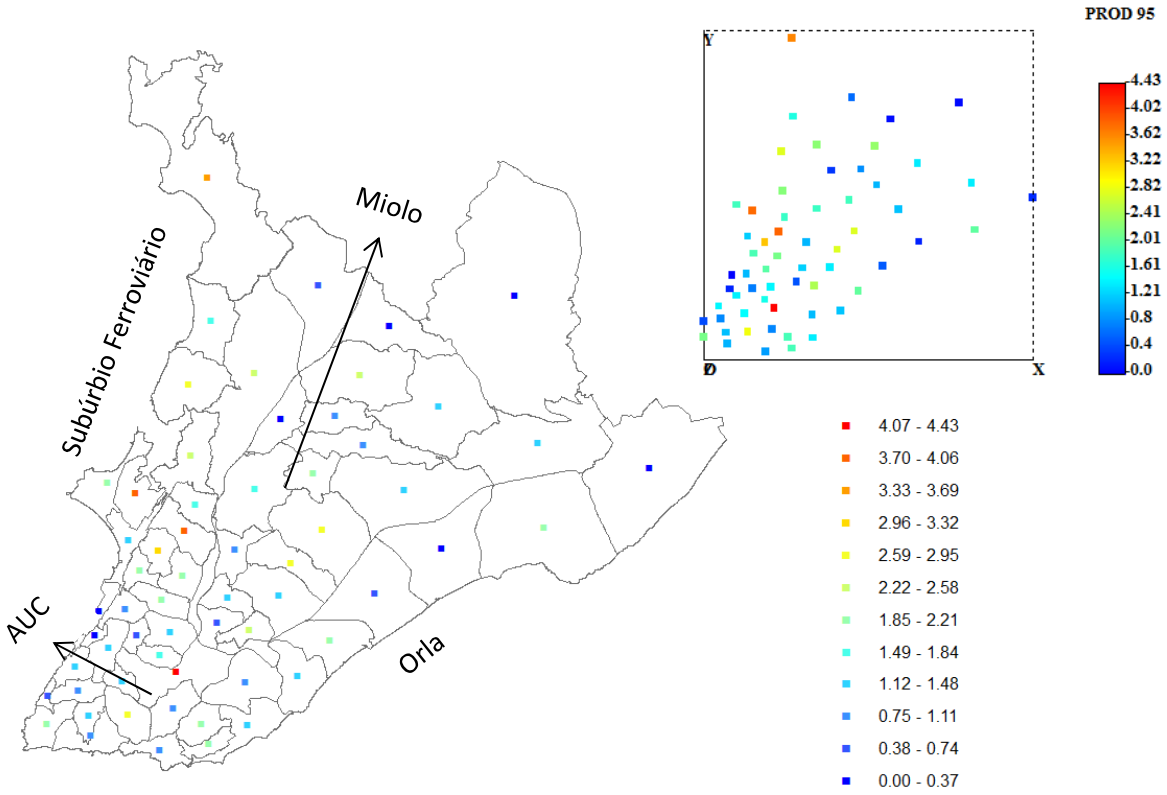


Figura 5.2: Mapa de pontos da variável produção de viagens por TCO (1995)

Foi obtido o diagrama de dispersão e coeficiente de *Pearson* entre as duas variáveis, bem como entre as variáveis empregos e atração de viagens por TCO. É ainda possível verificar que a correlação é maior para os valores baixos e intermediários, havendo uma pequena dispersão considerando os valores mais elevados. Os diagramas entre valores observado dos pares de variáveis população x produção de viagens e empregos x atração de viagens estão ilustrados no Apêndice A deste trabalho.

A partir da análise da estatística descritiva das variáveis população e produção de viagens por TCO, apresentadas na Tabela 5.1, pode-se notar a similaridade dos dados típicos entre o par de variáveis e a dispersão dos valores.

Tabela 5.1: Estatística descritiva das variáveis população e produção de viagens por TCO (1995)

Variáveis	Média	Variância	Desvio padrão	Correlação
Produção de viagens	1,57	0,98	0,99	0,93
População	1,62	0,98	0,99	

5.2.2 Empregos e Atração de viagens por TCO 1995 (amostragem original por centróides das MZs)

A partir da análise exploratória dos dados, foi possível verificar que os dados de empregos e atração de viagens também são altamente correlacionados, apresentando coeficiente de *Pearson* igual 0,89, além de distribuição espacial dos pontos similar (Figura 5.3 e Figura 5.4).

Inicialmente, foram visualizados os mapas de pontos para tais variáveis, e verificou-se que em 1995 as MZs com maior concentração de empregos em Salvador eram Campo Grande e Comércio, áreas conhecidas como Centro Antigo. As MZs Iguatemi, Pituba e Brotas também possuem certo destaque, apresentando-se como as áreas com maior quantidade de empregos (Figura 5.3). Já as demais áreas apresentam valores baixos em relação a empregos e também à atração de viagens por transporte coletivo.

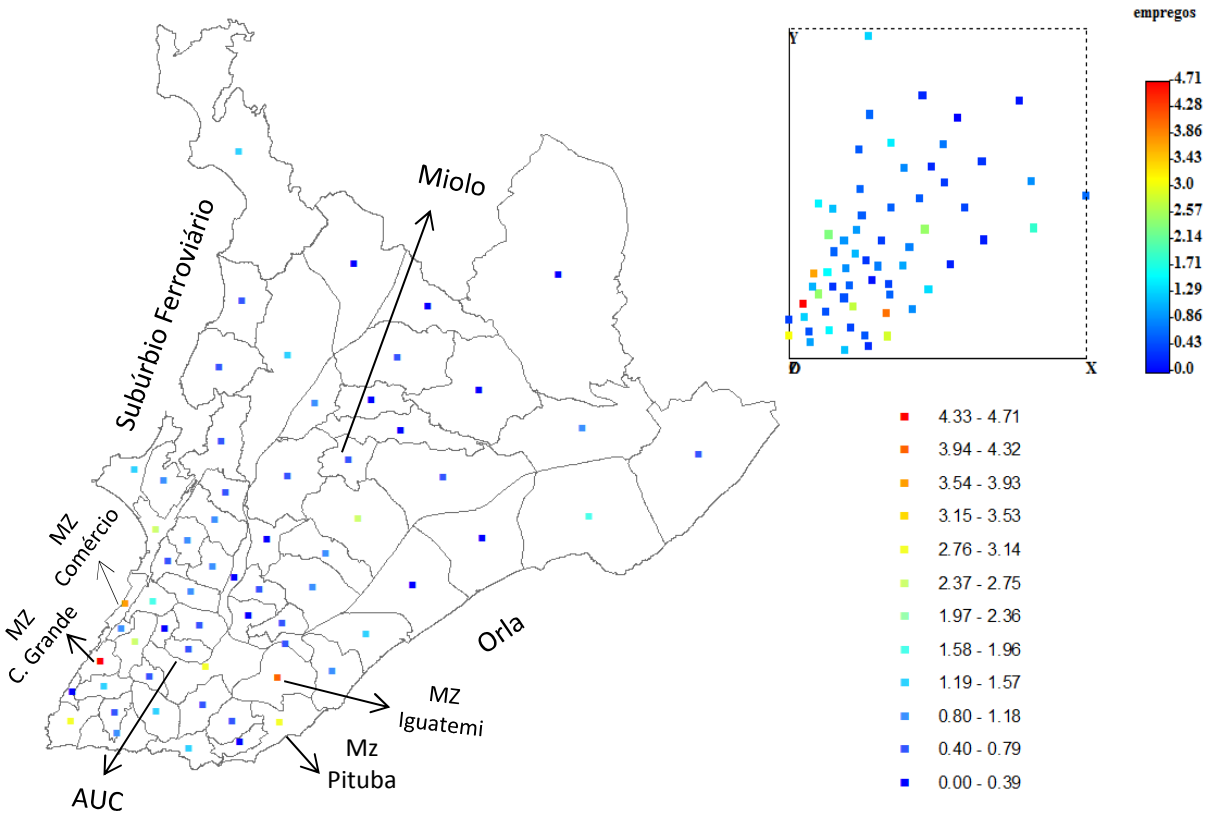


Figura 5.3: Mapa de pontos da variável empregos (1995).

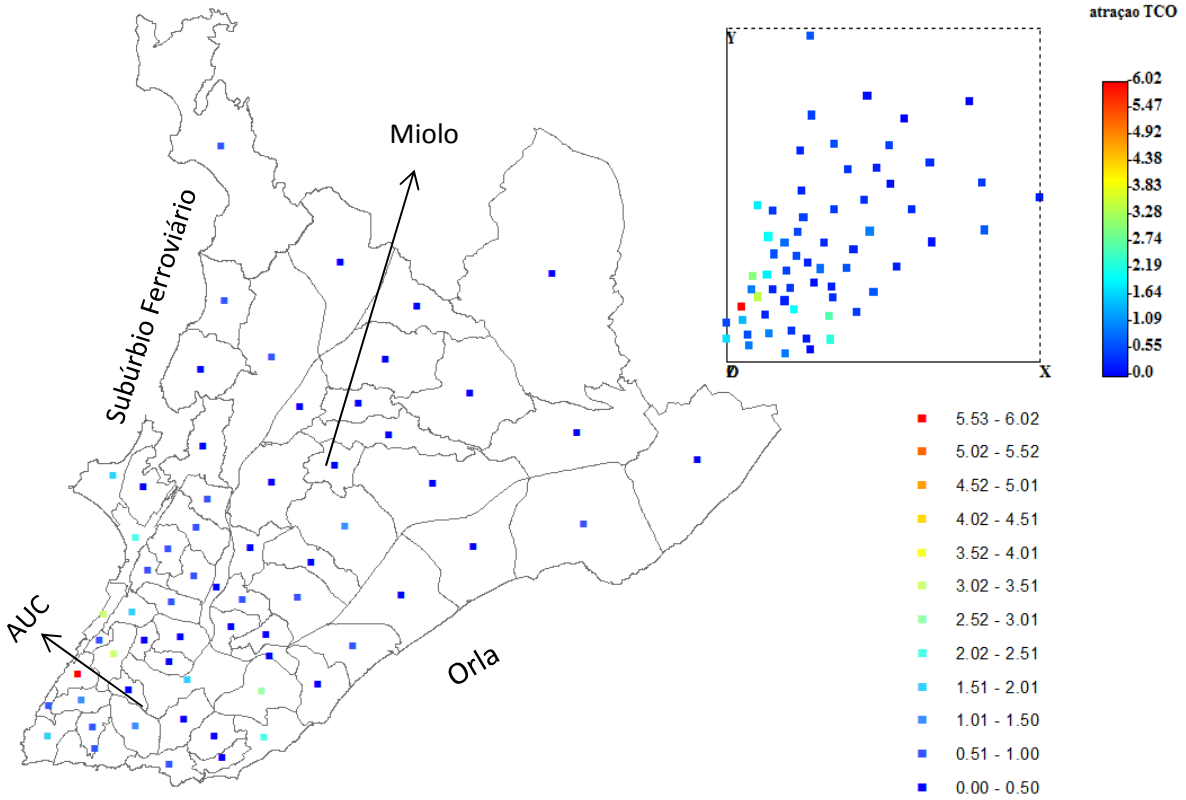


Figura 5.4: Mapa de pontos da variável atração de viagens por TCO (1995).

A análise do resumo estatístico permite deduzir que os dados apresentam pouca simetria, o que significa que há discrepância entre os mesmos, apesar da média aritmética estar em torno de 1,0 para as variáveis citadas. No entanto, há um ponto máximo e equivalente a 6,02 para atração de viagens, e 4,71 para empregos. A Tabela 5.2 apresenta o resumo das estatísticas descritivas. Apesar dos dados possuírem certa assimetria, esta tabela mostra a correlação entre as duas variáveis analisadas, já que possuem valores bem próximos.

Tabela 5.2: Estatística descritiva das variáveis empregos e atração de viagens por TCO (1995)

Variáveis	Média	Variância	Desvio padrão	Correlação
Atração de viagens	0,83	0,98	0,99	0,89
Empregos	1,07	0,98	0,99	

Os valores discrepantes estão muito acima da média dos dados nas duas variáveis. Segundo Yamamoto e Landim (2013), o tipo de distribuição e frequências dos dados influenciam os variogramas experimentais. Portanto, um variograma que venha a

apresentar uma falta de estruturação pode ser consequência de alta assimetria dos dados de uma variável.

5.2.3 População e Produção de viagens por TCO 1995 (amostragem sistemática artificial)

A amostra sistemática artificial foi realizada para tentar obter melhores resultados na estimação da geração de viagens, a partir de uma relativa “desagregação” dos dados. Esta subseção mostra as principais análises estatísticas para este tipo de amostragem, para as variáveis população e produção de viagens por TCO. As Figuras 5.5 e 5.6 mostram a distribuição espacial dos pontos de tais variáveis. Estas figuras revelam que a distribuição espacial das variáveis continua similar aos pontos analisados por centróides das MZ, apresentando, porém, uma quantidade maior de dados e diminuição da distância entre estes. Na Tabela 5.3 está a estatística descritiva dos dados.

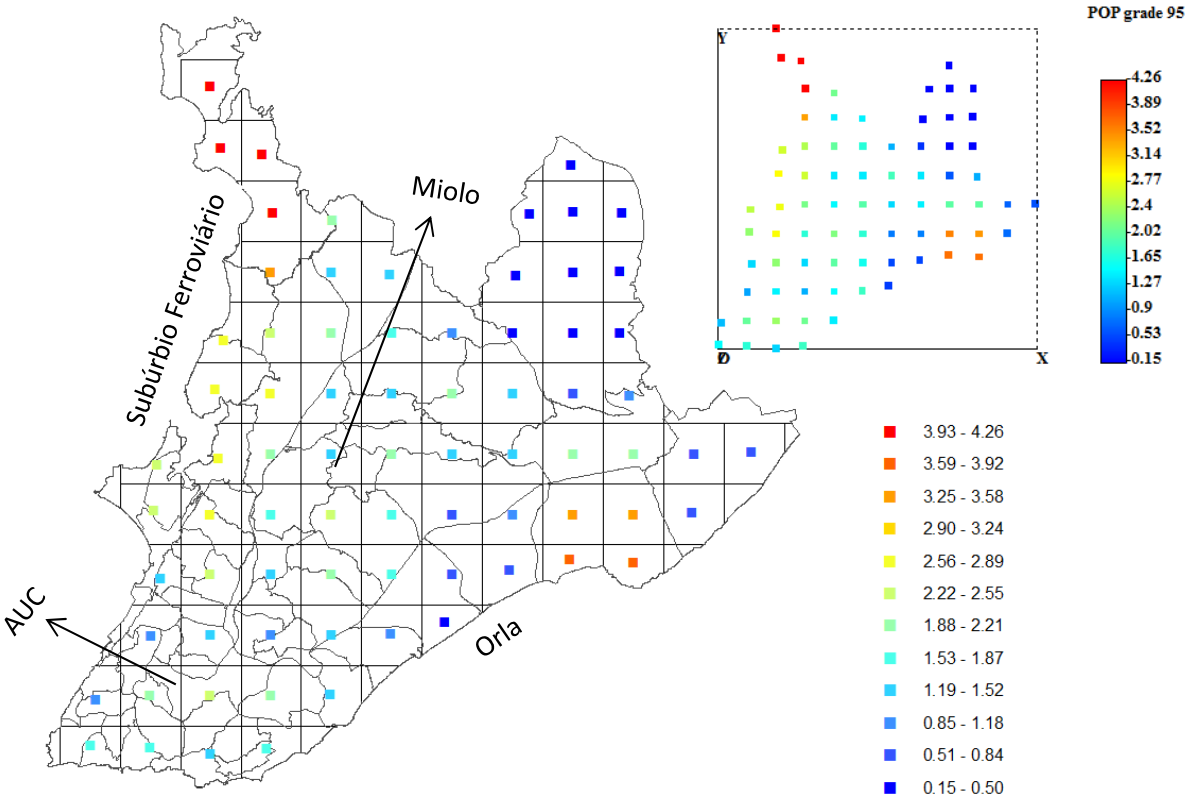


Figura 5.5: Mapa de pontos da variável população 1995 (amostra sistemática artificial)

A distribuição espacial das variáveis com base nos centróides das macrozonas é irregular, levando a espaçamentos entre amostras muito pequenos a sul e muito grandes a norte, que colocam lado a lado valores baixos e altos das variáveis em estudo. A utilização desta técnica permite uma homogeneização do espaçamento entre as amostras e permite suavizar as diferenças abruptas entre os valores.

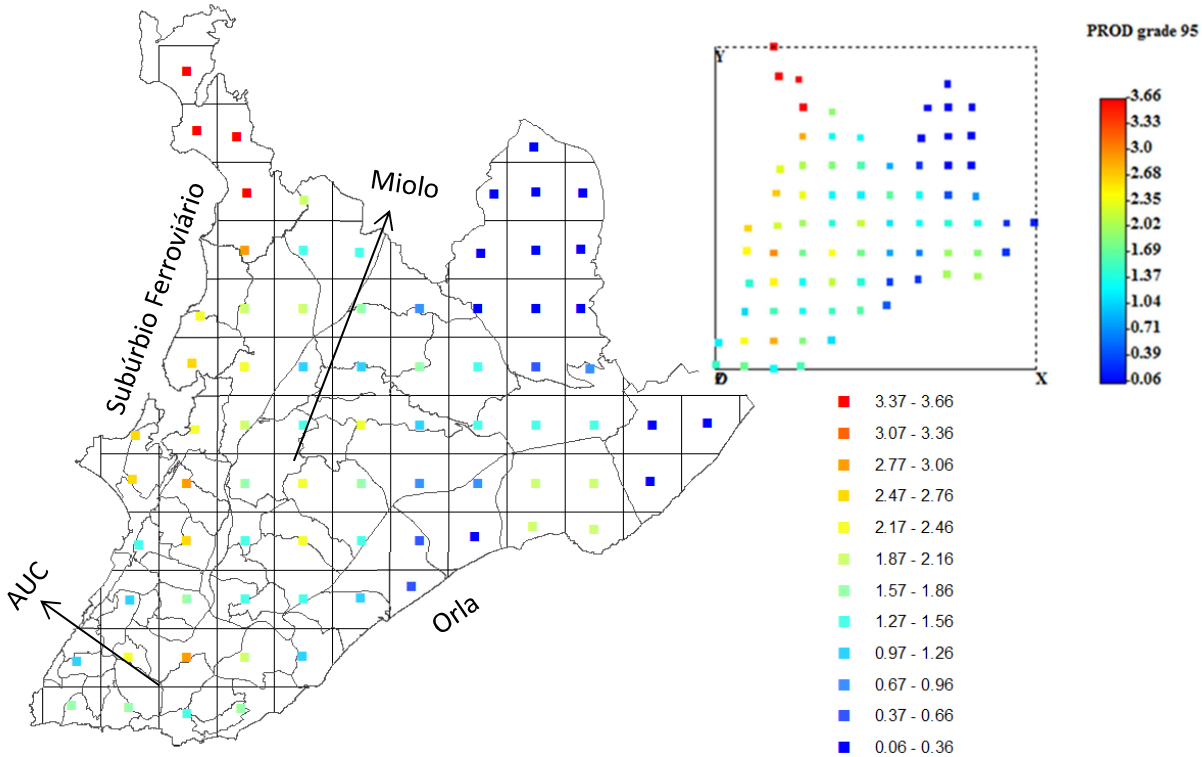


Figura 5.6: Mapa de pontos da variável produção de viagens por TCO 1995 (amostra sistemática artificial).

Tabela 5.3: Estatística descritiva das variáveis população e produção de viagens por TCO 1995 (malha regular)

Variáveis	Média	Variância	Desvio padrão	Correlação
Produção de viagens	1,52	0,88	0,94	0,98
População	1,74	1,17	1,08	

5.2.4 Empregos e Atração de viagens por TCO 1995 (amostragem sistemática artificial)

A amostra sistemática artificial demostrou, também neste caso, algumas diferenças em relação aos pontos obtidos através dos centróides das MZs. A Figura 5.7 e a

Figura 5.8 ilustram a dispersão espacial das variáveis empregos e atração de viagens por TCO, respectivamente.

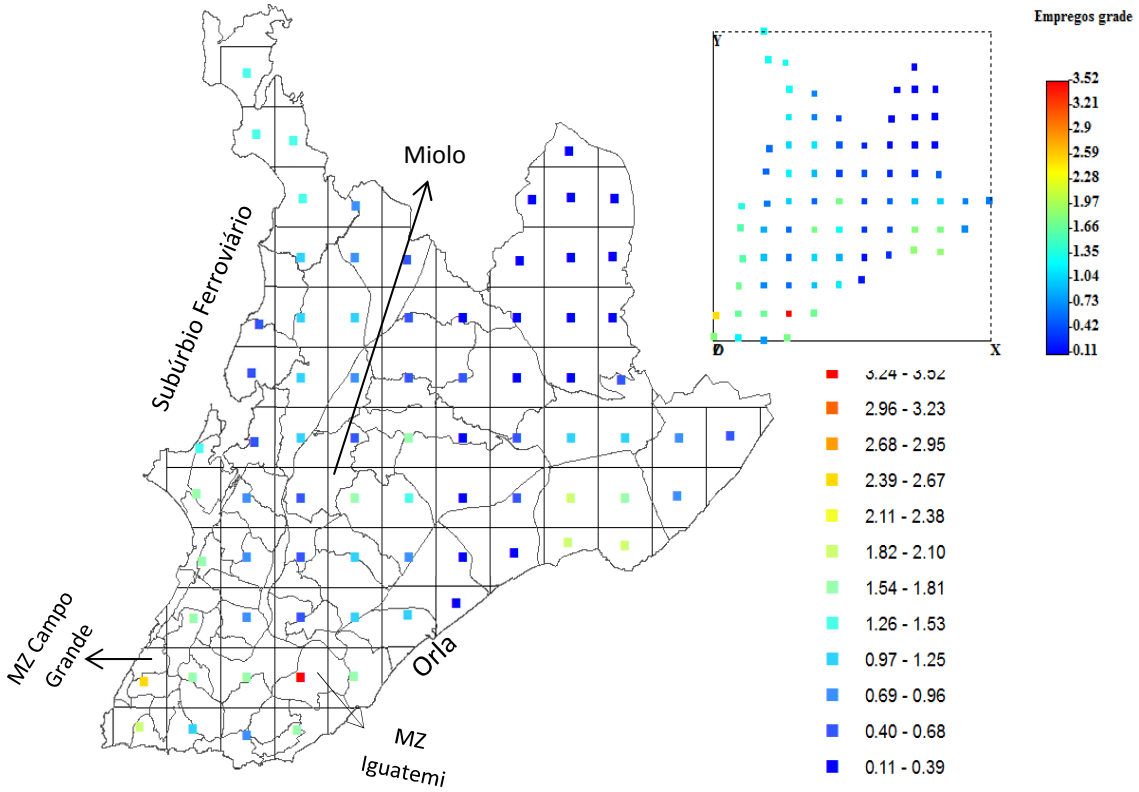


Figura 5.7: Mapa de pontos da variável empregos 1995 (malha regular)

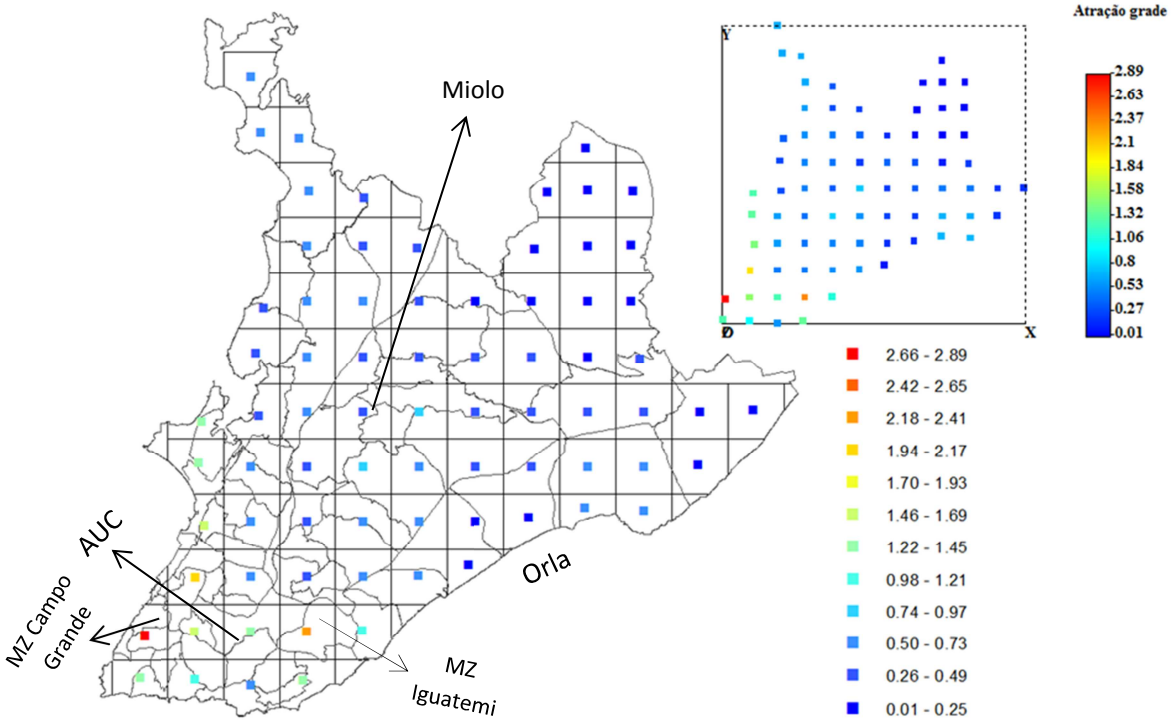


Figura 5.8: Mapa de pontos da variável atração de viagens por TCO (malha regular)

A Tabela 5.4 apresenta o resumo estatístico das variáveis empregos e atração de viagens. Apesar da elevada correlação entre elas, o valor da média, variância e desvio padrão da variável empregos são mais elevados, principalmente porque a taxa de empregos para algumas áreas, como já citados na subseção da amostragem por MZs, são bem mais elevadas que as taxas de atração de viagens.

Tabela 5.4: Estatística descritiva das variáveis empregos e atração de viagens por TCO (malha regular)

Variáveis	Média	Variância	Desvio padrão	Correlação
Atração de viagens	0,57	0,28	0,53	0,87
Empregos	0,93	0,44	0,66	

5.2.5 População e Produção de viagens por TCO 2010 (amostra por MZ)

Conforme descrito no capítulo 4, os dados do último censo foram compatibilizados com as MZs de tráfego da Pesquisa OD de 1995. Com isso foram obtidos os valores de população associados a áreas equivalentes a MZs para 2010 e estimados dados de produção de viagens para as referidas MZs, através de modelo linear de geração de viagens calibrado com dados da Pesquisa O/D de 1995. Tais dados podem ser visualizados nas Figuras 5.9 e 5.10.

A Tabela 5.5 apresenta a estatística descritiva destas variáveis. Nesta tabela é possível notar que os resultados são iguais para as duas variáveis, já que a variável produção de viagens foi estimada a partir dos dados de população e renda.

Tabela 5.5: Estatística descritiva das variáveis população e produção de viagens por TCO (2010)

Variáveis	Média	Variância	Desvio padrão
Produção de viagens	1,55	1,0	1,0
População	1,55	1,0	1,0

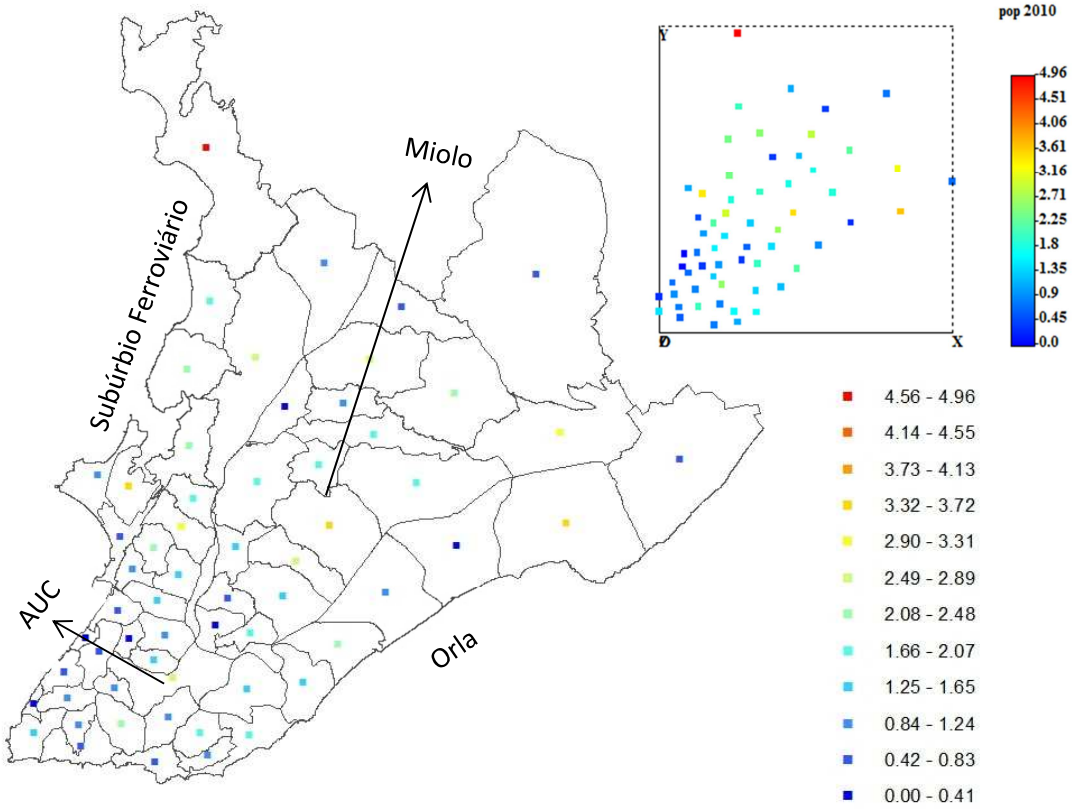


Figura 5.9: Mapa de pontos da variável população (2010).

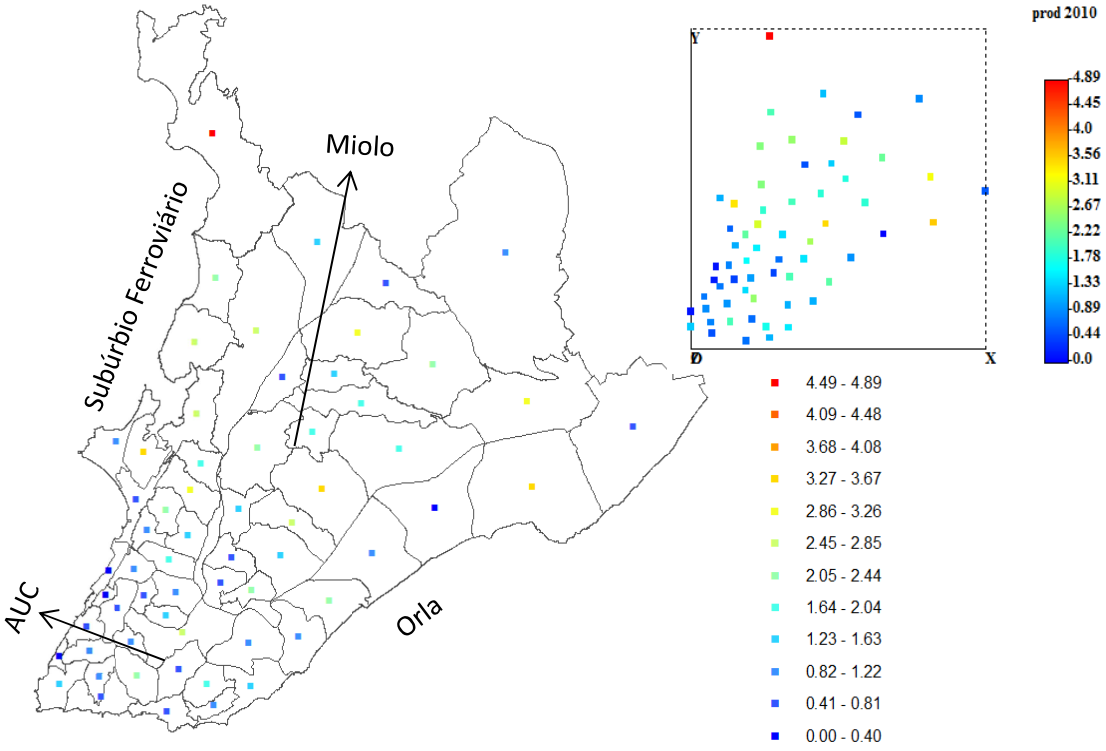


Figura 5.10: Mapa de pontos da variável produção de viagens (2010)

5.2.6 População e Produção de viagens por TCO 2010 (amostra sistemática artificial)

As Figuras 5.11 e 5.12 mostram os mapas de pontos obtidos através da amostragem sistemática artificial para os dados de 2010 (população e produção de viagens por TCO, respectivamente).

A análise destas figuras demostra que a distribuição espacial das variáveis continua similar aos pontos analisados por centróides das MZs para 2010. Os resultados da estatística descritiva apontam a alta correlação entre População e Produção de viagens para amostras sistemáticas artificiais, como foi verificado para os dados de 1995.

A descrição estatística destas duas variáveis (Tabela 5.6) demonstra a forte relação entre o número de residentes de uma MZ e a quantidade de viagens produzidas, bem como ocorreu com os dados observados na Pesquisa OD de 1995, para o mesmo tipo de amostragem. Os dados de produção de viagens para o ano de 2010 são provenientes de um modelo linear, sem constante e com alto coeficiente para a variável independente população.

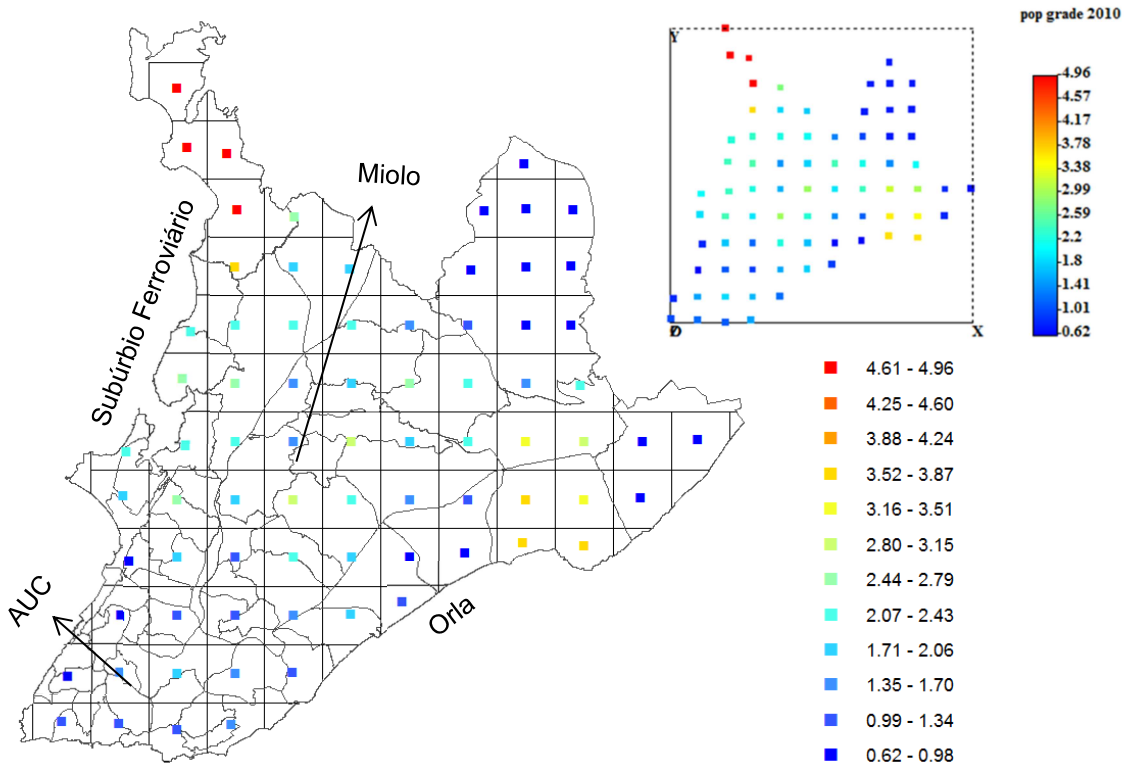


Figura 5.11: Mapa de pontos da variável população 2010 (amostra sistemática artificial).

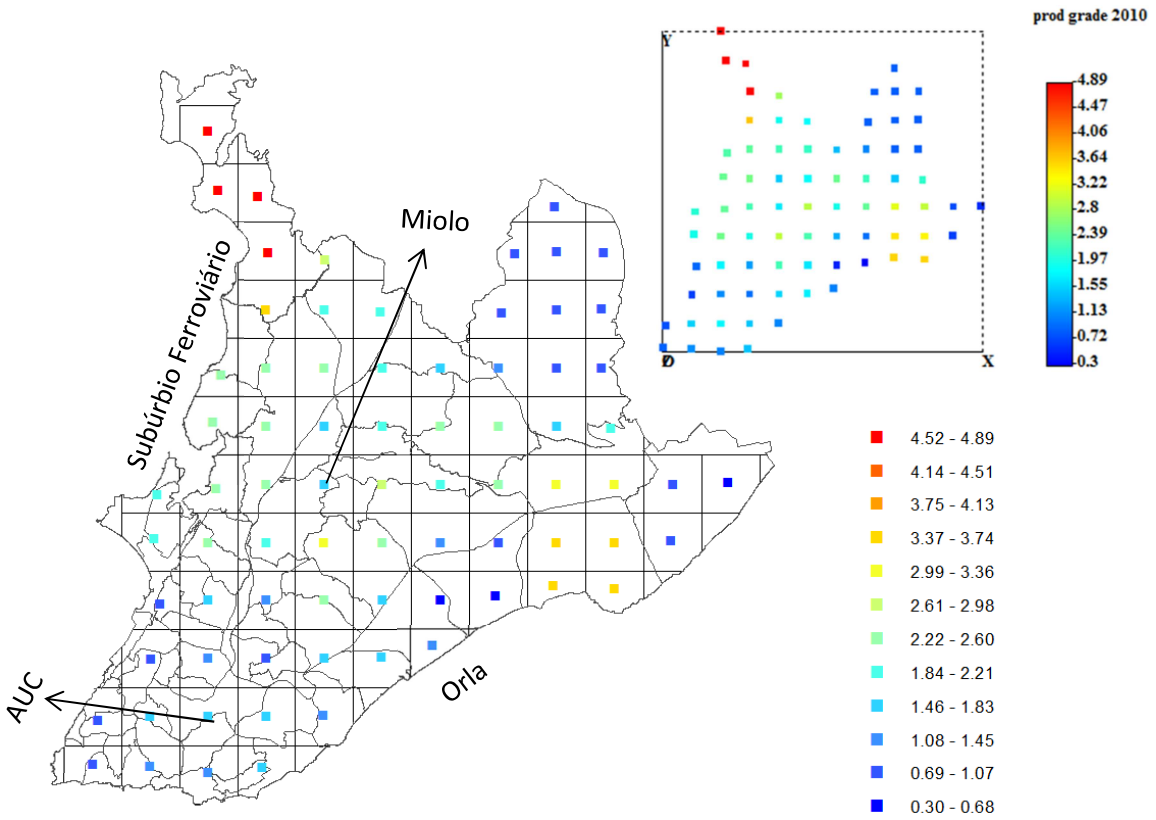


Figura 5.12: Mapa de pontos da variável produção de viagens por TCO 2010 (amostra sistemática artificial).

Tabela 5.6: Estatística descritiva das variáveis população e produção de viagens por TCO 2010 (malha regular)

Variáveis	Média	Variância	Desvio padrão
Produção de viagens	1,93	1,17	1,08
População	1,93	1,19	1,09

Apesar das melhoras na modelagem geoestatística, obtidas através da utilização de dados desagregados, é importante considerar que a amostra original pode ser considerada como sendo mais confiável, por se tratar de dados obtidos em campo. Por isso, pode ocorrer de a amostra sistemática artificial, que é produzida a partir da original, acumular erros primários ao processo de coleta de dados. No caso das análises para o ano de 2010, é importante considerar, inclusive que os dados de produção de viagens são oriundos de um modelo linear, não correspondendo a

dados pesquisados em campo. Portanto, os dados aqui apresentados referem-se a estimações que buscam representar a realidade a partir de métodos ainda pouco utilizados no planejamento de transportes.

Com a utilização da amostra sistemática artificial, o número de amostras aumentou, desagregando-se a amostra original e consequentemente 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, como apresentado a seguir.

5.3 DETERMINAÇÃO DE PARÂMETROS E MODELAGEM DOS VARIOGRAMAS EXPERIMENTAIS

A partir da análise exploratória descrita na seção anterior, foram determinadas algumas características dos dados sobre população, empregos, atração e produção de viagens pela amostra por MZ de tráfego e sistemática artificial, obtida através de uma malha regular.

Foi verificado que os dados não possuem uma direção principal de variabilidade, com exceção das variáveis população e produção de viagens (1995). Para estes dados se supôs que o ângulo “principal” estaria entre 45 e 60 graus. Segundo Camargo (1998) é importante analisar os ângulos que representam mínima e máxima continuidade espacial. Após visualização espacial e realização de alguns testes, optou-se por modelar os variogramas experimentais para a direção 50º e o seu respectivo ângulo ortogonal -40º, para os dados de População e Produção e viagens por centróides das MZs de 1995. Este ângulo representou melhor a anisotropia ou continuidade espacial das amostras.

As demais variáveis (empregos e atração de viagens) não apresentaram direção específica ou tendência espacial e foram modeladas com o ângulo máximo de abertura, ou seja, 180º. Neste caso, a modelagem do variograma experimental depende exclusivamente da distância entre as amostras, e não de uma direção (CÂMARA *et al.,* 2002).

Os parâmetros descritos abaixo foram utilizados para a modelagem dos variogramas experimentais das quatro variáveis analisadas.

Azimute *θ*: ângulo ou direção principal, com valores entre -90 e 90°. No caso de variáveis isotrópicas, o azimute adotado é 0º. Neste trabalho, produção e população, na amostragem original por centróide das MZs (1995) (50º), atração e empregos (0º);

Tolerância angular (*R*): limite de abertura angular para varredura dos dados a partir de um ponto amostral;

***Lag distance* (*h*):** distância entre as amostras (no caso analisado neste trabalho foi considerada a média das distâncias euclidianas dos centróides, devido a grande distâncias entre os pontos em algumas áreas);

***Cut distance*:** valor que representa a área máxima ou distância total da distribuição espacial dos pontos.

A seguir serão descritos os parâmetros que melhor descreveram o comportamento espacial das amostras de cada variável analisada. Foram modelados 14 variogramas. Dez para os dados da Pesquisa OD de 1995, sendo seis da amostragem original e quatro para a amostra obtida através da malha regular. Para os dados de 2010 foram modelados quatro variogramas para as variáveis População e Produção de viagens por TCO estimadas para o ano, também para os dois tipos de amostragem.

Os parâmetros utilizados para obtenção dos variogramas experimentais para as quatro variáveis e seus diferentes tipos de amostragem (para os dados da Pesquisa O/D de 1995) estão representados na Tabela 5.7.

Tabela 5.7: Parâmetros utilizados para a elaboração dos variogramas das variáveis empregos, população, produção e atração de viagens por TCO (1995)

Tipo de amostragem	Variáveis OD 1995	Direção
Centróides das MZs	População	50° -40º
	Produção de viagens	50° -40º
	Empregos	Omnidirecional
	Atração de viagens	
Malha Regular	População	Omnidirecional
	Produção de viagens	
	Empregos	
	Atração de viagens	

Os dados da Tabela 5.8 são referentes à compatibilização dos dados do censo 2010 (setores censitários) com as áreas das MZs da Pesquisa OD de 1995. A partir desta compatibilização foi possível obter amostragem por grade regular para os dados de população e os dados estimados de produção de viagens.

Tabela 5.8: Parâmetros utilizados para a elaboração dos variogramas das variáveis População e Produção de viagens por TCO (2010)

Tipo de amostragem	Variáveis	Direção
Centróides das MZs	População	Omnidirecional
	Produção de viagens	
Malha regular	População	
	Produção de viagens	

Na próxima subseção deste capítulo está demostrada a modelagem dos variogramas, a partir dos parâmetros mostrados anteriormente.

5.4 MODELAGEM DOS VARIOGRAMAS

A etapa de modelagem dos variogramas foi realizada no módulo geoMod do GeoMs. Este é módulo de escolha e ajustamento do modelo teórico do variograma, cujos critérios foram definidos na etapa anterior, através do módulo geoVar.

Nesta etapa são definidos os parâmetros que determinam a função do variograma, etapa principal para a interpolação dos dados em coordenadas desconhecidas. São eles: efeito pepita, alcance ou amplitude, patamar (*sill*) e o modelo teórico de ajuste do variograma (modelo esférico, gaussiano, exponencial). Portanto, nesta subseção serão apresentadas as modelagens dos variogramas de cada variável estudada neste trabalho e seus respectivos parâmetros.

Soares (2006) destaca que, para a modelagem dos variogramas experimentais, é necessário que se obtenha mesmo valor de patamar e de efeito de pepita (caso exista) em todas as direções, no entanto, amplitudes diferenciadas, com maior amplitude para a direção principal e amplitudes menores nas direções ortogonais.

5.4.1 Geração de Viagens - Pesquisa OD 1995 (amostragem por centróides das MZs)

A Tabela 5.9 mostra o resumo dos parâmetros para a modelagem dos variogramas das variáveis população e produção de viagens, empregos e atração e viagens. Estes dados são referentes à amostra por centróide das MZs de tráfego da Pesquisa OD de 1995. Os variogramas ajustados estão apresentados na Figura 5.13. Em todas as modelagens, o modelo teórico de variograma adotado é o modelo esférico.

Tabela 5.9: Parâmetros utilizados na modelagem dos variogramas das variáveis população, empregos e geração de viagens por TCO (1995)

Variáveis (por MZ) e Direções	Efeito Pepita (C0)	(C1)	Alcance	Patamar (C0+C1)
Produção de viagens por TCO (50º)	0,26	0,69	12882,68	0,95
Produção de viagens por TCO (-40º)	0,26	0,69	9832,59	0,95
População residente (50º)	0,29	0,67	14785,19	0,96

População residente (-40º)	0,29	0,67	8896,15	0,96
Atração de viagens por TCO (omnidirecional)	0,48	0,49	7400,27	0,98
Empregos (omnidirecional)	0,40	0,58	7559,26	0,98

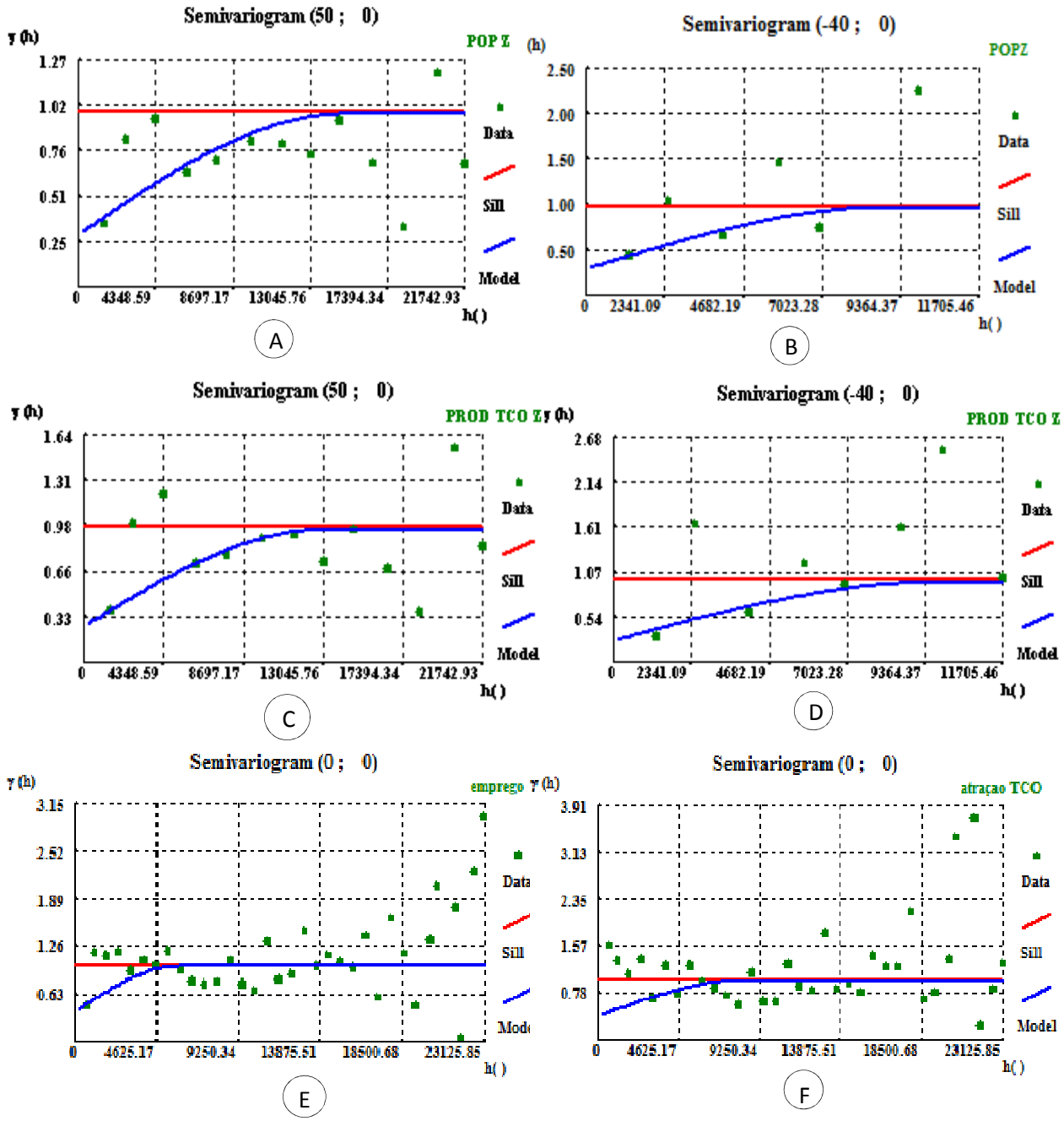


Figura 5.13: Modelagem dos variogramas. (A e B) população; (C e D) produção de viagens por TCO; (E) empregos; (F) atração de viagens por TCO.

Como pode ser observado através dos parâmetros e dos variogramas (Figura 5.13), e já conhecendo a alta correlação destas variáveis, é possível perceber que População e Produção de viagens por TCO apresentam parâmetros muito parecidos. Já os variogramas das variáveis empregos e atração de viagens não apresentam uma estruturação considerada ideal para a caracterização da variabilidade espacial, o que é analisado pela Geoestatística, podendo ser considerado como efeito pepita puro. A ocorrência destas características para estes modelos é recorrente, principalmente, da discrepância dos dados e, também, pela falta de estrutura espacial dos mesmos.

5.4.2 Geração de Viagens Pesquisa OD 1995 – Malha Regular

A seguir estão descritos os parâmetros utilizados na modelagem dos variogramas através da amostragem sistemática artificial. Todos os variogramas foram omnidirecionais. A Tabela 5.10 mostra os parâmetros adotados para a modelagem dos variogramas dos dados obtidos. A amostragem artificial gerou melhores variogramas devido à maior quantidade de pontos localizados na área de estudo. Esta amostragem resulta não apenas em uma distribuição mais regular dos pontos, como também atribui uma homogeneização em relação à distribuição de pontos pela área de estudo. Isto permite que o valor da função variograma em cada ponto não sofra tantas oscilações, e os primeiros pontos se apresentem abaixo da variância total dos dados.

Tabela 5.10: Parâmetros utilizados na modelagem dos variogramas (população e produção de viagens por TCO) - malha regular (1995)

Variáveis e Direções	Efeito Pepita (C0)	(C1)	Alcance	Patamar (C0+C1)
Produção de viagens por TCO (omnidirecional)	0,18	0,66	10845,02	0,85
População residente (omnidirecional)	0,21	0,92	9859,11	1,14
Atração de viagens por TCO (omnidirecional)	0,06	0,14	11336,76	0,21
Empregos (omnidirecional)	0,14	0,25	9599,38	0,39

A dependência espacial, representada pelo alcance na Tabela 5.10, é maior na variável Produção de viagens por TCO. No entanto, o efeito pepita é menor para esta variável, apesar da pequena diferença representada na Tabela 5.10 e nos variogramas (Figura 5.14). O Patamar da variável População é um pouco maior, assim como a Contribuição (C1).

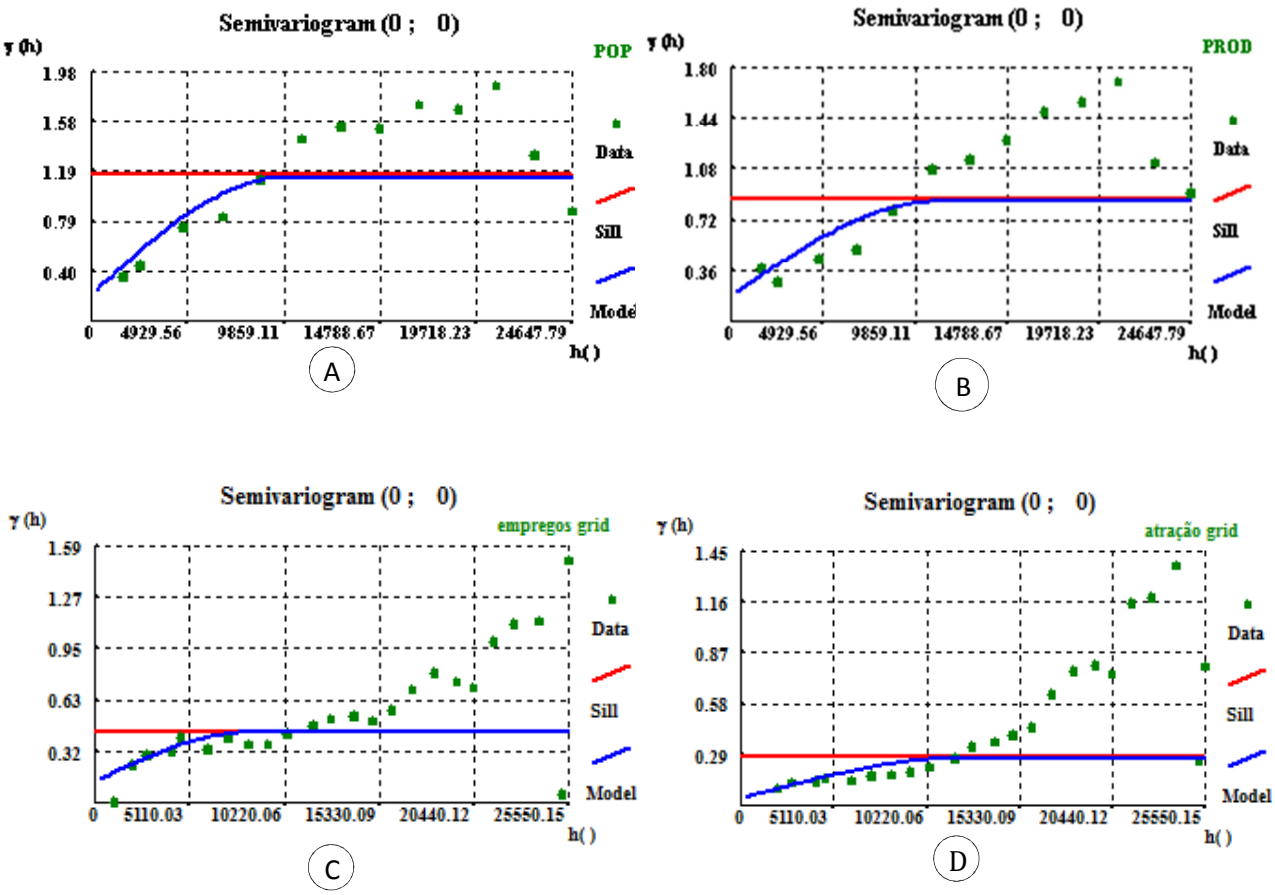


Figura 5.14: Modelagem dos variogramas. (A) população; (B) produção de viagens; (C) empregos; (D) atração de viagens (amostra sistemática artificial).

As variáveis empregos e atração de viagens também possuem comportamento isotrópico, e mesmo adaptados à malha regular de 2000 por 2000 metros, não demonstraram nenhuma continuidade ou variação aparente dos dados em alguma direção. No entanto, esta amostragem apresentou uma melhora significativa no ajustamento dos variogramas experimentais, em relação à amostragem pelos centróides das MZs (Figura 5.14 C e D).

5.4.3 Produção de Viagens 2010: Dados Estimados

A seguir, estão descritos os resultados da modelagem do variograma experimental das variáveis população e produção de viagens (dados 2010) por centróide de cada MZ, e posteriormente, através da malha regular.

A Tabela 5.11 mostra o resumo dos parâmetros dos variogramas experimentais representados pelos centróides das MZs e através da malha regular. Estas características da modelagem experimental podem ser vistas na Figura 5.15, que mostra o ajuste teórico dos variogramas a partir dos dados descritos na tabela acima. A partir desta figura, é possível perceber como os parâmetros experimentais influenciam os modelos teóricos dos variogramas, etapa muito importante da interpolação por krigagem.

Tabela 5.11: Parâmetros utilizados na modelagem dos variogramas (dados compatibilizados, IBGE 2010) - amostra por centróides das MZs.

Tipo de amostragem	Variáveis e Direções	Efeito Pepita (C0)	(C1)	Alcance	Patamar (C0+C1)
Centróides das MZs	Produção de viagens por TCO (omnidirecional)	0,27	0,70	9250,34	0,97
	População residente (omnidirecional)	0,30	0,68	9250,34	0,98
Malha regular	Produção de viagens por TCO (omnidirecional)	0,20	0,96	10799,72	1,16
	População residente (omnidirecional)	0,21	0,95	10799,72	1,16

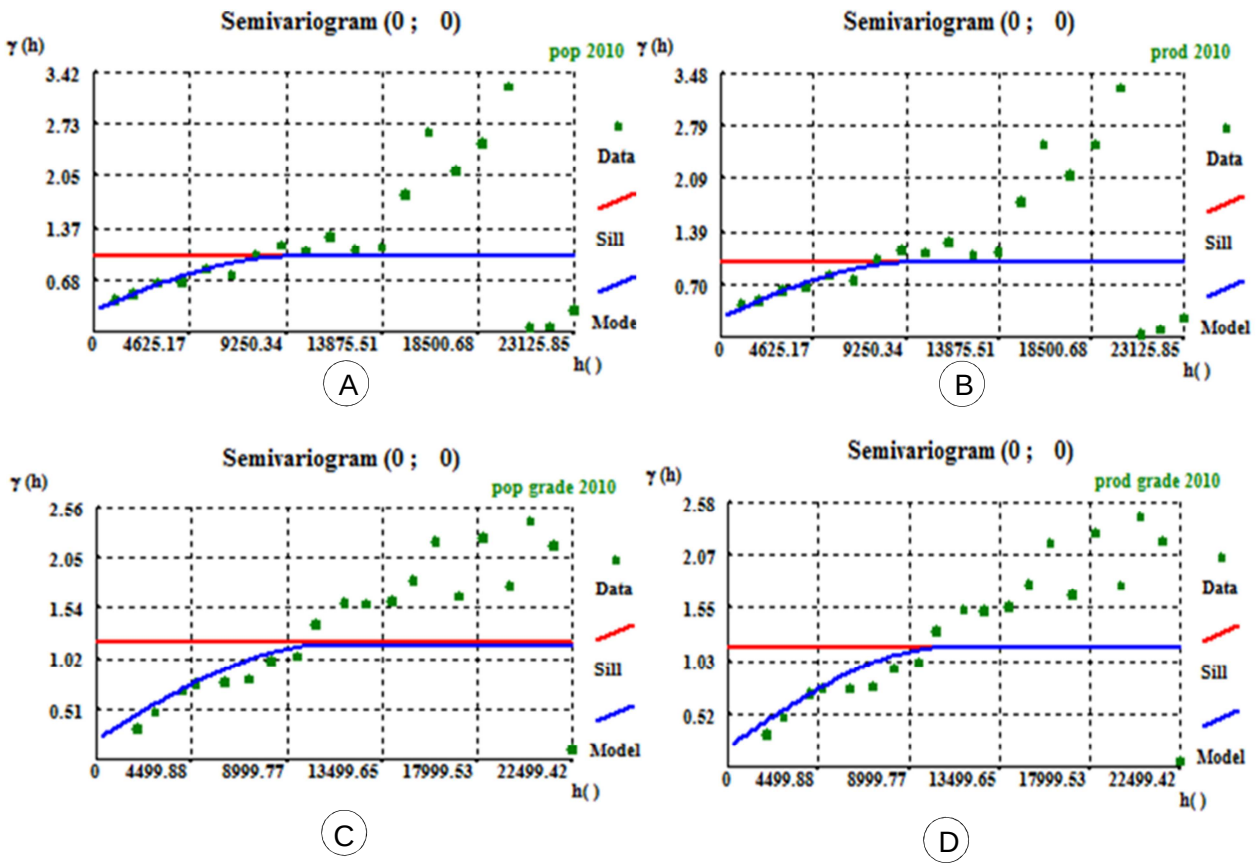


Figura 5.15: Modelagem dos variogramas das variáveis População (A e C) e produção de viagens (B e D) 2010 (amostragem por centróide das MZs e por malha regular)

A partir do ajuste dos modelos teóricos dos variogramas para as variáveis em estudo, a próxima etapa é definir alguns critérios necessários para a interpolação de dados não amostrados através da krigagem ordinária e da krigagem com deriva externa.

5.5 ESCOLHA DOS CRITÉRIOS PARA A KRIGAGEM - GeoKrig e GeoKed

Definidos os parâmetros dos variogramas experimentais, deve-se estabelecer alguns critérios necessários para a realização da estimação (Tabela 5.12). Estes critérios são definidos nos módulos GeoKrig ou GeoKed, de acordo o tipo de krigagem que se deseja realizar.

Nesta etapa também são definidas as coordenadas que abrangem toda área de estudo, o número de blocos a se estimar, o espaçamento entre as amostras ou blocos. Neste trabalho, onde foram estimados pontos, cada bloco é discretizado num só ponto segundo eixos principais (Tabela 5.12). Também é apresentado o resumo dos parâmetros do ajuste do variograma teórico, escolhidos na etapa anterior, como o efeito pepita, amplitude e as direções de anisotropia (caso ocorra).

Tabela 5.12: Parâmetros utilizados nas Krigagens

Coordenadas	Número de células	Número de células (malha regular)	Espaçamento entre células (m)	Discretização
X	107	110	200	1
Y	102	111	200	1

Definidos estes parâmetros, é necessário também definir alguns critérios importantes para a realização da krigagem. Abaixo estão descritos os principais deles.

- Número de amostras: o número máximo e mínimo de amostras necessárias para estimar um ponto da malha regular. Os valores permitidos variam entre 1 e 64. Neste trabalho o valor mínimo foi de 4 e máximo de 60 amostras.

- Tipo de busca das amostras: simples ou por quadrantes. O tipo de busca corresponde a um critério de seleção das amostras utilizadas para a estimação de um dado ponto. No caso da busca ser simples, são selecionadas amostras mais próximas do ponto a estimar (segundo qualquer direção) seguindo as restrições impostas pelo número máximo e mínimo de amostras a utilizar. No entanto, o sistema de equações de krigagem atribui pesos diferentes às amostras de acordo com os parâmetros da modelagem definidos nos variogramas. No caso da busca por quadrantes, divide-se o plano X Y em quadrantes, assumindo-se como origem dos quadrantes a posição do ponto a estimar, e dentro de cada quadrante aplicam-se novamente as restrições impostas pelo número máximo e mínimo de amostras a utilizar. A busca por octantes corresponde a uma generalização da busca por quadrantes a três dimensões. Neste trabalho foi utilizado o tipo simples.

- Ângulos do elipsóide de busca das amostras: Estes ângulos que definem as três direções dos eixos do elipsóide devem, quando possível, ser os mesmos que os definidos pelos variogramas, de modo a se efetuar uma busca por amostras mais próxima com a continuidade espacial do fenômeno analisado. As duas direções definidas devem ser perpendiculares entre si, para casos de anisotropia.

- Teste do ponto fictício (*Cross-Validation*): caso seja optado pela validação cruzada, o programa estima o valor nos pontos onde estão localizadas as amostras, ignorando o valor da mesma.

Como descrito no Capítulo 3, a Krigagem com Deriva Externa é feita a partir de dados de uma variável secundária. Os dados desta variável servem como entrada no geoKed para auxiliar na estimação da variável primária. Os dados utilizados como variável secundária foram população e empregos, estimados em coordenadas desconhecidas, e as variáveis primárias foram representadas pela produção e atração de viagens para o ano de 1995, e produção de viagens, estimada através do modelo de regressão linear, para o ano de 2010.

Após a krigagem, os resultados da estimação são verificados através da Validação Cruzada, que será vista no Capítulo 6. Vale ressaltar que esta é uma etapa fundamental da estimação geoestatística, pois ela permite mensurar a qualidade do modelo, definido a partir dos variogramas modelados anteriormente. Deve-se levar em conta que, especificamente neste trabalho, a modelagem e validação dos variogramas ajustados foi realizada para a mesma amostra de pontos, considerando a pouca quantidade de pontos amostrados (centróides de macrozonas). Um procedimento metodologicamente mais adequado seria a realização da modelagem dos variogramas com parte da amostra de pontos e validação cruzada com os pontos restantes.

Além da Validação Cruzada, outro resultado obtido a partir da definição dos critérios descritos acima é a interpolação de mapas krigados. Tais mapas permitirão visualizar espacialmente os resultados da estimação.

CAPÍTULO 6: RESULTADOS E DISCUSSÕES

Este capítulo apresenta os principais resultados das etapas da krigagem apresentadas no Capítulo 5. A partir da validação cruzada foi possível verificar a qualidade das estimações. Serão apresentados e discutidos também os mapas interpolados de cada variável analisada através dos dois métodos de krigagem utilizados. Serão apresentadas também as principais diferenças encontradas a partir das estimações da geração de viagens, empregos e população obtidas pela amostragem original por centróides das MZs e pela amostragem sistemática artificial.

6.1 VALIDAÇÃO CRUZADA

A validação cruzada (*cross validation*), ou teste do ponto fictício, compreende a análise dos erros, ou seja, mede a incerteza da estimação, sendo possível verificar sua confiabilidade. Nesta etapa é possível mensurar os erros da estimação e conhecer a variância dos resultados. A validação cruzada é realizada considerando os valores observados e estimados nas coordenadas conhecidas (PITOMBO e SOUSA, 2009). Com isso, é possível conhecer a qualidade do modelo do variograma (SOARES, 2006).

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.

No entanto, resultados da validação cruzada que apresentem erros com valores baixos não implica, necessariamente, que o modelo escolhido seja o mais adequado. Logo, subtende-se que, estando diante de médias, o modelo pode representar bem ou mal os padrões de continuidade espacial e variabilidade do conjunto de dados (SOARES, 2006).

A validação cruzada envolve a comparação de estimativas krigadas e suas variações (WEBSTER e OLIVER, 2001). Os mesmo autores afirmam que o erro médio ideal deve ser zero. A Tabela 6.1 mostra um exemplo dos resultados da validação cruzada.

Tabela 6.1: Exemplo de dados obtidos na etapa da validação cruzada.

X	Y	Valor observado	Valor estimado	Variância da estimação
552992	8562160	1,13	1,45	0,46
555434	8561660	0,89	1,64	0,49
557172	8561870	1,58	1,60	0,47
558542	8562530	1,94	1,61	0,49
551469	8563580	0,42	1,25	0,46
552551	8563750	0,93	1,14	0,41
552916	8562860	0,88	1,40	0,42
554321	8562900	2,79	1,28	0,43
555896	8563110	0,83	1,96	0,44
552445	8564590	1,31	0,79	0,42
553603	8565260	1,44	0,88	0,40
554106	8564090	0,98	1,50	0,42

A partir destes resultados são calculados outros atributos estatísticos, como a média de erros e a média do quadrado dos erros e coeficiente de correlação. As próximas seções mostram os resultados da validação cruzada para as estimações da geração de viagens, empregos e população para os dados dos anos de 1995 e 2010.

De acordo com o conjunto de resultados obtidos através da validação cruzada, pode-se considerar que a krigagem é uma ferramenta viável na estimação de viagens para coordenadas desconhecidas e que os parâmetros adotados para modelagem dos dados foram adequados. Vale ressaltar que um procedimento metodologicamente mais adequado seria a realização da modelagem dos variogramas com parte da amostra de pontos e validação cruzada com os pontos restantes. Neste trabalho tanto a modelagem quanto a validação foram realizados com a mesma amostra de pontos.

6.1.1 População e Produção de viagens por TCO 1995 (amostragem original por centróide das MZs)

A análise dos resultados da validação cruzada para técnicas de Krigagem Ordinária (KO) e Krigagem com Deriva Externa (KED) para as variáveis produção de viagens por TCO e população apresentou, no geral, bons resultados (Tabela 6.2). No entanto, o coeficiente de correlação entre valores observados e estimados das variáveis apresenta valores muito baixos. Isto, provavelmente, se deve ao fato das grandes distâncias entre os centróides das MZs e também poucas observações. Contudo, a técnica KDE apresentou maiores valores de coeficiente de correlação entre dados observados e estimados.

Tabela 6.2: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (1995)

Tipo de Krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média dos quadrados dos erros
KO	População	0,08	-0,0072	0,49	1,10
	Produção de viagens por TCO	0,20	0,0005	0,47	1,00
KDE	Produção de viagens por TCO	0,63	-0,0045	0,51	0,65

Apesar disso, deve-se levar em consideração que os dados utilizados para as estimações são dados sociais, não naturais (como utilizados comumente em geoestatística), logo, muitas vezes, não apresentam um padrão de continuidade espacial bem definido. Mas, esta hipótese não inviabiliza a utilização de técnicas de krigagem na estimação de viagens e de dados sobre população e empregos (como utilizados neste trabalho), nem de outros dados das Ciências Sociais, desde que se façam suposições e adequações.

Nota-se que o coeficiente de correlação para a estimação por krigagem com deriva externa melhorou consideravelmente. A variável população, altamente correlacionada com a variável produção de viagens por TCO, contribuiu para a estimação desta variável em coordenadas desconhecidas. Os demais resultados continuam bons, no entanto, a média do quadrado dos erros diminuiu, mas ainda

apresenta bom resultado, já que está próxima de 1. Segundo Almeida *et al.* (2007), isso demonstra o bom ajuste entre valores observados e estimados.

Outro resultado derivado da validação cruzada foram os diagramas de dispersão, que relacionam dados observados e estimados para cada variável, sendo possível analisar a relação de linearidade entre os resultados da krigagem. Os diagramas de dispersão, provenientes da validação cruzada para o emprego das técnicas KO e KDE para todas as variáveis em análise, estão apresentados nos apêndices deste trabalho. Pode-se notar, através das figuras, a relação entre as variáveis. Mesmo seus valores estimados por diferentes tipos de krigagem apresentam correlação, o que mostra que a krigagem pode ser um método eficiente na previsão de viagens urbanas.

A seguir, serão apresentados os mapas interpolados, resultantes dos modelos de variogramas e dos parâmetros de krigagem adotados para as variáveis citadas.

6.1.2 Interpolação de mapas: População e Produção de viagens (1995) amostragem original por centróides das MZs

A partir dos parâmetros adotados no processo de krigagem, foram estimadas 10.914 células em coordenadas desconhecidas para cada variável analisada. Os mapas resultantes do processo de estimação por Krigagem Ordinária apresentam o mesmo padrão espacial para população e produção de viagens, como pode ser visto na Figura 6.1 e Figura 6.2. Isso reafirma forte correlação espacial entre as variáveis em análise.

Através dos mapas, percebe-se que a área que possui maiores taxas de produção de viagens por TCO e também população na RMS é o Subúrbio Ferroviário de Salvador. Esta área é a mais pobre e possui maior vulnerabilidade socioeconômica.

Além das MZs correspondentes ao Subúrbio Ferroviário, MZs como Brotas, algumas MZs do Miolo de Salvador e Itapuã possuem elevado número de população residente. Tendência contrária percebe-se na região do Centro de Salvador, área destinada ao comércio e serviços, e com altas taxas de atração de viagens e empregos (como será visto posteriormente).

A partir dos resultados da krigagem e através dos mapas interpolados para as duas variáveis estimadas é possível afirmar que a estimação geoestatística da produção de viagens e da variável população correspondem à realidade da RMS, onde maiores valores da população e da produção de viagens por TCO se concentram na região da orla da Baía de Todos os Santos, mais precisamente a partir da MZ Ribeira até o Subúrbio Ferroviário.

Por outro lado, os dados da estimação obtidos pela técnica de Krigagem com Deriva Externa apresentaram valores inversos aos dados observados em alguns pontos (Figura 6.3). Esta diferença pode ser notada principalmente nas MZs Coutos e Itapuã, cujas áreas apresentaram os valores mais baixos de produção de viagens por TCO. Foi verificado que os dados estimados de população (que entram como variável auxiliar no processo de krigagem com deriva externa) apresentaram valores baixos (como pode ser verificado na Tabela 6.3).

Tabela 6.3: População observada e estimada

MZ	População observada	População estimada
Itapuã	3.66	0.827
Coutos	4.26	1.406

Pode-se, então, deduzir que os dados estimados da variável secundária influenciaram esta diferença, já que os valores estimados destas áreas adquiriram um valor menor na estimação utilizando a KDE.

Segundo Yamamoto (2009), nas estimativas por krigagem ordinária há um efeito de suavização nos dados estimados. O efeito de suavização pode ser conhecido através da validação cruzada, a partir do erro verdadeiro entre dados observados e estimados. Em erros abaixo de zero há uma subestimativa dos valores elevados, já erros maiores que zero indicam que houve uma superestimativa de valores baixos. Yamamoto (2009) apresenta a correção deste erro, utilizando uma nova variável padronizada denominada *número de desvios padrão de interpolação*, onde o erro verdadeiro é divido pelo desvio padrão da estimação, conforme a Equação 6.1.

$$NS_O = \frac{-ErroVerdadeiro}{S_o} \text{ , sendo } S_o \neq 0 \qquad \text{(eq 6.1)}$$

onde,

NS_0 é a variável adicional (erro verdadeiro padronizado);

S_0 é o desvio padrão da interpolação.

Assim, onde há subestimação (sinal negativo), uma quantidade de correção deve ser adicionada e, onde há superestimatimação uma quantidade de correção deve ser subtraída.

No processo de validação cruzada da krigagem com deriva externa retira-se o valor observado da variável primária, e mantém-se o valor da variável secundária, e passa a existir uma variável complementar NS_O. A partir desta nova variável, multiplicada pelo desvio padrão da estimação, obtém-se a correção do efeito de suavização e, consequentemente, da estimativa. No entanto, este método não é aplicado no presente trabalho. Apenas foi apresentado como uma justificativa para a diferença encontrada nos dois mapas, e também como uma alternativa para trabalhos futuros.

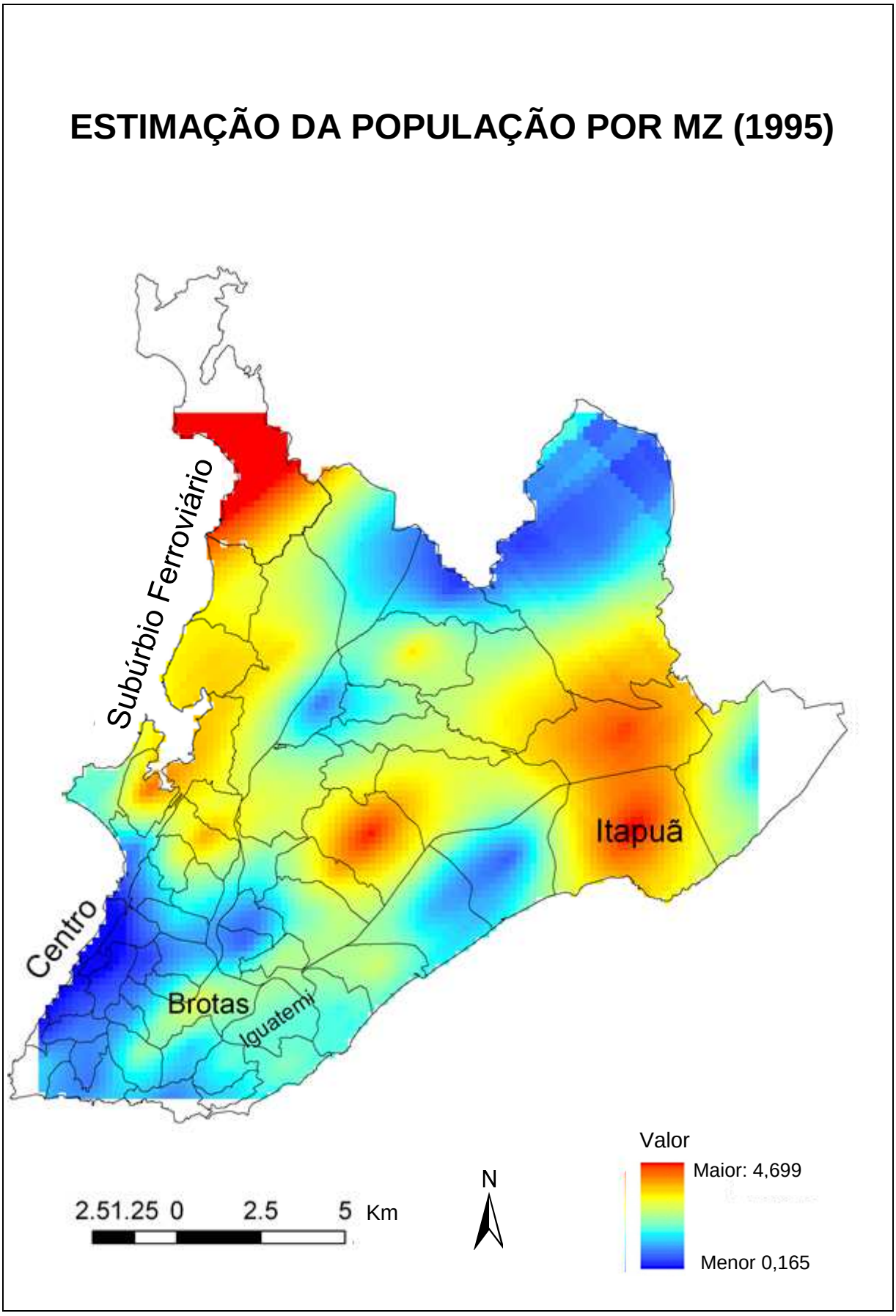


Figura 6.1: População estimada por MZ (1995) - amostra por centróide da MZ - KO

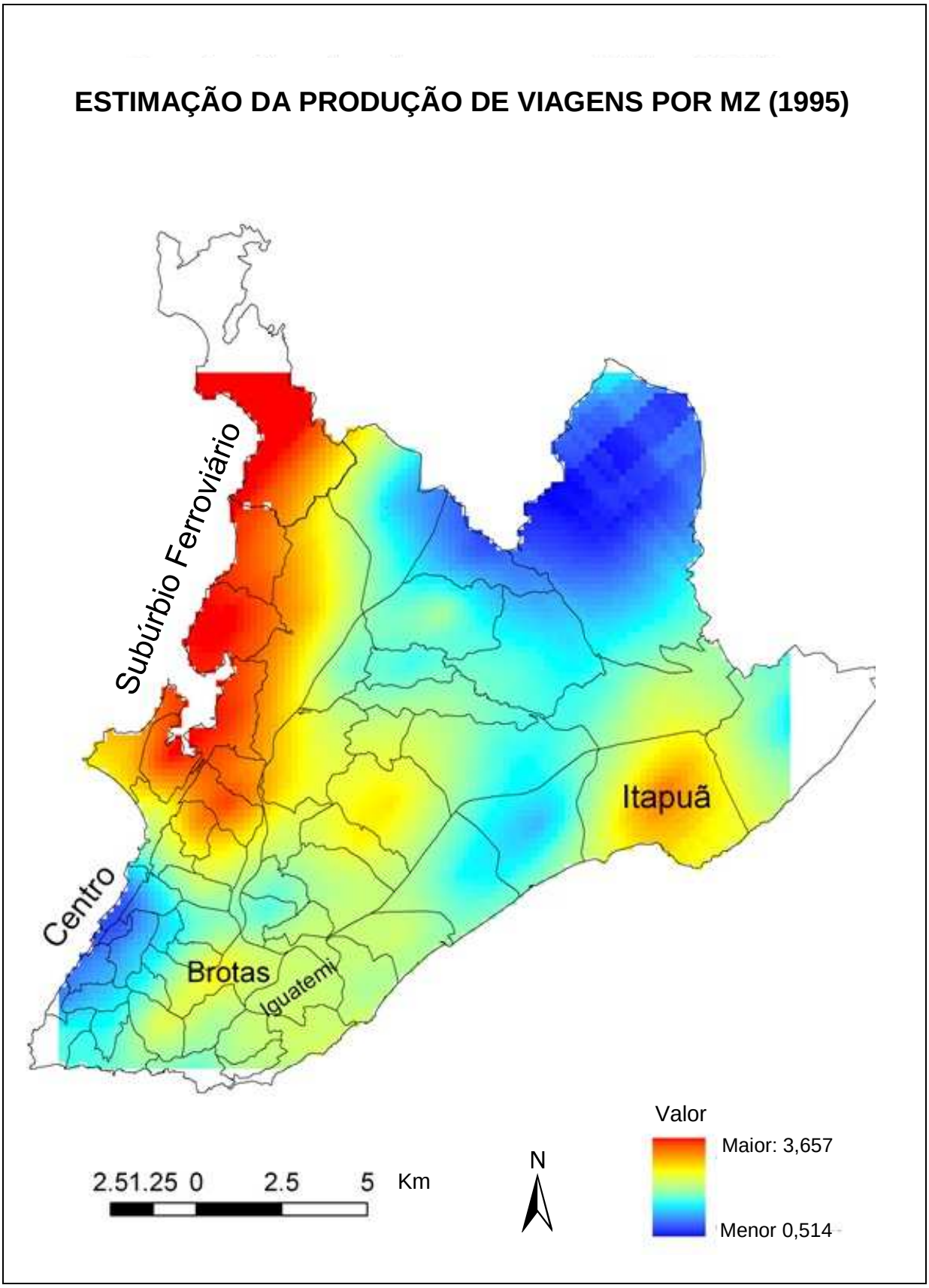


Figura 6.2: Produção de viagens por TCO estimada por MZ (1995) - amostra por centróide da MZ – KO

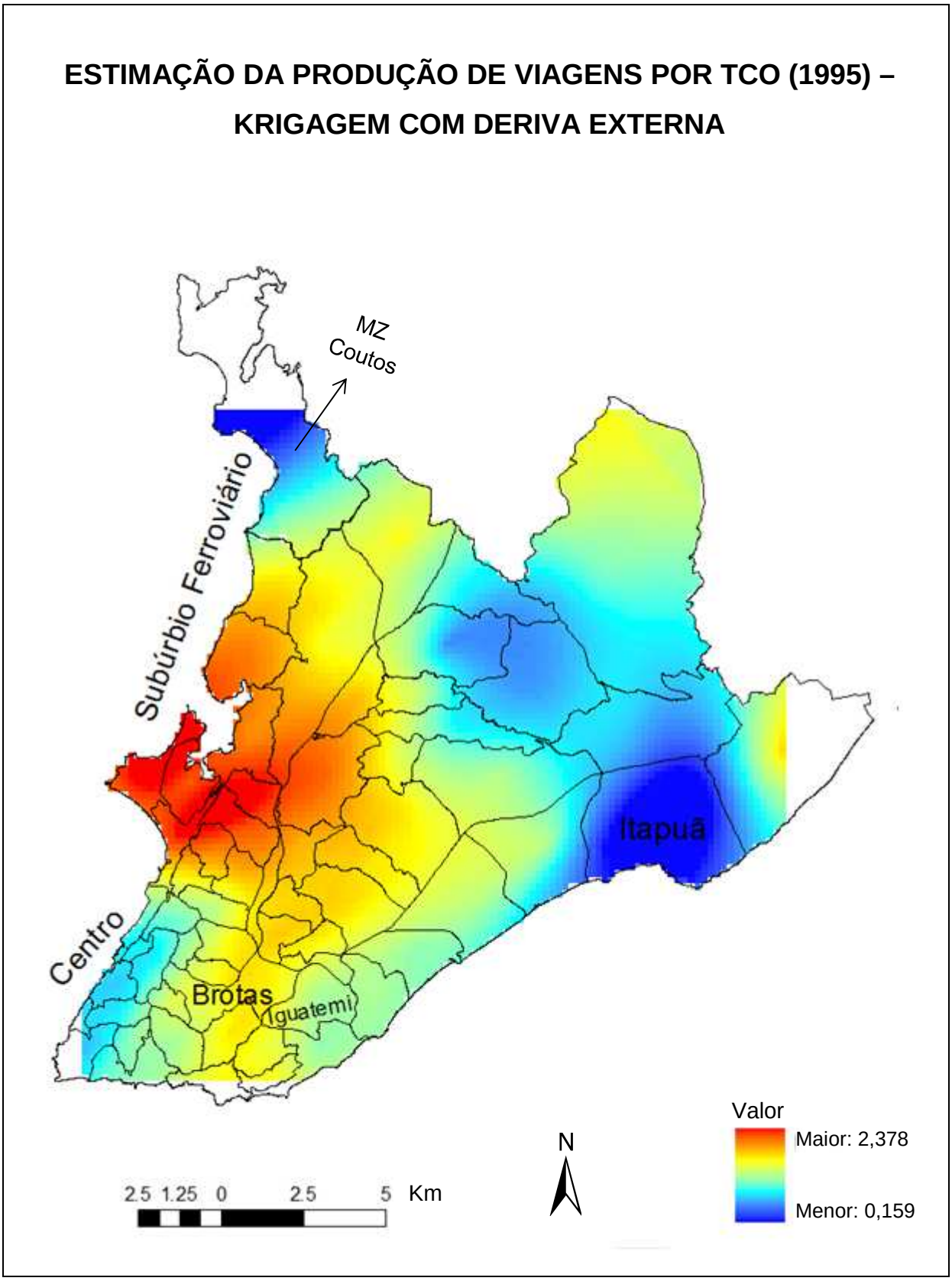


Figura 6.3: Produção de viagens por TCO estimada por MZ (1995) - amostra por centróide das MZs – KDE

6.1.3 Empregos e Atração de viagens por TCO 1995 (amostragem original por centróides das MZs)

A partir dos resultados da validação cruzada da KO e da KDE das variáveis atração de viagens por TCO e empregos (Tabela 6.4), pode-se considerar que os resultados estão de acordo os critérios de validação cruzada considerados satisfatórios. No entanto, assim como na estimação da produção de viagens e população, o coeficiente de correlação entre valores observados e estimados das variáveis apresenta valores considerados baixos. Da mesma forma, valores de correlação da KDE são maiores quando comparados aos valores da KO.

Tabela 6.4: Dados obtidos através da Validação Cruzada da estimação das variáveis empregos e atração de viagens (1995)

Tipo de Krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média do quadradro dos erros
KO	Empregos	0,19	-0,010	0,675	1,01
	Atração de viagens por TCO	0,40	-0,003	0,747	0,82
KDE	Atração de viagens por TCO	0,49	-0,001	0,774	0,75

Pode-se considerar que os resultados da validação cruzada destas variáveis apresentaram bons resultados.

Com a aplicação da Krigagem com Deriva Externa, o coeficiente de correlação de Pearson teve pouca alteração para a estimação da atração de viagens por TCO. Outras análises, como média de erros e da variância, apresentam bons resultados, assim como na krigagem ordinária.

6.1.4 Interpolação de mapas: Empregos e Atração de viagens (1995) amostragem original por centróides das MZs

Os resultados das krigagens para as variáveis empregos e atração de viagens podem ser vistos nas Figuras 6.4 e 6.5. A propensão ao aumento de viagens atraídas por TCO é contrária à tendência observada para viagens produzidas,

havendo crescimento em direção ao centro (AUC), área com maior número de empregos e atividades urbanas.

Outras áreas que apresentam taxas de empregos maiores são as MZs Iguatemi e Pituba. A MZ Itapuã apresenta uma variação em relação às MZs citadas, ainda assim se destacando em relação às demais MZs de seu entorno, que possuem baixa concentração de empregos.

Para atração de viagens por TCO no horário de pico (Figura 6.5) nota-se que apenas a área do centro de Salvador corresponde a valores elevados. Apesar das outras MZs possuírem maior quantidade de empregos, e mesmo considerando a alta correlação destas variáveis, o fluxo de viagens atraídas por TCO no horário de pico não se apresenta nas demais áreas com taxas de empregos elevadas, como visto na Figura 6.4. Possivelmente este resultado se dá, sobretudo, em função do horário (6 às 8 horas da manhã). Geralmente as atividades de comércio e serviços se iniciam em outros horários.

A Figura 6.6 apresenta o resultado da interpolação de viagens atraídas estimados pela krigagem com deriva externa. O mapa apresenta o mesmo padrão de distribuição do mapa obtido pela krigagem ordinária, com apenas uma área apresentando elevados valores de atração de viagens: a MZ Campo Grande, diminuindo à medida que se afasta dela.

Vale ressaltar que os variogramas destas variáveis não apresentaram uma estrutura adequada para estimação geoestatística, como discutido no capítulo anterior. No entanto, os resultados correspondem à realidade de Salvador, sobretudo no período analisado, onde as principais atividades se concentravam nas áreas do centro antigo e na região da MZ Iguatemi. Estes mapas também correspondem ao que foi discutido na análise exploratória dos dados (Capítulo 5), mais especificamente na espacialização apresentada nos mapas de pontos de ambas as variáveis.

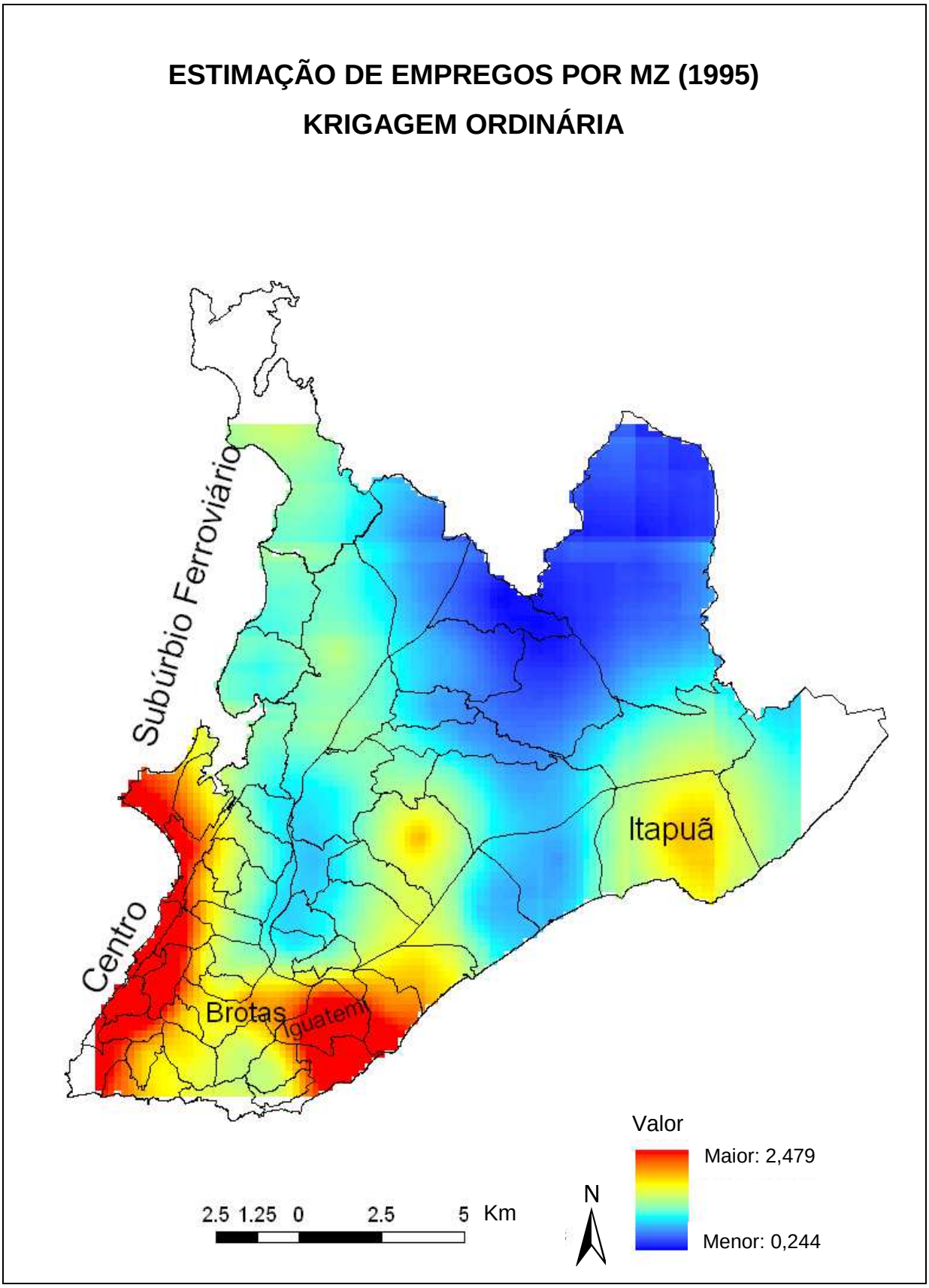


Figura 6.4: Empregos estimados por MZ (1995) - amostragem por centróide das MZs - KO

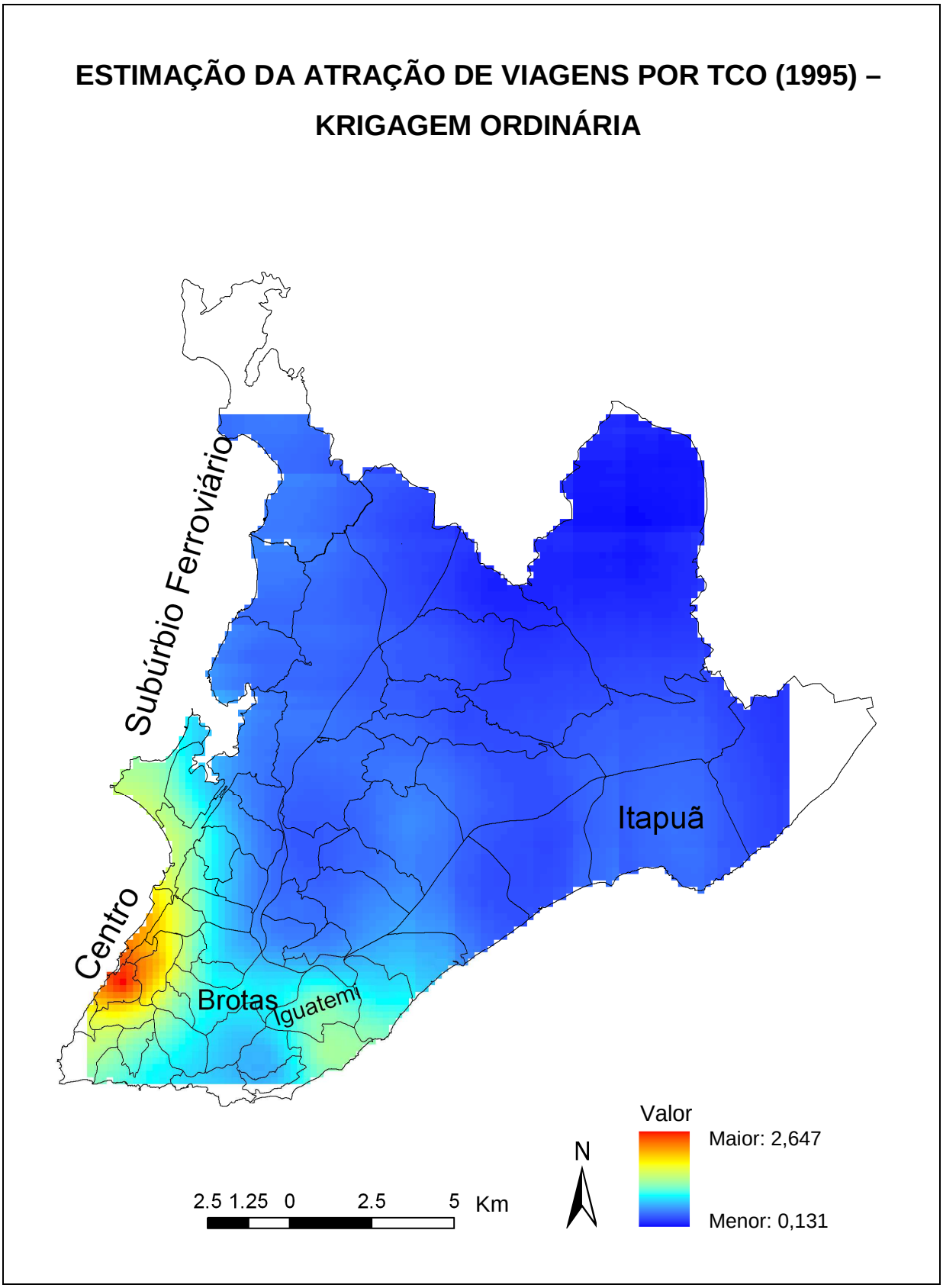


Figura 6.5: Atração de viagens por MZ (1996) - amostragem por centróide das MZs – KO

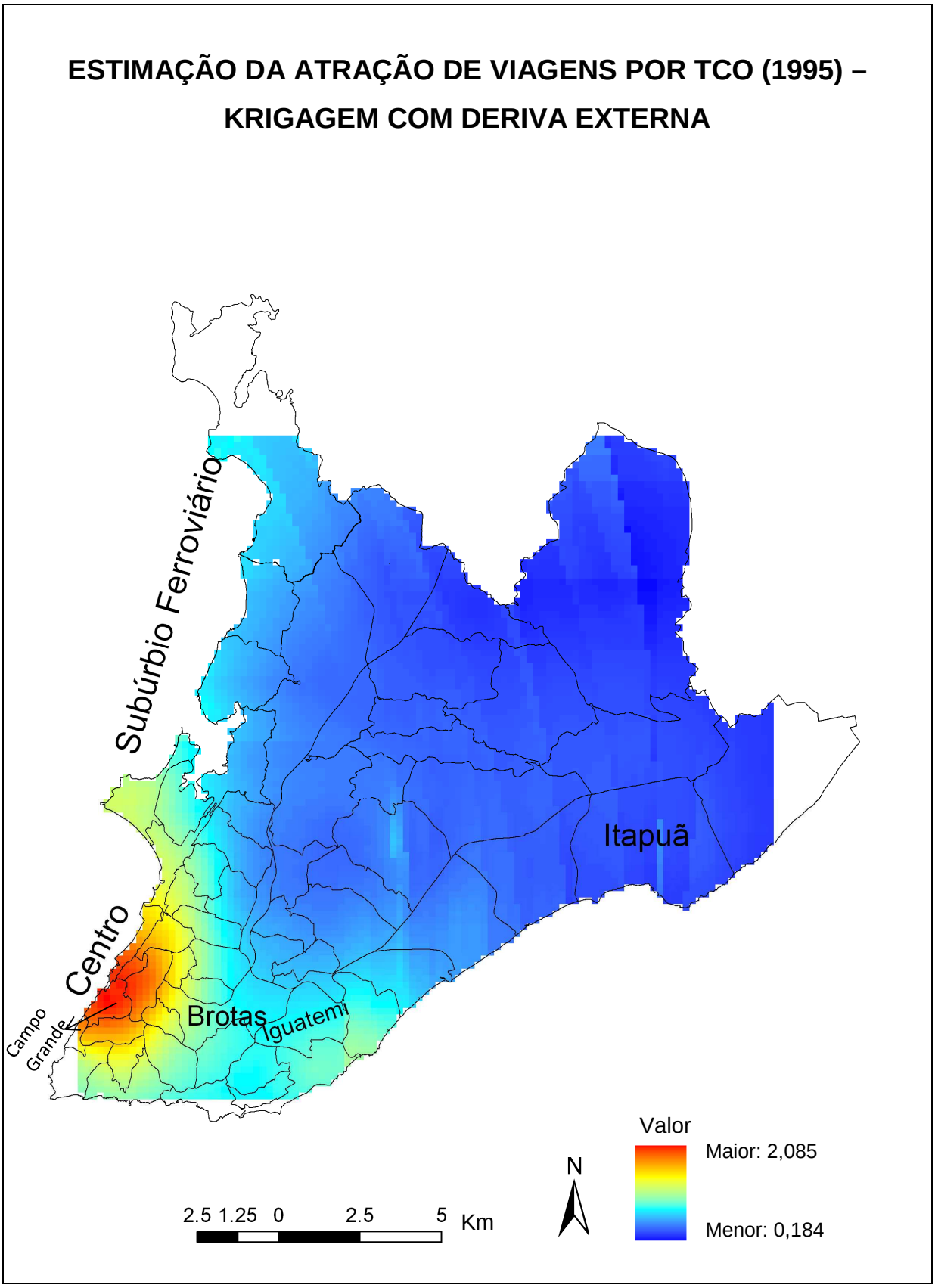


Figura 6.6: Atração de viagens (1995) - amostragem por centróide das MZs - KDE

6.1.5 População e Produção de viagens por TCO 1995 (amostragem sistemática artificial)

Os resultados obtidos a partir da validação cruzada da estimação a partir da amostragem sistemática artificial se encontram na Tabela 6.5. Nela estão apresentados os resultados da estimação da produção de viagens através da krigagem ordinária e da krigagem com deriva externa.

O primeiro ponto a ser considerado a partir da análise desta tabela é que o coeficiente de correlação de *Pearson* teve significativo aumento para as três estimações, se comparado com o mesmo dado obtido pela amostragem original por centróide das MZs. A média do quadrado dos erros é baixa para as três estimações, sobretudo para a estimação da produção de viagens pela KDE.

Tabela 6.5: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (1995) – amostragem sistemática artificial

Tipo de Krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média dos quadrados dos erros
KO	População	0,80	-0,004	0,51	0,41
	Produção de viagens por TCO	0,84	-0,004	0,38	0,25
KDE	Produção de viagens por TCO	0,86	0,006	0,35	0,22

Os demais resultados são satisfatórios, com média de erros e da variância próximas de zero. Com a aplicação da Krigagem com Deriva Externa, coeficiente de correlação de Pearson teve pouca alteração em relação à estimação da produção de viagens feita pela técnica de KO.

Esta melhora no coeficiente de correlação é perceptível através dos diagramas de dispersão entre dados observados e estimados para cada variável, apresentadas nos anexos deste trabalho.

Para esta análise, a amostragem sistemática demonstrou que seus resultados melhoram a estimação. Analisando os mesmos dados para a amostragem por centróide por MZs, percebe-se que os dados estimados por amostragem sistemática

artificial possuem maior relação entre si. A melhora significativa dos resultados é intuitiva, haja vista que, em geral, os variogramas experimentais apresentam maior estrutura espacial com o aumento de amostras observadas. Importante deixar claro que a desagregação de dados aqui proposta desconsidera, de certa forma, 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.

6.1.6 Interpolação de mapas de krigagem: População e Produção de viagens (1995) amostragem sistemática artificial

A Figura 6.7 apresenta o mapa de interpolação da variável de população 1995. Os mapas da variável produção de viagens por TCO por KO e pela KDE se encontram nas Figuras 6.8 e 6.9, respectivamente. Através destas figuras é possível notar que os dados estimados das duas variáveis possuem o mesmo padrão espacial verificado no mapa de pontos apresentado no Capítulo 5 desta dissertação. Isto demonstra que a estimação ganhou mais qualidade através da amostragem sistemática artificial. No mapa de interpolação através da técnica de Krigagem com deriva externa houve um notável ajuste do padrão espacial, sobretudo naquelas áreas que apresentaram problemas na estimação pela amostragem por centroides. Através desta nova amostragem foi possível obter resultados bastante satisfatórios, tanto através dos dados da validação cruzada quanto visualmente, através dos mapas de interpolação.

Nos mapas apresentados é possível perceber a relação entre a quantidade populacional e a produção de viagens, bem como a influência da renda na escolha do modo de transporte, já que a concentração de viagens produzidas por TCO ocorre nas áreas mais pobres.

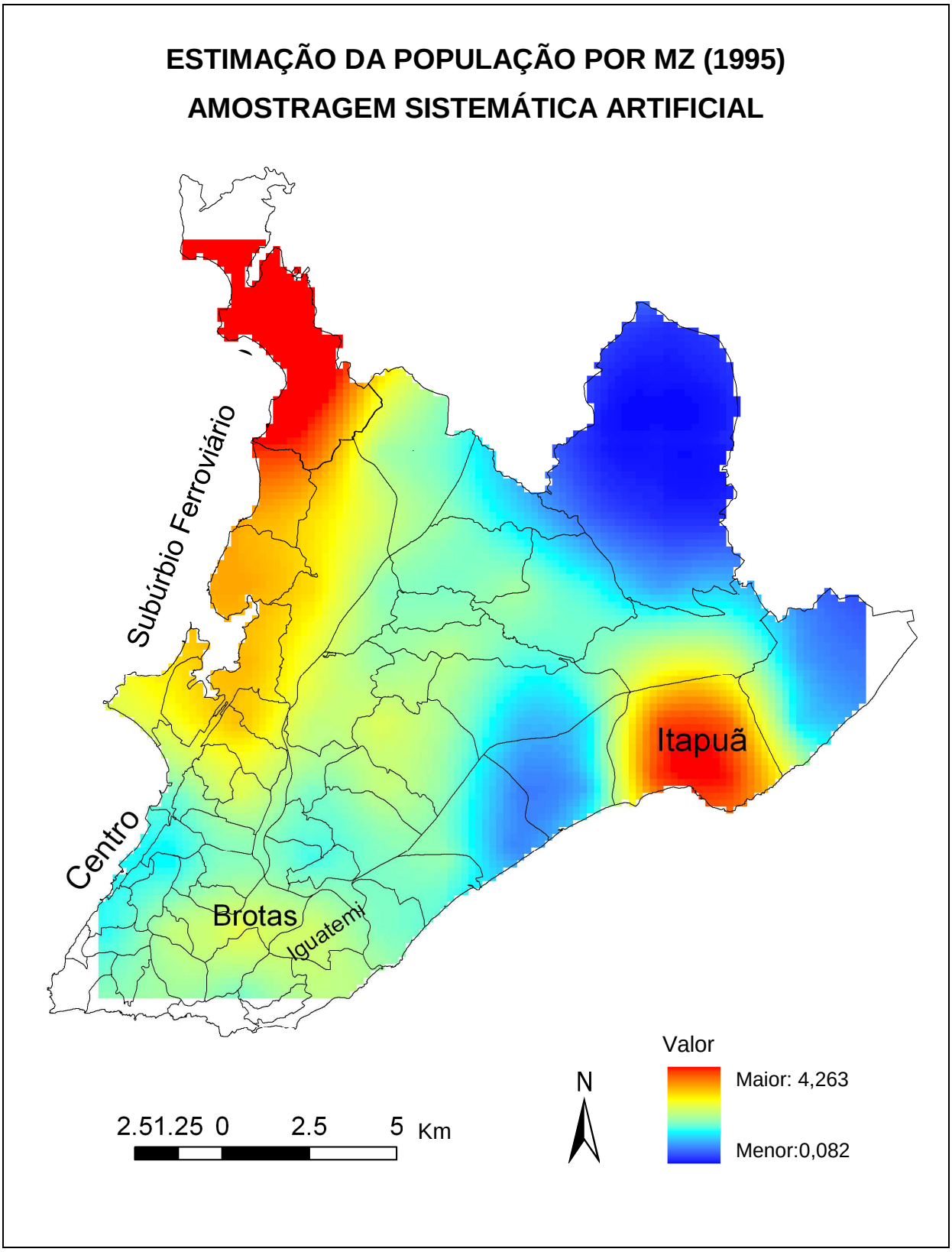


Figura 6.7: Mapa de interpolação população (1995) - amostragem sistemática artificial - KO

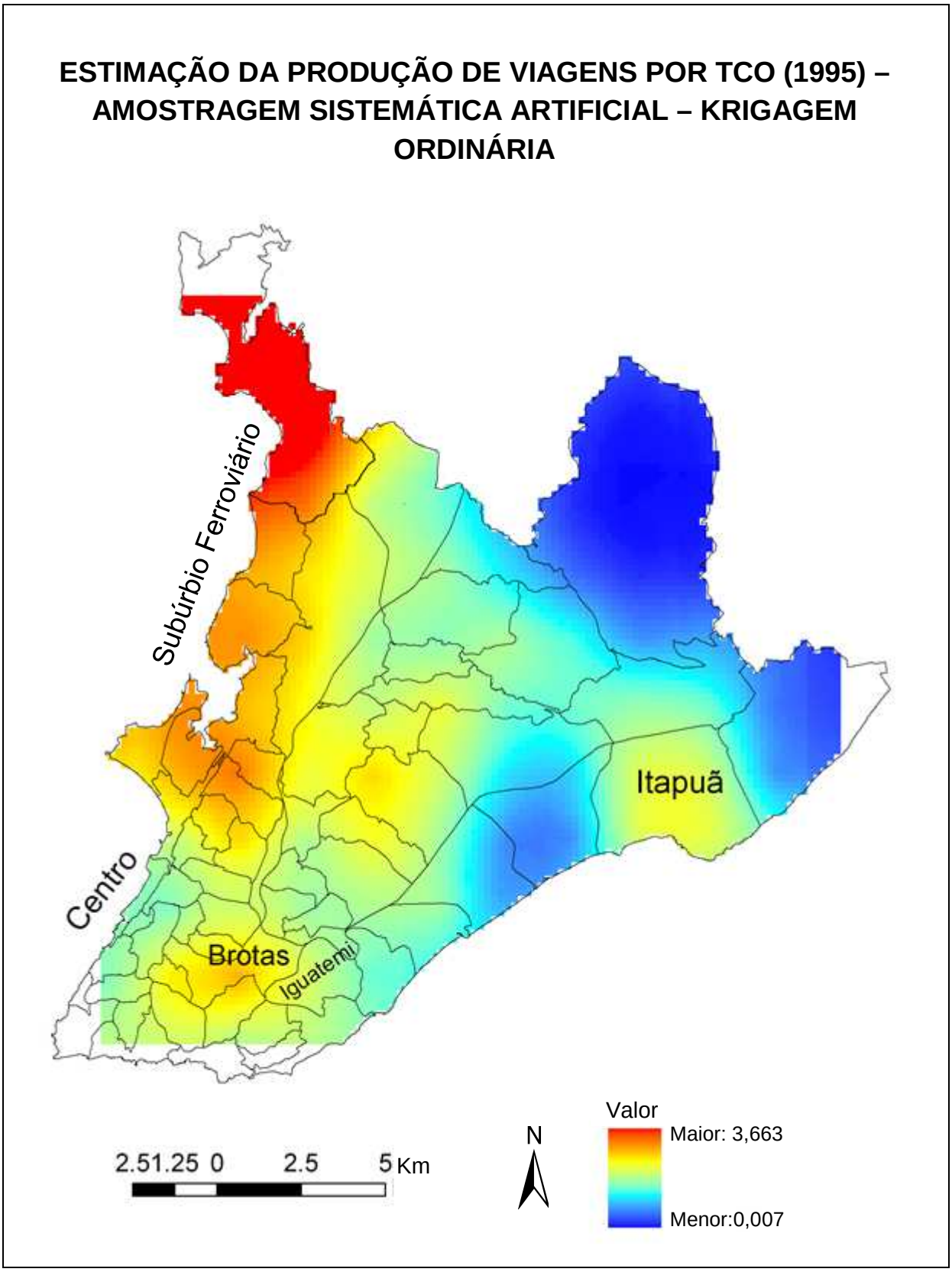


Figura 6.8: Mapa de interpolação da variável produção de viagens por TCO (1995) - amostragem sistemática artificial - KO

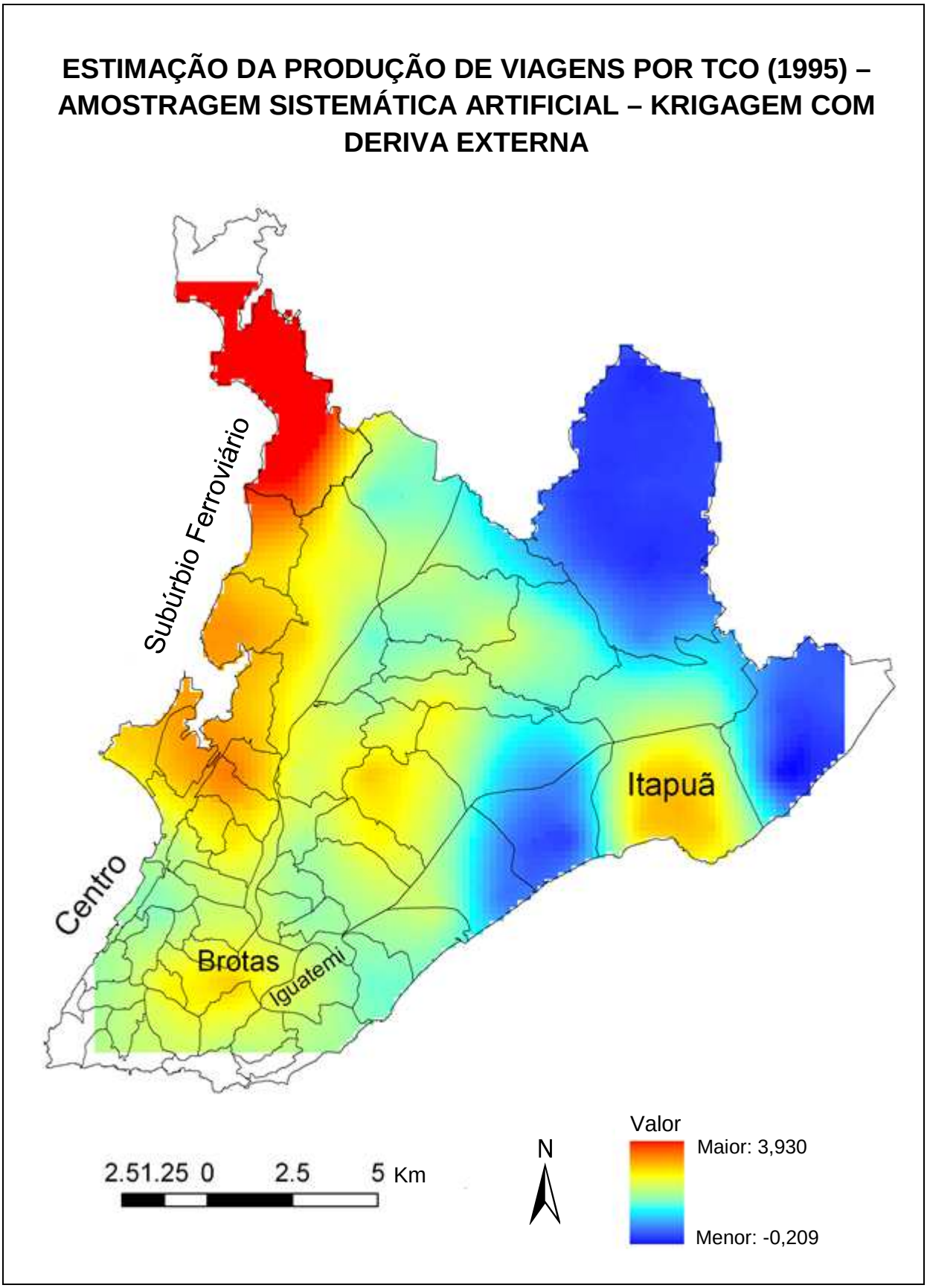


Figura 6.9: Mapa de interpolação da variável produção de viagens por TCO (1995) - amostragem sistemática artificial – KDE

6.1.7 Empregos e Atração de viagens por TCO 1995 (amostragem sistemática artificial)

A validação cruzada da estimação das variáveis empregos e atração de viagens por TCO a partir da amostragem sistemática artificial apresentou melhores resultados quando comparados àqueles obtidos através da amostra original por centroides das MZs das mesmas variáveis (Tabela 6.6). O coeficiente de correlação de Pearson aumentou consideravelmente nas três estimações. A média do quadrado dos erros diminuiu nas três estimações, se comparada à amostragem por MZs.

As outras médias apresentaram bons resultados, sendo que média da variância entre dados observados e estimados diminuiu para as três análises.

Tabela 6.6: Dados obtidos através da Validação Cruzada da estimação das variáveis empregos e atração de viagens (1995) – amostragem sistemática artificial

Tipo de Krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média dos quadrados dos erros
KO	Empregos	0,73	-0,010	0,27	0,20
	Atração de viagens por TCO	0,79	-0,006	0,11	0,10
KDE	Atração de viagens por TCO	0,77	-0,020	0,40	0,11

No geral, os resultados da validação cruzada obtidos pela amostragem sistemática artificial podem ser considerados bons, mesmo que para estas duas variáveis o ajuste do modelo teórico dos variogramas não tenham sido tão adequado para estimativas através da krigagem, devido à falta de padrão espacial e falta de simetria dos dados, como analisado no Capítulo 5.

6.1.8 Interpolação de mapas de krigagem: Empregos e Atração de viagens (1995) amostragem sistemática artificial

A partir dos diagramas de dispersão, apresentados nos anexos, e dos resultados da validação cruzada desta análise, pode-se considerar que houve uma melhora nos resultados em relação à amostragem original. Assim como nos mapas de pontos

apresentados no capítulo anterior, as maiores taxas de atração de viagens por TCO (horário de pico da manhã) e de empregos se apresentam na região da AUC, principalmente nas áreas do centro antigo e Iguatemi. Em outras áreas da RMS, a MZ Itapuã, por exemplo, apresenta-se como única MZ que apresenta certo destaque em número de empregos (Figura 6.10).

Tanto no mapa de interpolação obtido através da krigagem ordinária (Figura 6.11), quanto pela krigagem com deriva externa (Figura 6.12) as duas regiões que se destacam são o Centro e a área do Iguatemi. No entanto, observa-se que na área do centro os dados elevados extrapolam a área de outras regiões, representadas pela quadrícula da malha, que possuem valores de atração de viagens menores. Esta falta de continuidade ou de relação espacial acaba refletindo nos variogramas, e consequentemente, nos mapas interpolados.

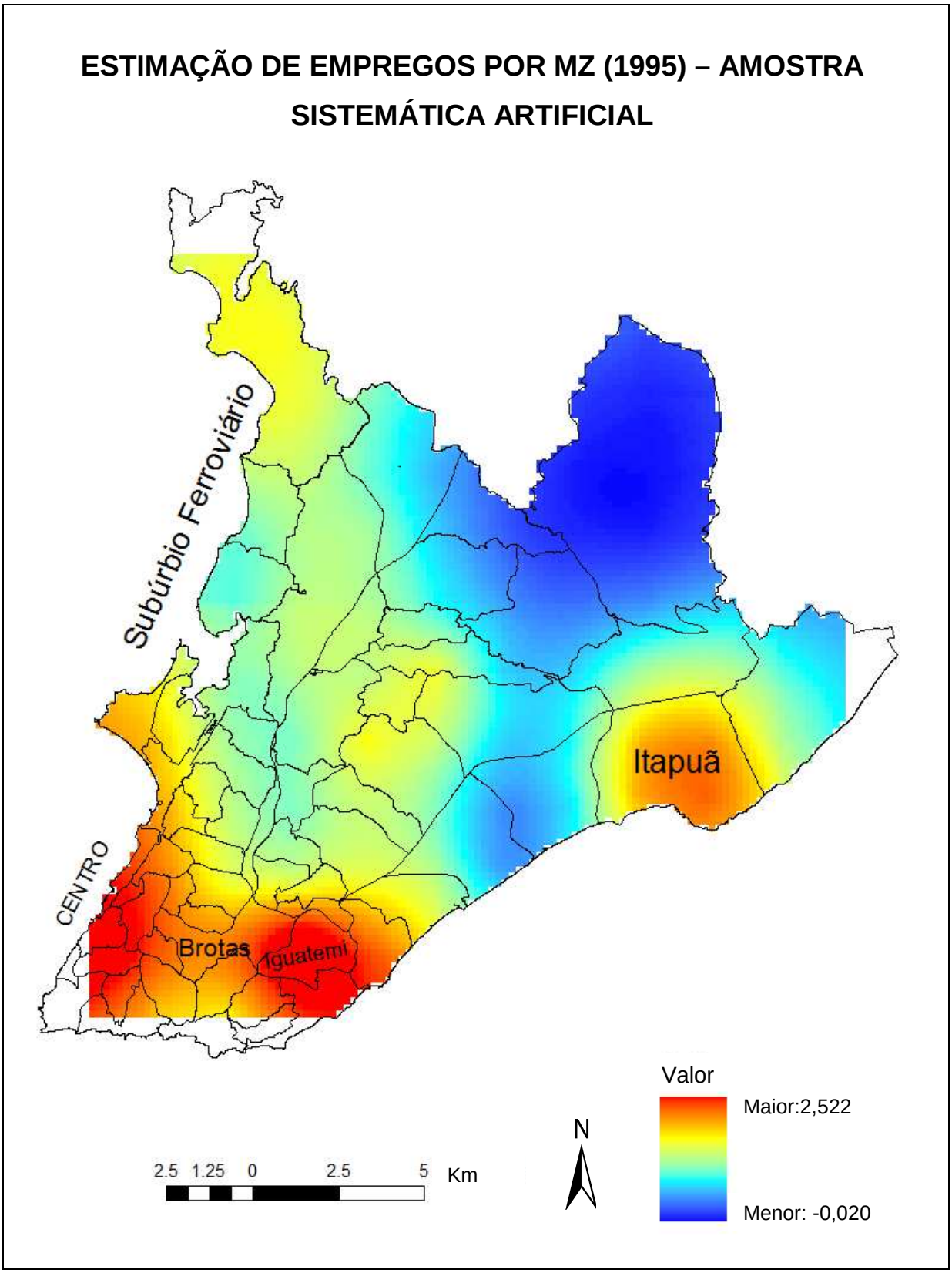


Figura 6.10: Mapa de interpolação da variável empregos (1995) - amostragem sistemática artificial - KO

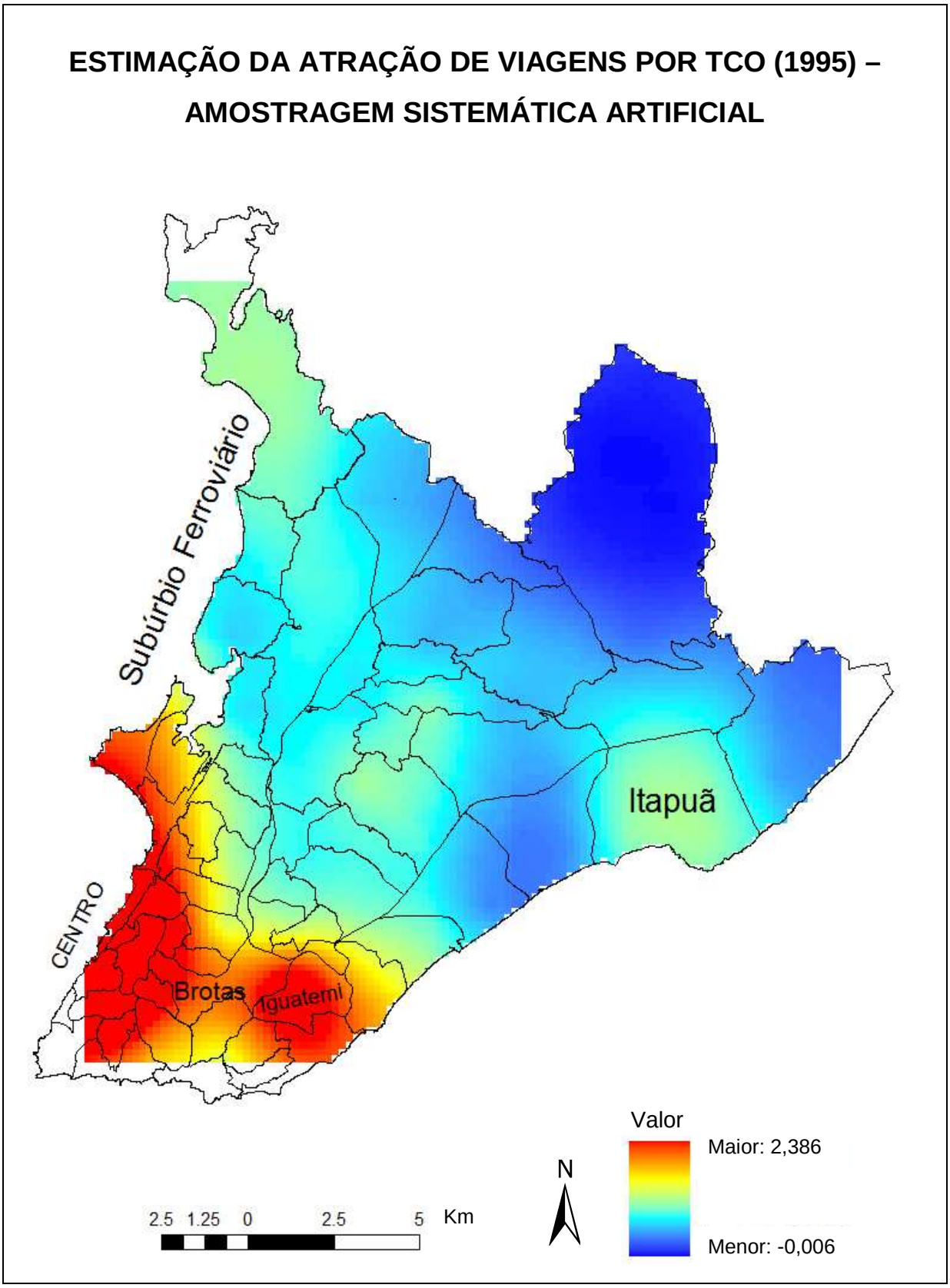


Figura 6.11: Mapa de interpolação da variável atração de viagens por TCO (1995) - amostragem sistemática artificial - KO

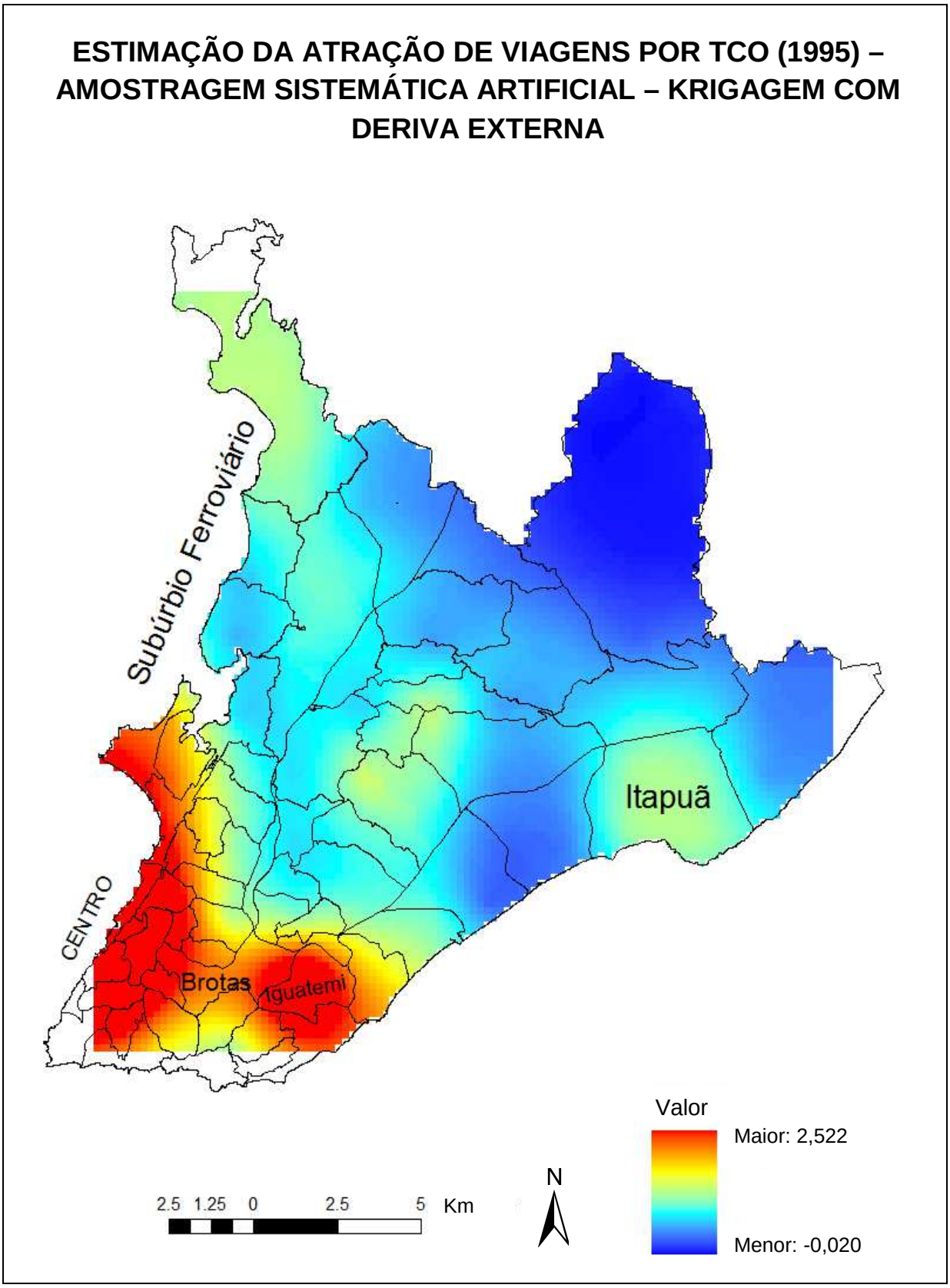


Figura 6.12: Mapa de interpolação da variável atração de viagens por TCO (1995) - amostragem sistemática artificial - KDE

6.1.9 População e Produção de viagens por TCO 2010 (amostra por MZs)

Esta seção compreende a análise dos resultados das krigagens dos dados de população (IBGE, 2010) e de produção de viagens por TCO estimados, a partir da compatibilização de setores censitários com o zoneamento de tráfego referente à Pesquisa OD 1995.

Os resultados obtidos através da validação cruzada estão na Tabela 6.7. O coeficiente de correlação é o mesmo para as variáveis população e produção de viagens, visto que este último dado foi estimado através de dados de população, levando-se em conta alta correlação entre estas variáveis (coeficiente de *Pearson* superior a 0,9). A estimação da produção de viagens a partir da krigagem com deriva externa apresentou significativo aumento de correlação entre dados observados e estimados (de 0,27 para 0,67). Os dados da variável secundária permitiram uma estimação mais precisa em coordenadas desconhecidas.

Tabela 6.7: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (2010)

Tipo de krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média dos quadrados dos erros
KO	População	0,27	-0,019	0,55	0,97
	Produção de viagens por TCO	0,27	-0,015	0,53	0,98
KDE	Produção de viagens por TCO	0,67	-0,020	0,55	0,59

A média de erros e da variância apresentaram bons resultados, assim como a média do quadrado dos erros, com valores próximos de 1 para a estimação da população e produção de viagens por TCO. Para a estimação por KDE o valor foi um pouco mais baixo 0,59.

6.1.10 Interpolação de mapas: População e Produção de viagens (2010) amostragem original por centróides das MZs

Os mapas resultantes da krigagem dos dados de 2010 estão nas Figuras 6.13, 6.14 e 6.15. Pode-se observar que, nestes três mapas, a estimação da população e de produção de viagens por TCO para o ano de 2010, possuem o mesmo padrão espacial. Além de já ter sido detectado que estas duas variáveis são altamente correlacionadas, os dados de produção de viagens são oriundos dos dados de população. Isto também influenciou os resultados, tanto a validação cruzada, quanto a interpolação dos mapas da KO e da KDE.

Portanto, há uma clara tendência de continuidade espacial que segue por todo Subúrbio Ferroviário até a MZ Coutos, que possui valores de população mais elevados da RMS. Estas áreas também possuem a menor renda média em Salvador, logo se subentende que há uma maior quantidade de viagens produzidas por transporte público, já que a renda influencia a escolha modal. O oposto observa-se na região da AUC, principalmente nas áreas do centro.

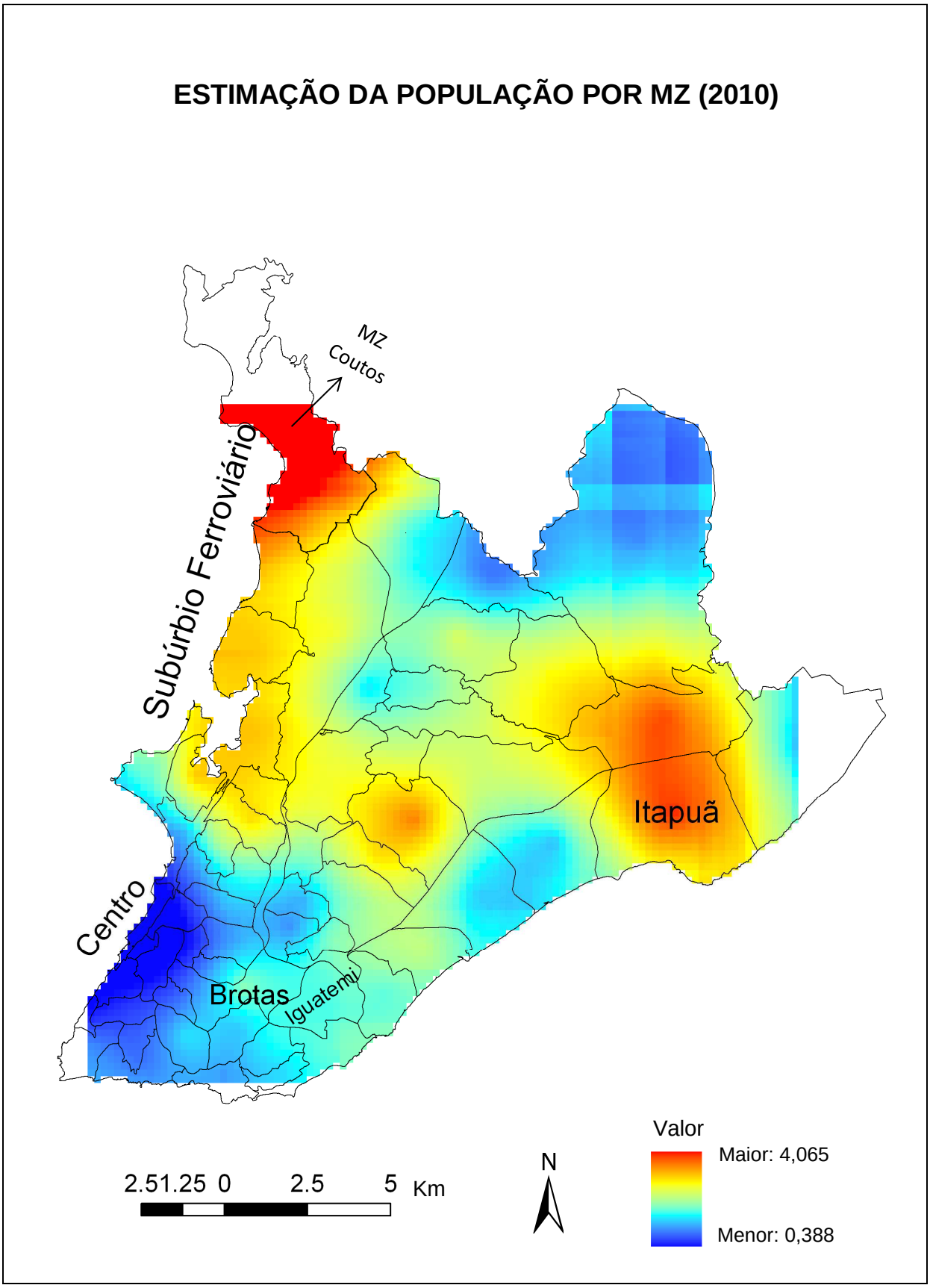


Figura 6.13: Mapa de interpolação da variável população (2010) - KO

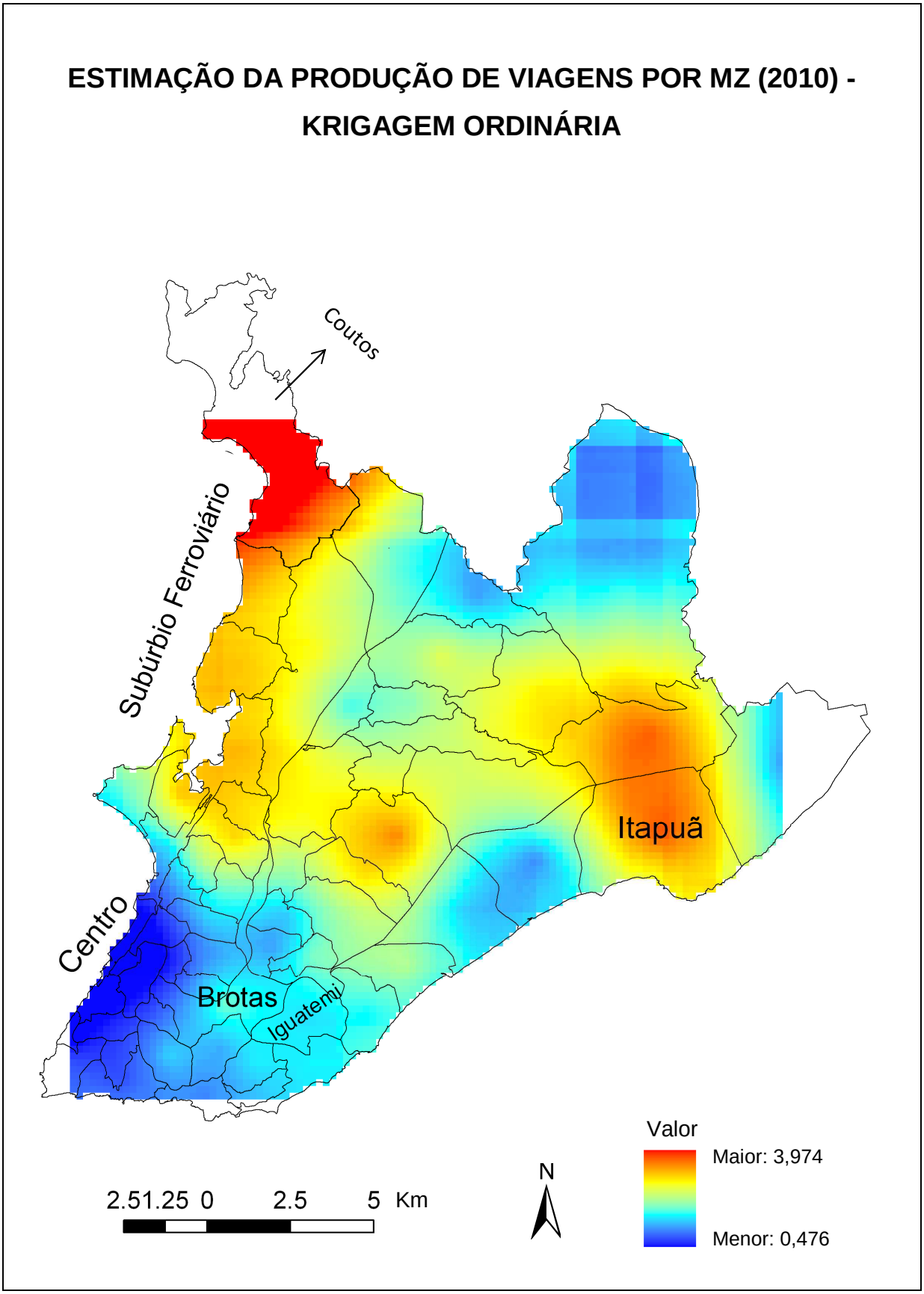


Figura 6.14: Mapa de interpolação da variável produção de viagens por TCO (2010) - KO

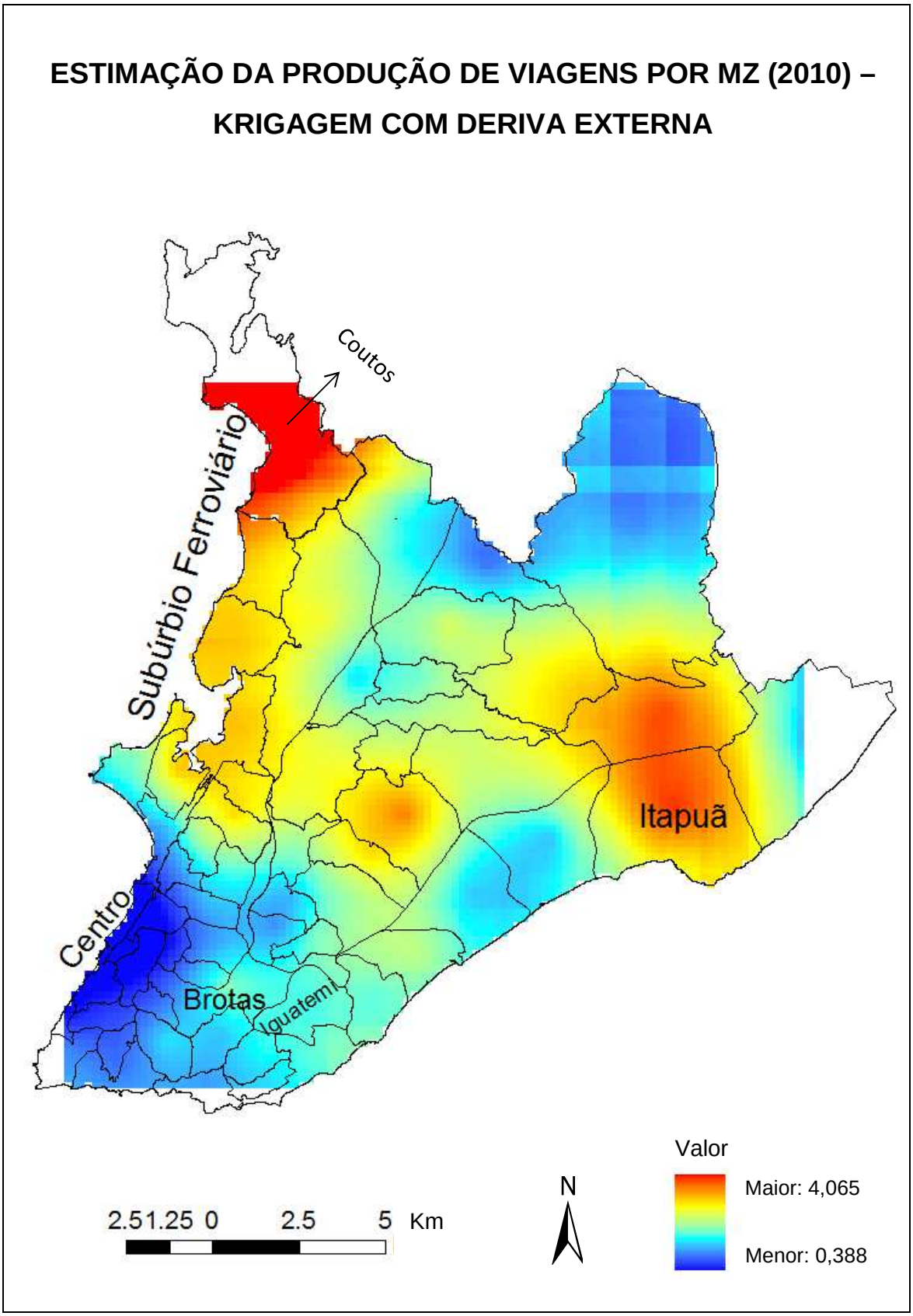


Figura 6.15: Mapa de interpolação da variável produção de viagens por TCO (2010) - KDE

6.1.11 População e Produção de viagens por TCO 2010 (amostra sistemática artificial)

Como observado nas análises anteriores feitas através da amostragem sistemática artificial, o coeficiente de correlação entre valores observados e estimados apresentou melhores resultados quando comparado aos resultados da amostragem por centróide das MZs. Na Tabela 6.8 estão os resultados da validação cruzada paras a krigagem ordinária dos dados de população e produção de viagens por TCO, e também para esta variável estimada a partir da krigagem com deriva externa. No entanto, a KDE não melhorou o resultado deste coeficiente, que apresenta um valor um pouco mais baixo do que o obtido pela krigagem ordinária, no entanto, superior ao mesmo coeficiente da análise por centróide das MZs. Os demais resultados da validação cruzada são bons, assim como foi para as demais análises. A média do quadrado dos erros ficou abaixo dos resultados obtidos de 1995 para a amostragem sistemática artificial, sendo que este resultado obtido através da KDE é um pouco maior.

Tabela 6.8: Dados obtidos através da Validação Cruzada da estimação das variáveis população e produção de viagens (2010) – amostragem sistemática artificial

Tipo de krigagem	Variável	Coeficiente de correlação	Média de erros	Média da variância	Média dos quadrados dos erros
KO	População	0,87	-0,003	0,48	0,29
	Produção de viagens por TCO	0,86	-0,002	0,47	0,30
KDE	Produção de viagens por TCO	0,84	-0,002	0,41	0,34

6.1.12 Interpolação de mapas: População e Produção de viagens (2010) amostragem sistemática artificial

Os mapas resultantes da interpolação da krigagem pela amostragem sistemática artificial possuem o mesmo padrão espacial. A Figura 6.16 representa a interpolação da estimação da população por MZ de tráfego em 2010. Este mapa mostra que as áreas que possuem maior número de população, assim como em 1995, continuam

sendo as áreas do Subúrbio Ferroviário, Itapuã e Miolo, e as áreas com menor quantidade de população são as áreas do centro e Iguatemi, como também a MZ Parque Senhor do Bonfim, no extremo norte de Salvador.

Os mapas de interpolação da produção de viagens por KO e KDE encontram-se nas figuras 6.17 e 6.18, respectivamente. Estes dois mapas apresentam as mesmas características do mapa de população, no entanto, na MZ Itapuã e em alguns pontos da região do Miolo, por provável influência da variável secundária, apresentam valores um pouco maiores (Figura 6.18).

Pode-se considerar que tanto a amostragem por centróide quanto pela malha regular representaram bem a realidade da RMS nos dois períodos analisados.

Vale ressaltar, sobretudo, que a dificuldade de obter bons variogramas, no geral, não fez com que os resultados sejam considerados necessariamente ruins. Além disso, dados sociais nem sempre possuem distribuição espacial ideal para análises geoestatísticas. O que se tentou mostrar através deste conjunto de análises foi uma busca para aprimorar uso de técnicas de geoestatística para estimações da geração de viagens (dados eminentemente sociais). No próximo capítulo estes fatores estão discutidos mais detalhadamente.

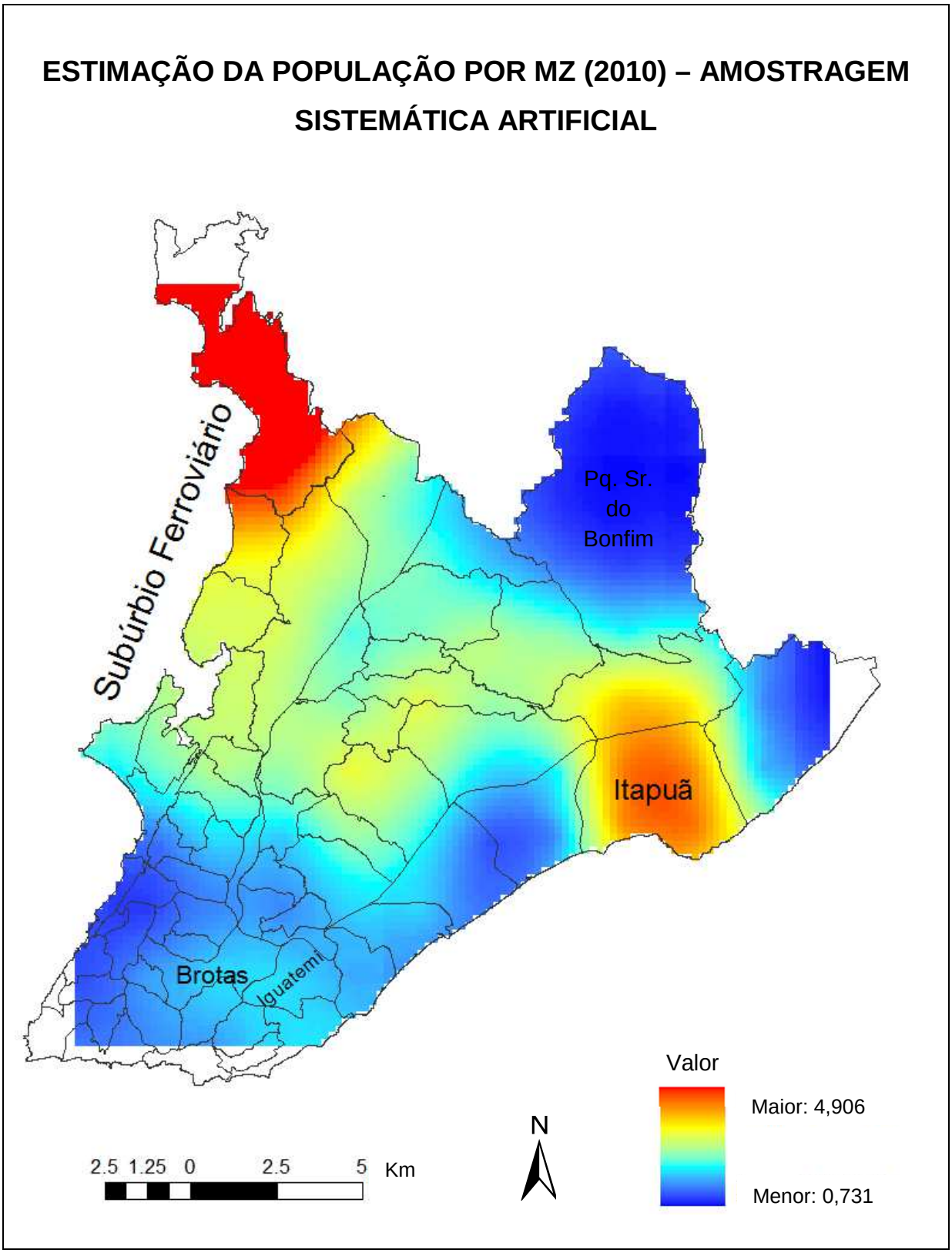


Figura 6.16: Mapa de interpolação da variável população (2010) - amostra sistemática artificial - KO

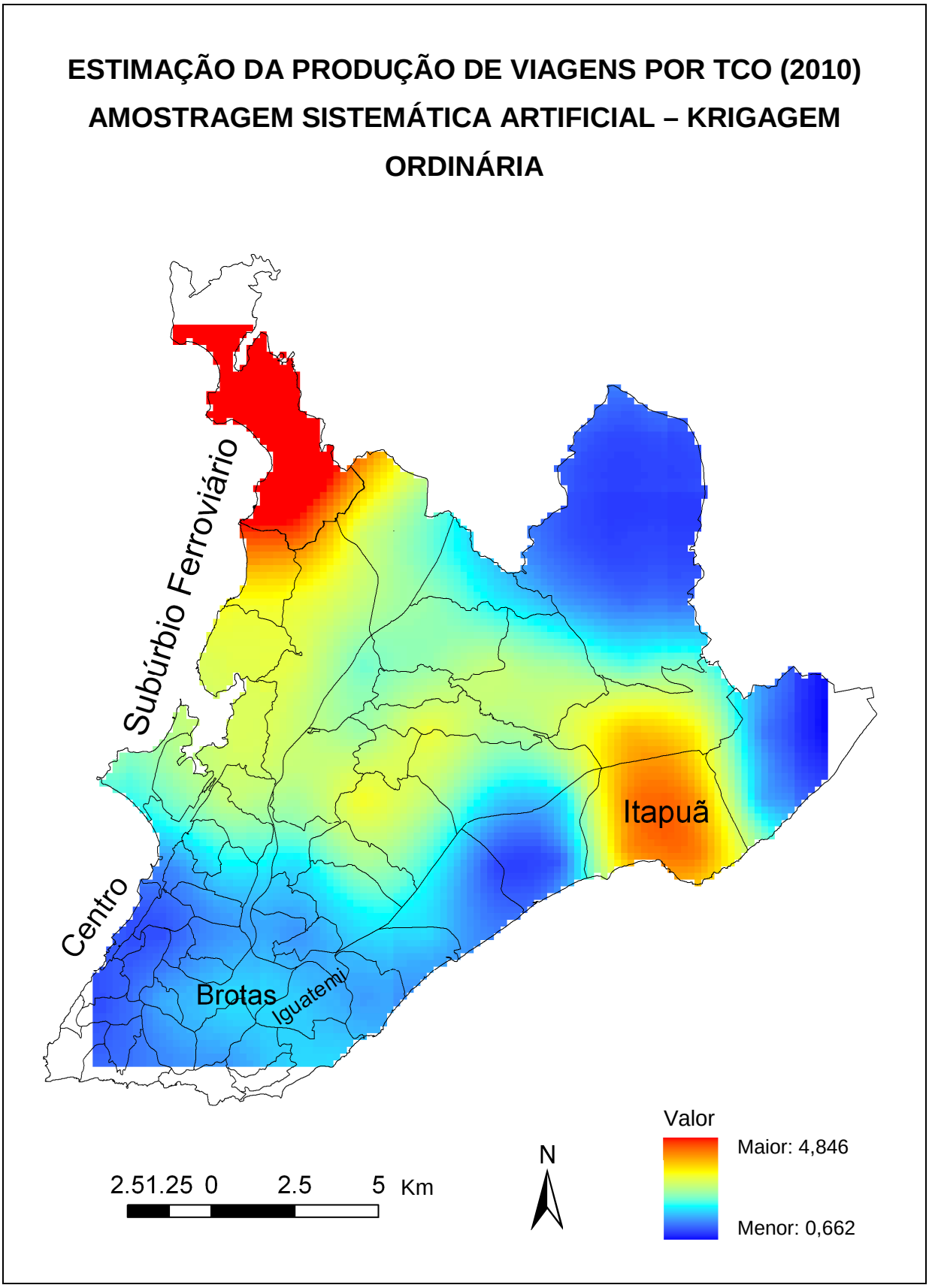


Figura 6.17: Mapa de interpolação da variável produção de viagens por TCO (2010) - amostragem sistemática artificial – KO

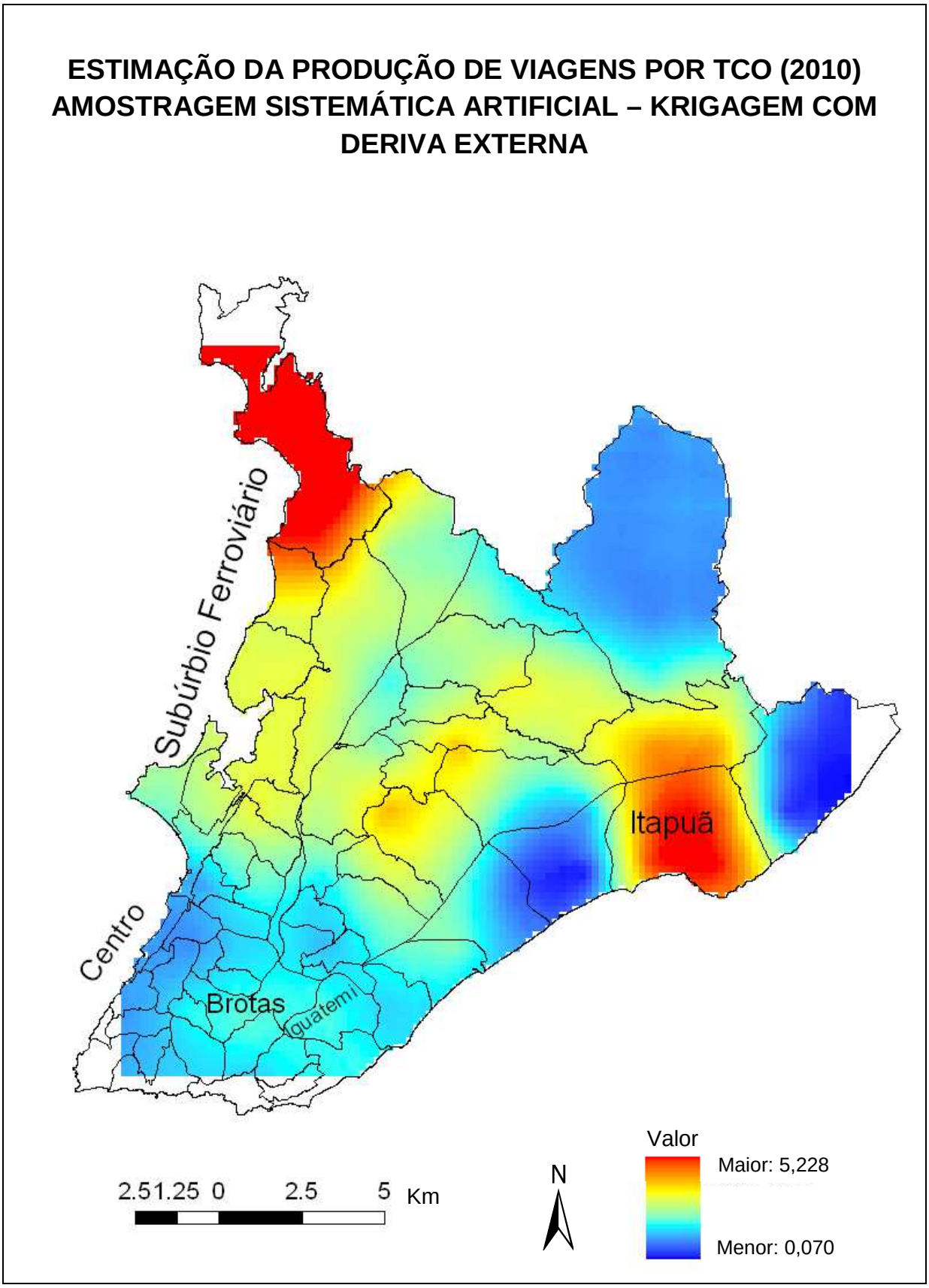
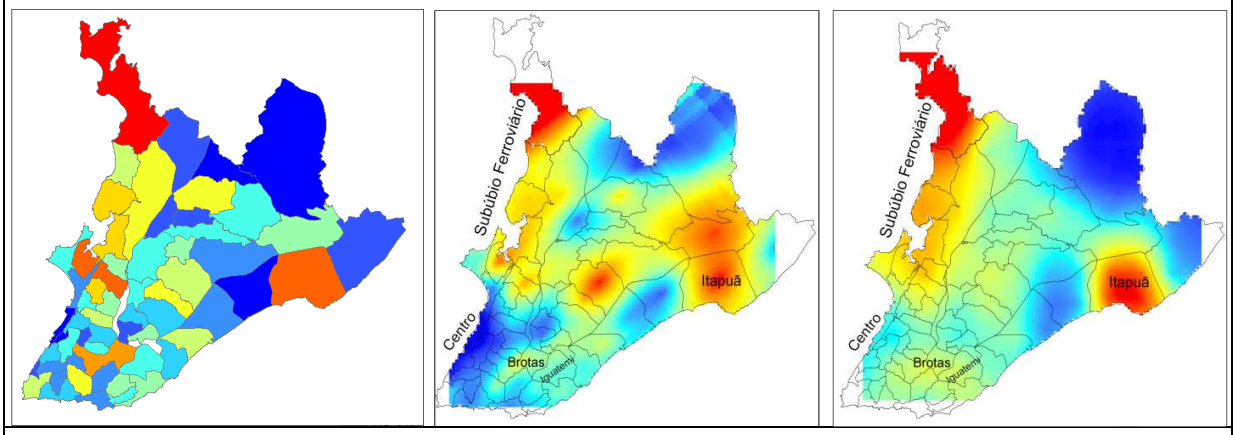


Figura 6.18: Mapa de interpolação da variável produção de viagens por TCO (2010) - amostragem sistemática artificial - KDE

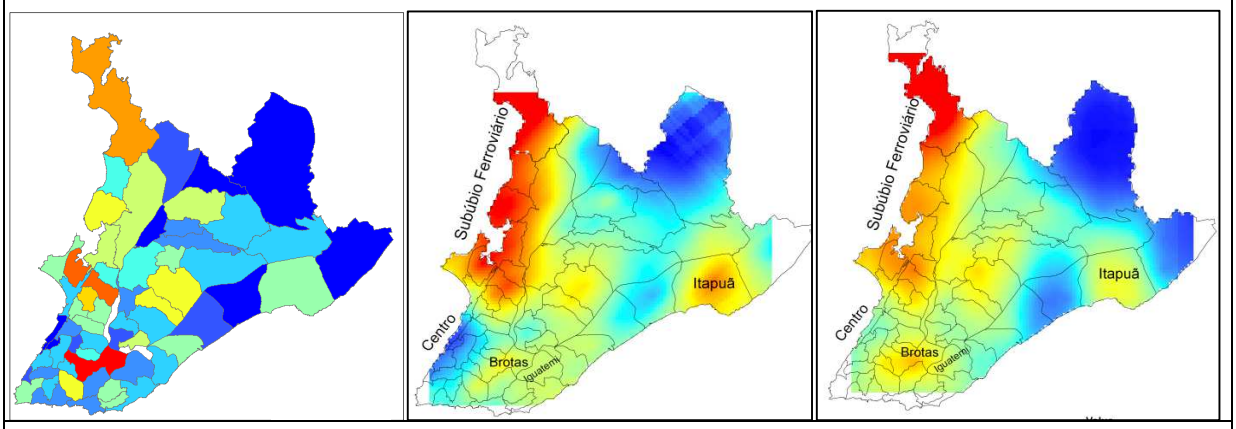
A Figura 6.19 mostra o resultado da interpolação dos mapas krigados pela KO utilizando os diferentes tipos de amostragem.

POPULAÇÃO 1995



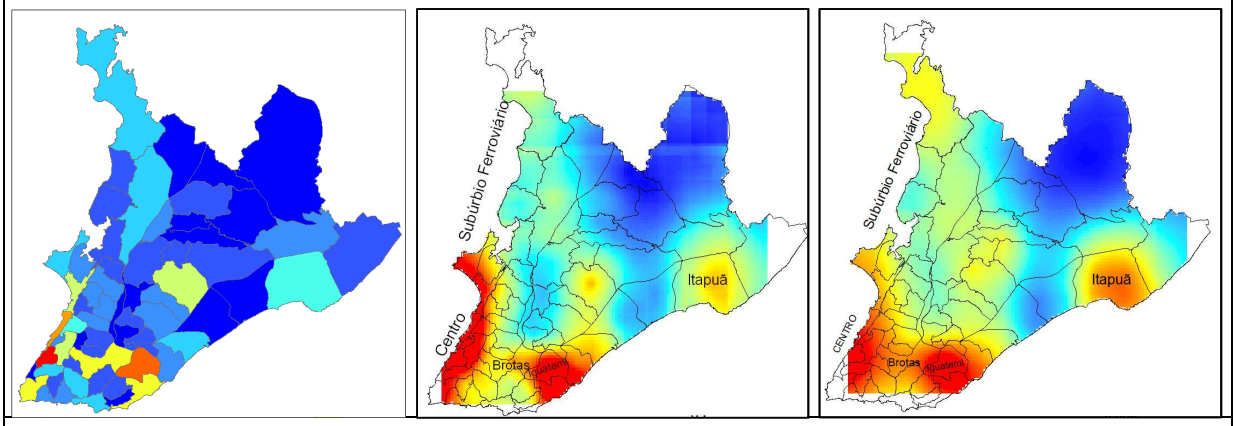
(a) observada por área (b) amostra por centróide (c) amostragem por grade regular

PRODUÇÃO DE VIAGENS POR TCO (1995)



(a) observada por área (b)amostra por centróide (c) amostra por grade regular

EMPREGOS 1995



(a) observada por área (b) amostra por centróide (c) amostra por grade regular

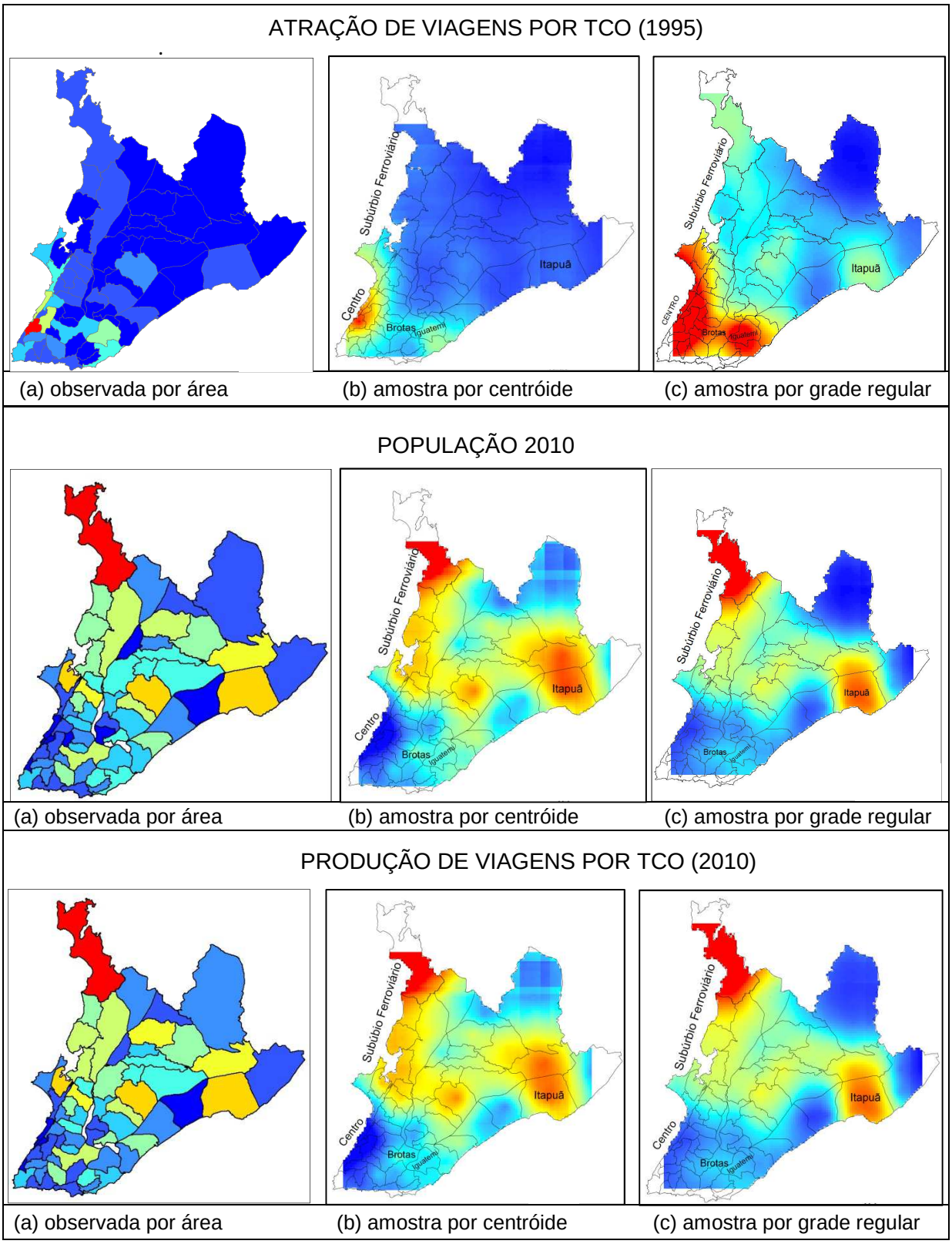


Figura 6.19: Mapas de interpolação krigados por KO: amostras por centróide x amostras por malha regular

A Figura 6.20 mostra os mapas interpolados através da krigagem com deriva externa, a partir do centróide das MZs e por amostragem sistemática artificial.

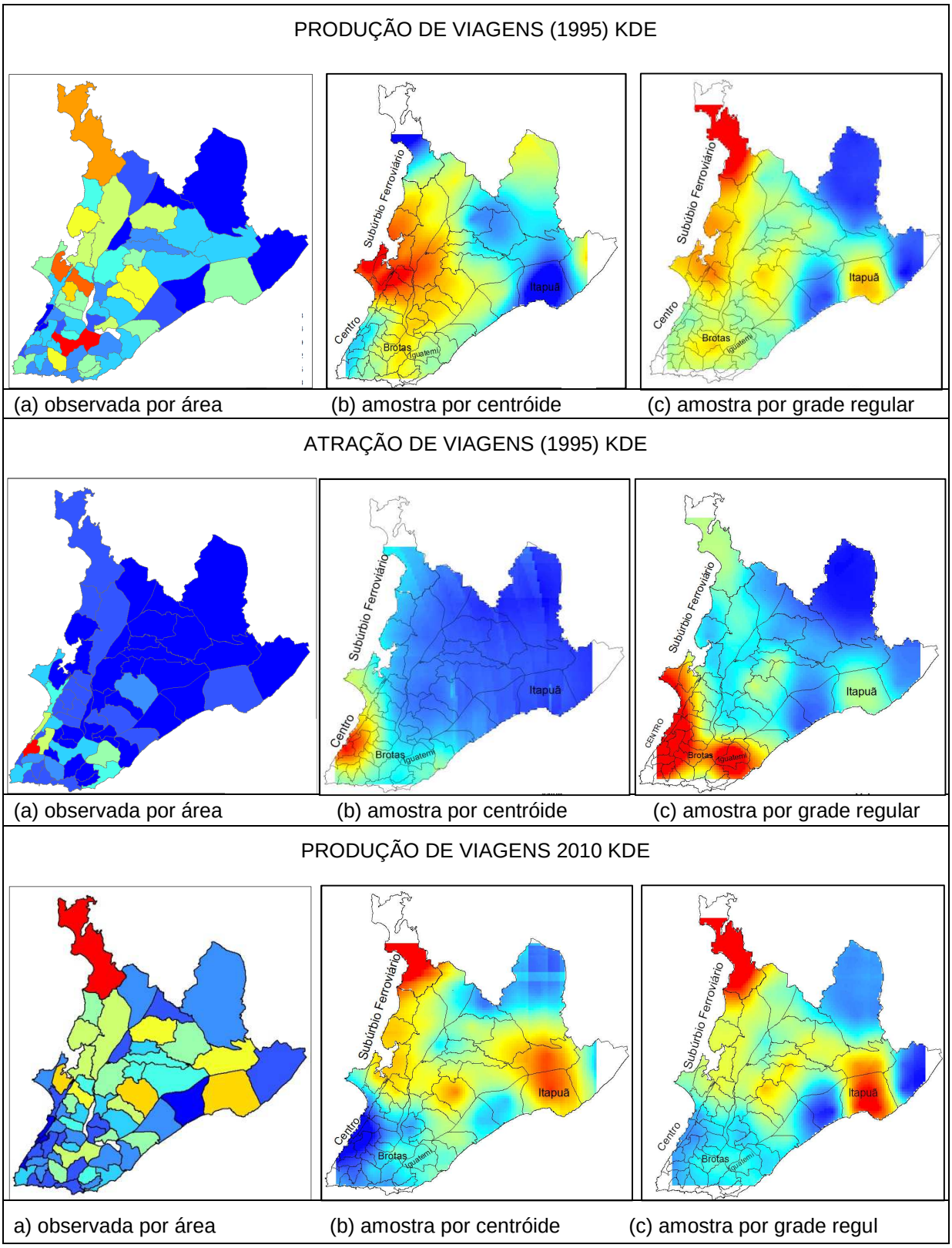


Figura 6.20: Mapas de interpolação krigados por KDE: amostras por centróide x amostras por malha regular

A partir da análise visual apresentada acima, entre as interpolações e os valores observados por área (Figuras “a”), 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. No entanto, neste trabalho, observou-se maior similaridade entre os mapas interpolados e os mapas de valores observados por áreas no caso da amostra agregada em comparação à amostra desagregada. Apesar disso, a validação dos dados desagregados foi mais satisfatória que a amostragem por centróide das MZs.

CAPÍTULO 7: CONCLUSÕES E RECOMENDAÇÕES

Este capítulo apresenta as principais conclusões relativas à aplicação de técnicas geoestatísticas na estimação da geração de viagens. Descreve também as principais vantagens e desvantagens verificadas ao longo do trabalho. Por fim, serão pontuadas algumas sugestões e recomendações para trabalhos futuros, a fim de explorar e aprimorar estas vantagens e/ou mitigar as dificuldades encontradas.

7.1 SÍNTESE DO TRABALHO

O objetivo principal do presente trabalho foi estimar geração de viagens por TCO em coordenadas conhecidas e desconhecidas, através de técnicas de interpolação espacial: Krigagem Ordinária e Krigagem com Deriva Externa. Esta ultima técnicas foi utilizada com o intuito de estimar dados de produção e atração de viagens através de variáveis que podem ser obtidas por outras fontes de dados, como caso de dados de população. O trabalho utilizou quatro diferentes amostras de pontos (dados da Pesquisa Origem/Destino da RMS de 1995 e dados do censo demográfico de 2010 – amostras agregadas e desagregadas), produzindo diversos resultados, mostrados pela Figura 7.1.

A proposta deste trabalho foi empregar técnicas pouco utilizadas no estudo de transportes, especificamente na geração de viagens. Os resultados obtidos a partir da aplicação da geoestatística para tal finalidade foram considerados satisfatórios, mesmo levando-se em conta que as variáveis relacionadas à demanda por transportes não são teoricamente apropriadas para estimação geoestatística. Vale ressaltar que esta técnica inicialmente foi desenvolvida no âmbito de estudos geológicos, e teve seu desenvolvimento vinculado a questões das ciências ambientais ou naturais.

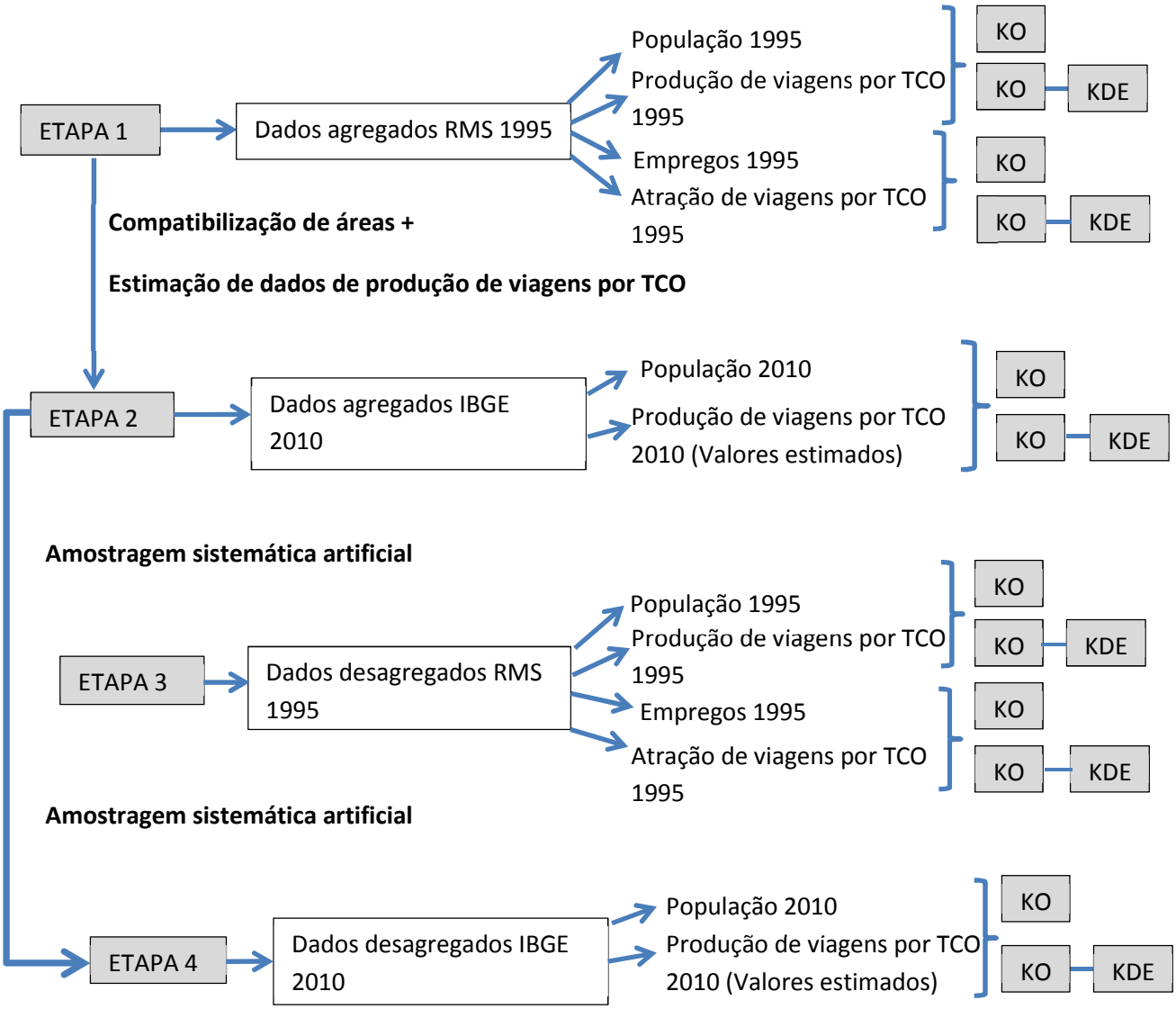


Figura 7.1: Resumo dos dados e resultados

Algumas dificuldades foram encontradas, sobretudo, na modelagem dos variogramas, principalmente aqueles que envolvem a previsão da atração de viagens. Diante destas dificuldades, e das várias tentativas de modelar os variogramas em diferentes direções, foi proposta uma análise através de uma amostragem diferente da utilizada inicialmente. Esta nova amostragem contribuiu para desagregar os dados analisados, que anteriormente eram representados por centróides de MZs de tráfego. Logo, verifica-se que a alta agregação de dados amostrais, como é o caso da primeira análise, prejudica a modelagem dos variogramas, devido ao alto espaçamento entre amostras, além da quantidade reduzida de observações. Considerando os resultados como o coeficiente de correlação entre dados observados e estimados, com esta nova amostragem houve

significativa melhora neste parâmetro para todas as estimativas realizadas por krigagem ordinária. No entanto, há ressalvas importantes em relação à amostragem sistemática artificial aqui proposta: (1) Dados de transportes e socioeconômicos podem sofrer alterações abruptas, não tendo uma continuidade espacial tal como solos ou água, por exemplo; (2) a produção da nova amostra seria mais adequada por meio de simulação, que depende da distribuição populacional correspondente aos dados amostrais; (3) o ideal seria o trabalho com amostras desagregadas realmente coletadas, tais como domicílios georreferenciados.

A krigagem com deriva externa foi outra forma de aprimorar as estimativas da geração de viagens através de variáveis altamente correlacionadas com produção e atração de viagens, respectivamente. Esta técnica demonstrou resultados bem melhores que a estimativa com krigagem ordinária. Por exemplo, o coeficiente de correlação aumentou bastante para as estimativas feitas pela amostragem original (considerando os centróides das MZs), e manteve-se alto para a amostragem sistemática artificial (em torno de 0.8). A Tabela 7.1 traz uma síntese da validação cruzada de todas as estimações realizadas.

Tabela 7.1: Quadro resumo da Validação Cruzada

Amostragem por centróides das MZs		
Variáveis	**Correlação KO**	**Correlação KDE**
População 95	0.08	0.62
Produção de viagens 95	0.21	
Empregos 95	0.19	0.49
Atração de viagens 95	0.41	
População 2010	0.27	0.67
Produção de viagens 2010	0.27	
Amostragem sistemática artificial		
Variáveis	**Correlação KO**	**Correlação KDE**
População 95	0.81	0.86
Produção de viagens 95	0.84	
Empregos 95	0.73	0.77
Atração de viagens 95	0.80	
População 2010	0.87	0.84
Produção de viagens 2010	0.87	

Para a estimação da produção de viagens (2010) através da amostragem sistemática artificial, a krigagem com deriva externa resultou em coeficientes de correlação um pouco menores do que os obtidos por krigagem ordinária. No entanto, isso não significa que tais estimações tenham piorado, visto que outros parâmetros da validação cruzada, como a média de erros e da variância continuaram bons. Além disso, deve-se considerar que os dados de produção de viagens por TCO para o ano de 2010 foram dados estimados através das variáveis independentes população e renda do IBGE. Portanto, vale ressaltar que não se tratam de dados observados, o que fez dos resultados uma estimativa derivada de dados também estimados. Considerando resultados da atração de viagens por malha regular e estimados pela KDE, considera-se, principalmente, como dificuldade na obtenção de melhores resultados, o problema da falta de uma aparente continuidade espacial dos dados e a existência de valores discrepantes.

7.2 PRINCIPAIS CONCLUSÕES E CONTRIBUIÇÕES

A aplicação da krigagem para estimação da geração de viagens atingiu os objetivos esperados neste trabalho. Foi possível estimar mais de 10 mil novas células em coordenadas desconhecidas, para a amostragem por centróide das MZs, e mais de 12 mil novas células para amostragem sistemática (malha regular), com média de erros e variância baixas. É importante considerar que, com a aplicação da krigagem, se deixou de conhecer as variáveis apenas pelo valor observado por pontos, e em locais definidos no espaço, e passou-se a conhecer uma superfície. Portanto, conclui-se que a Geoestatística, apesar de pouco utilizada no âmbito de estudos de transporte, é uma ferramenta viável para estimar dados de geração de viagens, podendo contribuir, desta forma, para pesquisas na área, bem como aperfeiçoar técnicas de planejamento e gestão do transporte urbano.

Outra importante contribuição deste trabalho foi a introdução do método de amostragem sistemática artificial. Contudo, esta técnica não foi introduzida no presente trabalho como já sendo um método consistente e de excelência. Como já foi explicitado anteriormente, ela foi uma tentativa de obter estimativas com melhores resultados de correlação entre dados observados e estimados. É

importante considerar que a amostragem sistemática artificial respondeu positivamente à proposta inicial de auxílio para a obtenção de dados mais precisos. Vale considerar que esta técnica ainda precisa ser aprimorada.

Este trabalho contribui para o meio científico e acadêmico, já que propõe a aplicação de uma técnica ainda pouco explorada na área de transportes, além de viabilizar a estimação de geração de viagens em coordenadas desconhecidas, inclusive em análise espaço-temporal. Outro fator a ser considerado é a importância da utilização do SIG, que permite a construção da superfície krigada das taxas de geração de viagens.

A utilização da krigagem com deriva externa possibilitou prever dados de viagens utilizando variáveis correlacionadas provenientes de outras fontes, que não a Pesquisa OD, o que geralmente não acontece. Isso pode facilitar a obtenção de dados de viagens urbanas através de dados facilmente disponíveis, como dados censitários.

7.3 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

A Geoestatística é uma técnica que pode ser mais intensamente explorada no planejamento de transportes, já que se considera que dados relacionados a viagens possuem autocorrelação espacial. Desta forma, dentro da Geoestatística há inúmeras técnicas que podem ser testadas e utilizadas, a fim de revelar qual delas melhor proporcionará estimativas que representem a realidade de uma área. Existem inúmeros métodos que podem proporcionar diferentes resultados, dependendo do tipo de dado e de amostragem.

Alguns métodos propõem a transformação de dados, dentre eles a krigagem lognormal, que pode ser utilizada para os casos de dados assimétricos, ou seja, quando o conjunto de dados possui grande quantidade de valores baixos e apenas alguns valores elevados, ou vice-versa. Este fato foi verificado na estimação da atração de viagens. Detectado este problema, pode ocorrer que valores altos influenciem nos resultados de áreas com valores baixos. Neste trabalho não foi possível estimar dados atuais de atração de viagens, apenas produção de viagens,

devido à disponibilidade da variável população em dados do censo. Neste tipo de pesquisa não são encontrados dados de empregos nem de matrículas escolares, o que inviabilizou a estimação da atração de viagens para 2010. Portanto, futuramente se faz necessário buscar dados relacionados à atração de viagens, como empregos (visto neste trabalho como variável altamente correlacionada) e/ou matriculas escolares. Possivelmente, estes dados podem ser encontrados em secretarias de educação dos municípios ou em subsecretarias especificas das prefeituras.

Não havendo dados de viagens atuais disponíveis até o momento, é importante verificar futuramente se os resultados encontrados neste trabalho correspondem aos dados obtidos através da pesquisa de mobilidade, realizada em Salvador em 2013. Poderá ser realizado também o ajuste do modelo de regressão utilizado neste trabalho.

Podem ser citadas ainda, diversas potenciais aplicações de técnicas geoestatísticas em estudos de transportes:

- Análise de escolha modal através de Krigagem Indicativa;
- Obtenção de amostras desagregadas através de simulação;
- Interpolação espacial através de resultados de modelos de escolha discreta;
- Interpolação espacial para previsão de acidentes de trânsito.

REFERÊNCIAS

ADJEMIAN, M. K.; LIN, C.; WILLIAMS, J. (2010) *Estimating spatial interdependence in automobile type choice with survey data.* Transportation Research Part A: Policy and Practice. Elsevier, v. 44, p. 661-675.

ALMEIDA, R.; BERNARDI, J. V. E.; BASTOS, W. R.; NASCIMENTO, E. L.; OLIVEIRA, R. C.; CARVALHO, D. P. (2007) *Método geoestatístico para modelagem ambiental de poluentes em sistemas lacustres – Amazônia Ocidental.* Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26 abril 2007, INPE, p. 2247-2253.

BARBETTA, P. A. (2005) *Estatística aplicada às ciências sociais*. 5 ed. Florianópolis: UFSC, 340 p.

BHAT, C.; ZHAO, H. (2002) *The spatial analysis of activity stop generation.* Transportation Research Part B: v. 36, p. 557-575.

BOLDUC, D.; DEGENAIS, M. G.; GAUDRY, M. J. (1989) *Spatially autocorrelated errors in origin-destination models: A new specification applied to aggregate mode choice.* Transportation Research Part B: Methodological, Elsevier, v. 23B, n. 5, p. 361-372.

BRUTON, M. J. (1979) *Introdução ao planejamento dos transportes.* Rio de Janeiro: Interciência.

CÂMARA, G.; CARVALHO, M.S.; CRUZ, O.G.; CORREIA, V. (2002) *Análise espacial de áreas*. Em: Análise espacial de dados geográficos, Eds. Fuks, S.D.; CARVALHO, M.S.; CÂMARA, G.; DRUCK, S.; CARVALHO, M.S.; CÂMARA, G.; MONTEIRO, A.V.N – Divisão de processamento de imagens – Instituto Nacional de Pesquisas Espaciais – São José dos Campos, 2002. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/cap5-areas.pdf>. Acesso em: 02 de agosto de 2013.

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, 2004. p.46-82.

CAMPOS, V. B. G. (2007). *Planejamento de transportes: Conceitos e modelos de análise.* (Apostila) 66 p.

CAMPOS, V. B. G. (2013) *Planejamento de transportes: conceitos e modelos*. 1. Ed. Rio de Janeiro: Interciencia, 188 p.

COSTA, A. S. G. (2013) *Proposta de um método para estimação de escolha modal através da geoestatística.* 143 f. Dissertação (Mestrado) – Universidade Federal da Bahia, Salvador, 2013.

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

DOUBLEDAY, C. (1976) *Some studies of the temporal stability of Person trip generation models*. England: University of Cambridge, v.11, p 255- 263.

EOM, J. K.; PARK, M. S.; HEO, T.; HUNTSINGER, L. F. (2006) *Improving the prediction of annual average daily traffic for nonfreeway facilities by applying a spatial statistical method.* Transportation Research Record: Journal of the Transportation Research Board, N. 1968, Transportation Research Board of the National Academies, Washigton, D. C., p. 20-29.

FELGUEIRAS, C. A. (1999) *Modelagem ambiental com tratamento de incertezas em sistemas de informação geográfica: o paradigma geoestatístico por indicação.* 215 p. Tese (Doutorado). São José dos Campos: INPE, 1999.

FERNANDES, J. A. (2009). *Krigagem com deriva externa aplicada à avaliação de recursos minerais de calcário de minério laterítico*. 75 f. Dissertação (Mestrado). USP: São Paulo, 2009.

FERREIRA, A. C. M.; RAIA, A. A. (2010) *Análise das taxas de mortalidade por atropelamento dos municípios paulistas utilizando ferramentas de estatística espacial*. Revista dos Transportes Públicos – ANTP, ano 32, p. 61-74.

GOES, M. A.; POSSA, M. V.; LUZ, A. B. (1991) *Amostragem de minérios*. Rio de Janeiro: CTEN, 48 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA - IBGE. (2000) *Censo Demográfico Brasileiro 2000*. Disponível em: < http://www.ibge.gov.br/home/estatistica/populacao/censo2000/default.shtm>. Acessado em: 25 de março 2014.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA – IBGE. (2010) *Censo Demográfico Brasileiro 2010.* Disponível em: < http://downloads.ibge.gov.br/downloads_top.php>. Acesso em 04 de novembro 2013.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – IPEA. (2014) *Malha digital dos municípios do Estado da Bahia.* Disponível em: <http://www.ipea.gov.br/ipeageo/malhas.html> Acesso em: 29 de março 2014.

JAKOB, A. A. E.; YOUNG, A. F. (2006) *O uso de métodos de interpolação espacial de dados nas análises sociodemográficas.* In: Encontro Nacional de Estudos Populacionais. Anais... Caxambu: XV ABEP, 2006.

JOURNEL, A. G. (1989) *Fundamentals of geoestatistic in five lessons*. Washington: American Geophysical Union, 40 p. (Short Course in Geology, v. 8).

KAWAMOTO, E. (1994) *Análise de sistemas de transporte.* 2 ed. (Apostila). Departamento de Transportes. Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos.

LANDIM, P. M. B.; STURARO, J.R.; MONTEIRO, R.C. (2002a) *Exemplos de aplicação da cokrigagem*. DGA, IGCE, UNESP/Rio Claro, Lab. Geomatemática, Texto Didático 09, 17 pp. 2002. Disponível em: <http://www.rc.unesp.br/igce/aplicada/textodi.html>. Acesso em: 12 de março 2014.

LANDIM, P.M.B. STURARO, J.R. (2002b) *Krigagem indicativa aplicada à elaboração de mapas probabilísticos de riscos.* DGA, IGCE, UNESP/Rio Claro, Lab. Geomatemática, Texto Didático 06, 19 pp. 2002. Disponível em: < http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/kindicativa.pdf>. Acesso em: 12 de março 2014.

LANDIM, P. M. B.; STURARO, J.R.; MONTEIRO, R.C. (2002c) *Krigagem ordinária para situações com tendência regionalizada.* DGA, IGCE, UNESP/Rio Claro, Lab. Geomatemática, Texto Didático 07, 12 pp. 2002. Disponível em: < http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/tkrigagem.pdf>. Acessado em: 09 de março 2014.

LANDIM, P.M.B. (2006) *Sobre Geoestatística e mapas*. Terra e Didática, 2(1): 19-33, 2006. Disponível em: <http://www.ige.unicamp.br/terraedidatica/v2/pdf-v2/t_didatica_2006_v02n01_p019-033_landim.pdf >. Acesso em: 12 de março 2014.

LANDIM, P.M.B. (2010) *Introdução à análise estatística de dados geológicos multivariados.* DGD, IGCE, UNESP/Rio Claro, Texto Didático 15, 229 p. 2010. Disponível em: <http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/MULTIVAR.pdf>. Acesso em 14 de março 2014.

LOPES, J. I. O. (2003) *Pós-avaliação da previsão de demanda por transportes no município de Fortaleza*. 179 f. Dissertação (Mestrado). Universidade Federal do Ceará, Fortaleza, 2003.

LOPES, S. B. (2005) *Efeitos da dependência espacial em modelos de previsão de demanda por transporte.* 153 f. Dissertação (Mestrado). Universidade de São Paulo, São Carlos, 2005.

LOPES, S. B.; BRONDINO, N. C. M.; SILVA, A. N. R. (2006) *Análise do desempenho de modelos de regressão espacial na previsão de demanda por transportes*. In: XIV Congresso Panamericano de Ingenharia de Tránsito y Transporte, 2006, Las Palmas de Gran Canaria. Anais... Las Palmas de Gran Canaria: XIV PANAM, 2006.

LOPES, S. B. (2010) *Uma ferramenta para planejamento da mobilidade sustentável com base em modelo de uso do solo e transportes.* 184 f. Tese (Doutorado). Universidade de São Paulo, São Carlos, 2010.

LOPES, S. B.; RODRIGUES DA SILVA, A. N. R. PFAFFENBICHLER, P. C. (2011) *Uma Aplicação do Modelo de Uso do Solo e Transportes MARS no Brasil*. Departamento de Engenharia Civil, Universidade do Minho, Guimarães, Portugal. Revista Engenharia Civil, n. 40, p. 109-120.

MATHERON, G. (1970) *La Théorie des Variables Régionalisées et ses Applications, Les Cahiers de Morphologie Mathématique de Fontainebleau, Fascicule 5*, Thome 1.

MELLO, J.C. (1975) *Planejamento dos transportes.* São Paulo, McGraw-Hill do Brasil.

MEURS, H. (1990) *Dynamic analysis of trip generation*. Transportation Research Part A. v. 24A, n. 6, p. 427-442.

MILLER, H. J. (1999) *Potential Contributions of Spatial Analysis to Geographic Information Systems for Transportation (GIS-T).* Geographical Analysis, The Ohio State University: v. 31, n.4, p. 373-399.

MISSIO, F.; JACOBI, L. F. (2007) *Variáveis dummy: especificações de modelos com parâmetros variáveis.* Curso de Ciências Econômicas/Universidade Federal do Mato Grosso do Sul.

MIURA, H. (2010) *A study of travel time prediction using universal kriging*. Sociedad de Estadística e Investigación Operativa, p. 257-270.

NOVAES, A. G. (1986) *Sistemas de Transportes*. Volume 1: Análise da Demanda. Ed. Edgard Blucher, São Paulo.

ORTÚZAR, J.D.; WILLUMSEN, L.G. (2011). *Modelling Transport*. Wiley, 4th Edition.

PITOMBO, C. S.; SOUSA, A. J. (2009) *Aplicação de Conceitos Geoestatísticos para Análise de Geração de Viagens Urbanas.* In: XXIII Congresso de Pesquisa e Ensino em Transportes - ANPET - Anais... 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. Proceeding sof the 12th WCTR. Lisboa, 2010, p. 25.

RIBEIRO, L. C. (2012) *Os Estados e as Regiões Metropolitanas constitutivas do Observatório das Metrópoles no Censo 2010.* Instituto Nacional de Ciência e Tecnologia CNPq/FAPERJ/CAPES***.*** Observatório das Metrópoles: Rio de Janeiro.

ROCHA, M. M. (2004) *Análise do impacto dos métodos de amostragem na reprodução da textura geológica com base em diferentes distribuições estatísticas.* 165 f. Tese (Doutorado). Universidade de São Paulo, São Paulo, 2004.

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.

SAID, G. M.; YOUNG, D. H. (1989) *A general linear model framework for estimating work trip rates for households in Kuwait*. Transportation Research Part A. v. 24 A, n. 3, p. 187-200.

SANTOS. M. (2009) *A urbanização brasileira*. 5 ed. São Paulo: Editora da Universidade de São Paulo, 176 p.

SETPS (1995) *Pesquisa Domiciliar Origem/Destino.* Salvador.

SHINOHARA, E. J.; RODRIGUES FILHO, O. S.; BERNUCCI, L. L. B.; QUINTANILHA, J. A. (2006). *Uso da ferramenta de análise geoestatística para estudo de atrito em pista de aeroporto.* In: XX ANPET - Congresso de Pesquisa e Ensino em Transportes, 2006, Brasília. Anais do XX ANPET - Congresso de Pesquisa e Ensino em Transportes, 2006. v. 2. p. 1377-1386.

SILVA, G. C. (2006) *Aplicação da geoestatística usando a predição espacial de monóxido de carbono pelo tráfego veicular urbano*. 157f. Dissertação (Mestrado). Programa de Pós-graduação em Engenharia Civil – PPGEC, Universidade Federal de Santa Catarina, Florianópolis, 2006.

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 Planeamento Mineiro.

SOUZA, V. R.; CAVENAGHI, S.; ALVES, J. E.D.; MAGALHÃES, M. A. M. (2008) *Análise espacial dos acidentes de trânsito com vítimas fatais: comparação entre o local de residência e de ocorrência do acidente no Rio de Janeiro.* Revista Brasileira de Estudos de População, São Paulo, v. 25, n. 2, p. 353-364.

SOUZA, L. A. (2013) *Análise Espacial em Modelos de Geração de Viagens*. 90f. Dissertação (Mestrado). Programa de Engenharia de Transportes, Rio de Janeiro: UFRJ/COPPE, 2013.

TEIXEIRA, G. L. (2003) *Uso de Dados Censitários pata 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, 2003.

TOBLER, W. R. (1970) *A computer movie simulating urban growth in the Detroit region.* Economic Geography, v. 26, p. 234-240.

TSUTSUMI, M.; SEYA, H. (2008) *Measuring the impact of large-scale transportation projects on land price using spatial statistical models.* Papers in Regional Science, v. 87, n. 3, p. 385-401.

TRIBUNAL REGIONAL DO TRABALHO – TRT. *Evolução do salário mínimo no Brasil.* Disponível em: < http://www.trt3.jus.br/informe/calculos/minimo.htm>. Acesso em: 14 de novembro 2013.

WACKERNAGEL, H. (2010) *Multivariate Geostatistics: An introduction with applications*. Springer, Third Edition.

WANG, C. *et al.* (2011) *Spatial models in transport: a review and assessment of methodological issues*. 91st Annual Meeting of the Transportation Research Board, Washington DC.

WANG, Y.; ZHUANG, D.; LIU, H. (2012) *Spatial Distribution of Floating Car Speed.* Journal of Transportation Systems Engineering and Information Technology. v.12, p. 36-41,.

WEBSTER, R.; OLIVER, M. (2001) *Geostatistics for environmental scientists*. England: John Wiley & Sons, p. 271.

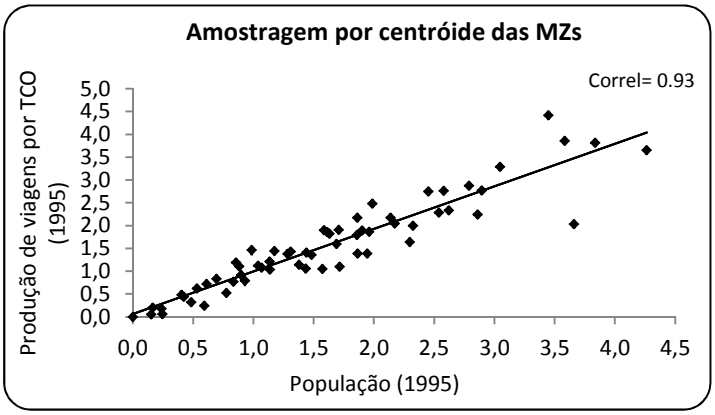
YAMAMOTO, J. K. (2009) *Pós-processamento das Estimativas de Krigagem com Deriva Externa.* São Paulo: Revista do Instituto de Geociências – USP. Série Científica, v.9, n.3, p. 89-101.

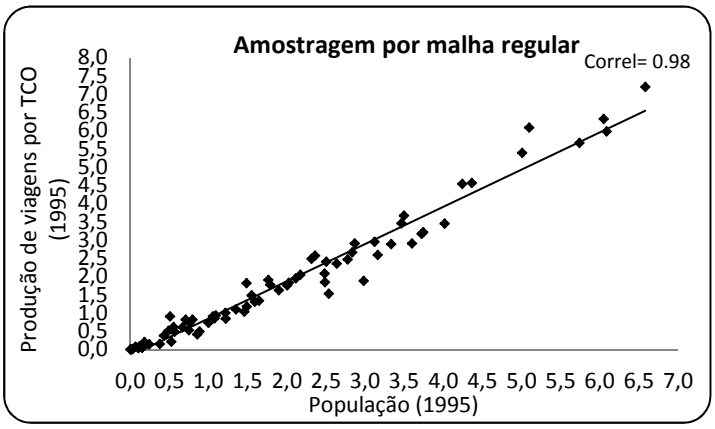
YAMAMOTO, J. K.; LANDIM, P. M, B. *(2013) Geoestatística: conceitos e aplicações.* São Paulo: Oficina de Textos.

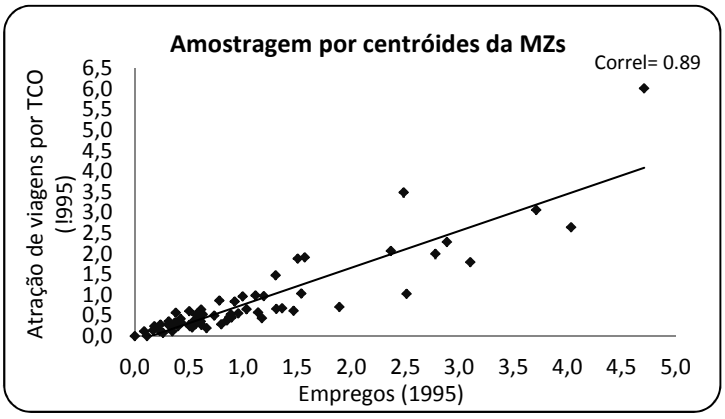
YAO, E.; MORIKAWA, T. (2005) *A study of an integrated intercity travel demand model.* Transportation Research Part A 39, p.367–381.

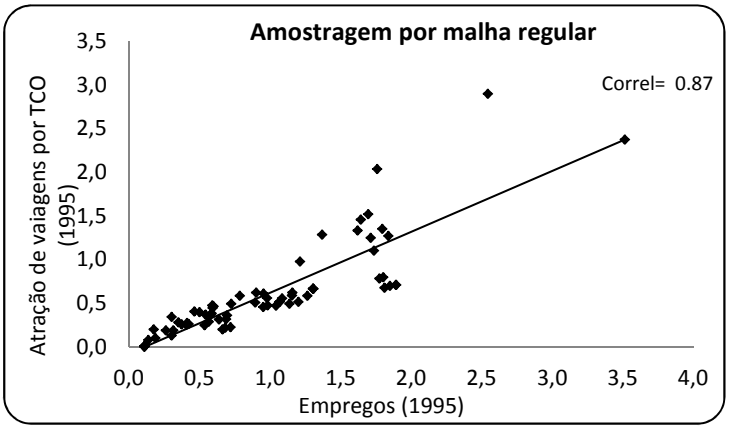
APÊNDICES

Apêndice A: Diagramas de dispersão dos pares de variáveis analisados: dados observados

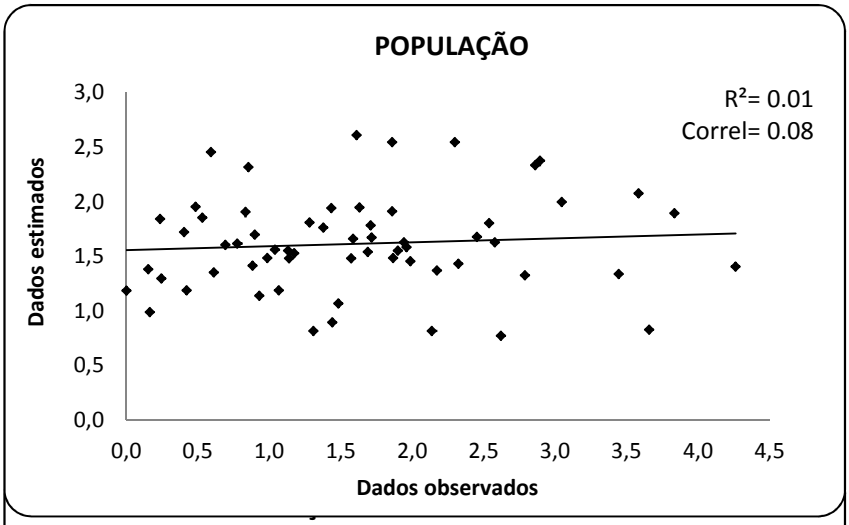


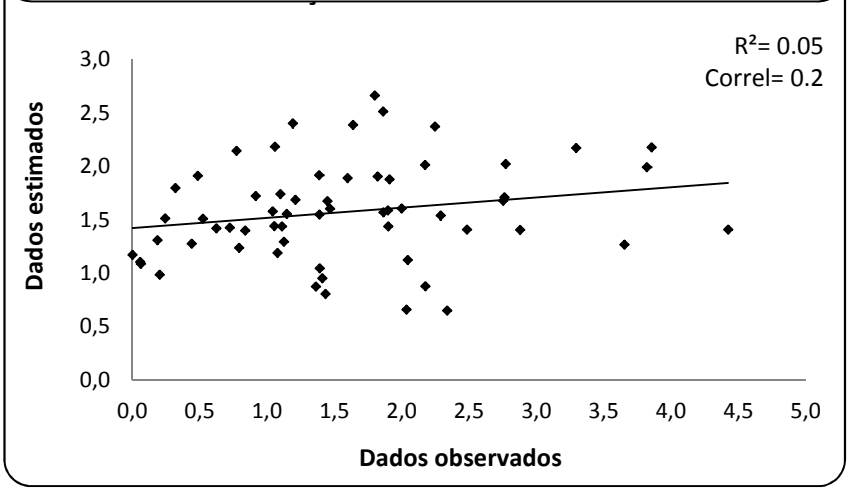


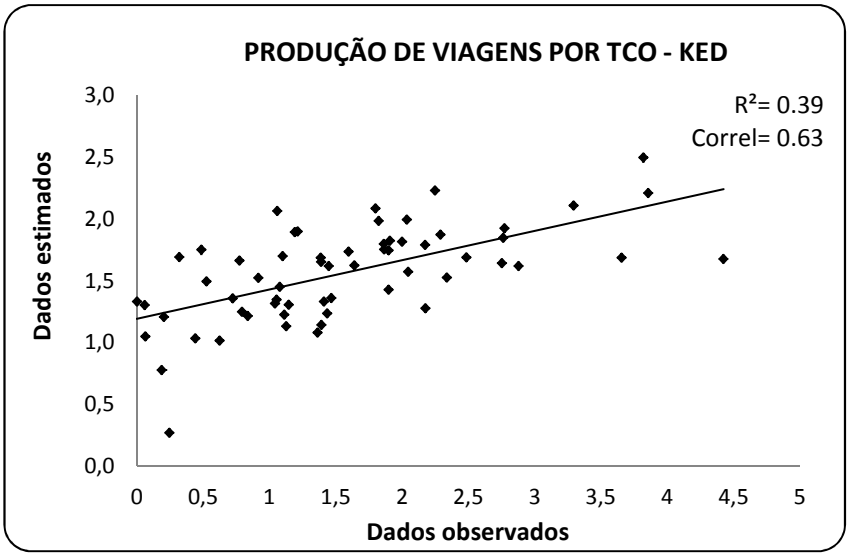




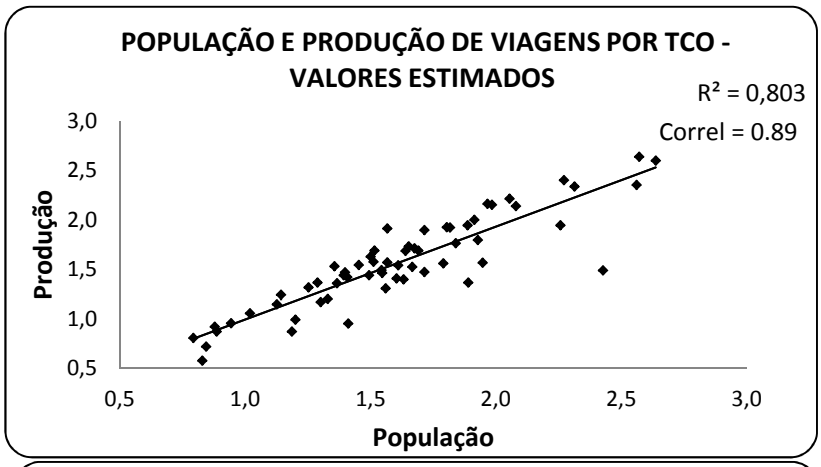
Apêndice B – Diagramas de dispersão população e produção de viagens - valores estimados x valores observados



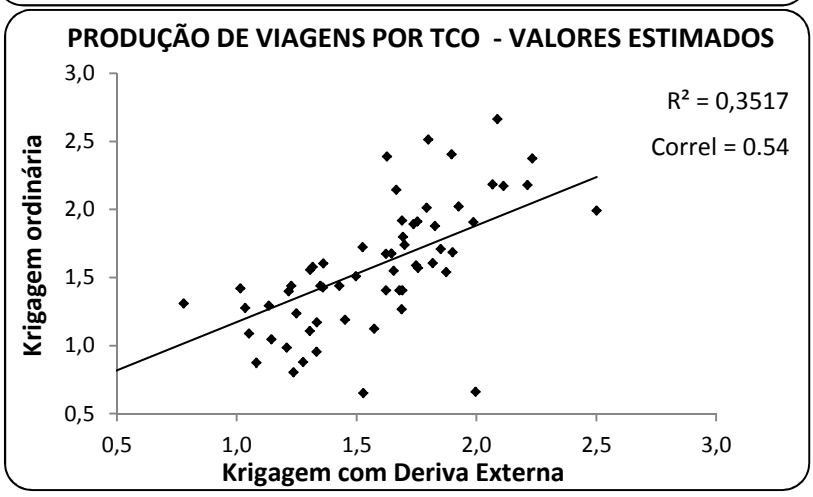




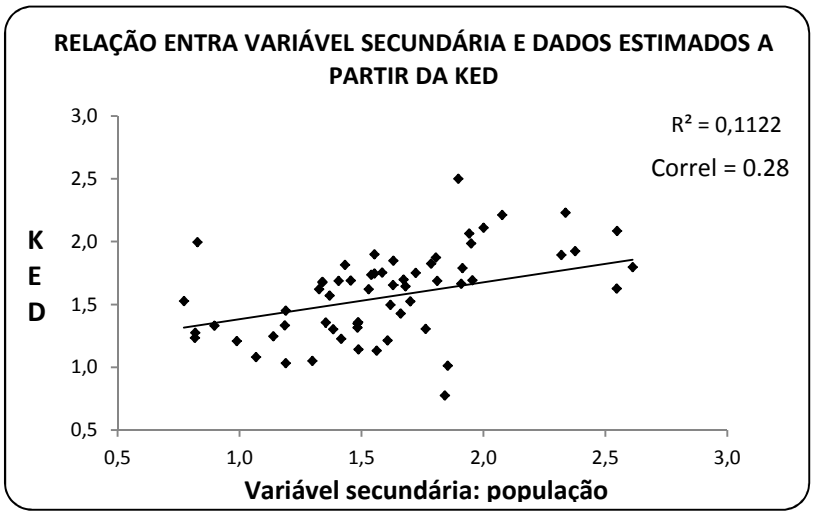
Apêndice C – Diagramas de dispersão valores estimados - (a) População x Produção de viagens, (b) Produção de viagens (KO) x Produção de viagens (KDE) e (c) Produção de viagens (KDE) x Variável Secundária.



(a)

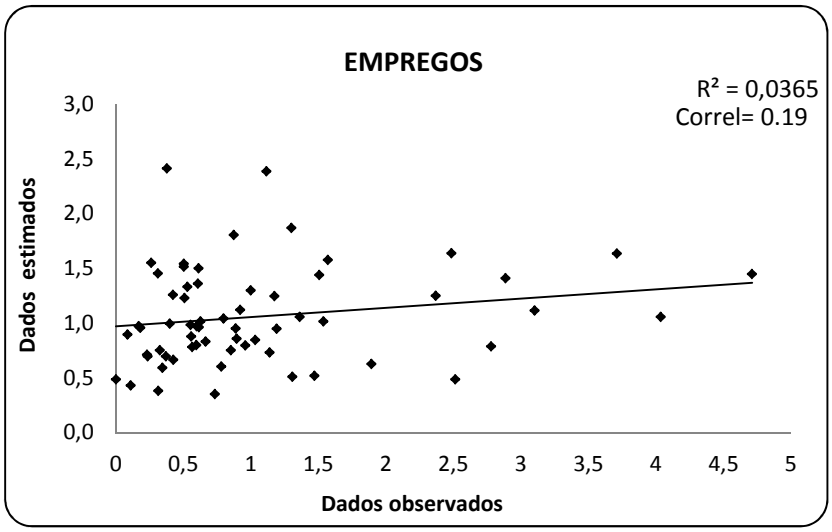


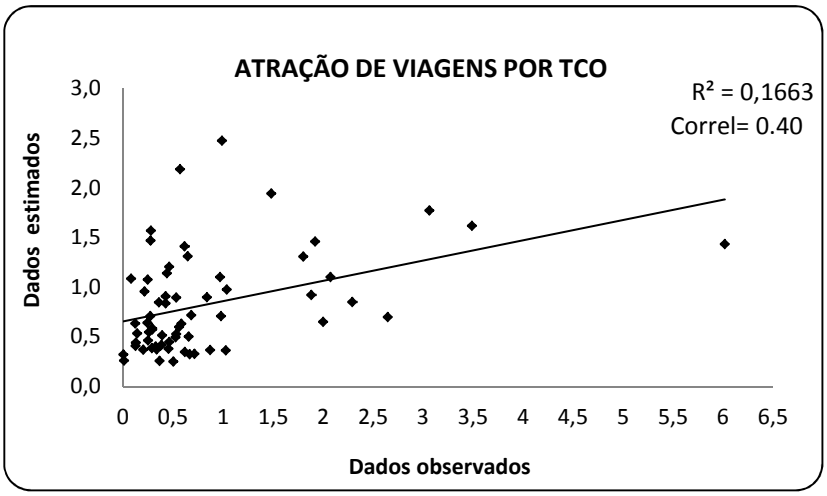
(b)

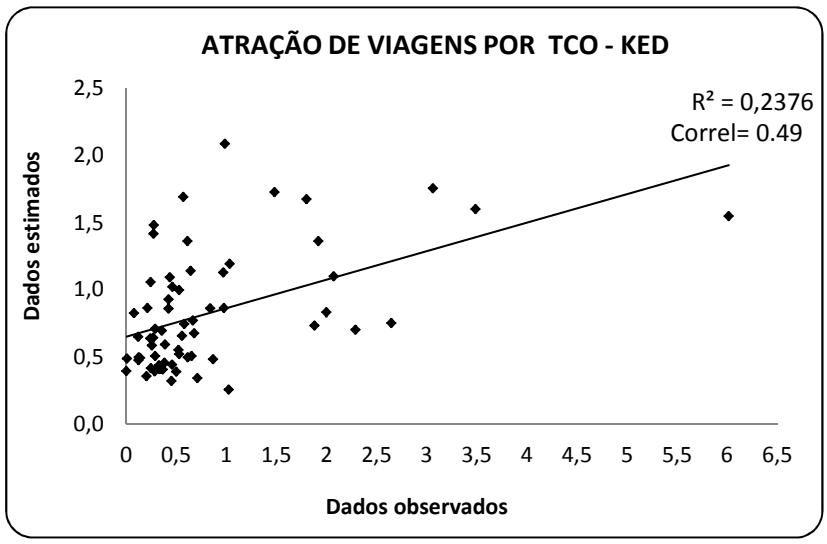


(c)

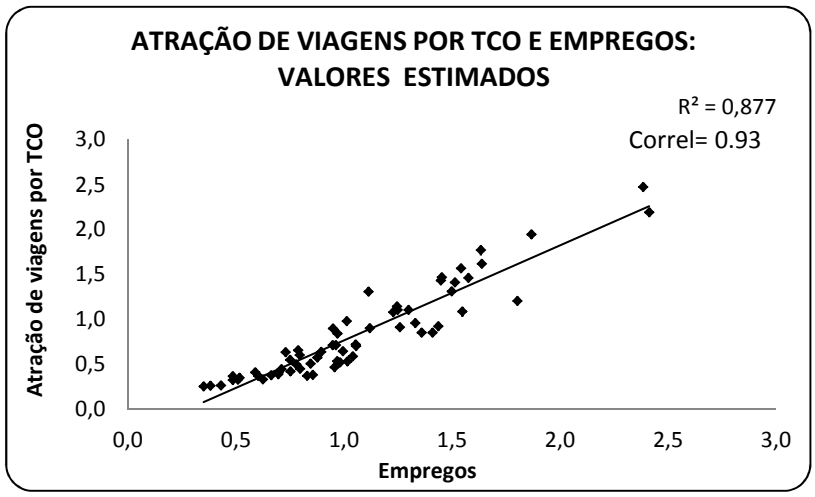
Apêndice D – Diagramas de dispersão - valores estimados x valores observados (Empregos e Atração de viagens)



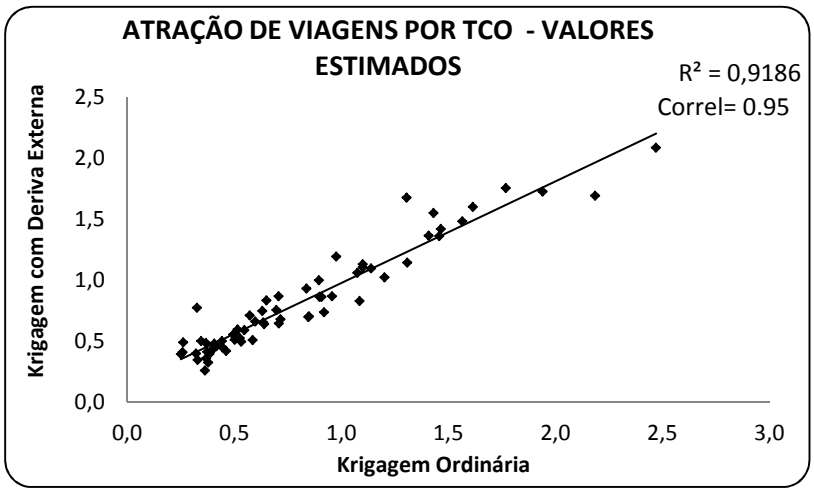




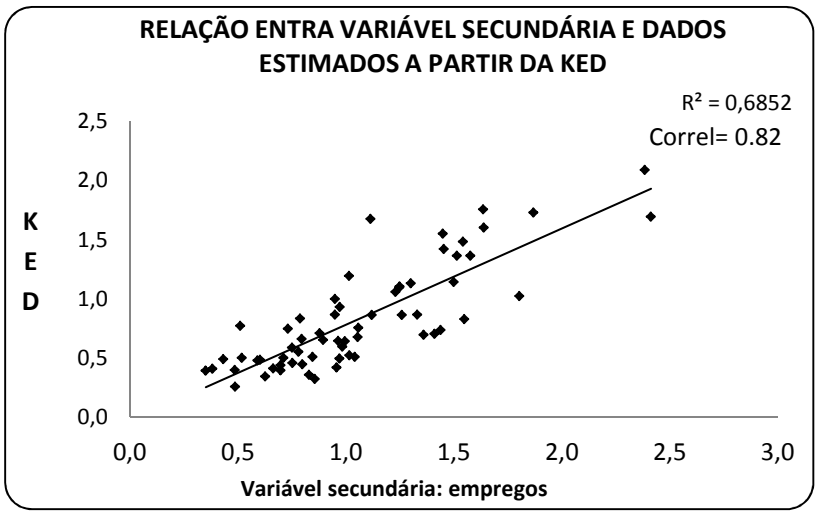
Apêndice E – Diagramas de dispersão valores estimados - (a) Empregos x Atração de viagens, (b) Atração de viagens (KO) x Atração de viagens (KDE) e (c) Atração de viagens (KDE) x Variável Secundária.



(a)

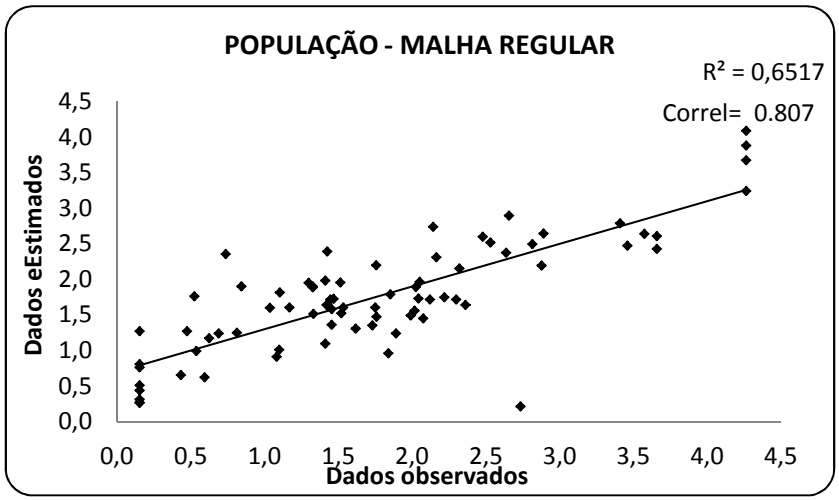


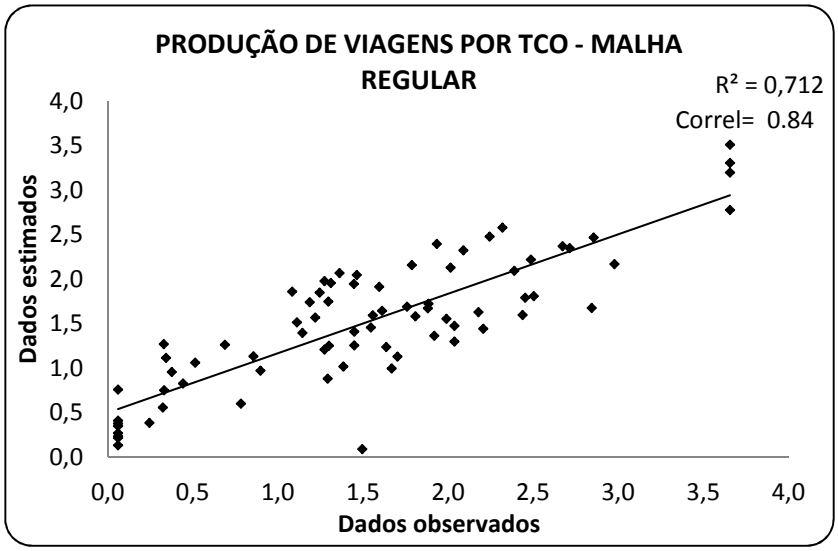
(b)

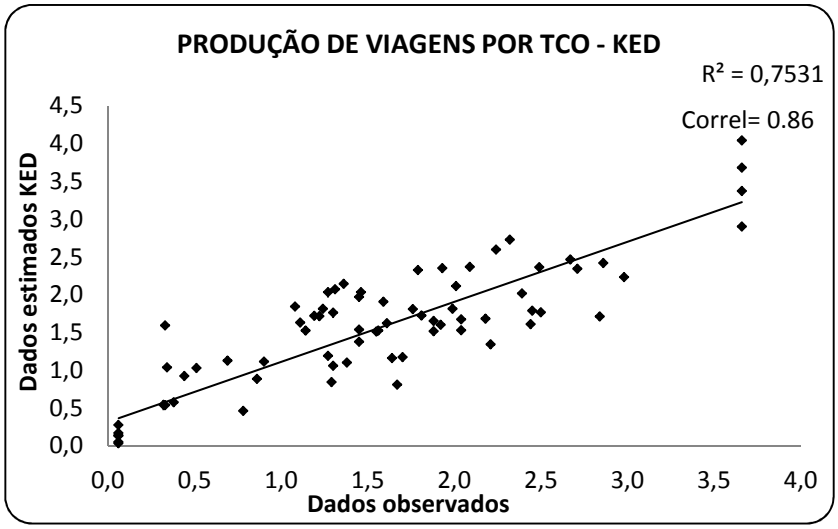


(c)

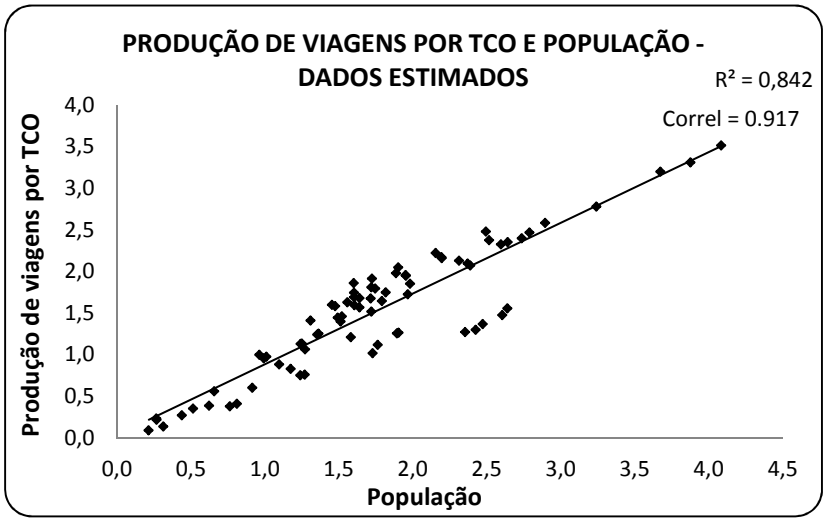
Apêndice F – Diagramas de dispersão produção de viagens por TCO e população (malha regular) 1995: dados observados x dados estimados



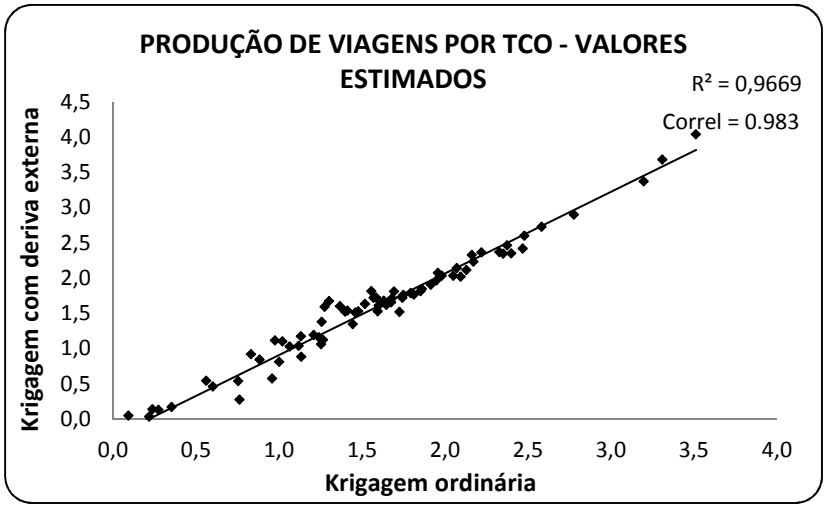




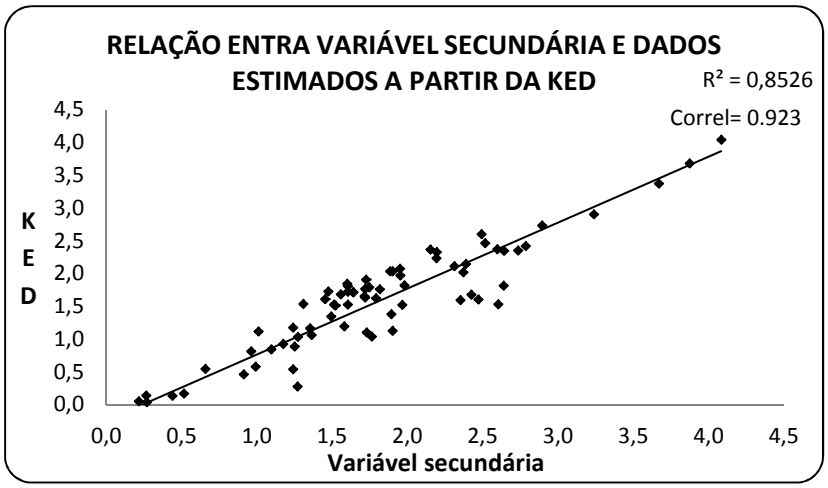
Apêndice G – Diagramas de dispersão valores estimados malha regular 1995 - (a) População x Produção de viagens, (b) Produção de viagens (KO) x Produção de viagens (KDE) e (c) Produção de viagens (KDE) x Variável Secundária.



(a)

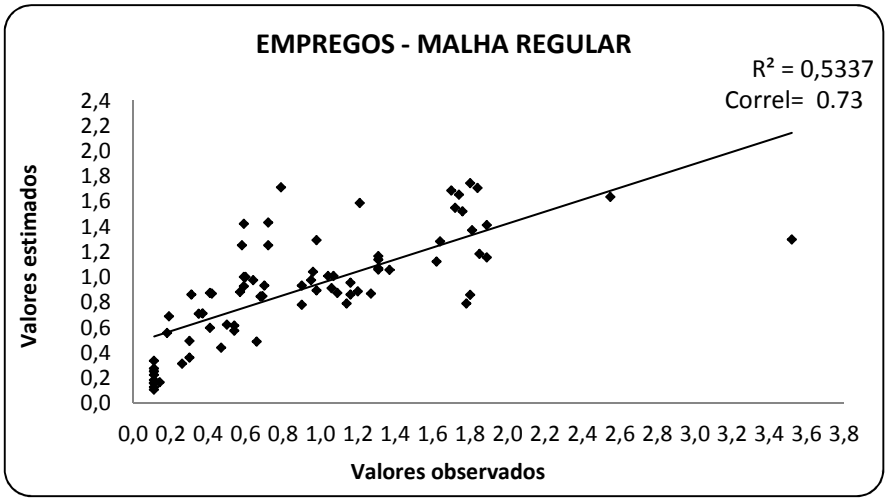


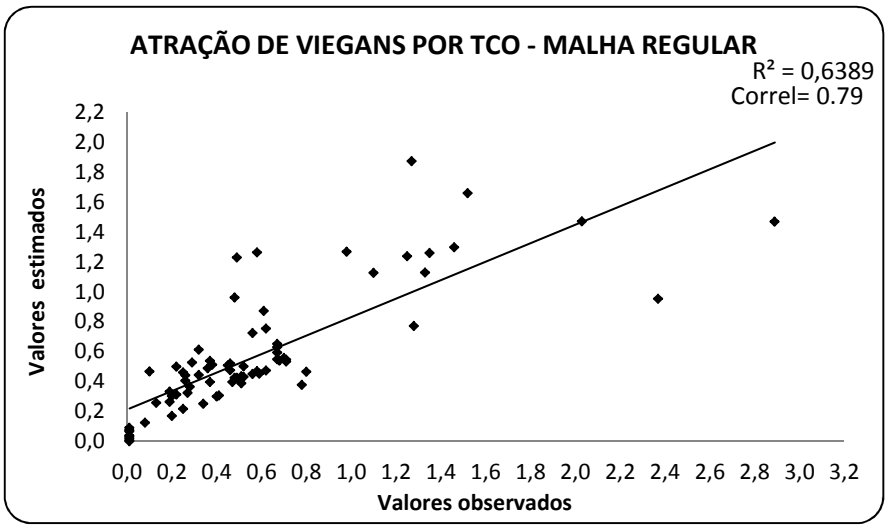
(b)

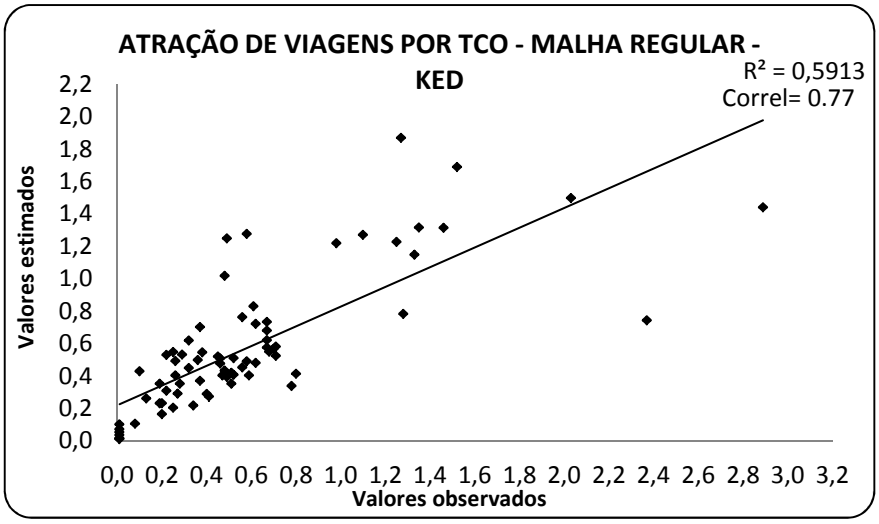


(c)

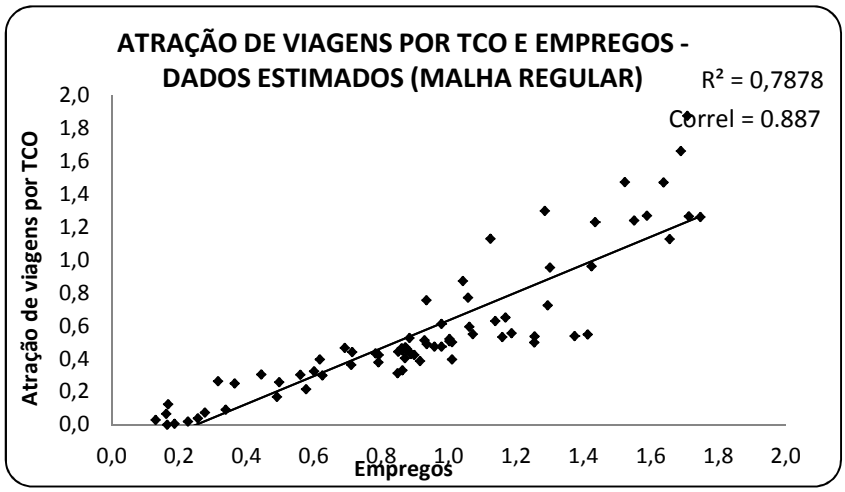
Apêndice H – Diagramas de dispersão empregos e atração de viagens por TCO (malha regular): dados estimados x observados



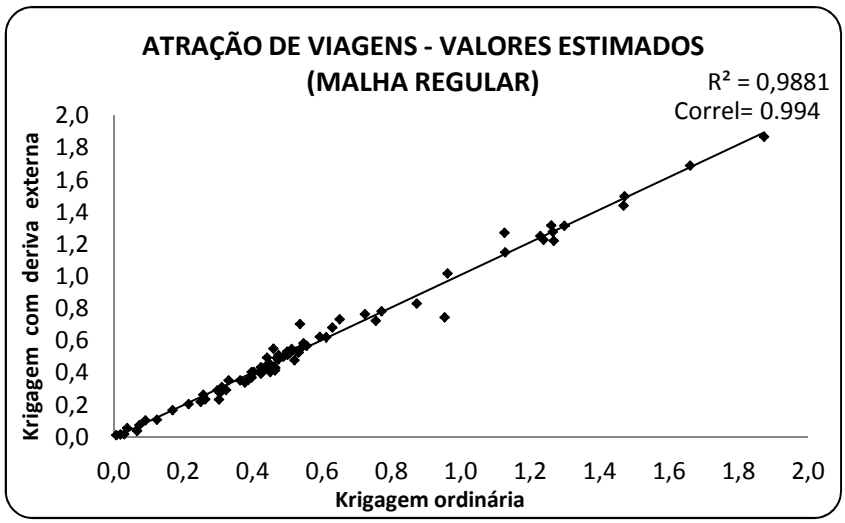




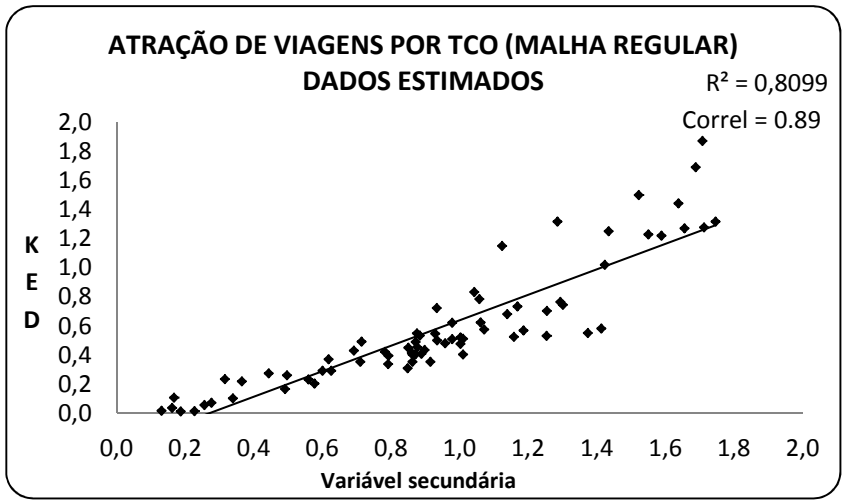
Apêndice I – Diagramas de dispersão valores estimados (malha regular) - (a) empregos x atração de viagens, (b) atração de viagens (KO) x atração de viagens (KDE) e (c) atração de viagens (KDE) x Variável Secundária.



(a)

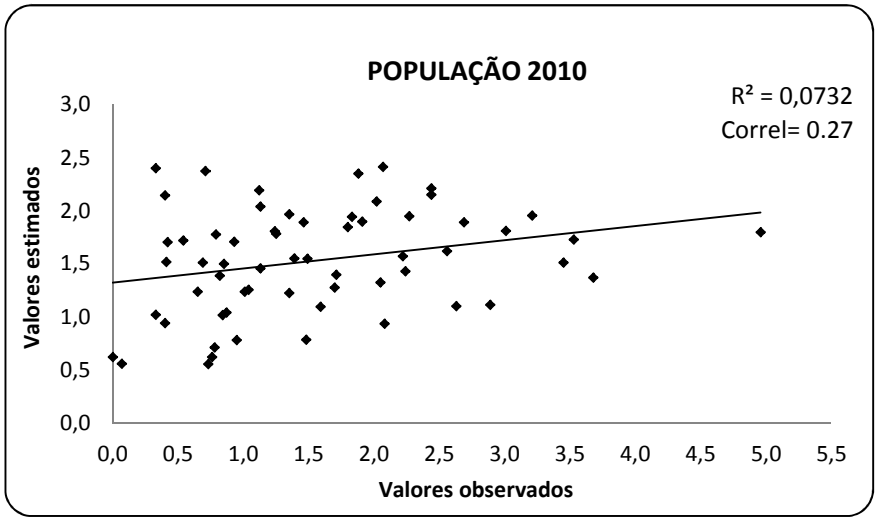


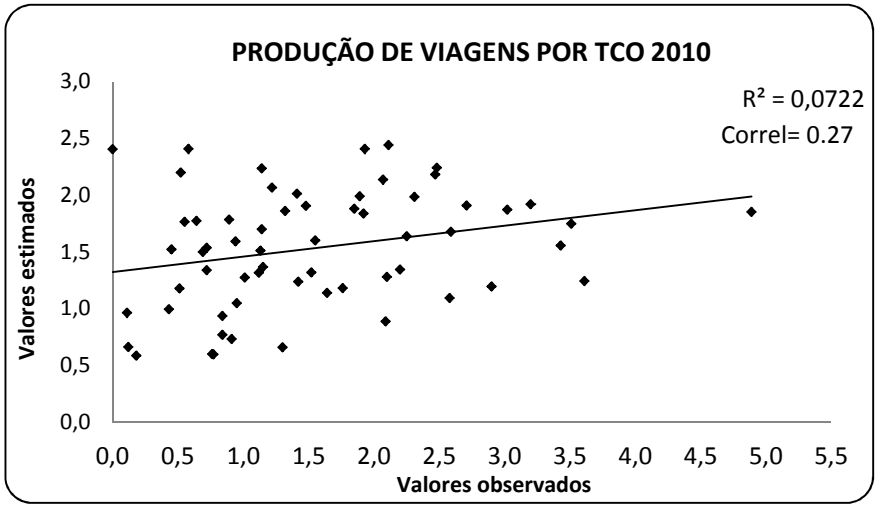
(b)

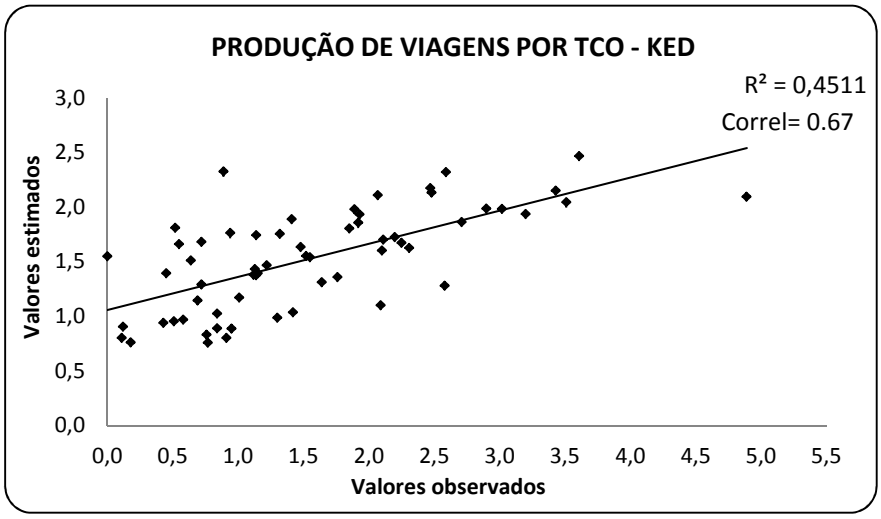


(c)

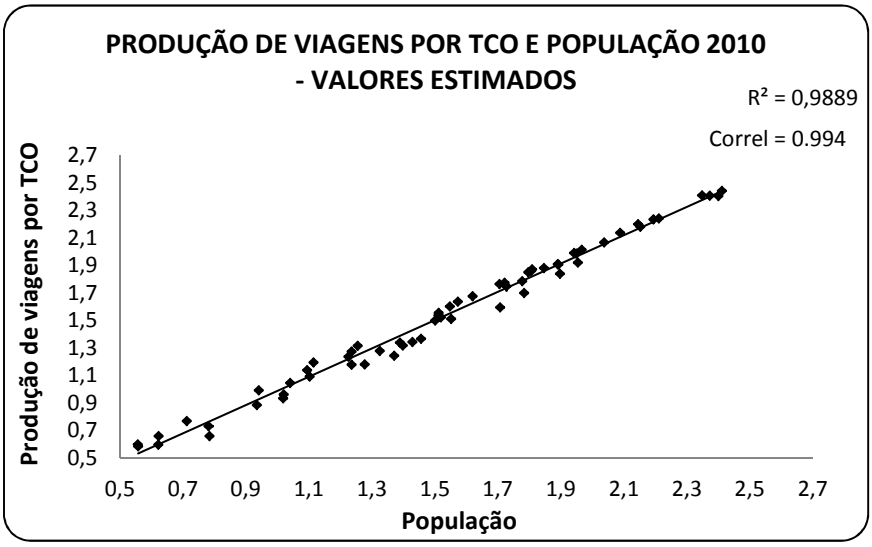
Apêndice J – Diagramas de dispersão população e produção de viagens por TCO 2010 - valores estimados x valores observados.



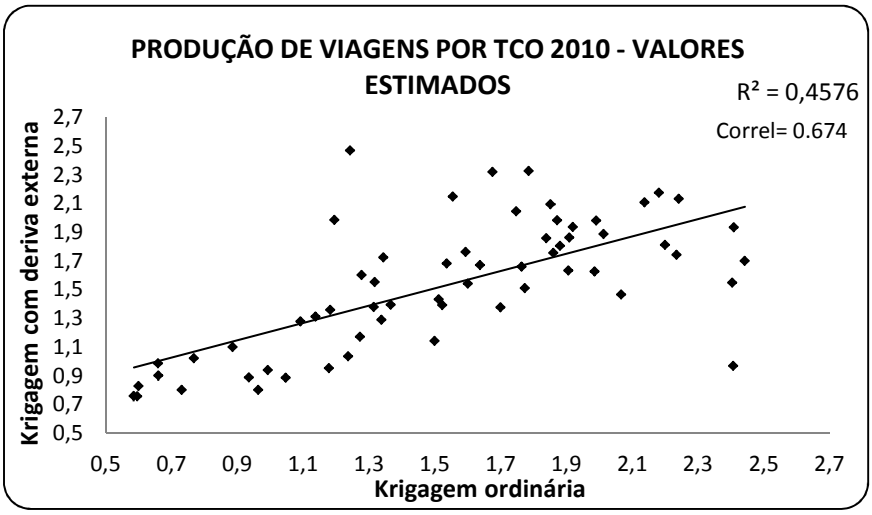




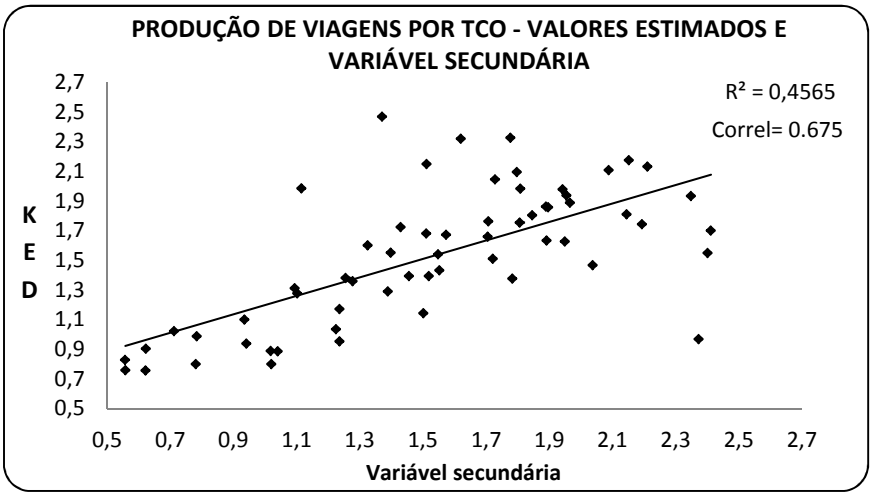
Apêndice K – Diagramas de dispersão entre valores estimados (2010) - (a) População x Produção de viagens, (b) Produção de viagens (KO) x Produção de viagens (KDE) e (c) Produção de viagens (KDE) x Variável Secundária.



(a)

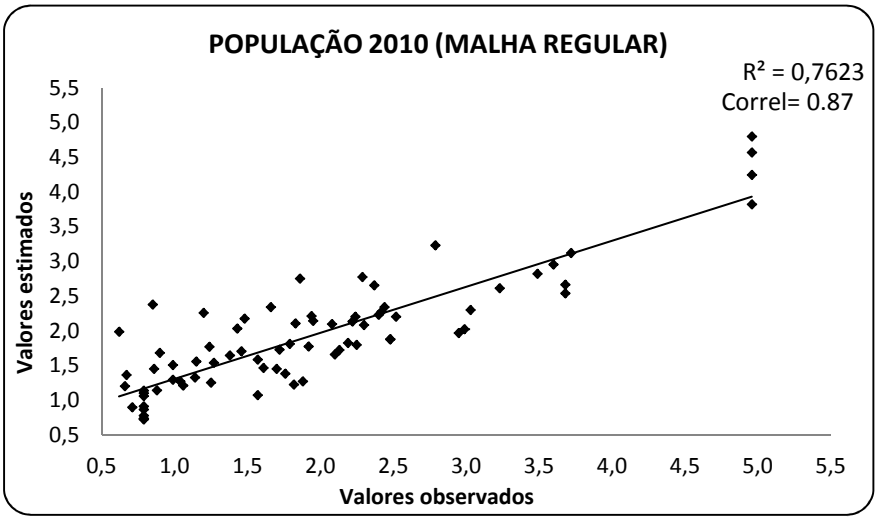


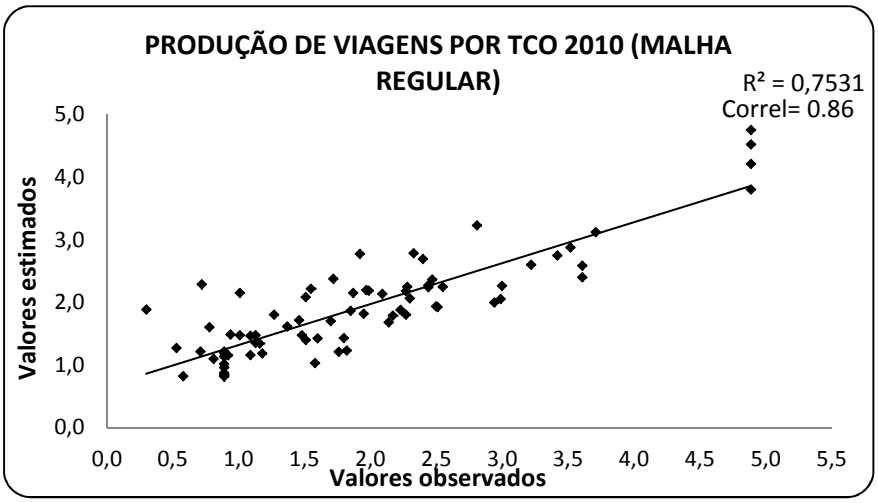
(b)

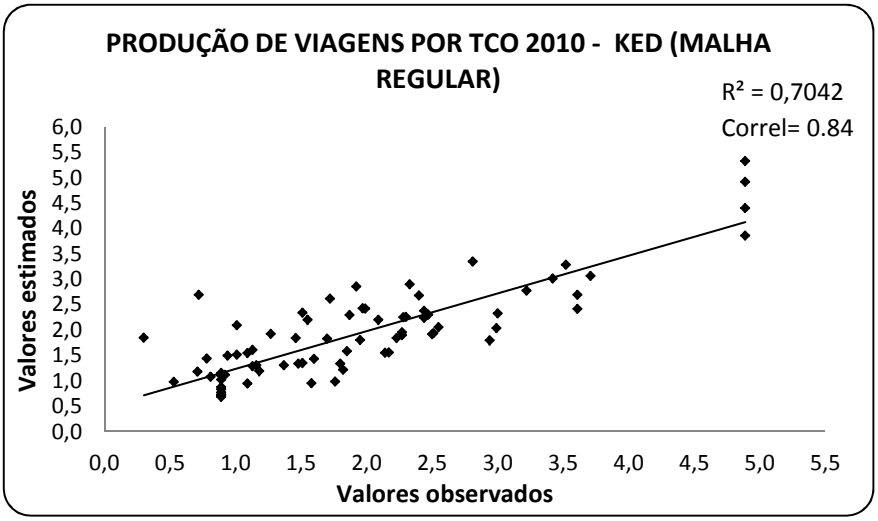


(c)

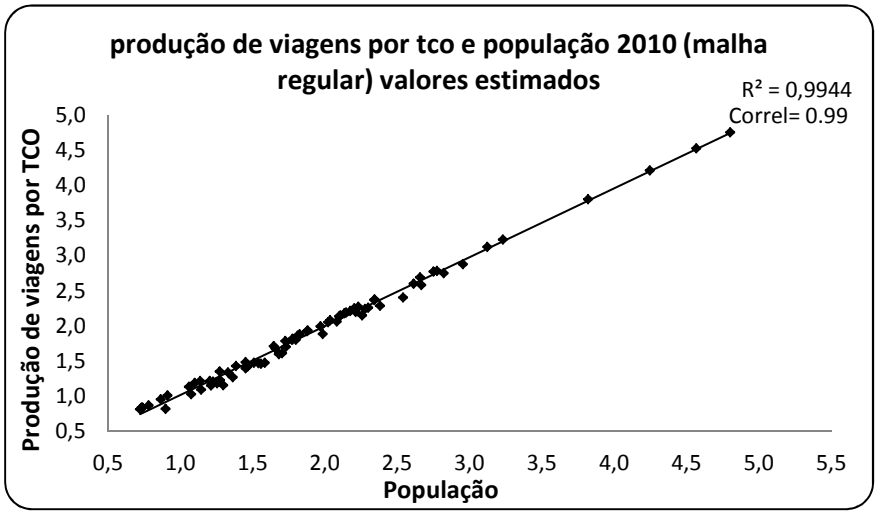
Apêndice L – Diagramas de dispersão população e produção de viagens por TCO 2010 (amostragem sistemática artificial) - valores estimados x valores observados.



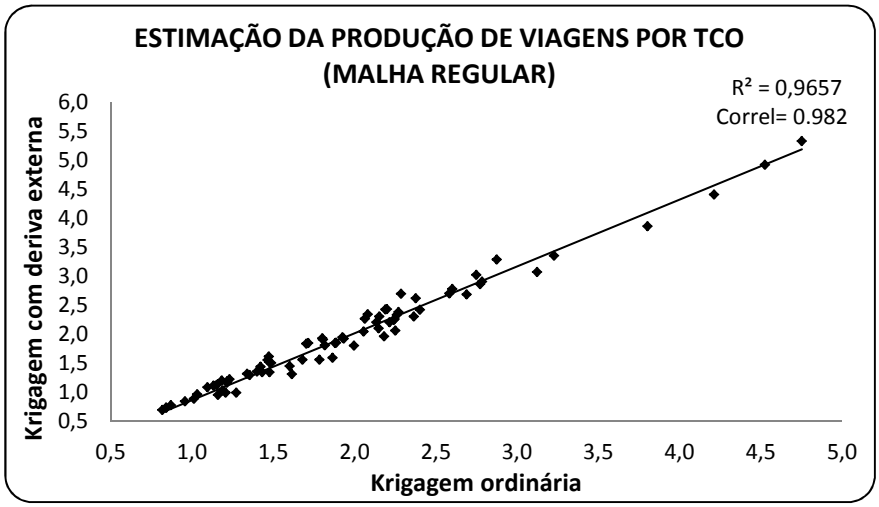




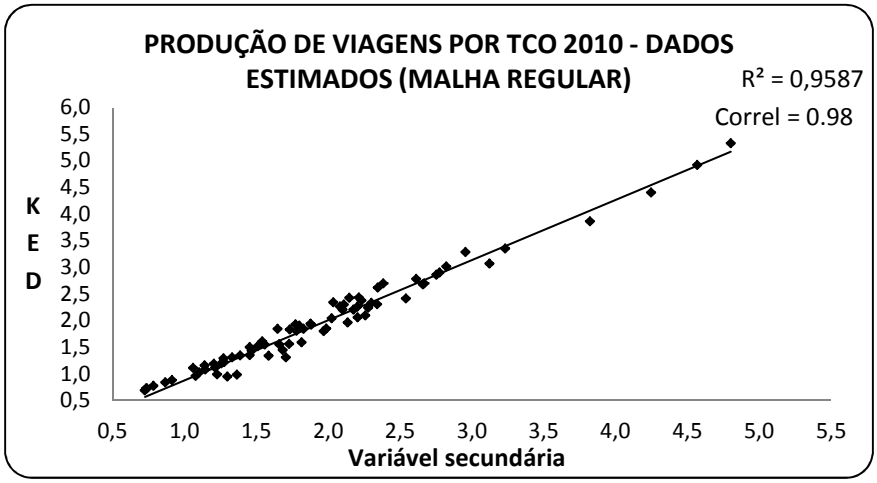
Apêndice M – Diagramas de dispersão entre valores estimados 2010 (amostragem sistemática artificial) - (a) População x Produção de viagens, (b) Produção de viagens (KO) x Produção de viagens (KDE) e (c) Produção de viagens (KDE) x Variável Secundária



(a)



(b)



(c)