

# **A utilização de Variáveis Sintomáticas para Estimativas de Populações Municipais do Estado Do Paraná**

## **Palavra-Chave**

Estimativas populacionais, Método de correlações, Partição do Crescimento.

## **Autores**

Luciano Gonçalves de Castro e Silva<sup>1</sup>: Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da Faculdade de Ciências Econômicas (FACE) da Universidade Federal de Minas Gerais (UFMG) / pesquisador demógrafo do Instituto Brasileiro de Geografia e Estatística – IBGE (licenciado)

Reinaldo Onofre dos Santos: Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da Faculdade de Ciências Econômicas (FACE) da Universidade Federal de Minas Gerais (UFMG).

---

<sup>1</sup> *O IBGE está isento de qualquer responsabilidade pelas opiniões, informações, dados e conceitos emitidos neste artigo, que são de exclusiva responsabilidade do autor.*

## 1. Introdução

Para os profissionais que se dedicam ao planejamento e gestão, público e privado, uma das questões cruciais para seus trabalhos é a resposta para a pergunta sobre quantos perguntado sobre “*quantos seremos e onde estaremos no futuro*”. A definição do tamanho populacional e sua distribuição, seja dentro dos estratos que compõe a sociedade, seja a composição espacialmente explícita, é fundamental para a definição de políticas, diretas e indiretas, de distribuição da população, bem como aquelas destinadas a tender a suas necessidades.

Nesse contexto, numa tentativa de entender o papel dos estudos sobre população no planejamento do desenvolvimento houve, em 1989, o Simpósio sobre População e Planejamento para o Desenvolvimento em Riga, na Letônia (UNITED NATIONS, 1993). Uma das recomendações do documento final foi a de inserção de variáveis demográficas no processo de planejamento, tendo em vista que, para se ter uma visibilidade maior sobre o futuro, é necessária a construção de projeções tanto do capital quanto da força de trabalho (UNITED NATIONS, 1993). Nesse mesmo período, segundo Rees (1993), as projeções populacionais são elementos-chave no planejamento para o desenvolvimento, sendo determinantes para a construção de políticas públicas. No Brasil<sup>2</sup>, alguns estudos têm se dedicado a este esforço, contribuindo com a criação, aplicações e adaptações de técnicas de projeções para pequenas áreas (WALDVOGEL, 1998; BRITO; CAVENAGHI; JANNUZZI, 2010; JANNUZZI, 2007; FÍGOLI et al., 2010; GOMES et al, 2010; GONZÁLEZ; TORRES, 2012). Estes trabalhos fazem parte de um esforço contínuo dos estudiosos sobre população para avaliar as técnicas correntemente empregadas, fazendo coro com a literatura internacional sobre a aplicabilidade no processo de planejamento para o desenvolvimento, em especial daqueles com forte base territorial.

No âmbito institucional, a construção de estimativas populacionais está incorporada no processo de planejamento e gestão do Estado brasileiro. Na Constituição Federal do Brasil de 1988 (Art. 159, I, b) (BRASIL, 1988), o Fundo de Participação dos Municípios, doravante FPM, é uma transferência constitucional da União para todas as cidades brasileiras, sendo composto por 22,5% da arrecadação total do Imposto de Renda (IR) e do Imposto sobre Produtos Industrializados (IPI). Essa distribuição de recursos guarda correlação direta com o número de habitantes da municipalidade onde, segundo a norma vigente, são consideradas 18

---

<sup>2</sup> Para um maior conhecimento da agenda recente de trabalhos sobre aplicações de técnicas de projeção na América Latina, ver Cavenaghi (2012).

(dezoito) faixas populacionais, cada uma delas possuindo um respectivo coeficiente individual.

O §3º do artigo nº 91 da Lei nº 5.172 de 25/10/1966 (BRASIL, 2010) determina que o Instituto Brasileiro de Geografia e Estatística – IBGE é o órgão responsável por apresentar anualmente os dados oficiais das estimativas de população dos atuais 5.570 municípios brasileiros. Todavia, dentro do planejamento do órgão está a efetivação do censo demográfico, atualmente com periodicidade decenal.

Nesse contexto, com marcos populacionais decenais a partir dos censos demográficos, o IBGE constrói estimativas populacionais anuais por meio de técnicas de projeção. A alternativa seguida pelo IBGE foi a de aplicar o método conhecido como “Partição do Crescimento” ou Aibi (OLIVEIRA; ALBUQUERQUE; LINS, 2004), considerando que a aplicação do método das componentes demográficas não seria possível devido à instabilidade das próprias componentes para pequenos domínios. Entretanto, o método aibi pressupõe uma relação constante entre o crescimento da área menor e a área menor ao qual é parte componente, o que, ao longo do horizonte de projeção, pode não se verificar. Mesmo assim, o método aibi mostra-se mais robusto que outros métodos de projeção tendencial da população (SANTOS; BARBIERI, 2015).

Destarte, em que pese a qualidade dos resultados do método, é possível verificar que o uso de estatísticas oficiais de outras fontes (matrículas, óbitos, nascimentos, veículos, dentre outros), quando correlacionados com o tamanho populacional, podem fornecer estimativas mais confiáveis se comparadas com as obtidas por técnicas tendenciais (BRITO, CAVENAGHI, JANNUZZI, 2010). Esse fato traz luz para a construção dessas estimativas, uma vez que calibrado um modelo de correlação dessas variáveis com a população, é possível utilizar as fontes de dados anuais das mesmas variáveis para estimar a população também de forma anualizada. Esse trabalho busca, nesse sentido, apresentar alguns resultados obtidos a partir da construção de estimativas municipais para o Estado do Paraná em 2010, lançando mão das fontes de dados nos anos de 2000 e 2007. A estratégia de trabalho será apresentada nos parágrafos que seguem.

## **2. Métodos de Projeções para pequenas áreas**

Neste trabalho, serão apresentados métodos de projeções para a população total de pequenas áreas. Existem técnicas que contribuem para a estimação da população por idade e

sexo, como o método *correlação de coortes* (DUCHESNE, 1987), porém que seguem o modelo tendencial.

Também chamados de pequenos domínios, as pequenas áreas aqui descritas não possuem limites definidos na literatura. Em geral, referem-se a unidades de planejamento que podem variar entre setores censitários e unidades de planejamento subnacionais que ultrapassam limites políticos. Segundo Smith e Morrison (2005, p. 3), pequenas áreas “*vary in size from less than an acre to thousands of square miles, and from a mere handful of residents (or none at all) to many millions*”. No presente artigo, pequena área define-se como unidade territorial na qual o número de habitantes é insuficiente para garantir resultados satisfatórios por métodos como o das componentes demográficas, mas que necessitam ser modeladas conjuntamente para a construção da condição de retorno, ou seja, a de que o crescimento do conjunto territorial maior seja resultante da soma das unidades menores.

A seguir serão apresentados os métodos utilizados para a projeção de pequenas áreas do Estado do Paraná. O Estado foi tomado como estudo de caso por possuir um número razoável de municípios (399) e que não se alterou ao longo da década estudada. O período entre 2000 e 2010 foi delimitado, por sua vez, pela disponibilidade de dados de fontes diversas como variáveis sintomáticas do crescimento populacional, algo difícil de se estimar antes dessa década devido à disponibilidade e grau de cobertura dessas informações. Os dados utilizados para as projeções foram:

- População enumeradas nos censos de demográficos de 2000 e 2010, bem como na contagem da população de 2007 (IBGE, 2000; IBGE, 2007; IBGE, 2010);
- Matrículas registradas pelo censo escolar nos períodos 1998-2000, 2005-2007, 2008-2010 (INEP; 2016);
- Número de veículos registrados no município nos períodos 1998-2000, 2005-2007, 2008-2010 (DENATRAN, 2016)
- Número de eleitores cadastrados no município em 2000, 2007 e 2010 (TSE; 2016);
- Emprego formal no município nos períodos 1998-2000, 2005-2007, 2008-2010 (TEM, 2016);
- Nascidos vivos no município nos períodos 1998-2000, 2005-2007, 2008-2010 (DATASUS, 2016);
- Óbitos registrados no município (residência) nos períodos 1998-2000, 2005-2007, 2008-2010 (DATASUS, 2016).

É importante ressaltar que, como a Contagem de 2007 é aplicada aos municípios com menos de 170 mil habitantes, apenas estes foram estimados<sup>3</sup>. Caso contrário, os coeficientes ganhariam vieses das estimativas populacionais elaboradas para municípios acima desse limiar.

### 3. Métodos de correlação: incorporando variáveis sintomáticas do crescimento populacional

O conjunto de métodos de correlações de variáveis são uma família de equações baseadas em um modelo de regressão estimado, em geral, pelo método de mínimos quadrados ordinários<sup>4</sup>. O objetivo, via de regra, é estimar um parâmetro  $\theta$ , que será utilizado para a estimação da relação da área menor a ser projetada e a área maior. Não se trata de um modelo de causalidade, mas de identificação de correlações entre cada uma das variáveis classificadas como sintomáticas do crescimento populacional e o próprio crescimento da população. Assim, tem-se que o modelo geral se expressa de forma linear, como na equação 1 (JARDIM, 2000; SMITH; TAYMAN; SWANSON, 2001; GONZÁLEZ; TORRES; CELTON, 2010).

$$\theta_i = a_0 + \sum_1^k a_k Z_i + u_i \quad (1)$$

Onde,

$\theta_i$  é o parâmetro para a projeção da população, baseado na diferença de razões;

$a_k$  é o parâmetro de correlação entre a variável sintomática  $k$  e o parâmetro de crescimento;

$Z_i$  é qualquer variável sintomática ajustada pela diferença de razões;

$u_i$  é um termo de erro.

A variação de cada método está em como o parâmetro  $\theta_i$  é construído. No método de *correlação de diferenças* estima-se que a variação da população é dada pela diferença das relações de proporção da população e cada variável sintomática em dois momentos no tempo (Equações 2 e 3) (ÁLVAREZ, 2001; SWANSON; TEDROW, 1984).

$$\theta_i = \frac{P_{i,t}}{P_{j,t}} - \frac{P_{i,0}}{P_{j,0}} \quad (2)$$

<sup>3</sup> Os municípios de Curitiba, Londrina, Maringá, Ponta Grossa, Foz do Iguaçu, Cascavel, São José dos Pinhais e Colombo foram excluídos da nossa análise pois tiveram suas populações estimadas (não contadas) em 2007.

<sup>4</sup> Testes de heterocedasticidade mostram que esse pode não ser o melhor método de estimação, apesar do uso frequente na literatura.

$$Z_i = \frac{X_{i,t}}{X_{j,t}} - \frac{X_{i,0}}{X_{j,0}} \quad (3)$$

Onde,

$P_{i,t}$  e  $X_{i,t}$  são a população e qualquer variável sintomática da área menor  $i$  no tempo  $t$ ;

$P_{i,0}$  e  $X_{i,0}$  são a população e qualquer variável sintomática do ano base da área menor  $i$ ;

$P_{j,t}$  e  $X_{j,t}$  são a população e qualquer variável sintomática da área maior  $j$  no tempo  $t$ .

$P_{j,0}$  e  $X_{j,0}$  são a população e qualquer variável sintomática do ano base da área maior  $j$ ;

Assim, define-se que o crescimento entre a área maior e a área menor irá progredir segundo um fator que expressa a diferença desse coeficiente de participação no tempo (Equação 4).

$$P_{i,t+n} = \left( \theta_i + \frac{P_{i,t}}{P_{j,t}} \right) * P_{j,t+n} \quad (4)$$

Em outras palavras, a população da área menor é produto da população total no ano de projeção  $t+n$  e da soma entre  $\theta$  e a relação entre a área menor e a área maior no início da projeção. É possível, no entanto, construir o parâmetro de crescimento de pequenas áreas de forma que ele tenha propriedades multiplicativas. O método de *correlação de razões* se aplica assumindo que a evolução da população pode ser definida pela correlação entre a razão de dois coeficientes de proporção em tempos distintos e essa mesma razão de proporções de um conjunto de covariáveis (ÁLVAREZ, 2001; BRITO; CAVENAGHI; JANUZZI, 2010) (equações 5 e 6).

$$\theta_i = \frac{P_{i,t}/P_{j,t}}{P_{i,0}/P_{j,0}} \quad (5)$$

$$Z_i = \frac{X_{i,t}/X_{j,t}}{X_{i,0}/X_{j,0}} \quad (6)$$

Diferente da razão de diferenças, a população da área menor é estimada a partir do produto do fator estimado e a relação de proporção da área menor na área maior (equação 7).

$$P_{i,t+n} = \left( \theta_i * \frac{P_{i,t}}{P_{j,t}} \right) * P_{j,t+n} \quad (7)$$

Swanson e Tedrow (1984) inseriram na discussão sobre o uso de modelos de regressão para estimar população dois problemas inerentes a aplicação do método, a saber, a violação do pressuposto de independência dos resíduos e a inconsistência temporal dos coeficientes. Os autores defendem que a resolução do segundo problema contribui para a correção do primeiro, porém sem demonstrar evidências. Em resumo, a questão da inconsistência temporal recai sobre o uso dos coeficientes de regressão, pois os mesmos são estimados a partir dos dados observados de população, em geral decenais, e utilizados para as estimativas anuais de população, dentro ou fora do período intercensitário de estimação dos mesmos (horizonte base). Para a resolução dessa inconsistência temporal, Swanson e Tedrow (1984) sugerem uma transformação nas variáveis a partir da aplicação do logaritmo natural e divisão pelo horizonte base  $h$  (equações 8 e 9).

$$\theta_i = \frac{1}{h} \ln \left( \frac{P_{i,t}/P_{j,t}}{P_{i,0}/P_{j,0}} \right) \quad (8)$$

$$Z_i = \frac{1}{h} \ln \left( \frac{X_{i,t}/X_{j,t}}{X_{i,0}/X_{j,0}} \right) \quad (9)$$

Destarte, denominada pelos autores como *correlação de taxas*, essa forma assemelha-se à correlação de razões, porém com uma adequação sobre o horizonte de projeção. Assim, o coeficiente  $\theta_i$  é anualizado por meio da linearização dos coeficientes, podendo ser aplicado em diferentes horizontes de projeção, ao contrário dos dois métodos anteriores. Para a estimação da população num determinado horizonte de projeção, toma-se o exponencial do coeficiente  $\theta_i$  e do comprimento do horizonte  $h$  (equação 10).

$$P_{i,t+n} = e^{\theta_i h} * \frac{P_{i,t}}{P_{j,t}} * P_{j,t+n} \quad (10)$$

#### 4. Métodos de extrapolação: ratio methods

A técnica denominada *apportionment method*, ou projeção da participação no crescimento, consiste em projetar a população da pequena área com base na sua contribuição no crescimento absoluto da população esperada na área maior (UNITED NATIONS, 1956; WALDVOGEL, 1998; BARBIERI et al., 2010; SZWARCOWALD; CASTILHO, 1989). No Brasil, este método é conhecido como *método dos coeficientes* ou simplesmente *AiBi* e foi utilizado primeiramente por Madeira e Simões (1972) para projetar os contingentes rural e urbano entre 1960 e 1980, segundo as Unidades da Federação.

Ao contrário das técnicas de extrapolação matemática e métodos estruturais, a partição do crescimento considera que o *crescimento* das áreas menores possui relação linear com o crescimento da área maior (SMITH et al., 2001). Assim, pode-se escrever da seguinte forma:

$$P_{i,t+n} = P_{i,t} + [(P_{i,t} - P_{i,0}) / (P_{j,t} - P_{j,0})] (P_{j,t+n} - P_{j,t}) \quad (11)$$

Onde  $P$  é a população,  $i$  corresponde ao índice da área menor,  $j$  representa a área maior,  $P_{t+n}$  é a população a ser projetada,  $P_0$  refere-se à população base e  $P_t$  é a população inicial de projeção. Outra maneira de descrever essa equação seria:

$$P_{i,t+n} = A_i P_{j,t+n} + B_i \quad (12)$$

Onde  $A_i$  é um coeficiente de proporcionalidade entre o crescimento da área menor e da área maior e  $B_i$  representa um coeficiente linear de correção (WALDVOGEL, 1998). Em termos algébricos, temos:

$$A_i = (P_{i,t} - P_{i,0}) / (P_{j,t} - P_{j,0}) \quad (13)$$

$$B_i = P_{i,t} - A_i P_{j,t} \quad (14)$$

Uma vantagem de usar a equação 12 é a interpretação de seus valores e a possibilidade de ajuste. A exposição dos coeficientes e seus resultados para um grupo de especialistas pode

contribuir para o ajuste de valores mais adequados para os parâmetros  $A_i$  e  $B_i$ , conjugando a esta técnica outras que chamamos anteriormente de subjetivas. Pode-se usar algum método de extrapolação, linear ou não, para aqueles municípios que tendem a apresentar ao longo do horizonte de projeção valores iguais ou menores que zero. Para evitar resultados improváveis, comumente divide-se essa partição do crescimento entre as unidades que possuem crescimento positivo e aquelas que apresentam crescimento negativo, ou mesmo agrupando conjuntos populacionais com taxas de crescimento e tamanho populacional similares (WALDVOGEL, 1998; FÍGOLI et al., 2010).

Para sanar a limitação de determinar resultados improváveis, via de regra pela diferença de ritmo ou direção de crescimento entre as áreas menores e as áreas maiores, alguns autores apresentam uma transformação do método original (FRIAS, 1987); Outra forma

$$P_{i,t} = A_i P_{j,t} + B_i \quad (12)$$

$$\frac{P_{i,t}}{P_{j,t}} = \frac{A_i P_{j,t} + B_i}{P_{j,t}} = A_i + B_i P_{j,t}^{-1} = \varphi_i \quad (15)$$

A relação entre o crescimento da população da área menor e da área maior deixa de ser linear. A proporção  $\varphi_i$ , por definição, possui apenas valores entre 0 e 1. Os valores combinados de  $A_i$  e  $B_i$  também oferecem limiares diferentes de variação de  $\varphi_i$ , sendo necessário avaliar cada um dos casos para definir os limiares aceitáveis para esse parâmetro. Isso surge como uma vantagem, uma vez que impede valores improváveis em casos de populações pequenas com crescimento negativo que desaparecem no horizonte de projeção. Frias (1987) desenvolve a proposta, apresentando que os quatro casos possíveis para definição dos limiares inferiores e superiores são<sup>5</sup>:

Se  $A_i > 0$  e  $B_i < 0$ ,  $L_1 = 0$  e  $L_2 = A_i$ ;

Se  $A_i > 0$  e  $B_i > 0$ ,  $L_1 = A_i$  e  $L_2 = 1$ ;

Se  $A_i < 0$  e  $B_i < 0$ ,  $L_1 = 0$  e  $L_2 = 0$ ;

Se  $A_i < 0$  e  $B_i > 0$ ,  $L_1 = 0$  e  $L_2 = 1$ ;

Basicamente, definidos os limiares aceitáveis de  $\varphi_i$ , aplica-se uma função que descreva a evolução temporal de cada conjunto de parâmetros  $A_i$  e  $B_i$ . FRIAS (1987) adota uma função logística para estimar o comportamento de  $\varphi_i$ .

---

<sup>5</sup> Uma descrição do método pode ser vista em Corrêa, Myrrha e Fígoli (2011).

$$\varphi_i = L_1 + \frac{L_2 - L_1}{1 + e^{\alpha + \beta t}} \quad (16)$$

Com a equação 16, estima-se a relação entre a área maior e a área menor e determina-se o tamanho populacional do período de interesse.

## 5. Limites de aceitação das estimativas

Existe uma dificuldade latente em definir ou justificar uma margem de erro<sup>6</sup> aceitável. Em geral, autores trabalham com uma margem de erro de 10%, para mais ou para menos (SMITH, 1987; ÁLVAREZ, 2001; GONZÁLEZ; TORRES; CELTON, 2010; GONZÁLEZ, 2010). Alguns autores, contudo, apontam que os erros de projeção devem possuir limites de aceitação variáveis, segundo o horizonte de projeção e o tamanho da população (SMITH; TAYMAN; SWANSON, 2001; BRITO; CAVENAGHI; JANNUZZI, 2010). Brito, Cavenaghi e Jannuzzi (2010) salientam que o limiar ideal para a margem de erro deve variar segundo o tamanho, uma vez que o crescimento de populações é mais estável se comparado com populações menores.

Quanto ao tempo, os autores acima afirmam, em concordância com Smith, Tayman e Swanson (2001), que em horizontes de aproximadamente dez anos a margem de erro pode variar entre 8% e 14%. Wilson e Rowe (2011), em sua análise sobre o Reino Unido, apontam que, independentemente do tamanho da população, erros inferiores a 5% são aceitáveis, entre 5% e 10% seriam considerados pequenos, entre 10% e 20% considerados moderados e superiores a 20% estariam classificados como grandes erros. Nesse trabalho, optou-se por destacar 5% e 10% como margens de erro relativo para cada unidade, buscando um equilíbrio entre o rigor aceitável para o horizonte de projeção, relativamente curto. Além disso, a avaliação das estimativas foi feita com base num conjunto de indicadores, a saber:

1) Raiz Quadrática Média dos Erros Absolutos:  $RQMEA = \sqrt[2]{\sum_{i=1}^N \frac{(P_{i,est} - P_{i,obs})^2}{n}}$

---

<sup>6</sup> Em que pese o uso corrente do termo "erro" na nomeação dos indicadores de avaliação aqui discutidos, o termo mais adequado seria o de "desvio". Isso se deve ao fato de que "erro" se refere ao desvio em relação a população real, ao passo que existem outras fontes de erro que não serão tratadas aqui, a saber, subenumeração censitária e subregistro dentre as bases que compõem as variáveis sintomáticas dos modelos. Assim, prefere-se o uso do termo "desvio" para interpretar as variações em relação a população enumerada no censo ou na contagem da população.

- 2) Raiz Quadrática Média dos Erros Relativos:  $RQMER = \sqrt{\sum_{i=1}^N \frac{(P_{i,est} - P_{i,obs})^2}{n * P_{i,obs}}}$
- 3) Erro médio Percentual:  $EMP = \frac{1}{n} \sum_{i=1}^N \frac{P_{i,est} - P_{i,obs}}{P_{i,obs}}$
- 4) Erro médio percentual em módulo:  $EMPM = \frac{1}{n} \sum_{i=1}^N \left| \frac{P_{i,est} - P_{i,obs}}{P_{i,obs}} \right|$
- 5) Erro de Distribuição:  $ED = \frac{1}{2} \sum_{i=1}^N \left| \frac{P_{i,est}}{P_{j,est}} - \frac{P_{i,obs}}{P_{j,obs}} \right|$
- 6) Erro Absoluto em módulo:  $EAM = \frac{1}{P_{j,obs}} \sum_{i=1}^N |P_{i,est} - P_{i,obs}|$
- 7) Índice de Autocorrelação Espacial de Moran:  $I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i^n [(u_i - \bar{u}) \sum_j^n w_{ij} (u_j - \bar{u})]}{\sum_i^n (u_i - \bar{u})^2}$
- 8) Índice Local de Autocorrelação Espacial de Moran:  $I_i = (u_i - \bar{u}) \sum_j^n w_{ij} (u_j - \bar{u})$

Cada um desses indicadores ressaltam um aspecto dos erros de projeção. O erro médio percentual e o erro médio percentual em módulo são medidas clássicas para avaliação, porém ocultam a variabilidade desses desvios. O chamado erro absoluto em módulo oferece a quantidade bruta de desvios em relação a população total observada. A raiz quadrática média penaliza os erros maiores em detrimento dos menores. O erro de distribuição, por seu turno, oferece uma medida de quantas pessoas deveriam ser realocadas dentre as unidades menores para adequar a distribuição da população estimada. Por fim, são avaliados os índices de autocorrelação espacial dos erros para identificar se há ou não algum padrão de agrupamento em regiões. A seguir, são analisados os resultados das estimativas.

## 6. Análise dos Resultados

Foi utilizada como base das estimativas populacionais as populações enumeradas no Censo Demográfico de 2000 e na Contagem Populacional de 2007<sup>7</sup>. Para os modelos de variáveis sintomáticas foram utilizadas as variáveis de nascimentos (SINASC/DATASUS),

<sup>7</sup> Apenas os municípios com menos de 170.000 habitantes.

óbitos (SIM/DATASUS), eleitores (TRE), vínculos de trabalho (CAGED/MT), frota de veículos (DENATRAN) e matrículas escolares do ensino fundamental (INEP/ME).

Para analisar a qualidade das estimativas vamos considerar que quanto mais próximas as mesmas estiverem dos resultados observados no Censo Demográfico de 2010, melhores serão essas estimativas. É fato que a existência de um mesmo nível de cobertura censitária (grau de omissão de pessoas), tanto no Censo Demográfico de 2000 e na Contagem Populacional de 2007, quanto no Censo Demográfico de 2010 são fundamentais para que a qualidade das estimativas municipais seja satisfatória. No caso do Paraná, a Pesquisa de Avaliação dos Censo Demográfico de 2000 e 2010 apontou um baixo grau de omissão de pessoas de menos de 5% em cada ano, indicando que a operação censitária no Estado do Paraná foi bem realizada pelo IBGE, o que ratifica e dá credibilidade aos bons resultados obtidos a partir dos modelos de variáveis sintomáticas, comentados a seguir.

A **Tabela** a seguir apresenta um resumo para o ano de 2010 com os principais resultados obtidos comparando as populações estimadas pelos 3 (três) modelos de variáveis sintomáticas (Correlação de Diferenças, Correlação de Razões e Correlação de Taxas) e para os 2 (dois) modelos de Tendência de Crescimento utilizados (AiBi Tradicional e AiBi Logístico), *versus* os quantitativos populacionais municipais efetivamente enumerados no Estado do Paraná a partir dos resultados do Censo Demográfico de 2010.

Pela análise da referida tabela percebemos claramente que as duas variantes do Método da Tendência de Crescimento apresentaram os piores resultados para todos os indicadores de avaliação utilizados nesse trabalho. Vamos tomar como referência para os nossos comentários o percentual de municípios com erro superior a 5% e a 10% em módulo. Dos 391 municípios com menos de 170.000 (cento e setenta mil) habitantes no Paraná, o Método AiBi Tradicional foi o que apresentou os piores resultados. O mesmo teve um desvio de mais de 5% (em módulo) da população em 2010 para 238 municípios (60,87%) e teve um desvio de mais de 10% (em módulo) da população em 2010 para 119 municípios (30,43%). Uma performance ruim também obteve o Método AiBi Logístico que teve um desvio de mais de 5% (em módulo) da população em 2010 para 125 municípios (31,97%) e teve um desvio de em mais de 10% (em módulo) da população em 2010 para 21 municípios (5,37%).

De longe os Modelos de Variáveis Sintomáticas apresentaram estimativas populacionais muito mais próximas à população efetivamente enumerada no Censo Demográfico de 2010. Os modelos de Correlação de Diferenças, Correlação de Razões e Correlação de Taxas apresentaram Coeficiente de Determinação ( $R^2$ ) de 0,6856, 0,6084 e 0,5965, respectivamente. Esses valores significam que 60% ou mais da variabilidade dos

dados consegue ser explicada pelos modelos ajustados. Curioso é o fato de que para esses 3 (três) modelos, quanto mais alto foi o  $R^2$  calculado pior foi a performance dos valores preditos para o percentual de municípios com erro relativo superior a 5% em módulo. O Modelo de Correlação de Diferenças teve um desvio de mais de 5% (em módulo) da população em 2010 para 81 municípios (20,72%), o Modelo de Correlação de Razões teve um desvio de mais de 5% (em módulo) da população em 2010 para 66 municípios (16,88%) enquanto que o Modelo de Correlação de Taxas teve um desvio de mais de 5% (em módulo) da população em 2010 para apenas 62 municípios (15,86%). Quando a análise é feita a partir dos valores preditos para o percentual de municípios com erro relativo superior a 10% em módulo, os 3 (três) modelos de variáveis sintomáticas apresentaram performance similar, tendo um desvio de mais de 10% (em módulo) da população em 2010 para 9 municípios (2,30%) no Modelo de Correlação de Diferenças, para 8 municípios (2,05%) no Modelo de Correlação de Razões e também para 8 municípios (2,05%) no Modelo de Correlação de Taxas.

Um outro indicador muito importante que corrobora os resultados encontrados é a raiz quadrática média do erro relativo. Mais uma vez os métodos de variáveis sintomáticas apresentaram as melhores *performances* na mesma ordem anterior, a saber: Correlação de Taxas (3,91%), Correlação de Razões (3,96%) e Correlação de Diferenças (4,21%). Seguindo esse indicador como critério, o AiBi tradicional também apresentou o pior desempenho com 10,66% de raiz quadrática média para o erro relativo.

Ao avaliar os padrões espaciais existentes nos erros relativos, observa-se que, mesmo a autocorrelação espacial sendo positiva e significativa (p-valor não apresentando, porém todos com significância a 0,01), o melhor desempenho é dado pelo método de correlação de taxas, como preconizado em Swanson e Tedron (1984). Além disso, o pior desempenho foi apresentado pelo AiBi logístico, apresentando maior autocorrelação espacial dos erros relativos.

A interpretação da figura 1, com a distribuição espacial dos erros relativos, permite suspeitar de padrões de agrupamentos, confirmando o índice global de Moran. Ao avaliar os resultados dos clusters a partir do indicador local de Moran<sup>8</sup>, percebe-se que existe um padrão de agrupamento que afeta, principalmente, o método AiBi logístico, sendo o método com o menor número de municípios com erros “não significantes” em termos de agrupamento e

---

<sup>8</sup> O agrupamento é feito pela significância do índice e comparando a média de cada unidade com média dos vizinhos, no caso de matriz padronizada como utilizado aqui. Assim, municípios classificados como “Alto-Alto” possuem valores altos e vizinhos de valores também altos, ao passo que “Baixo-Baixo” possuiriam valores baixos e seus vizinhos idem. Municípios com valores “Baixo-Alto” e “Alto-Baixo” possuem valores baixos com vizinhos detentores de valores elevados e valores altos com vizinhos detentores de valores baixos, respectivamente.

valores classificados como “Alto-Alto”. Os menores efeitos são apresentados pelo método AiBi, seguido pelo método de correlação de taxas (Figura 2 e Tabela 1).

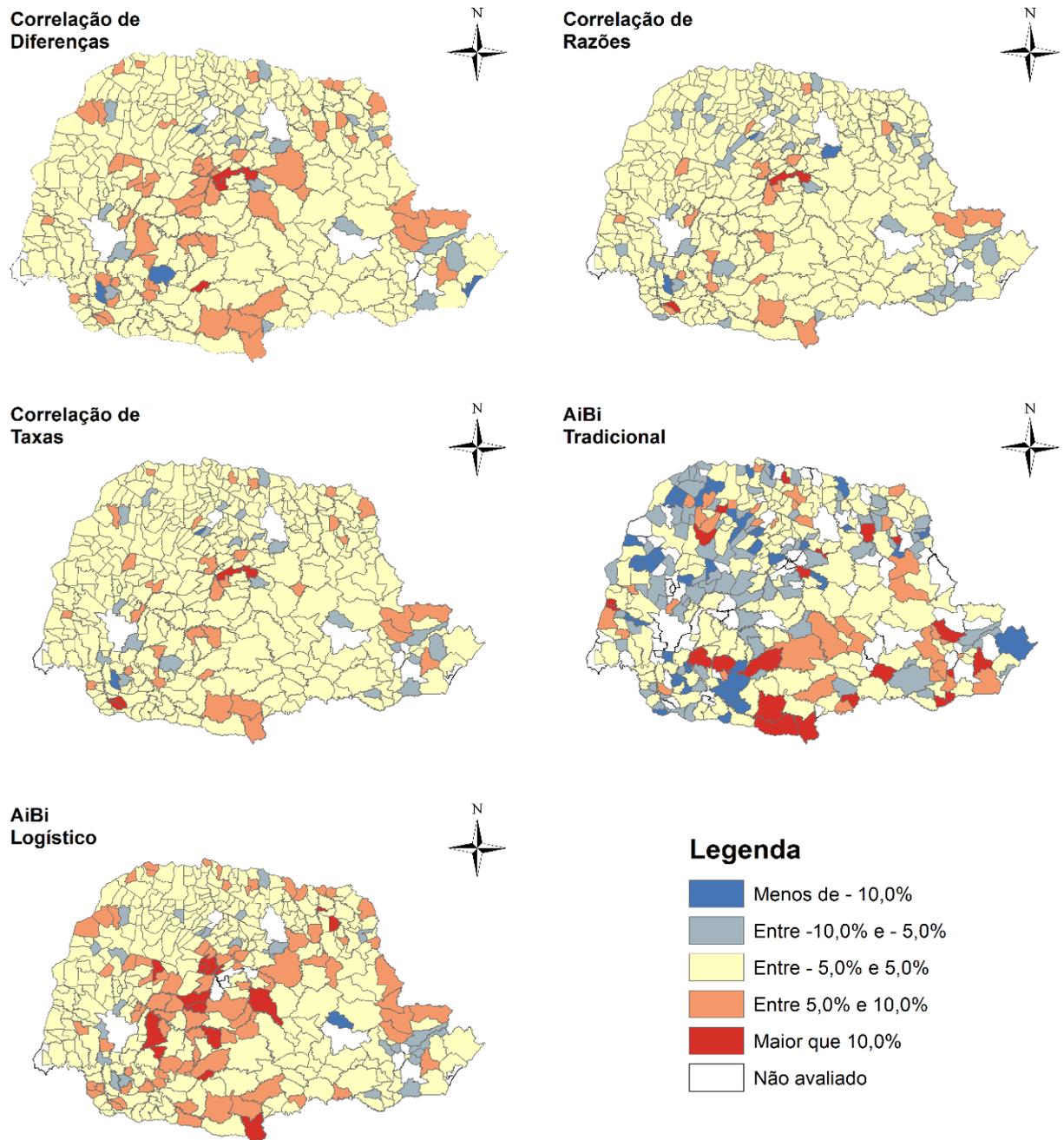
**Tabela 1 – Número de municípios classificados pelo Índice Global de Moran**

Agrupamento	Correlação de Diferenças	Correlação de Razões	Correlação de Taxas	AiBi Tradicional	AiBi Logístico
Não Significante	357	356	355	350	335
Alto - Alto	16	9	10	13	26
Alto - Baixo	2	3	4	5	2
Baixo - Alto	6	7	7	3	1
Baixo - Baixo	10	10	15	20	27

Fonte: elaborado pelos autores

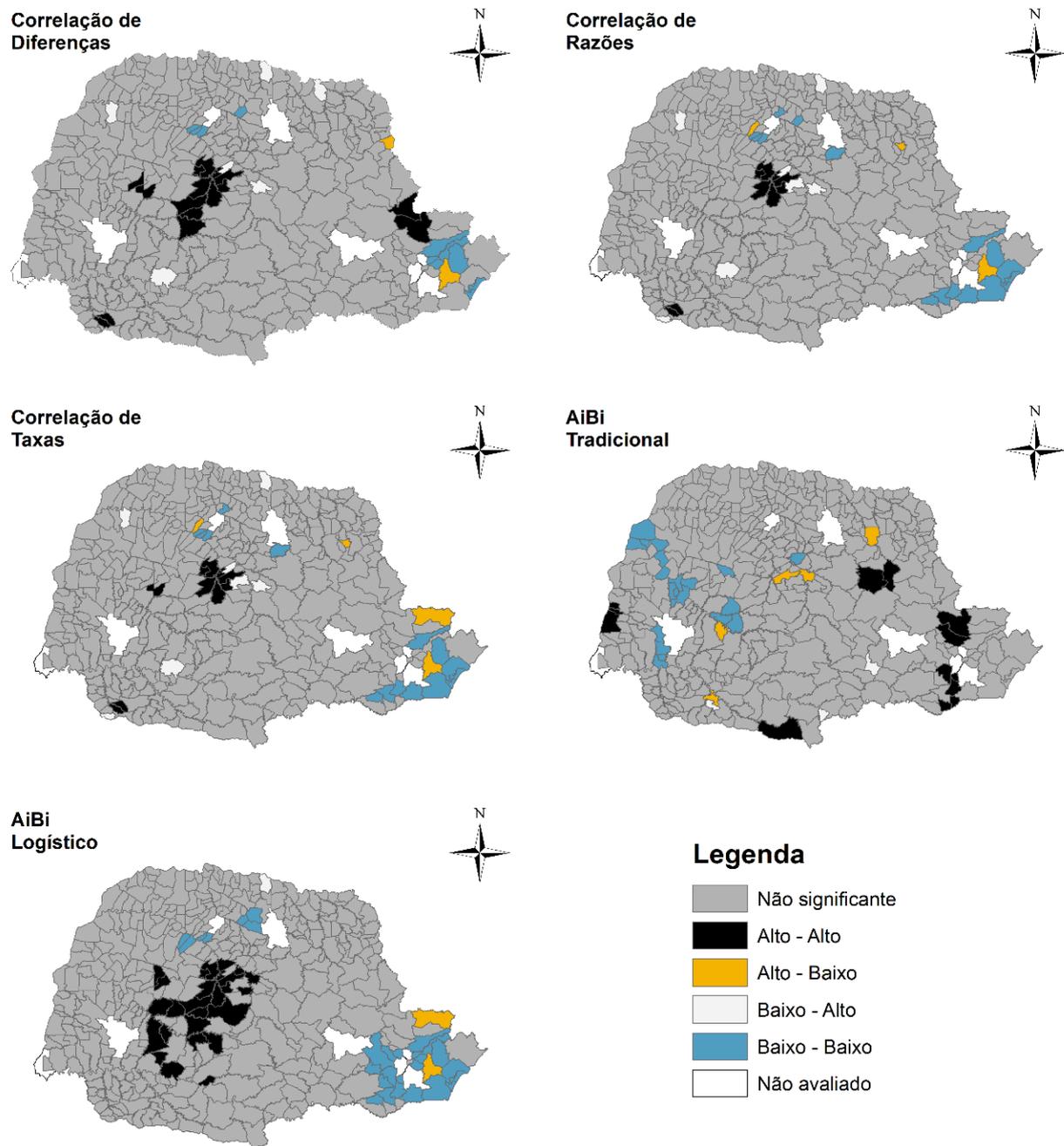
As piores *performances* dos indicadores da avaliação das estimativas para os municípios do Paraná com menos de 170 mil habitantes por técnica utilizada foram destacadas em vermelho na **Tabela** enquanto que as melhores performances foram destacadas em azul. Podemos perceber que nessa situação o Método da Correlação de Taxas foi o que apresentou os melhores resultados em quase todos os quesitos enquanto que o Método AiBi tradicional foi o pior de todos, em todos os quesitos, excetuando-se o Índice de Moran Global em que o AiBi logístico teve o pior desempenho.

**Figura 1 – Erros relativos segundo métodos de estimação populacional Paraná, 2010**



Fonte: IBGE, Censos Demográficos; MTE, 2016; INEP, 2016; DATASUS, 2016; TSE, 2016; DENATRAN, 2016.  
Sistema de projeções esférico  
Datum SIRGAS 2000

**Figura 2 – Agrupamentos classificados segundo o Índice Local de Autocorrelação Espacial de Moran - Paraná, 2010**



Fonte: IBGE, Censos Demográficos; MTE, 2016; INEP, 2016; DATASUS, 2016; TSE, 2016; DENATRAN, 2016.  
 Sistema de projeções esférico  
 Datum SIRGAS 2000

**Tabela 2 - Indicadores de avaliação das estimativas para os municípios do Paraná com menos de 170 mil habitantes por técnica utilizada - 2010\***

Indicador	Modelos de Variáveis Sintomáticas			AiBi	
	Correlação de Diferenças	Correlação de Razões	Correlação de Taxas	Tradicional	Logístico
Quantidade de Municípios	391	391	391	391	391
Coefficiente de Determinação (R <sup>2</sup> )	<b>0,6856</b>	0,6084	0,5965	-	-
Erro Absoluto	2,70%	2,87%	<b>2,52%</b>	<b>5,91%</b>	3,53%
Erro Médio Percentual	0,86%	-0,71%	<b>0,27%</b>	<b>-3,24%</b>	1,85%
Erro de Distribuição	1,35%	<b>1,26%</b>	<b>1,26%</b>	<b>2,95%</b>	1,77%
Erro Médio Percentual em Módulo	3,23%	3,01%	<b>2,97%</b>	<b>8,00%</b>	4,14%
Municípios com erro inferior a -5%	<b>24</b>	43	26	<b>170</b>	28
Municípios com erro superior a 5%	57	<b>23</b>	36	<b>68</b>	97
Municípios com erro superior a 5% em módulo	81	66	<b>62</b>	<b>238</b>	125
Municípios com erro inferior a -10%	5	5	<b>4</b>	<b>86</b>	3
Municípios com erro superior a 10%	<b>4</b>	3	<b>4</b>	<b>33</b>	18
Municípios com erro superior a 10% em módulo	9	<b>8</b>	<b>8</b>	<b>119</b>	21
Percentual de Municípios com Erro superior a 5% em módulo	20,72%	16,88%	<b>15,86%</b>	<b>60,87%</b>	31,97%
Percentual de Municípios com Erro superior a 10% em módulo	2,30%	<b>2,05%</b>	<b>2,05%</b>	<b>30,43%</b>	5,37%
Raiz quadrática média (Erro absoluto)	<b>775,71</b>	932,23	780,50	<b>1664,21</b>	1078,26
Raiz quadrática média (Erro relativo)	4,21%	3,96%	<b>3,91%</b>	<b>10,66%</b>	5,32%
Índice de Moran Global	0,149	0,151	<b>0,135</b>	0,167	<b>0,318</b>

*Fonte: desenvolvida pelos autores*

\* Para todas as projeções foi utilizado o período base de 2000-2007

## 7. Considerações Finais

A distribuição do Fundo de Participação dos Municípios (FPM) guarda correlação direta com o número de habitantes da municipalidade onde, segundo a norma vigente, são consideradas 18 (dezoito) faixas populacionais, cada uma delas possuindo um respectivo coeficiente individual. Segundo determinação legal, o Instituto Brasileiro de Geografia e Estatística – IBGE é o órgão responsável por apresentar anualmente os dados oficiais das estimativas de população dos atuais 5.570 municípios brasileiros.

Conforme visto, apesar do “Método da Partição de Crescimento”, também conhecido como AiBi, metodologia utilizada pelo órgão das estatísticas oficiais ser simples e parcimonioso, o mesmo tem se mostrado não muito adequado quando aplicado em períodos maiores que 5 (cinco) anos após o último Censo Demográfico.

O presente trabalho buscou apresentar uma alternativa metodológica de estimativa anual da população dos municípios brasileiros utilizando variáveis sintomáticas de diversas fontes diferentes, tais como: nascimentos, óbitos, eleitorado, vínculos trabalhistas, matrículas escolares e frota de veículos.

Foram realizadas estimativas para o Estado do Paraná utilizando como base os anos de 2000 e 2007 por 5 (cinco) metodologias diferentes: o AiBi, o AiBi logístico, variáveis sintomáticas com correlação de razões, com correlação de diferenças e com correlação de taxas. As estivas obtidas pelos métodos das variáveis sintomáticas apresentaram resultados muito promissores, acertando na maior parte das vezes (com 5 ou 10 % de erro relativo em módulo) os valores da população enumerada no Censo Demográfico de 2010. Os métodos de Tendência de Crescimento apresentaram os piores resultados, principalmente o AiBi tradicional.

Esse trabalho não tem a melhor pretensão de esgotar o tema da estimativa populacional de pequenas áreas, e visa tão somente agregar mais uma pesquisa na seara de discussão sobre quais as melhores formas de se estimar a população de pequenas áreas de forma a melhorarmos a qualidade da técnica e os resultados finais obtidos, visto que a variante população é a responsável pela maior parcela dos recursos recebidos pelos municípios brasileiros (FPM), e de onde sairá a maior parte dos investimentos para a saúde, educação, saneamento básico, infraestrutura, dentre muitos outros.

## 8. Referências Bibliográficas

ÁLVAREZ, G. **Estimación de población en áreas menores mediante variables sintomáticas**: una aplicación para los departamentos de la República Argentina (1991 y 1996). Santiago: Cepal, 2001. (Serie Población y Desarrollo).

BRASIL, **Constituição Federal**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/Constituicao.Constituicao.htm](http://www.planalto.gov.br/ccivil_03/Constituicao.Constituicao.htm), 1988.

BRASIL, CTN. **Lei nº 5.172, de 25 de Outubro de 1966**. Dispõe sobre o Sistema Tributário Nacional e institui normas gerais de direito tributário aplicáveis a União, Estados e Municípios. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis L, v. 5172, 2010](http://www.planalto.gov.br/ccivil_03/leis/L_5172_2010).

BRITO, L. P. G.; CAVENAGHI, S.; JANNUZZI, P. Estimativas e projeções populacionais para pequenos domínios: uma avaliação da precisão para municípios do Rio de Janeiro em 2000 e 2007. **Revista Brasileira de Estudos de População** vol.27 no.1 São Paulo Jan./Jun 2010.

BRITO, L. P. G.; CAVENAGHI, S.; JANNUZZI, P. Estimativas e projeções populacionais para pequenos domínios: uma avaliação da precisão para municípios do Rio de Janeiro em 2000 e 2007. **Revista Brasileira de Estudos de População**, v. 27, n. 1, p. 35-57, jan./jun. 2010.

CAVENAGHI, S. (org.) **Estimaciones y proyecciones de población en América Latina**. Desafíos de una agenda pendiente, Asociación Latinoamericana de Población-ALAP, 2012

CORRÊA, C. S. ; MYRRHA, L. J. D. ; FÍGOLI, M. G. B. . Métodos AiBi e Logístico para projeção de pequenas áreas: uma aplicação para a microrregião de Angicos ? RN. In: Seminário Internacional "Estimaciones y Proyecciones de Población: Metodologías, Inovaciones y Estimación, 2011, Rio De Janeiro. **Anais...** Seminário Internacional "Estimaciones y Proyecciones de Población: Metodologías, Inovaciones y Estimación, 2011.

DATASUS, **Sistema de informações sobre mortalidade**. Disponível em: <[www.datasus.gov.br](http://www.datasus.gov.br)> Acesso 4 de abril de 2016.

DATASUS, **Sistema de informações sobre nascidos vivos**. Disponível em: <[www.datasus.gov.br](http://www.datasus.gov.br)> Acesso em 4 de abril de 2016.

DENATRAN, **Frota de veículos**. Disponível em: <<http://www.denatran.gov.br/frota.htm>> Acesso em 1º de março de 2016.

DUCHESNE, L. **Proyecciones de poblacion por sexo y edad para areas intermedias y menores**. Santiago: Centro Latinoamericano de Demografia, 1987.

FÍGOLI, M. G. B.; ESPINOZA, L. L. R.; GONZAGA, M.; GOMES, M. M. F. Aspectos metodológicos para a projeção de localidades intra-urbanas uma aplicação a Minas Gerais. In: XVII ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS 2010. **Anais...** Caxambu: Abep, 2010

FRIAS, Luiz Armando de M. Projeções da população residente e do número de domicílios particulares ocupados por situação urbana e rural, segundo as unidades da Federação no período 1985-2020 In: WONG, Laura R; HAKKERT, Ralph; LIMA, Ricardo(Org) **Futuro da população brasileira: projeções, previsões e técnicas** Embu, São Paulo: ABEP, p148-172, 1987.

GOMES, M. M. F; GONZAGA, M.; UMBELINO, G. M.; FÍGOLI, M. G. B.; RODRIGUES, R. N. O "sertão nordestino" mineiro: caracterização e projeção populacional das microrregiões do Vale do Jequitinhonha, 2000-2030. In: XIV SEMINÁRIO SOBRE A ECONOMIA MINEIRA. **Anais...** Diamantina, 2010.

GONZÁLEZ, L. M.; TORRES, E. Estimaciones de población en áreas menores en América Latina: revisión de métodos utilizados. In: CAVENAGHI, S. (org.) **Estimaciones y proyecciones de población en América Latina**. Desafíos de una agenda pendiente, Asociación Latinoamericana de Población-ALAP, 2012

GONZÁLEZ, L. M.; TORRES, E.; CELTON, D. Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina). In: **IV CONGRESO DE LA ASOCIACIÓN LATINOAMERICANA DE POBLACIÓN**. Havana: Alap, 2010.

IBGE, **Censo demográfico de 2000**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em 15 de março de 2016

IBGE, **Censo demográfico de 2010**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em 15 de março de 2016

IBGE, **Contagem da população de 2007**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em 15 de março de 2016.

INEP, **Censo escolar**. Disponível em: <<http://www.inep.gov.br>>. Acesso em 15 de março de 2016

JANNUZZI, P. M. Cenários futuros e projeções populacionais para pequenas áreas: método e aplicação para distritos paulistanos 2000-2010. **Revista Brasileira de Estudos de População**, v. 24, n. 1, p. 109-137, 2007.

JARDIM, M. de L. T. Metodologias de estimativas e projeções populacionais para áreas menores: a experiência do Rio Grande do Sul. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 12. **Anais...** Caxambu: Abep, 2000.

MADEIRA, J. L.; SIMÕES, C. C. S. Estimativas preliminares da população urbana e rural segundo as unidades da Federação, 1960/1980: por uma nova metodologia. **Revista Brasileira de Estatística**, v. 33, n. 129, p. 3-11, jan./mar. 1972.

MTE, MINISTÉRIO DO TRABALHO E EMPREGO, **Relatório anual de informações sociais** Disponível em: <<http://www.mte.gov.br/frota.htm>> Acesso em 1º de janeiro de 2016.

OLIVEIRA, J. de C.; ALBUQUERQUE, F. R. P. C.; LINS, Ivan Braga. **Projeção da população do Brasil por sexo e idade para o período 1980-2050**—revisão 2004. Rio de Janeiro, IBGE, 2004.

REES, P. H. Population Projection for Development Planning. In: UNITED NATIONS. **Population and Development Planning**. New York: Department of Economic and Social Development, 1993.

SANTOS, Reinaldo Onofre dos; BARBIERI, Alisson Flávio. Projeções populacionais em pequenas áreas: uma avaliação comparativa de técnicas de extrapolação matemática. **Revista Brasileira de Estudos de População**, v. 32, n. 1, p. 139-163, 2015.

SMITH, S. K.; MORRISON, P. A. Small-area and business demography. In: POSTON, D.; MICKLIN, D. (Eds.). **Handbook of Population**. New York: Springer Publishers, 2005.

SMITH, S. K.; TAYMAN, J.; SWANSON, D. A. **State and local population projections: methodology and analysis**. New York: Kluwer: Plenum, 2001.

SWANSON, David A.; TEDROW, Lucky M. Improving the measurement of temporal change in regression models used for county population estimates. **Demography**, v. 21, n. 3, p. 373-381, 1984.

SZWARCWALD, C. L.; CASTILHO, E. A. Proposta de um modelo para desagregar projeções demográficas de grandes áreas em seus componentes geográficos. **Revista Saúde Pública**, v. 23, p. 269-276, ago. 1989.

TSE, **Repositório de dados eleitorais**. Disponível em: <<http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>> Acesso em 15 de março de 2016.

UNITED NATIONS. Department of Economic and Social Development. **Population and Development Planning**. New York, 1993.

WALDVOGEL, B. **Técnicas de projeção populacional para o planejamento regional**. Belo Horizonte: UFMG/Cedeplar, 1998. (Estudos Cedeplar, n. 1).

WILSON, T.; ROWE, F. The forecast accuracy of local government area population projections: a case study of Queensland. **Australasian Journal of Regional Studies**, v. 17, n. 2, p. 204-243, 2011.