



Ministério da
Ciência e Tecnologia



sid.inpe.br/mtc-m19/2011/02.04.16.51-TDI

**ESPAÇOS TOBLERIANOS COMPLETOS:
RECUPERANDO AS VISÕES DE TOBLER PARA A
REPRESENTAÇÃO DOS ESPAÇOS DE FLUXOS NA
ANÁLISE GEOGRÁFICA**

Marcio Azeredo

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelo Dr. Antônio Miguel Vieira Monteiro, aprovada em 04 de fevereiro de
2011

URL do documento original:

<<http://urlib.net/8JMKD3MGP7W/3959D5L>>

INPE
São José dos Campos
2011

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE (RE/DIR-204):**Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Membros:

Dr^a Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr^a Regina Célia dos Santos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

Dr. Horácio Hideki Yanasse - Centro de Tecnologias Especiais (CTE)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Vivéca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da
Ciência e Tecnologia



sid.inpe.br/mtc-m19/2011/02.04.16.51-TDI

**ESPAÇOS TOBLERIANOS COMPLETOS:
RECUPERANDO AS VISÕES DE TOBLER PARA A
REPRESENTAÇÃO DOS ESPAÇOS DE FLUXOS NA
ANÁLISE GEOGRÁFICA**

Marcio Azeredo

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelo Dr. Antônio Miguel Vieira Monteiro, aprovada em 04 de fevereiro de
2011

URL do documento original:

<<http://urlib.net/8JMKD3MGP7W/3959D5L>>

INPE
São José dos Campos
2011

Dados Internacionais de Catalogação na Publicação (CIP)

Az24e Azeredo, Marcio.
Espaços Toblerianos completos: recuperando as visões de
Tobler para a representação dos espaços de fluxos na análise geo-
gráfica / Marcio Azeredo. – São José dos Campos : INPE, 2011.
xxii+76 p. ; (sid.inpe.br/mtc-m19/2011/02.04.16.51-TDI)

Dissertação (Mestrado em Computação Aplicada) – Insti-
tuto Nacional de Pesquisas Espaciais, São José dos Campos, 2011.
Orientador : Dr. Antônio Miguel Vieira Monteiro.

1. Análise exploratória. 2. Espaço de fluxos. 3. Regressão
bidimensional. 4. Análise multidimensional . I.Título.

CDU 528.711.7:004

Copyright © 2011 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, reprográfico, de microfilmagem ou outros, sem a permissão escrita do INPE, com exceção de qualquer material fornecido especificamente com o propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2011 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, microfilming, or otherwise, without written permission from INPE, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

**Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de Mestre em
Computação Aplicada**

Dra. Lúbia Vinhas



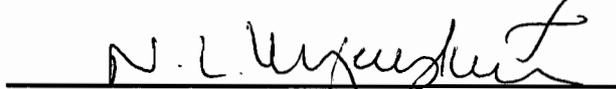
Presidente / INPE / São José dos Campos - SP

Dr. Antonio Miguel Vieira Monteiro



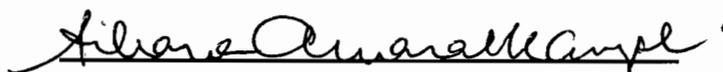
Orientador(a) / INPE / SJC Campos - SP

Dr. Nandamudi Lankalapalli Vijaykumar



Membro da Banca / INPE / SJC Campos - SP

Dra. Silvana Amaral Kampel



Membro da Banca / INPE / SJC Campos - SP

Dr. Julio D'alge



Membro da Banca / INPE / São José dos Campos - SP

Dr. Marcelo Tílio Monteiro de Carvalho



Convidado(a) / PUC/RJ / Rio de Janeiro - RJ

Aluno (a): Marcio Azeredo

São José dos Campos, 04 de fevereiro de 2011

“O grande desafio é redirecionar a atenção dos mapas que apresentam respostas, para os mapas que fomentam a busca pelas mesmas.”.

MACEachREN E MONMONIER
em *“Introduction”*, 1992

*A minha querida esposa Luciana, e adoráveis filhas Júlia e
Amanda.*

AGRADECIMENTOS

Ao Dr. Antônio Miguel Vieira Monteiro pelo conhecimento transmitido, assim como pelo apoio incondicional prestado durante todo o transcorrer do curso.

Aos amigos Emerson Xavier, Gilberto Ribeiro e Wladimir Meyer, pela cooperação diante das dificuldades apresentadas em diferentes momentos deste trabalho.

À Companhia do metropolitano de São Paulo (Metrô) por disponibilizar os dados referentes à Pesquisa O/D 2007. É um grande estímulo poder colaborar com a parceria entre o Metrô e o INPE (Programa Espaço e Sociedade), na busca de novas possibilidades analíticas de uso dos dados da pesquisa OD do Metrô.

À Diretoria de Serviço Geográfico (DSG) pela grande oportunidade que me foi dada, ao permitir a realização do mestrado em um Instituto de tamanha importância no cenário científico nacional.

RESUMO

Relacionamentos espaciais são dinâmicos, fortemente anisotrópicos, e passíveis de conexão por intermédio de diferentes tipos de redes (ex. redes sociais, de comunicações e de transportes). Essas redes, por sua vez, influenciam o surgimento de direções preferenciais de relacionamentos, que reorganizam os elementos nesse espaço, possibilitando novas formas de representação desses territórios em meio digital. Como consequência, requerem novos métodos exploratórios de análise, na tentativa de capturar essa dinâmica. Este trabalho procura resgatar a técnica da Regressão Bidimensional proposta por Tobler na década de 70, para juntamente com outras técnicas de representação oferecer uma nova visão do espaço em estudo, redesenhando o território em função dos fluxos em suas redes. Essa abordagem é parte do que chamamos de Espaços Toblerianos Completos (ETC), uma proposta que busca construir novas cartografias que possam ampliar as possibilidades de leitura para a geografia de fenômenos socioespaciais complexos.

FULL TOBLERIAN SPACES: RECOVERING TOBLER'S VIEWS ON THE REPRESENTATIONS NEEDED FOR ANALYSING A GEOGRAPHY OF THE FLUXES IN GISCIENCE

ABSTRACT

Spatial relationships are dynamic, highly anisotropic, and capable of connecting through different types of networks (ex. social, communication and transportation networks). These networks influence the appearance of preferential directions of relationships, providing different representations of space. As a result, requires new methods of exploratory analysis in an attempt to capture this dynamic. This work tries to recover the Bidimensional Regression technique proposed by Tobler in the 70's, together with other techniques of representation, offer a new vision of space study, representing the territory according to network flows. This approach is part of what we call Full Toblerian Spaces (FTS), our proposal for a new mapping in order to provide new opportunities for explanations of the geography of complex socio-spatial phenomena.

LISTA DE FIGURAS

| | <u>Pág.</u> |
|------------------------------------------------------------------------------------------|-------------|
| 2.1 Casos registrados de propagação do vírus Influenza A (H1N1) - Jun 2009. | 9 |
| 2.2 O conceito de proximidade em função do tempo. | 11 |
| 2.3 Diferentes possibilidades de representação do espaço. | 12 |
| 2.4 Os papéis dos fixos nos dados de origem e destino. | 13 |
| 2.5 Exemplo de representação tabular de dados de fluxo. | 13 |
| 2.6 Representação das baixas do Exército de Napoleão na Campanha na Rússia. | 14 |
| 2.7 Espaço anisotrópico (Dallas - EUA). | 15 |
| 2.8 Espaço em função do tempo (Salt Lake City - EUA). | 16 |
| 2.9 Variáveis Z e W | 17 |
| 2.10 Mapeamento de $Z(x, y) \rightarrow W(u, v)$ | 17 |
| 2.11 Vetores principais de distorção. | 18 |
| 2.12 Modelos de transformação. | 20 |
| 2.13 Comparação entre mapas da Península de Setúbal (1757). | 21 |
| 2.14 Estudo da hierarquia de lugares em Luxemburgo. | 22 |
| 2.15 Matriz de dissimilaridades com as distâncias aéreas. | 24 |
| 2.16 Configuração espacial obtida pelo CMDS. | 24 |
| 3.1 Arquitetura de funcionamento com a TerraLib. | 28 |
| 3.2 Metodologia da representação de fluxos utilizando a Regressão Bidimensional. | 29 |
| 3.3 Relacionamentos de similaridade. | 30 |
| 3.4 Relacionamentos de dissimilaridade. | 31 |
| 3.5 Exemplo de uma configuração espacial de fluxos. | 32 |
| 3.6 Matriz de dados de fluxos. | 33 |
| 3.7 Montagem da matriz simétrica com base na escolha dos relacionamentos. | 33 |
| 3.8 Montagem da matriz simétrica com base na diagonal superior. | 34 |
| 3.9 Base espacial composta por pontos. | 36 |
| 3.10 Base espacial composta por linhas. | 36 |
| 3.11 Base espacial composta por polígonos. | 37 |
| 4.1 Estrutura dos arquivos de dados de OD. | 41 |

| | | |
|------|------------------------------------------------------------------------|----|
| 4.2 | Rede total de relacionamentos presentes no conjunto de dados. . . . | 41 |
| 4.3 | Amostra selecionada dentre as 460 zonas de pesquisa da RMSP. . . | 42 |
| 4.4 | Amostra de estudo do experimento 1. | 43 |
| 4.5 | Rede de relacionamentos do experimento 1. | 44 |
| 4.6 | Vetores principais de distorção do experimento 1. | 48 |
| 4.7 | Interpolação do espaço do experimento 1. | 49 |
| 4.8 | Grade regular interpolada e vetores principais do experimento 1. . . . | 50 |
| 4.9 | Projeção das distorções sobre o espaço original do experimento 1. . . | 50 |
| 4.10 | Análise dos resultados do experimento 1. | 51 |
| 4.11 | Amostra de estudo do experimento 2. | 53 |
| 4.12 | Rede de relacionamentos do experimento 2. | 54 |
| 4.13 | Vetores principais de distorção do experimento 2. | 58 |
| 4.14 | Interpolação do espaço do experimento 2. | 59 |
| 4.15 | Grade regular interpolada e vetores principais do experimento 2. . . . | 60 |
| 4.16 | Projeção das distorções sobre o espaço original do experimento 2. . . | 60 |
| 4.17 | Análise dos resultados do experimento 2. | 61 |
| 4.18 | Sobreposição dos resultados dos experimentos 1 e 2. | 62 |

LISTA DE TABELAS

| | <u>Pág.</u> |
|----------------------------------------------------------------------|-------------|
| 4.1 Coordenadas dos centroides das zonas do experimento 1. | 44 |
| 4.2 Matriz de custo simétrica do experimento 1. | 46 |
| 4.3 Coordenadas do Sistema Imagem do experimento 1. | 47 |
| 4.4 Coordenadas do Sistema Imagem Ajustado do experimento 1. | 48 |
| 4.5 Coordenadas dos centroides das zonas do experimento 2. | 54 |
| 4.6 Matriz de quantidade de viagens do experimento 2. | 56 |
| 4.7 Coordenadas do Sistema Imagem do experimento 2. | 57 |
| 4.8 Coordenadas do Sistema Imagem Ajustado do experimento 2. | 58 |

LISTA DE ABREVIATURAS E SIGLAS

| | | |
|-------|---|---------------------------------------------|
| CMDs | – | Escalonamento Multidimensional Clássico |
| DPI | – | Divisão de Processamento de Imagens |
| EMQ | – | Erro Médio Quadrático |
| ETC | – | Espaços Toblerianos Completos |
| EUA | – | Estados Unidos da América |
| GPS | – | Sistema de Posicionamento Global |
| INPE | – | Instituto Nacional de Pesquisas Espaciais |
| MDS | – | Escalonamento Multidimensional |
| Metrô | – | Companhia do Metropolitano de São Paulo |
| MMQ | – | Método dos Mínimos Quadrados |
| OD | – | Origem e Destino |
| PDA | – | Assistentes Pessoais de Dados |
| RMSP | – | Região Metropolitana de São Paulo |
| SIGs | – | Sistemas de Informações Geográficas |
| SGBDs | – | Sistemas de Gerenciamento de Banco de Dados |
| SUS | – | Sistema Único de Saúde |
| WPS | – | Serviço de Processamento na Web |

SUMÁRIO

| | <u>Pág.</u> |
|---------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1 INTRODUÇÃO | 1 |
| 1.1 Motivação | 1 |
| 1.2 Objetivo | 3 |
| 1.3 Contribuição | 3 |
| 1.4 Prova de Conceito: novas cartografias para a Pesquisa Origem-Destino da Companhia do Metropolitano de São Paulo - Metrô | 4 |
| 1.5 Inserção na agenda científica | 5 |
| 1.6 Organização da dissertação | 5 |
| 2 FUNDAMENTAÇÃO TEÓRICA | 7 |
| 2.1 Os SIGs atuais e o conceito de proximidade | 7 |
| 2.2 A globalização dos relacionamentos espaciais | 9 |
| 2.3 As redes e os fluxos - uma alternativa | 10 |
| 2.4 A representação dos fluxos | 12 |
| 2.5 A anisotropia das relações espaciais | 14 |
| 2.6 A Regressão Bidimensional de Tobler | 16 |
| 2.6.1 Mapeando as transformações geométricas | 17 |
| 2.6.2 A Fase de Ajustamento | 19 |
| 2.6.3 A Fase de Interpolação | 20 |
| 2.6.4 Aplicações da Regressão Bidimensional | 21 |
| 2.7 O Escalonamento Multidimensional | 23 |
| 3 O FRAMEWORK ETC: UMA CONTRIBUIÇÃO PARA O DOMÍNIO DOS FLUXOS | 27 |
| 3.1 A biblioteca TerraLib | 27 |
| 3.2 A arquitetura do <i>framework</i> | 27 |
| 3.3 Metodologia de representação de dados de fluxo | 29 |
| 3.3.1 Identificando as relações de proximidade | 29 |
| 3.3.1.1 Dados de similaridade | 30 |
| 3.3.1.2 Dados de dissimilaridade | 30 |
| 3.3.2 Construindo a matriz de fluxos | 31 |

| | | |
|----------|-------------------------------------------------------------------------------------------------------------|-----------|
| 3.3.2.1 | O problema da matriz incompleta | 34 |
| 3.3.3 | Mapeando as redes de fluxos | 35 |
| 3.3.3.1 | Base de dados de pontos | 36 |
| 3.3.3.2 | Base de dados de linhas | 36 |
| 3.3.3.3 | Base de dados de polígonos | 37 |
| 3.3.4 | A análise multidimensional dos dados de origem e destino | 38 |
| 3.3.5 | A Regressão Bidimensional de Tobler | 38 |
| 4 | PROVA DE CONCEITO: O CASO DA PESQUISA ORIGEM-DESTINO PARA A RMSP DO METRÔ DE SÃO PAULO | 39 |
| 4.1 | A Pesquisa OD/2007 da Companhia do Metropolitano de São Paulo | 39 |
| 4.2 | Experimentos | 42 |
| 4.2.1 | Experimento 1: Tempo médio de duração das viagens | 43 |
| 4.2.1.1 | Base de dados | 43 |
| 4.2.1.2 | Construindo a matriz de fluxos | 44 |
| 4.2.1.3 | A análise multidimensional dos dados de origem e destino | 47 |
| 4.2.1.4 | A Regressão Bidimensional de Tobler | 47 |
| 4.2.1.5 | Reprojetando o espaço em função dos fluxos | 49 |
| 4.2.1.6 | Discussão dos resultados | 51 |
| 4.2.2 | Experimento 2: Número total de viagens | 53 |
| 4.2.2.1 | Base de dados | 53 |
| 4.2.2.2 | Construindo a matriz de fluxos | 54 |
| 4.2.2.3 | A análise multidimensional dos dados de origem e destino | 57 |
| 4.2.2.4 | A Regressão Bidimensional de Tobler | 57 |
| 4.2.2.5 | Reprojetando o espaço em função dos fluxos | 59 |
| 4.2.2.6 | Discussão dos resultados | 61 |
| 5 | CONCLUSÕES E PERSPECTIVAS | 65 |
| | REFERÊNCIAS BIBLIOGRÁFICAS | 67 |
| | APÊNDICE A - TRANSFORMAÇÕES GEOMÉTRICAS. | 73 |
| A.1 | A Transformação de Similaridade | 73 |
| A.2 | A Transformação Afim | 74 |

1 INTRODUÇÃO

1.1 Motivação

Quando os Sistemas de Informações Geográficas (SIGs) surgiram na década de 60, os dados eram escassos e o custo de processamento elevado (MILLER, 2000). Frente às limitações tecnológicas existentes, simplificar a representação dos relacionamentos espaciais, e a construção geográfica do espaço, mostrou-se uma solução necessária, conveniente e satisfatória.

Nesse contexto, muitos dos conceitos e métodos empregados na análise espacial foram impulsionados utilizando uma noção de proximidade baseada em uma métrica de distância euclidiana. Desse modo, partem do princípio de que os fenômenos espaciais apresentam um comportamento isotrópico, ou seja, que mantêm suas propriedades constantes em qualquer direção dentro do espaço de representação. No entanto, o mundo não é sempre regular e previsível (TOBLER, 2004), e o conceito de proximidade pode ter diferentes interpretações.

Sendo assim, os lugares não devem ser considerados simplesmente como um conjunto de coordenadas sem uma semântica associada. Em sua grande parte, relacionamentos espaciais são construídos baseados em fatores cognitivos, e não somente geométricos.

A escolha de uma agência bancária para a abertura de uma conta, por exemplo, não necessariamente implica na escolha da agência mais próxima do ponto de vista euclidiano. Fatores como a credibilidade do banco, infraestrutura disponível para o cliente, inovações tecnológicas, segurança, certamente serão considerados na decisão final. E para cada unidade de análise (ex. indivíduo, cidade ou país), existe um conjunto peculiar de características, que depois de analisadas, irão balizar o julgamento.

Desse modo, o conceito de proximidade euclidiano pode não ser suficiente para uma análise geográfica mais abrangente. Em muitos casos, os relacionamentos surgem com o objetivo de atender a necessidades específicas, ou a anseios pessoais, indicando que a influência de suas propriedades não deve ser desconsiderada. Em alguns casos, isso poderia restringir a capacidade de representação e análise de como ocorre a dinâmica de surgimento dos relacionamentos, e qual a sua influência na configuração do espaço.

O estudo dos atributos dos lugares passa a ser de grande importância na análise das relações espaciais. São essas diferenças de características existentes entre os lugares as responsáveis pela dinâmica de busca, ou transferência de bens, serviços, pessoas, tecnologia, entre outros (GOULD; WHITE, 1986), e que motivam o surgimento das relações.

Segundo Santos (2006), a construção geográfica do espaço pode ser compreendida como um conjunto de fixos, e fluxos. Os fixos representam os lugares, que condicionam e são condicionados em função das propriedades dos fluxos em seus relacionamentos. Dessa forma, fixos exercem uma função no espaço ao qual pertencem, e que pode ser analisada com base no custo associado ao estabelecimento das relações entre um lugar, e todos os outros lugares. Já os fluxos podem atravessar ou se instalarem nos fixos, sendo capazes de redefini-los em função de suas propriedades, ou de se realizarem sobre eles.

No estudo dos fluxos, um importante fator a ser considerado é a possibilidade de sua estruturação na forma de diferentes tipos de redes (ex. redes sociais, de comunicações ou de transportes). Essas redes, por sua vez, podem apresentar diferentes configurações em função dos atributos das relações espaciais considerados para sua construção.

Segundo Prager (2008), as redes construídas a partir das relações servem como dutos na condução dos fluxos. Mais do que isso, têm influência no surgimento de direções preferenciais de relacionamentos em função dos custos atribuídos às suas arestas. Isso ocorre, porque a frequência com que as interações surgem, tende a ser menor onde os custos associados sejam maiores.

No planejamento de uma viagem, por exemplo, os custos com pedágios e tarifas são normalmente considerados para efeito de escolha das rotas que serão percorridas. Desse modo, as redes são caracterizadas em função dos custos monetários envolvidos no trajeto. No entanto, outros atributos também podem ser considerados, como o tempo de deslocamento, o meio de transporte, o relevo no percurso, ou o período do dia. O que acaba por aumentar as possibilidades de análise.

Cada atributo utilizado para representar o espaço, acaba por inserir distorções na representação desse. Distorções essas, que têm a capacidade de oferecer

valiosas informações acerca do papel que o lugar exerce no espaço regional ao qual pertence.

O presente trabalho procura resgatar a técnica da Regressão Bidimensional aplicada por [Tobler \(1994\)](#) em um domínio geográfico, para juntamente com outras técnicas de representação, capturar as transformações geométricas do espaço em função dos atributos de uma rede de relacionamentos. Desse modo, o espaço é repensado a partir de uma representação que ele tenha em função dos fluxos em suas redes, ampliando a capacidade de análise de processos geográficos complexos. Nesse caso, o conceito de proximidade não é mais estabelecido pelo senso comum, mas sim por atributos descritivos dos elementos espaciais.

1.2 Objetivo

O presente estudo parte da premissa de que as transformações geométricas resultantes da aplicação da Regressão Bidimensional apresentada por [Tobler \(1994\)](#), em conjunto com outras técnicas de representação, oferecem uma nova visão do espaço em estudo, redesenhando o território em função dos fluxos em suas redes.

Sendo assim, o objetivo deste trabalho é apresentar uma leitura dos Espaços de Fluxos na análise espacial, a partir de instrumentos de representação já existentes. Com isso, espera-se fornecer novas possibilidades de estudo sobre fenômenos que envolvam dados de fluxo, ou seja, dados de origem e destino.

1.3 Contribuição

A principal contribuição do presente trabalho é apresentar uma forma de estudar e representar a dinâmica dos relacionamentos espaciais, segundo uma abordagem diferente da trivial. Uma abordagem que considera o conceito de proximidade como função dos atributos dos objetos geográficos em estudo, ou das propriedades das relações entre os mesmos.

Desse modo, agrega-se valor à discussão sobre a análise das ações que influenciam a dinâmica de surgimento, e desaparecimento, dos relacionamentos espaciais entre diferentes localizações geográficas dentro de um espaço regional.

Ao fim do trabalho, verifica-se que a metodologia apresentada possibilita realizar projeções sobre o território, oferecendo subsídios para a análise das possíveis causas, e impactos, da reconfiguração do espaço geográfico em função dos fluxos em suas redes, explorando por exemplo, as possíveis razões pelas quais eles surgem, ou desaparecem.

A abordagem aqui apresentada é uma contribuição para o que convenciou-se chamar neste trabalho de Espaços Toblerianos Completos (ETC). “Toblerianos”, em referência ao pesquisador W. R. Tobler, que dedicou anos de seu trabalho ao estudo de novas formas de representação de dados espaciais. “Completos”, por estar aberto a qualquer tipo de métrica, ou método, que contribua para uma cartografia que conduza a novas formas de explicação sobre a geografia dos relacionamentos espaciais, e da interpretação das relações entre os lugares e as redes. Um conceito ora proposto, sobre o qual é desenvolvida uma implementação computacional da metodologia apresentada.

1.4 Prova de Conceito: novas cartografias para a Pesquisa Origem-Destino da Companhia do Metropolitano de São Paulo - Metrô

A Pesquisa de Origem e Destino (METROPOLITANOS, 2008) realizada pela Companhia do Metropolitano de São Paulo tem periodicidade de dez anos, e visa caracterizar os deslocamentos diários na Região Metropolitana de São Paulo (RMSP) para a projeção de viagens em horizontes futuros. Neste caso, busca-se avaliar projetos de expansão e de reestruturação de redes de transporte por meio de modelos de simulação: por metrô, trem ou ônibus.

A metodologia aplicada na pesquisa consiste de uma coleta de informações atualizadas sobre as viagens realizadas pela população na RMSP em dia útil típico. Os resultados obtidos estão condensados em um conjunto de planilhas com informações de origem e destino, por modo principal de transporte (ex. trem, ônibus ou metrô), por tipo de transporte (ex. coletivo ou individual), por quantidade de viagens, e pelo tempo médio das mesmas.

Com o intuito de oferecer novas possibilidades de análise sobre dados de fluxo, foi conduzida uma prova de conceito com dois experimentos, onde a metodologia apresentada no Capítulo 3 é aplicada a um conjunto de dados selecionados do relatório da Pesquisa OD/2007.

As características dos experimentos e seus respectivos subconjuntos de dados são descritos a seguir:

- a) Experimento 1: Foram selecionadas 13 zonas de pesquisa localizadas na região central do conjunto de dados, considerando o atributo “tempo médio de duração das viagens” realizadas entre as mesmas.
- b) Experimento 2: Foram selecionadas 13 zonas de pesquisa localizadas na região central do conjunto de dados, considerando o atributo “número total de viagens” realizadas entre as mesmas.

Como resultado da metodologia, foi possível representar o espaço em função de algumas das variáveis produzidas pela pesquisa, o que ofereceu uma nova perspectiva de análise sobre a dinâmica de construção e modificação do espaço regional em estudo.

1.5 Inserção na agenda científica

O presente trabalho está inserido na linha de pesquisa Análise Espacial e Mineração de Dados Espaciais - Novas Cartografias, do Instituto Nacional de Pesquisas Espaciais (INPE), que busca desenvolver técnicas de representação de dados espaciais com implicações na análise exploratória de grandes bases de dados geoespaciais.

1.6 Organização da dissertação

A presente dissertação está organizada em cinco capítulos:

- a) O Capítulo 1 apresenta a motivação, o objetivo, a contribuição do trabalho, e introduz a prova de conceito utilizada nos experimentos desenvolvidos no transcurso do mesmo.
- b) O Capítulo 2 traz a fundamentação teórica com os principais conceitos discutidos ao longo deste trabalho.
- c) O Capítulo 3 apresenta o conceito do *Framework* ETC, e a metodologia computacional para a criação da representação de espaços de fluxos.

- d) O Capítulo 4 traz uma prova de conceito com dois experimentos, onde a teoria abordada no Capítulo 2, e o ambiente computacional do Capítulo 3, são aplicados aos dados da pesquisa OD/2007.
- e) O Capítulo 5 apresenta algumas conclusões acerca dos resultados obtidos, assim como são discutidas algumas perspectivas futuras de aplicação do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

O presente capítulo busca apresentar as principais teorias e conceitos que servem de base para este trabalho.

2.1 Os SIGs atuais e o conceito de proximidade

Desde a antiguidade, o homem busca representar as características físicas e sociais do espaço em que vive, utilizando desenhos e símbolos como forma de armazenar e compartilhar o conhecimento geográfico.

Criados para representar uma determinada porção do espaço, os mapas são abstrações da realidade que dão ênfase a objetos, e relações espaciais, escolhidos segundo um critério de interesse (MACEACHREN, 1994). Mais do que isso, são ferramentas indispensáveis para a compreensão de como o homem e os processos físicos agem, e interagem, sobre a superfície da Terra (GOODCHILD, 2004).

Durante séculos, o papel foi amplamente utilizado como meio de representação e disseminação de informações espaciais. No entanto, devido a sua natureza estática, impunha restrições à capacidade de análise e interpretação de como os processos evoluíam ao longo do tempo e do espaço.

Foi o advento do uso dos computadores como meio digital de representação, e conseqüentemente o aparecimento dos SIGs, que elevaram o estudo dos fenômenos espaciais a um outro patamar.

Nesse contexto, os SIGs surgiram em oposição à natureza estática dos mapas em papel, como uma evolução necessária na análise dos fenômenos geográficos. Por consequência, diversos avanços foram observados, como a possibilidade de integração de conjuntos de dados espaciais provenientes de diferentes fontes, e ainda a criação de uma variedade de produtos digitais combinando dados espaciais, e não espaciais.

Segundo Câmara et al. (2003), as estruturas de dados e algoritmos disponíveis em um determinado estágio de evolução tecnológica de um SIG condicionam as formas de expressão computacional dos problemas geográficos. E na década de 60, quando os SIGs surgiram, as tecnologias disponíveis eram limitadas. A

simplificação da representação espacial foi a solução encontrada para contornar o problema.

Foi movido por esse pensamento de simplificar a realidade geográfica, que o geógrafo suíço-americano Waldo R. Tobler propôs que o valor medido de uma variável em uma dada localização no espaço, tendia aos valores observados nas regiões mais próximas. Um princípio que ficou posteriormente conhecido como a “Primeira Lei da Geografia”, e dizia que: “Todas as coisas são parecidas, mas as coisas mais próximas se parecem mais do que coisas mais distantes” (TOBLER, 1970).

Muitos intérpretes de Tobler, no entanto, relacionam de forma direta o conceito de proximidade ao conceito de espaço com uma métrica de distância euclidiana. Ou seja, uma métrica definida como sendo uma função que associa a cada par de pontos $A \{a_1, a_2, a_3, \dots, a_n\}$ e $B \{b_1, b_2, b_3, \dots, b_n\}$, um número real d , não negativo, chamada de distância do ponto A ao ponto B , tal que:

$$d_{A,B} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (2.1)$$

Com base na repercussão da proposta de Tobler, muitos dos modelos geográficos foram desenvolvidos baseados na métrica euclidiana (TOBLER, 1993c. (Technical report 93-1)) e representados em um plano bidimensional (2D). Desse modo, dois pontos estão mais próximos quanto menor for o comprimento do segmento de reta $d_{A,B}$ que os une.

Segundo o próprio Tobler (1993b. (Technical report 93-1)), a responsabilidade pelo “pensamento euclidiano” é das escolas do ensino fundamental, que lecionam a geometria de Euclides ao invés da geometria elíptica, mais indicada para o estudo da Terra. Como consequência, grande parte dos nossos mapas são construídos em um sistema coordenado euclidiano, e nossas teorias sobre distribuição espacial são descritas cartograficamente, e matematicamente, em termos euclidianos (GOLLEDGE; HUBERT, 1982).

Apesar do conceito proposto por Tobler ser reconhecidamente de grande utilidade na análise de fenômenos espaciais (SUI, 2004), sofreu algumas críticas com o passar dos anos. Miller (2004), por exemplo, reflete sobre como o con-

ceito de proximidade apresentado é vago, e indefinido. No entanto, o próprio [Tobler \(2004\)](#) veio posteriormente esclarecer, que seu único intuito era aliviar a análise de possíveis fatores indesejáveis, concentrando-se nos efeitos locais.

2.2 A globalização dos relacionamentos espaciais

Considerando a aplicabilidade da “Primeira Lei da Geografia” nos efeitos locais, conclui-se que a mesma restringe a análise de certos fenômenos espaciais. Fato que vem sendo discutido no âmbito da comunidade científica por diferentes pesquisadores ([CÂMARA et al., 2003](#); [MILLER, 2004](#); [SUI, 2004](#); [TOBLER, 2004](#)).

No estudo da disseminação de doenças contagiosas, por exemplo, os relacionamentos espaciais podem surgir entre objetos não adjacentes, independentemente da contiguidade física. Como no caso das pandemias, que são epidemias de proporções globais.

Um fato recente que confirma o exposto anteriormente ocorreu no ano de 2009. Naquela oportunidade, o mundo viveu a experiência do vírus Influenza A (H1N1), que supostamente teve origem no México, e que rapidamente se propagou por todo o globo (figura 2.1).



Figura 2.1 - Casos registrados de propagação do vírus Influenza A (H1N1) - Jun 2009.
Fonte: [Brownstein et al. \(2009\)](#).

Como consequência, áreas desagregadas no espaço apresentaram os mesmos sintomas em seus habitantes, assim como o mesmo quadro de evolução da doença.

Os avanços experimentados nas últimas décadas pelos meios de comunicação e de transporte são apontados como os grandes “vilões”. Isso porque os mesmos facilitam e aceleram a dinâmica de surgimento desses relacionamentos.

As motivações que no passado determinaram a adoção de um modelo simplificado de representação já não são mais justificáveis, dado o aumento significativo da disponibilidade de dados geográficos. Além do que, há uma grande variedade de dispositivos móveis com recursos de captura de informações geográficas, como receptores GPS, PDAs e telefones celulares, cada vez mais ao alcance dos usuários de SIG.

Segundo [KRAAK e MACEACHREN \(2005\)](#), estima-se que 80% de todo dado digital produzido atualmente, inclua algum tipo de referência espacial (ex. coordenadas, endereço ou código postal). Por que então, os métodos de representação disponíveis em um ambiente de SIG ainda são os mesmos?

Com esse pensamento, algumas iniciativas de pesquisa entre grupos interdisciplinares foram criadas. Entre seus objetivos estão: a consideração de novas formas de representação de dados espaciais; a coordenação de diferentes trabalhos na área de geociências; a busca do desenvolvimento de novos métodos computacionais de representação; a criação de novas ferramentas de interação com dados espaciais; o estudo de mapas cognitivos; a integração e gerenciamento do acesso a dados distribuídos.

Muitos dos avanços significativos na compreensão de fenômenos espaciais complexos podem ser atribuídos aos avanços na representação dos mesmos ([PRAGER, 2008](#)). Desse modo, a busca por novas possibilidades de representação do espaço, e dos relacionamentos espaciais, deve ser um objetivo constantemente perseguido.

2.3 As redes e os fluxos - uma alternativa

O estudo do comportamento dos fluxos é de grande importância na análise de fenômenos geográficos, uma vez que muitas das mudanças que ocorrem no

mundo estão relacionadas com a busca, ou transferência, de diferentes tipos de recursos, ou serviços. Da mesma forma que o estudo dos atributos dos lugares, visto que são justamente as diferenças de características entre os mesmos as responsáveis pela dinâmica de surgimento desses relacionamentos (GOULD; WHITE, 1986).

A procura pelo atendimento de emergência no Sistema Único de Saúde (SUS) de um município qualquer, por exemplo, pode sobrecarregar a capacidade de um determinado hospital. Essa sobrecarga talvez esteja relacionada com as condições de acessibilidade do hospital em comparação com as demais opções disponíveis na região. É fácil compreender, que na procura por um serviço de emergência, a distância já não é tão importante, sendo o tempo um fator primordial de escolha (figura 2.2).

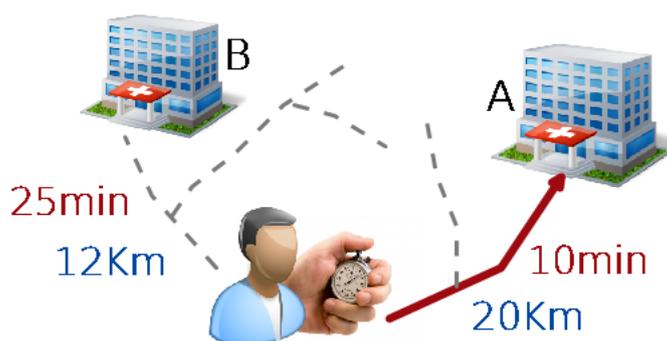


Figura 2.2 - O conceito de proximidade em função do tempo.

Uma análise dos fatores que influenciaram esse aumento de demanda pode ser de grande interesse na aplicação de políticas públicas na região.

Desse modo, caracteriza-se uma nova abordagem para o conceito de proximidade em função da acessibilidade de um ponto. No caso em questão, o mais próximo é o que se chega em menor tempo. No entanto, a escolha poderia ser influenciada por qualquer outra propriedade do espaço, ou derivada desta.

O exemplo apresentado chama a atenção para uma poderosa propriedade dos fluxos: a possibilidade de sua estruturação na forma de redes. Nesse caso, redes de transporte. Essas redes têm seu papel na condução dos fluxos de relaciona-

mentos, indicando direções preferenciais de conexão.

Cabe ressaltar que as redes podem apresentar configurações distintas em função de diferentes fatores: a variável considerada, os atores envolvidos no processo, o período de análise, entre outros. Sendo assim, diferentes propriedades produzem diferentes configurações de fluxos, que por sua vez geram diferentes representações do espaço (figura 2.3).

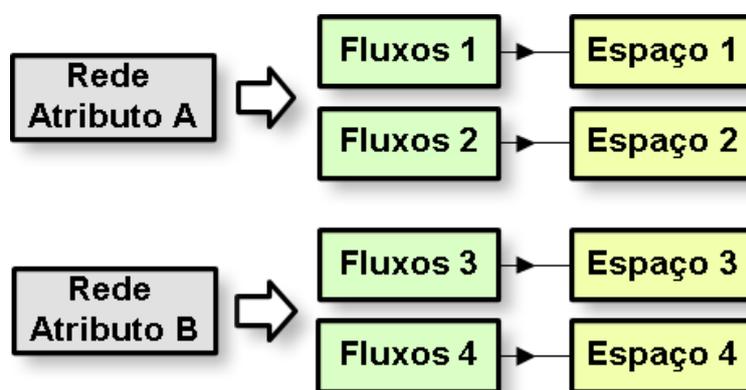


Figura 2.3 - Diferentes possibilidades de representação do espaço.

Desse modo, o conhecimento das características de construção das redes permite relacionar as diferentes localizações do espaço como uma função da conectividade entre as mesmas, ou seja, como decorrência da organização das relações sociais e econômicas, e da alta frequência de interações existentes.

2.4 A representação dos fluxos

Ao longo dos anos, vários trabalhos foram publicados sobre o estudo e a representação de diferentes tipos de fluxos. Fluxos de migração (CLARK, 1982; WORBOYS et al., 1997), fluxos de informação (WALMSLEY, 1983), de riquezas (TOBLER, 1986), entre outros. Alguns dando uma ênfase maior à simbolização e a representação, como Tobler, e outros à análise, como Walmsley.

No tocante à representação de um relacionamento com base nos fluxos, um fixo pode desempenhar diferentes papéis. Pode por exemplo, ser a origem do relacionamento, o destino, ou até mesmo os dois.

Para o caso em que o fixo é a origem do relacionamento, há um valor que quantifica o atributo que migra do ponto em questão para outro ponto no espaço (figura 2.4). O fixo como destino, o raciocínio é o inverso. Existe ainda o caso em que o fixo assume os dois papéis (autorrelacionamento), o que pode ocorrer quando pontos no espaço são agregados na forma de áreas, perdendo sua individualidade. Nesse caso, seus valores são analisados de forma conjunta.

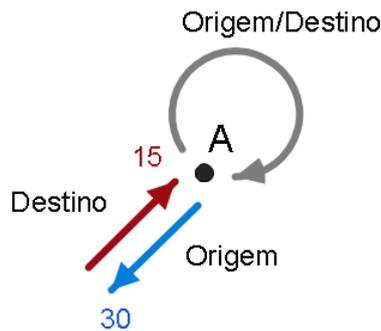


Figura 2.4 - Os papéis dos fixos nos dados de origem e destino.

Normalmente, a representação de grandes bases de dados de origem e destino é feita na forma de tabelas (figura 2.5).

| ORIGEM | DESTINO | | | | | | --- |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----|
| | P ₁ | P ₂ | P ₃ | P ₄ | P ₅ | P ₆ | |
| P ₁ | 27.09 | 39.82 | 38.63 | 44.82 | 25.60 | 20.98 | --- |
| P ₂ | 26.33 | 15.00 | 30.00 | 0.00 | 0.00 | 110.04 | --- |
| P ₃ | 25.00 | 30.00 | VALOR | | 44.38 | 15.00 | --- |
| P ₄ | 27.01 | 57.63 | VALOR | | 30.00 | 0.00 | --- |
| P ₅ | 24.28 | 0.00 | 30.47 | 25.48 | 25.00 | 0.00 | --- |
| P ₆ | 57.64 | 0.00 | 10.00 | 0.00 | 37.67 | 0.00 | --- |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |

Figura 2.5 - Exemplo de representação tabular de dados de fluxo.

Segundo Jonge (1992), a utilização de tabelas para a representação de dados de origem e destino tende a dificultar a análise e a identificação dos padrões de relacionamentos existentes entre os fixos. Visualmente, agrupar os dados pode ajudar a reduzir a complexidade da sua representação quando a quantidade de fluxos aumentar.

Uma característica da representação dos fluxos é a possibilidade de oferecer uma pronta análise visual. Em [Walmsley \(1983\)](#), por exemplo, a espessura dos fluxos é relacionada à quantidade de baixas nos exércitos de Napoleão, durante a Campanha da Rússia (figura 2.6).



Figura 2.6 - Representação das baixas do Exército de Napoleão na Campanha na Rússia.
Fonte: (WALMSLEY, 1983).

Existem ainda outras possibilidades de utilização dos fluxos, como a produção de espaços funcionais (MULLER, 1978; WALMSLEY, 1983; AHMED; MILLER, 2007; HIROSHI, 2007; L'HOSTIS, 2009).

2.5 A anisotropia das relações espaciais

Grande parte dos fenômenos geográficos envolvem relações espaciais fortemente anisotrópicas, e como visto anteriormente, passíveis de conexão por intermédio de redes. Isso ocorre, porque as propriedades dos elementos espaciais não são constantes, e variam no espaço de representação.

Nesse contexto, surgiram algumas iniciativas de pesquisa sobre a anisotropia do espaço geográfico e sua relação com a dinâmica de surgimento das relações espaciais (TOBLER, 1979; TOBLER, 1987; TOBLER, 1993c. (Technical report 93-1)). Como determinar em uma região, por exemplo, quais os pontos mais afastados

de um dado ponto central , em função do tempo de deslocamento (figura 2.7)?

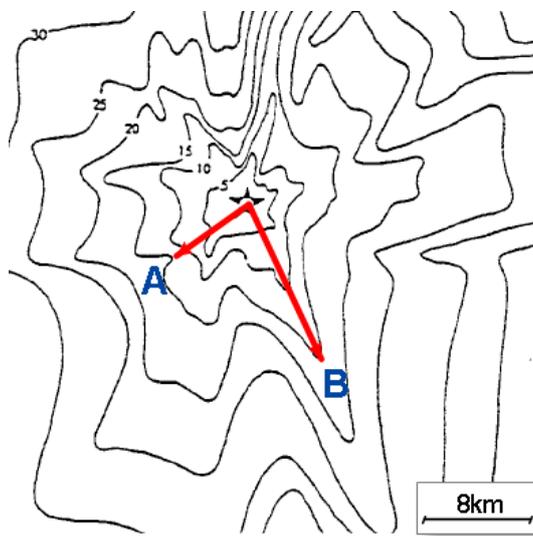


Figura 2.7 - Espaço anisotrópico (Dallas - EUA).

Fonte: Adaptado de Tobler (1993c. (Technical report 93-1)).

No exemplo da figura 2.7, são construídas isolinhas de tempo de 5 minutos sobre a rede de transporte da cidade de Dallas - EUA. Posteriormente, uma análise dos pontos destacados (*A* e *B*) permite concluir que para um deslocamento de 15 minutos, é possível chegar mais longe seguindo a direção do ponto *B*, do que a direção de *A*. O que pode indicar que as condições de deslocamento sobre a rede de transporte são mais favoráveis na direção daquele.

Nesse caso, são as características dos elementos representados que podem ter influenciado na determinação das direções preferenciais. Dentre essas características podem estar as condições de trafegabilidade da rede, ou qualquer outra propriedade que varie no espaço, e que acabe por retardar o deslocamento em determinada direção.

Com esse propósito, existem estudos semelhantes utilizando o tempo de deslocamento (figura 2.8) calculado sobre redes de transporte (MULLER, 1978; SPIEKERMANN; WEGENER, 1994; AHMED; MILLER, 2007; L'HOSTIS, 2009).

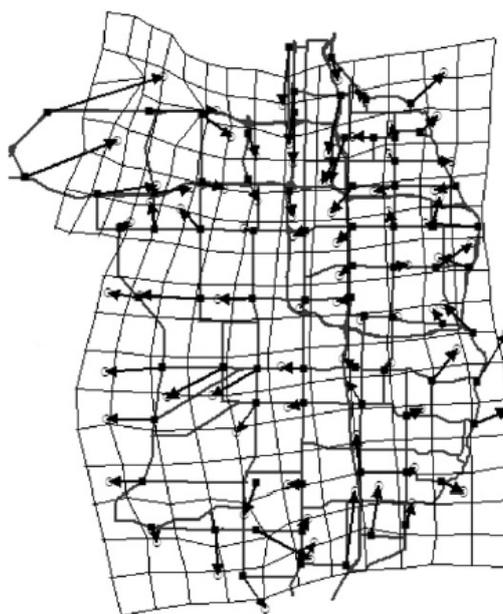


Figura 2.8 - Espaço em função do tempo (Salt Lake City - EUA).
Fonte: [Ahmed e Miller \(2007\)](#).

Observa-se que a utilização de uma nova variável na representação do espaço, acaba por inserir distorções na mesma. Distorções essas, que podem ser futuramente capturadas e analisadas por outras técnicas de representação, como a Regressão Bidimensional ([TOBLER, 1994](#)). Desse modo, procura-se representar essas distorções entre as configurações espaciais homólogas com base na comparação de formas ([DIAS M. H.; CAUVIN, 2000](#)).

2.6 A Regressão Bidimensional de Tobler

A técnica da Regressão Bidimensional foi aplicada por Tobler na década de 70, em um domínio geográfico, como uma extensão da regressão unidimensional para o caso em que, tanto a variável dependente Z , quanto a variável independente W , possuem duas dimensões ([TOBLER, 1994](#)). Originalmente, foi criada para avaliar o grau de semelhança entre duas, ou mais, configurações planares de um mesmo conjunto de pontos, cujas coordenadas são conhecidas em cada uma das representações ([FRIEDMAN; KOHLER, 2003](#)). Posteriormente, mostrou-se útil na comparação entre formas geométricas de uma maneira geral.

Na prática, cada uma das variáveis envolvidas representa um espaço bidimensional com n pontos (Figura 2.9)

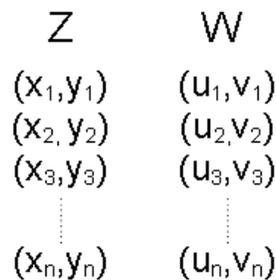


Figura 2.9 - Variáveis Z e W .

Um cuidado a ser tomado quando da utilização da Regressão Bidimensional é que os conjuntos de pontos devem ser homólogos, ou seja, as coordenadas em cada uma das representações devem estar associadas aos mesmos objetos no espaço geográfico. Uma das configurações deve ser tomada obrigatoriamente como “referência” (Z), ou seja, não paira dúvidas quanto aos valores das coordenadas dos seus pontos. As outras são denominadas de “imagem” (W), sendo comparadas com Z , aos pares.

2.6.1 Mapeando as transformações geométricas

Conhecidas as configurações espaciais homólogas Z e W , o mapeamento da transformação é realizado por uma função $W' = f(Z)$ (Figura 2.10).

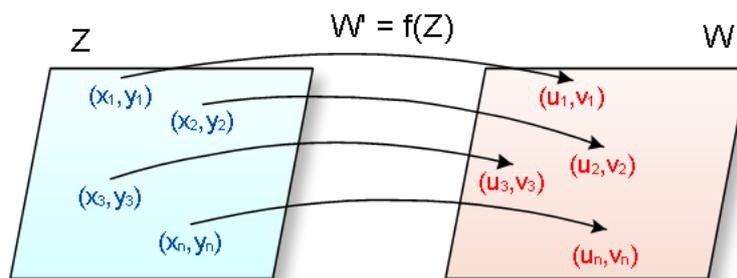


Figura 2.10 - Mapeamento de $Z(x, y) \rightarrow W(u, v)$.
 Fonte: Adaptado de: (TOBLER, 1994).

Desse modo, a similaridade entre as configurações é evidenciada pela magnitude das transformações geométricas necessárias para associar as variáveis de referência, e imagem.

Cabe ressaltar que a modelagem das transformações independe dos sistemas coordenados das configurações, e que o mapeamento $Z \rightarrow W'$ tem por objetivo ser o mais próximo possível de $Z \rightarrow W$ (TOBLER, 1994). Desse modo, é possível o desmembramento em duas outras funções (TOBLER, 1994):

$$\begin{aligned} u &= f(x, y) \\ v &= g(x, y) \end{aligned} \quad (2.2)$$

Determinadas as funções de transformação entre os sistemas, é possível calcular as coordenadas ajustadas de $W'(u', v')$ utilizando o Método dos Mínimos Quadrados (MMQ). Os vetores com origem nas coordenadas de referência $Z(x, y)$, e destino nas coordenadas $W'(u', v')$, são então chamados de vetores principais de distorção (Figura 2.11), e representam a direção e a magnitude das distorções associadas a cada par de pontos homólogos (SYMINGTON et al., 2002).

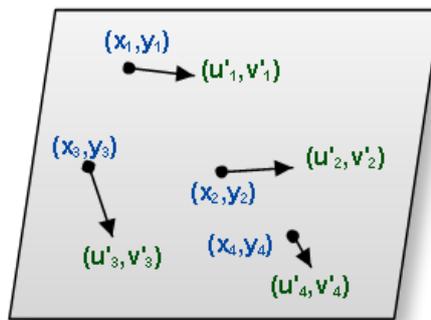


Figura 2.11 - Vetores principais de distorção.

Após a aplicação da Regressão Bidimensional, é possível verificar o quão boa é a estimativa de $W'(u', v')$ utilizando o coeficiente de correlação bidimensional de Tobler R^2 (equação 2.3), onde \bar{x} e \bar{y} correspondem à média dos valores de x e y , respectivamente (AHMED; MILLER, 2007).

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - u'_i)^2 + \sum_{i=1}^n (y_i - v'_i)^2}{\sum_{i=1}^n (x_i - \bar{x}_i)^2 + \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2.3)$$

O valor do coeficiente varia entre 0 e 1, e quanto mais alto, mais similares são as configurações espaciais. No entanto, por se tratar de dados espaciais, uma representação visual pode ser mais esclarecedora em alguns casos.

Segundo (CAUVIN, 2005), a técnica da Regressão Bidimensional pode ser dividida em duas fases principais: um ajustamento precedido de uma transformação geométrica de modo a fornecer a representação do Sistema Imagem Ajustado (W'); e uma interpolação com base nos vetores principais de distorção, de forma a generalizar os resultados para todo o espaço (W'').

2.6.2 A Fase de Ajustamento

Ao iniciar a fase do ajustamento, existem duas configurações espaciais de pontos. A variável $Z(x, y)$ representando o espaço de referência, e a variável $W(u, v)$ representando um sistema de coordenada qualquer. A transformação entre os sistemas é então dividida em duas fases distintas: definição dos parâmetros de transformação ajustados, e a transformação propriamente dita.

Para a definição dos parâmetros envolvidos no processo de transformação é preciso selecionar o modelo matemático que será utilizado para associar as coordenadas entre os referidos sistemas, ou seja, qual o tipo de transformação a ser utilizada (figura 2.12).

Embora Tobler (1994) acene com a possibilidade de utilização de transformações não lineares, as transformações mais empregadas no uso da Regressão Bidimensional são: a Transformação Linear de Similaridade (ou de Conformidade) e a Transformação Linear Afim (KITCHIN; FOTHERINGHAM, 1998; SYMINGTON et al., 2002; FRIEDMAN; KOHLER, 2003; KARE et al., 2010).

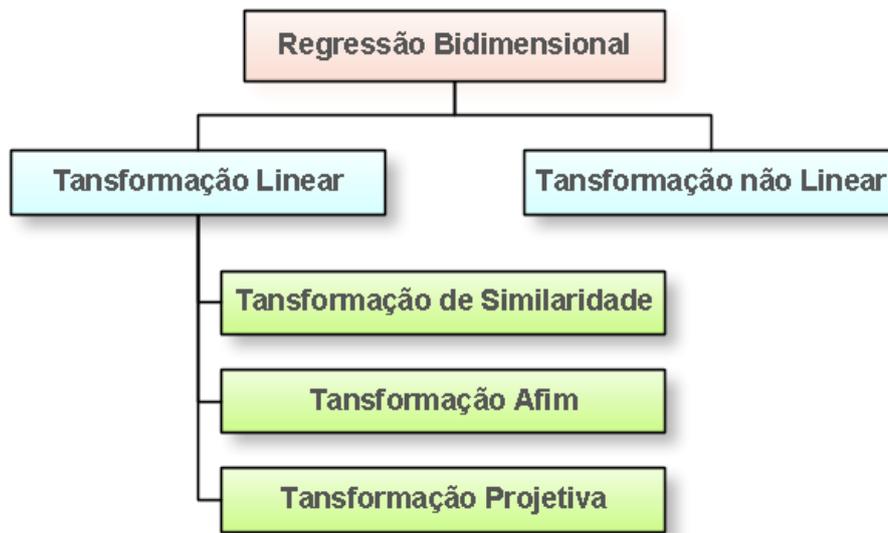


Figura 2.12 - Modelos de transformação.

As Transformações de Similaridade, e Afim, encontram-se melhor detalhadas no Apêndice A. As demais não são objeto deste estudo.

2.6.3 A Fase de Interpolação

O objetivo da interpolação é extrapolar os resultados obtidos para todos os pontos pertencentes ao espaço, considerando os vetores de distorção formados pela diferença entre os pontos $W'(u', v')$ e $Z(x, y)$ (figura 2.11).

Neste estudo foi utilizado o método de interpolação *Thin-Plate Splines* (TPS), por ser um método de interpolação que atende as necessidades do trabalho, minimizando a intensidade das distorções envolvidas na transformação entre as duas configurações geométricas bidimensionais (DRYDEN; MARDIA, 1998).

Segundo Yang e Lim (2005), dado um conjunto de n pontos de controle de $Z(x, y) \in R^2$, a interpolação TPS ($f(x^*, y^*)$) representa o mapeamento que torna a energia E_f mínima, tal que:

$$E_f = \int \int (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy \quad (2.4)$$

Desse modo, o valor interpolado de qualquer par ordenado (x^*, y^*) é fornecido

pela seguinte equação:

$$f(x^*, y^*) = a_1 + a_2x^* + a_3y^* + \sum_{i=1}^n G_i U(R) \quad (2.5)$$

onde

$$U(R) = r^2 \ln r^2 \quad (2.6)$$
$$r = \|Z_i - (x^*, y^*)\|$$

No conjunto de equações apresentadas, Z_i representa o conjunto de n pontos de controle pertencentes ao sistema de referência. Informações mais detalhadas sobre a aplicação da técnica podem ser encontradas em [Lim e Yang \(2005\)](#) e [Bookstein \(1989\)](#).

2.6.4 Aplicações da Regressão Bidimensional

Inicialmente, a Regressão Bidimensional surgiu com o propósito de comparar mapas elaborados em diferentes épocas, realizando inferências sobre as técnicas disponíveis e as metodologias utilizadas para tal (figura 2.13).

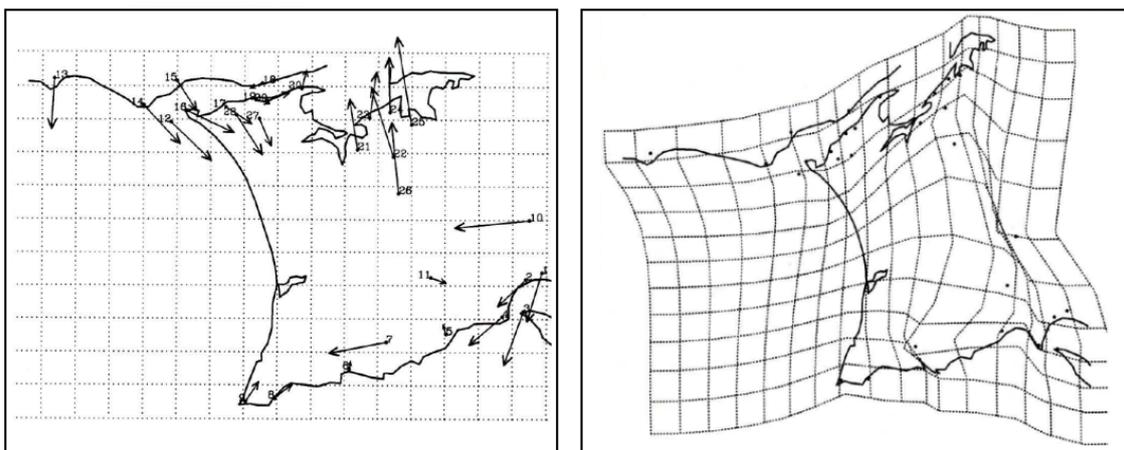


Figura 2.13 - Comparação entre mapas da Península de Setúbal (1757).
Fonte: Dias e Cauvin (2000).

Com o passar dos anos, áreas do conhecimento como a psicologia tem empregado a Regressão Bidimensional no estudo de mapas cognitivos, comparando diferentes percepções mentais sobre um mesmo espaço geográfico (CAUVIN, 2002; FRIEDMAN; KOHLER, 2003; PASKINS, 2005).

A técnica também é utilizada no estudo da acessibilidade de regiões (FORER P. C., 1981; MAYHEW, 1981), e na definição da hierarquia de lugares (Figura 2.14).

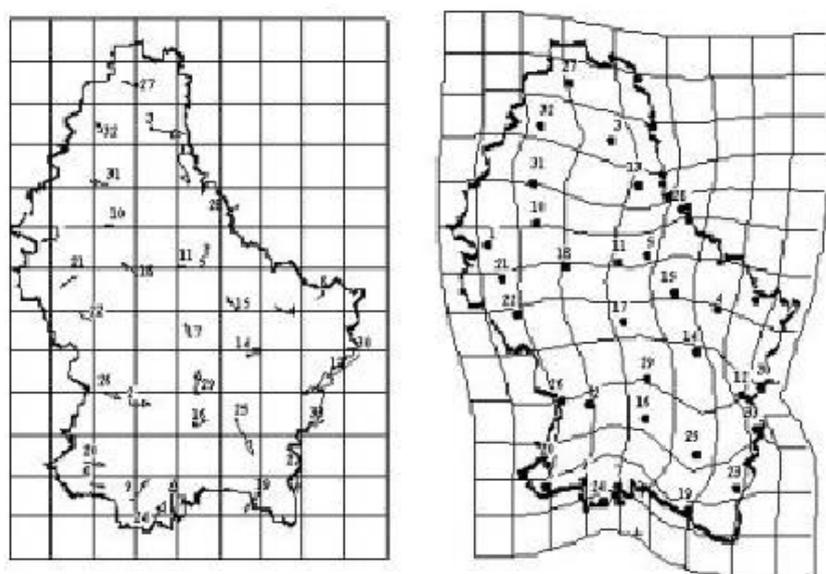


Figura 2.14 - Estudo da hierarquia de lugares em Luxemburgo.
Fonte: (CAUVIN, 2005).

Outra possibilidade de aplicação é apresentada em Kare et al. (2010), onde utiliza-se a Regressão Bidimensional na comparação entre rostos humanos. Para tal, as coordenadas dos conjuntos de pontos homólogos são determinadas com base em estudos antropológicos e técnicas de reconstrução cirúrgica da face. Existem ainda trabalhos utilizando a variável tempo na criação de espaços funcionais (MARCHAND, 1978; AHMED; MILLER, 2007; AXHAUSENA et al., 2008; LHOSTIS, 2009).

A aplicação da Regressão Bidimensional pressupõe dois conjuntos de dados (Z e W). No entanto, é possível que W seja desconhecido *a priori*. Nesse caso, o mesmo pode ser determinado utilizando-se alguma técnica de análise multidi-

mensional, como por exemplo, o Escalonamento Multidimensional (MDS).

2.7 O Escalonamento Multidimensional

Segundo [Kruskal e Wish \(1978\)](#), o MDS envolve uma classe de técnicas que utilizam como dados de entrada o conceito de proximidade entre um conjunto de objetos. Nesse caso, a proximidade é representada por um valor que indica o quão parecidos (similares), ou diferentes (dissimilares), são tais objetos. Como resultado, o MDS recupera a representação espacial multidimensional embutida em uma matriz de proximidade ([FRIEDMAN; KOHLER, 2003](#)).

Existem diferentes tipos de MDS de acordo com as características dos dados que se deseja trabalhar. O MDS métrico, por exemplo, é caracterizado pela comparação entre os valores de proximidade, enquanto o não métrico por um ordenamento dos pares baseado nas proximidades ([HAIR et al., 1998](#)). Neste estudo, será utilizado o escalonamento multidimensional clássico (CMDS), que é métrico (considera a proximidade em função do valor de uma variável), e utiliza uma única matriz de dissimilaridades.

Segundo [Hair et al. \(1998\)](#), a utilização do método pode ser realizada com as seguintes etapas:

- a) Identificar os pares a serem comparados, e determinar a dimensão desejada para a solução. No caso deste trabalho, duas (2D).
- b) Identificar que tipo de relacionamento a variável observada representa: similaridade ou dissimilaridade.
- c) Calcular os valores de proximidade entre os pares, e comparar os relacionamentos (observado x calculado). Para tal, é possível utilizar uma medida de ajuste.

Um exemplo prático é apresentado por [Kruskal e Wish \(1978\)](#), onde são calculadas inicialmente as distâncias referentes ao deslocamento de avião entre dez cidades americanas. O resultado é uma matriz de dissimilaridades simétrica ($n \times n$) em função da variável “distância aérea” (figura 2.15).

| CITIES | ATLA. | CHIC. | DENV. | HOUS. | L.A. | MIAMI | N.Y. | S.F. | SEAT. | WASH D.C. |
|-----------------|-------|-------|-------|-------|------|-------|------|------|-------|-----------|
| ATLANTA | | 587 | 1212 | 701 | 1936 | 604 | 748 | 2139 | 2182 | 543 |
| CHICAGO | 587 | | 920 | 940 | 1745 | 1188 | 713 | 1858 | 1737 | 597 |
| DENVER | 1212 | 920 | | 879 | 831 | 1726 | 1631 | 949 | 1021 | 1494 |
| HOUSTON | 701 | 940 | 879 | | 1374 | 968 | 1420 | 1645 | 1891 | 1220 |
| LOS ANGELES | 1936 | 1745 | 831 | 1374 | | 2339 | 2451 | 347 | 959 | 2300 |
| MIAMI | 604 | 1188 | 1726 | 968 | 2339 | | 1092 | 2594 | 2734 | 923 |
| NEW YORK | 748 | 713 | 1631 | 1420 | 2451 | 1092 | | 2571 | 2408 | 205 |
| SAN FRANCISCO | 2139 | 1858 | 949 | 1645 | 347 | 2594 | 2571 | | 678 | 2442 |
| SEATTLE | 2182 | 1737 | 1021 | 1891 | 959 | 2734 | 2408 | 678 | | 2329 |
| WASHINGTON D.C. | 543 | 597 | 1494 | 1220 | 2300 | 923 | 205 | 2442 | 2329 | |

Figura 2.15 - Matriz de dissimilaridades com as distâncias aéreas.
 Fonte: (KRUSKAL; WISH, 1978).

O emprego do CMDS fornece a configuração espacial embutida na matriz de dissimilaridades, o que pode ser observado na figura 2.16.

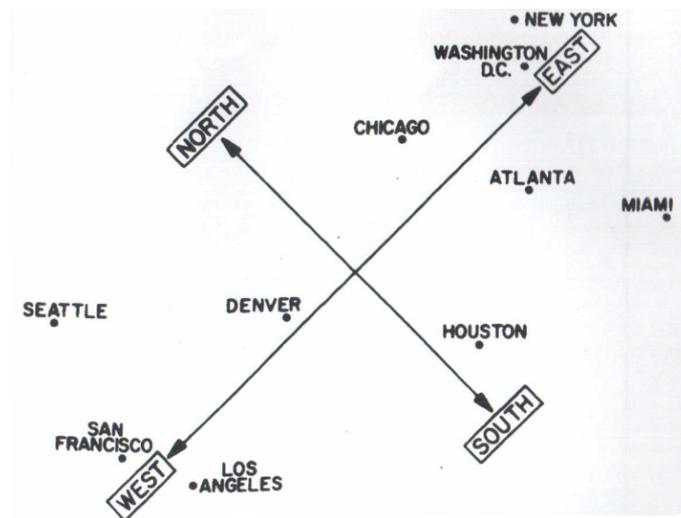


Figura 2.16 - Configuração espacial obtida pelo CMDS.
 Fonte: (KRUSKAL; WISH, 1978).

Verifica-se ainda, que as distâncias resultantes entre os pares de cidades (\hat{d}_{ij})

são diferentes das distâncias iniciais (d_{ij}). Segundo [Kruskal e Wish \(1978\)](#), essas diferenças são representadas por uma medida de ajuste chamada de *Stress* (equação 2.7). Essa medida é útil na avaliação da qualidade dos resultados obtidos, auxiliando na determinação da dimensão em que as distorções entre as configurações são mínimas.

$$Stress^2 = \frac{(d_{ij} - \hat{d}_{ij})^2}{(d_{ij} - \bar{d})^2} \quad (2.7)$$

Na equação 2.7, o termo d_{ij} representa a distância original entre cada par de objetos; \bar{d} a distância média original entre os pares de objetos; e \hat{d}_{ij} a distância calculada pelo MDS).

Quanto menor for o valor de *Stress* encontrado, mais semelhantes são as configurações espaciais. O ideal são Valores abaixo de 0.1 ([KRUSKAL; WISH, 1978](#)).

Segundo [Sturrock e Jorge \(2000\)](#), um alto valor de *Stress* indica que o número de dimensões escolhido para a representação do conjunto de dados não é adequado para capturar o complexo relacionamento entre os objetos em estudo.

A escolha das dimensões é um fator importante, e que influencia diretamente a capacidade de análise dos resultados. No entanto, não é determinante. Não adianta ter uma melhor adequação dos resultados calculados, se não for possível a sua interpretação. Por isso, os resultados são normalmente representados em duas (2D), ou três (3D) dimensões.

Dependendo do problema, pode ser feita uma análise subjetiva da solução para determinar qual dimensão oferece melhores condições de análise.

3 O FRAMEWORK ETC: UMA CONTRIBUIÇÃO PARA O DOMÍNIO DOS FLUXOS

Do ponto de vista computacional, o *framework* ETC é proposto neste trabalho para ser uma ferramenta de análise exploratória no estudo dos fluxos. É um repositório de classes escritas na linguagem C++, que funcionam em conjunto com a biblioteca geográfica TerraLib (CÂMARA et al., 2008).

No estágio atual de desenvolvimento em que se encontra, comporta as classes desenvolvidas para aplicar somente a metodologia de representação de dados de fluxo abordada neste trabalho. No entanto, com a capacidade de agregar no futuro outras funcionalidades de representação, ou adaptá-las para o funcionamento em conjunto com as já existentes.

3.1 A biblioteca TerraLib

A biblioteca geográfica TerraLib (CÂMARA et al., 2008) foi escolhida como suporte para o desenvolvimento do trabalho, por ser um projeto de *software* concebido para servir como base para o desenvolvimento colaborativo na comunidade de usuários, ou desenvolvedores de SIGs (CASANOVA et al., 2005).

Trata-se de uma biblioteca computacional com o código fonte aberto, distribuição livre, e com núcleo de desenvolvimento na Divisão de Processamento de Imagens (DPI) do Instituto Nacional de Pesquisas Espaciais (INPE).

A referida biblioteca encontra-se disponível para *download* no endereço eletrônico <<http://www.terralib.org>>, tanto na versão para desenvolvimento em ambiente Windows, quanto para Linux.

Cabe o registro, de que para a execução dos experimentos descritos no capítulo quatro deste trabalho, as classes desenvolvidas foram compiladas em ambiente Windows, com a versão 3.3.1 da mesma.

3.2 A arquitetura do *framework*

As classes desenvolvidas inicialmente neste trabalho têm o propósito de proporcionar a visualização de espaços de fluxos utilizando a técnica da Regressão Bidimensional. O principal objetivo é oferecer uma forma de representação de dados de fluxo diferente da normalmente encontrada nos SIGs disponíveis.

Conforme pode ser observado na figura 3.1, as classes desenvolvidas neste estudo formam camadas acima da TerraLib.

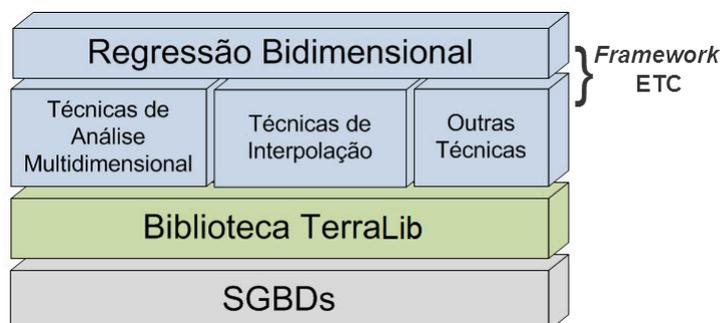


Figura 3.1 - Arquitetura de funcionamento com a TerraLib.

Desse modo, são aproveitadas diversas funcionalidades já disponíveis nas classes da própria TerraLib, principalmente para a execução de operações bem específicas, e julgadas como essenciais. São elas:

- conexão com os principais Sistemas Gerenciadores de Banco de Dados (SGBDs) disponíveis no mercado, utilizando para esse fim os *drivers* já implementados;
- recuperação e persistência no banco de dados das geometrias da base espacial utilizada no estudo (pontos, linhas e polígonos);
- realização de operações algébricas sobre matrizes, indispensáveis nos cálculos matemáticos presentes na metodologia (ex. produto, inversão e decomposição de matrizes);
- execução de operações que envolvam cálculo de coordenadas e projeção de dados espaciais.

Cabe ressaltar que a TerraLib possibilitou um ganho significativo no tempo de desenvolvimento, ao oferecer um valioso suporte aos algoritmos desenvolvidos neste trabalho. Desse modo, foi possível concentrar os esforços no desenvolvi-

mento das funcionalidades utilizadas na aplicação da metodologia em estudo, e na sua adequação à própria TerraLib.

3.3 Metodologia de representação de dados de fluxo

Com o propósito de oferecer uma nova possibilidade de representação de dados de fluxo em função de um conceito de proximidade baseado nos atributos dos objetos espaciais, é apresentada a metodologia da figura 3.2.

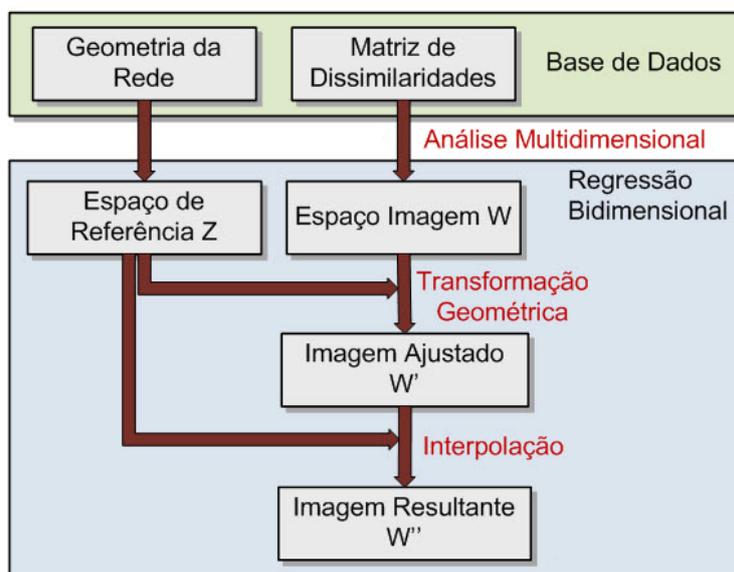


Figura 3.2 - Metodologia da representação de fluxos utilizando a Regressão Bidimensional.

Fonte: Adaptado de: Cauvin (2005).

3.3.1 Identificando as relações de proximidade

O primeiro passo para a aplicação da metodologia é definir no conjunto de dados o atributo representativo dos relacionamentos de origem e destino. Uma vez definido o atributo, é preciso analisá-lo sob o ponto de vista do conceito de proximidade, identificando-o como sendo um atributo de similaridade, ou de dissimilaridade.

A identificação do tipo de proximidade associado é uma tarefa simples, mas de fundamental importância posteriormente na análise multidimensional.

3.3.1.1 Dados de similaridade

No exemplo da figura 3.3 são apresentados três fixos (A , B e C) e dois relacionamentos distintos, ambos de similaridade. Em um deles, um investidor em B aplica R\$ 90.000,00 em um negócio imobiliário localizado em C . No outro, um paciente em A é consultado 15 vezes em um hospital, também em C .



Figura 3.3 - Relacionamentos de similaridade.

Observa-se que trata-se de duas variáveis distintas, ambas representando relações de similaridade. Desse modo, a relação entre os dados de similaridade e o conceito de proximidade podem ser compreendidos de uma forma simples: quanto maior o valor quantitativo do atributo, mais próximos são, ou estão esses objetos no contexto da análise.

De uma forma geral, atributos de similaridade são atributos que expressam quantidades de relacionamentos entre os pares de pontos analisados no espaço. Podem ser por exemplo, a quantidade de bens transferidos, de investimentos, de busca por serviços, ou de viagens. Desse modo, são quantificados os fluxos entre os pontos de origem e destino.

3.3.1.2 Dados de dissimilaridade

O raciocínio utilizado para a identificação de um atributo como sendo de dissimilaridade é o inverso do anterior. Neste caso, quanto maior o seu valor, menos próximos são, ou estão os objetos no espaço.

Um exemplo clássico de dissimilaridade é a distância. Outro bom exemplo utilizado no dia a dia é o *tempo de deslocamento* entre diferentes pontos no espaço.

Neste caso, quanto maior for o tempo para se chegar a um determinado destino, menos próximo esse ponto se encontra da origem.

No exemplo apresentado na figura 3.4, o atributo *tempo de viagem* é utilizado nos relacionamentos *A-C* e *B-C*. Verifica-se que o relacionamento *B-C* é concretizado mais rapidamente, pois quanto menor for o tempo de deslocamento entre dois pontos, mais próximos os mesmos estão.



Figura 3.4 - Relacionamentos de dissimilaridade.

Caso haja necessidade de uma análise utilizando atributos de dissimilaridade, as similaridades podem ser convertidas pela sua subtração de uma constante C . Nesse caso, a constante deve ser maior, que a maior das similaridades (DUNN-RANKIN et al., 2004), de forma que não hajam valores negativos, ou nulos.

Na verdade, dizer que uma dissimilaridade é nula pode ter algumas implicações indesejadas, pois é o mesmo que dizer que dois pontos que não estão relacionados, são o mesmo objeto no espaço.

3.3.2 Construindo a matriz de fluxos

Antes da construção da matriz de fluxos, é necessário verificar se a estrutura dos dados de origem e destino está em conformidade com os métodos de importação de dados das classes desenvolvidas. Isso implica em verificar se os dados vetoriais encontram-se em um dos formatos de importação suportados pela biblioteca TerraLib, e se os dados de origem-destino-valor estão no formato *.csv, utilizando o caractere “;” como separador.

Caso as coordenadas não sejam extraídas automaticamente da base espacial, podem ser obtidas a partir de uma tabela de coordenadas no mesmo formato dos dados de origem e destino, com três colunas. A primeira coluna é de identificação e contém os índices dos fixos envolvidos nos relacionamentos. As demais colunas apresentam as coordenadas de abscissa, e ordenada, respectivamente.

Desse modo, a base de dados utilizada na metodologia deve ser composta de duas partes:

- a) Os dados de origem-destino com os respectivos valores associados, estruturados na forma de planilhas em um formato de importação suportado pela biblioteca TerraLib.
- b) Os dados de geometria para a recuperação dos nós da rede, ou as planilhas contendo as coordenadas dos pontos envolvidos nos relacionamentos.

Seja um exemplo de configuração espacial de fluxos definida por um conjunto de fixos $P = (P_1, P_2, P_3, P_4, P_5)$ juntamente com os seus respectivos valores associados aos relacionamentos (figura 3.5).

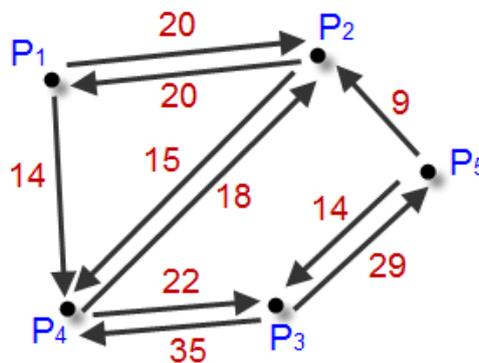


Figura 3.5 - Exemplo de uma configuração espacial de fluxos.

As relações de proximidade podem ser então estruturadas na forma de uma matriz M , onde a qualquer instante é possível recuperar a origem i , o destino

j , e o valor M_{ij} do atributo associado aos relacionamentos entre os fixos (figura 3.6).

| | P ₁ | P ₂ | P ₃ | P ₄ | P ₅ |
|----------------|----------------|----------------|----------------|----------------|----------------|
| P ₁ | 0 | 20 | 0 | 14 | 0 |
| P ₂ | 20 | 0 | 0 | 15 | 0 |
| P ₃ | 0 | 0 | 0 | 35 | 29 |
| P ₄ | 0 | 18 | 22 | 0 | 0 |
| P ₅ | 0 | 9 | 14 | 0 | 0 |

Figura 3.6 - Matriz de dados de fluxos.

Independente do atributo utilizado, o relacionamento com origem em P_2 , e destino em P_4 , possui o valor 15, e o relacionamento com origem em P_4 e destino em P_2 , tem o valor 18. Isso indica que a matriz não é simétrica. Como há a necessidade de que a matriz seja simétrica para a aplicação do CMDS, existe a opção de que os relacionamentos sejam escolhidos em função do objetivo da análise, e posteriormente utilizados para a construção de uma nova matriz, agora simétrica.

Na figura 3.7, observa-se uma matriz simétrica construída com base na matriz da figura 3.6, onde os círculos indicam os valores originais espelhados.

| | P ₁ | P ₂ | P ₃ | P ₄ | P ₅ |
|----------------|----------------|----------------|----------------|----------------|----------------|
| P ₁ | 0 | 20 | 0 | 14 | 0 |
| P ₂ | 20 | 0 | 0 | 18 | 9 |
| P ₃ | 0 | 0 | 0 | 35 | 29 |
| P ₄ | 14 | 18 | 35 | 0 | 0 |
| P ₅ | 0 | 9 | 29 | 0 | 0 |

Figura 3.7 - Montagem da matriz simétrica com base na escolha dos relacionamentos.

Uma outra opção mais simples, é duplicar os valores da matriz triangular inferior, ou superior. Na figura 3.8, por exemplo, observa-se uma matriz simétrica baseada na matriz diagonal superior da figura 3.6.

| | P ₁ | P ₂ | P ₃ | P ₄ | P ₅ |
|--------------------|----------------|----------------|----------------|----------------|----------------|
| P ₁ ... | 0 | 20 | 0 | 14 | 0 |
| P ₂ ... | 20 | 0 | 0 | 15 | 0 |
| P ₃ ... | 0 | 0 | 0 | 35 | 29 |
| P ₄ ... | 14 | 15 | 35 | 0 | 0 |
| P ₅ ... | 0 | 0 | 29 | 0 | 0 |

Figura 3.8 - Montagem da matriz simétrica com base na diagonal superior.

Quando o valor do relacionamento for nulo, deve-se tomar cuidado. Relacionamentos nulos indicam a inexistência do relacionamento entre os pares na base de dados, ou que o mesmo não fará parte da análise por algum motivo. No entanto, podem ser confundidos com um baixo grau de dissimilaridade.

3.3.2.1 O problema da matriz incompleta

Para um conjunto de n pontos, existe um total de $\frac{n(n-1)}{2}$ relacionamentos possíveis considerando uma matriz simétrica, e desconsiderando os autorrelacionamentos. Isso significa, que para um conjunto de 10 pontos, existem 45 possibilidades de relacionamentos, e que para 15 pontos, existem 105 relacionamentos. Isso pode se tornar um problema, pois quanto maior a quantidade de pontos, mais trabalhosa e indesejável é a coleta dos valores associados a tais relações (MALHOTRA et al., 1988).

Diante disso, grandes conjuntos de dados tem uma tendência a serem incompletos. Ou seja, em algum momento do processo de coleta de dados, ou não foi possível a sua identificação, ou simplesmente eles não existem. Como resultado é produzida uma tabela com vários valores de proximidade nulos.

Na literatura pesquisada, existem alguns trabalhos que oferecem alternativas para amenizar o problema da matriz incompleta. No entanto, algumas são aplicáveis ao escalonamento não métrico (SPENCE; DOMONEY, 1974) (GRAEF; SPENCE, 1979), e outras exigem que as proximidades representem a distância euclidiana (TSOGO et al., 1997) (YOUNG et al., 1982). Como este estudo utiliza o escalonamento métrico, e busca justamente expandir o conceito para outros atributos, as soluções não tiveram aplicabilidade. Tobler (1987) sugere ainda desconsiderar os dados abaixo da média dos valores do atributo estudado. Desse modo, busca-se reduzir o volume de dados nas grandes bases, sem no entanto fazer menção aos relacionamentos inexistentes.

Outra solução passível de adoção é a substituição dos valores nulos por valores superiores ao valor da maior das dissimilaridades. Com isso, procura-se “materializar” a dificuldade de construção do referido relacionamento na solução do escalonamento multidimensional.

Observou-se também que não há um consenso nos trabalhos pesquisados sobre a quantidade de pontos da amostra, e conseqüentemente da quantidade ideal de relacionamentos para a análise.

3.3.3 Mapeando as redes de fluxos

A forma tabular é a forma mais comum de apresentação de dados de origem e destino, e pode ser utilizada para estruturar os dados em um primeiro momento.

Caso o relacionamento não exista na tabela de dados original, o valor será preenchido com o valor nulo para posterior adoção de uma das soluções apresentadas anteriormente.

Conhecer as propriedades das redes que são utilizadas como dutos na condução desses fluxos, é de grande valia. Assim como a geometria da base de dados, visto que para cada tipo de geometria existe uma abordagem diferente.

De posse da matriz de relacionamentos, é necessário identificar as coordenadas dos nós que farão parte da rede. Isso pode ser feito com base em coordenadas previamente definidas, ou a partir das bases espaciais. Nesse caso, as mesmas devem estar em um formato de importação e armazenamento suportado pela TerraLib. Para essa importação, sugere-se o SIG TerraView, igual-

mente desenvolvido pelo INPE (2010), e disponível no endereço eletrônico <www.dpi.inpe.br/terraview>.

3.3.3.1 Base de dados de pontos

É o caso onde base de dados espacial é constituída de pontos. Uma vez identificados esses pontos, suas coordenadas são imediatamente associadas às coordenadas dos nós da rede (figura 3.9).

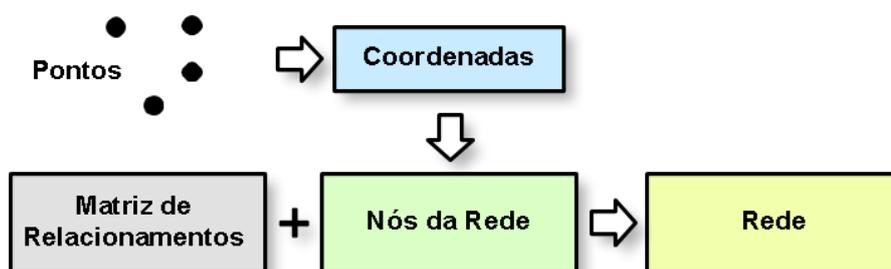


Figura 3.9 - Base espacial composta por pontos.

Definidos os nós, as arestas são diretamente associadas aos relacionamentos presentes na matriz de origem e destino, sendo que os valores nulos originais indicam a ausência de relacionamento.

3.3.3.2 Base de dados de linhas

Este é o caso onde, juntamente com a matriz de relacionamentos, existe uma base espacial composta por linhas, onde os pontos extremos são identificados, e associados às coordenadas dos nós da rede (figura 3.10).

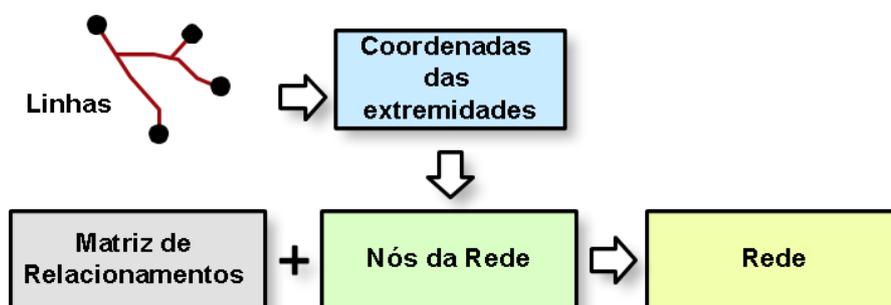


Figura 3.10 - Base espacial composta por linhas.

A base de linhas tem um papel fundamental, principalmente quando tratamos de atributos de custo, ou seja, atributos que podem ser acumulados ao longo da rede. Isso, porque sobre essa base podem ser utilizados algoritmos de caminho mínimo, como os algoritmos de [Floyd \(1959\)](#), e [Dijkstra \(1959\)](#).

3.3.3.3 Base de dados de polígonos

É o caso em que a base de dados espaciais é composta por polígonos, os valores de coordenadas dos nós da rede são associados diretamente aos centroides.

Neste caso, para a construção da rede existem duas possibilidades.

- a) Utilizar a matriz de relacionamentos onde as arestas da rede são representadas pelos próprios relacionamentos presentes. Desse modo, todo relacionamento representado na matriz fará parte da rede, interligando os pares de pontos relacionados.
- b) Utilizar os centroides dos polígonos adjacentes (figura 3.11). Esta é uma solução que propicia uma distribuição mais uniforme da rede no espaço em estudo, contribuindo para a solução da análise multidimensional a ser calculada posteriormente.

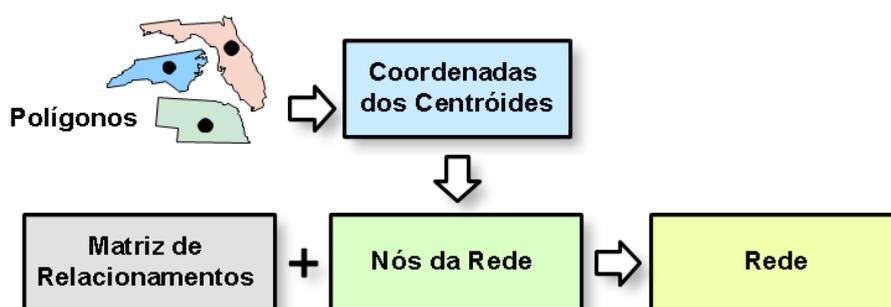


Figura 3.11 - Base espacial composta por polígonos.

Relacionamentos nulos presentes nos dados originais podem ser descartados, uma vez que as arestas entre os respectivos centroides não existam. Outras soluções podem ser adotadas e consideradas na análise dos resultados.

3.3.4 A análise multidimensional dos dados de origem e destino

Para a aplicação do escalonamento multidimensional, os dados de proximidade devem representar dissimilaridades, para que possam ser comparados com as distâncias originais. Caso haja necessidade, as similaridades devem ser previamente convertidas.

O objetivo do CMDS é recuperar a configuração espacial embutida na matriz de dissimilaridades, comparando para tal as dissimilaridades aos pares. Desse modo, é determinada uma nova configuração espacial de pontos (W), homóloga à configuração de referência (Z).

3.3.5 A Regressão Bidimensional de Tobler

A primeira fase da Regressão Bidimensional (Ajustamento) é realizada valendo-se de duas configurações espaciais homólogas, nesse caso, as coordenadas do Sistema de Referência (Z), e as coordenadas do Sistema Imagem (W).

Enquanto as coordenadas do Sistema de Referência são obtidas com base nos nós da rede de fluxos, as do Sistema Imagem são extraídas diretamente da matriz de relacionamentos utilizando o CMDS. Posteriormente, uma transformação linear determina as equações de transformação entre o Sistema de Referência e o Sistema Imagem.

As diferenças de coordenadas entre os pontos homólogos indicam as distorções do espaço em função da variável, mas somente nos pontos de referência. Desse modo, é necessária a utilização do método TPS para interpolar as distorções observadas para todos os pontos do espaço de representação (Interpolação).

Uma vez interpoladas as deformações locais para todo o espaço, é possível confrontá-las com a configuração original. Desse modo, os dados de OD são pré-processados em função dos objetivos da análise, submetidos ao CMDS, transformados e interpolados, de forma que as distorções observadas durante o processo sejam analisadas em função da variável de representação do espaço.

4 PROVA DE CONCEITO: O CASO DA PESQUISA ORIGEM-DESTINO PARA A RMSP DO METRÔ DE SÃO PAULO

O presente capítulo traz um estudo de caso com dois experimentos, onde a metodologia abordada no Capítulo 3 é aplicada aos dados da pesquisa OD/2007 (METROPOLITANOS, 2008), produzida pela Companhia do Metropolitano de São Paulo - Metrô.

4.1 A Pesquisa OD/2007 da Companhia do Metropolitano de São Paulo

A Pesquisa Origem e Destino, ou simplesmente Pesquisa O/D, é realizada desde 1967 na Região Metropolitana de São Paulo (RMSP), com periodicidade de dez anos, e tem por objetivo o levantamento de informações atualizadas sobre as viagens realizadas pela população da metrópole em dia útil típico.

Os resultados obtidos na pesquisa estão condensados em um conjunto de planilhas, sendo três dessas utilizadas neste trabalho. São elas: “Matriz de Modo Principal e Tempo Médio”, “Matriz de Tipo de Viagem e Tempo Médio” e “Matriz Total e Tempo Médio”.

A matriz “Modo Principal e Tempo Médio” contém ao todo 62.036 registros não nulos, e 05 variáveis. A saber:

- **ZONAO:** identifica a zona de origem do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **ZONAD:** identifica a zona de destino do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **MODOP_ON_A:** identifica o modo principal da viagem considerado no relacionamento dentre as opções: ônibus, ônibus fretado, ônibus escolar, automóvel, táxi, metrô, trem, moto, bicicleta, a pé e outros.
- **FE_VIA_SUM:** quantifica o total de viagens diárias entre as zonas envolvidas nos relacionamentos.
- **DURACAO_ME:** quantifica o tempo médio, em minutos, para a realização da viagem.

A matriz “Tipo de Viagem e Tempo Médio”, possui ao todo 55.856 registros não nulos, e 05 variáveis. A saber:

- **ZONAO:** identifica a zona de origem do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **ZONAD:** identifica a zona de destino do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **MODOP_ON_A:** identifica o modo principal de viagem considerado no relacionamento dentre as opções: coletivo (ônibus, ônibus fretado, ônibus escolar, metrô e trem), individual (automóvel, táxi, moto e outros), a pé e bicicleta.
- **FE_VIA_SUM:** quantifica o total de viagens diárias entre as zonas envolvidas nos relacionamentos.
- **DURACAO_ME:** quantifica o tempo médio, em minutos, para a realização da viagem.

A matriz “Matriz Total e Tempo Médio” condensa os dados anteriores, possuindo ao todo 40.831 registros não nulos, e 04 variáveis. A saber:

- **ZONAO:** identifica a zona de origem do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **ZONAD:** identifica a zona de destino do fluxo, ou seja, uma das 460 zonas de pesquisa.
- **FE_VIA_SUM:** quantifica o total de viagens diárias entre as zonas envolvidas nos relacionamentos.
- **DURACAO_ME:** quantifica o tempo médio, em minutos, para a realização da viagem.

Os dados de fluxo originais da pesquisa foram convertidos para o formato *.csv, o que permite uma maior flexibilidade de sua utilização, facilitando sua conversão para outros formatos. Já a base espacial encontra-se no *datum* de Córrego Alegre, sendo representada utilizando a projeção UTM, fuso 23S.

Um exemplo da estrutura dos dados de relacionamento utilizada é apresentada na figura 4.1.

```
0.000;10.130;31.290;34.540;22.410
10.130; 0.000;21.160;24.410;35.740
31.290;21.160; 0.000;15.000;40.830
34.540;24.410;15.000; 0.000;56.130
22.410;35.740;40.830;56.130; 0.000
26.970;40.300;45.390;60.690;11.550
26.270;39.600;44.690;59.990;36.890
```

Figura 4.1 - Estrutura dos arquivos de dados de OD.

Ao todo, a pesquisa OD relaciona 460 unidades de análise dentro da RMSP, o que implica em um total de 211.140 relacionamentos possíveis (dado que a matriz original é assimétrica). Desse montante, somente 19% estão efetivamente presentes, ou seja, não são nulos.

Na figura 4.2, é possível observar todos os relacionamentos presentes na “Matriz Total e Tempo Médio”.

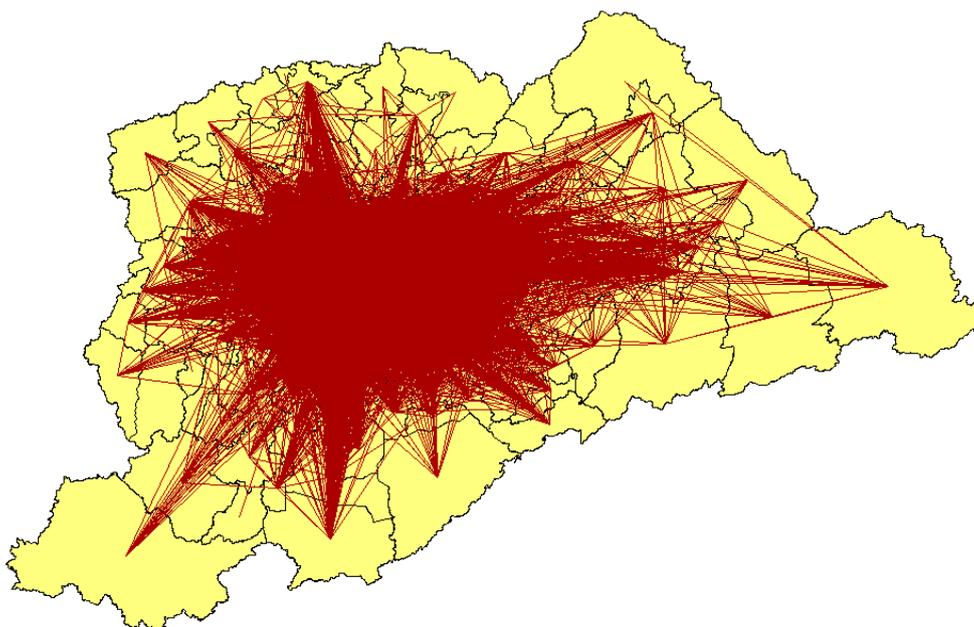


Figura 4.2 - Rede total de relacionamentos presentes no conjunto de dados.

Verifica-se uma dificuldade de análise visual dos fluxos representados, assim como identifica-se que a maior parte dos relacionamentos esta presente na região central da RMSP, existindo poucos relacionamentos entre as zonas periféricas. Esse “poder de atratividade” está possivelmente associado à oferta de trabalho, visto que os dados de viagem são referentes a um típico dia útil.

4.2 Experimentos

A divisão em 460 zonas de pesquisa foi uma decisão operacional do Metrô. No entanto, trata-se de um volume de dados considerável para ser utilizado em conjunto. Sendo assim, optou-se pela seleção de um subgrupo de dados, a ser analisado em dois experimentos distintos. Um dos experimentos utiliza o atributo “tempo médio de duração das viagens”, e o outro o “número total de viagens” entre as zonas da pesquisa.

A amostra utilizada no presente experimento possui 13 zonas de pesquisa e está localizada na região central da RMSP (figura 4.3).

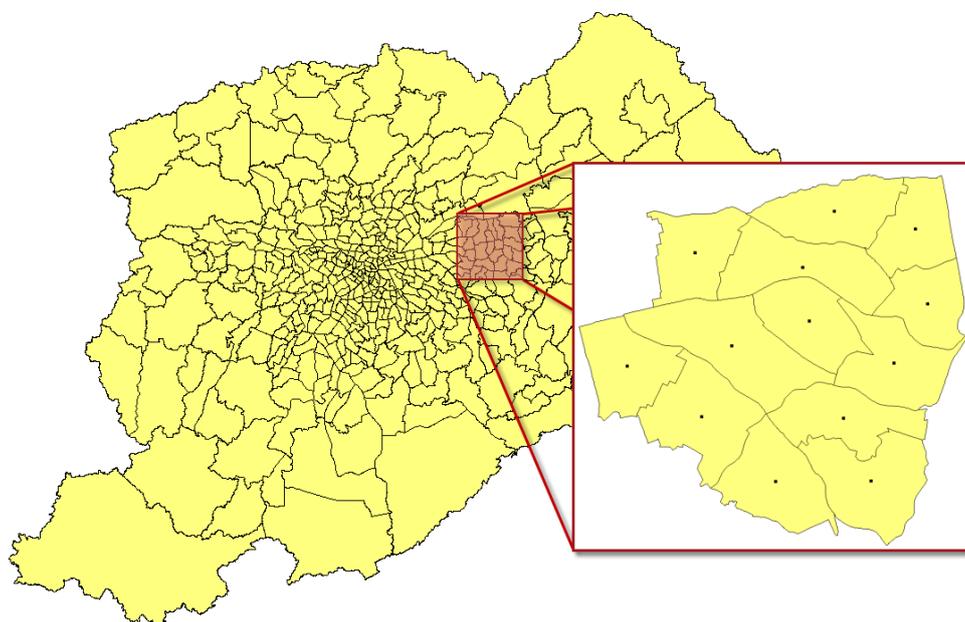


Figura 4.3 - Amostra selecionada dentre as 460 zonas de pesquisa da RMSP.

4.2.1 Experimento 1: Tempo médio de duração das viagens

4.2.1.1 Base de dados

A base de dados do experimento 1 refere-se a 13 zonas localizadas na região central da RMSP (figura 4.4), com seus respectivos relacionamentos em função da variável “tempo médio de duração das viagens”. Neste caso, a variável foi calculada considerando todos os modos principais de transporte disponíveis na pesquisa OD.

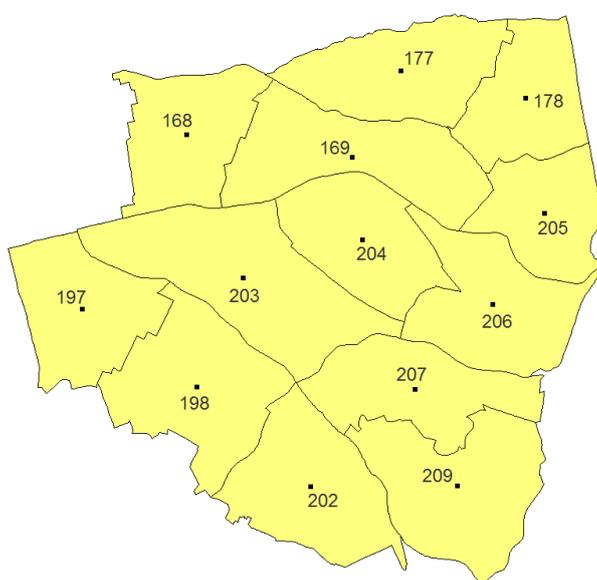


Figura 4.4 - Amostra de estudo do experimento 1.

Outro ponto a ser considerado é que a variável utilizada representa um atributo de dissimilaridade. Ou seja, quanto maior o seu valor, menos similares são as zonas relacionadas.

Desse modo, espera-se que após a aplicação da metodologia, os relacionamentos associados aos maiores valores na matriz, correspondam aos pontos mais afastados no espaço.

A identificação das zonas, juntamente com as coordenadas dos seus respectivos centroides, pode ser observada na tabela 4.1. Essas coordenadas têm um importante papel na construção da rede de relacionamentos.

Tabela 4.1 - Coordenadas dos centroides das zonas do experimento 1.

| Zona | Nome | N(m) | E(m) |
|------|----------------------|------------|-----------|
| 168 | Ticoatira | 7397843.98 | 343579.13 |
| 169 | Vila Esperança | 7397516.85 | 345987.99 |
| 177 | Ponte Rasa | 7398786.64 | 346699.99 |
| 178 | Águia de Haia | 7398377.92 | 348525.18 |
| 197 | Vila Califórnia | 7395274.51 | 342056.53 |
| 198 | Vila Carrão | 7394132.49 | 343724.85 |
| 202 | Aricanduva | 7392666.29 | 345385.32 |
| 203 | Vila Matilde | 7395733.94 | 344400.62 |
| 204 | Vila Guilhermina | 7396298.11 | 346138.60 |
| 205 | Cidade A.E. Carvalho | 7396692.59 | 348796.06 |
| 206 | Artur Alvin | 7395346.05 | 348037.33 |
| 207 | Cidade Líder | 7394101.29 | 346911.92 |
| 209 | Parque Savoy | 7392679.51 | 347522.44 |

4.2.1.2 Construindo a matriz de fluxos

Identificadas as zonas de análise, o próximo passo trata da construção da matriz de fluxos. Como a base de dados espacial é composta de polígonos, optou-se pela construção da rede de relacionamentos interligando-se os centroides das zonas adjacentes (figura 4.5).

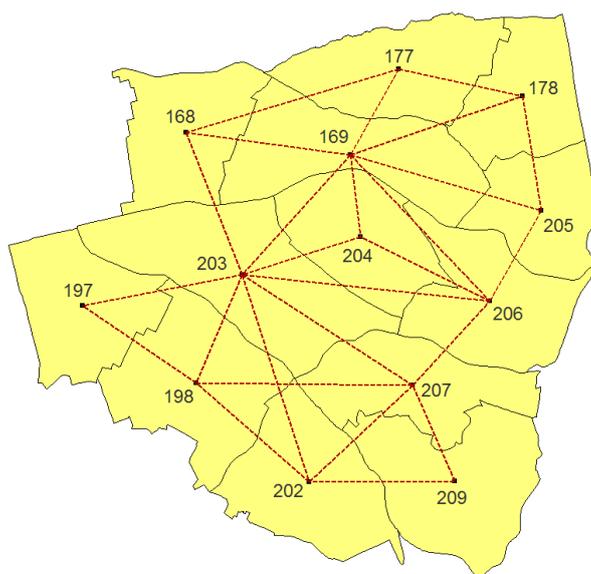


Figura 4.5 - Rede de relacionamentos do experimento 1.

Como consequência, os nós da rede correspondem aos centroides dos polígonos adjacentes, e as arestas representam os relacionamentos entre esses pares de nós. Neste ponto, algumas considerações devem ser feitas.

Como foram considerados somente os relacionamentos adjacentes para a construção da rede, aqueles que eventualmente ocorram entre polígonos não adjacentes são automaticamente desconsiderados. Da mesma forma, caso um relacionamento adjacente seja nulo na matriz de relacionamentos original da pesquisa, o mesmo não fará parte da construção da rede, e conseqüentemente não será utilizado nos cálculos realizados *a posteriori*.

As análises e ações que foram executadas sobre os dados originais selecionados para este experimento são resumidas a seguir:

- a) O tempo de deslocamento é um atributo de dissimilaridade, não havendo a necessidade de transformá-lo para sua utilização quando da aplicação do CMDS.
- b) A metodologia não considera os autorrelacionamentos. Desse modo, os mesmos são considerados nulos.
- c) Por se tratar de um atributo de custo que pode ser acumulado ao longo da rede, é utilizado um algoritmo de caminho mínimo para determinar os valores associados aos relacionamentos indiretos, ou seja, que contêm relações intermediárias. Desse modo, desde que todos os nós da rede sejam atingíveis, não há a possibilidade de existência de valores de relacionamentos nulos na matriz.
- d) Para a aplicação do escalonamento multidimensional, é necessário que a matriz de relacionamentos seja simétrica. Neste caso, optou-se por considerar como referência os valores contidos na matriz triangular superior. No entanto, a escolha poderia ser baseada em um outro critério pré-definido.
- e) Não foi realizado qualquer tipo de transformação na escala dos valores dos relacionamentos. Os valores apresentados na matriz correspondem aos valores de tempo de viagem originais da pesquisa (em minutos), e são apresentados na tabela 4.2.

Tabela 4.2 - Matriz de custo simétrica do experimento 1.

| Zona | 168 | 169 | 177 | 178 | 197 | 198 | 202 | 203 | 204 | 205 | 206 | 207 | 209 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 168 | 0.0 | 10.1 | 31.3 | 34.5 | 22.4 | 27.0 | 26.3 | 6.3 | 26.8 | 34.8 | 25.7 | 35.7 | 46.0 |
| 169 | 10.1 | 0.0 | 21.2 | 24.4 | 35.7 | 40.3 | 39.6 | 19.6 | 27.7 | 24.6 | 15.6 | 36.4 | 53.7 |
| 177 | 31.3 | 21.2 | 0.0 | 15.0 | 40.8 | 45.4 | 44.7 | 24.7 | 45.2 | 35.4 | 34.1 | 54.2 | 64.5 |
| 178 | 34.5 | 24.4 | 15.0 | 0.0 | 56.1 | 60.7 | 60.0 | 40.0 | 48.1 | 20.4 | 36.0 | 56.8 | 74.1 |
| 197 | 22.4 | 35.7 | 40.8 | 56.1 | 0.0 | 11.6 | 36.9 | 16.9 | 37.4 | 64.5 | 37.7 | 28.9 | 46.2 |
| 198 | 27.0 | 40.3 | 45.4 | 60.7 | 11.6 | 0.0 | 25.4 | 28.1 | 48.6 | 66.3 | 39.4 | 17.4 | 34.6 |
| 202 | 26.3 | 39.6 | 44.7 | 60.0 | 36.9 | 25.4 | 0.0 | 20.0 | 40.5 | 67.6 | 40.8 | 33.7 | 19.8 |
| 203 | 6.3 | 19.6 | 24.7 | 40.0 | 16.9 | 28.1 | 20.0 | 0.0 | 20.5 | 47.6 | 20.8 | 29.5 | 39.8 |
| 204 | 26.8 | 27.7 | 45.2 | 48.1 | 37.4 | 48.6 | 40.5 | 20.5 | 0.0 | 46.2 | 25.1 | 46.0 | 61.6 |
| 205 | 34.8 | 24.6 | 35.4 | 20.4 | 64.5 | 66.3 | 67.6 | 47.6 | 46.2 | 0.0 | 28.9 | 49.8 | 67.0 |
| 206 | 25.7 | 15.6 | 34.1 | 36.0 | 37.7 | 39.4 | 40.8 | 20.8 | 25.1 | 28.9 | 0.0 | 20.9 | 38.1 |
| 207 | 35.7 | 36.4 | 54.2 | 56.8 | 28.9 | 17.4 | 33.7 | 29.5 | 46.0 | 49.8 | 20.9 | 0.0 | 17.3 |
| 209 | 46.0 | 53.7 | 64.5 | 74.1 | 46.2 | 34.6 | 19.8 | 39.8 | 61.6 | 67.0 | 38.1 | 17.3 | 0.0 |

4.2.1.3 A análise multidimensional dos dados de origem e destino

A análise multidimensional dos dados de origem e destino buscou recuperar a configuração espacial embutida na matriz de relacionamentos. Neste caso, as coordenadas obtidas com a aplicação do CMDS sobre os dados de relacionamentos são apresentadas na tabela 4.3.

Tabela 4.3 - Coordenadas do Sistema Imagem do experimento 1.

| Zona | X(m) | Y(m) |
|------|--------|--------|
| 168 | -6.16 | 2.91 |
| 169 | -0.47 | 14.25 |
| 177 | -12.50 | 20.69 |
| 178 | -1.36 | 35.56 |
| 197 | -17.80 | -15.57 |
| 198 | -8.40 | -24.26 |
| 202 | -6.27 | -23.50 |
| 203 | -9.17 | -3.58 |
| 204 | -10.19 | 7.02 |
| 205 | 20.95 | 34.86 |
| 206 | 13.18 | 4.96 |
| 207 | 16.83 | -18.73 |
| 209 | 21.35 | -34.62 |

As coordenadas calculadas pelo CMDS encontram-se em um sistema coordenado qualquer, onde a origem está localizada no centro da configuração espacial dos pontos. No caso deste trabalho, houve a imposição de que a solução fosse fornecida em duas dimensões (2D), de forma que pudesse ser comparada com a configuração de referência. Por esse motivo, não houve a necessidade de cálculo do *Stress* para a determinação da melhor dimensão de representação.

4.2.1.4 A Regressão Bidimensional de Tobler

A Regressão Bidimensional foi realizada em duas fases: o cálculo das coordenadas do Sistema Imagem Ajustado e a interpolação dos vetores principais de distorção.

Para o cálculo das coordenadas do Sistema Imagem Ajustado foi utilizada a Transformação Linear de Similaridade, e a Transformação Linear Afim. A primeira

apresentou um Erro Quadrático Médio (EQM) de 2774 metros, e a segunda de 1122 metros. Desse modo, a Transformação Afim foi utilizada fornecendo as coordenadas da tabela 4.4.

Tabela 4.4 - Coordenadas do Sistema Imagem Ajustado do experimento 1.

| Zona | N(m) | E(m) |
|------|------------|-----------|
| 168 | 7396310.50 | 345172.26 |
| 169 | 7396935.74 | 346543.56 |
| 177 | 7397983.81 | 344995.34 |
| 178 | 7398637.93 | 347354.47 |
| 197 | 7395397.76 | 342576.38 |
| 198 | 7394293.49 | 343624.31 |
| 202 | 7394256.72 | 343982.05 |
| 203 | 7395940.84 | 344424.30 |
| 204 | 7396813.74 | 344739.62 |
| 205 | 7397570.92 | 350725.25 |
| 206 | 7395592.68 | 348212.40 |
| 207 | 7393580.04 | 347716.72 |
| 209 | 7392136.00 | 347699.29 |

Os vetores principais de distorção são apresentados na figura 4.6.

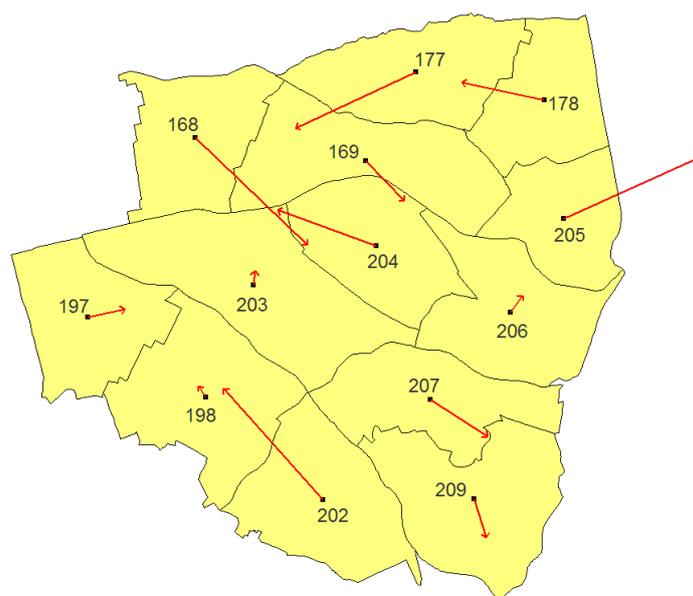


Figura 4.6 - Vetores principais de distorção do experimento 1.

Os vetores principais indicam a magnitude e as direções de distorção do espaço em função da variável de representação. Neste caso, o tempo médio de duração das viagens entre as zonas da amostra.

Posteriormente, os vetores principais de distorção foram interpolados utilizando-se o método de interpolação TPS. Desse modo, as distorções localizadas são então extrapoladas para todos os pontos dentro da região de estudo (figura 4.7).

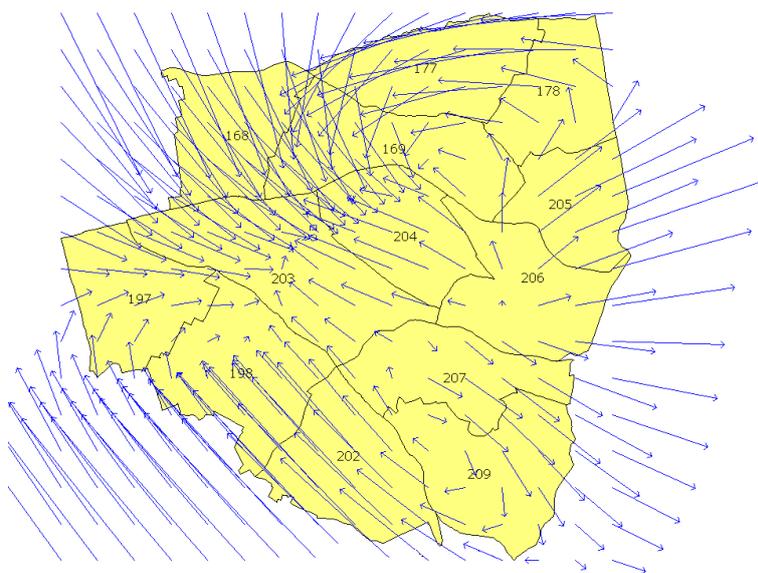


Figura 4.7 - Interpolação do espaço do experimento 1.

Observa-se que já é possível identificar as regiões onde há uma contração do espaço, ou seja, onde os vetores convergem para uma determinada região. Da mesma forma, as regiões onde os vetores divergem, indicando uma distensão no espaço.

4.2.1.5 Reprojetoando o espaço em função dos fluxos

Para facilitar a comparação entre a configuração resultante e a configuração original do espaço, foi utilizada uma grade regular de pontos abrangendo toda a região correspondente às zonas de pesquisa.

Desse modo, a grade regular apresenta-se como um instrumento útil de visualização, ainda mais se utilizada em conjunto com outros indicadores, como os

vetores principais de distorção (figura 4.8).

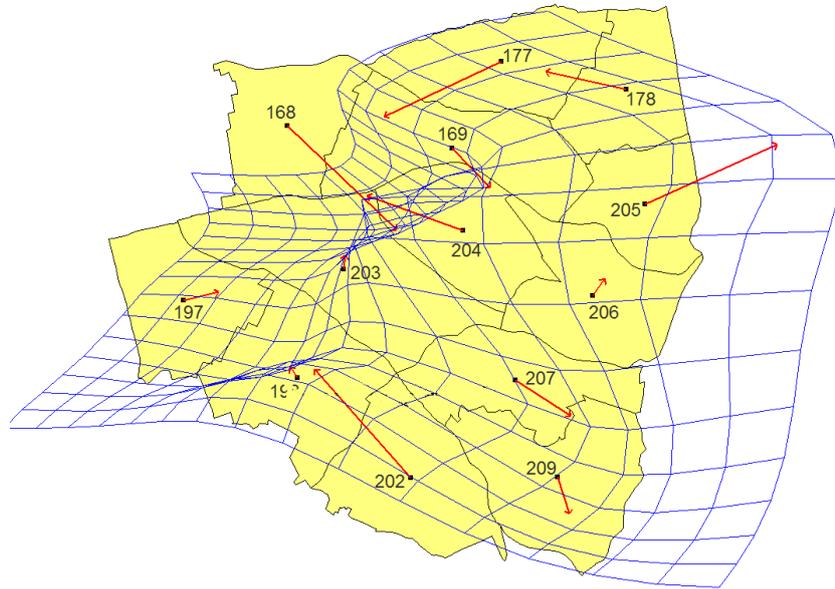


Figura 4.8 - Grade regular interpolada e vetores principais do experimento 1.

Outra possibilidade é calcular, e projetar, as distorções (D) das zonas sobre sua geometria original (O), antes da interpolação (figura 4.9).

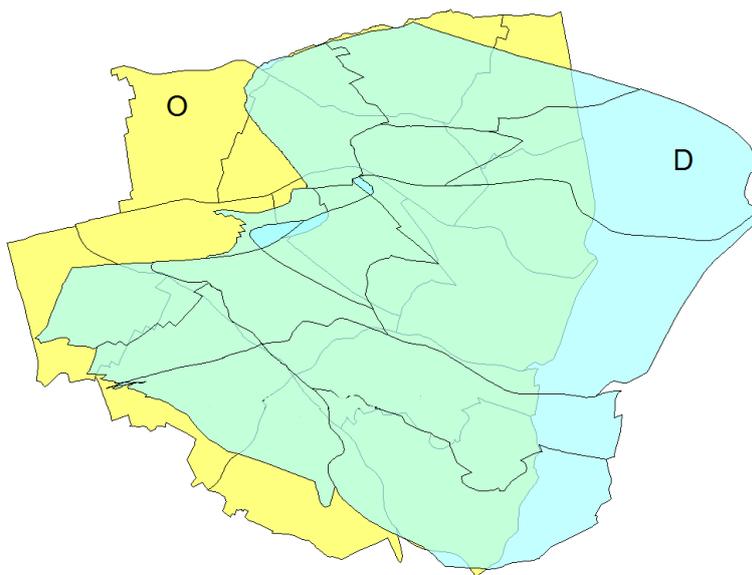


Figura 4.9 - Projeção das distorções sobre o espaço original do experimento 1.

4.2.1.6 Discussão dos resultados

Os resultados obtidos com a aplicação da metodologia representam na verdade um equilíbrio das distorções em função do atributo utilizado para a representação.

Em função dos resultados observados, uma análise é realizada sobre as regiões demarcadas pelas elipses *A*, *B* e *C* (figura 4.10).

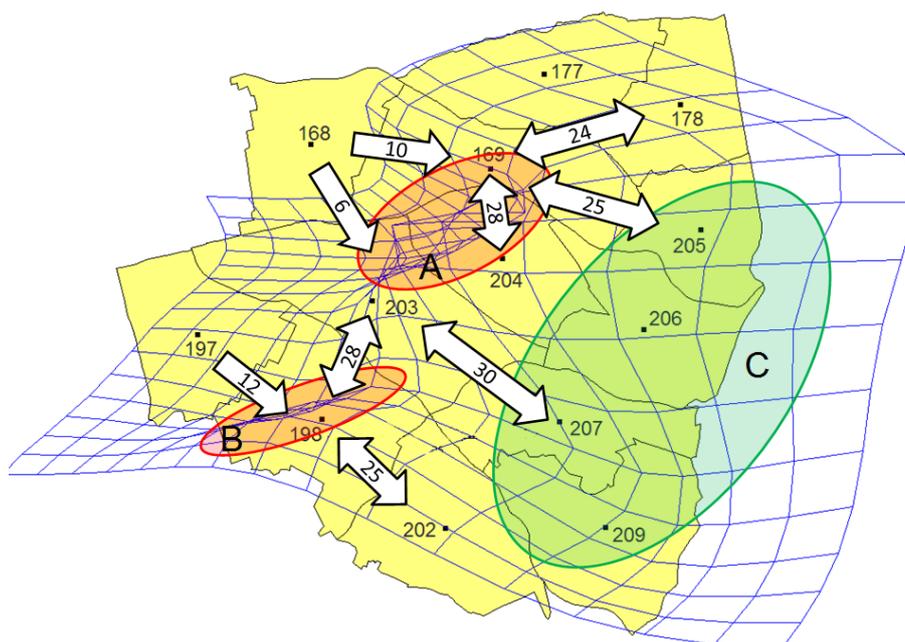


Figura 4.10 - Análise dos resultados do experimento 1.

Duas das regiões demarcadas estão localizadas a oeste da amostra (*A* e *B*), e representam onde a contração do espaço é maior; a outra região (*C*) está localizada a leste da amostra, e representa onde ocorre uma distensão no espaço. Alguns valores aproximados referentes aos tempos de viagem encontram-se indicados no interior das setas representadas na figura 4.10.

Verifica-se que as contrações do espaço observadas nas regiões *A* e *B* são compatíveis com os baixos tempos de deslocamentos observados nas mesmas. Como por exemplo, nos relacionamentos observados entre os pares 168-169, 168-203 e 197-198. Essas regiões de contração podem indicar onde o desloca-

mento é mais rápido, ou onde a rede de transporte é mais eficiente. O que pode ser justificado pela disponibilidade de meios de transporte mais velozes, ou por melhores condições de trafegabilidade. Em uma projeção futura, é possível que essas regiões sejam responsáveis por uma maior demanda pelo surgimento de relacionamentos. Fato que deve ser acompanhado para evitar que a infraestrutura não entre em colapso.

Em contrapartida, observa-se uma distensão na região C , possivelmente relacionada com os maiores tempos de deslocamento encontrados nos pares de relacionamentos 198-202, 203-207, 169-205 e 169-178, indicando que os deslocamentos nessa região é mais demorado. Entre as possíveis causas, estão: uma deficiência na rede de transporte, a existência de problemas localizados, ou pouca disponibilidade de meios de transporte mais velozes.

Desse modo, dependendo do bem, ou serviço, que esteja “trafegando” pela rede, as zonas devem estar preparadas para absorver esses impactos, e a influência dessa dinâmica analisada em conjunto com outras variáveis, pode oferecer novas opções de análise.

Cabe ressaltar, que a rede utilizada neste experimento não corresponde a rede de transporte física real, mas sim uma rede criada segundo um critério pré-definido (centroides adjacentes), o que condiciona a análise dos resultados obtidos.

4.2.2 Experimento 2: Número total de viagens

4.2.2.1 Base de dados

A base de dados do experimento 2 refere-se as mesmas 13 zonas utilizadas no experimento 1 (figura 4.11). No entanto, os relacionamentos são função da variável “número total de viagens”. Cabe ressaltar, que a variável foi calculada considerando todos os modos principais de transporte previstos na variável MO-DOP_ON_A da pesquisa OD.

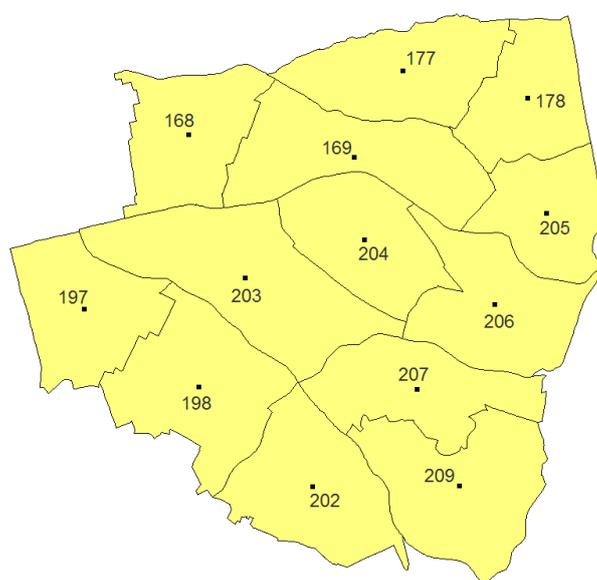


Figura 4.11 - Amostra de estudo do experimento 2.

No caso deste experimento, a variável utilizada representa um atributo de similaridade. Ou seja, quanto maior o seu valor, mais similares são as zonas relacionadas.

Desse modo, será necessária sua transformação em dissimilaridades, onde espera-se que após a aplicação da metodologia, os maiores valores estejam associados aos maiores afastamentos.

A identificação das zonas, juntamente com as coordenadas dos seus respectivos centroides, pode ser observada na tabela 4.5.

Tabela 4.5 - Coordenadas dos centroides das zonas do experimento 2.

| Zona | Nome | N(m) | E(m) |
|------|----------------------|------------|-----------|
| 168 | Ticoatira | 7397843.98 | 343579.13 |
| 169 | Vila Esperança | 7397516.85 | 345987.99 |
| 177 | Ponte Rasa | 7398786.64 | 346699.99 |
| 178 | Águia de Haia | 7398377.92 | 348525.18 |
| 197 | Vila Califórnia | 7395274.51 | 342056.53 |
| 198 | Vila Carrão | 7394132.49 | 343724.85 |
| 202 | Aricanduva | 7392666.29 | 345385.32 |
| 203 | Vila Matilde | 7395733.94 | 344400.62 |
| 204 | Vila Guilhermina | 7396298.11 | 346138.60 |
| 205 | Cidade A.E. Carvalho | 7396692.59 | 348796.06 |
| 206 | Artur Alvin | 7395346.05 | 348037.33 |
| 207 | Cidade Líder | 7394101.29 | 346911.92 |
| 209 | Parque Savoy | 7392679.51 | 347522.44 |

4.2.2.2 Construindo a matriz de fluxos

Neste experimento, não foram utilizados os centroides para a construção da rede de relacionamentos, e sim os próprios relacionamentos (não nulos) presentes na matriz de relacionamentos original (figura 4.12).

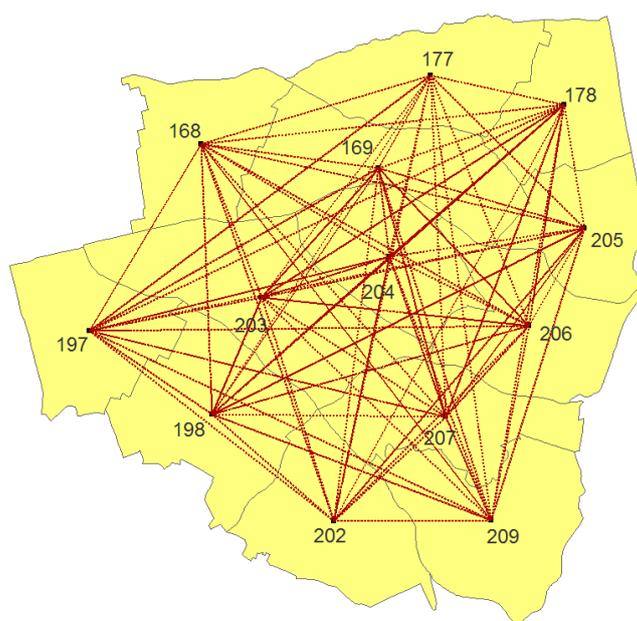


Figura 4.12 - Rede de relacionamentos do experimento 2.

Como consequência, os nós da rede correspondem aos centroides dos polígonos, e as arestas representam os relacionamentos existentes entre os mesmos. Neste caso, os relacionamentos são diretos, e não existe a possibilidade de relações intermediárias.

Desse modo, as seguintes análises e ações foram executadas sobre os dados originais da pesquisa selecionados para este experimento:

- a) A quantidade de viagens é um atributo de similaridade, havendo a necessidade de transformá-lo em dissimilaridade para sua utilização quando da aplicação do CMDS. O que é realizado com a subtração dos valores de uma constante C .
- b) A metodologia não considera os autorrelacionamentos. Desse modo, os mesmos são nulos.
- c) Para a aplicação do escalonamento multidimensional, é necessário que a matriz de relacionamentos seja simétrica. Neste caso, optou-se por considerar como referência os valores contidos na matriz triangular superior. No entanto, a escolha poderia ser baseada em um outro critério.
- d) Os valores nulos de relacionamentos eventualmente encontrados foram substituídos por um valor superior a maior das dissimilaridades, de modo que os valores resultantes são apresentados na tabela 4.6.

Tabela 4.6 - Matriz de quantidade de viagens do experimento 2.

| Zona | 168 | 169 | 177.00 | 178 | 197 | 198 | 202 | 203 | 204 | 205 | 206 | 207 | 209 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 168 | 0 | 87368 | 99958 | 100448 | 101167 | 99521 | 101526 | 99303 | 100847 | 101526 | 101167 | 101526 | 101253 |
| 169 | 87368 | 0 | 94049 | 96534 | 101526 | 100932 | 101526 | 98096 | 96559 | 96786 | 100073 | 100615 | 101526 |
| 177 | 99958 | 94049 | 0 | 93357 | 101068 | 101526 | 101526 | 100499 | 100689 | 101526 | 101526 | 100812 | 101526 |
| 178 | 100448 | 96534 | 93357 | 0 | 101526 | 101526 | 101526 | 101526 | 100162 | 98137 | 99483 | 101526 | 101526 |
| 197 | 101167 | 101526 | 101068 | 101526 | 0 | 95671 | 100070 | 99538 | 101526 | 101230 | 100409 | 100734 | 101275 |
| 198 | 99521 | 100932 | 101526 | 101526 | 95671 | 0 | 98367 | 100129 | 100397 | 101226 | 99560 | 99977 | 99963 |
| 202 | 101526 | 101526 | 101526 | 101526 | 100070 | 98367 | 0 | 100933 | 100768 | 101526 | 100288 | 100301 | 95880 |
| 203 | 99303 | 98096 | 100499 | 101526 | 99538 | 100129 | 100933 | 0 | 96726 | 101199 | 97165 | 95682 | 100363 |
| 204 | 100847 | 96559 | 100689 | 100162 | 101526 | 100397 | 100768 | 96726 | 0 | 101230 | 95975 | 101071 | 100361 |
| 205 | 101526 | 96786 | 101526 | 98137 | 101230 | 101226 | 101526 | 101199 | 101230 | 0 | 100302 | 101526 | 99761 |
| 206 | 101167 | 100073 | 101526 | 99483 | 100409 | 99560 | 100288 | 97165 | 95975 | 100302 | 0 | 94673 | 99367 |
| 207 | 101526 | 100615 | 100812 | 101526 | 100734 | 99977 | 100301 | 95682 | 101071 | 101526 | 94673 | 0 | 90641 |
| 209 | 101253 | 101526 | 101526 | 101526 | 101275 | 99963 | 95880 | 100363 | 100361 | 99761 | 99367 | 90641 | 0 |

4.2.2.3 A análise multidimensional dos dados de origem e destino

A análise multidimensional dos dados de origem e destino busca recuperar a configuração espacial embutida na matriz de relacionamentos. Neste caso, as coordenadas obtidas com a aplicação do CMDS sobre os dados de relacionamentos são apresentadas na tabela 4.7.

Tabela 4.7 - Coordenadas do Sistema Imagem do experimento 2.

| Zona | X(m) | Y(m) |
|------|-----------|-----------|
| 168 | -1316.31 | -29749.94 |
| 169 | 8860.69 | -38913.50 |
| 177 | 159.56 | -27555.41 |
| 178 | 402.55 | -24538.15 |
| 197 | -41652.65 | 10958.35 |
| 198 | -39131.93 | 13019.75 |
| 202 | -20933.08 | 19863.52 |
| 203 | 19015.74 | 8291.48 |
| 204 | 18244.01 | -1509.59 |
| 205 | -3421.41 | -8769.63 |
| 206 | 22057.34 | 16550.92 |
| 207 | 26459.78 | 31217.41 |
| 209 | 11255.71 | 31134.79 |

As coordenadas calculadas pelo CMDS encontram-se em um sistema coordenado qualquer, e são fornecidas em duas dimensões (2D), como imposição para a comparação com a configuração de referência. Desse modo, não houve a necessidade do cálculo do *Stress*.

4.2.2.4 A Regressão Bidimensional de Tobler

Para a execução da primeira fase da Regressão Bidimensional (cálculo das coordenadas do Sistema Imagem Ajustado) foi utilizada uma Transformação Linear de Similaridade, e uma Transformação Linear Afim. A primeira apresentou um Erro Quadrático Médio (EQM) de 1641 metros, e a segunda de 1566 metros. Desse modo, a Transformação Afim foi utilizada fornecendo as coordenadas da tabela 4.8.

Tabela 4.8 - Coordenadas do Sistema Imagem Ajustado do experimento 2.

| Zona | N(m) | E(m) |
|------|------------|-----------|
| 168 | 7398099.25 | 346006.65 |
| 169 | 7398954.80 | 346596.07 |
| 177 | 7397949.28 | 346080.64 |
| 178 | 7397718.01 | 346084.16 |
| 197 | 7394366.92 | 343656.17 |
| 198 | 7394241.97 | 343788.04 |
| 202 | 7393965.12 | 344766.30 |
| 203 | 7395425.87 | 347000.45 |
| 204 | 7396177.34 | 346989.99 |
| 205 | 7396437.91 | 345822.49 |
| 206 | 7394826.17 | 347140.75 |
| 207 | 7393747.25 | 347334.96 |
| 209 | 7393540.27 | 346499.30 |

De posse das coordenadas do Sistema de Referência Z , e do Sistema Imagem Ajustado W' , é possível identificar os vetores principais de distorção do espaço em função da variável (figura 4.13).

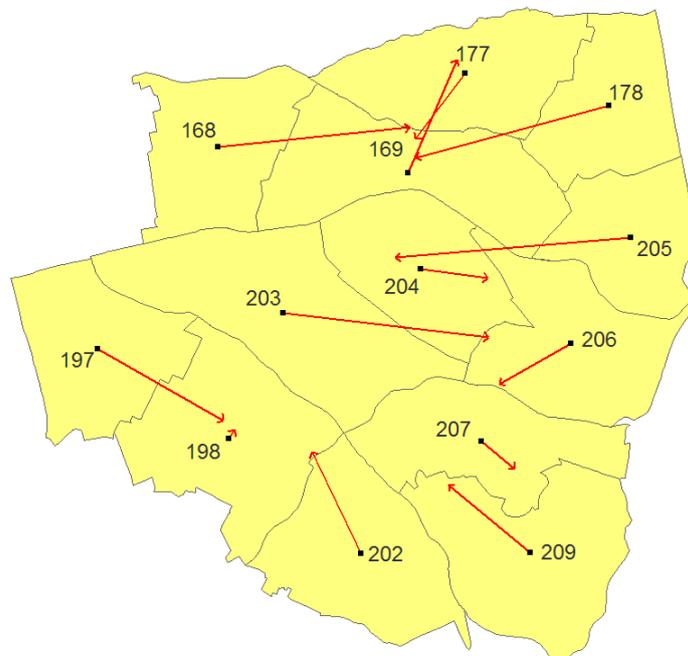


Figura 4.13 - Vetores principais de distorção do experimento 2.

Aplicando a interpolação TPS, os vetores são então utilizados para generalizar as distorções obtidas para todo o espaço em estudo (figura 4.14).

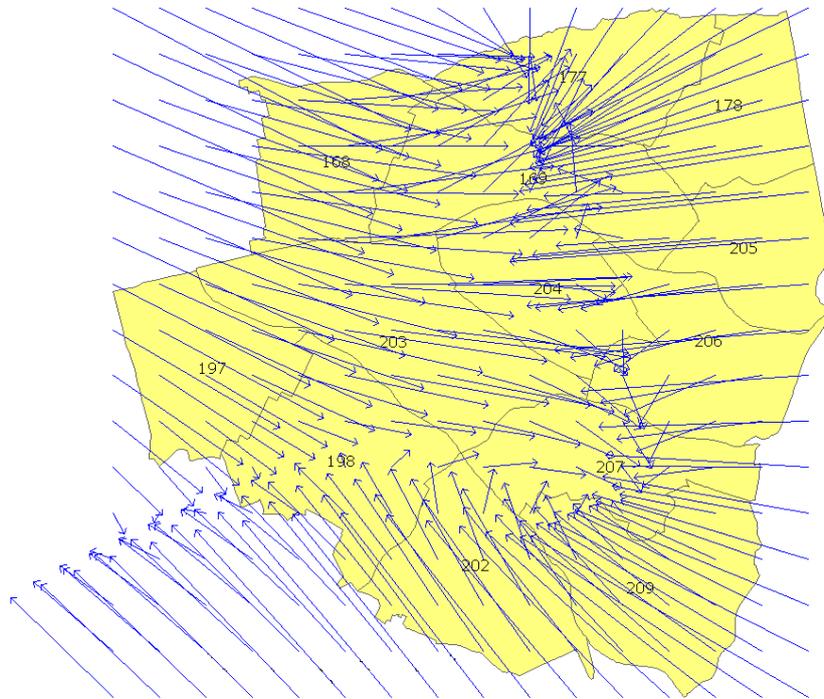


Figura 4.14 - Interpolação do espaço do experimento 2.

Observa-se em uma primeira análise, que é possível identificar as regiões para onde os vetores convergem, ou seja, onde existe uma contração no espaço.

Da mesma forma, são identificadas regiões onde os vetores divergem, o que indica onde ocorre uma distensão no espaço.

4.2.2.5 Reprojetoando o espaço em função dos fluxos

Para facilitar a comparação entre a configuração resultante e a configuração original do espaço, é utilizado o mesmo mecanismo do experimento anterior, ou seja, uma grade regular em conjunto com os vetores principais de distorção (figura 4.15).

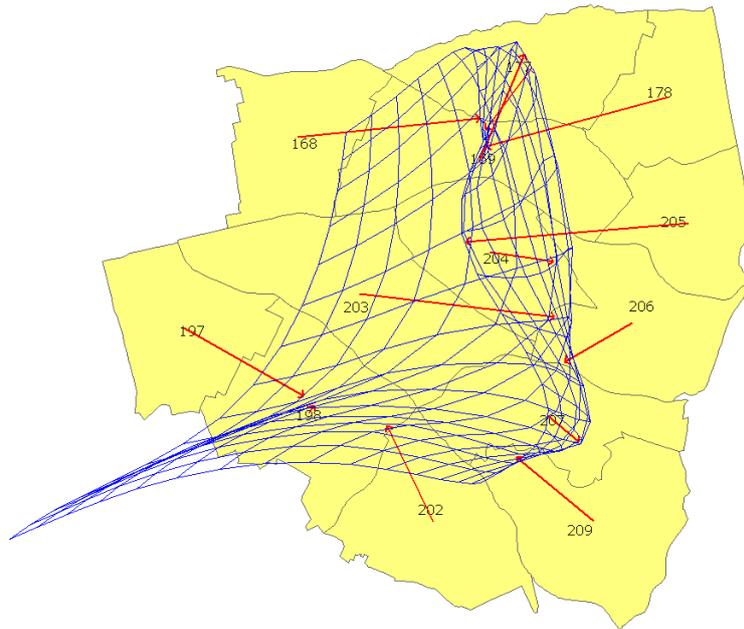


Figura 4.15 - Grade regular interpolada e vetores principais do experimento 2.

Outra possibilidade é projetar as distorções (D) sobre a origem (O) (figura 4.16).

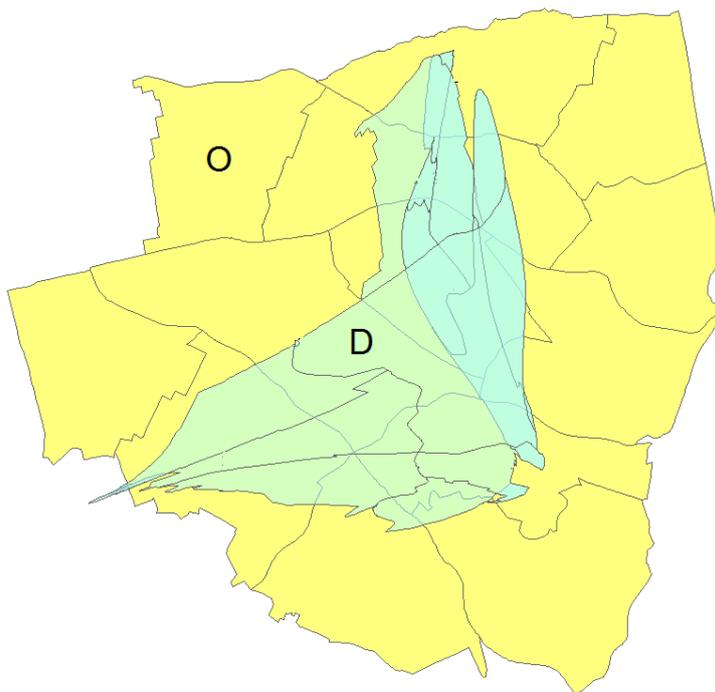


Figura 4.16 - Projeção das distorções sobre o espaço original do experimento 2.

4.2.2.6 Discussão dos resultados

Uma análise é realizada sobre as regiões demarcadas pelas elipses *A* e *B* (figura 4.17).

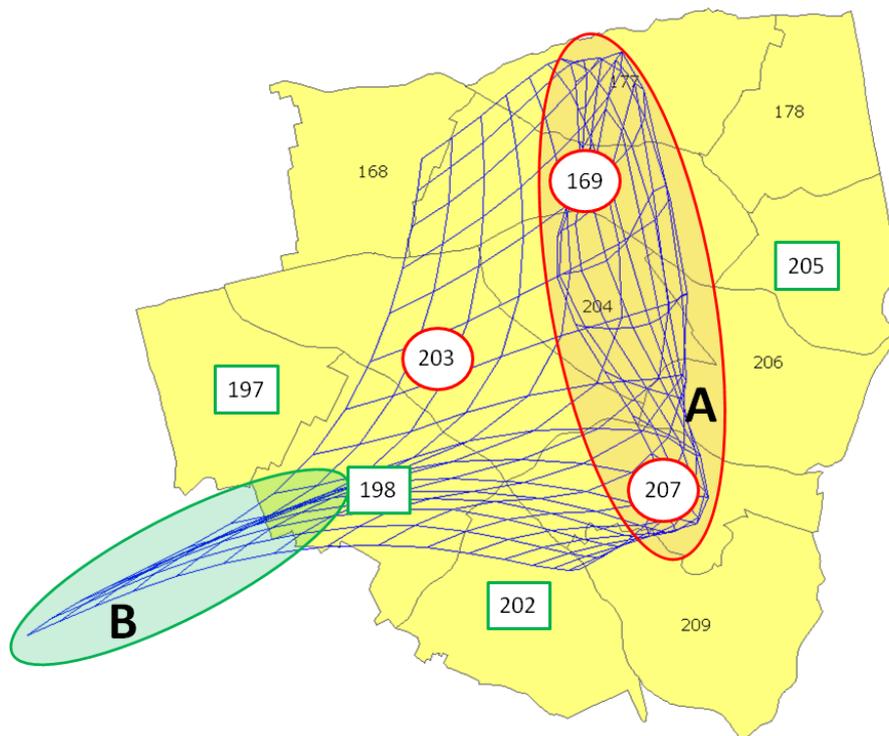


Figura 4.17 - Análise dos resultados do experimento 2.

A região demarcada *A* está localizada a leste da amostra, alonga-se no sentido norte-sul, e representa onde a contração no espaço é maior; já a região *B* está localizada a oeste da amostra, e representa onde a distensão no espaço é maior.

Os círculos identificam as zonas participantes dos relacionamentos com maiores quantidades de viagens envolvidas, e por retângulos as zonas com as menores quantidades. Desse modo, espera-se que as zonas envolvidas nos relacionamentos com maiores quantidades de viagens estejam nas regiões de maior contração do espaço.

Verifica-se ainda, que as contrações do espaço observadas na região *A* são compatíveis com as altas quantidades de viagens relacionadas às zonas 169, 203 e 207. Isso indica que essa região possui um fluxo maior de viagens, configurando-

se um núcleo de atração. Sendo assim, deve existir a preocupação de que no futuro, não haja sobrecarga do sistema utilizado para esse deslocamento.

Em contrapartida, observa-se uma distensão na região *B*, possivelmente relacionada com as baixas quantidades de viagens relacionadas às zonas 197, 198 e 202. Isso indica, que há uma quantidade menor de viagens nessa região. Outro fato a ser considerado, é que apesar de possuir um baixa quantidade de viagens, a zona 205 não é suficiente para se contrapor a contração da região *A*.

Outro tipo de análise que pode ser realizada é a sobreposição dos resultados obtidos para ambas as variáveis utilizadas (figura 4.18).

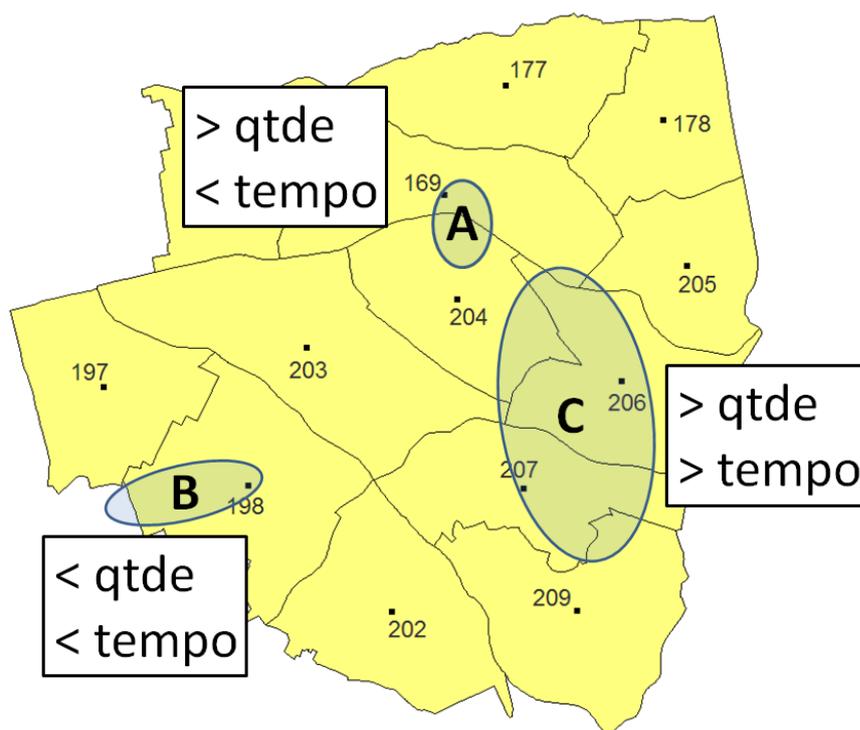


Figura 4.18 - Sobreposição dos resultados dos experimentos 1 e 2.

Na região *A* da figura 4.18, há uma grande quantidade de fluxo de viagens, mas como o deslocamento é feito em pouco tempo, pressupõe-se que a infraestrutura seja eficiente, e possivelmente não exista problemas de tráfego na rede.

Na região *B*, a quantidade de viagens é pequena, e as mesmas são realizadas em pouco tempo. Desse modo, não há grandes preocupações aparentes com

essa região.

A região C é a mais crítica da amostra, e a qual deve ser dada uma atenção especial. Existe uma grande quantidade de fluxos de viagens, que por sua vez são demoradas. Talvez esses dados estejam diretamente relacionados. Quanto mais viagens, mais o sistema fica sobrecarregado, e conseqüentemente mais lento. Existe ainda a possibilidade de que os meios de transporte mais velozes não sejam acessíveis nas quantidades necessárias.

Observa-se, que as variáveis presentes nos dados de origem e destino podem ser estudadas de forma isolada, mas também de forma conjunta, o que acaba por agregar valor no momento de sua análise.

5 CONCLUSÕES E PERSPECTIVAS

Esta dissertação teve como objetivo principal utilizar algumas das ideias disponíveis na literatura para propor uma nova leitura dos Espaços de Fluxos na análise espacial, apresentando uma forma de estudar, e representar, a dinâmica dos relacionamentos espaciais. Com esse intuito, foi conduzida uma prova de conceito aplicando uma metodologia baseada na técnica da Regressão Bidimensional a um conjunto de dados produzido pela Companhia do Metropolitano de São Paulo, considerando duas variáveis selecionadas da Pesquisa OD/2007.

Durante a aplicação da metodologia apresentada, algumas limitações foram observadas, principalmente na fase de construção das matrizes de relacionamentos, onde há uma grande dificuldade de preencher os dados quando se trabalha com grandes bases. A solução adotada para contornar esse problema foi a análise de pequenas bases de dados por vez.

Os resultados dos experimentos, mesmo que conduzidos com algum grau de simplificação, permitem identificar algumas regiões críticas dentro das redes de fluxos estudadas. Desse modo, a metodologia utilizada apresenta-se como uma ferramenta com potencial de utilização no apoio à decisão, principalmente na implantação de políticas públicas que envolvam a análise de fluxos.

Os experimentos conduzidos indicam que as transformações geométricas geradas pela Regressão Bidimensional proposta por Tobler, em conjunto com outras técnicas de representação, oferecem uma nova visão do espaço em estudo, redesenhando o território em função dos fluxos em suas redes. Além do que, têm as seguintes possibilidades de aplicação:

- realização de projeções sobre o território, analisando possíveis causas e impactos da reconfiguração do espaço geográfico em função dos fluxos em suas redes;
- avaliação por intermédio de modelos de simulação, de projetos de expansão e reestruturação da redes, principalmente as de transportes;
- avaliação da qualidade da infraestrutura de diferentes tipos de redes, identificando possíveis pontos críticos que possam interferir nos fluxos, ou na construção dos relacionamentos;

- utilização como subsídio para a implantação de políticas públicas regionais;
- auxílio da análise temporal dos processos espaciais em função da reconfiguração dos fluxos ao longo do tempo;
- identificação tendências regionais em função das características da variável de análise.

Cabe ressaltar que algumas propostas podem ser futuramente agregadas ao *framework* ETC com o objetivo de prover novas funcionalidades, ou otimizar as já existentes. A saber:

- utilização de redes multimodais, que podem ser interligadas e analisadas a partir de uma mesma variável;
- estudo de novos métodos de interpolação e de transformações não lineares;
- disponibilização de um serviço de processamento pela rede mundial de computadores (WPS), possibilitando a utilização da metodologia para a produção dos resultados remotamente;
- desenvolvimento de um *plugin* para o aplicativo TerraView, de forma a otimizar a entrada de parâmetros por parte dos usuários.

Ao fim do presente estudo, acredita-se que o mesmo tenha atingido seu objetivo ao agregar valor à discussão sobre a análise dos fluxos em um espaço regional.

REFERÊNCIAS BIBLIOGRÁFICAS

AHMED, N.; MILLER, H. J. Time-space transformations of geographic space for exploring, analysing and visualizing transportation systems. **Journal of Transport Geography**, n. 206, p. 2–17, 2007. 14, 15, 16, 18, 22

AXHAUSENA, K. W.; DOLCIB, C.; FROUMLHLICHA, P.; SCHERERA, M.; CAROSIO, A. Constructing time-scaled maps: Switzerland from 1950 to 2000. **Transport Reviews**, v. 28, n. 3, p. 391–413, 2008. 22

BOOKSTEIN, F. L. Principal warps: Thin plate splines and the decomposition of deformations. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 6, p. 567–585, 1989. 21

BROWNSTEIN, J.; FREIFELD, C.; MADOFF, L. **Influenza A (H1N1) virus, 2009 - online monitoring**. 2009. Disponível em: <<http://www.healthmap.org/swineflu/>>. Acesso em: 01 março 2011. 9

CÂMARA, G.; MONTEIRO, A. M. V.; MEDEIROS, J. S. Representações computacionais do espaço: Fundamentos epistemológicos da ciência da geoinformação. **Geografia**, v. 28, n. 1, p. 83–96, 2003. 7, 9

CÂMARA, G.; VINHAS, L.; QUEIROZ, G.; FERREIRA, K.; MONTEIRO, A. M. V.; CARVALHO, M. T.; CARTAXO, R.; CASANOVA, M. A.; FREITAS, U. M. Terralib: An open-source gis library for large-scale environmental and socio-economic applications. In: HALL, G. B.; LEAHY, M. (Ed.). **Open source approaches to spatial data handling**. Berlin: Springer-Verlag, 2008. 27

CASANOVA, M. A.; CÂMARA, G.; JR, C. A. D.; VINHAS, L.; QUEIROZ, G. **Bancos de Dados Geográficos**. Curitiba: MundoGeo, 2005. 27

CAUVIN, C. Cognitive and cartographic representations: Towards a comprehensive approach. **Cybergeo**, n. 206, 2002. 22

_____. A systemic approach to transport accessibility: A methodology developed in strasbourg: 1982-2002. **Cartography and Geographic Information Science**, n. 311, 2005. 19, 22

CLARK, W. A. V. A revealed preference analysis of intraurban migration choices. In: GOLLEDGE, R. G.; RAYNER, J. N. (Ed.). **Proximity and preference**:

problems in the multidimensional analysis of large data sets. Estados Unidos da América: University of Minnesota Press, 1982. [12](#)

DIAS M. H.; CAUVIN, C. A. M. F. Comparação de configurações cartográficas através da regressão bidimensional. **Finisterra**, v. 35, n. 69, p. 95–107, 2000. [16](#)

DIJKSTRA, E. W. A note on two problems in connexion with graphs. **Numerische Mathematik**, v. 1, p. 269–271, 1959. [37](#)

DRYDEN, I.; MARDIA, K. V. **Statistical shape analysis**. Londres: John Wiley and Sons, 1998. ISBN 0-471-95816-6. [20](#)

DUNN-RANKIN, P.; KNEZEK, G. A.; WALLACE, S.; ZHANG, S. **Scaling methods**. Londres: Lawrence Erlbaum Associates, 2004. ISBN 0-8058-1802-2. [31](#)

FLOYD, R. W. Algorithm 97: Shortest path. **Communications of the ACM**, v. 5, n. 6, p. 345, 1959. [37](#)

FORER P. C., K. H. The effects of spatial and locational cueing on the analysis of aggregate cognitive mapping data. **Environment and Planning A**, v. 13, n. 4, p. 497–509, 1981. [22](#)

FRIEDMAN, A.; KOHLER, B. Bidimensional regression: Assessing the configural similarity and accuracy of cognitive maps and other two-dimensional data sets. **Psychological Methods**, v. 8, n. 4, p. 468–491, 2003. [16](#), [19](#), [22](#), [23](#)

GHILANI, C. D.; WOLF, P. R. **Adjustment Computations: Spatial data analysis**. Estados Unidos da América: John Wiley and Sons, 2006. ISBN 0-47169728-1. [73](#)

GOLLEDGE, R. G.; HUBERT, L. J. Some comments on non-euclidean mental maps. **Environment and Planning**, v. 14, p. 107–118, 1982. [8](#)

GOODCHILD, M. F. Giscience, geography, form, and process. **Annals of the Association of American Geographers**, v. 94, n. 4, p. 709–714, 2004. [7](#)

GOULD, P.; WHITE, R. **Mental Maps**. Estados Unidos da América: ROUTLEDGE, 1986. ISBN 0-41508482-2. [2](#), [11](#)

GRAEF, J.; SPENCE, I. Using distance information in the design of large multidimensional scaling experiments. **Psychological Bulletin**, v. 86, n. 1, p. 60–66, 1979. 35

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Multivariate data analysis**. Estados Unidos da América: PRENTICE HALL, 1998. ISBN 0-13894858-5. 23

HIROSHI, I. **Space-time accessibility measures for evaluating mobility related social exclusion of the elderly**. Canadá: [s.n.], 2007. Disponível em: <<http://www.tc.gc.ca/eng/policy/transed2007-pages-1160-1848.htm>>. Acesso em: 01 março 2011. 14

INPE. **TerraView 3.3.1**. São José dos Campos - SP: INPE, 2010. Disponível em: <www.dpi.inpe.br/terraview>. Acesso em: 01 março 2011. 35

JONGE, E. **Weighted flow diagrams for statistical output**. 1992. Disponível em: <http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/documents/S10P2_WEIGHTED_FLOW_DIAGRAMS_FOR_STATISTICAL_OUTPUT_DE_J.pdf>. Acesso em: 01 março 2011. 13

KARE, S.; SAMAL, A.; MARX, D. Using bidimensional regression to assess face similarity. **Machine Vision and Applications**, v. 21, p. 261–274, 2010. 19, 22

KITCHIN, R. M.; FOTHERINGHAM, A. The effects of spatial and locational cueing on the analysis of aggregate cognitive mapping data. **Environment and Planning A**, v. 30, p. 2245–253, 1998. 19

KRAAK, M.; MACEACHREN, A. M. Geovisualization and giscience. **Cartography and Geographic Information Science**, v. 32, n. 2, p. 67–68, 2005. 10

KRUSKAL, J. B.; WISH, M. (Ed.). **Multidimensional Scaling**. Estados Unidos da América: SAGE University, 1978. 23, 24, 25

L'HOSTIS, A. The shriveled usa: Representing time-space in the context of metropolitanization and the development of high speeds. **Journal of Transport Geography**, v. 17, p. 433–439, 2009. 14, 15, 22

LIM, J.; YANG, M. A direct method for modeling non-grid motion with thin plate spline. **Computer Vision and Pattern Recognition**, v. 5, p. 1196–1202, 2005. 21

MACEACHREN, A. M. **Some Truth With Maps**: A primer on symbolization and design. Washington DC.: American Geographers, 1994. ISBN 0-89291-214-6. 7

MALHOTRA, N. K.; JAIN, A. K.; PINSON, C. The robustness of mds configurations in the case of incomplete data. **Journal of Marketing Research**, v. 25, n. 1, p. 95–102, 1988. 34

MARCHAND, B. Deformation of a transportation surface. **Annals of the Association of American Geographers**, v. 63, n. 4, p. 507–521, 1978. 22

MAYHEW, L. D. Automated isochrones and the locations of emergency medical services in cities. **The Professional Geographer**, v. 33, n. 4, p. 423–428, 1981. 22

METROPOLITANOS, S. de T. (Ed.). **Pesquisa Origem e Destino 2007**: Região metropolitana de são paulo. São Paulo, Brasil: Companhia do Metropolitano de São Paulo, 2008. 83 p. Disponível em:
<http://www.metro.sp.gov.br/empresa/pesquisas/od_2007/teod.asp>. Acesso em: 01 março 2011. 4, 39

MILLER, H. J. Geographic representation in spatial analysis. **Journal of geographical systems**, v. 2, n. 1, p. 55–60, 2000. 1

_____. Toblers first law and spatial analysis. **Annals of the Association of American Geographers**, v. 94, n. 2, p. 284–289, 2004. 8, 9

MULLER, J. C. The mapping of travel time in edmonton, alberta. **Canadian Geographer**, v. 22, n. 3, p. 195–210, 1978. 14, 15

PASKINS, J. Investigating the effects of a car culture on a childs spatial skills. In: INTERNATIONAL CONFERENCE ON WALKING IN THE 21ST CENTURY, 6., Sept. 22-23 2005, Zurich. **Proceedings...** Zurich, Switzerland, 2005. p. 22–23. 22

PRAGER, S. D. Complex networks for representation and analysis of dynamic geographies. In: HORNSBY, K. S.; YUAN, M. (Ed.). **Understanding dynamics of geographic domains**. Estados Unidos da América: CRC Press, 2008. 10

SANTOS, M. **A natureza do espaço**: Técnica e tempo, razão e emoção. São Paulo: EdUSP, 2006. 2

SPENCE, I.; DOMONEY, D. W. Single subject incomplete designs for nonmetric multidimensional scaling. **Psychometrika**, v. 39, n. 4, p. 469–490, 1974. 35

SPIEKERMANN, K.; WEGENER, M. The shrinking continent: new time - space maps of europe. **Environment and Planning B: Planning and Design**, v. 21, n. 6, p. 653–673, 1994. 15

STURROCK, K.; JORGE, R. A multidimensional scaling stress evaluation table. **Field Methods**, v. 12, n. 1, p. 49–60, 2000. 25

SUI, D. Z. Tobler's first law of geography: A big idea for a small world? **Annals of the Association of American Geographers**, v. 94, n. 2, p. 269–277, 2004. 8, 9

SYMINGTON, A.; CHARLTON, M. E.; BRUNSDON, C. F. Using bidimensional regression to explore map lineage. **Computers, Environment and Urban Systems**, v. 26, n. 2-3, p. 201–218, 2002. 18, 19

TOBLER, W. R. A computer movie simulating urban growth in the detroit region. **Economy Geography**, v. 46, n. 2, p. 234–240, 1970. 8

_____. A transformational view of cartography. **The American Cartographer**, v. 6, n. 2, p. 101–106, 1979. 14

_____. The geographic movement of wealth in the united states. **MappedMonde**, v. 3, p. 37–40, 1986. 12

_____. Experiments in migration mapping by computer. **The American Cartographer**, v. 14, n. 2, p. 155–163, 1987. 14, 35

_____. Global spatial analysis. In: **Three presentations on geographical analysis and modeling**. Santa Barbara: National Center for geographic Information and Analysis, 1993b. (Technical report 93–1). 8

_____. Non-isotropic geographic modeling. In: **Three presentations on geographical analysis and modeling**. Santa Barbara: National Center for geographic Information and Analysis, 1993c. (Technical report 93–1). 8, 14, 15

_____. Bidimensional regression. **Geographical Analysis**, v. 26, n. 3, p. 187–212, 1994. 3, 16, 17, 18, 19

_____. On the first law of geography: A reply. **Annals of the Association of American Geographers**, v. 94, n. 2, p. 304–310, 2004. 1, 8, 9

TSOGO, L.; MASSON, M. H.; BARDOT, A. A metric multidimensional scaling method for large objects sets and its monte carlo evaluation. **IEEE International Conference on Systems Mans and Cybernetics**, v. 4, p. 3807–3812, 1997. 35

WALMSLEY, D. J. Public information flows in rural austrália. **Environment and Planning A**, v. 15, p. 255–263, 1983. 12, 13, 14

WORBOYS, M.; MASON, K.; LINGHAM, J. **Computational techniques for non-Euclidean planar spatial data applied to migrant flows**. Reino Unido: Keele University, 1997. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.4042&rep=rep1&type=pdf>>. Acesso em: 01 março 2011. 12

YOUNG, F. W.; NULL, C. H.; SARLE, W. S.; HOFFMAN, D. L. Interactively ordering the similarities among a large set of stimuli. In: GOLLEDGE, R. G.; RAYNER, J. N. (Ed.). **Proximity and preference: problems in the multidimensional analysis of large data sets**. Estados Unidos da América: University of Minnesota Press, 1982. 35

APÊNDICE A - TRANSFORMAÇÕES GEOMÉTRICAS

Neste apêndice são abordadas as Transformações de Similaridade, e Afim, realizadas entre dois sistemas de coordenadas bidimensionais: $Z(x, y)$ e $W(u, v)$. Caso haja necessidade, é possível encontrar maiores informações sobre o assunto em [GHILANI e WOLF \(2006\)](#).

A.1 A Transformação de Similaridade

Segundo ([GHILANI; WOLF, 2006](#)), a determinação das equações de transformação de similaridade consiste na definição dos seguintes parâmetros:

- T_x : um parâmetro de translação do eixo x .
- T_y : um parâmetro de translação do eixo y .
- θ : um parâmetro de rotação do sistema de eixos.
- S : um parâmetro de distorção em escala para ambos os eixos.

As equações de distorção em escala podem ser definidas como:

$$\begin{aligned}x' &= S.x \\ y' &= S.y\end{aligned}\tag{A.1}$$

Assim como as equações de translação:

$$\begin{aligned}u &= u' + T_y \\ v &= v' + T_x\end{aligned}\tag{A.2}$$

E ainda as equações de rotação:

$$\begin{aligned}u' &= x'.\cos\theta - y'.\sen\theta \\ v' &= x'.\cos\theta + y'.\cos\theta\end{aligned}\tag{A.3}$$

Substituindo a equação [A.3](#) em [A.2](#), tem-se:

$$\begin{aligned}u &= x'.\cos\theta - y'.\sen\theta + T_x \\v &= x'.\sen\theta + y'.\cos\theta + T_y\end{aligned}\tag{A.4}$$

Substituindo a equação A.1 em A.4, tem-se:

$$\begin{aligned}u &= (S.\cos\theta).x - (S.\sen\theta).y + T_x \\v &= (S.\sen\theta).x + (S.\cos\theta).y + T_y\end{aligned}\tag{A.5}$$

Substituindo os termos constantes por α_1 , α_2 , β_1 e β_2 , tem-se:

$$\begin{aligned}u &= \beta_1.x - \beta_2.y + \alpha_1 \\v &= \beta_2.x + \beta_1.y + \alpha_2\end{aligned}\tag{A.6}$$

A substituição dos valores de coordenadas dos pontos homólogos possibilita a determinação das constantes α_1 , α_2 , β_1 e β_2 .

A.2 A Transformação Afim

Enquanto que na transformação de similaridade são considerados quatro parâmetros, na transformação afim são acrescentados mais dois:

- T_x : um parâmetro de translação do eixo x .
- T_y : um parâmetro de translação do eixo y .
- θ : um parâmetro de rotação do sistema de eixos.
- φ : um parâmetro de cisalhamento do sistema de eixos.
- S_x : um parâmetro de distorção em escala no eixo x .
- S_y : um parâmetro de distorção em escala no eixo y .

No caso da Transformação Afim, as equações finais de transformação são apresentadas a seguir:

$$\begin{aligned} u &= \beta_1 \cdot x + \beta_2 \cdot y + \alpha_1 \\ v &= \beta_3 \cdot x + \beta_4 \cdot y + \alpha_2 \end{aligned} \tag{A.7}$$

Dado um conjunto de n pontos, é possível montar os seguintes sistemas de equações:

Transformação de Similaridade:

$$\begin{aligned} u_1 &= \beta_1 x_1 - \beta_2 y_1 + \alpha_1 \\ v_1 &= \beta_2 x_1 + \beta_1 y_1 + \alpha_2 \\ u_2 &= \beta_1 x_2 - \beta_2 y_2 + \alpha_1 \\ v_2 &= \beta_2 x_2 + \beta_1 y_2 + \alpha_2 \\ &\quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ u_n &= \beta_1 x_n - \beta_2 y_n + \alpha_1 \\ v_n &= \beta_2 x_n + \beta_1 y_n + \alpha_2 \end{aligned} \tag{A.8}$$

Transformação afim:

$$\begin{aligned} u_1 &= \beta_1 x_1 + \beta_2 y_1 + \alpha_1 \\ v_1 &= \beta_3 x_1 + \beta_4 y_1 + \alpha_2 \\ u_2 &= \beta_1 x_2 + \beta_2 y_2 + \alpha_1 \\ v_2 &= \beta_3 x_2 + \beta_4 y_2 + \alpha_2 \\ &\quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ u_n &= \beta_1 x_n + \beta_2 y_n + \alpha_1 \\ v_n &= \beta_3 x_n + \beta_4 y_n + \alpha_2 \end{aligned} \tag{A.9}$$

Na solução dos sistemas de equações A.9 e A.10, é desejável que se tenha um mínimo de três pontos na transformação de similaridade, e quatro pontos na transformação afim. No entanto, uma maior quantidade de pontos implica em uma melhor determinação dos parâmetros, que podem ser ajustados utilizando o Método dos Mínimos Quadrados (MMQ). Este método é utilizado para o cálculo da melhor estimativa de solução dos sistemas apresentados, tornando mínima a soma dos quadrados dos resíduos das observações.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o International Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.