



Universidade Federal de Pernambuco - UFPE

Josemir Araújo Neves

Um Índice de Susceptibilidade ao Fenômeno da Seca para o Semi-árido Nordestino



PERCY FRAGOSO. Cadeia de aridez. 2008. Cadeia em papel e tela, 0,44 x 0,70 m.

Recife - PE

2010

Josemir Araújo Neves

Um Índice de Susceptibilidade ao Fenômeno da Seca para o Semi-árido Nordestino

Tese apresentada ao Programa de Pós-Graduação em Matemática Computacional da Universidade Federal de Pernambuco, como parte dos requisitos exigidos para a obtenção do título de Doutor em Matemática Computacional.

Orientador: Prof. Dr. Silvio de Barros Melo

Co-orientador: Prof. Dr. Everardo Valadares de Sá Barreto Sampaio

Recife - PE

2010

Neves, Josemir Araújo.

Um índice de susceptibilidade ao fenômeno da seca para o semi-árido nordestino / Josemir Araújo Neves. - Recife: O Autor, 2010.

xxvi, 399 p. : il., fig., tab., quadros

Tese (doutorado) - Universidade Federal de Pernambuco, CCEN. Matemática Computacional, 2010.

Inclui bibliografia e apêndices.

1. Estatística matemática. 2. Preenchimento múltiplo. 3. Índice de seca. 4. Semi-árido nordestino. I. Título.

519.5

CDD(22. ed)

MEI2010-087



UNIVERSIDADE FEDERAL DE PERNAMBUCO
Centro de Ciências Exatas e da Natureza
Doutorado em Matemática Computacional

PARECER DA BANCA EXAMINADORA DE DEFESA DE TESE DE DOUTORADO

JOSEMIR ARAÚJO NEVES

*“Um índice de susceptibilidade ao fenômeno da seca
para o semi-árido nordestino”*

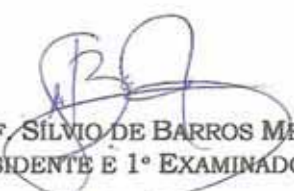
A Banca Examinadora composta pelos Professores: SÍLVIO DE BARROS MELO (Orientador), do Centro de Informática da UFPE; MARCÍLIA ANDRADE CAMPOS, do Centro de Informática da UFPE; JOSENILDO DOS SANTOS, do Departamento de Ciências Contábeis da UFPE; EVERARDO VALADARES DE SÁ BARRETO SAMPAIO, do Departamento de Energia Nuclear da UFPE; e HUMBERTO ALVES BARBOSA, do Instituto de Ciências Atmosféricas da UFAL, considera a Tese do candidato:

☒ APROVADA

☐ EM EXIGÊNCIA

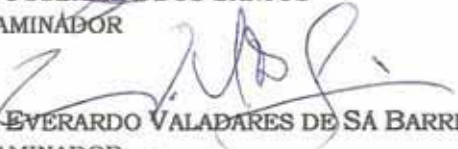
☐ REPROVADA

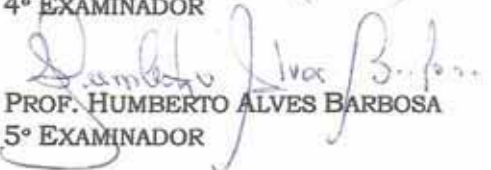
Secretaria do Programa de Pós-Graduação em Matemática Computacional do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, aos 14 dias do mês de junho de 2010.


PROF. SÍLVIO DE BARROS MELO
PRESIDENTE E 1º EXAMINADOR


PROF.ª MARCÍLIA ANDRADE CAMPOS
2º EXAMINADOR


PROF. JOSENILDO DOS SANTOS
3º EXAMINADOR


PROF. EVERARDO VALADARES DE SÁ BARRETO SAMPAIO
4º EXAMINADOR


PROF. HUMBERTO ALVES BARBOSA
5º EXAMINADOR

*A Deus, a quem devemos tudo,
e aos meus pais, que como pessoas simples do sertão,
criaram os seus filhos sob ideais de integridade e
honestidade.*

Agradecimentos

Inicialmente agradeço a Deus, criador supremo do universo, por ter me presenteado com a vida e as condições necessárias para executar esse trabalho.

À minha família, especialmente a minha irmã Jaércia, que sempre estiveram presentes ao longo dessa minha jornada na terra.

Ao Prof. Sílvio, orientador e amigo, por ter aceitado o convite para fazer minha orientação, e por toda a contribuição durante a pesquisa.

Ao Prof. Everardo, co-orientador e amigo, pela importante contribuição dispendida nesse trabalho, através do seu amplo conhecimento sobre o tema e a sua competência.

À EMPARN e a Embrapa, que viabilizaram a execução e conclusão desse trabalho.

À minha namorada Lenira, pelo amor, paciência e compreensão.

À minha amiga Ângela, companheira de muitas jornadas durante o curso, por toda a sua ajuda, principalmente nas disciplinas teóricas.

À minha amiga Marlúcia, sempre colaboradora para o meu sucesso.

Em memória de Rommel Mesquita de Faria (Ex-Diretor da EMPARN).

Aos queridos e prestimosos amigos de ideal espírita Raquel Soares, Jaciana Medeiros, Liliam Medeiros, Otávio Pereira, Bete Pereira, Ednar Santos, os amigos do grupo de estudos Joanna de Ânglis e os amigos do ESDE/EADE, pelo apoio carinho e amizade dispensados durante os últimos quatro anos.

Ao Programa de Pós-graduação em Matemática Computacional pelo financiamento dos dados das estações meteorológicas do Rio Grande do Norte adquiridos junto ao INMET.

À todas as pessoas que de alguma forma incentivaram e contribuíram para o desenvolvimento desse trabalho, meu muito obrigado.

“[...] O Senhor não te identificará pelos tesouros que ajuntaste, pelas bênçãos que retiveste, pelos anos que viveste no corpo físico. Reconhecer-te-á pelo emprego dos teus dons, pelo valor de tuas realizações e pelas obras que deixaste, em torno dos teus próprios pés. ”

Emmanuel.

Resumo

A seca é o principal fenômeno climático que atinge o semi-árido nordestino de forma recorrente, com repercussões negativas extremas de déficit hídrico e dimensão de catástrofe sócio-econômico-cultural-ambiental. Nesse contexto os índices de seca apresentam-se como ferramentas de auxílio à tomada de decisão dos gestores públicos e à aplicação correta e otimizada dos recursos, possibilitando o direcionamento das ações a fim de que efetivamente alcancem a população atingida. Geralmente, os índices de seca utilizados no semi-árido nordestino consideram aspectos apenas de natureza meteorológica e/ou hidrológica. Uma das exceções é o Índice Funceme de Severidade à Seca (IFS), criado pela Fundação Cearense de Meteorologia, em 2000, mas que não chegou, a ser implementado, devido a dificuldades na operação e obtenção dos dados. Há, portanto, a necessidade do desenvolvimento de índices que possibilitem uma melhor compreensão do fenômeno, usando dados acessíveis e confiáveis, e que sejam facilmente manipulados. O objetivo dessa tese foi construir um Índice de Susceptibilidade ao Fenômeno da Seca (ISFS), que considera além da climatologia, aspectos da natureza física dos solos, econômicos, sociais e os riscos de perda da safra agrícola e de falta de água para consumo humano e animal. O ISFS foi elaborado seguindo metodologia proposta pelo JRC/OECD, utilizando técnicas de preenchimento de dados com falhas e winsorização. Ele foi testado e validado com os dados reais dos municípios do Rio Grande do Norte, área objeto de estudo, em três cenários climatológicos: ano seco, normal e chuvoso. Como parte do processo de validação, realizou-se uma análise de sensibilidade do ISFS utilizando uma abordagem de simulação de Monte Carlo, cujos resultados validaram as classificações dos municípios obtidos nos testes com dados reais e a consistência dos pesos adotados na sua composição. Por fim, foi desenvolvido na linguagem C Sharp um aplicativo computacional para a sua operacionalização em 8 módulos. Esse aplicativo utiliza bibliotecas estatísticas do R e apresenta uma série de facilidades de importação e entrada de dados, cálculo de estatísticas, geração de mapas e relatórios.

Palavras-chaves: 1 Índice de Severidade de Seca - 2 Semi-árido Nordeste - 3 Indicadores Compostos - 4 Preenchimento Múltiplo de Dados com Falhas - 5 Winsorização.

Abstract

Drought is the major climatic phenomenon that affects the semi-arid region of Northeast Brazil on a recurring basis, with extreme negative consequences on water availability, resulting in environmental, economic, social and cultural catastrophes. Drought indices are tools that aid decisions of public managers, guiding actions to effectively reach the affected population. Generally, indices used in Northeast Brazil are based solely on meteorological or hydrological data, with the exception of the FUNCEME's Drought Severity Index (FDSI), created by Fundação Cearense de Meteorologia, in 2000, that was not implemented due to operational difficulties. Therefore, there is a lack of an index that allows a better understanding of the phenomenon, using available and reliable data, and that can be easily manipulated. The objective of this thesis was to construct an Index of Susceptibility to Drought (ISD), which considers not only climatic data, but also soil, economic and social aspects, plus risks of harvest loss and shortage of human and animal drinking water. The ISD was constructed following the method proposed by JRC/OECD, using techniques for multiple imputation and data winsorization. It was tested and validated with real data from all Rio Grande do Norte municipalities, in three climatic scenarios: dry, normal and wet years. The ISD classification of the municipalities and the consistency of the weights adopted in the ISD construction were validated with a sensitivity analysis, using Monte Carlo simulations. A software written in C Sharp and composed of eight modules was developed to obtain and operate the ISD. This applicative uses libraries statistics on R and has import and data entry facilities, performs statistical calculations, and generates maps and reports.

Keywords: Drought Index - Semi-arid Northeast - Composite Indicators - Multiple Imputation - Winsorization.

LISTA DE FIGURAS

2.1 Áreas susceptíveis a desertificação no Nordeste	036
2.2 Repercurssões do fenômeno El Niño ao redor do mundo	038
2.3 Delimitação do semi-árido brasileiro	040
2.4 Indicadores Sociais, Brasil	044
2.5 Indicador de Inclusão Previdenciária (Brasil) e de Transferências Governamentais(NE)	045
2.6 Evolução de IDH por Regiões, Brasil, Período: 1991 a 2005	046
2.7 Variação de IDH e de seus componentes entre 1991 e 2005, por Estados, Brasil	046
2.8 Índice de Desenvolvimento Humano Municipal dos Municípios do Nordeste	047
2.9 Desvios de precipitação em relação a média climatológica do período chuvoso do Semi-árido do Nordeste	060
2.10 Messorregiões do Rio Grande do Norte	063
2.11 Microrregiões do Rio Grande do Norte	064
2.12 Densidade demográfica (hab/ km^2) dos municípios do Rio Grande do Norte	065
2.13 População rural (hab) dos municípios do Rio Grande do Norte	066
2.14 Altimetria do Rio Grande do Norte	067
2.15 Classificação climática do Rio Grande do Norte - Segundo Thornthwaite	067
2.16 Saldo do Emprego Formal da Agropecuária - RN - Período: 12/2000 a 02/2009	070
2.17 Índice de Desenvolvimento Humano Municipal, Rio Grande do Norte	071
2.18 Bacias Hidrográficas do Rio Grande do Norte	072
2.19 Percentual de pessoas que vivem em domicílios com água encanada	073
2.20 Evolução do número de projetos, municípios beneficiados pelos assentamentos rurais do INCRA (A) e área média por famílias assentadas (B)	074
2.21 Processo de planejamento e gestão dos riscos das secas	080

2.22	Árvore dos Aspectos Fundamentais (IFS)	086
2.23	Detalhe dos Parâmetros PV1, PV2, PV3 e PV4 (IFS)	087
2.24	Detalhe dos Parâmetros PV5, PV6 e PV7 (IFS)	088
2.25	Árvore dos Aspectos Fundamentais (IMERG)	090
2.26	Esquema Interativo do MACBETH	095
2.27	Escala de pontuação - Efeito sobre a população rural	098
2.28	Escala de pontuação - Condições municipais de atenuação dos efeitos da seca	099
3.1	Pirâmide de Informações	103
3.2	Evolução do número de índices desenvolvidos no mundo	108
3.3	Distribuição dos índices desenvolvidos por tipo de instituição	108
3.4	Distribuição das instituições por quantidades de índices desenvolvidos	108
3.5	Exemplo de decomposição em um gráfico de barras - Índice de Avanço Tecnológico ..	126
3.6	Exemplo de decomposição em um gráfico de radar - Índice de Avanço Tecnológico ...	126
3.7	Exemplo de decomposição em um gráfico semáforo - Índice de Avanço Tecnológico ..	127
3.8	Padrões de não respostas dos dados. (a) Padrão univariado, (b) Padrão Monotônico e (c) Padrão arbitrário	129
3.9	Representação gráfica dos mecanismos de não resposta dos dados: (a) MCAR, (b) MAR e (c) MNAR	130
3.10	Representação esquemática dos passos de MI utilizados no MICE	142
4.1	Mapa de Aptidão Agrícola de Solos do Rio Grande do Norte- Sistema de Manejo Pouco Desenvolvido	148
4.2	Distribuição do sub-indicador Natureza Física - PV11	150
4.3	Gráfico da escala de pontuação utilizada para o sub-indicador Natureza Física - PV11	151
4.4	Localização dos Postos Pluviométricos e das Estações Meteorológicas do Rio Grande do Norte	155

4.5 Quantidade de Postos Pluviométricos e Estações Meteorológicas utilizados no cálculo dos quantis	156
4.6 Grade Retangular gerada a partir da Médias Pluvimétricas.	161
4.7 Gráfico da escala de pontuação - Natureza Climatológica - PV12	163
4.8 Distribuição da População Rural do Rio Grande do Norte. Fonte: IBGE, Censo 2000	164
4.9 Gráfico da escala de pontuação - Efeito sobre a Base Econômica - PV132	165
4.10 Gráfico da escala de pontuação - Condições Municipais de Atenuação à Seca - PV133	166
4.11 Distribuição da Receita per capita média dos municípios do Rio Grande do Norte (Período: 1994-2006). Fonte: Ministério da Fazenda, Secretaria Tesouro Nacional	166
4.12 Gráfico da escala de pontuação - Risco de perda da produção agrícola - PV21	169
4.13 Distribuição do Parâmetro: Risco de falta de água para consumo humano e animal - PV22	172
4.14 Gráfico da escala de pontuação - Risco de falta de água para consumo humano e animal - PV22	172
4.15 Estrutura do ISFS - Fluxograma de Cálculo	180
5.1 Índice de Distribuição de Chuvas (IDC) - Ano: 1998 - Período: janeiro-abril	186
5.2 Índice de Distribuição de Chuvas (IDC) - Ano: 1998 - Período: fevereiro-maio	186
5.3 Índice de Distribuição de Chuvas (IDC) - Ano: 1998 - Período: março-junho	187
5.4 Índice de Distribuição de Chuvas (IDC) - Ano: 1998 - Período: abril-julho	187
5.5 Índice de Distribuição de Chuvas (IDC) - Ano: 2000 - Período: janeiro-abril	188
5.6 Índice de Distribuição de Chuvas (IDC) - Ano: 2000 - Período: fevereiro-maio	188
5.7 Índice de Distribuição de Chuvas (IDC) - Ano: 2000 - Período: março-junho	189
5.8 Índice de Distribuição de Chuvas (IDC) - Ano: 2000 - Período: abril-julho	189
5.9 Índice de Distribuição de Chuvas (IDC) - Ano: 2005 - Período: janeiro-abril	190
5.10 Índice de Distribuição de Chuvas (IDC) - Ano: 2005 - Período: fevereiro-maio	190

5.11 Índice de Distribuição de Chuvas (IDC) - Ano: 2005 - Período: março-junho	191
5.12 Índice de Distribuição de Chuvas (IDC) - Ano: 2005 - Período: abril-julho	191
5.13 Índice de Distribuição de Chuvas (IDC) - Ano: 2006 - Período: janeiro-abril	192
5.14 Índice de Distribuição de Chuvas (IDC) - Ano: 2006 - Período: fevereiro-maio	192
5.15 Índice de Distribuição de Chuvas (IDC) - Ano: 2006 - Período: março-junho	193
5.16 Índice de Distribuição de Chuvas (IDC) - Ano: 2006 - Período: abril-julho	193
5.17 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 1998	195
5.18 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2000	195
5.19 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2005	196
5.20 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2006	196
5.21 Distribuição dos dados do sub-indicador PV132, antes e depois da transformação - Ano: 1998	200
5.22 Distribuição dos dados do sub-indicador PV131, antes e depois da transformação - Ano: 2000	201
5.23 Distribuição dos dados do sub-indicador PV132, antes e depois da transformação - Ano: 2000	201
5.24 Distribuição dos dados do sub-indicador PV131, antes e depois da transformação - Ano: 2005	202
5.25 Distribuição dos dados do sub-indicador PV132, antes e depois da transformação - Ano: 2005	202
5.26 Distribuição dos dados do sub-indicador PV131, antes e depois da transformação - Ano: 2006	203
5.27 Distribuição dos dados do sub-indicador PV132, antes e depois da transformação - Ano: 2006	203
5.28 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 1998 - Resultado Final ...	213
5.29 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2000 - Resultado Final ...	215

5.30 Situação dos sub-indicadores do ISFS no ano de 1998, para os municípios de Mossoró (A), Assu (B), Governador Dix Sept Rosado (C) e Jandaíra (D)	216
5.31 Situação dos sub-indicadores do ISFS no ano de 1998, para os municípios de Angicos (A), Cerro Corá (B) e Santa Cruz (C)	217
5.32 Situação dos sub-indicadores do ISFS no ano de 1998, para os municípios de Santana do Matos (A), Macau (B), Ielmo Marinho (C), Campo Grande (D), Severiano Melo (E) e Marcelino Vieira (F)	218
5.33 Situação dos sub-indicadores do ISFS no ano de 1998, para os municípios de Carnaubais (A), Serra de São Bento (B) e Apodi (C)	219
5.34 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2005 - Resultado Final ...	220
5.35 Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) - 2006 - Resultado Final ...	221
5.36 Situação dos sub-indicadores do ISFS no ano de 2006, para os municípios de Mossoró (A), Gov. Dix. Spt. Rosado (B), Assu (C) e Jandaíra (D)	222
5.37 Situação dos sub-indicadores do ISFS no ano de 2006, para os municípios de Itaú (A), Lajes (B) e Pedra Grande (C)	223
5.38 Situação dos sub-indicadores do ISFS (valores médios) para o ano de 1998	227
5.39 Situação dos sub-indicadores do ISFS (valores médios) para o ano de 2000	228
5.40 Situação dos sub-indicadores do ISFS (valores médios) para o ano de 2005	229
5.41 Situação dos sub-indicadores do ISFS (valores médios) para o ano de 2006	230
5.42 Situação dos sub-indicadores do ISFS - Brejinho	236
5.43 Situação dos sub-indicadores do ISFS - Jundiá	237
5.44 Situação dos sub-indicadores do ISFS - Santo Antônio	238
5.45 Situação dos sub-indicadores do ISFS - Almino Afonso	239
5.46 Situação dos sub-indicadores do ISFS - Doutor Severiano	240
5.47 Situação dos sub-indicadores do ISFS - Patu	241
5.48 Situação dos sub-indicadores do ISFS - Currais Novos	242

5.49 Situação dos sub-indicadores do ISFS - São Tomé	243
5.50 Classificação Pluviométrica pelo DI - 1998	248
5.51 Classificação Pluviométrica pelo DI - 2000	249
5.52 Classificação Pluviométrica pelo DI - 2005	249
5.53 Classificação Pluviométrica pelo DI - 2006	250
5.54 Índice de Anomalia de Precipitação (RAI) - 1998	251
5.55 Índice de Anomalia de Precipitação (RAI) - 2000	252
5.56 Índice de Anomalia de Precipitação (RAI) - 2005	253
5.57 Índice de Anomalia de Precipitação (RAI) - 2006	254
5.58 Índice Anual Regional (RI)	255
6.1 Boxplot da classificação do ISFS resultado da primeira simulação	263-264
6.2 Dominância estocástica da classificação do ISFS de todos os municípios do RN de acordo com a indexação da Tabela 6.1	265
6.3 Boxplot da classificação do ISFS resultado da segunda simulação - Ano: 1998 ...	268-269
6.4 Boxplot da classificação do ISFS resultado da segunda simulação - Ano: 2000 ...	270-271
6.5 Boxplot da classificação do ISFS resultado da segunda simulação - Ano: 2005 ...	272-273
6.6 Boxplot da classificação do ISFS resultado da segunda simulação - Ano: 2006 ...	274-275
6.7 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, primeira abordagem da segunda simulação - Ano: 1998	276
6.8 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, primeira abordagem da segunda simulação - Ano: 2000	277
6.9 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, primeira abordagem da segunda simulação - Ano: 2005	278
6.10 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, primeira abordagem da segunda simulação - Ano: 2006	279

6.11 Boxplot da class. do ISFS, segunda abordagem da segunda simulação-Ano:1998	282-283
6.12 Boxplot da class. do ISFS, segunda abordagem da segunda simulação-Ano:2000	284-285
6.13 Boxplot da class. do ISFS, segunda abordagem da segunda simulação-Ano:2005	286-287
6.14 Boxplot da class. do ISFS, segunda abordagem da segunda simulação-Ano:2006	288-289
6.15 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, segunda abordagem da segunda simulação - Ano: 1998	290
6.16 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, segunda abordagem da segunda simulação - Ano: 2000	291
6.17 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, segunda abordagem da segunda simulação - Ano: 2005	292
6.18 Dominância estocástica dos municípios do RN de acordo com a classificação do ISFS, segunda abordagem da segunda simulação - Ano: 2006	293
7.1 Diagrama esquemático dos componentes do ISFS_Sist	295
7.2 Tela principal do ISFS_Sist	297
7.3 Menus - Módulos de entradas de dados do ISFS_Sist	298
7.4 Menu do módulo ISFS	298
7.5 Menu do módulo Relatórios/Gráficos	299
7.6 Menu do módulo Estatísticas/Análises	299
7.7 Menu dos módulos Parâmetros e Rotinas Administração	299
7.8 Tela da Configuração dos Parâmetros do ISFS	300
7.9 Exemplo de tela de ajuda	301
7.10 Tela da entrada de dados da precipitação diária dos municípios	302
7.11 Tela da rotina de importação de dados dos PIBs municipais	303
7.12 Exemplo de arquivo de importação de dados dos PIBs municipais	303
7.13 Tela da tabela de microrregiões	304

7.14 Tela do cálculo do IDC	305
7.15 Módulo de geração dos mapas do IDC (opção: Estado)	306
7.16 Módulo de geração dos mapas do IDC (opção: Microrregiões)	306
7.17 Módulo de entrada de dados do sub-indicador PV133	308
7.18 Módulo de cálculo do ISFS (tela mostrando o cálculo em andamento)	309
7.19 Módulo de cálculo do ISFS (tela final)	310
7.20 Módulo de geração dos mapas do ISFS para todo o Estado	311
7.21 Módulo de geração dos mapas do ISFS por microrregião	311
7.22 Exemplo de relatório da classificação geral do ISFS	312
7.23 Exemplo de relatório dos perfis municipais	314
7.24 Módulo Padrão de Dados	315
7.25 Módulo Análise das Estimativas - Modelos de Regressão	316

LISTA DE QUADROS

2.1 Distribuição da área e da população do novo semi-árido por Estado da Federação ...	041
2.2 Número de imóveis rurais e área agrícola	042
2.3 Registro histórico da ocorrência de secas no Nordeste	048
2.4 Número de açudes públicos e em cooperação construídos pelo DNOCS no período de 1909 a 1981	051
2.5 Variação da produção das principais lavouras no Nordeste no período de 1969 a 1970	056
2.6 Evolução da área plantada e produção das principais culturas temporárias e permanentes do NDS	069
2.7 Evolução das receitas correntes a preços de 2007	070
2.8 Distribuição do número e da área dos estabelecimentos rurais do Rio Grande do Norte	074
2.9 Características dos principais índices de seca utilizados no mundo	077
2.10 Plano de preparação para a seca do NDMC.	079
2.11 Esquema de julgamento do MACBETH	094
2.12 Limites dos descritores definidos para os PVs do IMERG	096
2.13 Aptidão agrícola das terras.	097
3.1 Características necessárias de um bom indicador ou indicador composto	106
3.2 Vantagens e desvantagens dos indicadores compostos	109
3.3 Algoritmos de agrupamento utilizados na AA	114
3.4 Métodos de normalização	118
3.5 Aplicativos computacionais utilizados no preenchimento único e múltiplo dos dados com falhas	139
3.6 Métodos de preenchimento das falhas disponíveis no MICE	142
4.1 Descrição das classes de aptidão agrícola utilizadas no Mapa da Figura 4.1	147

4.2 Convenções utilizadas na construção do sub-indicador Natureza Física - PV11	149
4.3 Propriedades desejáveis para quantis amostrais	158
4.4 Classificação utilizada para o Índice de Distribuição de Chuvas	168
5.1 Fontes de Informações utilizadas na construção do ISFS	183
5.2 Padrões de não-respostas do ISFS obtidos com o MICE	199

LISTA DE TABELAS

2.1 Escala de Pontuação - Aptidão Agrícola das Terras	097
2.2 Escala de Pontuação - Climatologia	098
2.3 Escala de Pontuação - Efeito sobre a base econômica	099
2.4 Escala de Pontuação - Risco de perda da safra agrícola	100
2.5 Escala de Criticidade - Falta para consumo humano e animal d'água - população rural	101
4.1 Resultado da classificação da aptidão agrícola de solos do RN	152
4.2 Resultado da classificação do sub-indicador Natureza Física - PV11	153
4.3 Classificação pluviométrica para a quadra chuvosa das microrregiões do Rio Grande do Norte, a partir da técnica dos quantis	159
4.4 Classificação pluviométrica para as microrregiões do Rio Grande do Norte, a partir da técnica dos quantis	160
4.5 Médias pluviométricas dos municípios do Rio Grande do Norte, obtidas pelo cálculo das medianas	162
5.1 Grau de susceptibilidade ao fenômeno da seca	194
5.2 Consistência dos dados dos sub-indicadores do ISFS	197
5.3 Valores de assimetria dos sub-indicadores do ISFS	200
5.4 Análise das estimativas dos modelos de regressão utilizados no preenchimento múltiplo - Ano: 1998	206
5.5 Análise das estimativas dos modelos de regressão utilizados no preenchimento múltiplo - Ano: 2000	207
5.6 Análise das estimativas dos modelos de regressão utilizados no preenchimento múltiplo - Ano: 2005	208
5.7 Análise das estimativas dos modelos de regressão utilizados no preenchimento múltiplo - Ano: 2006	209
5.8 Testes de significâncias dos modelos de regressão no preenchimento múltiplo	210

5.9 Fração da informação ausente e eficiência relativa	211
5.10 Resultados da winsorização	211
5.11 Avaliação das diferenças entre as categorias dos gráficos de radar, ano a ano	231
5.12 Amplitude de variação do ISFS	232-235
5.13 Resultados da Análise de Componentes Principais do ISFS	245
5.14 Matriz de fatores de cargas rotacionados para os sub-indicadores do ISFS	246
6.1 Estatísticas da classificação do ISFS resultado da primeira simulação	260-262

ABREVIATURAS

AA	Análise de Agrupamentos
AC	Análise Conjunta
ACP	Análise de Componentes Principais
AF	Análise Fatorial
Agritempo	Sistema de Monitoramento Agrometeorológico
AHP	Análise Hierárquica de Processos
AMNC	Abordagem Multicritério Não Compensatória
BLR	<i>Bayesian Linear Regression</i>
BNDES	Banco Nacional do Desenvolvimento
BOD	<i>Benefit of the Doubt</i>
CEPAGRI	Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
DI	<i>Deciles Index</i>
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
EMPARN	Empresa de Pesquisa Agropecuária do Rio Grande do Norte
ENOS	El Niño-Oscilação Sul
ESI	<i>Environmental Sustainable Index</i>
FUNCEME	Fundação Cearense de Meteorologia e Recursos Hídricos
GTDN	Grupo de Trabalho para o Desenvolvimento do Nordeste
IDC	Índice de Distribuição de Chuvas
IFS	Índice FUNCEME de Severidade de Seca
IMA	Índice Municipal de Alerta
IMERG	Indicador de Emergência
INPE	Instituto de Pesquisa Espaciais
ISFS	Índice de Suscetibilidade ao Fenômeno da Seca

ISFS_Sist	Software de implementação do ISFS
JRC	<i>Joint Research Centre</i>
MACBETH	<i>Measuring Attractiveness by a Categorical Based Evaluation Technique</i>
MAR	<i>Missing at Random</i>
MCAR	<i>Missing Completely at Random</i>
MCDA	<i>Multicriteria Decision Aiding</i>
MCMC	<i>Markov Chain Monte Carlo</i>
MDMC	<i>National Drought Mitigation Center</i>
MI	<i>Multiple Imputation</i>
MICE	<i>Multiple Imputation by Chained Equations</i>
MNAR	<i>Missing not at Random</i>
MNT	Modelo Numérico de Terreno
NDS	Núcleo de Desertificação do Seridó
OECD	<i>Organization for Economic Co-operation and Development</i>
PAO	Processo de Alocação de Orçamentos
PMM	<i>Predictive Mean Matching</i>
PV	Ponto de Vista Fundamental
PV11	Sub-indicador Natureza Física
PV12	Sub-Indicador Natureza Climatológica
PV131	Sub-Indicador Efeito sobre a Base Rural
PV132	Sub-Indicador Efeito sobre a Base Econômica
PV133	Sub-Indicador Condições Municipais de Atenuação dos efeitos da Seca
PV21	Sub-Indicador Risco de Perda da Produção Agrícola
PV22	Sub-Indicador Risco de Falta de Água para Consumo Humano e Animal
RAI	<i>Rainfall Anomaly Index</i>
RI	<i>Regional Annual Index</i>

SNIS	Sistema Nacional de Informações sobre Saneamento
SUDENE	Superintendência de Desenvolvimento do Nordeste
UNICAMP	Universidade Estadual de Campinas
ZCIT	Zona de Convergência Intertropical

Sumário

1	INTRODUÇÃO	27
1.1	Relevância do Estudo	29
1.2	Hipótese da Tese	31
1.3	Objetivos	31
1.3.1	Geral	31
1.3.2	Específicos	31
1.4	Estrutura da Tese	32
2	A PROBLEMÁTICA DA SECA	34
2.1	Introdução	34
2.2	Características do Semi-árido Nordeste	35
2.3	Aspectos das Secas no Nordeste Brasileiro	47
2.4	Descrição da Área Foco do Estudo	62
2.5	Índices de Severidade de Seca	75
2.5.1	Considerações Iniciais	75
2.5.2	Principais Índices de Severidade de Seca Utilizados no Mundo	76
2.5.3	Principais Índices de Severidade de Seca Utilizados no Brasil, Nordeste e Rio Grande do Norte	82
2.6	Índice FUNCEME de Severidade de Seca (IFS) ou Indicador de Emergência (IMERG)	86
2.6.1	Histórico	86
2.6.2	Metodologia Empregada	90
2.6.3	Características do IMERG	95
3	A CONSTRUÇÃO DE INDICADORES	102
3.1	Introdução	102
3.2	Componentes e Características dos Indicadores	104

3.3	A Construção de Indicadores Compostos - Um Guia para Utilização de Melhores Práticas	107
3.3.1	Ambiente Conceitual	110
3.3.2	Seleção das Variáveis	111
3.3.3	Análise Multivariada	111
3.3.4	Preenchimento de Dados com Falhas	115
3.3.5	Normalização	117
3.3.6	Pesos e Métodos de Agregação	119
3.3.7	Análise de Sensibilidade e Robustez	122
3.3.8	Correlação com Outras Variáveis	125
3.3.9	Decomposição do Indicador	125
3.3.10	Apresentação e Visualização dos Resultados	127
3.4	Métodos de Preenchimento de Dados	127
3.4.1	Padrões e Mecanismos de “não resposta” dos Dados	128
3.4.2	Métodos de Preenchimento Único dos Dados	131
3.4.3	Métodos de Preenchimento Múltiplo dos Dados	133
3.4.4	Aplicativos Computacionais que Implementam o Preenchimento Múltiplo dos Dados	138
3.4.5	MICE - Multiple Imputation by Chained Equations	139
3.5	A Técnica de Winsorização dos Dados	143
4	A CONSTRUÇÃO DO ÍNDICE DE SUSCEPTIBILIDADE AO FENÔMENO DA SECA (ISFS)	144
4.1	Considerações Iniciais	144
4.2	Seleção e Definição das Variáveis e Sub-indicadores	145
4.2.1	Sub-indicador: Natureza Física - PV11	145
4.2.2	Sub-indicador: Natureza Climatológica - PV12	154
4.2.3	Sub-indicador: Natureza Sócio-econômica - Efeito sobre a Base Rural - PV131	163
4.2.4	Sub-indicador: Natureza Sócio-econômica - Efeito sobre a Base Econômica - PV132	164
4.2.5	Sub-indicador: Natureza Sócio-econômica - Condições Municipais de Ate- nuação dos efeitos da Seca - PV133	165
4.2.6	Sub-indicador: Risco de Perda da Produção Agrícola - PV21	166

4.2.7	Sub-indicador: Risco de Falta de Água para Consumo Humano e Animal	
	- PV22	169
4.3	Pesos e Método de Agregação Utilizados no Cálculo do ISFS	173
4.4	O Preenchimento dos Dados com Falhas e a Winsorização no ISFS	173
4.5	Formulação Matemática e Estrutura do ISFS	175
5	VALIDAÇÃO E AJUSTE DO ISFS	181
5.1	Considerações Iniciais	181
5.2	Resultados Preliminares do ISFS - Anos: 1998, 2000, 2005 e 2006.	184
5.3	Preenchimento Múltiplo dos Dados com Falhas e Winsorização	197
5.4	Resultados Finais do ISFS e Discussão - Anos: 1998, 2000, 2005 e 2006.	211
5.5	Análise Multivariada dos Dados	243
5.6	Avaliação de Desempenho - Testes Comparativos	247
6	ANÁLISE DE SENSIBILIDADE E ROBUSTEZ	256
6.1	Considerações Iniciais	256
6.2	Resultados das Simulações	257
7	IMPLEMENTAÇÃO COMPUTACIONAL DO ISFS	294
7.1	Características Gerais	294
7.2	Descrição dos Módulos	296
8	CONCLUSÕES E SUGESTÕES PARA FUTUROS TRABALHOS	317
8.1	Considerações Finais e Conclusões	317
8.2	Sugestões para Futuros Trabalhos	324
	Bibliografia	326
	Apêndices	345
A	Caracterização Pluviométrica	345
B	Mapas da Classificação Pluviométrica	352
C	Tabelas	382
D	Tabelas	387

Capítulo 1

INTRODUÇÃO

O fenômeno da seca é caracterizado pela ausência parcial, total ou pela distribuição irregular das chuvas durante o período chuvoso. É cíclico e característico das regiões semi-áridas espalhadas pelo mundo. Ao contrário de outros desastres naturais, tais como enchentes, terremotos e furações, a seca tem um efeito silencioso, porém não menos grave que esses fenômenos. A explicação está no seu efeito lento, na sua periodicidade, nas medidas políticas assistencialistas que mantêm a população local presa à ajuda governamental e nas migrações dos trabalhadores para áreas de maior fluxo econômico, gerando diversos problemas sociais, como o aumento das favelas nas grandes cidades. Dessa forma, a gravidade de uma seca também está associada às condições sócio-econômico-culturais das populações atingidas, embora seja a ambiental, pela escassez de água, o fator ecológico limitante para plantas, animais e o próprio homem durante esses eventos [56, 58, 70, 78].

No Brasil, a região semi-árida nordestina é a mais afetada pela seca datando do século XVI o primeiro registro histórico de sua ocorrência [43, 56, 78]. Desde então, a seca tem sido registrada repetidamente na região, com efeitos danosos principalmente na agricultura, na pecuária e no suprimento hídrico das populações urbanas e rurais. O principal fenômeno meteorológico associado à ocorrência das secas no semi-árido nordestino é o El Niño-Oscilação Sul (ENOS). Ele tanto causa secas no Nordeste brasileiro e em muitas regiões ao redor do mundo, como inundações em outras, exemplo das regiões Sul e Sudeste do Brasil. Durante a fase positiva do ENOS, a região semi-árida nordestina experimenta uma diminuição sazonal das chuvas e conseqüentemente um aumento de intensidade das secas [7, 10, 34, 109].

Nos últimos anos ocorreram avanços significativos no monitoramento de eventos climáticos no semi-árido nordestino. Exemplo disso é o monitoramento de tempo e clima realizado pelo CPTEC/INPE, com a participação dos Centros de Meteorologia Estaduais do Nordeste, entre

eles a Fundação Cearense de Meteorologia (FUNCEME). Um dos serviços resultante desse monitoramento é a previsão de ocorrência de secas. Ela possibilita o planejamento, na esfera governamental executiva, de ações de mitigação dos efeitos das secas, o que, na maioria das vezes, só ocorre quando a seca já está completamente configurada, a agropecuária em declínio e a população em estado de calamidade, com falta de água até para consumo humano.

Esse acompanhamento também é realizado através dos índices de secas, que são instrumentos que delineam as informações relevantes através da sua agregação, quantificação e simplificação para o entendimento e compreensão do fenômeno. Em resumo, os índices de secas constituem ferramentas essenciais para um diagnóstico mais preciso da sua dimensão nas regiões atingidas, existindo para isso vários modelos de avaliação. Os índices de seca mais simples consideram somente variáveis meteorológicas e/ou hidrológicas. Os mais sofisticados associam a precipitação a parâmetros econômicos, sociais, culturais, ambientais e políticos das regiões atingidas [38]. Essa maior complexidade de informações, inter e intra-relacionadas ou não, resulta em respostas mais adequadas para a mitigação, com estratégias de preparação e a redução de impactos associados a esses eventos.

Em 2000, a FUNCEME propôs, para o Ceará, o Índice Funceme de Severidade à Seca (IFS), que considerava além da climatologia, aspectos da física dos solos, econômicos, sociais, os riscos de perda da safra agrícola e a situação do abastecimento de água urbano e rural dos municípios [64,72,157]. Este índice não foi implementado, possivelmente devido à complexidade da coleta de dados e à difícil operação, estacionando na fase de protótipo.

Como conclusão, observa-se a necessidade do desenvolvimento de índices com dados e informações acessíveis, em um software de fácil manuseio, que possibilitem sua implementação para uma melhor compreensão do fenômeno da seca na região semi-árida nordestina. Eles podem vir a ser uma ferramenta para dar suporte aos tomadores de decisão, tanto na aplicação de recursos, como no direcionamento de ações efetivas para o enfrentamento desses eventos.

Toda essa problemática da seca, sua ocorrência no Brasil, seus desdobramentos no semi-árido nordestino estão no Capítulo dois dessa tese. Outros pontos nele abordados são a origem, tipos e formulação dos índices, com um item especial dedicado ao IFS, desenvolvido pela FUNCEME. Esta referência se dá porque ele serviu de modelo inicial e apresenta contribuição conceitual e informações para a construção do índice proposto nessa tese.

1.1 Relevância do Estudo

Atualmente, a área do semi-árido brasileiro foi redefinida pelo Ministério da Integração Nacional em 62% do território nordestino, apresentando um total de 982.563 km². A população atual é de 22 milhões de habitantes sendo 9,6 milhões na zona rural. Isso implica em uma densidade demográfica de 22,4 hab/km², considerada uma das mais elevadas, quando comparada às de outras regiões semi-áridas do mundo [111]. Dentro de suas peculiaridades climatológicas, o semi-árido nordestino apresenta temperaturas médias elevadas (superiores a 26°C), precipitações pluviométricas anuais entre 300 e 800mm, insolação elevada e altas taxas de evapotranspiração. Apresenta, ainda, solos pobres e rasos, sendo a Caatinga a vegetação predominante e diversas áreas com processo acentuado de desertificação, devido à ação antrópica. Quanto aos recursos hídricos superficiais, a maioria de seus rios é intermitente, mas há grande quantidade de açudes, com capacidade total de armazenamento próxima a 25 bilhões de metros cúbicos, utilizados principalmente para o abastecimento humano e animal. A reserva hídrica subterrânea não é favorecida pela formação geológica, sendo escassa nas áreas dos cristalino e com boa parte salinizada ou em processo de salinização [43,59,147].

A precipitação pluviométrica é determinada pela ação da Zona de Convergência Tropical (ZCIT), o principal sistema atmosférico atuante no semi-árido nordestino, principalmente no período mais chuvoso: os meses de fevereiro a maio. Esse sistema recebe influência direta do sistema ENOS que, durante a sua fase positiva, influencia o posicionamento da ZCIT para uma localização mais ao norte do Equador, diminuindo substancialmente a quantidade de chuvas na região [10,121].

Historicamente, a economia da região sempre esteve associada à disponibilidade de água, o que influenciou no processo de povoamento e ocupação, resultando assim, na atual estrutura fundiária e produtiva [127]. Nessa estrutura, prevalece a desigualdade na distribuição de terras, exemplificada pela grande quantidade de estabelecimentos rurais com menos de 50 ha (75%) e que correspondem a apenas 12% da área total dos estabelecimentos rurais [59]. Essas propriedades, em função das condições climáticas e ambientais da região, não são auto-sustentáveis, podendo ter, nos períodos de secas, perda completa da lavoura e dos rebanhos, levando seus proprietários a se somar à grande massa de flagelados. Por outro lado, nas médias e grandes propriedades, é frequente o uso do trabalho temporário, relação que é facilmente desfeita nos períodos de seca, aumentando a massa de desempregados e flagelados. Nos últimos anos, a economia do semi-árido nordestino apresentou diversas transformações, principalmente tecnológica, tendo se modernizado. No entanto, essas mudanças não foram suficientes para

modificar o quadro social-econômico-cultural das populações mais vulneráveis.

A seca, portanto, constitui-se no principal fenômeno climático que atinge o semi-árido nordestino de forma recorrente e com repercussões negativas na dimensão sócio-econômica-cultural-ambiental da região. Apesar do primeiro registro histórico ser antigo, foi somente no último século que os governos federal, estadual e municipal deram a devida atenção para as repercussões desse fenômeno. No entanto, as ações governamentais de longo prazo para enfrentamento dos efeitos das secas, ao longo do século, deram-se de forma muito heterogênea, marcada por épocas de grandes intensidades de obras e outras de completa escassez. Essa situação foi motivada em grande parte pela falta de comprometimento dos governantes, pela corrupção e desvios de verbas e, pelas constantes mudanças na diretriz das políticas públicas durante as transições de governos. Por outro lado, as ações de curto prazo sempre ocorreram, em função da gravidade do fenômeno e da sua repercussão na mídia. Como exemplos, podem-se citar os programas de frentes de emergência, que embora permeados por obras paliativas e de caráter assistencialista, constituíram-se na principal e, muitas vezes, na única alternativa da população flagelada para sobreviver durante esses períodos [113].

Nas últimas décadas, em função dos avanços tecnológicos nas plataformas orbitais (satélites), dos modelos de mapeamento atmosférico e de previsão climática disponibilizados pelos centros de pesquisas mundiais, ampliou-se o conhecimento e a capacidade de previsão desses eventos. No Brasil, esse avanço foi dado pelo Instituto de Pesquisas Espaciais (INPE) [12,113,146]. Todo esse desenvolvimento permitiu ao poder executivo coordenar ações de forma antecipada para o atendimento das populações atingidas durante períodos de secas.

O monitoramento das secas no semi-árido nordestino, embora ainda em estado inicial, tem contribuído para minimizar os efeitos das secas, disponibilizando produtos para pesquisadores, produtores, gestores públicos e o público em geral.

Em resumo, sabe-se da função, da dimensão e da gravidade das secas no semi-árido nordestino, da grande vulnerabilidade da região a esses eventos e da necessidade crescente de ferramentas que dêem suporte à tomada de decisão. Para contribuir nessa opção de monitoramento surgiu a necessidade do desenvolvimento de um índice de severidade de seca, que, como o IFS, considerasse outras dimensões além da climática, e utilizasse em sua constituição dados mais usuais e disponíveis e com facilidade de implementação.

Como conclusão, foi com base na sistematização, definições e necessidades expostas no Capítulo 2 e no estado da arte na construção de indicadores compostos descrita no Capítulo 3, que se elaborou a proposta deste trabalho de tese, descrita nos Capítulos 4, 5 e 6, para atender ao anseio da criação de um índice de severidade de seca que apresente a possibilidade de ser

utilizado por todos os estados do semi-árido nordestino. Ele se caracteriza como um índice de fácil manuseio em função do aplicativo computacional desenvolvido e descrito no Capítulo 7, permitindo coletas de dados e informações já existentes, e constitui-se em um instrumento de auxílio para a tomada de decisão, contemplando as diferenças econômicas, sociais, de semi-aridez, ambientais e culturais que ocorrem no semi-árido nordestino. O Rio Grande do Norte foi a área inicial escolhida para a sua implementação e validação, por razões de natureza prático funcional.

1.2 Hipótese da Tese

É possível representar a complexidade dos problemas da seca nordestina em um índice de forma simples e transparente, e que forneça subsídios aos gestores públicos na tomada de decisão quando da ocorrência de uma seca na região.

1.3 Objetivos

1.3.1 Geral

Representar a complexidade das questões presentes no fenômeno da seca nordestina de uma forma simples e transparente, através do desenvolvimento de um índice de susceptibilidade à seca com dados mais usuais e disponíveis, e que possibilite aos tomadores de decisão e ao público em geral um quadro mais abrangente da situação dos municípios atingidos por esse fenômeno no semi-árido nordestino.

1.3.2 Específicos

1. Apresentar uma revisão de literatura da problemática da seca nordestina; dos índices de severidade de seca utilizados no mundo, no Brasil e no semi-árido nordestino; e, das principais técnicas e métodos de construção de indicadores compostos;
2. Identificar e selecionar as técnicas e os métodos que melhor se adequem à composição do modelo de índice proposto nesta tese;
3. Desenvolver o modelo do índice, testar e validar com dados reais de três cenários climatológicos: anos secos, normais e chuvosos;
4. Avaliar o intervalo de confiança, a dominância estocástica e a consistência do modelo desenvolvido a partir de uma análise de sensibilidade e robustez;

5. Desenvolver um aplicativo computacional final que vai operacionalizar o modelo de índice proposto, possibilitando seu uso e implementação.

1.4 Estrutura da Tese

Esta tese foi organizada em oito capítulos incluindo a presente introdução mais bibliografia e apêndices.

O Capítulo dois aborda a problemática da seca no Nordeste e no Rio Grande do Norte, área objeto de estudo. É apresentada a caracterização climática da região; o registro histórico de ocorrência das secas; as principais vulnerabilidades econômicas e sociais; e as ações governamentais que norteiam os programas de atendimento às populações atingidas. São também objetos de estudo nesse capítulo os índices de severidade de secas utilizados no mundo, no Brasil, no Nordeste e no Rio Grande do Norte; e o Índice FUNCEME de Severidade de Seca (IFS) ou Indicador de Emergência (IMERG), que inspirou a criação do Índice de Suscetibilidade ao Fenômeno da Seca (ISFS), proposto neste trabalho de tese. O IFS/IMERG é apresentado em detalhes incluindo todo o processo de desenvolvimento, a descrição da metodologia empregada e as características das dimensões e variáveis constitutivas desse índice.

O Capítulo três apresenta a fundamentação metodológica da construção do ISFS, descrevendo as técnicas e métodos utilizados na construção de indicadores compostos. Primeiro, é dada ênfase às características e componentes principais desses indicadores; depois a uma abordagem de construção de indicadores compostos proposta pelo JRC/OECD, baseada na utilização das melhores práticas. Um outro tema incluído com ênfase são os procedimentos e técnicas estatísticas escolhidas para a sua composição.

No Capítulo quatro são abordados os métodos e procedimentos utilizados na construção do ISFS, a partir da seleção e definição das variáveis que compõem os seus sub-indicadores. São apresentados também a formalização matemática do modelo proposto, a sua estrutura e o fluxograma de cálculo. Ademais, são apresentados os resultados de estudos que foram identificados como necessários durante a construção de alguns sub-indicadores, tais como: a classificação da aptidão agrícola dos municípios e a caracterização de períodos secos e chuvosos do Rio Grande do Norte.

No Capítulo cinco são apresentados os procedimentos e convenções utilizados no cálculo e na validação do ISFS, os resultados para os diversos cenários climatológicos testados e uma avaliação de desempenho do ISFS em relação a outros índices de severidade de seca utilizados no semi-árido nordestino.

O Capítulo seis apresenta a análise de sensibilidade e robustez do ISFS, em função de uma abordagem baseada em uma simulação de Monte Carlo. O objetivo dessa análise foi avaliar o intervalo de confiança do modelo desenvolvido, a dominância estocástica da classificação dos municípios e a consistência dos pesos adotados na sua composição.

No Capítulo sete são apresentadas as principais características do software de implementação do ISFS. A ênfase está na descrição dos módulos de entrada de dados, cálculos, relatórios, gráficos e mapas que constituem a apresentação dos resultados desse aplicativo. Esse capítulo também aborda as tecnologias utilizadas para a construção do software, com destaque para as bibliotecas estatísticas e gráficas do R.

No Capítulo oito são apresentadas as conclusões da pesquisa, mediante uma discussão geral e algumas recomendações para futuros trabalhos.

Por último, nos apêndices, são apresentados os seguintes resultados calculados para o Rio Grande do Norte: os gráficos da caracterização de períodos secos e chuvosos pelo método dos quantis; os mapas e a análise pluviométrica da quadra chuvosa e do ano para o período de 1963 a 2006; e os resultados da classificação do ISFS para os anos testados: 1998, 2000, 2005 e 2006.

Capítulo 2

A PROBLEMÁTICA DA SECA

Este capítulo aborda a problemática da seca no Nordeste Brasileiro e no Rio Grande do Norte - área objeto de estudo - a partir da caracterização climática da região, do histórico das ocorrências, das vulnerabilidades econômicas e sociais, da ação governamental para atender as populações atingidas e dos índices de severidade de seca que possibilitam um diagnóstico mais preciso dos efeitos danosos do fenômeno na região. É também objeto desse estudo o Índice Funceme de Severidade de Seca (IFS) ou Indicador de Emergência (IMERG), que inspirou a criação do índice proposto nesse trabalho, o Índice de Susceptibilidade ao Fenômeno da Seca (ISFS).

2.1 Introdução

A seca é um fenômeno climático natural que atinge diversas regiões no mundo, e diferencia-se das demais catástrofes climáticas por ter uma ação lenta e silenciosa, longa duração e atingir de forma cíclica extensas áreas de clima árido e semi-árido [13, 70, 112].

Existem diversas definições para caracterizar um período de seca, e essas definições dependem das diferenças regionais, necessidades e perspectivas disciplinares. O *National Drought Mitigation Center*, nos Estados Unidos, por exemplo, apresenta uma caracterização para a seca baseada em fatores meteorológicos, hidrológicos, agrícolas e sócio-econômicos. A seca meteorológica baseia-se na intensidade e na duração do fenômeno, sendo definida como um período em que a precipitação apresenta-se abaixo da sua média normal. A seca agrícola configura-se quando há deficiência de água no solo afetando o desenvolvimento das culturas agrícolas. A hidrológica quando há deficiência no suprimento de água superficial ou subterrânea, e a sócio-econômica quando afeta a produção de bens de consumo e a economia de uma região [119].

Em países como Indonésia, Bali e Inglaterra um período de ausência de chuvas de vários

dias seguidos, já pode ser caracterizado como seca, devido a frequência em que ocorre chuvas nessas regiões. Na Líbia e particularmente no semi-árido brasileiro, em que ocorre um período de 8 meses secos e 4 meses com chuvas durante o ano, a seca está caracterizada quando há déficit hídrico por mais de um ano [70, 112, 119].

O fenômeno da seca é característico nas regiões semi-áridas, porém o seu efeito depende das condições sócio-econômicas e culturais da população atingida [70]. No Brasil, é a região Nordeste, na sua área semi-árida, a mais afetada por esse fenômeno, datando do século XVI o primeiro registro histórico de sua ocorrência [56, 58, 78].

À exceção do Maranhão, os outros Estados nordestinos apresentam parte do seu território no semi-árido, variando este percentual em relação à área total do Estado ou do total geral da área semi-árida, como será detalhado na caracterização da região. O Rio Grande do Norte, Estado objeto desse estudo, apresenta 93,4% de sua área no semi-árido, sendo portanto, o Estado com maior área semi-árida proporcional. A sua caracterização será apresentada de forma detalhada no item 2.4.

Aspectos históricos do fenômeno das secas nordestinas serão abordados e discutidos a partir dos fenômenos climáticos, da estrutura fundiária e das características agropecuárias e sócio-econômicas da região. Os índices de severidade de seca serão definidos e comparados e o Índice Funceme de Severidade de Seca (IFS) ou Indicador de Emergência (IMERG) destacado e apresentado em detalhes.

2.2 Características do Semi-árido Nordestino

O semi-árido nordestino é caracterizado pela insuficiência e irregularidade de chuvas, com médias pluviométricas anuais que oscilam de um mínimo de 300 a 400 mm e um máximo de 700 a 800 mm, com um período seco com duração de seis a oito meses e um período chuvoso com duração de quatro meses no Sertão e até seis meses no Agreste. A insolação é elevada, com altas temperaturas (mínimas com média de 20°C e máximas com média de 30°C), evapotranspiração potencial elevada (aproximadamente 3.000 mm anuais) e ventos secos [59].

Apresenta também solos pobres e rasos que, associados às características climáticas anteriores, à devastação das matas ciliares e das Caatingas, sua principal vegetação, pela indústria ceramista, pecuária e outras atividades agro-industriais, conduz grandes áreas do semi-árido nordestino à salinização e/ou sodicidade dos solos e à desertificação.

A desertificação, segundo a Convenção das Nações Unidas, é a degradação de terras nas zonas áridas, semi-áridas e sub-úmidas secas, resultante de vários fatores, incluindo as variações

climáticas e as atividades humanas e que vem se agravando pelas mudanças climáticas atuais. No Brasil, segundo essa definição, as áreas susceptíveis a desertificação estão localizadas na região Nordeste e no Norte de Minas Gerais. Os principais núcleos estão localizados nos Estados do Piauí (Gilbués), Ceará (Irauçuba), Rio Grande do Norte (Seridó) e Pernambuco (Cabrobó), embora as áreas do seu entorno, semi-áridas e sub-úmidas secas também sejam susceptíveis (Figura 2.1) [147,148].

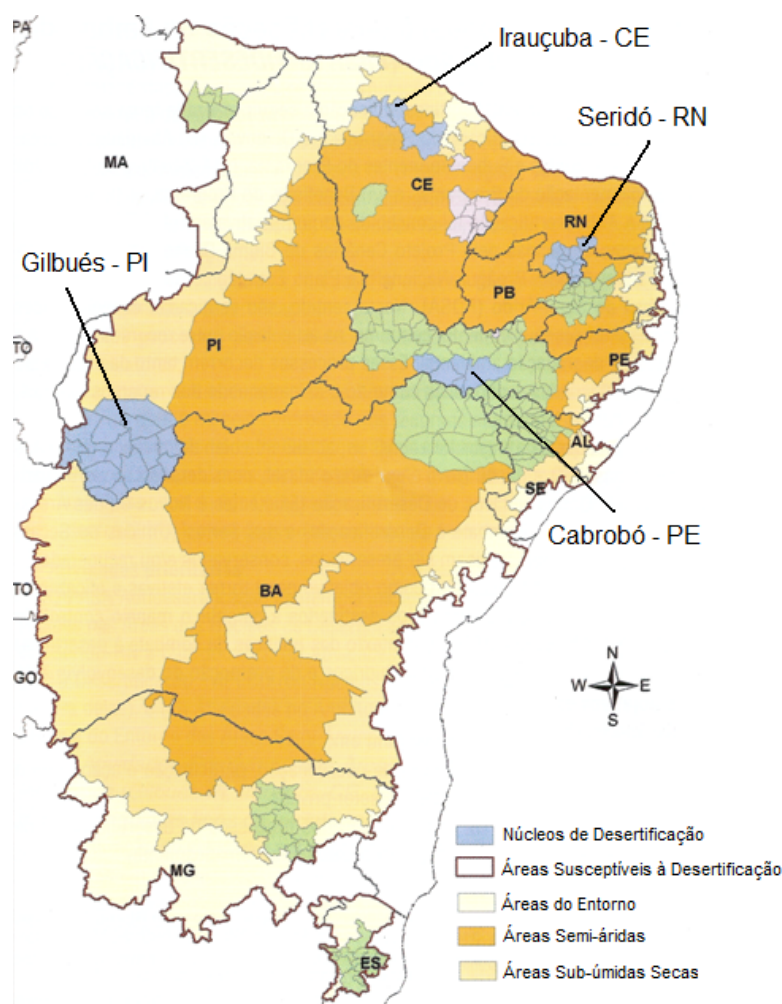


Figura 2.1 Áreas susceptíveis a desertificação no Nordeste. Fonte: Adaptado de Carvalho [44].

Quanto aos recursos hídricos o Nordeste está subdividido em seis regiões hidrográficas, sendo as principais: a região hidrográfica do São Francisco, que abrange 521 municípios nos Estados da Bahia, Minas Gerais, Pernambuco, Alagoas, Sergipe e Goiás, além do Distrito Federal; a região hidrográfica do Parnaíba, que abrange os Estados do Piauí e parte dos Estados do Maranhão e Ceará; e a região hidrográfica do Atlântico Nordeste Oriental que abrange os Estados do Rio Grande do Norte, Paraíba, Ceará e partes dos Estados de Pernambuco e Alagoas [4]. A quantidade atual de açudes existentes no semi-árido nordestino supera os 100 mil, com capacidade de armazenamento de 25 bilhões de metros cúbicos [59].

Quanto à água subterrânea, 70% da área do semi-árido estão no cristalino. Essa formação geológica apresenta solos rasos, alta capacidade de trocas catiônicas, portanto com baixa permeabilidade, capacidade limitada de retenção natural de água, e 60% da reserva hídrica salinizada. As bacias sedimentares, que possibilitam o acúmulo de reserva hídrica subterrânea na região, estão distribuídas de forma dispersa. A maior concentração dessas reservas ocorre no Maranhão e Piauí; e, no Rio Grande do Norte, Ceará e Bahia, em menores proporções [59].

A vegetação predominante na região é a Caatinga, áreas de Cerrados, fragmentos de Mata Atlântica, Biomas Costeiros e Insulares, e Mata de transição e Floresta Equatorial no Meio Norte (Maranhão e parte do Piauí).

Dentre os principais sistemas atmosféricos que atuam e influenciam a precipitação pluviométrica no Nordeste, destacam-se os Sistemas Frontais, os Vórtices Ciclônicos de Ar Superior (VCAS), as Perturbações Ondulatórias dos Alísios (POA), as linhas de instabilidade associadas a brisa marítima, a Oscilação de Madden-Julian (30-60 dias) e a Zona de Convergência Intertropical (ZCIT) [7, 8, 10, 121, 122, 139, 169, 187].

Os Sistemas Frontais, constituídos principalmente pelas frentes frias, atuam principalmente no sul do Nordeste, causando chuvas nos meses de novembro a fevereiro, período em que atuam também em todo o Nordeste, os VCAS, produzindo tanto chuvas intensas como estiagens.

As Perturbações Ondulatórias dos Alísios, constituídas por massas de ar instáveis, profundas e úmidas, são sistemas que provocam chuvas no litoral do Nordeste, atuando desde o Norte do Rio Grande do Norte até a Bahia, no período de maio a agosto [122, 187]. Também no litoral atuam o sistema de brisas marítimas, decorrentes das diferenças térmicas entre as superfícies terrestres e marítimas, ocasionando aumento de precipitação no período noturno [121, 169].

A Oscilação 30-60 dias consiste basicamente da propagação de uma onda atmosférica no sentido zonal ao longo do globo terrestre em torno do equador e representa a mais significativa forma de variabilidade atmosférica nos trópicos em escala intra-sazonal. Esse sistema possui grande influência no padrão de nebulosidade sobre a América do Sul e afeta o regime de precipitação sobre o Nordeste modulando-o [121, 139].

Contudo, a Zona de Convergência Intertropical (ZCIT), constituída de uma pequena zona em torno do Equador, formada pela confluência dos ventos alísios de nordeste e de sudeste, é o principal sistema atmosférico atuante no semi-árido nordestino, e responsável direto pelas chuvas nessa região, no período de fevereiro a maio.

A ZCIT recebe influência dos padrões de Temperatura da Superfície do Mar (TSM). Anomalias positivas de TSM no Atlântico Norte e negativas no Atlântico Sul, próximo à costa africana, configuram um padrão de dipolo de TSM desfavorável às chuvas no Nordeste, influenciando a

ZCIT para sua posição mais ao norte do Equador. Na situação inversa, tem-se um padrão de dipolo de TSM favorável às chuvas no Nordeste, com a ZCIT localizada em uma posição mais ao sul do Equador [10,121]. Durante episódios de El Niño-Oscilação Sul (ENOS), que é o principal fenômeno de acoplamento atmosférico com influências significativas na variabilidade climática de diversas áreas ao redor do mundo (Figura 2.2), há um deslocamento da ZCIT para sua posição mais ao norte, e portanto não favorecendo a formação de chuvas na região [7,10].

O El Niño é um fenômeno oceânico caracterizado pelo aquecimento incomum das águas superficiais nas porções central e leste do oceano Pacífico, próximo à costa do Peru, com repercussões climáticas em várias regiões do mundo [109]. A situação inversa, La Niña, que é o resfriamento atípico das águas do pacífico, também desempenha considerável impacto no clima global [109]. Eventos de La Niña são associados ao aumento das chuvas na região semi-árida nordestina [7,19].

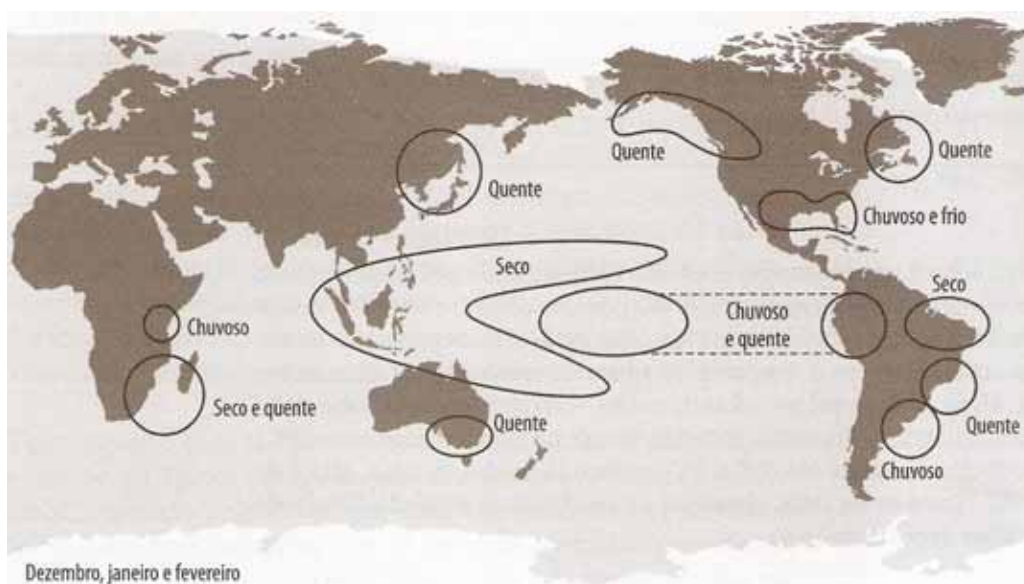


Figura 2.2 Repercursões do fenômeno El Niño ao redor do mundo. Fonte: MENDONÇA & DANNI-OLIVEIRA [109].

Na geoeconomia, o Nordeste como um todo apresenta dois grandes subsistemas: as periferias úmidas e o interior semi-árido. As periferias úmidas constituem a Zona do Litoral, Zona da Mata, e a Zona do Meio-norte, e o interior semi-árido é formado pelo Agreste e pelas Manchas Férteis, e pelas zonas fisiográficas do Sertão e do Cerrado [113].

No Litoral e Zona da Mata, que se estendem do Rio Grande do Norte à Bahia e correspondem a 7% da área total da região, predominam a cana-de-açúcar e o cacau, e, os maiores níveis de concentração populacional e de densidade demográfica. Nessa região, localiza-se também a maior parte das zonas industriais do Nordeste [113].

A Zona do Meio-norte compreende o Maranhão e parte do Piauí, sendo uma região de transição, passando do clima semi-árido, na parte oriental, para a floresta úmida na parte ocidental. Corresponde a 22,4% do território nordestino e nela predominam o cultivo do arroz, o extrativismo, a pecuária, a exploração de minérios e a soja, na parte sul [113].

O Agreste constitui uma zona de transição entre a Zona da Mata e o Sertão, e vai desde o Rio Grande do Norte à Bahia, correspondendo a 10,8% da área nordestina. Apresenta clima úmido ou sub-úmido na parte oriental e semi-árido na parte ocidental, grande diversidade na distribuição de terras e certo grau de especialização agrícola [113].

As Manchas Férteis são microclimas dispersos pelo Agreste e Sertão que apresentam maiores índices de umidade do que as áreas circunvizinhas e melhores índices de produtividade na agricultura. Correspondem a 2,4% da área total nordestina e exercem a função de abrigar os rebanhos da região quando da ocorrência de uma grande seca [113].

O Sertão e o Cerrado, são áreas muito dependentes do ciclo de chuvas onde predominam a agricultura de baixo rendimento, solos pobres e rasos, altas temperaturas, forte insolação, pluviosidade baixa e irregular e a vegetação de Caatinga [113]. Abrange 57% do território nordestino, com uma densidade demográfica relativamente baixa ($20 \text{ hab}/\text{km}^2$), mas que se torna elevada quando comparadas às outras regiões semi-áridas do mundo. Estes últimos constituintes do semi-árido estão no denominado “Polígono das Secas”, que delimitou o semi-árido brasileiro até 2005, englobando áreas parciais de todos os Estados do Nordeste (exceto o Maranhão) e o norte de Minas Gerais, totalizando uma área de 945.000 km^2 [58, 113, 127].

Em março de 2005, através das portarias de N° 1 e N° 89, o Ministério da Integração Nacional instituiu uma nova delimitação do semi-árido brasileiro, abolindo a utilização do Polígono das Secas como instrumento legal de delimitação das áreas do Nordeste sujeitas às secas. Essa atualização aumentou a área do semi-árido brasileiro para $982.563,3 \text{ km}^2$ (Figura 2.3). Ela engloba 1133 municípios que se beneficiam do bônus de adimplência de 25% dos recursos do Fundo Constitucional de Financiamento do Nordeste (FNE) [111]. No Quadro 2.1 é apresentada a distribuição da área e da população da nova região semi-árida por Estado da Federação e os totais para o Nordeste. O maior percentual, em área do semi-árido, está localizado na Bahia (40%), seguido pelo Piauí (15,3%) e Ceará (12,9%), e o menor percentual está localizado em Sergipe (1,1%). No entanto, ao observar a predominância da semi-aridez, o Estado que apresenta o maior percentual da sua área no Semi-árido é o Rio Grande do Norte com 93,4% da sua área total, seguido por Pernambuco (88%), Ceará (86,8%) e Paraíba (86,6%). Em termos populacionais, Bahia, Ceará e Pernambuco concentram os maiores contingentes com 30,9%, 20,2% e 15,5%, respectivamente.

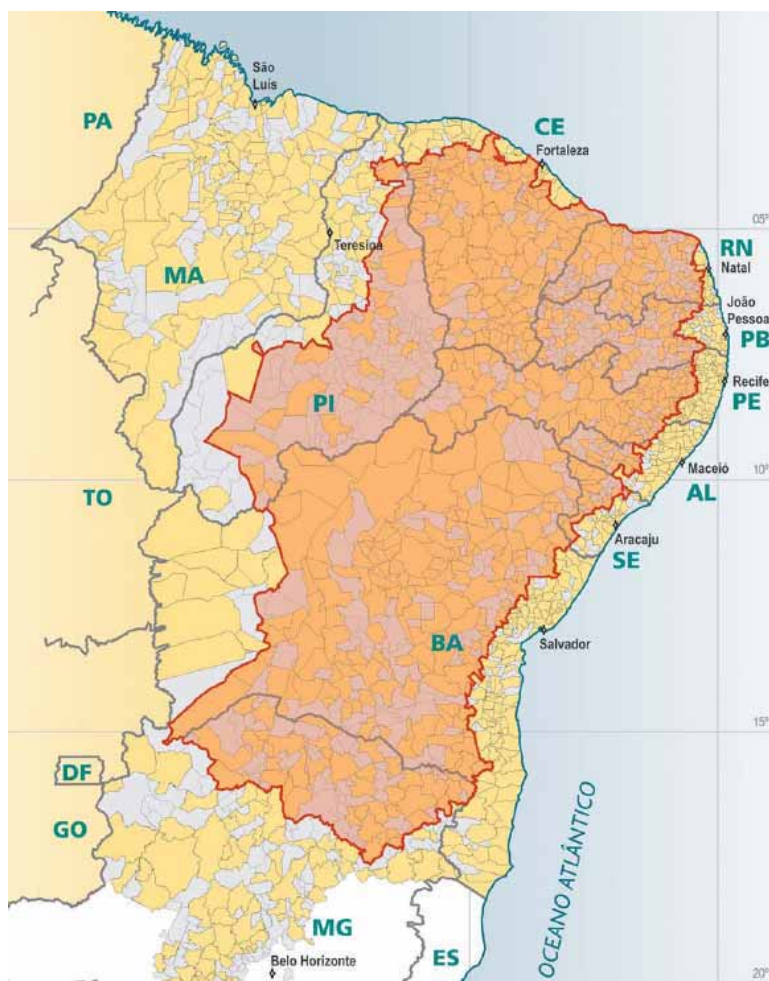


Figura 2.3 Delimitação do semi-árido brasileiro. Fonte: ANA [4].

Estado	Área				População				
	Total (km^2)	RSA ¹ (km^2)	% total RSA	% RSA no Estado	Urbana RSA	Rural RSA	Total RSA	Total	% RSA
PI	251.311,5	150.454,3	15,3	59,9	437.508	531.891	969.399	2.843.278	4,6
CE	145.711,8	126.514,9	12,9	86,8	2.451.214	1.760.078	4.211.292	7.430.661	20,2
RN	53.077,1	49.589,9	5,0	93,4	1.061.296	539.874	1.601.170	2.776.782	7,7
PB	56.340,9	48.785,3	5,0	86,6	1.232.095	734.618	1.966.713	3.443.825	9,4
PE	98.526,6	86.710,4	8,8	88,0	1.896.082	1.340.659	3.236.659	7.918.344	15,5
AL	27.818,5	12.686,9	1,3	45,6	424.132	391.172	838.740	2.822.621	4,0
SE	21.962,1	11.175,6	1,1	50,9	208.908	185.310	396.399	1.784.475	1,9
BA	564.273,0	393.056,1	40,0	69,7	3.398.156	3.055.127	6.453.283	13.070.250	30,9
MG ³	200.221,0	103.590,0	10,5	51,7	637.990	546.537	1.184.527	2.773.232	5,7
NE	1.419.242,5	878.973,3	89,5	61,9	11.747.381	9.085.266	20.858.264	44.858.264	46,5

¹Região Semi-árida, ²Fonte: IBGE - Censo Populacional 2000, ³Área de atuação da ADENE em MG.

Quadro 2.1 Distribuição da área e da população do novo semi-árido por Estado da Federação. Fonte: MI [111].

A economia da região semi-árida nordestina está historicamente associada a disponibilidade de água, fator preponderante na orientação do processo de povoamento e ocupação da região, o que resultou na atual estrutura fundiária e produtiva [127].

A pobreza em que vive a maioria da população do semi-árido decorre de fatores que vão além dos condicionantes climáticos e é justificada em grande parte pela posse e uso das terras. Em 1992, os estabelecimentos rurais com menos de 50 hectares correspondiam a 75% do número total, ocupando somente 12% da área total. Em contrapartida os imóveis com 200 hectares ou mais, representavam 7% do número total e ocupavam 68,6% da área total (Quadro 2.2). Observa-se ainda, quanto à utilização das terras, que a intensidade era maior quanto menor fosse a área do imóvel; dessa forma apresentando uma razão área explorada por área aproveitável de 66,3 para propriedades com área inferior a 10 hectares e de 45,2 para propriedades de 500 hectares ou mais (INCRA, 1992 *apud* DUARTE [59]).

No Nordeste, essa situação de desigualdade na distribuição de terras é também comprovada pelo índice Gini de distribuição de terras, que apresentava um patamar de 0,79, em 1992, portanto, superior ao das regiões Sul e Sudeste com 0,69 e 0,74, respectivamente. Nesse mesmo ano, a instabilidade ocupacional do trabalho assalariado temporário, predominante nas propriedades com áreas de 200 hectares ou mais, contabilizava 2,5 milhões de trabalhadores rurais. Esse é um

fator preponderante e corrobora a situação de pobreza e vulnerabilidade em que se encontram essas populações. Essas relações de trabalho têm um papel importante em períodos de seca, já que são facilmente desfeitas pelos proprietários, aumentando rapidamente o contingente de desempregados, tornando-se essa a massa trabalhista mais vulnerável a seca [59].

Classe de Área (ha)	Quantidade de Imóveis		Área Total (ha)	
	Valor Absoluto	%	Valor Absoluto	%
0,1 a menos de 10	661.784	38,3	2.695.880,9	1,8
10 a menos de 50	632.054	36,6	15.230.842,2	10,2
50 a menos de 100	195.635	11,3	13.286.705,5	8,9
100 a menos de 200	117.567	6,8	15.733.404,6	10,5
200 a menos de 500	76.461	4,4	23.041.113,8	15,4
500 e mais	43.462	2,6	79.490.709,2	53,2
Total	1.726.963	100,0	149.478.656,2	100,0

Quadro 2.2 Número de imóveis rurais e área agrícola. Fonte: INCRA, 1992 *apud* DUARTE [59].

Essa problemática das propriedades rurais, da agropecuária e das susceptibilidades sócio-econômicas, faz o Nordeste brasileiro ser uma região de contrastes, pois, se por um lado, sua economia exibe um certo porte, com PIB de 161,2 bilhões de dólares [88], semelhante ao PIB de países como Israel, Colômbia, Chile e Romênia, e mais expressivo que o PIB de países como Nova Zelândia, Egito, e Paquistão; apresenta índices comparados a países pobres, como por exemplo, o PIB per capita regional [6]. A densidade econômica regional é de US\$ 103,5 mil/km², o que corresponde a 67% da brasileira (US\$ 154,3 mil/km²) e apenas 13% da atingida pelo Sudeste (US\$ 752,5 mil/km²). O grau de industrialização da economia da região ainda é baixo, 28% em 1998, quando comparado com o do Brasil e o do Sudeste, 37,1% e 40,4% em 1998, respectivamente.

Quanto aos indicadores sociais, o Nordeste apresenta em geral os piores índices, em relação as outras regiões do país, à exceção da região Norte. Esse é o caso do Índice de Inclusão Educacional (Figura 2.4 A), que contabiliza o ensino médio e superior, e onde o Nordeste em conjunto com a região Norte e o Centro-Oeste, apresentam os piores resultados do País. Os Estados do Rio Grande do Norte (0,13), Pernambuco (0,1) e Ceará (0,1) apresentam uma situação mais favorável que os outros Estados da região Nordeste, mas, mesmo assim, muito inferiores ao Paraná (0,36), Santa Catarina (0,37), Rio Grande do Sul (0,38), Distrito Federal

(0,4) e São Paulo (0,42), que apresentaram os melhores índices no País, e também ao índice de países como Argentina (0,4 à 0,7) e Chile ($> 0,7$).

O Índice de Inclusão em Saúde (Figura 2.4 B), composto pela oferta de leitos e pela quantidade de médicos e enfermeiros por mil habitantes, também apresenta o Nordeste com o segundo pior índice em relação a média nacional e do Sudeste que apresenta uma situação acima da média nacional. A Bahia, o Piauí e o Maranhão são os Estados do Nordeste que aparecem nas piores colocações. Nesse indicador, a pior situação é a da Região Norte.

Na Figura 2.4 C é apresentada a participação de pobres na população brasileira. Foi considerado, na constituição desse índice, o critério de “pobreza relativa” que são todas as pessoas que, embora tenham superado as necessidades básicas, vivem com um nível de renda inferior a renda média ou mediana do país. Dessa forma, esse índice é composto por todas as pessoas que vivem em famílias com renda inferior a 50% da renda média nacional familiar *per capita* [134]. Nesse mapa percebe-se que o Nordeste concentra o maior percentual de pobres do país, em conjunto com o Estado de Tocantins. Esse aspecto aparece de forma mais proeminente no semi-árido, e é um dos determinantes da grande vulnerabilidade da região ao fenômeno da seca, seja pelo longo período de estiagem anual, seja quando é agravado pela irregularidade hídrica ou estiagem no período chuvoso.

Na Figura 2.4 D é apresentado o déficit de trabalho decente no país. Nesse indicador são computados o desemprego aberto (pessoas desocupadas e procurando trabalho), pessoas não-remuneradas, pessoas subocupados com renda inferior a meio salário mínimo e pessoas subocupados com jornada de trabalho inferior a 14 horas semanais [134]. Nota-se claramente que o Nordeste em conjunto com os Estados do Pará e Tocantins, concentram os maiores índices de déficit de trabalho decente no País [58,59,113].

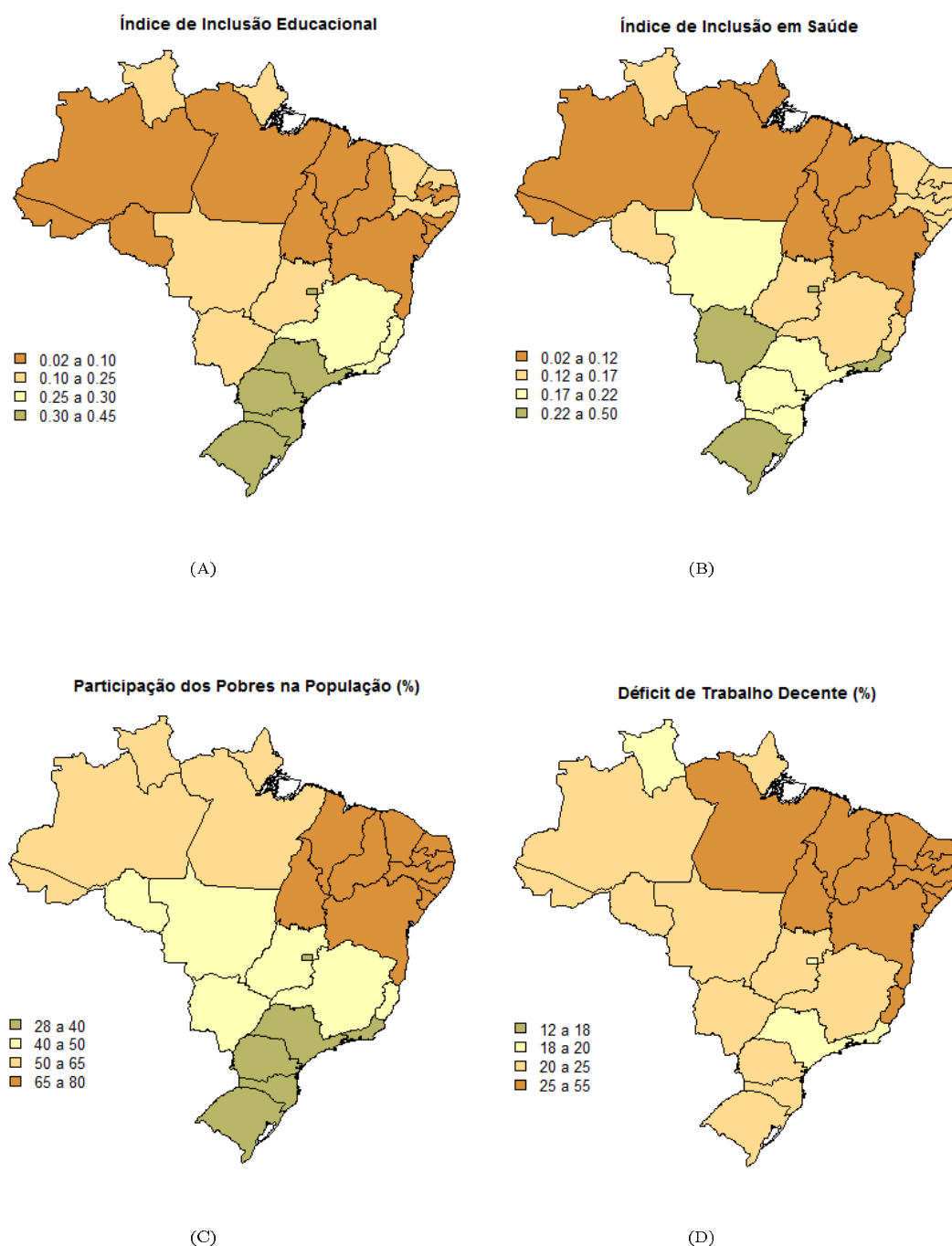


Figura 2.4 Indicadores Sociais, Brasil. Fonte: POCHMANN et al. [134]

Na Figura 2.5 A tem-se o índice de inclusão previdenciária, que considera a cobertura previdenciária nos segmentos etários de 16 a 59, e, 60 anos ou mais. Esse indicador mostra que um pouco mais da metade do Nordeste apresenta valores na mesma categoria que os Estados do Sudeste e mostra-se interessante devido ao fato das aposentadorias constituírem uma parte considerável da renda das famílias do semi-árido nordestino, o que significa muito nos períodos de seca, e é corroborado pelo quadro social apresentado na Figura 2.5 B, que caracteriza o percentual de pessoas com mais de 50% de sua renda provenientes de transferências

governamentais. Nota-se claramente para este indicador, que a ação do executivo se dá de forma mais sistemática na região semi-árida nordestina. Essa ação, por seu alcance limitado e caráter não estrutural, não diminuiu até hoje a grande vulnerabilidade das populações atingidas pelo fenômeno da seca [58,59,113].

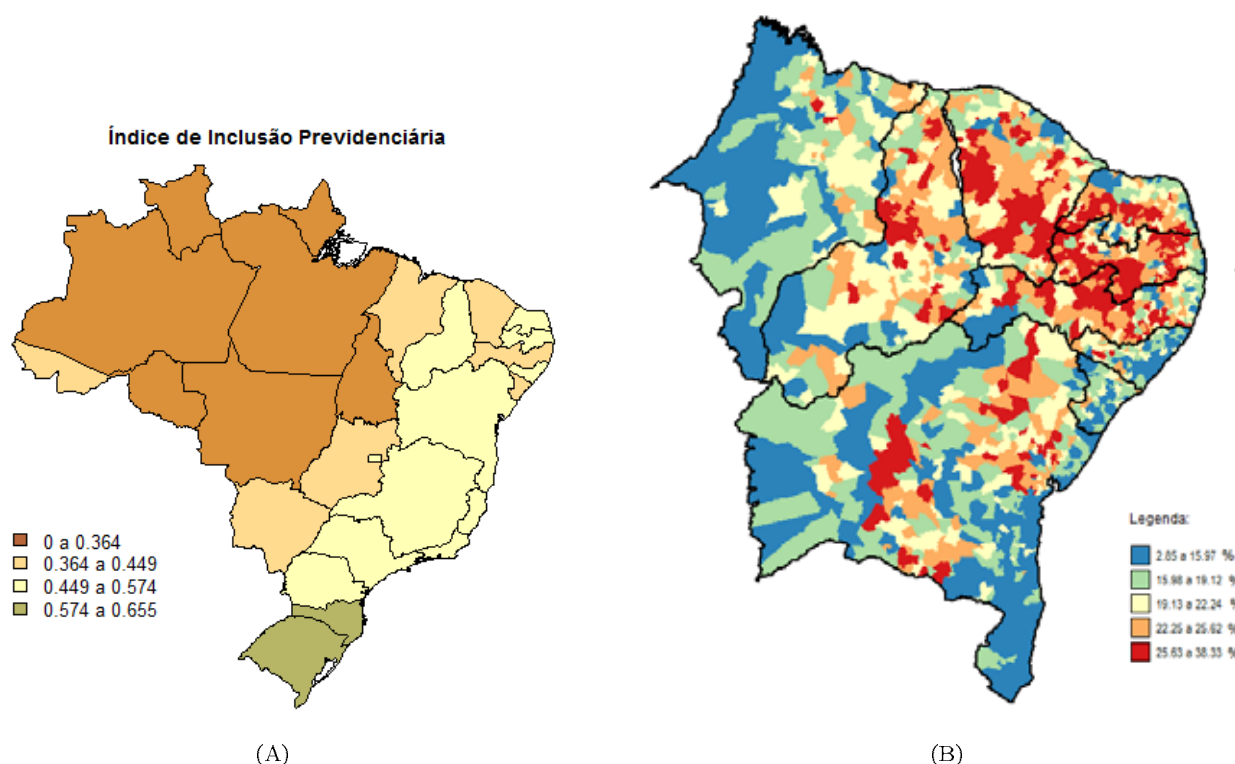


Figura 2.5 Indicador de Inclusão Previdenciária (Brasil) e de Transferências Governamentais (NE). Fonte: POCHMANN et al. [134] e PNUD [133].

Com relação ao Índice de Desenvolvimento Humano (IDH), embora o Nordeste tenha seguido a tendência de evolução positiva do Brasil desde o início dos anos 90, a região apresenta os piores índices em relação ao restante do País (Figura 2.6). Esses índices se assemelham aos de Países como Honduras, Bolívia, Vietnã, Cabo Verde, Guatemala, Nicarágua e Congo. Dentre as dimensões do IDH (renda, educação e longevidade) a educação foi o que mais cresceu em todos os Estados Brasileiros, sendo que o Rio Grande do Norte figura com o terceiro maior crescimento do Nordeste (Figura 2.7) [55], mesmo assim apresenta números comparados a Honduras, Vietnã, Cabo Verde e Namíbia.

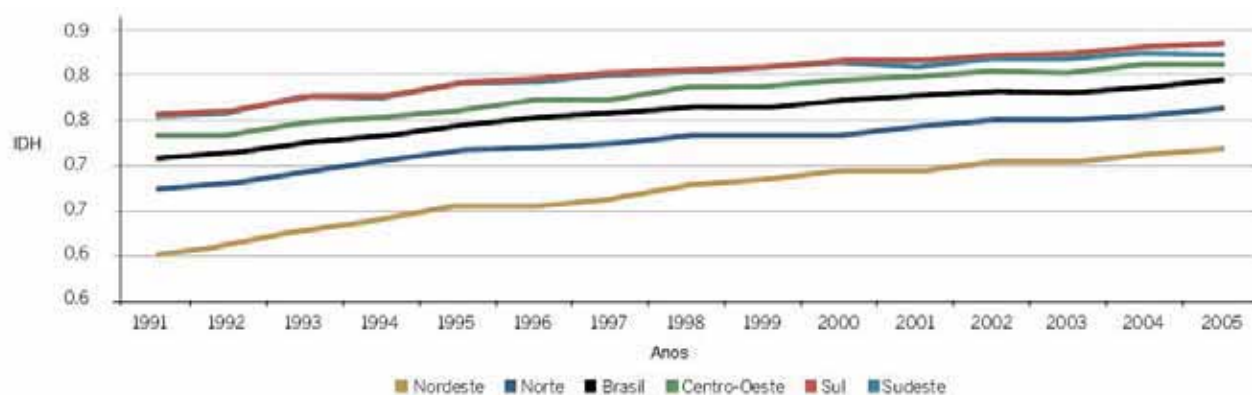


Figura 2.6 Evolução de IDH por Regiões, Brasil, Período: 1991 a 2005. Fonte: CEPAL/PNUD/OIT [55].

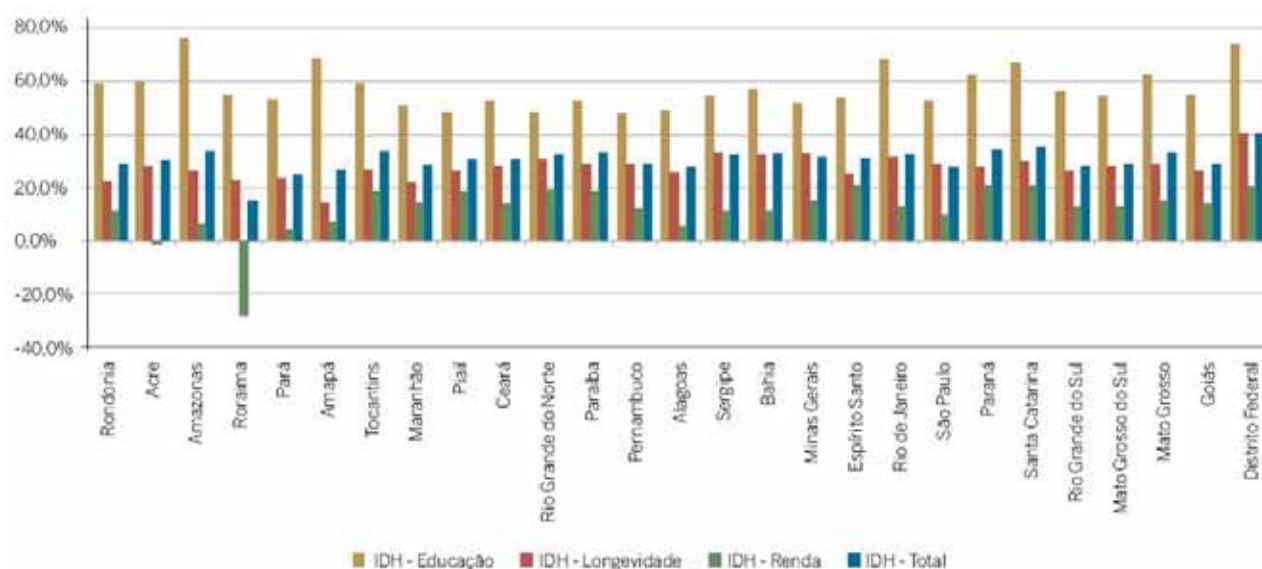


Figura 2.7 Variação de IDH e de seus componentes entre 1991 e 2005, por Estados, Brasil. Fonte: CEPAL/PNUD/OIT [55].

Em uma avaliação da distribuição do Índice de Desenvolvimento Humano Municipal (IDH-M), Maranhão, Piauí, Alagoas e Paraíba concentram a maior quantidade de municípios que apresentam IDH-M mais baixo. Em contraste, o Rio Grande do Norte e o Ceará apresentam IDH-M mais elevado na maioria de seus municípios (Figura 2.8). Observa-se também que nas regiões próximas aos grandes centros urbanos (capitais) os municípios apresentam IDH-M mais elevado. É interessante destacar a dualidade do Semi-árido Nordestino com relação a este indicador, pelo fato de regiões com situação climática crítica terem tanto IDH-M médio como IDH-M baixo. Tal fato deve-se, em grande parte, às diferenças existentes entre as prioridades de ação governamental nas esferas municipal e estadual a medida que os governos se sucedem [133].

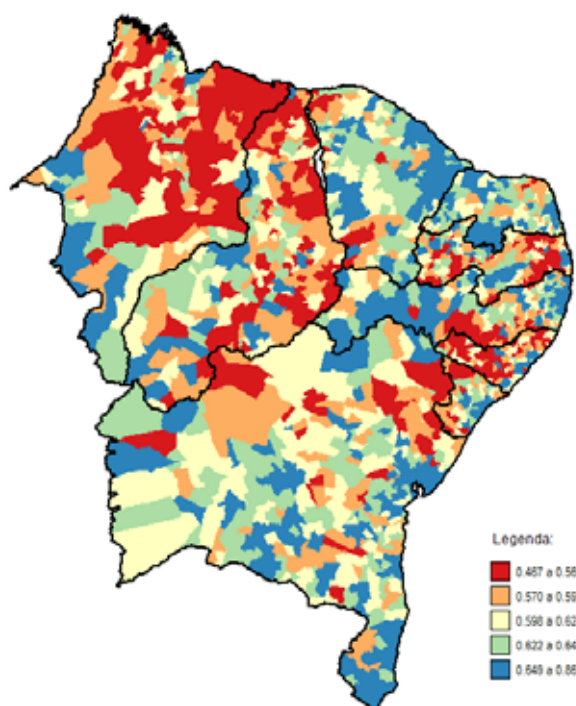


Figura 2.8 Índice de Desenvolvimento Humano Municipal dos Municípios do Nordeste. Fonte: PNUD [133].

2.3 Aspectos das Secas no Nordeste Brasileiro

O Nordeste é uma região de grande heterogeneidade em termos agroecológicos e econômicos. A diversidade edáfica, a estrutura fundiária e as relações de trabalho que imperam na região são fatores que contribuem mais decisivamente para o agravamento dos períodos de secas [58,113].

As condições edafo-climáticas do semi-árido nordestino não são uniformes, e as características intrínsecas da região não apresentam similaridade quando comparadas a outras regiões semi-áridas no mundo (DUQUE, 1993 apud DUARTE [59]). Porém, é possível apontar algumas características dominantes, embora variando de região para região, como a ocorrência da seca.

O relato histórico mais remoto de ocorrência de secas no Nordeste data do Século XVI na Bahia e em Pernambuco, pelo padre jesuíta Serafim Leite na obra: “História da Companhia de Jesus no Brasil” [56,78]. No Quadro 2.3 é apresentada uma panorâmica das secas ocorridas no Nordeste, segundo referência de vários autores.

A seca ocorrida nos anos de 1877-79 merece destaque especial, pois, segundo os relatos históricos teriam perecido aproximadamente quinhentos mil Nordestinos, vitimados pela fome, sede, epidemias, falta de condições sanitárias, e ausência de infraestrutura e de intervenção governamental para assistir as populações atingidas [43,78]. Existe uma certa divergência com relação a esse número, porém, é fato inegável que o Nordeste nunca tinha anteriormente sofrido

uma catástrofe com essa dimensão [43]. Foi essa grande seca que motivou o Governo Imperial Brasileiro, pela primeira vez, a adotar medidas de longo prazo para combater os efeitos das secas no Nordeste [78].

Século	Ano ou Período	Área de Ocorrência	Bibliografia
XVI	1559, 1564, 1590 e 1592	Bahia, Pernambuco	Serafim Leite, 1945
XVII	1600, 1614, 1691 e 1692	Nordeste	Dantas, 1961 [56]
	1603, 1614, 1645 e 1692	Nordeste	Guerra, 1981 [78]
XVIII	1710-11, 1723-27, 1736-37, 1744-46, 1766, 1777-78, 1782, 1784, 1790-93	Nordeste	Dantas, 1961 [56]
	1711, 1721-25, 1736-37, 1745-46, 1754, 1760, 1766, 1772, 1777-78, 1791-93	Nordeste	Guerra, 1981 [78]
XIX	1804, 1810, 1824-25, 1844-45, 1877-79, 1888-89.	Nordeste	Guerra, 1981 [78]
	1808-09, 1814, 1816-17, 1824-26, 1833, 1835, 1837, 1844-46, 1860, 1868-69, 1877-79, 1885, 1888-89, 1891-92, 1898	Nordeste	Dantas, 1961 [56]
XX	1900, 1903-04, 1908	Rio Grande do Norte	Guerra, 1981 [78]
	1915, 1919, 1930-32, 1942, 1953, 1958, 1970, 1976, 1979-83, 1993, 1998	Nordeste	Dantas, 1961 [56] Guerra, 1981 [78] Carvalho, 1988 [43] Cavalcante, 2002 [46] Duarte, 2002 [58] Duarte, 2002 (b) [59] Moreira Filho, 2002 [113] Pessoa, 2002 [127]

Quadro 2.3 Registro histórico da ocorrência de secas no Nordeste.

Neste início de século XXI, apesar da ocorrência de anos com déficit de precipitação pluviométrica, como nos anos de 2001 e 2005, não se verificou a ocorrência de uma grande seca no Nordeste.

Os efeitos danosos das secas mostram-se bastantes evidentes sobre a agropecuária e são mais proeminentes e significativos sobre a atividade agrícola, principalmente nas culturas de subsistência, que concentram o maior percentual de ocupação no campo, levando rapidamente

ao desemprego centenas de milhares de trabalhadores e de pequenos produtores rurais. As culturas econômicas, tais como a fruticultura irrigada, por constituírem uma atividade com maior grau de tecnologia agregada, apresentam uma menor vulnerabilidade a seca. Embora o algodão, que durante muito tempo foi a principal cultura econômica do semi-árido nordestino, tenha sido praticamente dizimado, pela praga do bicudo, em meados da década de 80.

A pecuária apresenta uma menor vulnerabilidade em relação a atividade agrícola, devido a expansão da oferta de água no semi-árido, depois de anos de obras de infraestrutura na construção de açudes e adutoras, e da facilidade de transporte de água e alimentos. Entre as várias estratégias utilizadas pelos pecuaristas para enfrentar períodos de secas estão: a possibilidade de transferência dos rebanhos para as áreas de serras, vales úmidos ou para outras regiões do País, e o sacrifício da lavoura para garantir o alimento e a sobrevivência dos rebanhos. Mesmo assim, o excesso de animais nas propriedades, sem a condição necessária para sustentá-los, mesmo no período seco dos anos normais, e a inexistência do hábito de acumular reservas de água e ração para esse período, tornam a pecuária nordestina vulnerável aos efeitos das estiagens. Adicionem-se a isso os prejuízos diretos decorrentes da transferência dos animais, que já se encontram em condições precárias e o atraso mínimo de dois anos para a recuperação do rebanho, decorrente da perda de crias, atrofia dos animais sobreviventes e morte de matrizes [46, 58, 59].

As secas que ocorrem no semi-árido nordestino não atingem de forma homogênea toda a população e sua intensidade varia de acordo com a classe social a que pertence o grupo de pessoas atingidas. DUARTE [58] cita três grupos principais, que constituem as camadas sociais do semi-árido nordestino, caracterizando o nível de intensidade dos efeitos das secas sobre esses grupos. O primeiro grupo é constituído pelas pessoas que tem renda fixa, como os funcionários públicos, empregados do comércio, serviços e transportes. Estas pessoas recebem a influência das secas de modo indireto através da elevação dos preços de alguns produtos agrícolas. O segundo grupo é constituído pelos médios e grandes proprietários rurais, que recebem um impacto mais direto, pela desvalorização das propriedades e redução do volume de negócios. Porém, sendo a classe mais abastada, consegue enfrentar melhor a fase crítica. E o terceiro grupo, o mais numeroso e mais vulnerável, é constituído pelos pequenos proprietários, meeiros, parceiros, arrendatários, agregados e trabalhadores avulsos. Esses não tem reserva financeira e sofrem os efeitos diretos da falta de chuvas, através da perda da lavoura, morte dos poucos animais que possuem e da perda dos seus empregos. Na maioria das vezes chega à situação de indigência, que só é sanada pela intervenção governamental através dos programas de auxílio às populações atingidas, como as frentes de emergência.

A intervenção governamental em ações de enfrentamento dos efeitos das secas iniciou-se a partir da seca de 1877-79. Essas ações tornaram-se mais sistemáticas no início do século XX, condicionadas à intensidade do fenômeno e à presença de representantes da região na esfera do Governo Federal. Exemplos dessas atuações ocorreram nas secas de 1919 e 1931-32, no governo do Presidente Epitácio Pessoa, e com José Américo de Almeida, à frente do Ministério de Obras Públicas, ambos paraibanos [127]. Foi sob a influência de Epitácio Pessoa que foi aprovada, em 25 de dezembro de 1919, a Lei 3.965, que autorizava a construção de obras necessárias à irrigação de terras cultiváveis no Nordeste Brasileiro. Essa lei também continha normas para a criação de um caixa especial para financiar essas obras, previa o loteamento de terras para a venda ou arrendamento, a assistência técnica, a revenda de implementos e fertilizantes, e, a comercialização dos produtos [78]. Posteriormente, José Américo de Almeida, apoiado nas linhas de ação política de Epitácio Pessoa, incentivou a irrigação como elemento essencial para a transformação da agricultura do semi-árido, partindo da acumulação de água em grandes açudes para abastecimento nas épocas de estiagem e mananciais preventivos à seca. Essas medidas caracterizaram, na época, o padrão da política anti-seca de longo prazo, haja vista que as medidas de curto prazo sempre ocorreram, independentemente do governante, em decorrência do apelo social da dimensão catastrófica das secas nordestinas [43].

De uma maneira geral, pode caracterizar-se as políticas governamentais de ações anti-seca de longo prazo em 3 fases distintas: 1^a) Fase das soluções e obras de engenharia; 2^a) Fase da perspectiva sócio-econômica; e 3^a) Fase do desenvolvimento planejado [127].

A **primeira fase**, estendeu-se de 1877-79 até 1945. Foi caracterizada pela tônica da açudagem e de obras de infra-estrutura, e pela criação, em 1909, da Inspetoria de Obras contra a Seca (IOCS), que ao longo do tempo foi mudando de nome. Em 1919, esse órgão passou a se chamar de Inspetoria Federal de Obras Contra a Seca (IFOCS), e em 1945, Departamento Nacional de Obras Contra a Seca (DNOCS), designação que perdura até hoje [127].

Esse órgão foi criado, inicialmente, com o objetivo de centralizar e unificar a direção dos serviços em toda a área assolada pelas secas e executar um plano de combate aos efeitos das irregularidades climáticas na região. Nesse período foram realizadas algumas obras de vulto, como a construção de grandes, médios e pequenos açudes e de estradas, ainda que a dinâmica de execução dessas obras não tenha se dado de forma contínua, pelas constantes interrupções nas mudanças de governo [127].

Nessa fase inicial até 1945, o DNOCS construiu 50,2% dos açudes públicos e 53,2% dos açudes particulares existentes no Nordeste (Quadro 2.4). Embora esses números representem somente o quantitativo e não a reserva hídrica construída, constata-se que o Ceará concentrou

o maior proporção dos açudes construídos, tanto públicos como particulares, principalmente no período de 1909 a 1950. Em segundo lugar, figura o Rio Grande do Norte, com maior quantitativo em açudes públicos [127].

A ênfase na difusão dos açudes pelo semi-árido era fundamentada na esperança de que as barragens serviriam para fortalecer as práticas agropecuárias e reduzir as causas que levavam as populações pobres à situação de indigência na época das secas, e de que bastava armazenar a água que a irrigação se daria naturalmente. Não se levaram em consideração aspectos importantes como as altas perdas por evaporação e a forma correta de utilização desses recursos hídricos para irrigação, evitando assim a salinização dos solos.

Apesar desses aspectos, é fato que a construção desses açudes aliviou a demanda hídrica da região e os efeitos danosos da falta d'água tanto para a população humana como para os rebanhos. Associado a isso, a infraestrutura de transporte, construída também nessa fase, facilitou tanto o abastecimento de víveres como o movimento migratório de pessoas e animais. Por fim, na Constituição de 1934, foi estabelecida a dotação de 4% da receita tributária anual da união para o financiamento das atividades do DNOCS, o que fortaleceu a atuação do órgão nas diversas ações desenvolvidas [43,78,127].

Estados	Número de Açudes											
	Públicos						Em cooperação					
	1909-50	1951-58	1959-63	1964-78	1979-81	Total	1909-50	1951-58	1959-63	1964-78	1979-81	Total
PI	9	1	-	2	-	12	-	-	-	-	-	-
CE	44	6	3	6	3	62	256	81	86	20	-	443
RN	31	7	2	9	1	50	27	28	7	-	-	62
PB	21	10	2	5	-	38	26	19	12	2	-	59
PE	10	7	7	8	-	32	-	11	-	-	-	11
AL	2	12	1	9	-	24	-	-	-	-	-	-
SE	2	5	1	3	-	11	1	-	-	-	-	1
BA	14	8	-	8	2	32	7	10	3	-	-	20
MG ¹	-	-	4	-	-	4	-	-	-	-	-	-
NE	133	56	20	50	6	265	317	149	108	22	-	596

¹ Área mineira do Polígono das Secas.

Na **segunda fase**, iniciada por volta de 1945, e classificada como de aproveitamento dos recursos d'água, foram tomadas medidas importantes para a região, como a criação da Companhia Hidrelétrica do São Francisco (Chesf), em 1945, e a criação da Comissão do Vale do São Francisco (CVSF), em 1948. Comparativamente, a CVSF tinha um escopo de atividades maior que o do DNOCS, para a área de sua atuação, desde a regularização do curso do rio até iniciativas de crédito agrícola e empreendimentos nas áreas de educação e saúde. No entanto a sua atuação possibilitou que se desenvolvessem clientelas políticas a sua sombra, o que diminuiu a sua eficiência na região. A Chesf, criada para atender às necessidades energéticas do Nordeste inteiro, sobretudo dos grandes centros urbanos da zona costeira, manteve um nível eficiente de desempenho [127].

O modelo analítico dessa fase reconhecia a frustração da perspectiva das obras da fase anterior, face às dificuldades encontradas na utilização das águas mobilizáveis do Nordeste e centrava na utilização dessas águas a base da política de longo prazo [127]. Em 1952 foi criado o Banco do Nordeste do Brasil (BNB), com a finalidade de viabilizar o aproveitamento da infraestrutura social instalada na região, depois de anos de obras de engenharia. A criação do BNB implantava uma nova diretriz na região através da perspectiva de financiamento às atividades privadas, tanto primárias como secundárias, alicerçadas nos princípios da economia moderna [43, 78, 127]. Essa fase, que foi considerada uma fase de transição, termina em 1958 quando o governo, pressionado com uma nova e grande seca, lança mão de modernas técnicas de planejamento regional e dá início a uma fase mais abrangente de soluções para o fortalecimento do Nordeste [127].

A **terceira fase**, caracterizada como de desenvolvimento planejado, tem início em 1959, com a criação da Superintendência de Desenvolvimento do Nordeste (SUDENE). Na década de 1950 foram elaborados dois dos mais importantes documentos sobre a economia do Nordeste [58]. O primeiro foi um estudo caracterizando a economia nordestina, feito pelo economista Hans H. Singer e apresentado em 1953 ao Banco Nacional do Desenvolvimento (BNDES); o segundo, foi um documento feito pelo Grupo de Trabalho para o Desenvolvimento do Nordeste (GTDN), criado em 1956, pelo então presidente Juscelino Kubitschek. Esses documentos comprovaram o atraso relativo do Nordeste em relação ao Sul e o Sudeste do País e muitas coincidências na abordagem do problema a que estava sujeita a economia do semi-árido nordestino. Dentre os problemas citados nesses dois documentos, comparativamente à região Sudeste, estavam os menores quocientes no Nordeste, entre população, terras aráveis e capital, e a evasão de renda: fruto do comércio triangular da região com o exterior - para onde exportava a preços internacionais, e o Centro-Sul do País - de onde importava bens a preços protegidos [127].

Esses documentos também apresentavam sugestões para atenuar os efeitos danosos das secas: o deslocamento da fronteira agrícola do Nordeste para áreas menos afetadas pelas secas, particularmente para o Maranhão; a migração populacional; a irrigação; o cultivo de lavouras xerófilas; o melhoramento genético das plantas a fim de torná-las mais resistentes as estiagens prolongadas; e, a necessidade de industrialização, partindo-se das matérias-primas existentes na região [43,46,58,127]. A própria criação da SUDENE e a sua atuação no planejamento e coordenação de ações no semi-árido nordestino, previstas em lei, foram inspiradas nos estudos do GTDN.

O que se observou na prática foi que a função coordenadora sobre a atuação do setor público no Nordeste, que cabia à SUDENE, fragmentou-se na fase operativa, isso devido à resistência de órgãos públicos e privados a essa ação coordenadora, e a deficiência do aparelhamento dos órgãos públicos e das agências de desenvolvimento regional [59,127]. Esse enfraquecimento da SUDENE fez com que a instituição tivesse o seu papel reduzido praticamente à administração dos incentivos fiscais concedidos à indústria e à pecuária nordestina e culminou com a sua extinção em 2001. Em 2007, a SUDENE foi novamente recriada porém com orçamento e ação limitados.

Essa terceira fase também marcou uma mudança na orientação das políticas públicas, deixava-se de lado a concepção de que o problema do subdesenvolvimento do Nordeste tinha causas climáticas e admitia-se que o problema não era a seca, mas a falta de adaptação da economia e da exploração da terra às condições climáticas. Assim sendo, a partir de 1970 o governo tomou importantes decisões sobre os rumos da economia nordestina. Dentre os principais programas criados a partir desta data pode-se citar: o Programa de Desenvolvimento do Nordeste - POLONORDESTE, em 1974; o Programa Especial de Apoio ao Desenvolvimento da Região Semi-árida do Nordeste - PROJETO SERTANEJO, em 1976; o Programa de Aproveitamento de Recursos Hídricos do Nordeste - PROHIDRO, em 1979, o PROJETO NORDESTE, em 1982, que foi incorporado ao Programa de Apoio ao Pequeno Produtor Rural - PAPP, em 1985; e, o projeto ÁRIDAS de 1992 [59].

Esses programas tiveram as suas ações prejudicadas, em maior ou menor grau, pelos cortes ou irregularidades na liberação de verbas, ingerência política na gestão, corrosão do valor das verbas pela inflação, falta de determinação para implementação efetiva e eficiente dos programas e a inexistência de uma política de reorganização fundiária que ampliasse o volume de créditos à produção agropecuária no semi-árido nordestino [59].

Essas mudanças também estão relacionadas a uma mudança de paradigma que expressam a passagem da concepção mecanicista para uma visão holística e ecológica, reconciliando homem e

natureza. A concepção holística tem por base que inter-relação e interdependência são elementos essenciais em todos os fenômenos físicos, biológicos, culturais e sociais. Dentro dessa visão passa a ter destaque o conceito de convivência com o semi-árido, com base na sustentabilidade ambiental, na qualidade de vida das famílias sertanejas e no incentivo às atividades econômicas apropriadas ao semi-árido [159].

Nesse contexto, no final do governo de Fernando Henrique Cardoso, na estiagem de 2001/2002, foi lançado o programa Sertão Cidadão: Convívio com o semi-árido e inclusão social. Além das ações imediatas de assistência à população sertaneja que esse programa contemplava, foi proposta a criação de um Sistema de Planejamento e Gestão do Semi-árido, com a finalidade de monitorar a dinâmica espacial e temporal de sistemas ecológicos e socioeconômicos. Uma outra proposta nesse período foi o Programa de Disseminação de Tecnologias Apropriadas para o Semi-árido, que visava a promoção de mudanças no padrão tecnológico e alternativas de inserção no mercado [159].

No governo Lula, em março de 2003, foi lançado o Programa Conviver - Desenvolvimento Sustentável do Semi-árido, com ações voltadas para a melhoria de vida dos agricultores familiares da região, cuja principais ações foram: seguro safra, compra de alimentos pelo Governo Federal, acesso ao crédito para ações de manejo e captação de recursos hídricos, investimentos em culturas forrageiras e manejo da caatinga, cartão alimentação, assistência técnica e educação para desenvolvimento de metodologias e tecnologias de convivência com o semi-árido [159].

Destacam-se também os programas de transferência de renda (Bolsa Família, Benefício de Prestação Continuada e Programa de Erradicação do Trabalho Infantil) que correspondiam a cerca de 65% do total de recursos aplicados; o Fome Zero, com ações voltadas para a construção de cisternas no semi-árido, dentro das prioridades de segurança alimentar; uma linha especial de crédito no Programa Nacional de Agricultura Familiar (Pronaf); o Projeto Água Doce, que apoia a implantação de pequenas usinas de dessalinização; e o Plano Estratégico de Desenvolvimento Sustentável do Semi-árido (PDSA), com ações voltadas para: a) revitalização da Bacia do Rio São Francisco, como base para a integração de bacias hidrográficas, b) hidrovias do São Francisco, c) Ferrovia Transnordestina, d) agricultura irrigada, e) produção de energia alternativa (biodiesel, gás natural) e f) mineração e refinaria de petróleo [159].

Não obstante toda a ação governamental proveniente da política de longo prazo no Nordeste, persistem os problemas crônicos da região, embora alguns tenham sido atenuados. A própria sobrevivência das frentes de trabalho, como foi observado na última grande seca, em 1998, é o atestado mais eloquente do insucesso das políticas de longo prazo na região [59].

As frentes de emergência ou de trabalho, constituem historicamente a principal política

governamental de curto prazo para lidar com os efeitos da seca no Nordeste. Começaram a ser implantadas desde a seca de 1877-79, porém, ao longo de mais de um século, os benefícios sociais permanentes advindos das obras executadas têm sido extremamente modestos. Problemas como a ingerência política no alistamento dos trabalhadores, desvios de verbas e de víveres destinados aos flagelados, execução de benfeitorias em propriedades privadas, atraso e pagamento de salários em locais distantes do município onde residem os trabalhadores, falta de plano para as obras, entre outros, contribuíram para a ineficiência desses resultados. Dentre as principais obras executadas pelas frentes de trabalho estão a implantação e manutenção de trechos rodoviários, construção e melhoramento de açudes e aguadas, perfuração de poços tubulares e o estabelecimento e limpeza de canais de irrigação. Não obstante os seus resultados negativos, estas medidas aliviam a situação de castrastrofe da região em épocas de seca, concedendo remuneração com dignidade ao homem do campo e dinamizando a economia com a injeção financeira através do pagamento de salários. Associado a isto está o apego à terra do nordestino, que prefere manter-se nas proximidades do local onde vive, na esperança de que, na primeira chuva, possa retornar a sua gleba [59,127].

Observam-se grandes impactos econômicos e sociais como consequências das secas no semi-árido nordestino. Esses impactos sócio-econômicos, referentes às últimas secas ocorridas no período de 1958-1998, serão comentados de forma mais detalhada a seguir.

A seca de 1958 atingiu de forma mais intensa áreas parciais de seis Estados: Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco e Bahia, abrangendo uma área de aproximadamente 650.000 km² com uma população de 7 milhões de habitantes. Cerca de 2 milhões de pessoas foram vítimas diretas dessa seca, segundo os dados da pesquisa sistemática feita pelos técnicos do BNB. Excetuando-se as áreas de serras, os vales úmidos e as áreas de irrigação, houve uma redução de 40% na safra agrícola dessas áreas com prejuízo de 132 milhões de dólares, à cotação de 1957, com a agricultura contabilizando um prejuízo da ordem de 106 milhões de dólares e a pecuária de 26 milhões. As ações governamentais no socorro às populações atingidas se restringiram principalmente à implantação de frentes de emergência para execução de obras de infraestrutura em rodovias, ferrovias e açudes [58].

Em 1970 o quadro social demonstrou-se mais calamitoso, motivado pela intermitência na precipitação dos meses do período chuvoso (seca verde), o que teve um efeito descapitalizador sobre o agricultor, pelas sucessivas tentativas de plantios improdutivos. Essa alternância na precipitação é particularmente nociva à agricultura, o que evidencia a importância de se utilizar mecanismos de análise que caracterizem e quantifiquem a distribuição das precipitações dentro do ano ou do período chuvoso [127]. No Quadro 2.5 são apresentadas as perdas relacionadas

às principais lavouras do semi-árido nordestino, no período de 1969 a 1970. Nota-se que os Estados do Piauí, Ceará, Rio Grande do Norte e Paraíba foram os mais afetados pela estiagem. Os dados das culturas alimentares referentes a Alagoas foram caracterizados como erro estatístico [127]. De uma forma geral, na seca de 1970, o Nordeste contabilizou 32% de perdas nas culturas alimentares e industriais, mas a pecuária praticamente não sofreu. Os efetivos dos rebanhos ou mantiveram a situação de estabilidade ou chegaram até a evoluir. Isso foi consequência da estratégia utilizada pelos produtores de sacrificarem as culturas permanentes, nesse caso o algodão, para manterem o efetivo dos seus rebanhos. Quanto às ações governamentais, não se restringiram às frentes de emergência, tendo o governo adotado outras medidas, com destaque para: o controle do abastecimento de água; a abertura de uma linha de crédito de emergência para os proprietários; a antecipação da liberação das cotas de participação dos municípios, o Plano de Integração Nacional (PIN) - baseado na concessão de 30% de incentivos fiscais ao programa de irrigação no Nordeste; e o Programa de Redistribuição de Terras e de Estímulo à Agro-Indústria do Norte e Nordeste (Proterra) - que absorvia 20% das importâncias deduzidas do imposto de renda. Quanto às frentes de trabalho, foram inscritos cerca de 500 mil trabalhadores de um total de 2,5 milhões de flagelados, com uma remuneração que correspondia, na época, a um quarto da renda familiar média do setor rural nordestino [127].

Estados	Culturas Alimentares (A) ¹	Culturas Industriais (B) ²	(A) e (B)
PI	-42,4	-67,6	-47,6
CE	-52	-48	-50,2
RN	-44,3	-36,5	-40,1
PB	-37,1	-38,8	-37,9
PE	-27,2	-35,4	-29,3
AL	-66,7	-37	-60,2
SE	-14,7	-11,9	-14,4
BA	-7,9	-2,2	-6,2
NE ³	-32	-32,8	-32,3

¹ Arroz, batata-doce, feijão, mandioca e milho, ² Algodão, amendoim, fumo, mamona e sisal, ³ Excluindo o Maranhão.

Quadro 2.5 Variação percentual da produção das principais lavouras do Nordeste no período de 1969 a 1970.

Fonte: PESSOA [127].

A seca de 1979-83 caracterizou-se principalmente pelo longo período de anos consecutivos com déficit na precipitação, embora a intensidade tenha variado de ano para ano. Em 1979 a

extensão da calamidade já atingia 539 mil km² do território nordestino, com 0,8 flagelados por km², porém os efeitos no abastecimento d'água ainda eram pouco sentidos, devido aos bons níveis das reservas hídricas. Na pesquisa realizada em 1979, pelos técnicos do BNB, pouco mais de 7% dos alistados nas frentes de emergência abasteciam-se com carro pipa. Os anos consecutivos de seca agravaram a situação, a ponto que, no ano de 1983, a SUDENE necessitou mobilizar a ação de 4 mil carros pipas em todo o Nordeste, para o suprimento de água à população flagelada [46].

As culturas de subsistência como milho e o feijão tiveram uma redução de 82 e 72%, respectivamente, no período de 1978-80, e a principal cultura comercial na época, o algodão, teve um declínio de produção de 69%. Um exemplo desses déficit de produção é que, em 1981, a quantidade produzida de milho no Nordeste correspondia a um terço da alcançada em 1978.

A pecuária, por sua vez, manteve a característica de estabilidade no efetivo dos rebanhos das secas anteriores, comprovando a afirmativa de tratar-se de uma atividade mais resistente aos efeitos das secas.

Em 1979, a ação do Estado no atendimento aos flagelados da seca recebeu um novo direcionamento, pelo fato de prever a execução de obras financiadas em propriedades particulares, assim beneficiando os proprietários rurais e buscando a organização e o fortalecimento de pequenas e médias unidades de produção agropecuária. Foram beneficiados com o programa de emergência, 5 Estados, 513 municípios e 430 mil trabalhadores rurais com uma série de obras na infra-estrutura das propriedades. O programa previa para propriedades de 0 a 20 hectares, 100% de mão-de-obra a fundo perdido; para propriedades de 21 a 100 hectares, 80% da mão-de-obra e 20% financiados por meio de empréstimos; para os imóveis de 101 a 500 hectares, 70% a fundo perdido e 30% financiado; e para as propriedades acima de 500 hectares, 50% a fundo perdido e 50% financiados. Um ponto negativo desse programa foi a exigência de cadastramento da propriedade no Instituto Nacional de Colonização e Reforma Agrária (INCRA), pois, essa medida excluiu um contingente considerável de pequenas propriedades que em decorrência de partilhas entre herdeiros não possuíam a área mínima necessária. Associado a isto vinha a restrição ao alistamento de mais de uma pessoa da mesma família, as interpretações equivocadas por parte dos técnicos responsáveis pelo alistamento e a falta de divulgação do alistamento da forma adequada [46].

Na seca de 1993, o Nordeste contava com o Programa de Monitoramento de Tempo e Clima, criado em 1992, pelo Ministério de Ciência e Tecnologia, com o objetivo de estudar em toda a sua plenitude as variações climáticas da região. Com esse programa foram criados Centros de Meteorologia nos Estados do Nordeste sob a coordenação do Centro de Previsão de

Tempo e Estudos Climáticos (CPTEC) do Instituto de Pesquisa Espaciais (INPE). Apesar das informações atualizadas, da reputação científica desse órgão e dos métodos de previsão baseados em séries estatísticas, a previsão em janeiro de 1993, era de uma quadra chuvosa normal a ligeiramente seca, com um total de chuvas entre 0 a 20% abaixo da média (CLIMANALISE, 1993 *apud* MOREIRA FILHO [113]). Isso não foi o que ocorreu e no Boletim Climanálise do CPTEC de junho de 1993, é apresentada uma redução nas precipitações registradas da ordem de 63% abaixo da média climatológica. Nesse mesmo documento afirma-se que foi uma situação anômala, pois o fenômeno El Niño não se configurava como evento forte e as condições de temperaturas da superfície do mar no Atlântico encontravam-se normais, não favorecendo, portanto, a ocorrência de seca na região (CLIMANALISE, 1993 *apud* MOREIRA FILHO [113]). Indiferente às querelas de acerto ou não da previsão climática, o quadro social da seca de 1993 apresentou 12 milhões de flagelados, com 2 milhões de inscritos nas frentes de trabalho, o que corrobora a redução drástica na precipitação observada e a dimensão de catástrofe do fenômeno na região. A maioria dos açudes de Pernambuco e da Paraíba encontrava-se com percentuais extremamente baixos, em torno de 10% de suas capacidades máximas, ou simplesmente secaram, situação vivenciada pelos outros Estados da região [113].

Na seca de 1998, as previsões climáticas foram acertadas. Os parâmetros oceânicos e atmosféricos demonstravam-se desfavoráveis às chuvas no Nordeste: o El Niño estava configurado como forte; no Atlântico Norte predominaram as anomalias positivas de TSM e no Atlântico Sul anomalias negativas de TSM, constituindo um padrão de dipolo desfavorável às precipitações; e, a ZCIT permaneceu ao norte do Equador durante todo o período chuvoso, não influenciando as chuvas na região [48–51]. Os desvios de precipitação (mm) em relação à média climatológica ocorreram em toda região, em praticamente todos os meses do período chuvoso (Figura 2.9 A-D), sendo que em muitos locais a precipitação acumulada em abril totalizava apenas 20% do valor climático da quadra chuvosa (fevereiro a maio) [50]. Corroborando a situação climática estimou-se em 10 milhões de flagelados o contingente da população atingida nessa seca, para um número de alistados nas frentes de emergência ou produtivas de 1,2 milhões, o que corresponde somente a 12% da população atingida. Isso demonstra a pouca amplitude do programa governamental de socorro às vítimas nesse ano, seguindo a tendência das secas anteriores.

Há que se registrar que embora as previsões do INPE já apontassem para a possibilidade de uma seca no Nordeste, houve demora do Governo Federal em prestar socorro à população de flagelados, sendo que as primeiras iniciativas do executivo só chegaram às populações atingidas a partir do mês de junho, e possivelmente motivadas pelas imagens de privação dos flagelados na imprensa nacional. Outro ponto de desarticulação foi a ênfase descentralizadora da admi-

nistração federal, que, na época, transferiu para as comissões municipais a atribuição de alistar os trabalhadores e de indicar e administrar as obras a serem executadas, o que contribuiu para que grande parte dos trabalhadores alistados ficassem ociosos. Apesar destes fatos, as ações implementadas trouxeram também algumas inovações, como o programa de alfabetização de jovens e adultos, as frentes ecológicas e culturais e destinou 450 milhões de reais ao Programa Especial de Financiamento para Combate aos Efeitos da Estiagem, administrado pelo BNB [59].

As perdas na agricultura e pecuária, segundo pesquisa com os alistados das frentes de trabalho e grandes proprietários da região, totalizaram 77% para o milho e o feijão, 42% para o arroz, 92% no algodão e 83% para a castanha-de-caju, que constituem os cinco produtos de maior importância econômica na agricultura do semi-árido nordestino. Na pecuária ocorreu uma redução do contingente dos rebanhos de ordem 42% para os bovinos, 37% para os caprinos, 46% para os suínos, 41% para os ovinos e 51% para as aves. Essas reduções se deram principalmente pela venda, morte e abate para consumo [59].

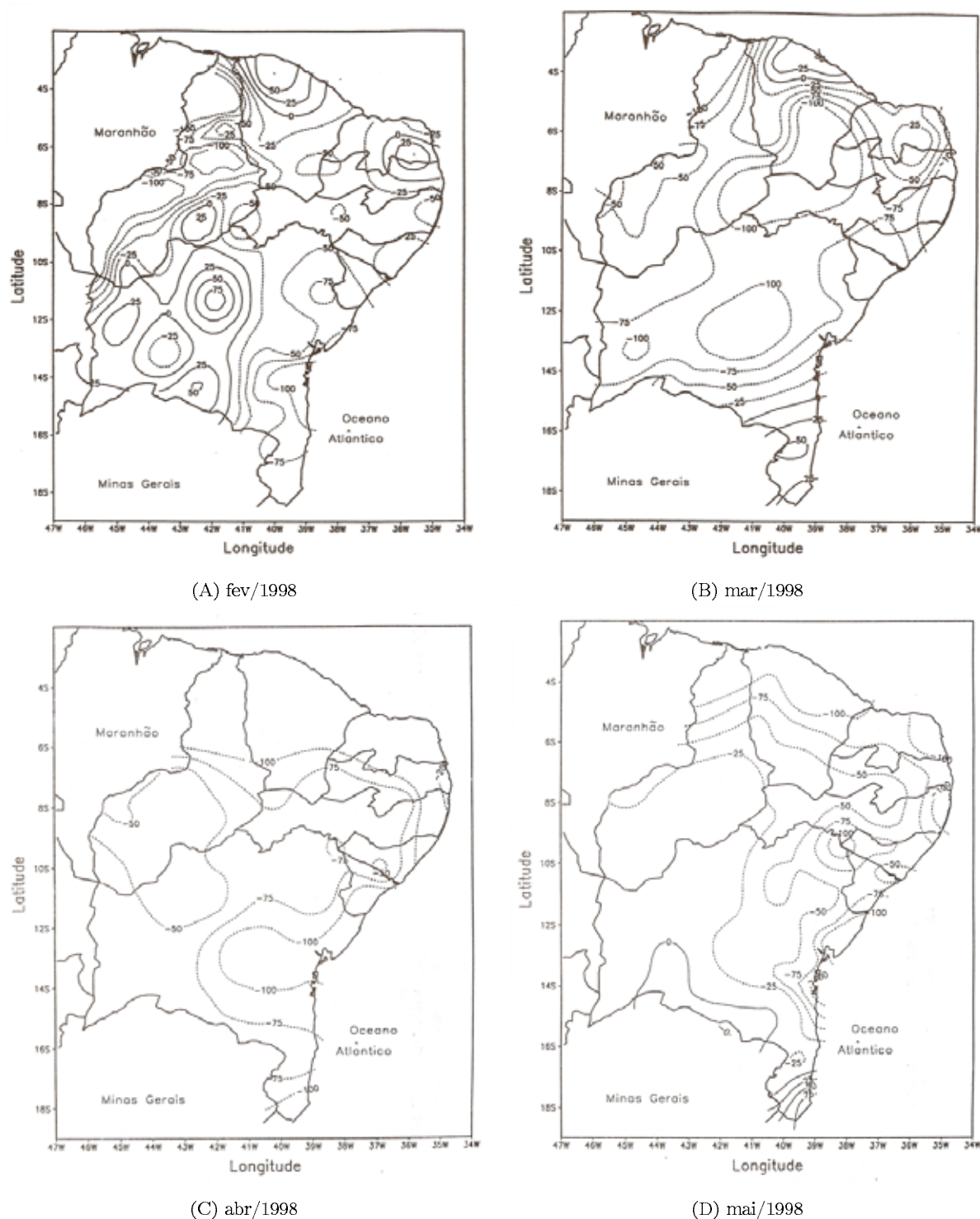


Figura 2.9 Desvios de precipitação em relação a média climatológica do período chuvoso do Semi-árido do Nordeste. As linhas tracejadas representam desvios negativos. Fonte: INPE/CPTEC [48-51].

Ao longo do século passado, diversas tecnologias agropecuárias foram desenvolvidas por instituições de pesquisas, com o objetivo de convivência com a seca. Na década de 70 foi fundada a Empresa de Pesquisa Agropecuária - Embrapa, que implantou centros de pesquisas nos Estados Nordestinos, tais como, o Centro de Pesquisa Agropecuária do Trópico Semi-árido - CPATSA, visando à convivência do homem com o semi-árido. No final da década de

1970, início da década de 1980, foi fomentada a criação das Empresas Estaduais de Pesquisa Agropecuária no Nordeste, como a Empresa de Pesquisa Agropecuária da Paraíba - EMEPA e a Empresa de Pesquisa Agropecuária do Rio Grande do Norte - EMPARN. Outros centros de pesquisas com foco no semi-árido são as Universidades e outras instituições como a Companhia de Desenvolvimento do Vale do São Francisco - CODEVASF.

Um exemplo desse esforço é o sistema idealizado pela Embrapa e a CODEVASF para o desenvolvimento integrado do Nordeste a partir de dois eixos de especialização da atividade agropecuária. O primeiro eixo consistiria na criação de perímetros de irrigação com tecnologia modernas em terras de boa qualidade, e o segundo contemplaria a agropecuária de sequeiro. Esta última através de duas frentes: a primeira seria a substituição do algodão arbóreo pelo algodão herbáceo e a segunda consistiria na introdução, em propriedades do semi-árido, do sistema desenvolvido no CPATSA/Embrapa denominado CBL (caatinga, capim buffel e leucena) [59].

Esse sistema consiste na utilização da caatinga como alimento para os rebanhos na época das chuvas, e na utilização do feno e forragem do capim buffel, da leucena, da maniçoba ou palma forrageira para alimentação dos rebanhos no período seco. No período de grandes secas as famílias poderiam se valer da venda organizada de parte do rebanho para a sua manunção. Esse sistema preconiza, para uma área carente de recursos hídricos, o mínimo de 100 hectares, o que implicaria, para a sua adoção em larga escala, a necessidade de uma reforma agrária, dadas as condições do atual sistema fundiário vigente na região semi-árida nordestina [59].

Diversas outras tecnologias foram desenvolvidas, principalmente na década de 80, com foco voltado para a criação e/ou adaptação de tecnologias de baixo custo para a agropecuária nordestina. A partir da década de 90, o foco estava direcionado para a convivência da população do semi-árido nordestino com a seca.

Embora ainda persistam muitos dos problemas do Nordeste, pois a natureza do subdesenvolvimento da região não está fundamentada somente nas causas climáticas e sim também na adaptação da sua economia e do modelo de exploração da terra, a região apresenta atualmente mais condições e possibilidades que permitem o convívio com as adversidades do clima. Durante muito tempo acreditou-se que o problema resumia-se à questão climática e isso serviu de diretriz para a ação governamental de longo prazo na região. Nesse período, foram executadas diversas obras de infra-estrutura hídrica e de transportes, entre outras, e apesar da não existência de um planejamento estratégico que previsse o seu uso sustentável, modificaram o panorama da região.

Os estudos desenvolvidos por diversos órgãos e pesquisadores sobre o clima, recursos na-

turais e atividades econômicas e sociais, dotaram a região de ferramentas que possibilitam um maior controle dessas variáveis. Um exemplo disso é o monitoramento ostensivo das variáveis climáticas feito na região pelos Centros de Meteorologia Estaduais e o INPE, e que possibilitou, por exemplo, a previsão antecipada da ocorrência das secas. A partir da década de 1990, pode-se citar também as instituições de pesquisa que desenvolveram tecnologias na adaptação e melhoramento genético de cultivares mais resistentes à seca e de ciclo mais precoce; no desenvolvimento e adaptação de sistemas de cultivo e produção; e, na adaptação e introdução de raças mais produtivas nos rebanhos da região.

Muito ainda há que se fazer, apesar das tentativas de desenvolvimento tecnológico, pois persistem na região as práticas arcaicas de produção, a grande vulnerabilidade da agricultura e em menor proporção da pecuária e o subemprego agravado pelo aumento do trabalho temporário. A isto somam-se os mesmos programas governamentais quando da ocorrência de uma grande seca, tais como as frentes de emergência ou trabalho. Isso demonstra que as soluções encontradas até hoje não foram suficientes para resolver o problema do semi-árido nordestino, independente das razões, se técnicas ou políticas. Dentro deste escopo, além das soluções definitivas, torna-se importante o desenvolvimento de ferramentas que possibilitem ao executivo otimizar os recursos e as suas ações, quando da ocorrência de uma grande seca na região. Uma opção pode ser a adoção de ferramentas que possibilitem um diagnóstico mais preciso dos efeitos das secas nas populações atingidas, tais como os índices de severidade de seca.

2.4 Descrição da Área Foco do Estudo

O Estado do Nordeste com maior proporção de área semi-árida é o Rio Grande do Norte (RN), com 93,4% dos 53.077 km² da área total. É o sétimo Estado em área no Nordeste, mas apresenta uma localização estratégica no segmento oriental do Nordeste Brasileiro, entre os paralelos 4°49'53" e 6°58'57" de latitude sul e os meridianos de 34°58'06" e 38°34'54" de longitude a oeste de Greenwich. A sua população é de 2,77 milhões de habitantes e está dividido politicamente em 167 municípios, agrupados em 4 mesorregiões (Figura 2.10) e 19 microrregiões geográficas (Figura 2.11).

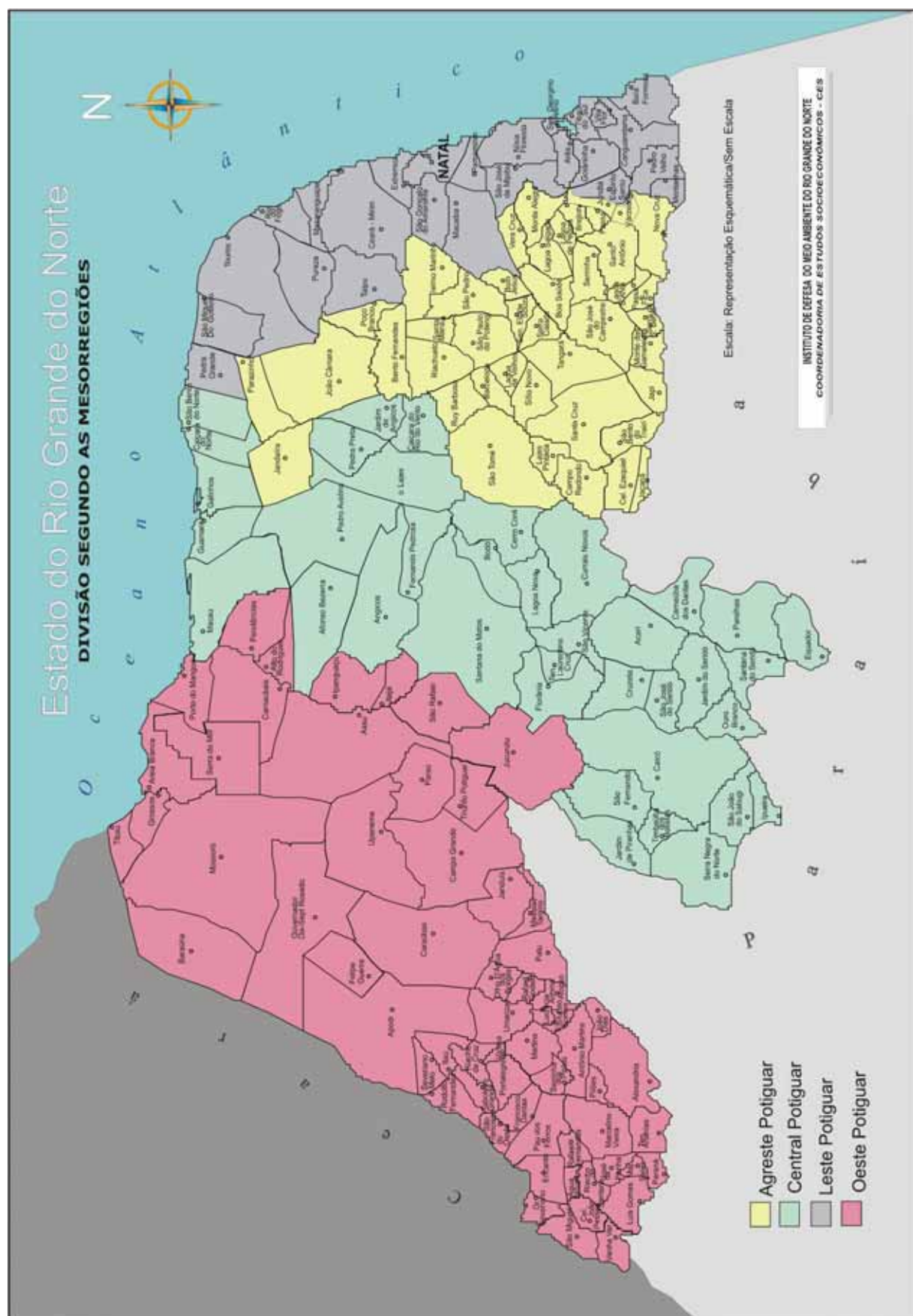


Figura 2.10 Mesorregiões do Rio Grande do Norte. Fonte: IDEMA[90]

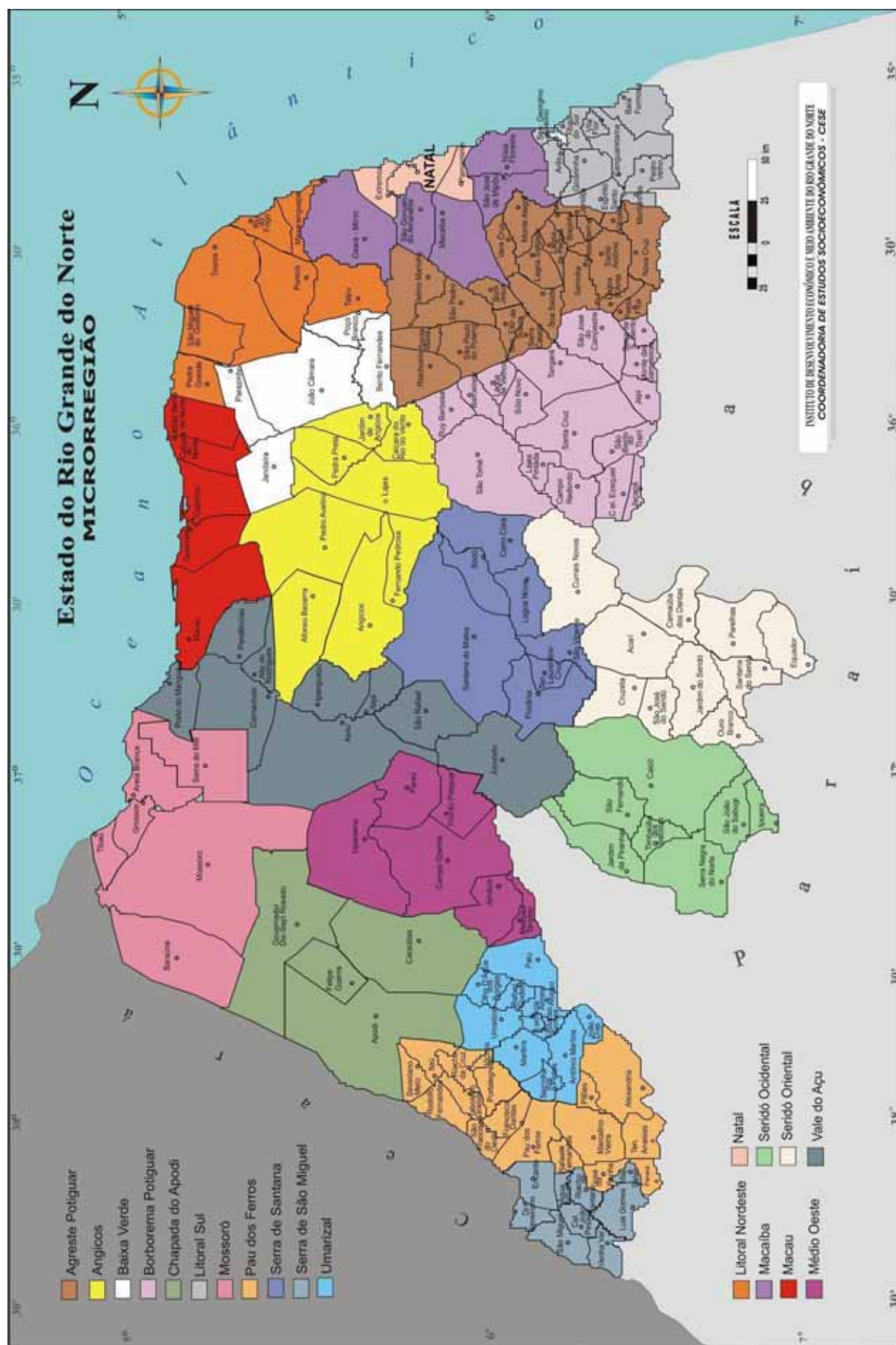


Figura 2.11 Microrregiões do Rio Grande do Norte. Fonte: IDEMA [90].

Comparativamente o RN, em população, está em 8º lugar no Nordeste, e apresenta uma densidade populacional no semi-árido de 32,28 hab/km². Este fato faz do semi-árido do RN o 6º em densidade populacional em relação aos outros Estados semi-áridos, considerado elevado quando comparado à densidade total do semi-árido nordestino (23,73 hab/km²), e, mais marcadamente quando comparado à densidade populacional média das áreas semi-áridas do mundo (26 hab/km²) [54].

Os municípios com maior densidade demográfica (Figura 2.12), localizam-se principalmente na faixa litorânea, com alguns municípios nas microrregiões de Serra de São Miguel, Pau dos Ferros, Umarizal, e Mossoró. As microrregiões do Seridó Oriental e Ocidental apresentam municípios com densidade que variam de 17,5 a 38,4 hab/km². As menores densidades demográficas ocorrem na parte central do Estado e na Chapada do Apodi e Médio Oeste.

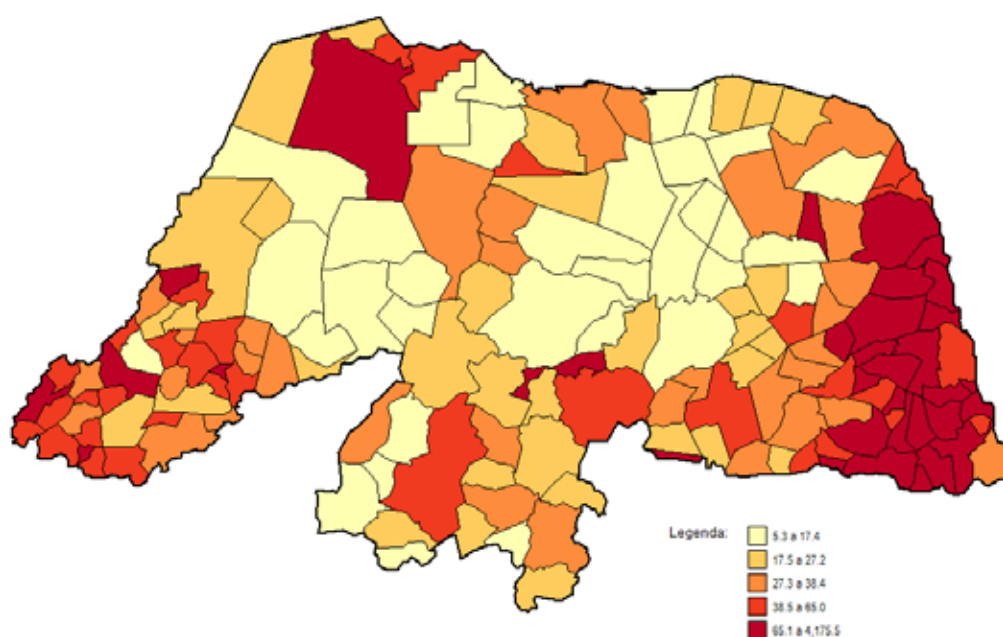


Figura 2.12 Densidade demográfica (hab/km²) dos municípios do Rio Grande do Norte. Fonte: PNUD [133].

Quanto à população rural (Figura 2.13), as maiores concentrações localizam-se na mesorregião Leste, com contingentes variando de 6.932 a 59.637 habitantes, e no Oeste Potiguar, particularmente nas microrregiões de Mossoró, chapada do Apodi, e Assu, onde destacam-se os municípios de Mossoró, Assu e Apodi. A mesorregião Central apresenta uma situação intermediária, com predominância de contingentes populacionais rurais variando de 2.458 a 6.932 habitantes.

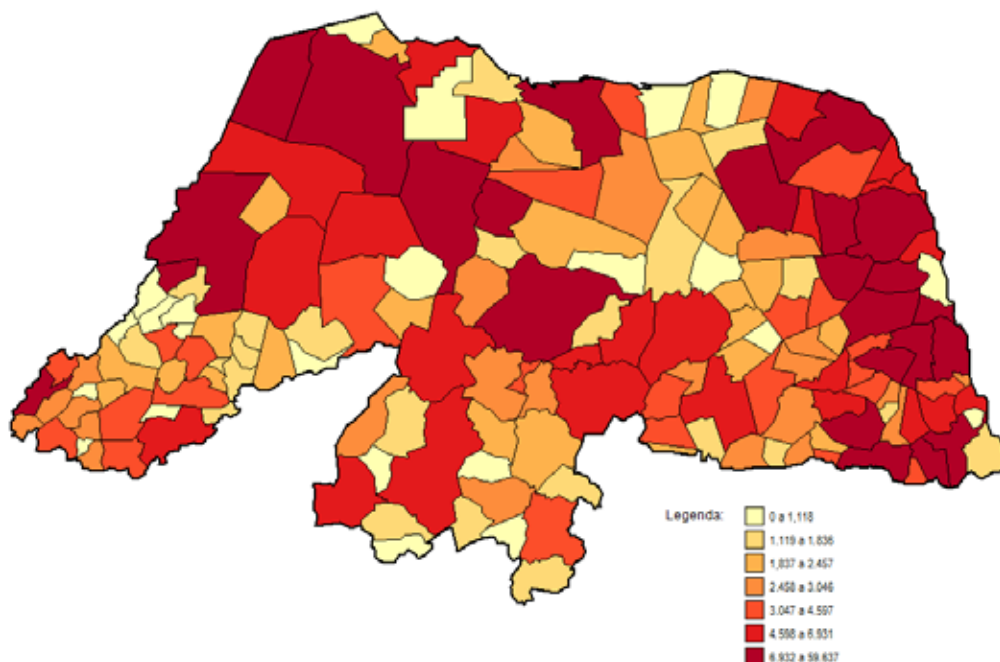


Figura 2.13 População rural (hab) dos municípios do Rio Grande do Norte. Fonte: PNUD [133].

Quanto aos aspectos físico-ambientais apresenta as características de regiões próximas ao equador, com temperaturas elevadas durante todo o ano, baixas amplitudes térmicas, forte insolação e altas taxas de evaporação. A semi-aridez destaca-se pela acentuada variabilidade espacial e temporal da pluviometria, taxas negativas de balanço hídrico, predominância de solos cristalinos de baixa profundidade e reduzida capacidade de retenção de água [4].

O relevo apresenta as seguintes formações: planícies fluviais e costeira, depressões, tabuleiros costeiros, chapadas e planaltos. As maiores altitudes (serras), com variação de aproximadamente 580 m a 830 m, localizam-se nas microrregiões Serra de São Miguel, Pau dos Ferros, Umarizal, Médio Oeste, Serra de Santana, Seridó Oriental e Borborema Potiguar (Figura 2.14).

Quanto à classificação climática apresenta os seguintes tipos climáticos: sub-úmido, úmido seco, árido e semi-árido (Figura 2.15).

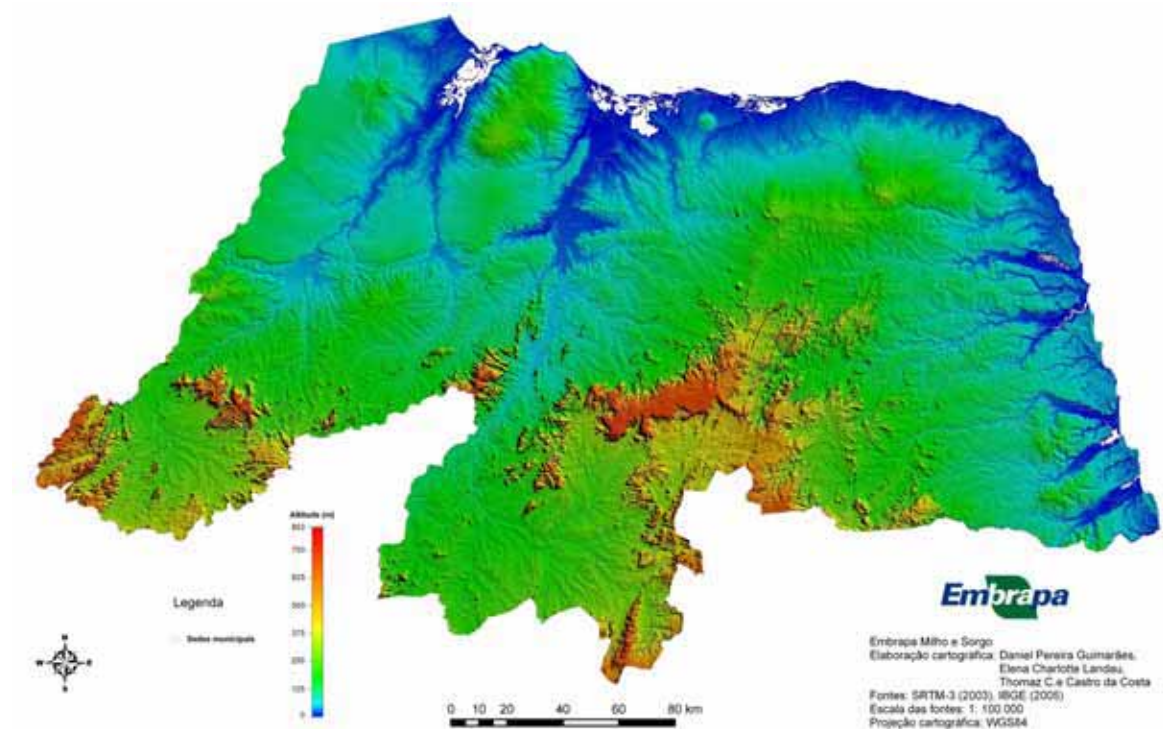


Figura 2.14 Altimetria do Rio Grande do Norte. Fonte: Embrapa.

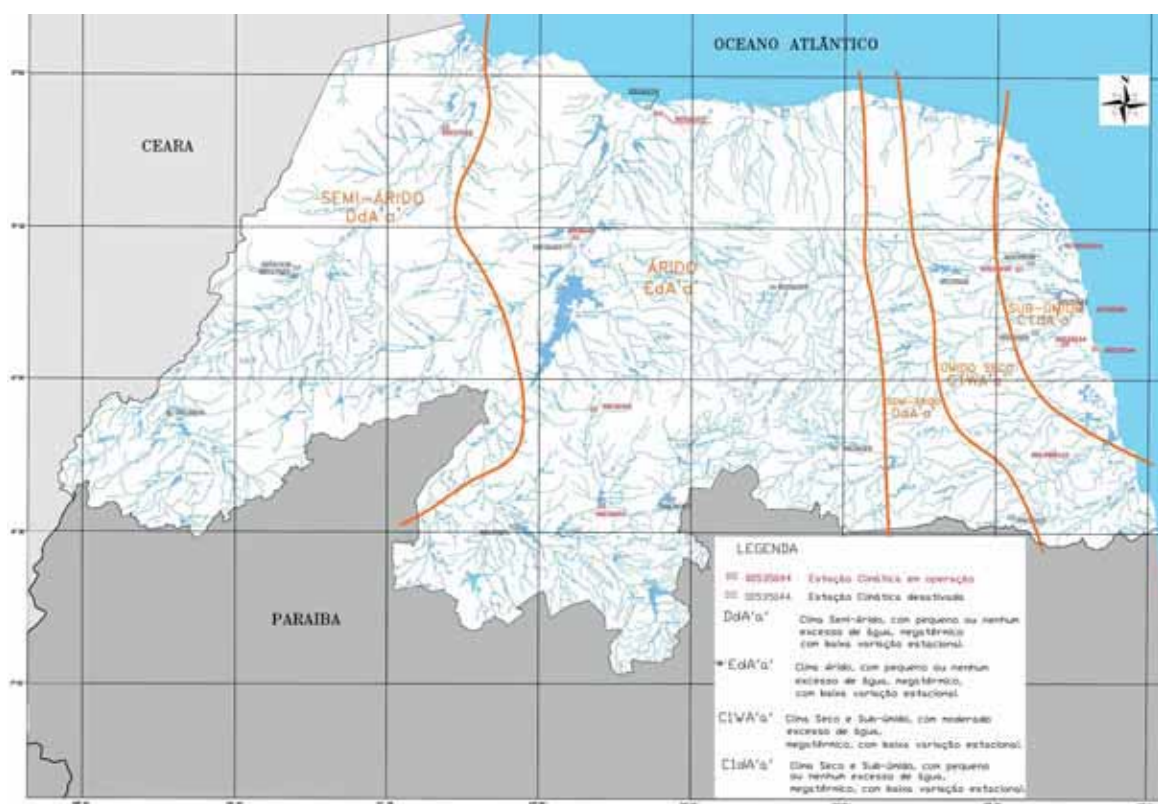


Figura 2.15 Classificação climática do Rio Grande do Norte - Segundo Thornthwaite

O regime pluviométrico do RN é determinado pelo deslocamento da ZCIT, atuando no período de fevereiro a maio na região semi-árida central, oeste e norte do Estado. A ausência de

deslocamento da ZCIT para a posição mais ao sul do equador contribui para as irregularidades de chuvas e favorece os períodos de estiagem. Na faixa litorânea Leste e no Agreste são os Distúrbios Ondulatórios de Leste que se apresentam como o principal sistema causador de chuvas, atuando principalmente de maio a agosto [8,10,121].

A desertificação é outro fator que contribui para aumentar a vulnerabilidade do semi-árido do RN à seca, particularmente, nos municípios das microrregiões Seridó Oriental e Ocidental que constituem o Núcleo de Desertificação do Seridó (NDS) (Figura 2.1), e nas áreas do seu entorno. O NDS abrange os municípios de Currais Novos, Cruzeta, Equador, Carnaúba dos Dantas, Acari, Paralhas, Caicó, Jardim do Seridó, Ouro Branco, Santana do Seridó e São José do Seridó. Dentre os quatro núcleos de desertificação do Nordeste, o NDS é o que apresenta maior densidade populacional (33,9 hab/km²), muito elevada quando comparada com outras áreas semi-áridas no mundo, e, principalmente, com os outros núcleos de desertificação: Gilbués (3,2 hab/km²), Irauçuba (8,5 hab/km²) e Cabrobó (4,8 hab/km²) [147,148].

As principais causas da desertificação no NDS são: o desmatamento da Caatinga para a extração de lenha e argila, patrocinado principalmente pela indústria ceramista; e, o uso intensivo dos recursos naturais e sobrepastejo, consequência da atividade agropecuária na região. Isso é agravado pelas condições climáticas vigentes na região (precipitação média variando de 400 a 680 mm anuais), topografia acidentada, solos rasos e pedregosos e pela vegetação arbustiva baixa e muito aberta [147,148].

Em uma análise do impacto da ação humana no uso da terra e na produção do NDS, SAMPAIO et al. [148] demonstraram que as lavouras permanentes e temporárias apresentaram uma contração de área e produção no período 1985-1996, voltando a se recuperar no período de 1996-2006, caso do milho e do feijão, embora o milho em termos de área não tenha conseguido chegar aos patamares de 1985 (Quadro 2.6). A mandioca manteve a tendência de retração do período anterior também para o período de 1996-2006. Isso sugere, de certa forma, um fator de agravamento sobre essa atividade, além das secas periódicas que atingem a região, e que pode ser a desertificação. Quanto à produtividade, houve declínio somente em alguns municípios, com destaque para os municípios de Parelhas e Equador, com queda na área e produtividade para mais de um produto. Na pecuária não ocorreram mudanças muito significativas, o que sugere que a desertificação não se mostra ainda limitante do ponto de vista dos rebanhos [148].

Produto	1970	1975	1980	1985	1996	2006
Área Plantada (ha) ⁽¹⁾						
Alg.arbóreo	-	27.522	27.125	11.363	327	-
Feijão	1.052	3.955	1.299	6.436	4.182	9.512 ⁽³⁾
Mandioca	24	14	3	142	44	2 ⁽³⁾
Milho	408	3.668	565	6.143	4.174	4.932 ⁽³⁾
Produção (ton.) ⁽²⁾						
Alg.arbóreo	1.576	4.692	1.487	1.560	65	-
Feijão	173	1.012	218	1.188	816	4.408
Mandioca	78	79	25	593	84	2
Milho	78	1.199	93	1.626	1.511	7.843

(1), (2) Fonte: IBGE. Censos Agropecuários de 1970, 1975, 1980, 1985 e 1995/96; (3) Área colhida(ha).

Quadro 2.6 Evolução da área plantada e produção das principais culturas temporárias e permanentes do NDS.

Fonte: Adaptado SAMPAIO et al. [148].

Os principais produtos agrícolas do RN são banana, castanha de caju, mamão, cana-de-açúcar, melão e mandioca. Os principais produtos de origem animal comercializados no Estado são camarão, pescado, leite, ovos de galinha e mel de abelha, que correspondem a aproximadamente 30%, 11%, 6,5%, 10% e 9% do valor da produção do Nordeste. Convém observar que o Estado é o maior produtor de camarão em cativeiro do Brasil e o quarto em pesca marinha no Nordeste. Quanto ao petróleo é o primeiro lugar em produção terrestre e o terceiro em produção total, e também é o quinto em gás natural [3].

O valor da produção das lavouras permanentes e temporárias de todo o Estado, corresponde a aproximadamente 4% do valor da produção dessas lavouras no Nordeste, e 0,8% e 0,5% do Brasil, respectivamente.

Quanto aos indicadores econômicos, o Rio Grande do Norte em 2006 apresentava um PIB de aproximadamente vinte bilhões de reais, o quinto maior do Nordeste, e o terceiro maior PIB per capita do Nordeste, demonstrando uma certa robustez econômica, em relação aos outros Estados do Nordeste [3].

Embora o saldo do emprego formal, acumulado em 12 meses, apresente-se positivo para os setores do comércio varejista e atacadista, construção civil e indústria têxtil, na agropecuária, que é o setor mais vulnerável à seca, apresenta-se negativo desde o mês de julho de 2005 (Figura 2.16) [3].

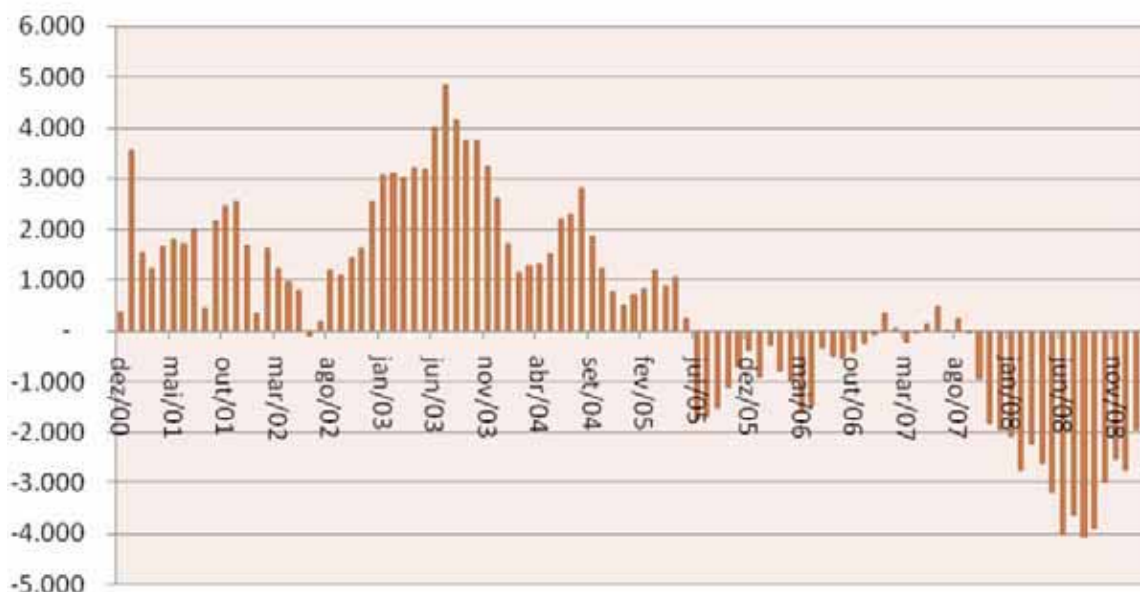


Figura 2.16 Saldo do Emprego Formal da Agropecuária - RN - Período: 12/2000 a 02/2009. Fonte: AGN [3].

Quanto às finanças públicas, as receitas correntes do Estado, em 2007, totalizaram 5 bilhões de reais, o que dá um crescimento de 47,6% em relação a 2002, superior aos crescimentos ocorridos no Nordeste e no Brasil, nesse mesmo período, sendo o quarto maior crescimento do Nordeste (Quadro 2.7) [3].

Área Geográfica	Valores corrigidos pelo IPCA (Em R\$ 1.000.000)						Variação (%)
	2002	2003	2004	2005	2006	2007	2007/2002
AL	2.579	2.618	2.866	3.368	3.658	4.107	59,25
PB	3.361	3.508	3.553	4.308	4.752	5.074	50,97
MA	4.373	4.111	4.461	5.235	5.839	8.541	49,58
RN	3.527	3.551	3.931	4.508	5.052	5.208	47,66
PI	2.456	2.210	2.694	3.215	3.570	3.593	46,30
SE	2.861	2.811	2.995	3.526	3.863	4.047	41,45
BA	13.376	13.913	15.515	16.307	17.596	17.534	31,09
CE	7.185	7.212	7.357	7.914	8.718	9.199	28,03
PE	8.873	8.680	9.501	10.841	10.409	11.275	27,07
Nordeste	48.591	48.614	52.873	59.222	63.457	66.578	37,02
Brasil	251.938	258.667	279.493	304.415	324.001	346.278	37,45

Quadro 2.7 Evolução das receitas correntes a preços de 2007. Fonte: AGN [3].

Com relação aos indicadores sociais, o Rio Grande do Norte apresenta um IDH de 0,73, que o classifica na vigésima primeira posição em relação aos outros Estados Brasileiros, na terceira posição em relação ao restante do Nordeste, sendo semelhante aos IDHs de países como Turquemenistão, Palestina, Indonésia e Honduras [55]. Observa-se que o IDH-M não é uma informação que seja determinante na dinâmica da seca. Este fato se torna visível na distribuição desse índice no RN. Os municípios com maiores IDH-M (Figura 2.17) concentram-se em três regiões distintas: a região úmida, com as microrregiões de Natal e Macaíba; e as regiões semi-áridas compreendendo as microrregiões Seridó Oriental e Ocidental, Mossoró e Assu. Essa característica de distribuição espacial do IDH-M demonstra a ausência de relação entre os valores de áreas semi-áridas e a maior vulnerabilidade à seca, em relação a áreas mais úmidas do Estado. Isso sugere uma não correlação do IDH-M com índices de severidade de seca. Um outro fato que corrobora esta observação é que os municípios com piores IDH-M estão distribuídos pelas diversas microrregiões do Estado, havendo uma maior concentração no Litoral Nordeste.

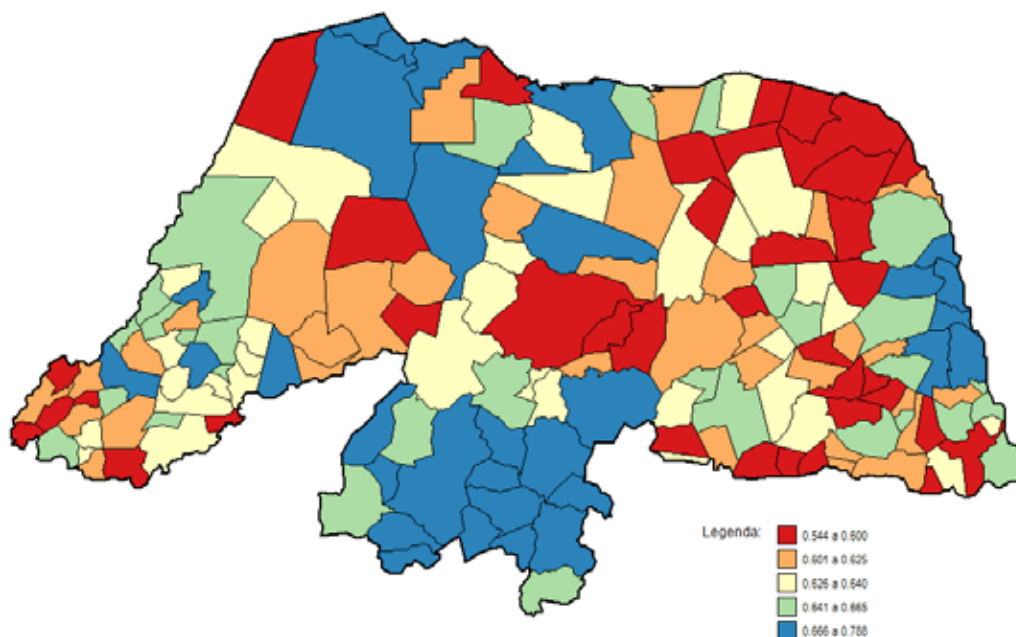


Figura 2.17 Índice de Desenvolvimento Humano Municipal, Rio Grande do Norte. Fonte: PNUD [133].

O Estado conta com 16 bacias hidrográficas (Figura 2.18). Duas são de longo curso e perenizadas pela construção de barragens, a bacia de Apodi/Mossoró com 14.276 km² e a Piranhas/Assu com 17.498,5 km². Elas cobrem aproximadamente 60% da área do RN e no Vale do Assu é onde encontra-se o maior polo de fruticultura irrigada do Estado.

Os recursos hídricos superficiais do RN apresentam uma reserva máxima de aproximadamente 5 bilhões de metros cúbicos, com um grande número de reservatórios d'água, constituídos

por açudes, barragens e lagoas. A distribuição de água à população do semi-árido é feita por uma rede de sistemas de adutoras e canais, com cerca de 1200 km, que se destina principalmente ao abastecimento humano das regiões mais carentes em recursos hídricos do Estado.

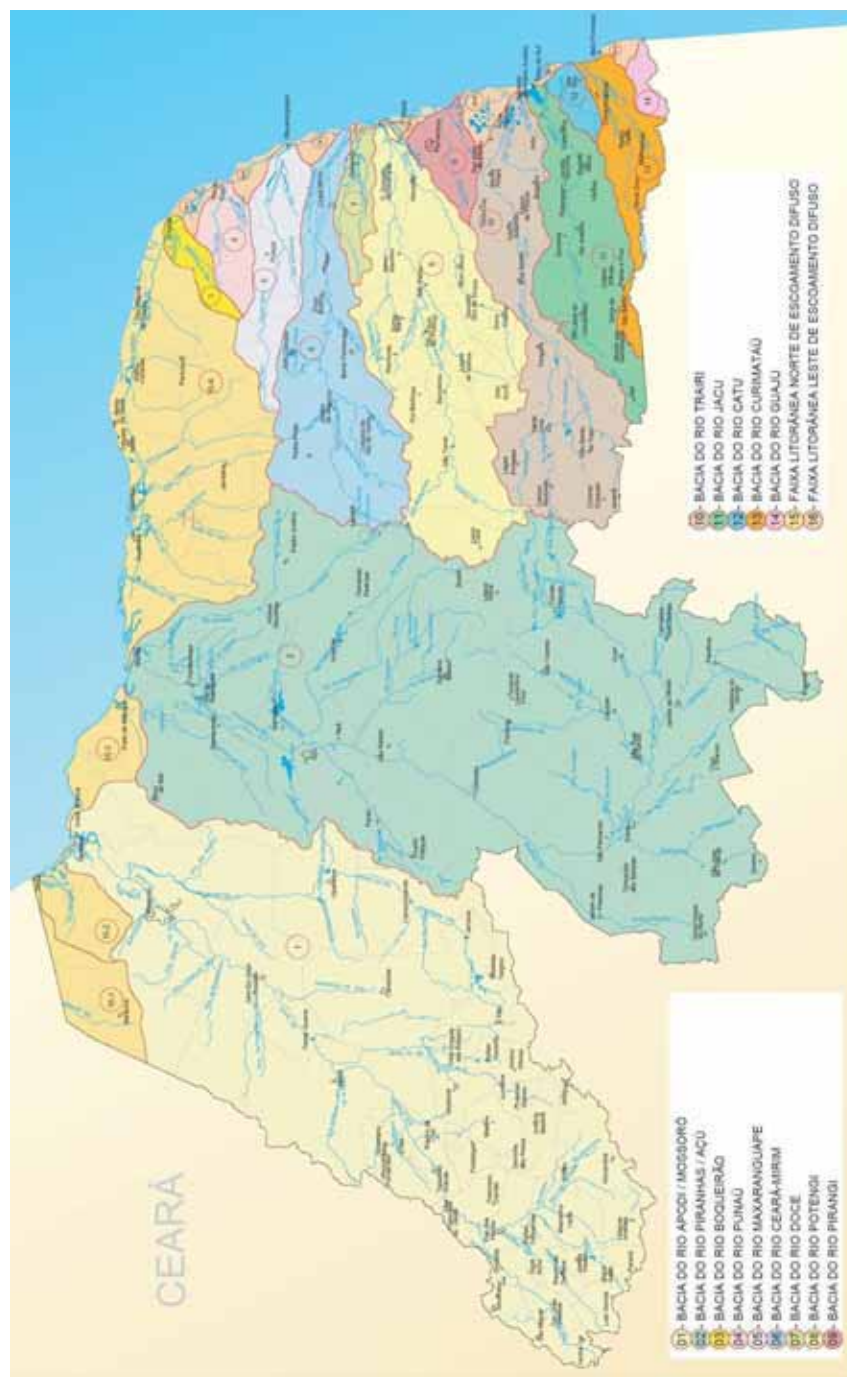


Figura 2.18 Bacias Hidrográficas do Rio Grande do Norte. Fonte: SEMARH/RN [152].

Com relação ao percentual da população atendida pela rede de abastecimento d'água (Figura 2.19), as microrregiões Seridó Oriental e Ocidental, Assu, Pau dos Ferros, o município de Mossoró e os municípios da faixa litorânea, apresentam os maiores percentuais. Os municípios com menor cobertura estão localizados na parte central e norte da mesorregião Central Potiguar, na mesorregião Agreste e na microrregião do Médio Oeste.

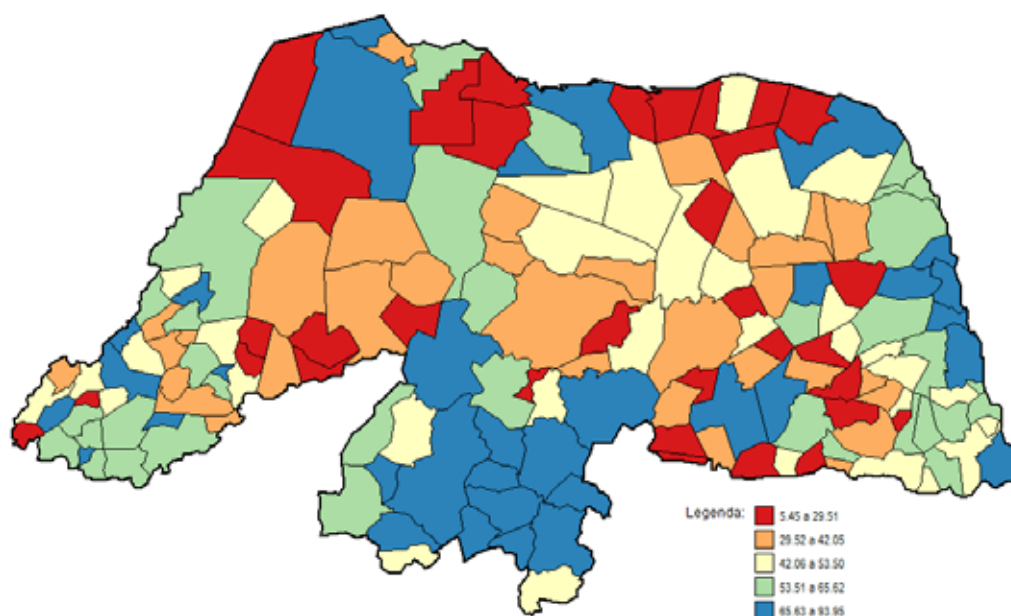


Figura 2.19 Percentual de pessoas que vivem em domicílios com água encanada. Fonte: PNUD [133].

Embora o Rio Grande do Norte apresente alguns indicadores econômicos e sociais relativamente bons, quando comparados com os dos outros Estados do Nordeste, em contrapartida, as condições climáticas do semi-árido se fazem presente em quase a totalidade de sua área. Permanecem, portanto, os problemas crônicos referentes à região semi-árida nordestina e à grande vulnerabilidade da região aos efeitos danosos das secas, principalmente sobre a agricultura, a pecuária e o suprimento hídrico da população e dos rebanhos.

A estrutura fundiária do Rio Grande do Norte segue a mesma tendência verificada no Nordeste, concentração de terras e grande quantidade de minifúndios. Em 1996, 86,19% das propriedades rurais tinham área inferior a 50 ha e abrangiam somente 16,63% da área total de todos os estabelecimentos rurais do RN. Em contrapartida, 1,44% das propriedades com área igual ou superior a 500 ha abrangiam 43% da área total dos estabelecimentos, o que demonstra a assertiva anterior (Quadro 2.8).

	Quantidade de Imóveis		Área Total (ha)	
Classe de Área (ha)	Valor Absoluto	%	Valor Absoluto	%
0,1 a menos de 10	57.958	63,71	155.525,50	4,17
10 a menos de 50	20.451	22,48	465.4545,03	12,46
50 a menos de 100	5.904	6,49	386.009,97	10,34
100 a menos de 200	3.030	3,33	408.651,58	10,95
200 a menos de 500	2.335	2,57	710.792,86	19,04
500 e mais	1.298	1,44	1.607.096,25	43,04
Total	90.976	100	3.733.521,19	100

Quadro 2.8 Distribuição do número e da área dos estabelecimentos rurais do Rio Grande do Norte. Fonte: IBGE [89].

Apesar da intensificação do número de projetos de assentamento rurais realizados pelo INCRA (Figura 2.20 A), principalmente no período de 1996 a 1999, o impacto dessas ações foi pouco significativo na diminuição da concentração da posse da terra [69]. A área média desses assentamentos por família de 28 ha (Figura 2.20 B), representa um outro aspecto negativo, pois, é bastante inferior aos 100 ha mínimos propostos pelo CPATSA para o tamanho de uma propriedade sustentável no semi-árido.

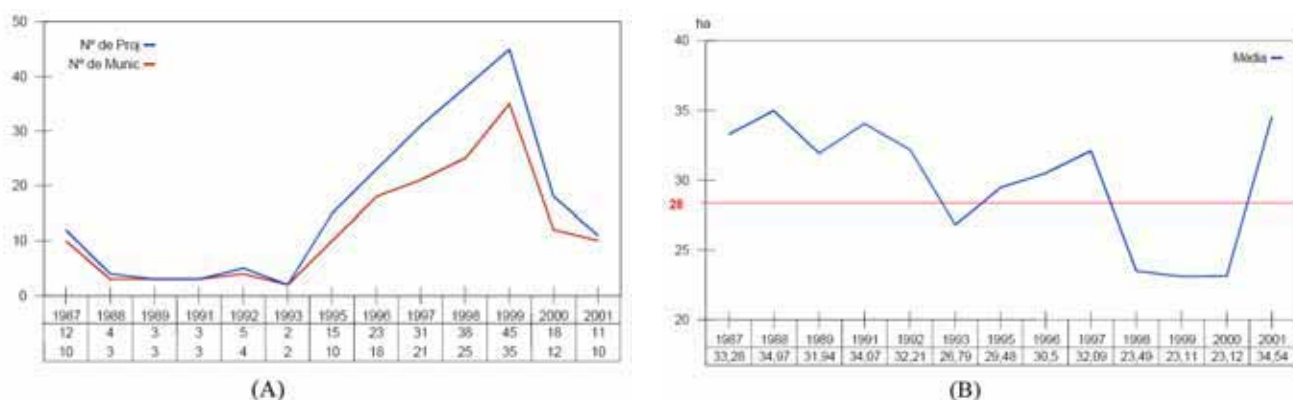


Figura 2.20 Evolução do número de projetos, municípios beneficiados pelos assentamentos rurais do INCRA (A) e área média por famílias assentadas (B). Fonte: FRANÇA & MEDEIROS [69].

Portanto, a questão fundiária, devido a desigualdade na distribuição das terras; a precariedade das relações de trabalho com o agravamento da situação pelo decréscimo do número de empregos no setor agropecuário nos últimos 5 anos; e o processo de desertificação por que passam diversas áreas no Estado, particularmente, o NDS, estão entre os principais problemas que contribuem para aumentar a vulnerabilidade do semi-árido do Estado ao fenômeno da seca.

Essa situação é passível de ser intensificada pelas mudanças climáticas que estão ocorrendo no mundo.

Embora existam atualmente algumas ações quanto ao monitoramento da ocorrência das secas no Estado, através de alguns indicadores, estas ações ainda são muito incipientes e estão longe do ideal. Detalhes desse monitoramento através dos índices de severidade de seca serão apresentados no tópico a seguir.

Em síntese, a própria concepção de convivência com o semi-árido demanda medidas de gestão ambiental sustentável, com iniciativas sociais que resultem em melhoria das condições de vida das populações [159].

É necessário, pois, que se desenvolvam ferramentas de uso fácil, e que permitam integrar as informações ambientais e sociais em modelos únicos de análises, a partir da utilização da tecnologia da informação, a fim de conjugar os conhecimentos na busca de diagnósticos mais precisos e de estratégias eficazes para tratar da seca e dos problemas a ela associados. Esse é um dos motivos que justificam o desenvolvimento do índice proposto neste trabalho de tese.

2.5 Índices de Severidade de Seca

2.5.1 Considerações Iniciais

Um índice ou indicador é definido como uma medida que resume as informações relevantes de um fenômeno particular, cujo objetivo principal é agregar, quantificar e simplificar essas informações com o intuito de melhorar o seu entendimento e compreensão [25].

Os índices de seca são normalmente funções contínuas de chuva e/ou temperatura, descargas de rios ou outras variáveis mensuráveis. Os dados de chuvas são as observações mais largamente utilizadas para composição desses índices, devido ao registro histórico de longo prazo frequentemente disponível; porém, eles sozinhos não refletem o espectro total da dimensão de uma seca [161].

Segundo BYUN & WILHITE [38] o estudo da seca pode ser classificado em quatro categorias:

A primeira categoria trata das causas das secas visando melhorar a compreensão dos fenômenos atmosféricos associados com a sua ocorrência [38]. Trabalhos de pesquisa nessa categoria foram feitos por diversos autores, tais como: BYUN [39], BYUN et al. [37], CHU et al. [53], NAMIAS [118] e TRENBERTH & BRANSTATOR [171]. Pode-se citar também como exemplos de resultados desses tipos de estudos os modelos numéricos de simulação atmosférica utiliza-

dos pelos centros de pesquisas meteorológicas mundiais, tais como: o NCEP/NCAR (*National Center for Environmental Prediction/National Center of Atmospheric Research*) e o ECMWF (*European Center for Medium-Range Weather Forecasts*) [35]. No Brasil, tem-se os modelos de previsão de tempo e clima globais e regionais, como o MCGA (Modelo de Circulação Geral da Atmosfera) e o modelo Eta, ambos utilizados pelo CPTEC/INPE [12].

A segunda categoria é dirigida para o entendimento da frequência e severidade da seca, caracterizando a probabilidade de ocorrência em suas várias magnitudes [38]. Pode-se citar como exemplos de resultados de estudos nessa categoria os seguintes índices: Quantis [132], Decis (GIBBS & MAHER, 1967 *apud* BYUN & WILHITE [38]), SPI (Standardized Precipitation Index) (MACKEE et al. 1993 *apud* BYUN & WILHITE [38]), entre outros. No Brasil, a literatura apresenta uma extensa gama de trabalhos com aplicações desses índices que serão descritos mais adiante.

A terceira categoria tenta descrever e compreender os impactos das secas, baseada nos custos e perdas associados ao fenômeno [38]. O Índice de Severidade de Seca Palmer (*Palmer Drought Severity Index-PDSI*), o Índice FUNCEME de Severidade de Seca (IFS), que será descrito no item 2.5, e o Índice de Susceptibilidade ao Fenômeno da Seca (ISFS), proposto nesse trabalho, são exemplos de resultados nessa categoria de estudo.

E a quarta e última categoria centra-se na resposta, mitigação adequada, estratégias de preparação para eventos dessa natureza e na redução dos impactos associados ao fenômeno da seca. Esta categoria tem progredido rapidamente em recentes anos e o estabelecimento de centros como o NDMC (*National Drought Mitigation Center*) é um exemplo da implantação dessa filosofia pró-ativa, em vez da tradicional filosofia reativa com as quais as instituições sempre enfrentaram os problemas provenientes das secas [38].

2.5.2 Principais Índices de Severidade de Seca Utilizados no Mundo

Uma avaliação objetiva das condições de déficit hídrico ou seca de uma região é o primeiro passo para o planejamento dos recursos d'água a fim de prevenir e mitigar impactos de futuras ocorrências desse fenômeno [35]. Visando atender a essa demanda, diversos índices tem sido propostos no mundo, baseados, principalmente, em variáveis meteorológicas ou hidrológicas (Quadro 2.9) [13,38].

A literatura apresenta uma extensa gama de trabalhos com aplicações diretas ou adaptações desses índices que caracterizam a segunda e terceira categorias propostas por BYUN & WILHITE [38]. Por exemplo: RAI, RI, PDSI, BMI, SPI e HSI na região Saheliana e Oeste da

África, e, na África do Sul [23,81,102,123]; SPI e BMDI na Índia [26,27,112]; SPI e EDI no Iran [5,14]; EDI na Korea [99]; PDSI, CMI, SPI, EDI, SWSI, DI e Desvio da Normal no sul da Ásia [161]; RDI, SPI e DI na Grécia [173,174]; SPI, RAI e RDI na Turquia [67,170,188]; PDSI e SPI em Bangladesh e Portugal [126,128,155]; SPI na península Ibérica [181]; e, PDSI, SPI, DI, quantis e EDI nos Estados Unidos [38,93,132,135].

Índice	Conceito Principal	Fatores p/ cálculo (Escala de Tempo)	Origem
PDSI (<i>Palmer Drought Severity Index</i>)	Baseado na entrada, saída e armazenamento de umidade no solo. Constitui-se no balanço de umidade de solo simplificado.	r,t,et,sm,rf (m, 2s)	Palmer(1965)
RAI (<i>Rainfall Anomaly Index</i>)	Quantifica anomalias positivas e negativas em função dos dez maiores e dos dez menores valores de precipitação.	r (m,yr)	Rooy (1965)
HSI (<i>Herbst Severity Index</i>)	Baseado na precipitação mensal, possibilita o cálculo da duração, severidade, início e fim do período seco.	r (m)	Herbst et al. (1966)
Quantis	Divide a ocorrência histórica de precipitação em faixas de quantis.	r (m,yr)	Pinkayan (1966)
DI (<i>Deciles Index</i>)	Divide a ocorrência histórica de precipitação em seções de 10%.	r (m,yr)	Gibbs e Maher (1967)
CMI (<i>Crop Moisture Index</i>)	Semelhante ao PDSI exceto que considera a umidade disponível nos 5 primeiros metros do perfil do solo.	r,t (w)	Palmer (1968)
BMDI (<i>Bhalme and Mooley Drought Index</i>)	Desvios percentuais da média histórica de precipitação.	r (m,yr)	Bhalme and Mooley (1980)
SWSI (<i>Surface Water Supply Index</i>)	Peso médio das anomalias padronizadas dos principais elementos do balanço hídrico.	P,sn (m)	Shater e Dezman (1982)
SMDI (<i>Soil Moisture Drought Index</i>)	Somatório da umidade de solo diária para um ano.	sm (yr)	Holinger et al. (1993)
RI (<i>Regional Annual Index</i>)	Caracteriza padrões e anormalidades de precipitação em escala regional.	r (yr,c)	Lamb et al. (1986)
CSDI (<i>Crop-specific Drought Index</i>)	Somatório da et calculada dividido em possíveis durações da et no crescimento de culturas específicas.	et (s)	Meyer et al. (1993)
SPI (<i>Standardized Precipitation Index</i>)	Quantifica o déficit ou o excesso de precipitação para diferentes escalas de tempo.	r (3m, 6m, 12m, 24m e 48m)	McKee et al. (1993)
RI _b (<i>National Rainfall Index</i>)	Caracteriza padrões e anormalidades de precipitação em escala continental.	r (yr,c)	Gomes e Petrassi (1994)
RDI (<i>Reclamation Drought Index</i>)	Elemento de fornecimento - elemento de demanda	r,t,sn,st,rs (m)	Weghorst (1996)
EDI (<i>Effective Drought Index</i>)	É baseado na identificação do início e final do período de seca e no cálculo da sua duração.	r (m)	Byun e Wilhite (1999)
NDVI (<i>Normalized Difference Vegetation Index</i>)	É uma relação entre as reflectâncias (ρ) das bandas de infravermelho próximo (IVP) e do vermelho (V).	$\rho IVP, \rho V, V$ (w)	-
VCI (<i>Vegetation Condition Index</i>)	É um refinamento do NDVI, baseado em uma padronização em relação ao NDVI máximo e mínimo de vários anos.	$\rho IVP, \rho V, V$ (w)	-

Legenda: P - fator usado no PSDI, r - precipitação, et - evapotranspiração, t - temperatura, sm - umidade do solo, rf - escoamento, sn-altura de neve, st - vazões, rs - armazenagem no reservatório, w - semana, m - mês, s - estação, yr - ano, c - século, 3m - 3 meses.

Quadro 2.9 Características dos principais índices de seca utilizados no mundo. Fonte: Adaptado de BYUN & WILHITE [38].

QUIRING [135] estudou os pontos fortes e fracos de seis desses índices (PDSI, SPI, EDI, VCI, Desvio da Normal, e DI), utilizando uma avaliação qualitativa baseada nos seguintes parâmetros: robustez, tratabilidade, transparência, sofisticação, estendabilidade e dimensionalidade. Entenda-se por: robustez - a capacidade do índice medir e quantificar a seca através de uma gama de condições climáticas; tratabilidade - a facilidade para cálculo e obtenção dos dados que alimentarão o índice; transparência - o nível de compreensão dos seus resultados; sofisticação - o mérito conceitual ou científico do índice; estendabilidade - capacidade do cálculo ser extensivo ao passado; e, dimensionalidade - a conexão entre o índice e o mundo físico, ou seja, é preferível que um índice tenha uma unidade com significado físico (por exemplo: mm de água no solo, percentual da normal), do que, uma unidade estritamente adimensional ou categórica.

Esse estudo apontou o SPI e o DI como os mais apropriados para o monitoramento das secas meteorológicas nos EUA, com a ressalva de que um só índice não pode representar todos os aspectos inerentes às secas meteorológicas, sendo mais indicado, portanto, o uso de uma abordagem multi-índice [135].

Os índices baseados nas plataformas orbitais (satélites), como o NDVI e o VCI, também utilizados para o monitoramento de secas, apresentam algumas vantagens em relação aos outros tipos de índices já citados, porém, também apresentam limitações [14, 19, 67, 84, 135, 146, 160, 181]. Uma das principais vantagens desses tipos de índices é o fornecimento de informações em tempo real de todas as áreas da terra, com alta resolução espacial. Essas informações possibilitam detectar o início de um período de seca, intensidade, duração e os impactos nas regiões atingidas [135]. As limitações estão associadas à não capacidade de detecção do vigor da vegetação em períodos de dormência, caso do período de inverno com neve, no Hemisfério Norte, cujo stress na vegetação pode ocasionar pragas, doenças e falta de nutrientes; a influência da cobertura das nuvens; a mudanças nos satélites e seus sensores; e, a ocorrência de viés nos dados devido a degradação desses sensores. Como conclusão, esses tipos de índices são mais apropriados para o monitoramento da seca meteorológica e agrícola do que para o monitoramento da seca hidrológica [135].

Nos Estados Unidos, o monitoramento dos eventos de secas é realizado sistematicamente por diversos órgãos como o CPC (*Climate Prediction Center*), o NOAA Drought Information Center, o NCDC (*National Climate Data Center*), o TWDB (*Texas Water Development Board*), e do já citado NDMC. Todos utilizam grande diversidade de índices, tanto isoladamente como em conjunto, a partir de múltiplos índices, que caterizam a quarta categoria proposta por BYUN & WILHITE [38, 135].

O NDMC concentra as suas ações no planejamento para lidar com a ocorrência e os efeitos das secas baseado no monitoramento, na avaliação do risco e dos impactos e nas medidas de mitigação dos seus efeitos. Para tal, o NDMC sugere um planejamento para mitigar os efeitos das secas composto por dez passos que estão descritos no Quadro 2.10 [184]. Ao estudá-los observa-se que os passos de 1 a 4 têm seu foco na maneira correta de reunir as pessoas, no entendimento do processo, no conhecimento do plano a ser executado e no fornecimento de dados suficientes para tornar justas e equitativas as propostas formuladas no plano de preparação para a seca. O passo 5 descreve o processo de desenvolvimento de uma estrutura organizacional para a conclusão das tarefas necessárias para preparar o plano. Os passos 6 e 7 detalham as necessidades de investigação em curso e de coordenação entre cientistas e decisores políticos. Os passos 8 e 9 salientam a importância de promover e testar o plano antes que a seca ocorra. E o passo 10 enfatiza a revisão do plano para mantê-lo atual e avaliar a sua eficácia no período pós-seca [184].

Passo	Descrição
1º	Nomear uma força tarefa da seca
2º	Definir o propósito e os objetivos do plano de preparação para a seca
3º	Buscar a participação dos interessados e resolver os conflitos
4º	Inventário dos recursos e identificação dos grupos de risco
5º	Preparar/Escrever o plano
6º	Identificar necessidades de pesquisa e preencher as lacunas institucionais
7º	Integrar a ciência com a política
8º	Divulgar o plano de preparação para a seca e desenvolver a consciência pública
9º	Desenvolver programas educacionais
10º	Avaliar e rever plano de preparação para a seca

Quadro 2.10 Plano de preparação para a seca do NDMC. Fonte: WHILITE [184].

Também dentro desse escopo, ROSSI et al. [142] apresentaram um estudo de classificação das medidas de mitigação à seca baseado nas propostas dessas medidas, na abordagem de implementação, na janela de tempo de duração dos eventos e na proposição do Grupo Europeu de Escassez de Água, além de uma panorâmica da legislação para administração dos eventos dessa natureza na Espanha e na Itália. Eles propuseram nesse estudo uma nova legislação, um ambiente institucional para administração das secas e um sistema de monitoramento efetivo para avaliação dos riscos das secas (Figura 2.21).

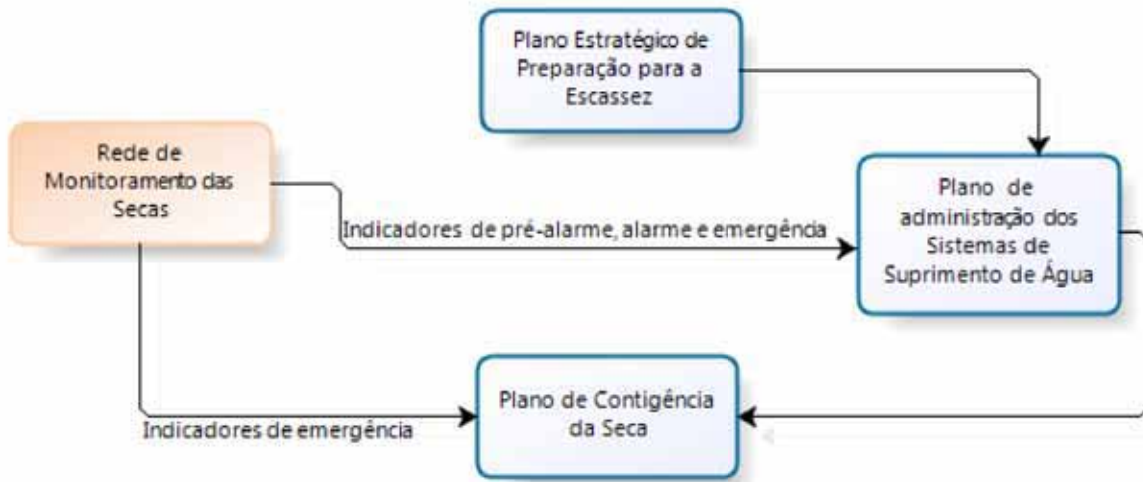


Figura 2.21 Processo de planejamento e gestão dos riscos das secas. Fonte: Adaptado de ROSSI et al. [142]

Esses dois últimos modelos caracterizam uma visão sistêmica e pró-ativa para enfrentamento dos problemas decorrentes das secas e constituem uma realidade ainda não presente no semi-árido nordestino, inclusive no Rio Grande do Norte.

Neste trabalho de tese para comparação e análise de desempenho com o ISFS foram escolhidos os índices RAI, Quantis, DI e RI (Quadro 2.9). A seguir é apresentada uma descrição mais detalhada das características e da forma de cálculo de cada um desses índices.

O RAI computa a anomalia de precipitação em relação à normal de precipitação de um ponto específico ou de uma região [38, 70, 188]. É obtido a partir das seguintes equações:

$$RAI = 3 \times \left[\frac{N - \bar{N}}{\bar{M} - \bar{N}} \right], \quad \text{para anomalias positivas} \quad (2.1)$$

$$RAI = -3 \times \left[\frac{N - \bar{N}}{\bar{X} - \bar{N}} \right], \quad \text{para anomalias negativas} \quad (2.2)$$

onde,

N = Precipitação mensal atual;

$\bar{N}$ = Precipitação média mensal da série histórica;

$\bar{M}$ = Média das dez maiores precipitações mensais históricas;

$\bar{X}$ = Média das dez menores precipitações mensais históricas;

Logo, se $N \geq \bar{N}$ utiliza-se a equação 2.1, caso contrário a equação 2.2.

Seus pontos positivos são a dupla aptidão de cálculo (para um ponto específico ou para uma região), a facilidade de cálculo e a sua efetividade na detecção de períodos de seca. Quando

comparado a índices mais complexos como o PDSI e o BMDI, as diferenças encontradas são consideradas desprezíveis (OLADIPO, 1985 *apud* YUREKLI & ANLI [188]). A limitação se dá pelo fato de utilizar somente dados de chuvas [161].

O método dos quantis teve sua origem no trabalho pioneiro de Pinkayan [132], em 1966, que investigou a probabilidade de ocorrência de anos chuvosos e secos em diversas áreas do oeste dos Estados Unidos, utilizando uma abordagem estocástica, através da classificação dos anos dentro de cinco categorias, em faixas de quantis: Muito Seco - MS ($0 < F(x) \leq 0,15$); Seco - S ($0,15 < F(x) \leq 0,35$); Normal - N ($0,35 < F(x) \leq 0,65$); Chuvoso - C ($0,65 < F(x) \leq 0,85$) e Muito Chuvoso - MC ($0,85 < F(x) \leq 1,0$), com $F(x)$ sendo a função de distribuição de probabilidade da precipitação anual. O objetivo desse estudo foi responder a duas questões: 1) Se existia um padrão sistemático de distribuição de anos secos e chuvosos nas áreas em questão; e 2) Admitindo-se que esse padrão fosse sistemático, que tipo de regularidade ele seguia.

Esse método apresenta os seguintes pontos positivos: extrema simplicidade de cálculo; interpretação precisa em termos probabilísticos; podem ser calculados ou estimados diretamente de dados empíricos; e, o emprego é seguro e é imune a qualquer assimetria na função de densidade de probabilidade [185]. O ponto negativo é o fato de utilizar somente dados de chuvas [161].

O DI é calculado classificando-se o total de precipitação mensal do registro histórico do menor ao maior valor, para assim, construir uma distribuição de frequência acumulada. Depois essa distribuição é dividida em 10 partes ou decis. O primeiro decil corresponde ao valor de precipitação não excedido pelas 10% menores precipitações de todo o registro. O segundo corresponde ao valor de precipitação não excedido pelas 20% menores precipitações e assim por diante. Para composição usual desse índice utilizam-se cinco classes: “Muito abaixo da normal” (decis 1 e 2), “Abaixo da normal” (decis 3 e 4), “Próximo à normal” (decis 5 e 6), “Acima da normal” (decis 7 e 8) e “Muito acima da normal” (decis 9 e 10) [38, 85, 161, 162, 185].

Esse índice apresenta os mesmos pontos positivos e negativos dos quantis.

O RI é baseado em um procedimento de normalização dos desvios médios da precipitação de diversos postos pluviométricos de uma determinada região [70, 97, 102]. É obtido a partir da seguinte equação:

$$RI_y = \frac{1}{n} \sum_{s=1}^n S_{sy} \quad (2.3)$$

onde,

$$S_{sy} = \frac{R_{sy} - \overline{R_s}}{\sigma} \quad (2.4)$$

com,

R_{sy} = Precipitação anual da estação ou posto pluviométrico s para o ano y ;

$\overline{R_s}$ = Média da precipitação anual da estação ou posto pluviométrico s ;

σ = Desvio padrão da precipitação da estação ou posto pluviométrico s ;

n = Número de estações ou postos pluviométricos com informações para o ano y ;

Uma vantagem capital desse índice é que ele admite série de dados com falhas, principalmente por ter aplicação regional [70]. A principal limitação deve-se ao fato de utilizar somente dados de chuvas, que é a mesma encontrada nos três índices anteriores [161].

2.5.3 Principais Índices de Severidade de Seca Utilizados no Brasil, Nordeste e Rio Grande do Norte

No Brasil, alguns trabalhos têm sido desenvolvidos principalmente com a utilização dos índices meteorológicos e hidrológicos [28–30,60,115,149,158,165,166]. No Nordeste, a literatura apresenta alguns trabalhos com a utilização desses índices, em Alagoas [100], Pernambuco [150], Paraíba [13,96,141], Ceará [8,20,70,185] e Rio Grande do Norte [11,42].

LIMA et al. [100] utilizaram o PDSI e outros índices para estudar as secas em Alagoas e SANTOS & ANJOS [150] utilizaram o SPI para monitorar as secas em Pernambuco.

Na Paraíba, AZEVEDO & SILVA [13] utilizaram RAI, BMDI e PDSI para caracterizar o grau de severidade de seca e o teor de umidade do solo na microrregião do Agreste da Borborema, RODRIGUES DA SILVA & AZEVEDO [141] o BMDI para cada microrregião do Estado e KUMAR & SOUZA [96] o SPI para estudar as secas na cidade de Campina Grande.

No Ceará, XAVIER & XAVIER [185] utilizaram o método dos quantis para a caracterização de períodos secos ou extremamente chuvosos, considerando, além da precipitação anual, a quadra chuvosa, quadrimestre de maior pluviosidade no ano, e a estação chuvosa, semestre de maior pluviosidade no ano, além de reavaliarem anos historicamente considerados muito secos ou muito chuvosos. ALVES et al. [8] utilizaram também esse método na verificação de prognósticos sazonais de precipitação também para o Ceará. BARRA et al. [20] caracterizaram as secas de 21 localidades do Ceará utilizando o PDSI e FREITAS [70] desenvolveu um sistema de suporte a decisão com base no RAI, BMDI, HSI e RI.

No ano de 2000, a Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME) em conjunto com a Universidade Federal de Santa Catarina (UFSC), devolveram o Índice FUNCEME de Severidade de Seca (IFS) também chamado de Indicador de Emergência (IMERG), que correlaciona aspectos de natureza física, climatológica, sócio-econômica e fatores de risco

associados à perda da produção agrícola e à falta de água para consumo humano e animal [16,64,157].

O IFS foi construído como protótipo para ser aperfeiçoado e calibrado pela FUNCEME [16], e até agora, 2010, não o utilizaram de forma sistemática no Ceará. Como foi dito, esse índice situa-se na terceira categoria da classificação de BYUN & WILHITE [38] e foi o modelo inspirador do índice proposto nessa tese. O processo de sua construção e os detalhes das suas principais características serão expostas no item 2.6.

No Rio Grande do Norte, ARAÚJO FILHO [11] utilizou o PDSI para a caracterização das secas em regiões homogêneas do Estado com dados de 137 localidades para o período de 1910 a 1991 e CAMPOS et al. [42] analisaram a variação espaço-temporal do PDSI em todo o Estado. Outros trabalhos também foram realizados, porém, utilizando-se índices de aridez [32,33,163,164], que diferem dos índices de seca, por se restringirem a regiões com baixa quantidade de chuvas, e pelo fato da aridez, diferentemente da seca, ser uma característica permanente do clima [35].

O monitoramento das secas no Brasil, embora ainda em um estágio inicial, é uma atividade que tem evoluído nos últimos anos, em grande parte impulsionada pelos avanços na área de monitoramento e previsão de tempo e clima no País [12,34,129]. Algumas ações estão sendo desenvolvidas para minimizar os efeitos danosos das secas, como por exemplo: o zoneamento de risco climático, coordenado pelo Ministério da Agricultura e com abrangência em todo o País [34]; e o Sistema de Monitoramento Agrometeorológico (Agritempo), criado pelo Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (CEPAGRI/UNICAMP) e a Embrapa Informática. Ele fornece, via internet, informações meteorológicas e agrometeorológicas de diversos municípios e estados brasileiros [129].

Dentre os principais produtos oferecidos pelo Agritempo estão: boletins agrometeorológicos, mapas de monitoramento climático e de previsão, mapas de índices de seca, sumários e mapas de séries históricas, produtos do zoneamento agrícola e mapas de probabilidade de precipitação [129]. O índice de seca utilizado no Agritempo é o SPI com escala de tempo de 5 dias, 10 dias, 20 dias, 30 dias, 1 mês, 2 meses, 3 meses e 6 meses. Essa informação está disponível também para o Estado do Rio Grande do Norte.

Em 2004, o Governo do Estado do Ceará desenvolveu um projeto piloto coordenado pelo Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), cujo objetivo era orientar os tomadores de decisão no âmbito do Grupo Interinstitucional Permanente de Convivência e Desenvolvimento Sustentável do Semi-árido no Ceará a partir da utilização do Índice Municipal de Alerta (IMA). O propósito era estabelecer uma estrutura destinada a atender as populações

afetadas pelas irregularidades climáticas e a instabilidade econômica e social nas localidades afetadas por esses eventos, congregando diversas instituições, principalmente quanto ao fornecimento das informações básicas para a composição do IMA. O trabalho consolidou-se em 2005, mediante o uso do IMA em ações desenvolvidas pelos diversos órgãos do Ceará, no sentido de reduzir a vulnerabilidade climática e melhorar a convivência do homem do campo com o semi-árido.

O IMA é um indicador composto, síntese de doze outros indicadores, utilizado como instrumento de monitoramento socioeconômico e ambiental para os municípios do Ceará. Os indicadores utilizados são: produtividade agrícola por hectare, produção agrícola por habitante, utilização da área colhida com culturas de subsistência, perda de safra, proporção de famílias beneficiadas com bolsa família, número de vagas do seguro safra por 100 habitantes rurais, climatologia, desvio normalizado das chuvas, escoamento superficial, índice de distribuição de chuvas, índice de aridez e taxa de cobertura de abastecimento urbano de água. É obtido a partir do somatório da padronização desses indicadores multiplicado pelo peso de $\frac{1}{12}$ [91,92].

Uma visão mais abrangente dessa temática é dada no Programa de Ação Nacional de Combate à Desertificação e Mitigação dos Efeitos da Seca (PAN-Brasil), que foi formulado, em 2004, com a colaboração de instituições governamentais (da esferal federal e das esferas estaduais) e de instituições não-governamentais; e pelos Pontos Focais Estaduais, que constituem a célula operativa a nível estadual desse plano [24,31,44].

Os Pontos Focais Estaduais apresentam três linhas de ações: Legislativo, Governamental (executivo) e Sociedade Civil. Dentre as suas funções estão a coordenação no desenho, execução, avaliação e o monitoramento dos Programas Estaduais de Combate à Desertificação (PAE). Esses programas visam implementar, a nível estadual, as diversas ações desenvolvidas no PAN-Brasil, destacando-se o monitoramento e a avaliação da desertificação e a implementação de um sistema de alarme precoce de eventos críticos de seca a partir de indicadores.

O PAN-Brasil reconhece que, apesar dos avanços dos modelos de previsão de tempo e clima e de monitoramento, incluindo os índices de severidade de seca, muito ainda há de se fazer para melhorar a capacidade e precisão desses modelos e índices [31].

No Rio Grande do Norte, o monitoramento das secas é uma atividade ainda incipiente, restringindo-se basicamente a três ações: o monitoramento e previsão de tempo e clima realizado pela EMPARN, o monitoramento do nível dos principais açudes feito pela Secretaria de Recursos Hídricos e o monitoramento da qualidade da água superficial e subterrânea coordenado pelo Instituto de Gestão das Águas do Rio Grande do Norte (IGARN).

O principal índice utilizado pela EMPARN no monitoramento do déficit pluviométrico no

Estado é o desvio em relação à normal pluviométrica. Também são elaborados boletins diários com a previsão de tempo para até 48 horas e da precipitação pluviométrica ocorrida nas últimas 24 horas [22].

As informações advindas desses órgãos são fornecidas às outras esferas do governo estadual, auxiliando no apoio à decisão para adoção de medidas necessárias na mitigação dos efeitos das secas. Não há ainda um programa de monitoramento que utilize índices de severidade de seca mais sofisticados, operacionais e que possibilitem diagnosticar a gravidade do fenômeno da seca nas diversas regiões do RN, nem um planejamento antecipado, integrado e sistemático para atuação desses órgãos quando da ocorrência de uma grande seca.

Portanto, ainda existe um descompasso muito grande entre o monitoramento e ações efetivas de mitigação dos efeitos das secas no Brasil, no Nordeste e, conseqüentemente, no Rio Grande do Norte. Isto porque a maioria das agências governamentais só age após a seca estar caracterizada [34]. Um exemplo disso, já citado, foi o ocorrido na última grande seca que atingiu o Nordeste, em 1998, quando a atuação governamental para atendimento à população flagelada foi atrasada e deficitária [59]. Um outro ponto é a ausência de uma legislação específica que trate do problema em todas as suas dimensões e que congregue os diversos órgãos governamentais de forma sistemática, cada um na sua competência e responsabilidade, com um planejamento antecipado e uma gestão pró-ativa de risco dos eventos de secas na região.

O Índice de Susceptibilidade ao Fenômeno da Seca (ISFS), proposto neste trabalho, pode contribuir fornecendo um diagnóstico mais preciso da situação e da vulnerabilidade dos municípios durante os eventos de seca. O modelo foi desenvolvido para ser aplicado em todo semi-árido nordestino, embora tenha sido validado com dados do Rio Grande do Norte.

Em uma análise comparativa com os índices de seca atualmente utilizados no RN, o SPI e o desvio em relação a normal pluviométrica, o ISFS apresenta um substancial diferencial por considerar aspectos de natureza física, climatológica, sócio-econômica e de abastecimento d'água dos municípios.

No tópico a seguir será detalhado o processo de construção e as características principais do IFS, que inspirou a criação do índice desenvolvido neste trabalho de tese, conforme já foi citado anteriormente.

2.6 Índice FUNCEME de Severidade de Seca (IFS) ou Indicador de Emergência (IMERG)

2.6.1 Histórico

O Governo do Ceará, através da FUNCEME, e a UFSC desenvolveram, no ano de 2000, o Índice FUNCEME de Severidade de Seca (IFS). Ele surgiu da necessidade de fornecer informações mais precisas e coerentes sobre a vulnerabilidade das diversas regiões do Estado e sua relação com o fenômeno das secas [64,72]. O índice foi construído em duas etapas, a primeira no começo do ano de 2000, que resultou em um protótipo (Figuras 2.22 a 2.24), e a segunda no final desse ano, resultando em um modelo mais reduzido, mais viável e mais adequado a realidade do Estado (Figura 2.25). Em um trabalho posterior, no qual descrevem o processo de construção e as características desse índice, SILVA & FORTE [157] o denominaram de Indicador de Emergência (IMERG), denominação adotada também nesse trabalho de tese.

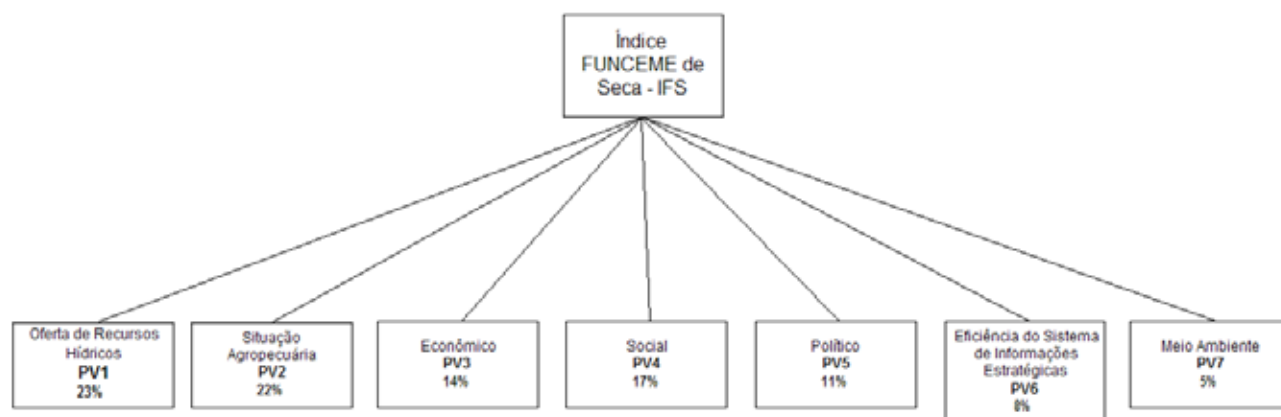


Figura 2.22 Árvore dos Aspectos Fundamentais (IFS). Fonte: ENSSLIN et al. [64]

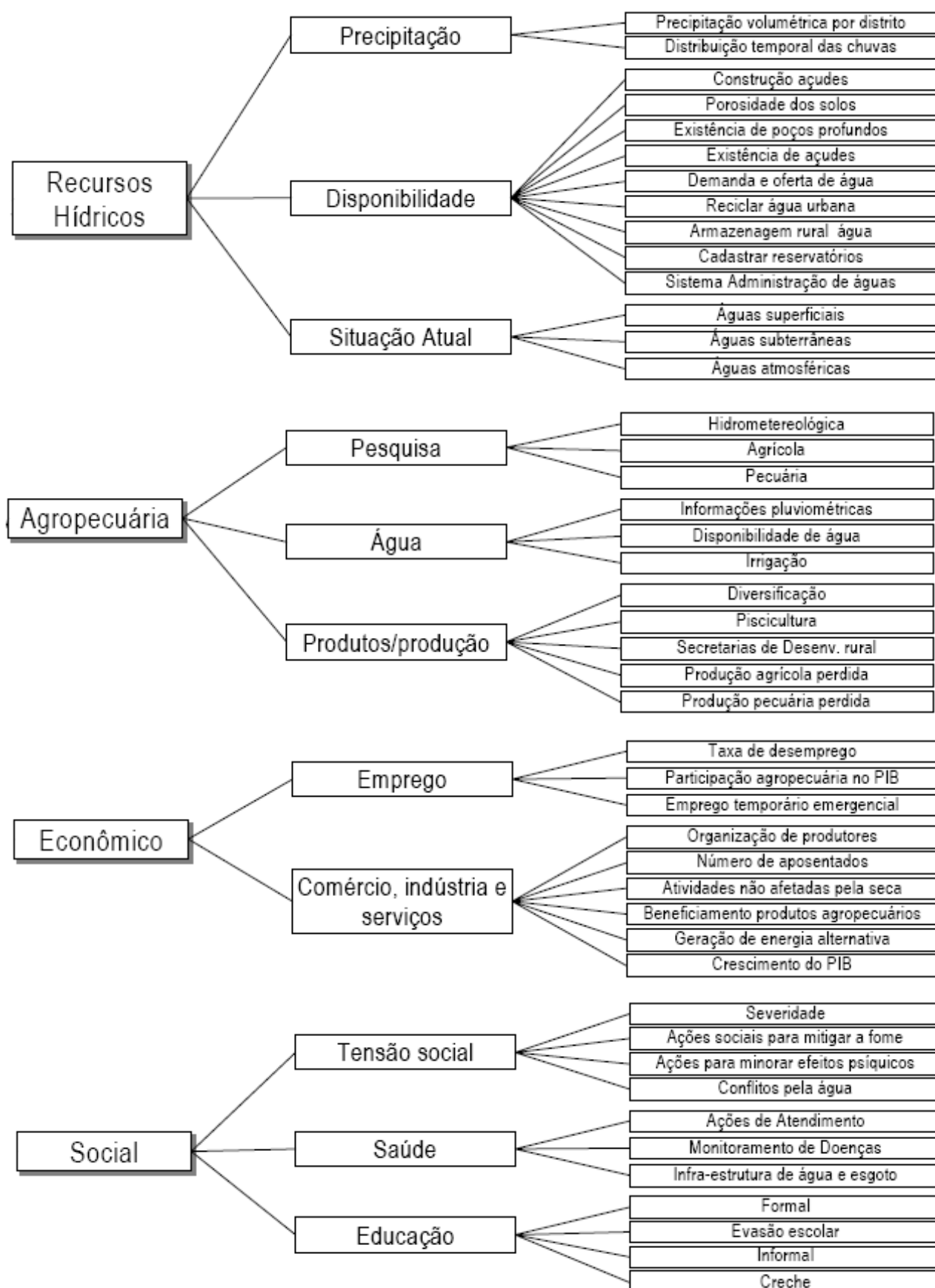


Figura 2.23 Detalhe dos Parâmetros PV1, PV2, PV3 e PV4 (IFS). Fonte: ENSSLIN et al. [64]

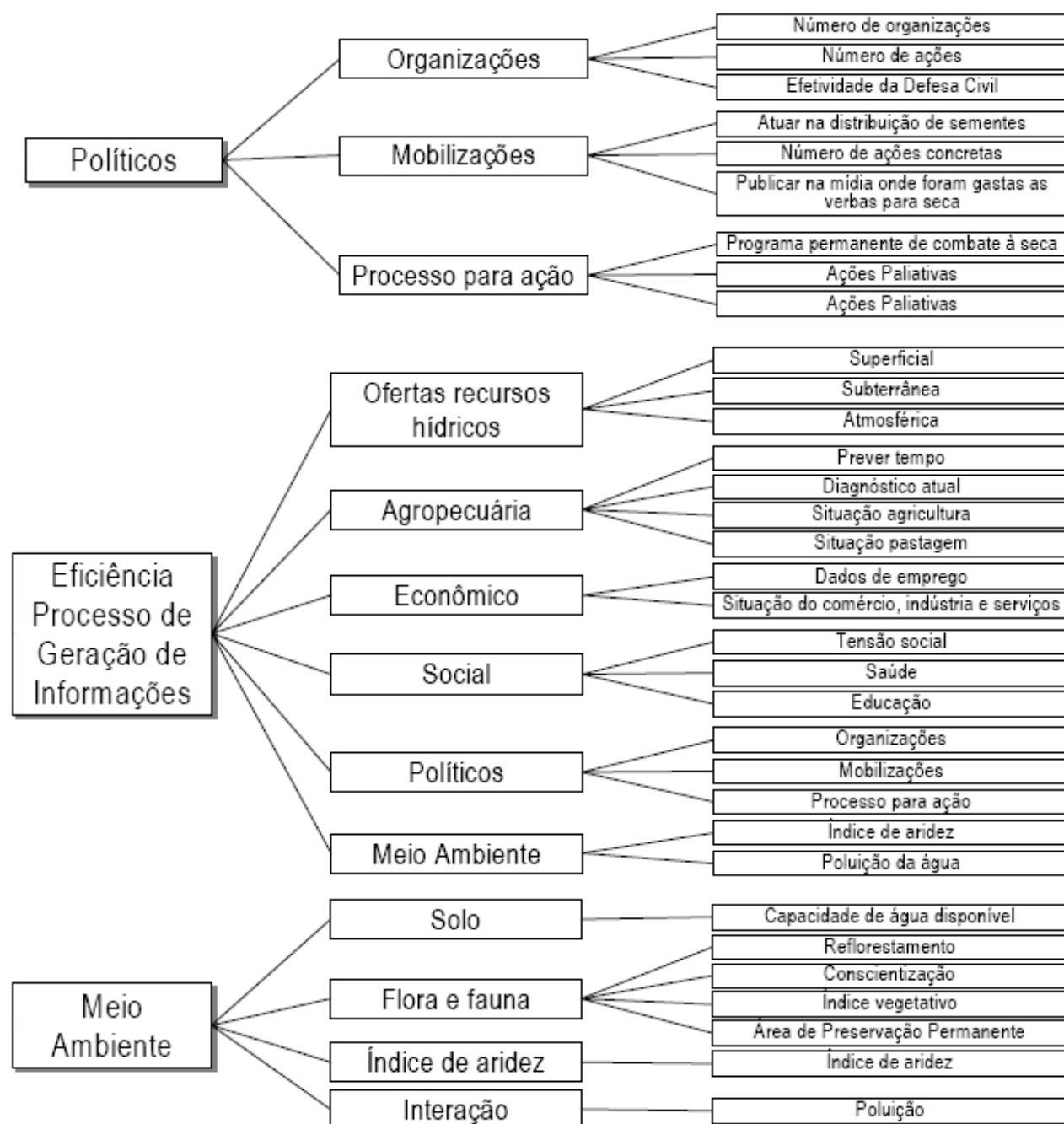


Figura 2.24 Detalhe dos Parâmetros PV5, PV6 e PV7 (IFS). Fonte: ENSSLIN et al. [64]

O protótipo do índice foi construído em três etapas. Na primeira etapa, foi inicialmente definida a equipe de decisores, composta de especialistas e técnicos da própria FUNCEME e outros órgãos do Estado, e esses decisores foram solicitados a expor por email o seu conhecimento e preocupações a respeito da problemática da seca. Na segunda etapa, os decisores foram estimulados a apresentar sugestões para resolver as questões levantadas na etapa anterior, resultando em um conjunto de áreas prioritárias de ação. E, na terceira etapa, foi realizada

uma Conferência de Decisão (*Decision Conference*), onde foram debatidos e expandidos os conceitos anteriores, uniformizando-se os conhecimentos dos integrantes do grupo com respeito ao problema de como avaliar a severidade das secas no Estado do Ceará. Desse trabalho resultou: 1) Indentificação dos Pontos de Vistas Fundamentais (PVs); 2) Construção dos descritores; 3) Construção das funções de valor; 4) Indentificação das taxas de compensações do modelo; 5) Construção do modelo; 6) Testes e validação do modelo; e, 7) Calibração do modelo. A calibração do modelo se deu com dados de 4 municípios do Estado, sendo que dois deles (Irauçuba e Aiuaba) tinham sofrido com a seca no último ano e dois deles não (Crateús e Granja) [64,72].

O segundo modelo foi desenvolvido em quatro etapas: 1^a) Realização de uma entrevista, semi-estruturada com técnicos da Defesa Civil do Ceará, utilizando-se uma técnica de mapeamento cognitivo para obter as percepções e valores envolvidos na sistemática utilizada quando da decretação de um estado emergencial nos municípios do Ceará; 2^a) Entrevista com técnicos da Empresa de Extensão Rural do Ceará (EMATERCE), Secretaria dos Recursos Hídricos do Estado do Ceará (SRH), Companhia de Gestão dos Recursos Hídricos do Estado do Ceará (COGERH) e da Fundação Cearense de Meteorologia (FUNCEME) buscando-se inserir no sistema a visão dos segmentos: recursos hídricos, meteorologia e agropecuária; 3^a) Aplicação de um questionário via internet com 90 técnicos que atuavam na coordenação de programas emergenciais no Ceará, com o objetivo de ampliar a percepção do mapa cognitivo desenvolvido nas etapas anteriores; e, 4^a) Realização de uma Conferência de Decisão, em dezembro de 2000, quando se reestruturou e obteve-se o modelo consensual final do IMERG (Figura 2.25). Os percentuais em destaque (vermelho) na Figura 2.25, são o pesos ou níveis de importância atribuídos pelos decisores a cada PV.

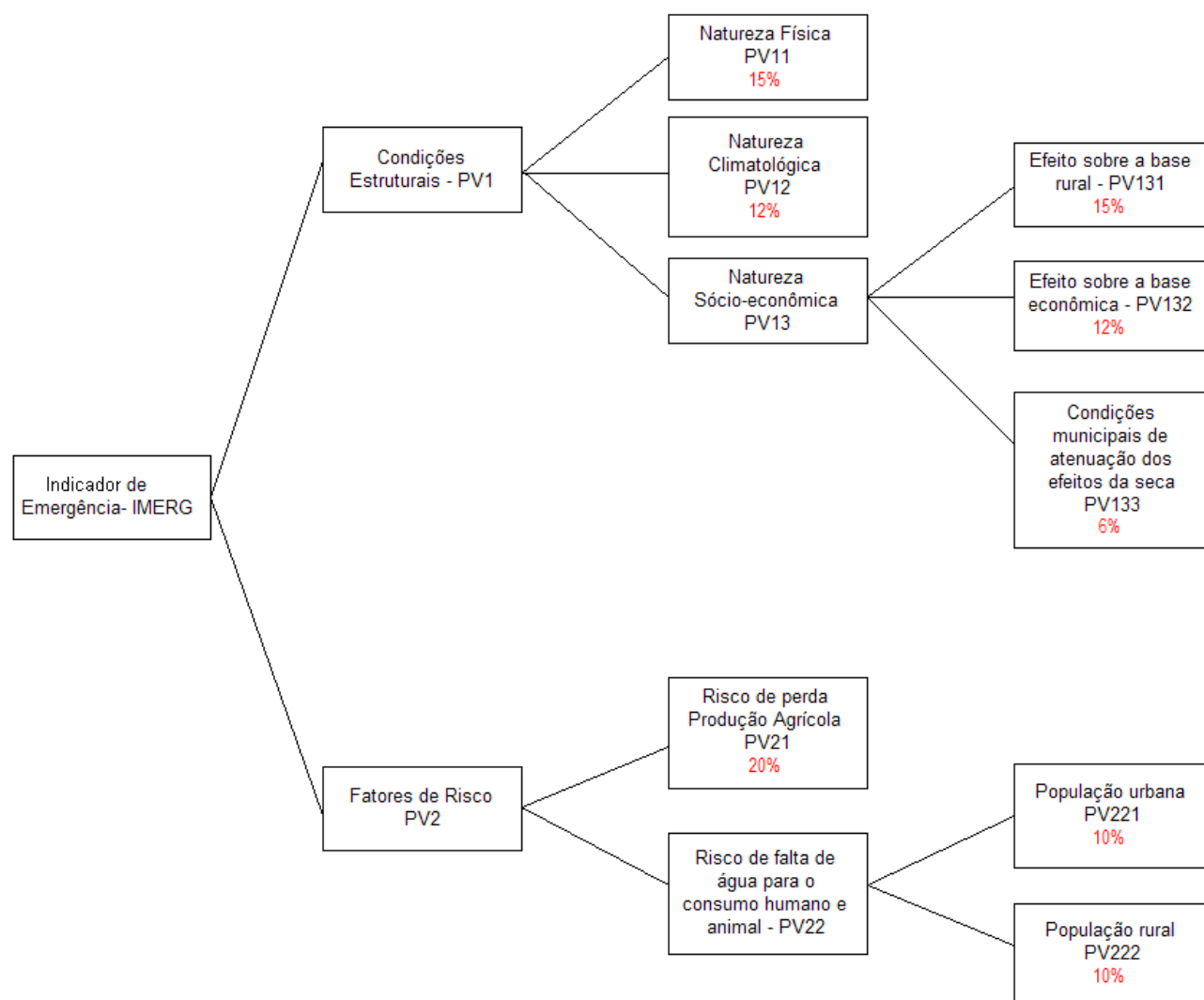


Figura 2.25 Árvore dos Aspectos Fundamentais (IMERG) - Fonte: SILVA & FORTE [157].

2.6.2 Metodologia Empregada

Na construção do IMERG foi utilizada a metodologia multicritério de apoio à decisão *Multicriteria Decision Aiding* (MCDA).

Em geral, o apoio à decisão se dá em um contexto de múltiplos critérios, mesmo quando a decisão cabe a um único decisor. Essa abordagem elimina alguns problemas característicos da abordagem monocritério, quanto à negligência de certos aspectos da realidade, da natureza fictícia que permanecem invisíveis e da tendência de apresentar características de um sistema de valores particular como objetivo [144].

Quando se trabalha com MCDA, três conceitos básicos assumem um papel fundamental para a análise e estruturação do processo de decisão [144]:

O primeiro são as ações potenciais ou alternativas. Entende-se como ação potencial toda

ação que é julgada possível de implementar, e alternativa a ação potencial que de maneira alguma pode ser colocada em operação conjuntamente com outra.

O segundo conceito são os critérios ou famílias de critérios. Um critério g é uma ferramenta construída para avaliar e comparar ações potenciais de acordo com um ponto de vista bem definido. Define-se como X_g o conjunto de todas as possíveis avaliações que este critério pode conduzir e $(<_g, X_g)$ a escala do critério g , com $<_g$ uma ordem completa em X_g . Elementos $x \in X_g$ são chamados graus ou escores da escala. Diversos tipos de escalas são utilizadas para comparar as ações potenciais, como por exemplo: escalas puramente ordinais (verbal ou numérica), escalas quantitativas, entre outras.

O terceiro conceito a se considerar diz respeito à problemática de decisão. No apoio à decisão o resultado pretendido em determinado problema pode ser identificado entre quatro tipos de problemática de referência [74,143]:

1) Problemática de Descrição ($P.\delta$) - Objetiva esclarecer a decisão pela elaboração de qualquer prescrição ou recomendação. O resultado pretendido é uma descrição ou procedimento cognitivo.

2) Problemática de Escolha ($P.\alpha$) - É orientada para esclarecer a decisão para a escolha de um subconjunto tão restrito quanto possível, tendo em vista a escolha final de uma única ação. O resultado pretendido é uma escolha ou um procedimento de seleção.

3) Problemática de Classificação ($P.\beta$) - Objetiva esclarecer a decisão pela atribuição de cada ação a uma categoria (ou classe) predefinida. O resultado pretendido é uma triagem ou um procedimento de classificação.

4) Problemática de Ordenação ($P.\gamma$) - Objetiva esclarecer a decisão pelo reagrupamento de todas ou parte das ações em classes de equivalência. O resultado pretendido é um arranjo ou procedimento de ordenação.

O desenvolvimento do IMERG é um típico problema pertencente a problemática $P.\gamma$, pois tinha-se como objetivo principal o agrupamento e ordenação dos municípios do Estado do Ceará em classes de vulnerabilidade em relação à seca.

De uma maneira geral, pode-se definir um problema de decisão multicritério da seguinte forma [74,95]:

Seja,

G = Objetivo e/ou objetivos a serem alcançados;

F_D = Função multicritério de decisão;

$A = \{a_1, a_2, \dots, a_M\}$, conjunto das alternativas que serão avaliadas, comparadas, analisadas e/ou priorizadas;

$I = \{i\}_{i=1,N}$, conjunto de índices dos critérios;

$J = \{j\}_{j=1,m}$, conjunto de índices dos atributos;

$K = \{k\}_{k=1,K}$, conjunto de índices dos membros do grupo de decisão;

$C = \{c_i\}$, conjunto de critérios a serem usados para avaliação das alternativas;

$O = \{o_j\}$, conjunto de atributos a serem usados para avaliação das alternativas;

$U = \{u_i^k\}$, onde u_i representa a informação individual do k -ésimo decisor pela consequência da seleção de uma alternativa quando o critério i é dado;

$W = \{w_1, \dots, w_k\}$, conjunto de pesos atribuídos aos critérios ou atributos, onde w_k é função da preferência do decisor por c_i ou o_j .

R = Restrições a serem observadas;

Ω = Possíveis estados da natureza ou possíveis eventos;

$p(\Omega)$ = Probabilidade ou possibilidade associada ao evento Ω ;

$H(A, \Omega, p(\Omega))$ = Benefícios esperados da alternativa.

$$\Rightarrow F_D = G \otimes [A \otimes R] \otimes [W \otimes C \otimes O] \otimes [\Omega \otimes H \otimes p(\Omega)] \otimes [K \otimes U] \quad (2.5)$$

onde, $\otimes$ significa “associado e/ou sob influência”, e, a relação $A \otimes R$, busca eliminar as alternativas que não superam as restrições; $W \otimes C \otimes O$ modela as preferências do decisor; $\Omega \otimes H \otimes p(\Omega)$ quantifica a incerteza e a agregação de valor das alternativas; e, $K \otimes U$ gera um valor maior que a simples soma das partes, uma vez que ocorre a interação do grupo.

Na MCDA o apoio à decisão se desenvolve em duas fases: a primeira trata da construção dos critérios de avaliação que serão usados no modelo e a segunda dos parâmetros que serão usados para agregação desses critérios de forma a auxiliar a decisão [63,64].

Na primeira fase, que corresponde a estruturação do problema, propriamente dita, são definidos os atores da decisão, mapas cognitivos, Pontos de Vista Fundamentais (PV's) e os descritores de impactos para cada PV.

Os atores da decisão englobam os decisores, facilitadores e analistas. O decisor é uma pessoa ou um grupo de pessoas responsáveis pela tomada de decisão. O facilitador é um líder experiente que deve focalizar a sua atenção na resolução do problema coordenando os pontos de vistas dos decisores. E o analista auxilia o facilitador e os decisores na estruturação do problema consistindo a maior parte do seu trabalho da formulação do problema [74]. Os mapas cognitivos são representações gráficas das interpretações mentais que um consultor faz das discussões

feitas com os decisores, tendo como resultado final a identificação dos PV's e da árvore de PV's (Árvore de Valor) (exemplo: Figuras 2.22 e 2.23). Os PV's identificados na árvore de valor são as variáveis (dimensões, critérios ou atributos) considerados como relevantes para resolução do problema. A operacionalização dos PV's se dá através de funções denominadas de descritores de impactos das ações. Os descritores medem o grau em que um determinado objetivo do decisor é alcançado e podem ser quantitativo ou qualitativo, discreto ou contínuo, direto, indireto ou construído [63,64]. A construção do mapa cognitivo, dos PV's e dos descritores são atividades recursivas [63].

Na segunda fase são definidos os parâmetros que permitem a obtenção de informações quantitativas a respeito das preferências dos decisores. Esta fase envolve a escolha do método ou abordagem de agregação, a construção das funções de valor, a identificação das taxas de compensação ou pesos e a avaliação dos impactos das ações potenciais de todos os PV's a partir de uma análise de sensibilidade [63,64,157].

Os métodos de auxílio à decisão mais frequentemente utilizados usam procedimentos de agregação multicritério explicitamente matemáticos. Tais métodos advêm de dois tipos de abordagens principais: a abordagem baseada em um critério de síntese e a abordagem baseada em um sistema relacional de preferências.

A primeira abordagem é a mais tradicional sendo caracterizada, mais frequentemente, por uma fórmula matemática que conduz a uma definição explícita de um único critério, síntese dos n critérios considerados no problema. Exemplos dessa abordagem são encontrados nos métodos MAUT (*Multiple Attribute Utility Theory*), SMARTS (*Simple Multiattribute Rating Technique using Swings*), TOPSIS (*Technique for Order Preference by Similarity to Ideal Solution*), MACBETH (*Measuring Attractiveness by a Categorical Based Evaluation Technique*), AHP (*Analytic Hierarchy Process*), entre outros. Essa abordagem não permite qualquer incomparabilidade entre as alternativas.

A segunda abordagem não trabalha com cada ação potencial separadamente das outras, mas sucessivamente compara-as duas a duas. São exemplos dessa abordagem os métodos de sobreclassificação (*outranking methods*), tais como: PROMETHE (*Preference Ranking Organisation Method for Enrichment Evaluations*), ELECTRE (*Elimination and Choice Translating Algorithm*) e suas derivações, entre outros. Essa abordagem já possibilita que se considere incomparabilidades entre alternativas [144].

Na construção do IMERG, o método de agregação escolhido foi o baseado em um critério único de síntese. Para tal foi utilizado o MACBETH na obtenção das funções de valor dos descritores e na determinação das taxas de substituição ou pesos do modelo [64,157].

O MACBETH é uma abordagem de análise de decisão multicritério que requer somente julgamentos qualitativos sobre diferenças de valores, para auxiliar um indivíduo ou um grupo a quantificar a relativa atratividade das opções [17]. O método possibilita a construção de funções de valor intracritérios e, por meio de uma análise intercritério agregar os diversos critérios de avaliação em um critério único de síntese, pela atribuição de pesos aos critérios [15, 75, 76].

Para construir uma função de valor intracritério, as preferências são representadas por uma função v e por funções limiares s_k , de modo que $a P^{(k)} b \Leftrightarrow s_k < v(a) - v(b) < s_{k+1}$, em que $P^{(k)}$ representa uma relação de preferência tanto mais forte quanto k for grande. A partir da comparação par a par da atratividade das alternativas o método propõe que o decisor exprima os seus juízos absolutos de diferença de atratividade, de acordo com a escala semântica dada pelo Quadro 2.11 abaixo [15, 17, 63, 64, 75, 76].

Categoria	Descrição
C_0	Nenhuma diferença de atratividade (indiferença)
C_1	Diferença de atratividade muito fraca
C_2	Diferença de atratividade fraca
C_3	Diferença de atratividade moderada
C_4	Diferença de atratividade forte
C_5	Diferença de atratividade muito forte
C_6	Diferença de atratividade extrema

Quadro 2.11 Esquema de julgamento do MACBETH. Fonte: BANA E COSTA & VANSNICK [15].

Com base nessas informações constrói-se uma matriz denominada Matriz Semântica, que contém esquematicamente as respostas do(s) decisor(es). Com esse conjunto de julgamentos o MACBETH é executado, primeiramente para a verificação das eventuais inconsistências, e depois para a determinação de uma escala de valor cardinal que represente os julgamentos de valor do decisor. A escala obtida é então normalizada, fornecendo os valores dos pesos para as alternativas em avaliação, e, possibilitando assim, o uso do modelo de agregação, em geral, o modelo aditivo. O processo é iterativo, consistindo basicamente na resolução de quatro problemas de programação linear sequenciais (Mc1, Mc2, Mc3 e Mc4) (Figura 2.26). O software que implementa computacionalmente o método ainda possibilita ao decisor ajustar graficamente os valores atribuídos, dentro dos intervalos permitidos (análise de sensibilidade). Somente após este ajuste, com a introdução dos conhecimentos dos especialistas, é que se obtém a escala cardinal final de valores [15, 64, 75, 76].

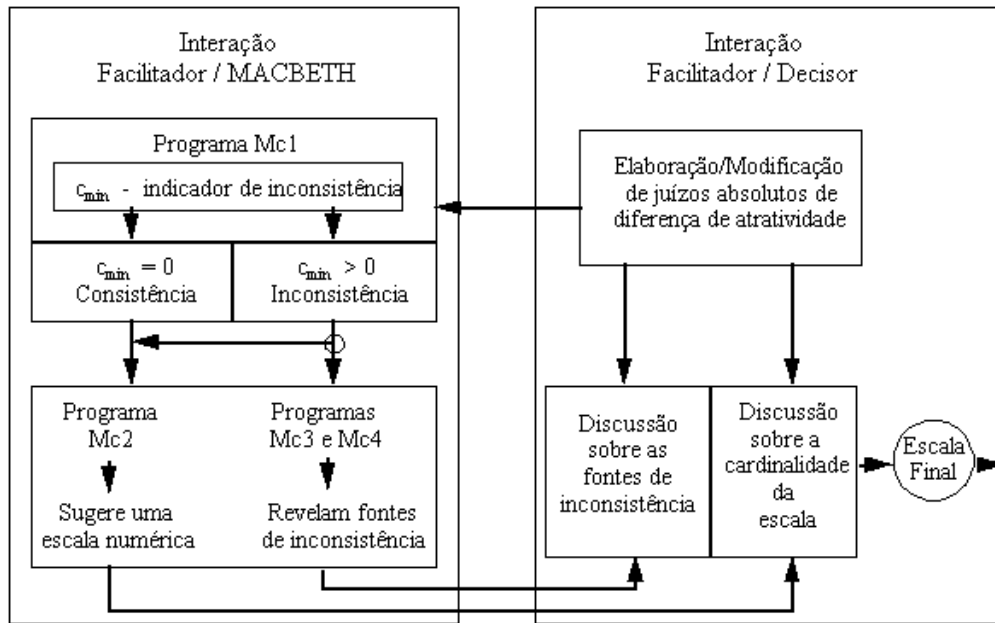


Figura 2.26 Esquema Interativo do MACBETH - Fonte: BANA E COSTA & VANSNICK [15].

No tópico a seguir são descritas as principais características do segundo protótipo de índice desenvolvido pela FUNCEME e que foi a base do modelo de índice proposto neste trabalho de tese.

2.6.3 Características do IMERG

O IMERG foi construído com sete PVs operacionais, que, de acordo com os especialistas da equipe de decisão, são características determinantes para medir as vulnerabilidades dos municípios do Ceará ao fenômeno da seca. Esses PVs são agregados a dois outros PVs principais: “Condições Estruturais” e “Fatores de risco” (ver Figura 2.25). As “Condições Estruturais” apresentam a situação física, pluviométrica, demográfica e financeira encontrada nos municípios, e os “Fatores de Risco” quantificam os riscos da perda de safras em função da distribuição de chuvas, e, os riscos associados à falta de água para as populações humana e animal [16,157].

O modelo de agregação utilizado para obtenção do índice de vulnerabilidade final de cada município foi o aditivo. A expressão para cálculo foi definida como:

$$IMERG_m = \sum_{i=1}^7 k_i \times v_i(m) \quad (2.6)$$

onde:

$IMERG_m$ = Índice de vulnerabilidade à seca do município m .

k_i = Peso do i -ésimo PV ($k_1 = 0,15$; $k_2 = 0,12$; $k_3 = 0,15$; $k_4 = 0,12$; $k_5 = 0,06$; $k_6 = 0,2$; $k_7 = 0,2$).

$v_i(m)$ = vulnerabilidade à seca do município com respeito ao i -ésimo PV.

em que:

$$\sum_{i=1}^7 k_i = 1 \quad e \quad \begin{cases} (IMERG_m)_{max} = 100 \\ (IMERG_m)_{min} = 0 \\ (v_i(m))_{max} = 100, \forall i \\ (v_i(m))_{min} = 0, \forall i \end{cases}$$

No Quadro 2.12 abaixo, são apresentados os limites de variação dos descritores definidos para cada um dos PVs constitutivos do IMERG.

PV	Descritor	Limites de Variação	
		Vulnerabilidade Máxima	Vulnerabilidade Mínima
PV11	Classe de capacidade de uso do solo	Classe 6	Classe 1
PV12	Média pluviométrica anual (30 anos)	2000 mm	500 mm
PV131	População rural residente no município	40.000	0
PV132	Participação da Agropecuária no PIB do município	50%	0%
PV133	Receita <i>per capita</i> recebida pelo município	R\$ 900/ano	R\$ 150/ano
PV21	Índice de distribuição de chuvas	1000	0
PV221	Risco de falta d'água para a população urbana	Não-Satisfatório	Satisfatório
PV222	Risco de falta d'água para a população rural	Nível 5	Nível 0

Quadro 2.12 Limites dos descritores definidos para os PVs do IMERG. Fonte: SILVA & FORTE [157].

A seguir serão apresentados os detalhes dos descritores, das variáveis e as escalas de pontuação utilizadas em cada um dos PVs do IMERG.

a) Condições Estruturais de Natureza Física (PV11)

A abordagem metodológica da aptidão agrícola das terras utilizada foi baseada em RAMALHO & BEEK [137], ela procura uma relação custo/benefício favorável que represente a média das possibilidades dos agricultores, numa tendência econômica a longo prazo, considerando-se o nível tecnológico a ser adotado. Esta classificação utiliza os níveis de manejo A (primitivo), B (pouco desenvolvido) e C (desenvolvido), 4 classes de aptidão agrícola (boa, regular, restrita e inapta) e 6 grupos de classificação, conforme o quadro abaixo [16,157].

Grupo	Caracterização
1	Terras com aptidão boa para lavouras em pelo menos um dos níveis de manejo A, B ou C
2	Terras com aptidão regular para lavouras em pelo menos um dos níveis de manejo A, B ou C
3	Terras com aptidão restrita para lavouras em pelo menos um dos níveis de manejo A, B ou C
4	Terras com aptidão boa, regular ou restrita para pastagem plantada.
5	Terras com aptidão boa, regular ou restrita para pastagem silvicultura e/ou pastagem natural
6	Terras sem aptidão agrícola.

Quadro 2.13 Aptidão agrícola das terras. Fonte: SILVA & FORTE [157].

A escala de pontuação para este PV foi obtida pela aplicação do MACBETH (Tabela 2.1), e demonstra o resultado dos julgamentos de diferença de atratividade entre as classes de aptidão agrícola dada pelos decisores.

Tabela 2.1 Escala de Pontuação - Aptidão Agrícola das Terras.

Classe	Pontos
6	100
5	90
4	75
3	35
2	15
1	0

Fonte: SILVA & FORTE [157].

b) Condições Estruturais de Natureza Climatológica (PV12)

Os aspectos de natureza climatológica foram representados pelo volume de precipitação histórica média de cada município do Ceará, sendo o julgamento da diferença de atratividade entre os níveis de precipitação, em mm/ano, construído pela aplicação do MACBETH. Na Tabela 2.2 é apresentada a escala de pontuação resultante. Os pontos intermediários dessa escala são obtidos por interpolação linear. Para qualquer município com níveis de precipitação abaixo de 500 mm/ano é atribuída a pontuação máxima (100 pontos), e para valores acima de 2.000 mm/ano a pontuação mínima (0 pontos) [16,157].

Tabela 2.2 Escala de Pontuação - Climatologia.

Precipitação	Pontos
500	100
1000	37,5
1500	12,5
2000	0

Fonte: SILVA & FORTE [157].

c) Efeitos sobre a População Rural (PV131)

Para a composição da escala de pontuação desse PV, utilizou-se uma escala linear, em função da população rural dos municípios, obtida pela equação: $y = 0,025 \times População_{mun.}$. Na Figura 2.27 é mostrada a escala construída. Foi convencionado que para municípios com populações rurais acima de 40.000 habitantes é atribuída a pontuação máxima (100 pontos) [16, 157].

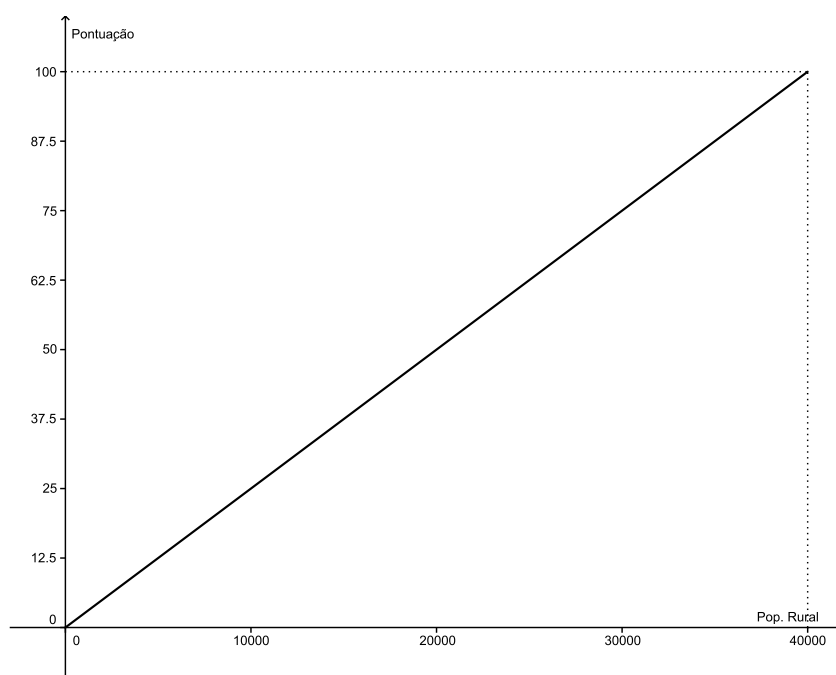


Figura 2.27 Escala de pontuação - Efeito sobre a população rural - Fonte: SILVA & FORTE [157].

d) Efeitos sobre a Base Econômica (PV132)

A escala de pontuação para esse PV também foi construída pelo MACBETH. Na Tabela 2.3 é apresentada a escala resultante dos julgamentos de atratividade entre os níveis de participação do PIB rural no PIB total dos municípios do Ceará em percentual. Os pontos intermediários

Tabela 2.3 Escala de Pontuação - Efeito sobre a base econômica.

PIB rural / PIB total (%)	Pontos
50	100
37,5	60
25	30
12,5	10
0	0

Fonte: SILVA & FORTE [157].

são obtidos por interpolação linear, e para qualquer município com valores acima de 50% é atribuída a pontuação máxima (100 pontos) [16, 157].

e) Condições Municipais de Atenuação dos Efeitos da Seca (PV133)

As condições municipais para atenuação dos efeitos das secas foram representadas pelas receitas *per capita* recebidas pelos municípios de acordo com os limites mínimo e máximo dados no Quadro 2.12. Foi utilizada uma escala de pontuação linear para esse PV, obtida a partir do MACBETH (Figura 2.28). Para receitas per capita inferiores a 150 reais é atribuída a pontuação máxima (100 pontos) e para valores maiores que 900 reais a pontuação mínima (0 pontos) [16, 157].

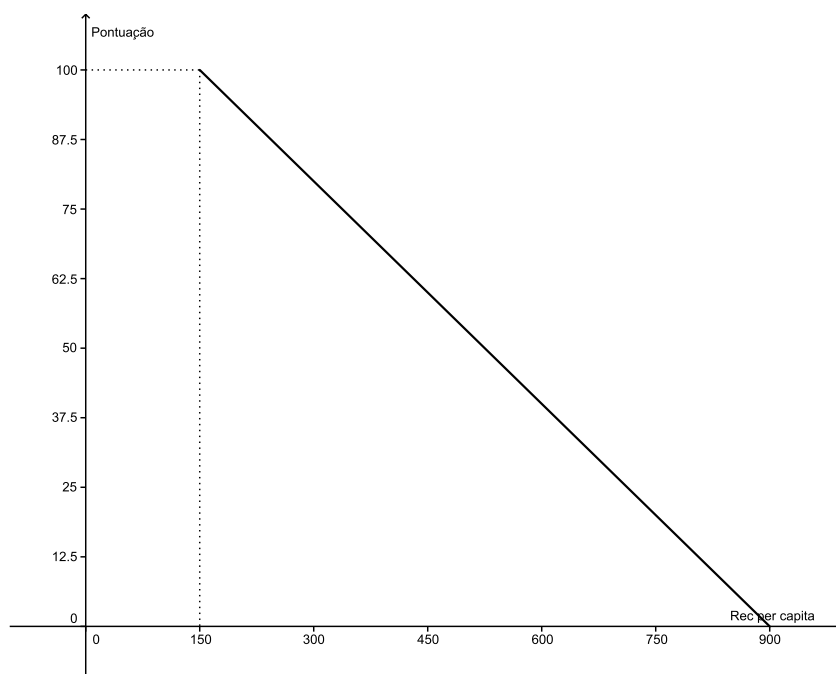


Figura 2.28 Escala de pontuação - Condições municipais de atenuação dos efeitos da seca - Fonte: SILVA & FORTE [157].

f) Risco de Perda da Produção Agrícola (PV21)

Para quantificar o risco de perdas das safras agrícolas utilizou-se o Índice de Distribuição de Chuva (IDC), que é um indicador desenvolvido pela FUNCEME e que tem sido utilizado com bastante êxito no Ceará. O IDC representa as variações temporais e espaciais da chuva, o que permite com antecedência e boa acuidade estimar a ocorrência de perdas nas safras agrícolas. o IDC varia de 0 a 1000, representando diversos níveis de condição da distribuição de precipitação, desde a condição mais crítica (0) até a condição mais favorável (1000) [16,157]. Maiores detalhes da constituição e do cálculo desse indicador serão dados no capítulo 4 dessa tese.

Na Tabela 2.4 é apresentado a escala de pontuação utilizada para esse PV, obtida pela aplicação do MACBETH. Os pontos intermediários são obtidos por interpolação linear, e para qualquer município com valor de IDC acima de 1000 é atribuída a pontuação mínima (0 pontos).

Tabela 2.4 Escala de Pontuação - Risco de perda da safra agrícola.

IDC	Pontos
0	100
50	95
150	45
250	10
1000	0

Fonte: SILVA & FORTE [157].

g) Risco de Falta D'água para o Consumo Humano e Animal - População Urbana(PV221)

Os riscos associados à falta de água para a população humana foram representados por três condições categóricas: 1) não satisfatório; 2) satisfatório, com tendência a passar a não satisfatório; e, 3) satisfatório. A escala de pontuação resultante da aplicação do MACBETH apresentou a pontuação de 100, 62,5 e 0 pontos para cada uma dessas categorias, respectivamente [16,157].

h) Risco de Falta D'água para o Consumo Humano e Animal - População Rural (PV222)

Para quantificar os riscos associados à falta de água para o abastecimento rural utilizou-se um quadro elaborado periodicamente pela Defesa Civil do Ceará, que apresenta em diversos níveis de criticidade a situação atual desse recurso (Tabela 2.5).

Tabela 2.5 Escala de Criticidade - Falta d'água para consumo humano e animal - população rural.

Nível	Descrição da Situação
5	Situação crítica
4	Risco de falta d'água iminente
3	Alto risco de falta d'água
2	Médio risco de falta d'água
1	Baixo risco de falta d'água
0	Não há risco de falta d'água

Fonte: SILVA & FORTE [157].

Por fim, a pontuação final do PV22 “Risco de falta de água para consumo humano e animal” é obtida pela média ponderada do PV221 e PV222.

Capítulo 3

A CONSTRUÇÃO DE INDICADORES

Este capítulo aborda a temática da construção de indicadores em um sentido mais amplo. Primeiro, são apresentados os seus componentes e características principais. Depois, uma abordagem de construção de indicadores compostos com base na utilização das melhores práticas. Essa foi a metodologia utilizada na construção do ISFS e, nesse capítulo, é dada ênfase aos procedimentos e métodos estatísticos escolhidos para a sua composição.

3.1 Introdução

A palavra indicador vem do verbo latino *indicare*, que significa descobrir, apontar, anunciar, estimar [79]. Além da sua etimologia e significado linguístico outros conceitos, características, funções e sistematizações serão abordados nesse capítulo de maneira mais restritiva às ciências exatas e naturais.

Indicador pode ser entendido como uma medida que resume as informações relevantes de um fenômeno particular ou um substituto dessa medida (MCQUEEN & NOAK, 1998 *apud* BELLEN [25]). Pode significar um parâmetro ou valor derivado de parâmetros que apontam e fornecem informações sobre o estado de um fenômeno (OECD, 1993 *apud* BELLEN [25]) ou, ainda, uma variável que está relacionada hipoteticamente com outra variável estudada que não pode ser diretamente observada (CHEVALIER et al. *apud* BELLEN [25]).

A característica mais importante do indicador, quando comparado com outros tipos de informação, é ser a resultante de uma série de variáveis ou parâmetros interrelacionados. Daí a sua relevância para o processo de tomada de decisão. O indicador, portanto, necessita ser significativo e representativo como descritor de certo fenômeno, tanto para os tomadores de decisão como para o público em geral (GALLOPIN, 1996 *apud* BELLEN [25]).

Entre as suas principais funções estão: a avaliação das condições e tendências em relação às

suas metas e aos seus objetivos; a comparação das diferenças e/ou semelhanças entre lugares e situações; o fornecimento de informações de advertência e a antecipação de futuras condições e tendências (TUNSTALL, 1994 *apud* BELLEN [25]).

No que se refere ao conjunto de informações a serem coletadas e disponibilizadas em um sistema de gestão, HAMMOND et al. [79] apresentou uma abordagem para o desenvolvimento de indicadores em forma de pirâmide, que ilustra bem a relação entre dados primários, estatísticos, indicadores e índices (Figura 3.1). Ela é um exemplo do conjunto de informações a serem coletadas e disponibilizadas em um sistema de gestão.

Na base da pirâmide estão os dados primários, que necessitam ser processados, validados e ajustados para serem tratados e assim transformados em dados estatísticos. A diferença entre indicadores e dados estatísticos é que os últimos descrevem os fenômenos necessitando de uma interpretação para os seus resultados, enquanto que os indicadores, que estão em um plano mais acima na pirâmide, representam um valor que é a interpretação direta do próprio resultado. No topo da pirâmide estão os índices ou indicadores compostos, descritos como a agregação sob diversas formas de inter ou intra-relações dos indicadores [79, 131].

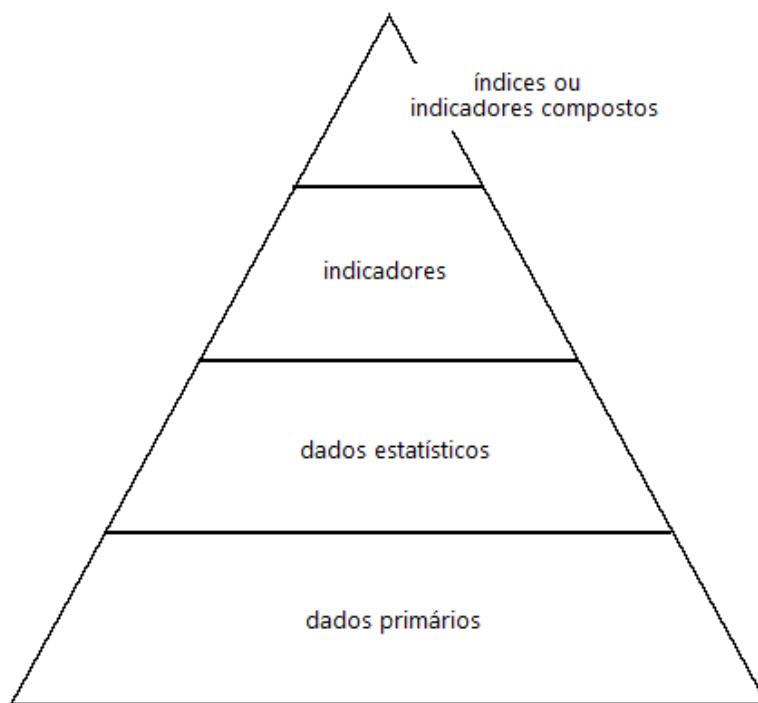


Figura 3.1 Pirâmide de Informações. Fonte: HAMMOND et al. [79].

Em resumo, os indicadores compostos são indicadores ou índices síntese de outros indicadores ou sub-indicadores, tomando como ponto de partida um método de agregação das informações.

Os índices são modelos da realidade, mas não a própria realidade. São pedaços de informação que apontam características dos sistemas, realçando e descrevendo passos indicativos do que está acontecendo. Portanto, devem ser analiticamente legítimos e construídos detalhadamente dentro de uma metodologia coerente e mensurável [25].

Como a proposta da pesquisa dessa tese é a construção de um índice, será apresentado nesse capítulo um guia para a construção de indicadores compostos que foi lançado no final de um workshop organizado pelo JRC (*Joint Research Centre*) e a OECD (*Organization for Economic Co-operation and Development*), em fevereiro de 2004, em Paris [120]. Esse guia apresenta um conjunto de recomendações de como projetar, desenvolver e disseminar um indicador composto com base em uma série de metodologias e métodos estatísticos e foi utilizado na construção do ISFS. Diversos aspectos referentes a essa metodologia serão abordados e discutidos com ênfase nos métodos estatísticos de preenchimento dos dados e na winsorização.

3.2 Componentes e Características dos Indicadores

A maioria dos indicadores existentes foi desenvolvido e utilizado para aplicações específicas: econômicas, sociais, ambientais e culturais, entre outras. Problemas mais complexos envolvendo a interconexão de diversas áreas do conhecimento requerem o desenvolvimento de indicadores inter-relacionados ou agregados com construções complexas através de indicadores e índices [25].

A abordagem de construção a partir da agregação de dados e variáveis a nível individual (Figura 3.1), não deve ser vista como exclusiva, uma vez que pode descartar considerações importantes sobre a potencialidade de outras metodologias [25].

Os indicadores podem ser classificados como escalares ou vetoriais. Os escalares apresentam-se na forma de um simples número gerado da agregação de dois ou mais indicadores ou sub-indicadores; os vetoriais apresentam os seus componentes simultaneamente e de forma não agregada (DAHL, 1997 *apud* BELLEN [25]). Há uma certa controvérsia dos autores na preferência de utilização de um ou do outro tipo. Os defensores dos indicadores vetoriais argumentam que a complexidade de um sistema pode ser melhor compreendida a partir de medidas vetoriais. Assim é possível, no vetorial, utilizar-se ferramentas matemáticas da álgebra abstrata e análise vetorial para melhor compreender o todo. Os defensores dos indicadores escalares, em contrapartida, justificam a sua utilização na simplificação e facilidade de uso, e na sua fácil compreensão pelo público em geral [25].

Os indicadores também podem ser divididos em dois grupos: os sistêmicos e os de performance. Os indicadores sistêmicos ou descritivos traçam um grupo de medidas individuais para

diferentes questões e comunicam as informações relevantes aos tomadores de decisão. Os de performance, apresentam ferramentas para comparação, que incorporam indicadores descritivos e referenciais, focados em um objetivo específico (HARDI & BARGI, 1997 *apud* BELLEN [25]).

Um aspecto essencial na construção de indicadores, que necessita ser levado em consideração, é que os julgamentos de valor estão sempre presentes nos sistemas de avaliação, nos diferentes níveis e dimensões existentes. Esses julgamentos podem ser explícitos e implícitos. Os julgamentos explícitos são aqueles tomados conscientemente e compreendem uma parte fundamental do processo de construção de indicadores. Eles podem aparecer diretamente no processo de observação e medição, adicionados na medida observada; ou, dentro do sistema agregado, a partir dos pesos atribuídos a diferentes indicadores. Os julgamentos implícitos são mais difíceis de ser observados, pois eles são inconscientes e relacionados a características pessoais e de uma determinada sociedade ou cultura [25].

Outro ponto observado na construção de indicadores é a forma da agregação dos dados. Embora indicadores altamente agregados sejam necessários para aumentar o grau de conhecimento e consciência a respeito dos problemas sob foco, os desagregados são essenciais para que se possa tomar iniciativas de ações específicas. Essa questão é particularmente importante em indicadores altamente agregados que não tem subestrutura de informação desagregada (WALL et al., 1995 *apud* BELLEN [25]). Essa limitação não ocorre em indicadores com estrutura de informação hierárquica em forma de árvore, exemplo do ISFS, índice construído nessa tese e apresentado nos capítulos subsequentes.

Indicadores altamente agregados têm também maior probabilidade de possuir problemas conceituais, principalmente se a agregação resulta em índices que condensam esferas de avaliação totalmente distintas [25]. Entretanto, a necessidade de indicadores com um certo grau de agregação é imprescindível. Esses problemas podem ser contornados pelo uso de metodologias de construção de indicadores, com base científica robusta, que possibilite a transparência dos resultados.

A transparência é um dos pré-requisitos fundamentais para a aceitação e utilização dos indicadores e que muitas vezes é negligenciada pelos construtores. Dentro dessa perspectiva sugere-se que os indicadores atendam a cinco pré-requisitos essenciais (GALLOPIN, 1996 *apud* BELLEN [25]):

1. Os valores dos indicadores devem ser mensuráveis ou observáveis;
2. Necessidade da existência e disponibilidade dos dados;
3. A metodologia necessita ser transparente e padronizada, tanto para a coleta como para o

processamento dos dados;

4. Disponibilidade dos meios para construir e monitorar os indicadores, incluindo a capacidade humana, financeira e técnica;
5. Viabilidade financeira com levantamento das informações, construção, implementação e atualização dos indicadores.

Em resumo, para a construção de um bom indicador são necessárias as seguintes características descritas no Quadro 3.1:

1 - Os valores devem ser claros; as incertezas são indesejadas.
2 - Seu conteúdo deve ser claro, entendível, com unidades que façam sentido.
3 - Deve ser factível, isto é, mensurável dentro de um custo razoável.
4 - Deve ser um descritor eficiente do fenômeno, ou seja, um meio termo entre o excesso de informações e as informações insuficientes, disponibilizando um quadro adequado do fenômeno descrito.
5 - A sua compilação necessita ser ágil e o tempo de processamento não deve ser excessivo.
6 - Deve apresentar uma escala apropriada, situada dentro de limites que nem estejam super nem subagregados.
7 - Deve ser democrático, dando possibilidade às pessoas de ter acesso às informações resultantes da aplicação da ferramenta.
8 - Deve ser hierárquico, permitindo descer na pirâmide das informações, ao mesmo tempo que transmite a informação rapidamente.
9 - Deve ser um suporte à decisão, ou seja, fornecer informações que conduzam à ação.
10 - Deve ser provocativo, levando à discussão, ao aprendizado e à mudança.
11 - Necessita de elaboração criteriosa para impulsionar a ação política.
12 - Necessita de sistemas de informações coerentes, adequados e disponíveis, dos quais esse indicador possa ser derivado.

Quadro 3.1 Características necessárias de um bom indicador ou indicador composto. Fonte: MEADOWS, 1998 *apud*

BELLEN [25].

Os indicadores são parte de um sistema que deve coletar e gerenciar as informações, além de fornecê-las para a ferramenta de avaliação (MEADOWS, 1998 *apud* BELLEN [25]). Indicadores mal formulados ou baseados em modelos não confiáveis podem levar a sociedade a uma percepção errônea da realidade. Por isso, é de suma importância, no processo de construção de indicadores, a utilização de metodologias com forte rigor científico e que viabilizem a sua implementação e a transparência dos resultados. Isso pode ser alcançado com o uso de metodologias de construção de indicadores compostos, descritas no próximo item.

Em resumo, os indicadores não são a realidade, não são completos, e nem contém todos os elementos da realidade, sendo uma aproximação desta. Eles expressam um comportamento e,

apesar da imprecisão, fazem parte do processo de compreensão das relações entre o homem e o meio ambiente em qualquer de suas dimensões ou na união de todas: sócio-econômica-cultural-ambiental [25].

3.3 A Construção de Indicadores Compostos - Um Guia para Utilização de Melhores Práticas

Os indicadores compostos são medidas que agregam diversas variáveis associadas a um fenômeno observável, multidimensional, compreendido sobre um determinado quadro teórico, através de um índice [71,138].

Um indicador composto é construído a partir de indicadores individuais que são compilados dentro de um simples índice em função de um modelo básico. O indicador mede conceitos multidimensionais que não são capturados por um simples indicador sozinho, como por exemplo: vulnerabilidade a fenômenos climáticos, competitividade, industrialização, sustentabilidade, integração de mercados, entre outros [120]. No contexto da análise política, os indicadores compostos ou índices são úteis na identificação de tendências, na definição de prioridades políticas e em *benchmarking* ou monitoramento de performances [120].

Os indicadores compostos têm se difundido largamente nos últimos anos, principalmente em organizações econômicas e estatísticas nacionais e internacionais, com o objetivo de classificar países ou regiões [71,117,120]. Segundo o Escritório de Estudos em Desenvolvimento das Nações Unidas (EED/NU), há uma clara tendência de crescimento tanto na quantidade de índices desenvolvidos no mundo quanto da variedade das instituições (públicas ou privadas) que elaboram esses índices [18].

De acordo com a pesquisa realizada pelo EED/NU, 83% dos índices da amostra foram criados no período de 1991-2006 e 50% no período de 2001-2006 (Figura 3.2). A maior parte desses índices foram elaborados por instituições privadas sem fins lucrativos, seguidas pelas instituições públicas e instituições privadas com fins lucrativos, respectivamente (Figura 3.3). Dessas instituições, 22% foram responsáveis pela construção de dois ou mais índices, mas a maior parte desenvolveu um único índice (Figura 3.4).

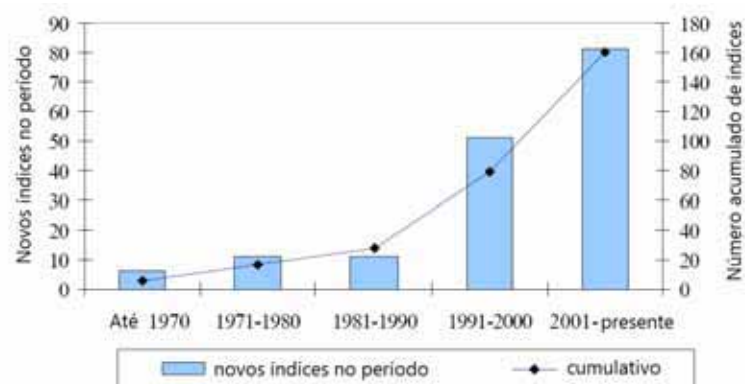


Figura 3.2 Evolução do número de índices desenvolvidos no mundo. Fonte: BANDURA [18].

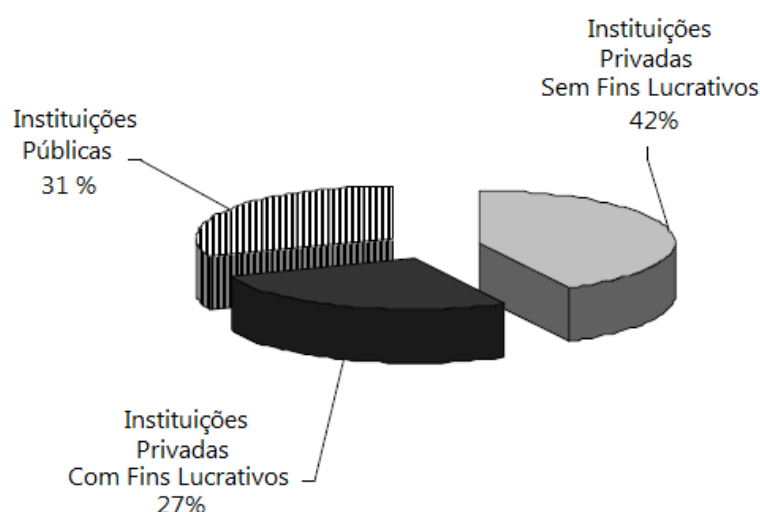


Figura 3.3 Distribuição dos índices desenvolvidos por tipo de instituição. Fonte: BANDURA [18].

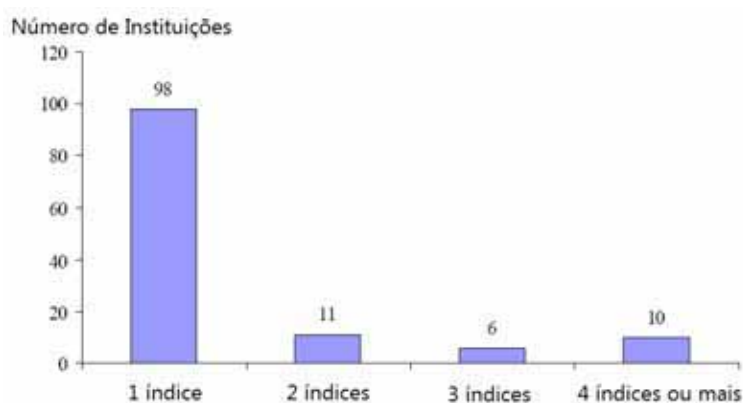


Figura 3.4 Distribuição das instituições por quantidades de índices desenvolvidos. Fonte: BANDURA [18].

Na prática, é extremamente difícil integrar variáveis individuais de uma forma que corretamente reflitam a realidade econômica, social, ambiental e cultural de uma determinada região. Esse fato torna os indicadores compostos sujeitos a usos inadequados, particularmente quando

utilizados para eleger, classificar e/ou ordenar alternativas como países, regiões e setores, entre outros [71].

Portanto, o usuário desse tipo de informação deve estar ciente que existem vantagens e desvantagens associados aos índices, conforme descrição do Quadro 3.2.

Vantagens	Desvantagens
• Pode resumir questões complexas ou multi-dimensionais com o intuito de dar suporte ao processo de tomada de decisão;	• Possibilidade de embasar recomendações inadequadas para a formulação de políticas públicas, se construídos com pouco rigor ou interpretados indevidamente;
• É mais fácil de interpretar que tentar achar uma tendência em indicadores vetoriais;	• Possibilidade de incitar conclusões simplistas com relação a políticas públicas;
• Facilita a tarefa de classificação (ordenamento) de alternativas, informações e dados que ajudam a descrever o fenômeno;	• Possibilidade de ser indevidamente utilizado, se o processo de construção não for transparente, não houver embasamento conceitual e tratamento estatístico rigoroso;
• De acordo com o conjunto de indicadores que o compõem, encerra um mesmo volume de informações em um espaço menor ou mais informações no mesmo espaço;	• Possibilidade de mascarar sérias falhas em algumas dimensões, aumentando a dificuldade de identificar ações corretivas apropriadas;
• Facilita a comunicação com a sociedade em geral promovendo a quantificação de importantes fenômenos ligados ao crescimento econômico ou ao desenvolvimento social, cultural e ambiental.	• Possibilidade de indicar delineamento de políticas públicas não apropriadas, se as dimensões de execução, difícil de ser medidas, forem ignoradas na composição do indicador.

Quadro 3.2 Vantagens e desvantagens dos indicadores compostos. Fonte: NARDO et al. [120].

Tendo em vista tanto as vantagens como as limitações no uso desses indicadores, há a recomendação que o indicador seja desagregado, de modo que os analistas possam investigar o desempenho de cada dimensão ou componente em separado [71]. Essa é a meta do ISFS, que foi construído em uma estrutura hierárquica para viabilizar, facilitar e direcionar as ações dos gestores públicos nesse sentido.

A OECD, organização internacional criada com o intuito de ajudar na reconstrução da Europa após a Segunda Guerra Mundial, e que congrega os 31 países que concentram mais da metade da riqueza mundial, tem apoiado sistematicamente o desenvolvimento de indicadores compostos [138]. Essa instituição, em conjunto com o JRC, órgão ligado à Comissão Europeia, elaboraram em 2004 um guia indicativo de melhores práticas metodológicas para construção de indicadores compostos. Esse trabalho surgiu do crescente interesse em indicadores compostos pelos círculos acadêmicos, pela mídia e pelo meio político; da existência de uma grande quantidade de abordagens metodológicas; e da necessidade clara de um guia que direcionasse as ações na elaboração desses indicadores [120]. O objetivo desse guia foi fornecer aos construtores de indicadores compostos um conjunto de recomendações de como delinear, desenvolver e disseminar um indicador composto, baseado em uma sequência de dez passos: 1) ambiente conceitual,

2) seleção das variáveis, 3) análise multivariada, 4) preenchimento de dados com falhas, 5) normalização, 6) pesos e agregação, 7) análise de sensibilidade e robustez, 8) correlação com outras variáveis, 9) decomposição do indicador, 10) apresentação e visualização dos resultados [120].

A construção do ISFS foi realizada seguindo esses passos. Nos tópicos a seguir, será apresentado em detalhes cada um deles, e a posteriori, um aprofundamento nos métodos estatísticos eleitos e utilizados na composição do ISFS.

3.3.1 Ambiente Conceitual

Uma abordagem teórica rigorosa é o ponto inicial para a construção de um indicador composto. A definição teórica do problema deverá abordar todo o fenômeno a ser medido, os seus componentes, indicadores individuais selecionados, pesos que reflitam a sua importância relativa e as dimensões de toda a composição. O ideal é que o processo seja baseado no que é desejável medir e não em que indicadores estão disponíveis. Portanto, a complexidade de informações e a sua transparência são essenciais na construção de indicadores confiáveis, e isto requer [120]:

1. Definição do conceito - A definição deverá dar ao usuário um claro sentido do que está sendo medido pelo indicador composto. O construtor deve estar ciente que alguns conceitos complexos são difíceis de definir e medir precisamente ou poderão estar sujeitos a controvérsias, afetando a credibilidade do indicador.
2. Determinação de sub-grupos - Conceitos multi-dimensionais podem ser divididos em vários sub-grupos. Estes sub-grupos não necessitam ser estatisticamente independentes; mas se existirem dependências, estas necessitam ser descritas teoricamente ou empiricamente.
3. Identificação dos critérios de seleção - O critério de seleção deverá trabalhar como um guia na definição dos indicadores individuais e, na definição se eles deverão ou não ser incluídos no indicador composto.

Após essa fase deseja-se que o construtor tenha um claro entendimento e definição do fenômeno multidimensional a ser medido, uma lista de critérios de seleção para as variáveis e uma clara documentação do processo [120].

3.3.2 Seleção das Variáveis

Um indicador composto é uma soma de partes, interrelacionadas ou não, fornecendo uma resultante indicadora do comportamento do fenômeno a que se propõe medir. A força ou a fraqueza da composição deriva fortemente da qualidade das variáveis que o compõem. As variáveis deverão ser selecionadas com base na sua relevância quanto ao fenômeno a ser medido e na solidez analítica, oportunidade, acessibilidade e mensurabilidade, entre outras características [71,120].

Os indicadores compostos são povoados por problemas de falhas nos dados, a exemplo da maioria das séries estatísticas. Isto limita bastante a escolha das variáveis, porém medidas aproximadas podem ser utilizadas quando os dados desejados estão indisponíveis ou a comparabilidade sofre limitação pela ausência de algumas componentes [71,120]. Quando se utilizam medidas aproximativas deve-se verificar a precisão dessas medidas através de análise de correlação e sensibilidade.

Em resumo, a qualidade e precisão dos indicadores compostos deve evoluir em paralelo com a melhoria na coleção de dados e o desenvolvimento dos indicadores que o compõem [71,120]. Cumprida essa fase, espera-se que o construtor verifique a qualidade dos indicadores disponíveis, tenha discutido a força e a fraqueza dos indicadores selecionados, e crie uma tabela resumo das características dos dados, elencando disponibilidade, origem e tipo, por exemplo [120].

3.3.3 Análise Multivariada

Muitas vezes, na construção de indicadores compostos, as variáveis e os indicadores individuais que os compõem são selecionados e escolhidos de forma arbitrária, com pouca atenção às interrelações entre eles [120]. Esta é uma limitação durante a fase de construção do indicador, que necessita de atenção, o que desafortunadamente não ocorre com frequência.

A natureza subjacente dos dados necessita ser cuidadosamente analisada antes da construção final do indicador e, conseqüentemente, a avaliação da adequação do conjunto de dados fornece uma melhor compreensão das implicações das escolhas metodológicas feitas. O analista deve primeiro decidir se a estrutura do indicador composto está bem definida e se o conjunto de sub-indicadores disponíveis é suficiente e apropriado para retratar o fenômeno. Esta decisão pode ser baseada na opinião de um pesquisador, grupo de excelência ou na estrutura estatística do conjunto de dados [120].

Diferentes abordagens analíticas multivariadas podem ser utilizadas para esse fim, tais como: Análise de Componentes Principais, Análise Fatorial, Análise de Agrupamentos, entre outras.

A Análise de Componentes Principais (ACP) é um método estatístico bastante utilizado para identificar as dimensões principais de um modelo multivariado, objetivando-se revelar como as diferentes variáveis mudam em relação às outras e como elas estão associadas [2, 65, 120, 186]. Logo, a meta é reduzir um conjunto p -dimensional dentro de um pequeno número q , preservando a variação dos dados na medida do possível. Essas q novas dimensões são construídas de forma que: 1º) Elas sejam combinações lineares das variáveis originais; 2º) Sejam independentes umas das outras; e 3º) Cada dimensão capture sucessivamente uma pequena quantidade do total da variação nos dados.

Teoricamente, o número de combinações lineares pode variar de nenhuma a todas as p variáveis, mas, o objetivo é encontrar q ($0 < q < p$) combinações lineares de p , que melhor resumam as informações contidas nos dados [186]. Portanto, as p variáveis originais são combinadas dentro de q combinações lineares que formam as novas componentes principais do sistema.

Sejam Z_i as combinações lineares dos vetores de dados originais X_j , com $i, j = 1, 2, \dots, p$, definidas como:

$$\begin{aligned} Z_1 &= a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Z_2 &= a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Z_p &= a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \end{aligned} \tag{3.1}$$

Sendo que os pesos a_{ij} (também chamados componentes ou fatores de cargas) aplicados as variáveis X_j nas equações acima, são escolhidos de forma que as componentes principais Z_i satisfaçam as seguintes condições:

1) Elas não estão correlacionadas, ou seja, são ortogonais ou estatisticamente independentes, e,

2) A primeira componente principal contabiliza o máximo possível da proporção da variância do conjunto de X_i 's, a segunda componente contabiliza o máximo da variância restante, e assim por diante, até a última componente absorver todo o restante da variância não encontrada pelas componentes precedentes, e, $a_{i1}^2 + a_{i2}^2 + \dots + a_{ip}^2 = 1$, com $i = 1, 2, \dots, p$.

O próximo passo envolve encontrar os autovalores λ_j , $j = 1, \dots, p$, da matriz da covariância da amostra CM :

$$CM = \begin{bmatrix} cm_{11} & cm_{12} & \cdots & cm_{1p} \\ cm_{21} & cm_{22} & \cdots & cm_{2p} \\ \vdots & \vdots & & \vdots \\ cm_{p1} & cm_{p2} & \cdots & cm_{pp} \end{bmatrix} \quad (3.2)$$

onde o elemento da diagonal cm_{ii} é a variância de X_i e cm_{ij} é a covariância das variáveis X_i e X_j . Os autovalores da matriz CM são as variâncias das componentes principais e podem ser encontrados resolvendo-se a equação $|CM - \lambda I| = 0$, onde I é a matriz identidade de mesma ordem de CM , e λ é o vetor de autovalores.

Esse método foi escolhido para utilização no ISFS, conforme será visto nos capítulos subsequentes, com o objetivo principal de avaliar se as dimensões do modelo estão bem balanceadas e identificar quais os sub-indicadores que são mais significativos.

A Análise Fatorial (AF) é um método estatístico multivariado similar a ACP. Seu objetivo é descrever um conjunto de Q variáveis $x_1, x_2, \dots, x_Q$ em termos de um pequeno número de m fatores, e destacar a relação entre estas variáveis.

Enquanto a ACP é baseada em combinações lineares dos dados, a AF é baseada em um modelo especial. A AF assume que os dados estão baseados em fatores subjacentes do modelo, e a sua variância pode ser decomposta e representada por fatores comuns e únicos, ao contrário da ACP [2,120].

O modelo é dado por:

$$\begin{aligned} x_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + e_1 \\ x_2 &= a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + e_2 \\ &\vdots \\ x_Q &= a_{Q1}F_1 + a_{Q2}F_2 + \dots + a_{Qm}F_m + e_Q \end{aligned} \quad (3.3)$$

onde $x_i (i = 1, \dots, Q)$ representa a variável original padronizada com média zero e variância unitária; $a_{i1}, a_{i2}, \dots, a_{im}$ são as cargas fatoriais relacionadas com as variáveis X_i ; $F_1, F_2, \dots, F_m$ são os m fatores comuns não correlacionados e e_i os erros residuais.

Existem várias abordagens para lidar com o modelo dado no conjunto de Equações 3.3, como por exemplo: comunalidades, fatores de máxima versossimilhança, método do centróide, método do eixo principal, etc. O mais comum é utilizar a ACP para extrair os primeiros m componentes principais, considerá-los como fatores e negligenciar os restantes [2,120]. A combinação AF/ACP também é utilizada como método principal na construção de indicadores,

tais como, os indicadores de sustentabilidade [73,103].

A Análise de Agrupamentos (AA) é outra ferramenta estatística multivariada para a classificação de grande quantidade de informação dentro de conjuntos administráveis. Pode ser utilizada como: 1) Um método puramente estatístico de agregação de indicadores; 2) Uma ferramenta de diagnóstico para explorar o impacto das escolhas metodológicas feitas durante a fase de construção; 3) Um método de disseminação da informação do indicador composto, sem perda das dimensões dos sub-indicadores; e, 4) Um método para selecionar grupos alternativos para preenchimento de falhas nos dados com o objetivo de diminuir a variância dos dados [120].

A AA pode ser hierárquica se a classificação tem um aumento no número de classes aninhadas, por exemplo, agrupamento em árvore; ou não hierárquica, quando o número de grupos é decidido antecipadamente, como por exemplo, agrupamento pelo método *k-means* [120].

Grupos homogêneos ou distintos podem ser delineados, baseados na avaliação da distância ou, como no caso do método de Ward, em um teste F. A distância medida é uma avaliação da similaridade ou dissimilaridade entre os elementos do conjunto. Na AA, uma pequena distância significa uma grande similaridade [120].

Os algoritmos de agrupamentos definem as regras de como as distâncias são medidas entre os grupos [120]. No Quadro 3.3 são listados os algoritmos mais comumente utilizados.

Algoritmo	Características
Vizinho mais próximo	A distância entre dois grupos é determinada pela distância entre os dois elementos mais próximos nos diferentes grupos.
Vizinho mais distante	A distância entre dois grupos é determinada pela maior distância entre quaisquer dois objetos pertencentes a diferentes grupos.
Média não ponderada	A distância entre dois grupos é calculada pela média da distância entre todos os pares de objetos em dois grupos.
Média ponderada	Similar ao anterior exceto que o tamanho do grupo, ou seja número de objetos, é usado como peso para a distância média. Este método é útil quando os tamanhos dos grupos são muito diferentes.
Método de Ward	O membro de um grupo é determinado pela variância dos elementos, ou seja, a soma dos desvios quadrados da média dos grupos.

Quadro 3.3 Algoritmos de agrupamento utilizados na AA. Fonte: NARDO et al. [120].

Em resumo, no final dessa fase, espera-se que o construtor tenha checado a estrutura subjacente dos dados ao longo das suas dimensões; aplicado uma metodologia de análise de dados multivariada adequada; identificado sub-grupos de indicadores que são estatisticamente similares; analisado se a estrutura do conjunto de dados está bem definida e adequada ao fenômeno

que propõe medir; e, por fim tenha documentado o resultado da análise multivariada e a interpretação dos componentes e fatores [120].

3.3.4 Preenchimento de Dados com Falhas

A questão de como lidar com falhas ou observações incompletas que surgem comumente em todos os ambientes de coleção de dados, está entre os mais persistentes e complicados problemas da gestão de dados e informações nas pesquisas científicas. O grau de incerteza associado a falhas nos dados afeta diretamente a habilidade de tirar conclusões precisas e, em muitos casos, aumenta com o nível de agregação dos dados [186]. Esta ausência ou inconsistência nas séries de dados, como os dados meteorológicos, por exemplo, aumenta a dificuldade de se trabalhar de forma consistente, principalmente em países como o Brasil, de dimensão continental.

A ausência ou falha de dados muitas vezes impede o desenvolvimento de indicadores compostos ou restringem o conjunto de alternativas a se avaliar. Isto pode diminuir a confiabilidade do indicador desenvolvido e distorcer a posição relativa das alternativas em relação a todo o universo de avaliação do indicador [71,120].

Para lidar com essas ausências de dados, três métodos gerais são apresentados como possível solução: 1) exclusão dos dados com falhas, 2) preenchimento único e 3) preenchimento múltiplo.

O método de exclusão omite da análise os dados com falhas. A utilização desse método é impeditiva quando o objetivo é realizar comparações onde não é possível excluir nenhuma das alternativas. Como regra de ouro, os autores sugerem que se uma variável contém mais de 5% dos dados com falhas então não se deve usar a exclusão (LITTLE & RUBIN, 2002 *apud* NARDO et al. [120]). Esse é o caso do índice desenvolvido nesta tese, o ISFS, no qual deseja-se comparar todos os municípios do Estado, especificamente o Rio Grande do Norte nesta etapa, sem exclusão de nenhum deles.

Para as outras duas abordagens existem uma série de métodos que podem ser utilizados dependendo da percentagem de falhas, do mecanismo ou do padrão de não resposta dos dados. Para o preenchimento único, apresentam-se os seguintes métodos: substituição pela média, mediana ou moda, último valor observado, hot deck, regressão, estimativa de máxima verossimilhança, entre outros. Para o preenchimento múltiplo, os métodos mais utilizados são: Regressão Linear Bayesiana (*Bayesian Linear Regression - BLR*); Método da Média Preditiva (*Predictive Mean Matching - PMM*); ou Método de Monte Carlo baseado em Cadeias de Markov (*Markov Chain Monte Carlo - MCMC*). Maiores detalhes sobre o preenchimento dos dados com falhas serão vistos no item 3.4. O ISFS foi construído com a opção de utilização de qualquer um

desses métodos de preenchimento múltiplo, dependendo do padrão de não resposta dos dados, conforme será detalhado nos capítulos seguintes dessa tese. Antes, porém, de ser efetuado o preenchimento dos dados com falhas talvez seja necessário transformar algumas das variáveis, com o objetivo de satisfazer aos pressupostos dos métodos de preenchimento dos dados. Esse é o caso do MCMC, cuja aplicação exige o pressuposto da normalidade multivariada dos dados.

Em resumo, o primeiro passo é testar a condição de normalidade para todas as variáveis ou sub-indicadores, que em muitos casos exibem substancial assimetria. A maioria dessas variáveis também apresenta padrões de heterocedasticidade, o que significa que a variância aumenta com a magnitude dos dados. Nessas condições, a transformação das variáveis se faz necessária, e o construtor pode optar por uma série de métodos de transformação de dados. O mais comum é a transformação logarítmica. Porém, a transformação pode mudar a distribuição das variáveis de forma a afetar o resultado final do indicador composto. Os autores também afirmam que uma distribuição simétrica normal apresenta a maioria das observações com desvio padrão menor ou igual a dois, o que inclui 95% das observações, e consequentemente, os valores extremos ocorrem com menor probabilidade [186].

Logo, a fim de estabelecer um equilíbrio entre a distribuição das variáveis ou sub-indicadores e minimizar o impacto das transformações na pontuação final do indicador composto sugere-se o seguinte procedimento [186]:

1. Antes de executar o preenchimento múltiplo, transformar todas as variáveis ou sub-indicadores com assimetria maior que dois,
2. Após o preenchimento dos dados, transformar novamente as variáveis para sua escala original exceto as que apresentarem assimetria maior que quatro.

Esse procedimento garante que as variáveis com valores extremos ou desvio-padrão maior que quatro terão suas assimetrias corrigidas.

Uma consequência do preenchimento dos dados com falhas é a possibilidade de gerar muitos valores extremos no indicador composto, o que pode ocasionar a geração de *outliers* [186]. Esse problema pode ser atenuado com a utilização da técnica de winsorização nos dados, que corresponde a uma suavização nas caudas da distribuição para percentis específicos. Esse procedimento será detalhado mais adiante no item 3.5. Tanto o procedimento de transformação dos dados descrito acima, quanto a winsorização foram utilizados na construção do ISFS, conforme será detalhado nos capítulos seguintes dessa tese.

Como conclusão, ao final dessa fase espera-se que o construtor obtenha um conjunto de dados completos e sem falhas, uma medida de segurança de cada valor imputado para explorar os impactos do preenchimento no indicador composto e apresente a documentação do procedimento de preenchimento utilizado [120].

3.3.5 Normalização

A normalização é uma etapa necessária antes de qualquer agregação dos indicadores, principalmente em conjuntos de dados com unidades de medidas diferentes. Existe uma série de métodos à disposição conforme a descrição do Quadro 3.4. A escolha do método mais adequado, porém, não é trivial e merece especial atenção [120]. É recomendado que se considerem, na escolha do método de normalização, as propriedades dos dados e os objetivos do indicador, pois diferentes métodos de normalização produzirão diferentes resultados. Testes de robustez também podem ser necessários para avaliar os impactos dos seus resultados [120].

Método	Equação	Características Principais/Vantagens/Desvantagens
Classificação pelo indicador bruto	$I_{qc}^t = Rank(x_{qc}^t)$	Esse método não é afetado por <i>outliers</i> e permite que o desempenho das alternativas seja avaliado ao longo do tempo em função da sua posição relativa. O desempenho das alternativas em termos absolutos, no entanto, não pode ser avaliada quando ocorrem falhas nos dados.
Padronização ou <i>z-score</i>	$I_{qc}^t = \frac{x_{qc}^t - x_{qc=\bar{c}}^t}{\sigma_{qc=\bar{c}}^t}$	Esse método converte indicadores para uma escala comum, com média zero e desvio-padrão unitário. Indicadores com valores extremos têm maior efeito no indicador composto. Isto pode não ser desejável se a intenção é recompensar comportamento excepcional, o que pode ser corrigido na metodologia de agregação excluindo o melhor e o pior escores individuais.
Min-Max	$I_{qc}^t = \frac{x_{qc}^t - \min_c(x_q^{t0})}{\max_c(x_q^{t0}) - \min_c(x_q^{t0})}$	Normaliza os indicadores no intervalo [0, 1]. Valores extremos ou <i>outliers</i> podem distorcer o indicador transformado. Por outro lado, a normalização Min-Max pode ampliar o leque de indicadores que se situam num pequeno intervalo, aumentando o efeito sobre o indicador composto mais do que a transformação <i>z-score</i> .
Distância a um referencial	$I_{qc}^t = \frac{x_{qc}^t}{x_{qc=\bar{c}}^t}$ ou $I_{qc}^t = \frac{x_{qc}^t - x_{qc=\bar{c}}^{t0}}{x_{qc=\bar{c}}^{t0} - x_{qc=\bar{c}}^t}$	Traduz a posição relativa de um dado indicador em relação a uma medida de referência, que poderia ser a meta a ser atingida em um dado momento. Essa alternativa recebe 1 e os outros são percentuais de distância em relação a essa medida.
Escala de Categorias	$I_{qc}^t = \begin{cases} 25, & \text{se } x_{qc}^t \in \{p^{25th} \} \text{ percentil} \\ 50, & \text{se } x_{qc}^t \in \{p^{50th} - p^{25th} \} \text{ percentil} \\ 75, & \text{se } x_{qc}^t \in \{p^{75th} - p^{50th} \} \text{ percentil} \\ 100, & \text{se } x_{qc}^t \in \{p^{100th} - p^{75th} \} \text{ percentil} \end{cases}$	Atribui uma pontuação para cada indicador. As escalas podem ser numéricas ou qualitativas, tais como: "plenamente atingido", "parcialmente alcançados" ou "não alcançado". Frequentemente os escores são baseados nos percentis da distribuição do indicador. As escalas de categoria excluem uma grande quantidade de informação sobre a variância do indicador transformado.
Acima ou abaixo da média	$I_{qc}^t = \begin{cases} 1, & \text{se } w > (1 + p) \\ 0, & \text{se } (1 - p) \leq w \leq (1 + p) \\ -1, & \text{se } w < (1 - p) \end{cases}$	Essa transformação atribui zero para valores próximos a média, 1 ou -1 para valores acima ou abaixo da média, respectivamente. Esse método é simples e não é afetado por <i>outliers</i> ; contudo, a arbitrariedade do limiar pode omitir diferenças substanciais entre os indicadores.
Cíclico	$I_{qc}^t = \frac{x_{qc}^t - E_t(x_{qc}^t)}{E_t(x_{qc}^t - E_t(x_{qc}^t))}$	Este método implicitamente dá menos peso para uma série mais irregular do indicador composto, a não ser que seja feita uma suavização "ad hoc" antes.
Percentual em relação ao período anterior	$I_{qc}^t = \frac{x_{qc}^t - x_{qc}^{t-1}}{x_{qc}^{t-1}}$	Representa o crescimento ou decréscimo em relação ao valor do indicador no período anterior. Esta transformação só pode ser utilizada quando os indicadores estão disponíveis para os períodos anteriores.

Legenda: x_{qc}^t é o valor do indicador para a alternativa c no período t ; $\bar{c}$ é a alternativa de referência; E = Esperança.

Quadro 3.4 Métodos de normalização. Fonte: NARDO et al. [120].

Como conclusão final dessa fase, espera-se que o construtor tenha selecionado o método de normalização mais adequado, tomando por referência o ambiente conceitual do indicador proposto e as propriedades intrínsecas dos seus dados; com isso, documentado e explicado o procedimento de normalização selecionado e os seus resultados [120].

3.3.6 Pesos e Métodos de Agregação

Os pesos tem um significativo efeito sobre o indicador composto, pois definem a importância relativa de cada um dos sub-indicadores ou dimensões em relação ao todo. Existe uma série de métodos que podem ser utilizados na definição e composição dos pesos de um indicador composto. Alguns são derivados de modelos estatísticos, tais como: Análise Fatorial, Análise Envoltória de Dados, Benefício da Dúvida (*Benefit of the Doubt - BOD*) e Modelos de Componentes não Observados. Outros são derivados de métodos participativos, tais como: Processo de Alocação de Orçamentos (PAO), Análise Hierárquica de Processos (AHP) e Análise Conjunta (AC). Não importa o método usado, os pesos são essencialmente valores de julgamentos de decisores [120]. Estes necessitam apresentar, além da visão de pontos específicos, uma visão sistêmica envolvendo a complexidade do indicador.

A maioria dos indicadores utiliza pesos iguais nas suas dimensões. Isso quer dizer que essas dimensões tem o mesmo valor na sua composição, mas também podem “disfarçar a ausência de estatísticas ou base empírica”. Isto ocorre quando há um conhecimento insuficiente das relações de casualidade ou uma falta de consenso sobre uma alternativa [120].

Ao se utilizarem pesos iguais, pode acontecer que um elemento de dupla contagem seja introduzido no índice. Isto ocorre através da combinação de variáveis com um elevado grau de correlação. Se dois indicadores colineares são incluídos no indicador composto com pesos w_1 e w_2 , então os dois indicadores deverão ter peso $(w_1 + w_2)$ na composição. Para resolver esse problema, uma alternativa é testar a correlação desses indicadores e escolher os indicadores com baixo grau de correlação ou ajustar os pesos correspondentemente. Em resumo é dar menos peso aos indicadores correlacionados [120].

Observa-se aqui, que existe sempre alguma correlação positiva entre as medidas na mesma agregação, e uma regra que poderá ser adotada é definir um limiar além do qual a correlação é sintoma de dupla contagem [120].

Os pesos também podem ser escolhidos em função da qualidade estatística dos dados. Um peso maior pode ser atribuído a dados estatisticamente confiáveis, com ampla cobertura. No entanto, esse método pode ser um viés para os indicadores disponíveis, penalizando as informações que são estatisticamente mais problemáticas de identificar e medir [120].

A situação ideal é que os pesos reflitam a contribuição de cada indicador na composição do índice. Métodos estatísticos como ACP e AF também podem ser utilizados para definir os pesos de um indicador composto. Esses métodos representam a maior variação no conjunto de dados, utilizando um pequeno número de fatores que refletem a dimensão estatística subjacente

do conjunto de dados. Entretanto, os pesos não podem ser estimados se não houver correlação entre os indicadores. [120].

Outros métodos estatísticos, como BOD, são extremamente parcimoniosos sobre os pressupostos dos pesos, pois permitem que os dados decidam os pesos e são sensíveis as prioridades das alternativas.

Na abordagem BOD, o indicador composto é definido em razão da performance atual de cada alternativa sobre a performance de um *benchmark* (teste):

$$CI_i = \frac{\sum_{q=1}^M I_{qi} w_{qi}}{\sum_{q=1}^M I_{qi}^* w_{qi}} \quad (3.4)$$

onde, I_{qi} é o valor normalizado do q -ésimo sub-indicador ($q = 1, \dots, Q$) para a alternativa i ($i = 1, \dots, M$) e w_{qi} é o peso correspondente, e,

$$I^* = I^*(w) = \arg \max_{I_k, k \in \{1, \dots, M\}} \left(\sum_{q=1}^Q I_{qk} w_q \right) \quad (3.5)$$

Sendo que I^* é um valor de alternativa hipotético que maximiza toda a performance, dado um conjunto de pesos desconhecidos w . Operacionalmente, a equação acima pode ser reduzida a um problema de programação linear da seguinte forma:

$$\begin{aligned} CI_i^* &= \arg \max \sum_{q=1}^Q I_{qc} w_{qc} \\ \text{sujeito a :} \\ \sum_{q=1}^Q I_{qk} w_{qk} &\leq 1 \\ w_{qk} &\geq 0 \\ \forall k &= 1, \dots, M; \forall q = 1, \dots, Q. \end{aligned} \quad (3.6)$$

Uma outra alternativa é a utilização de métodos participativos que incorporem vários *stakeholders* - especialistas, gestores, políticos ou o público em geral - na atribuição de pesos. Existe uma vasta gama de métodos que utilizam essa abordagem e que podem ser utilizados pelos construtores de indicadores compostos: PAO, AHP, AC, entre outros. Na PAO os especialistas são dotados de um orçamento com n pontos, para ser distribuído ao longo de um certo número de indicadores individuais, "pagando" mais por aqueles indicadores cuja importância querem privilegiar. Essa abordagem é limitada a um máximo de 10 ou 12 sub-indicadores. A AHP e a AC são técnicas largamente utilizadas na tomada de decisão multicritério. A primeira através da comparação par a par dos atributos e a segunda pela comparação dos diferentes níveis de atributos.

Essas técnicas permitem a derivação da importância do atributo global, baseada em uma série de comparações dos atributos por rotação, ao invés de simplesmente atribuir arbitrariamente um determinado peso. Os pesos resultantes da aplicação dessas abordagens são menos susceptíveis a erros de julgamentos [120]. Um exemplo de método usando essa abordagem foi o utilizado na composição do IMERG, o MACBETH (ver Capítulo 2, item 2.6.2).

Na composição do ISFS foram incorporados os mesmos pesos utilizados no IMERG, que foram obtidos pela aplicação da metodologia multicritério MACBETH a uma gama de decisores (*stakeholders*), conforme foi visto no Capítulo 2, item 2.6, Figura 2.25. Com a adoção dessa estratégia nesse trabalho de tese houve uma apropriação de todo o conhecimento elaborado e desenvolvido na composição do IMERG e dos seus referidos pesos.

Um fator importante na construção de um indicador composto é a escolha do método de agregação a se utilizar. O método de agregação linear (LIN) é útil quando todos os indicadores individuais têm a mesma unidade de medida, desde que algumas propriedades matemáticas sejam respeitadas. Já o método de agregação geométrica (GME) é mais adequado a um modelo onde se deseja algum grau de não compensabilidade entre os indicadores individuais ou dimensões que o compõem. As agregações lineares recompensam seus indicadores proporcionalmente aos seus pesos, enquanto que as geométricas recompensam as alternativas com maior pontuação [120].

Em ambas as abordagens de agregação, os pesos refletem um conflito de escolha entre os indicadores individuais. Um déficit, em uma dimensão, pode ser compensado por um excedente em outra. Isso implica em uma inconsistência entre como os pesos são concebidos e o seu significado real, quando as agregações geométricas ou lineares são utilizadas. Em uma agregação linear, a compensabilidade é constante, enquanto na geométrica a compensabilidade é menor para o indicador composto com menor valor [120].

Para garantir que os pesos continuem a ser uma medida de importância, outros métodos de agregação podem ser utilizados, em particular os métodos que não permitam compensabilidade. Isso pode ser alcançado com a utilização de uma Abordagem Multicritério Não Compensatória - AMNC, que pode garantir a não compensabilidade, buscando um compromisso entre dois ou mais objetivos legítimos. Porém, métodos com essa abordagem podem aumentar em demasia o tempo computacional se o número de alternativas for muito grande [116, 117].

Dado um conjunto de sub-indicadores $\mathbf{G} = \{x_q\}$, $q = 1, \dots, Q$, e um conjunto finito $\mathbf{M} = \{i\}$, $i = 1, \dots, M$ de alternativas, assume-se que a avaliação de cada alternativa i com respeito a um indicador x_q é baseada em uma escala intervalar ou de razão de medidas. Assume-se também a existência de um conjunto de pesos $w = \{w_q\}$, $q = 1, \dots, Q$, com $\sum_{q=1}^Q w_q = 1$,

interpretados como coeficientes de importância. Então, a aplicação da AMNC é dada por:

1. Comparação dos pares de alternativas para o conjunto completo de sub-indicadores;
2. Classificação das alternativas em uma pré-ordem completa.

Do primeiro passo resulta uma matriz de classificação E com dimensão $M \times M$. Qualquer elemento de E , e_{jk} , $j \neq k$ é resultado da comparação entre j e k , de acordo com todos os sub-indicadores. Essa comparação global dos pares é obtida por:

$$e_{jk} = \sum_{q=1}^Q (w_q(Pr_{jk}) + \frac{1}{2}w_q(In_{jk})) \quad (3.7)$$

onde, $w_q(Pr_{jk})$ e $w_q(In_{jk})$ são os pesos dos sub-indicadores apresentando uma relação de preferência e indiferença, respectivamente.

Para o ISFS, foi escolhido como método de agregação o método aditivo linear (LIN), por dois motivos principais: 1) Os sub-indicadores, após a normalização, têm a mesma escala de medida; e 2) É permitido compensabilidade, uma vez que os pesos atribuídos a cada sub-indicador apresentam dimensões diferentes, em função da sua significância na descrição e quantificação da severidade da seca.

Ao concluir essa etapa, o construtor deve definir os pesos e o procedimento de agregação que mais se adequem à descrição do fenômeno que se propõe medir e apresentar a documentação e as justificativas das escolhas dos pesos e do procedimento de agregação.

3.3.7 Análise de Sensibilidade e Robustez

Nenhum índice pode ser melhor do que os dados que utiliza; porém, esse argumento em prol da melhoria dos dados deve se estender também para a melhoria da estrutura do próprio índice [120]. Na construção de um indicador composto, julgamentos são realizados durante todo o processo: na seleção das variáveis, na normalização dos dados, na escolha dos pesos e do método de agregação, entre outros. Esse fato permite que a robustez do indicador e os seus resultados sejam questionados pelo público alvo (gestores, políticos e o público em geral). Para lidar com esses questionamentos durante o processo de institucionalização dos indicadores uma combinação de análise de sensibilidade e de incertezas pode ajudar na avaliação do indicador e na transparência dos seus resultados [120].

A análise de sensibilidade avalia a contribuição das origens individuais de incertezas a partir da análise da variância dos resultados dos indicadores. A análise de incertezas concentra-se em

como os fatores de incertezas na entrada se propagam através da estrutura do indicador e afetam os seus resultados finais [120].

Uma condição ideal na construção dos indicadores compostos é que todas as possíveis origens de incertezas possam ser abordadas.

NARDO et al. [120] apresentaram uma abordagem para avaliar as incertezas durante o processo de construção de um indicador composto, baseado nos seguintes passos:

1. Inclusão e exclusão de sub-indicadores;
2. Modelagem dos erros dos dados com base nas informações disponíveis sobre a estimativa da variância;
3. Uso de sistemas alternativos para preenchimento de falhas nos dados, por exemplo: preenchimento simples ou múltiplo;
4. Uso de sistemas alternativos de normalização;
5. Uso de diferentes sistemas para escolha dos pesos e métodos de agregação

Logo, seja CI o valor do indicador composto para uma alternativa i , $i = 1, \dots, M$, dado por:

$$CI_i = f_{rs}(I_{1,i}, I_{2,i}, \dots, I_{Q,i}, w_{s,1}, w_{s,2}, \dots, w_{s,Q}) \quad (3.8)$$

onde $r = 1, 2, 3$ é o método de agregação utilizado (LIN, GME, AMNC) e $s = 1, 2, 3$ é o esquema de pesos utilizados (PAO, AHP, BOD).

O cálculo desse índice é baseado nos Q sub-indicadores normalizados $(I_{1,i}, I_{2,i}, \dots, I_{Q,i})$ e são dependentes dos esquemas de pesos dos sub-indicadores $(w_{s,1}, w_{s,2}, \dots, w_{s,Q})$.

São sugeridos os seguintes métodos de normalização: Min-Max (Eq. 3.9), padronização (Eq. 3.10) ou o valor do indicador bruto (Eq. 3.11).

$$I_{q,i} = \frac{x_{q,i} - \min(x_q)}{\max(x_q) - \min(x_q)} \quad (3.9)$$

$$I_{q,i} = \frac{x_{q,i} - \text{mean}(x_q)}{\text{std}(x_q)} \quad (3.10)$$

$$I_{q,i} = x_{q,i} \quad (3.11)$$

O método de normalização Min-Max pode ser usado em todos os esquemas de pesos (PAO, AHP, BOD) e sistemas de agregação (LIN, GME, AMNC). O método de padronização nos esquemas de pesos (PAO, AHP) e sistemas de agregação (LIN, AMNC). E o valor do indicador bruto para os esquemas de pesos (PAO, AHP) e sistemas de agregação (GME, AMNC).

Nessa abordagem, a classificação atribuída a cada indicador $Rank(CI_i)$ é uma saída da análise de sensibilidade/incertezas.

A mudança média na classificação de uma alternativa também pode ser explorada, em função da média da diferença entre a classificação absoluta de uma alternativa e da classificação de referência sobre todas as M alternativas:

$$\overline{R_s} = \frac{1}{M} \sum_{i=1}^M |Rank_{ref}(CI_i) - Rank(CI_i)| \quad (3.12)$$

Uma outra abordagem para a análise de sensibilidade foi sugerida por BUTLER et. al [36] e aplicada a um modelo para escolha da melhor localização para a instalação de uma usina termelétrica, envolvendo treze opções de localização. Eles utilizaram uma simulação de Monte Carlo para proporcionar simultaneamente a flexibilização de todos os pesos do modelo multi-critério e investigaram o impacto da variação na forma funcional da agregação multi-atributo. Três tipos de simulação foram utilizados: 1) aleatorização total dos pesos, 2) aleatorização dos pesos mantendo-se a classificação original dada pelos decisores e 3) pesos obtidos como resposta de uma distribuição escolhida a priori. As simulações foram baseadas em 5000 iterações e os testes de robustez foram obtidos em função da dominância estocástica dos resultados das simulações.

Uma das vantagens da utilização dessa abordagem em relação aos métodos tradicionais de análise de sensibilidade é a possibilidade de hierarquizar os pesos, bem como a inclusão de testes de robustez [36].

Essa abordagem foi utilizada nesse trabalho de tese, com o intuito principal de avaliar o intervalo de confiança do ISFS, a dominância estocástica da classificação dos municípios obtidas nos testes e a consistência dos pesos utilizados na composição do ISFS.

Ao final dessa etapa o construtor deve ter realizado uma análise de sensibilidade, identificado a origem da incertezas e avaliado os seus impactos nos resultados finais do indicador que está em desenvolvimento. É interessante também que a análise de sensibilidade e os seus resultados sejam documentados.

3.3.8 Correlação com Outras Variáveis

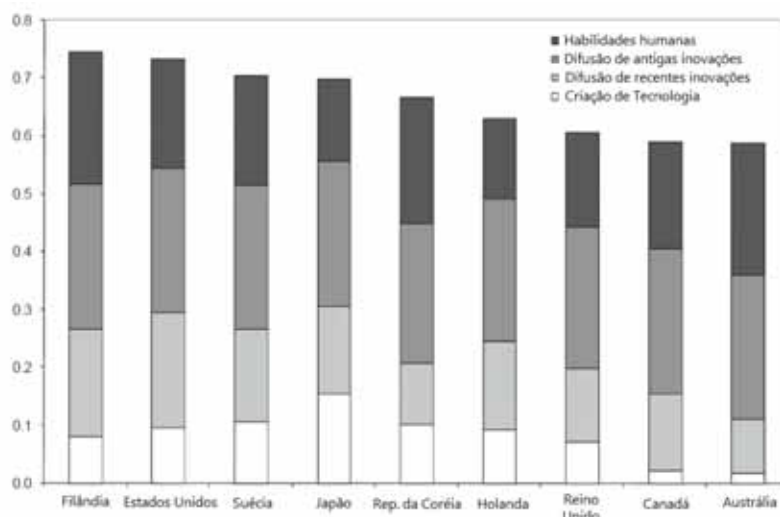
Indicadores compostos geralmente medem conceitos que estão ligados a fenômenos conhecidos e mensuráveis. A correlação com esses conceitos pode dar uma idéia do poder explanatório do indicador que está em desenvolvimento. Um simples gráfico bidimensional pode ser a melhor forma de ilustrar essa correlação, ou uma análise de regressão. Uma alta correlação sugere alta qualidade do indicador composto, embora a análise de correlação não deva ser confundida com análise de causalidade. A correlação simplesmente indica que a variação nos dois conjuntos de dados é similar e uma mudança em um indicador não necessariamente conduz a uma mudança no indicador composto ou vice e versa [120].

Correlacionar o indicador composto com outros indicadores que descrevem bem o fenômeno em questão e documentar essa correlação são atribuições que deve ser dominadas pelo construtor ao final dessa fase.

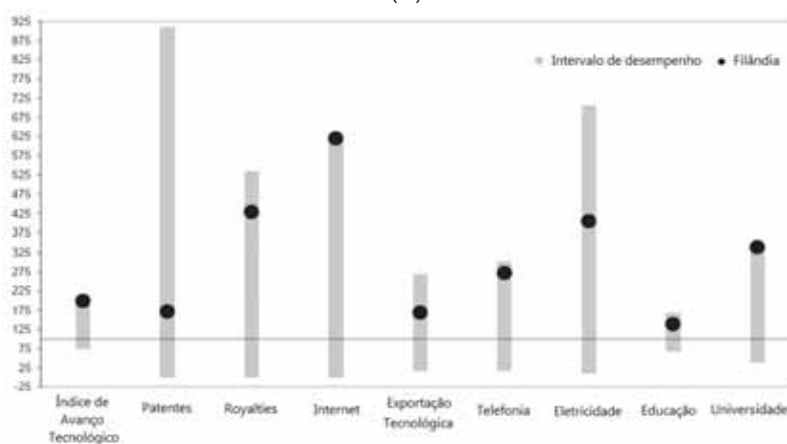
3.3.9 Decomposição do Indicador

Indicadores compostos fornecem um ponto de partida para a análise. Enquanto eles podem ser usados como indicadores de síntese, a sua decomposição possibilita avaliar a contribuição dos componentes e indicadores individuais, possibilitando a extensão da avaliação dos resultados. A decomposição do indicador pode lançar luz sobre o desempenho global de uma determinada alternativa, indicando pontos fortes ou fracos que devam ser trabalhados para melhorar o seu desempenho em uma futura avaliação [120].

A decomposição pode ser ilustrada por uma série de gráficos tais como: gráfico de barras (Figuras 3.5 - A e B), gráfico de radar (Figura 3.6) ou gráfico semáforo (Figura 3.7).



(A)



(B)

Figura 3.5 Exemplo de decomposição em um gráfico de barras - Índice de Avanço Tecnológico. Fonte: NARDO et al. [120].

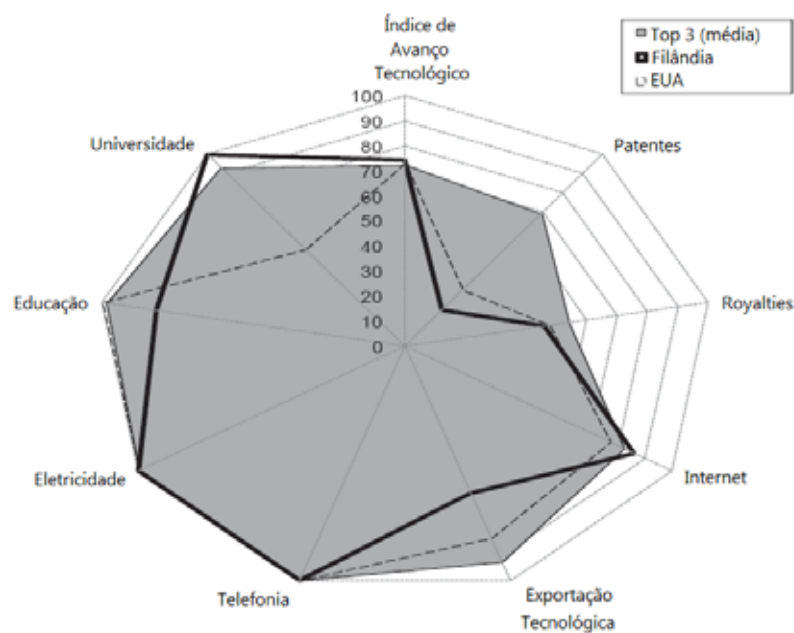


Figura 3.6 Exemplo de decomposição em um gráfico de radar - Índice de Avanço Tecnológico. Fonte: NARDO et al. [120].

	Muito abaixo da média (abaixo 20)	Abaixo da média (20-40)	Média (40-60)	Acima da média (60-80)	Muito acima da média (acima 80)
Filândia					
IAT	74			X	
Patentes	19	X			
Royalties	46		X		
Internet	86				X
Exportação Tecnológica	63			X	
Telefonia	100				X
Eleticidade	100				X
Educação	82				X
Universidade	100				X
Japão					
IAT	70			X	
Patentes	100				X
Royalties	24	X			
Internet	21	X			
Exportação Tecnológica	100				X
Telefonia	100				X
Eleticidade	100				X
Educação	78			X	
Universidade	36	X			

Figura 3.7 Exemplo de decomposição em um gráfico semáforo - Índice de Avanço Tecnológico. Fonte: NARDO et al. [120].

No ISFS foram utilizados gráficos de radar como ferramentas de descrição do impacto de cada uma das dimensões que o constituem. Essa ferramenta foi incorporada ao aplicativo de operacionalização final do ISFS e constitui uma importante opção de análise dos seus resultados.

3.3.10 Apresentação e Visualização dos Resultados

A apresentação dos resultados é uma parte importante do processo de construção de um indicador composto, pois não é uma questão trivial. Os indicadores devem ser capazes de comunicar todo o quadro do fenômeno que se propõem medir, de forma rápida e precisa, sem obscurecer ponto de dados individuais. Existem interessantes formas de mostrar e visualizar os resultados dos indicadores, desde uma simples tabela aos mais complicados gráficos multi-dimensionais e softwares interativos. Os resultados podem ser mostrados através de tabelas, gráficos de barras, linhas, mapas, diagramas de tendência entre outros tipos de representações [120].

No ISFS isto é realizado através de mapas, gráficos e tabelas disponíveis no aplicativo final.

Espera-se que o construtor defina uma coerente forma de representação dos resultados do indicador desenvolvido, de forma clara e transparente, e que não suscite dúvidas aos usuários.

3.4 Métodos de Preenchimento de Dados

A falta de dados é uma ocorrência comum em muitas pesquisas científicas. Determinar a abordagem apropriada na presença de observações incompletas é a maior questão para os

analistas de dados e o desenvolvimento de métodos para a análise de dados incompletos tornou-se nas últimas décadas uma ativa área de pesquisa [57, 82, 83, 124, 125].

As causas das falhas nos dados são numerosas: mau planejamento da pesquisa; impossibilidade de coleta em todo o universo da pesquisa; recusa dos entrevistados de fornecer a informação desejada; confidencialidade que faz com que sejam propositadamente escondidas; medições imperfeitas nos testes ou questionários (variáveis latentes); entre outras [152].

Quando os dados com falhas ocorrem por razões fora de controle, devem-se fazer suposições acerca do processo que os gerou. Essas suposições devem ser explicitadas e a sensibilidade à mudança nos resultados por violação dessas suposições deve ser investigada [152].

O problema das ausência de dados ou “não resposta” é que os objetivos pretendidos pela pesquisa podem não ser alcançados, o que diminui a eficiência nas estimativas por causa da redução do tamanho da amostra. Em muitos casos, os métodos tradicionais de análise de dados completos não podem ser utilizados diretamente, devido a magnitude do número de falhas nos dados. Além disso, podem ocorrer vieses devido a diferenças entre os dados observados e não observados [82, 152].

Portanto, é de suma importância que os dados com falhas não sejam considerados apenas como um problema de análise de dados, mas também como uma questão do planejamento da pesquisa e da interpretação dos resultados, e que os padrões e mecanismos que os geraram sejam investigados e identificados [124, 152].

3.4.1 Padrões e Mecanismos de “não resposta” dos Dados

Muitos conjuntos de dados podem ser arranjados na forma de uma matriz onde as linhas correspondem às unidades observacionais ou participantes da pesquisa e as colunas correspondem aos itens ou variáveis investigadas. Com esse tipo de arranjo é possível identificar vários padrões de “não resposta” nos dados.

Na Figura 3.8 são mostrados todos os possíveis padrões de “não respostas” dos dados. Quando as falhas ocorrem em uma variável, por exemplo, a variável Y , com todas as outras variáveis $X_1, X_2, \dots, X_p$ com dados completos, tem-se o padrão univariado (Figura 3.8 (a)). Quando as falhas ocorrem nas variáveis $Y_1, \dots, Y_p$ que podem ser ordenadas de forma que se Y_j está faltando para uma unidade então $Y_{j+1}, \dots, Y_p$ também está faltando, tem-se o padrão de “não resposta” monotônico. Note que o padrão univariado é um caso particular do padrão monotônico. Se, porém, as falhas ocorrem em qualquer conjunto de variáveis e para qualquer unidade, onde não é possível ordenar as variáveis como nos padrões anteriores, tem-se o padrão

de “não resposta” arbitrário [152].

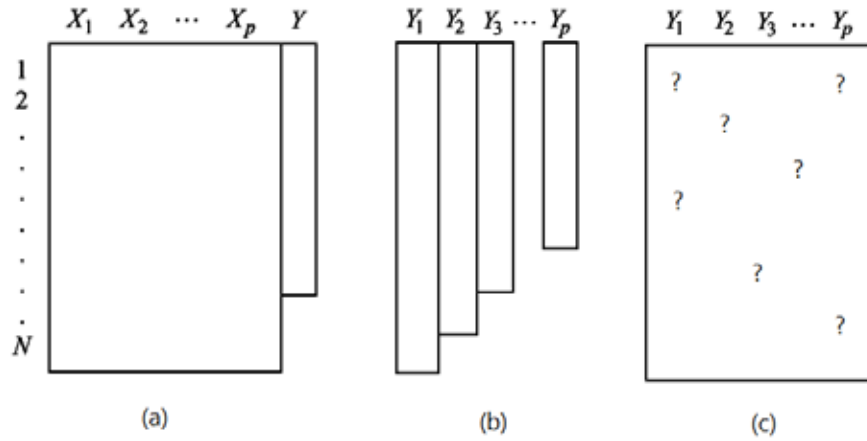


Figura 3.8 Padrões de “não respostas” dos dados. (a) Padrão univariado, (b) Padrão Monotônico e (c) Padrão arbitrário. Fonte: SCHAFFER & GRAHAM [152].

Para qualquer conjunto de dados pode-se definir uma variável indicadora R que identifica o que é conhecido (observado) e o que é desconhecido (não observado). A distribuição dessa variável depende da complexidade do padrão de “não resposta” dos dados.

Então, por exemplo, na Figura 3.8 (a), R pode ser definida como uma variável binária simples indicando se Y é observado ($R = 1$) ou se Y não é observado ($R = 0$). Na Figura 3.8 (b), R pode ser representada por uma variável inteira ($R = 1, 2, \dots, p$) indicando o maior valor do índice j para o qual Y_j é observado. E na Figura 3.8 (c), R pode ser uma matriz binária de mesma dimensão da matriz de dados, cujos elementos são representados por 1 ou 0, dependendo se os correspondentes valores dos dados são observados ou não observados.

Na literatura estatística, a distribuição de R é algumas vezes chamada de mecanismo de não resposta, o que pode levar a confusão, uma vez que o termo mecanismo sugere um processo do mundo real nos quais existem dados que são registrados e outros que não. Sabe-se que a descrição precisa de todas as causas potenciais ou razões para falhas nos dados é uma condição não realista. Logo, R será melhor compreendida como um dispositivo matemático para descrever as taxas e padrões das falhas nos dados, bem como para capturar as possíveis relações entre essas falhas [145, 152].

A partir dessas considerações e, principalmente, em função da natureza das relações entre as falhas, pode-se classificar R nos seguinte tipos de mecanismos de “não resposta” [152]:

1. Falhas completamente ao acaso (*MCAR - Missing Completely at Random*) - Quando as razões para as falhas não são relacionadas a quaisquer respostas, incluindo o próprio valor

que falta. Esse é o mecanismo mais restritivo pois impõe que a probabilidade de “não resposta” seja a mesma para diversas situações.

2. Falhas ao acaso (*MAR* - *Missing at Random*) - Quando o padrão das falhas é previsível a partir das outras variáveis e não é devido à variável específica na qual os dados estão faltando.
3. Falhas não aleatórias (*MNAR* - *Missing not at Random*) - Quando os dados que faltam estão relacionados com os valores não observados, mesmo quando controlados por outras variáveis da análise.

Denotando por Y_{com} o conjunto de dados completos, de forma que $Y_{com} = (Y_{obs}, Y_{mis})$ onde Y_{obs} e Y_{mis} são as partes observadas e com falhas desse conjunto, respectivamente, então os dados com falhas serão MAR se a distribuição não depende de Y_{mis} , ou seja:

$$P(R | Y_{com}) = P(R | Y_{obs}) \quad (3.13)$$

Por outro lado, os dados com falhas pertencerão ao mecanismo MCAR, se a distribuição não depende de Y_{mis} nem de Y_{obs} , ou seja:

$$P(R | Y_{com}) = P(R) \quad (3.14)$$

Quando a Equação 3.13 é violada e a distribuição depende de Y_{mis} então os dados com falhas são ditos serem MNAR.

Na Figura 3.9 é apresentada uma visão gráfica desses mecanismos. Nela, X representa as variáveis completamente observadas; Y as variáveis que apresentam falhas; Z representa a componente das causas da omissão não relacionadas a X e Y ; e R representa a distribuição das falhas.

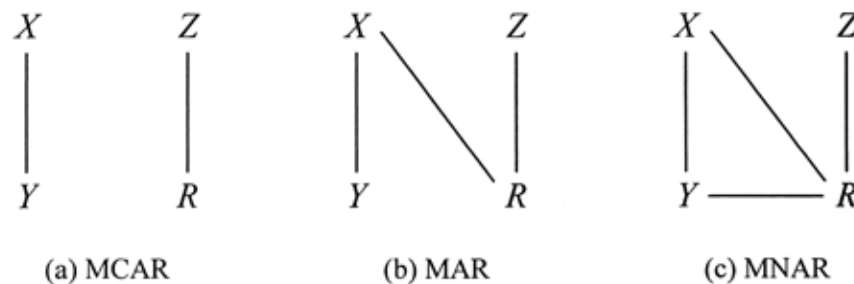


Figura 3.9 Representação gráfica dos mecanismos de “não resposta” dos dados: (a) MCAR, (b) MAR e (c) MNAR. Fonte: SCHAFFER & GRAHAM [152].

Dados MCAR e MAR também são chamados de *dados com falhas ignoráveis* enquanto que os dados NMAR são chamados de *não-ignoráveis* [82,124,152]. Dados NMAR são mais difíceis de tratar em uma análise, enquanto que os dados MCAR são mais fáceis (HAREL, 2001 *apud* NUNES [124]). Quando os dados são NMAR, uma estimação eficiente com dados com falhas não ignoráveis requer um bom conhecimento prévio do mecanismo que gerou a não resposta, pois os dados não contêm informações sobre qual modelo não ignorável seria adequado. Por isso, os resultados seriam sensíveis aos modelos assumidos (HAREL, 2007 *apud* NUNES [124]).

Muitos dos métodos disponíveis para tratar dados com falhas assumem que os dados são MAR. Essa suposição ocupa uma importante posição no ambiente de dados com falhas, não porque seja mais plausível na prática, mas porque representa uma condição mais geral sob a qual inferências válidas podem ser obtidas sem fazer referência ao mecanismo de não-resposta [94]. Embora a suposição MAR torne a análise dos dados com falhas exequível e mais simples, essa suposição não pode ser testada, como se faz com a suposição da normalidade dos dados. Porém, tem sido apontado na literatura, que se são incluídas variáveis suficientes no modelo de preenchimento dos dados, mais próximo se chega dessa suposição [151].

3.4.2 Métodos de Preenchimento Único dos Dados

Uma das primeiras abordagens utilizadas para lidar com os dados com falhas foram os métodos de preenchimento único ou simples dos dados (ENGELS, 2003 *apud* NUNES [124]). Estes métodos constituem na substituição das falhas nos dados uma única vez por um valor obtido em um dos métodos de preenchimento único: média/mediana/moda, “hot deck”, regressão linear, entre outros.

No entanto, esses métodos podem reduzir a variabilidade amostral por imputarem valores centrais da distribuição para todos os dados que faltam. Embora não tornem os resultados viesados, superestimam a precisão, isto é, produzem erros padrão muito pequenos, o que não ocorre com a abordagem de preenchimento múltiplo dos dados. Uma outra desvantagem desses métodos é não levar em consideração as incertezas associadas ao processo de preenchimento dos dados (LITTLE & RUBIN *apud* NUNES [124]).

A seguir serão descritos os métodos de preenchimento único dos dados mais utilizados.

a) Média/mediana/moda

Nesse método os dados com falhas são substituídos pela média, mediana ou moda. Quando se utiliza a média pode-se optar pela média de todos os valores observados ou de um grupo

similar ao do dado com falha. A mediana é uma medida usada principalmente se os dados da amostra contiverem *outliers*. E a moda é indicada se a variável for categórica não ordinal [9,124].

b) “Hot Deck”

O princípio deste método é o uso de dados observados (doadores) para fornecer os valores que preencherão os dados com falhas. O procedimento consiste na procura de um doador que se aproxima do dado com falha, em função de variáveis auxiliares e a substituição dessa falha por esse valor. A menor distância adjacente ou “vizinho mais próximo” também pode ser utilizada como critério para escolha do doador. Esse método também é utilizado na abordagem de preenchimento múltiplo dos dados [9,124,145].

c) Regressão Linear

Nesse método, os dados com falhas são preditos através de uma regressão linear simples ou múltipla. Pode ser utilizada uma ou mais variáveis existentes para prever os dados com falhas de outra variável altamente correlacionada com as anteriores [124].

d) Estimativa de Máxima Verossimilhança

Esse método utiliza o Algoritmo de Máxima Expectância (*EM - Expectation Maximization*), para computar a estimativa de Máxima Verossimilhança para os dados com falhas. O algoritmo EM consiste de dois passos: *E* (Estimação) e *M* (Maximização). No passo *E*, os dados com falhas são estimados para completar a matriz de dados. No passo *M*, a função de verossimilhança é maximizada sob a suposição que as falhas dos dados são conhecidas. Nesse passo, ocorre a estimativa dos parâmetros da função de verossimilhança e esses parâmetros são usados para fazer inferência no passo *E*, e assim, sucessivamente até que haja convergência [101,124].

e) Métodos de Preenchimento Único para Dados Longitudinais

Esses métodos envolvem o preenchimento das falhas de dados com observações repetidas no tempo (Dados Longitudinais). Os principais métodos são:

1. Média/mediana - utiliza a média/mediana dos valores prévios da variável para substituir o valor que está faltando.
2. Último valor observado (*LOCF - Last Observation Carried Forward*) - O último valor observado da variável é utilizado para substituir o valor que está faltando.

3. Próximo valor observado (*NOCB - Next Observation Carried Backward*) - O valor posterior ao que está faltando é utilizado como valor de substituição.
4. Último e próximo valores - Utiliza a média do último e do próximo valor ao que está faltando como valor de substituição.

3.4.3 Métodos de Preenchimento Múltiplo dos Dados

Até a década de 1970 os dados com falhas eram tratados por abordagens tradicionais de preenchimento dos dados, tais como os métodos de preenchimento único citados anteriormente. Em 1987, RUBIN introduziu a idéia do preenchimento múltiplo dos dados (*MI - Multiple Imputation*), que consiste na substituição do valor do dado que está faltando por m ($m > 1$) valores simulados antes da análise. A partir dessa abordagem são obtidos m banco de dados completos e cada conjunto de dados pode assim ser analisado utilizando-se os procedimentos de análise estatísticas tradicionais que são usados para dados completos [82, 83, 94, 124, 152].

Nos últimos anos, a MI tem sido muito utilizada, principalmente por sua implementação computacional tanto em softwares comerciais quanto em softwares livres [82, 83, 94, 124, 152].

Uma das principais características da MI é a possibilidade de explorar as incertezas associadas tanto aos valores que estão preenchendo as falhas, como às do próprio processo de preenchimento das falhas. Essas incertezas são de dois tipos: 1) a variabilidade amostral, assumindo que as razões para a não resposta são conhecidas; e, 2) a variabilidade sobre as razões para a não resposta dos dados [124, 145].

RUBIN [145] afirmou que existem três vantagens extremamente importantes da MI em relação ao preenchimento único das falhas nos dados:

1. O aumento da eficiência da estimação, quando o preenchimento é realizado aleatoriamente;
2. As inferências válidas são obtidas simplesmente combinando-se as inferências dos dados completos, após os m preenchimentos, de uma maneira relativamente simples;
3. É possível o estudo da sensibilidade das inferências para vários modelos de não resposta considerados.

Podem-se também citar três desvantagens da MI: 1) É necessário mais trabalho para produzir os dados a serem utilizados na substituição dos dados com falhas, 2) É necessário mais espaço para armazenar o conjunto de dados produzidos e 3) É necessário mais esforço na análise dos bancos de dados completos provenientes do preenchimento múltiplo. No entanto, quando m

não é grande, essas desvantagens não são significativas e também podem ser minimizadas com a utilização dos métodos de MI disponíveis nos pacotes de softwares comerciais ou livres [124,145].

A MI é constituída, basicamente, de três passos [82,124]:

1. Obtenção de m bancos de dados completos através de técnicas de preenchimento múltiplo adequadas ao problema em questão;
2. Análise dos m bancos de dados gerados por um método estatístico tradicional;
3. Combinação dos m resultados encontrados no passo anterior de um modo simples e adequado para se obter a chamada inferência de preenchimento repetido.

O primeiro passo é o mais crítico, pois a técnica de preenchimento utilizada tem que preservar a relação entre os dados com falhas e os observados e tem que levar em conta o mecanismo de não resposta (MCAR, MAR ou NMAR) e o padrão de não resposta (monotônico ou não monotônico) dos dados. No passo dois, os m bancos de dados gerados são analisados separadamente por um método tradicional de análise estatística dos dados. E, no passo três, esses resultados são combinados utilizando-se as regras propostas por RUBIN [82,124].

A seguir serão detalhados os principais métodos de MI utilizados no passo um, em função do padrão de não resposta dos dados, e mais adiante as regras de RUBIN e um roteiro para escolha da abordagem de preenchimento das falhas a se utilizar.

- Métodos de MI para o padrão de não resposta monotônico

a) Método da Regressão Linear Bayesiana (*BLR - Bayesian Linear Regression*)

Quando os dados com falhas seguem o padrão de não resposta monotônico (Figura 3.8 (a) e (b)), a função conjunta de verossimilhança dos dados observados pode ser fatorada num produto de funções de verossimilhança independentes, ou seja:

$$L(\phi_1, \dots, \phi_p \mid Y_{obs}) = \prod_{j=1}^p L(\phi_j \mid Y_{obs}) \quad (3.15)$$

onde,

$$L(\phi_j \mid Y_{obs}) \propto \prod_{i=1}^{n_j} P(y_{ij} \mid y_{i,1}, \dots, y_{i,j-1}, \phi_j) \quad (3.16)$$

e, $P(Y_j \mid Y_1, \dots, Y_{j-1}, \phi_j)$ é a distribuição condicional de Y_j dados $Y_1, \dots, Y_{j-1}$ e ϕ_j é a distribuição condicional dos parâmetros. Portanto, os dados com falhas com padrão monotônico

podem ser preenchidos a partir de distribuições univariadas independentes, dados os valores prévios observados [189].

Se a normalidade multivariada é assumida para as variáveis respostas $Y_1, \dots, Y_p$, a função de verossimilhança $L(\phi_j | Y_{obs})$ torna-se uma regressão linear de Y_j em $Y_1, \dots, Y_{j-1}$ baseada nas n_j primeiras observações, e a distribuição condicional dos parâmetros ϕ_j torna-se os coeficientes de regressão e a variância residual [189].

Portanto, os valores com falhas de Y_j podem ser preenchidos pelos valores preditos deste modelo de regressão linear, dados os valores observados $Y_1, \dots, Y_{j-1}$ e os parâmetros de regressão simulados que são delineados aleatoriamente da distribuição dos dados observados a *posteriori* [189].

Pode-se descrever formalmente esse método da seguinte forma [189]:

Seja X a matriz de dados para $Y_1, \dots, Y_{j-1}$ aumentada de uma coluna correspondente ao intercepto, e X_{obs} e X_{mis} as linhas correspondentes a $Y_{obs,j}$ e $Y_{mis,j}$, respectivamente. Desde que os dados com falhas tenham um padrão de não resposta monotônico, as primeiras n_j observações de Y_j são observadas e o restante das $(n_j + 1)$ até n observações estão faltando. Assumindo que o modelo de probabilidade de Y_j dados $Y_1, \dots, Y_{j-1}$ é uma distribuição normal univariada $Y_j \sim N_1(\mu_j, \sigma_j^2)$, onde $\mu_j = \beta_0 + \beta_1 Y_1 + \dots + \beta_{j-1} Y_{j-1}$, a função de verossimilhança dos dados observados dos parâmetros de regressão $\phi_j = (\beta_0, \beta_1, \dots, \beta_{j-1}, \sigma_j^2)$ é

$$L(\mu_j, \sigma_j^2 | Y_{obs}) \propto \sigma_j^{-n_j} \exp \left\{ -\frac{1}{2\sigma_j^2} \sum_{i=1}^{n_j} (y_{ij} - X_{obs(i)}\beta)^2 \right\}, \quad (3.17)$$

onde $\beta = (\beta_0, \beta_1, \dots, \beta_{j-1})$ é um vetor de coeficientes de regressão.

Suponha que uma distribuição não informativa a *priori* $\pi(\phi_j) \propto \sigma_j^{-1}$ é assumida para (β, σ_j^2) . Após algumas manipulações, a distribuição posterior dos dados observados de (β, σ_j^2) pode ser expressa como:

$$\sigma_j^{-\frac{p}{2}} \exp \left\{ -\frac{1}{2\sigma_j^2} (\beta - \hat{\beta})^T (X_{obs}^T X_{obs}) (\beta - \hat{\beta}) \right\} \times \sigma_j^{-\left(\frac{n_j - p}{2}\right) - 1} \exp \left\{ -\frac{1}{2\sigma_j^2} \sum_{i=1}^{n_j} (y_{ij} - X_{obs(i)}\hat{\beta})^2 \right\} \quad (3.18)$$

que é o produto de uma distribuição normal multivariada e uma distribuição qui-quadrado escalada invertida,

$$\beta | Y_{obs}, \sigma_j^2 \sim N_j(\hat{\beta}, \sigma_j^2 (X_{obs}^T X_{obs})^T), \quad (3.19)$$

$$\sigma_j^2 \mid Y_{obs} \sim \hat{\varepsilon}^T \hat{\varepsilon} \chi_{n_j-p}^{-2}, \quad (3.20)$$

onde $\hat{\beta} = (X_{obs}^T X_{obs})^{-1} X_{obs}^T Y_{obs,j}$ é o estimador de máxima verossimilhança de β das funções de verossimilhança dos dados observados, e $\hat{\varepsilon} = Y_{obs,j} - X_{obs} \hat{\beta}$ é o vetor residual.

Os valores de (β, σ_j^2) podem ser simulados da distribuição posterior dos dados observados (Equações 3.19 e 3.20) fazendo $\tilde{\sigma}_j^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^*}$, onde σ^* é uma variável aleatória $\chi_{n_j-p}^2$ e, então, simulando $\tilde{\beta}$ de (3.19) dados $\tilde{\sigma}_j^2$. Após a simulação de $(\tilde{\beta}, \tilde{\sigma}_j^2)$ ter sido feita da distribuição posterior dos dados observados, os valores com falhas de Y_j são preenchidos aleatoriamente da distribuição preditiva condicional $P(Y_{mis,j} \mid Y_{obs}, \tilde{\beta}, \tilde{\sigma}_j^2)$, que é uma distribuição normal univariada $N_1(X_{mis} \tilde{\beta}, \tilde{\sigma}_j^2)$. Esse procedimento é repetido até se obterem os m conjuntos de dados preenchidos, ou seja, $(\tilde{\beta}^{(t)}, \tilde{\sigma}_j^{2(t)})$, com $t = 1, \dots, m$.

b) Método da Média Preditiva (*PMM - Predictive Mean Matching*)

Este método é semelhante ao método anterior (BLR), entretanto, há uma modificação no último procedimento de obtenção dos valores com falhas. Para cada valor com falha de Y_j predito, por exemplo, Y_{j*} , encontra-se o valor de Y_j observado que seja mais próximo ao valor de Y_{j*} , e esse valor é utilizado para o preenchimento do dado com falha. Pode-se dizer que neste método há um componente “*hot-deck*”, já que somente são usados para o preenchimento das falhas valores observados.

- Método de MI para o padrão de não resposta não monotônico

a) Método de Monte Carlo baseado em Cadeias de Markov (*MCMC - Markov Chain Monte Carlo*)

O método MCMC constrói uma Cadeia de Markov a partir da simulação de amostras pseudo-aleatórias de uma distribuição condicional conjunta de Y_{mis} e θ dado Y_{obs} , ou seja,

$$P(Y_{mis}, \theta \mid Y_{obs}). \quad (3.21)$$

O MCMC é indicado para dados com padrão de não-resposta não monotônico ou arbitrário e, semelhante aos métodos BLR e PMM é requerida a suposição da normalidade multivariada dos dados [82, 189]. Embora em muitos casos a suposição da normalidade multivariada dos dados não seja realista, testes simulados tem demonstrado relativa robustez para pequenos desvios dessa suposição (LITTLE & RUBIN, 1987 *apud* YALE [186]).

O método MCMC consiste dos seguintes passos:

- Passo Inicial: Trocar os dados com falhas Y_{mis} por alguns valores assumidos aleatoriamente, e simular θ da distribuição posterior completa resultante $P(\theta | Y_{obs}, Y_{mis})$;
- Passo I (*Imputation-step*): Simular $Y_{mis}^{(t+1)} \sim P(Y_{mis} | Y_{obs}, \theta^{(t)})$;
- Passo P (*Parameter-step*): Simular $\theta^{(t+1)} \sim P(\theta | Y_{obs}, Y_{mis}^{(t+1)})$.

O procedimento é iterativo e na t -ésima iteração tem-se a seguinte Cadeia de Markov:

$$\{(\theta^{(1)}, Y^{(1)}), \dots, (\theta^{(t+1)}, Y^{(t+1)})\}, \quad (3.22)$$

que converge para a distribuição posterior de interesse. Quando t for suficientemente grande, tem-se, na iteração final, $Y_{mis}^{(t)}$ representando os valores que devem preencher os dados com falhas, formando portanto, o conjunto de dados completo.

- As regras de RUBIN

As regras de RUBIN constituem o passo três da MI e são normas simples para combinar as análises estatísticas realizadas no passo dois. São utilizadas independentemente do método de preenchimento das falhas adotado [124].

A ideia é que, a partir de cada análise, obtenham-se estimativas para o parâmetro de interesse Q_j , com $j = 1, \dots, m$. Q_j pode ser qualquer medida escalar a ser estimada: média, correlação, coeficiente de regressão, razão de chance, entre outras. Portanto a estimativa combinada será a média das estimativas individuais:

$$\bar{Q} = \frac{1}{m} \sum_{j=1}^m \hat{Q}_j \quad (3.23)$$

E a variância combinada é obtida em função da variância dentro dos preenchimentos ($\bar{U} = \frac{1}{m} \sum_{j=1}^m U_j$) e a variância entre os preenchimentos ($B = \frac{1}{m-1} \sum_{j=1}^m (\hat{Q}_j - \bar{Q})^2$) da seguinte forma:

$$T = \bar{U} + (1 + \frac{1}{m})B. \quad (3.24)$$

- Um roteiro para escolha da abordagem de preenchimento das falhas

É possível definir linhas gerais para escolha da abordagem e do método de preenchimento de dados com falhas, em função da percentagem das falhas nos dados (PF_{dados}) em qualquer uma das variáveis, da seguinte forma (HARREL, 2001 *apud* NUNES [124]):

1. Se $PF_{dados} \leq 5\%$ - Neste caso pode-se utilizar o preenchimento único ou analisar somente os dados completos;
2. Se $5\% < PF_{dados} < 15\%$ - O preenchimento único poderá ser utilizado, entretanto o preenchimento múltiplo também é indicado;
3. Se $PF_{dados} \geq 15\%$ - O preenchimento múltiplo é indicado na maior parte dos casos.

3.4.4 Aplicativos Computacionais que Implementam o Preenchimento Múltiplo dos Dados

No início da aplicação da MI, na década de 1980, por razões práticas usavam-se valores muito pequenos para o m . Essa limitação desapareceu com os avanços computacionais das últimas décadas, tanto de hardware como de software, sendo atualmente viável para alguns tipos de problemas utilizar-se m com valores entre 100 e 200 [94]. Porém, a experiência prática demonstra que para quantidades modestas de dados com falhas, ou seja, menos de 30%, valores para m entre 5 e 10 são suficientes para fornecer estimativas razoáveis dos parâmetros de interesse [94, 145, 186].

A implementação computacional do ISFS (*ISFS_Sist*), foi construída de forma que o usuário defina o valor do m . Isto possibilita que se trabalhe com percentuais de falhas nos dados superiores a 30%. Maiores detalhes serão dados nos capítulos posteriores dessa tese.

Atualmente, muitos softwares, tanto os comerciais como de licença livre, disponibilizam a MI. Isto facilitou e aumentou a utilização desta técnica em ambientes permeados de falhas nos dados. Dentre os mais utilizados podem-se citar: Amélia, CAT, EMCOV, NORM, Hmisc, MICE, MIXED, IVEware e PAN (softwares livres) e HLM, Mplus, SAS, SOLAS, S-Plus, SPSS e Stata (Softwares comerciais) [1, 82, 83].

Uma análise de desempenho de diversos aplicativos computacionais que disponibilizam a MI é feita em HORTON & LIPSITZ [82], ACOCK [1] e HORTON & KLEINMAN [83]. Esses autores afirmam que nenhum aplicativo é claramente superior, porém existem algumas características que devem ser levadas em consideração quando da utilização destes softwares. A primeira é o fator custo, sendo que nesse quesito, os softwares livres levam vantagem em relação aos softwares comerciais. A segunda é a possibilidade de integração desses aplicativos aos programas dos próprios usuários. O MICE é um dos aplicativos computacionais que apresentam esses dois diferenciais: é um software livre, acoplado à plataforma do software R; e, possibilita que o usuário ou construtor integre as funções do MICE ao seu ambiente de trabalho ou pro-

gramas. O ISFS foi implementado utilizando-se o MICE, principalmente devido a essas duas facilidades, conforme será detalhado nos capítulos posteriores dessa tese.

No Quadro 3.5, os softwares citados acima são listados destacando-se a sua homepage e a sua capacidade de aplicação: preenchimento único, múltiplo ou ambos.

Aplicativo	Homepage	Preenchimento único	Preenchimento Múltiplo
Softwares Livres			
Amelia	http://gking.harvard.edu/amelia/		✓
CAT	http://www.stat.psu.edu/~jls/misoftwa.html#aut		✓
EMCOV	http://www.smallwaters.com/weblinks/	✓	
NORM	http://www.stat.psu.edu/~jls/misoftwa.html#aut	✓	✓
Hmisc	http://cran.r-project.org/web/packages/Hmisc/index.html (Livre com R, comercial com S-Plus)		✓
MICE	http://www.multiple-imputation.com (Livre com R, comercial com S-Plus)		✓
MIXED	http://www.stat.psu.edu/~jls/misoftwa.html#aut (Livre com R, comercial com S-Plus)	✓	✓
IVEware	http://www.isr.umich.edu/src/smp/ive/	✓	✓
PAN	http://www.stat.psu.edu/~jls/misoftwa.html#aut (Livre com R, comercial com S-Plus)	✓	✓
Softwares Comerciais			
HLM	http://www.scienceplus.nl/hlm		✓
Mplus	http://www.statmodel.com		✓
SAS	http://www.sas.com		✓
SOLAS	http://www.statsol.ie/solas/imputationtechniques.htm	✓	✓
S-Plus	http://www.stat.psu.edu/~jls/misoftwa.html#aut	✓	✓
SPSS	http://www.spss.com (módulo opcional)	✓	
Stata	http://www.stata.com		✓

Quadro 3.5 Aplicativos computacionais utilizados no preenchimento único e múltiplo dos dados com falhas. Fonte: Adaptado de ACOCK [1].

3.4.5 MICE - Multiple Imputation by Chained Equations

Espera-se que um modelo eficiente de preenchimento de falhas atenda aos seguintes princípios: considere o processo de criação dos dados com falhas, preserve a relação nos dados e a

incerteza sobre essas relações. A esperança é que a adesão a esses princípios renda preenchimentos que são estatisticamente corretos para uma ampla gama de parâmetros Q [145]. No entanto, quando se está lidando com dados com falhas podem surgir alguns problemas, tais como [175,177]:

- Para um dado Y_j , os preditores $Y_1, \dots, Y_{j-1}$ usados no modelo de preenchimento podem eles próprios ser incompletos;
- Pode ocorrer dependência circular, ou seja, Y_1 depende de Y_2 e Y_2 depende de Y_1 devido a correlação entre essas duas variáveis, mesmo considerando outras variáveis;
- As variáveis podem ser de tipos diferentes, por exemplo, binárias, contínuas, ordenadas, não ordenadas, entre outras; fazendo com que a aplicação de modelos teoricamente convenientes seja inadequada;
- A relação entre Y_j e $Y_1, \dots, Y_{j-1}$ pode ser complexa;
- O preenchimento pode criar combinações impossíveis ou destruir relações determinísticas dos dados;
- O preenchimento pode ser absurdo;
- O modelo para Q que será aplicado nos dados preenchidos pode ser ainda desconhecido.

Partindo-se da complexidade real dos dados, destacada acima, é conveniente especificar um modelo de preenchimento das falhas nos dados para cada variável em separado, o que levou ao desenvolvimento do MICE. Nesse modelo a especificação ocorre em um nível bem compreendido pelo usuário, isto é, ao nível da variável, o que possibilita a utilização de técnicas de preenchimento de falhas já bem desenvolvidas [175–177].

Formalmente o MICE pode ser assim definido:

Seja $Y = (Y_1, \dots, Y_p)$ um conjunto de p variáveis aleatórias, onde cada variável $Y_j = (Y_j^{obs}, Y_j^{mis})$ com Y_j^{obs} representando as informações observadas e Y_j^{mis} as informações ausentes ou falhas, com $j = 1, \dots, p$. Assuma que a distribuição multivariada de Y é completamente especificada por θ , um vetor de parâmetros desconhecidos. O problema é como obter uma distribuição multivariada de θ ou explicitamente ou implicitamente. As equações encadeadas se propõem a obter uma distribuição posterior de θ pela amostragem iterativa de distribuições condicionais:

$$\begin{aligned}
&P(Y_1 \mid Y_{-1}, \theta_1) \\
&\vdots \\
&P(Y_p \mid Y_{-p}, \theta_p).
\end{aligned} \tag{3.25}$$

Os parâmetros $\theta_1, \dots, \theta_p$ são específicos para as respectivas densidades condicionais e não são necessariamente o produto da fatoração da distribuição conjunta $P(Y \mid \theta)$. Partindo de um simples delineamento das distribuições marginais observadas, a t -ésima iteração das equações encadeadas é uma amostragem de Gibbs que sucessivamente obtém:

$$\begin{aligned}
\theta_1^{*(t)} &\sim P(\theta_1 \mid Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}) \\
Y_1^{*(t)} &\sim P(Y_1 \mid Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_1^{*(t)}) \\
&\vdots \\
\theta_p^{*(t)} &\sim P(\theta_p \mid Y_p^{obs}, Y_1^{(t-1)}, \dots, Y_{p-1}^{(t-1)}) \\
Y_p^{*(t)} &\sim P(Y_p \mid Y_p^{obs}, Y_1^{(t-1)}, \dots, Y_{p-1}^{(t-1)}, \theta_p^{*(t)})
\end{aligned} \tag{3.26}$$

onde $Y_j^{(t)} = (Y_j^{obs}, Y_j^{*(t)})$ é a j -ésima variável preenchida na iteração t . Observe que nos preenchimentos anteriores $Y_j^{*(t-1)}$ somente consideram $Y_j^{*(t)}$ através da relação com outras variáveis, e não diretamente. A convergência é portanto relativamente rápida, ao contrário de outros métodos como o MCMC. É importante monitorar a convergência; no entanto, em muitas situações, um número de iterações relativamente pequeno, entre 10 e 20, é suficiente [175, 177].

Basicamente, o MICE é um software para preenchimento múltiplo de dados com falhas que disponibiliza uma série de métodos de preenchimento de dados e contém funções para as três fases da MI: geração do preenchimento múltiplo, análise dos dados preenchidos e a combinação dos resultados das análises. O software surgiu no ano de 2000 no pacote do software S-PLUS e, no ano de 2001, no pacote do software R [1, 82, 83, 175, 177].

O procedimento utilizado pelo MICE inicia-se com um conjunto de dados incompletos na estrutura de dados *data frame*, que são preenchidos para gerar k conjuntos de dados completos com a função *mice*. Cada conjunto de dados é então analisado separadamente por um método estatístico padrão, por exemplo, regressão linear, a partir da função *with*, e, os k resultados analisados em separado são então reunidos em um único resultado, utilizando-se as regras de RUBIN a partir da função *pool*, culminando em um modelo de regressão linear final, que adequadamente reflete a quantidade de incerteza nas estimativas. O pacote ainda apresenta classes especiais para armazenamento dos dados em cada fase do processo (*mids*, *mira* e *mipo*) (Figura 3.10).

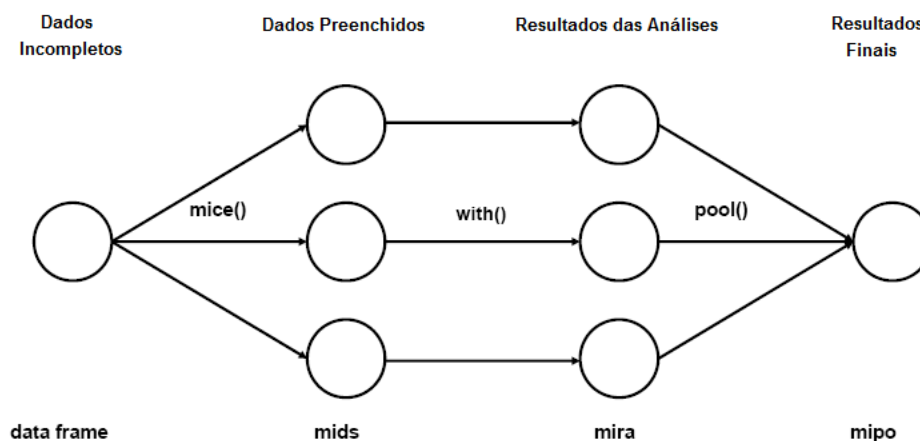


Figura 3.10 Representação esquemática dos passos do preenchimento múltiplo utilizados no MICE. Fonte: VAN BUUREN & OUDSHOORN [175].

O MICE possibilita a utilização de uma série de métodos de preenchimento de falhas nos dados (Quadro 3.6), de acordo com o padrão de não resposta e o tipo de variável que se objetiva preencher. O padrão (*default*) do MICE é método PMM se a variável for numérica e regressão logística ou politômica não ordenada se a variável for categórica com dois ou mais níveis, respectivamente. A função *md.pattern* permite a investigação do padrão de não resposta dos dados e, assim, proporciona ao construtor/analista subsídios para a escolha do método mais adequado ao seu problema.

Método	Descrição	Tipo Variável	Default
pmm	Método da média preditiva (PMM)	Numérica	✓
norm	Regressão linear bayesiana (BLR)	Numérica	
norm.nob	Regressão linear não bayesiana	Numérica	
mean	Preenchimento pela média incondicional	Numérica	
2l.norm	Modelo linear em dois níveis	Numérica	
logreg	Regressão logística	Categórica, 2 níveis	✓
polyreg	Regressão politômica não ordenada	Categórica, > 2 níveis	✓
lda	Análise discriminante linear	Categórica	
sample	Amostragem aleatória dos dados observados (MCMC)	Qualquer	

Quadro 3.6 Métodos de preenchimento das falhas disponíveis no MICE. Fonte: VAN BUUREN & OUDSHOORN [176,177].

Muitas vezes existe a necessidade de transformar, combinar e recodificar as várias versões dos dados. No caso de dados incompletos, pode-se preencher o original e transformar o original preenchido posteriormente ou transformar primeiro e preencher depois. Se, no entanto,

tanto os dados originais como os transformados forem necessários no algoritmo de preenchimento, nenhuma dessas abordagens funciona, porque não se tem certeza qual é a verdadeira transformação entre os valores originais e os transformados. O MICE apresenta um mecanismo especial chamado preenchimento passivo (*passive imputation*) para lidar com essas situações. O *passive imputation* mantém a consistência entre as diferentes transformações do mesmo dado, garantindo que a transformação sempre dependa do mais recente preenchimento gerado nos dados não transformados. Isso é conseguido especificando-se um $\sim$ (til) no primeiro caracter do método de preenchimento [177].

Outra funcionalidade de destaque do MICE é a possibilidade do usuário construir as suas próprias funções de preenchimento das falhas. Isso possibilita que se utilizem funções de preenchimento específicas para certos tipos de variáveis, chamando essas funções no algoritmo de amostragem de Gibbs.

3.5 A Técnica de Winsorização dos Dados

A winsorização é uma técnica estatística que foi originalmente proposta pelo bioestatístico C. P. Winsor e que permite a suavização das caudas de uma distribuição [21, 186].

O procedimento consiste na substituição dos valores extremos (acima ou abaixo de percentis mínimos e máximos pré-definidos), substituindo-se pelos valores desses limites [21, 186].

É pratica comum que a substituição seja feita na mesma proporção, portanto se for definido $\alpha = 0,05$, tem-se duas substituições, uma no quantil 0,975 e outra no quantil 0,025. Esse método objetiva, principalmente, a eliminação de *outliers*, que podem ter sido originados no procedimento do preenchimento anterior ou serem provenientes dos dados originais.

Capítulo 4

A CONSTRUÇÃO DO ÍNDICE DE SUSCEPTIBILIDADE AO FENÔMENO DA SECA (ISFS)

Neste capítulo são apresentados os procedimentos para construção do ISFS a partir da definição conceitual dos sub-indicadores, a seleção das variáveis constitutivas e o método de normalização utilizado para a padronização da pontuação de cada um. Ademais é apresentada a formalização matemática e a estrutura do modelo proposto.

4.1 Considerações Iniciais

O Índice de Susceptibilidade ao Fenômeno da Seca (ISFS) foi construído a partir do Indicador de Emergência (IMERG) desenvolvido para o Ceará (Figura 2.25), tomando como base metodológica os procedimentos para construção de indicadores compostos, descritos no item 3.3 do capítulo anterior. Na construção do ISFS levou-se em consideração principalmente as variáveis disponíveis e/ou alternativas que representassem uma melhoria qualitativa em relação ao parâmetro originalmente utilizado no IMERG e a disponibilidade da informação.

Durante o desenvolvimento do trabalho constatou-se a escassez de informações e de trabalhos de pesquisa desta natureza para o Rio Grande do Norte. Os trabalhos que foram encontrados após uma exaustiva busca bibliográfica ou têm foco restrito a determinadas regiões ou não cobrem todo o Estado.

O sub-indicadores PV131, PV132, PV133 e PV21 foram os que, relativamente, sofreram as menores alterações em relação à definição utilizada no IMERG, dada a disponibilidade das

informações para o RN. Os demais parâmetros sofreram alterações em grau e natureza diversos, inclusive com a inclusão de novas variáveis que qualificam melhor o resultado final obtido com o índice.

4.2 Seleção e Definição das Variáveis e Sub-indicadores

4.2.1 Sub-indicador: Natureza Física - PV11

Na composição deste sub-indicador no IMERG, os autores consideraram a aptidão agrícola das terras baseada na metodologia de RAMALHO & BEEK [137] (ver Capítulo 2, item 2.6.3 (a)), e obtiveram a escala de pontuação a partir da aplicação da metodologia MACBETH (ver Capítulo 2, Tabela 2.1). Na adaptação deste sub-indicador para o ISFS, primeiro foi realizado um extenso levantamento bibliográfico dos trabalhos existentes sobre solos e aptidão agrícola das terras do Rio Grande do Norte. Foi encontrada a seguinte relação de trabalhos, em ordem cronológica de sua realização:

1. Documento sobre Levantamento Exploratório - Reconhecimento de Solos, realizado pela Divisão de Pesquisa Pedológica do Departamento Nacional de Pesquisa Agropecuária (DNPEA), do Ministério da Agricultura (MA), e pela Divisão de Agrologia do Departamento de Recursos Naturais (DRN), da Superintendência de Desenvolvimento do Nordeste (SUDENE), convênio MA/DNPEA-SUDENE/DRN, apresentando um mapa exploratório-reconhecimento de solos na escala de 1:500.000 (1971) [105];
2. Documento sobre a Aptidão Agrícola dos Solos do Estado, baseado na interpretação do levantamento exploratório - reconhecimento de solos anterior, apresentando um mapa de aptidão agrícola dos solos do Rio Grande do Norte para o sistema de manejo pouco desenvolvido e outro para o sistema de manejo desenvolvido, na escala de 1:500.000 (1973) [140];
3. Documento sobre Aptidão Agrícola das Terras do Rio Grande do Norte, realizado pela Secretaria Nacional de Planejamento Agrícola (SUPLAN-MA), baseado também no levantamento exploratório - reconhecimento de solos já citado, com um mapa exploratório-reconhecimento de solos na escala de 1:500.000 (1978) [106];
4. Mapa Exploratório dos Solos realizado pela Divisão de Pedologia do projeto RADAMBRASIL, Ministério das Minas e Energia, na escala de 1:250.000 (Jaguaribe/Natal, Folhas SB.24/25) (1981) [136];

5. Zoneamento Edafoclimático realizado pela Fundação Estadual de Planejamento Agrícola (CEPA/RN), apresentando mapas de aptidão climática, edáfica e edafoclimática, para diversas culturas na escala de 1:1.000.000 (1981) [47].

A partir da análise das informações encontradas acima, optou-se por utilizar o documento descrito no item 2, principalmente, por apresentar mapas em uma escala de informação mais precisa. O mapa utilizado neste trabalho foi o Mapa de Aptidão Agrícola dos Solos do Rio Grande do Norte - Sistema de Manejo Pouco Desenvolvido, na escala de 1:500.000, por ser menos restritivo que o mapa de aptidão para o sistema de manejo desenvolvido, constante no mesmo documento.

O mapa, no formato digital, foi obtido no portal *European Digital Archive on Soils Maps of the World - EuDASM* (http://eusoils.jrc.it/esdb_archive/eudasm/EUDASM.htm. Acesso em: 11/05/2008), mantido pelo ISRIC (*Institute of Environment and Sustainability - World Soil Information*). Este banco de dados contém arquivos digitais de mapas de solos e de aptidão agrícola de solos de todo o mundo.

O mapa de Aptidão Agrícola de Solos foi digitalizado no SPRING, utilizando-se o módulo de edição vetorial para a digitalização dos polígonos correspondentes as classes de aptidão e as convenções cartográficas constantes no mapa.

O SPRING é um Sistema de Informações Geográficas (SIG), de 2ª geração, contendo funções que integram técnicas de Sensoriamento Remoto e de Geoprocessamento, funções de análise espacial, modelagem numérica de terreno, composição de mapas temáticos, cadastrais e de redes, entre outras funcionalidades [40,66]. Como resultado obteve-se um mapa ou Plano de Informação (PI) da aptidão agrícola de Solos do Rio Grande do Norte (Figura 4.1). A descrição detalhada das referidas classes de aptidão, utilizadas no mapa, encontra-se no Quadro 4.1.

Classe	Descrição
I 2	Boa para culturas de ciclo curto. Regular para culturas de ciclo longo.
II 2	Regular para culturas de ciclo curto. Regular para culturas de ciclo longo.
II 3	Regular para culturas de ciclo curto. Restrita para culturas de ciclo longo.
III 3	Restrita para culturas de ciclo curto. Restrita para culturas de ciclo longo.
III 4	Restrita para culturas de ciclo curto. Inapta para culturas de ciclo longo.
IV 4	Inapta para culturas de ciclo curto. Inapta para culturas de ciclo longo.
A	Área apta para cultura do algodoeiro arbóreo.
C	Área apta para cultura do cajueiro.
S	Área apta para cultura do sisal ou agave.

Quadro 4.1 Descrição das classes de aptidão agrícola utilizadas no Mapa da Figura 4.1.

Concluída esta etapa, foi utilizada a ferramenta de recorte de PIs do SPRING, para obter o PI de aptidão agrícola para cada município do Estado, utilizando como máscara a Malha Digital de Municípios do Rio Grande do Norte, obtido no portal do IBGE, também na escala de 1:500.000 (ftp://geofp.ibge.gov.br/mapas/malhas_digitaes/municipio_2005/E500. Acesso em: 12/05/2008). Em seguida foi realizado, também no SPRING, o cálculo das áreas das classes de aptidão agrícola para cada município. A classificação se deu em função da classe com maior área no município e a partir da tabela de convenção apresentada no Quadro 4.2. Adotou-se o mesmo esquema de pontuação utilizado para o IMERG.

IMERG	ISFS		
Classes de Aptidão	Classe de Aptidão (Mapa)	Descrição da Aptidão	Pontuação
Classe 1	I 2	Boa	0.01
Classe 2	II 2, II 3	Regular	15.0
Classe 3	III 3, III 4	Restrita	35.0
Classe 4	IV 4 (II 2, II 3)	Inapta	75.0
Classe 5	IV 4 (III 3, III 4)	Inapta	90.0
Classe 6	IV 4	Inapta	100.0

Quadro 4.2 Convenções utilizadas na construção do sub-indicador “Natureza Física - PV11”.

Concluída a classificação de todos os 167 municípios do RN, de acordo com a sua aptidão agrícola, conforme descrito acima, optou-se por considerar também na construção deste sub-indicador a variável “*uso da terra*”. Portanto, municípios com aptidão agrícola desfavorável e alto índice de utilização de suas terras para agricultura e pecuária são mais susceptíveis à seca do que aqueles que não tem uma agricultura e pecuária fortes, já que essas duas atividades, nessa ordem, são muito susceptíveis ao fenômeno climático da seca.

Não foi encontrado mapa de uso da terra que abrangesse todo o estado do RN. Existem trabalhos de uso e ocupação do solo para alguns municípios do RN [45,61,68,130,179]. Portanto, optou-se por considerar as informações sobre as áreas ocupadas com lavouras permanente, temporária e em descanso e as áreas ocupadas com pastagens naturais e plantadas, contidas nos Censos Agropecuários do IBGE, como um indicativo do uso da terra para a agricultura e pecuária desses municípios.

Considerou-se para a construção deste sub-indicador um fator multiplicativo, para cada município, dado pela equação:

$$PV11_i = \frac{PAA_i \times (ALP_i + ALT_i + ALTD_i + APN_i + APP_i)}{1000} \quad (4.1)$$

onde,

PAA_i = Peso(pontuação) da Aptidão agrícola do solo do município i (Quadro 4.2).

ALP_i = Área de Lavoura Permanente do município i (ha).

ALT_i = Área de Lavoura Temporária do município i (ha).

$ALTD_i$ = Área de Lavoura Temporária em Descanso do município i (ha).

APN_i = Área de Pastagem Natural do município i (ha).

APP_i = Área de Pastagem Plantada do município i (ha).

com, $i = 1, \dots, 167$.

Na Figura 4.2 é apresentada a distribuição para este sub-indicador usando como referência as informações do Censo Agropecuário do IBGE - 2006 e na Figura 4.3 a escala de pontuação utilizada. Optou-se como método de normalização uma escala de categorias dadas pelas Equações 4.2 à 4.5 (segmentos de retas A, B, C e D). Para qualquer valor do sub-indicador acima de 4000 é atribuída a pontuação 100.

Dada a natureza e atualização das informações contidas na variável “uso da terra”, a periodicidade de atualização deste sub-indicador na operacionalização do ISFS, será de 10 anos, com respeito as informações provenientes do Censo Agropecuário, e, no caso das informações de aptidão agrícola, quando for possível um refinamento do mapa de aptidão agrícola de solos para todo o RN.

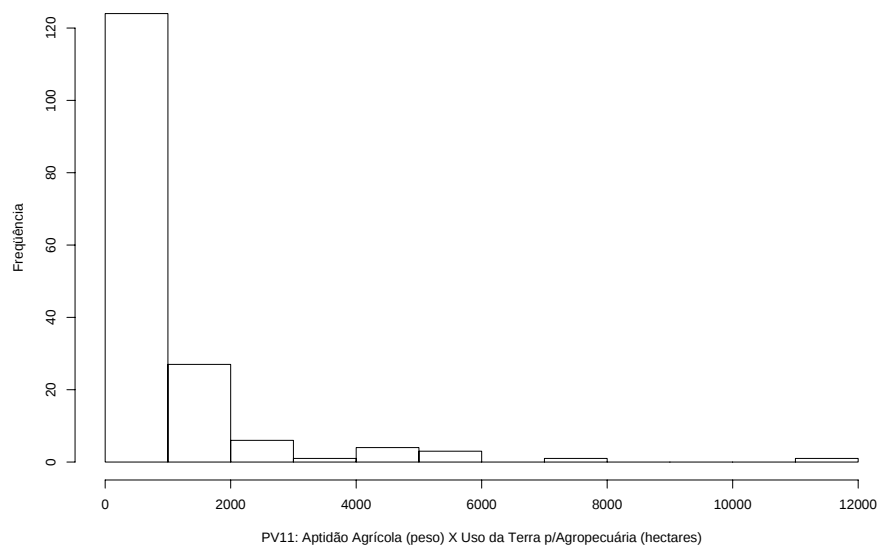


Figura 4.2 Distribuição do sub-indicador Natureza Física - PV11

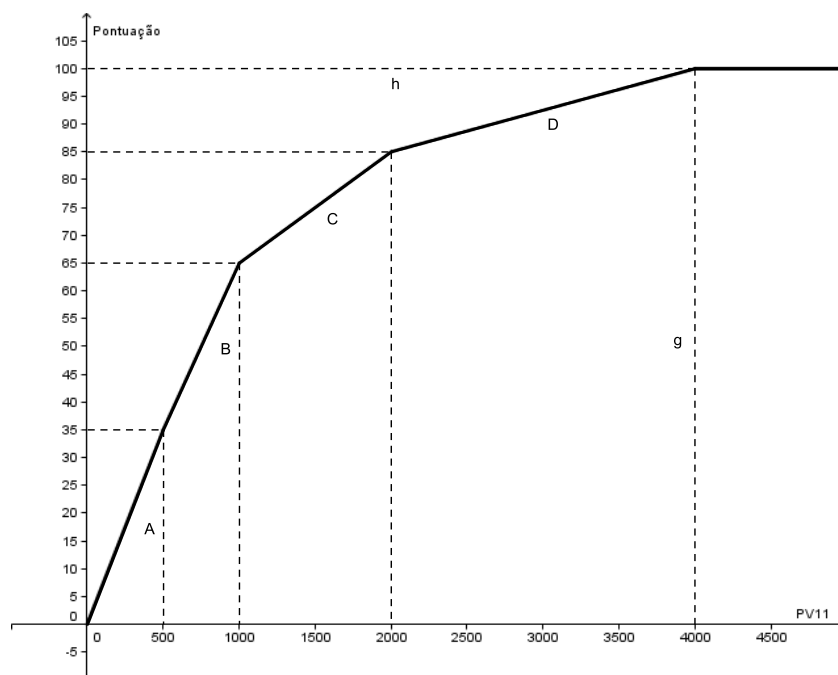


Figura 4.3 Gráfico da escala de pontuação utilizada para o sub-indicador “Natureza Física - PV11”

$$A : \quad y = 0,07x \quad (4.2)$$

$$B : \quad y = 0,06x + 5 \quad (4.3)$$

$$C : \quad y = 0,02x + 45 \quad (4.4)$$

$$D : \quad y = 0,0075x + 70 \quad (4.5)$$

Na Tabela 4.1, são apresentados os resultados da classificação da aptidão agrícola de solos para os 167 municípios do RN e na Tabela 4.2 a classificação do sub-indicador, em ordem decedente.

Tabela 4.1 Resultado da classificação da aptidão agrícola de solos do RN.

Município	Class.	Município	Class.	Município	Class.	Município	Class.
Acari	5	Frutuoso Gomes	3	Nisia_Floresta	5	São José do Seridó	6
Assu	6	Galinhos	5	Nova_Cruz	3	São Miguel	2
Afonso Bezerra	3	Goianinha	5	O. D'Água dos Borges	6	Sao M. do Gostoso	3
Água Nova	5	Gov.Dix.Sept.Ros.	6	Ouro Branco	6	Sao Paulo Potengi	3
Alexandria	5	Grossos	3	Paraná	6	Sao_Pedro	3
Almino Afonso	3	Guamaré	5	Paraú	6	Sao_Rafael	6
Alto do Rodrigues	3	Ielmo Marinho	4	Parazinho	3	Sao Tome	6
Angicos	6	Ipanguassu	2	Parelhas	5	Sao Vicente	6
Antonio Martins	3	Ipueira	6	Parnamirim	5	Senador Eloi de Souza	3
Apodi	6	Itajá	6	Passa e Fica	3	Senador G. Avelino	5
Areia Branca	3	Itaú	6	Passagem	3	Serra Caiada	3
Arez	5	Jacana	2	Patu	3	Serra de Sao Bento	5
Baia Formosa	5	Jandaira	3	Pau dos Ferros	5	Serra do Mel	3
Barauna	3	Janduis	3	Pedra Grande	2	Serra Negra do Norte	6
Barcelona	3	Japi	3	Pedra Preta	3	Serrinha	3
Bento Fernandes	2	Jardim_de_Angicos	2	Pedro Avelino	3	Serrinha dos Pintos	6
Boa_Saude	3	Jardim_de_Piranhas	6	Pedro Velho	3	Severiano Melo	5
Bodo	6	Jardim_do_Serido	5	Pendencias	3	Sítio Novo	3
Bom Jesus	3	Joao_Camara	3	Pilões	5	Tabuleiro Grande	5
Brejinho	3	Joao_Dias	6	Poco Branco	2	Taipu	2
Caiçara do Norte	5	Jose_da_Penha	5	Portalegre	6	Tangara	3
Caicara do R. Ventos	3	Jucurutu	6	Porto_do_Mangue	3	Tenente Ananias	5
Caico	6	Jundia	3	Pureza	3	Tenente L. Cruz	6
Campo Grande	5	Lagoa_D'Anta	3	Rafael Fernandes	5	Tibau	3
Campo_Redondo	3	Lagoa_de_Pedra	3	Rafael Godeiro	3	Tibau do Sul	5
Canguaretama	5	Lagoa_de_Velhos	3	Riacho da Cruz	5	Timbauba dos Batista	6
Caraubas	6	Lagoa_Nova	3	Riacho de Santana	5	Touros	3
Carnauba dos Dantas	6	Lagoa_Salgada	3	Riachuelo	3	Triunfo Potiguar	6
Carnaubais	3	Lajes	6	Rio do Fogo	6	Umarizal	3
Ceara_Mirim	3	Lajes_Pintadas	6	Rodolfo Fernandes	5	Upanema	3
Cerro_Cora	6	Lucrecia	3	Ruy_Barbosa	3	Varzea	3
Coronel Ezequiel	3	Luis_Gomes	2	Santa_Cruz	3	Venha Ver	2
Coronel J. Pessoa	2	Macaiba	3	Santa Maria	3	Vera Cruz	3
Cruzeta	6	Macau	5	Santana do Matos	6	Viçosa	6
Currais_Novos	6	Major_Sale	6	Santana do Seridó	5	Vila Flor	5
Doutor_Severiano	2	Marcelino_Viera	5	Santo Antônio	3		
Encanto	3	Martins	6	São Bento do Norte	5		
Equador	6	Maxaranguape	6	São Bento do Trairi	3		
Espírito Santo	3	Messias_Targin	3	São Fernando	6		
Extremoz	5	Montanhas	5	Sao F. do Oeste	3		
Felipe Guerra	6	Monte_Alegre	3	Sao G. do Amarante	3		
Fernando Pedroza	6	Monte das Gam.	5	Sao_Joao_do_Sabugi	6		
Floriania	6	Mossoro	6	Sao_Jose_de_Mipibu	3		
Francisco Dantas	6	Natal	5	São J. do Campestre	2		

Tabela 4.2 Resultado da classificação do sub-indicador “Natureza Física - PV11”

Município	PV11	Município	PV11	Município	PV11	Município	PV11
1.Rio do Fogo	11138,4	45.Rodolfo Fernandes	977,9	89.Bodó	435,2	133.Vera Cruz	213,2
2.Mossoro	7546,2	46.João Câmara	956,5	90.Riacho de Santana	433,6	134.Ruy Barbosa	211,8
3.Apodi	5713,1	47.Itaú	948,1	91.Boa Saúde	427,1	135.Brejinho	199,2
4.Currais Novos	5568,2	48.Baraúna	876,3	92.Japi	426,8	136.Tenente L. Cruz	198,4
5.Caico	5069,0	49.Pau dos Ferros	864,2	93.Campo Redondo	421,6	137.Parnamirim	196,8
6.Campo Grande	4825,2	50.Santana do Seridó	843,5	94.Paraná	416,3	138.Jardim de Angico	196,5
7.Caraubas	4777,6	51.Sao Vicente	779,4	95.Felipe Guerra	413,7	139.Água Nova	196,3
8.Santana do Matos	4181,2	52.Sao Paulo Potengi	764,2	96.Barcelona	409,3	140.Lagoa Salgada	195,8
9.Angicos	4088,2	53.Macafra	711,2	97.Lagoa Nova	408,4	141.Ipanguassu	191,4
10.Lajes	3926,0	54.Jardim de Piranhas	686,6	98.Bom Jesus	399,9	142.Maxaranguape	190,1
11.Sao Tome	2585,2	55.Santa Maria	686,3	99.Pedro Velho	388,6	143.Porto do Mangue	187,4
12.Acari	2572,5	56.São José do Seridó	682,3	100.Sítio Novo	386,7	144.Poço Branco	180,8
13.Serra N. do Norte	2397,9	57.Ecuador	680,4	101.Serra Caiada	385,9	145.Espírito Santo	165,6
14.Florânia	2332,0	58.Patu	664,6	102.Antonio Martins	385,4	146.Sao F. do Oeste	165,0
15.São Fernando	2108,1	59.Upanema	654,0	103.Rafael Fernandes	384,0	147.Encanto	164,3
16.São J. do Sabugi	2001,5	60.Touros	647,1	104.Pureza	377,0	148.Lagoa D'anta	157,4
17.Assu	1923,7	61.Santo Antônio	644,1	105.Senador E. de Souza	365,8	149.Grossos	153,9
18.Jucurutu	1881,7	62.Ouro Branco	643,2	106.Portalegre	364,5	150.São Miguel	136,8
19.Alexandria	1861,6	63.Pilões	642,7	107.Triunfo Potiguar	351,9	151.A. do Rodrigues	131,7
20.Fernando Pedroza	1852,1	64.Guamaré	634,4	108.Major Sales	350,0	152.Rafael Godeiro	130,8
21.São B. do Norte	1832,9	65.Monte Alegre	633,6	109.Caicara do R. Ventos	340,3	153.Luis Gomes	125,3
22.Paraú	1815,2	66.José da Penha	621,6	110.Lagoa de Pedras	312,7	154.Passagem	116,5
23.Goiânia	1656,4	67.O. D'Água dos Borges	618,7	111.M. das Gameleiras	305,6	155.Passa e Fica	92,1
24.Cerro Corá	1623,9	68.Serra de Sao Bento	601,6	112.São M.I do Gostoso	298,1	156.Pendências	84,8
25.Ielmo Marinho	1616,3	69.Serrinhas dos Pintos	580,4	113.Coronel Ezequiel	296,6	157.Pedra Grande	82,2
26.Marcelino Vieira	1603,4	70.Tangará	578,8	114.Taipu	294,9	158.Frutoso Gomes	72,6
27.Arez	1598,3	71.Riacho da Cruz	568,5	115.São J. do Campestre	290,9	159.Sen. G. Avelino	62,8
28.Gov.Dix.Sept.Ros.	1583,0	72.Riachuelo	562,6	116.Serrinha	289,3	160.Cel. João Pessoa	58,8
29.Martins	1492,2	73.Lajes Pintadas	550,0	117.Itajá	284,9	161.Jundiá	57,5
30.São Rafael	1482,8	74.Ipueira	540,3	118.Extremoz	277,3	162.Caicara do Norte	57,3
31.Severiano Melo	1444,0	75.Tibau do Sul	520,7	119.Bento Fernandes	268,8	163.Doutor Severiano	56,2
32.Santa Cruz	1365,3	76.Afonso Bezerra	516,4	120.Messias Targino	254,4	164.Jacaná	49,7
33.Tenente Ananias	1292,6	77.São José do Mipibu	516,0	121.Várzea	253,5	165.Lucrecia	35,6
34.Pedro Avelino	1288,5	78.Tabuleiro Grande	506,7	122.Parazinho	251,6	166.Natal	18,6
35.Francisco Dantas	1232,4	79.Carnaubais	504,2	123.Montanhas	251,0	167.Venha Ver	15,0
36.Baía Formosa	1207,9	80.Nova Cruz	499,6	124.Areia Branca	250,1		
37.Jardim do Seridó	1151,1	81.Galinhas	475,3	125.São Bento do Trairi	249,5		
38.Canguaretama	1105,6	82.Timbaúba dos Batistas	473,2	126.Pedra Preta	246,5		
39.Parelhas	1103,9	83.Nísia Floresta	471,3	127.Almino Afonso	246,2		
40.Cruzeta	1093,5	84.Jandaíra	458,9	128.Viçosa	237,7		
41.Serra do Mel	1059,8	85.Lagoa de Velhos	447,4	129.São G. do Amarante	237,6		
42.C. dos Dantas	1043,9	86.Umarizal	444,0	130.Vila Flor	236,4		
43.Macau	1031,6	87.São Pedro	439,1	131.João Dias	234,0		
44.Ceará Mirim	987,5	88.Janduí	435,7	132.Tibau	230,7		

4.2.2 Sub-indicador: Natureza Climatológica - PV12

Na construção do IMERG este sub-indicador foi representado pelo volume de precipitação histórica de cada município, obtido da classificação da precipitação através da técnica dos quantis [185]. O Rio Grande do Norte não dispunha de estudos dessa natureza, com abrangência em todo o Estado e que possibilitasse o fornecimento dessas normais climatológicas. Para atender a essa necessidade foi realizado um estudo visando a caracterização pluviométrica das 19 microrregiões do Estado através da técnica dos quantis.

A primeira etapa do trabalho constituiu no levantamento, coleta e organização dos dados mensais de chuva dos postos pluviométricos da Superintendência de Desenvolvimento do Nordeste – SUDENE, do Departamento Nacional de Obras contra a Seca – DNOCS, e do Departamento Nacional de Águas e Energia Elétrica – DNAEE. Estes dados foram obtidos, sem nenhum ônus, no Departamento de Meteorologia da Empresa de Pesquisa Agropecuária do Rio Grande do Norte - EMPARN, que foi a instituição que assumiu a rede de postos pluviométricos da SUDENE a partir de 1992, após a criação dos Centros Estaduais de Meteorologia pelo Programa de Monitoramento de Tempo e Clima para os Estados do Nordeste do Brasil [22]. Estes dados foram consistidos e avaliados tomando-se como referência as séries de dados pluviométricos mensais até o ano de 1990, obtidas em documento impresso da SUDENE [168], e foram complementados com os dados do Banco de Dados Pluviométrico da EMPARN.

Foram utilizados também dados das estações meteorológicas de Mossoró, Macau, Ceará-Mirim, Florânia e Cruzeta adquiridos junto ao Instituto Nacional de Meteorologia (INMET). Estas estações apresentaram no mínimo 30 anos de dados sem falhas. Não foram consideradas as estações de Apodi, Caicó (Seridó) e de Natal por apresentarem menos de 30 anos de dados sem falhas. O Programa de Pós-graduação em Matemática Computacional da Universidade Federal de Pernambuco financiou 77,63% do valor destes dados, sendo o restante custeado pela Empresa Brasileira de Pesquisa Agropecuária - Embrapa.

Portanto, para a realização deste estudo, foram utilizados dados de 93 postos pluviométricos da SUDENE/DNOCS/DNAEE/EMPARN, e de 5 estações meteorológicas do INMET, distribuídos em 85 municípios nas 19 microrregiões do Estado (Figura 4.4).

A Organização Mundial de Meteorologia recomenda, para se trabalhar com normais climatológicas, a utilização de séries de dados com 30 anos consecutivos, porém, dada a grande quantidade de falhas nas séries de dados dos postos pluviométricos do Rio Grande do Norte, optou-se por estender o período trabalhado, garantindo com isso o mínimo de 30 anos de dados sem falhas em cada série. As séries consideradas no estudo cobrem o período de 1963 a 2006

perfazendo 44 anos, com o mínimo de 30 anos e o máximo de 43 anos de informações sem ausência de dados. Porém, para satisfazer o critério de 30 anos pré-estabelecido, foi necessário estender o período de 1954 a 2006 em 7 postos pluviométricos.

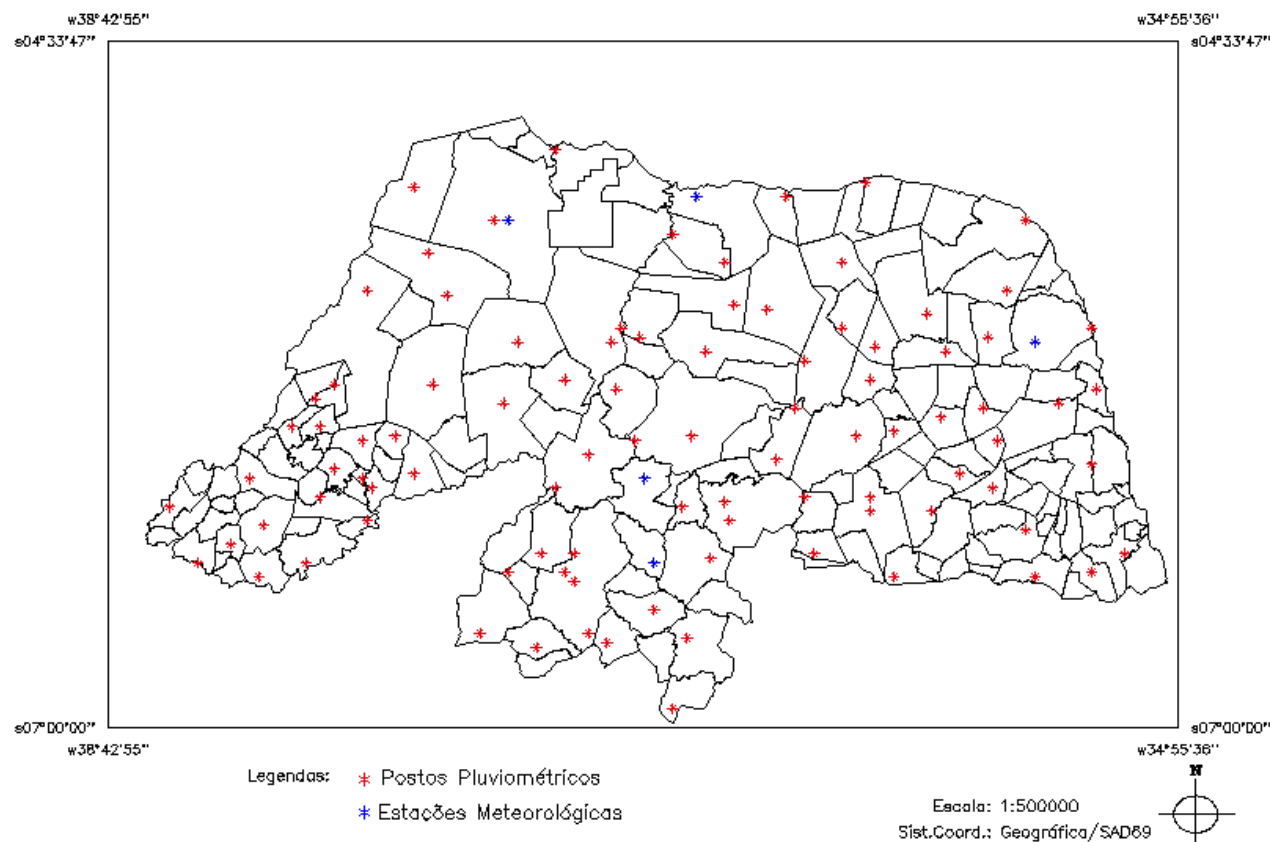


Figura 4.4 Localização dos Postos Pluviométricos e das Estações Meteorológicas do Rio Grande do Norte.

Na Figura 4.5 é apresentada a quantidade de postos utilizados por ano no cálculo dos quantis. A partir da análise deste gráfico ficou evidente uma queda brusca ocorrida na quantidade de informações disponíveis (dados sem falhas) no período compreendido entre 1990 e 1992, período de transição que correspondeu a transferência da rede de postos pluviométricos da SUDENE para a EMPARN. Pode-se verificar também, após essa transferência, uma maior variabilidade na quantidade de postos sem falhas de ano para ano, o que evidencia uma queda na qualidade do monitoramento com respeito a ocorrência de falhas no monitoramento destes postos oriundos da SUDENE, apesar da EMPARN ter aumentado nos últimos anos a capilaridade da rede pluviométrica do Estado [22]. Este fato demonstra a necessidade de uma maior aplicação de recursos financeiros por parte da EMPARN e do Governo do Estado no programa de monitoramento e na manutenção da rede pluviométrica do Rio Grande do Norte.

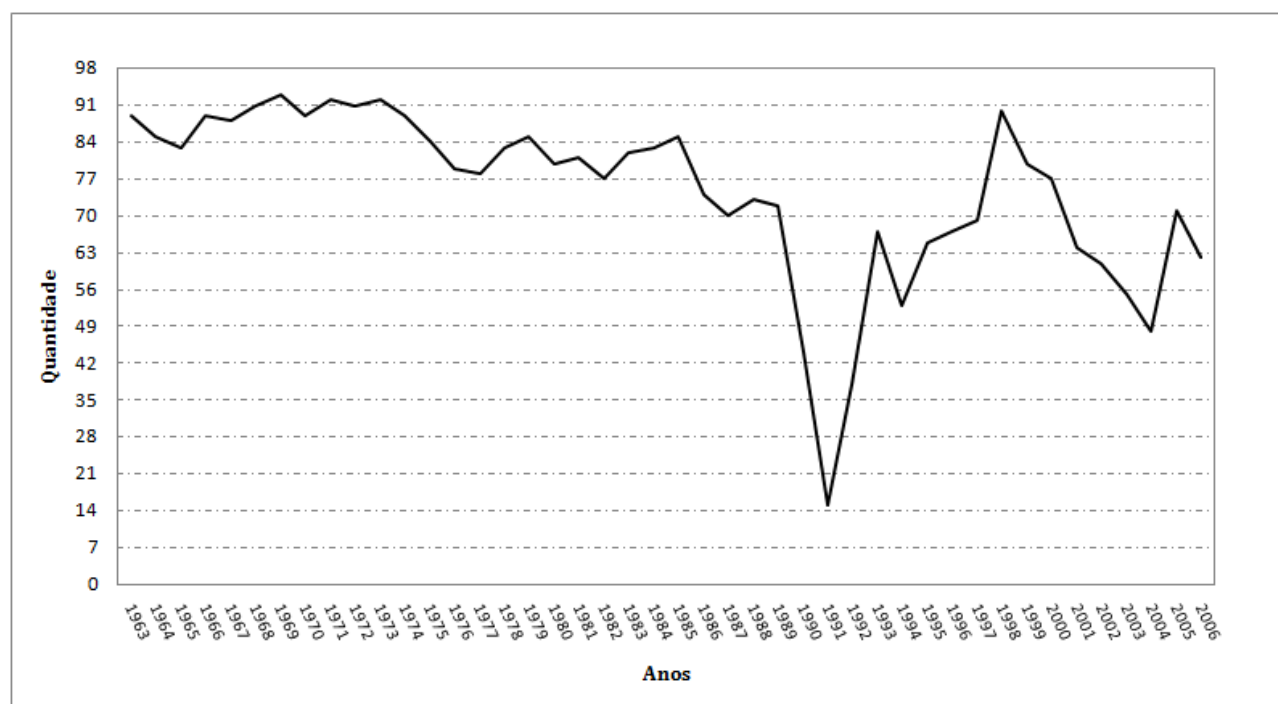


Figura 4.5 Quantidade de Postos Pluviométricos e Estações Meteorológicas utilizados no cálculo dos quantis.

As normais climáticas das séries foram obtidas através da técnica dos quantis. Trabalhos com essa abordagem foram primeiramente desenvolvidos por PINKAYAN [132], e esta técnica tem sido utilizada por diversos autores em trabalhos no Brasil e mais especificamente no Nordeste (ver Capítulo 2, item 2.5.3). Portanto, o objetivo principal deste estudo foi o de caracterizar períodos secos e chuvosos no Rio Grande do Norte, através da técnica dos quantis, a fim de possibilitar a sua implementação no ISFS. Utilizaram-se os parâmetros propostos por XAVIER & XAVIER [185], excetuando-se as reavaliações dos anos muitos secos e muitos chuvosos por não serem necessárias ao estudo. Foram calculados os quantis para todo o Estado do Rio Grande do Norte e para cada uma das suas 19 microrregiões.

São várias as formas de se calcular quantis amostrais, que dependem do livro texto ou do pacote computacional que se está utilizando [86, 98]. HYNDMAN & FAN [86] apresentaram nove formas para se calcular os quantis amostrais que estão implementadas no pacote do software R. Já LANGFORD [98] avaliou 15 formas distintas, sendo algumas destas as mesmas do trabalho de HYNDMAN & FAN, e que estão implementadas nos pacotes de softwares estatísticos comerciais tais como SAS, JMP, MINITAB, Excel e TI-83 PLUS, além de sugerir um novo método que é estatisticamente forte e fácil de aplicar.

De uma forma geral, os quantis amostrais provêm um estimador não paramétrico da sua população baseados em similaridades no conjunto de observações independentes $\{X_1, \dots, X_n\}$ de uma distribuição F, sendo, portanto, definidos como:

$$Q(p) = F^{-1}(p) = \inf\{x : F(x) \geq p\}, 0 < p < 1, \quad (4.6)$$

onde, $F(x)$ é a função de distribuição de probabilidades [86].

Sendo $\{X_{(1)}, \dots, X_{(n)}\}$ a estatística de ordem de $\{X_1, \dots, X_n\}$, e sendo $\hat{Q}_i(p)$ a definição do i -ésimo quantil amostral. Pode-se, portanto, definir quantis amostrais como:

$$\hat{Q}_i(p) = (1 - \gamma)X_{(j)} + \gamma X_{(j+1)}, \quad \text{com} \quad \frac{j - m}{n} \leq p \leq \frac{j - m + 1}{n} \quad (4.7)$$

para algum $m \in \mathbb{R}$ e $0 \leq \gamma \leq 1$, γ sendo uma função de $j = \lfloor pn + m \rfloor$ e $g = pn + m - j$, com $\lfloor u \rfloor$ denotando o maior inteiro não maior que u . Por outro lado, assume-se também que $\lceil u \rceil$ denota o menor inteiro não menor que u .

HYNDMAN & FAN [86] apresentaram também um conjunto de seis propriedades que são desejáveis a um bom método de obtenção de quantis amostrais (Quadro 4.3). Neste trabalho foi utilizado o método que é baseado em um estimador não-viciado para a mediana. Este método está disponível no software R (tipo 8), e a seguir é apresentada a sua definição matemática.

Primeiramente considera-se a inversa da função de distribuição empírica obtida fazendo $m = 0$ e,

$$\gamma = \begin{cases} 1, & \text{se } g > 0, \\ 0, & \text{se } g = 0. \end{cases}$$

Isto resulta em uma função constante por partes. Assume-se também que a $Freq(X_k \leq x)$ denota o número de observações menores ou iguais a x .

Logo, obtem-se os quantis amostrais a partir de

$$Freq(X_k \leq \hat{Q}(p)) = \lfloor pn \rfloor \quad (4.8)$$

e,

$$Freq(X_k \geq \hat{Q}(1 - p)) = \lceil pn + 1 \rceil \quad (4.9)$$

Sendo L uma medida local tal como média, mediana ou moda, existem duas classes de definições de quantis que são derivadas com base na estimativa dos argumentos. A primeira abordagem escolhe $p_k = LF(X_{(k)})$ e a segunda abordagem escolhe $p_k = F(LX_{(k)})$. Se F for uma distribuição uniforme ou L for a mediana e F for estritamente monotônica, as duas abordagens são equivalentes.

Portanto, faz-se, então,

$$p_k = \frac{(k - \frac{1}{3})}{(n + \frac{1}{3})} \approx MF(X_{(k)}), \text{ c/ } M \text{ sendo a mediana.} \quad (4.10)$$

Este método apresenta as vantagens de fornecer um estimador não-viciado para a mediana de ordem $o(n^{-1/2})$ e de ser independente da distribuição, e a desvantagem de não satisfazer a 3ª propriedade das desejáveis a um bom método de obtenção de quantis, mas que é compensada por satisfazer a propriedade 4, que também garante a simetria nas caudas da distribuição [86].

Prop.	Definição	Descrição
1	$\hat{Q}_i(p)$ é contínuo	Baseada na suposição de que $\hat{Q}(p)$ é uma função contínua de p .
2	$Freq(X_k \leq \hat{Q}_i(p)) \geq pn$	Análogo amostral do resultado $F(Q(u)) \geq u$ (com igualdade quando F é contínua).
3	$Freq(X_k \leq \hat{Q}_i(p)) = Freq(X_k \geq \hat{Q}_i(1-p))$	Propriedade que garante a simetria, ou seja, que as caudas da distribuição em questão sejam tratadas igualmente.
4	Seja $\hat{Q}_i^{-1}(x)$ definido unicamente, $\Rightarrow \hat{Q}_i^{-1}(x_{(k)}) + \hat{Q}_i^{-1}(x_{(n-k+1)}) = 1$, $p/ k = 1, \dots, n$.	Ídem a propriedade anterior.
5	Seja $\hat{Q}_i^{-1}(x)$ definido unicamente, $\Rightarrow \hat{Q}_i^{-1}(x_{(1)}) > 0$ e $\hat{Q}_i^{-1}(x_{(n)}) < 1$.	Garante que, para uma função de distribuição contínua, espera-se que existam probabilidades positivas para os valores, no intervalo de dados.
6	$\hat{Q}_i^{-1}(0.5)$ é igual à mediana amostral, definida por, $\begin{cases} \frac{[X_{(l)} + X_{(l+1)}]}{2} & \text{se } n = 2l. \\ X_{(l+1)} & \text{se } n = 2l + 1. \end{cases}$	Garante que a mediana seja igual a $\hat{Q}_i^{-1}(0.5)$.

Quadro 4.3 Propriedades desejáveis para quantis amostrais. Adaptado de HYNDMAN & FAN [86].

O cálculo dos quantis foi realizado utilizando o software R, que é uma linguagem de programação e um ambiente de software livre para computação estatística e geração de gráficos (disponível em <<http://www.r-project.org>>. Acesso em: 20/05/2008). Foram calculados os quantis para todo o Rio Grande do Norte e para cada uma das suas 19 microrregiões, considerando-se a precipitação acumulada anual e de todos os quadrimestres do ano, neste caso, objetivando-se determinar historicamente os quadrimestres mais chuvosos (quadra chuvosa), para cada uma das regiões consideradas. Esta informação, por quadra chuvosa, será

útil na definição do sub-indicador “*Risco de Perda da Produção Agrícola - PV21*”, detalhado mais adiante.

As Figuras A.1 a A.19 (Apêndice A), apresentam a evolução dos quantis, por quadrimestres do ano, para cada uma das 19 microrregiões consideradas. A partir da análise dos resultados apresentados nos gráficos verificou-se que as microrregiões de Natal, Macaíba e Litoral Sul apresentam a quadra chuvosa caracterizada no período de abril a julho. As microrregiões Agreste Potiguar, Baixa Verde e Litoral Sul, no período de março a junho, e o restante do Estado, com excessão do Seridó Oriental, no período de fevereiro a maio. A quadra chuvosa para o Seridó Oriental configurou-se no período de janeiro a abril. Na tabela 4.3, abaixo, são apresentados os resultados desta classificação e na Tabela 4.4 os resultados dos quantis considerando-se a precipitação anual.

Tabela 4.3 Classificação pluviométrica para a quadra chuvosa das microrregiões do Rio Grande do Norte, a partir da técnica dos quantis.

Microrregião	Quadra	Categorias				
		Muito Seco	Seco	Normal	Chuvoso	Muito Chuvoso
Agreste Potiguar	mar-jun	$P \leq 236, 25$	$236, 25 < P \leq 327, 45$	$327, 45 < P \leq 491, 90$	$491, 90 < P \leq 631, 39$	$P > 631, 39$
Angicos	fev-mai	$P \leq 191, 35$	$191, 35 < P \leq 285, 98$	$285, 98 < P \leq 452, 96$	$452, 96 < P \leq 625, 66$	$P > 625, 66$
Baixa Verde	mar-jun	$P \leq 228, 33$	$228, 33 < P \leq 301, 23$	$301, 23 < P \leq 482, 11$	$482, 11 < P \leq 621, 67$	$P > 621, 67$
Borborema Potiguar	fev-mai	$P \leq 184, 64$	$184, 64 < P \leq 282, 07$	$282, 07 < P \leq 406, 19$	$406, 19 < P \leq 573, 86$	$P > 573, 86$
Chapada do Apodi	fev-mai	$P \leq 321, 32$	$321, 32 < P \leq 420, 73$	$420, 73 < P \leq 623, 85$	$623, 85 < P \leq 784, 51$	$P > 784, 51$
Litoral Nordeste	mar-jun	$P \leq 377, 55$	$377, 55 < P \leq 536, 56$	$536, 56 < P \leq 728, 40$	$728, 40 < P \leq 1060, 33$	$P > 1060, 33$
Litoral Sul	abr-jul	$P \leq 499, 10$	$499, 10 < P \leq 659, 72$	$659, 72 < P \leq 839, 43$	$839, 43 < P \leq 1018, 13$	$P > 1018, 13$
Macaíba	abr-jul	$P \leq 505, 73$	$505, 73 < P \leq 625, 90$	$625, 90 < P \leq 928, 60$	$928, 60 < P \leq 1151, 06$	$P > 1151, 06$
Macau	fev-mai	$P \leq 220, 80$	$220, 80 < P \leq 334, 18$	$334, 18 < P \leq 494, 40$	$494, 40 < P \leq 726, 01$	$P > 726, 01$
Médio Oeste	fev-mai	$P \leq 335, 35$	$335, 35 < P \leq 467, 50$	$467, 50 < P \leq 690, 35$	$690, 35 < P \leq 890, 00$	$P > 890, 00$
Mossoró	fev-mai	$P \leq 278, 20$	$278, 20 < P \leq 452, 50$	$452, 50 < P \leq 643, 60$	$643, 60 < P \leq 816, 43$	$P > 816, 43$
Natal	abr-jul	$P \leq 740, 16$	$740, 19 < P \leq 893, 11$	$893, 11 < P \leq 1279, 29$	$1279, 29 < P \leq 1517, 27$	$P > 1517, 27$
Pau dos Ferros	fev-mai	$P \leq 373, 73$	$373, 73 < P \leq 507, 40$	$507, 40 < P \leq 664, 28$	$664, 28 < P \leq 818, 58$	$P > 818, 58$
Seridó Ocidental	fev-mai	$P \leq 290, 77$	$290, 77 < P \leq 440, 85$	$440, 85 < P \leq 615, 41$	$615, 41 < P \leq 782, 13$	$P > 782, 13$
Seridó Oriental	jan-abr	$P \leq 230, 68$	$230, 68 < P \leq 365, 60$	$365, 60 < P \leq 487, 90$	$487, 90 < P \leq 660, 66$	$P > 660, 66$
Serra de Santana	fev-mai	$P \leq 241, 65$	$241, 65 < P \leq 361, 50$	$361, 50 < P \leq 576, 15$	$576, 15 < P \leq 784, 55$	$P > 784, 55$
Serra de São Miguel	fev-mai	$P \leq 429, 95$	$429, 95 < P \leq 564, 80$	$564, 80 < P \leq 720, 69$	$720, 69 < P \leq 936, 28$	$P > 936, 28$
Umarizal	fev-mai	$P \leq 389, 73$	$389, 73 < P \leq 560, 03$	$560, 03 < P \leq 764, 48$	$764, 48 < P \leq 942, 45$	$P > 942, 45$
Vale do Assu	fev-mai	$P \leq 288, 97$	$288, 97 < P \leq 420, 68$	$420, 68 < P \leq 615, 25$	$615, 25 < P \leq 781, 79$	$P > 781, 79$

P - Precipitação acumulada no período (mm).

Este estudo também compreendeu a avaliação pluviométrica dos anos do período considerado no estudo (1963-2006), a partir da classificação nas categorias dos quantis obtidos (Tabelas 4.3 e 4.4) para a quadra chuvosa e o acumulado anual de cada um dos municípios do Rio Grande do Norte. Essa informação possibilitou a escolha dos anos em que foi ajustado o ISFS, dando-se preferência a escolha de anos cuja quadra chuvosa representasse a condição seca, normal

Tabela 4.4 Classificação pluviométrica para as microrregiões do Rio Grande do Norte, a partir da técnica dos quantis.

Microrregião	Categorias				
	Muito Seco	Seco	Normal	Chuvoso	Muito Chuvoso
Agreste Potiguar	$P \leq 448,96$	$448,96 < P \leq 572,24$	$572,24 < P \leq 781,58$	$781,58 < P \leq 1071,51$	$P > 1071,51$
Angicos	$P \leq 287,38$	$287,38 < P \leq 397,75$	$397,75 < P \leq 611,90$	$611,90 < P \leq 829,78$	$P > 829,78$
Baixa Verde	$P \leq 389,21$	$389,21 < P \leq 528,12$	$528,12 < P \leq 740,72$	$740,72 < P \leq 1005,71$	$P > 1005,71$
Borborema Potiguar	$P \leq 316,05$	$316,05 < P \leq 435,70$	$435,70 < P \leq 628,95$	$628,95 < P \leq 849,05$	$P > 849,05$
Chapada do Apodi	$P \leq 426,12$	$426,12 < P \leq 562,55$	$562,55 < P \leq 867,77$	$867,77 < P \leq 1068,80$	$P > 1068,80$
Litoral Nordeste	$P \leq 713,47$	$713,47 < P \leq 889,72$	$889,72 < P \leq 1216,17$	$1216,17 < P \leq 1.642,12$	$P > 1.642,12$
Litoral Sul	$P \leq 932,07$	$932,07 < P \leq 1131,02$	$1131,02 < P \leq 1438,09$	$1438,09 < P \leq 1.679,21$	$P > 1.679,21$
Macaíba	$P \leq 826,66$	$826,66 < P \leq 1101,90$	$1101,90 < P \leq 1478,40$	$1478,40 < P \leq 1779,23$	$P > 1779,23$
Macau	$P \leq 311,58$	$311,58 < P \leq 444,48$	$444,48 < P \leq 681,44$	$681,44 < P \leq 898,84$	$P > 898,84$
Médio Oeste	$P \leq 442,15$	$442,15 < P \leq 602,31$	$602,31 < P \leq 885,96$	$885,96 < P \leq 1115,77$	$P > 1115,77$
Mossoró	$P \leq 365,06$	$365,06 < P \leq 551,70$	$551,70 < P \leq 885,00$	$885,00 < P \leq 1046,53$	$P > 1046,53$
Natal	$P \leq 1185,77$	$1185,77 < P \leq 1581,70$	$1581,70 < P \leq 2015,20$	$2015,20 < P \leq 2263,51$	$P > 2263,51$
Pau dos Ferros	$P \leq 490,34$	$490,34 < P \leq 654,74$	$654,74 < P \leq 891,12$	$891,12 < P \leq 1096,79$	$P > 1096,79$
Seridó Ocidental	$P \leq 379,25$	$379,25 < P \leq 574,40$	$574,40 < P \leq 801,75$	$801,75 < P \leq 1020,20$	$P > 1020,20$
Seridó Oriental	$P \leq 323,95$	$323,95 < P \leq 463,45$	$463,45 < P \leq 678,50$	$678,50 < P \leq 865,40$	$P > 865,40$
Serra de Santana	$P \leq 317,81$	$317,81 < P \leq 509,20$	$509,20 < P \leq 784,50$	$784,50 < P \leq 1011,13$	$P > 1011,13$
Serra de São Miguel	$P \leq 595,26$	$595,26 < P \leq 759,31$	$759,31 < P \leq 1061,37$	$1061,37 < P \leq 1271,48$	$P > 1271,48$
Umarizal	$P \leq 492,41$	$492,41 < P \leq 760,72$	$760,72 < P \leq 1014,21$	$1014,21 < P \leq 1289,03$	$P > 1289,03$
Vale do Assu	$P \leq 375,67$	$375,67 < P \leq 527,32$	$527,32 < P \leq 802,20$	$802,20 < P \leq 1003,92$	$P > 1003,92$

P - Precipitação acumulada no ano (mm).

e chuvosa na maioria do Estado. Essa avaliação foi realizada a partir do desenvolvimento de um programa em C, e o resultado foi trabalhado no SPRING/SCARTA para a geração dos mapas de classificação. Nas Figuras B.1 a B.88 (*Apêndice B*) são apresentados os mapas com os resultados dessa classificação e na Tabela C.1 (*Apêndice C*) a avaliação pluviométrica para cada ano considerado no estudo.

Por fim, para a caracterização deste parâmetro foram utilizados os quantis $\hat{Q}(0.50)$ correspondentes à média histórica dos municípios onde existiam postos pluviométricos ou estações meteorológicas. Para os demais municípios, as médias históricas foram obtidas por interpolação, a partir de um Modelo Numérico de Terreno (MNT), ferramenta disponível nos SIGs, em particular no SPRING, para a representação matemática da distribuição espacial de um fenômeno vinculado a uma superfície qualquer [167]. Foi gerada uma grade retangular a partir destes pontos cotados (médias históricas), usando-se o método de interpolação média ponderada por quadrante (Figura 4.6). A partir daí, os pontos correspondentes às médias pluviométricas históricas foram obtidos das coordenadas da sede de cada município (Tabela 4.5).

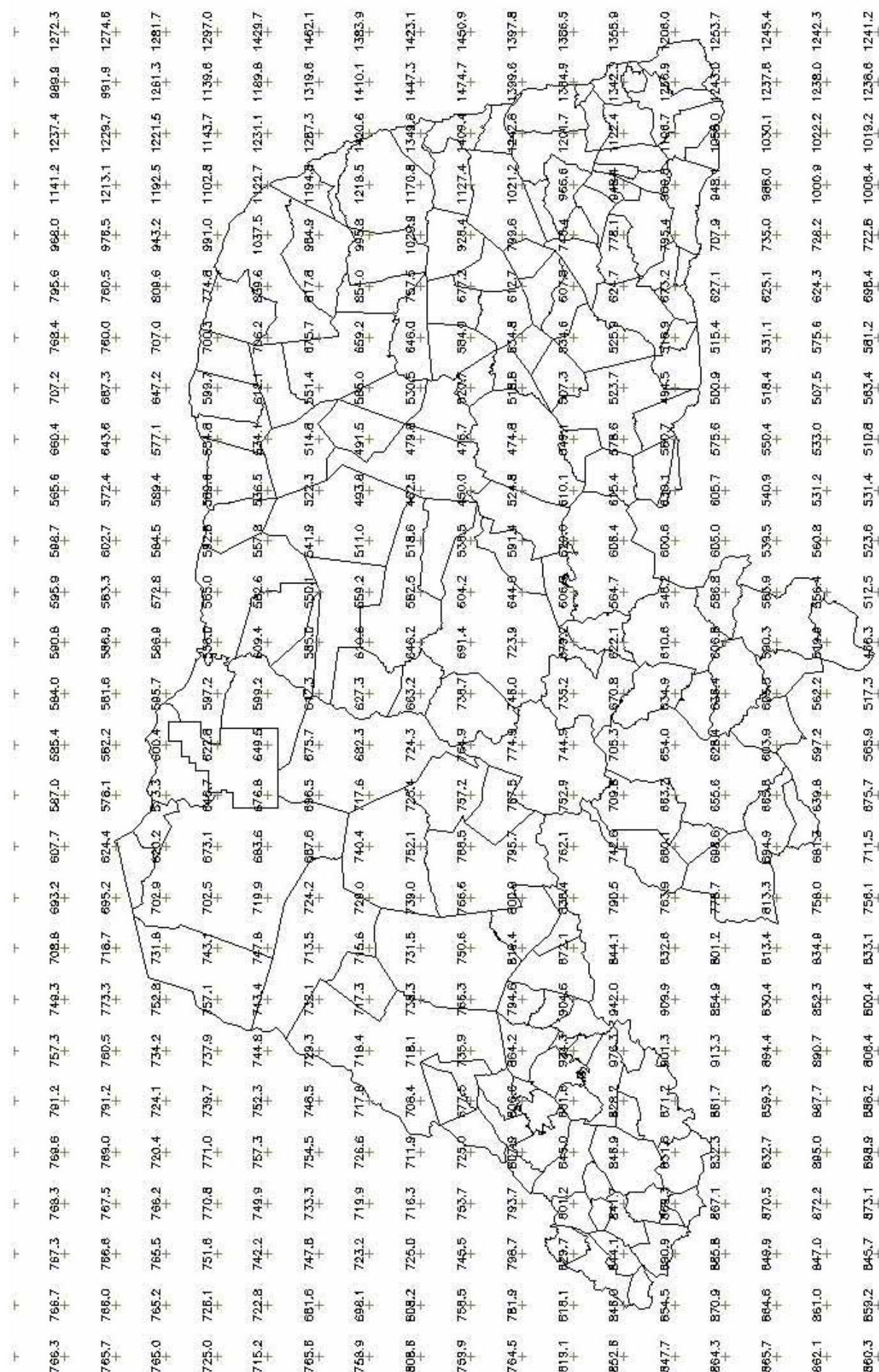


Figura 4.6 Grade retangular gerada a partir das médias pluviométricas ($\hat{Q}_{0.5}$).

Tabela 4.5 Médias pluviométricas dos municípios do Rio Grande do Norte, obtidas pelo cálculo das medianas ($\hat{Q}_{0.5}$).

Município	Prec.	Município	Prec.	Município	Prec.	Município	Prec.
Acari	492,3	Frutuoso Gomes	935,2*	Nisia_Floresta	1288,6*	São José do Seridó	637,3*
Assu	699,8	Galinhos	622,5*	Nova_Cruz	742,9	São Miguel	814,7
Afonso Bezerra	594,6	Goianinha	1151,4*	O. D'Agua dos Borges	744,2	Sao M. do Gostoso	794,2*
Agua Nova	859,2*	Gov.Dix.Sept.Ros.	629,3	Ouro Branco	581,8	Sao Paulo Potengi	547,0
Alexandria	886,6	Grossos	576,2	Paraná	866,7*	Sao_Pedro	585,5
Almino Afonso	966,1	Guamaré	701,2	Paraú	684,5	Sao_Rafael	814,0
Alto do Rodrigues	603,3*	Ielmo Marinho	879,2*	Parazinho	670,9*	Sao Tome	382,9
Angicos	601,5	Ipanguassu	575,3	Parelhas	606,7	Sao Vicente	631,0
Antonio Martins	696,0	Ipueira	655,9*	Parnamirim	1268,4*	Senador Eloi de Souza	607,8*
Apodi	736,9	Itajá	651,6*	Passa e Fica	687,6*	Senador G. Avelino	1302,6*
Areia Branca	546,5	Itaú	646,8	Passagem	957,5*	Serra Caiada	580,2
Arez	1216,2*	Jacana	632,6*	Patu	912,7	Serra de Sao Bento	658,7*
Baia Formosa	1238,8*	Jandaira	473,5	Pau dos Ferros	801,4	Serra do Mel	630,3*
Barauna	794,7	Janduis	802,0*	Pedra Grande	479,9	Serra Negra do Norte	847,0
Barcelona	569,9	Japi	447,3	Pedra Preta	488,2*	Serrinha	826,9*
Bento Fernandes	650,8*	Jardim_de_Angicos	509,5	Pedro Avelino	483,7	Serrinha dos Pintos	948,0*
Boa_Saude	665,8	Jardim_de_Piranhas	762,0*	Pedro Velho	1142,9	Severiano Melo	723,7
Bodo	589,8*	Jardim_do_Serido	679,1	Pendencias	652,8	Sitio Novo	528,3*
Bom Jesus	692,2	Joao_Camara	641,7	Piloes	859,3*	Tabuleiro Grande	716,5
Brejinho	1036,8*	Joao_Dias	1043,9	Poco Branco	714,4	Taipu	995,9
Caiçara do Norte	551,8*	Jose_da_Penha	901,4	Portalegre	848,5*	Tangara	475,9
Caicara do R. Ventos	487,6	Jucurutu	831,8	Porto_do_Mangue	589,8*	Tenente Ananias	807,8
Caico	663,9	Jundia	1017,6*	Pureza	870,0	Tenente L. Cruz	685,4*
Campo Grande	806,4	Lagoa_D'Anta	707,8*	Rafael Fernandes	811,8*	Tibau	622,1*
Campo_Redondo	692,0	Lagoa_de_Pedra	916,6*	Rafael Godeiro	856,4*	Tibau do Sul	1370,2*
Canguaretama	1342,5	Lagoa_de_Velhos	531,3*	Riacho da Cruz	746,2	Timbauba dos Batista	679,7
Caraubas	717,0	Lagoa_Nova	624,8*	Riacho de Santana	843,3*	Touros	1076,3
Carnauba dos Dantas	591,3*	Lagoa_Salgada	776,0*	Riachuelo	563,37*	Triunfo Potiguar	772,7*
Carnaubais	600,4*	Lajes	462,1	Rio do Fogo	1122,6*	Umarizal	852,2
Ceara_Mirim	1176,6	Lajes_Pintadas	553,3*	Rodolfo Fernandes	702,2*	Upanema	803,0
Cerro_Cora	648,7	Lucrecia	884,4	Ruy_Barbosa	537,8*	Varzea	970,6*
Coronel Ezequiel	681,6	Luis_Gomes	906,5	Santa_Cruz	510,7	Venha Ver	854,6*
Coronel J. Pessoa	838,7*	Macaiba	1086,1	Santa Maria	680,0*	Vera Cruz	962,8*
Cruzeta	647,1	Macau	495,9	Santana do Matos	736,5	Viçosa	852,1*
Currais_Novos	521,5	Major_Sales	872,7*	Santana do Seridó	562,7*	Vila Flor	1247,8*
Doutor_Severiano	817,3*	Marcelino_Viera	821,0	Santo Antônio	831,1		
Encanto	802,6*	Martins	1245,7	São Bento do Norte	541,2		
Equador	412,10	Maxaranguape	1330,1*	São Bento do Trairi	575,9*		
Espirito Santo	1068,4*	Messias Targino	820,2*	São Fernando	671,8		
Extremoz	1294,6*	Montanhas	1061,4*	Sao F. do Oeste	806,5*		
Felipe Guerra	725,4*	Monte_Alegre	1141,2*	Sao G. do Amarante	1255,83*		
Fernando Pedroza	579,9*	Monte das Gam.	563,8*	Sao_Joao_do_Sabugi	659,6		
Florania	744,2	Mossoro	695,7	Sao_Jose_de_Mipibu	1380,2		
Francisco Dantas	827,4*	Natal	1699,5	São J. do Campestre	557,1*		

* Valor Interpolado

A Figura 4.7 apresenta o gráfico da escala de pontuação adaptada às condições do Rio

Grande do Norte, para ser utilizada na aplicação deste sub-indicador. Também foi definido como método de normalização uma escala de categorias dada pelas equações 4.11 a 4.13. Para qualquer município cuja média histórica da precipitação fosse igual ou inferior a 500mm atribuíram-se 100 pontos, e para qualquer valor acima de 1500mm 0 pontos. Nos pontos intermediários correspondentes aos seguimentos de retas B , C e D , a pontuação foi obtida pelas referidas equações.

$$B : y = 100 - 0,125 \times (x - 500) \quad (4.11)$$

$$C : y = 37,5 - 0,1 \times (x - 1000) \quad (4.12)$$

$$D : y = 12,5 - 0,05 \times (x - 1250) \quad (4.13)$$

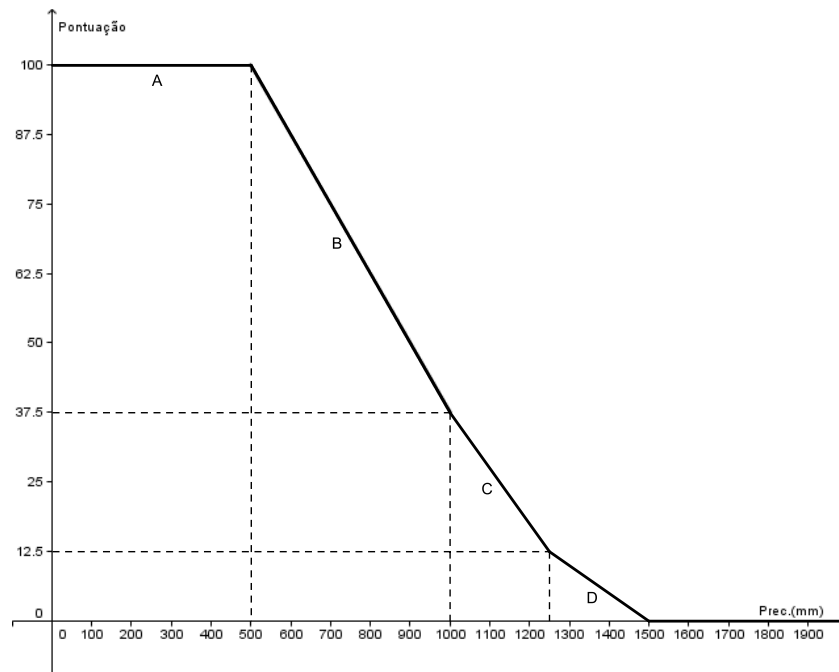


Figura 4.7 Gráfico da escala de pontuação - "Natureza Climatológica - PV12"

4.2.3 Sub-indicador: Natureza Sócio-econômica - Efeito sobre a Base Rural - PV131

A adaptação desse sub-indicador para o ISFS seguiu a definição original utilizada no IMERG, ou seja, seguindo uma escala de pontuação linear função do número de habitantes

na zona rural do município. A função de valor foi obtida de acordo com a distribuição da população rural (Figura 4.8) dada pela equação $y = c \times x$, onde x é a população rural do município, $c = \frac{100}{pop_{max}} = 0,005$ e $pop_{max} = 20000$ o limite máximo da população rural utilizada nos cálculos. Para populações acima de 20000 habitantes atribui-se 100 pontos. As populações rurais foram obtidas no Censo 2000 do IBGE. A atualização desse sub-indicador ocorrerá de acordo com a periodicidade de atualização das informações populacionais dos Censos, o que se dá a cada dez anos.

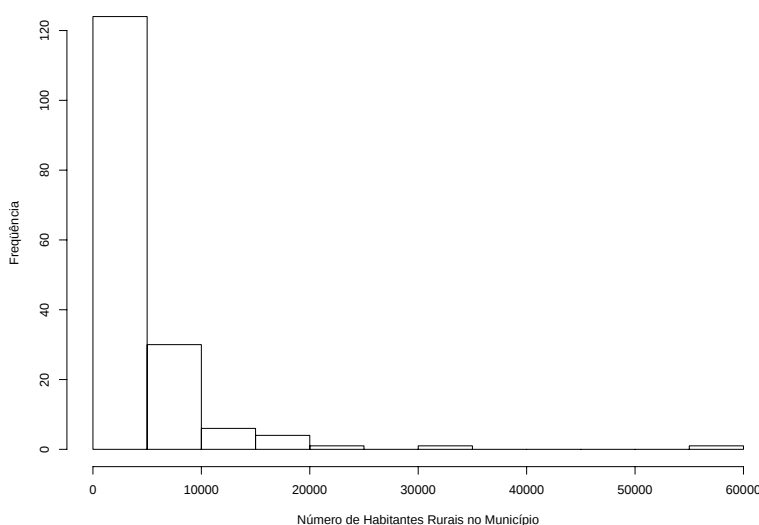


Figura 4.8 Distribuição da População Rural do Rio Grande do Norte. Fonte: IBGE, Censo 2000.

4.2.4 Sub-indicador: Natureza Sócio-econômica - Efeito sobre a Base Econômica - PV132

Para este sub-indicador também foi utilizada a definição do IMERG, sendo portanto obtido a partir do cálculo da participação do PIB rural em relação ao PIB total de cada município. O método de normalização, seguindo o utilizado no IMERG, foi definido a partir de uma escala de categorias dadas pelas equações 4.14 a 4.17 (Figura 4.9). Este sub-indicador terá uma periodicidade de atualização anual, pois o PIB dos municípios apresenta variação de ano para ano.

$$A : \quad y = 0,8x \quad (4.14)$$

$$B : y = 1,6x - 10 \quad (4.15)$$

$$C : y = 2,4x - 30 \quad (4.16)$$

$$D : y = 3,2x - 60 \quad (4.17)$$

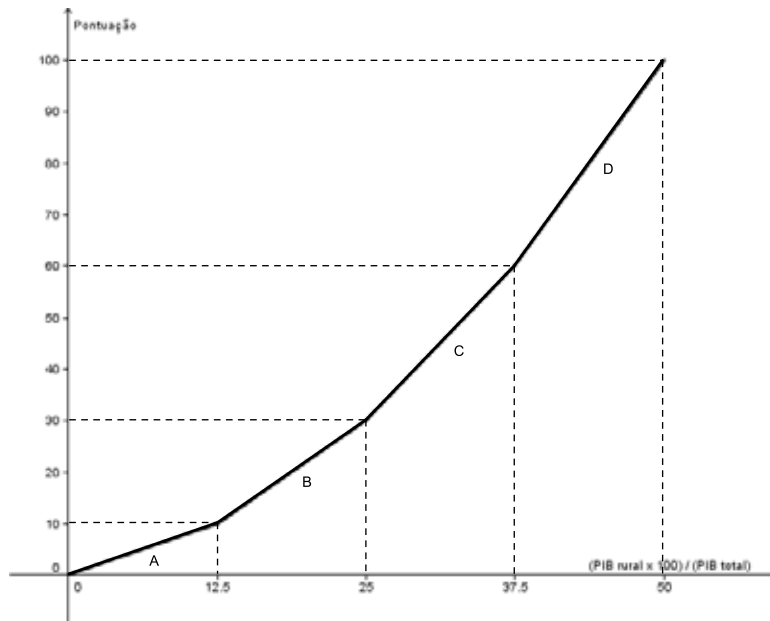


Figura 4.9 Gráfico da escala de pontuação - “Efeito sobre a Base Econômica - PV132”

4.2.5 Sub-indicador: Natureza Sócio-econômica - Condições Municipais de Atenuação dos efeitos da Seca - PV133

Para este sub-indicador, também como no IMERG, foi definido uma normalização obtida por uma função linear dependente da quantidade de receita *per capita* (*Recpc*) recebida pelos municípios:

$$Recpc = \text{Receitas Correntes}_{mun} / \text{População}_{mun} \quad (4.18)$$

Como a receita corrente do município é variável de ano para ano, esse sub-indicador também apresentará essa variação. Na Figura 4.10 é apresentada a escala de pontuação utilizada para a normalização deste sub-indicador, cujos limites foram obtidos em função da distribuição de frequência das receitas per capita médias dos municípios no período de 1994 a 2006 (Figura

4.11). Os dados de receitas dos municípios foram obtidos na Secretaria de Tesouro Nacional do Ministério da Fazenda e os de população no censo 2000 do IBGE.

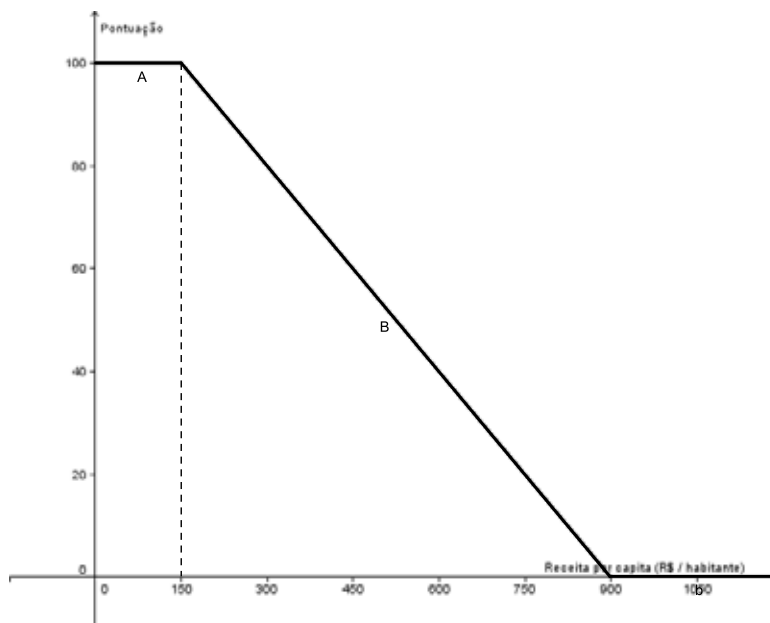


Figura 4.10 Gráfico da escala de pontuação - “Condições Municipais de Atenuação à Seca - PV133”

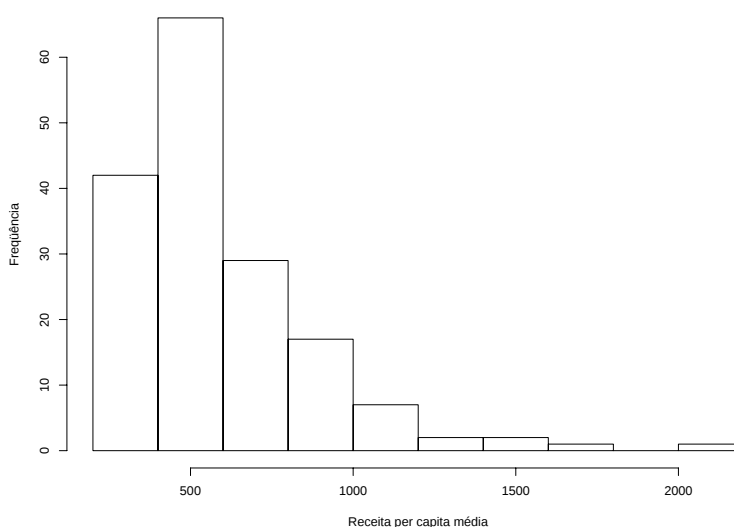


Figura 4.11 Distribuição da Receita per capita média dos municípios do Rio Grande do Norte (Período: 1994-2006). Fonte: Ministério da Fazenda, Secretaria Tesouro Nacional.

4.2.6 Sub-indicador: Risco de Perda da Produção Agrícola - PV21

Neste sub-indicador foi mantida a definição original do IMERG com respeito a utilização do Índice de Distribuição de Chuvas (IDC), desenvolvido pela FUNCEME, e o método de

normalização obtido a partir de uma escala de categorias de pontuação usada para obtenção do valor do sub-indicador no cálculo do IMERG [16,64,72,157].

O IDC tem sido utilizado pela FUNCEME, desde 1993, para classificar os municípios quanto à variação da precipitação pluviométrica no Estado do Ceará [108]. Este parâmetro tem sido utilizado também pela Defesa Civil do Ceará no planejamento da assistência aos flagelados com os constantes períodos de estiagens na região, e a partir de 2004, na elaboração do Índice Municipal de Alerta (IMA) pelo Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Este último congrega o esforço de várias instituições do Estado do Ceará, dentre elas a FUNCEME, com intuito de criar um instrumento de monitoramento socioeconômico e ambiental, voltado às atividades da agropecuária, e que possibilite uma intervenção do Estado em convênio com as prefeituras municipais, em tempo hábil, onde ações emergenciais se façam necessárias [91,92].

O IDC associa variações pluviométricas, temporais e espaciais da precipitação, levando-se em consideração o período escolhido para a análise e é obtido a partir das seguintes equações:

$$IS = \frac{\psi_1 \times \psi_3 \times P_{obs}}{ND \times (ND - \psi_2)} \quad (4.19)$$

$$IDC = \frac{1000 \times IS}{P} \quad (4.20)$$

onde,

IS = Índice de Seca;

ψ_1 =Número de dias com chuva;

ψ_2 =Duração em dias do maior período sem chuva;

ψ_3 =Número de dias do maior período de chuvas, no qual não ocorreu um verânico igual ou superior a 15 dias;

P_{obs} =Precipitação observada no período para o posto pluviométrico ou município;

IDC =Índice de Distribuição de Chuvas;

P =Precipitação média observada no período escolhido para todo o Estado;

ND =Número total de dias do período escolhido para cálculo do IDC.

A partir daí, os resultados obtidos são classificados nas categorias dadas pelo quadro abaixo:

Categoria	Variação do IDC	Classificação	Descrição
1 ^a	≥ 301	Muito Bom	A precipitação apresenta-se acima da normal climatológica e/ou nas categorias Chuvoso e Muito Chuvoso.
2 ^a	201 - 300	Bom	A precipitação apresenta-se em torno da normal climatológica e/ou na categoria normal .
3 ^a	101 - 201	Regular	A precipitação apresenta-se no limite inferior da categoria normal com quebra na safra agrícola.
4 ^a	000 - 100	Crítica	A precipitação apresenta-se entre as categorias seco e muito seco.

Quadro 4.4 Classificação utilizada para o Índice de Distribuição de Chuvas.

Na Figura 4.12 é apresentada a escala de pontuação para a normalização desse sub-indicador, dada pelas equações 4.21 a 4.24, na qual são prestigiados municípios com IDCs baixos e desprestigiados municípios com IDCs altos. O período considerado para cálculo do IDC é a quadra chuvosa de cada microrregião onde está inserido o município em questão, obtida a partir do estudo realizado para o sub-indicador “Natureza Climatológica” (PV12), porém esse período poderá variar. Este é um sub-indicador cuja atualização se dará de ano para ano.

$$A : \quad y = 100 - 0,1x \quad (4.21)$$

$$B : \quad y = 120 - 0,5x \quad (4.22)$$

$$C : \quad y = 97,5 - 0,35x \quad (4.23)$$

$$D : \quad y = 13,33 - 0,0133x \quad (4.24)$$

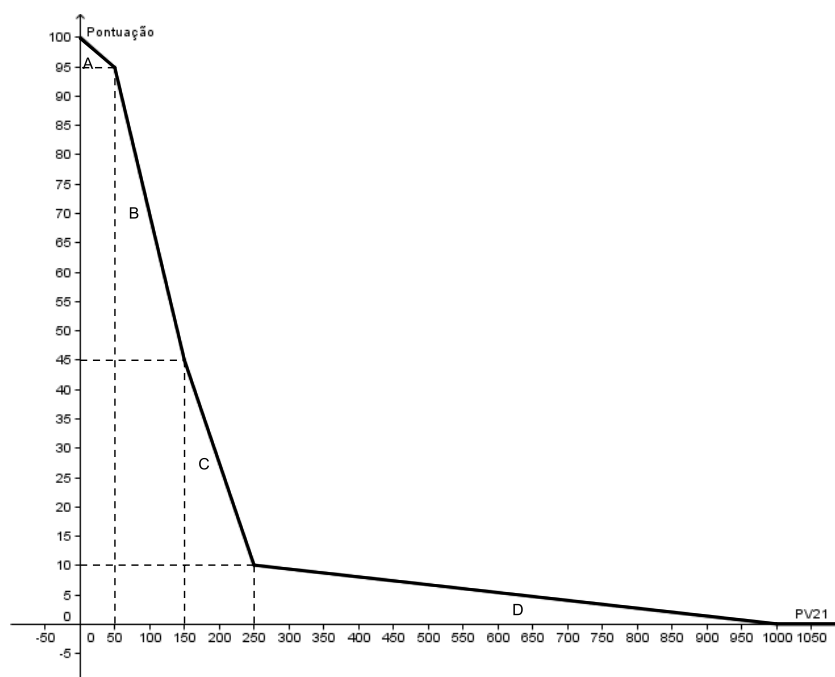


Figura 4.12 Gráfico da escala de pontuação - “Risco de perda da produção agrícola - PV21” . Fonte: SILVA & FORTE [157].

4.2.7 Sub-indicador: Risco de Falta de Água para Consumo Humano e Animal - PV22

Por fim, para compor esse sub-indicador primeiramente foi realizado um levantamento dos tipos de informações disponíveis e que poderiam qualificar de forma mais precisa o risco de falta de água para as populações dos municípios no semi-árido nordestino, particularmente, no RN. Foram localizadas quatro fontes de informações: 1) o documento “Atlas Nordeste: abastecimento urbano de água: alternativas de oferta de água para as sedes municipais da Região Nordeste do Brasil e do norte de Minas Gerais” da Agência Nacional de Águas (ANA) [4]; 2) um Relatório Administrativo Interno com o número de economias (Água e Esgoto) cadastradas da Companhia de Água e Esgotos do Rio Grande do Norte (CAERN); 3) Informações sobre o abastecimento de água contidas no Censo 2000 - IBGE; e 4) diversos tipos de indicadores sobre água e esgotos contidos no Sistema Nacional de Informações sobre Saneamento (SNIS) do Ministério das Cidades.

O documento da ANA, que foi publicado no ano de 2006, apresenta os resultados de um estudo de avaliação da situação atual dos recursos hídricos e a tendência futura para os Estados do Nordeste e municípios do Norte de Minas Gerais [4]. Neste estudo foram considerados municípios com população superior ou igual a 5 mil habitantes, de acordo com o censo 2000 do IBGE. No Rio Grande do Norte foram considerados 102 municípios, ou seja, 59% do total de

municípios do Estado. Foram avaliadas a disponibilidade hídrica dos mananciais superficiais e subterrâneos, a qualidade da água desses mananciais, a capacidade dos sistemas produtores e de distribuição e a demanda para o abastecimento humano. A partir destes critérios foram classificadas as sedes municipais nas seguintes categorias: sedes com abastecimento satisfatório, sedes em situação crítica por sistema, sedes em situação crítica por manancial e sedes em situação crítica por manancial e sistema. Esse estudo, dada a natureza das informações nele contidas, poderia ser escolhido para definir o sub-indicador em questão, porém a sua abrangência reduzida inviabilizaria a sua utilização, pois ficariam muitos municípios sem informação.

Nesse mesmo sentido, as informações fornecidas pela CAERN e o IBGE apesar de apresentarem uma cobertura abrangendo todos os municípios do RN, só representam o nível de atendimento do abastecimento de água.

Por outro lado, as informações contidas no SNIS fornecem um grande número de indicadores relevantes para todos os segmentos envolvidos com o setor de saneamento a nível nacional, congregando informações das diversas prestadoras de serviço de abastecimento de água e de esgotamento sanitário no Brasil, além de abranger quase a totalidade dos municípios do RN. Nesse sistema, pode-se encontrar informações administrativo-financeiras, de balanço contábil, operacionais (água e esgoto) e de qualidade dos sistemas envolvidos (água e esgoto). A coleta de dados e a elaboração dos diagnósticos dos serviços de água e esgotos iniciaram-se em 1995 e desde então estão sendo realizadas anualmente [107].

Portanto, para compor esse sub-indicador, optou-se pela utilização de três indicadores, disponíveis no SNIS: o Índice de Atendimento Total de Água (I_{055}), o Índice Bruto de Perdas Lineares (I_{050}) e o Consumo Médio Per Capita de Água (I_{022}), que são calculados pelas equações 4.25 a 4.27. Nesse estudo, a título de compatibilidade em nomenclatura, optou-se por utilizar os mesmos nomes para os índices dos parâmetros usados no SNIS.

$$I_{055} = \frac{A_{01}}{G_{12a}} \quad (4.25)$$

$$I_{050} = \frac{(A_{06} + A_{18} - A_{24}) - A_{10}}{A_{05}} \quad (4.26)$$

$$I_{022} = \frac{A_{10} - A_{19}}{A_{01}} \quad (4.27)$$

onde,

A_{01} = População total atendida com abastecimento de água;

G_{12a} = População total do município (IBGE);

A_{05} = Extensão da rede de água;

A_{06} = Volume de água produzido;

A_{10} = Volume de água consumido;

A_{18} = Volume de água tratada importada;

A_{19} = Volume de água tratada exportada;

A_{24} = Volume de água de serviço.

Portanto, o sub-indicador $PV22$ será obtido pela seguinte equação:

$$PV22 = I_{055} \times \frac{1}{(1 + I_{050})} \times I_{022} \quad (4.28)$$

Objetiva-se com isso identificar municípios com alta vulnerabilidade hídrica traduzida pelo baixo atendimento de abastecimento, alto índice de perdas e baixo consumo médio per capita de água. Quanto às perdas de água optou-se por considerar as perdas lineares, pois, segundo MORAIS [114], os indicadores de perdas de água percentuais têm sido considerados inadequados para a avaliação de desempenho, por serem fortemente influenciados pelo consumo e imprimirem uma característica de homogeneidade aos sistemas. Isto na prática não se verifica, pois, para um mesmo volume de água perdida, quanto maior o consumo menor será o índice de perdas em percentual e os sistemas se apresentam de forma bastante heterogênea.

Na Figura 4.13 é apresentada a distribuição desse parâmetro para os 167 municípios do RN e na Figura 4.14 a escala de pontuação utilizada para a normalização desse sub-indicador, dada pelas equações 4.29 a 4.31 que representam os segmentos de retas A, B e C, respectivamente. Para valores maiores que 1000 é atribuída a pontuação zero. Para a definição dos limites que constituem a escala de pontuação desse sub-indicador utilizaram-se os dados do ano de 2006, pois, apesar da série de dados contida no SNIS ter sido iniciada em 1995, somente a partir do ano de 2006 esta apresentou dados significativamente consistentes.

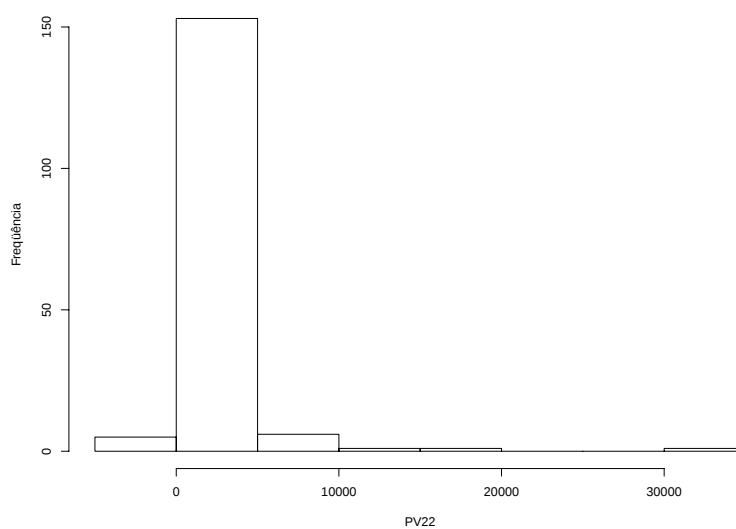


Figura 4.13 Distribuição do Parâmetro: “Risco de falta de água para consumo humano e animal - PV22”.

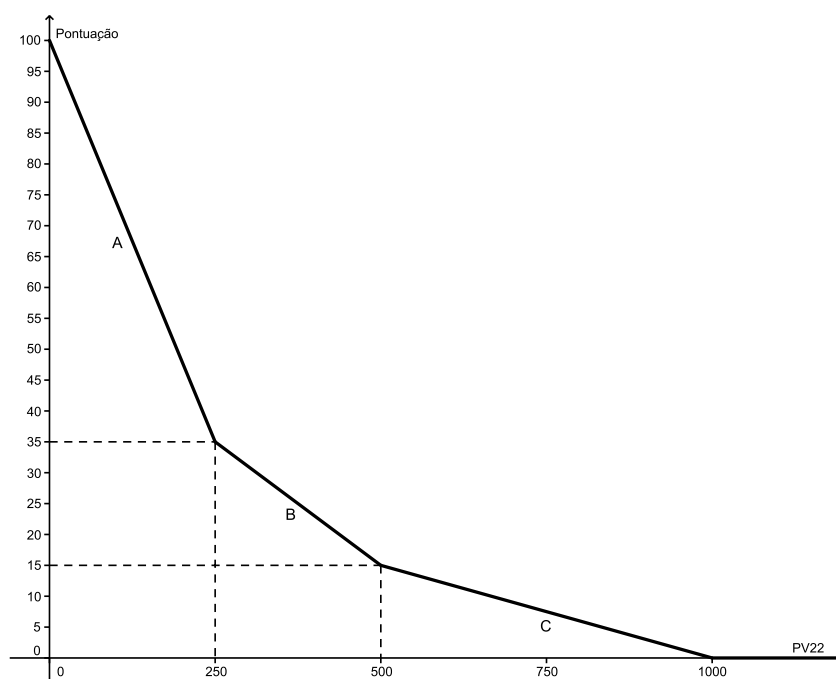


Figura 4.14 Gráfico da escala de pontuação - “Risco de falta de água para consumo humano e animal - PV22” .

$$A : \quad y = 100 - 0,26x \quad (4.29)$$

$$B : \quad y = 55 - 0,08x \quad (4.30)$$

$$C : \quad y = 30 - 0,03x \quad (4.31)$$

4.3 Pesos e Método de Agregação Utilizados no Cálculo do ISFS

Os pesos em um indicador composto refletem a importância relativa associada a cada sub-indicador ou dimensão [120]. Os pesos definidos para o ISFS foram os mesmos utilizados no IMERG (Figura 2.25), que foram resultado da aplicação de uma metodologia multicritério de apoio à decisão junto a uma equipe de experts sobre o Nordeste Brasileiro, em particular, sobre a Seca Nordestina.

O método de agregação escolhido foi o aditivo (linear), no qual a compensabilidade entre os sub-indicadores é constante. Embora esse método apresente a característica indesejável de performances pobres de alguns sub-indicadores poderem ser compensadas por performances suficientemente altas de outros sub-indicadores [120], esse efeito é atenuado em função dos diferentes pesos utilizados no ISFS e a depender da combinação destes com as performances dos sub-indicadores. Ademais, por admitir o valor 0 (zero) na pontuação final dos sub-indicadores o método de agregação geométrico ou multiplicativo torna-se inviável para utilização no ISFS.

4.4 O Preenchimento dos Dados com Falhas e a Winsorização no ISFS

Em ambientes que envolvem uma grande variedade de coleções de dados, provenientes de diversas fontes de informações, é comum ocorrer ausências de dados ou observações incompletas nessas coleções [82, 120, 124, 125, 151, 186]. Esse é o caso do ISFS, cuja elaboração envolve uma grande quantidade de dados e uma gama considerável de fontes de informações.

Conforme descrito no capítulo anterior, item 3.4, existem muitos métodos estatísticos para se tratar com as falhas de dados, e a sua aplicação dependerá do mecanismo e do padrão de não-resposta, e da percentagem dos dados com falhas em relação aos dados totais. Tanto o preenchimento simples quanto o múltiplo se aplicam ao ISFS, conforme será detalhado no próximo capítulo, porém com o intuito de minimizar o grau de incerteza associado a ausência dos dados e ao processo de preenchimento destes, foi escolhido o preenchimento múltiplo e o

método de Monte Carlo baseado em Cadeias de Markov, esse último devido ao padrão de não resposta identificado nos dados.

O detalhamento de todo o processo de preenchimento efetuado nos testes, tanto a parte conceitual quanto a computacional, incluindo as estimativas dos modelos de regressões preliminares e os modelos finais serão apresentados no próximo capítulo.

Antes porém de ser processado o preenchimento múltiplo dos dados é necessário testar a normalidade multivariada destes. O procedimento adotado para o ISFS foi o descrito no Capítulo 3, item 3.3.4, com uma pequena modificação no passo dois desse procedimento. Logo o procedimento adotado foi:

1. Antes de executar o preenchimento múltiplo transformar todas as variáveis ou sub-indicadores com assimetria maior que dois,
2. Após o preenchimento dos dados transformar novamente as variáveis para sua escala original.

Note que, no procedimento original essa transformação de retorno só ocorre se a assimetria dos dados transformados e preenchidos for menor ou igual a quatro, porém, como o ISFS tem uma fase posterior de winsorização nos sub-indicadores, supõe-se que com a aplicação dessa técnica haverá uma redução substancial na assimetria, dispensando, portanto, o teste da assimetria e a manutenção dos dados transformados. Com isso evita-se também, o impacto da utilização de uma escala diferente no computo final do ISFS, após o procedimento de agregação.

A winsorização é uma técnica estatística que permite a suavização das caudas de uma distribuição (ver Capítulo 3, item 3.5). No ISFS esta técnica foi utilizada para dois fins: primeiro com intuito de diminuir a assimetria nos dados dos sub-indicadores, principalmente pelo motivo descrito acima; e, segundo objetivando-se a eliminação de *outliers* provenientes dos próprios dados ou do procedimento de preenchimento múltiplo dos dados com falhas.

Foi escolhido dois patamares para a winsorização: 1) $\alpha = 0,1$, o que significa, um corte no quantil 0,95 e outro no quantil 0,05; e, 2) $\alpha = 0,05$, o que significa, um corte no quantil 0,975 e outro no quantil 0,025. Essas opções foram disponibilizadas para escolha do usuário no software de operacionalização do ISFS, conforme será visto no Capítulo 7. No entanto, para os testes de validação do ISFS, que serão descritos no próximo capítulo, foi adotado $\alpha = 0,1$.

Maiores detalhes da aplicação desse método serão apresentados no próximo capítulo.

4.5 Formulação Matemática e Estrutura do ISFS

O Índice de Susceptibilidade ao Fenômeno da Seca será, portanto, obtido a partir da seguinte equação:

$$ISFS_m^t = \sum_{q=1}^7 w_q I_{q,m}^t, \quad \text{com } w_1 + w_2 + \dots + w_7 = 1 \text{ e } I_{q,m}^t \neq -1 \quad (4.32)$$

onde,

$ISFS_m^t$ = valor do indicador composto para o município m no período t , com $m = 1, 2, \dots, L$ e L representando o número total de municípios considerados no cálculo.

$I_{q,m}^t$ = valor normalizado do subindicador q para o município m no tempo t .

w_q = peso associado ao subindicador q com $w_1 = 0,15$; $w_2 = 0,12$; $w_3 = 0,15$; $w_4 = 0,12$; $w_5 = 0,06$; $w_6 = 0,2$ e $w_7 = 0,2$.

sendo que,

$$I_{q,m}^t = f_q(x_{q,m}^t) \in \{[0, 100] \cup \{-1\}\} \quad (4.33)$$

com,

$x_{q,m}^t$ = valor bruto do subindicador q para o município m no período t .

$f_q(\)$ = função de normalização do subindicador q .

Logo, tem-se,

Para $q = 1$ (*Sub-indicador - PV11*):

$$x_{1,m}^t = \frac{PAA_m \times (ALP_m + ALT_m + ALTD_m + APN_m + APP_m)}{1000} \quad (4.34)$$

com,

PAA_m = peso (pontuação) da aptidão agrícola do solo do município m .

ALP_m = área da lavoura permanente do município m (ha).

ALT_m = área da lavoura temporária do município m (ha).

$ALTD_m$ = área da lavoura temporária em descanso do município m (ha).

APN_m = área de pastagem natural do município m (ha).

APP_m = área de pastagem plantada do município m (ha).

$$\Rightarrow f_1(x_{1,m}^t) = \begin{cases} 0,07x_{1,m}^t, & se\ x_{1,m}^t \leq 500 \\ 0,06x_{1,m}^t + 5, & se\ 500 < x_{1,m}^t \leq 1000 \\ 0,02x_{1,m}^t + 45, & se\ 1000 < x_{1,m}^t \leq 2000 \\ 0,0075x_{1,m}^t + 70, & se\ 2000 < x_{1,m}^t \leq 4000 \\ 100, & se\ x_{1,m}^t > 4000 \\ -1, & se\ \nexists x_{1,m}^t \end{cases} \quad (4.35)$$

Para $q = 2$ (*Sub-indicador - PV12*):

$x_{2,m}^t =$ média pluviométrica histórica do município m .

$$\Rightarrow f_2(x_{2,m}^t) = \begin{cases} 100, & se\ x_{2,m}^t \leq 500 \\ 100 - 0,125 \times (x_{2,m}^t - 500), & se\ 500 < x_{2,m}^t \leq 1000 \\ 37,5 - 0,1 \times (x_{2,m}^t - 1000), & se\ 1000 < x_{2,m}^t \leq 1250 \\ 12,5 - 0,05 \times (x_{2,m}^t - 1250), & se\ 1250 < x_{2,m}^t \leq 1500 \\ 0, & se\ x_{2,m}^t > 1500 \\ -1, & se\ \nexists x_{2,m}^t \end{cases} \quad (4.36)$$

Para $q = 3$ (*Sub-indicador - PV131*):

$x_{3,m}^t =$ população rural do município m .

$$\Rightarrow f_3(x_{3,m}^t) = \begin{cases} 0,005 \times x_{3,m}^t, & \text{se } 0 < x_{3,m}^t \leq 20000 \\ 100, & \text{se } x_{3,m}^t > 20000 \\ -1, & \text{se } \nexists x_{3,m}^t \end{cases} \quad (4.37)$$

Para $q = 4$ (*Sub-indicador - PV132*):

$$x_{4,m}^t = \frac{(PIB_{rural})_m \times 100}{(PIB_{total})_m} \quad (4.38)$$

onde PIB_{rural} e o PIB_{total} referem-se ao ano anterior ao do cálculo.

$$\Rightarrow f_4(x_{4,m}^t) = \begin{cases} 100, & \text{se } x_{4,m}^t > 50 \\ 3,2x_{4,m}^t - 60, & \text{se } 37,5 < x_{4,m}^t \leq 50 \\ 2,4x_{4,m}^t - 30, & \text{se } 25 < x_{4,m}^t \leq 37,5 \\ 1,6x_{4,m}^t - 10, & \text{se } 12,5 < x_{4,m}^t \leq 25 \\ 0,8x_{4,m}^t, & \text{se } 0 \leq x_{4,m}^t \leq 12,5 \\ -1, & \text{se } \nexists x_{4,m}^t \end{cases} \quad (4.39)$$

Para $q = 5$ (*Sub-indicador - PV133*):

$x_{5,m}^t =$ receita per capita recebida pelo município m no ano anterior ao do cálculo.

$$\Rightarrow x_{5,m}^t = \frac{(ReceitasCorrentes)_m}{(Populacao)_m} \quad (4.40)$$

$$\Rightarrow f_5(x_{5,m}^t) = \begin{cases} 100, & \text{se } x_{5,m}^t \leq 150 \\ 120 - 0,1334x_{5,m}^t, & \text{se } 150 < x_{5,m}^t < 900 \\ 0, & \text{se } x_{5,m}^t \geq 900 \\ -1, & \text{se } \nexists x_{5,m}^t \end{cases} \quad (4.41)$$

Para $q = 6$ (*Sub-indicador - PV21*):

$x_{6,m}^t =$ Índice de Distribuição de Chuvas(IDC) para o município m no período t .

$$\Rightarrow x_{6,m}^t = \frac{1000 \times IS}{P} \quad (4.42)$$

e,

$$IS = \frac{\psi_1 \times \psi_3 \times P_{obs}}{ND \times (ND - \psi_2)} \quad (4.43)$$

onde,

IS = Índice de Seca;

ψ_1 = Número de dias com chuva;

ψ_2 = Duração em dias do maior período sem chuva;

ψ_3 = Número de dias do maior período de chuvas, no qual não ocorreu um verânico igual ou superior a 15 dias;

P_{obs} = Precipitação observada no período;

P = Precipitação média observada no período escolhido para todo o Estado;

ND = Número total de dias do período escolhido para cálculo do IDC.

$$\Rightarrow f_6(x_{6,m}^t) = \begin{cases} 100 - 0,1x_{6,m}^t, & se \ x_{6,m}^t \leq 50 \\ 120 - 0,5x_{6,m}^t, & se \ 50 < x_{6,m}^t \leq 150 \\ 97,5 - 0,35x_{6,m}^t, & se \ 150 < x_{6,m}^t \leq 250 \\ 13,33 - 0,0134x_{6,m}^t, & se \ 250 < x_{6,m}^t \leq 1000 \\ 0, & se \ x_{6,m}^t > 1000 \\ -1, & se \ \nexists \ x_{6,m}^t \end{cases} \quad (4.44)$$

Para $q = 7$ (Sub-indicador - PV22):

$$x_{7,m}^t = I_{055} \times \frac{1}{1 + I_{050}} \times I_{022} \quad (4.45)$$

onde,

I_{055} = Índice de Atendimento Total de Água;

I_{050} = Índice Bruto de Perdas Lineares;

I_{022} = Consumo Médio Per Capita de Água.

$$\Rightarrow f_7(x_{7,m}^t) = \begin{cases} 100 - 0,26x_{7,m}^t, & \text{se } x_{7,m}^t \leq 250 \\ 55 - 0,08x_{7,m}^t, & \text{se } 250 < x_{7,m}^t \leq 500 \\ 30 - 0,03x_{7,m}^t, & \text{se } 500 < x_{7,m}^t \leq 1000 \\ 0, & \text{se } x_{7,m}^t > 1000 \\ -1, & \text{se } \nexists x_{7,m}^t \end{cases} \quad (4.46)$$

Na Figura 4.15 é apresentado o fluxograma de cálculo para obtenção do ISFS. O procedimento é realizado no máximo em 6 fases, sendo as principais F-1, F-3, F-5 e F-6, constituindo o cálculo inicial do índice, o preenchimento múltiplo das falhas, a winsorização e a geração de mapas, tabelas e gráficos com os resultados finais, respectivamente. Se, no cálculo do ISFS para um ano qualquer, não ocorrerem falhas nos dados, o procedimento se completará com o mínimo de duas fases. Na F-1 são computadas todas as variáveis que constituem os 7 sub-indicadores do ISFS, conforme as equações descritas anteriormente. É interessante ressaltar que, quando da ausência ou falha em uma das variáveis que compõem um sub-indicador, este é computado como falha. A F-2 é opcional, somente ocorrerá se houver sub-indicador com assimetria maior ou igual a dois. Este procedimento é realizado com o intuito de melhorar as características da distribuição dos dados e aproximar-se da condição de normalidade multivariada exigida pelo método de preenchimento múltiplo utilizado na próxima fase. A F-3 constitui a rotina de preenchimento múltiplo das falhas nos dados ocorridas na fase inicial. Os detalhes desse procedimento serão apresentados no capítulo 5. A F-4 só ocorrerá se a F-2 for executada, e a F-5 constitui o procedimento de winsorização, que também será detalhado no capítulo 5. Por fim, a F-6 engloba as rotinas de geração do mapa final com os resultados do ISFS, a tabela com o resultado da classificação dos municípios em relação a susceptibilidade à seca e os gráficos de radar apresentando o impacto de cada um dos sub-indicadores na composição dos ISFS's dos municípios.

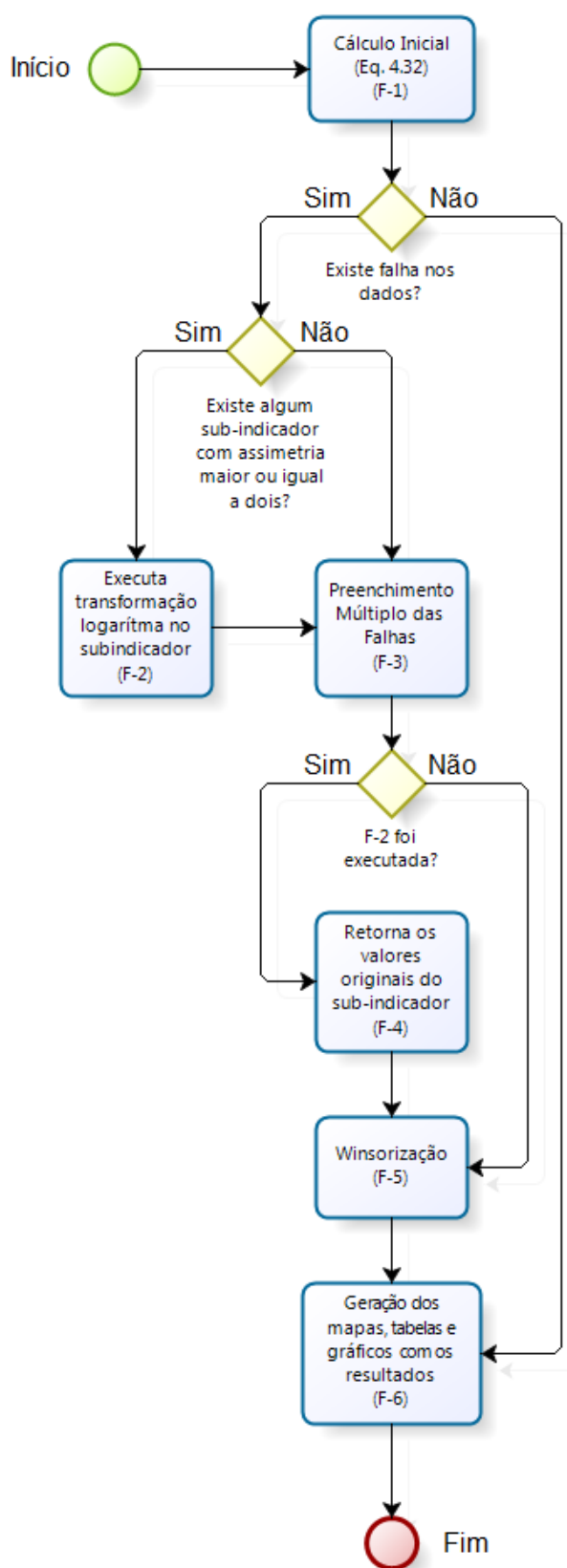


Figura 4.15 Estrutura do ISFS - Fluxograma de Cálculo.