



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
PROGRAMA DE PÓS-GRADUAÇÃO EM DEMOGRAFIA



Antonino Melo dos Santos

**MORTALIDADE INFANTIL E CONDIÇÕES SOCIOECONÔMICAS NAS
MICRORREGIÕES DO NORDESTE BRASILEIRO**

Natal
2013

Antonino Melo dos Santos

**MORTALIDADE INFANTIL E CONDIÇÕES SOCIOECONÔMICAS NAS
MICRORREGIÕES DO NORDESTE BRASILEIRO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Demografia da Universidade Federal do Rio Grande do Norte, como requisito parcial para a obtenção do título de Mestre em Demografia.

Orientador:

Prof. Dr. Paulo Cesar Formiga Ramos

Co-orientadora:

Prof^a. Dr^a. Maria Célia de Carvalho Formiga

Natal

2013

Antonino Melo dos Santos

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Demografia da Universidade Federal do Rio Grande do Norte, como requisito parcial para a obtenção do título de Mestre em Demografia.

Aprovada em: 07/06/2013

BANCA EXAMINADORA

Prof. Dr. Paulo Cesar Formiga Ramos (UFRN)

(Orientador)

Prof.^a Dr.^a Maria Célia de Carvalho Formiga (UFRN)

(Examinadora Interna)

Prof.^a Dr.^a Lára de Melo Barbosa Andrade (UFRN)

(Examinadora Interna)

Prof. Dr. Mardone Cavalcante França (UFRN)

(Examinador Interno)

Prof. Dr. Kleber Fernandes de Oliveira (UFS)

(Examinador Externo)

Dedico aos meus Pais, estes que muito suaram e se esforçaram me dando atenção, amor e, sobretudo, educação, para que um dia eu chegasse ao patamar onde me encontro.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me concedido energia e sabedoria para terminar essa batalha árdua, que foi a conclusão desse mestrado.

Aos meus pais, Antonio e Rosimar, meu irmão Artur, pelo amor, compreensão, ajuda, paciência, força e incentivo durante todo esse período de provações.

À minha família (tios, tias, primos e primas), que muito atenciosa, me apoiou e incentivou a seguir em frente nos meus objetivos. Em especial, ao primo Rodrigo Lucas, que assim como eu, nos tornamos mestrandos, porém em áreas diferentes, mas sempre conversando sobre assuntos acadêmicos e científicos, enriquecendo nossas experiências dentro da universidade.

Dirijo, também, este agradecimento, à professora Maria Célia e ao professor Paulo Cesar Formiga (casal Formiga), que considero meus “Pais acadêmicos”, estes que me acolheram e, deram a oportunidade de experimentar e trabalhar uma área chamada Demografia, ainda na época da graduação. Área esta, que estou concluindo o mestrado e que a cada dia me identifico mais.

Aos colegas de curso, pelo empenho, trabalhos em conjunto, dedicação ao grupo, animação, alegria, entusiasmo, e principalmente, pelo companheirismo em todos os momentos do nosso curso. Creio que sem estas qualidades este grupo não seria o mesmo.

Ao professor Moises Alberto Calle, pelas conversas e palavras incentivadoras e idealizadoras, que tanto contribuíram para o enriquecimento dos meus trabalhos. Com sua forma simples de ser e seu jeito alegre, conseguia levantar a autoestima, não só a minha, como também da turma.

Agradeço à professora Lára de Melo, nossa coordenadora, pelos direcionamentos e incentivos nas tomadas de decisões acerca da minha dissertação.

A Fabiola, minha namorada, amiga e confidente, que sempre me estimulou e que me deu muito apoio, carinho e amor em momentos, tanto de dificuldades como também de alegrias.

Enfim, agradeço a todos que de forma direta e indireta, contribuíram para o meu crescimento e desenvolvimento acadêmico.

***“Todo guerreiro já ficou com medo de entrar em
combate.***

[...]

Todo guerreiro já perdeu a fé no futuro.

***Todo guerreiro já trilhou um caminho que não era
dele.***

Todo guerreiro já sofreu por bobagens.

[...]

Todo guerreiro já achou que não era guerreiro.

Todo guerreiro já falhou em suas obrigações.

***Todo guerreiro já disse "SIM" quando queria dizer
"NÃO".***

Todo guerreiro já feriu alguém que amava.

***Por isso é um guerreiro; porque passou por estes
desafios, e não perdeu a esperança de ser melhor
do que era.”***

(Paulo Coelho)

RESUMO

O estudo propõe-se responder à seguinte questão: quais os diferentes perfis da mortalidade infantil, segundo variáveis demográficas, socioeconômicas, de infraestrutura e de assistência à saúde, das microrregiões do Nordeste brasileiro? Assim, o objetivo principal é analisar os perfis ou tipologias de mortalidade associados aos níveis de condições sociodemográficas das microrregiões, no ano de 2010. Para tanto, fez-se uso das bases de dados do SIM e SINASC (DATASUS/MS), dos microdados do Censo populacional de 2010 e do SIDRA/IBGE. Utilizou-se como variável resposta, a mortalidade infantil e, como independentes, variáveis demográficas, socioeconômicas, de infraestrutura e de assistência à saúde das microrregiões. Como metodologias de análise, foram empregadas: a regressão linear múltipla ponderada, para encontrar as variáveis mais significantes na explicação da mortalidade infantil, para o ano de 2010 e a análise de clusters, buscando encontrar indícios, inicialmente, de grupos homogêneos de microrregiões, a partir das variáveis significantes. Utilizou-se como variável resposta, o logito da taxa de mortalidade infantil e, como independentes, variáveis demográficas, socioeconômicas, de infraestrutura e de assistência à saúde das microrregiões. A técnica de estimação Bayesiana Empírica foi aplicada às informações de óbitos e nascimentos, devido ao fato inconveniente da subnotificação e das flutuações aleatórias de pequenos números existentes nas pequenas áreas. As técnicas de Estatística Espacial foram usadas para apurar espacialmente o comportamento da distribuição das taxas a partir de mapas temáticos. Concluindo, empregou-se o método *GoM (Grade of Membership)*, para encontrar tipologias de mortalidade associadas às variáveis, selecionadas por microrregião, buscando responder à questão principal do estudo. Os resultados apontam para a formação de três perfis: o perfil 1, de alta mortalidade infantil e condições sociais de vida desfavoráveis; o perfil 2, de baixa mortalidade infantil, com medianas condições sociais de vida; e o perfil 3, de mediana mortalidade infantil e altas condições sociais de vida. Com esta classificação, encontrou-se que, das 188 microrregiões, 20 (10%) enquadraram-se ao perfil extremo 1, 59 (31,4%) caracterizaram-se no perfil extremo 2, 34 (18,1%) caracterizaram-se no perfil extremo 3 e apenas 9 (4,8%) classificaram-se como perfil amorfo. As demais microrregiões enquadraram-se nos perfis mistos. Tais perfis

sugerem a necessidade de diferentes intervenções em termos de políticas públicas voltadas para a redução da mortalidade infantil na região.

Palavras-chave: Mortalidade infantil; Perfis de mortalidade infantil; Condições sociodemográficas.

ABSTRACT

The study aims to answer the following question: what are the different profiles of infant mortality, according to demographic, socioeconomic, infrastructure and health care, for the micro-regions at the Northeast of Brazil? Thus, the main objective is to analyze the profiles or typologies associated mortality levels sociodemographic conditions of the micro-regions, in the year 2010. To this end, the databases of birth and death certificates of SIM and SINASC (DATASUS/MS), were taken from the 2010 population Census microdata and from SIDRA/IBGE. As a methodology, a weighted multiple linear regression model was used in the analysis in order to find the most significant variables in the explanation child mortality for the year 2010. Also a cluster analysis was performed, seeking evidence, initially, of homogeneous groups of micro-regions, from of the significant variables. The logit of the infant mortality rate was used as dependent variable, while variables such as demographic, socioeconomic, infrastructure and health care in the micro-regions were taken as the independent variables of the model. The Bayesian estimation technique was applied to the database of births and deaths, due to the inconvenient fact of underreporting and random fluctuations of small quantities in small areas. The techniques of Spatial Statistics were used to determine the spatial behavior of the distribution of rates from thematic maps. In conclusion, we used the method GoM (Grade of Membership), to find typologies of mortality, associated with the selected variables by micro-regions, in order to respond the main question of the study. The results points out to the formation of three profiles: Profile 1, high infant mortality and unfavorable social conditions; Profile 2, low infant mortality, with a median social conditions of life; and Profile 3, median and high infant mortality social conditions. With this classification, it was found that, out of 188 micro-regions, 20 (10%) fits the extreme profile 1, 59 (31.4%) was characterized in the extreme profile 2, 34 (18.1%) was characterized in the extreme profile 3 and only 9 (4.8%) was classified as amorphous profile. The other micro-regions framed up in the profiles mixed. Such profiles suggest the need for different interventions in terms of public policies aimed to reducing child mortality in the region.

Keywords: Infant mortality, infant mortality profiles; sociodemographic conditions.

LISTA DE GRÁFICOS

Gráfico 01 – Comportamento da estrutura etária demonstrada na pirâmide etária Brasil, 1980, 1991, 2000 e 2010.....	23
Gráfico 02 – Comportamento da estrutura etária demonstrada na pirâmide etária da região Nordeste, 1980, 1991, 2000 e 2010.....	24
Gráfico 03 – Evolução da mortalidade proporcional CMD, para os óbitos infantis (menores de 1 ano), segundo região brasileira, 2000-2010.....	43
Gráfico 04 – Evolução da mortalidade proporcional CMD, para os óbitos infantis (menores de 1 ano), segundo UF da região Nordeste brasileira, 2000-2010.....	44
Gráfico 05 – Avaliação bivariada do comportamento do cruzamento das variáveis em análise - <i>Scatterplot</i> (Matriz).....	49
Gráfico 06 – Gráfico de passos (Step) versus o nível de distância do agrupamento (Linkge Distance) das 188 microrregiões, segundo as variáveis analisadas.....	52
Gráfico 07 – Dendograma de ligação simples, microrregiões da região Nordeste do Brasil, 2010.....	82

LISTA DE MAPAS

Mapa 01 – Comparativo das taxas de mortalidade infantil para o conjunto de informações corrigidas segundo estimador bayesiano empírico e cálculos diretos para os anos de 2000 e 2010, Nordeste, Brasil.....	46
Mapa 02 – Microrregiões do Nordeste – Tipologias dos níveis de mortalidade infantil e condições socioeconômicas 2010.....	67
Mapa 03 – Distribuição das microrregiões segundo a variável “Anos médio de estudo materno quando notificado o óbito infantil”, Região Nordeste, Brasil, 2010.....	83
Mapa 04 – Distribuição das microrregiões segundo a variável “Proporção de domicílios com abastecimento de água adequado da microrregião”, Região Nordeste, Brasil, 2010.....	83
Mapa 05 – Distribuição das microrregiões segundo a variável “Proporção de domicílios com existência de cobertura de esgotamento sanitário”, Região Nordeste, Brasil, 2010.....	84

Mapa 06 – Distribuição das microrregiões segundo a variável “Densidade domiciliar”, Região Nordeste, Brasil, 2010.....	84
Mapa 07 – Distribuição das microrregiões segundo a variável “Proporção de domicílio com existência de coleta de lixo adequada”, Região Nordeste, Brasil, 2010.....	85
Mapa 08 – Distribuição das microrregiões segundo a variável “Grau de Urbanização”, Região Nordeste, Brasil, 2010.....	85
Mapa 09 – Distribuição das microrregiões segundo a variável “Idade média das mães quando notificado o óbito infantil”, Região Nordeste, Brasil, 2010.....	86
Mapa 10 – Distribuição das microrregiões segundo a variável “Número médio de consultas Pré-Natal”, Região Nordeste, Brasil, 2010.....	86
Mapa 11 – Distribuição das microrregiões segundo a variável “Proporção de pessoas com baixa renda (< ½ S.M)”, Região Nordeste, Brasil, 2010.....	87
Mapa 12 – Distribuição das microrregiões segundo a variável “Renda Média domiciliar Per capita”, Região Nordeste, Brasil, 2010.....	87
Mapa 13 – Distribuição das microrregiões segundo a variável “Taxa de analfabetismo feminino”, Região Nordeste, Brasil, 2010.....	88

LISTA DE TABELAS

Tabela 01 – Sumário da regressão linear múltipla para a variável resposta Logit da Taxa de mortalidade infantil 2010, <i>método forward stepwise</i> , microrregiões da região Nordeste.....	51
Tabela 02 – Distribuição das microrregiões segundo UF e grupos formados na análise de <i>Clusters</i>	53
Tabela 03 – Frequências absolutas, relativas e estimativas de λ_{1j} , segundo perfis extremos.....	63
Tabela 04 – Distribuição das microrregiões segundo a predominância dos perfis extremos e mistos.....	66
Tabela 05 – Distribuição das microrregiões segundo a sua UF e perfil alocado....	68
Tabela 06 – Análise descritiva das variáveis estudadas por UF.....	89

LISTA DE QUADROS

Quadro 01 – Relação dos códigos e descrição das variáveis utilizadas nas análises, informações do ano de 2010, Microrregiões da região Nordeste.....	27
Quadro 02 – Descrição dos nomes simplificados das variáveis envolvidas no modelo de regressão linear múltipla ponderada.....	48
Quadro 03 – Análise (gráfica) dos resíduos do modelo de regressão linear múltipla ponderada.....	50
Quadro 04 – Análise gráfica (box-plots) da distribuição das variáveis: Taxa de mortalidade infantil (Tx_Mort_inf_ObC_2010), renda média domiciliar per capita (Rend_Pc), proporção de pessoas com baixa renda (Prop_Bx_rend) e taxa de analfabetismo feminino (Tx_Analf_M), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.....	55
Quadro 05 – Análise gráfica (box-plots) da distribuição das variáveis: idade média das mães (Id_Md_M), anos médio de estudo materno (A_Md_Est), Densidade domiciliar (Dens_Domic) e Número médio de consultas pré-natal (Num_Md_Cont_PN), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.....	57
Quadro 06 – Análise gráfica (box-plots) da distribuição das variáveis: Grau de urbanização (Gr_Urb), proporção de domicílios com abastecimento de água (Cobert_Agua), proporção de domicílios com cobertura de esgotamento sanitário (Cobert_Esg_Sanit) e proporção de domicílio com coleta de lixo adequada (Exist_C_Lixo), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.....	59
Quadro 07 – Grupos de microrregiões geradas com a análise de <i>cluster</i> e perfis encontrados com a utilização do Método <i>GoM</i> , região Nordeste, Brasil, 2010.....	77

LISTA DE ABREVIATURAS E SIGLAS

IPEA – Instituto de Pesquisa Econômica Aplicada;

TMI – Taxa de Mortalidade Infantil;

IBGE – Instituto Brasileiro de Geografia e Estatística;

CMD – Causas Mal Definidas;

UF – Unidade da Federação;

GoM – Grade of Membership;

DATASUS – Departamento de informática do Sistema Único de Saúde do Brasil;

SIM – Sistema de Informação sobre Mortalidade;

SINASC – Sistema de Informação sobre Nascidos Vivos;

SIDRA – Sistema IBGE de Recuperação Automática;

CID – Classificação Internacional de Doenças;

RIPSA – Rede Interagencial de Informações para a Saúde;

n.v. – Nascidos Vivos;

S.M. – Salários Mínimos;

MRLMP – Modelo de regressão linear múltipla ponderada.

SUMÁRIO

DEDICATÓRIA.....	III
AGRADECIMENTOS.....	IV
EPÍGRAFE.....	VI
RESUMO.....	VII
ABSTRACT.....	IX
LISTA DE GRÁFICOS.....	X
LISTAS DE MAPAS.....	X
LISTA DE TABELAS.....	XI
LISTA DE QUADROS.....	XII
LISTA DE ABREVIATURAS E SIGLAS.....	XIII
1. INTRODUÇÃO.....	16
1.1. DEFINIÇÃO DO PROBLEMA.....	19
1.2. JUSTIFICATIVA DA PESQUISA.....	20
1.3. OBJETIVO DO ESTUDO.....	21
1.3.1. <i>Objetivo Geral</i>	21
1.3.2. <i>Objetivos Específicos</i>	21
1.4. BRASIL E REGIÃO NORDESTE: COMPORTAMENTO FRENTE À MORTALIDADE INFANTIL E DINÂMICA DEMOGRÁFICA.....	22
1.5. SUB-REGISTRO DE ÓBITOS INFANTIS E NASCIDOS VIVOS.....	25
2. DADOS, MÉTODOS E OPERACIONALIZAÇÃO.....	26
2.1. ANÁLISE DE <i>CLUSTERS</i>	30
2.1.1. <i>Medida de similaridade e dissimilaridade</i>	30
2.1.2. <i>Técnica para construção de conglomerados (clusters)</i>	31
2.1.3. <i>Técnica Hierárquica Aglomerativa</i>	31
2.1.4. <i>Método para encontrar o Número g de Clusters da Partição Final</i>	33
2.2. ESTIMADOR BAYESIANO EMPÍRICO.....	33
2.3. MODELO DE REGRESSÃO LINEAR MÚLTIPLA PONDERADA (MRLMP).....	36
2.4. GRADE OF MEMBERSHIP - <i>GoM</i> (GRAU DE SIMILARIDADE).....	38
2.4.1. <i>Considerações gerais</i>	38
2.4.2. <i>Modelagem dos conjuntos nebulosos</i>	40

2.4.3. <i>Estimação dos parâmetros do modelo GoM</i>	41
3. RESULTADOS	42
3.1. AVALIAÇÃO DA QUALIDADE DOS DADOS COM BASE NA ANÁLISE DAS CAUSAS MAL DEFINIDAS REFERENTES À MORTALIDADE INFANTIL DO BRASIL E REGIÃO NORDESTE.....	42
3.2. CORREÇÃO DO SUB-REGISTRO DO SIM.....	44
3.3. SELEÇÃO DAS VARIÁVEIS PARA POSTERIOR ANÁLISE DE AGRUPAMENTOS.....	47
3.4. ANÁLISE DE <i>CLUSTERS</i> – FORMAÇÃO DE GRUPOS HOMOGÊNEOS..	51
3.5. CONSTRUÇÃO DOS PERFIS DE MORTALIDADE (<i>GoM</i>).....	53
3.5.1. <i>Características das UF e microrregiões da região Nordeste do Brasil</i>	53
3.5.2. <i>Tipologias das microrregiões segundo as variáveis trabalhadas</i>	59
4. DISCUSSÃO	68
5. CONSIDERAÇÕES FINAIS	69
6. REFERÊNCIAS	72
APÊNDICES	77

1. INTRODUÇÃO

Como um dos principais componentes da dinâmica demográfica, a mortalidade vem sendo estudada, não só por demógrafos, mas também por epidemiologistas e outros pesquisadores. Mas, por que estudar a mortalidade? Como resposta, tem-se que, a partir dela pode-se inferir sobre o estado de saúde e condições de vida, além da forma de morrer, o que, por fim, resultará em nível de sobrevivência de uma população. Demograficamente, este componente, que sendo, historicamente, o primeiro a ser estudado, foi o ponto de partida para a avaliação da dinâmica e transição demográfica.

No que diz respeito ao estudo da mortalidade infantil, há bastante tempo vem sendo trabalhada, devido à sua grande importância como indicador na avaliação das condições de vida e de saúde da população, sofrendo grande influência do status socioeconômico e meio envolvido. É, também, um elemento de forte contribuição nos ganhos gerais de sobrevivência de uma população, visto pela forma de esperança de vida ao nascer. Os melhores exemplos para os indicadores de mortalidade e sobrevivência são os coeficientes de mortalidade infantil e a esperança de vida ao nascer, ambos considerados os mais fiéis, por sua sensibilidade em expressar melhorias ocorridas nas condições de vida e saúde das comunidades, como apontam Simões (2002) e Silva (1991).

Nas últimas décadas, as taxas representativas da mortalidade de crianças menores de um ano, vêm apresentando um declínio significativo, devido à queda das incidências de doenças infantis, preveníveis a partir de melhores condições de alimentação, higiene e meios ambientais, presentes no período pós-neonatal, componente da mortalidade infantil tardia. Um fator que contraria as baixas taxas de mortalidade infantil, em algumas regiões do Brasil, ocorre quando as condições de vida estão associadas a uma maior ocorrência de óbitos pós-neonatais, em relação aos neonatais, possivelmente por serem esses primeiros os mais sensíveis às variações socioambientais (VERMELHO e MONTEIRO, 2002).

Apesar da diminuição significativa das taxas de mortalidade infantil, ainda existem regiões no Brasil (Norte e Nordeste), que continuam sofrendo por problemas associados aos níveis socioeconômicos, ou seja, setores menos favorecidos da sociedade brasileira, que se tivessem contado, no passado, com efetivos

investimentos por parte de seus governantes, poderiam, na atualidade, contornar tais problemas.

A relação dessas taxas com os fatores socioeconômicos já vem sendo trabalhada, no Brasil, por vários pesquisadores (Formiga et al (2000), Rios Neto e Moreira (1994), Fernandes (1984, p. 643), Formiga e Ramos (2002), Formiga (2003), Simões (1998), dentre outros) comprovando a estreita associação de condições socialmente favoráveis com redução da mortalidade infantil, assim como a correlação existente entre as variáveis socioeconômicas.

Para Carvalho e Sawyer (1978), dentre os indicadores de qualidade de vida de uma população, os índices de mortalidade, particularmente os da mortalidade infantil, são amplamente utilizados para estudar os efeitos das variações socioeconômicas sobre a população.

Dentro dessa discussão, Wood e Carvalho (1994), apontam para duas linhas de pensamento que se destacam na literatura. Uma, que coloca como causa fundamental da queda da mortalidade, as melhorias do padrão de vida da população (melhoria na nutrição, qualidade da água, higiene pessoal, limpeza, melhores condições habitacionais, cuidados na alimentação e outros), e a outra, “modelo da difusão”, que destaca os avanços médico-tecnológicos, importação de conhecimentos, procedimentos e equipamentos médicos dos países desenvolvidos como os principais responsáveis pelo declínio da mortalidade.

Com a melhoria ocorrida nas condições sociais da população brasileira e em países de terceiro mundo (países em desenvolvimento), o que se percebe, atualmente, é o consenso final de que os fatores que influenciaram na queda da mortalidade formam um conjunto de variáveis, não dando créditos apenas a um segmento. Como aponta Sawyer (1980), referindo-se à queda da mortalidade nos países desenvolvidos (ocorrida desde o século passado), os fatores responsáveis por ela poderiam ser agrupados em quatro grandes categorias, que interagem entre si: socioeconômicas, sanitárias, políticas e os avanços técnicos da medicina.

Segundo Formiga (2003), em estudos de mortalidade segundo fatores socioeconômicos, é necessário definir, inicialmente, os fatores a serem considerados. Pode-se observar que os mais comumente utilizados são a renda e a educação, mas outras variáveis disponíveis podem ser, também, utilizadas.

A educação materna, primordialmente, apresenta-se como ponto central nas análises de mortalidade infantil, devido ao fato de que uma mãe que tenha uma bagagem educacional, tenha, claramente, condições de fazer uso dos meios acessíveis à melhoria das condições da vida de uma criança, tais como acesso à assistência à saúde, condições alimentares e de higiene.

Estudo de Monteiro (1990, p. 74), conclui que é altamente significativo o efeito da educação materna sobre o risco de mortalidade infantil, independentemente do nível de renda. Assim como Monteiro, em estudo realizado para o estado do Ceará, Terra de Souza et al. (1999), chegou à conclusão de que o incremento na renda, quando não acompanhado de melhoria na educação materna, pode não reduzir substancialmente a mortalidade infantil.

Na utilização da variável renda, deve-se ter uma atenção maior, devido ao fato da colinearidade com as demais variáveis estudadas. Romero e Szwarcwald (2000) mostraram, baseados em dados para nove países latino-americanos, uma evidente associação entre crise econômica e mortalidade infantil, durante as últimas décadas. Os fatores econômicos, por sua vez, apresentam-se como elementos fundamentais, vistos pela forma de acesso a bens de consumo e qualidade domiciliar, revelando um padrão de bem estar e de sobrevivência da criança.

Analisando o fator assistência à saúde, pela variável *número de consultas pré-natal*, Hernandez (2011) conclui, em estudo realizado para a cidade de Porto Alegre (RS), uma relação inversa entre a tendência da taxa de mortalidade infantil e gestantes que realizam consultas pré-natal.

Segundo o relatório do IPEA (2010), o grande progresso na assistência à gestação, também serviu de homogeneizador, à medida que as mulheres que vivem em meio rural passaram a ter acesso a este tipo de assistência, isto revelado por dados da região Nordeste do Brasil. Em 1996, o número de crianças que tinham mães que não haviam realizado uma consulta pré-natal era de 36% e, atualmente, é inferior a 4%.

Avaliando, também, questões relacionadas a serviços de provisão de água e saneamento básico (infraestrutura), variáveis também relevantes, em estudos associados à mortalidade infantil, observa-se que ao longo da última década, não houve variação acentuada na porcentagem de crianças de zero a cinco anos sem acesso adequado a, pelo menos, um desses serviços básicos (aproximadamente

13%). A porcentagem de crianças vivendo em domicílios sem acesso adequado à água é atualmente dois terços da porcentagem dos que não têm acesso adequado a saneamento (IPEA, 2010).

Portanto, a partir do conhecimento da importância do estudo da mortalidade e suas principais variáveis preditoras e associações existentes, este trabalho trata de um estudo que pretende avaliar o comportamento da relação mortalidade infantil versus fatores socioeconômicos, nas microrregiões da região Nordeste. A escolha da região Nordeste foi devido à busca por desafios, dado que, na maioria dos estudos com esta temática, as escolhas têm sido direcionadas a regiões brasileiras que apresentam melhores condições econômicas e dados de excelente qualidade. A proposta deste estudo parte do princípio de que, a melhoria da qualidade de condições de vida apresentada pela população, não só nas regiões desenvolvidas do país, mas também no Nordeste brasileiro, trouxe consigo, uma melhoria substancial da qualidade das informações, podendo-se, então, com adequadas técnicas de correção, apresentar informações de confiança e resultados que possam refletir a realidade e que possam, até, aproximar-se das estimativas oficiais do IBGE.

Assim posto, serão desenvolvidos agrupamentos das microrregiões segundo suas variáveis estudadas, tentando mostrar perfis de regiões e suas configurações perante as variáveis socioeconômicas, educacionais maternas, demográficas, assistência à saúde e de infraestrutura.

1.1. Definição do problema

Como ponto de partida, o presente estudo propõe-se a responder às questões:

- Quais os perfis de mortalidade infantil, segundo variáveis demográficas, socioeconômicas, de infraestrutura e de assistência à saúde das microrregiões do Nordeste brasileiro?
- A partir do conhecimento dos níveis de mortalidade por microrregião, será possível encontrar configurações ou grupos homogêneos reveladores da estrutura real da região?

O presente estudo propõe-se a discutir sobre esses questionamentos, assim como encontrar soluções cientificamente plausíveis, procurando, então, trazer sua

contribuição para um melhor entendimento da mortalidade infantil na região, gerando subsídios para políticas públicas.

1.2. Justificativa da pesquisa

Desde muito tempo, estudar a relação entre mortalidade infantil e condições de vida de uma população, toma espaço notório nos segmentos de pesquisas sobre mortalidade e morbidade.

Conhecer o comportamento da mortalidade infantil torna-se de grande importância, para que se tenha o controle do evento e para a formulação de estratégias e medidas que possam trazer benefícios para a população, que ainda padece de acesso à saúde. Este conhecimento trás, como contribuição maior, a identificação da área e/ou grupo populacional mais desassistido, podendo possibilitar a implantação de medidas de serviços, a médio e longo prazo, estimulando a continuidade deste trabalho. Desta forma, os índices e tipologias encontradas neste trabalho poderão servir de subsídios mais eficientes para as políticas públicas voltadas a população infantil.

Somado ao conhecimento do evento mortalidade, Bird e Bauman (1998), apontam que o nível de mortalidade resulta da interação de três conjuntos de fatores que afetam o bem-estar da população: serviços que podem levar à melhoria do nível de saúde, como, por exemplo, a disponibilidade da água potável; serviços públicos de saúde, que influenciam a mortalidade, independente de decisões individuais e, por fim, uma série de características diretamente ligadas ao indivíduo, tais como a renda, a qual afeta a saúde através da nutrição, a moradia e a educação, associadas à rapidez e eficiência com as quais os indivíduos respondem aos serviços de saúde e às ameaças ambientais. Desta forma, o entendimento destes fatores complementa a efetividade do trabalho a ser apresentado.

Atrelada à busca por estabelecer a associação da mortalidade aos fatores demográficos e socioeconômicos, somados, também, à dificuldade de se trabalhar a mortalidade infantil no Brasil por unidades de análise mais reduzidas, principalmente na região Nordeste, devido a questões relacionadas à qualidade das informações, o presente estudo tem como propósito, ultrapassar essas dificuldades, assim como fornecer subsídios para as autoridades competentes, para a melhoria da qualidade das informações. Ao mesmo tempo, serão apresentados resultados que revelam a

estrutura real das microrregiões da região Nordeste e suas interações, segundo as condições encontradas em cada uma delas. Por fim, estimular a criação de programas e projetos que possam dar melhor visibilidade a essas pequenas regiões que padecem de assistência, não só em saúde, mas também de infraestrutura, educação e demais variáveis capazes de impactar na mortalidade infantil.

1.3. Objetivo do estudo

A partir da formulação do problema da pesquisa e da sua justificativa, são definidos os objetivos que serão norteadores das análises.

1.3.1. Objetivo geral

Identificar os perfis de mortalidade de crianças menores de um ano de idade das microrregiões do Nordeste brasileiro, segundo características socioeconômicas, demográficas, de infraestrutura e de assistência à saúde, referentes ao ano 2010.

1.3.2. Objetivos específicos

1. Investigar a qualidade das informações, a partir das Causas Mal Definidas (CMD) de mortalidade, por UF, no período de 2000 a 2010;
2. Estimar e corrigir os sub-registros da mortalidade infantil das microrregiões do Nordeste, para os anos de 2000 e 2010;
3. Avaliar os níveis de mortalidade infantil nas Microrregiões da região Nordeste do Brasil, nos anos de 2000 e 2010;
4. Identificar, dentre as variáveis selecionadas para o estudo, quais apresentam maior significância explicativa da mortalidade infantil, para o ano de 2010;
5. Formar grupos de microrregiões com características semelhantes, segundo suas respectivas taxas de mortalidade infantil e demais variáveis selecionadas;
6. Elaborar perfis/tipologias de mortalidade infantil (taxas de mortalidade de menores de um ano), a partir das variáveis selecionadas, contribuindo para a possível identificação da relação entre as características da região e os níveis de mortalidade infantil.

1.4. Brasil e Região Nordeste: comportamento frente à mortalidade infantil e dinâmica demográfica

No Brasil, a mortalidade infantil tem apresentado importante queda nas últimas décadas, reflexo das melhorias apresentadas em todos os indicadores sociodemográficos brasileiros em, praticamente, todos os seus estados. Assim, de Norte a Sul do país, as taxas de mortalidade infantil dos últimos 40 anos caíram de forma generalizada, conforme mostram diversos estudos (Simões, 1998, Barros et. al. 2010).

Segundo o IBGE, o Brasil, em 1980, apresentava uma taxa de mortalidade infantil de 82,8 / 1000 n.v. e, seguindo esse comportamento, exibindo uma taxa superior às outras regiões (117.6 / 1000 n.v.), a região Nordeste se destacou.

O IBGE (1999) descreve que:

“A queda consistente da mortalidade que se observa, a partir da década de 70, parece estar fortemente dependente do modelo de intervenção na área das políticas públicas, então adotado principalmente no campo da medicina preventiva, curativa, de saneamento básico e, mais recentemente, na ampliação dos programas de saúde materno-infantil, sobretudo os voltados para o pré-natal, parto e puerpério; a ampliação da oferta de serviços médico-hospitalares em áreas do País, até então bastante carentes, as campanhas de vacinação, os programas de aleitamento materno e reidratação oral. Agreguem-se a estes fatores as grandes mudanças nos padrões reprodutivos, com quedas acentuadas nos níveis de fecundidade e teremos um quadro explicativo da evolução da queda da mortalidade, principalmente, a partir dos anos 80, até o presente momento.”

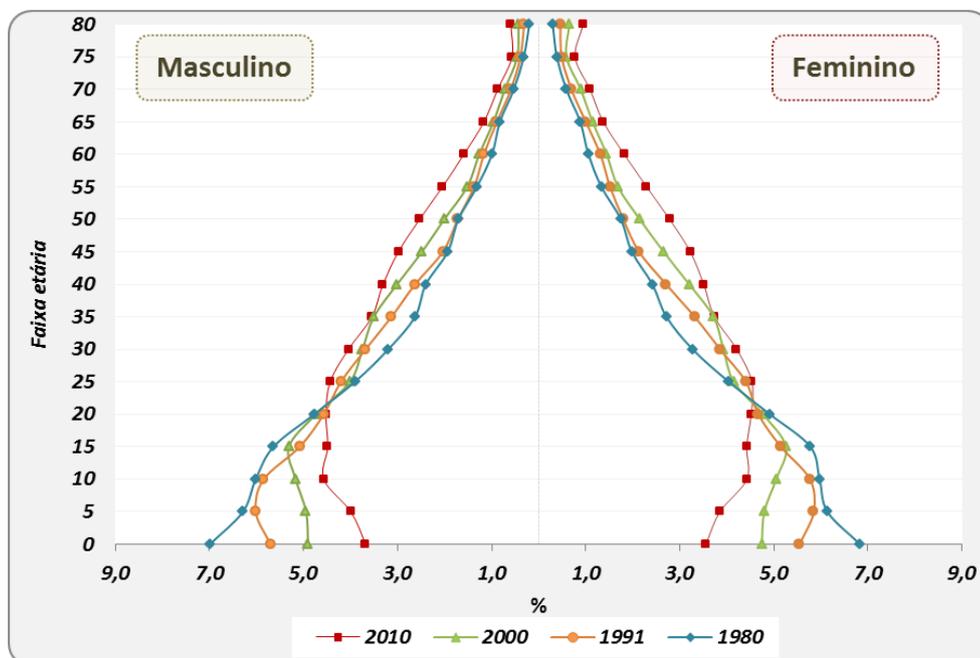
Como ressaltado, a queda na taxa de fecundidade no Brasil ocorreu de forma acelerada, não só nas classes mais privilegiadas, mas também nas classes de menor poder aquisitivo. Formiga (2003), em estudo realizado para o Estado do Rio Grande do Norte, aponta que:

“Muito embora não se tenha conhecimento de estudos específicos para o Estado, mostrando diretamente a relação entre a queda da fecundidade e o declínio da mortalidade na infância, o comportamento decrescente de ambos os indicadores permitem inferir que ela vem se processando ao longo do tempo.”

Desta forma, o declínio da fecundidade tem impactos e implicações em diversas fases da pirâmide etária, seja no Brasil ou em qualquer outro país.

Como se pode observar no gráfico 01, a evolução das estruturas etárias para o Brasil (1980, 1991, 2000 e 2010), a redução da fecundidade foi ponto fundamental nas mudanças do comportamento dessas estruturas, evidenciado pela redução da base (população de 0 a 5 anos), com o passar dos anos e um conseqüente envelhecimento populacional. Certamente, a queda da fecundidade tem seu impacto na redução das taxas de mortalidade infantil. À medida que nascem menos crianças, a quantidade de expostos ao evento mortalidade é reduzida, atrelado também, à melhoria das condições do País.

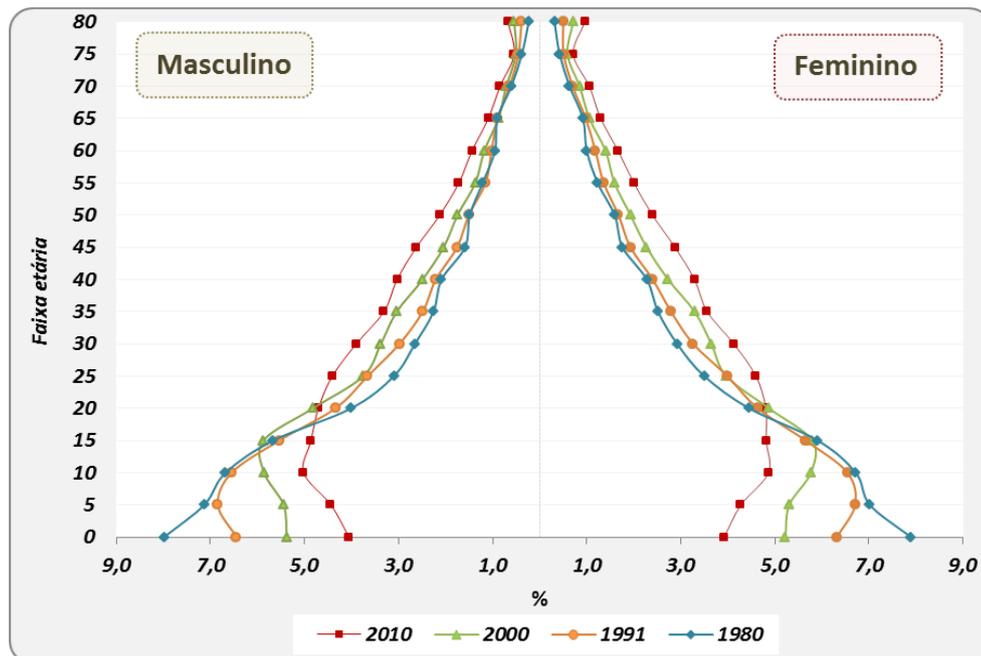
Gráfico 01 – Comportamento da estrutura etária demonstrada nas pirâmides etária Brasil, 1980, 1991, 2000 e 2010.



Fonte: Elaborado com base nos dados disponíveis no DATASUS - Ministério da Saúde.

No gráfico 02, é apresentada a evolução das estruturas etárias para a região Nordeste. Percebe-se a tendência de comportamento semelhante à estrutura apresentada para o Brasil. Porém, do ano de 1980 ao ano 2000, a estrutura da base apresentava-se com proporções superiores às do Brasil, refletindo as grandes taxas de fecundidade apresentadas pela região Nordeste. Entretanto para o ano de 2010, o comportamento mostrou-se bem semelhante aos padrões gerais mostrados nas pirâmides para o Brasil.

Gráfico 02 – Comportamento da estrutura etária demonstrada nas pirâmides etária da região Nordeste, 1980, 1991, 2000 e 2010.



Fonte: Elaborado com base nos dados disponíveis no DATASUS - Ministério da Saúde.

Ferreira (1990), analisou o efeito da fecundidade sobre a queda da mortalidade infantil em São Paulo. Essa influência, mais incidente sobre a mortalidade infantil, já foi observada em vários países em épocas distintas, podendo-se citar os trabalhos desenvolvidos por Woods (1988, 1989), para a Inglaterra e Gales, entre 1861 e 1921.

Talvez a grande influência da queda da fecundidade, na queda de mortalidade infantil, esteja associada à redução do número de mulheres com alta parturição, ampliação nos intervalos de nascimentos, mudanças na idade da concepção pelo parto, favorecendo a saúde da mulher e à sobrevivência dos filhos (Beltrão e Sawyer, 1990).

Associado ao fator fecundidade, as grandes transformações percebidas na sociedade brasileira nas últimas décadas, em segmentos políticos, sociais e econômicos, tornaram possíveis uma grande mobilidade populacional para os centros urbanos em busca de melhor qualidade de vida e, conseqüentemente, mais acesso à saúde e condições favoráveis de vida. Acompanhando estas transformações, os dados de óbitos infantis apresentaram melhorias qualitativas e quantitativas. Estas últimas, traduzidas pelo sub-registro de óbitos infantis e

nascidos vivos, continuam presentes e traduzem sérios problemas nos estudos analíticos, conforme mostrado na seção seguinte.

1.5. Sub-registro de óbitos infantis e nascidos vivos

Avaliar o comportamento da mortalidade de uma população, de forma geral, é uma tarefa que requer grande atenção por parte dos gestores. Tal atividade possibilita um melhor entendimento do comportamento do modo de viver e morrer de um conjunto de indivíduos, como já comentado.

Segundo Simões (2002), estudos têm sido desenvolvidos, mostrando que países onde as desigualdades sociais são profundas, como ainda é a situação do Brasil, apresentam, da mesma forma, desigualdades frente à morte, por parte dos distintos estratos sociais. Apesar da diminuição significativa dessas taxas, ainda existem regiões no Brasil (Norte e Nordeste), que continuam padecendo de problemas associados aos seus níveis de desenvolvimento econômico e social, ou seja, setores socialmente menos favorecidos.

Associados às questões de baixo desenvolvimento, encontram-se, também, problemas ligados às notificações de informações. Segundo Paes (2007), o Brasil possui duas fontes oficiais de acesso público, responsáveis pela produção contínua das estatísticas de óbitos. A Fundação IBGE, que publica seus resultados nas “Estatísticas do Registro Civil” e o Ministério da Saúde, que disponibiliza seus dados pelo Sistema de Informação sobre Mortalidade (SIM) em “Estatísticas de Mortalidade”, em seus respectivos sítios institucionais. Há discrepâncias nos resultados dessas duas fontes, mas somente o Ministério da Saúde processa e divulga os dados classificados por causas de morte.

Também, um dos grandes problemas encontrados, que ainda limitam as análises de mortalidade, é o sub-registro de óbitos, com magnitudes expressivas observadas nas Regiões Norte e Nordeste do Brasil e, especificamente, mais forte na faixa de idade de crianças menores de um ano (Szwarcwald et al, 2002).

A partir da avaliação de Mello-Jorge (1983), o sub-registro de óbitos diz respeito, principalmente, à ocorrência de sepultamentos sem a exigência da certidão de óbito, nos denominados cemitérios clandestinos, prevalentemente na área rural e associados à pobreza.

Sabe-se que a dificuldade em estimar a mortalidade infantil é proporcionalmente maior quanto mais problemas socioeconômicos apresenta a Região. Consciente destas limitações de informações, o IBGE, por ser o órgão responsável por elaborar as estimativas de mortalidade infantil no Brasil, vem recorrendo a cálculos ou técnicas indiretas, substituindo os cálculos diretos, de modo a contornar o problema do sub-registro em algumas regiões do país (IBGE, 2001).

Segundo o Manual X (Nações Unidas, 1986), o termo “indireta”, usado para classificar algumas das técnicas de estimação demográfica, originou-se do fato de que tais técnicas produzem estimativas de um certo parâmetro, baseando-se em informações que só indiretamente estão relacionadas com o mesmo.

É com base em novas abordagens, para contornar problemas do sub-registro, que o presente estudo usa como ferramenta de correção (suavização), a técnica de estimadores bayesianos de James-Stein, proposta por Cavalini e Leon (2007), utilizada para áreas geográficas com grandes flutuações aleatórias devido aos pequenos números das áreas menores.

2. DADOS, MÉTODOS E OPERACIONALIZAÇÃO

As fontes das bases de dados utilizadas no estudo foram: SIM, SINASC, SIDRA/IBGE, Microdados do Censo 2010 (IBGE).

Do Sistema de Informações sobre Mortalidade (SIM) e do Sistema de Informações sobre os Nascidos Vivos (SINASC), ambos hospedados no DATASUS/MS, foram retiradas as informações, por microrregião, de óbitos infantis (menores de um ano) e nascidos vivos, dos anos de 2000 e 2010, para o cálculo das taxas de mortalidade infantil e correção dos óbitos e nascidos vivos. Para a avaliação das proporções de Causas Mal Definidas, foram tomadas as informações de óbitos por causa, segundo o capítulo do CID-10, por UF e região do Brasil, sendo estas dispostas numa série histórica dos anos de 2000 a 2010.

Segundo a divisão territorial proposta pelo IBGE, a região Nordeste apresenta nove Unidades de Federação (Maranhão, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Piauí, Bahia, Sergipe), divididas em 188 microrregiões, que compreendem 1794 municípios. Todas as análises envolveram as microrregiões como unidade de análise.

O quadro 01 apresenta a descrição das variáveis trabalhadas no estudo e seus respectivos códigos e fontes.

Quadro 01 – Relação dos códigos e descrição das variáveis utilizadas nas análises, informações do ano de 2010, Microrregiões da região Nordeste.

1. Variáveis Socioeconômicas		
Cód.	Descrição	Fonte
Rend_Pc	<i>Renda Média domiciliar per capita</i>	DATASUS/TABNET
Prop_Bx_rend	<i>Proporção de pessoas com baixa renda (< ½ S.M)</i>	DATASUS/TABNET
2. Variáveis Demográficas		
Cód.	Descrição	Fonte
Tx_Md_Analf_M	<i>Taxa de analfabetismo feminino</i>	DATASUS/TABNET
Id_Md_M	<i>Idade média das mães quando notificado o óbito infantil</i>	DATASUS/SIM
A_Md_Est	<i>Anos médio de estudo materno quando notificado o óbito infantil</i>	DATASUS/SIM
Dens_Domic	<i>Densidade domiciliar</i>	DATASUS/TABNET
3. Variável de Assistência à Saúde		
Cód.	Descrição	Fonte
Num_Md_Cont_PN	<i>Número médio de consultas Pré-Natal</i>	DATASUS/SINASC
4. Variáveis de Infraestrutura		
Cód.	Descrição	Fonte
Gr_Urb	<i>Grau de Urbanização</i>	DATASUS/TABNET
Cobert_Agua	<i>Proporção de domicílios com abastecimento de água adequado</i>	Microdados IBGE/2010
Exist_C_Lixo	<i>Proporção de domicílio com existência de coleta de lixo adequada</i>	Microdados IBGE/2010
Cobert_Esg_Sanit	<i>Proporção de domicílios com existência de cobertura de esgotamento sanitário</i>	Microdados IBGE/2010

Fonte: Elaboração própria.

As variáveis Taxa de Analfabetismo feminino (Tx_Md_Analf_M), Renda média per capita (Rend_Pc), Proporção de pessoas com baixa renda (Prop_Bx_rend), Idade média das mães quando notificado o óbito infantil (Id_Md_M), Anos médio de estudo quando notificado o óbito (A_Md_Est), Número médio de consultas pré-natal

(Num_Md_Cont_PN) e Grau de Urbanização (Gr_Urb), foram retiradas da base de dados disponível no DATASUS/SIM/SINASC (informações demográficas e socioeconômicas). Já as variáveis, Proporção de domicílios com abastecimento de água adequado (Cobert_Agua), Proporção de domicílios com existência de coleta de lixo adequada (Exist_C_Lixo) e Proporção de domicílios com existência de cobertura de esgotamento sanitário (Cobert_Esg_Sanit), foram gerados a partir dos microdados do Censo 2010 do IBGE.

Algumas variáveis tiveram que passar por um processo de tratamento (médias, redistribuição proporcional de informações ignoradas – método pró-rata, cálculos de proporções e percentuais), para se chegar ao formato ideal para o uso, assim definido:

Médias: este tratamento foi desenvolvido com o intuito de apresentar as médias por microrregião, visto que os dados, quando captados eram apresentados em formato de intervalos de classes. As variáveis *Id_Md_M*, *A_Md_Est* e *Num_Md_Cont_PN*, passaram por este tratamento;

Tratamento de redistribuição proporcional de informações ignoradas: a intenção principal foi para que não exista perda de informações, principalmente nas estatísticas de óbitos infantis e nascidos vivos, dessa forma possibilitando fazer as correções indicadas para a taxa de mortalidade infantil por microrregião. O método utilizado foi o *pró-rata*, método que identifica o peso das informações ignoradas na UF e as redistribui entre as microrregiões envolvidas no estudo.

Cálculo de proporção: tratamento utilizado quando a variável que seria utilizada no estudo não era encontrada no formato adequado, necessitando do cálculo de proporção ou taxa, de acordo com as definições originais da variável. As variáveis *Gr_Urb* e *Dens_Domic*, passaram por este tratamento;

Foram tomadas algumas definições e conceitos segundo as notas técnicas utilizadas pela Rede Interagencial de Informações para a Saúde – RIPSA, para se encontrar as proporções referentes à cobertura (abastecimento da água, coleta de lixo e cobertura de esgotamento sanitário):

Cálculo de proporção – Abastecimento da água: define-se como cobertura de abastecimento d'água, o percentual da população residente servida por rede geral de abastecimento, com ou sem canalização domiciliar, em determinado espaço geográfico, no ano considerado. Portanto, dos itens que compõem esta variável no

censo populacional, apenas fizeram parte do cálculo que caracteriza esta variável, foram: (i) com canalização interna, quando o domicílio possui canalização em pelo menos um cômodo e a água utilizada for proveniente de rede geral de distribuição; (ii) sem canalização, quando a água utilizada no domicílio for proveniente de rede geral de distribuição para o terreno ou propriedade onde está localizado o domicílio.

Cálculo de proporção – Coleta de lixo: define-se como cobertura de lixo, o percentual da população residente atendida, direta ou indiretamente, por serviço regular de coleta de lixo domiciliar, em determinado espaço geográfico, no ano considerado. Portanto, dos itens que compõem esta variável no censo populacional, apenas fizeram parte do cálculo que caracteriza esta variável, foram: (i) direto, quando a coleta é realizada no domicílio, por empresa de limpeza urbana; (ii) indireto, quando o lixo é depositado em caçamba, tanque ou outro depósito, sendo posteriormente coletado por serviço ou empresa de limpeza urbana.

Cálculo de proporção – Cobertura de Esgotamento sanitário: define-se como cobertura de esgotamento sanitário, o percentual da população residente que dispõe de escoadouro de dejetos através de ligação do domicílio à rede coletora ou fossa séptica, em determinado espaço geográfico, no ano considerado. Portanto, dos itens que compõem esta variável no censo populacional, apenas fizeram parte do cálculo que caracteriza esta variável, foram: (i) rede coletora de esgoto ou pluvial; (ii) fossa séptica ligada à rede coletora de esgoto ou pluvial; (iii) fossa séptica não ligada à rede coletora de esgoto ou pluvial.

Os *softwares* utilizados no estudo foram:

- **Microsoft Excel 2010** – Para a organização das bases de dados, elaboração de quadros, tabelas e confecção dos gráficos;
- **Microsoft Word 2010** – Para edição e elaboração do trabalho;
- **SPSS 19** – Para trabalhar os microdados do Censo 2010, modelagem da regressão linear múltipla ponderada;
- **TerraView 4.2.0** – Para a elaboração dos mapas temáticos;
- **GoM 3.4** – Para desenvolver as tipologias/perfis das microrregiões segundo a Taxa de mortalidade infantil e demais variáveis;
- **STATISTICA 7.0** – Desenvolvimento de estatísticas descritivas, gráficos box-plots e análise de *Clusters*.

2.1. Análise de *Clusters*

A Análise de Agrupamentos, também conhecida como Análise de Conglomerados, Classificação ou *Cluster*, tem como objetivo dividir os elementos de uma amostra ou população, em grupos, de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si, com respeito às variáveis (características) que neles foram medidas e os elementos em grupos diferentes sejam heterogêneos em relação a estas mesmas características.

2.1.1. Medida de Similaridade e Dissimilaridade

- **Distância Euclidiana**

Considerando um conjunto de dados constituídos de n elementos amostrais, tendo-se medido p -variáveis aleatórias em cada um deles, o objetivo é agrupar esses elementos em g grupos. Para cada elemento amostral j , tem-se, portanto, o vetor de medidas X_j definido por:

$$X_j = [X_{1j} \ X_{2j} \ \dots \ X_{pj}]', \quad j = 1, 2, \dots, n,$$

Onde X_{ij} representa o valor observado da variável i medida no elemento j . Para se proceder ao agrupamento de elementos, é necessário que se decida, *a priori*, a medida de similaridade ou dissimilaridade que será utilizada. Existem várias medidas diferentes que podem ser adotadas e cada uma delas produz um determinado tipo de agrupamento. Uma das medidas mais comuns, apropriada para variáveis quantitativas, é a distância euclidiana, apresentada a seguir. Para essa medida de dissimilaridade, quanto menor o seu valor, mais similares serão os elementos que estão sendo comparados.

A distância Euclidiana entre dois elementos X_j e X_k , $j \neq k$, é definida por:

$$d(X_j, X_k) = [(X_j - X_k)'(X_j - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{ij} - X_{ik})^2 \right]^{1/2}$$

Ou seja, os dois elementos amostrais são comparados em cada variável pertencente ao vetor de observações.

2.1.2. Técnica para construção de conglomerados (*Clusters*)

As técnicas de conglomerados ou *cluster* são frequentemente classificadas em dois tipos: técnicas hierárquicas e não hierárquicas, sendo que as hierárquicas são classificadas em aglomerativas e divisivas. As técnicas hierárquicas, na maioria das vezes, são utilizadas em análises exploratórias dos dados com o intuito de identificar possíveis agrupamentos e o valor provável do número de grupos g . Já para o uso de técnicas não hierárquicas, é necessário que o valor do número de grupos já esteja pré-especificado pelo pesquisador. Neste estudo, tomou-se como técnica a metodologia hierárquica, com o sistema de ligação simples (*Single Linkage*).

2.1.3. Técnica Hierárquica Aglomerativa

- **Método de Ligação Simples (*Single Linkage*)**

As técnicas Hierárquicas Aglomerativas partem do princípio de que, no início do processo de agrupamento, tem-se n conglomerados, ou seja, cada elemento do conjunto de dados observados é considerado como sendo um conglomerado isolado. Em cada passo do algoritmo, os elementos amostrais vão sendo agrupados, formando novos conglomerados, até o momento no qual todos os elementos considerados estão num único grupo. Portanto, no estágio inicial do processo de agrupamento, cada elemento amostral é considerado um *cluster* de tamanho 1 e no último estágio de agrupamento tem-se apenas um único *cluster* constituído de todos os n elementos amostrais. Em termos de variabilidade, no estágio inicial tem-se a partição com menor dispersão interna possível, já que todos os conglomerados têm um único elemento e, em consequência, a variância de cada um deles é igual a zero. No estágio final, tem-se a maior dispersão interna possível, já que todos os elementos amostrais estão num único *cluster*. Em cada estágio do procedimento de agrupamento, os grupos são comparados através de alguma medida de similaridade (ou dissimilaridade), previamente definida. Neste trabalho, utilizou-se a Distância Euclidiana (julgada a mais apropriada), devido à propriedade hierárquica, permitindo a construção de um gráfico, chamado de Dendograma, que representa a “árvore” ou a história do agrupamento. Para tanto, no eixo horizontal são marcados os elementos amostrais, numa ordem conveniente, relacionadas à história de agrupamento. As linhas verticais, partindo dos elementos amostrais agrupados, têm

altura correspondente ao nível em que os elementos foram considerados semelhantes, isto é, a distância do agrupamento ou nível de similaridade.

O método hierárquico utilizado foi o de Ligação simples ou vizinho mais próximo. Para melhor entendimento considere o seguinte exemplo:

Considere 4 objetos (A, B, C, D), para os quais a matriz D de distâncias entre os objetos é apresentada a seguir.

$$D = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} 0 & & & \\ 3 & 0 & & \\ 7 & 9 & 0 & \\ 8 & 6 & 5 & 0 \end{bmatrix} \end{matrix}$$

Utilizando o método da ligação simples, os objetos menos distantes devem, inicialmente, ser fundidos. Então, $\min(d_{h,1}) = d_{A,B} = 3$. O próximo passo é fundir A com B formando o grupo (AB) e em seguida calcular as distâncias deste grupo e os objetos remanescentes. As distâncias dos vizinhos mais próximos são:

$$d_{(AB),C} = \min\{d_{AC}, d_{BC}\} = \min\{7, 9\} = 7$$

$$d_{(AB),D} = \min\{d_{AD}, d_{BD}\} = \min\{8, 6\} = 6$$

A nova matriz D para o próximo passo é:

$$D = \begin{matrix} & \begin{matrix} AB & C & D \end{matrix} \\ \begin{matrix} AB \\ C \\ D \end{matrix} & \begin{bmatrix} 0 & & \\ 7 & 0 & \\ 6 & 5 & 0 \end{bmatrix} \end{matrix}$$

A menor distância é entre D e C, com $d_{DC} = 5$, os quais foram fundidos formando o subgrupo DC, no nível 5. Recalculando as distâncias tem-se,

$$d_{(DC),(AB)} = \min\{d_{D(AB)}, d_{C(AB)}\} = \min\{6, 7\} = 6$$

A nova matriz D fica,

$$D = \begin{matrix} & DC & AB \\ \begin{matrix} DC \\ AB \end{matrix} & \begin{bmatrix} 0 & \\ 6 & 0 \end{bmatrix} \end{matrix}$$

Então, a partir da matriz final, pode-se observar os grupos formados: **Grupo 1:** Objetos (A e B); **Grupo 2:** Objetos (D e C).

2.1.4. Método para encontrar o Número g de *Clusters* da Partição Final

Uma questão de grande importância é como se deve proceder para escolher o número final de grupos que define a partição do conjunto de dados ou, de outra forma, em qual passo o algoritmo de agrupamentos deve ser interrompido.

Um critério conveniente utilizado é o da análise do comportamento do chamado nível de fusão (distância). À medida que se avança no algoritmo de agrupamentos, ou seja, passa-se do estágio k para o estágio $k+1$, a similaridade entre os conglomerados que estão sendo combinados nos respectivos passos vai decrescendo. Conseqüentemente, a distância entre eles vai aumentando. Desse modo, se for feito um gráfico de passos (ou número de grupos) versus o nível de distância (nível de fusão) do agrupamento de cada estágio do processo, pode-se visualizar se há “pontos de salto” relativamente grandes, em relação aos demais valores de distância. Estes pontos indicam os momentos ideais de parada do algoritmo, isto é, o número de conglomerados final e a composição final dos grupos. Logo, se a função apresentar vários “pontos de salto”, pode-se definir uma região de prováveis valores do número de grupos que deveriam ser melhor investigados por algum outro procedimento. Quando n não é muito grande, o dendograma também pode ser utilizado como uma ferramenta para visualização dos “pontos de salto” das distâncias (perda acentuada de similaridade) e dos grupos formados.

Uma análise mais detalhada do método pode ser encontrada em Mingoti (2005).

2.2. Estimador Bayesiano Empírico

Para a estimação dos óbitos gerais por microrregião, empregou-se a técnica de estimadores bayesianos de James-Stein, proposta por Cavalini e Leon (2007). Este modelo é indicado quando existe o envolvimento de áreas de diferentes

magnitudes, por exemplo, apresentando grandes flutuações de tamanhos populacionais, característicos de pequenas áreas (Freire, 2001; Freire e Assunção 2002). Neste trabalho, as microrregiões trabalhadas por UF (totalizando cento e oitenta e oito), foram tomadas como unidade de análise, as quais, por sua vez foram aglutinadas, formando as chamadas mesorregiões.

Os contingentes populacionais envolvidos nos cálculos das taxas (nascidos vivos das microrregiões e óbitos de menores de um ano, por exemplo), estão sujeitos às chamadas oscilações dos pequenos números, pelo que se fez opção pela aplicação de uma metodologia capaz de contornar tais problemas. Dessa forma, os dados foram submetidos ao processo de suavização, propiciada pelo emprego do modelo bayesiano empírico de James-Stein (Efron e Morris, 1975), recomendado por diversos autores (Assunção et al 1998; Freire e Assunção, 2002; Cavalini, 2005), como forma de reduzir as referidas flutuações. Como resultado da aplicação do método, obtém-se uma média ponderada (taxa de mortalidade infantil ajustada – TMIA) entre a taxa bruta da localidade (área menor, no caso, microrregião) e a taxa da área maior (região de vizinhança, no caso, mesorregião) tomada como referência.

As estimativas são operacionalizadas através da seguinte expressão:

$$\theta_i = T_m + C_i \cdot (t_i - T_m)$$

Para,

$$C_i = \frac{s^2 - \left(\frac{T_m}{P_m}\right)}{\left(s^2 - \left(\frac{T_m}{P_m}\right)\right) + \left(\frac{T_m}{P_i}\right)}$$

Onde:

θ_i - é a taxa suavizada;

T_m - é a taxa média global ou a taxa média dos vizinhos;

t_i - é a taxa da área i ;

s^2 - é a variância da taxa a ser medida;

P_m - é a população média global ou a população média dos vizinhos;

P_i - é a população da área i .

Pode-se observar, na expressão acima, que o multiplicador C_i será próximo de 1 se a população (nascidos vivos) P_i for grande, o que fará a taxa suavizada (θ_i) tender para o mesmo valor da taxa (t_i) calculada inicialmente. Por outro lado, se a área em tela possuir uma população (nascidos vivos) muito pequena, então C_i será também muito pequeno (próximo de zero), fazendo com que a taxa suavizada da pequena área aproxime-se da taxa média de referência. Essa taxa de referência deve ser selecionada em função dos dados observados, com o critério de vizinhança podendo ser o de uma região maior ou da média de áreas vizinhas.

Do exposto, fica claro que a taxa de mortalidade ajustada poderá ser maior ou menor do que a taxa da grande área, tomada como região de referência, podendo ser menor do que a taxa de mortalidade observada para a microrregião, o que implicaria numa supressão de óbitos informados pelo SIM. Nesses casos, foi adotado o mesmo ajuste utilizado por Cavalini e Leon (2007) e Cavalini (2005), ou seja, desprezar essa taxa ajustada e adotar a taxa original da microrregião. Conforme observado pelos autores, como resultado final deste ajuste, o número total de óbitos da UF tende a aumentar, pois somente serão permitidos acréscimos, e nunca reduções, ao número de óbitos em cada microrregião. De posse dos óbitos estimados ou finais, será possível estimar o sub-registro de óbitos para cada microrregião, tomado como a diferença entre a unidade e a chamada cobertura dos óbitos. Este, por sua vez, é obtido da razão entre os óbitos observados e estimados (ou finais), de acordo com a seguinte expressão:

$$\textit{Sub - registro} = 1 - \left(\frac{\textit{Óbitos observados}}{\textit{Óbitos estimados}} \right)$$

Pode-se encontrar, também, o fator de correção aplicado a cada microrregião, como mostra a expressão a seguir:

$$\textit{Fator de correção} = \frac{1}{\textit{Grau de cobertura}}$$

Vale salientar que vão existir microrregiões onde a cobertura terá representatividade de 100%, ou seja, o sub-registro será nulo. Isso ocorrerá quando o número de óbitos estimados for inferior ao número de óbitos observados, dado que, quando ocorre isto na estimativa, é adotado o valor de óbitos observado, fazendo com que não seja possível fazer reduções nos números e sim acréscimos.

2.3. Modelo de Regressão Linear Múltipla Ponderada (MRLMP)

A proposta da utilização do método vem do princípio de redução do número de variáveis envolvidas na formação dos grupos homogêneos (*clusters*), tentando minimizar o efeito de muitas variáveis que venham a causar confundimento, devido à grande variação de informações das variáveis. Sendo assim, a regressão linear múltipla ponderada foi implementada, de forma a encontrar as variáveis mais significantes na explicação da variável “taxa de mortalidade infantil 2010” (variável dependente - Y). Por se tratar de uma taxa, torna-se necessário, para satisfazer às restrições do modelo de regressão, que se utilize, como variável resposta, o “logito da taxa de mortalidade infantil 2010”. Foram tomadas como variáveis independentes (explicativas - X_i), todas as variáveis apresentadas no quadro das descrições e códigos das variáveis (Quadro 01).

Esta metodologia, portanto, implicará, para este estudo, em conhecer a importância relativa das variáveis independentes envolvidas ou de suas interações, a magnitude dos coeficientes de regressão (β_i), bem como do próprio coeficiente de determinação (R^2), para o qual não necessita ser dada muita importância, devido ao fato de que não serão feitas previsões ou estimativas, após encontrar o modelo ou variáveis mais explicativas (Hall e Rao, 1992). O que definirá a variável independente como significativa será o p-valor do teste *t de Student* parcial para os coeficientes de regressão β_i da variável em questão, tomado a um nível de significância de 5%.

A expressão geral do modelo de regressão linear múltipla, considerando as variáveis explanatórias, é dada por:

$$\text{Logit}(Y_i) = \log\left(\frac{Y_i}{1 - Y_i}\right) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i$$

O modelo assume que a variável resposta $\text{logit}(Y_i)$ deve ser contínua e estar, provavelmente, correlacionada com p variáveis explicativas X_i . Apenas os resíduos ε_i são considerados aleatórios, com $\beta_0, \beta_1, \dots, \beta_p$ constantes (coeficientes) desconhecidas (parâmetros do modelo). Os ε_i 's são aleatórios, para os quais, supõe-se distribuição Normal, com média zero e variância constante σ^2 e independentes. Estes erros representam a parte da formação da variável resposta $\text{logit}(Y)$ que não é explicada pelo modelo de regressão ajustado aos dados. As variáveis explicativas X_{ij} são controladas e medidas com erro negligenciável (ou desprezível), portanto, os valores de X_{ij} são pré-fixados à priori.

A estimação dos parâmetros β_i , $i = 1, 2, \dots, p$ é feita através da minimização da soma dos quadrados dos erros (resíduos), ou seja, pelo chamado método de mínimos quadrados considerando-se n observação da variável resposta.

Também, como forma de minimizar as distintas variabilidades existentes entre as variáveis envolvidas, fez-se a opção por incluir uma variável que pudesse fazer o papel ponderador, neste caso, usou-se o número de nascidos vivos em cada microrregião (Halli e Rao, 1992).

A seleção das variáveis regressoras mais importantes, dado uma variável Y e um conjunto de variáveis explicativas regressoras (X_1, X_2, \dots, X_k), foi realizada neste estudo pela técnica "*Stepwise Regression*", com o processo de seleção "*Forward Selection*".

Assim procedendo, adotou-se um nível de significância α a ser empregado nos testes de entrada e saída de variáveis do modelo. Em cada passo desenvolvido, serão obtidos dois valores da distribuição F com $(1, n-p-1)$ graus de liberdade, onde p é o número de parâmetros do modelo de regressão completo ajustado, em cada estágio do processo de seleção. Estes valores designados por F_E e F_S , serão utilizados como critérios para entrada e saída de variáveis do modelo, respectivamente. Na sequência, ajusta-se um modelo de regressão linear simples para cada uma das variáveis regressoras candidatas, para cada um dos modelos de regressão linear simples ajustado (encontrando o valor F parcial), calculando-se o valor da estatística F_0 . O critério de entrada da variável acontece se $\max(F_0) > F_E$, levando a variável correspondentes a ser adicionada ao modelo. Caso contrário, o procedimento é interrompido e verifica-se se há alguma variável em condições de sair do modelo (critério de saída). O maior valor de F_0 é comparado com o respectivo

valor F_E , se $F_0 > F_E$, a correspondente variável X_j é adicionada ao modelo, caso contrário, o procedimento de entrada é interrompido, e verifica-se a possibilidade da variável X_i sair do modelo, em caso negativo, conclui-se que o modelo de regressão final deve conter apenas a variável X_i (Lucio, 2007).

2.4. Grade of Membership - GoM (Grau de similaridade)

2.4.1. Considerações gerais

A estrutura metodológica apresentada no desenvolvimento desta parte do trabalho denominada *Grade of Membership (GoM)*, guia-se por uma sequência composta de três etapas: (i) a construção da tipologia onde as microrregiões estão inseridas (UF), (ii) a construção da tipologia ou perfil de grupos selecionados a partir dos indicadores definidos para o estudo e (iii) verificação das possíveis associações entre estes perfis extremos e mistos, com a incidência da taxa de mortalidade infantil.

Diferentemente da análise nos conjuntos discretos, em que se baseia a maioria dos métodos estatísticos, onde um elemento ou objeto pode pertencer ou não a um determinado conjunto com características específicas, nos conjuntos nebulosos um elemento ou objeto pode pertencer parcialmente (com determinado grau de pertinência) a múltiplos conjuntos (Cerqueira, 2004).

Como apresentado por Caetano e Machado (2009), no Brasil, o método *GoM* vem se tornando uma ferramenta de grande utilidade na área de Demografia e de Saúde Pública. Uma das vantagens desse método é que se caracteriza por considerar que os eventos ou indivíduos não têm contornos absolutamente definidos, ou seja, não pertencem particularmente a um conjunto, mas sim, podem pertencer, parcialmente, a múltiplos conjuntos, com distintos atributos.

Este método, que tem como base a teoria dos conjuntos dos números nebulosos de Zadeh (1965), é utilizado na modelagem multidimensional de dados discretos (Manton, 1994).

Desta forma, o método apresenta vantagens, à medida que permite às microrregiões pertencerem a mais de um perfil, possibilitando que a heterogeneidade entre as mesmas possa ser expressa como um componente do modelo, o que remete às descrições mais naturais dos grupos gerados. Outra vantagem é a possibilidade de lidar com grande número de casos e variáveis. O

método é baseado nos procedimentos estatísticos de máxima verossimilhança, apresentando, assim, o rigor científico necessário (Cerqueira, 2002).

Para aplicar o método GoM, os dados devem apresentar características discretas, com um número finito de categorias de respostas para a j -ésima variável, ou seja, variáveis como: classe econômica, sexo, cidade, etc. Nos casos em que as variáveis sejam de natureza contínua, estas devem passar por um tratamento de codificação, para gerar variáveis categóricas.

Para entender o conjunto nebuloso que, no caso deste trabalho, é composto pelas microrregiões, existe um escore de pertinência, denotado por g_{ik} , que indica o grau de pertinência do i -ésimo elemento, ao k -ésimo conjunto ou perfil. Este escore não possui interpretação probabilística e pode ser interpretado como um atributo individual, variando de 0 a 1, onde, quanto mais próximo de 1, indica que realmente este elemento possui mais características do k -ésimo perfil (Caetano e Machado, 2009).

A modelagem que provém desta heterogeneidade consiste em identificar variadas características da função de densidade multivariada que descreve a distribuição dos escores do grupo de interesse.

A partir do conhecimento dos chamados perfis extremos ou puros encontra-se um certo número de conjuntos. Estes conjuntos, formados pelos scores, são chamados de partições nebulosas e estão sujeitos às restrições:

$$a) \quad g_{ik} \geq 0 \quad \forall i, j \quad i=1, \dots, 188 \quad j=1, \dots, 12$$

$$b) \quad \sum_{k=1}^K g_{ik} = 1 \quad \forall i \quad i=1, \dots, 188$$

A construção dos perfis, neste estudo, foi feita com base no método programado *GOM3.exe*, sendo utilizada a versão 3.4 do software. Este software foi desenvolvido em ambiente MS-DOS para estimar os modelos *Grade of Membership* (GoM). Possui código fonte aberto e livre e foi escrito originalmente por Woodbury e Clive em 1975, sendo posteriormente modificado por Peter Charpentier e Burt Singer, no Departamento de Epidemiologia e Saúde Pública da Escola de Medicina da Universidade de Yale – EUA, em 1996, e disponível em <http://lib.stat.cmu.edu/DOS/general/.index.html>.

Porém, neste estudo, foi utilizada uma nova versão que foi adaptada por Rafael Kelles V. Laje, para as plataformas LINUX e UNIX, disponível, mediante solicitação, em groupgom@gmail.com.

2.4.2. Modelagem dos conjuntos nebulosos

A aplicação dos conjuntos nebulosos pode ter utilidade em análises estatísticas, no entanto, seus princípios devem ser consistentes aos princípios matemáticos dos modelos de conjuntos nebulosos (Manton, 1994).

Para a formulação do modelo, cada variável j referente ao indivíduo i (microrregião) seja a resposta representada por um conjunto de L_j variáveis aleatórias binárias Y_{ijl} . Então, são requeridas as seguintes suposições:

- a) Variáveis aleatórias Y_{ijl} são independentes para diferentes valores de i , ou seja, as respostas das diferentes unidades independentes;
- b) Os valores de g_{ik} , $k= 1,2,\dots,K$ são realizações das componentes do vetor aleatório $\xi_i = (\xi_{i1}, \xi_{i2}, \dots, \xi_{ik})$ com função de distribuição acumulada $H(x) = P(\xi_i \leq x)$;
- c) Se o grau de pertinência g_{ik} é conhecido, os índices de cada microrregião i , para as várias variáveis (Y_{ijl}) são independentes para as categorias de cada variável;
- d) A probabilidade da resposta 1, para a j -ésima variável, pela microrregião com k -ésimo perfil extremo é denotada por λ_{kjl} . Por pressuposto do modelo, existe pelo menos uma unidade que é membro bem definido do k -ésimo perfil. Este pressuposto dá a probabilidade de resposta, para esta unidade, para os vários níveis de cada variável. Os valores de λ_{kjl} obedecem às seguintes restrições:

$$i) \lambda_{kjl} \geq 0 \quad \forall k, j, l$$

$$ii) \sum_{l=1}^{L_j} \lambda_{kjl} = 1 \quad \forall k, j$$

- e) A probabilidade de uma resposta 1 para a j -ésima variável pela microrregião i , condicional aos escores g_{ik} , é dada por:

$$P(Y_{ijl} = 1) = \sum_{k=1}^K g_{ik} \lambda_{kjl}$$

onde cada g_{ik} é suposto conhecido e satisfaz às restrições “a” e “b”, apresentadas anteriormente.

Com base nos pressupostos, o modelo de probabilidade para a construção do procedimento de estimação de máxima verossimilhança é formulado. O modelo de probabilidade, para uma amostra aleatória é o produto do modelo multinomial.

Considerando os pressupostos acima, o modelo de verossimilhança pode ser escrito como:

$$L(Y) = \prod_{i=1}^I \prod_{j=1}^J \prod_{l=1}^{L_j} \left(\sum_{k=1}^K g_{ik} \lambda_{kjl} \right)^{y_{ijl}}$$

onde: I é o número de microrregiões; J é o número de variáveis incluídas a dimensão do espaço de observações; L_j é o número de categorias de cada uma das J ; K é o número de perfis extremo.

2.4.3. Estimação dos parâmetros do modelo GoM

Nesta etapa, diferentemente da anterior, onde foi apresentada a função de verossimilhança do modelo GoM e suas suposições e restrições que permitiram a sua construção, é apresentado o processo de estimação dos parâmetros do referido modelo, que utiliza princípios da Máxima Verossimilhança, combinado com métodos numéricos iterativos, dado que os estimadores dos referidos parâmetros não têm forma fechada.

O método utilizado consiste em maximizar a expressão:

$$L(Y) = \prod_{i=1}^I \prod_{j=1}^J \prod_{l=1}^{L_j} \left(\sum_{k=1}^K g_{ik} \lambda_{kjl} \right)^{y_{ijl}}$$

onde, a maximização é feita sob cada um dos conjuntos de parâmetros g_{ik} e λ_{kjl} , mantendo o outro constante. Esta estimação é feita de forma iterativa.

Uma análise mais detalhada do método pode ser encontrada em Manton (1994).

3. RESULTADOS

Apresenta-se a seguir, os resultados encontrados com base nos objetivos específicos que nortearam o desenvolvimento do estudo, sob a ótica das opções metodológicas adotadas.

3.1. Avaliação da qualidade dos dados com base na análise das Causas Mal Definidas referentes à mortalidade Infantil do Brasil e região Nordeste

A preocupação com a qualidade dos dados no Brasil vem se tornando cada vez maior, à medida que se necessita de dados com maior confiabilidade e qualidade, para uso em estudos diversos.

Paes (2007), aponta que um sério problema emerge quando se deseja caracterizar o perfil da mortalidade quanto à distribuição por causas de morte, uma vez que estas, no Brasil, são registradas com certa imprecisão. Um dos mecanismos para avaliá-las é se recorrer à análise dos óbitos com Causas Mal Definidas. Uma definição de “causas desconhecidas”, também é apresentada e representa o resultado da combinação de dois fatores: óbitos não registrados, cujas causas de morte, conseqüentemente, não são conhecidas e os óbitos registrados, cujas causas básicas foram mal definidas e que, também, não se conhecem.

Na sequência, Santo (2000), também aponta que a proporção de óbitos por causas mal definidas (CMD) é um indicador clássico de qualidade das informações sobre mortalidade, não devendo ultrapassar o limiar de 10%, para se considerar boa a qualidade das informações.

Destaca-se, também a classificação apresentada por Paes (2007), num estudo que avaliou dados de óbitos por causas desconhecidas, para a população adulta dos estados brasileiros de 1990 a 2000, no qual, para quantificar a magnitude das causas mal definidas, calculou-se seu percentual em relação ao total de morte de cada estado. No estudo ele utilizou a classificação proposta por Chackiel (1986), que estabeleceu quatro categorias para os percentuais de causas mal definidas: adequada, sendo menor que 10%; pouco adequada para valores entre 10% e 15%; inadequada para percentuais entre 16% e 30% e altamente inadequada se superior ou igual a 30%.

Os dados de óbitos analisados neste estudo foram aqueles classificados no capítulo XVIII – **Sintomas, sinais e achados anormais em exames clínicos e de**

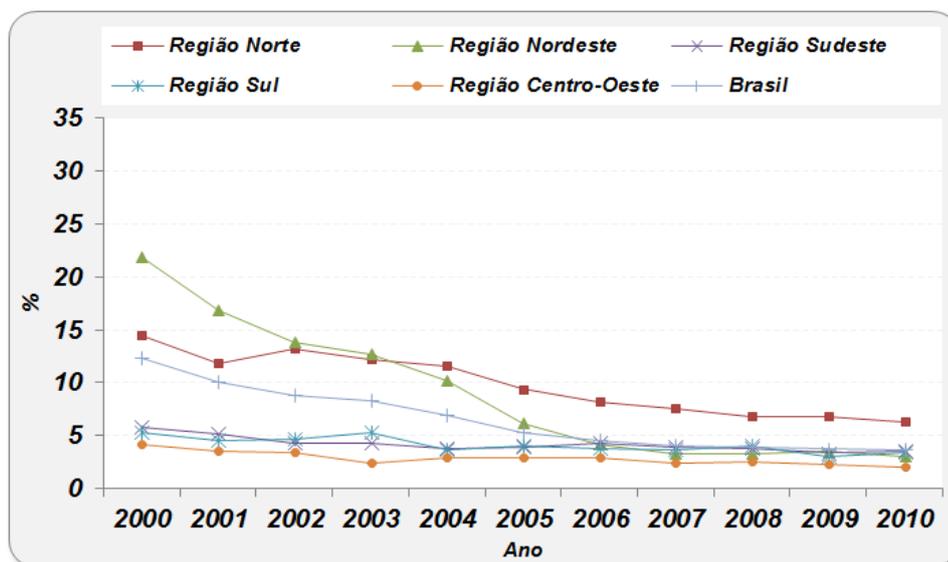
laboratório não classificados em outra parte, da Classificação Internacional de Doenças, 10ª revisão – CID-10.

No ano de 2010, foram registrados 39.870 óbitos infantis, sendo 1.440 destes, registrados como Causa de Morte Mal Definida (CMD), o que representa 3,6% do total de óbitos para este ano. Resultado bastante satisfatório, se comparado com o ano de 2000, onde foram registrados 68.199 óbitos de menores de um ano e, deste total, 8.388 (12,3%) foram classificados como CMD.

Para avaliar o comportamento da CMD por óbitos de crianças menores de um ano, nas grandes regiões brasileiras, tomou-se a série histórica do período de 2000 a 2010, da proporção de óbitos por essas causas.

Com base no gráfico 03, pode-se perceber uma significativa queda, em termos percentuais, nas causas de morte por região brasileira. A região Nordeste, que em 2000 apresentava 21,8%, caiu para proporções próximas das regiões Sul e Sudeste, em 2010, atingindo o patamar de 3,1% de CMD.

Gráfico 03 – Evolução da mortalidade proporcional CMD, para os óbitos infantis (menores de 1 ano), segundo região brasileira, 2000-2010.



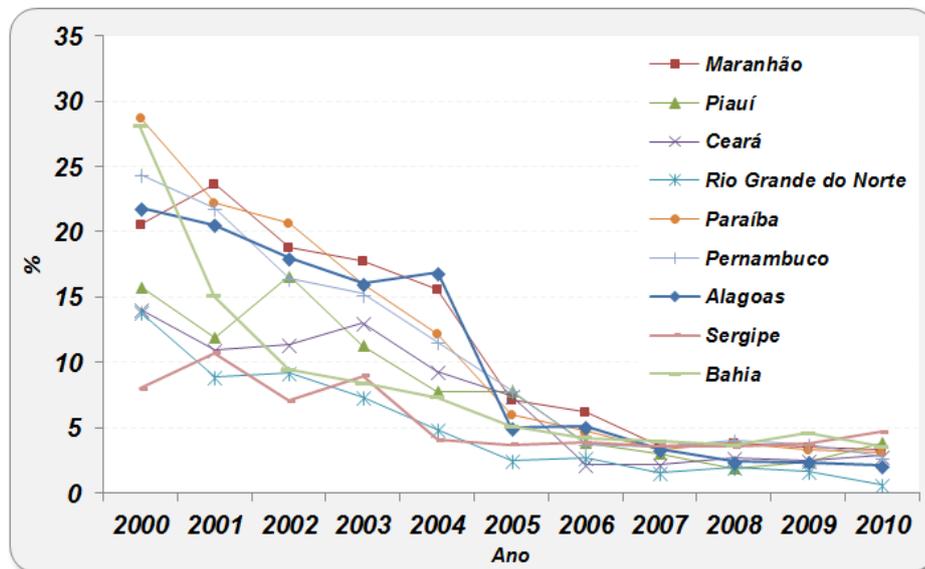
Fonte: Elaborado com base nos dados disponíveis no DATASUS - Ministério da Saúde.

Ponto notório é a distribuição das proporções para a região Centro-Oeste, onde, desde o ano de 2000, já apresentava índices abaixo de 5%, mostrando,

assim, o cuidado com a declaração referente aos óbitos de menores de um ano e suas classificações.

O gráfico 04 apresenta a forte tendência de queda nas proporções de causas mal definidas nas UF da região Nordeste. Percebe-se que, até o ano de 2006, as UF vinham em um constante declínio, onde flutuavam em torno de 15% a 30% e caíram para percentuais abaixo de 5%, em 2010.

Gráfico 04 – Evolução da mortalidade proporcional por CMD, para os óbitos infantis (menores de 1 ano), segundo UF da região Nordeste brasileira, 2000-2010.



Fonte: Elaborado com base nos dados disponíveis no DATASUS - Ministério da Saúde.

A região Nordeste, que apresentava 5.357 óbitos infantis classificados como CMD, em 2000, caiu para apenas 406 óbitos, ou seja, uma redução de quase 13 vezes, com todas as UF seguindo esta mesma tendência.

Tais informações demonstram o crescente interesse na melhoria das estatísticas de óbitos, relacionadas à mortalidade infantil, também acompanhado pela mortalidade em outras faixas etárias.

3.2. Correção do Sub-registro do SIM

Nesta etapa do trabalho foi desenvolvida a correção dos dados de mortalidade fornecidos pelo SIM. Como proposto pelos objetivos específicos, foi

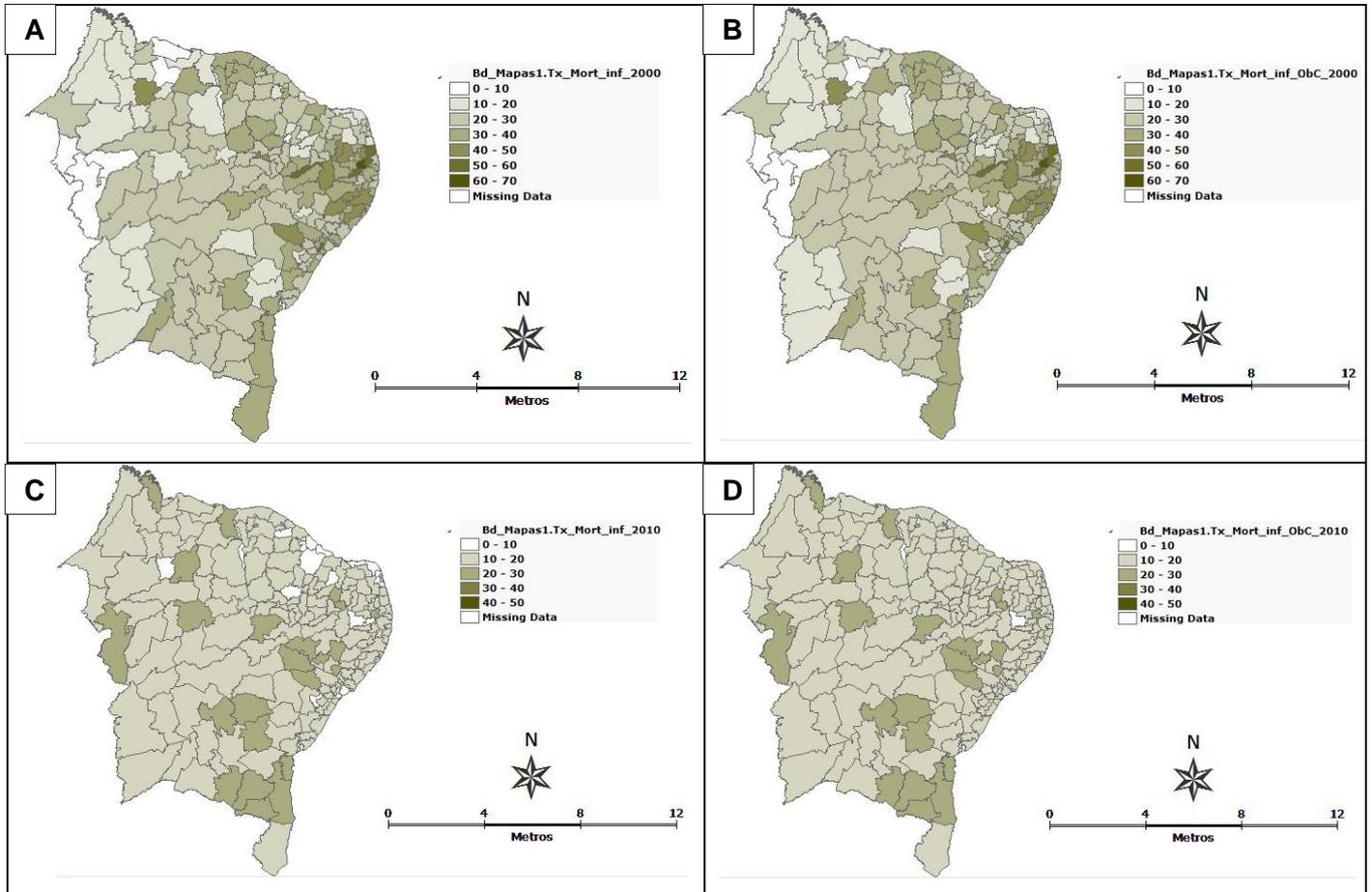
desenvolvida uma correção no número de óbitos infantis, para se encontrar uma posterior taxa de mortalidade infantil corrigida, a partir da metodologia do estimador Bayesiano empírico.

Conforme apresentado na metodologia do estudo, o modelo Bayesiano empírico foi empregado para suavização das taxas das microrregiões que sofrem grandes flutuações, em função dos reduzidos tamanhos populacionais, no caso os óbitos infantis e/ou número de crianças nascidas vivas. O emprego do modelo Bayesiano gera os estimadores de Bayes empíricos ou taxas suavizadas, que permitem uma distribuição espacial com menor efeito de flutuações aleatórias.

Por estar trabalhando com pequenas áreas (microrregiões), um problema que pode ocorrer no uso das taxas, é a instabilidade nos resultados, pois se trata do risco de um determinado fenômeno, o qual ocorre quando a população de interesse é pequena, fato ocorrido com os componentes formadores da taxa de mortalidade infantil, podendo resultar em taxas irreais. Nesses casos, as variações ou flutuações aleatórias (causais) de pequenos números associadas a um ou dois eventos em regiões que apresentam uma população pequena, podem fornecer variações significantes nas taxas brutas. E, contrário a isto, em regiões onde a taxa é representada por uma população grande, porém, com a mesma quantidade de eventos de uma população pequena, podem trazer taxas totalmente diferentes.

Segundo Freire (2001), um estimador desta natureza aproxima uma taxa observada em uma determinada área pequena a uma taxa média global ou, ainda, a uma taxa média dos vizinhos, levando em consideração o tamanho da população da área em questão. No estudo, a área pequena utilizada foi a microrregião e a área maior foi a mesorregião. Este exercício foi realizado em cada UF, de forma isolada, segundo as suas microrregiões e mesorregiões, para informações dos anos de 2000 e 2010, os quais têm suas taxas apresentadas em mapas temáticos nas situações com e sem correção (Mapa 01).

Mapa 01 – Comparativo das taxas de mortalidade infantil para o conjunto de informações corrigidas segundo estimador bayesiano empírico e cálculos diretos (taxas observadas) para os anos de 2000 e 2010, Nordeste, Brasil.



Fonte: Elaboração própria com base em dados disponíveis no DATASUS/MS.

LEGENDA:

MAPA A: Taxa de Mortalidade Infantil (Cálculo direto) – 2000

MAPA B: Taxa de Mortalidade Infantil (Estimador Bayesiano) – 2000

MAPA C: Taxa de Mortalidade Infantil (Cálculo direto) – 2010

MAPA D: Taxa de Mortalidade Infantil (Estimador Bayesiano) – 2010

Com base no comparativo (Mapa 1), percebe-se claramente uma significativa melhora relacionada às taxas de mortalidade infantil na região Nordeste, entre 2000 e 2010, pois o mapa B comparado ao D, apresentou um clareamento em níveis que antes eram representados por cores mais escuras, segundo a escala de taxas de mortalidade do mapa. Na primeira linha do quadro comparativo acima, estão expostas as taxas de mortalidade para o ano de 2000, fazendo um comparativo das taxas calculadas de forma direta (quadrante esquerdo) e as estimadas, a partir do

método bayesiano empírico (quadrante direito). O que se percebe, é uma pequena suavização (uniformização), especialmente nas microrregiões dos estados Maranhão, Piauí e Ceará, em decorrência de um aumento em taxas provocado pelo método.

Atuação mais notável do estimador ocorreu na segunda linha do quadro comparativo, onde são comparadas as taxas por microrregião para o ano de 2010. À esquerda está o mapa das taxas a partir do cálculo direto e ao lado direito encontra-se o mapa com as taxas corrigidas (suavizadas). O estimador mostrou-se mais atuante nas microrregiões dos Estados do Ceará, Rio Grande do Norte, Paraíba e Pernambuco.

Houve, em média, uma queda considerável na taxa de mortalidade da região Nordeste, a partir das microrregiões. Em 2000, a taxa de mortalidade encontrada (informações corrigidas), era de aproximadamente 28,8 óbitos infantis para cada 1000 nascidos vivos e, em 2010, passou para (em média) 16,5 óbitos para cada 1000 nascidos vivos. Ponto notável, também, é a queda de variabilidade, o que indica que as microrregiões estão se tornando mais homogêneas com relação às suas taxas de mortalidade infantil, ou seja, em 2000, o desvio padrão dessas taxas era de 9,6 óbitos/1000 n.v., já em 2010 passou para 3,7 óbitos/1000 n.v.

3.3. Seleção das variáveis para posterior análise de agrupamentos

No modelo de regressão linear múltipla ponderada tomou-se como variáveis independentes ou explicativas: Rend_Pc, Prop_Bx_rend, Tx_Md_Analf_M, Id_Md_M, A_Md_Est, Dens_Domic, Num_Md_Cont_PN, Gr_Urb, Cobert_Agua, Exist_C_Lixo, Cobert_Esg_Sanit. Como variável resposta, tomou-se o logito da Taxa de Mortalidade Infantil para o ano de 2010 (Tx_Mort_Inf_Obc_2010_LOGIT). A transformação Logit foi utilizada na variável explicativa (Y_i) para evitar o problema restritivo de que os valores de probabilidades são números no intervalo de 0 a 1. Esta transformação foi definida como:

$$\log\left(\frac{Y_i}{1 - Y_i}\right) = \beta_0 + \beta_1 X_{j1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i$$

No quadro 02, são apresentadas as variáveis que foram selecionadas na regressão linear múltipla ponderada, podendo ser consideradas, como variáveis que melhor explicam o logito da taxa de mortalidade infantil nas microrregiões da região Nordeste.

Quadro 02 – Descrição dos nomes simplificados das variáveis envolvidas no modelo de regressão linear múltipla ponderada.

Código das variáveis	Descrição da variável
*** Rend_Pc <small>(Selecionada)</small>	<i>Renda Média domiciliar per capita</i>
Prop_Bx_rend	<i>Proporção de pessoas com baixa renda (< ½ S.M)</i>
Tx_Md_Analf_M	<i>Taxa de analfabetismo feminino</i>
Id_Md_M	<i>Idade média das mães quando notificado o óbito infantil</i>
*** A_Md_Est <small>(Selecionada)</small>	<i>Anos médio de estudo materno quando notificado o óbito infantil</i>
*** Dens_Domic <small>(Selecionada)</small>	<i>Densidade domiciliar</i>
*** Num_Md_Cont_PN <small>(Selecionada)</small>	<i>Número médio de consultas Pré-Natal</i>
Gr_Urb	<i>Grau de Urbanização</i>
Cobert_Agua	<i>Proporção de domicílios com abastecimento de água adequado</i>
Exist_C_Lixo	<i>Proporção de domicílio com existência de coleta de lixo adequada</i>
Cobert_Esg_Sanit	<i>Proporção de domicílios com existência de cobertura de esgotamento sanitário</i>

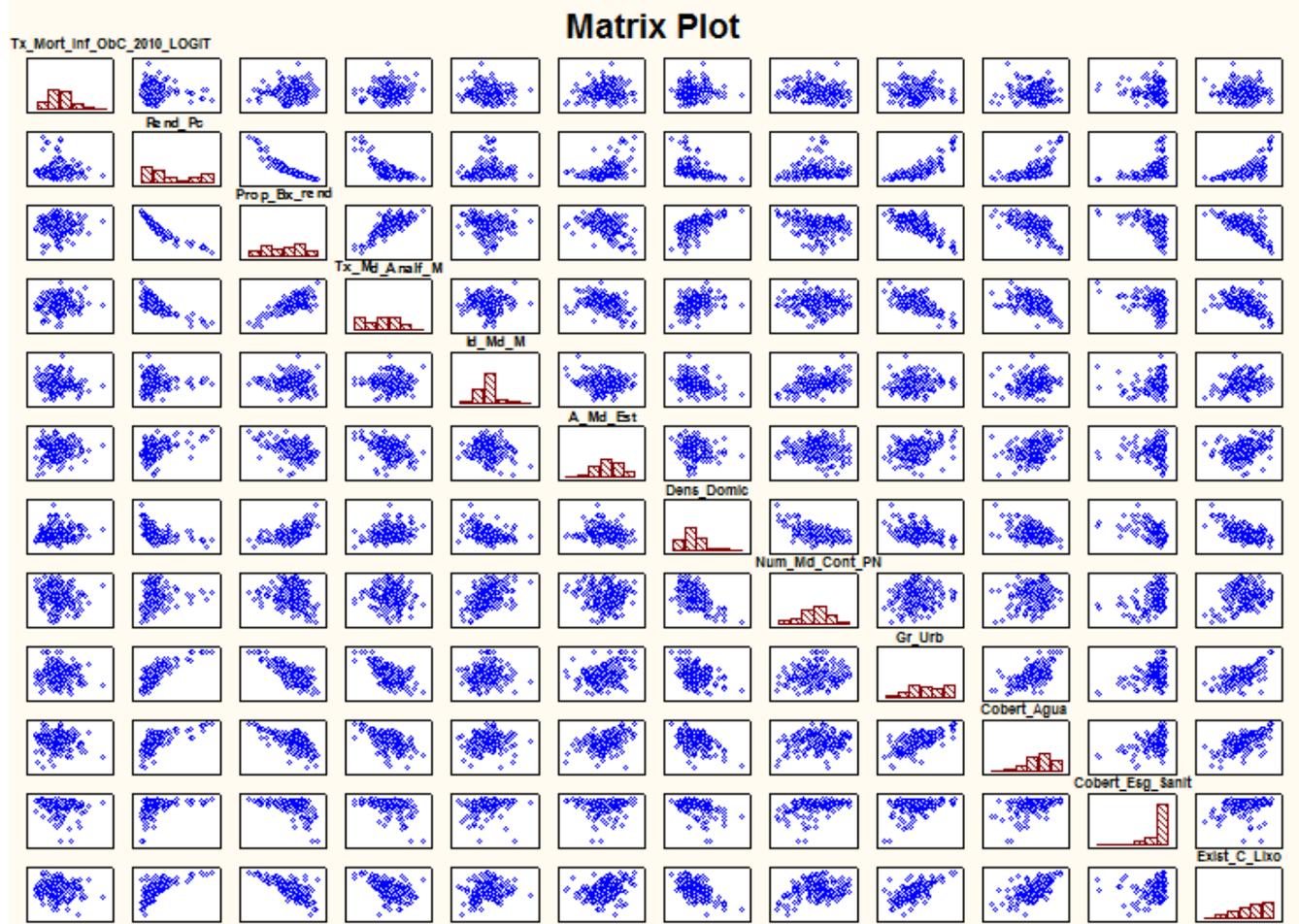
Fonte: Elaboração própria.

*** Variável selecionada para o modelo

O gráfico a seguir (gráfico 05), foi desenvolvido no intuito de avaliar o comportamento das variáveis, entre si, permitindo visualizar a relação bivariada, através da análise gráfica.

Das correlações (análise bivariada) encontradas, várias são identificadas e abordadas em outros estudos, como: a taxa média de analfabetismo materno com a renda média per capita (relação inversa); proporção de baixa renda com taxa média de analfabetismo (relação direta); proporção de pessoas com baixa renda com densidade domiciliar (relação direta); densidade domiciliar com número médio de consultas pré-natal (relação inversa).

Gráfico 05 – Avaliação bivariada do comportamento das variáveis em análise - Scatterplot (Matriz).



Fonte: Elaboração própria a partir da seleção do modelo MRL (Modelo de regressão linear múltipla).

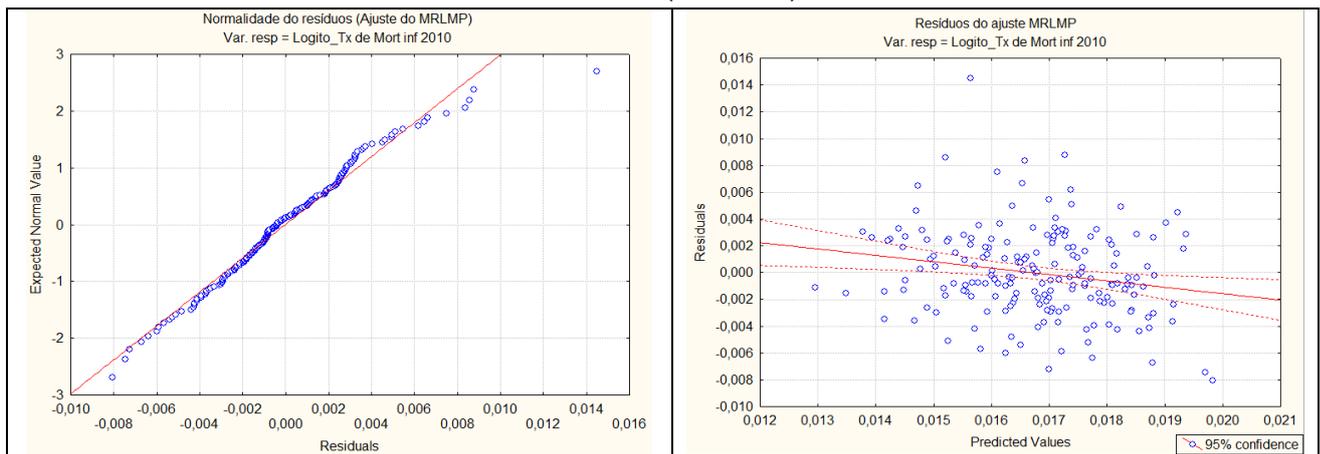
O modelo de regressão foi ajustado pelo método “*Forward / Stepwise*”, de forma que foram utilizadas as onze variáveis independentes (X_i) e, apenas, Rend_Pc, A_Md_Est, Dens_Domic e Num_Md_Cont_PN foram selecionadas para fazer parte do modelo final, ambas apresentando-se estatisticamente significantes, ao nível de 5%, p-valor <0,022 (Tabela 01).

Como se sabe, é necessário checar algumas suposições para a validação do modelo de regressão linear múltipla. As suposições do modelo ajustado precisam ser avaliadas para que os resultados sejam confiáveis. Para tanto, se fez uma análise de resíduos (gráfica) para investigar a adequabilidade do modelo de regressão.

Com esta análise, foi possível verificar a presença da microrregião de Fernando de Noronha/PE, comportando-se de forma diferente. Portanto, optou-se por excluir esta unidade da análise, por suas características diferenciadas, evitando possível distorção dos resultados.

No quadro 03, observa-se a análise de resíduos para o modelo de regressão, com a exclusão da microrregião citada, mostrando, desta forma que a hipótese de normalidade é plenamente satisfeita pelos dados, com os resíduos distribuindo-se de forma aleatoriamente esperada em torno de zero. A distribuição dos resíduos e de suas probabilidades normais comprovam as pressuposições de aleatoriedade e normalidade dos dados.

Quadro 03 – Análise (gráfica) dos resíduos do Modelo de Regressão Linear Múltipla Ponderada (MRLMP).



Fonte: Elaboração própria a partir da seleção do modelo MRLMP (Modelo de regressão linear múltipla ponderada).

Dos resultados encontrados pode-se concluir que as variáveis que melhor explicam o logito da taxa de mortalidade infantil por microrregião para o ano de 2010, para o conjunto de informações do estudo, são relacionadas à renda per capita, à educação materna, assistência à saúde, representada pelo número médio de consultas pré-natal e densidade domiciliar (Tabela 01). Ou seja, exercem um maior efeito nas taxas de mortalidade infantil, que as demais variáveis consideradas, expressando a grande importância desses fatores na redução dessas taxas.

Tabela 01 – Sumário da regressão linear múltipla para a variável resposta Logit da Taxa de mortalidade infantil 2010, *método forward stepwise*, microrregiões da região Nordeste.

Variáveis	b padronizado	B	Erro padrão de B	t	p-valor
Intercepto		0,044	0,006	6,993	0,0001
<i>Rend_Pc</i>	-0,407	0,000005	0,0001	-2,314	0,022
<i>A_Md_Est</i>	0,235	0,001	0,0001	2,934	0,004
<i>Dens_Domic</i>	-0,317	-0,004	0,001	-3,827	0,0001
<i>Num_Md_Cont_PN</i>	-0,358	-0,002	0,0001	-4,649	0,0001

Fonte: Elaboração própria a partir da seleção do modelo MRLM (Modelo de Regressão Linear Múltipla).

Os resultados encontrados corroboram com os achados de Wood e Carvalho (1994, p. 105), que constataram a relação entre escolaridade da mãe e mortalidade. Esta relação tem grande impacto entre educação da mãe e a probabilidade de que seus filhos sobrevivam até o quinto aniversário.

Também, segundo constatação de Formiga e Ramos (2002), a definição dos fatores socioeconômicos a serem considerados em estudos de mortalidade assume especial importância. Dentre eles, os mais comumente utilizados são a renda e a educação, mas outras variáveis disponíveis podem ser também utilizadas.

3.4. Análise de *Clusters* – formação de Grupos homogêneos

Nesta seção, foram determinados grupos de microrregiões, dentro de uma estrutura hierárquica, a partir de uma análise de *cluster*, tomando por base o indicador “taxa de mortalidade infantil do ano 2010”, segundo os fatores selecionados (renda per capita, anos de estudo materno, densidade domiciliar e número médio de consultas pré-natal), que demonstraram uma maior influência sobre esse indicador (identificado pela aplicação do modelo de regressão linear múltipla ponderada), identificando-se grupos de microrregiões mais ou menos explicados por este.

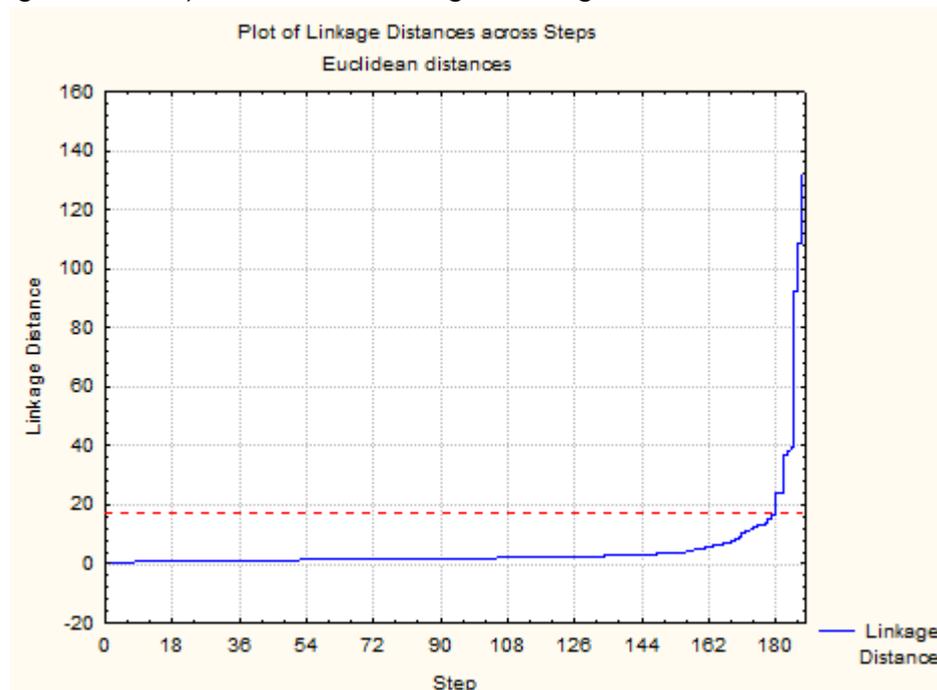
Para a análise de *cluster*, foram consideradas as 188 microrregiões da região Nordeste e optou-se por estabelecer os grupamentos a partir delas, utilizando-se a regra da ligação simples (*single linkage*) e medidas de distâncias Euclidianas

(*Euclidean distances*), descritas em Johnson e Wichern (1988). Para que o método fosse realizado de forma aceitável, fez-se uso de uma adequação nos dados das variáveis (Rend_Pc, A_Md_Est, Dens_Domic e Num_Md_Cont_PN), devido ao fato de apresentarem unidades de escalas diferentes. Com a adequação realizada, as variáveis ficaram parametrizadas em uma escala variando de 0 a 1, tornando-se possível a análise.

A partir da análise do dendograma das Microrregiões, foi possível identificar 11 grupos homogêneos, a partir do corte feito na distância, aproximada, de ligação 16 da medida euclidiana (gráfico 06). Entretanto, devido ao fato da Região Nordeste apresentar comportamento semelhante entre as microrregiões (segundo as variáveis estudadas), o grupo 11 obteve mais destaque com 165 microrregiões com comportamentos semelhantes. Em seguida, aparece o grupo 10 com um total de 6 microrregiões. Os grupos 2, 6, 7 e 8 cada um apresentou 3 microrregiões e os demais grupos, cada um, apresentou uma microrregião.

Na tabela 02, é apresentado um resumo dos agrupamentos encontrados segundo os 11 grupos criados na análise de Clusters.

Gráfico 06 – Gráfico de passos (Step) versus o nível de distância do agrupamento (Linkge Distance) das 188 microrregiões, segundo as variáveis analisadas.



Fonte: Elaboração própria, com base na Análise de Clusters.

Tabela 02 – Distribuição das microrregiões segundo UF e grupos formados na análise de *Clusters*.

Grupos Formados	Frequência de microrregiões por UF									Total
	MA	PI	CE	RN	PB	PE	AL	SE	BA	
1	0	0	0	0	0	1	0	0	0	1
2	0	0	0	0	1	1	0	1	0	3
3	0	0	0	0	0	0	0	0	1	1
4	0	0	0	1	0	0	0	0	0	1
5	0	1	0	0	0	0	0	0	0	1
6	1	0	1	0	0	0	1	0	0	3
7	0	0	0	1	1	0	0	0	1	3
8	0	0	0	1	0	0	0	0	0	1
9	0	0	0	0	0	1	0	0	2	3
10	0	0	0	1	0	1	0	0	4	6
11	20	14	32	15	21	15	12	12	24	165
Total de Microrregiões	21	15	33	19	23	19	13	13	32	188

Fonte: Elaboração própria, com base na Análise de Clusters.

O dendograma (gráfico 07) e os grupos homogêneos (quadro 07) encontrados, encontram-se no Apêndice.

Esta análise apresentou seu mérito a partir do momento em que proporcionou uma primeira visão dos agrupamentos das microrregiões com comportamento homogêneo, segundo as variáveis envolvidas na avaliação. Portanto, este recurso metodológico serviu para dar uma primeira visão do que poderia ser encontrado na análise utilizando a ferramenta *GoM*.

3.5. Construção dos Perfis de mortalidade (*GoM*)

3.5.1. Características das UF e microrregiões da região Nordeste do Brasil

Nesta sessão serão apresentados os resultados das análises descritivas de cada UF, com base nas variáveis utilizadas. É de suma importância entender o comportamento de cada variável em cada estado, pois possibilita um melhor entendimento do fenômeno, assim como acompanhar o desenvolvimento comparativo de cada região.

Com o intuito de identificar possíveis variações de indicadores entre as UF, foram construídos gráficos *box-plots*, que fornecem uma visualização de importantes aspectos descritivos dos dados, fornecendo uma síntese de informações (distribuição, variabilidade e assimetria).

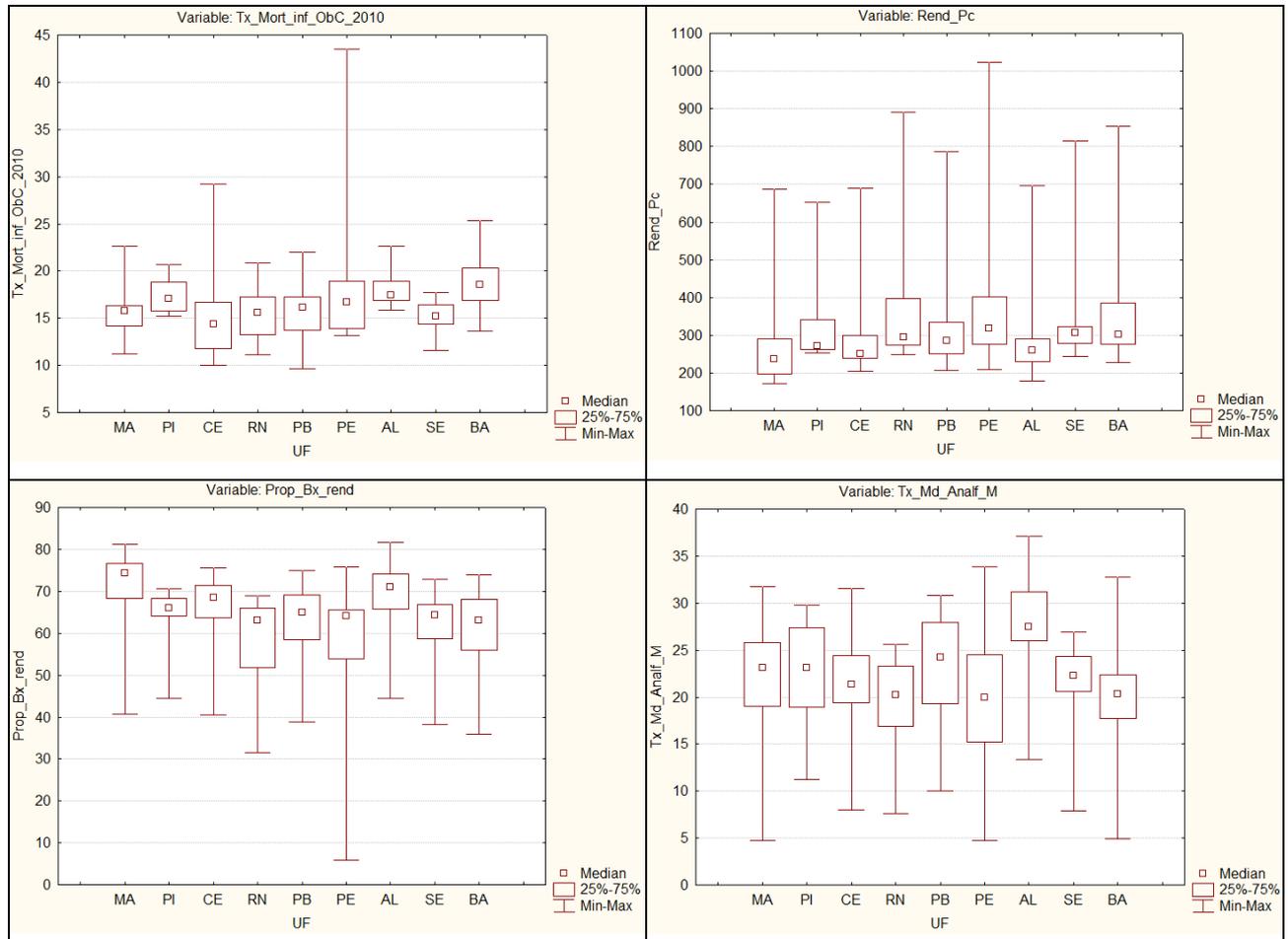
No quadro 04, são apresentados os gráficos comparativos referentes a quatro variáveis: taxa de mortalidade infantil, renda média domiciliar per capita, proporção de pessoas com baixa renda e taxa de analfabetismo feminino. Na primeira (Tx_Mort_Inf_ObC_2010), percebe-se clara discrepância no estado de Pernambuco, devido à microrregião de Fernando de Noronha, apresentando-se, segundo a estimativa, como a maior taxa de mortalidade infantil entre as microrregiões da Região Nordeste (43,5/1000 n.v.), seguido do estado do Ceará, com a segunda maior taxa (29,2/1000 n.v), mais precisamente a microrregião de Meruoca. Avaliando como um todo, a média da taxa de mortalidade infantil para os estados da Região Nordeste é de aproximadamente 16,6/1000 n.v. Percebe-se, também, que os estados com menor variabilidade nas taxas de mortalidade, em suas microrregiões, são os estados do Piauí, Rio Grande do Norte, Sergipe e Alagoas, mostrando, assim, uma concentração das taxas de mortalidade em torno do valor mediano.

Avaliando a segunda variável, renda per capita (Rend_Pc), existe um destaque, ou topo máximo, em cada estado da região Nordeste, onde os pontos de máximo nos gráficos representam exatamente as microrregiões onde estão localizadas as capitais estaduais, destacando-se, no entanto, a microrregião de Fernando de Noronha, como a que apresenta melhor renda per capita, com R\$ 1.024,00. Os estados do Maranhão e Alagoas são os que apresentam os menores valores de renda per capita por microrregião.

A terceira variável analisada é a proporção de pessoas com baixa renda (Prop_Bx_rend), onde novamente evidencia-se a microrregião de Fernando de Noronha (PE), com apenas 5,9% da sua população com menos de $\frac{1}{2}$ S.M. Os estados do Maranhão e Alagoas apresentam as maiores proporções de pessoas com baixa renda. A variação desta proporção entre as microrregiões dos estados do Nordeste gira em torno de 63,8%.

A quarta variável, taxa de analfabetismo feminino (Tx_Md_Analf_M), apresenta o estado de Alagoas destacando-se como principal, com taxas mais elevadas, entre as suas microrregiões. O estado com as menores taxas de analfabetismo feminino é o Rio Grande do Norte, com 19,3%, avaliando também em termos de variabilidade entre as microrregiões nela inseridas.

Quadro 04 – Análise gráfica (box-plots) da distribuição das variáveis: Taxa de mortalidade infantil (Tx_Mort_inf_ObC_2010), renda média domiciliar per capita (Rend_Pc), proporção de pessoas com baixa renda (Prop_Bx_rend) e taxa de analfabetismo feminino (Tx_Md_Analf_M), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.



Fonte: Elaboração própria, com base nos dados do DATASUS/MS e Microdados do Censo 2010.

Dando sequência às análises, o quadro 5, apresenta, também, mais quatro variáveis: idade média das mães quando notificado o óbito, anos médio de estudo materno, densidade domiciliar e número médio de consultas pré-natal.

Do exposto, avaliando a primeira variável, idade média das mães quando notificado o óbito (Id_Md_M), destaca-se o estado do Ceará, com a microrregião de Serra do Pereiro, com idade média de 31,3 anos, sendo o maior valor encontrado em 2010, para a região Nordeste. As microrregiões do estado do Maranhão destacam-se com as menores idades, aproximadamente 23,2 anos, com um desvio padrão de 1,2 anos. Ao todo, a média de idade das microrregiões da região

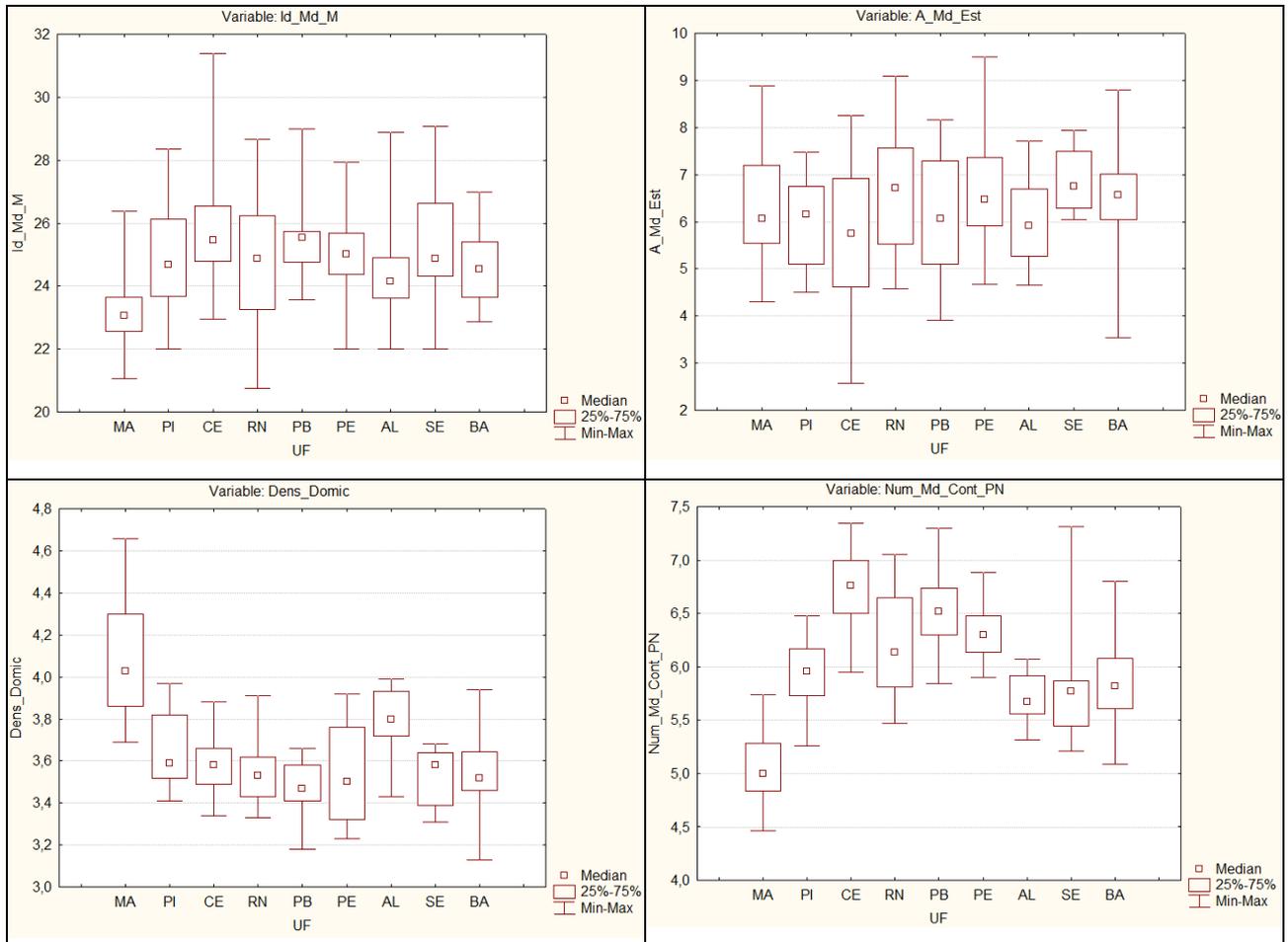
Nordeste é de 24,3 anos, com a microrregião de Umarizal (RN) apresentando a menor idade registrada (20,7 anos).

A segunda variável, anos médio de estudo materno (A_Md_Est), apresenta destaque para o estado de Sergipe, onde aparece uma maior concentração de microrregiões com nível mais alto em anos de estudo. O estado de Pernambuco é o que apresenta a microrregião com o maior índice (Fernando de Noronha – 9,5 anos de estudo) e o Estado do Ceará é que apresenta o menor índice, representado pela microrregião de Cascavel. O indicador médio de anos de estudo, para as microrregiões da região Nordeste, é de aproximadamente 6,3 anos.

Ao avaliar a variável densidade domiciliar (Dens_Domic), apresenta-se em destaque o estado do Maranhão, com uma média das suas microrregiões de 4,05 habitantes por domicílio. Já o estado da Bahia desponta com um dos menores índices de densidade domiciliar, representado pela microrregião de Salvador (3,1 habitantes por domicílio).

Comportamento, também de destaque, para o estado do Maranhão, quando se avalia a variável número médio de consultas pré-natal (Num_Md_Cont_PN), onde o estado desponta, de uma forma geral, com um índice de 4,05 consultas, bem abaixo da média da região Nordeste, que é de 6,08 consultas. O estado do Ceará apresenta-se como o estado com melhor índice (6,7 consultas) de mães que realizaram consulta pré-natal na região Nordeste, no ano de 2010.

Quadro 05 – Análise gráfica (box-plots) da distribuição das variáveis: idade média das mães (Id_Md_M), anos médio de estudo materno (A_Md_Est), Densidade domiciliar (Dens_Domic) e Número médio de consultas pré-natal (Num_Md_Cont_PN), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.



Fonte: Elaboração própria, com base nos dados do DATASUS/MS e Microdados do Censo 2010.

Por fim, no quadro 06, são apresentados os gráficos comparativos referentes às quatro últimas variáveis estudadas: grau de urbanização, proporção de domicílios com abastecimento de água adequado, proporção de domicílios com existência de coleta de lixo adequada e proporção de domicílios com existência de cobertura de esgotamento sanitário adequado.

Avaliando a primeira variável, grau de urbanização (Gr_Urb), percebe-se que os estados da região Nordeste, segundo suas microrregiões, apresentam grau de urbanização próximo de 61,3%. Entretanto, o estado de Pernambuco apresenta uma média de suas microrregiões em torno de 70%, ou seja, um número de pessoas residentes em áreas urbanas bem superior à média regional. O estado de Alagoas

detém o menor índice, representado pela microrregião Serrana do Sertão Alagoano, com apenas 28,8% dos habitantes morando em setor urbano.

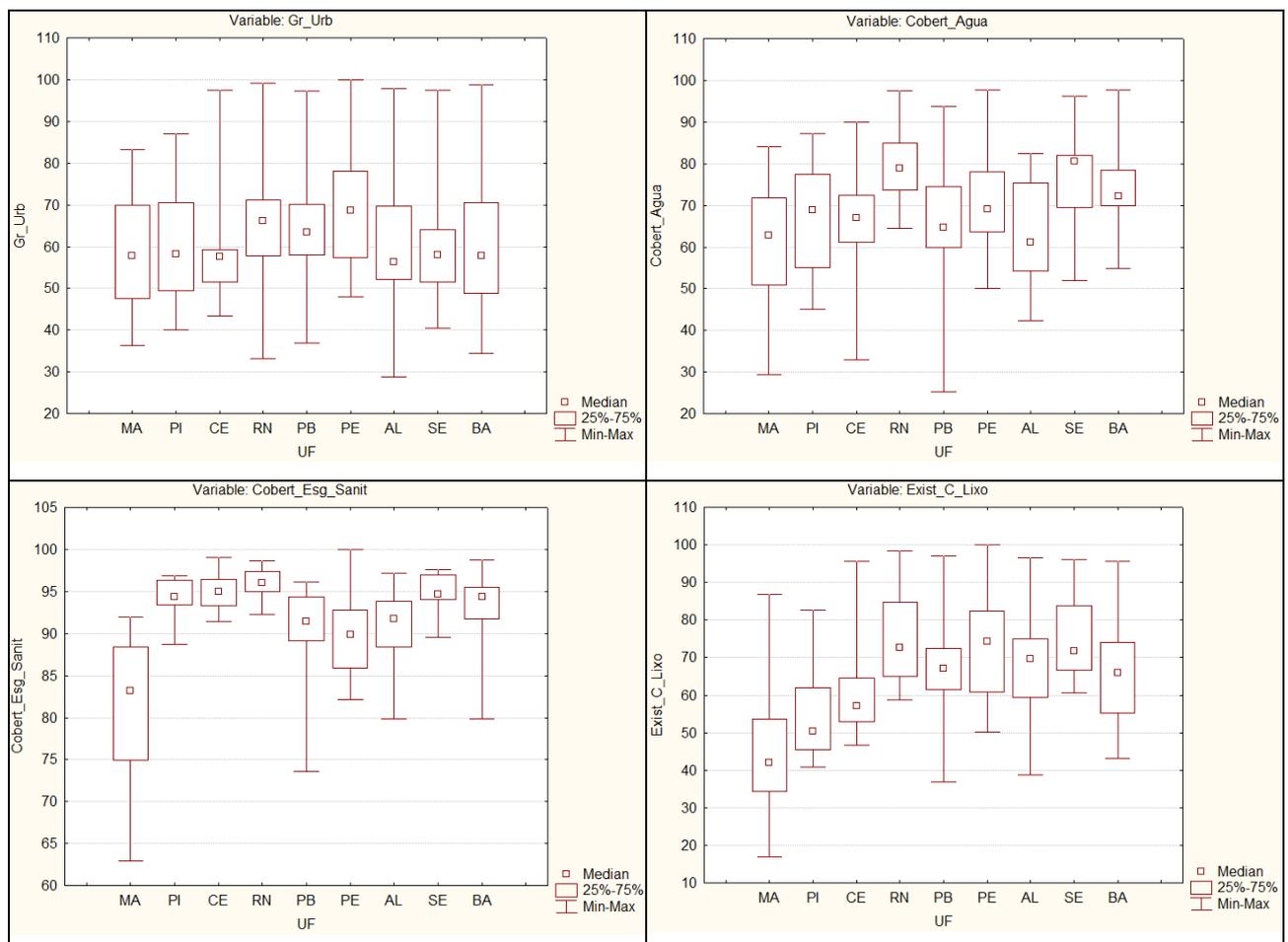
A segunda variável, proporção de domicílios com abastecimento adequado de água (Cobert_Agua), apresenta uma grande variabilidade entre as UF, porém o estado do Rio Grande do Norte destaca-se como um dos estados com melhores indicadores em suas microrregiões, com taxa média de 79,4% de cobertura, muito além da taxa de média para a região Nordeste (69,07%). Os estados do Maranhão e Paraíba apresentam os menores índices de cobertura, representados, respectivamente, pelas microrregiões de Lençóis Maranhenses (29,3%) e Umbuzeiro (25,2%).

Estudando o comportamento da variável proporção de domicílios com existência de coleta de lixo adequada (Exist_C_Lixo), percebe-se clara deficiência no estado do Maranhão, despontando novamente como o estado com piores coberturas de coleta de lixo em suas microrregiões, seguido do estado do Piauí. Novamente, as microrregiões do estado do Rio Grande do Norte despontam como as melhores coberturas neste quesito. A taxa média de cobertura para a região Nordeste é de 64,4%, porém, na microrregião do Litoral Ocidental Maranhense chega a 17,03 %, mostrando um ponto de grande vulnerabilidade nesta região.

A quarta e última variável analisada, proporção de domicílios com existência de cobertura de esgotamento sanitário adequado (Cobert_Esg_Sanit), é a que apresenta as maiores disparidades inter-regionais. O estado do Rio Grande do Norte apresenta as melhores taxas de cobertura referente a saneamento entre suas microrregiões (média de 96% de cobertura). Contrariamente, o estado do Maranhão, tem sua variação entre 62,9% e 92,9%. O estado de Alagoas também se apresenta, na sequência, como um dos piores, em termos de cobertura (79,8%). No estado de Pernambuco, destaca-se a microrregião de Fernando de Noronha com 100% de cobertura de esgotamento sanitário adequado.

Portanto, os mapas com a distribuição espacial de cada variável por microrregião e UF, assim como a tabela com as medidas estatísticas (média, desvio padrão, máximo e mínimo), encontram-se localizados no apêndice deste trabalho.

Quadro 06 – Análise gráfica (box-plots) da distribuição das variáveis: Grau de urbanização (Gr_Urb), proporção de domicílios com abastecimento de água (Cobert_Agua), proporção de domicílios com cobertura de esgotamento sanitário (Cobert_Esg_Sanit) e proporção de domicílio com coleta de lixo adequada (Exist_C_Lixo), segundo UF da região Nordeste, em termos de duas microrregiões, 2010.



Fonte: Elaboração própria, com base nos dados do DATASUS/MS e Microdados do Censo 2010.

3.5.2. Tipologias das microrregiões segundo as variáveis trabalhadas

Inicialmente, convém apresentar uma breve explicação no que diz respeito à construção das tipologias das microrregiões. Na construção, utilizou-se doze variáveis que foram categorizadas em função dos seus quartis, atribuindo ao menor quartil a indicação de categoria 1 e assim, na sequência até a maior indicação de quartil, com categoria 4. Dessa forma, todas as variáveis passaram a ter quatro categorias, tendo-se uma frequência marginal de cada variável de 25%, aproximadamente. Para avaliar quantos perfis deveriam ser utilizados, testou-se

inicialmente duas possibilidades no número de perfis definidos, uma com 3 perfis e outra com 4 perfis. Finalmente, optou-se por 3 perfis, onde no total formaram 9 perfis, entre extremos e mistos, mais o perfil amorfo. Assim, a interpretação dos resultados comportou-se de forma mais adequada e satisfatória à realidade dos dados.

Barbosa (2006), em seu estudo, analisando o panorama da fecundidade e condições socioeconômicas das microrregiões do Nordeste brasileiro, utilizou uma significância de estimativa de λ_{kij} para a construção de um perfil, da seguinte forma: se o λ_{kij} fosse maior que 1,3 (130%) vezes a respectiva frequência relativa marginal, então essa categoria pertencente à variável avaliada seria considerada como característica dos elementos do perfil extremo k. Porém, para este estudo utilizou-se um nível de 1,1 (110%), como valor de seleção, onde caracterizou a pertinência da categoria da variável ao perfil k. Entretanto, vale salientar que este procedimento não caracteriza completamente as unidades de análise (microrregiões), mas sim, os perfis extremos.

Diante do conhecimento da caracterização dos perfis extremos, estabeleceram-se os resultados relativos ao grau de pertinência, para que possa expandir o número de perfis, pois algumas microrregiões podem pertencer, parcialmente, a mais de um perfil extremo. A definição dos perfis mistos foi possível com o uso de expressões *booleanas*, disponíveis em Sawyer et al. (2000).

Com o intuito de adequar todas as microrregiões a algum perfil, dividiu-se os escores em quartis e considerou-se predominante do perfil k aqueles com 75% ou mais de características de um perfil k, ou seja $\{g_{1k} \geq 0,75\}$. Também, foram considerados como predominantes ao perfil k aqueles cujos escores situavam-se entre 50% e 74%, como mostra a expressão: $\{0,5 \leq g_{i1} < 0,75\} \cap \{g_{i2} < 0,25\} \cap \{g_{i3} < 0,25\}$.

Assim, expressões *booleanas* estabelecidas por Sawyer et al. (2000) definiram os tipos considerados mistos:

I. Perfil misto 1 e 2 (PM12) com predominância do perfil extremo 1:

$$\text{Se } \{0,5 \leq g_{i1} < 0,75\} \cap \{0,25 \leq g_{i2} < 0,5\} \cap \{g_{i3} < 0,25\}$$

II. Perfil misto 1 e 3 (PM13) com predominância do perfil extremo 1:

$$\text{Se } \{0,5 \leq g_{i1} < 0,75\} \cap \{g_{i2} < 0,25\} \cap \{0,25 \leq g_{i3} < 0,5\}$$

- III. Perfil misto 2 e 1 (PM21) com predominância do perfil extremo 2:
 Se $\{0,25 \leq g_{i1} < 0,5\} \cap \{0,5 \leq g_{i2} < 0,75\} \cap \{g_{i3} < 0,25\}$
- IV. Perfil misto 2 e 3 (PM23) com predominância do perfil extremo 2:
 Se $\{g_{i1} < 0,25\} \cap \{0,5 \leq g_{i2} < 0,75\} \cap \{0,25 \leq g_{i3} < 0,5\}$
- V. Perfil misto 3 e 1 (PM31) com predominância do perfil extremo 3:
 Se $\{0,25 \leq g_{i1} < 0,5\} \cap \{g_{i2} < 0,25\} \cap \{0,5 \leq g_{i3} < 0,75\}$
- VI. Perfil misto 3 e 2 (PM32) com predominância do perfil extremo 3:
 Se $\{g_{i1} < 0,25\} \cap \{0,25 \leq g_{i2} < 0,5\} \cap \{0,5 \leq g_{i3} < 0,75\}$

Foram considerados “amorfo”, ou sem definição, as microrregiões cujos escores de pertinência aos três perfis foram inferiores a 0,5 (Cerqueira, 2004).

A tabela 03 apresenta os resultados no processo de estimação dos λ_{kjl} , bem como as frequências marginais e respectivas razões entre estes, resultando nos fatores. Para a definição das características dominantes de cada perfil, definiu-se o valor indicado anteriormente (1,1), lembrando que a seleção desse valor também se baseou no grau de heterogeneidade que se pretende captar e quanto maior o seu valor mais restritivo será o processo de inclusão de categorias de determinada variável a um perfil.

O delineamento das características dos perfis encontrados no estudo, é apresentado a seguir, com base nas variáveis estudadas:

PERFIL 1 (Extremo)

- a) O perfil extremo 1, em relação à variável mortalidade infantil, apresenta um “elevado” índice (quartil 4). Com relação às demais variáveis, apresenta “elevada” proporção de pessoas com baixa renda (quartil 4); “baixa” idade média das mães a maternidade (quartil 1); “baixo” anos médio de estudo materno; “elevada” densidade domiciliar; “baixo” número de consultas pré-natal; “baixo” grau de urbanização; “baixa” proporção de domicílios com abastecimento de água adequado; “baixa” proporção de domicílios com existência de coleta de lixo adequada; “baixa” proporção de domicílios com existência de cobertura de

esgotamento sanitário; “baixa” renda média domiciliar per capita; “elevada” taxa de analfabetismo feminino. Ou seja, os piores indicadores socioeconômicos de condições de vida.

PERFIL 2 (Extremo)

b) O perfil extremo 2 é caracterizado por microrregiões que apresentam “baixa” taxa de mortalidade infantil (quartil 1), entretanto apresentam variáveis com comportamento mediano para mais e para menos, conforme apresentados a seguir: “mediana” proporção de pessoas com baixa renda (quartil 2 e 3); “mediana superior” idade média das mães a maternidade (quartil 3 e 4); “mediano inferior” anos médio de estudo materno; “mediana” densidade domiciliar; “mediano superior” número de consultas pré-natal; “mediano” grau de urbanização; “mediana” proporção de domicílios com abastecimento de água adequado; “mediana” proporção de domicílios com existência de coleta de lixo adequada; “mediana” proporção de domicílios com existência de cobertura de esgotamento sanitário; “mediana” renda média domiciliar per capita; “mediana” taxa de analfabetismo feminino. Resumidamente, este perfil encontra-se num perfil mediano de desenvolvimento de vida.

PERFIL 3 (Extremo)

c) O perfil extremo 3 é caracterizado como o perfil com melhores qualidades de vida, porém com a taxa de mortalidade infantil “mediana”. As demais variáveis apresentam-se neste perfil da seguinte maneira: “baixa” proporção de pessoas com baixa renda (quartil 1); “elevada” idade média das mães a maternidade (quartil 1); “elevado” anos médio de estudo materno; “baixa” densidade domiciliar; “mediano inferior” número de consultas pré-natal; “elevado” grau de urbanização; “elevada” proporção de domicílios com abastecimento de água adequado; “elevada” proporção de domicílios com existência de coleta de lixo adequada; “elevada” proporção de domicílios com existência de cobertura de esgotamento sanitário; “elevada” renda média domiciliar per capita; “baixa” taxa de analfabetismo feminino.

Tabela 03 – Frequências absolutas, relativas e estimativas de λ_{1j} , segundo perfis extremos.

VARIÁVEIS	Categorias	Frequência			Lambdas			Fatores		
		Abs.	%	Milhar	λ_{1j}	λ_{2j}	λ_{3j}	Perfil 1	Perfil 2	Perfil 3
TxMI10	< 14,35	47	25,0	0,25	0,164	0,309	0,223	0,6568	1,2364	0,8912
	14,35 – 16,4	47	25,0	0,25	0,313	0,227	0,234	1,2528	0,9064	0,9368
	16,41 – 18,2	47	25,0	0,25	0,245	0,224	0,301	0,9792	0,8964	1,2052
	> 18,2	47	25,0	0,25	0,278	0,240	0,242	1,1116	0,9608	0,9668
PrBxRen	< 58,86	47	25,0	0,25	0,000	0,000	1,000	0,0000	0,0000	4,0000
	58,87 – 65,8	48	25,5	0,255	0,000	0,516	0,000	0,0000	2,0227	0,0000
	65,81 – 70,21	45	23,9	0,239	0,000	0,484	0,000	0,0000	2,0259	0,0000
	> 70,21	48	25,5	0,255	1,000	0,000	0,000	3,9216	0,0000	0,0000
IdMdM	< 23,71	47	25,0	0,25	0,716	0,135	0,000	2,8640	0,5400	0,0000
	23,71 – 24,89	49	26,1	0,261	0,143	0,272	0,361	0,5460	1,0425	1,3843
	24,90 – 25,81	46	24,5	0,245	0,057	0,290	0,354	0,2310	1,1849	1,4445
	> 25,81	46	24,5	0,245	0,085	0,303	0,285	0,3469	1,2351	1,1624
AMdEst	< 5,5	48	25,5	0,255	0,306	0,364	0,000	1,2004	1,4286	0,0000
	5,5 – 6,43	46	24,5	0,245	0,565	0,147	0,130	2,3069	0,5988	0,5298
	6,44 – 7,04	48	25,5	0,255	0,000	0,367	0,283	0,0000	1,4380	1,1110
	> 7,04	46	24,5	0,245	0,129	0,122	0,587	0,5253	0,4992	2,3955
DensDom	< 3,46	50	26,6	0,266	0,000	0,250	0,558	0,0000	0,9410	2,0962
	3,46 – 3,58	49	26,1	0,261	0,050	0,381	0,246	0,1920	1,4613	0,9437
	3,59 – 3,75	43	22,9	0,229	0,000	0,368	0,196	0,0000	1,6083	0,8563
	> 3,75	46	24,5	0,245	0,950	0,000	0,000	3,8771	0,0000	0,0000
NuMdConsPn	< 5,68	47	25,0	0,25	0,909	0,000	0,126	3,6376	0,0000	0,5052
	5,68 – 6,12	48	25,5	0,255	0,091	0,260	0,398	0,3553	1,0200	1,5608
	6,13 – 6,52	48	25,5	0,255	0,000	0,368	0,272	0,0000	1,4435	1,0659
	> 6,52	45	23,9	0,239	0,000	0,372	0,204	0,0000	1,5556	0,8531
GrUrb	< 53,06	48	25,5	0,255	0,701	0,172	0,000	2,7490	0,6745	0,0000
	53,07 – 59,68	46	24,5	0,245	0,299	0,362	0,000	1,2204	1,4788	0,0000
	59,69 – 69,89	47	25,0	0,25	0,000	0,466	0,124	0,0000	1,8628	0,4948
	> 69,89	47	25,0	0,25	0,000	0,000	0,876	0,0000	0,0000	3,5052
CobertAgua	< 61,34	47	25,0	0,25	1,000	0,000	0,000	4,0000	0,0000	0,0000
	61,35 – 69,59	47	25,0	0,25	0,000	0,575	0,000	0,0000	2,2980	0,0000
	69,60 – 78,18	47	25,0	0,25	0,000	0,426	0,213	0,0000	1,7020	0,8524
	> 78,18	47	25,0	0,25	0,000	0,000	0,787	0,0000	0,0000	3,1476

VARIÁVEIS	Categorias	Frequência			Lambdas			Fatores		
		Abs.	%	Milhar	λ_{1j}	λ_{2j}	λ_{3j}	Perfil 1	Perfil 2	Perfil 3
CobertEsgS	< 89,44	47	25,0	0,25	1,000	0,000	0,000	4,0000	0,0000	0,0000
	89,45 – 93,47	47	25,0	0,25	0,000	0,350	0,299	0,0000	1,3988	1,1952
	93,48 – 95,71	49	26,1	0,261	0,000	0,399	0,240	0,0000	1,5284	0,9199
	> 95,71	45	23,9	0,239	0,000	0,251	0,461	0,0000	1,0519	1,9293
ExColLixo	< 53,88	47	25,0	0,25	0,984	0,000	0,000	3,9340	0,0000	0,0000
	53,89 – 64,63	47	25,0	0,25	0,017	0,524	0,000	0,0660	2,0944	0,0000
	64,64 – 74,12	47	25,0	0,25	0,000	0,476	0,091	0,0000	1,9056	0,3620
	> 74,12	47	25,0	0,25	0,000	0,000	0,910	0,0000	0,0000	3,6380
RendPc	< 249,43	47	25,0	0,25	1,000	0,000	0,000	4,0000	0,0000	0,0000
	249,44 – 284,5	47	25,0	0,25	0,000	0,500	0,000	0,0000	2,0000	0,0000
	284,51 – 338,2	47	25,0	0,25	0,000	0,500	0,000	0,0000	2,0000	0,0000
	> 338,2	47	25,0	0,25	0,000	0,000	1,000	0,0000	0,0000	4,0000
TxMdAnalf	< 18,73	47	25,0	0,25	0,000	0,000	0,914	0,0000	0,0000	3,6548
	18,74 – 22,05	47	25,0	0,25	0,201	0,373	0,086	0,8028	1,4932	0,3452
	22,06 – 25,70	49	26,1	0,261	0,085	0,509	0,000	0,3268	1,9487	0,0000
	> 25,70	45	23,9	0,239	0,714	0,118	0,000	2,9874	0,4941	0,0000

Fonte: Elaboração própria, com base nos resultados do GoM.

Estabelecidos os perfis extremos, faz-se necessário averiguar de forma aprofundada, a investigação das microrregiões que podem pertencer a mais de um perfil extremo, assim como a nenhum perfil, ou seja, microrregiões “amorfas”.

A tabela 04 apresenta os perfis extremos e mistos encontrados e as frequências absoluta e relativa, bem como uma descrição resumida de cada perfil.

Das 188 microrregiões estudadas, um total de 113, aproximadamente 60%, apresentou perfil de predominância extrema. O perfil que apresentou maior predominância entre as microrregiões foi o perfil 2 (baixa taxa de mortalidade infantil e medianas condições de vida) com 31,4% da frequência, seguido do perfil 3 (mediana taxa mortalidade infantil e melhores condições sociais de vida) com 18,1% e, por fim, o perfil 1 com (elevada taxa de mortalidade infantil e piores condições sociais de vida) com 10,6%.

As demais microrregiões apresentaram grau de pertinência a pelo menos dois tipos de perfis e apenas 9 (4,8%) apresentaram características voltadas para o grupo amorfo.

Das microrregiões que apresentaram características mistas, destacam-se o PM21 (misto, predominância de P2 com característica de P1) com 16%, o que significa que 30 microrregiões têm predominância para baixa mortalidade infantil e medianas condições sociais de vida, porém com características de elevadas taxas de mortalidade e piores condições sociais de vida. Fato de possível explicação deste resultado é uma combinação de condições de infraestrutura, condições sociais e econômicas destas microrregiões, que apresentam variáveis com bons índices em alguns segmentos e condições baixas em outros.

Seguindo a sequência de análise, o perfil PM23 (misto, predominância de P2 com características de P3) apresenta-se com 6,9%, um total de 13 microrregiões têm predominância de baixa taxa de mortalidade infantil e mediana condições sociais de vida, entretanto, com características de mediana mortalidade infantil e melhores condições sociais de vida.

De forma geral, os resultados apontam que 49,5% do total das microrregiões estão classificadas como detentoras da característica predominante dos tipos 2 e 3, que são aquelas que se caracterizam por apresentar indicadores mais modestos em termos de condições sociais de vida e média e baixa taxa de mortalidade infantil. Dessa forma, os resultados da utilização/aplicação do método *GoM*, permitem associar índices de melhores condições sociais de vida às taxas de mortalidade baixas. E essas características são encontradas principalmente nos perfis 2 e 3.

Tabela 04 – Distribuição das microrregiões segundo a predominância dos perfis extremos e mistos.

Predominância de Perfil	Frequência	%
Predominante 1 (<i>Elevada Mortalidade infantil + Piores condições sociais de vida</i>)	20	10,6
PM12 (Misto, predominância de 1 com característica de 2)	12	6,4
PM13 (Misto, predominância de 1 com característica de 3)	1	0,5
Total	33	17,6
Predominante 2 (<i>Baixa Mortalidade infantil + Mediana condições sociais de vida</i>)	59	31,4
PM21 (Misto, predominância de 2 com característica de 1)	30	16,0
PM23 (Misto, predominância de 2 com característica de 3)	13	6,9
Total	102	54,3
Predominante 3 (<i>Mediana Mortalidade infantil + Melhores condições de vida</i>)	34	18,1
PM31 (Misto, predominância de 3 com característica de 1)	9	4,8
PM32 (Misto, predominância de 3 com característica de 2)	1	0,5
Total	44	23,4
Amorfo	9	4,8
Total	188	100,0

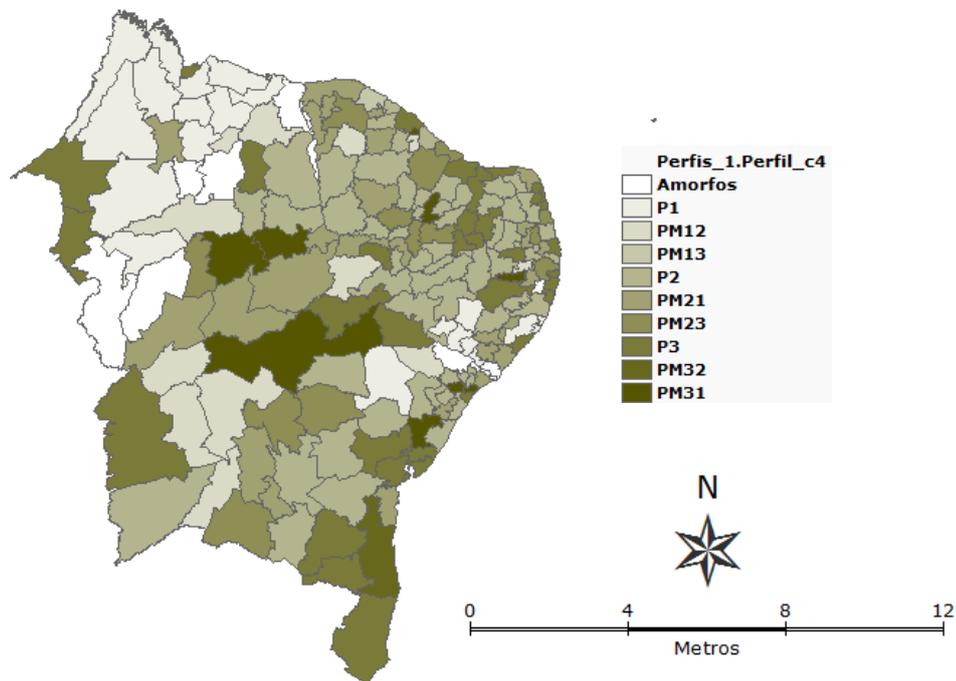
Fonte: Elaboração própria, com base nos resultados do GoM.

O mapa 02 apresenta a distribuição espacial da tipologia dos níveis de mortalidade infantil encontrado nas microrregiões associados às características das condições sociais.

Como comentado anteriormente, os resultados indicam uma ampla associação entre as altas taxas de mortalidade infantil e as piores condições sociais.

O padrão espacial encontrado a partir das tipologias mostra uma ampla diferenciação intra-regional, onde os perfis com características mais precárias tendem a se concentrar mais ao interior na região. As microrregiões situadas mais ao litoral apresentam perfis de melhores condições sociais. Os estados do Maranhão, Bahia, Piauí, Alagoas e Sergipe, apresentam microrregiões com perfil de precárias condições sociais (perfil 1). Além disso, o estado do Maranhão é aquele onde se concentra uma grande mancha associada ao perfil 1 (elevada taxa de mortalidade infantil e baixas condições sociais de vida). As microrregiões que apresentaram características “amorfas” estão situadas nos estados do Maranhão, Pernambuco, Sergipe, Alagoas e Piauí.

Mapa 02 – Microrregiões do Nordeste – Tipologias dos níveis de mortalidade infantil e condições socioeconômicas 2010.



Fonte: Elaboração própria, com base nos resultados do GoM.

Como forma de ilustrar um melhor detalhamento da distribuição das microrregiões por UF e seus perfis alocados, apresenta-se a tabela 05.

O estado do Maranhão apresenta, das 21 microrregiões, 12 (57%) com característica do perfil 1 (elevada mortalidade infantil e piores condições sociais de vida). Esta UF também apresenta 3 (14%) microrregiões que foram alocadas ao grupo de amorfos.

As demais UF apresentam grande parte de suas microrregiões situadas entre o perfil 2 e o perfil 3, perfis que foram caracterizados por exibirem média e alta condições de sociais de vida.

Outro destaque é encontrado nos estados de Alagoas e Maranhão, onde 46% e 57%, respectivamente, das suas microrregiões, foram alocadas às características de perfil 1.

Tabela 05 – Distribuição das microrregiões segundo a sua UF e perfil alocado.

UF / Microrregião		Perfis										Total
		Amorfos	P1	PM12	PM13	P2	PM21	PM23	P3	PM32	PM31	
Maranhão	Freq.	3	12	2	0	0	1	0	3	0	0	21
	%	14%	57%	10%	0%	0%	5%	0%	14%	0%	0%	100%
Piauí	Freq.	2	0	2	0	3	4	1	1	0	2	15
	%	13%	0%	13%	0%	20%	27%	7%	7%	0%	13%	100%
Ceará	Freq.	0	0	2	1	12	11	4	2	0	1	33
	%	0%	0%	6%	3%	36%	33%	12%	6%	0%	3%	100%
Rio Grande do Norte	Freq.	0	0	0	0	9	1	1	7	0	1	19
	%	0%	0%	0%	0%	47%	5%	5%	37%	0%	5%	100%
Paraíba	Freq.	0	0	1	0	12	3	3	4	0	0	23
	%	0%	0%	4%	0%	52%	13%	13%	17%	0%	0%	100%
Pernambuco	Freq.	1	1	1	0	7	1	1	6	0	1	19
	%	5%	5%	5%	0%	37%	5%	5%	32%	0%	5%	100%
Alagoas	Freq.	1	6	0	0	1	4	0	1	0	0	13
	%	8%	46%	0%	0%	8%	31%	0%	8%	0%	0%	100%
Sergipe	Freq.	2	0	0	0	6	2	0	1	0	2	13
	%	15%	0%	0%	0%	46%	15%	0%	8%	0%	15%	100%
Bahia	Freq.	0	1	4	0	9	3	3	9	1	2	32
	%	0%	3%	13%	0%	28%	9%	9%	28%	3%	6%	100%
Total	Freq.	9	20	12	1	59	30	13	34	1	9	188
	%	5%	11%	6%	1%	31%	16%	7%	18%	1%	5%	100%

Fonte: Elaboração própria, com base nos resultados do GoM.

Para melhor entendimento da classificação dos perfis por microrregião, os resultados no apêndice podem ser consultados.

4. DISCUSSÃO

O estudo das taxas de mortalidade infantil e suas variáveis preditoras é de suma importância que sejam realizados em níveis mais desagregados e, a cada dia, torna-se maior a exigência pelas mesmas, uma vez que metas de atenção à saúde materno infantil são estabelecidas e necessitam de subsídios para torná-las mais eficientes, em determinadas áreas.

Alguns resultados deste estudo, apontam para um claro desenvolvimento da região Nordeste, em termos de qualidade das informações, assim como de indicadores sociais e econômicos. Claro, ainda faltam grandes passos para se chegar a patamares como os das regiões Sul e Sudeste, mas percebe-se que a se

encaminha para tal. Os dados da região Norte e Nordeste do país, infelizmente devido ao seu lento desenvolvimento, ainda sofrem de questões relacionadas aos aspectos qualitativos e quantitativos. Entretanto, o avanço das técnicas demográficas e estatísticas, está possibilitando conhecer melhor os fenômenos relacionados a estas regiões.

A análise de regressão linear múltipla ponderada, também veio somar com seus resultados, apontando para o que já vem sendo apresentado em vários estudos, comprovando a estreita relação das taxas de mortalidade infantil na região Nordeste com níveis educacionais maternos, econômicos familiares e domiciliares.

Da análise de perfis descritivos, encontraram-se regiões com extrema precariedade, que fazem correlação com altas taxas de mortalidade infantil. Destaque maior para o estado do Maranhão, em quase sua totalidade de variáveis analisadas: esta UF apresentou indicadores com médias e altas precariedades. Tal fato recai num ponto de alerta, necessitando de uma atenção especial por parte das autoridades competentes. O elemento considerado como ponto de atenção também está voltado para a microrregião de Fernando de Noronha (PE), que é uma região com desenvolvimento social de vida elevado, porém apresentando uma alta taxa de mortalidade infantil, devido a questões de pequenos números envolvidos no cálculo (nascidos vivos e óbitos de menores de 1 ano).

Os resultados do método *GoM* também possibilitaram conhecer melhor e identificar a relação das áreas estudadas com suas variáveis analisadas. Ponto a se destacar é a possibilidade de inclusão de várias outras variáveis na análise, podendo-se ter um panorama mais amplo dos perfis destas regiões. Entretanto, recai-se na questão da qualidade das informações, as quais precisam ser cuidadosamente investigadas e validadas para que possam ser empregadas no modelo.

5. CONSIDERAÇÕES FINAIS

Do exposto, fica evidente, a partir das informações e resultados apresentados, a significativa melhoria na qualidade das informações referentes aos óbitos infantis para as microrregiões no Nordeste brasileiro, a partir da análise das causas mal definidas, para os anos de 2000 a 2010.

A estimação e correção do sub-registro se mostraram consistentes, à medida que foi possível fazer a suavização proposta pelo método, contribuindo para uma estimativa da taxa de mortalidade infantil, possível de ser considerada como corrigida.

Além disso, a significativa redução na mortalidade infantil, não só no Brasil, mas também na região Nordeste, associadas às melhorias nas condições sociais da população, continua sendo ponto marcante na transição da estrutura da população. Entretanto, a região Nordeste ainda necessita de maiores investimentos para se chegar a melhores níveis de qualidade nas condições sociais, econômicas, estruturais e assistenciais.

Dos dados selecionados, a partir da modelagem da regressão linear múltipla ponderada, buscando a associação (explicação) da mortalidade infantil para o ano de 2010, ficou evidenciada a relação com a educação materna, renda per capita, densidade domiciliar e número médio de consultas pré-natal, como sendo variáveis de maiores contribuições no comportamento da mortalidade infantil.

Destaca-se, o comportamento de homogeneização das microrregiões com relação às taxas de mortalidade infantil e às variáveis envolvidas. Com a Análise de *Clusters* (preliminar), foi possível verificar um grupo maior (dominante), talvez por efeito das melhorias de acesso a melhores condições de vida da população nordestina, através das transferências de renda do governo federal como o Bolsa família. À medida que a mortalidade infantil vem em trajetória de declínio, as microrregiões tenderão a apresentar um comportamento de estabilidade mínima.

Por fim, pode-se concluir, de modo geral, através dos resultados da aplicação do método *GoM*, uma clara associação entre as altas taxas de mortalidade infantil com as piores condições socioeconômicas, assim como a participação de microrregiões em outros perfis (mistos), caracterizando a composição real da região, de forma a identificar o impacto de cada variável nas regiões estudadas. Tal fato possibilitou conhecer a real imagem da área em estudo e, também, das variáveis envolvidas.

Desta forma, os objetivos gerais e específicos do estudo, foram alcançados com êxito, contribuindo para o conhecimento do comportamento da mortalidade frente às condições atuais das microrregiões do Nordeste Brasileiro, podendo servir

de importante subsídio para a elaboração de políticas públicas voltadas para a população infantil na região.

6. REFERÊNCIAS

ASSUNÇÃO, R. M.; BARRETO, S.A.; GUERRA, H. L.; Sakurai, E. Figuras de taxas epidemiológicas: uma abordagem Bayesiana. **Cadernos de Saúde Pública**, 14: 713-723, 1998.

BARROS, R. P. de. et. al. **Determinantes do desenvolvimento na primeira infância**. Texto para discussão nº 1478. Brasília: Instituto de Pesquisa Econômica Aplicada. Março de 2010.

BARBOSA, Lára de Melo (Org.). **Questões contemporâneas da dinâmica populacional do Nordeste**. Natal: Edufrn, 2006. 254 p.

BELTRÃO, I. K.; SAWYER, D. O. **Medidas de mortalidade**: um estudo sobre os efeitos das mudanças da escolaridade da mãe e da estrutura de fecundidade em quatro áreas brasileiras. Rio de Janeiro: ENCE/IBGE, 1990. Relatórios Técnicos, 4. 52 p.

BIRD, Sheryl Thorburn & Bauman, Karl E. State-Level infant, neonatal, and postneonatal mortality: the contribution of selected structural socioeconomic variables. Intern. *Journal of Health*, 28 (1): 13-27, 1998.

CAVALINI, L.T.; Leon, A. C. P. Correção de sub-registros de óbitos e proporção de internações por causas mal definidas. *Rev. De Saúde Pública. FSP*, (1): 85-93, 2007.

CAVALINI, L.T. Morbi-mortalidade nos municípios brasileiros: associação com indicadores sócio-econômicos e de assistência à saúde. Um estudo multinível (Tese Doutorado). Instituto de Medicina Social – IMS/UERJ, Rio de Janeiro, 2005.

CERQUEIRA, C. A. Uma tipologia dos municípios da região Nordeste. Belo Horizonte, 2002. Mimeografado

CERQUEIRA, C. A. Tipologia e características dos estabelecimentos escolares brasileiros. Tese de doutorado em Demografia. Belo Horizonte: Cedeplar, 2004.

CAETANO, A. J.; MACHADO, C. J. Consistência e identificabilidade no modelo grade of membership: uma nota metodológica. **Rebep**. Rio de Janeiro, ano 2009, v. 26, n. 1, 01 jun. 2009.

CARVALHO, J. A. M.; SAWYER, D.R.T.O. **Diferenciais de mortalidade no Brasil**. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 1, 1978, Campos do Jordão. **Anais...** Campos do Jordão, 1978, v. 1. p. 231-259.

CHACKIEL J. Studies of Causes of Death in Latin America Current Situation and Future Perspectives. In: Committee on Comparative Mortality Changes: proceedings of the IUSSP and Institute of Statistics University of Siena; 1986 Jul 7-12; Siena, Itália.

EFRON, B.; MORRIS, C. Data analysis using Stein's estimation rule and its competitors: an empirical Bayes approach. *J Am Stat Assoc.*;70:311-9, 1975.

FORMIGA, M.C. de C., RAMOS, P.C.F., MONTEIRO, M. F. G. **Fatores socioeconômicos e mortalidade: uma associação observada nas mesorregiões do Rio Grande do Norte**. ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 12, 2000, Caxambú, Minas Gerais, **Anais...**, CD-ROM.

FORMIGA, M. C. de C. **O efeito da desigualdade socioeconômica na sobrevivência de crianças menores de 5 anos de idade no estado do Rio Grande do Norte**. 2003. 314 f. Tese (Doutorado) - Curso de Saúde Coletiva, Departamento de Instituto De Medicina Social, Universidade Do Estado Do Rio De Janeiro, Rio de Janeiro, 2003.

FORMIGA, M.C. de C., RAMOS, P.C.F. **Influência dos Fatores Socioeconômicos Maternos na Sobrevivência de Crianças Menores de 5 Anos de Idade no Estado do Rio Grande do Norte**. In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS POPULACIONAIS, 13, 2002, Ouro Preto, Minas Gerais, **Anais...**, Minas Gerais: ABEP, 2002.

FERREIRA, C. E. C. Os efeitos da queda da fecundidade sobre a mortalidade infantil em São Paulo. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 7, 1990, Caxambu. **Anais...** Belo Horizonte: ABEP, p. 435-447, 1990.

FERNANDES, D. M. **Diferenciais de mortalidade segundo instrução: regiões metropolitanas - Brasil - 1970**. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 4, 1984, Águas de São Pedro. **Anais...** São Paulo: ABEP, 1984. v. 2, p. 643-660.

FREIRE, F. H. M. de A.. **Projeção populacional para pequenas áreas pelo método das componentes demográficas usando estimadores bayesianos espaciais**. 2001. 100 f. Tese (Doutorado) - Cedeplar, Belo Horizonte, MG, 2001.

FREIRE, F. H. M. A; ASSUNÇÃO, R. M. Estimadores bayesianos para níveis de mortalidade e fecundidade em pequenas áreas. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 15, 2002, Águas de Lindóia, SP, Resumos... São Paulo, SP, ABE, v. 1, 2002

HALLI, S. S., RAO, K. V. Advanced techniques of population analysis. New York, Plenum Press, 1992. p. 63-71.

HERNANDEZ, A. R. et al. Análise de tendências das taxas de mortalidade infantil e de seus fatores de risco na cidade de Porto Alegre, Rio Grande do Sul, Brasil, no período de 1996 a 2008. **Caderno de Saúde Pública**, Rio de Janeiro, v. 11, n. 27, p.188-2196, nov. 2011.

IBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 1999. *Evolução e perspectivas da mortalidade infantil no Brasil*. <<http://www.ibge.gov.br>>

IBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 2001. *População. Indicadores Sociais*. <<http://www.ibge.gov.br>>

IPEA. Instituto de Pesquisa Econômica Aplicada. **Determinantes do desenvolvimento na primeira infância no Brasil**. Brasília, 2010. <<http://www.ipea.gov.br/portal/>>

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 2 ed., London: Prentice-Hall, 1988. p. 554-559.

MANTON, K. G.; WOODBURY, M. A.; TOLLEY, H. D. **Statistical applications using fuzzy sets**. First edition. New York: Willey Interscience, 1994.

MELLO-JORGE, M. H. P., 1983. Sub-registro dos eventos vitais. *Revista de Saúde Pública*, 17:148-151.

MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada**. UFMG. ISBN: 857041451X.

MONTEIRO, M. F. G. **O efeito da educação materna sobre o risco da mortalidade infantil**. In: *Rev. Bras. Estudos de População*, v. 7. n. 1, p. 74-85, jan/jun 1990.

NACIONES UNIDAS. **Manual X: Técnicas indiretas de estimación demográfica**. New York, 1986. (Estudios de población, 81). 318 p.

PAES, N. A. Qualidade das estatísticas de óbitos por causas desconhecidas dos Estados brasileiros. **Revista de saúde pública**. São Paulo, ano 2007, v. 41, n. 3, 2007.

LUCIO, P. S. – Notas de Aula da disciplina “Análise de Regressão usando o R”, Universidade Federal do Rio Grande do Norte, CCET/DEST – Natal, 2007.

ROMERO, D. E.; SWARCWALD, C. L. Economic crisis and infant mortality in Latin America since the 1980's. **Cad. Saúde Pública**. v. 3, n. 16, p. 799-814, 2000.

RIOS NETO, E.L.G; MOREIRA, M.R. **Crise econômica e seus impactos na mortalidade infantil**: um estudo comparativo das diferenças sócio-econômicas regionais. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 9, 1994, Caxambú. **Anais...** Belo Horizonte: ABEP, 1994. v. 1, p.99-118.

SANTO, A. H. Equivalência entre revisões da Classificação Internacional de Doenças: causas de morte. **Revista de Saúde Pública**, São Paulo, v.34, n.1, p. 21-8, 2000.

SAWYER, D. O. Mortalidade. In: SANTOS, J.L.F. et al. (Org.), **Dinâmica da População**: teoria métodos e técnicas de análise. p. 209-235. São Paulo, T.A. Queiroz, 1980.

SAWYER, D. O. et al. Caracterização dos tipos de doadores de sangue em Belo Horizonte: heterogeneidade do homogêneo. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS,12, 2000, CAXAMBU. **Anais...** Belo Horizonte: ABEP, 2000 (CD-ROM).

SIMÕES, C. C. da S. Perfis de saúde e de mortalidade no Brasil: uma análise de seus condicionantes em grupos populacionais específicos. 1. ed. Brasília: Organização Pan-americana da Saúde, 2002. 141 p.

SIMÕES, C. C. S. **Estimativas da mortalidade infantil para os estados das regiões Norte e Nordeste por distintos níveis de desagregação espacial**: as limitações nas informações básicas e possíveis metodologias. In: Encontro sobre População do Norte/Nordeste, 1, 1998, Recife. (Digitado). Recife, Fundação Joaquim Nabuco, Associação Brasileira de Estudos Populacionais,1998. p.46.

SZWARCWALD, C. L. et al. Estimação da mortalidade infantil no Brasil: o que dizem as informações sobre óbitos e nascimentos do Ministério da Saúde? **Cad. Saúde Pública**, Rio de Janeiro, v. 6, n. 18, p.1725-1736, 1 dez. 2002.

SILVA, M. G. C. da. **Técnicas para definir prioridades em saúde**: análise da mortalidade por causas em Fortaleza, Fortaleza: Edições IOCE, 1991. 225 p.

TERRA DE SOUZA, et al. Variations in infant mortality rates among municipalities in the state of Ceara, Northeast Brazil: an ecological analysis. **Int. J. Epidemiology**, v. 2, n. 28, p 267-275, 1999.

VERMELHO, L. L.; MONTEIRO, M. F. G. Transição demográfica e epidemiológica. In: MEDRONHO, R. A., et al. (Ed.). **Epidemiologia**. São Paulo, Atheneu, 2002, cap. 6, p. 91-103.

WOOD, C.H. e CARVALHO, J.A.M. de. **A demografia da desigualdade no Brasil**, Rio de Janeiro, IPEA,1994.

ZADEH, L. A. Fuzzy sets: information and control, v. 8, p. 338-353, 1965.

APÊNDICE A – Quadro 07. Grupos de microrregiões geradas com a análise de *cluster* e perfis encontrados com a utilização do Método *GoM*, região Nordeste, Brasil, 2010.

Código Microrregião	Nomes das Microrregiões	Classificação dos grupos (Clusters)	Classificação dos perfis (GoM)
210101	LITORAL OCIDENTAL MARANHENSE	11	Perfil 1
210102	AGLOMERACAO URBANA DE SAO LUIS	6	Perfil 3
210103	ROSARIO	11	Perfil 1
210104	LENCOIS MARANHENSES	11	Perfil 1
210105	BAIXADA MARANHENSE	11	Perfil 1
210106	ITAPECURU MIRIM	11	Perfil 1
210207	GURUPI	11	Perfil 1
210208	PINDARE	11	Perfil 1
210209	IMPERATRIZ	11	Perfil 3
210310	MEDIO MEARIM	11	Perfil Misto 21
210311	ALTO MEARIM E GRAJAU	11	Perfil 1
210312	PRESIDENTE DUTRA	11	Amorfos
210413	BAIXO PARNAIBA MARANHENSE	11	Perfil 1
210414	CHAPADINHA	11	Perfil 1
210415	CODO	11	Perfil 1
210416	COELHO NETO	11	Perfil Misto 12
210417	CAXIAS	11	Amorfos
210418	CHAPADAS DO ALTO ITAPECURU	11	Perfil Misto 12
210519	PORTO FRANCO	11	Perfil 3
210520	GERAIS DE BALSAS	11	Amorfos
210521	CHAPADAS DAS MANGABEIRAS	11	Perfil 1
220101	BAIXO PARNAIBA PIAUIENSE	11	Perfil Misto 12
220102	LITORAL PIAUIENSE	11	Amorfos
220203	TERESINA	5	Perfil 3
220204	CAMPO MAIOR	11	Perfil 2
220205	MEDIO PARNAIBA PIAUIENSE	11	Perfil 2
220206	VALENCA DO PIAUI	11	Perfil 2
220307	ALTO PARNAIBA PIAUIENSE	11	Amorfos
220308	BERTOLINIA	11	Perfil Misto 23
220309	FLORIANO	11	Perfil Misto 31
220310	ALTO MEDIO GURGUEIA	11	Perfil Misto 21
220311	SAO RAIMUNDO NONATO	11	Perfil Misto 21
220312	CHAPADAS DO EXTREMO SUL PIAUIENSE	11	Perfil Misto 12
220413	PICOS	11	Perfil Misto 31
220414	PIO IX	11	Perfil Misto 21
220415	ALTO MEDIO CANINDE	11	Perfil Misto 21
230101	LITORAL DE CAMOCIM E ACARAU	11	Perfil Misto 21
230102	IBIAPABA	11	Perfil Misto 21
230103	COREAU	11	Perfil Misto 21
230104	MERUOCA	11	Perfil Misto 21

Código Microrregião	Nomes das Microrregiões	Classificação dos grupos (Clusters)	Classificação dos perfis (GoM)
230105	SOBRAL	11	Perfil Misto 23
230106	IPU	11	Perfil Misto 21
230107	SANTA QUITERIA	11	Perfil Misto 12
230208	ITAPIOCA	11	Perfil Misto 13
230209	BAIXO CURU	11	Perfil 2
230210	URUBURETAMA	11	Perfil Misto 21
230211	MEDIO CURU	11	Perfil 2
230212	CANINDE	11	Perfil Misto 21
230213	BATURITE	11	Perfil 2
230214	CHOROZINHO	11	Perfil Misto 12
230215	CASCADEL	11	Perfil 2
230317	FORTALEZA	6	Perfil 3
230418	PACAJUS	11	Perfil Misto 31
230419	SERTAO DE CRATEUS	11	Perfil 2
230420	SERTAO DE QUIXERAMOBIM	11	Perfil 2
230421	SERTAO DE INHAMUNS	11	Perfil 2
230522	SERTAO DE SENADOR POMPEU	11	Perfil Misto 21
230523	LITORAL DE ARACATI	11	Perfil Misto 23
230524	BAIXO JAGUARIBE	11	Perfil Misto 23
230525	MEDIO JAGUARIBE	11	Perfil 2
230626	SERRA DO PEREIRO	11	Perfil 2
230627	IGUATU	11	Perfil Misto 23
230628	VARZEA ALEGRE	11	Perfil 2
230729	LAVRAS DA MANGABEIRA	11	Perfil 2
230730	CHAPADA DO ARARIPE	11	Perfil Misto 21
230731	CARIRIACU	11	Perfil Misto 21
230732	BARRO	11	Perfil Misto 21
230733	CARIRI	11	Perfil 3
240101	BREJO SANTO	11	Perfil 2
240102	MOSSORO	7	Perfil 3
240103	CHAPADA DO APODI	11	Perfil 2
240104	MEDIO OESTE	11	Perfil 2
240105	VALE DO ACU	11	Perfil 3
240106	SERRA DE SAO MIGUEL	11	Perfil 2
240107	PAU DOS FERROS	11	Perfil Misto 31
240208	UMARIZAL	11	Perfil 2
240209	MACAU	11	Perfil 3
240210	ANGICOS	11	Perfil 2
240211	SERRA DE SANTANA	11	Perfil 2
240212	SERIDO OCIDENTAL	8	Perfil 3
240313	SERIDO ORIENTAL	10	Perfil 3
240314	BAIXA VERDE	11	Perfil 2
240315	BORBOREMA POTIGUAR	11	Perfil 2

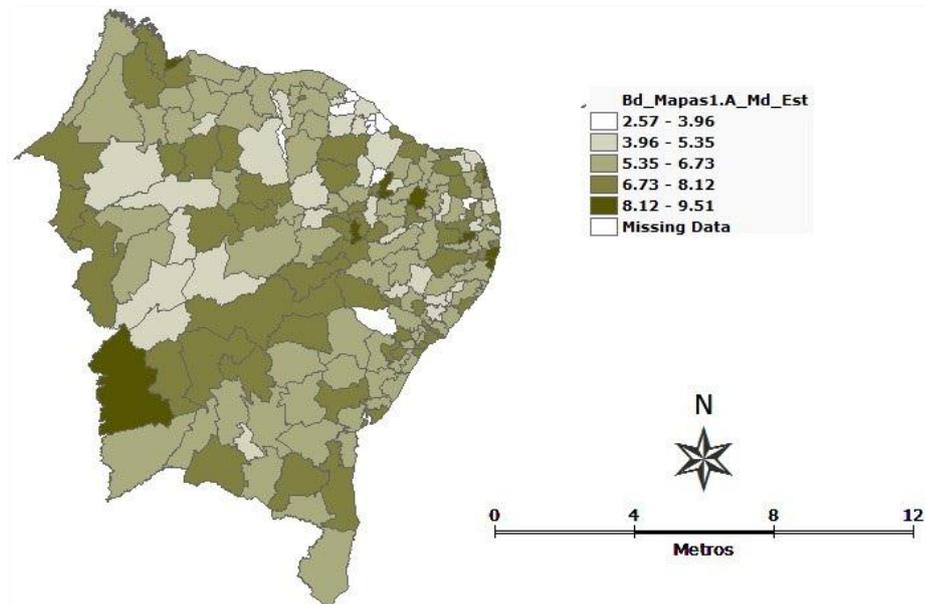
Código Microrregião	Nomes das Microrregiões	Classificação dos grupos (Clusters)	Classificação dos perfis (GoM)
240416	AGRESTE POTIGUAR	11	Perfil 2
240417	LITORAL NORDESTE	11	Perfil Misto 21
240419	MACAIBA	11	Perfil 3
250101	NATAL	4	Perfil 3
250102	LITORAL SUL	11	Perfil Misto 23
250103	CATOLE DO ROCHA	11	Perfil 2
250104	CAJAZEIRAS	11	Perfil Misto 23
250105	SOUSA	11	Perfil Misto 23
250106	PATOS	11	Perfil 3
250107	PIANCO	11	Perfil 2
250208	ITAPORANGA	11	Perfil 2
250209	SERRA DO TEIXEIRA	11	Perfil 2
250210	SERIDO OCIDENTAL PARAIBANO	11	Perfil 3
250211	SERIDO ORIENTAL PARAIBANO	11	Perfil 2
250312	CARIRI OCIDENTAL	11	Perfil 2
250313	CARIRI ORIENTAL	11	Perfil 2
250314	CURIMATAU OCIDENTAL	11	Perfil 2
250315	CURIMATAU ORIENTAL	11	Perfil Misto 21
250316	ESPERANCA	11	Perfil 2
250317	BREJO PARAIBANO	11	Perfil Misto 21
250318	GUARABIRA	11	Perfil Misto 23
250319	CAMPINA GRANDE	7	Perfil 3
250420	ITABAIANA	11	Perfil 2
250421	UMBUZEIRO	11	Perfil Misto 12
250423	LITORAL NORTE	11	Perfil 2
260101	SAPE	11	Perfil Misto 21
260102	JOAO PESSOA	2	Perfil 3
260103	LITORAL SUL	11	Perfil 2
260104	ARARIPINA	11	Perfil Misto 12
260205	SALGUEIRO	11	Perfil 2
260206	PAJEU	11	Perfil 2
260308	SERTAO DO MOXOTO	11	Perfil 2
260309	PETROLINA	9	Perfil 3
260310	ITAPARICA	11	Perfil 2
260311	VALE DO IPANEMA	11	Perfil 1
260312	VALE DO IPOJUCA	11	Perfil 3
260413	ALTO CAPIBARIBE	11	Perfil Misto 31
260414	MEDIO CAPIBARIBE	11	Perfil Misto 21
260415	GARANHUNS	11	Perfil 2
260516	BREJO PERNAMBUCANO	11	Perfil 2
260518	MATA SETENTRIONAL PERNAMBUCANA	11	Perfil Misto 23
270101	VITORIA DE SANTO ANTAO	11	Amorfos
270102	MATA MERIDIONAL PERNAMBUCANA	11	Perfil 2

Código Microrregião	Nomes das Microrregiões	Classificação dos grupos (Clusters)	Classificação dos perfis (GoM)
270103	ITAMARACA	11	Perfil 3
270104	RECIFE	2	Perfil 3
270205	SUAPE	10	Perfil 3
270206	FERNANDO DE NORONHA	1	Perfil 3
270308	SERRANA DO SERTAO ALAGOANO	11	Perfil 1
270309	ALAGOANA DO SERTAO DO SAO FRANCISCO	11	Perfil 2
270310	SANTANA DO IPANEMA	11	Perfil 1
270311	BATALHA	11	Perfil 1
270312	PALMEIRA DOS INDIOS	11	Perfil Misto 21
270313	ARAPIRACA	11	Perfil Misto 21
280101	TRAIPIU	11	Perfil 1
280203	SERRANA DOS QUILOMBOS	11	Perfil Misto 21
280204	MATA ALAGOANA	11	Perfil 1
280205	LITORAL NORTE ALAGOANO	11	Perfil 1
280206	MACEIO	6	Perfil 3
280307	SAO MIGUEL DOS CAMPOS	11	Perfil Misto 21
280308	PENEDO	11	Amorfos
280309	SERGIPANA DO SERTAO DO SAO FRANCISCO	11	Amorfos
280310	CARIRA	11	Perfil 2
280312	NOSSA SENHORA DAS DORES	11	Perfil 2
280313	AGRESTE DE ITABAIANA	11	Perfil Misto 31
290101	TOBIAS BARRETO	11	Perfil 2
290102	AGRESTE DE LAGARTO	11	Perfil 2
290103	PROPRIA	11	Amorfos
290204	COTINGUIBA	11	Perfil 2
290205	JAPARATUBA	11	Perfil Misto 21
290206	BAIXO COTINGUIBA	11	Perfil Misto 31
290207	ARACAJU	2	Perfil 3
290308	BOQUIM	11	Perfil Misto 21
290309	ESTANCIA	11	Perfil 2
290310	BARREIRAS	7	Perfil 3
290311	COTEGIPE	11	Perfil Misto 12
290312	SANTA MARIA DA VITORIA	11	Perfil 2
290413	JUAZEIRO	11	Perfil Misto 31
290414	PAULO AFONSO	10	Perfil 3
290415	BARRA	11	Perfil Misto 12
290416	BOM JESUS DA LAPA	11	Perfil Misto 12
290417	SENHOR DO BONFIM	11	Perfil 2
290418	IRECE	11	Perfil Misto 23
290519	JACOBINA	11	Perfil Misto 23
290520	ITABERABA	11	Perfil 2
290622	FEIRA DE SANTANA	9	Perfil 3

Código Microrregião	Nomes das Microrregiões	Classificação dos grupos (Clusters)	Classificação dos perfis (GoM)
290623	JEREMOABO	11	Perfil Misto 12
290624	EUCLIDES DA CUNHA	11	Perfil 1
290625	RIBEIRA DO POMBAL	11	Perfil 2
290626	SERRINHA	11	Perfil 2
290627	ALAGOINHAS	10	Perfil Misto 31
290628	ENTRE RIOS	11	Perfil 2
290629	CATU	11	Perfil 3
290730	SANTO ANTONIO DE JESUS	10	Perfil 3
290731	SALVADOR	3	Perfil 3
290732	BOQUIRA	11	Perfil Misto 21
230316	SEABRA	11	Perfil 2
250422	JEQUIE	11	Perfil 2
260517	LIVRAMENTO DO BRUMADO	11	Perfil Misto 21
280311	GUANAMBI	11	Perfil Misto 23
240418	BRUMADO	11	Perfil 2
260307	VITORIA DA CONQUISTA	11	Perfil 3
260519	ITAPETINGA	11	Perfil 3
270207	VALENCA	11	Perfil Misto 21
280102	ILHEUS-ITABUNA	10	Perfil Misto 32
290521	PORTO SEGURO	9	Perfil 3

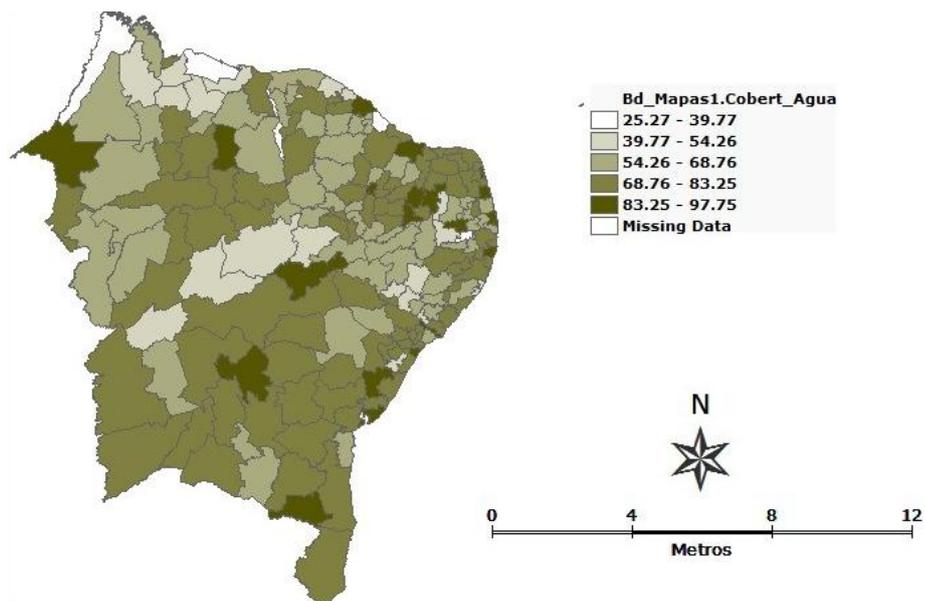
Fonte: Elaboração própria com base nos resultados da análise de Clusters.

APÊNDICE C – Mapa 03. Distribuição das microrregiões segundo a variável “Anos médio de estudo materno quando notificado o óbito infantil”, Região Nordeste, Brasil, 2010.



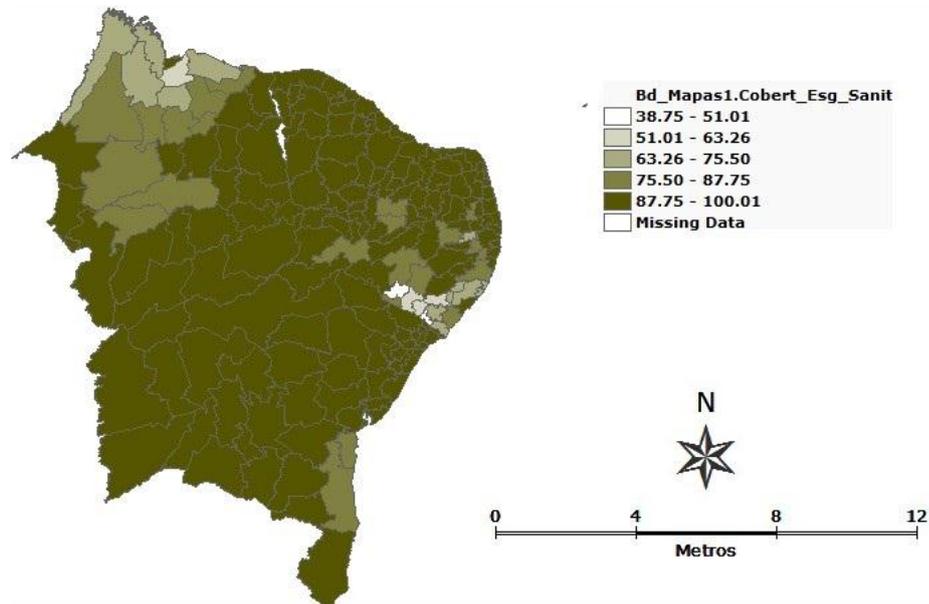
Fonte: Elaboração própria com base nos dados do DATASUS/MS, 2010.

APÊNDICE D – Mapa 04. Distribuição das microrregiões segundo a variável “Proporção de domicílios com abastecimento de água adequado da microrregião”, Região Nordeste, Brasil, 2010.



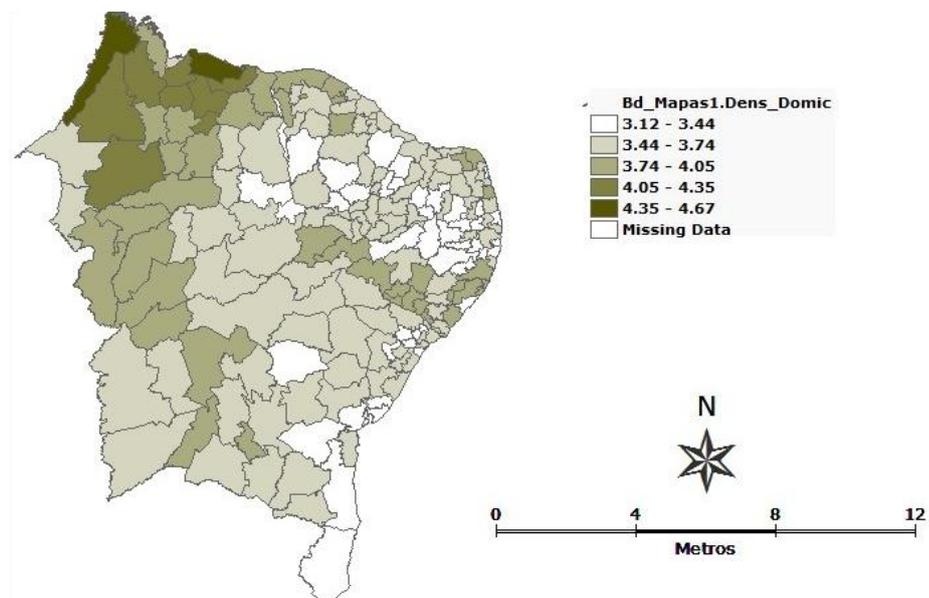
Fonte: Elaboração própria com base nos dados microdados do censo/IBGE, 2010.

APÊNDICE E – Mapa 05. Distribuição das microrregiões segundo a variável “Proporção de domicílios com existência de cobertura de esgotamento sanitário”, Região Nordeste, Brasil, 2010.



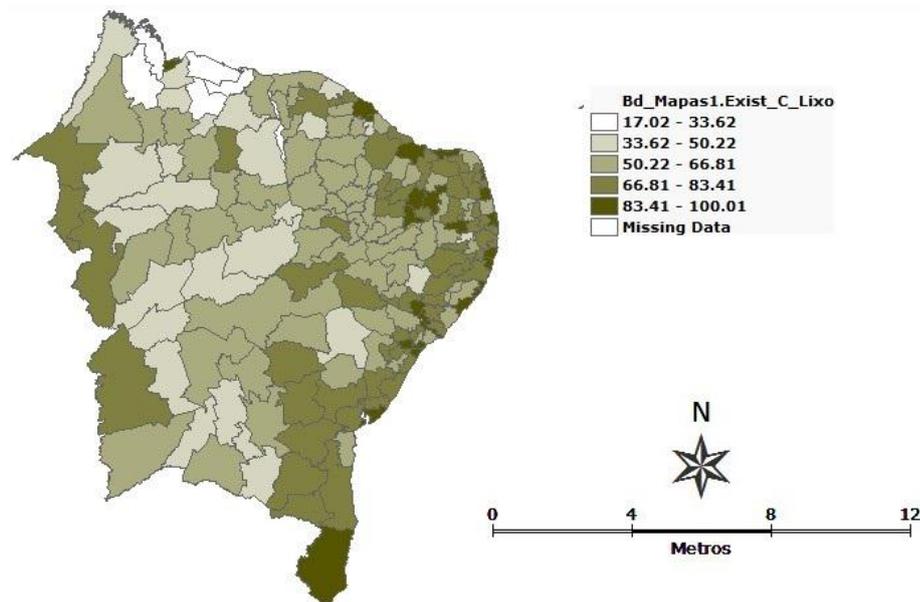
Fonte: Elaboração própria com base nos dados microdados do censo/IBGE, 2010.

APÊNDICE F – Mapa 06. Distribuição das microrregiões segundo a variável “Densidade domiciliar”, Região Nordeste, Brasil, 2010.



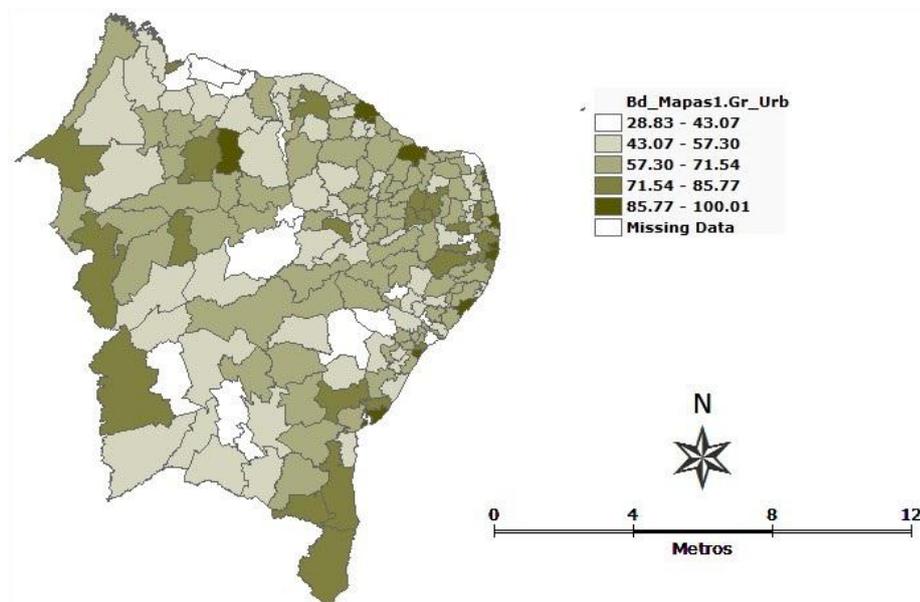
Fonte: Elaboração própria com base nos dados microdados do censo/IBGE, 2010.

APÊNDICE G – Mapa 07. Distribuição das microrregiões segundo a variável “Proporção de domicílio com existência de coleta de lixo adequada”, Região Nordeste, Brasil, 2010.



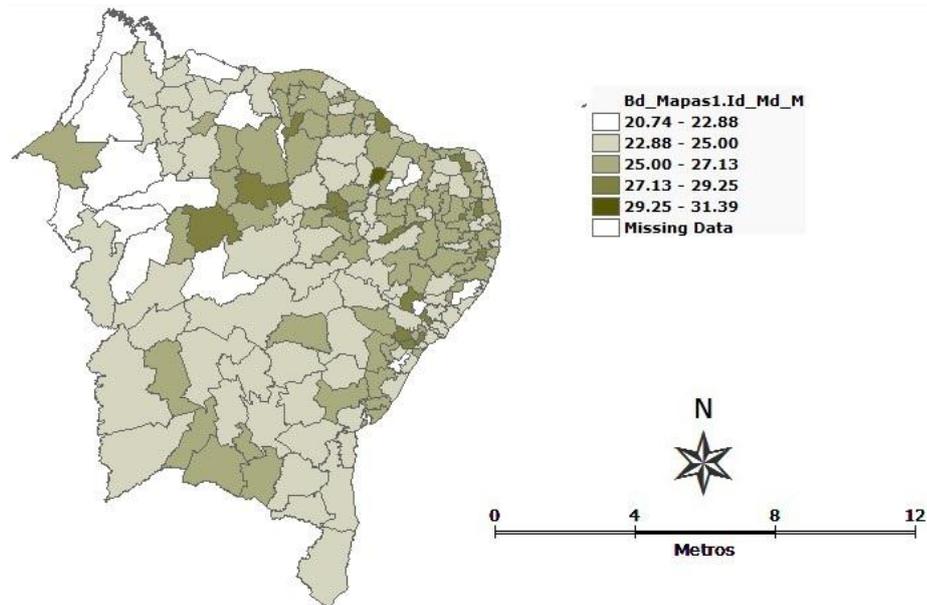
Fonte: Elaboração própria com base nos dados microdados do censo/IBGE, 2010.

APÊNDICE H – Mapa 08. Distribuição das microrregiões segundo a variável “Grau de Urbanização”, Região Nordeste, Brasil, 2010.



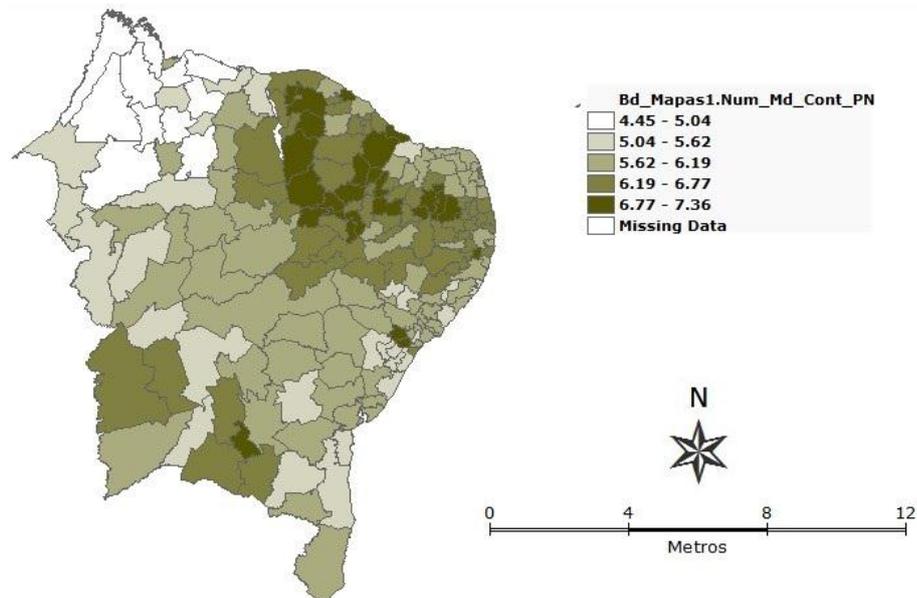
Fonte: Elaboração própria com base nos dados microdados do censo/IBGE, 2010.

APÊNDICE I – Mapa 09. Distribuição das microrregiões segundo a variável “Idade média das mães quando notificado o óbito infantil”, Região Nordeste, Brasil, 2010.



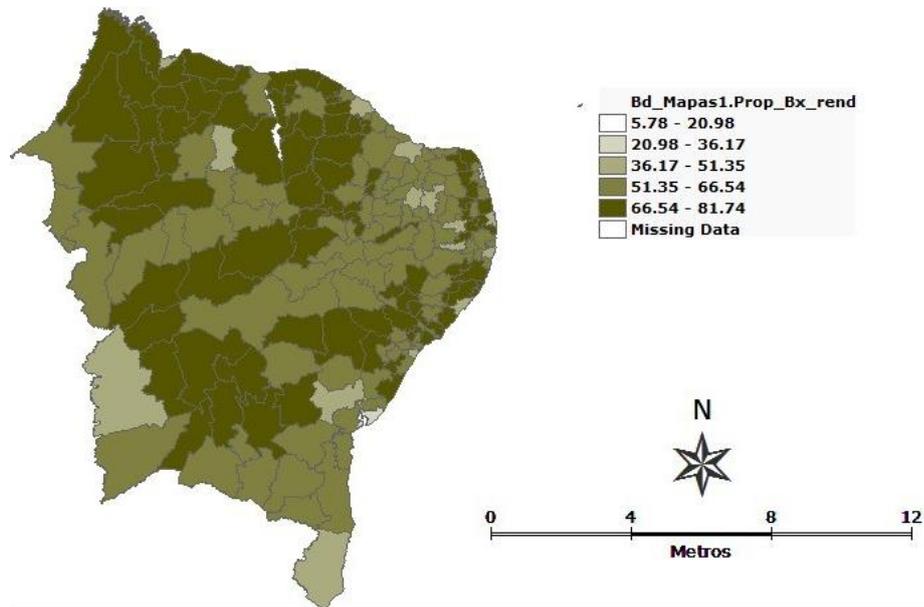
Fonte: Elaboração própria com base nos dados do DATASUS/MS, 2010.

APÊNDICE J – Mapa 10. Distribuição das microrregiões segundo a variável “Número médio de consultas Pré-Natal”, Região Nordeste, Brasil, 2010.



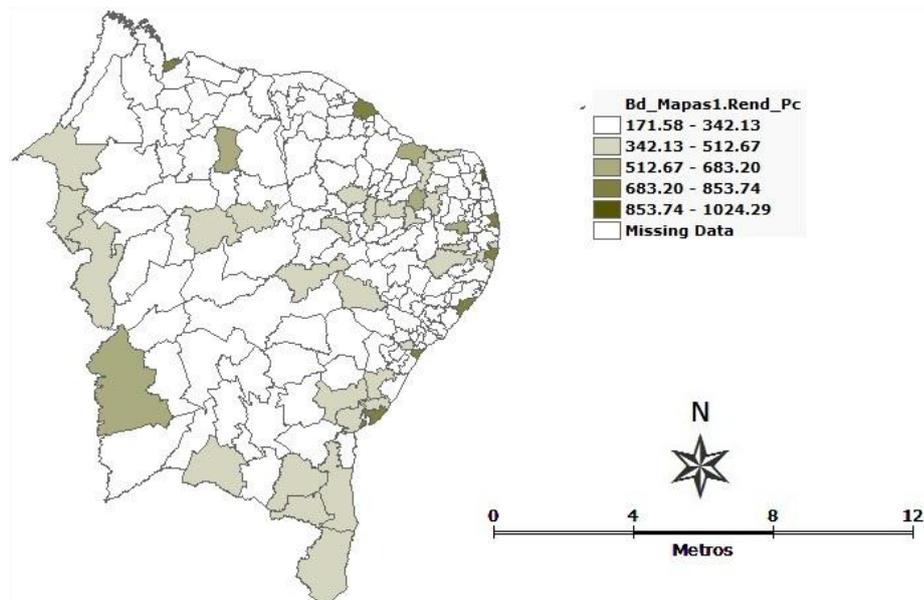
Fonte: Elaboração própria com base nos dados do DATASUS, 2010.

APÊNDICE K – Mapa 11. Distribuição das microrregiões segundo a variável “Proporção de pessoas com baixa renda (< ½ S.M)”, Região Nordeste, Brasil, 2010.



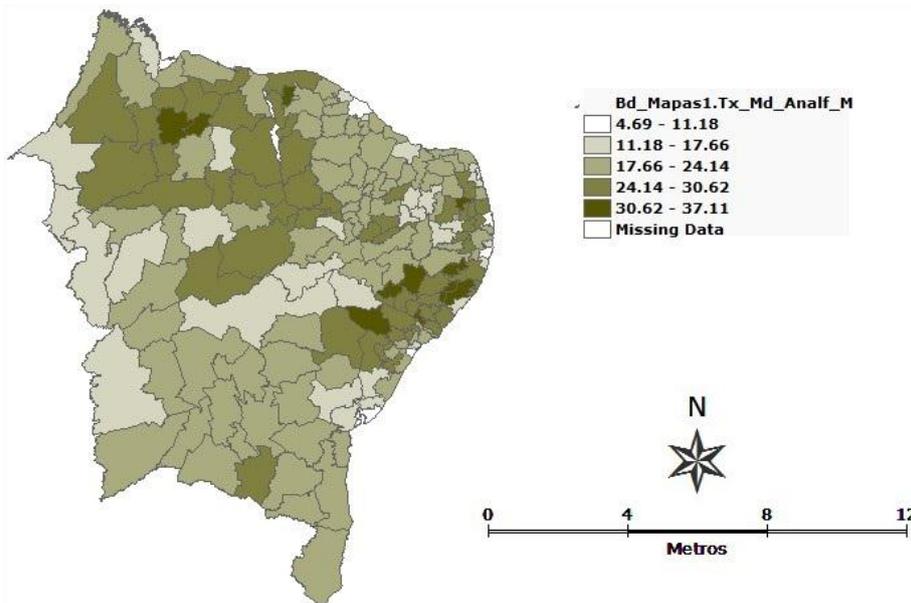
Fonte: Elaboração própria com base nos dados do DATASUS, 2010.

APÊNDICE L – Mapa 12. Distribuição das microrregiões segundo a variável “Renda Média domiciliar Per capita”, Região Nordeste, Brasil, 2010.



Fonte: Elaboração própria com base nos dados do DATASUS, 2010.

APÊNDICE M – Mapa 13. Distribuição das microrregiões segundo a variável “Taxa de analfabetismo feminino”, Região Nordeste, Brasil, 2010.



Fonte: Elaboração própria com base nos dados do DATASUS, 2010.

APÊNDICE N – Tabela 06. Análise descritiva das variáveis estudadas por UF.

ANÁLISE DESCRITIVA DAS VARIÁVEIS DO ESTUDO, SEGUNDO AS UF DA REIGIÃO NORDESTE												
MARANHÃO	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	22,65	688,24	81,16	31,70	26,38	8,88	4,66	5,74	83,38	84,07	91,95	86,82
MÍNIMO	11,25	171,59	40,77	4,70	21,06	4,31	3,69	4,46	36,22	29,35	62,95	17,03
MÉDIA	16,12	276,03	70,18	21,82	23,20	6,41	4,05	5,09	58,65	61,16	81,13	46,54
DESVIO P.	2,942	124,355	10,616	6,441	1,244	1,134	0,262	0,366	13,023	14,885	8,412	18,019
PIAUI	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	20,72	651,60	70,53	29,80	28,36	7,48	3,97	6,48	87,09	87,16	96,91	82,64
MÍNIMO	15,25	253,12	44,52	11,20	22,00	4,50	3,41	5,26	40,06	45,06	88,76	40,92
MÉDIA	17,44	316,14	64,07	22,79	24,96	5,99	3,65	5,91	59,26	67,29	94,06	54,32
DESVIO P.	1,692	103,375	6,902	5,201	1,919	0,965	0,186	0,364	13,098	12,623	2,374	11,379
CEARA	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	29,25	689,37	75,66	31,60	31,38	8,25	3,88	7,35	97,46	89,93	99,01	95,72
MÍNIMO	10,02	205,19	40,50	8,00	22,94	2,58	3,34	5,95	43,46	33,04	91,44	46,77
MÉDIA	14,69	282,37	66,91	21,54	25,83	5,62	3,59	6,72	59,01	65,70	94,94	60,64
DESVIO P.	3,777	83,987	7,105	4,245	1,609	1,521	0,141	0,343	11,387	11,900	1,897	11,317
RIO G. DO NORTE	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	20,88	891,94	68,86	25,60	28,67	9,09	3,91	7,05	99,15	97,59	98,66	98,37
MÍNIMO	11,17	249,33	31,50	7,60	20,75	4,59	3,33	5,47	33,24	64,58	92,22	58,70
MÉDIA	15,47	363,53	58,25	19,32	24,81	6,78	3,55	6,20	66,88	79,47	96,00	75,49
DESVIO P.	2,655	152,299	10,645	4,720	1,947	1,341	0,168	0,448	14,498	8,019	1,672	11,453

PARAIBA	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	22,00	787,78	75,05	30,80	29,00	8,17	3,66	7,30	97,22	93,73	96,13	97,04
MÍNIMO	9,67	206,56	38,89	10,00	23,56	3,92	3,18	5,84	36,92	25,28	73,62	37,01
MÉDIA	15,64	321,32	62,50	23,04	25,62	6,10	3,49	6,57	64,69	65,98	90,73	67,58
DESVIO P.	2,793	125,335	8,350	5,679	1,320	1,251	0,115	0,330	12,831	14,126	5,111	12,230
PERNAMBUCO	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	43,56	1024,28	75,82	33,90	27,95	9,50	3,92	6,88	100,00	97,60	100,00	100,00
MÍNIMO	13,19	208,59	5,79	4,70	22,00	4,68	3,23	5,90	47,95	50,12	82,19	50,10
MÉDIA	18,34	387,32	58,16	20,29	25,03	6,71	3,51	6,30	70,85	70,20	89,83	73,20
DESVIO P.	7,058	198,967	15,298	7,091	1,162	1,132	0,224	0,239	14,690	11,609	4,406	13,932
ALAGOAS	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	22,65	695,53	81,73	37,10	28,88	7,71	3,99	6,07	97,93	82,45	97,13	96,65
MÍNIMO	15,91	179,82	44,51	13,40	22,00	4,66	3,43	5,32	28,84	42,45	79,84	38,76
MÉDIA	17,97	284,33	69,72	27,70	24,49	6,01	3,79	5,72	58,50	63,31	91,16	67,28
DESVIO P.	1,955	129,781	9,105	5,440	1,953	0,919	0,179	0,229	17,684	13,261	4,652	16,769
SERGIPE	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	17,68	814,20	72,96	26,90	29,08	7,94	3,68	7,31	97,45	96,18	97,61	96,21
MÍNIMO	11,62	243,40	38,24	7,90	22,00	6,04	3,31	5,21	40,56	51,89	89,51	60,62
MÉDIA	15,13	338,25	62,42	21,52	25,52	6,90	3,51	5,90	60,44	76,68	94,70	74,78
DESVIO P.	1,790	147,746	9,003	5,154	2,090	0,696	0,136	0,676	14,527	10,731	2,391	10,860
BAHIA	Tx_Mort_inf_ObC_2010	Rend_Pc	Prop_Bx_rend	Tx_Md_Analf_M	Id_Md_M	A_Md_Est	Dens_Domic	Num_Md_Cont_PN	Gr_Urb	Cobert_A_gua	Cobert_Es_g_Sanit	Exist_C_Lixo
MÁXIMO	25,37	853,62	73,84	32,80	27,00	8,79	3,94	6,80	98,74	97,74	98,79	95,64
MÍNIMO	13,66	228,31	35,83	4,90	22,88	3,54	3,13	5,09	34,32	54,87	79,82	43,09
MÉDIA	18,77	344,69	62,03	20,03	24,60	6,53	3,54	5,84	59,54	73,21	93,26	65,30
DESVIO P.	2,648	121,938	8,650	4,923	1,163	0,989	0,165	0,417	15,242	8,869	3,814	13,611

