

Paulo Fernando Braga Carvalho

SUBSISTEMAS E TIPOLOGIAS DA MIGRAÇÃO POPULACIONAL NO BRASIL A PARTIR DE DADOS CENSITÁRIOS: uma proposta metodológica

Tese apresentada ao Programa de Pós-Graduação em Geografia - Tratamento da Informação Espacial da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do título de doutor em Geografia.

Área de concentração: Análise Espacial

Orientador: Prof. Dr. José Irineu Rangel Rigotti

Co-Orientador: Prof. Dr. Leônidas Conceição Barroso

**Belo Horizonte
2009**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

FICHA CATALOGRÁFICA

Elaborada pela Biblioteca da Pontifícia Universidade Católica de Minas Gerais

C331s Carvalho Paulo Fernando Braga
Subsistemas e tipologias da migração populacional no Brasil a partir de dados censitários: uma proposta metodológica / Paulo Fernando Braga Carvalho. Belo Horizonte, 2009.
168f. : il. + 1 CD-ROM

Orientador: José Irineu Rangel Rigotti

Co-Orientador: Leônidas Conceição Barroso

Tese (Doutorado) – Pontifícia Universidade Católica de Minas

Gerais.

Programa de Pós-Graduação em Tratamento da Informação

Espacial.

1. Migração - Brasil. 2. Métodos estatísticos. 3. Análise fatorial. 4. Métodos interativos (Matemática). 5. Análise espacial. I. Rigotti, José Irineu Rangel. II. Barroso, Leônidas Conceição. III. Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Tratamento da Informação Espacial. III. Título.

CDU: 325(81)

FOLHA DE ASSINATURAS

Paulo Fernando Braga Carvalho

SUBSISTEMAS E TIPOLOGIAS DA MIGRAÇÃO POPULACIONAL NO BRASIL A PARTIR DE DADOS CENSITÁRIOS: uma proposta metodológica

Programa de Pós-Graduação em Geografia - Tratamento da Informação Espacial da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do título de doutor em Geografia, Belo Horizonte, 2009.

Prof. Dr. José Irineu Rangel Rigotti (Orientador) - PUC-Minas

Prof. Dr. Leônidas Conceição Barroso (Co-orientador) - PUC-Minas

Prof. Dr. José Flávio Morais Castro – PUC-Minas

Prof. Dr. André Junqueira Caetano – PUC-Minas
(Ciências Sociais)

Prof. Dr. Edson Paulo Domingues – UFMG
(CEDEPLAR Economia)

Prof. Dr. José Alberto Magno de Carvalho – UFMG
(CEDEPLAR Demografia)

DEDICATÓRIA

Dedico este trabalho às minhas filhas, Marina e Gabriela, pelos momentos que precisei ficar ausente.

AGRADECIMENTOS

À minha esposa Nannda, pelo apoio e pelas concessões necessárias que esta caminhada exigiu.

Às minhas filhas, Marina e Gabriela, que, com suas brincadeiras e o carinho incondicional, me ajudaram a suportar os momentos mais difíceis.

Agradeço aos meus pais por terem me incentivado e sempre valorizado a busca pelo conhecimento.

Ao Professor Leônidas, responsável por minha inserção no mundo da Informação Espacial, que nos momentos de maior ansiedade consegue transmitir o equilíbrio necessário.

Ao Professor Irineu, pela presença segura nas orientações e por acreditar em meu potencial.

Aos alunos Jonas Antonio Vieira Junior e Maria Natália Valério Bittencourt que, voluntariamente, prestaram inestimável apoio na produção cartográfica.

À Pontifícia Universidade Católica de Minas Gerais que, por intermédio do Fundo de Incentivo à Pesquisa-FIP 2008/2552-S1, do Programa Permanente de Capacitação Docente-PPCD e da concessão de bolsa investiu em minha formação.

À Maria Carmem, bibliotecária da PUC Contagem, pelas orientações relacionadas à normalização.

À amiga Ana Márcia que, por estar vivenciando o mesmo momento acadêmico, tornou-se uma importante interlocutora intelectual.

Ao Professor José Flávio pela orientação na produção cartográfica.

Aos demais professores do Programa de Pós-Graduação em Tratamento da Informação Espacial.

Aos funcionários Fátima e Délio pela prontidão quando solicitados.

A Tom de Jong e Roland Goetgeluk pelas esclarecedoras respostas aos emails.

EPÍGRAFE

*“É melhor tentar e falhar,
que preocupar-se e ver a vida passar;
é melhor tentar, ainda que em vão,
que sentar-se fazendo nada até o final.
Eu prefiro na chuva caminhar,
que em dias tristes em casa me esconder.
Prefiro ser feliz, embora louco,
que em conformidade viver...”*

Martin Luther King

RESUMO

A proposta desta tese é apresentar e discutir uma contribuição ao estudo da análise espacial de fluxos migratórios com a explicitação de procedimentos metodológicos que possam ser aplicados a um conjunto de dados censitários. Tal abordagem metodológica tem como suporte dois métodos quantitativos: a Análise Fatorial de Ordem Superior e o Método Intramax e, ainda, a incorporação de um Sistema de Informação Geográfica. Enquanto a Análise Fatorial explora a similaridade existente no padrão de fluxos populacionais, o Método Intramax evidencia a interação existente entre as unidades espaciais. A metodologia é aplicada sobre os microdados dos Censos de 1991 e 2000 do Instituto Brasileiro de Geografia e Estatística-IBGE, tomando as 558 microrregiões brasileiras como unidades espaciais básicas.

Palavras-chave: Migração. Métodos Quantitativos. Análise Fatorial de Alta Ordem. Intramax. Análise Espacial.

ABSTRACT

The purpose of this thesis is to present and discuss a contribution to the study of spatial analysis related to migration flows and detail methodological procedures to be applied to a set of data from censuses. The approach presented herein relies on the quantitative methods of Higher-Order Factor Analysis and Intramax and comprises a Geographical Information System. While the Factor Analysis examines the similarities among patterns of migration flows, Intramax provides evidence of interaction among spatial units. The methodology is applied to microdata census samples from Instituto Brasileiro de Geografia e Estatística-IBGE referring to 1991 and 2000 and considers 558 micro-regions in Brazil as basic spatial units.

Key-words: Migration. Higher-Order Factor Analysis. Intramax. Quantitative Methods. Spatial Analysis.

LISTA DE FIGURAS

Figura 1: Padrões espaciais destacados com aplicações dos métodos IPFP e Intramax	44
Figura 2: Dendrograma de agrupamento baseado no padrão de migração	47
Figura 3: Divisão espacial após agrupamento.....	48
Figura 4: Comparação da regionalização funcional para o mercado imobiliário/2002, em linhas pretas espessas, com delimitação oficial em cores	50
Figura 5: Regiões de migração para profissionais e gestores de manufaturados, baseado no padrão de emigração, 1975-80.....	55
Figura 6: Áreas de migração para Califórnia nos períodos 1940, 1960 e 1980	56
Figura 7: Tipologia para imigrantes nos EUA/1940-1980. Os estados não hachurados são aqueles que não estão fortemente orientados para nenhuma das tipologias	57
Figura 8: Exemplo de dendrograma elaborado no software Flowmap 7.2	71
Figura 9: Etapas seguidas na metodologia proposta com uso do método Intramax	73
Figura 10: Etapas seguidas na metodologia proposta com uso da Análise Fatorial de Alta Ordem	79
Figura 11: Três primeiros subsistemas gerados pelo Método Intramax/1986-1991 .	83
Figura 12: Evolução dos subsistemas determinados/1986-1991	86
Figura 13: Subsistemas após aplicação de 500 passos do Método Intramax/1986-1991	91
Figura 14: Subsistemas após aplicação de 540 passos do Método Intramax/1986-1991	92
Figura 15: Subsistemas após aplicação de 550 passos do Método Intramax/1986-1991	93
Figura 16: Subsistemas após aplicação de 540 passos do Método Intramax/1995-2000	96
Figura 17: Subsistemas após aplicação de 550 passos do Método Intramax/1995-2000	97
Figura 18: Estudo de caso para o período 1986-1991 – Microrregião de Chapecó/SC.....	98
Figura 19: Estudo de caso para o período 1995-2000 – Microrregião de Chapecó/SC.....	100

Figura 20: Estudo de caso para o período 1986-1991 – Microrregião de Belo Horizonte/MG	102
Figura 21: Estudo de caso para o período 1995-2000 – Microrregião de Belo Horizonte/MG	104
Figura 22: Tipologias determinadas nas três primeiras ordens: Matriz Origem-Destino/1986-1991	108
Figura 23: Tipologias determinadas nas três primeiras ordens: Matriz Destino-Origem/1986-1991	111
Figura 24: Tipologias determinadas nas três primeiras ordens: Matriz Origem-Destino/1995-2000	113
Figura 25: Tipologias determinadas nas três primeiras ordens: Matriz Destino-Origem/1995-2000	114
Figura 26: Tela com informações sobre o software Flowmap 7.3	136
Figura 27: Caminho de acesso ao método Intramax no software Flowmap 7.3.....	136
Figura 28: Exemplo de mapa produzido com aplicação do método Intramax no software Flowmap	137
Figura 29: Tipos de variância na Análise Fatorial.....	139
Figura 30: Exemplo de Scree plot para determinação de número de fatores	154
Figura 31: Rotações ortogonal e oblíqua de fatores.....	159
Figura 32: Uma análise de alta ordem geográfica.....	163
Figura 33: Estágios no procedimento de Análise Fatorial de Ordem Superior.....	165
Figura 34: Representação de uma matriz de dados.....	166

LISTA DE QUADROS

Quadro 1: Determinantes da desconcentração econômica, segundo o autor.....	32
Quadro 2: Principais características dos métodos abordados.....	119
Quadro 3: Elementos fundamentais da Análise de Componentes Principais.....	157

LISTA DE SIGLAS E ABREVIATURAS

ACP – Análise de Componentes Principais

FLOWMAP – Software dedicado a análise e apresentação de dados de fluxo elaborado por University of Utrecht- Holanda

IBGE – Instituto Brasileiro de Geografia e Estatística

IPEA – Instituto de Pesquisa Econômica Aplicada

IPFP - Iterative Proportional Fitting Procedure

INTRAMAX – Método iterativo para avaliação de interações entre unidades espaciais

MATLAB – MATrix LABoratory – Software matemático da Mathworks

SPSS - Statistical Package for the Social Sciences

UNICAMP – Universidade Estadual de Campinas

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Considerações iniciais	14
1.2 Objetivos e justificativa	17
1.3 Estrutura da tese	19
2 UMA SÍNTESE DAS MIGRAÇÕES BRASILEIRAS A PARTIR DA SEGUNDA METADE DO SÉCULO XX	20
2.1 Contextualização: o processo produtivo brasileiro	20
2.2 A desconcentração do processo produtivo brasileiro e os reflexos sobre os movimentos migratórios	30
3. O ESTUDO DE FLUXOS COM APLICAÇÃO DE MÉTODOS QUANTITATIVOS	41
3.1 Estudos de fluxos a partir do Método Intramax	41
3.2 Estudos de fluxos a partir da Análise Fatorial de Ordem Superior	50
4 MÉTODOS QUANTITATIVOS PARA ANÁLISE GEOGRÁFICA DE FLUXOS MIGRATÓRIOS	59
4.1 A importância e a evolução dos métodos quantitativos para os estudos geográficos	59
4.2 Dado Espacial	60
4.3 Métodos multivariados e dados geográficos	62
4.4 Proposta metodológica	64
4.4.1 <i>Introdução</i>	64
4.4.2 <i>Delimitação de subsistemas de migração com uso do Método Iterativo Intramax</i>	65
4.4.3 <i>Delimitação de tipologias de migração com uso da Análise Fatorial de Ordem Superior</i>	74
5 A IDENTIFICAÇÃO DE SUBSISTEMAS E TIPOLOGIAS ESPACIAIS DOS FLUXOS MIGRATÓRIOS BRASILEIROS	80
5.1 <i>Introdução</i>	80
5.2 <i>Subsistemas de migração</i>	81
5.3 <i>Tipologias de migração</i>	104
5.4 <i>Uma síntese dos resultados</i>	115
6 CONSIDERAÇÕES FINAIS	117
REFERÊNCIAS	122

APÊNDICE A - FUNDAMENTOS MATEMÁTICOS E COMPUTACIONAIS DO MÉTODO INTRAMAX	131
APÊNDICE B – FUNDAMENTOS MATEMÁTICOS DA ANÁLISE FATORIAL	138
APÊNDICE C – Mapas de localização	168
APÊNDICE D – Mapas operacionais produzidos com base nos resultados da aplicação do método Intramax para o período 1986-1991	168
APÊNDICE E – Mapas operacionais produzidos com base nos resultados da aplicação do método Intramax para o período 1995-2000	168
APÊNDICE F – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1986-1991/Matriz Origem-Destino... 168	168
APÊNDICE G – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1986-1991/Matriz Destino-Origem.... 168	168
APÊNDICE H – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1995-2000/Matriz Origem-Destino.... 168	168
APÊNDICE I – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1995-2000/Matriz Destino-Origem.... 168	168
APÊNDICE J – Dendrogramas	168

1 INTRODUÇÃO

1.1 Considerações iniciais

O movimento de pessoas entre países, estados, regiões, municípios, etc. é um fenômeno antigo, bastante complexo e de interesse de muitos pesquisadores, gestores governamentais e de instituições privadas, dados os reflexos que tais deslocamentos provocam sobre a estrutura da população e nos interesses da sociedade.

Dados mostram que o Brasil apresenta importantes alterações no ritmo de crescimento populacional que impactam seus processos de planejamento e gestão e a avaliação de políticas de educação e saúde, dentre outras. Segundo registros do IBGE, enquanto na década de 1960 sua população cresce, em média, a taxas próximas de 2,8% ao ano, na década de 1970 este valor se aproxima de 2,4% ao ano e na década de 1980 são observadas taxas inferiores a 2% ao ano, apontando para um quadro de reversão confirmado na década de 1990.

Além disso, as alterações se mostram dispersas heterogeneamente pelo território brasileiro. Enquanto na década de 1980 a região Norte cresce a uma taxa de 4,01% ao ano e a região Centro-Oeste a 3,02%, as regiões Sudeste e Sul apresentam acréscimos populacionais a taxas inferiores a 2% ao ano.

Quando se avalia essa dinâmica populacional, algumas componentes podem ser estudadas: a variação natural - apontada pela mortalidade e pela natalidade, sendo essa última diretamente ligada à fecundidade - e a migração. Esses elementos do comportamento demográfico podem, de modo complementar, esclarecer muitas alterações detectadas.

A primeira tábua de mortalidade representativa da população brasileira, data de 1980 (IBGE, 2009a). Enquanto a expectativa de vida ao nascer do brasileiro nascido em 1980 era de 62,6 anos, para o nascido em 2000 este valor atinge a marca de 70,5 anos. Aliada a esta componente, a taxa de mortalidade no Brasil apresenta forte queda de 1950 a 1970, passando de 21 óbitos por 1.000 habitantes ao ano para 9 óbitos, chegando à média de 7,7 óbitos em 1991.

Quanto à taxa de fecundidade da população brasileira, definida como o número médio de filhos nascidos vivos, tidos por mulher ao final do seu período reprodutivo, conforme dados do IBGE (2009b) observa-se que, a partir da década de 1960, ocorre um importante declínio. Enquanto nos anos 1970 essa taxa está próxima de 5,8 filhos por mulher, na década de 1990 esse indicador passa para 2,9 filhos. A taxa de fecundidade pode ser apontada como um importante determinante da dinâmica demográfica, e a queda observada indica o início de um processo de envelhecimento da população brasileira.

Para Rigotti (2008), as alterações observadas nestes indicadores são conseqüências de um conjunto complexo de transformações sociais, econômicas e culturais que tiveram início em meados do século XX.

De acordo com os indicadores apresentados, a migração assume um caráter estratégico no entendimento da dinâmica populacional.

A diminuição sustentada da fecundidade implica em abrupta diminuição do crescimento natural. Sendo assim, as migrações passam a ser a componente mais importante para a identificação das tendências da distribuição espacial da população. (RIGOTTI, 2008,p.9)

Assim, neste trabalho, admite-se a importância da migração para o entendimento da dinâmica populacional brasileira, dada a significativa redução da fecundidade e as melhorias dos indicadores de mortalidade e de esperança de vida ao nascer.

Os debates motivados pela desconcentração ou reconcentração produtiva no Brasil, principalmente a partir da década de 1970, buscam esclarecer, também, as estruturas que norteiam o fluxo migratório populacional pelo país. Este fluxo acompanha ou é influenciado pelas transformações ocorridas nas atividades econômicas que passaram a ocupar novos espaços do território brasileiro.

A distribuição espacial do trabalho no Brasil passou por fortes alterações em relação às atividades do campo e da cidade ou entre atividades rurais e urbanas, desde os anos 1970. Em especial, estas mudanças atingem aquelas regiões mais dinâmicas economicamente e se caracterizam, mais recentemente, como migrações urbanas,

deixando transparecer a existência de uma correlação entre a rede urbana brasileira e as migrações nas últimas décadas.

As mudanças no processo migratório nacional tiveram, a partir dos anos 1970, o deslançar de suas transformações.” (BAENINGER,2008).

Assim, admite-se que a questão da migração no Brasil, que ocorre heterogeneamente pelo território nacional, esteja relacionada às grandes transformações da estrutura produtiva e, conseqüentemente, à distribuição espacial da oferta de trabalho observada, principalmente, a partir da década de 1970.

O estudo da migração, como proposto neste trabalho, é desenvolvido a partir de duas linhas complementares: a quantitativa e a geográfica. A primeira baseia-se no uso de um método estatístico multivariado, a Análise Fatorial de Ordem Superior, e de outro método matemático iterativo denominado Intramax. A segunda linha é representada pelo uso de um Sistema de Informação Geográfica que incorpora a qualidade espacial do dado e gera representações cartográficas auxiliares.

Segundo Pandit (1994, p.1), dois tipos de migração regional podem ser definidos:

- i. Regiões de migração como grupos de unidades territoriais que estão altamente conectadas com outras, denominadas **subsistemas de migração**.
- ii. Regiões de migração como grupos de unidades territoriais que apresentam padrões similares de trocas com outras unidades, denominadas **tipologias de migração**.

Entende-se, assim, que um subsistema de migração representa um subconjunto de unidades territoriais com altos níveis de trocas internas e baixos níveis de interação com outros subconjuntos. Ou seja, ao estudar a migração, baseando-se nos subsistemas de migração, o que se pretende é a regionalização do espaço total, de acordo com as interações existentes entre as diversas unidades.

Quanto ao estudo das tipologias de migração, busca-se regionalizar o espaço total formando subconjuntos de unidades territoriais que apresentem maiores similaridades dos seus fluxos com outras unidades territoriais.

Dois exemplos didáticos podem ser usados para mostrar a diferença entre subsistemas de migração e tipologia de migração. Considere dois estados distantes, A e B, com altas porcentagens de população idosa e que tenham entre si baixos fluxos migratórios. Estes estados A e B podem apresentar fortes conexões com um terceiro estado C, que representaria um destino importante para seus migrantes. Estes dois estados, neste caso, poderiam fazer parte de uma mesma tipologia de migração, sem pertencerem a um mesmo subsistema de migração. A escolha pelo estado C, pode ser motivada por questões culturais ou mesmo caracterizar uma migração de retorno.

Outra situação hipotética que ilustra a diferença é a de dois estados, A e B, que passam por graves dificuldades econômicas e que, provavelmente, apresentam baixos níveis de interação entre si. A e B podem apresentar-se conectados pela característica de serem estados que estejam expulsando parte de sua população para outro estado mais próspero C, sem constituir um subsistema de migração. Neste caso, o fator que motiva a migração, provavelmente estará relacionada às questões econômicas ou de mercado de trabalho.

Deve-se destacar, ainda, que os dados de migração analisados neste trabalho estão relacionados ao conceito de migrante de data-fixa, ou seja, aquele indivíduo que apresenta município de residência em 1986 ou 1995 diferente daquele registrado em 1991 ou 2000, respectivamente. Trata-se, portanto, de migrantes sobreviventes dos respectivos períodos, independentemente de terem ou não realizado movimentos migratórios posteriores.

1.2 Objetivos e justificativa

O que se busca com este trabalho de tese de doutorado é proporcionar uma contribuição à análise espacial de fluxos migratórios, com a explicitação de procedimentos metodológicos que possam ser aplicados a um conjunto de dados censitários, nos níveis de microrregiões, buscando identificar, para o caso do Brasil:

- i. As tipologias das regiões receptoras e perdedoras de população
- ii. Os subsistemas de migração populacional;

- iii. As intensidades, direções e sentidos dos fluxos;

Objetivos específicos:

- Avaliar a consistência matemática de cada uma das metodologias propostas, detalhando os fundamentos envolvidos na Análise Fatorial e no método iterativo Intramax;
- Aplicar as metodologias propostas na base de microdados dos Censos de 1991 e 2000, fornecidos pelo IBGE e disponíveis no Programa de Pós-Graduação em Geografia-Tratamento da Informação Espacial da PUC Minas, com o intuito de avaliar a aplicabilidade da metodologia proposta e avaliar possíveis mudanças ou estabilidade no contexto migratório brasileiro;
- Discutir os relacionamentos entre as regionalizações obtidas empiricamente e aquelas definidas pelo contexto econômico-produtivo brasileiro;
- Incorporar um Sistema de Informação Geográfica na análise e representação gráfica dos resultados empíricos obtidos, com a adoção de mapeamentos usando o ARCGIS 9.0, com suporte do software Flowmap 7.2.

O interesse pelo estudo da migração no Brasil é justificado, pois este fenômeno é um reflexo dos fenômenos naturais, políticos, econômicos e de produção, dentre outros relacionados à vida em sociedade, causando impactos sobre a gestão dos recursos ligados à educação, saúde, emprego, etc. Mas, também, pode ser entendida como um medidor da eficiência das medidas estatais tomadas para ocupação do território, redistribuição de atividades produtivas e de emprego, que busquem atrair melhorias para a população de determinadas regiões.

O enfoque geográfico com o suporte da quantificação, como apresentado neste trabalho, pode se tornar uma importante contribuição metodológica complementar no estudo da migração, ao evidenciar os relacionamentos espaciais inerentes a este fenômeno.

1.3 Estrutura da tese

O trabalho está organizado em cinco capítulos: o primeiro apresenta uma breve introdução sobre a relevância de estudos sobre migração e os objetivos da tese. O capítulo 2 contextualiza a questão da migração no Brasil, relacionando-a com o processo produtivo brasileiro, e apresenta uma revisão das referências sobre os métodos quantitativos abordados. No terceiro capítulo, faz-se uma discussão sobre o uso de métodos quantitativos na Geografia, o dado espacial e são apresentadas as propostas metodológicas para estudo da migração. As fundamentações matemáticas encontram-se nos Apêndices A e B.

No quarto capítulo, são analisados os resultados obtidos com a aplicação das metodologias propostas no capítulo anterior sobre os dados de fluxo migratório por microrregião brasileira para os períodos de 1986-1991 e 1995-2000.

O capítulo cinco é dedicado a algumas considerações finais sobre as metodologias propostas, suas características principais, similaridades, diferenças e potencialidades. Além disso, são registradas algumas propostas para estudos futuros.

Em um CD – Compact Disc – estão os demais Apêndices, de C a J, contendo toda a cartografia produzida com mais de 600 mapas complementares, em arquivos com a extensão .pdf. Como sugestão, ao abrir cada arquivo, deve-se clicar “Ctrl + L” para melhor visualização em tela cheia.

2 UMA SÍNTESE DAS MIGRAÇÕES BRASILEIRAS A PARTIR DA SEGUNDA METADE DO SÉCULO XX

Tendo em vista os objetivos precípuos desta tese, faz-se uma síntese do processo produtivo brasileiro a partir da segunda metade do século XX e seus reflexos sobre os movimentos migratórios.

2.1 Contextualização: o processo produtivo brasileiro

A reestruturação produtiva observada no Brasil até os anos 70 provocou impactos na rede urbana e, com isso, os principais movimentos migratórios observados nas décadas de 80 e 90 apresentaram características distintas daquelas observadas até a década de 70, – quando prevalecia o rápido crescimento das metrópoles e de outras grandes aglomerações urbanas. De fato, este forte crescimento das cidades não foi acompanhado por um crescimento industrial distribuído homogeneamente pelo país. Como conseqüência, nos períodos mais recentes, observa-se uma expansão da rede urbana brasileira. Esta expansão acaba realimentando o novo processo de reorganização espacial da população brasileira.

Com o objetivo de analisar o processo migratório observado no Brasil, da segunda metade da década de 80 até o ano 2000, faz-se necessário entender o processo de redistribuição das atividades produtivas no País e suas conseqüências quanto às alterações regionais provocadas. Pretende-se, assim, fazer uma revisão, ainda que sintética, de fatos importantes observados na economia nacional, bem como sobre sua evolução industrial, e o relacionamento entres estes fatos, redes urbanas e os movimentos migratórios.

Segundo Magnoli e Araújo (1996), na segunda metade do século XIX, o comércio do café era o principal responsável pela relação do Brasil com o comércio exterior. Pode-se dizer que o Brasil abastecia o mundo com seu café e esta demanda provocou uma grande revolução na vida econômica e social, principalmente no sudeste do país. Segundo Baer (1996, p.37), a extensão de ferrovias cresceu de 14 quilômetros, em 1854, para 474, em 1864; 3.302, em 1884; 16.306, em 1904.

Com a proximidade e, posteriormente, o fim da escravidão, a burguesia cafeeira foi obrigada a encontrar uma solução para a falta de mão-de-obra. Com isto, na década de 1880, registrou-se o início de uma das primeiras e mais importantes migrações no Brasil, enquanto as ferrovias ampliavam as limitações espaciais do país e a população européia emigrava, buscando novas opções para contornar a crise econômica e social por que passava o “Velho Continente”.

Como consequência natural do momento histórico, o Brasil é impelido para uma revolução em sua produção industrial. Para Magnoli e Araújo (1996), ao final do século XIX e início do século XX, a região sudeste do Brasil se apresenta como o berço de sua industrialização, baseada na economia cafeeira. Além do café, predominaram as produções de bens de consumo não-duráveis como têxteis, vestuário, alimentos e calçados. Rapidamente, o estado de São Paulo, que gerou um mercado consumidor interno relevante, assume a liderança tanto em quantidade de empresas, como em número de trabalhadores empregados. Neste período, as taxas de crescimento demográfico, até então estagnadas, assumem valores próximos de 2%.

De acordo com Magnoli e Araújo (1996), nos momentos de crise (ou oscilações), provocados pela superprodução cafeeira, que derrubava os preços do café e fazia recuar a plantação em novos cafezais, o êxodo rural provocava maior ocupação das cidades por trabalhadores em busca de emprego. E, com a queda das exportações de café, os investidores, em busca de novas aplicações, favoreciam o florescimento da indústria. Ou seja, o processo de industrialização no Brasil foi promovido à custa de várias crises nos cafezais. Entre 1900 e 1940, a população brasileira salta de aproximadamente 18 milhões de habitantes para 41 milhões.

Com a crise da Bolsa de Nova Iorque, em 1929, a economia brasileira, assim como a do resto do mundo, foi fortemente atingida. As exportações agrícolas, base da economia, eram fortemente reduzidas, assim como as importações de manufaturados. Neste período, a produção cafeeira encontra sua pior crise, ocasionada pelas reduções bruscas em sua cotação. É o fim de uma época histórica para o Brasil, pois o café nunca mais retomaria sua posição dominante.

Deve-se observar que Minas Gerais e o Nordeste sediaram importantes atividades econômicas até metade do século XIX e, portanto, acumularam grandes massas populacionais. E, como até 1930 não foram registrados significativos fluxos migratórios, fora o provocado pela borracha na região Norte, começa-se a perceber disparidades entre a parcela relativa de população e de produção industrial, quando se faz a comparação com São Paulo. Em 1920, aproximadamente 56% da população brasileira se encontrava na região Nordeste ou em Minas Gerais, mas estas mesmas regiões eram responsáveis por apenas 22% da produção industrial.

Como consequência, após 1930, Minas Gerais e região Nordeste registram altos índices de emigração. Este fenômeno também atingiu a região Sul e parte do Centro-Oeste, ainda que em menor escala. Entre a crise da década de 1930 e a Segunda Grande Guerra, ocorrida entre 1939 e 1945, o Brasil passa por uma Revolução Industrial. A indústria nacional encontrou espaços para crescimento, por exemplo, na produção de manufaturados, pois não estava submetida à mesma concorrência internacional de tempos atrás e o êxodo rural, provocado pela crise cafeeira, disponibilizava mão de obra para as fábricas. Na década de 1930, mais de 12.000 novas fábricas, mais do que as aqui existiam até o momento, foram criadas.

Para Magnoli e Araújo (1996), a Segunda Guerra Mundial, apesar de todos os seus malefícios, de certa forma apresentou aspectos positivos para a industrialização brasileira. Com a grande queda nas importações de equipamentos, a economia nacional sofreu um forte baque, ficando praticamente estagnada. Isto mostrou a grande, porque não dizer excessiva, dependência do Brasil em relação aos mercados estrangeiros e estimulou a produção nacional para atender a demanda, com incentivos governamentais, principalmente, à indústria de base.

Ainda, de acordo com Magnoli e Araújo (1996), deve-se ressaltar que a Companhia Siderúrgica Nacional (CSN) foi criada em 1941, com financiamento e transferência de tecnologia americana em troca de cessões de uso de portos no Rio Grande do Norte, tornando-se um marco da industrialização brasileira. Isso levou a uma redução radical nas importações, com a produção de 50% do ferro nacional numa única fábrica, embora sejam contestáveis a qualidade da tecnologia transferida e o

reflexo causado na dívida externa brasileira pelos altos custos para a implantação da CSN.

A partir da década de 1940, o crescimento da produção e do consumo energético registrados no Brasil refletem as profundas transformações estruturais em andamento na economia brasileira e a modernização da agricultura. Também neste período, a taxa de fecundidade contribuiu bastante para o crescimento demográfico, que ficou acima de 2%, favorecido pelas melhorias na urbanização, no sistema de saúde e com o desenvolvimento da indústria química farmacêutica.

Até então o Brasil era considerado, pelos países desenvolvidos, basicamente, como mercado consumidor de manufaturas e fornecedor de matérias-primas e produtos agrícolas. No pós-guerra, o país começa a receber investimentos diretos no setor industrial e inicia a mudança de um modelo agroexportador para um modelo urbano-industrial.

A partir de 1950, com a diversificação da base industrial, principalmente com o crescimento das indústrias de bens de consumo duráveis, criaram-se condições para a emergência de economias de aglomeração, onde regiões mais industrializadas (como, por exemplo, São Paulo) atraíam outras indústrias. E, como consequência, os serviços ofertados são favorecidos e ampliados e atraem mão-de-obra. Com tudo isto, acelerou-se o processo de urbanização.

Nos primeiros anos da década de 1960, a economia brasileira vinha perdendo seu dinamismo, apresentando taxas de crescimento do PIB muito pequenas. Enquanto em 1961, o PIB havia registrado um aumento de 10,3%, nos anos seguintes apresentou taxas muito menores. A partir de 1964, os governantes adotaram políticas de estabilização e de reformas estruturais no mercado financeiro, além de investimentos estatais na economia, que propiciaram grande crescimento nos períodos seguintes.

O Brasil passou, até meados da década de 1970, por um período de grande crescimento industrial e internacionalização da economia, paralelo à formação de um mercado consumidor interno significativo, inclusive de bens duráveis. Neste

período, são observadas mudanças significativas na economia brasileira e no comportamento de sua população, inclusive registrando grandes deslocamentos da população rural para a cidade.

A unificação do mercado nacional foi um processo que só avançou efetivamente com a industrialização. Em sua primeira etapa, entre 1930 e 1960, essa unificação foi essencialmente mercantil e conduzida pelo capital sediado em São Paulo... Essa articulação comercial, com o aumento do fluxo inter-regional de mercadorias, acabaria por sancionar a liderança da indústria paulista... É apenas com a industrialização pesada, ao longo dos ciclos expansivos dos anos 60 e 70, que a integração do mercado nacional avançaria de fato. (PACHECO, 1998, 52-53).

Na década de 1970, o Brasil registrou importante expansão em suas fronteiras agrícolas, principalmente na região Centro-Oeste. Ao todo, foram acrescentados mais de 70 milhões de hectares às áreas dedicadas à agropecuária. Com esta expansão, veio também a introdução de novos produtos e melhorias na produtividade e na rentabilidade.

Economicamente, o país cresceu drasticamente no período que vai de 1969 a 1973, conhecido como época do “Milagre Econômico”, com o PIB apresentando taxas de crescimentos próximas a 12% ao ano. Grandes obras eram patrocinadas pelo governo, como a Rodovia Transamazônica e a Ponte Rio-Niterói. Mas, estes investimentos deixaram uma grande conta a ser paga no futuro, provocando a geração de uma dívida externa incompatível com o padrão econômico brasileiro.

Em 1975, o então presidente Geisel apresenta o Segundo Plano Nacional de Desenvolvimento-1975/1979 (II PND), que trazia um imenso programa de investimentos, buscando promover uma verdadeira Integração Nacional, com incentivos significativos para a região Nordeste do Brasil. Neste período, observa-se o estímulo à iniciativa privada e a intervenção governamental com investimentos públicos.

O II PND apresentava, explicitamente, como um de seus objetivos na política industrial, a redução das desigualdades regionais em termos de localização de suas unidades produtivas, ofertando incentivos aos projetos industriais que, de alguma forma viessem a proporcionar a integração do desenvolvimento regional. Buscava-

se atender a questão de um maior equilíbrio das forças entre São Paulo, Minas Gerais e Rio de Janeiro e, ao mesmo tempo, reduzir o peso político de São Paulo.

Grande parte da modificação espacial da indústria foi um resultado da estratégia do II PND, que previa implantar projetos como: ferro e bauxita no Pará; cobre na Bahia; fosfato em Minas Gerais; Potássio em Sergipe; petróleo e gás natural nas plataformas nordestinas e fluminense; carvão e pirita em Santa Catarina e no Rio Grande do Sul; sal no Nordeste; siderurgia no Maranhão, no Espírito Santo e em Minas Gerais; cloroquímica em Alagoas e Sergipe; papel e celulose e cimento no Nordeste e no Espírito Santo; e diversas hidrelétricas como Itaipu, Paulo Afonso IV, Xingó e Tucuruí... O resultado foi a redução dos níveis de concentração econômica. (PACHECO, 1998, p.68)

Não que São Paulo tenha reduzido sua produção, apenas sua participação relativa diminuiu. Por exemplo, na década de 1970, sua participação caiu de 58,1% para 54,4% na composição do PIB nacional. No período de 1975 a 1980, o PIB do estado de São Paulo cresceu, em média, 5,9% a.a., em contraposição aos 10,4% a.a. observados no período de 1970/1975, enquanto Minas Gerais e o Nordeste brasileiro mantiveram crescimentos próximos de 10% a.a.

O estado de São Paulo manteve seu poderio industrial. Observe-se que os quase 5% de crescimento no PIB de São Paulo significam mais que os 10% alcançados, por exemplo, por Minas Gerais, dadas as disparidades de grandezas entres estes valores. Mas, ainda assim, vale salientar que ocorreu uma importante desconcentração referente à Região Metropolitana de São Paulo e ao interior do Estado. Enquanto São Paulo perdia participação relativa no total nacional, seu interior registrava um aumento, de 14,7% para 20,2%, da produção industrial do país. Ou seja, os incentivos apresentados pelos governos também contribuíram para o fortalecimento do interior de São Paulo.

Com isto, nota-se uma desconcentração da produção industrial e, conseqüentemente, os fluxos migratórios são influenciados. O trabalhador é atraído para outros municípios do estado de São Paulo, embora ainda próximos da região metropolitana de São Paulo, e outros centros industriais pelo Brasil, em Minas Gerais e na região Sul.

Merece destaque o fato de este crescimento ter sido impulsionado pelo bom desempenho, principalmente, das indústrias de bens duráveis e indústrias básicas, dependentes dos recursos naturais.

Pacheco (1998) aponta indicadores do avanço da indústria em praticamente todas as regiões e nos principais estados brasileiros:

- a) o Norte, em função da Zona Franca de Manaus e do complexo de mineração e de Siderurgia do Programa Grande Carajás;
- b) o Nordeste, em que se destacam o Pólo Petroquímico de Camaçari, na Bahia, a cloroquímica de Alagoas e Sergipe e a chamada “nova indústria” nordestina, constituída com apoio dos mecanismos de incentivos fiscais do sistema FINOR;
- c) o Centro-Oeste, especialmente pelo crescimento da agroindústria do complexo soja;
- d) no Sudeste, os estados de Minas Gerais e Espírito Santo ampliaram suas participações no Valor da Transformação Industrial-VTI nacional, devido a crescimento da siderurgia e da indústria de papel e celulose, além da instalação da FIAT em Betim;
- e) o Sul vai gradativamente ampliando sua participação, devido ao crescimento da indústria metal-mecânica no conjunto da região, especialmente da indústria de equipamentos agrícolas do Rio Grande do Sul, do Pólo Petroquímico de Triunfo e da indústria de calçados também neste estado, da têxtil de Santa Catarina, da indústria de papel e celulose do Paraná e do Rio Grande do Sul e, por fim, pelo acelerado processo de integração da agropecuária com a agroindústria nos três estados. (PACHECO, 1998, P.70)

Ou seja, as implantações dos PND's, ocorridas na década de 1970, favoreceram a desconcentração industrial, proporcionando o surgimento de novos pólos industriais fora do eixo Rio - São Paulo. Outro fator importante é o registro da modernização da produção agrícola. Com isto, o êxodo rural ganha força e a população passa a buscar as cidades médias de maior porte.

Apesar destes indicadores, o crescimento industrial nas diversas regiões continuava na dependência, principalmente, de São Paulo. Ou seja, as fronteiras industriais foram ampliadas, mas, em grande parte, para atender ou complementar suas necessidades, em termos de insumos, bens de capital e produtos finais. Sendo assim, já se pode questionar o surgimento de uma despolarização industrial e pensar em um início de desconcentração espacial das atividades produtivas no Brasil, baseado nas indústrias básicas e duráveis de consumo, iniciando “o desenvolvimento poligonal cujos vértices abrangiam uma área bem maior do que o

estado de São Paulo. Este polígono seria delimitado por Belo Horizonte, Uberlândia, Londrina, Porto Alegre, Florianópolis, Belo Horizonte” (Diniz, 1993).

Nos anos 1970, observou-se a tendência de aumento do número e do porte das cidades médias. Deve-se atentar, também, que esta nova configuração na estrutura de produção e de distribuição da população provocou grandes fluxos migratórios do espaço rural para as cidades. Observa-se uma interiorização da indústria, em especial para o interior paulista.

Os investimentos realizados na década de 1970 ainda provocaram repercussões no início dos anos 1980, inclusive quanto à tendência à desconcentração.

Na década de 1980, a dívida externa, que foi acelerada na década anterior, começou a ser um grave problema para as orientações políticas governamentais, devido à alta elevação dos juros internacionais, somada ao fim do fornecimento de crédito externo em condições atraentes. Este foi um período em que o Brasil apresentou desequilibrados saldos positivos na balança comercial, possivelmente favorecidos pela baixa na cotação do petróleo e por políticas econômicas voltadas para inversão dos resultados da balança comercial verificados na década de 1970.

Segundo Magnoli e Araújo (1997), os superávits alcançados na balança comercial não foram utilizados para sustentar as importações necessárias ao desenvolvimento industrial, mas para cobrir os também sucessivos rombos gerados nas transações invisíveis (juros e serviço da dívida). Com isto, o país perdeu a capacidade de importar e a base produtiva instalada nas décadas anteriores conheceu um processo de defasagem tecnológica. Não por acaso, os anos 80 são conhecidos como a “década perdida”.

Desde o fracasso do plano Cruzado, em 1986, até 1992, instalou-se no Brasil um período com baixíssimos crescimentos do PIB, (crescimento negativo em 1988, 1990 e 1992) e inflação exorbitante, chegando a superar a marca de 1000%, que mostram o grau de estagnação na economia.

Quanto à questão agropecuária, entre 1980-1985 registrou-se baixo crescimento no setor, inclusive nos investimentos na mecanização. Apesar disto, a fronteira agrícola continuou em expansão, apesar de ser em ritmo menor. Até 1985 este avanço ocorria dentro da região Norte, no Mato Grosso, parte do cerrado baiano e Paraná.

A pobreza no meio rural se agrava, cresce a mão-de-obra familiar sem remuneração e estabelece-se um processo de divisão das propriedades de menor porte, que não ajuda a diminuir as desigualdades.

Na segunda metade dessa década, com incentivos governamentais, entre eles a política de preços mínimos, ao invés de disponibilização de créditos, a questão agrícola melhora, permitindo aos produtores a recomposição de equipamentos e instalações, favorecendo o crescimento dos cultivos para exportação e para o mercado interno.

Segundo Pacheco (1998, p.102), o avanço da produção no Centro-Oeste, e no conjunto dos cerrados (Bahia e Minas Gerais, inclusive), é o aspecto que mais chama a atenção acerca das transformações regionais na agricultura, nesse período.

Com o deslocamento do cultivo da soja, principalmente para a região Centro-Oeste, também observa-se a presença da agroindústria processadora. Esta região reafirma, na década de 1980, sua vocação agroindustrial. No início da década registrava 14,4% da produção nacional de grãos e no período de 1989/1991 já registrava 23,4%, enquanto Minas Gerais permanecia com uma participação em torno de 9% e São Paulo via sua participação relativa estabilizada em torno de 9,3%.

De modo geral, analisando a questão agrícola para a década de 1980, como um todo, observa-se um resultado melhor que o industrial, mas, apesar disto, o crescimento se deveu mais ao crescimento da produção por área do que pela expansão das fronteiras. Ou seja, não houve ganhos significativos de empregos no meio rural e nem de qualidade nas relações de trabalho.

Portanto, a crise que tomou conta da indústria brasileira na década de 1980, não ocorreu de modo homogêneo. O que se observa são piores resultados nas indústrias de bens de capital e bens de consumo, que foram responsáveis pelo crescimento na década de 1970, e uma pequena expansão da fronteira agrícola, que favorece a desconcentração da atividade industrial, principalmente para o interior, em especial o paulista.

Em síntese, um balanço do desempenho industrial na década de 80 indica que “a produção de bens intermediários foi dinamizada pela exportação, principalmente das indústrias de papel e celulose, química petro-química, de produtos de borracha (pneumáticos) e metalúrgica. O crescimento da produção dos não-duráveis de consumo, por sua vez, foi impulsionado pela demanda interna, menos elástica em relação à renda. Entretanto, outros gêneros industriais produtores de bens não-duráveis de consumo tiveram um desempenho negativo (tais como têxtil, vestuário e calçados). (PACHECO, 1998,p.112)

Apesar dos fracassos nas tentativas de estabilização no período de 1990 a 1992, segundo Anjos (2002), a proposta de retomada de crescimento do governo Collor, alinhada ao ideário neoliberal, significou uma alteração profunda na forma de inserção da economia brasileira na economia internacional, no relacionamento do setor público com o setor privado e na atuação do Estado na economia.

Em busca de maior competitividade econômica, o governo apresentou um programa de privatizações, reduziu alíquotas que incidiam sobre as importações, eliminou-se a reserva de mercado, especialmente a relativa a computadores, promovendo a abertura comercial. Acreditava-se que, com a concorrência e o aumento de investimentos estrangeiros, a economia nacional se tornaria mais eficiente.

A estrutura de produção no Brasil sofreu importantes modificações a partir da década de 1990, que refletiram no mercado de trabalho. No início, uma das graves consequências foi a redução na oferta de emprego.

A abertura comercial propiciou o aumento de opções de bens para consumo e, além disso, trouxe a oportunidade de troca de produtos nacionais por importados. A indústria nacional começou a enfrentar a concorrência internacional, sendo obrigada a ser mais eficiente e, assim, buscou o ganho de produtividade, inclusive com a redução nos custos relativos à mão-de-obra.

A queda dos juros no mercado internacional e as renegociações da dívida externa brasileira aliviaram o peso do endividamento sobre o Produto Interno Bruto. E, ainda, a abertura dos mercados financeiros e a privatização de várias estatais criaram um grande mercado para investimentos internacionais.

O setor industrial busca avanços tecnológicos que otimizem o processo produtivo e, assim, exige também mão-de-obra mais especializada. Com isso, as indústrias são atraídas por centros mais desenvolvidos (que, não por acaso, são os mais urbanizados), que apresentam maiores ofertas de trabalhadores qualificados e centros de pesquisas que possam colaborar para o uso e o desenvolvimento de novas tecnologias.

Quanto à estrutura de distribuição da população brasileira, vale atentar para o fato de o número de cidades com mais de um milhão de habitantes passar de duas, em 1960, para cinco, em 1970, e 12, em 1991. Isto mostra os efeitos da metropolização na distribuição espacial da população brasileira, processo que se tornou mais forte a partir da década de 1970.

2.2 A desconcentração do processo produtivo brasileiro e os reflexos sobre os movimentos migratórios

Como visto, a divisão espacial do trabalho pelo território brasileiro vem passando por fortes alterações, desde os anos 1970, em relação às atividades do campo e da cidade ou entre atividades rurais e urbanas. Em especial, estas mudanças atingem aquelas regiões mais dinâmicas economicamente.

Do texto *Configuração atual e tendências da rede urbana*, publicado pelo IPEA em 2002, pode-se extrair que o processo de desconcentração da economia está sujeito a duas linhas de análises distintas, mas complementares: a primeira indica a ampliação e diversificação das funções desempenhadas pelos centros urbanos, que assumiram maior complexidade em suas atividades e a segunda, a ampliação das necessidades de articulação e integração entre eles. Isto é um reflexo da

desconcentração das atividades produtivas, em especial, pela interiorização do desenvolvimento.

Vários estudos indicam que a distribuição da produção industrial no Brasil tem mudado significativamente nos últimos anos, sendo observado um processo de desconcentração ou polarização reversa em andamento. Diniz (1993) admite como mais apropriado considerar este processo como um caso de desenvolvimento poligonal, onde um limitado número de novos pólos de crescimento ou regiões têm capturado a maior parte das novas atividades econômicas.

Houve nos anos 80 um relativo descolamento das economias regionais, com o surgimento de *Ilhas de produtividade*, muitas voltadas para o comércio exterior, o que estimulou maior heterogeneidade interna da estrutura produtiva nacional, aprofundando as desigualdades inter e intra-regionais do país. (CONFIGURAÇÃO...,2002,p.35).

Pode-se observar que o surgimento destes novos pólos não vem ocorrendo de modo homogêneo pelo território brasileiro. Fora alguns centros em regiões específicas e com funções particulares, o que se observa é que a maioria dos novos centros está próxima de ou no próprio estado de São Paulo.

Segundo Diniz (1993), este desenvolvimento, denominado por ele de poligonal, é resultado de um conjunto de forças, dentre elas:

- a) as deseconomias de aglomeração na área Metropolitana de São Paulo e a criação de economias de aglomeração em vários outros centros urbanos e regiões;
- b) o papel da política econômica ou ação do Estado em termos de investimentos diretos, incentivos fiscais e construção de infra-estrutura;
- c) a busca de recursos naturais, que estimulou a abertura de novas regiões para o desenvolvimento;
- d) grande concentração social e espacial da renda, e o conseqüente poder de compra e de pesquisa, que retém o crescimento em regiões com maior base econômica;
- e) a unificação do mercado através da infra-estrutura de transportes e comunicações.

A mudança da configuração regional brasileira, ocorrida deste a década de 1970 até os últimos anos da década de 1990, para Pacheco (1998), está sujeita a alguns determinantes:

os determinantes dessa desconcentração relativa foram muitos: o deslocamento da fronteira agrícola e mineral; o processo de integração produtiva do mercado nacional; o perfil relativamente desconcentrado do sistema urbano brasileiro; o surgimento de “deseconomias de aglomeração” e pressões ambientais nas áreas mais intensamente industrializadas; as próprias políticas de governo e o investimento do setor produtivo estatal; e, por fim, na década de 1990, os impactos diferenciados que a crise econômica e a orientação exportadora, bem como o ajuste microeconômico das novas formas de organização da grande empresa, assumiram em cada região (Pacheco, 1998, p.208).

Veja, no Quadro 1, alguns determinantes da desconcentração econômica, segundo alguns autores.

Carlos A. Pacheco	Wilson Cano	Clélio Campolina Diniz	Leonardo Guimarães Neto
Integração produtiva do mercado nacional	Processo de integração do mercado nacional	Unificação do mercado e concorrência interempresarial	Possibilidade de as empresas ocuparem espaços fora da área central, visando posições oligopolísticas
Deslocamento da fronteira agrícola mineral	Marcha para U oeste	Distribuição espacial da base dos recursos naturais	Vantagens da exploração de recursos naturais, criadas pela infra-estrutura
Políticas de governo: econômica Diretrizes setoriais Políticas de desenvolvimento regional Investimento do setor produtivo estatal	Políticas de desenvolvimento regional Políticas de incentivo e exportações e crédito rural Investimento do II PND	Ação do Governo Federal e dos Estados: Investimentos diretos Incentivos fiscais Infra-estrutura básica	Políticas setoriais Políticas de desenvolvimento regional (incentivos fiscais e financeiros) Política econômica
Efeitos da crise nas áreas mais industrializadas	Crise da década de 80, que teria afetado mais a economia de São Paulo		Desaceleração da economia, com efeitos negativos concentrados nas regiões mais industrializadas
Deseconomias de aglomeração e pressões ambientais nas áreas mais industrializadas		Contradição entre economias e deseconomias de aglomeração	Aspectos microeconômicos ligados às deseconomias de aglomeração em alguns espaços nacionais
Perfil diferenciado da orientação exportadora			
Novas formas de organização da grande empresa			
		A concentração regional da pesquisa e da renda	

Quadro 1: Determinantes da desconcentração econômica, segundo o autor

Fonte: Pacheco (1998,p.221)

Frente a toda esta discussão sobre a desconcentração produtiva no Brasil, fomentada, principalmente, durante a década de 1970, deve-se atentar para o fato de esta desconcentração ter ocorrido em grande parte dentro do estado de São Paulo, com a interiorização da indústria.

Outros pólos, ainda que de menor porte, se comparados com São Paulo, foram estabelecidos, configurando, na realidade, uma reconcentração ou aglomeração do setor produtivo dentro do polígono, definido por Diniz (1993). Não por acaso, os principais centros de produção e desenvolvimento de tecnologia encontram-se inseridos neste polígono, bem como a concentração da produção industrial e do mercado de trabalho, gerados a partir da década de 1970.

Segundo Diniz (1993), a concentração regional da pesquisa e do mercado de trabalho profissional, assim como a concentração pessoal e regional da renda, agem como obstáculo a um intenso processo de desconcentração industrial. Pode-se dizer que um dos grandes fatores capazes de frear esta desconcentração industrial esteja ligado aos grandes avanços da tecnologia aplicada no setor, que favorece o estabelecimento de indústrias em meios mais próximos de centros de pesquisa e com melhores resultados no setor educacional. Com isto, a região delimitada pelo polígono sugerido por Diniz apresenta-se como a mais adequada para localização industrial, descaracterizando uma possível reversão da polarização industrial.

Apostando no relacionamento entre localização de atividades econômicas e movimentos migratórios, e de acordo com o que foi discutido preliminarmente sobre a desconcentração industrial, deve-se acreditar em uma reorganização da ocupação do espaço brasileiro. Os dados mostram que as profundas modificações observadas nas atividades econômicas brasileiras provocaram mudanças, também, no processo migratório dentro do território nacional.

Estes fatos parecem ter implicado a intensificação da migração de curta distância ou de caráter intra-regional, e redução da migração em direção às metrópoles do Sudeste, sobretudo São Paulo...

O próprio desempenho das atividades agrícolas no período 1980-90 pode ter contribuído para maior retenção da população nas pequenas cidades. Isso é especialmente válido para o Sudeste, onde, ao lado do menor crescimento

das aglomerações urbanas metropolitanas, e em particular de suas cidades centrais, ocorreu um crescimento da população dos pequenos municípios bem superior ao esperado. (CONFIGURAÇÃO...,2002,p.35).

Quanto à organização espacial da população brasileira, segundo Rigotti (2006,p.240), de acordo com os dados preliminares do Censo 2000, quase um quarto da população vive em capitais, enquanto municípios com menos de 20 mil habitantes apresentaram diminuição da população. Os municípios de tamanho demográfico intermediário foram aqueles que apresentaram o mais expressivo ritmo de crescimento da população. Outra conclusão é sobre o papel de destaque que as cidades médias assumiram a partir dos anos 1990.

Alguns dos grandes centros urbanos (principalmente as capitais) vêm perdendo o poder de atração frente à população, inclusive apresentando saldos migratórios negativos, por motivos como violência excessiva, trânsito caótico, alta valorização dos bens imóveis e taxas de desemprego desanimadoras. Mas, apesar disto, esta mesma população não encontra incentivos para o desbravamento de novas fronteiras agrícolas, como ocorrido, por exemplo, em décadas anteriores para o Centro-Oeste. Assim, a rede urbana brasileira amadurece, em especial o interior paulista, as cidades médias e aquelas aglomerações urbanas não-metropolitanas, além de algumas capitais das regiões Centro-Oeste, Norte e Nordeste.

Vale lembrar, ainda, que a população em busca de trabalho tem encontrado posições no mercado que exigem mão-de-obra mais qualificada, dados os avanços tecnológicos em andamento no setor industrial.

De acordo com os novos fatores que influenciam a localização do setor industrial, os centros urbanos com melhores recursos para pesquisa e formação de base educacional tornam-se atrativos para os migrantes. Com a perda de poder de atração dos grandes centros, como citado anteriormente, os centros urbanos médios, que começam a apresentar novo potencial produtivo, tendem a assumir o papel de receptores, determinando novos pólos intra e inter-regionais.

Segundo Baeninger (2008, p.4), a análise dos movimentos migratórios, em anos recentes, indica o reforço da tendência de configuração de novos espaços da

migração, agora, no entanto, muito mais relacionados ao âmbito de suas próprias regiões.

Assim, o processo migratório brasileiro observado nas últimas décadas não pode ser explicado apenas do ponto de vista da economia. O estudo da densificação da rede urbana brasileira pode trazer importantes elucidações sobre a dinâmica existente e a integração do território.

O deslançar dos processos migratórios recentes tem suas raízes de transformações desde os anos 80, quando as clássicas interpretações da migração via o desempenho econômico das áreas alcançaram seus limites... (BAENINGER, 2008, p.1)

Rigotti (2008) adota uma abordagem desenvolvida a partir de dois ângulos complementares: o demográfico e o geográfico, no estudo da (re)distribuição espacial da população brasileira e possíveis impactos sobre a metropolização. Para esse autor, o Brasil, país de dimensões continentais, apresenta grande heterogeneidade em sua rede urbana, com algumas regiões extremamente dinâmicas e o surgimento de novos pólos regionais.

Segundo Matos e Braga (2002, p.1), ações do Estado voltadas à descompressão espacial existem no Brasil desde os anos 50, quando se elaboravam projetos, programas e diretrizes de políticas, visando ou o desenvolvimento das regiões “atrasadas”, ou, um pouco depois, o desenvolvimento das cidades de porte médio, de modo a interiorizar mais efetivamente o crescimento econômico.

Outro fenômeno relacionado com a ocupação espacial do território é a taxa de fecundidade, que desde meados da década de 60, encontra-se em fase de redução. Enquanto na década de 60 cada mulher tinha em média 6,3 filhos, em 1991 esta média fica próxima de 2,9 filhos (IBGE, 2009b). Assim, a migração se torna, relativamente, um fenômeno mais significativo para a explicação da dispersão da população no espaço brasileiro que o seu crescimento natural.

A diminuição sustentada da fecundidade implica em abrupta diminuição do crescimento natural. Sendo assim, as migrações passam a ser a componente mais importante para

a identificação das tendências da distribuição espacial da população. (RIGOTTI, 2008,p.9)

Enquanto, na década de 60, o Brasil ainda podia ser considerado país agrícola, pois 55% de sua população era rural, em 2000 aproximadamente 82% da população residia em zonas urbanas. Este forte crescimento das cidades não foi acompanhado no mesmo ritmo pelo crescimento industrial e o desemprego tornou-se um grave problema. Com isso, a distribuição espacial da população e o padrão de urbanização brasileiro se alteram.

Até certo momento da história, a urbanização do país era gerada pelos intensos movimentos migratórios da população rural para os centros urbanos em fase de crescimento, o êxodo rural. Mas, nas últimas décadas, o que se observa é o impacto das migrações urbanas, ou seja, da população que migra de um centro urbano para outro centro urbano.

Boa parte da expansão da rede urbana nacional nas últimas décadas deriva dos efeitos de “espraiamento” originários da histórica concentração urbano-industrial do Sudeste, onde afinal, estava quase metade dos centros urbanos em 1980, entre os quais as três principais áreas metropolitanas.

A associação entre a descompressão espacial e a rede de localidades centrais em expansão é, por definição, bastante evidente. Isto porque baseia-se no pressuposto da existência de uma estrutura espacial organizada (só aparentemente desorganizada), na qual as relações de interação são inerentes à circulação de bens e serviços que a urbanização produz. (MATOS; BRAGA, 2002, p.2).

Usando outras palavras, o que se observa é a existência de uma relação entre a rede urbana brasileira e as migrações internas nas últimas décadas.

No texto *Configuração Atual e Tendências da Rede Urbana* (IPEA, IBGE e UNICAMP, p.35) encontra-se a pressuposição de que as tendências da urbanização brasileira e o próprio desenho do sistema urbano do país incorporam as transformações espaciais da economia. A partir dos anos 80, começa-se a verificar o surgimento de ilhas de produtividade, principalmente voltadas para a exportação.

Os dados evidenciam, em grande medida, a intensificação do processo de urbanização das últimas décadas, e sugerem a existência de um processo de redistribuição da população vis-

à-vis o aprofundamento da integração físico-territorial, tanto no Sudeste, quanto em vários outros subespaços urbanos articulados diretamente a ele... Todas as grandes regiões brasileiras aumentaram sua participação na população urbana do País, em especial o Centro-Oeste, o Norte e também o Nordeste, áreas com características predominantemente rurais até bem pouco tempo atrás. (MATOS; BRAGA, 2002, p.10)

Para Domenach e Picouet, (1990), citados por Baeninger,

O descolamento da relação migração-industrialização, migração-fronteira agrícola, migração-desconcentração industrial, migração-emprego, migração-mobilidade social no contexto atual da economia e da reestruturação produtiva, em anos recentes, induziu um novo dinamismo às migrações no Brasil, onde os fluxos mais volumosos são compostos de idas-e-vindas, refluxos, re-emigração, outras etapas – que pode ser mesmo o próprio local de origem antes do próximo refluxo para o último destino –, onde as migrações assumem um caráter mais reversíveis. (DOMAENCH e PICOUET apud BAENINGER,2008,p.5)

Sobre a urbanização induzida pelas regiões do agronegócio, pode-se destacar a região Centro-Oeste que, nas décadas de 1960 e 1970, teve fortalecida a introdução de migrantes nas fronteiras de ocupação, causando reflexos na rede urbana. Na realidade,

esse período representa o corolário das transformações que tiveram início na década de 1930, com a implementação das políticas nacionais de colonização, integração e interiorização da economia, capitaneadas pelo Estado brasileiro, e, ainda, com a implantação, algum tempo depois, dos dois principais núcleos urbanos, que cumprem funções polares na região, respectivamente, Goiânia e Brasília. (CONFIGURAÇÃO...,2002,p.311).

Mesmo na década de 50, muitos nordestinos migram como consequência da forte seca que assola a região, sendo atraídos pelas oportunidades geradas na construção civil no Centro-Oeste, especialmente para Brasília e Goiânia. Na década de 1980, o Centro-Oeste ainda apresenta importante crescimento, relacionado às questões da fronteira agrícola, com significativa produção de grãos e carne.

Quanto ao estado do Maranhão, também denotado como fronteira agrícola, deve-se destacar expressiva mudança na década de 1950, com o fortalecimento da cultura do babaçu e a inserção da produção do arroz, além da instalação da hidrelétrica de

Boa Esperança. Nesse período, o número de habitantes apresenta uma taxa de crescimento de 57,4%, passando de 1.583.248, em 1950, para 2.492.139, em 1960. A partir da década de 1970, migrantes provenientes do Sul do país, trazendo uma agricultura, com uso de tecnologias modernas, às áreas do cerrado e aprimoramento tecnológico da pecuária, valorizam a produção regional. Deve-se chamar a atenção para o fato de o estado estar localizado como região de passagem para muitos migrantes para as terras de mineração no estado do Pará. Assim, muitos acabam se estabelecendo por lá. São registradas taxas decenais de crescimento populacional de 34,9% e 20,3% nas décadas de 1970 e 1980, impactando a rede urbana do estado.

Sobre a fronteira de expansão no estado do Piauí, destaca-se a pecuária como primeira atividade econômica praticada no estado e, ainda, uma das mais importantes, sendo a criação de caprinos o maior destaque. Em seus vales úmidos, a extração vegetal é caracterizada pelas matas de babaçu e carnaúba e, mais recentemente, a introdução das culturas de mamona e grãos, que geram grandes latifúndios. Segundo Alves (2005), a instalação de migrantes com origem no Sul do Brasil ou de sulistas que se deslocam de outras regiões brasileiras é observada a partir da década de 1980 e fortalecida na década de 1990, sendo os migrantes com o mesmo perfil daqueles que foram para o oeste baiano e o sul maranhense em busca de uma agricultura capitalista.

Ainda sobre áreas de expansão, é importante evidenciar os estados da Bahia e do Paraná. A partir da década de 1980, o estado da Bahia tem sua dinâmica econômica alterada em decorrência do fortalecimento do Pólo Petroquímico de Camaçari, do incentivo às atividades relacionadas ao turismo e da produção de frutas para exportação, tornando-se importante região receptora de migrantes de seu entorno, mas, possivelmente, também daqueles que voltavam de São Paulo e do Rio de Janeiro.

A região do Paraná, no período de 1940 a 1960, é favorecida com a transferência de grande parte da produção cafeeira do estado de São Paulo, atraindo migrantes, principalmente dos outros estados do Sul do Brasil e de São Paulo. Este quadro sofre uma forte reversão durante a década de 1970, tornando-se uma importante

fonte de migrantes para os grandes centros da região Sudeste, além de provocar a migração do tipo rural-urbana, pela forte concentração de terra observada no período anterior. E, a partir da década de 1980, passa a compor a denominada corrente sulista de migrantes para os estados do Maranhão e Piauí, ainda que em menor escala que o estado do Rio Grande do Sul.

Outra importante questão é a complexidade espacial dos fluxos de pessoas altamente qualificadas/escolarizadas, brevemente citada anteriormente. Sobre esse tema, Rigotti (2006) apresenta resultados de um estudo em que dois grupos de migrantes são avaliados para os períodos 1986-1991 e 1995-2000: pessoas com mais de 25 anos de idade, com quinze anos ou mais de estudo, e outro com até quatro anos de estudo, na mesma faixa etária.

Rigotti destaca que, nos dois períodos estudados, os migrantes mais qualificados são originários de uma região metropolitana ou uma capital de seu próprio estado. Não coincidentemente, são essas as regiões com maior concentração de universidades e centros de pesquisa. Entretanto, os grandes centros também representam os principais destinos desses migrantes. Esse fato apresenta certa coerência, tendo em vista a importância desse migrante para os pólos regionais em expansão.

Exemplificando,

a Região metropolitana de São Paulo desempenha o duplo papel de oferecer e receber migrantes qualificados, mantendo uma grande interação com o interior do Estado e regiões limítrofes dos Estados vizinhos. Nesse sentido, a Mesorregião de Campinas se destaca, pois sempre aparece como área de atração para aqueles que deixam a mesorregião da capital. Aliás, esse mesmo papel é desempenhado por algumas localidades no Sul do país, como no Vale do Itajaí, destino para aqueles que saem de Florianópolis; enquanto Londrina e Maringá atraem migrantes das áreas vizinhas. (Rigotti, 2006, p.250)

Rigotti também destaca a importância do migrante mais qualificado, particularmente daquele proveniente do Rio de Janeiro, para outras regiões muito distantes, como aquelas localizadas no Norte do Brasil, apesar de chamar a atenção para a tendência à regionalização desse tipo de fluxo.

Sobre os migrantes com menor escolaridade, o estudo mostra que as mesorregiões que contêm aglomerações metropolitanas ou capitais também exercem o papel tanto de origens quanto destinos principais, mas os grandes centros se destacam mais fortemente como receptores desse tipo de migrante do que como emissores.

Como exemplo, é citado o caso de Campinas como

um centro de recepção de pessoas menos qualificadas; mas, ao contrário do que ocorre em relação aos mais escolarizados, sua mesorregião exerceu uma função centrípeta muito mais extensa, abrangendo pessoas de praticamente todo o este do Estado. (Rigotti, 2006, p.251)

Fato semelhante é observado no Sul do Brasil, onde as regiões metropolitanas continuam sendo os destinos preferidos da população menos qualificada, enquanto as regiões do interior atraem os mais qualificados.

Quanto ao migrante do Nordeste brasileiro, observa-se que os mais escolarizados optam por ficar na própria região, ao contrário dos menos qualificados, que se arriscam percorrendo distâncias mais longas.

O estudo dos impactos que a escolaridade do migrante exerce sobre o padrão do fluxo migratório merece ser feito com maior profundidade, assim como a questão da faixa etária, mas fica além das pretensões deste trabalho, servindo apenas para mostrar que existem vários fatores que influenciam a escolha do migrante além da distância entre a origem e o destino.

Encerra-se, assim, uma breve contextualização sobre a migração no Brasil, enfatizada em sua relação com os fatores econômicos que, pelo menos de modo complementar, influenciam as opções do migrante.

3. O ESTUDO DE FLUXOS COM APLICAÇÃO DE MÉTODOS QUANTITATIVOS

Neste momento, expõem-se as técnicas que são utilizadas na proposta metodológica, apresentando suas principais características, bem como suas limitações. Posteriormente, discute-se como o uso simultâneo dessas técnicas pode contribuir para o aprofundamento no estudo do padrão espacial das migrações no Brasil.

3.1 Estudos de fluxos a partir do Método Intramax

Uma alternativa metodológica para a abordagem do estudo da identificação de padrões em dados de interação espacial é o método Intramax. É possível observar sua aplicação no estudo de áreas de mercado delimitadas por dados de fretes, na identificação de padrões na comunicação via telefone, na delimitação de regiões funcionais em termos de viagens ao trabalho ou análise da migração populacional, dentre outras.

Neste item, faz-se a apresentação de alguns desses casos, procurando, com isso, apresentar as principais características e limitações do método Intramax.

Brown e Pitfield (1990) discutem as dificuldades existentes na tentativa de distinguir e sumarizar as similaridades e diferenças que são observadas nos trabalhos que buscam determinar uma configuração de áreas de mercado de produtos. Os autores aplicam o método Intramax em um banco de dados que compreende um registro de fluxos, em termos de toneladas de mercadorias, que foram observados, em 1972, entre 134 áreas da Grã-Bretanha, com respeito a 30 grupos de mercadorias. No referido artigo, são apresentados os resultados obtidos para 5 desses grupos de mercadorias.

Um das conclusões apresentadas nesse trabalho é a não existência de consenso ou regra para decidir qual número de partições¹ ou áreas de mercado deve ser adotado.

¹ Define-se como partição de um conjunto A qualquer subconjunto do conjunto das partes de A, tal que nenhum dos elementos dessa partição seja o conjunto vazio, a interseção entre dois elementos da partição seja vazia e a união das partes seja o conjunto A.

Ou seja, o número de agrupamentos a ser determinado. Segundo os autores, isso varia de acordo com a proposta da análise e a natureza do conjunto de dados. Vários critérios podem ser aplicados na determinação deste número, tais como, por exemplo, escolher um ponto de descontinuidade no aumento progressivo de fluxos intrazonais.

Nesse trabalho, fez-se a opção por determinar 20 áreas para todos os 5 grupos de mercadorias e comparar os resultados obtidos. Foi importante o registro e análise da proporção de fluxos que se apresentam como intrazonais, após a determinação das partições sugeridas para a unidade espacial.

Fischer e outros (1993) apresentam e comparam duas alternativas metodológicas para a abordagem do problema da identificação de padrões na interação espacial: IPFP (Iterative Proportional Fitting Procedure), originalmente proposto por Slater em 1976, e Intramax, proposto por Masser e Brown em 1975. Faz-se breve discussão sobre as principais vantagens e desvantagens dessas metodologias e apresentam-se os resultados obtidos no estudo de caso referente a dados da telecomunicação medidos pelo Austrian PTT (responsável pela infraestrutura de telecomunicação pública na Áustria), em 1991.

O artigo chama a atenção para as dificuldades associadas com a identificação de padrões nos dados de fluxo de comunicações por telefone e os resultados empíricos obtidos ilustram a superioridade da abordagem pelo método Intramax quando comparado com o método IPFP. De acordo com os autores, os resultados obtidos via Intramax são mais facilmente interpretáveis em termos de regiões funcionais.

O estudo empírico serviu para avaliar o desempenho das duas abordagens metodológicas na prática, mais em termos qualitativos do que em termos quantitativos. Os dados se referem ao tráfego total de telecomunicação entre 32 distritos representando o segundo nível da estrutura hierárquica da rede telefônica austríaca (denominado regional). Logo, os dois métodos foram aplicados a uma matriz de fluxo de tamanho 32x32 com zeros na diagonal. Os autores aplicaram o método IPFP disponível no software estatísticos SAS e o Flowmap versão 2.1 para o caso do Intramax.

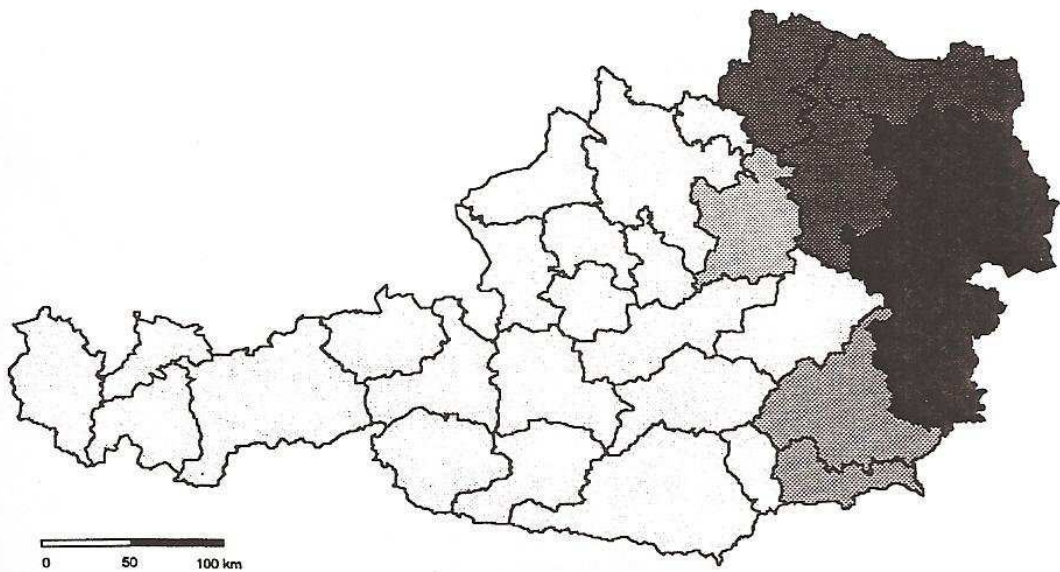
No IPFP mais de dois grupos são unidos num mesmo passo e algumas unidades não são unidas com outras, até os últimos estágios do processo de agregação. De acordo com Slater, citado por Fischer e outros (1993, as entradas de unidades espaciais simples no final do processo de agregação indicam unidades que têm mais identidade no nível nacional do que no regional no processo de comunicação telefônica. Segundo os autores, isso se mostra razoável para Viena, mas não para outras unidades que ficaram isoladas até quase o final do processo, como Feldkirch e Amstetten.

Em contraste, no método Intramax não existem entradas posteriores de unidades espaciais simples. Sua aplicação deixa clara a dicotomia norte-sul no padrão de comunicação observado na Áustria. De acordo com os autores, esta divisão tem alguma correlação com várias diferenças nas geografias física e humana e, a maior parte dos grupos formados é contígua, compactuando com os relatos observados na literatura, que indicam a existência da relação entre distância e interação, em geral, e entre distância e comunicação telefônica em particular.

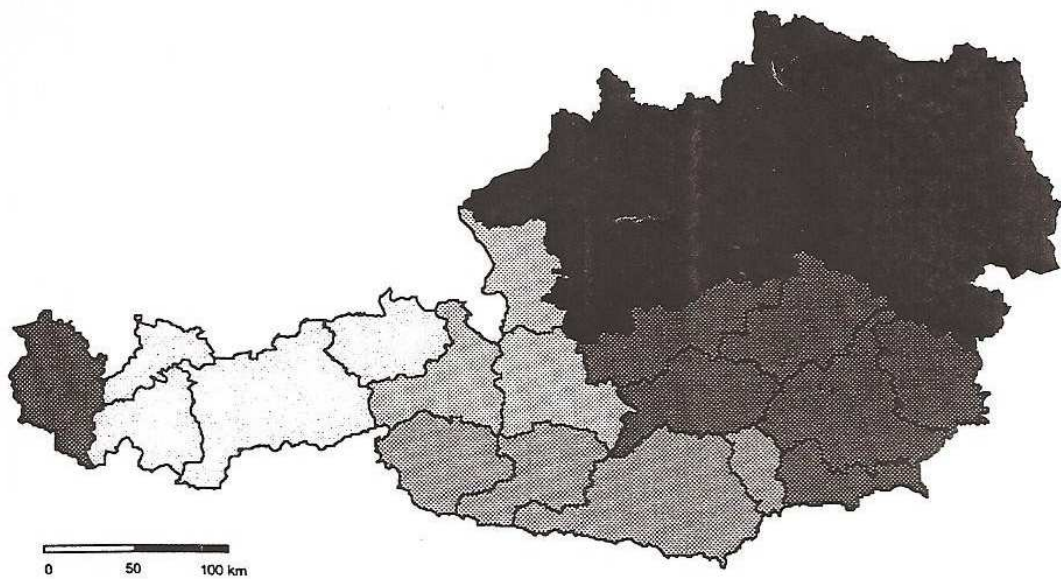
Evidently, geographical distance in general and geographical contiguity in particular is a major factor in understanding communication pattern. (Fischer *et al.*, 1993, p.231)

As configurações espaciais mostram que aplicação do IPFP leva a uma grande região constituída de 21 das 32 unidades espaciais e 5 pequenas regiões localizadas no leste da Áustria, enquanto o resultado obtido pelo Intramax exhibe essencialmente um padrão regional compacto de unidades espaciais, que mostram mais interações entre si que com outras unidades espaciais. Esses resultados são apresentados na Figura 1.

As fronteiras definidas indicam discontinuidades nas intensidades de comunicações telefônicas e apontam para algum tipo de barreira para a comunicação.



(a) Abordagem pelo Método IPFP



(b) Abordagem pelo método INTRAMAX

Figura 1: Padrões espaciais destacados com aplicações dos métodos IPFP e Intramax
Fonte: Fischer *et al.* (1993)

Feldman e outros (2005) apresentam os resultados obtidos na determinação de áreas funcionais com altos níveis de contenção no padrão de viagens para o trabalho, para a região delimitada por Inglaterra, Gales e Escócia. Para tanto, é aplicado o método de agrupamento hierárquico Intramax incorporado no software Flowmap. A análise de tais áreas funcionais provê um estudo de possíveis sistemas de zoneamentos para uso da terra, modelagem de transporte e planejamento.

Os dados usados neste estudo foram extraídos do banco de dados do Censo do Reino Unido de 2001 e consistem de estatísticas de jornadas para o trabalho, resultando em uma matriz de fluxos com pares de origens e destinos.

Para os autores, um fator decisivo para a definição de região funcional é a força de interação entre as unidades espaciais, o que implica na definição de região funcional como uma área ou entidade locacional que tem mais interação ou conexão entre si do que com outras áreas.

No censo, os participantes foram perguntados sobre seu local de trabalho usual, seu principal modo de acesso ao trabalho, dentre outros detalhes sobre o trabalho. Na Escócia, além desses questionamentos, perguntou-se sobre o local de estudo. No total, são aproximadamente 13.000 registros, formando uma matriz de fluxos com pares de origens e destinos.

Ao analisar 376 áreas administrativas na Inglaterra e Gales e comparar seus limites com aqueles determinados pelas áreas funcionais, importantes observações foram feitas. Dentre elas, destaca-se que vários distritos administrativos apresentam fronteiras que se sobrepõem perfeitamente às fronteiras das regiões funcionais e que áreas administrativas urbanas tendem a ser maiores que as áreas urbanas funcionais, enquanto as áreas administrativas rurais tendem a ser menores que as áreas rurais funcionais.

Ainda, é possível identificar uma forte tendência a produzir áreas que consistem de partes de uma área urbana e sua área suburbana/rural adjacente. Os autores consideram este resultado natural, tendo em vista que o Intramax não identifica diferenças entre áreas urbanas e rurais, pois lida apenas com fluxos e números de comutações entre zonas. De acordo com os objetivos do trabalho, deve-se atentar para esta situação, que pode influir na análise do contexto.

São apresentadas regionalizações funcionais em três níveis: nacional, regional ou sub-regional, sugerindo que a escolha do modelo mais adequado deve ficar a critério do gestor ou analista, de acordo com as decisões locacionais a serem tomadas.

Goetgeluk e Jong (2005) mostram como um simples algoritmo para análise espacial e a visualização num Sistema de Informação Geográfica de padrões de migração intermunicipal reabriram as negociações para uma fusão formal, chamada *Holland Rijnland*, de 6 municípios da área urbanizada da região de Leiden e 10 municípios da região rural adjacente Bulbs, localizadas na Randstad Holland (composta por Amsterdam, Rotterdam, The Hague, Utrecht e outras cidades menores). O mercado imobiliário regional era apenas uma questão nas negociações, mas a discussão política sobre isso emperrava todo o processo de fusão.

Os autores usaram dados de migração do *Central Bureau of Statistic* e o método Intramax de agrupamento no programa Flowmap, na busca de respostas a questões como: poderia a nova fusão funcionar como uma região de mercado imobiliário dentro do contexto amplo da *Randstad Holland*? As duas regiões originais interagem com todas ou a interação entre as duas é restrita?

Para responder às questões propostas, utiliza-se a soma de todos os movimentos dentro e entre todos os municípios da Holanda, para o período de 1999 a 2002, desconsiderando os migrantes que chegam ou saem da Holanda. A unidade de referência é a pessoa, que foi classificada em vários grupos etários, em anos: 0 a 17, 18 a 29, 30 a 44, 45 a 64 e 65 ou mais.

Estes dados foram dispostos em uma matriz quadrada, com a diagonal indicando fluxos intramunicipais, com os triângulos acima e abaixo da diagonal refletindo a migração e a emigração entre pares de municípios. As somas dos elementos das linhas e das colunas resultam no total de entradas e saídas por municípios.

Os resultados mostram que *Holland Rijnland* realmente é uma região de mercado imobiliário regional, o que levou os debatedores a continuar o processo de fusão. É importante ressaltar que análises como essa são estáticas, relativas a um determinado período de tempo, podendo sofrer alterações de acordo com políticas de habitação adotadas.

O dendrograma da Figura 2 mostra o resultado da análise. Observa-se que a migração interna na área de estudo (evidenciada no retângulo), é suficientemente

forte para determinar-se como uma região em separado. A análise não suporta a idéia de subdividir esta área.

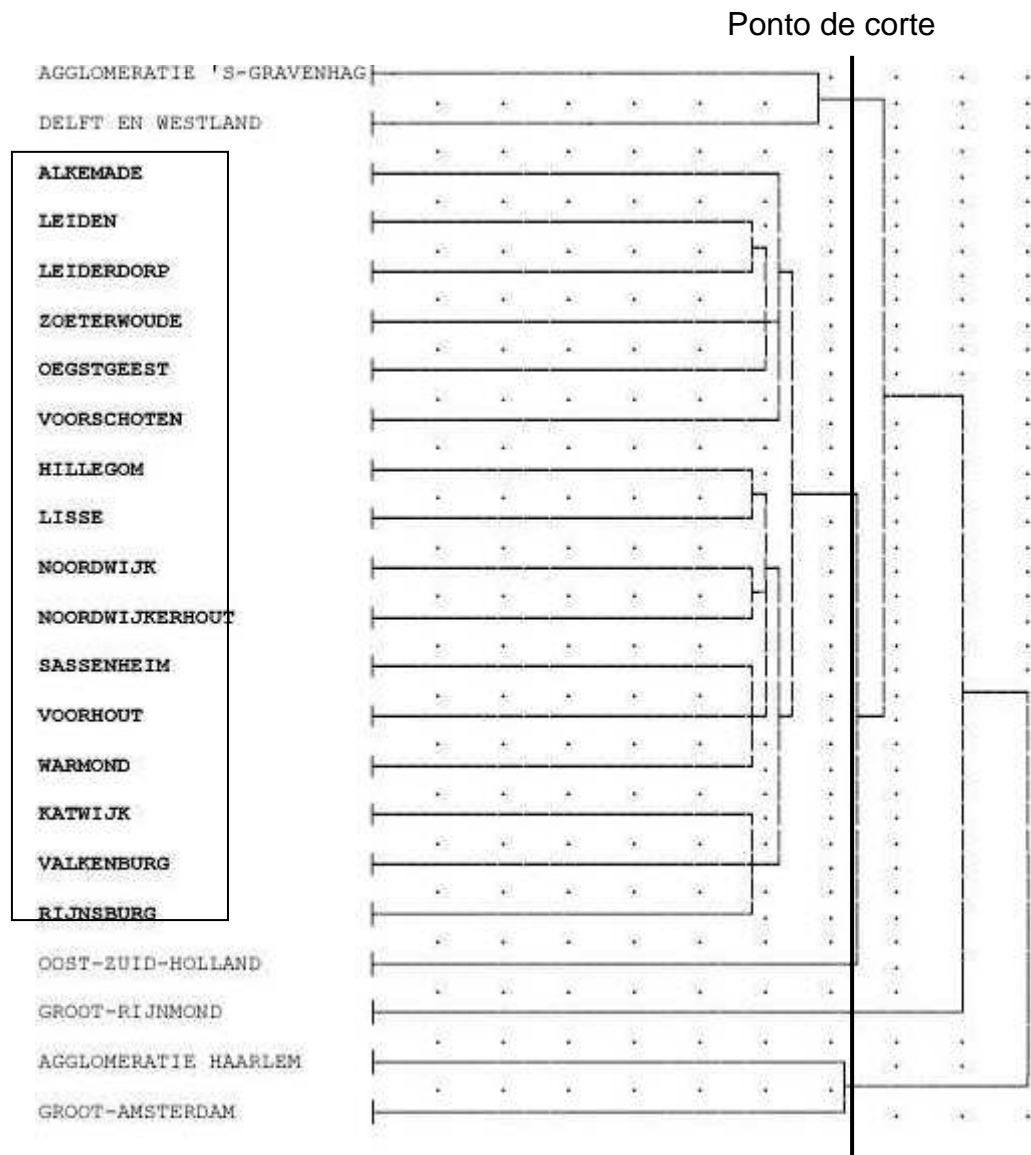


Figura 2: Dendrograma de agrupamento baseado no padrão de migração

Fonte: GOETJELUK; JONG (2005, p.8)

O resultado sugerido pelo dendrograma foi mapeado, conforme Figura 3.

Esta análise foi feita para a população em sua totalidade, mas, quando repetida para diferentes grupos etários o resultado é, de acordo com os autores, surpreendentemente estável, reforçando que a área em estudo é uma região funcional bem definida.

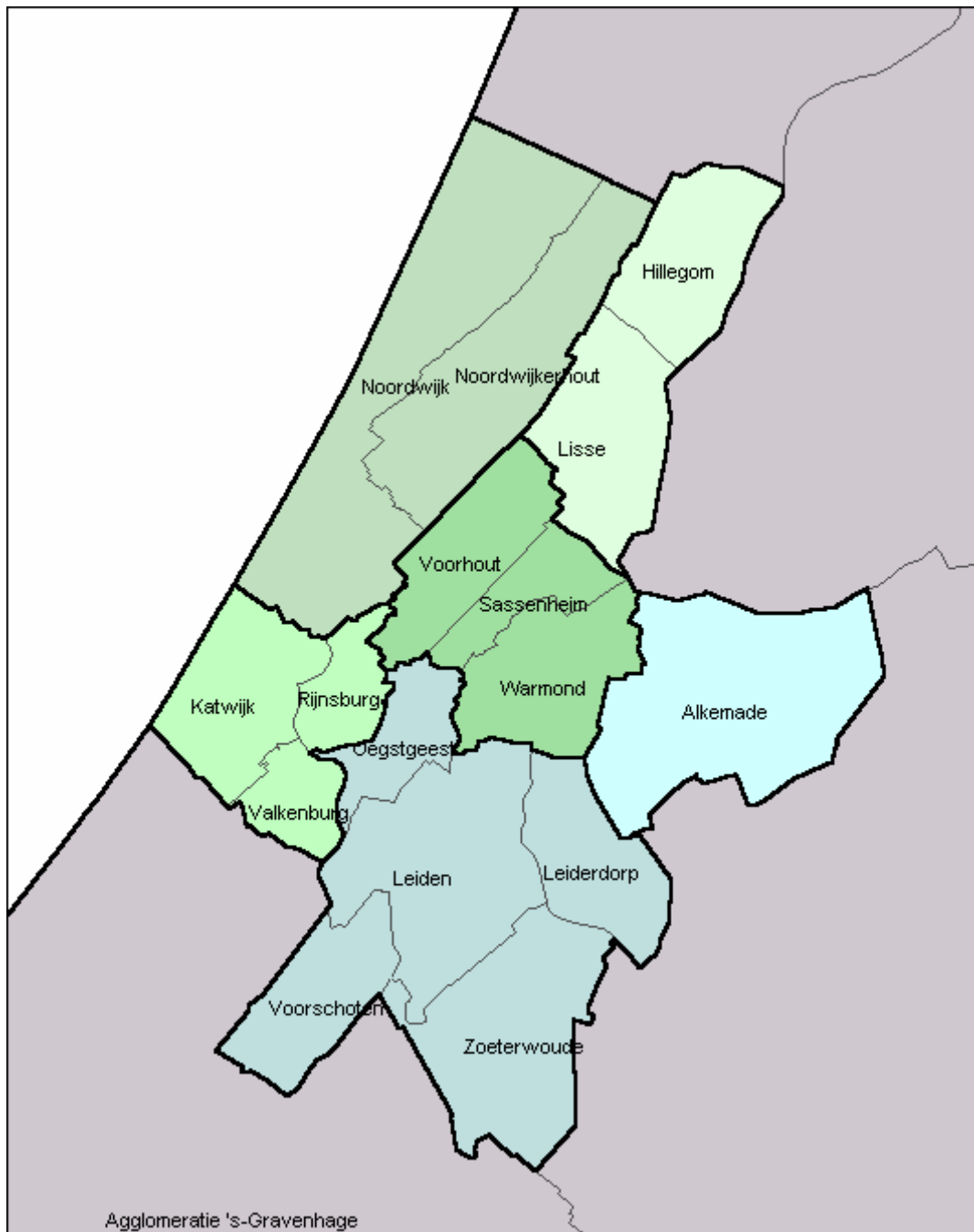


Figura 3: Divisão espacial após agrupamento

Fonte: GOETJELUK; JONG (2005, p.9)

Em outro trabalho de Goetgeluk e Jong (2007), observa-se nova aplicação do método Intramax, agora na identificação de área de mercado imobiliário, área esta definida como a área total de busca, na qual o consumidor encontra vagas sem perda de emprego, família ou amigos.

Os autores avaliam o relacionamento entre distância, idade e motivo para mudança, sendo observado que, por ano, aproximadamente 10% das famílias, ou outros arranjos domésticos, mudam de residência, anualmente.

Na aplicação do método Intramax, o processo foi interrompido quando aproximadamente 75% dos fluxos haviam sido computado nos dois períodos: 1990 e 2002. Este ponto coincide justamente com o momento em que ocorre um aumento abrupto no agrupamento dos fluxos, o que significa a união de duas regiões não muito homogêneas.

Assim, para 1990, com 75,55% dos fluxos, as 495 unidades espaciais foram agrupadas em 28 regiões de mercado imobiliário, enquanto para 2002 são obtidas 25 regiões.

Na Figura 4, em cores estão as regionalizações oficiais e de mercado de habitação. As linhas vermelhas indicam os limites das províncias enquanto as linhas pretas mais finas referem-se aos municípios. As linhas mais espessas referem-se aos resultados obtidos pela aplicação do Intramax. Observa-se que as regiões oficiais, em cores, não refletem o comportamento de escolha da habitação. Com isto, os autores sugerem uma regionalização considerada mais adequada, sendo esta considerada mais adequada que outras divisões obtidas em critérios que não os relacionados com distâncias funcionais.

Embora os casos ora apresentados tenham sido abordados de modo superficial, acredita-se ter mostrado as principais características do Método Intramax, sua superioridade sobre o método IPFP e a aplicabilidade no estudo da migração.



Figura 4: Comparação da regionalização funcional para o mercado imobiliário/2002, em linhas pretas espessas, com delimitação oficial em cores

Fonte: GOETJELUK; JONG (2007, p.11)

No Apêndice A, a matemática envolvida no Método Intramax é discutida com maior profundidade, além dos parâmetros necessários para sua aplicação e os critérios de parada mais comumente adotados. Também é feita a apresentação de um software dedicado a análises geográficas, o Flowmap, que tem o método em questão implementado, tornando-se um ótima ferramenta de apoio apesar da baixa qualidade da cartografia produzida internamente.

3.2 Estudos de fluxos a partir da Análise Fatorial de Ordem Superior

A Análise Fatorial, desenvolvida originalmente como um método para redução da complexidade presente em um grande conjunto de variáveis, apresenta-se, também, como uma alternativa metodológica auxiliar para identificação de padrões em dados de interação espacial como, por exemplo, aqueles observados em problemas que envolvem redes de transporte aéreo, trocas comerciais, fluxos dos mais diversos

tipos (de pessoas, dinheiro, comunicação, etc.), hierarquizações e agrupamento de unidades espaciais de acordo com seus potenciais, etc.

Faz-se, a seguir, a apresentação de alguns casos relacionados com o estudo de fluxos migratórios, procurando, com isso, apresentar as principais características e limitações da técnica de Análise de Componentes Principais no estudo da migração.

Clayton (1977b) apresenta o uso da técnica de Análise de Componentes Principais como um meio de reduzir a dimensão do conjunto de dados através da análise de sua estrutura e como um procedimento classificatório. Em seu estudo, a aplicação da Análise de Componentes Principais o auxilia a identificar diferentes dimensões da variação contida na matriz de interação (de fluxos) e determina de que maneira os elementos do sistema de migração se agrupam em função de suas interações.

Nesta abordagem, as unidades com mais altos *scores* em qualquer componente são vistos como lugares centrais, ou maiores destinos, quando se dispõem os destinos nas linhas da matriz de interação. Portanto, ao investigar os *loadings* e *scores*, Clayton busca a identificação de grupos interdependentes.

Ao aplicar a Análise de Componentes Principais no modo R, em uma matriz origem-destino quadrada de ordem 49, mas não simétrica, para os estados americanos da região oeste do país, Clayton (1977b) extrai 14 dimensões (vetores ou componentes) independentes, com autovalores maiores que 1, que acumularam 63% da variância total da matriz original e mapeia os resultados obtidos com a regionalização sugerida, identificando os estados mais importantes que agem como origem de migrantes e relaciona-os com outros conjuntos de estados que recebem estes migrantes. Os tamanhos dos agrupamentos variam de 2 a 8 estados cada. O padrão ilustrado cartograficamente indica que o país, Estados Unidos, é composto por um conjunto de áreas de migração espacialmente independentes.

Ainda neste trabalho, o autor trabalha com a matriz transposta daquela adotada na primeira abordagem, modo Q da Análise de Componentes Principais, e analisa o padrão encontrado no estudo dos principais destinos.

Em outro trabalho, Clayton (1982) usa a Análise Fatorial de ordem superior em grandes matrizes de interação. Esta técnica permite que a complexidade da tabela de interação seja reduzida sucessivamente com a aplicação de alguns passos hierárquicos. Com isto, é possível observar quais unidades de área de um determinado nível estão ligadas com grupos de níveis mais altos. Neste trabalho, o autor usa a rotação oblíqua sobre as componentes obtidas com o intuito de realçar as correlações existentes.

Para Clayton (1982,p.1), a estrutura de sistemas geográficos humanos é frequentemente de natureza hierárquica. Sistemas de migração populacional podem, de modo útil, serem concebidos como uma série de níveis de áreas de migração hierarquicamente relacionados. As áreas de um nível podem alojar-se dentro de áreas do próximo nível, mais alto. Tais áreas de migração e os relacionamentos entre diferentes níveis podem ser extraídos das grandes matrizes origem-destino de migração populacional, com o auxílio da Análise Fatorial de ordem superior.

As matrizes de interação, neste caso tabelas de fluxos migratórios, fornecem informações indispensáveis, concebendo ligações e relacionamentos que existem entre unidades de área em um sistema espacial. Mas, em geral, estes sistemas são muitos grandes e faz-se necessário reduzir a complexidade das matrizes envolvidas, mas, retendo a estrutura básica de relacionamentos interna ao sistema. Para Clayton, isso pode ser feito com a aplicação da Análise Fatorial de ordem superior que permite reduzir sucessivamente a complexidade da matriz de dados.

Ao estudar o padrão de migração nos Estados Unidos no período 1955-1960, partindo de uma matriz quadrada de ordem 72, Clayton (1982) extraiu treze fatores e, assim, com auxílio da representação cartográfica, definiu treze grupos mutuamente excludentes de unidades de área, que constituíam regiões funcionais, ou seja, que estavam relacionadas de modo similar com as outras unidades do mesmo grupo. Neste caso, os treze fatores captaram 78% da variância contida na matriz original de fluxos. Vale ressaltar que a identificação obtida é funcional, sugerindo uma tipologia de migração.

Ao impor outra aplicação fatorial à matriz de correlação determinada pelos treze fatores iniciais, foram extraídos quatro fatores secundários que captam 52% da variância dos dados. Com isto, identificaram-se regiões mais fortes na estrutura de fluxos.

Este processo de obtenção de outras ordens de fatores secundários pode ser continuado enquanto se fizer necessário e possibilita a identificação de níveis superiores da estrutura de migração, favorecendo a identificação de unidades de área como receptoras ou perdedoras de população ou mesmo sua função dentro do processo migratório como ponto intermediário para futura migração, área de retorno, etc.

Uma questão levantada pelo autor refere-se ao relacionamento existente entre diferentes níveis de fatores. Cattell, citado por Clayton (1982,p.13), sustenta que os relacionamentos entre os fatores de diferentes níveis podem ser descritos como uma rede e que a estrutura desta rede pode ser, mas não necessariamente, de natureza hierárquica. Uma estrutura hierárquica pode ser imposta na rede fazendo certas suposições arbitrárias, tais como escolher para examinar apenas as maiores cargas fatoriais de cada fator. Esta foi a abordagem adotada neste estudo.

Holsman (1980) apresenta um trabalho de abordagem geográfica com aplicação da análise fatorial de ordem superior sobre redes de transporte. O autor manifesta a consideração de que a interpretação dos fatores obtidos depende da natureza do conjunto original de dados e do problema que se tem em mãos.

Para Cattell (1965), citado por Holsman (1980,p.192), soluções com fatores de alta ordem resultam da fatoração da matriz de correlação, de onde os fatores são extraídos via Análise de Componentes Principais e uma rotação oblíqua é aplicada.

Nesse trabalho, Holsman trabalha com dados de uma matriz de “caminho mais curto” da rede aérea australiana, composta de 162 localidades. Este “caminho mais curto” mede a distância entre duas localidades em termos de números de pontos intermediários (escalas ou conexões) entre eles.

Ao adotar a interpretação de que os vetores, representados pelas linhas e colunas da matriz, podem ser interpretados como indicadores do perfil de acessibilidade de um centro a outros e dos outros centros a este, o que se obtém com a aplicação da Análise de Componentes Principais é uma regionalização baseada na similaridade de acessibilidades ou de perfis de destinos.

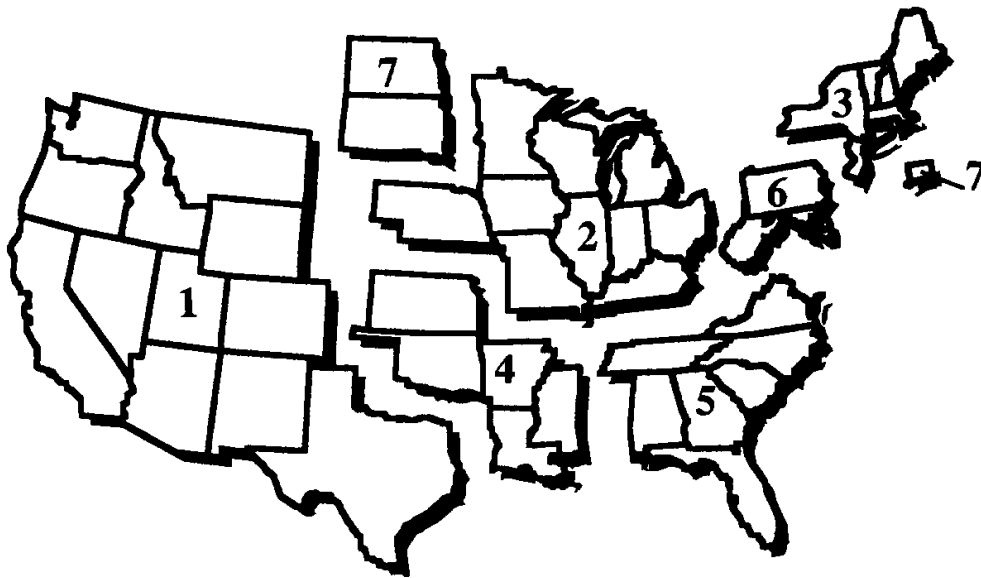
No primeiro estágio da análise, o de primeira ordem, foram identificados seis agrupamentos regionais, com retenção de 72% da variância total, em que foram identificados os seis centros mais acessíveis, seja como ponto final da viagem ou como ponto intermediário dentro do percurso.

O autor discute os resultados obtidos, inclusive com o apoio da representação cartográfica, chegando até a aplicação da análise fatorial de segunda ordem. O uso da análise fatorial de alta ordem permitiu uma comparação entre estrutura de fluxos do transporte aéreo e a rede aérea, servindo de suporte para os gestores das empresas aéreas que buscam analisar e se estruturar frente a demanda de serviços para as rotas de interesse.

Clark, Davies e Johnston (1974) fazem um estudo sobre a aplicação da Análise Fatorial na Geografia Humana, mostrando uma grande variedade de usos. Preocupa-se, nesse trabalho, em discutir algumas dificuldades encontradas ao adotar esta técnica e possíveis soluções para os problemas.

São apresentados alguns exemplos relacionados com as áreas da Geografia Urbana e da Geografia Regional, chegando à conclusão de que a análise fatorial pode ser adotada como um caminho para resolver problemas enfrentados pelos geógrafos ao lidar com muitas variáveis ou várias unidades espaciais.

Ellis, Barff e Renard (1993) examinaram as tipologias da migração para diferentes grupos ocupacionais, usando os microdados de 1983, publicados pelo Censo dos Estados Unidos da América. Os resultados (ver Figura 5) confirmam análises anteriores, que apontavam diferenças de comportamento em movimentos migratórios, de acordo com as ocupações dos migrantes e das indústrias.



COMPONENT	EIGENVALUE	PERCENT OF VARIANCE
1	22 147	45 2
2	6 693	13 7
3	4 576	9 3
4	2 623	5 4
5	2 521	5 1
6	1 499	3 1
7	1 177	2 4
total		84 7

Figura 5: Regiões de migração para profissionais e gestores de manufaturados, baseado no padrão de emigração, 1975-80

Fonte: Ellis, Barff e Renard 1993, p.176

Pandit (1994), na análise de regiões de migração nos Estados Unidos nos períodos de 1940, 1960 e 1980, aplica a Análise de Componentes Principais nos modos R e Q. As duas abordagens são importantes, tendo em vista que se estuda a tipologia de migração para os imigrantes e para os emigrantes, respectivamente.

No modo R, ao lidar com dados dos imigrantes, foram obtidos 11 componentes significativas para o ano de 1940 e 10 para 1960 e 1980. Cada componente representa uma tipologia de migração distinta, identificada pelo conjunto de estados com cargas fatoriais mais altas em cada fator, o que sugere serem destinos de migrantes de origens similares. Ao aplicar o modo Q, são obtidas 10 componentes/tipologias para 1940 e nove para 1960 e 1980. Neste caso, são identificados estados que enviam migrantes para outro conjunto de estados comuns.

Os resultados são mapeados, ver Figura 6, e revelam uma estabilidade temporal, principalmente na porção leste do país, identificando que as regiões de New England, North Atlantic (Penssylvania, New York, Delaware), Midwest (Michigan, Indiana, Ohio, Kentucky) e Southeast constituem tipologias persistentes de migração.

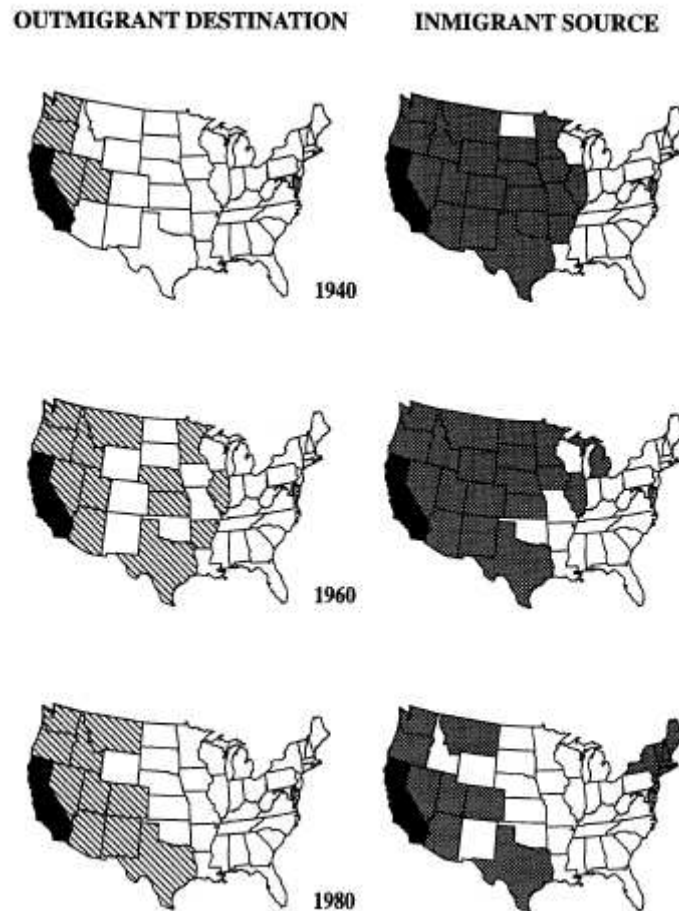


Figura 6: Áreas de migração para Califórnia nos períodos 1940, 1960 e 1980
 Fonte: PANDIT, 1994, p.342

É possível observar que a técnica permite inferir que, apesar de o tamanho e a direção dos fluxos terem sofrido alterações consideráveis entre 1940 e 1980, conforme pode ser visto na Figura 7, os estados dentro de cada tipologia foram similarmente influenciados pelas mudanças.

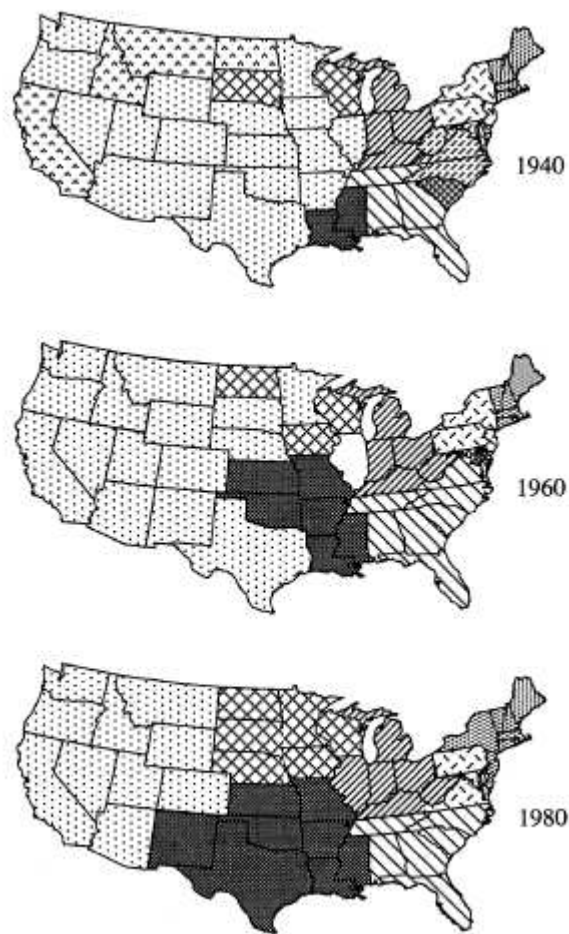


Figura 7: Tipologia para imigrantes nos EUA/1940-1980. Os estados não hachurados são aqueles que não estão fortemente orientados para nenhuma das tipologias

Fonte: PANDIT,1994,p.341

Pandit cita outro estudo sobre migração feito com o uso da Análise de Componentes Principais, de Plane e Isserman (1983). Este trabalho, sobre os dados migratórios entre estados dos EUA para os períodos 1960-65, 1965-70 e 1970-75, sugere a existência de um conjunto de subsistemas de migração notavelmente estável, apesar das mudanças observadas nos padrões migratórios de áreas não metropolitanas para metropolitanas e do sul para o norte.

Em verdade, os autores trabalham a identificação de subsistemas de migração interestadual da força de trabalho nos Estados Unidos. Outra conclusão extraída com aplicação da técnica é a importância da curta distância na redistribuição da força de trabalho.

Apesar da superficialidade com que os casos foram aqui abordados, espera-se ter mostrado a aplicabilidade, a viabilidade e algumas limitações da Análise de Componentes Principais no estudo da migração.

No Apêndice B, faz-se a discussão dos fundamentos matemáticos envolvidos na Análise Fatorial e, por conseqüência, na Análise de Componentes Principais. Nas Seções 3.1 a 3.3 avalia-se o uso da quantificação na Geografia, abordando especificamente o uso de métodos multivariados.

4 MÉTODOS QUANTITATIVOS PARA ANÁLISE GEOGRÁFICA DE FLUXOS MIGRATÓRIOS

4.1 A importância e a evolução dos métodos quantitativos para os estudos geográficos

As primeiras aplicações da quantificação, em Geografia, podem ser encontradas em trabalhos dos séculos passados, porém, sua utilização assumiu maior importância e foi intensificada a partir da década de 1950, com a denominada *Revolução Teorético-quantitativa* (GERARDI, 2008), quando os mapas passam a apresentar, também, uma função analítica. Com a rápida evolução dos recursos tecnológicos a Análise Espacial recebeu e vem recebendo novo impulso, o que implica na necessidade de maiores conhecimentos de métodos quantitativos.

Em 1970, Cole já afirmava que a *Revolução Teorético-quantitativa* na Geografia havia passado e a aplicação dos recursos de quantificação já era uma realidade tanto na Geografia como em todas as áreas de conhecimento.

Desde a Segunda Guerra Mundial, quase todos os países do mundo passaram por uma explosão de informação. A introdução de computadores, equipamentos microeletrônicos, e serviços de telecomunicação abriram o caminho para uma avalanche de informações, não apenas para a pesquisa científica, mas também para uso de um público mais amplo e por aqueles que definem planos e políticas. (ABREU, 1995)

Várias ciências lançam mão dos instrumentos quantitativos, porém, a Geografia pede uma atenção diferenciada, pois lida, também, com a variável espacial, o que leva a um tratamento particular dos dados na busca da relação entre os fenômenos estudados e sua distribuição no espaço.

Nas ciências que se ressentem da falta de um campo experimental, há inúmeros modelos que pretendem interpretar a realidade. Em muitos casos a associação das técnicas estatísticas com as informações disponíveis não é suficiente para se decidir em favor de uma tese ou de sua competidora. A probidade científica exige, no caso, que se diga que não há informações capazes de rejeitar qualquer das teses. O avanço extra científico consiste em exibir estatísticas como base de comprovação desta ou daquela hipótese. (SIMONSEN, 1983)

O grande volume de dados trabalhado pelo geógrafo precisa ser organizado e apresentado de uma maneira sintética, fácil de ser manipulada e compreendida. Logo, o uso da quantificação em Geografia é muito influenciado pela aplicação de técnicas estatísticas.

A quantificação se apresenta como um meio auxiliar para que a pesquisa geográfica esteja baseada em formulações seguras do problema concreto que pretende estudar, tendo fundamentação teórica consistente e compreensível, como um instrumento promotor da interdisciplinaridade e promovendo o enriquecimento teórico-conceitual com suas formulações.

Para O'Brien (1992), a preocupação com a quantificação é relevante para o geógrafo, porque a Geografia é uma disciplina altamente dependente de dados. Uma grande parte da pesquisa geográfica está baseada na coleta, análise e interpretação de dados.

A abordagem quantitativa, como técnica complementar na explicação e descrição de diversos fenômenos de caráter geográfico, tem sido bastante aceita e freqüente. Isso pode ser visto com facilidade nas dissertações e teses aprovadas no Programa de Pós-Graduação em Geografia - Tratamento da Informação Espacial da PUC MG, e em diversas outras referências bibliográficas, como as citadas ao final deste trabalho.

Compreender a distribuição espacial de dados oriundos de fenômenos ocorridos no espaço constitui um grande desafio para a elucidação de questões centrais em diversas áreas do conhecimento ... (CÂMARA et al, 2004).

4.2 Dado Espacial

As particularidades dos dados espaciais pedem o desenvolvimento ou adaptação de um conjunto de técnicas voltadas para as necessidades específicas da Análise Espacial, que, de acordo com Camargo (1998), é o estudo quantitativo de fenômenos que são localizados no espaço. Utiliza-se a expressão análise de dados espaciais em oposição à análise de dados em geral, quando as técnicas utilizadas consideram explicitamente a localização espacial.

Neste trabalho, é dada atenção especial ao uso da Estatística Multivariada no estudo da migração populacional, pela aplicação da Análise Fatorial via Componentes Principais, e do método iterativo Intramax .

Cada migrante está ligado a pelo menos um ponto de origem e a um ponto de destino que podem ser determinados em um sistema de coordenadas, por exemplo, de longitudes e latitudes. Também é possível avaliar a migração em termos da distância coberta pelo migrante.

Quando observamos que a Microrregião de Belo Horizonte recebeu 184.818 migrantes das outras microrregiões de Minas Gerais entre 1995 e 2000, deve-se considerar que cada migrante traz consigo a componente espacial. Vários migrantes apresentam origens iguais ou de microrregiões contíguas; outros se assemelham por saírem de microrregiões próximas a Belo Horizonte. Assim, ao pensar em termos de unidades espaciais, pode-se identificar microrregiões que se comportam de modo semelhante por apresentarem destinos comuns para seus emigrantes ou por receberem migrantes das mesmas microrregiões.

Partindo do princípio de que existe o interesse em analisar e sintetizar as diferenças e similaridades entre unidades espaciais, neste caso microrregiões, e interpretar a organização espacial quanto ao fluxo migratório, deve-se levar em consideração que se trabalhará com dados espaciais.

Para tanto, toma-se como definição de dado espacial aquele cujos elementos ficam definidos pela localização espacial, fornecida pelas componentes x e y, relacionadas a algum sistema de coordenadas, e pela componente z, denominada atributo, que pode representar saldo migratório, número de imigrantes ou emigrantes, etc.

Os dados espaciais trazem consigo a peculiaridade da dependência espacial que, segundo Anselin (1989), representa a tendência de a vizinhança influenciar uma determinada localização e possuir atributos semelhantes.

4.3 Métodos multivariados e dados geográficos

Na geografia, o estudo da Interação Espacial é de grande importância. A distância entre centros, por exemplo, consumidores e produtores, receptores ou expulsos de população, nem sempre pode ser medida apenas metricamente, pois a interação entre os centros dependerá, além de outros fatores, da forma como esta pode ocorrer. Por exemplo, no caso da existência de dois centros, A e B, e de outro centro, C, equidistante dos dois primeiros, as facilidades de acesso podem fazer com que o emigrante de C opte por aquele com melhores condições de acessibilidade. O potencial de atração de cada centro, de acordo com os serviços oferecidos, será determinante na escolha do emigrante em C. Em certos casos, mesmo tendo que enfrentar maiores distâncias, o migrante pode optar por um centro que ofereça melhores condições ou oportunidades.

Segundo Fotheringham e Haynes (1982,p.5), o termo Interação Espacial é bastante amplo, e pode ser considerado como qualquer movimento no espaço provocado por relações humanas. Isto significa dizer que relações como migração, movimentos de ida ao trabalho ou a um hospital, transmissão de dados ou conhecimentos, acesso a serviços públicos ou privados, e vários outros, podem ser tratados por modelos de Interação Espacial.

Pode-se encontrar na bibliografia diversos estudos que ajudam a compreender como funciona a Interação Espacial dentro de uma região ou entre diversas regiões, como índices de autocorrelação espacial, teorias da localização, estatística espacial, dentre outros, além das técnicas de análise multivariada aplicadas neste estudo. Exemplos podem ser vistos em Fotheringham e Haynes (1982), Fotheringham e O'Keily (1988), Clark(1973), Druck e outros (2004), etc.

Fenômenos que sofrem influência, ou que estão relacionados com um grande número de variáveis sempre foram de difícil abordagem pelas dificuldades operacionais inerentes. Estas dificuldades foram sendo superadas, dados os avanços registrados na tecnologia computacional, permitindo que a Estatística incorporasse recursos importantes no tratamento e análise de dados. Uma de suas

áreas que mais se beneficiaram é a da Estatística Multivariada. Atualmente, a complexidade operacional não é um empecilho para sua aplicação.

De acordo com Hair e outros (2005a), a Estatística Multivariada pode ser vista como um conjunto de métodos estatísticos aplicáveis em situações cuja análise envolve um grande número de variáveis aleatórias e inter-relacionadas. Em uma rápida consulta às fontes bibliográficas, pode-se observar que as mais diversas áreas do conhecimento humano têm pedido suporte de técnicas que consigam lidar com esse tipo de situações.

A necessidade de trabalho com dados psicológicos e do comportamento, provavelmente, foi o que provocou o desenvolvimento e/ou surgimento das técnicas estatísticas para análise de conjuntos de dados mais amplos e complexos.

Os métodos de análise multivariada predominarão no futuro e resultarão em drásticas mudanças na maneira como profissionais de pesquisa pensam em problemas e planejam sua pesquisa. Esses métodos tornam possível levantar questões específicas e precisas de considerável complexidade em cenários naturais. Isso viabiliza a condução de pesquisas teoricamente importantes e a avaliação dos efeitos de variações paramétricas que naturalmente ocorrem no contexto em que elas normalmente aparecem. Dessa maneira, as correlações naturais entre as múltiplas influências de comportamento podem ser preservadas e os efeitos separados dessas influências, estudados estatisticamente sem causar um isolamento comum de qualquer indivíduo ou variável.

(Hardyck e Petrinovich apud Hair et al.,2005a)

Atualmente, pode-se encontrar referências ao uso da Estatística Multivariada em diversos trabalhos nas áreas de Geografia, Demografia, Ecologia, Economia, Geologia, Educação, Ergonomia, Engenharia, etc.

Uma das grandes vantagens das técnicas multivariadas é sua capacidade de sintetizar as informações disponíveis em um grande número de variáveis em um conjunto menor de variáveis, sem grandes perdas para análises.

A Estatística Multivariada se divide em técnicas exploratórias e técnicas de inferência estatística. Neste trabalho, aborda-se a Análise Fatorial, via Componentes Principais, que se encontra no primeiro grupo. Os softwares estatísticos mais importantes apresentam rotinas específicas da Estatística Multivariada, que tornam possível ao pesquisador o trabalho com grandes quantidades de variáveis e enormes bancos de dados.

4.4 Proposta metodológica

4.4.1 Introdução

Como visto no Capítulo 3, a Análise Fatorial de Ordem Superior, com extração de fatores via Análise de Componentes Principais, e o Método Intramax são aplicados na identificação de tipologias e subsistemas de migração, respectivamente.

Ao adotar o método da Análise Fatorial de Ordem Superior, considera-se cada unidade espacial uma variável. Assim, na matriz origem-destino, cada coluna, ou seja cada destino, é tratado como uma variável e as linhas, as origens, são os casos. Estudando a variabilidade presente em cada variável, promove-se o agrupamento de unidades espaciais que se comportam como destinos semelhantes, e identificam-se as principais origens para cada um desses agrupamentos.

Ao lidar com a matriz transposta, identificam-se as unidades espaciais que são origens com comportamentos semelhantes e os principais destinos para cada agrupamento sugerido.

Pela própria formulação matemática, discutida no Apêndice A, observa-se que o método Intramax aborda o problema de regiões de migração de modo distinto ao da Análise Fatorial de Ordem Superior. Nesse caso, são agrupadas as unidades espaciais que interagem com maior intensidade proporcional.

O que se faz nas seções seguintes é a apresentação de propostas metodológicas para identificação de tipologias de migração e de subsistemas de migração, tomando como unidades espaciais fundamentais as 558 microrregiões brasileiras

definidas pelo IBGE. Os dados relativos à migração utilizados na análise empírica foram obtidos dos microdados dos Censos de 1991 e 2000. Considera-se o conceito de migrante de data-fixa, ou seja, aquele indivíduo que apresenta município de residência em 1986 ou 1995 diferente daquele registrado em 1991 ou 2000, respectivamente. Trata-se, portanto, de migrantes sobreviventes dos respectivos períodos, independentemente de terem ou não realizado movimentos migratórios intermediários, dentro do quinquênio em questão.

Os conceitos de subsistema de migração e tipologia de migração adotados são aqueles apresentados por Pandit(1994,p.1) e indicados no Capítulo 1 deste trabalho:

- i. Regiões de migração como grupos de unidades territoriais que estão altamente conectadas com outras, denominadas **subsistemas de migração**.
- ii. Regiões de migração como grupo de unidades territoriais que tem padrões similares de trocas com outras unidades, denominadas **tipologias de migração**.

4.4.2 Delimitação de subsistemas de migração com uso do Método Iterativo Intramax

Nesta Seção, faz-se a apresentação do método iterativo Intramax aplicado na identificação de padrões em dados de interações espaciais e do software Flowmap, dedicado à análise e apresentação de dados de interação espacial ou de fluxo. Os fundamentos matemáticos são discutidos no Apêndice A.

O dado de interação espacial ou fluxo possui a particularidade de, além de apresentar um atributo ou intensidade de ocorrência do fenômeno, ser formado por mais duas componentes geográficas indicando a origem, onde o fluxo começa, e o destino, onde o fluxo termina. O fluxo ou a interação espacial pode ter como atributo o número de pessoas que migram de um ponto a outro, de ligações telefônicas entre duas localidades, quantidade de dinheiro movimentado entre dois pontos, quantidade de mercadoria transportada de uma localidade a outra, etc.

Neste trabalho, admite-se que a interação entre diferentes unidades espaciais pode ser vista como uma medida de distância funcional, ou seja, aquela não propriamente física. Assim, como consequência, interpreta-se que quanto mais fortes forem as interações, menor será a distância entre as unidades espaciais.

Agrupar regiões que estão próximas em termos de sua distância funcional leva à criação de regiões funcionais, que se tornam alternativas geográficas para divisões administrativas ou análises mais apropriadas do espaço. Pode-se identificar, por exemplo, se os fluxos mais significativos estão relacionados às curtas distâncias absolutas ou se estas não se apresentam como um fator relevante para a opção do migrante.

O método iterativo Intramax, apresentado originalmente por Masser e Brown, em 1975, é estruturado especificamente para revelar padrões em distâncias funcionais, como as implícitas nos dados geográficos de fluxos. Este método é iterativo e de natureza hierárquica.

Segundo Masser e Scheurwater, apud Feldman et al. (2005), *“the objective of Intramax procedure is to maximize the proportion within the group interaction at each stage of the grouping process, while taking account of the variations in the row and the column totals of the matrix”*.

Neste método, as quantidades de interações proporcionais são calculadas para todos os pares de unidades espaciais existentes. Aquelas duas com maior força de relacionamento são agrupadas, formando-se, assim, uma nova unidade espacial. Essas unidades espaciais terão as quantidades de interações combinadas e designadas para a nova unidade criada, assim, a nova unidade substitui as duas unidades originais na próxima etapa. Então, no caso de N unidades espaciais, após (N-1) etapas, todas as unidades espaciais estarão agrupadas em uma única unidade e todas as interações serão internas.

A função objetiva do método Intramax leva a agrupamentos de unidades espaciais que apresentam mais interações entre si do que com outras unidades espaciais.

A matriz de interações espaciais deve ser quadrada, as somas dos totais das linhas e colunas iguais e os elementos da diagonal representam as migrações internas. Assim, neste trabalho, os valores contidos na diagonal da matriz origem-destino indicam os fluxos que ocorreram de municípios da microrregião para outros dentro da mesma microrregião.

Ao adotar as somas das linhas e colunas, observa-se uma forte relação entre a fórmula característica do método Intramax com as medidas utilizadas no teste Qui-Quadrado, adotado para testar a associação entre duas variáveis

Para identificação de subsistemas de migração, ou seja, regiões formadas por microrregiões que apresentam forte intensidade de interação migratória, faz-se a proposta de aplicação do método iterativo Intramax, especificamente projetado para revelar padrões de interações, podendo ser considerado como uma técnica de cluster.

Este método encontra-se implementado no software Flowmap 7.2. Ao aplicá-lo, dois arquivos são gerados: um indicando qual o agrupamento efetuado em cada um dos passos e outro arquivo contendo o gráfico dendrograma, muito utilizado em técnicas de *cluster* e, neste trabalho, útil para determinação do ponto de interrupção do processo de iteração. Além dos arquivos gerados, o software gera novas colunas no arquivo *dBase* com os dados de fluxos, úteis para o mapeamento dos resultados.

A seguir, algumas questões são tratadas com o intuito de definir os critérios propostos para a aplicação do método Intramax sobre os dados de fluxos migratórios do censo. Discute-se sobre a organização dos dados, a definição do ponto de parada a ser adotado na aplicação desse método iterativo, o tratamento dado à contigüidade e a viabilidade de acompanhamento do processo de agrupamento para alguma unidade espacial de interesse;

Na tentativa de responder tais questões, além das leituras já citadas anteriormente e de outras cujas referências estão listadas ao final do trabalho, houve uma importante troca de emails com Tom de Jong, professor da Utrecht University, um dos idealizadores do software Flowmap, citado anteriormente, e com Roland Goetgeluke,

pesquisador sênior do Research Institute for Housing, Urban and Mobility Studies, da Delft University of Technology, Holanda².

Como pode ser visto no Apêndice A, o objetivo do método Intramax é de maximizar a proporção de interação dentro de um conjunto de unidades espaciais, em cada estágio do processo de agrupamento. Neste método, as quantidades de interações proporcionais são calculadas para todos os pares de unidades espaciais existentes e aquelas duas com maior força de relacionamento são agrupadas, formando-se, assim, uma nova unidade espacial.

Para tanto, trabalha-se com duas matrizes quadradas origem-destino de ordem 558, não-simétrica, sendo uma para o período 1986-1991 e outra para o 1995-2000, contendo os fluxos migratórios entre todas as microrregiões brasileiras, conforme definição do IBGE, em que os elementos das diagonais principais representam os fluxos intramicrorregionais, ou seja, entre municípios de uma mesma microrregião.

De acordo com a formulação matemática própria do modelo, discutida no Apêndice A, ao trabalhar com a matriz transposta, ou seja, destino-origem, os resultados obtidos serão idênticos, pois a função objetiva do método Intramax leva a agrupamentos de unidades espaciais que apresentam mais interações entre si do que com outras unidades espaciais, não importando o sentido dessas interações.

Como explicitado anteriormente, para uma matriz de ordem N o número máximo de passos a executar é $(N-1)$, quando todos os fluxos passam a ser intrazonais. No primeiro passo, duas microrregiões serão unidas, escolhidas por apresentarem a maior intensidade proporcional dentre todas as 155.403 interações³ calculadas pela função objetiva do método. Os fluxos migratórios entre elas se tornam intrazonais e os fluxos dessas microrregiões com outras microrregiões são somados e atribuídos a essa nova unidade espacial, restando, neste momento, 557 unidades espaciais. Ou seja, a matriz origem-destino original é transformada em outra de ordem 557.

² Tais contatos foram motivados pela leitura de alguns trabalhos desses autores, com aplicação do método INTRAMAX, que chamaram a atenção pela qualidade dos resultados obtidos e por ambos utilizarem o software Flowmap.

³ Valor obtido pela combinação de 558 unidades duas a duas, o que representa o número de todas as interações possíveis, excluindo as de interações de uma microrregião com ela mesma.

Continuando o processo iterativo, a função objetiva é aplicada sobre a nova matriz obtida, e uma nova fusão é realizada. Mas, desta vez, pode-se obter o agrupamento de uma microrregião com outra microrregião ou de uma microrregião com o agrupamento feito no primeiro passo. É gerada uma nova matriz de ordem 556.

Nos passos seguintes, pode-se observar a fusão de uma microrregião com outra microrregião, de uma microrregião com um agrupamento de microrregiões ou de um agrupamento com outro agrupamento. O número máximo de passos como esses, no caso da matriz de ordem 558, é igual a 557, quando todo o espaço é coberto por um único agrupamento, ou seja, o Brasil como uma única unidade espacial.

De acordo com Slater, citado por Fischer e outros (1993,p.231), a entrada tardia de uma unidade espacial simples (neste caso, uma microrregião) no processo de agregação deve ser interpretada em termos de unidade que tem uma identidade mais nacional do que regional.

Isso reforça a interpretação de uma propriedade discutida no Apêndice A, sugerindo que nos passos iniciais devem ser observados agrupamentos das menores unidades espaciais (ou seja, com menor representatividade sobre o total de fluxos), mas com fortes ligações; na seqüência pequenas unidades espaciais são unidas com unidades maiores e, nos passos finais, observam-se uniões de grandes áreas com outras grandes áreas.

Neste ponto, cabe a discussão sobre a determinação do ponto de parada do método iterativo, pois, não existe uma propriedade matemática que indique o momento ideal para interrupção do processo. Esta decisão dependerá muito dos objetivos e das análises parciais feitas pelo pesquisador sobre os agrupamentos sugeridos.

O que se observa é que esta decisão é tomada empiricamente e o ponto, como sugerido por Goetgeluk (2006,p.11), é definido como aquele antes de grandes aumentos nos fluxos intrazonais, o que implicaria no agrupamento de duas unidades espaciais que guardam menor homogeneidade no padrão dos fluxos e em significativa perda de detalhes. Neste mesmo trabalho, Goetgeluk interrompeu o processo de agrupamento quando 75,55% dos fluxos estavam computados e,

portanto, se tornaram intrazonais, para os dados de 2002, enquanto, para 1990, o ponto de parada ocorreu quando foram atingidos 73,47%.

Brown e Pitfield (1990) destacam que este ponto de parada varia de acordo com a proposta de análise e a natureza dos dados e vários critérios podem ser aplicados na determinação deste número, como, por exemplo, com a identificação de um ponto de descontinuidade no aumento progressivo de fluxos intrazonais no procedimento de agregação.

Um recurso auxiliar adequado para a determinação do ponto de parada do processo iterativo é a representação gráfica denominada dendrograma, veja Figura 8. Este é um gráfico do tipo árvore, onde cada passo de um processo hierárquico de *clustering* é representado como uma fusão de “galhos” de uma “árvore” em outro maior. Nele, é possível visualizar bidimensionalmente a similaridade ou dissimilaridade de todo o conjunto. O corte no dendrograma deve ser efetuado onde houver uma alteração brusca no ritmo de agrupamentos.

Quanto à contigüidade, em momento algum a função objetiva do método Intramax impõe exigências ou restrições. Ou seja, os agrupamentos podem ser realizados entre microrregiões contíguas ou não.

Nesta proposta, sugere-se a não imposição da contigüidade para definição dos agrupamentos, para que estes sejam determinados exclusivamente pela força de interação entre as unidades espaciais. Deve-se ressaltar que os dados de fluxos migratórios trazem, intrinsecamente, a natureza da informação espacial e esta será evidenciada na determinação dos grupos. Caso a contigüidade seja observada como fator recorrente nos agrupamentos, uma importante conclusão deve ser registrada: o fato de unidades espaciais mais próximas interagirem mais intensamente entre si do que com unidades distantes e, conseqüentemente, que as migrações de curta ou média distância estejam predominando.

Jong (2008) relata, em uma das mensagens eletrônicas recebidas, que, pela experiência com dados confiáveis e sistemas não muito pequenos, caso deste trabalho, não é necessário impor restrição de contigüidade.

INTRAMAX ANALYSIS by Flowmap 7.1

Origin data from: D:\DATA\zeedata\zeeland1.dbf
 Destination data from: D:\DATA\zeedata\zeeland1.dbf
 Flow data from: D:\DATA\zeedata\Migr9502.dbf

Total interaction: 36238
 Intrazonal interaction: 28099
 Percentage intrazonal: 77.54%

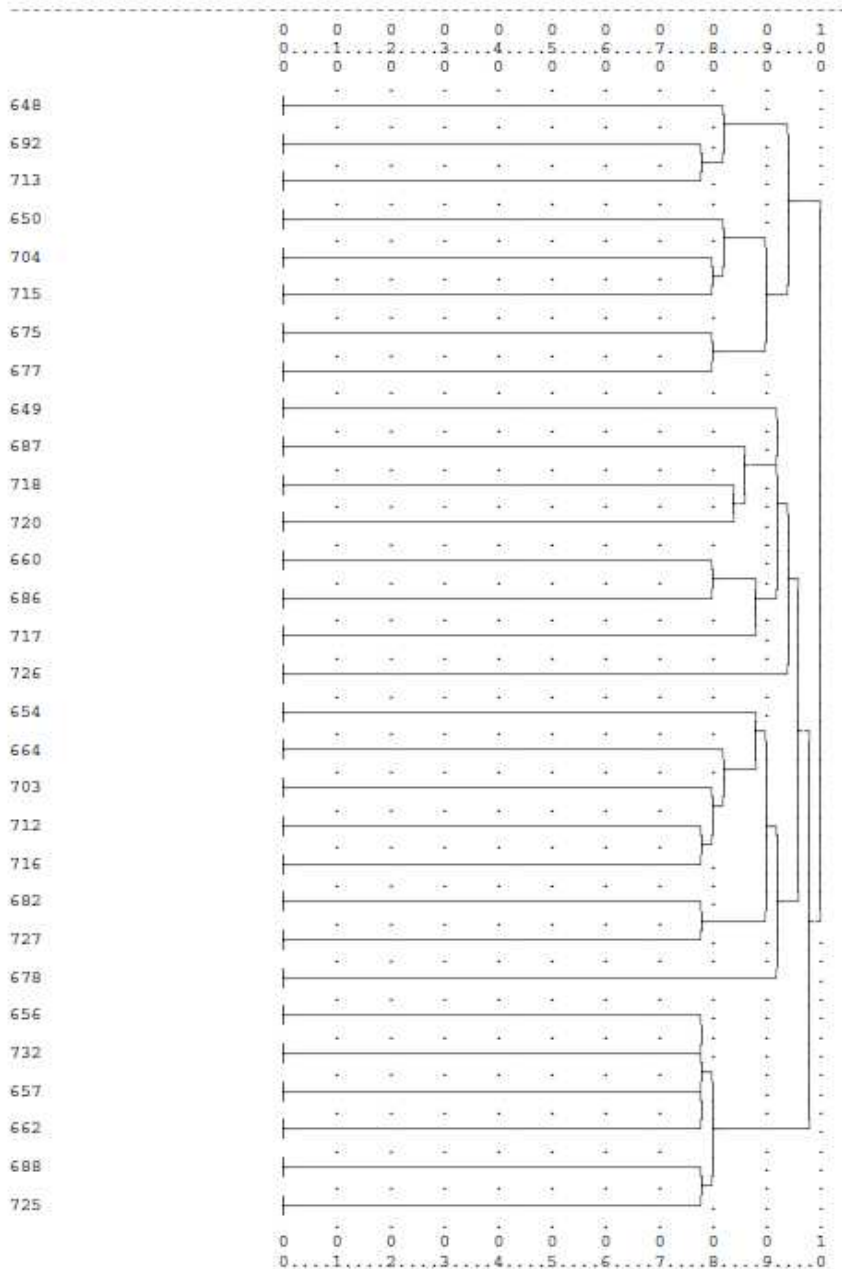


Figura 8: Exemplo de dendrograma elaborado no software Flowmap 7.2
 Fonte: ZWAN et al. (2005, p.78)

Após aplicação do método Intramax sobre o conjunto de dados, propõe-se que duas linhas de análise sejam seguidas. A primeira trata de discutir os resultados sugeridos para divisão do território brasileiro em subsistemas de migração.

Na segunda, acompanha-se a situação de qualquer unidade espacial do início ao fim do método iterativo, passo a passo. Isso leva à segunda linha de análise, ou seja, identificar e acompanhar o processo de agrupamento das unidades espaciais que mereçam destaque. Tais unidades devem ser selecionadas de acordo com os resultados observados para os subsistemas de migração detectados no primeiro momento ou aquelas que, por algum motivo, peçam uma atenção especial, de acordo com os as necessidades da pesquisa ou interesse do pesquisador.

Um fluxograma para a metodologia proposta encontra-se na Figura 9.

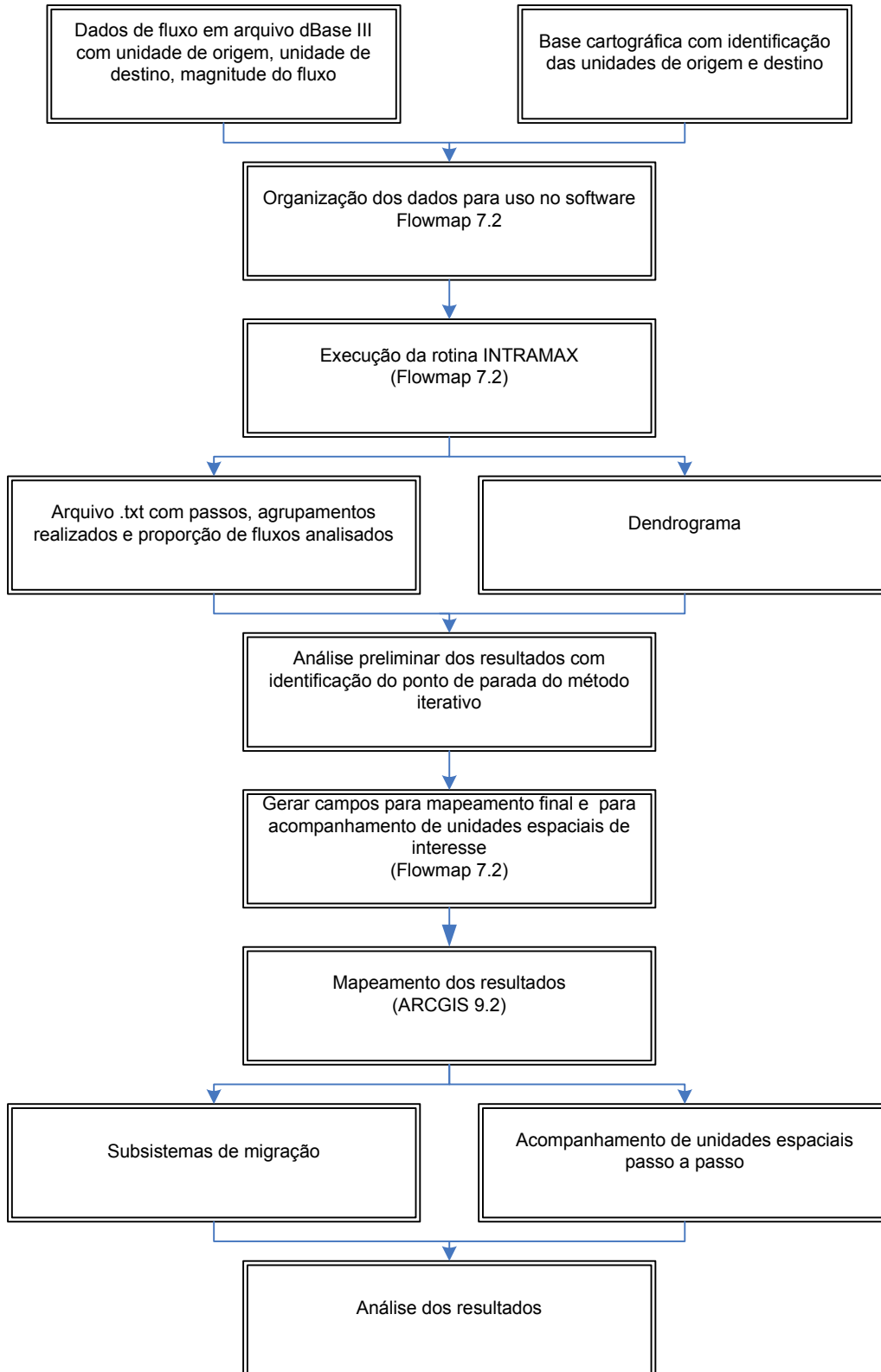


Figura 9: Etapas seguidas na metodologia proposta com uso do método Intramax

4.4.3 Delimitação de tipologias de migração com uso da Análise Fatorial de Ordem Superior

Para identificação de tipologias de migração, definidas como regiões de migração em que os grupos de unidades territoriais apresentam padrões similares de trocas com outras unidades, faz-se a proposta de uso da Análise Fatorial de alta ordem, com extração inicial de fatores por Análise de Componentes Principais e aplicação da rotação Promax. Os fundamentos matemáticos são discutidos no Apêndice B.

Ao apresentar esta proposta metodológica, discute-se: a forma de organização das matrizes de dados; a aplicação da técnica de Análise de Componentes Principais sobre a matriz de correlação entre as variáveis; a necessidade de aplicação de uma rotação oblíqua sobre os resultados; o significado de cada elemento obtido com a aplicação da técnica; sua interpretação na análise da migração e os critérios adotados como ponto de corte – para definição dos agrupamentos e para identificação de principais origens ou principais destinos do agrupamento sugerido.

Os dados de fluxos são organizados em matrizes quadradas não-simétricas de ordem 558. Tendo em vista que, com esta metodologia, busca-se a identificação de unidades espaciais com padrões similares de trocas de migrantes, a diagonal da matriz é nula, ou seja, o fluxo intramicrorregional não é considerado.

Para aplicação da metodologia no modo R, nas linhas da matriz encontram-se as origens e nas colunas os destinos. Assim, identificam-se os destinos com comportamentos similares e as principais origens para cada um dos grupos sugeridos. No modo Q, trabalha-se a transposta da matriz anterior, com as linhas sendo ocupadas pelos destinos e as colunas pelas origens. Neste caso, identificam-se as origens que apresentam comportamentos semelhantes e os principais destinos para cada um dos grupos sugeridos.

Conforme discussão matemática feita no Apêndice B, a técnica de Análise de Componentes Principais está diretamente relacionada à matriz de correlação entre as variáveis e os fatores gerados são linearmente independentes, ou seja, não-correlacionados. Deve-se reforçar ainda que a Análise de Componentes Principais é

um caso particular da Análise Fatorial, que, quando aplicada, permite a redução da dimensionalidade apresentada pelo conjunto de variáveis retendo uma variabilidade significativa da informação, eliminando parte da redundância registrada.

Os fatores extraídos do conjunto de dados no estado original (dados brutos) são referidos como de primeira ordem ou fatores primários e, como citado acima, quando obtidos via Análise de Componentes principais, são independentes, não apresentando correlação entre os fatores. Mas, se os fatores primários são submetidos a uma rotação oblíqua, é possível identificar uma estrutura de novos fatores inter-relacionados e, portanto, gerar uma nova matriz de correlação (veja discussão adiante). Sobre essa matriz de correlação, aplica-se novamente a técnica de Análise de Componentes Principais gerando um conjunto secundário de fatores, ou fatores de segunda ordem, em geral menos complexo e mais generalizante que o primeiro. Para entender melhor este significado, veja o item 4 do Apêndice B. Aplicando-se sucessivamente a mesma seqüência de passos, obtém-se os fatores de terceira ordem, quarta ordem, etc. Este processo se segue até que seja gerado apenas um fator ou todos os fatores gerados sejam iguais ou independentes ou, ainda, até o ponto em que o pesquisador considerar adequado, tendo em vista que, quanto maior a ordem, mais generalizador é o fator.

Clayton (1982), Gorsuch (1983) e Holsman (1980) sugerem uma técnica para avaliar o peso das variáveis originais presente nos fatores obtidos em ordens superiores. A demonstração da validade desta técnica pode ser vista em Gorsuch (1983). O que se faz é multiplicar a matriz de cargas fatoriais de uma ordem pela matriz de cargas fatoriais do nível imediatamente superior. Isso será possível, pois, o número de colunas da primeira matriz é sempre igual ao número de linhas da segunda. Desta maneira, pode-se calcular a comunalidade associada às variáveis originais ao trabalhar com os fatores de ordens superiores e o pesquisador terá como avaliar qual o poder de explicação perdido na generalização.

Portanto, propõe-se que seja feito, como exposto acima, o acompanhamento da força de explicação dos fatores obtidos em cada ordem e do grau de ligação com os dados originais para tomada de decisão do melhor momento de parada, na geração

de novos fatores, ou seja, de resultados de ordens superiores ainda mais generalizantes.

A seguir, faz-se uma breve discussão sobre a rotação oblíqua. Os fatores gerados, via Análise de Componentes Principais, são independentes (vide Apêndice B) e, portanto, independentes. Assim, a matriz de correlação entre os fatores será sempre idêntica à matriz identidade, composta de elementos iguais a 1 (um) na diagonal principal e zeros fora dessa diagonal, impossibilitando a reaplicação desta técnica sobre a matriz gerada. Mas, ao aplicar uma rotação oblíqua sobre os fatores obtidos via Análise de Componentes Principais, dois importantes aspectos podem ser verificados:

- i. Em geral, a estrutura do conjunto fica evidenciada e de mais fácil análise;
- ii. Os fatores obtidos não são independentes e, portanto, geram uma matriz de correlação não idêntica à matriz identidade.

Portanto, com os objetivos de evidenciar a estrutura de relacionamento entre as unidades espaciais e de reaplicação da técnica de Análise de Componentes Principais, faz-se a opção pelo uso de uma rotação oblíqua sobre os fatores obtidos em cada ordem. A rotação oblíqua proposta é a PROMAX, pois, conforme discutido e citado por Gorsuch (1970,pp.861-872), este método é o mais rápido e pode ser aplicado de modo aceitável em um grande volume de dados, sendo o mais eficiente dentre os demais e o que apresenta resultados com uma estrutura interna mais simples de ser analisada.

Tendo sido definida e justificada a proposta de aplicação da Análise Fatorial de ordem superior com fatores extraídos via Análise de Componentes Principais com rotação PROMAX, inicia-se a discussão sobre o significado de cada elemento obtido com a aplicação da técnica e respectiva interpretação na análise da migração.

Em um conjunto com N variáveis, N fatores são extraídos. Desses, serão considerados relevantes para a análise proposta aqueles com autovalor maior ou igual a um – critério da raiz latente – o que sugere que este fator consegue captar a variância total equivalente a de pelo menos uma variável. Além disso, exigir-se-á que o conjunto de fatores retenha o mínimo de 65% da variância total – critério de

porcentagem da variação – valor adotado empiricamente via observação de resultados em trabalhos do autor desta tese ou em outros trabalhos na área da migração desenvolvido por Clayton (1977b), Clayton (1982), Holsman (1980), Elis, Barff e Renard (1993) e outros, não existindo regra matemática específica para determinação deste número.

Na Análise de Componentes Principais, as cargas fatoriais indicam a correlação de cada variável com o fator, mas, com a aplicação da rotação oblíqua esta relação deixa de ser verdadeira e as cargas fatoriais passam a representar o grau de correspondência entre a variável e o fator, sendo que cargas com maiores módulos fazem a variável mais representativa no fator.

Selecionados os fatores mais significativos, passa-se a estudar suas respectivas cargas fatoriais, que indicam o grau de correspondência entre a variável e o fator. Vale lembrar que cargas fatoriais maiores tornam a variável mais significativa no fator. Ou seja, partindo do princípio de que quanto maior a carga fatorial de uma microrregião brasileira, maior a relação desta microrregião com o fator e, por consequência, com a unidade regional definida por esse fator, serão identificadas as tipologias de migração ou, ainda, os agrupamentos de microrregiões que se comportam de modo similar na troca de migrantes.

As cargas fatoriais, na Análise de Componentes Principais, assumem valores no intervalo de -1 a 1 , mas, com a aplicação da rotação oblíqua, esses valores podem fugir deste intervalo. O que se observa é uma pequena variação nestes limites. Gorsuch (1983) mostra que a carga fatorial deve ser tomada como significativa dentro do fator, em um universo de 100 registros, quando for maior ou igual a $0,4$, com significância de 5% , enquanto, para Rummel (1970), a carga fatorial deve atingir pelo menos o valor $0,5$. Sendo assim, baseando em observações empíricas feitas pelo próprio autor e nos autores citados acima, propõe-se como ponto de corte para definir se uma microrregião é significativa para aquele fator e, por conseguinte, para o agrupamento de microrregiões gerado, quando a carga fatorial tiver módulo maior ou igual a $0,5$.

Além disso, os agrupamentos de microrregiões deverão ser mutuamente excludentes, ou seja, cada microrregião deverá pertencer a apenas um agrupamento. Quando uma microrregião se mostrar significativa em mais de um fator, essa será alocada no grupo em que apresentar maior carga fatorial. Análises complementares devem ser feitas sobre a particularidade dessa unidade espacial.

No modo R, quando as linhas são ocupadas pelas origens e as colunas pelos destinos, as maiores cargas fatoriais indicarão destinos que se comportam de modo semelhante, ou seja, que recebem migrantes de origens comuns.

No modo Q, quando as linhas são ocupadas pelos destinos e as colunas pelas origens, as maiores cargas fatoriais indicarão origens que se comportam de modo semelhante, ou seja, que enviam migrantes para destinos comuns.

Definidos os agrupamentos, parte-se para a definição das principais origens ou destinos das microrregiões que os compõem. De acordo com Clayton (1982,p.12), discutindo os resultados da aplicação da Análise Fatorial de alta ordem no modo R, aquelas origens que apresentam mais altos escores em um mesmo componente identificam origens que estão relacionadas de modo mais significativo com o grupo de destinos identificado previamente pelas cargas fatoriais.

Este fato pode ser melhor visualizado quando se lembra que o fator obtido, usado para definição do agrupamento, nada mais é que uma combinação linear das variáveis, neste caso, dos destinos e os registros são as origens. Como os escores são obtidos com a substituição dos valores dos registros (migrantes) na combinação linear, teremos o escore total gerado por uma origem e os escores parciais dessa mesma origem para cada fator.

Quanto maior o escore parcial, o escore de uma origem em determinado fator, maior a contribuição daquela origem para o agrupamento de destinos.

De modo análogo, ao trabalhar no modo Q, pode-se identificar os principais destinos para cada agrupamento de microrregiões, de acordo com a similaridade de emigração.

Um fluxograma para a metodologia proposta encontra-se na Figura 10.

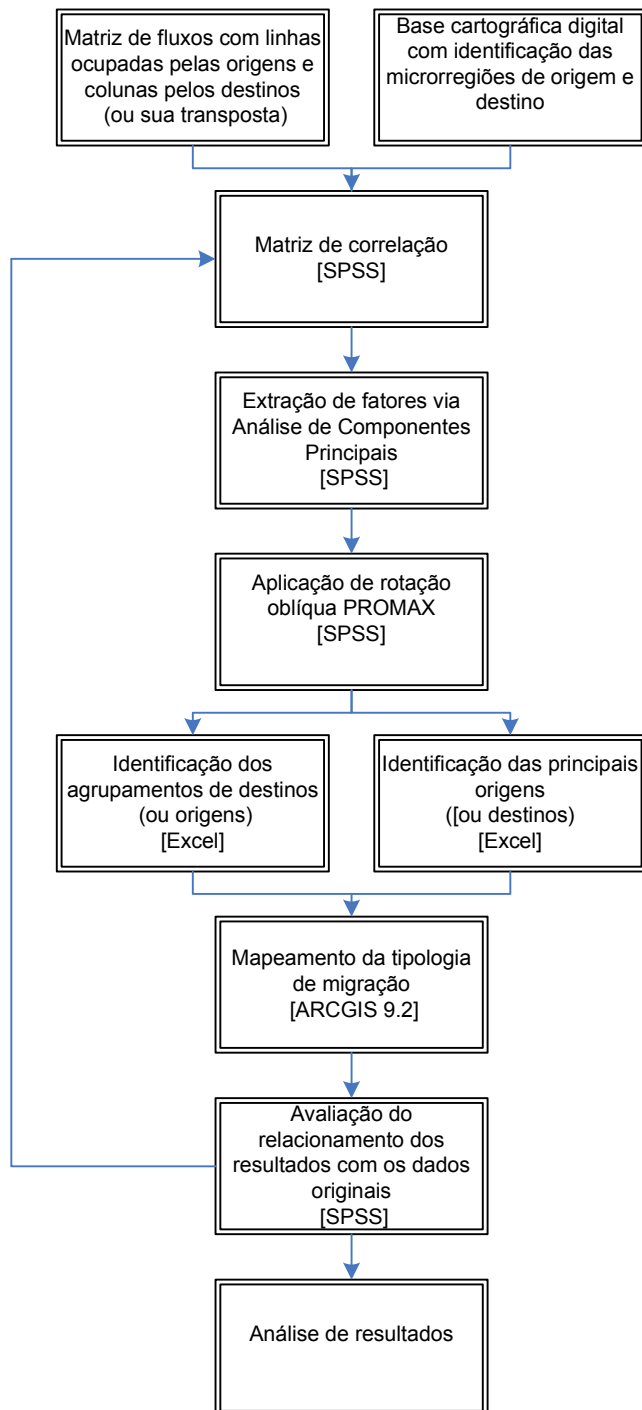


Figura 10: Etapas seguidas na metodologia proposta com uso da Análise Fatorial de Alta Ordem

5 A IDENTIFICAÇÃO DE SUBSISTEMAS E TIPOLOGIAS ESPACIAIS DOS FLUXOS MIGRATÓRIOS BRASILEIROS

5.1 Introdução

Neste capítulo, faz-se a análise dos resultados obtidos com a aplicação das metodologias propostas sobre os microdados dos Censos de 1991 e 2000, identificando subsistemas e tipologias de migração, tomando as 558 microrregiões brasileiras como unidades espaciais básicas. Essa análise é feita para cada uma das metodologias em separado e os resultados são comparados e discutidos.

Os dados trabalhados estão de acordo com a definição de migrante de data-fixa, ou seja, aquele indivíduo que apresenta município de residência em 1986 ou 1995 diferente daquele registrado em 1991 ou 2000, respectivamente. Trata-se, portanto, de migrantes sobreviventes dos respectivos períodos, independentemente de terem ou não realizado movimentos migratórios intermediários no quinquênio estudado.

No caso em que é feita a aplicação da Análise Fatorial de Alta Ordem, via Análise de Componentes Principais, para identificação de tipologias de migração entre as microrregiões brasileiras, são avaliados os resultados obtidos para quatro matrizes quadradas não-simétricas de ordem 558, sendo duas relativas ao período de 1986 a 1991 e outras duas para o período de 1995 a 2000. Para cada período, uma das matrizes é do tipo Origem-Destino e a outra do tipo Destino-Origem.

No caso de aplicação do Método Intramax, para identificação de subsistemas de migração, duas matrizes não-simétricas de ordem 558 são avaliadas, sendo uma para cada período estudado. Faz-se a análise dos agrupamentos obtidos com a metodologia e mostra-se a evolução do processo de agrupamento para duas microrregiões selecionadas, Belo Horizonte-MG e Chapecó-SC.

Os resultados obtidos para os agrupamentos são representados cartograficamente, usando os softwares ARCGIS 9.2 e Flowmap 7.2. Alguns mapas são inseridos no próprio texto, revelando padrões espaciais no comportamento do migrante e características inerentes aos passos metodológicos adotados. Outros são

disponibilizados digitalmente em um CD contendo os Apêndices, tendo em vista que foram produzidos mais de 600 mapas.

No Apêndice C são apresentados quatorze mapas para a identificação das unidades espaciais básicas adotadas neste trabalho, ou seja, das microrregiões brasileiras. O critério adotado para apresentação desses mapas é o de viabilização de leitura, não tendo relação com os agrupamentos determinados pelas metodologias propostas.

5.2 Subsistemas de migração

Conforme exposto anteriormente, na Seção 3.4.2, faz-se a aplicação do Método Iterativo INTRAMAX, para identificação de subsistemas de migração, ou seja, regiões de migração como grupos de microrregiões brasileiras que estão fortemente conectadas entre si.

Para os dois períodos, o método iterativo é interrompido e as regiões de migração determinadas após 550 passos dos 557 possíveis. Esse ponto de corte é definido pela análise do gráfico dendrograma (no Apêndice J) e pela observação de que 76,2% dos fluxos, para 1991, e 76,1%, para 2000, são considerados intrazonais, ou seja, ocorrem dentro dos agrupamentos definidos. Portanto, menos de 25% dos fluxos ocorrem entre as regiões definidas. Caso o processo iterativo tivesse continuidade, seriam formadas regiões excessivamente generalizantes, com aumentos exagerados nos fluxos intrazonais, trazendo pouca ou nenhuma contribuição para o estudo do fluxo migratório.

Como auxiliares para melhor entendimento da metodologia e, também, para a análise dos resultados, são apresentadas as Figuras 12,13,14 e 15, para o período 1986-1991, e as Figuras 16 e 17, para o período 1995-2000, com as regionalizações obtidas após aplicação de 500, 540 e 550 passos do método iterativo.

Além da análise das regiões determinadas, faz-se a apresentação de uma abordagem viabilizada pelo método, que é o estudo particularizado da evolução no processo de agrupamento de alguma unidade espacial básica de interesse. Nesse caso, para exemplificação, foram selecionadas as microrregiões de Belo Horizonte-

MG e Chapecó-SC. A primeira por interesse particular do autor e a segunda por apresentar importante mudança no padrão de fluxos migratórios entre os períodos analisados, exemplificando uma das relevantes características da metodologia proposta.

A Figura 11 apresenta os agrupamentos determinados após a aplicação de 3 passos do método Intramax para o período de 1986-1991. No primeiro passo são unidas as microrregiões de Tefé-AM e de Japurá-AM, no segundo Oiapoque-AP e Amapá-AP e no terceiro Purus-AM e Madeira-AM. Deve-se observar que, tomando como universo todos os fluxos observados entre as microrregiões brasileiras, são microrregiões com fluxos proporcionalmente pequenos.

Acompanhando o processo de agrupamento passo a passo, para os dados do período 1986-1991, pode-se observar que, dos 100 maiores acréscimos de volumes migratórios, 95% ocorrem a partir do 329^o passo, sendo que 85% ocorrem após o 406^o passo, ou seja, já na fase final do processo e 80% dos 100 menores acréscimos ocorrem antes do 203^o passo, ainda na fase inicial, onde os agrupamentos ainda se apresentam bastante dispersos no território brasileiro.

Observa-se, ainda, que as primeiras entradas das microrregiões de São Paulo, Rio de Janeiro e Belo Horizonte ocorrem, respectivamente, nos 532^o, 477^o e 482^o passos, de um total de 557 passos.

Analogamente, analisando os dados referentes ao período de 1995-2000, observa-se que, dos 100 maiores acréscimos de volumes migratórios, 95% ocorrem a partir do 302^o passo, sendo que 85% ocorrem após o 403^o passo, ou seja, já na fase final do processo, e 79% dos 100 menores acréscimos ocorrem antes do 225^o passo, ainda na fase inicial.

Para os dados do Censo de 2000, as primeiras participações das microrregiões de São Paulo, Belo Horizonte e Rio de Janeiro ocorrem nos 520^o, 480^o e 482^o passos. Observa-se uma inversão no posicionamento das microrregiões de Rio de Janeiro e Belo Horizonte, o que pode sugerir uma perda de importância relativa da

microrregião do Rio de Janeiro nos fluxos migratórios, apesar de ainda ser bastante significativa.

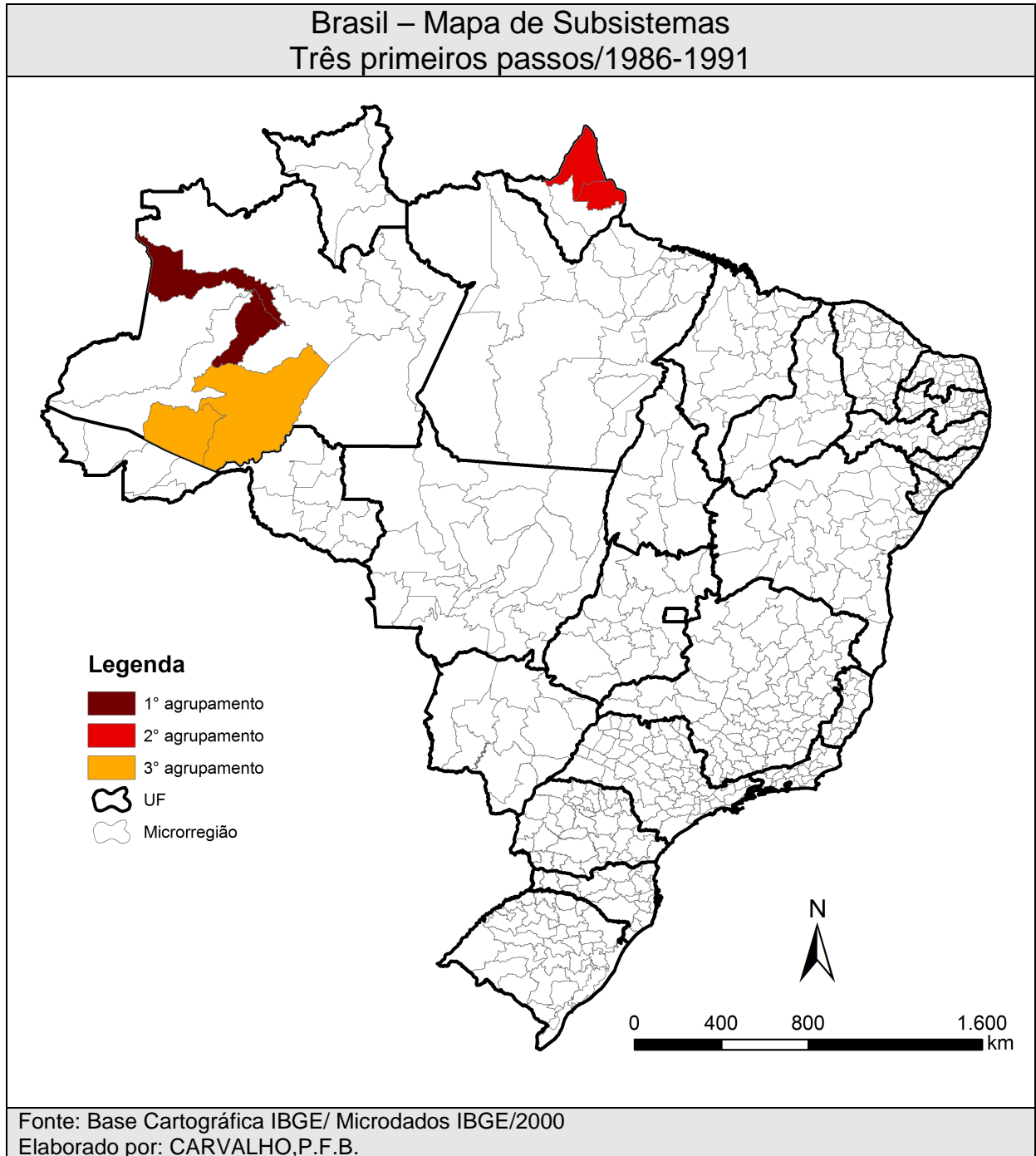


Figura 11: Três primeiros subsistemas gerados pelo Método Intramax/1986-1991

Com o exposto nos parágrafos anteriores, confirma-se o fato esperado de, na aplicação do método Intramax, os agrupamentos das unidades espaciais com menores fluxos, mas com fortes interações, ocorrerem nos passos iniciais, enquanto nos passos intermediários são observados agrupamentos de pequenas unidades

com outras unidades maiores e as uniões de grandes unidades com outras grandes ficando para os passos finais.

Outra importante constatação se refere à questão da contigüidade. Apesar de o método Intramax não admiti-la explicitamente em sua formulação, observa-se a tendência, nos dois períodos estudados, para a predominância quase absoluta de delimitações de subsistemas de migração onde a contigüidade é predominante, fortalecendo o entendimento corrente da importância da migração em distâncias mais curtas, como citado no texto “Configuração atual e tendências da rede urbana” (CONFIGURAÇÃO...,2002,p.35), publicado por IPEA, IBGE e UNICAMP.

Com a regionalização determinada por aproximadamente 65% de fluxos intrarregionais – 540 passos –, observa-se perda da contigüidade apenas no norte da Bahia, para os dois períodos, particularmente para as microrregiões de Jeremoabo e Feira de Santana, situando-se como “ilhas” dentro de dois agrupamentos distintos, mas vizinhos (ver Apêndice D), e para o período 1995-2000 o estado do Piauí estando mais fortemente ligado com o estado de Goiás, sul do Tocantins e oeste de Mato Grosso, exceto pela microrregião de Valença do Piauí pertencendo ao agrupamento definido pelo Triângulo Mineiro/MG, Alto Paranaíba/MG e uma porção do norte de São Paulo, formada por microrregiões contíguas a Franca e Barretos (ver Apêndice E).

Analisando os resultados da regionalização determinada por aproximadamente 75% dos fluxos intrarregionais – 550 passos –, mantém-se a não contigüidade apenas de Feira de Santana e Jeremoabo, na Bahia, nos dois períodos, e de Valença do Piauí/PI para o período 1995-2000.

Também deve ser observada a questão das fronteiras interestaduais não se mostrarem altamente significativas na maior parte dos subsistemas definidos, fato que pode ser constatado nas análises subseqüentes. Isso sugere a existência de centros regionais que não são necessariamente apenas as capitais.

Essa constatação reforça a discussão apresentada na seção 2.1, que versa sobre o setor industrial buscar mão-de-obra especializada em outros centros

educacionalmente mais desenvolvidos, de o migrante buscar regiões mais dinâmicas economicamente, ou, ainda, que dada a maior complexidade das atividades, surge a necessidade de maior integração entre os centros urbanos.

O município de Uberlândia pode ser apresentado como um exemplo típico, tendo assumido a posição de maior cidade do interior do estado de Minas Gerais com população de 501.214 habitantes no ano 2000, estando entre aquelas que mais geram emprego no estado, junto com a região do entorno de Ipatinga, atraindo migrantes do próprio estado, de Goiás, São Paulo, Bahia, dentre outros. No período 1995-2000, essa microrregião recebeu 132.840 migrantes provenientes de outros estados e 27.688 de outras microrregiões de seu próprio estado. São várias instituições de ensino superior, dentre elas a Universidade Federal de Uberlândia, e outras tecnológicas.

"A importância econômica desta região está pautada nas atividades ligadas à modernização do setor agropecuário e agroindustrial, que provocou uma expressiva evasão populacional do campo para as cidades grandes e médias; bem como na ampliação e diversificação do setor terciário... É possível também verificar um movimento migratório interurbano entre as cidades menores e as cidades maiores. Neste contexto, o município de Uberlândia destaca-se enquanto importante centro atrativo de população e de investimentos econômicos, projetando-se a nível regional e nacional". (SOARES; BESSA, 1999)

O fato desta ser uma região que atrai mais que expulsa migrantes pode ser destacada quando se registra que, para o mesmo período de 1995-2000, a microrregião de Uberlândia perdeu 37.110 migrantes, sendo 23.471 para microrregiões de outros estados. Esse total significa 23% do total de imigrantes.

A Figura 12 apresenta os resultados obtidos após os passos de números 1, 200, 300, 400, 500, 520, 540 e 550, mostrando a evolução na determinação dos agrupamentos com a aplicação do método iterativo para o período de 1986-1991.

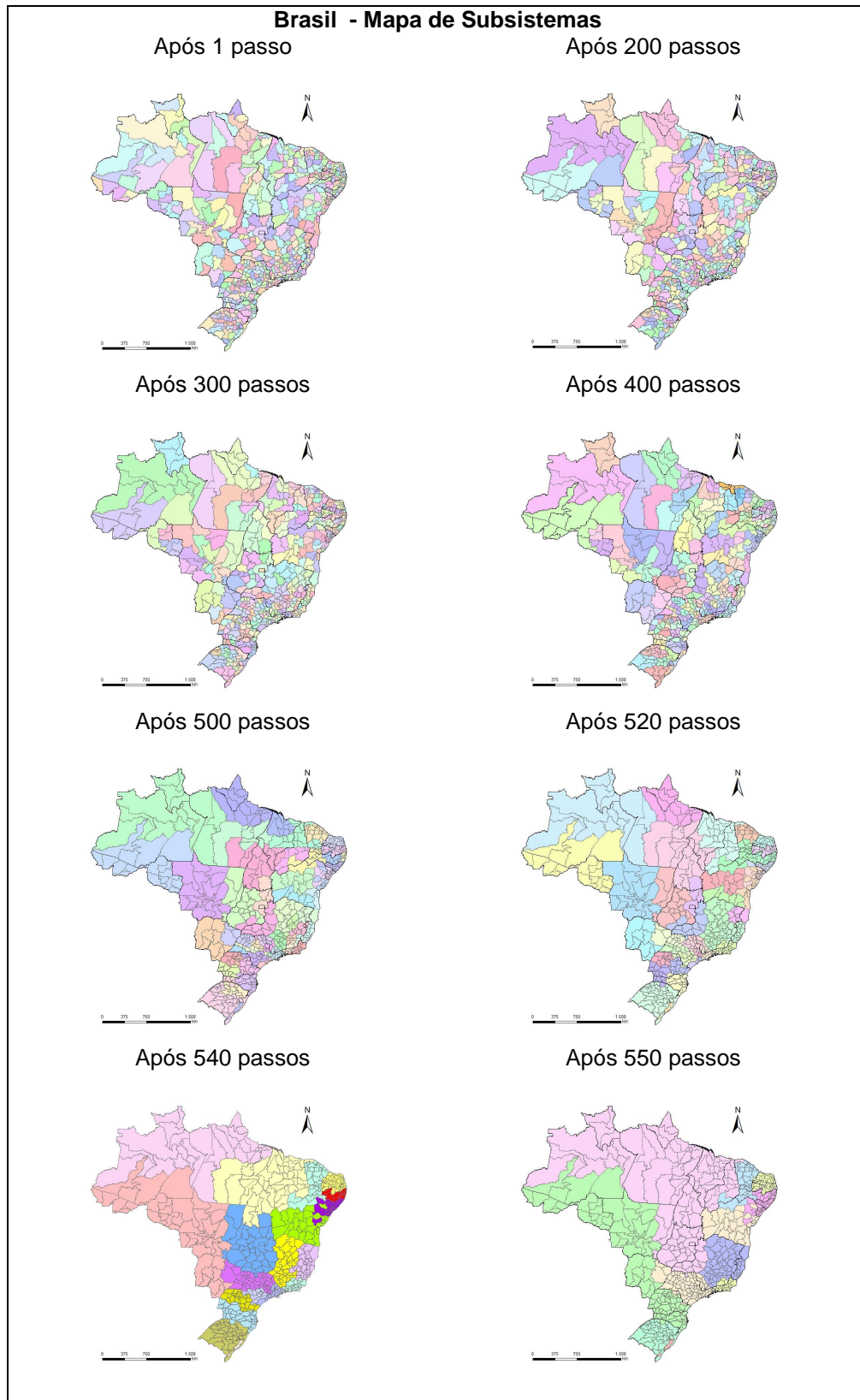


Figura 12: Evolução dos subsistemas determinados/1986-1991

Quanto aos subsistemas identificados para o período 1986-1991, faz-se a análise principal sobre os resultados obtidos após 540 passos do método iterativo,

doravante denominado primeiro nível, complementada com aqueles obtidos após 550 passos, denominado segundo nível da análise e que apresenta um menor grau de detalhamento, chegando aos 76,2% de fluxos intrarregionais. A representação cartográfica pode ser vista nas Figuras 13, 14 e 15.

Inicia-se a apresentação dos subsistemas pelas regiões Norte e Centro-Oeste do Brasil que, de acordo com dados do IBGE, são aquelas com maiores taxas de crescimento geométrico da população nas décadas de 1980 e 1990: 3,85% e 2,86% ao ano, respectivamente, para a região Norte e 3,01% e 2,39% ao ano para a região Centro-Oeste, enquanto a população brasileira cresce a uma taxa média de 1,93% ao ano, na década de 1980, e de 1,64%, na década de 1990.

Observa-se a região Norte do Brasil fragmentada em 4 subsistemas. Um subsistema identificado é formado pelas totalidades dos estados do Acre e Rondônia, porção sul do Amazonas, portanto, sem a microrregião de Manaus, mas com as microrregiões de Juruá, Boca do Acre, Purus e Madeira, menos urbanizadas, ligadas com a maior parte do estado do Mato Grosso, excetuando sua porção leste, e centro-oeste do Mato Grosso do Sul. A porção norte deste subsistema recebeu um importante fluxo de migrantes na década de 80, motivado pela mineração e pela pecuária, de acordo com a abordagem da Seção 2.1. No segundo nível, observa-se a interação deste subsistema com os estados do Paraná e Santa Catarina que atuam, dentro deste subsistema, principalmente, com saldo migratório negativo, mas com um importante fluxo de retorno ou mesmo de outros migrantes do norte para esses.

Nesse subsistema, Rondônia se destaca ao apresentar taxa geométrica de crescimento populacional de 7,89% ao ano, na década de 1980, apesar da exorbitante taxa de 16,03%, referente à década de 1970, quando ainda se podia apontá-lo como fronteira agrícola em ocupação. Em 1991, Rondônia conta com mais de 40% de população rural.

As porções central e norte do estado do Amazonas, incluindo a microrregião de Manaus, mais urbanizadas, formam outro subsistema de migração que incorpora as totalidades dos estados de Roraima e Amapá e as porções norte e oeste do Pará, incluindo a microrregião de sua capital, Belém. Destes, Roraima é o estado com

maior taxa de crescimento anual na década de 1980, 9,63%, seguido pelo Amapá, com 4,67%. Os principais centros deste subsistema são Manaus e Parintins no Amazonas; Belém, Ananindeua e Santarém, no Pará; Boa Vista, em Roraima, e Macapá, no Amapá.

Outro subsistema que engloba parte da região Norte é formado pelo centro-sul do Pará, com destaque para as microrregiões de Altamira, Marabá, Paragominas e Tucuruí; a totalidade do estado do Maranhão; o estado do Tocantins, excetuando-se a microrregião de Dianópolis, ao sul do estado, que apresenta maior interação com o estado de Goiás e entorno de Brasília, e o estado do Piauí, menos a microrregião de São Raimundo Nonato, com maior ligação com norte da Bahia, Ceará e oeste de Pernambuco.

Ainda partindo do Norte e Centro-Oeste, outro subsistema fica definido pelo leste do Mato Grosso, com as microrregiões de Médio Araguaia, Norte Araguaia e Canarana, junto com todo o estado de Goiás e o Distrito Federal; e, ainda, o Triângulo Mineiro/Alto Paranaíba e oeste de Minas Gerais. Esse subsistema tem como alguns de seus importantes centros: o entorno de Brasília; a região do Triângulo Mineiro com Uberlândia e Uberaba, além de Paracatu e Unaí, no oeste mineiro, e Goiânia, no estado de Goiás.

O estado de Minas Gerais tem seu território dividido em seis subsistemas. O primeiro, já mencionado acima, formado pelo Triângulo Mineiro, Alto Paranaíba e noroeste, interagindo com o estado de Goiás, entorno de Brasília, parte leste do Mato Grosso, sul do Tocantins e Barretos em São Paulo. O maior e mais importante subsistema desse estado é formado por um corredor que corta o estado de norte a sul, tendo como principais centros sua capital, Belo Horizonte, responsável por grande parte da distribuição populacional no estado, Montes Claros, ao norte, e Lavras, Varginha, Campo Belo e São Lourenço, ao sul.

Outro subsistema fica determinado pelos municípios do Vale do Rio Doce, tendo Governador Valadares como maior centro, e outros pólos importantes como Ipatinga, Timóteo e Coronel Fabriciano, mais o Vale do Mucuri, com Teófilo Otoni e Nanuque, interagindo com o estado do Espírito Santo e sul da Bahia.

O quarto subsistema que corta parte do estado de Minas Gerais é formado, na realidade, pela totalidade do estado do Rio de Janeiro mais as microrregiões mineiras de Juiz de Fora, Ubá e Cataguases, que tradicionalmente sempre mostraram grande interação com o estado fluminense.

O penúltimo subsistema, que engloba microrregiões do estado de Minas Gerais, com Poços de Caldas, Alfenas, Passos e São Sebastião do Paraíso, é caracterizado, principalmente, pela composição das porções oeste, norte e parte da região central do estado de São Paulo, além das microrregiões de Três Lagoas, Paranaíba e Cassilândia, do oeste do Mato Grosso do Sul.

O sexto e último subsistema, complementado pelas microrregiões mineiras de Pouso Alegre, Itajubá e Santa Rita do Sapucaí, é delimitado pela região leste do estado de São Paulo, percorrendo desde Registro em sua porção mais ao sul até Bananal, Guaratinguetá, Campos do Jordão e, principalmente, a microrregião de São Paulo e entorno.

Esse subsistema, no segundo nível de iteração, com aplicação dos 550 passos, será ampliado, assumindo toda a extensão do estado de São Paulo, o sul de Minas Gerais, a região do Mato Grosso do Sul, citada acima, e quase todo o estado da Bahia, mostrando resquícios da forte ligação do estado de São Paulo com o Nordeste brasileiro.

Ainda no estado de São Paulo, observa-se um importante subsistema, formado, dentre outras, pelas microrregiões de Sorocaba, Itapeva, Capão Bonito, Ourinhos e Botucatu, com outras do norte do estado do Paraná: Cornélio Procópio, Jacarezinho e Cerro Azul.

O estado do Paraná está basicamente dividido entre dois subsistemas. O primeiro, que ocupa maior extensão, o percorre de leste a oeste, desde as microrregiões de Curitiba e Paranaguá até Umuarama, Cianorte e Paranaíba. O outro subsistema, formado por sua parcela mais a sudoeste, interage mais fortemente com todo o estado de Santa Catarina.

No extremo sul brasileiro, o estado do Rio Grande do Sul delimita um único subsistema, sugerindo que interação migratória mais forte é a intraestadual, exceto pela microrregião de Erechim, na divisa com Santa Catarina.

Quanto ao Nordeste brasileiro, observa-se os estados do Maranhão e Piauí interagindo com o estado do Tocantins e o sudeste do Pará, como dito acima, e o Ceará interagindo com a porção oeste do Pernambuco e extremo norte da Bahia. A região central e leste de Pernambuco formam um único subsistema, capitaneado por sua capital, Recife, enquanto Rio Grande do Norte e Paraíba formam outro subsistema de migração. E, finalmente, Alagoas e Sergipe apresentam interação significativa com a porção nordeste da Bahia.

A região Nordeste ainda apresenta um saldo negativo de aproximadamente 500.000 migrantes em seu relacionamento com a região Sudeste do Brasil. Sua interação é muito mais forte com São Paulo do que com os outros estados. Aproximadamente 700.000 pessoas migraram do Nordeste para o estado de São Paulo no período 1986-1991.

Avaliando os resultados para o período de 1995-2000, observa-se a manutenção do panorama geral apresentado pelos subsistemas formados para o período anterior. Portanto, faz-se a discussão das alterações detectadas.

A região Norte fica dividida em apenas dois subsistemas: o subsistema formado pelos estados do Acre, Rondônia e porção sul do Amazonas tem sua interação reduzida com o estado do Mato Grosso do Sul, que, por sua vez, passa a formar um único subsistema com as microrregiões do noroeste paulista de Jales, Andradina, Araçatuba, Auriflama, Fernandópolis, Birigui, Nhandeara, Votuporanga e São José do Rio Preto.

A região norte do Amazonas, mais a totalidade dos estado de Roraima, Amapá Pará formam o outro subsistema, complementado pelo estado do Maranhão e quase a todo o Tocantins.

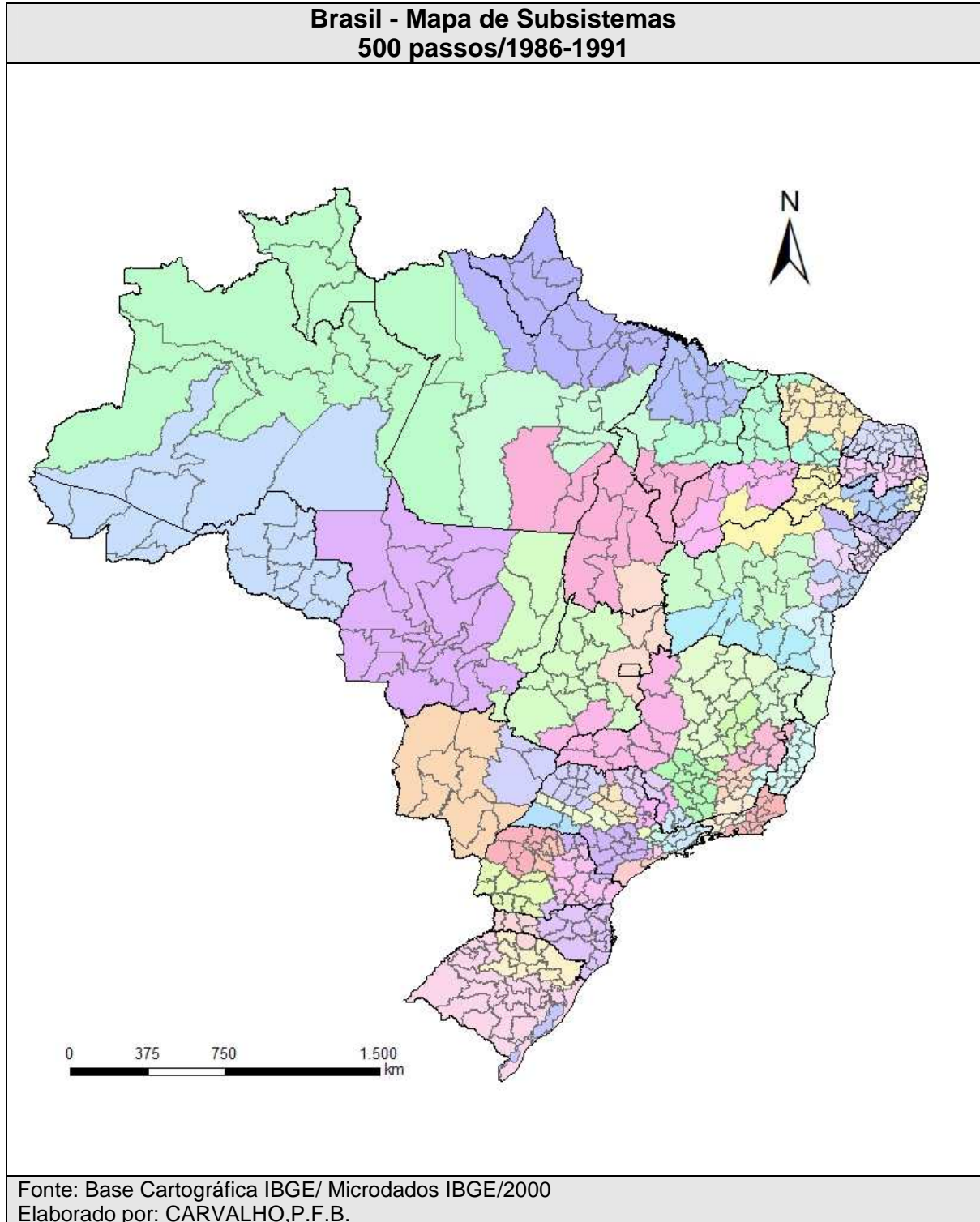


Figura 13: Subsistemas após aplicação de 500 passos do Método Intramax/1986-1991

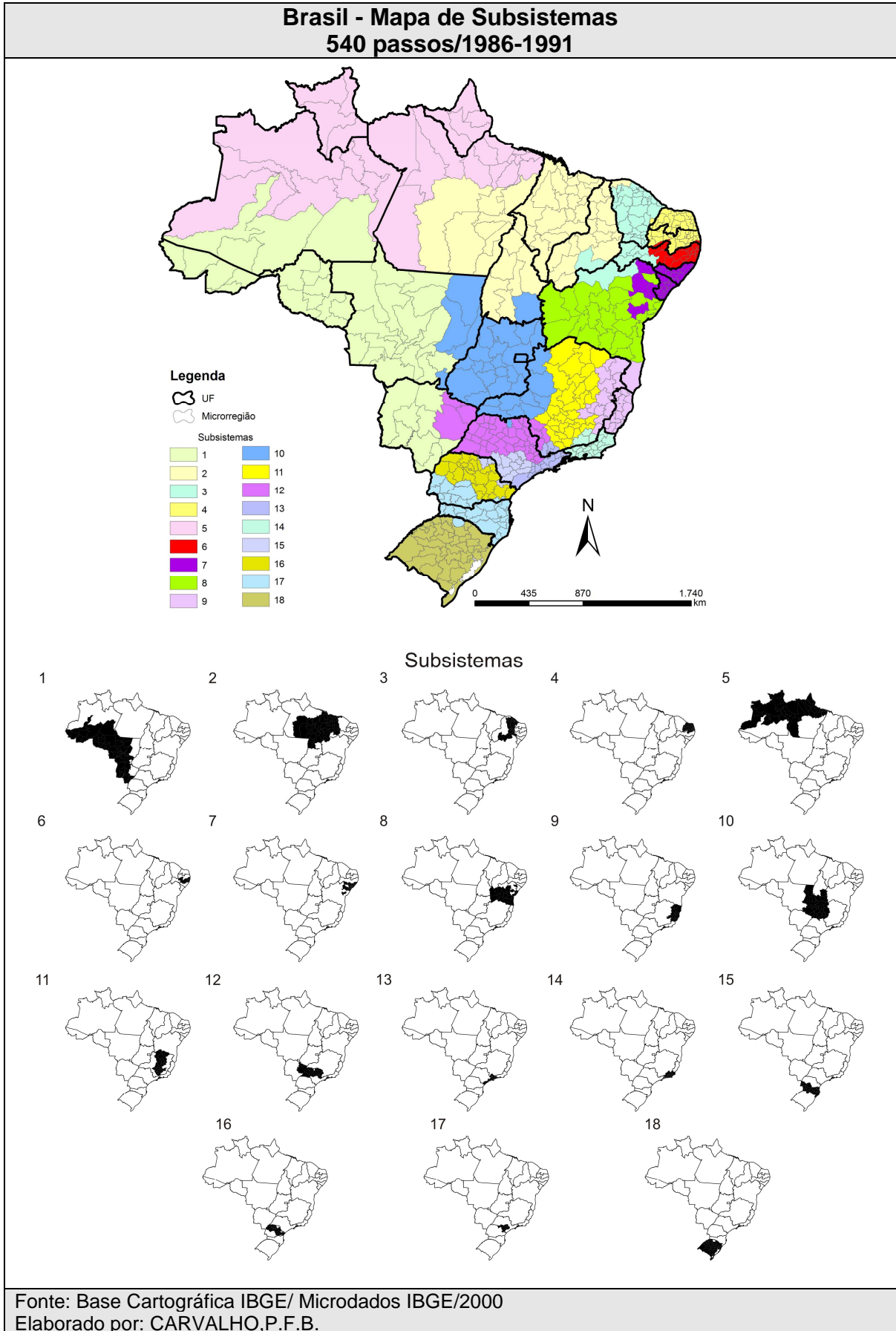


Figura 14: Subsistemas após aplicação de 540 passos do Método Intramax/1986-1991

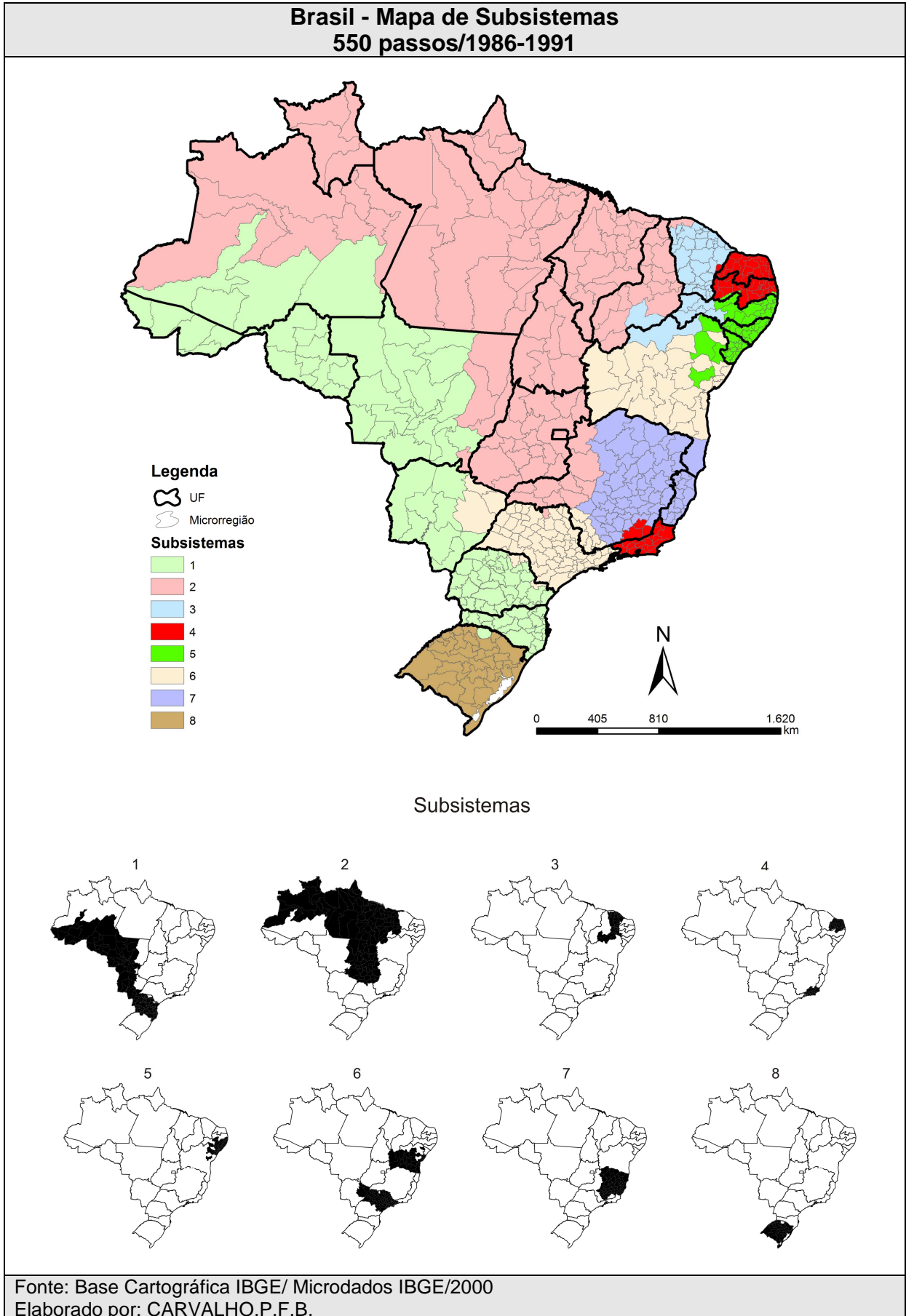


Figura 15: Subsistemas após aplicação de 550 passos do Método Intramax/1986-1991

O estado de Goiás e o entorno de Brasília formam outro subsistema, complementado por parte da região sul do Tocantins e leste do Mato Grosso, e com aumento na interação com o estado do Piauí, exceto pela microrregião de Valença do Piauí. Observa-se redução na força de interação desse subsistema com o Triângulo Mineiro, Alto Paranaíba e oeste de Minas Gerais que passam a compor, junto com Passos, outro subsistema com as microrregiões do norte de São Paulo como Barretos, São Joaquim da Barra, Franca, Ituverava, Novo Horizonte, Araraquara, Catanduva, Jaboticabal, Batatais, Ribeirão Preto e São Carlos.

As microrregiões mineiras de Alfenas, Poços de Caldas e São Sebastião do Paraíso continuam mostrando forte interação com o estado de São Paulo, que passa a ser subdividido em 5 subsistemas, ao invés dos 3 do período anterior, indicando o fortalecimento de pólos regionais no estado e redução do poder de atração de sua capital.

O Paraná, quase em sua totalidade, passa a formar um único subsistema com o estado de Santa Catarina, enquanto o estado do Rio Grande do Sul continua em outro subsistema isolado.

Quanto ao Nordeste brasileiro, observa-se manutenção da configuração registrada para o período anterior, exceto pelo estado do Piauí, que passa a interagir mais fortemente com parte da Região Centro-Oeste. Uma importante observação pode ser feita quanto ao relacionamento da Bahia com o estado de São Paulo. Ao comparar os resultados dos dois períodos, no segundo nível de análise, com 550 passos, nota-se que o estado da Bahia passa a formar um novo subsistema com o estado de Minas Gerais e o Espírito Santo, ao invés do estado de São Paulo. Uma vez mais, isso mostra a relevância dos fluxos de mais curta distância no período atual.

De fato, os dados apontam para uma redução no saldo migratório negativo da Bahia com o estado de São Paulo de 9.216 migrantes, e aumento neste saldo migratório negativo, de 7.037 migrantes, na relação com Minas Gerais entre os dois períodos. Vale destacar que este fenômeno também é observado quando tomamos como

universo todo o Nordeste do Brasil, que apresenta redução de 84.745 migrantes em seu saldo negativo, na relação com o estado de São Paulo.

O número de migrantes que sai de São Paulo com destino à Bahia, no período de 1986-1991, equivalia a 26% do fluxo no sentido contrário, enquanto que, para o período seguinte, este percentual passa a ser de 39%. Tomando a região Nordeste como um todo, este percentual passa de 33% para 47%, indicando mudança no comportamento do migrante desta região com o estado paulista.

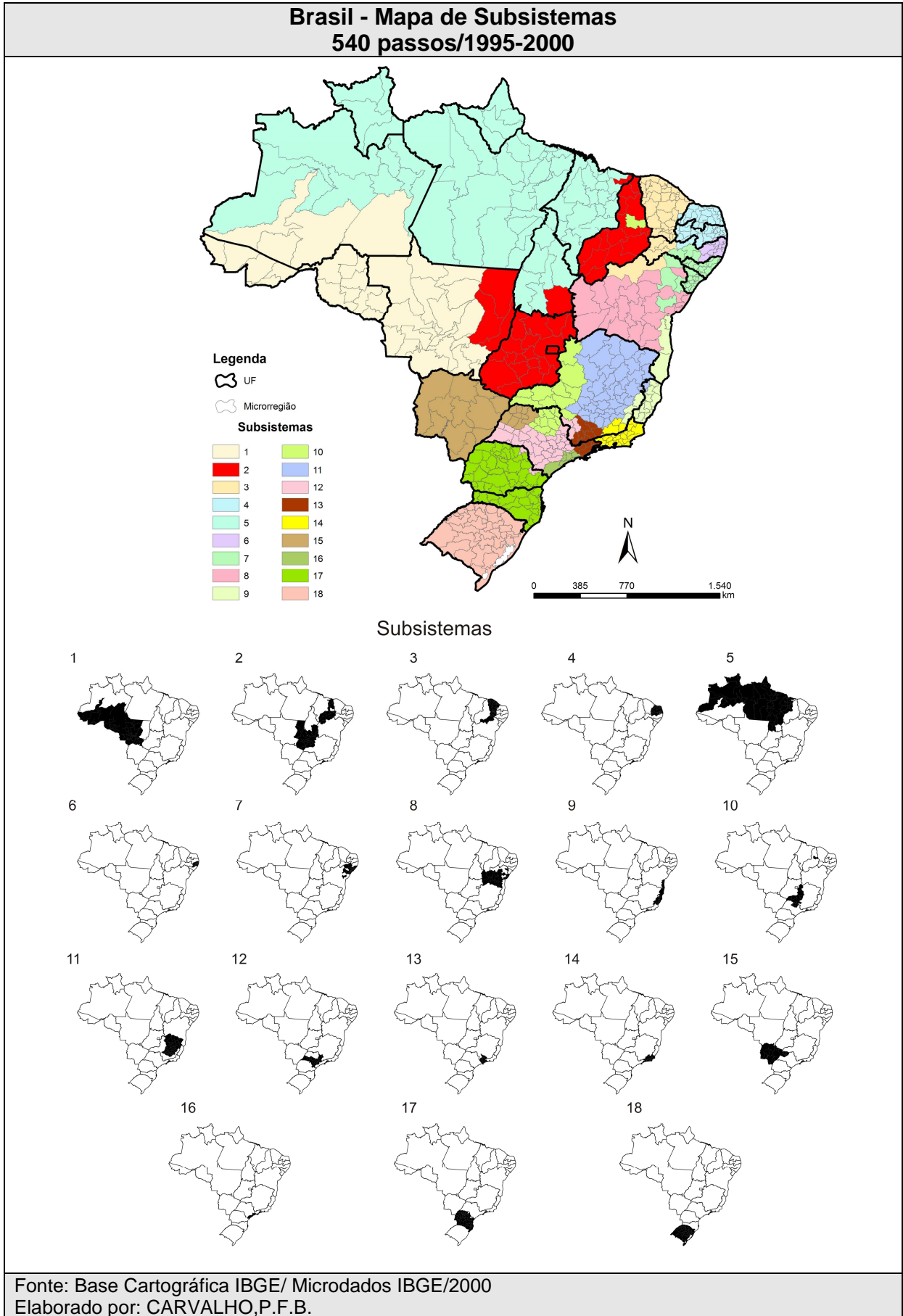
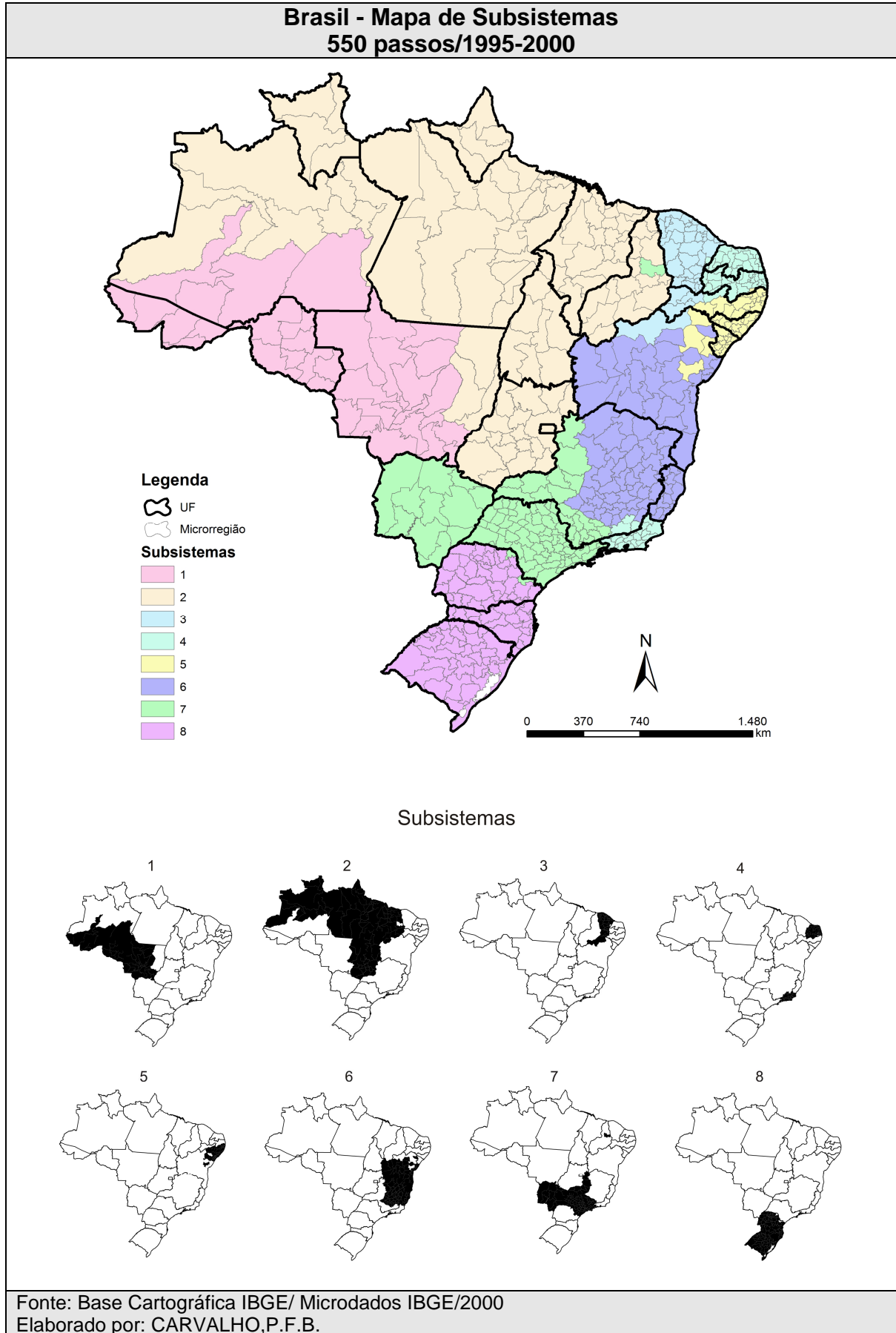


Figura 16: Subsistemas após aplicação de 540 passos do Método Intramax/1995-2000



Outra possibilidade de análise viabilizada com a aplicação do método Intramax refere-se ao acompanhamento do processo de agrupamento de uma determinada unidade espacial de interesse passo a passo. A seguir, faz-se a análise dos resultados obtidos para o estudo de duas microrregiões: Belo Horizonte/MG e Chapecó/SC.

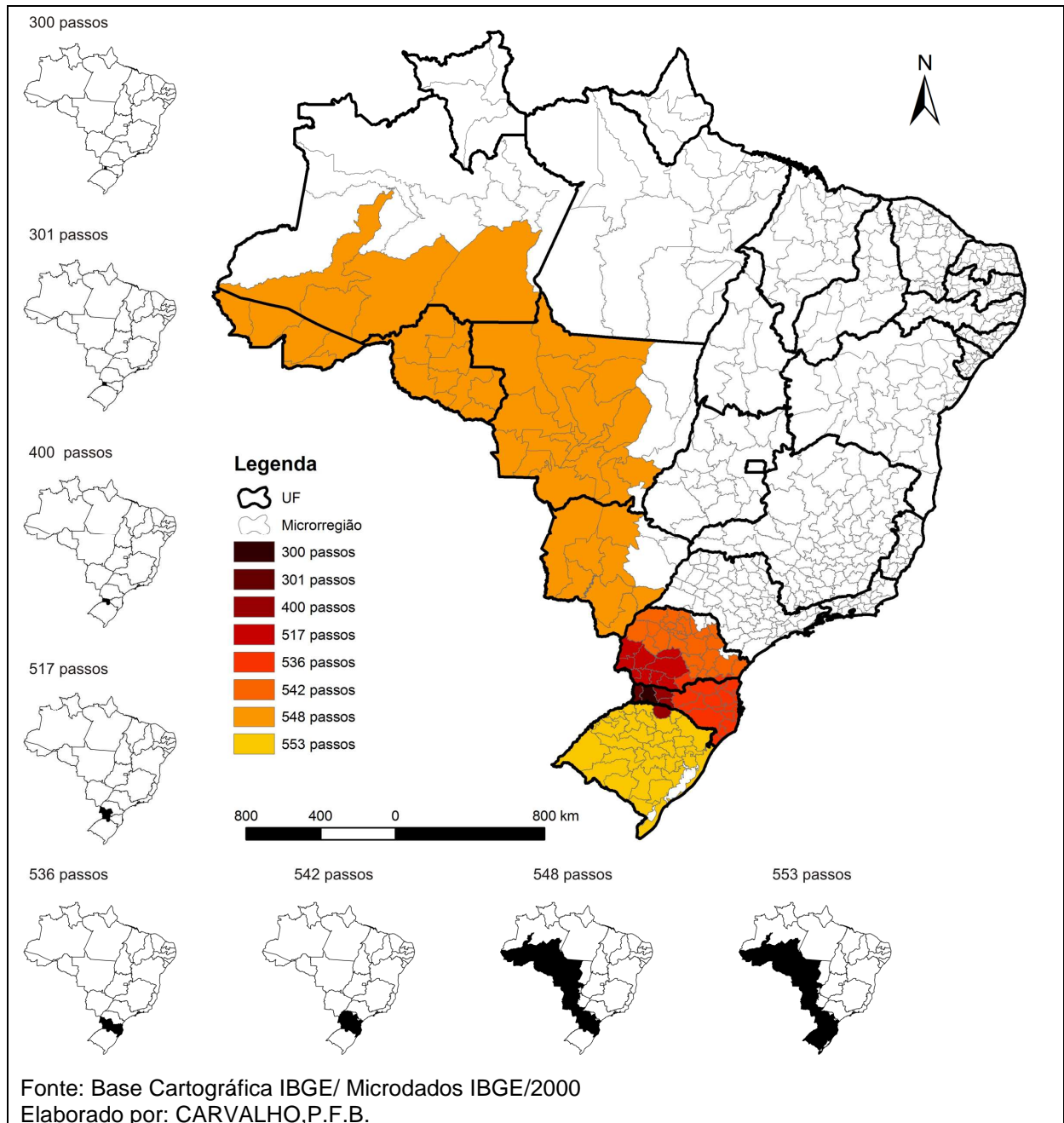


Figura 18: Estudo de caso para o período 1986-1991 – Microrregião de Chapecó/SC

As Figuras 18 e 19 mostram as representações cartográficas das incorporações da microrregião de Chapecó/SC aos diversos subsistemas de migração em que está

envolvida. O número de passos indica cada um dos momentos em que a microrregião ou o subsistema é incorporado a outro subsistema.

Observa-se que, no período de 1986-1991, a microrregião apresenta interação mais forte com as microrregiões de seu próprio estado, com a microrregião de Erechim, no Rio Grande do Sul, e com o estado do Paraná. Na sequência, este subsistema é incorporado ao grande subsistema no corredor oeste brasileiro, determinado por parte do estado do Mato Grosso do Sul, Mato Grosso, Rondônia, parte do Amazonas e Acre. Só então, é ligado com o estado do Rio Grande do Sul, mostrando os impactos ainda decorrentes da expansão da fronteira agrícola ao norte do país, apesar de denominada “Fronteira Consolidada” por Martine e Camargo (1984, p.121), já na década de 1970.

A microrregião de Chapecó expulsou 37.360 migrantes no período de 1986-1991, sendo 15.780 para outras microrregiões do próprio estado e 21.580 para outros estados. Apresenta saldo migratório negativo de quase 3.000 migrantes, em sua interação com o estado do Paraná, e de pouco menos que 3.500 migrantes, com os estados do Mato Grosso e Mato Grosso do Sul, no mesmo período. A troca de migrantes com os estados mais ao norte, como Rondônia, Acre e Amazonas, não é tão significativa, expulsando menos de 300 migrantes e recebendo menos de 100, apesar de o estado, como um todo, ter enviado mais de 2150 migrantes para Rondônia. Nesse período, o estado de Santa Catarina apresenta taxas de crescimento geométrico de sua população em redução. Essa taxa, que já foi de 3,2% ao ano, na década de 1960, é de 2,06% ao ano, na década de 1980, quase igual à média nacional de 1,93%.

A Figura 19 mostra as representações cartográficas para o período de 1991-2000, deixando bastante clara a mudança no panorama geral. Observa-se, num primeiro momento, a interação da microrregião de Chapecó com o restante de seu estado, depois com o estado do Paraná e fechando o subsistema com Rio Grande do Sul. A interação com as regiões Centro-Oeste e Norte já não são tão significativas, mostrando a tendência à migrações de curta distância.

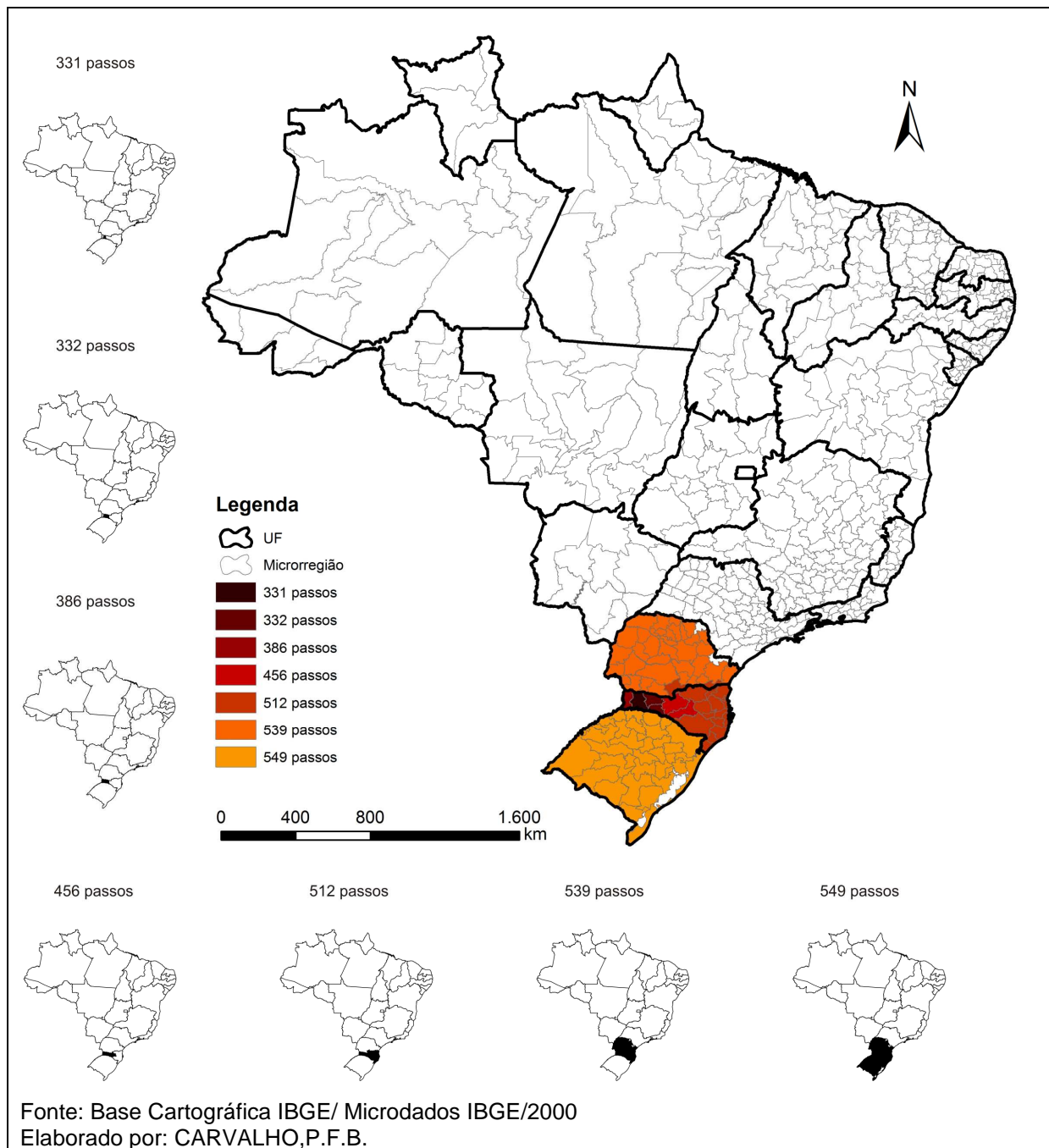


Figura 19: Estudo de caso para o período 1995-2000 – Microrregião de Chapecó/SC

O fluxo da microrregião de Chapecó para as microrregiões de Mato Grosso e Mato Grosso do Sul reduz-se em 40%, e do total de migrantes, 33% têm como destino o Paraná, ficando com saldo migratório negativo de 2.185 migrantes na interação com esse estado, ao passo que 37% migram para outras microrregiões do próprio estado de Santa Catarina.

Na década de 1990, a taxa de crescimento geométrico da população residente em Santa Catarina mantém a tendência de redução, chegando a 1,87% ao ano, ficando próxima da média nacional de 1,64%.

Passando para o estudo da microrregião de Belo Horizonte, considerada como a principal região de distribuição populacional do estado de Minas Gerais, observa-se, na Figura 20, a representação cartográfica da evolução de sua incorporação dentro dos subsistemas no período de 1986-1991.

Inicialmente, a microrregião de Belo Horizonte forma um subsistema não contíguo com as microrregiões de Conceição do Mato Dentro, Guanhães, Peçanha, Diamantina e Capelinha, seguido pela incorporação da porção centro-sul do estado com Bom Despacho e Piuí, mais à oeste, Ponte Nova e Viçosa, a leste, e Varginha, Andrelândia e São Lourenço, ao sul.

Observa-se que as microrregiões do extremo sul do estado só são somadas a um mesmo subsistema com a microrregião de Belo Horizonte, quando o estado de São Paulo também é adicionado, reforçando o que já foi dito anteriormente sobre a influência preponderante deste estado sobre a porção sul do estado de Minas Gerais.

Algo semelhante ocorre com a Zona da Mata mineira, que só é incorporada a um mesmo subsistema, quando os estados do Rio de Janeiro e Espírito Santo são adicionados, reforçando a tese do predomínio da atração do estado fluminense sobre as microrregiões de Juiz de Fora, Ubá e Cataguases.

Observa-se a incorporação das microrregiões dos estados do Rio Grande do Norte e da Paraíba, seguida pelo restante do Nordeste, antes da incorporação do estado de São Paulo ao subsistema que contém a microrregião de Belo Horizonte.

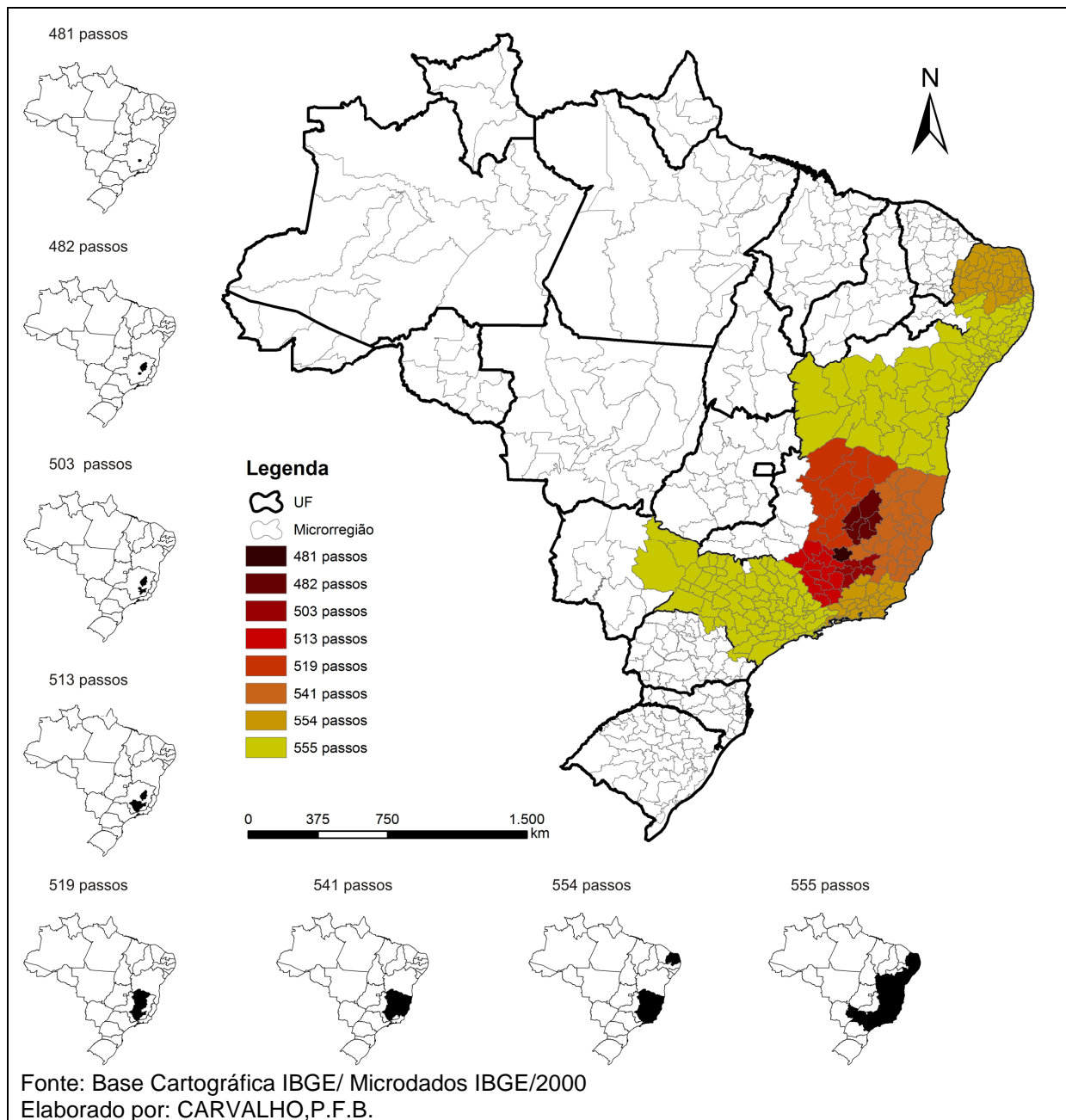


Figura 20: Estudo de caso para o período 1986-1991 – Microrregião de Belo Horizonte/MG

As regiões do Triângulo Mineiro, Alto Paranaíba e Oeste mineiro só são incorporadas ao mesmo subsistema da microrregião de Belo Horizonte nos passos finais, mostrando maior interação desta região com o Centro-Oeste brasileiro do que com o estado de Minas Gerais e, particularmente, Belo Horizonte, grande pólo do estado.

A Figura 21 apresenta os resultados obtidos na incorporação da microrregião de Belo Horizonte nos diversos subsistemas para o período de 1995-2000, mostrando um panorama bastante semelhante ao obtido para o período anterior.

Observa-se um padrão quase idêntico nos três quadros iniciais (até 498^o passo), com pequena diferença nos dois quadros seguintes, quando se detecta um corredor que parte do norte de Minas Gerais, abrangendo as microrregionais do centro do estado e entorno de Belo Horizonte e alcançando, também, o Vale do Mucuri, com destaque para Teófilo Otoni e Nanuque.

Evidencia-se, neste período, a relação das fronteiras ao Sul de Minas e da Zona da Mata com os estados de São Paulo e Rio de Janeiro, respectivamente. Observe que as microrregiões de Cataguases, Juiz de Fora e Ubá só são somadas ao subsistema quando a relação com o estado do Rio de Janeiro se torna significativa, como mostrado na figura referente ao 554^o passo.

O subsistema formado pelo Triângulo Mineiro, Alto Paranaíba e o Noroeste mineiro só é incorporado a um mesmo subsistema da microrregião de Belo Horizonte quando ocorre a adição do estado do Mato Grosso do Sul, reforçando, mais uma vez, a maior interação desta região com o Centro-Oeste brasileiro do que com o restante do estado de Minas Gerais.

Outro destaque evidenciado nesta análise fica para o descolamento dos estados da região sul do Brasil, que formam um subsistema isolado do restante do Brasil.

Com esta análise, espera-se ter mostrado a aplicabilidade e o nível de contribuição do Método Intramax na análise de fluxos migratórios, mas que pode ser utilizada no estudo de outros tipos de fluxos, como aqueles motivados por questões educacionais, por coorte, de mercadorias, para atendimento hospitalar, etc.

Destaca-se a viabilidade de uma análise global, avaliando o relacionamento entre muitas unidades espaciais, detectando o padrão de interação entre elas, ou de caráter específico, ao acompanhar a evolução do agrupamento e relacionamento de uma unidade espacial específica.

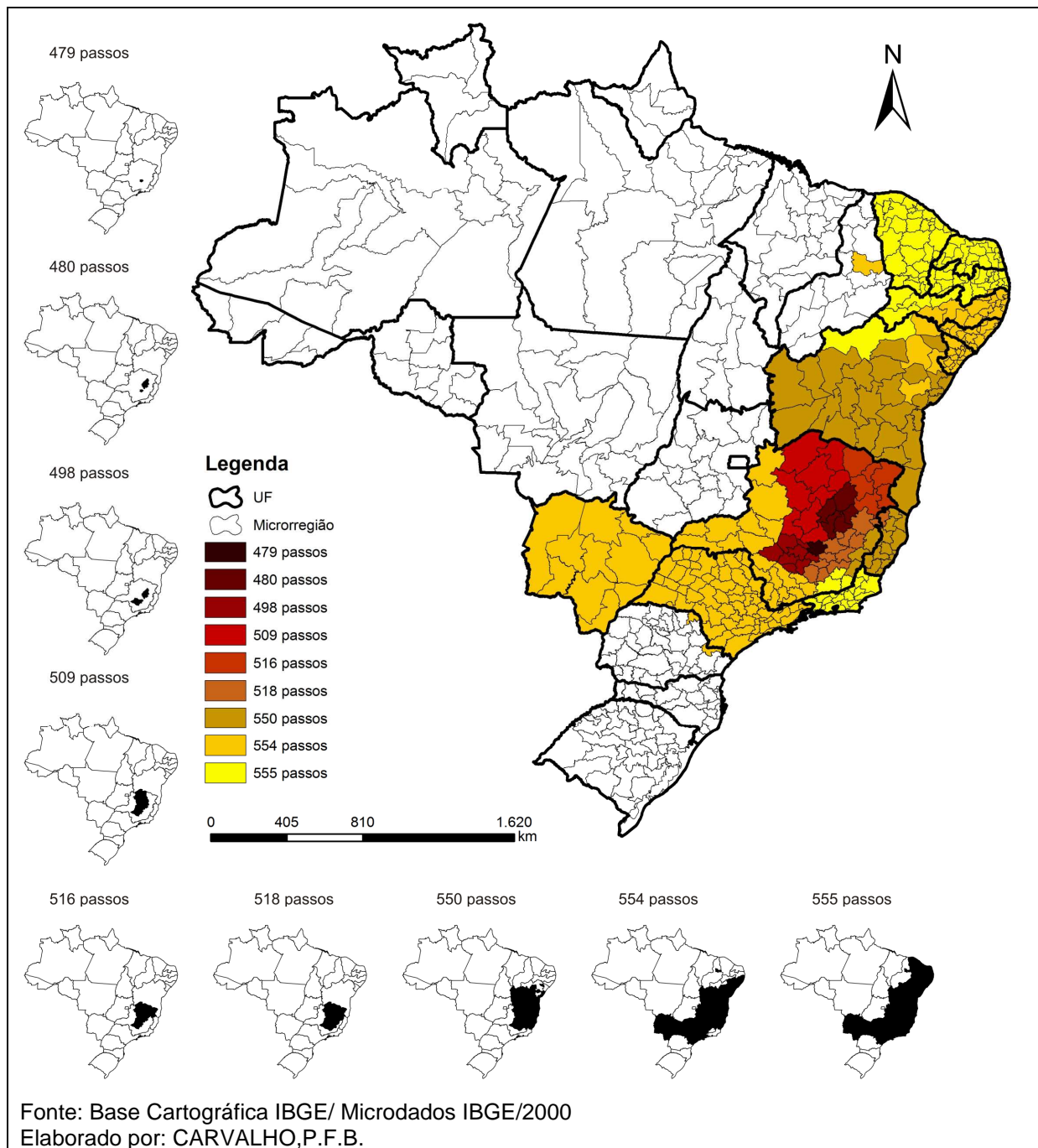


Figura 21: Estudo de caso para o período 1995-2000 – Microrregião de Belo Horizonte/MG

5.3 Tipologias de migração

Conforme exposto anteriormente, na Seção 3.2, faz-se a aplicação da Análise Fatorial de Alta Ordem com extração de fatores por Análise de Componentes Principais e aplicação da rotação Promax, para identificação de regiões de migração

como grupos de unidades territoriais que apresentam padrões similares de trocas com outras unidades, denominadas **tipologias de migração**.

Com a aplicação no modo R, onde as origens são dispostas nas linhas da matriz de fluxos e os destinos, nas colunas, são determinadas tipologias que representam agrupamentos de microrregiões que são destinos comuns e suas principais origens. No modo Q, as tipologias determinadas são formadas por microrregiões que são origens comuns e seus principais destinos.

Para os dois períodos estudados, observa-se a forte tendência à contigüidade na determinação das tipologias de migração, o que confirma um fenômeno bastante conhecido em estudos nessa área, ou seja, a tendência de os migrantes escolherem, pelo menos em um primeiro momento, áreas mais próximas.

O número de fatores considerados significativos na primeira ordem fatorial do modo R, para o período 1986-1991, é bastante superior àquele obtido para o período de 1995-2000: são 104, para o primeiro período, e 88, para o segundo. Isso sugere maior heterogeneidade, ou seja, diversidade de destinos, no padrão migratório no final da década de 1980. Este fracionamento mais intenso do território fica mais evidente nas regiões Norte, Centro-Oeste e parte do Nordeste brasileiro. Isso deve estar indicando que a consolidação da fronteira agrícola provavelmente foi um fenômeno bem menos sincrônico do que sugere a literatura.

Nos resultados de segunda ordem, quando os agrupamentos são um pouco maiores, a diferença se torna menos significativa, sendo de 30 tipologias para o primeiro período, e 27, para o segundo.

A extração de fatores é interrompida na 3ª ordem da Análise Fatorial, já com um menor nível de detalhamento, quando são registradas, coincidentemente, 10 tipologias, no modo R, e 9, no modo Q, para cada um dos períodos.

A discussão feita neste momento está baseada, principalmente, nos resultados obtidos na Análise Fatorial de segunda e terceira ordens. Apenas quando necessário, ou seja, para melhor compreensão do perfil de alguma tipologia em

especial, recorre-se aos resultados de primeira ordem, tendo em vista que esses resultam em elevado número de agrupamentos, apresentando, portanto, excesso no detalhamento da rede migratória, o que dificulta a identificação do padrão existente. Mas, caso exista o interesse em uma análise mais pormenorizada, deve-se usar a ordem inicial.

Na identificação de destinos similares para o período de 1986-1991, são registrados dez grandes agrupamentos na Análise Fatorial de terceira ordem. Merece atenção o fato de as microrregiões de São Paulo e Rio de Janeiro se destacarem como importantes origens em quase todas as tipologias definidas. Os resultados estão representados cartograficamente na Figura 22 e pela coleção de mapas no Apêndice F.

Ao aumentar o nível de detalhamento, quando são avaliados os resultados de segunda e primeira ordens, fica mais clara a maior influência, não só da microrregião de São Paulo, mas de boa parte de seu estado como um todo, na função de redistribuição populacional para o território brasileiro.

Ao norte, destaca-se a tipologia determinada pelos estados do Acre, Amapá, quase a totalidade do Amazonas, exceto pela microrregião de Madeira, Itaituba e Obidos, no leste do Pará, acrescido de sua porção norte e quase todo o estado do Ceará. Esse agrupamento recebe migrantes, principalmente, das microrregiões das capitais Belém, Manaus, Fortaleza e Macapá. As microrregiões de São Paulo e Santarém também se destacam como importantes origens, além de outras microrregiões dos estados do Pará, Rondônia, Maranhão e Paraná e de Brasília.

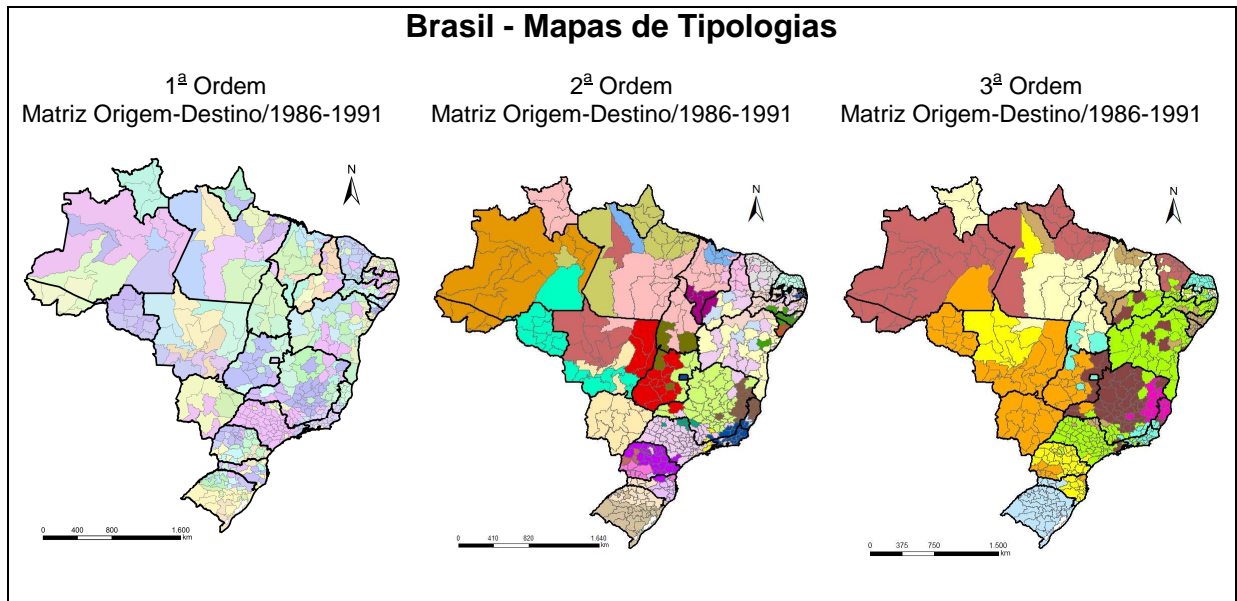
O estado de Rondônia, a microrregião de Madeira/AM, todo o Mato Grosso do Sul, as porções leste e sul do Mato Grosso, quase todo o estado de Goiás, a microrregião de Uberlândia e parte do oeste do Paraná e Joinville, em Santa Catarina, formam outra tipologia, que recebe migrantes principalmente das microrregiões que a compõem, mostrando uma redistribuição interna da população dos estados de São Paulo, Rio de Janeiro, e do Triângulo Mineiro, particularmente Ituiutaba, Uberlândia e Frutal.

A porção centro-sul do Pará, incluindo a microrregião de sua capital, forma outra tipologia, com ramificações no norte do Mato Grosso, todo o estado de Roraima, norte do Tocantins, região central do Maranhão, norte do Piauí e algumas microrregiões não contíguas da Paraíba, com migrantes originando especialmente do Rio de Janeiro, São Paulo, Fortaleza, Teresina, Brasília e João Pessoa.

Uma tipologia, com significativa descontinuidade, é formada pelo noroeste do Mato Grosso, Santarém no Pará, quase todo o estado do Paraná, porção centro-leste de Santa Catarina, recebendo migrantes principalmente dos estados de Santa Catarina e do Paraná e da microrregião de São Paulo. Provavelmente esta tipologia fica assim definida pela similaridade das origens, pois algumas microrregiões de Santa Catarina e do Paraná, além de redistribuírem população para o próprio estado, ainda expulsam migrantes para parte do Norte e Centro-Oeste, especialmente para atividades do agronegócio.

Ao analisar o fator mais forte na segunda ordem, ou seja, com maior autovalor, observa-se a tipologia determinada por quase todo o estado de São Paulo, o extremo sul de Minas Gerais, ramificações no norte de Minas Gerais e no Nordeste brasileiro, especialmente sudoeste da Bahia. Ampliando o agrupamento, os resultados de terceira ordem mostram que o entorno da cidade de São Paulo continua fortemente ligado com o restante do Nordeste brasileiro. Essa tipologia registra, como principais origens, as microrregiões de São Paulo, das capitais Salvador, Recife, Aracajú, Brasília e com Belo Horizonte, contribuindo, mais fortemente, para as microrregiões de seu próprio estado.

Analisando os resultados obtidos no Modo R, com a matriz na forma Destino-Origem, as microrregiões são agrupadas por origens com comportamentos semelhantes no fluxo migratório e são identificados os principais destinos de cada tipologia. Inicia-se a discussão dos resultados para o período 1986-1991, apresentados na Figura 23 e no Apêndice G.



Fonte: Base Cartográfica IBGE/ Microdados IBGE/2000
Elaborado por: CARVALHO,P.F.B.; VIEIRA JUNIOR, J.A.

Figura 22: Tipologias determinadas nas três primeiras ordens: Matriz Origem-Destino/1986-1991

É mais evidente a forte tendência à contigüidade na formação dos agrupamentos, sugerindo que as microrregiões mais próximas tendem a ser origens para destinos comuns.

Além disso, o número de agrupamentos determinados na aplicação de primeira ordem reduz-se de 104 para 79, o que sugere que os agrupamentos determinados pelos destinos escolhidos apresentam maior variabilidade, são mais heterogêneos que os agrupamentos definidos pelas origens dos migrantes. Este fato também pode ser observado nas ordens superiores, caindo de 30 para 26 tipologias na segunda ordem e de 10 para 9 na terceira.

Observa-se um grande agrupamento de origens de migrantes ao norte do território brasileiro, com a totalidade dos estados do Amazonas, Acre, Roraima, Amapá, parte de Rondônia e porções norte e oeste do Pará. Essa tipologia envia migrantes para os grandes centros da própria região, como Belém, Santarém, Manaus, Porto Velho e Boa Vista, mas, também, para as microrregiões da Aglomeração Urbana de São Luís, Rio de Janeiro e Belo Horizonte. Do estado de São Paulo, destaca-se como importante destino apenas a microrregião de Santos.

A tipologia determinada por quase todo o estado do Mato Grosso é complementada pelas microrregiões de Vacaria, Sananduva e Guaporé, ao norte do estado do Rio Grande do Sul. Esse agrupamento apresenta como principais destinos algumas microrregiões do Rio Grande do Sul, Santa Catarina e Paraná, como Caxias do Sul, Chapecó, Foz do Iguaçu, Cascavel, Francisco Beltrão e Curitiba, além daquelas do próprio Centro-Oeste, como SINOP e Entorno de Brasília, Salvador na Bahia e Rio de Janeiro.

Essa tipologia sugere a necessidade de se aprofundar estudos sobre a migração de retorno referente a, anteriormente denominada, Fronteira Agrícola em expansão, pois esse agrupamento pode estar enviando migrantes para as microrregiões da própria tipologia e, também, para aquelas que lhe abasteceram de migrantes em períodos anteriores, como os estados do sul do Brasil e do entorno de Brasília.

O centro-oeste do Pará, todo o estado do Tocantins, quase a totalidade de Goiás, e do Nordeste brasileiro formam outra tipologia complementada pelo Triângulo Mineiro e noroeste de Minas Gerais, além de Presidente Prudente e Assis, no estado de São Paulo. Esse agrupamento registra como principais destinos Brasília e a microrregião de São Paulo, além dos principais centros do Nordeste como Teresina, Fortaleza, Natal, Imperatriz, no Maranhão, e João Pessoa. O Centro-Oeste brasileiro se apresenta como importante destino, além da microrregião do Rio de Janeiro.

Para a tipologia definida por quase a totalidade dos estados de Alagoas, Sergipe e as porções oeste e leste do Pernambuco, observam-se fluxos mais significativos para as microrregiões de São Paulo e Santos, no estado de São Paulo, grandes centros do Nordeste, como Recife, Maceió, Aracajú, Petrolina-PE, João Pessoa e Natal, e mais Brasília e Rio de Janeiro.

A Paraíba forma uma tipologia com todo estado do Espírito Santo, exceto pela microrregião de Vitória, a microrregião de Porto Seguro, no extremo sul da Bahia, e as microrregiões fluminenses de Cantagalo-Cordeiro e Santa Maria Madalena. Como principais destinos se destacam Vitória, no Espírito Santo, Rio de Janeiro, Brasília, algumas microrregiões do Nordeste, especialmente João Pessoa e

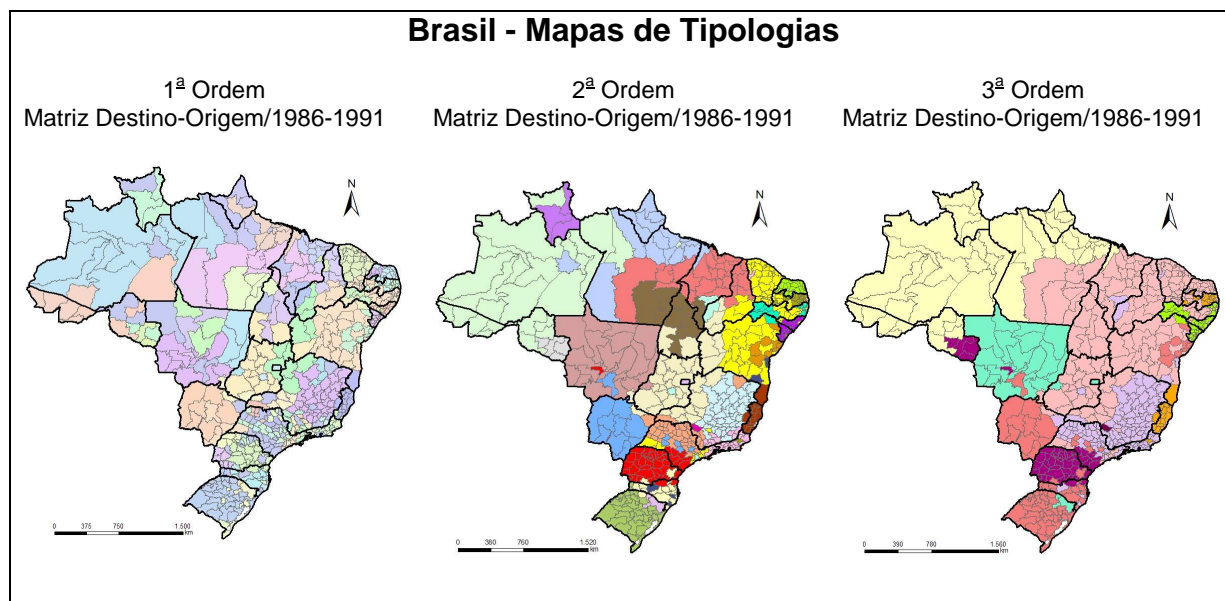
Salvador, além de outras dos estados de Minas Gerais e Rio de Janeiro, com destaque para Belo Horizonte, e a microrregião de São Paulo.

A tipologia com maior autovalor e, portanto, estatisticamente mais significativa, é delimitada pelo estado de Minas Gerais, sem o Triângulo Mineiro e o oeste, a porção centro-norte do estado de São Paulo e algumas poucas microrregiões não contíguas espalhadas pelo território brasileiro, com destaque para Goiânia-GO e Valença-BA. Os principais destinos para os migrantes desse agrupamento são, essencialmente, microrregiões do estado de São Paulo, com destaque para São Paulo, Campinas, Santos, Sorocaba, Ribeirão Preto e São José dos Campos, e outras de Minas Gerais, em especial, Belo Horizonte, Montes Claros e Varginha, Rio de Janeiro e Brasília.

O estado do Paraná e o sul do estado de São Paulo formam uma tipologia complementada por microrregiões não contíguas em parte do estado de Rondônia e Alto Paraguai, no Mato Grosso. Os principais destinos são Campinas, Sorocaba, Osasco, Moji das Cruzes e outras no estado de São Paulo, Curitiba, Londrina e Maringá, no Paraná, Porto Velho, Cacoal, Vilhena e Alvorada do Oeste em Rondônia e outras do Centro-Oeste, em especial em Mato Grosso e Mato Grosso do Sul.

Finalmente, uma tipologia que se apresenta bastante fechada, determinada pela quase totalidade das microrregiões dos estados do Rio Grande do Sul e de Santa Catarina, provendo migrantes para o próprio agrupamento. Destaca-se que o estado do Rio Grande do Sul envia migrantes para as microrregiões de São Paulo, Rio de Janeiro, Brasília, Criciúma e Tubarão, em Santa Catarina, mas, principalmente, para o próprio estado.

Ao analisar os dados referentes ao período 1995-2000, confirma-se a forte tendência à contigüidade na determinação das tipologias de migração, o que sugere que microrregiões mais próximas tendem a apresentar padrões semelhantes no comportamento migratório. Os resultados podem ser vistos na Figura 24 e na coleção de mapas do Apêndice H.



Fonte: Base Cartográfica IBGE/ Microdados IBGE/2000
Elaborado por: CARVALHO,P.F.B.; VIEIRA JUNIOR, J.A.

Figura 23: Tipologias determinadas nas três primeiras ordens: Matriz Destino-Origem/1986-1991

No modo R da Análise Fatorial, em que as origens ocupam as linhas e os destinos figuram nas colunas, os agrupamentos de microrregiões que são destinos com comportamentos similares apresentam configuração muito parecida com o observado para o período anterior.

Os estados de Rondônia, Mato Grosso e Mato Grosso do Sul passam a delimitar uma grande tipologia com microrregiões contíguas, com ramificações no Amazonas, Acre e Pará. No período anterior, boa parte desse agrupamento apresentava significativa similaridade de comportamento, compondo uma tipologia complementada pelas microrregiões de Goiás e do Paraná, mas sem parte do Mato Grosso. Um fato importante a ser considerado é a mudança no perfil das principais origens dos migrantes desse agrupamento, que passa a interagir mais fortemente com microrregiões do próprio grupo, agora mais limitado, e perde força na interação com outros estados, principalmente aqueles do Sul.

Como complemento dos resultados expostos no parágrafo anterior, observa-se que os estados do sul do país também se mostram mais fechados, com Paraná e Santa Catarina formando uma tipologia com origens quase essencialmente dos próprios estados. No período anterior era observada uma maior interação com estados do Centro-Oeste.

O Rio Grande Sul continua praticamente isolado, mostrando, como principal característica, a redistribuição interna de população e baixa interação com outros estados, exceto ao seu norte, com o estado de Santa Catarina.

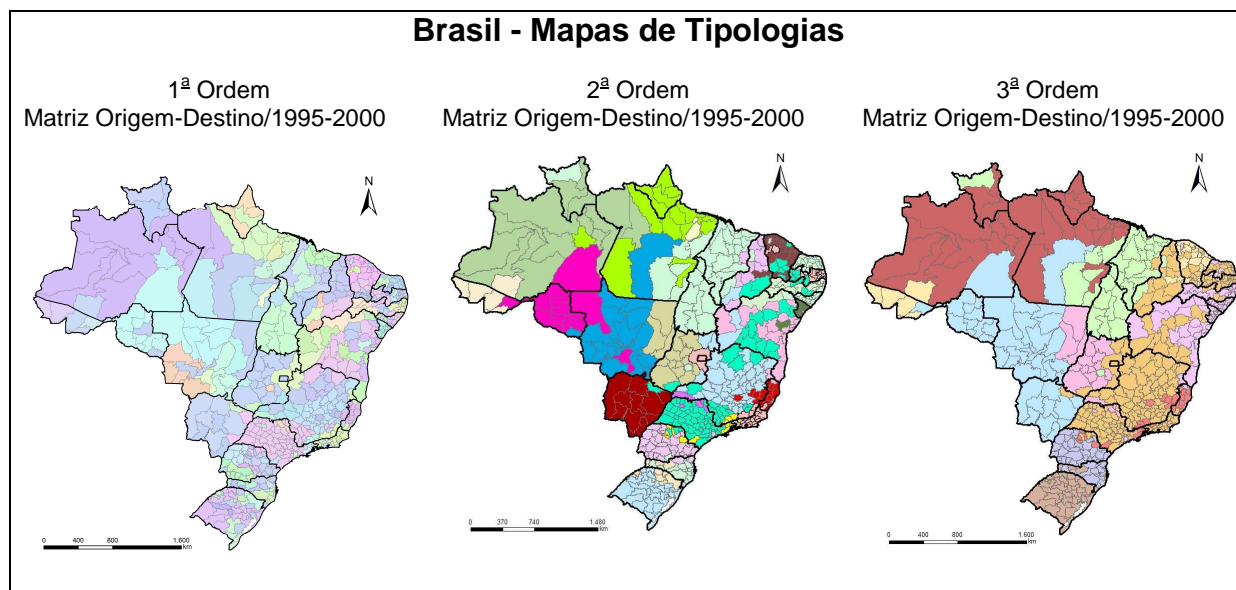
Voltando os olhos para o Norte do país, observa-se Roraima com comportamento similar ao das microrregiões do Amazonas, Amapá e parte do Pará, registrando movimento de imigrantes do Norte do país, ou seja, com redistribuição interna, e com São Paulo destacando-se como uma das principais origens do grupo.

O estado de Goiás se isola daquele grande grupo formado por Rondônia, quase todo o Mato Grosso e Mato Grosso do sul, se dividindo dentro de duas tipologias. A porção oeste e sul do estado mostra maior similaridade de comportamento de destino apenas com o leste do Mato Grosso, formando uma tipologia abastecida por imigrantes, principalmente, do estado de São Paulo, microrregiões próximas do Triângulo Mineiro e do nordeste mineiro, como Teófilo Otoni, Capelinha e Araçuaí.

A porção leste de Goiás, incluindo aí o entorno do Distrito Federal, apresenta-se agrupada com quase todo o estado de Minas Gerais, São Paulo, Rio de Janeiro e microrregiões do Nordeste brasileiro, principalmente sul da Bahia, Piauí, Rio Grande do Norte e Paraíba. Observa-se que essa configuração é conseqüência da forte interação de sua microrregiões com os migrantes do Distrito Federal, São Paulo e Rio de Janeiro.

A Figura 25 apresenta os resultados obtidos com a aplicação da Análise Fatorial no modo Q para o período de 1995-2000, ou seja, com os agrupamentos determinados pelas similaridades no status de origens de migrantes. Veja coleção de mapas no Apêndice I.

Nesta abordagem, os resultados da Análise Fatorial de Terceira Ordem são praticamente idênticos aos observados para o período anterior, registrando fortíssima tendência à contigüidade na delimitação das tipologias e sugerindo que microrregiões mais próximas tendem a ser origens com características semelhantes de expulsão de migrantes.



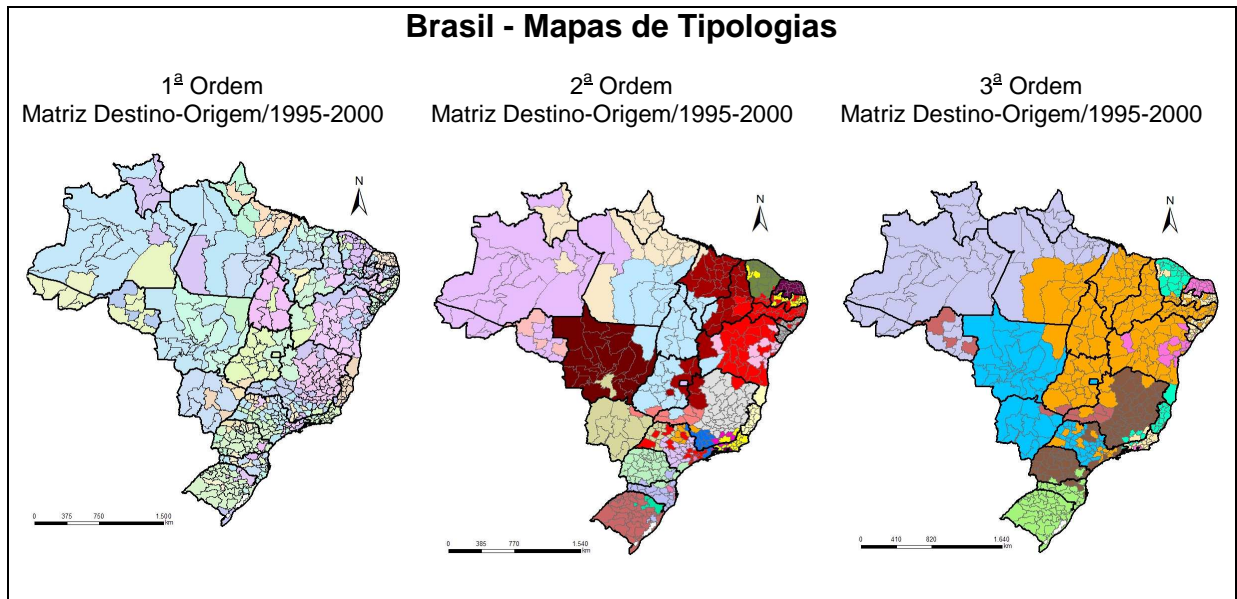
Fonte: Base Cartográfica IBGE/ Microdados IBGE/2000
Elaborado por: CARVALHO, P.F.B.; VIEIRA JUNIOR, J.A.

Figura 24: Tipologias determinadas nas três primeiras ordens: Matriz Origem-Destino/1995-2000

Destaca-se a configuração delimitada pelos estados do Mato Grosso e Mato Grosso do Sul, com ramificações pelo estado de São Paulo. Essa tipologia mostra perda de força no relacionamento com os estados de Santa Catarina e Rio Grande Sul.

Destaca-se, para este período, o desprendimento do estado do Ceará do restante do Nordeste, para formar uma tipologia com o estado do Espírito Santo, extremo sul da Bahia, as microrregiões da Zona da Mata, no estado de Minas Gerais, e Barra do Piraí e Três Rios, no estado do Rio de Janeiro. Esse agrupamento apresenta como principais destinos as microrregiões do Rio de Janeiro, Vale do Paraíba Fluminense e de Juiz de Fora, em Minas Gerais. Provavelmente, esse descolamento ocorre pela semelhança no padrão de relacionamento das microrregiões cearenses com aquelas do estado do Rio de Janeiro, ou seja, o significativo fluxo de migrantes para a microrregião do Rio de Janeiro.

Vale destacar a tipologia com microrregiões não contíguas composta por uma porção do estado de Rondônia e o Triângulo Mineiro. Esse agrupamento registra como principais destinos para seus migrantes Goiânia e Sudoeste de Goiás, microrregiões do Mato Grosso como Parecis, Rondonópolis, Alto Pantanal, Aripuana e Médio Araguaia, além do próprio Triângulo Mineiro, com destaque para Uberlândia.



Fonte: Base Cartográfica IBGE/ Microdados IBGE/2000
Elaborado por: CARVALHO,P.F.B.; VIEIRA JUNIOR, J.A.

Figura 25: Tipologias determinadas nas três primeiras ordens: Matriz Destino-Origem/1995-2000

Com a discussão dos resultados, espera-se ter deixado clara a contribuição que a Análise Fatorial de Ordem Superior oferece ao estudo de fluxos. Destaca-se a qualidade de esta metodologia apontar não só as unidades espaciais que se assemelham no comportamento de envio ou recebimento de migrantes, mas, também, de mostrar as unidades espaciais com as quais a interação ocorre com maior intensidade.

Também deve-se ressaltar que, de acordo com a qualidade dos dados, a análise pode ser feita em níveis diferenciados. Desde um nível mais detalhado, com agrupamentos menores, até um nível global, mais generalizante, com agrupamentos maiores. De acordo com os objetivos do trabalho, o pesquisador deve optar pelo nível de detalhamento mais adequado.

5.4 Uma síntese dos resultados

Os resultados obtidos com a aplicação das metodologias propostas permitem que o território brasileiro seja regionalizado em níveis diferenciados de detalhamento e por princípios que levam em conta a similaridade de comportamento no envio ou recebimento de migrantes ou o grau de interação entre as unidades espaciais. Também é possível identificar o nível de interação de uma determinada unidade espacial com as demais.

Algumas observações relevantes podem ser observadas nos resultados analisados. A participação do estado de São Paulo na redistribuição populacional pelo território brasileiro ainda se mostra muito significativa, apesar de apresentar alterações no padrão. O interior do estado ganha importância, tanto como fonte de migrantes como áreas de ocupação e reorganização espacial com pólos fora da Grande São Paulo. Por exemplo, o norte de São Paulo mostra fortes interações com o sul de Minas Gerais e com o Centro-Oeste.

O estado de São Paulo apresenta a característica de ser origem de emigrantes num grau mais elevado do que observado em períodos anteriores, seja pela possível migração de retorno, a ser avaliado em outros estudos, ou como formação de mão-de-obra mais especializada, dado o maior número de instituições de formação.

O estado do Rio de Janeiro perde importância no cenário migratório brasileiro, mas, ainda, não pode ser desconsiderado, principalmente pela dimensão de sua região metropolitana, que foi destino de muitos migrantes em décadas passadas.

Os demais grandes centros urbanos espalhados pelo Brasil, como, por exemplo, Brasília, Belo Horizonte, Porto Alegre, Curitiba, Recife, etc. continuam capitaneando a redistribuição espacial da população em seus estados, apesar de novos centros surgirem e se tornarem pólos regionais, formando delimitações regionais de grande influência. Mais uma vez, um exemplo típico é observado pela microrregião de Uberlândia, no Triângulo Mineiro, que polariza os movimentos de seu entorno, incluindo de microrregiões dos estados vizinhos, tanto como emissor como receptor de migrantes. No período de 1995-2000, essa microrregião recebeu quase 10.000

migrantes do estado de São Paulo, mais de 13.000 do estado de Goiás, aproximadamente 2.200 da Bahia e mais 2.100 dos estados do Mato Grosso e Mato Grosso do Sul.

As regiões antes denominadas de “Fronteiras de Expansão” mostram-se mais consolidadas, ainda recebendo muitos migrantes, mas, também distribuindo migrantes internamente ou para outras regiões. Como exemplo, o saldo migratório do conjunto dos estados de Rondônia, Mato Grosso e Mato Grosso do Sul, que foi de 166.379, no período 1986-1991, cai para 42.787, no período 1995-2000. Além disso, esses estados receberam quase 120.000 migrantes provenientes de Santa Catarina e Paraná no primeiro período e quase a metade no segundo período.

Finalmente, o Nordeste brasileiro continua mostrando forte ligação com a região Sudeste, em especial com São Paulo e Rio de Janeiro, mas, seus pólos regionais, principalmente as capitais, ganham importância neste cenário.

6 CONSIDERAÇÕES FINAIS

A análise dos resultados alcançados mostra que as metodologias propostas para o estudo de fluxos, ou seja, a do Método Iterativo Intramax e a da Análise Fatorial de Alta Ordem são complementares, mas, também, podem ser exploradas independentemente no estudo da migração ou de fluxos de modo geral.

São apresentadas considerações sobre as metodologias propostas, suas características principais (ver Quadro 2), similaridades, diferenças e potencialidades. Além disso, são registradas algumas propostas para estudos futuros.

Sobre a característica da base de dados que é evidenciada, observa-se que o método da Análise Fatorial promove agrupamentos baseando-se na similaridade de fluxos, criando tipologias que enviam migrantes para destinos semelhantes ou tipologias que recebem migrantes de origens comuns, não necessariamente determinadas pela dimensão do volume da interação, enquanto o método Intramax baseia-se na força de interação proporcional entre as unidades espaciais para determinar os subsistemas.

A respeito do número e da forma dos agrupamentos que são determinados, constata-se que a Análise Fatorial os promove em blocos de unidades espaciais, pois cada fator determina uma tipologia composta por unidades espaciais com alta carga fatorial e os escores indicam as unidades espaciais que interagem mais fortemente com a tipologia determinada. Portanto, o número de agrupamentos fica limitado ao número de fatores significativos obtidos.

O método Intramax delimita os agrupamentos passo a passo. A cada passo iterativo uma unidade espacial é unida a outra, ou a outro agrupamento definido nos passos anteriores. Sendo assim, o número máximo de subsistemas é igual ao número de unidades espaciais básicas menos um. No estudo de caso avaliado, como existem 558 microrregiões, no máximo 557 subsistemas podem ser determinados. Esta escolha depende do grau de detalhamento desejado pelo pesquisador.

Os principais critérios para identificação do número de agrupamentos na Análise Fatorial são baseados no autovalor, na porcentagem da variação retida ou pelo

gráfico Scree Plot, enquanto, no Intramax, o ponto de parada na delimitação dos agrupamentos é determinado pela porcentagem do fluxo computada e pelo grau de detalhamento desejado da rede migratória desejado.

Ao definir as unidades espaciais que determinam uma tipologia, com a aplicação da Análise Fatorial de Alta Ordem, é possível identificar as unidades espaciais que se relacionam no sentido inverso ao do agrupamento. Ou seja, ao determinar um grupo de origens com comportamentos similares, determinam-se os seus principais destinos pelos escores e vice-versa. Essa relação não é explicitada pelos resultados gerados com a aplicação do método Intramax.

Quanto à aplicação dos métodos, deve ser observado que a Análise Fatorial de Alta Ordem pede apoio de softwares estatísticos com pacote de funções de estatística multivariada e exige muito trabalho para geração de uma representação cartográfica, principalmente no caso de um grande número de unidades espaciais básicas. Para elaboração dos mapas, foram utilizados, de forma complementar, os recursos dos softwares Excel, SPSS e ArcGIS.

Sobre este último aspecto, o método Intramax leva vantagem, por apresentar uma lógica computacional simples, fácil de ser implementada com o apoio de qualquer linguagem de programação de alto nível e seus resultados são facilmente representados cartograficamente, tendo em vista a lógica de agrupamento, natural do processo iterativo. Simples rotinas computacionais podem gerar relatórios compatíveis com os principais Sistemas de Informação Geográfica. Neste trabalho, seus resultados foram representados cartograficamente apenas com o uso dos softwares ArcGIS e Flowap.

Análise Fatorial de Alta Ordem	Método Intramax
Agrupamento realizado baseando-se na similaridade de fluxos (unidades espaciais que enviam ou recebem fluxo de unidades semelhantes)	Agrupamento realizado baseando-se na interação entre unidades espaciais
Agrupamento realizado em blocos	Agrupamento realizado unidade espacial básica por unidade espacial básica
Número de agrupamentos limitado pelo número de fatores que atendem aos critérios	Até N-1 grupos, onde N é o número de unidades espaciais básicas
Principais critérios para identificação do número de agrupamentos: do autovalor, % de variação, screeplot.	Principais critérios para identificação do ponto de parada nos agrupamentos: % do fluxo computado, grau de generalização do agrupamento (subjetivo)
É possível identificar principais origens ou principais destinos dos grupos	Pode-se acompanhar a evolução de agrupamento passo-a-passo para uma unidade espacial
Pode ser aplicado por categorias de interesse, por exemplo, por coorte, nível de escolaridade, raça, outros fluxos.	Pode ser aplicado por categorias de interesse, por exemplo, por coorte, nível de escolaridade, raça, outros fluxos.
Pede apoio de um software estatístico	Lógica computacional simples, podendo ser implementado em qualquer linguagem de alto nível. Está disponível no software Flowmap
Representação cartográfica laboriosa para um grande número de unidades espaciais	Representação cartográfica facilitada pelos relatórios fornecidos pelo software Flowmap

Quadro 2: Principais características dos métodos abordados

Sobre as potencialidades das metodologias propostas, avalia-se que podem ser aplicadas como auxiliares ao estudo da migração, tanto numa abordagem exploratória inicial, quanto como suporte para confirmação de hipóteses. Os resultados discutidos nas Seções 4.2 e 4.3 sugerem a supremacia da migração de curta distância sobre a de longa distância no Brasil, corroborando com o entendimento de autores já citados neste trabalho nas Seções 2.1 e 4.2, pois a redução no número de fatores significativos do primeiro período para o segundo, aliada à observação da tendência à contigüidade dos agrupamentos, provoca esse tipo de interpretação. Também fica evidenciada a importância da microrregião de São Paulo e do interior do estado paulista na redistribuição da população pelo território brasileiro, como fonte e/ou destino de migrantes.

Os resultados encontrados referentes ao relacionamento entre os estados das regiões Sul e Centro-Oeste e aos estados do Nordeste sugerem a necessidade de estudos mais aprofundados sobre a relevância, ou não, da migração de retorno e sobre o perfil do migrante: escolaridade, faixa etária, etc.

A importância de centros urbanos como pólos regionais relevantes para a redistribuição da população, por exemplo, Uberlândia/MG, Campinas/SP, Entorno de Brasília/DF e outros, também pode ser observada nos resultados e deve ser estudada mais detalhadamente. Estes novos pólos são derivados, em sua maioria, de cidades médias ou centros especializados beneficiados pela desconcentração ou reorganização espacial das indústrias, apesar de concentradas, principalmente, no eixo Belo Horizonte-Porto Alegre, segundo Diniz (1993, p.57). O fortalecimento de municípios do interior de São Paulo é beneficiado pela realocação de indústrias, provavelmente provocada pelas “deseconomias de aglomeração” típicas dos grandes centros.

A possibilidade de representação cartográfica dos resultados obtidos para períodos diferentes é um interessante recurso, pois pode mostrar a mudança, ou não, no padrão espacial de comportamento do migrante. Caso típico foi o discutido para a microrregião de Chapecó-SC, que tem sua interação com a região Centro-Oeste claramente reduzida no período 1995-2000 e fica evidenciada com a aplicação do método Intramax.

Nos mapas obtidos com a aplicação do método Intramax, é possível visualizar uma mudança no padrão de comportamento dos migrantes dos estados do Sul do Brasil, em especial Santa Catarina e Paraná, pois, no período 1995-2000, passam a compor um agrupamento mais fechado dentro da própria grande região, enquanto no período anterior, 1986-1991, observa-se uma significativa interação destes dois estados com o Centro-Oeste brasileiro.

Finalmente, pensando em tarefas futuras para o próprio autor deste trabalho ou para outros que queiram se aventurar na complementação do que foi discutido, sugere-se o desenvolvimento de um sistema computacional que automatize as atividades mais laboriosas, em especial aquelas relacionadas à cartografia, pois não foi observada a

existência de um sistema que agrupe as metodologias, da forma como proposto, com representações gráficas de qualidade. Um ambiente a ser avaliado é o do Visual Basic for Applications (VBA), tendo em vista a existência de algumas aplicações complementares interessantes para uso no software ArcGIS, como o Flow Data Model Tools, desenvolvido por J. Alan Glenon, Mike Goodchild e Waldo Tobler, para uso no ArcGIS 9.0 ou versões posteriores.

Outra tarefa proposta é a de aplicação dessas metodologias em estudos de casos mais específicos. Pelo observado nos resultados analisados neste trabalho, acredita-se que importantes questões poderão ser explicitadas no estudo da migração. Por exemplo, em discussões sobre o padrão de fluxo migratório, quando os dados são filtrados por nível de escolaridade, ou faixa etária, os impactos do gênero do migrante no padrão de fluxo, etc., ou mesmo para universos de análises menores, como por estado ou por grandes regiões, acompanhados de um aprofundamento teórico sobre temas relacionados com os impactos das redes urbanas e sociais.

REFERÊNCIAS

ABREU, João Francisco de. Sistemas de informações Geográficas e manufatura integrada por computador; GIS e CIM uma análise exploratória. In: TERRA, L.D.B (editor). **Manufatura integrada por computador**. Belo Horizonte: Fundação CEFETMINAS, 1995. 249p.

ANDRADE, Thompson A. Métodos estatísticos e econométricos aplicados à Análise Regional. In FERREIRA, Carlos Maurício de Carvalho; HADDAD Paulo Roberto; BOISIER, Sérgio (Org.). **Economia Regional: teorias e métodos de análise**. Fortaleza: BNB, 1989. Cap. 9, p. 427-507.

ALVES, Vicente Eudes Lemos. A mobilidade sulista e a expansão da fronteira agrícola brasileira. **Agrária**. n.2,p. 40-68, 2005. Disponível em: <http://www.geografia.fflch.usp.br/revistaagraria/revistas/2/alves_vel.pdf>. Acesso em 20 fev. 2009.

ANJOS, Maria Anita. Retrospectiva da economia brasileira nos últimos 45 anos. **Revista FAE Business**, n.4, p.18-21, 2002. Blumenau, Disponível em: <http://www.fae.edu/publicacoes/pdf/revista_fae_business/n4_dezembro_2002/>. Acesso em 17 jan. 2007.

ARCGIS: Versão 9.2. Environmental Systems Research Institute, Inc., 1999-2006.

BAENINGER, Rosana. A nova configuração urbana no Brasil: desaceleração metropolitana e redistribuição da população. In: **Encontro Nacional de Estudos Populacionais**, ABEP, 11, Caxambu, 1998, p.729-772.

BAENINGER, Rosana. Deslocamentos populacionais, urbanização e regionalização. In: **Encontro Nacional sobre migração**, 2, 1999, Ouro Preto, **Anais...** p. 67-78.

BAENINGER, Rosana. Migrações internas no Brasil: municípios metropolitanos e não-metropolitanos. In: **Encontro Nacional sobre migração**, 2, Belo Horizonte, 2000, p.535-562.

BAENINGER, Rosana. Rotatividade migratória: um novo olhar para as migrações no século XXI. In: **Encontro Nacional de Estudos Populacionais**, ABEP, 16, Caxambu, 2008.

BAER, Werner. **A economia brasileira**. São Paulo: Nobel, 1996. 416 p.

BAUER, Kenneth W.; KOCHAR, Bipin; TALAVAGE, Joseph J. **Simulation model decomposition by factor analysis**. Disponível em: < <http://delivery.acm.org/10.1145/260000/253146/p185-auer.pdf?key1=253146&key2=4955749021&coll=GUIDE&dl=ACM&CFID=65854538&CFTOKEN=91297983> >. Acesso em: 10 jan. 2008.

BRITO, Fausto; MARQUES, Denise. As grandes metrópoles e as migrações internas: um ensaio sobre o seu significado recente. In: ENCONTRO NACIONAL SOBRE MIGRAÇÕES, IV, 2005, Rio de Janeiro. **Anais...** Rio de Janeiro: Associação Brasileira de Estudos Populacionais, 2005, p. 1-15. Disponível em: <<http://www.abep.nepo.unicamp.br/docs/anais/outros/4EncNacSobreMigracao/ST5-1.pdf>>.

Acesso em: 10 jan. 2008.

BROWN, Peter J.B.;PITFIELD, David E. An Intramax derivation of commodity market áreas from freight flow data. **Transportation Planning and Technology**, v.15, p.59-81, 1990.

CAMARA, Gilberto et al. **Análise espacial de dados geográficos**. São José dos campos, 2004. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>>. Acesso em 15 jul 2002.

CARROL, J. Douglas;GREEN, Paul E. **Mathematical tools for applied multivariate analysis**. San Diego: Academic Press,1997, 376p.

CARVALHO, José Alberto Magno de; SAWER, Diana Oya; RODRIGUES, Roberto do Nascimento. **Introdução a alguns conceitos básicos e medidas em demografia**. 2 ed. rev. São Paulo: ABEP, 1994, reimpr. 1998. Disponível em: <http://www.ifch.unicamp.br/pos/selecao/2008/texto_carvalho.pdf>. Acesso em: 10 jan. 2008.

CLARK, D.; DAVIES, W. K. D.; JOHNSTON, R. J. The application of factor analysis in Human Geography. **The Statistician**. v. 23, n. 3/4, p. 259-281, September/December 1974. Disponível em: <[http://www.jstor.org/sici?sici=0039-0526\(197409%2F12\)23%3A3%2F4%3C259%3ATAOFAI%3E2.0.CO%3B2-S](http://www.jstor.org/sici?sici=0039-0526(197409%2F12)23%3A3%2F4%3C259%3ATAOFAI%3E2.0.CO%3B2-S)>. Acesso em: 10 jan. 2008.

CLARK, David. The formal and functional structure of Wales. **Annals of the Association of American Geographers**. v. 63, n. 1, p. 71-84, mar. 1973.

CLAYTON, C. Hierarchically organized migration fields: the application of higher order factor analysis to population migration tables. **Annals of Regional Science**, 16, p.11-20, 1982.

CLAYTON, C. Interstate population migration process and structure in the United States, 1935 to 1970. **Professional Geographer**, 29, p.177-181,1977a.

CLAYTON, C. The structure of interstate and inter-regional migration: 1965-1970. **Annals of Regional Science**, 11, p.109-122, 1977b.

CLAYTON, Christopher. Interdependence in urban systems and its application to political reorganization. **Geografiska Annaler**. v. 62, n. 1, p. 11-20, 1980.

COLE, John P., KING, Cuchlaine A.M. **Quantitative Geography; Techniques and Theories in Geography**. 3.ed. London: John Wiley & Sons, 1970. 692p.

COLLINS, Lyndhurst; DREWETT, Roy; FERGUSON, Rob. Markov models in geography. **The Statistician**. v. 23, n. 3/4, p. 179-209, September/December 1974. Disponível em: <[http://www.jstor.org/sici?sici=0039-0526\(197409%2F12\)23%3A3%2F4%3C179%3AMMIG%3E2.0.CO%3B2-O](http://www.jstor.org/sici?sici=0039-0526(197409%2F12)23%3A3%2F4%3C179%3AMMIG%3E2.0.CO%3B2-O)>. Acesso em: 10 jan. 2008.

CONFIGURAÇÃO atual e tendências da rede urbana. Brasília: IPEA, IBGE, UNICAMP,

2002.

CORRAR, Luiz J.; PAULO, Edílson; DIAS FILHO, José Maria (Coords.). **Análise multivariada**: para os cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007. 541 p.

CORRÊA, Roberto Lobato. **Região e organização espacial**. 5.ed São Paulo: Atica, 1995. 93p.

DAMIANI, Amélia Luisa. **População e geografia**. 8. ed. São Paulo: Contexto, 2004. 107 p.

DINIZ, Clélio Campolina. Desenvolvimento poligonal no Brasil: nem desconcentração, nem contínua polarização. **Nova Economia**, Belo Horizonte, UFMG/FACE/DCE, v.31 n.11, p.35-64, set.1993.

DINIZ, Clélio Campolina; LEMOS, Mauro Borges (Org.). **Economia e território**. Belo Horizonte: Editora UFMG, 2005.

DORIGO, Guido; TOBLER, Waldo. Push pull migration laws. **Annals of the Association of American Geographers**. v. 73, n. 1, p. 1-17, 1983. Disponível em: <http://www.geog.ucsb.edu/~tobler/publications/pdf_docs/movement/migration/PushPull.pdf>. Acesso em: 10 jan. 2008.

DRUCK, S. *et al.* **Análise especial de dados geográficos**. Brasília: EMBRAPA, 2004. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>>. Acesso em: 23 mar 2006.

ELLIS, Mark; BARFF, Richard; RENARD, Beverly. Migration regions and interstate labor flows by occupation in the United States. **Growth and Change**. v. 24, p. 166-190, 1993.

EMRICK, Harry W. A method for inter-regional comparisons (Factor Analysis). **The Professional Geographer**. v.17, issue 3, p.18-19,1965.

FABRIKANT, Sara Irina *et al.* **The first law of cognitive Geography: distance and similarity in semantic space**. Disponível em: <http://www.geog.ucsb.edu/~sara/html/research/pubs/fabrikant_et_al_gis02.pdf>. Acesso em: 10 jan. 2008.

FABRIKANT, Sara Irina. Visualizing region and scale in information spaces. In: International Cartographic Conference, 20, 2001, Beijing, China. **Proceedings...** Beijing: 6-10 August, 2001, p. 2522-2529. Disponível em: <http://geog.ucsb.edu/~sara/html/research/pubs/fabrikant_ica01.pdf>. Acesso em: 10 jan. 2008.

FELDMAN, Olga *et al.* Creation of a system of functional áreas for England and Wales and for Scotland. In: EUROPEAN TRANSPORT CONFERENCE, 2005, Strasbourg. **Proceedings...**, Association for European Transport, London, 2005. Disponível em: <http://www.mvaconsultancy.com/papers>. Acesso em out. 2007.

FERREIRA, Carlos Maurício de Carvalho; HADDAD Paulo Roberto; BOISIER, Sérgio. **Economia Regional: teorias e métodos de análise**. Fortaleza: BNB, 1989.

FIELDS, Gary S. Place-to-place migration: some new evidence. **The Review of**

Economics and Statistics. v. 61, n. 1, p. 21-32, February 1979. Disponível em: <[http://www.jstor.org/sici?sici=0034-6535\(197902\)61%3A1%3C21%3APMSNE%3E2.0.CO%3B2-6](http://www.jstor.org/sici?sici=0034-6535(197902)61%3A1%3C21%3APMSNE%3E2.0.CO%3B2-6)>. Acesso em: 10 jan. 2008.

FISCHER *et al.* Telephone Communication patterns in Áustria: A comparison of the IPFP-based Graph-Theoretic and the Intramax approaches. **Geographical Analysis**, v.25,n.3, p.224-233, Jul.1993.

FLOWMAP: Versão 7.3. Faculty of Geographical Sciences, Utrecht University, the Netherlands, 2000.

FOTHERINGHAM, A.S.; O'KELLY, M.E. **Spatial interaction models: formulations and applications.** Dordrecht: Kluwer Academic Publishers,1989.

GERARDI, Lúcia Helena de Olivera; SILVA, Barbara-Christine Nentwig. **Quantificação em Geografia.** São Paulo: Difel,1981.

GERARDI, Lúcia Helena de Olivera. Memórias, histórias e interpretações da “revolução” Teorético-quantitativa em rio claro. In: **Colóquio Brasileiro de História do Pensamento Geográfico**, 1, Uberlândia, 2008. Disponível em: <<http://www.ig.ufu.br/coloquio/textos/Lucia%20Helena%20de%20Oliveira%20Gerardi.pdf>>. Acesso em 26 jun. 2009.

GODDARD, J. B. Functional regions within the city centre: a study by factor analysis of taxi flows in central London. **Transactions of the Institute of British Geographers.** n. 49, p. 161-182, Mar. 1970.

GOETGELUK, Roland; JONG, Tom de. Dynamic clusters in migration patterns: Intramax-analyses of inter-municipal migration flows between 1990 en 2004. Workshop 3 – Migration, Residential Mobility, and Housing Policy. In: **ENHR International Conference “Housing in an expanding Europe: theory, policy, participation and implementation”**, 2006, Ljubljana, Slovenia, 2-5 July 2006. p. 01-21. Disponível em: <http://enhr2006-ljubljana.uirs.si/publish/W03_Goetgeluk.pdf>. Acesso em: 10 jan. 2008.

GOETGELUK, Roland; JONG, Tom de. Migration analysis as a political instrument: the case of the Leiden and Bulb Regions in the Randstad Holland. **Open house International**, Netherlands v.30,n.3,p.75-82,2005.

GOETGELUK, Roland; JONG, Tom de. What about the spatial dimensiono f subsidiarity in housing policy? In: INTERNATIONAL CONFERENCE ON SUSTAINABLE URBAN AREAS, Rotterdam, 2007, Delft University of Technology, Rotterdam, 2007. Disponível em: <http://www.enhr2007rotterdam.nl/documents/W02_paper_Goetgeluk_DeJong.pdf> Acesso em Out.2007.

GOETGELUK, Roland. **Intramax.** [mensagens pessoais] . Mensagens recebidas por <professor@paulofernando.mat.br> em Jul. 2008.

GOLGHER, André Braz. **Fundamentos da Migração.** Belo Horizonte: UFMG/Cedeplar, 2004. 49 p. Disponível em: <<http://www.cedeplar.ufmg.br/pesquisas/td/TD%20231.pdf>>. Acesso em: 10 jan. 2008.

GORSUCH, L. Richard. **Factor Analysis**. New Jersey: Lawrence Erlbaum Associate, 1983.

GORSUCH, Richard L. **A comparison of BIQUARTIMIN, MAXPLANE, PROMAX, AND VARIMAX**. *Educational and Psychological Measurement*. p. 861-872, 1970.

GRAY, B. Thomas. **Higher-Order Factor Analysis**. In: Annual meeting of the Southwest Educational Research Association, Austin, TX, January 24, 1997.

GULDMANN, Jean-Michael. **Spatial analysis of telecommunication flows**.

Disponível em: <http://facweb.knowlton.ohio-state.edu/jguldmann/crpinfo/research/CHAP_JMG2.pdf>. Acesso em: 10 jan. 2008.

HAIR, Joseph F. et. al. **Análise Multivariada de dados**. 5 ed. Porto Alegre: Bookman, 2005a.

HAIR, Joseph F. et. al. **Métodos de pesquisa em administração**. Porto Alegre: Bookman, 2005b.

HAKKERT, Ralph. **Fontes de dados demográficos**. Belo Horizonte: ABEP, 1996. 72 p. Disponível em: <<http://www.abep.nepo.unicamp.br/docs/outraspub/textosdidaticos/tdv03.pdf>>. Acesso em: 10 jan. 2008.

HAYNES, Kingsley E.; FOTHERINGHAM, A. Stewart. **Gravity and spatial interaction models**. Beverly Hills: Sage, 1984. p. 5-13.

HOLLOWAY, Steven R.; PANDIT, Kavita. The disparity between the level of economic development and human welfare. **Professional Geographer**. v. 44, n. 1, p. 57-71, February, 1992.

HOLSMAN, Andy. Higher-order factor analysis and its application to transport networks. **Professional Geographer**. v. 32, n. 2, p. 192-198, May, 1980.

IBGE. Censo Demográfico 2000. Microdados dos Resultados da amostra. Rio de Janeiro: IBGE. 2003.

IBGE. Censo Demográfico de 1991: Microdados dos Resultados da amostra, Rio de Janeiro: IBGE. 1996.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Malha Municipal digital do Brasil de 2001. Rio de Janeiro: IBGE, 2001.

IBGE. Tábuas completas de mortalidade. Disponível em: http://www.ibge.gov.br/home/presidencia/noticias/noticia_impresao.php?id_noticia=266. Acesso em 20 jun. 2009a.

IBGE. Taxa de fecundidade total, por situação do domicílio, segundo as Grandes Regiões e Unidades da Federação 1970/2000. Disponível em: http://www.ibge.gov.br/home/estatistica/populacao/censo2000/tendencias_demograficas/tabela07.pdf. Acesso em 20

jun. 2009b.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Caracterização e tendências da rede urbana no Brasil: Configuração atual e tendências da rede urbana**. Brasília: IPEA, IBGE, UNICAMP, 2001.

JOHNSON, Richard A., WICHERN, Dean W. **Applied Multivariate Statistical Analysis**. 5 ed. New Jersey: Prentice Hall, 2002.

JONG, Tom de. **Intramax**. [mensagens pessoais] . Mensagens recebidas por <professor@paulofernando.mat.br> em Jul. 2008.

LEMONS, Mauro Borges; DINIZ, Clelio Campolina; GUERRA, Leonardo Pontes. **A nova configuração regional brasileira e sua Geografia Econômica**. Disponível em: <http://www.econ.fea.usp.br/novo_site/publicacoes/estudos_economicos/33_4/borges-dinizmoro.pdf>. Acesso em: 10 jan. 2008.

LIMA AMARAL, Ernesto Friedrich de; FÍGOLI, Moema Gonçalves Bueno. Métodos e técnicas de mensuração e caracterização de movimentos migratórios. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, XIV, 2004, Caxambu. **Anais...** Caxambu: Associação Brasileira de Estudos Populacionais, 2004, p. 1-25. Disponível em: <http://www.abep.nepo.unicamp.br/site_eventos_abep/PDF/ABEP2004_583.pdf>. Acesso em: 10 jan. 2008.

MAGNOLI, Demétrio; ARAÚJO, Regina. **A nova geografia: estudos de geografia do Brasil**. 2 ed. São Paulo: Moderna, 1996. 378p.

MARTINE, George. Migrações internas e alternativas de fixação produtiva: experiências recentes de colonização no Brasil. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, I, 1978, Campos do Jordão. **Anais...** Campos do Jordão: Associação Brasileira de Estudos Populacionais, 1978, p. 51-85. Disponível em: <<http://www.abep.nepo.unicamp.br/docs/anais/pdf/1978/T78V01A02.pdf>>. Acesso em: 10 jan. 2008.

MARTINE, George; CAMARGO, Lísio. Crescimento e distribuição da população brasileira: tendências recentes. **Revista Brasileira de Estudos de População**, v.1, n.1/2, p.99-143, jan/dez/1984.

MASSER, I.; SCHEURWATER, J. Functional regionalization of spatial interaction data: an evaluation of some suggested strategies. **Environment and Planning A**, v.12, p.1357-382, 1980.

MASSER, I; BROWN, P.J.B. Hierarchical aggregation procedures for interaction data. **Environment and Planning A**, v.7, p.509-23, 1975.

MATHER, P.; OPENSHAW S. Multivariate methods and geographical data. **The Statistician**. v.23, n.3/4, p.283-308, set-dez. 1974.

MATOS, Ralfo. Crescimento populacional e migração interna nos principais municípios de Minas Gerais. In: ENCONTRO NACIONAL SOBRE MIGRAÇÃO, I, 1997, Curitiba. **Anais...** Curitiba: Associação Brasileira de Estudos Populacionais, 1997, v. 1, p. 329-347. Disponível em: <<http://www.abep.nepo.unicamp.br/docs/>

[anais/outros/1EncNacSobreMigracao/AnaisENSMigracaoCuritiba1997p329a347.pdf](http://www.abep.nepo.unicamp.br/docs/anais/outros/1EncNacSobreMigracao/AnaisENSMigracaoCuritiba1997p329a347.pdf)>. Acesso em: 10 jan. 2008.

MATOS, Ralfo. Migração e Desconcentração Demográfica nas Principais áreas de Atração Populacional de Minas Gerais. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, XI, 1998, Caxambu. **Anais...** Caxambu: Associação Brasileira de Estudos Populacionais, 1998, v. 1, p. 713-728. Disponível em: <<http://www.abep.nepo.unicamp.br/docs/anais/PDF/1998/a149.pdf>>. Acesso em: 10 jan. 2008.

MATOS, Ralfo; BRAGA, Fernando Gomes. Migração e Rede Urbana: procedências e inserção ocupacional. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, XIII, 2002, Ouro Preto. **Anais...** Ouro Preto: Associação Brasileira de Estudos Populacionais, 2002, v. 1, p. 1-35. Disponível em: <http://www.abep.nepo.unicamp.br/docs/anais/pdf/2002/GT_MIG_ST1_Matos_texto.pdf>. Acesso em: 10 jan. 2008.

MATOS, Ralfo; STEFANI, João; LOBO, Carlos Fernando Ferreira. Inversão no balanço migratório mineiro: as trocas populacionais entre Minas e São Paulo. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, XIV, 2004, Caxambu. **Anais...** Caxambu: Associação Brasileira de Estudos Populacionais, 2004, p. 1-17. Disponível em: <http://www.abep.nepo.unicamp.br/site_eventos_abep/PDF/ABEP2004_337.pdf>. Acesso em: 10 jan. 2008.

MINGOTTI, Sueli Aparecida. **Análise de dados através de métodos de estatística multivariada**. Belo Horizonte: Editora UFMG, 2005.

MOITA NETO, José Machado; MOITA, Graziella Ciaramella. Uma introdução à análise exploratória de dados multivariados. **Química Nova**. Teresina, v. 21, n. 4, p. 467-469, 1998. Disponível em: <<http://www.scielo.br/pdf/qn/v21n4/3193.pdf>>. Acesso em: 10 jan. 2008.

NIELSEN, Thomas Alexander Sick; HOVGESSEN, Henrik Harder. Exploratory mapping of commuter flows in England and Wales. **Journal of transport geography**, v.16, p.90-99, 2008.

NORMAN, Paul. **Putting iterative proportional fitting on the researcher's desk**. Disponível em: <<http://citeseer.ist.psu.edu/cache/papers/cs/13985/http:zSzzSzwww.geog.leeds.ac.ukzSzwpaperszSz99-3.pdf/putting-iterative-proportional-fitting.pdf>>. Acesso em: 10 jan. 2008.

NORONHA, Valerian T.; GOODCHILD, Michael F. Modeling interregional interaction: implications for defining functional regions. **Annals of the Association of American**

Geographers. v. 82, n. 1, p. 86-102, 1992.

O'BRIEN, Larry. **Introducing Quantitative Geography; Measurement, methods and generalised linear models.** London: Routledge, 1992. 356p.

PACHECO, Carlos Américo. **Fragmentação da nação.** Campinas: UNICAMP. IE, 1998.

PANDIT, Kavita. Differentiating between subsystems and typologies in the analysis of migration regions: A U.S. example. **Professional Geographer**, n.46, p. 331-345, 1994.

PAPPS, Kerry L.; NEWELL, James O. **Identifying functional labour market areas in New Zealand: a reconnaissance study using travel-to-work data.** Disponível em: <<http://opus.zbw-kiel.de/volltexte/2003/976/pdf/dp443.pdf>>. Acesso em: 10 jan. 2008.

PLANE, David A.; ISSERMAN, Andrew M. **U.S. interstate labor force migration: An analysis of trends, net exchanges, and migration subsystems.** Socio-Economic Planning Sciences, Elsevier, vol. 17(5-6), 1983, p. 251-266.

PLANE, David A.; ROGERSON, Peter A. Dynamic flow modeling with interregional dependency effects: an application to structural change in the U.S. migration system. **Demography.** v. 23, n. 1, p. 91-104, Feb. 1986. Disponível em: <<http://www.jstor.org/stable/view/2061411>>. Acesso em: 10 jan. 2008.

POOLER, James. Structural spatial interaction. **Professional Geographer.** v. 45, n. 3, p. 297-305, August, 1993.

RIGOTTI, José Irineu Rangel. A (re)distribuição espacial da população brasileira e possíveis impactos sobre a metropolização. In: **Encontro Anual da Associação Nacional de Pós-Graduação em Ciências Sociais**, 32, Caxambu, 2008.

RIGOTTI, José Irineu Rangel. **A distribuição espacial da população brasileira e das atividades econômicas nas décadas de 80 e 90.** Relatório de pesquisa. 2004.

RIGOTTI, José Irineu Rangel. Estimativas de saldos e fluxos migratórios a partir do Censo Demográfico de 1991: uma aplicação para as mesorregiões de Minas Gerais. **Revista Brasileira de Estudos da População**, São Paulo, v.17, n.1/2, jan./dez. 2000.

RIGOTTI, José Irineu Rangel. Geografia dos fluxos populacionais segundo níveis de escolaridade dos migrantes. **Estudos Avançados**, São Paulo, n.20, p.237-254, 2006.

RIGOTTI, José Irineu Rangel. O uso dos quesitos censitários para o estudo das migrações. In: **XII Encontro de Estudos Populacionais da ABEP**, Caxambu, 2000. v.1. Anais... Disponível em http://www.abep.nepo.unicamp.br/docs/anais/pdf/2000/Todos/migt4_1.pdf. Acesso em 25 dez. 2006.

RIGOTTI, José Irineu Rangel. **Projeção da população por sexo, idade e situação do domicílio – Brasil – 2000/2050.** Rio de Janeiro. 2001. 206p.

RIGOTTI, José Irineu Rangel; VASCONCELLOS, Idamila R.P. Uma análise espacial

exploratória dos fluxos populacionais brasileiros nos períodos 1986-1991 e 1995-2000. In: **Encontro Nacional sobre Migração**, 4, Rio de Janeiro, p.1-20,2005.

ROGERSON, Peter A. New directions in the modelling of interregional migration. **Economic Geography**. v. 60, n. 2, p. 111-121, Apr. 1984.

ROGERSON, Peter. **Statistical methods for geography**. London: Sage, 2001. 236 p.

RUMMEL, R. J. **Applied factor analysis**. USA: Northwestern University, 1970.

SCHINDLER, Eric. **Iterative proportional fitting in the Census 2000 dress rehearsal**. Disponível em: < http://www.amstat.org/Sections/Srms/Proceedings/papers/1999_077.pdf >. Acesso em: 10 jan. 2008.

SCREMIN, Marcos Antônio Antonello. **Método para a seleção do número de componentes principais com base na lógica difusa**. 2003. 124 f. Tese (Doutorado) – Universidade Federal de Santa Catarina, Programa de Pós-graduação em Engenharia de Produção, Florianópolis. Disponível em: <<http://teses.eps.ufsc.br/defesa/pdf/3826.pdf>>. Acesso em: 10 jan. 2008.

SIMONSEN, Mário Henrique. Prefácio. In: BUESCU, Mircea. **Métodos quantitativos em História**. Rio de Janeiro: Livros Técnicos e Científicos, 127p.1983.

SLATER, Paul B. A hierarchical regionalization of state economic areas based upon migration flows. **Public data use**. v. 4, n. 6, p. 32-56, November 1976.

SLOCUM, t.. **The use of quantitative methods in major geographical journals, 1956-1986**. *Professional Geographer* 42:84-94. 1990

SOARES, Beatriz. R., BESSA, Kelly. C. F. de O. Uberlândia e sua região: caracterizando uma cidade média. In: **IV Congresso de Ciências Humanas, Letras e Artes**, Viçosa, 1999, Cd rom, Videolar, 1999.

VALENTIN, Jean Louis. **Ecologia numérica**. Rio de Janeiro: Interciência, 2000.

VASCONCELLOS, Idamila R.P.; RIGOTTI, José Irineu Rangel. **Migrações entre os municípios brasileiros, a partir das informações dos censos demográficos de 1991 e 2000**. Belo Horizonte, 2003. Disponível em <<http://www.abep.nepo.unicamp.br/docs/anais/outros/4EncNacSobreMigracao/SCII-6.pdf>>. Acesso em 15 jan. 2007.

WISEMAN, Robert. Interpretation of factored flow data: a comparison of two traffic systems. **The Professional Geographer**. v. 28, n. 2, p. 172-177, may 1976.

ZWAN, J. van der *et al.* **Flowmap 7.2: manual**. Netherlands: Faculty of Geographical Sciences, 2005. Disponível em: <<http://flowmap.geo.uu.nl>>. Acesso em 10 jul. 2007.

APÊNDICE A - FUNDAMENTOS MATEMÁTICOS E COMPUTACIONAIS DO MÉTODO INTRAMAX

1 Introdução

Com o intuito de embasar matematicamente a proposta metodológica de uso do Método Intramax no estudo da migração são apresentados os conceitos iniciais, a fundamentação matemática intrínseca, um *script* desenvolvido no ambiente MATLAB, e o software Flowmap. Opta-se por numerar as seções deste Apêndice de modo independente do corpo da tese.

Considere:

- T_{ij} a interação existente entre a unidade de origem i e a unidade de destino j
- E_{ij} valor esperado para a interação entre a unidade de origem i e a unidade de destino j
- N o total de interações registrado na matriz de fluxos
- $O_i = \sum_j T_{ij}$ o total marginal da linha i na matriz de fluxos
- $D_j = \sum_i T_{ij}$ o total marginal da coluna j na matriz de fluxos
- I_{ij} a função objetiva do método Intramax

Assim, o valor esperado para a intensidade de interação entre a unidade espacial de origem i e a unidade de destino j é dado por $E_{ij} = \frac{O_i \cdot D_j}{N}$. (1)

Calcula-se a função auxiliar S_{ij} , para $i \neq j$, como a soma das proporções dos fluxos de i para j e de j para i pelos respectivos valores esperados:

$$S_{ij} = \frac{T_{ij}}{\frac{O_i \cdot D_j}{N}} + \frac{T_{ji}}{\frac{O_j \cdot D_i}{N}} \quad (2)$$

Logo, $S_{ij} = \frac{N \cdot T_{ij}}{O_i \cdot D_j} + \frac{N \cdot T_{ji}}{O_j \cdot D_i}$ e pode-se colocar o fator comum, N, em evidência,

$$\text{obtendo-se } S_{ij} = N \cdot \left(\frac{T_{ij}}{O_i \cdot D_j} + \frac{T_{ji}}{O_j \cdot D_i} \right). \quad (3)$$

Busca-se o valor máximo de S_{ij} , ou seja, identificar quais unidades espaciais apresentam maior soma das proporções entre fluxo observado e esperado, mas, como N é uma constante (valor total dos fluxos registrados na matriz de interação) e, portanto, estará presente para todos os valores obtidos para a função auxiliar S_{ij} , divide-se S_{ij} por N, chegando à função objetiva do método Intramax:

$$\text{Max}_{i \neq j} I_{ij} = \frac{T_{ij}}{O_i \cdot D_j} + \frac{T_{ji}}{O_j \cdot D_i} \quad (4)$$

Segundo Feldman et al. (2005), nos estágios iniciais da aplicação do método Intramax as menores unidades (em termos dos totais de entradas e saídas) com relativamente fortes ligações serão unidas (alto T_{ij} comparado com pequenos valores de O_i e D_j). Nos estágios intermediários, pequenas áreas são unidas com grandes áreas (alto T_{ij} comparado com pequeno O_i e grande D_j ou grande O_i e pequeno D_j) e nos estágios finais grandes unidades são unidas com outras grandes unidades (alto T_{ij} comparado com altos valores para O_i e D_j).

Uma consequência da formulação, como proposto para a função objetiva do método Intramax, é a da formação de agrupamentos de unidades espaciais que apresentam mais interações entre si do que com outras unidades espaciais.

A representação gráfica mais adequada para apresentar os resultados obtidos pelo método Intramax é o diagrama de similaridade denominado dendrograma.

2 Script do Intramax a ser rodado no software MATLAB

O *script* visto adiante representa a rotina de passos que devem ser seguidos para execução do método Intramax no software MATLAB. Para tanto, o usuário deve fornecer a matriz de fluxos origem-destino adicionada por uma coluna, que deve ser a primeira, composta pelos rótulos numéricos de cada unidade espacial, conforme modelo:

Rótulo 1	$F_{1,1}$	$F_{1,2}$	$F_{1,3}$
Rótulo 2	$F_{2,1}$	$F_{2,2}$	$F_{2,3}$
Rótulo 3	$F_{3,1}$	$F_{3,2}$	$F_{3,3}$

$F_{i,j}$ indica o fluxo da unidade espacial i para a unidade espacial j .

O resultado fica armazenado na variável “grupos” e é exportado para um arquivo de mesmo nome no atual diretório de trabalho. Esse arquivo pode ser visualizado em uma planilha eletrônica ou no próprio software MATLAB.

```

% Script para aplicação do método iterativo Intramax
% Desenvolvido por Paulo Fernando Braga Carvalho
loops=length(MATRIZ)-1;
grupos=[];
for k=1:loops-2
    p=length(MATRIZ)-1;
    I=[];
    for i=1:p
        for j=(i+1):p
            INTAUX=MATRIZ(i,j+1)/( sum(MATRIZ(i,2:p+1))*sum(MATRIZ(:,j+1))
)+MATRIZ(j,i+1)/(sum(MATRIZ(j,2:p+1))*sum(MATRIZ(:,i+1)));
            I=[I; i j+1 MATRIZ(i,1) MATRIZ(j,1) INTAUX];
        end
    end
    I=sortrows(I,5); % ordena linhas pelo resultado da
função objetiva
    tamanho=size(I); % fornecerá a posição do maior
resultado
    posicao_maior1=I(tamanho(1),1); % indica a linha do primeiro elemento
%selecionado
    posicao_maior2=I(tamanho(1),2)-1; % indica a linha do segundo elemento
%selecionado
    rotulo_maior1=I(tamanho(1),3); % rótulo do primeiro elemento
selecionado
    rotulo_maior2=I(tamanho(1),4); % rótulo do segundo elemento
selecionado

    %Montagem da matriz resultado com agrupamentos
    %Na matriz grupos, a terceira coluna indica o nome do novo grupo
    grupos=[grupos; rotulo_maior1 rotulo_maior2 rotulo_maior1];
    [linhas colunas]=size(MATRIZ);

    MATRIZ(posicao_maior1,2:colunas)=MATRIZ(posicao_maior1,2:colunas)+MATRIZ(po
sicao_maior2,2:colunas);
    MATRIZ(posicao_maior2,:)=[]; %excluindo linha

    MATRIZ(:,posicao_maior1+1)=MATRIZ(:,posicao_maior1+1)+MATRIZ(:,posicao_maio
r2+1);
    MATRIZ(:,posicao_maior2+1)=[]; %excluindo coluna
end
grupos
dlmwrite('grupos',grupos,',' );

```


3 O software Flowmap e o método Intramax

O Flowmap é um software dedicado a análises geográficas, com sua primeira versão desenvolvida em 1990 por Tom de Jong⁴ da *Faculty of Geographical Sciences* pertencente à *Utrecht University*, Holanda, em parceria com *Gadja Mada University* em Yogyakarta, Indonésia. Seu desenvolvimento começou com a elaboração de um simples programa para apresentação de fluxos de mercadorias e pessoas em um mapa.

É possível baixar a versão 7.3, a mais atual, seguindo o link <http://flowmap.geog.uu.nl> e, conforme informações obtidas neste site, os equipamentos requeridos para sua instalação e operação são modestos e encontrados em qualquer computador padrão com sistema operacional Windows. Neste mesmo link, pode-se obter manual, na forma de tutorial, em inglês, com todos os passos a serem seguidos para sua utilização bem como a base de dados utilizada nos exemplos.

Existem duas versões do software Flowmap: Edição Educacional e Edição Profissional. A versão profissional, naturalmente, apresenta mais recursos que a educacional, entretanto, essa última é gratuita e possui a maior parte das funções ativas, podendo ser copiada e distribuída livremente, não sendo permitida a sua comercialização. O autor pede, apenas, que sejam dados os créditos devidos ao software e à *Faculty of Geographical Sciences* nos trabalhos.

A opção pelo uso do Flowmap neste trabalho se justifica pelo fato de este ser um software especializado na manipulação e apresentação de dados de interação como trocas e fluxos migratórios, tendo implementado o método Intramax. Mas, além dessa função, é possível realizar outras análises de interação como análises de acessibilidade, de redes e de modelos de interação como o potencial.

⁴ Dr. Tom de Jong é pesquisador sênior da Faculty of Geosciences da Utrecht University of Technology, Holanda, tdj@geog.uu.nl

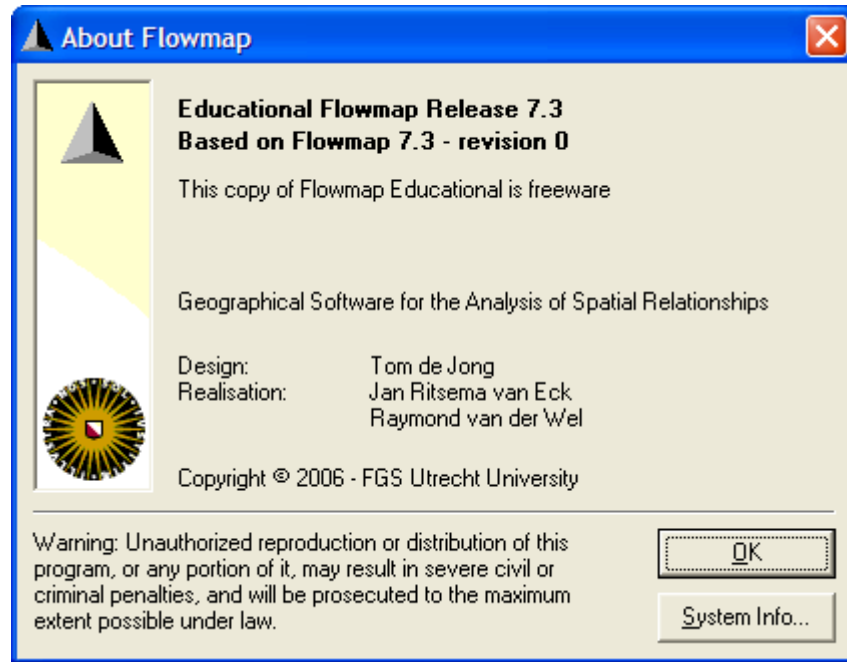


Figura 26: Tela com informações sobre o software Flowmap 7.3

No menu Analysis e submenu Spatial Aggregation, veja Figura 27, encontra-se a opção para uso do método Intramax. Ao aplicá-lo, o software gera dois arquivos do tipo .txt: um relatório indicando o par de unidades espaciais agrupadas em cada passo iterativo inerente ao método e o dendrograma, diagrama de similaridades adequado para visualização dos resultados.

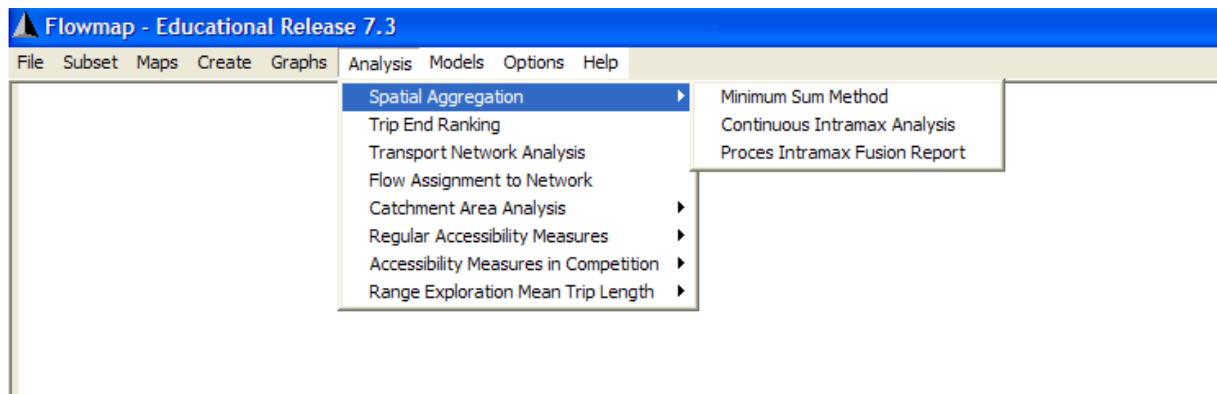


Figura 27: Caminho de acesso ao método Intramax no software Flowmap 7.3

Além disso, os resultados podem ser representados cartograficamente, sendo possível visualizar as unidades espaciais que são agrupadas em cada passo iterativo na área de mapas do software ou imprimir o resultado final com os agrupamentos obtidos. Infelizmente, não é possível realizar grandes alterações

estéticas ou complementares nos mapas, diretamente no Flowmap, sendo necessário tratá-los em outros ambientes.

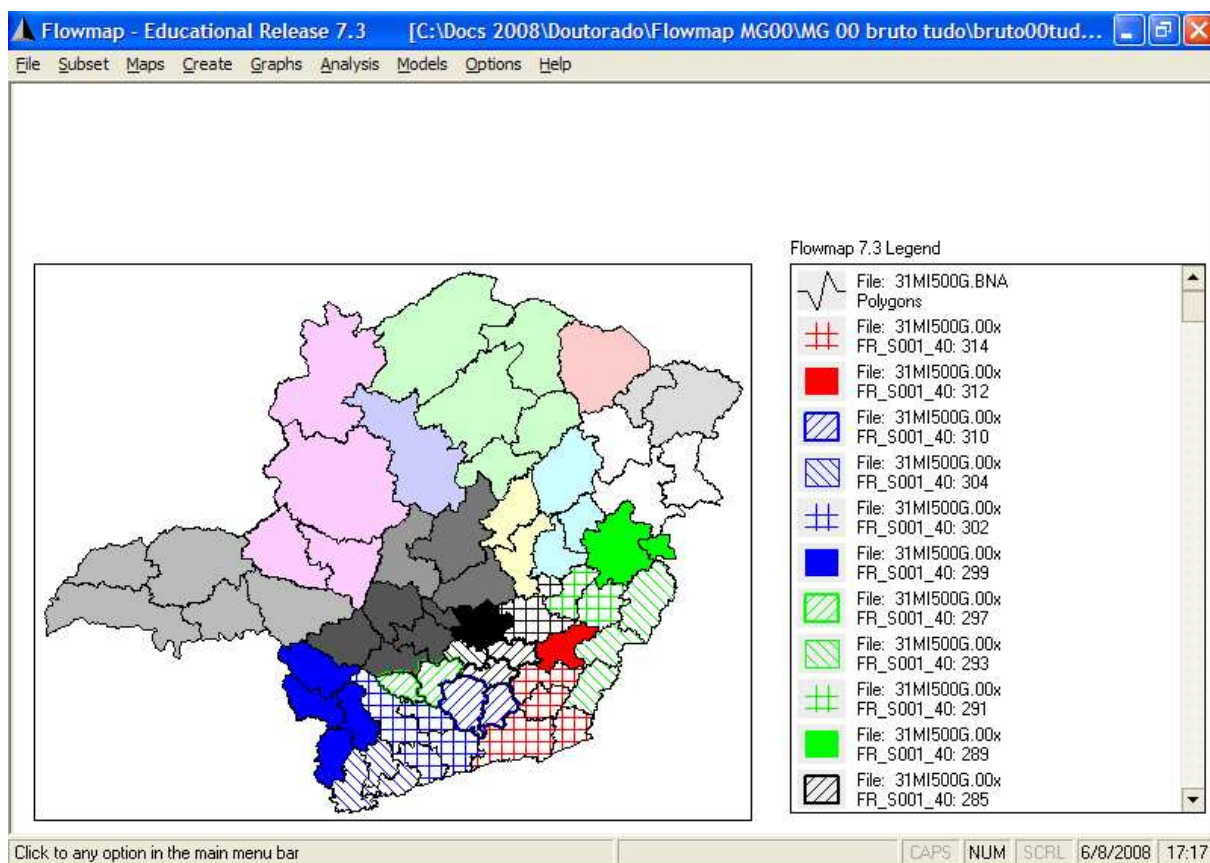


Figura 28: Exemplo de mapa produzido com aplicação do método Intramax no software Flowmap

Na legenda da Figura 28, lê-se o nome do arquivo base usado na elaboração da cartografia, o nome da coluna criada nesse arquivo pela rotina do Intramax, indicando que o estado atual é obtido após a aplicação de 40 passos dos 65 possíveis, pois em Minas Gerais existem 66 microrregiões. Cada simbologia indica um agrupamento de microrregiões com fortes interações.

APÊNDICE B – FUNDAMENTOS MATEMÁTICOS DA ANÁLISE FATORIAL

1 Introdução

Com o intuito de embasar matematicamente a proposta metodológica do uso de Componentes Principais no estudo da migração, faz-se necessário apresentar alguns conceitos e propriedades fundamentais da Álgebra Linear e da Estatística bem como a definição de Componente Principal e a demonstração do processo de obtenção de tal Componente. Opta-se por numerar as seções deste Apêndice de modo independente do corpo da tese.

Inicia-se, portanto, com definição de Análise Fatorial apresentada por Hair et al.:

Análise Fatorial é um nome genérico dado a uma classe de métodos estatísticos multivariados cujo propósito principal é definir a estrutura subjacente em uma matriz de dados. Em termos gerais a Análise Fatorial aborda o problema de analisar a estrutura das inter-relações (correlações) entre um grande número de variáveis, definindo um conjunto de dimensões latentes comuns chamadas de fatores. Com a análise fatorial, o pesquisador pode primeiro identificar as dimensões separadas da estrutura e então determinar o grau em que cada variável é explicada por cada dimensão.
(Hair et al., 2005a,p.91)

Para uso da Análise Fatorial o pesquisador deve ter consciência de como as variáveis estão inter-relacionadas. Para isso, lança mão da matriz de correlação.

Existem basicamente duas técnicas de Análise Fatorial: Análise de Fatores Comuns e Análise de Componentes Principais. As diferenças entre estas técnicas estão relacionadas à variância.

A variância é dividida em 3 tipos: comum, específica e do erro. A variância comum é aquela variância compartilhada com todas outras variáveis da análise. A variância específica (ou única) é aquela relativa a uma única variável, enquanto que a variância do erro é a variância provocada pela coleta dos dados, cálculos dentro do processo, etc. A Análise de Componentes Principais-ACP usa todos os três tipos de variância para extrair soluções fatoriais enquanto a Análise Fatorial Comum, utiliza somente a variância comum para determinação dos fatores. Veja Figura 29.

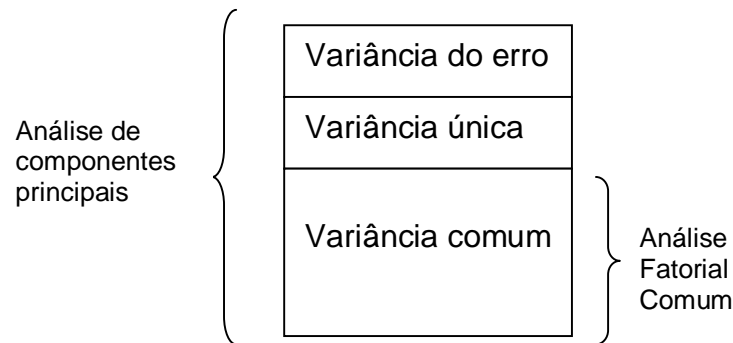


Figura 29: Tipos de variância na Análise Fatorial
 Fonte: Hair et al.,2005b,p.391

Ao considerar a variância total, a Análise de Componentes Principais, apresenta em sua solução fatores que trazem pequenas proporções de variância única e de variância do erro. Mas, apesar disto, os primeiros fatores, os principais, não contêm variância de erro ou única suficiente para distorcer a estrutura fatorial geral.

Matematicamente falando, ao se trabalhar com Análise de Componentes Principais, a diagonal principal da matriz de correlação é preenchida com valores unitários, implicando no uso da variância total para obtenção da matriz fatorial. Já na Análise Fatorial Comum, a diagonal principal é composta pelas comunalidades, ou seja, estruturas da variância comum.

A escolha entre uma técnica e outra deve se basear nos objetivos da Análise Fatorial e do conhecimento prévio sobre as variâncias das variáveis.

Pode-se dizer que, o objetivo geral das técnicas de Análise Fatorial é concentrar a grande quantidade de informações fornecidas por um grande conjunto de variáveis, em outro conjunto menor, com o mínimo de perda de informação. Com isso, o pesquisador pode optar pela Análise Fatorial quando precisa identificar a estrutura do conjunto de variáveis ou reduzir o número de variáveis a analisar.

A Análise Fatorial pode ser usada como um método de seleção de variáveis. Por exemplo, se um pesquisador tem acesso a dados relativos a 30 variáveis, mas deseja selecionar as 12 mais importantes, poderá realizar tal tarefa analisando os maiores coeficientes obtidos em cada componente, em valor absoluto.

2 Conceitos e propriedades fundamentais

Para melhor fluência na leitura, alguns conceitos da Álgebra Linear são recuperados.

2.1 Notação fundamental

X_j : uma variável, que é um vetor de dados para um número de casos.

x_{ij} : o valor da variável para o caso i na variável X_j .

$\sum_{i=1}^n x_{ij}$: a soma dos valores para n casos na variável X_j , onde,

$$\text{se } i=1, \sum_{i=1}^n x_{ij} = x_{1j} + x_{2j} + x_{3j} + \dots + x_{nj}.$$

2.2 Média

A média de uma variável é a soma de seus n elementos dividido por n .

$$\bar{X}_j: \text{ média da variável } j. \quad \bar{X}_j = \frac{\sum_{i=1}^n x_{ij}}{n} \quad (1)$$

$$\bar{X}: \text{ média da variável } X, \quad \bar{X} = \frac{\sum_{i=1}^n x_i}{n} \text{ quando estiver tratando de uma única variável.} \quad (2)$$

2.3 Variância

A Variância de uma variável é a média dos quadrados dos desvios de seus valores em relação à média. É uma medida estatística que descreve a variabilidade na distribuição dos dados.

$$s_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{X}_j)^2}{n} \text{ é a variância da variável } j \quad (3)$$

2.4 Desvio Padrão

É a raiz quadrada da variância.

$$s_j = \sqrt{s_j^2} \text{ é o desvio padrão da variável } j. \quad (4)$$

2.5 Variável padronizada

Quando se trabalha com várias variáveis, é possível que essas apresentem variabilidades muito diferentes, muitas vezes por terem grandezas ou unidades originais distintas. Nesses casos, uma opção é fazer a padronização da variável, trabalhando com sua posição dentro do conjunto, tomando como referência a distância, em número de desvios padrões, da média da variável.

Z_j é a variável j padronizada.

$$z_{ij} = \frac{x_{ij} - \bar{X}_j}{s_j} \text{ é o valor padronizado para o caso } i \text{ da variável } j. \quad (5)$$

A variável padronizada dessa maneira tem duas propriedades de interesse:

- i. Média igual a zero;
- ii. Desvio Padrão igual a 1 ($s_j = s_j^2 = 1$)

2.6 Covariância

A covariância é uma medida de relacionamento entre duas variáveis aleatórias, sendo dada pela razão do produto de seus desvios em relação à média pelo número de casos.

$$s_{ij} = \frac{\sum_{i=1}^n (x_{ij} - \bar{X}_j)(x_{ik} - \bar{X}_k)}{n} \quad (6)$$

A dificuldade no uso da covariância se encontra no fato de não existirem valores de referência máximo ou mínimo para comparação. Por isso, a medida de correlação se torna mais útil.

- i. Deve-se observar que, como consequência da definição, a covariância de uma variável com ela mesma é igual à variância total da variável, ou seja, $s_{ii} = s_i^2$.

2.7 Correlação linear

A correlação linear é, também, uma medida adequada para avaliar a força ou grau de relacionamento entre duas variáveis aleatórias. Quando se diz que duas variáveis estão fortemente correlacionadas, não se deve concluir que exista, necessariamente, dependência de uma em relação à outra.

Um dos coeficientes de correlação linear mais usados é o de Pearson:

$$r_{jk} = \frac{n \sum_{i=1}^n x_{ij} y_{ik} - \sum_{i=1}^n x_{ij} \cdot \sum_{i=1}^n x_{ik}}{\sqrt{\left[n \sum_{i=1}^n x_{ij}^2 - \left(\sum_{i=1}^n x_{ij} \right)^2 \right] \cdot \left[n \sum_{i=1}^n x_{ik}^2 - \left(\sum_{i=1}^n x_{ik} \right)^2 \right]}} \quad (-1 \leq r \leq 1) \quad (7)$$

r_{jk} é a correlação entre as variáveis j e k.

- i. $r_{jk} = -1$ indica correlação linear negativa perfeita, sugerindo que quanto maior o valor da variável j, menor será o da variável k. Em um gráfico de dispersão tem-se todos os pontos perfeitamente alinhados em uma reta decrescente.
- ii. $r_{jk} = 0$ indica correlação linear nula.
- iii. $r_{jk} = 1$ indica correlação linear positiva perfeita, sugerindo que quanto maior o valor da variável j, maior será o da variável k. Em um gráfico de dispersão tem-se todos os pontos perfeitamente alinhados em uma reta crescente.

Se dividimos a covariância entre as variáveis aleatórias X e Y pelo produto dos desvios padrões de X e Y, obtemos, então, o coeficiente de correlação linear de Pearson entre X e Y. Portanto,

$$r_{jk} = \frac{\text{cov}(X_j, X_k)}{s_j \cdot s_k} = \frac{s_{jk}}{s_j \cdot s_k} \quad (8)$$

2.7.1 Matriz de correlação

É a matriz quadrada com as correlações entre as variáveis.

$$r = \begin{pmatrix} r_{11} & r_{12} & \Lambda & r_{1p} \\ r_{21} & r_{22} & \Lambda & r_{2p} \\ \Lambda & \Lambda & \Lambda & \Lambda \\ r_{p1} & r_{p2} & \Lambda & r_{pp} \end{pmatrix} \text{ onde } r_{jk} \text{ é a correlação entre as variáveis } j \text{ e } k. \quad (9)$$

- i. Deve-se atentar para o fato de que uma matriz de correlação é simétrica;
- ii. $r_{ii} = 1$, para todo i (correlação de uma variável com ela mesma);
- iii. Como consequência da propriedade da covariância de uma variável com ela mesma tem-se que, $r_{ii}=1$ indica a variância total da variável padronizada Z_i ;
- iv. A matriz de correlação pode ser obtida pela operação matricial

$$r = \frac{Z^T \cdot Z}{n} \quad (10)$$

onde Z é a matriz com as variáveis padronizadas e n é o número de variáveis (colunas).

2.8 Traço de uma matriz

Se A é uma matriz quadrada de ordem n , então o traço de A é definido como a soma de todos os elementos da diagonal principal de A .

Notação: $\text{Tr}(A)$

$$\text{Sendo } A = \begin{pmatrix} \mathbf{a}_{11} & & & \mathbf{a}_{1p} \\ & \mathbf{M} & & \mathbf{a}_{2p} \\ & & \dots & \\ \mathbf{a}_{p1} & & & \mathbf{a}_{pn} \end{pmatrix}, \text{ o traço de } A \text{ é dado por } \text{Tr}(A) = \mathbf{a}_{11} + \mathbf{a}_{22} + \dots + \mathbf{a}_{pp}.$$

- i. Como consequência da definição de traço, o traço da matriz de correlação é igual ao número de variáveis, pois nessa matriz todos elementos da diagonal principal

são iguais a 1 (correlação de qualquer variável com ela mesma). Logo, para variáveis padronizadas, o número de variáveis é a variância total.

2.9 Autovalores e Autovetores

2.9.1 Seja A uma matriz $n \times n$. O número real λ é um **autovalor** de A se existe um vetor não-nulo \vec{v} em \mathbb{R}^n tal que

$$A \cdot \vec{v} = \lambda \cdot \vec{v} \quad \text{ou} \quad (A - \lambda \cdot I) \cdot \vec{v} = \vec{0}, \quad (11)$$

onde I é a matriz identidade.

Neste caso, \vec{v} é denominado **autovetor** de A associado ao autovalor λ .

Outras denominações: Autovalores são também chamados de valores próprios, valores característicos ou valores latentes e os autovetores de vetores próprios, vetores característicos ou vetores latentes.

Exemplo:

$$\text{Seja } A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

$$\text{i. } A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad \text{e}$$


$$\text{ii. } \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

Então, como o resultado obtido em i) é igual ao resultado obtido em ii), $\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ é um

autovetor de A associado ao autovalor $\lambda = \frac{1}{2}$.

Pode-se verificar que $\vec{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ é um autovetor de A associado ao autovalor $\lambda = -\frac{1}{2}$.

- i. Observe que \vec{x}_1 e $A \cdot \vec{x}_1$, são paralelos. O mesmo ocorrendo para \vec{x}_2 e $A \cdot \vec{x}_2$.
- ii. Um autovalor λ de A tem uma infinidade de autovetores diferentes associados. De fato, se \vec{v} é um autovetor de A associado ao autovalor λ (isto é, se $A \cdot \vec{v} = \lambda \cdot \vec{v}$) e se r é qualquer número real diferente de zero, então

$$A(r\vec{v}) = r(A\vec{v}) = r(\lambda\vec{v}) = \lambda(r\vec{v}).$$


Logo, $r\vec{v}$ também é um autovetor associado a λ .

2.9.2 O polinômio $P(\lambda) = \det(\lambda \cdot I_n - A)$, onde I_n é a matriz identidade de ordem n , denomina-se polinômio característico de A .

2.9.2.1 Teorema Os autovalores de A são as raízes do polinômio característico de A . Como conseqüência, para encontrar os autovalores de A , basta calcular as raízes de $\det(\lambda I_n - A) = 0$.

2.10 Combinação linear de vetores

2.10.1 Combinação linear de vetores

Y é combinação linear dos vetores X_1, X_2, \dots, X_p , se existem números reais $\alpha_1, \alpha_2, \dots, \alpha_p$ tais que $Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_p X_p$. (12)

Os números $\alpha_1, \alpha_2, \dots, \alpha_p$ são denominados coeficientes ou pesos da combinação linear ou, ainda, cargas fatoriais (*loadings*).

2.10.2 Média da combinação linear

A média da combinação linear é

$$\bar{Y} = \alpha_1 \bar{X}_1 + \alpha_2 \bar{X}_2 + \dots + \alpha_p \bar{X}_p \quad \text{ou} \quad \bar{Y} = \sum_{i=1}^p \alpha_i \bar{X}_i \quad (13)$$

2.10.3 Variância da combinação linear

A variância da combinação linear é

$$\text{Var}(Y) = \sum_{j=1}^p \alpha_j^2 \cdot s_j^2 + 2 \cdot \sum_{j=1}^{p-1} (\alpha_j \cdot s_j \cdot \alpha_k \cdot s_k \cdot r_{jk}) \quad (14)$$

Onde

α_j : coeficientes da combinação linear

s_j^2 : variância de X_j

s_j e s_k : desvios-padrão das variáveis X_j e X_k

r_{jk} : correlação entre as variáveis X_j e X_k .

De modo equivalente, pode-se obter a variância de Y via operações matriciais da seguinte maneira:

$$\text{Var}(Y) = \alpha^T \cdot S \cdot \alpha = (\alpha_1 \quad \alpha_2 \quad \Lambda \quad \alpha_p)_{1 \times p} \cdot \begin{pmatrix} s_1^2 & s_{12} & \Lambda & s_{1p} \\ s_{21} & s_2^2 & \Lambda & s_{2p} \\ M & M & M & M \\ s_{p1} & s_{p2} & \Lambda & s_p^2 \end{pmatrix}_{p \times p} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1} \quad (15)$$

Onde

α : é a matriz (vetor) dos coeficientes da combinação linear, ou seja, $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1}$

α^T : matriz (vetor) transposta de α .

S : matriz de covariâncias $\begin{pmatrix} s_1^2 & s_{12} & \Lambda & s_{1p} \\ s_{21} & s_2^2 & \Lambda & s_{2p} \\ M & M & M & M \\ s_{p1} & s_{p2} & \Lambda & s_p^2 \end{pmatrix}_{p \times p}$

2.11 Tamanho do vetor

O tamanho do vetor α , indicado por $\|\alpha\|$, é dado por $\|\alpha\| = \sqrt{\alpha_1^2 + \alpha_2^2 + K + \alpha_p^2}$. (16)

Diz-se que um vetor está normalizado quando seu tamanho é igual a 1

3 COMPONENTE PRINCIPAL

Neste momento, são apresentados o conceito de Componente Principal e suas principais propriedades.

3.1 Obtenção da Componente Principal

Matematicamente, o problema a que se propõe é o de determinar os coeficientes da combinação linear $Y = \alpha_1 X_1 + \alpha_2 X_2 + K + \alpha_p X_p$ sujeita a duas hipóteses:

- i) A variância de Y seja máxima;
- ii) a matriz-coluna (vetor) α , composta pelos coeficientes da combinação

linear, esteja normalizada, ou seja, $\|\alpha\| = \alpha^T \alpha = 1$, ou ainda que $\sum_{j=1}^p \alpha_j^2 = 1$;

Partindo de (15), tem-se que

$$\text{Var}(Y) = \alpha^T \cdot S \cdot \alpha = (\alpha_1 \quad \alpha_2 \quad \Lambda \quad \alpha_p)_{1 \times p} \cdot \begin{pmatrix} s_1^2 & s_{12} & \Lambda & s_{1p} \\ s_{21} & s_2^2 & \Lambda & s_{2p} \\ M & M & M & M \\ s_{p1} & s_{p2} & \Lambda & s_p^2 \end{pmatrix}_{p \times p} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1}.$$

Multiplicando os dois primeiros termos, obtém-se

$$\text{Var}(Y) = (\alpha_1 s_1^2 + \alpha_2 s_{12} + \alpha_3 s_{31} + K + \alpha_p s_{p1} \quad \alpha_1 s_{21} + \alpha_2 s_2^2 + \alpha_3 s_{23} + K + \alpha_p s_{2p} \\ \alpha_1 s_{31} + \alpha_2 s_{32} + \alpha_3 s_3^2 + K + \alpha_p s_{3p} \quad \Lambda \quad \alpha_1 s_{p1} + \alpha_2 s_{p2} + \alpha_3 s_{p3} + K + \alpha_p s_p^2)_{1 \times p} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1}.$$

Multiplicando as matrizes obtidas chega-se a:

$$\text{Var}(Y) = \alpha_1^2 s_1^2 + \alpha_1 \alpha_2 s_{12} + \alpha_1 \alpha_3 s_{13} + K \alpha_1 \alpha_p s_{1p} + \alpha_2 \alpha_1 s_{21} + \alpha_2^2 s_2^2 + \alpha_2 \alpha_3 s_{23} + K + \alpha_2 \alpha_p s_{2p}$$

$$\alpha_3 \alpha_1 s_{31} + \alpha_3 \alpha_2 s_{32} + \alpha_3^2 s_3^2 + K + \alpha_3 \alpha_p s_{3p} + K + \alpha_p \alpha_1 s_{p1} + \alpha_p \alpha_2 s_{p2} + \alpha_p \alpha_3 s_{p3} + K + \alpha_p^2 s_p^2.$$

Neste momento, faz-se o agrupamento dos termos equivalentes e a separação por grau do monômio:

$$\begin{aligned} \text{Var}(Y) = \alpha^T S \alpha = & \alpha_1^2 s_1^2 + \alpha_2^2 s_2^2 + \alpha_3^2 + K + \alpha_p^2 s_p^2 + 2\alpha_1 \alpha_2 s_{12} + 2\alpha_1 \alpha_3 s_{13} + K + 2\alpha_1 \alpha_p s_{1p} + \\ & + 2\alpha_2 \alpha_3 s_{23} + K + 2\alpha_2 \alpha_p s_{2p} + 2\alpha_3 \alpha_4 s_{34} + K + 2\alpha_3 \alpha_p s_{3p} + K \end{aligned}$$

Como algebrismo auxiliar, faz-se a subtração do termo $\lambda(\alpha^T \alpha - 1)$, onde $\alpha^T \alpha = 1$, pela segunda hipótese. Dessa forma o termo subtraído é igual a zero, não alterando o valor de $\text{Var}(Y)$.

$$\text{Var}(Y) = \alpha^T S \alpha - \lambda(\alpha^T \alpha - 1)$$

$$\text{Var}(Y) = \alpha^T S \alpha - \lambda(\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + K + \alpha_p^2 - 1)$$

Os pontos candidatos a maximizarem a expressão $\text{Var}(Y)$ são os pontos críticos, obtidos quando as derivadas parciais de $\text{Var}(Y)$ são nulas.

Logo, calculando e igualando as derivadas parciais de $\text{Var}(Y)$, cujas variáveis são $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_p$, a zero, obtém-se:

$$\frac{\partial \text{Var}(Y)}{\partial \alpha_1} = 2\alpha_1 s_1^2 + 2\alpha_2 s_{12} + 2\alpha_3 s_{13} + K + 2\alpha_p s_{1p} - 2\alpha_1 \lambda = 0$$

$$\frac{\partial \text{Var}(Y)}{\partial \alpha_2} = 2\alpha_2 s_2^2 + 2\alpha_1 s_{12} + 2\alpha_3 s_{23} + K + 2\alpha_p s_{2p} - 2\alpha_2 \lambda = 0$$

$$\frac{\partial \text{Var}(Y)}{\partial \alpha_3} = 2\alpha_3 s_3^2 + 2\alpha_1 s_{13} + 2\alpha_2 s_{23} + \alpha_4 s_{34} + K + 2\alpha_p s_{3p} - 2\alpha_3 \lambda = 0$$

N

N

N

$$\frac{\partial \text{Var}(Y)}{\partial \alpha_p} = 2\alpha_p s_p^2 + 2\alpha_1 s_{1p} + 2\alpha_2 s_{2p} + 2\alpha_3 s_{3p} + K + 2\alpha_{p-1} s_{p-1,p} - 2\alpha_p \lambda = 0$$

Essas equações podem ser divididas por 2, formando um sistema de equações lineares que pode ser escrito na forma matricial como se segue:

$$\begin{pmatrix} s_1^2 & s_{12} & \Lambda & s_{1p} \\ s_{21} & s_2^2 & \Lambda & s_{2p} \\ M & M & M & M \\ s_{p1} & s_{p2} & \Lambda & s_p^2 \end{pmatrix}_{p \times p} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1} - \lambda \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} 0 \\ 0 \\ M \\ 0 \end{pmatrix}$$

Colocando o termo comum em evidência,

$$\left(\begin{pmatrix} s_1^2 & s_{12} & \Lambda & s_{1p} \\ s_{21} & s_2^2 & \Lambda & s_{2p} \\ M & M & M & M \\ s_{p1} & s_{p2} & \Lambda & s_p^2 \end{pmatrix}_{p \times p} - \lambda \begin{pmatrix} 1 & 0 & \Lambda & 0 \\ 0 & 1 & \Lambda & 0 \\ M & M & M & M \\ 0 & 0 & \Lambda & 1 \end{pmatrix}_{p \times p} \right) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ M \\ \alpha_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} 0 \\ 0 \\ M \\ 0 \end{pmatrix}.$$

Logo, $(S - \lambda \cdot I) \cdot \alpha = \vec{0}$. Como, pela segunda hipótese $\|\alpha\| = 1$, tem-se que $\alpha \neq \vec{0}$, portanto, $(S - \lambda \cdot I) = 0$ e, ainda, resolvendo esse sistema, de acordo com a equação (11), obtém-se um autovetor associado à matriz de covariância S e ao autovalor λ . Ou ainda, para que a variância da Combinação Linear Y seja máxima, deve-se trabalhar com os coeficientes $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_p$ que são as componentes do autovetor, associado ao autovalor λ .

3.2 Autovalor e Variância

Mostra-se, a seguir, que $\text{Var}(Y) = \lambda$. Partindo da igualdade vista anteriormente $(S - \lambda I)\alpha = \vec{0}$, multiplicada membro a membro por α^T , chega-se a $\alpha^T(S - \lambda I)\alpha = \alpha^T\vec{0}$. Aplicando a propriedade distributiva tem-se $(\alpha^T S - \alpha^T \lambda \cdot I) \cdot \alpha = \vec{0}$. Aplicando a propriedade distributiva novamente, encontra-se $\alpha^T S \alpha - \alpha^T \lambda \cdot I \cdot \alpha = \vec{0}$. Mas, como visto em (15), $\alpha^T S \alpha = \text{Var}(Y)$, logo, $\text{Var}(Y) - \lambda \alpha^T \alpha = \vec{0}$, sendo, por hipótese, $\alpha^T \alpha = 1$. Portanto, $\text{Var}(Y) - \lambda \cdot 1 = 0$, ou seja, $\text{Var}(Y) = \lambda$.

Conclui-se, assim, que a variância da Combinação Linear Y é dada pelo autovalor associado ao autovetor determinado pelos coeficientes dessa combinação linear e à matriz de covariâncias.

3.3 Teorema da Decomposição Espectral

Outra interpretação matemática para Componente Principal, pode ser dada via Teorema da Decomposição Espectral, que é a escrita de uma matriz quadrada em fatores formados por seus autovetores e autovalores.

3.3.1 Teorema Todas as raízes do polinômio característico de uma matriz simétrica são reais.

3.3.2 Teorema Se A é uma matriz simétrica, então os autovetores associados a autovalores distintos são ortogonais. Como consequência, os autovetores associados à matriz simétrica A são independentes.

Esse teorema nos garantirá a ortogonalidade dos vetores iniciais encontrados na Análise de Componentes principais.

3.4 Matriz ortogonal

Uma matriz A invertível é dita ortogonal se sua transposta é igual a sua inversa, ou seja, se

$$P^T = P^{-1} \quad (17)$$

3.5 Matriz ortonormal

Uma matriz é dita ortonormal se, além de ser ortogonal, suas colunas formarem vetores de normas (tamanhos) iguais a 1.

3.5.1 Teorema Uma matriz A , $p \times p$, é ortogonal se, e somente se, suas colunas (e linhas) formam um conjunto ortonormal de vetores em \mathbb{R}^n

3.5.2 Teorema da Decomposição Espectral Se A é uma matriz simétrica $p \times p$, então existe uma matriz ortogonal P tal que $P^{-1} \cdot A \cdot P = D$, uma matriz diagonal. Os

autovalores de A são os elementos da diagonal de D e os autovetores de A compõem a matriz P.

Como consequência, multiplicando os dois membros da equação por P, obtém-se $P \cdot P^{-1} \cdot A \cdot P = P \cdot D$, mas $P \cdot P^{-1} = I$, onde I é a matriz identidade.

Logo, $A \cdot P = P \cdot D$. Multiplicando, pela direita, ambos os membros por P^{-1} , chega-se a $A \cdot P \cdot P^{-1} = P \cdot D \cdot P^{-1}$ e, portanto, como P é uma matriz ortogonal,

$$A = P \cdot D \cdot P^T \quad (18)$$

Ou seja, a matriz A pode ser decomposta em função de seus autovalores e autovetores.

3.5.2.1 Exemplo Considere a matriz simétrica $A = \begin{pmatrix} 1 & 0,5 & 0,8 \\ 0,5 & 1 & 0,2 \\ 0,8 & 0,2 & 1 \end{pmatrix}$.

Os autovalores e respectivos autovetores normalizados são

$$\lambda_1 = 0,1371 \text{ e } \vec{v}_1 = \begin{pmatrix} 0,7351 \\ -0,2832 \\ -0,6159 \end{pmatrix}, \lambda_2 = 0,8219 \text{ e } \vec{v}_2 = \begin{pmatrix} 0,0996 \\ -0,8536 \\ 0,5113 \end{pmatrix} \text{ e } \lambda_3 = 2,0410 \text{ e}$$

$$\vec{v}_3 = \begin{pmatrix} 0,6706 \\ 0,4372 \\ 0,5993 \end{pmatrix}. \text{ Logo, pode-se formar as matrizes } P = \begin{pmatrix} 0,7351 & 0,0996 & 0,6706 \\ -0,2832 & -0,8536 & 0,4372 \\ -0,6159 & 0,5113 & 0,5993 \end{pmatrix}$$

$$\text{e } D = \begin{pmatrix} 0,1371 & 0 & 0 \\ 0 & 0,8219 & 0 \\ 0 & 0 & 2,0410 \end{pmatrix}, \text{ tais que, a matriz A pode ser reescrita na forma}$$

fatorada $A = P \cdot D \cdot P^T$.

3.5.2.2 O autovalor, que pode ser obtido pela soma das cargas fatoriais (os coeficientes do fator) ao quadrado representa, assim, a quantia de variância explicada pelo respectivo fator.

3.6 Cargas fatoriais (*Loadings*)

Os valores α_{ij} são denominados cargas fatoriais ou *loadings* relativos à variável i no fator j . Como visto anteriormente, essas cargas fatoriais também podem ser vistas como coeficientes das combinações lineares obtidas com a Análise de Componentes Principais ou, ainda, como os elementos dos fatores (autovetores) obtidos com a Análise de Componentes Principais. A Carga fatorial Indica a correlação de cada variável com o respectivo fator.

Dispostos na matriz α , tem-se

$$\alpha = \begin{matrix} & \begin{matrix} F_1 & F_2 & \Lambda & F_m \\ \downarrow & \downarrow & \Lambda & \downarrow \end{matrix} & \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \Lambda & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \Lambda & \alpha_{2m} \\ M & M & \Lambda & M \\ \alpha_{p1} & \alpha_{p2} & \Lambda & \alpha_{pm} \end{pmatrix} & \begin{matrix} \leftarrow X_1 \\ \leftarrow X_2 \\ M \\ \leftarrow X_p \end{matrix} & \end{matrix} \quad (19)$$

Cada coluna da matriz α representa um fator (ou autovetor) que é uma nova variável (componente) obtida via Análise de Componentes Principais e cada linha representa uma variável original X_p .

O estudo de tais cargas fatoriais é um bom caminho para determinar qual a relevância de cada variável para cada um dos fatores. Por exemplo, α_{23} indica o peso da variável original X_2 no fator F_3 (ou na nova variável F_3). Segundo Hair et al. (2005, p.103), as cargas fatoriais indicam o grau de correspondência entre a variável e o fator, sendo que cargas maiores fazem a variável representativa no fator.

Essas cargas fatoriais serão usadas, mais adiante, como referência para agrupamento de microrregiões brasileiras de acordo com a matriz de fluxos populacionais, pois, quanto maior a carga fatorial de uma microrregião, maior a relação desta microrregião com o fator e, por conseqüência, com a unidade regional definida por esse fator.

3.6.1 Como consequência do que foi visto no item 1.2.3.2.5, e de acordo com Hair et al. (2005, p.91), as cargas fatoriais ao quadrado indicam qual percentual da variância em uma variável original é explicado pelo fator.

3.7 Número de fatores a extrair

Uma discussão importante na Análise Fatorial é a da determinação de fatores a serem extraídos. Uma analogia feita por Hair et al. (2005a, p.101) sobre o assunto, compara a escolha do número de fatores com a focalização utilizada em um microscópio onde, um ajuste muito alto ou muito baixo prejudicará a identificação da estrutura do conjunto, detalhando demais poucas características ou abrangendo grandes áreas e, com isso, perdendo a capacidade de enxergar características fundamentais do corpo em estudo.

Neste momento, faz-se a apresentação das três técnicas mais comuns para determinação do número de fatores a extrair em uma Análise Fatorial: Critério da raiz latente, Critério de porcentagem de variação e critério do teste *scree*.

3.7.1 Critério da raiz latente

Partindo da propriedade em que a proporção da variância total explicada pela j -ésima componente é igual a $\frac{\lambda_j}{p}$, onde p é o número de variáveis, é comum adotar o critério de aceitar um fator apenas se seu autovalor for maior que 1, o que sugere que este fator consegue captar a variância total de pelo menos uma variável.

3.7.2 Critério de porcentagem de variação

Com o objetivo de conseguir a explicação de uma determinada porcentagem de variação do conjunto de variáveis, selecionam-se fatores (ordenados com autovalores decrescentes) até que tal porcentagem de explicação seja alcançada.

Ou seja, até que $\left(\sum_{i=1}^j \frac{\lambda_i}{p} \right) \times 100$, alcance a significância desejada.

3.7.3 Critério do teste scree

O teste *scree* é feito com uso do gráfico *scree-plot*, Figura 30, que mostra os valores numéricos dos autovalores λ_i de acordo com a respectiva ordem i . A curva obtida é usada para determinar o número de fatores a selecionar. O ponto em que o gráfico começa a ficar na horizontal (e os autovalores ficarem próximos de zero) indica o número máximo de fatores a serem extraídos.

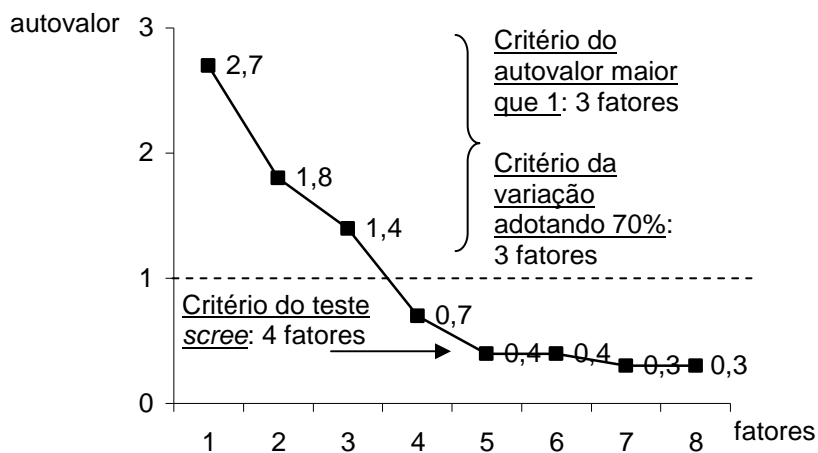


Figura 30: Exemplo de Scree plot para determinação de número de fatores

Observe que

- i) $2,7 + 1,8 + 1,4 + 0,7 + 0,4 + 0,4 + 0,3 + 0,3 = 8$;
- ii) $\frac{2,7}{8} + \frac{1,8}{8} + \frac{1,4}{8} = 0,74$, ou seja, 74%.

3.8 Escores (scores)

Os escores, também denominados escores fatoriais ou ponderações, são os resultados obtidos com a aplicação de cada caso (linha) em cada fator (coluna). Ou seja, substituem-se os valores relativos a um determinado caso nas combinações lineares sugeridas por cada um dos fatores. Para calcular os escores, em geral, usa-se a matriz de dados padronizados.

Por exemplo, considere a matriz Z de dados padronizados em que cada coluna representa uma variável e cada linha um caso.

$$Z = \begin{pmatrix} z_{11} & z_{12} & \Lambda & z_{1p} \\ z_{21} & z_{22} & \Lambda & z_{2p} \\ \Lambda & \Lambda & \Lambda & \Lambda \\ z_{p1} & z_{p2} & \Lambda & z_{pp} \end{pmatrix} \quad (20)$$

Considere, também, a matriz de cargas fatoriais, como visto em (19)

$$\alpha = \begin{matrix} & F_1 & F_2 & \Lambda & F_m \\ & \downarrow & \downarrow & \Lambda & \downarrow \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \Lambda & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \Lambda & \alpha_{2m} \\ M & M & \Lambda & M \\ \alpha_{p1} & \alpha_{p2} & \Lambda & \alpha_{pm} \end{pmatrix} & \leftarrow X_1 \\ & & & & \leftarrow X_2 \\ & & & & M \\ & & & & \leftarrow X_p \end{matrix}$$

O escore para o primeiro caso, no primeiro fator, será dado por

$$e_{11} = \alpha_{11}z_{11} + \alpha_{21}z_{12} + \dots + \alpha_{p1}z_{1p} \quad (21)$$

Segundo Hair (2005, p.112), conceitualmente, o escore fatorial representa o grau em que cada indivíduo tem escore elevado no grupo de itens que têm cargas elevadas em um fator. Assim, valores mais altos nas variáveis com cargas elevadas em um fator resultam em um escore fatorial superior.

Como sugerido por Andrade (1989, p.474) é importante ressaltar que os escores calculados para as componentes principais são sempre medidos em uma escala ordinal e, por isso, só podem indicar a posição relativa de cada caso. Um escore igual a 4 não indica que o caso em questão é duas vezes melhor (ou maior) que outro caso com escore igual a 2.

Pensando na matriz de fluxos populacionais, o escore será uma medida de acessibilidade de uma microrregião para o grupo regional definido pelas maiores cargas fatoriais do fator, portanto, auxiliará na identificação das principais origens e principais destinos dos agrupamentos sugeridos.

3.9 Comunalidade

Segundo Hair (2005, p.90) comunalidade é a quantidade total de variância que uma variável compartilha com todas as outras variáveis incluídas na análise. A Análise Fatorial Comum, também conhecida apenas como Análise Fatorial, se baseia apenas nessa variância comum, enquanto a Análise de Componentes Principais lida com a variância total. Para um grande número de variáveis ou se as comunalidades forem altas (maiores que 0,60) os resultados obtidos via Análise de Componentes Principais e Análise Fatorial Comum são muito parecidos.

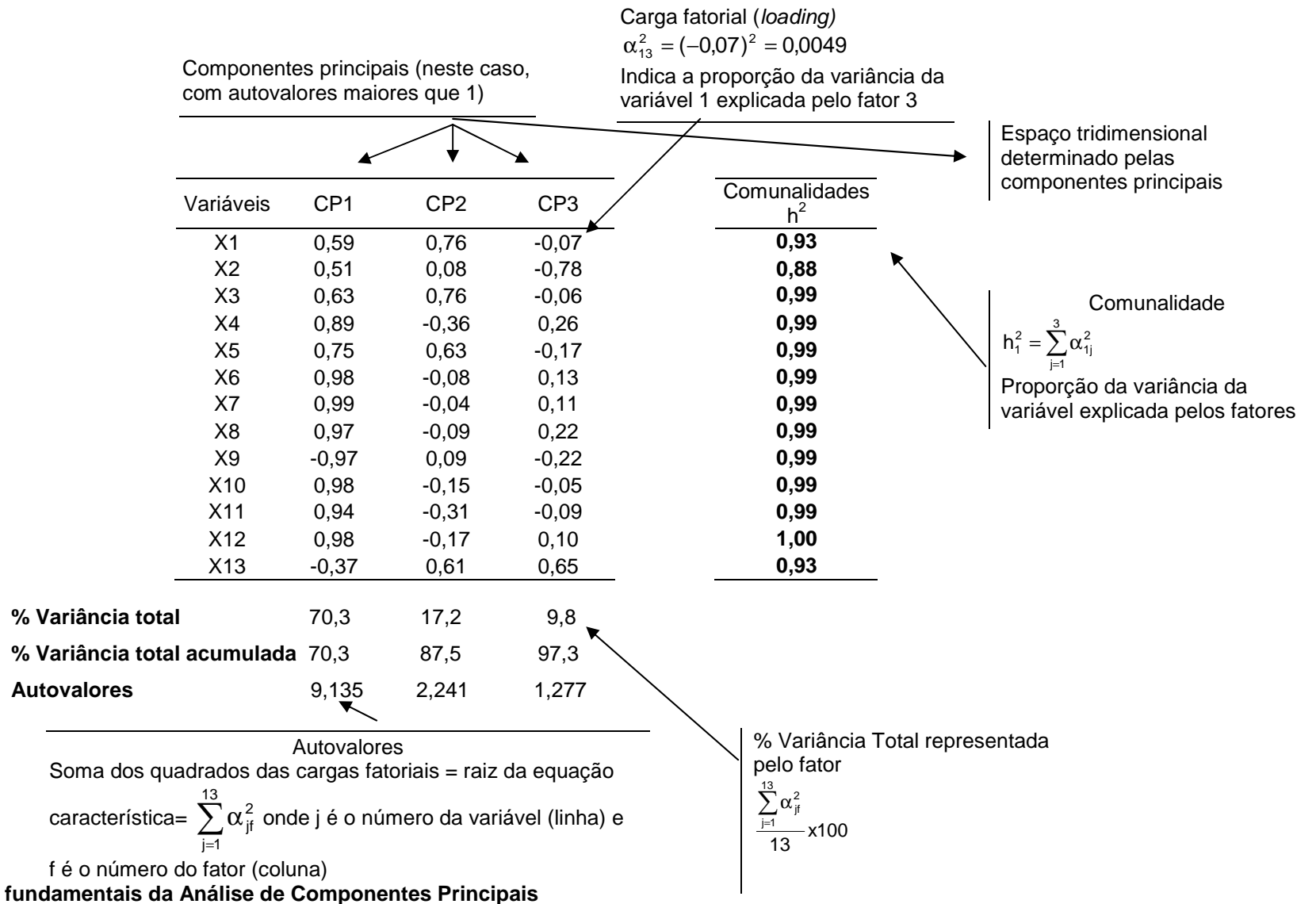
As comunalidades mostram o quanto os fatores juntos são capazes de explicar uma variável. Para que os fatores apresentem grande capacidade de explicação de uma variável, a comunalidade deve estar próxima de 1. Quanto menor o valor da comunalidade, menos os fatores conseguem explicar uma variável.

i) A comunalidade de uma variável é obtida pela soma dos quadrados das cargas

fatoriais da variável, ou seja, $h_m^2 = \sum_{i=1}^p \alpha_{i,m}^2$. (22)

O Quadro 3 apresenta os elementos fundamentais discutidos até este ponto sobre a Análise de Componentes Principais.

Quadro 3 – Elementos fundamentais da Análise de Componentes Principais



Quadro 3: Elementos fundamentais da Análise de Componentes Principais

3.10 Rotação

Para auxiliar na interpretação dos fatores obtidos, usa-se uma importante estratégia: a rotação dos fatores em torno da origem até que uma posição mais conveniente seja alcançada. A rotação pode ser ortogonal ou oblíqua.

Um importante objetivo a ser alcançado com as rotações é a geração de matrizes mais simples, ou seja com o maior número possível de valores iguais ou próximos de zero. Ao simplificar uma coluna da matriz fatorial, ou seja, gerar um novo fator com o máximo de valores próximos de zero, ter-se-á um menor número de cargas fatoriais altas, o que deixará mais clara a estrutura do fator.

Ao aplicar a rotação dos fatores o que se consegue é a redistribuição da variância pelos fatores buscando visualizar um padrão mais simples onde se consiga identificar com mais clareza a distribuição dos fatores. Com a rotação ortogonal, os novos fatores obtidos são não-correlacionados (independentes), enquanto na rotação oblíqua os novos fatores podem ser correlacionados uns com os outros. Portanto, dependendo das intenções do pesquisador, deve-se optar pelo tipo mais adequado de rotação.

Os métodos mais comuns de rotação ortogonal são: Quartimax, Varimax e Equimax. A rotação Quartimax, que trabalha sobre as linhas da matriz, busca fortalecer um fator e reduzir a representatividade dos demais, apresentando uma estrutura mais simples para análise. Nem sempre esta é uma boa opção, principalmente quando se pretende trabalhar a estrutura de distribuição das variáveis. A rotação Varimax promove a simplificação operando sobre as colunas da matriz de fatores. De acordo com Hair (2005a), essa rotação parece apresentar uma distribuição mais clara dos fatores, sendo Equimax um método de rotação intermediário, pois busca simplificações com operações sobre as linhas e colunas da matriz.

Quanto aos métodos de rotação oblíqua, que permitem a existência de fatores correlacionados, destacam-se as técnicas OBLIMIN, PROMAX e ORTHOBLIQUE.

A Figura 31 mostra, em situação hipotética, aplicações das rotações ortogonal e oblíqua. Neste caso em particular, percebe-se que, com a aplicação da rotação oblíqua, foram obtidos fatores mais próximos dos agrupamentos de pontos (variáveis), representando melhor sua estrutura de distribuição.

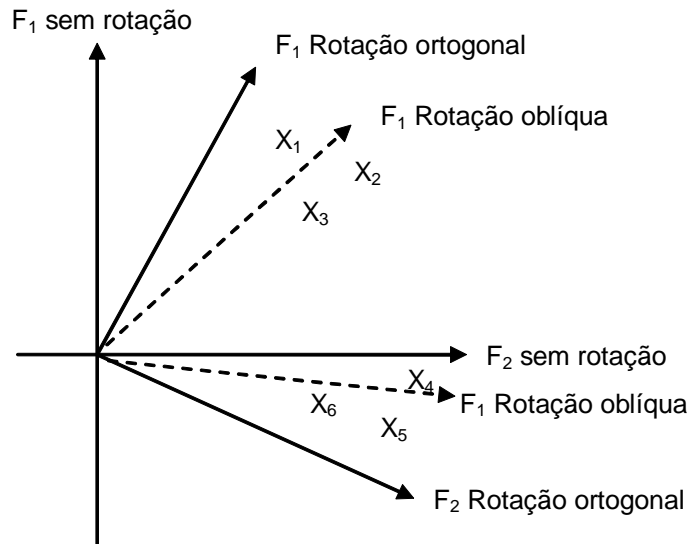


Figura 31: Rotações ortogonal e oblíqua de fatores
 Fonte: Hair et al., 2005b,p.393 (Adaptado)

Quanto à escolha do método de rotação a ser utilizado, deve-se pensar nos objetivos da pesquisa. Se o objetivo for de reduzir o número de variáveis independentes, mas mantendo a independência entre elas para uso em outras técnicas estatísticas, como nas regressões, deve-se optar pela rotação ortogonal. Mas, se o principal interesse é o de avaliar a estrutura do conjunto de variáveis, a melhor opção é pela rotação oblíqua.

Neste trabalho, faz-se a opção pela rotação oblíqua tendo em vista a intenção de uso de fatores de ordens superiores, que será explicado adiante, e do objetivo de identificação da estrutura de fluxos populacionais entre as microrregiões brasileiras.

3.10.1 Método de rotação oblíqua PROMAX

Alguns critérios devem ser observados na escolha do melhor método de rotação a ser aplicado na Análise Fatorial. 1) Parcimônia e simplicidade na própria estrutura do

método; 2) A facilidade e agilidade para obtenção dos resultados; 3) Replicabilidade de fatores (que destacam as mesmas características ou mostram estruturas semelhantes às aquelas obtidas com os dados originais); 4) Eficiência e 5) Adequação aos propósitos da pesquisa.

Gorsuch (1970), fez uma comparação entre quatro métodos de rotação: BIQUARTIMIN, MAXPLANE, PROMAX e VARIMAX, levando em consideração alguns dos critérios acima e outros que considerou importantes. Em geral, os quatro métodos se mostraram eficientes, mas, dentre os métodos de rotação oblíqua o que apresentou melhores resultados e agilidade foi o PROMAX.

Neste trabalho, faz-se a opção pelo método de rotação oblíqua PROMAX, desenvolvido por Hendrickson and White, em 1964. Acredita-se que a escolha por um método de rotação oblíqua contempla melhor a realidade original dos dados, tendo em vista que os agrupamentos de microrregiões sugeridos de acordo com o padrão de migração não são totalmente sem correlação e, ainda, esse método proporcionará novos fatores (e por conseqüência, novos agrupamentos) com alguma correlação, o que viabilizará a obtenção dos fatores de ordens superiores, que serão propostos na metodologia.

Outra qualidade importante encontra-se no fato de o método PROMAX poder ser aplicado com maior rapidez que outro método de rotação oblíqua encontrado nos softwares estatísticos, o DIRECT OBLIMIN, sendo, portanto, mais usado para grandes matrizes de dados, como ocorre no caso das 558 microrregiões brasileiras. Além disso, os fatores obtidos via método PROMAX são mais simples que os obtidos via método DIRECT OBLIMIN, pois este último procura alcançar autovalores maiores o que aumenta a complexidade dos fatores ao captar maior variância em cada fator.

Na prática, o Método PROMAX é aplicado sobre os resultados obtidos via rotação ortogonal VARIMAX. Ou seja, prioriza a simplificação das colunas, ou ainda, dos fatores.

3.11 Análise Fatorial dos tipos R e Q

A Análise Fatorial pode ser classificada em dois tipos, R e Q, dependendo da disposição dos dados na matriz original.

Neste momento, vale apresentar dois tipos de Análise Fatorial e, portanto, também da Análise de Componentes Principais: R e Q. A Análise Fatorial do tipo Q é aplicada quando se deseja identificar grupos homogêneos, por exemplo, de municípios, clientes, etc., Nesse caso, os elementos que serão agrupados ficam nas colunas e as variáveis nas linhas da matriz de dados. Enquanto a Análise do tipo R favorece o estudo da estrutura do conjunto de variáveis e, com isso, as variáveis ocupam as colunas da matriz de dados.

Portanto, se um pesquisador estiver interessado em identificar municípios com perfis semelhantes, de acordo com um conjunto de variáveis socioeconômicas, e sugerir novos agrupamentos para análises em separado, deve optar pelo tipo Q de Análise Fatorial, onde as variáveis ocuparão as linhas e os municípios as colunas da matriz de dados.

Se o objetivo do pesquisador for o de identificar variáveis com comportamentos e variabilidades semelhantes e criar indicadores, por exemplo, de potencial econômico, de qualidade de vida, de urbanização, etc. dentro seu conjunto maior de variáveis, deve optar pela Análise do tipo R, onde os municípios devem ocupar as linhas e as variáveis as colunas da matriz de dados.

Neste trabalho, ao lidar com a matriz de fluxos populacionais, o que identificará o tipo de Análise de Componentes Principais, R ou Q, será a disposição das microrregiões de origem nas linhas ou nas colunas.

4. Análise Fatorial de Ordem Superior

Segundo Clayton (1982, p.11), a técnica de Análise Fatorial de ordem superior permite reduzir sucessivamente a complexidade da matriz de dados com uma série de passos hierárquicos relacionados.

Como visto anteriormente, a técnica de Análise de Componentes Principais está diretamente relacionada à matriz de correlação entre as variáveis e os fatores gerados são linearmente independentes, ou seja, não-correlacionados. Deve-se reforçar ainda que a Análise de Componentes Principais é um caso particular da Análise Fatorial que, quando aplicada, permite a redução da dimensionalidade apresentada pelo conjunto de variáveis retendo uma variabilidade significativa da informação, eliminando parte da redundância registrada.

Os fatores extraídos do conjunto de dados no estado original (dados brutos) são referidos como de primeira ordem ou fatores primários e, como citado acima, quando obtidos via Análise de Componentes principais, são independentes, não apresentando correlação entre os fatores. Mas, se os fatores primários são submetidos a uma rotação oblíqua é possível identificar uma estrutura de novos fatores inter-relacionados e, portanto gerar uma nova matriz de correlação. Sobre essa matriz de correlação, aplica-se novamente a técnica de Análise de Componentes Principais gerando um conjunto secundário de fatores, ou fatores de segunda ordem, em geral, menos complexo que o primeiro. Aplicando-se sucessivamente a mesma seqüência de passos, obtém-se os fatores de terceira ordem, quarta ordem e assim por diante. Este processo se segue até que seja gerado apenas um fator ou todos os fatores gerados sejam iguais ou independentes ou, ainda, até o ponto em que o pesquisador considerar adequado, tendo em vista que, quanto maior a ordem, mais generalizador é o fator.

A Análise Fatorial de Ordem Superior caminha de um baixo grau de generalização para outros com maiores níveis de generalização. Gorsuch (1983, p.240) apresenta uma interessante analogia que ilustra bem a idéia dessa técnica:

Considere o planeta Terra como sendo o mais alto nível fatorial e, portanto, com maior grau de generalização de uma hipotética análise de dados. Em um nível mais específico, onde o grau de generalização é menor, mas a acurácia aumenta, pode-se dividir o planeta em áreas de terra ou água. Em outro nível, divide a terra entre continentes e ilhas, enquanto a água é dividida em oceanos, lagos e demais fluxos d'água. Cada uma dessas novas categorias, neste novo nível, pode ser subdividida:

continentes são subdivididos (Figura 32). Assim, o Planeta seria o fator de maior nível e, portanto, com maior grau de generalização enquanto os continentes apresentariam maior detalhamento e, portanto, complexidade.

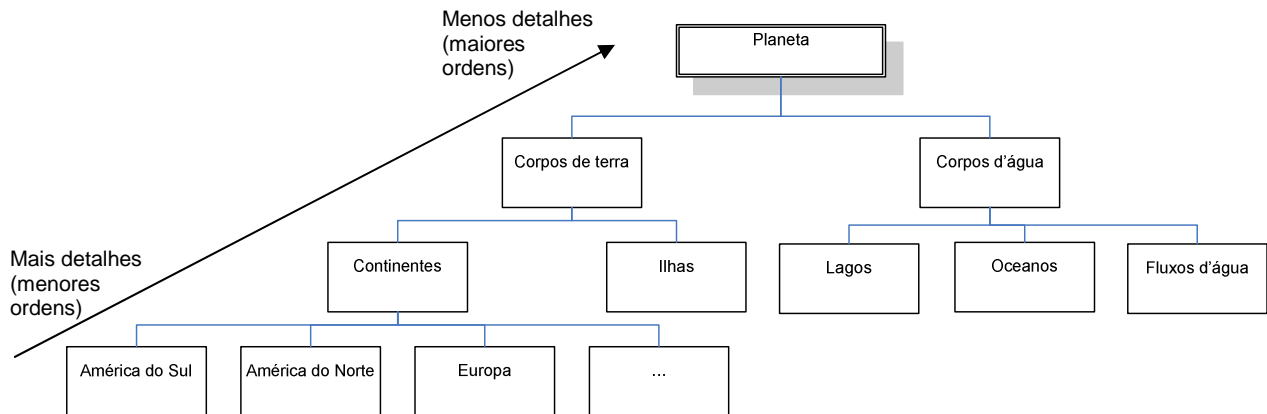


Figura 32: Uma análise de alta ordem geográfica

Fonte: Gorsuch, 1983, p.240 (Adaptado)

Outra analogia interessante é apresentada por Gray (1997, p.6). Considere uma matriz derivada de um grande conjunto de dados como sendo algo como a visão obtida por um observador posicionado em uma das ruas da cidade de Nova York. Os fatores primários poderiam fornecer uma visão da cidade do topo do Empire State Building. Os fatores secundários dariam a perspectiva de um piloto sobrevoando a cidade e os de terceira ordem, algo como um panorama visto do espaço.

Um aspecto importante e que deve ser levado em consideração é sobre o relacionamento entre as variáveis originais e os vários níveis de fatores extraídos. Cattell, citado por Clayton (1982, p.13), sustenta que os relacionamentos entre os fatores de diferentes níveis podem ser descritos como uma rede e que a estrutura desta rede pode, mas não precisa ser, de natureza hierárquica. Uma estrutura hierárquica pode ser imposta na rede fazendo certas suposições arbitrárias tais como escolher para examinar apenas as maiores cargas fatoriais de cada variável (ou novo fator).

Clayton (1982), Gorsuch (1983) e Holsman (1980) sugerem uma técnica para avaliar o peso das variáveis originais presente nos fatores obtidos em ordens superiores. A

demonstração da validade desta técnica pode ser vista em Gorsuch (1982). O que se faz é multiplicar a matriz de cargas fatoriais de uma ordem pela matriz de cargas fatoriais do nível imediatamente superior. Isso será possível, pois, o número de colunas da primeira matriz é sempre igual ao número de linhas da segunda. Desta maneira, pode-se calcular a comunalidade associada às variáveis originais ao trabalhar com os fatores de ordens superiores e o pesquisador terá como avaliar qual o poder de explicação perdido na generalização.

A Figura 33, com diagrama criado por Clayton (1982, p. 14) apresenta os estágios do procedimento de Análise Fatorial de Ordens Superiores.

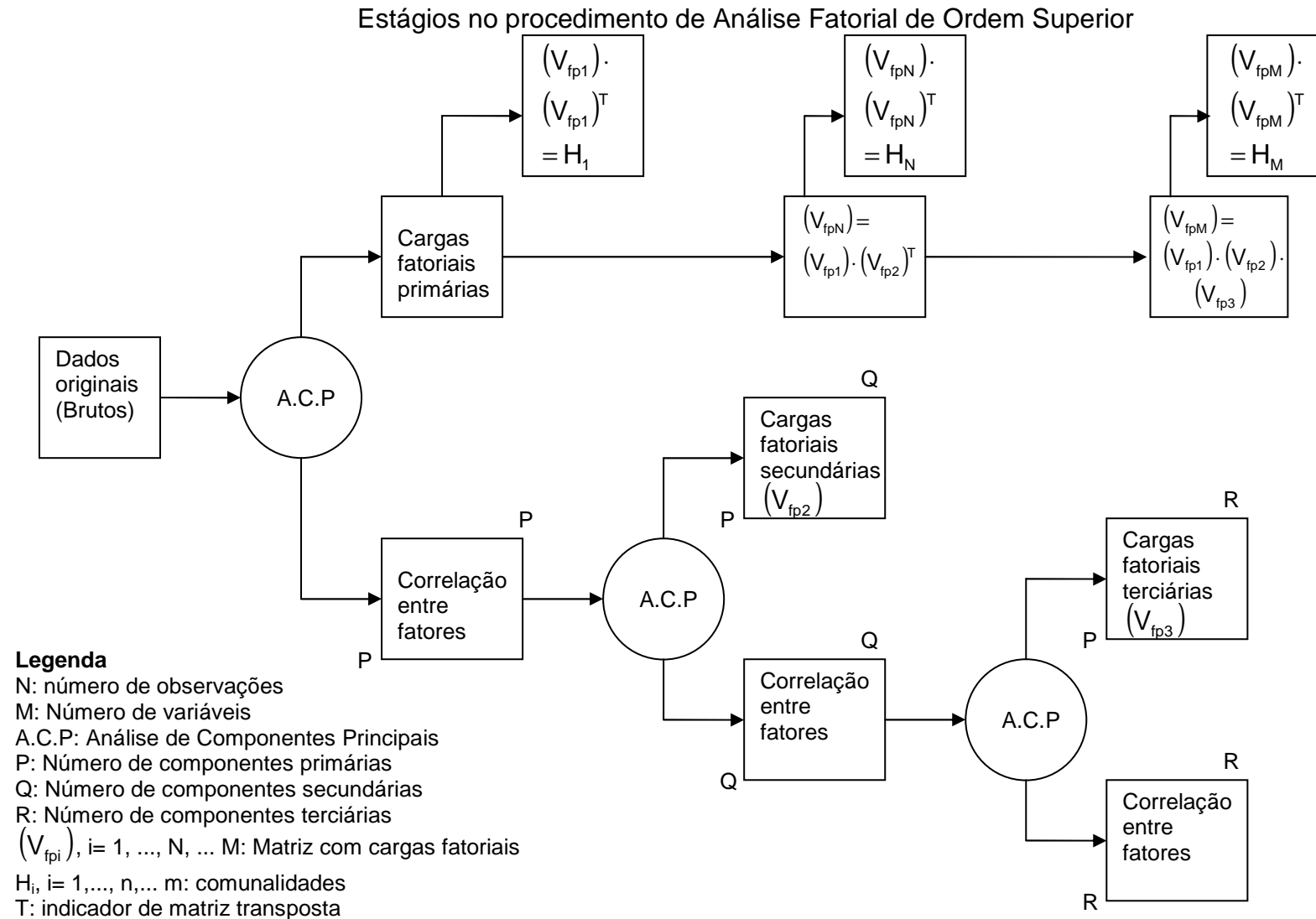


Figura 33: Estágios no procedimento de Análise Fatorial de Ordem Superior

Fonte: Clayton, 1982, p. 14 (Adaptado)

5 Matriz de fluxos populacionais e Análise de Componentes Principais

Vários fenômenos podem ser representados por um conjunto de células organizadas em uma caixa de dados tridimensional onde, as dimensões são denominadas entidades, atributos e ocasiões (ou momentos). As duas primeiras dimensões podem ser analisadas simultaneamente e sua evolução temporal fica dirigida pela terceira dimensão. É comum posicionar cada entidade (ou caso) em uma linha da matriz enquanto os atributos (características) são posicionados nas colunas, com cada coluna indicando uma variável. Neste trabalho, cada coluna da matriz é denominada fator.

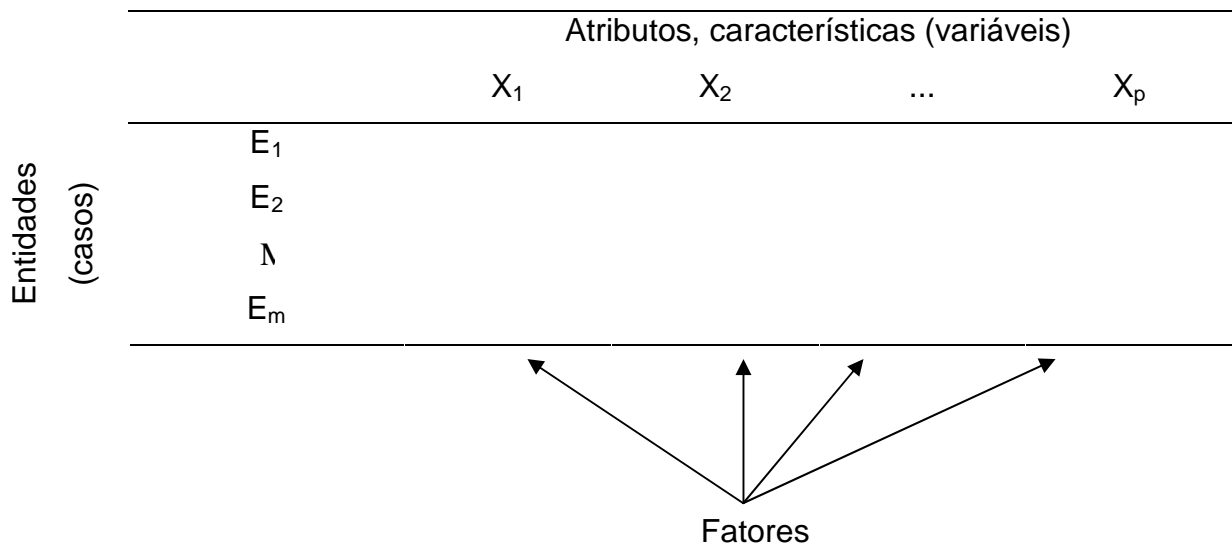


Figura 34: Representação de uma matriz de dados

Para cada ocasião (ou momento) pode-se organizar uma tabela como essa, obtendo-se, assim, a terceira dimensão.

No contexto geográfico, cada célula, identificada pelo cruzamento de uma linha com uma coluna, traz consigo, implicitamente, a grandeza espacial, a localização, componente fundamental para a Análise Espacial e que permitirá mapeamento dos resultados.

A Análise de Componentes Principais e a Análise Fatorial têm sido aplicadas na investigação de matrizes de fluxos na Geografia e nas Ciências Políticas. Este método de análise permite a substantiva redução da dimensionalidade do banco

de dados enquanto, simultaneamente, retém uma grande porção da informação original. (Clayton, 1977b, p.111)

Para a análise de fluxos, em geral, a matriz é quadrada com as linhas e colunas representando as mesmas unidades espaciais. Essa matriz, comumente denominada matriz origem-destino, dispõe em suas linhas as unidades espaciais que enviam, ou seja, são origens da informação, enquanto as colunas são as unidades espaciais que recebem as informações ou são os destinos. Portanto, cada unidade espacial respeita duas funções diferentes na análise espacial.

Esta matriz origem-destino pode ser submetida à Análise de Componentes Principais. Quando as linhas representarem as origens e as colunas os destinos, aplicar-se-á uma análise do Tipo R (vide Apêndice B) que permitirá ao pesquisador o agrupamento de destinos com variabilidades semelhantes e das principais origens relacionadas a esse agrupamento.

Na Análise de Componentes Principais do Tipo R os agrupamentos de destinos (que assumem o papel das variáveis) são determinados pelos fatores. Em cada fator, são identificadas as unidades espaciais com maiores cargas fatoriais e que, portanto, apresentam variabilidades semelhantes e os escores indicarão as origens mais fortemente relacionadas a cada agrupamento. Mais adiante, esta questão será abordada mais detalhadamente.

Trabalhando com a transposta da matriz origem-destino tradicional, tem-se nas linhas os destinos e nas colunas as origens dos fluxos. Nesse caso, faz-se a aplicação da Análise de Componentes Principais do tipo Q, quando será possível identificar agrupamentos de origens que apresentam variabilidades semelhantes e os principais destinos relacionados a cada agrupamento.

APÊNDICE C – Mapas de localização

APÊNDICE D – Mapas operacionais produzidos com base nos resultados da aplicação do método Intramax para o período 1986-1991

APÊNDICE E – Mapas operacionais produzidos com base nos resultados da aplicação do método Intramax para o período 1995-2000

APÊNDICE F – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1986-1991/Matriz Origem-Destino

APÊNDICE G – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1986-1991/Matriz Destino-Origem

APÊNDICE H – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1995-2000/Matriz Origem-Destino

APÊNDICE I – Mapas operacionais produzidos com base nos resultados da aplicação da Análise Fatorial para o período 1995-2000/Matriz Destino-Origem

APÊNDICE J – Dendrogramas

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)