

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS

EDUARDO FELIX JUSTINIANO

**Análise da eficiência de zoneamentos do desempenho de alunos
no Exame Nacional do Ensino Médio (2012-2018)**

Versão Corrigida

São Paulo
2020

EDUARDO FELIX JUSTINIANO

**Análise da eficiência de zoneamentos do desempenho de alunos
no Exame Nacional do Ensino Médio (2012-2018)**

Versão Corrigida

(Versão original encontra-se na unidade que aloja
o Programa de Pós-graduação)

Dissertação apresentada à Faculdade de Filosofia, Letras
e Ciências Humanas da Universidade de São Paulo para
obtenção do título de Mestre em Ciências.

Área de Concentração: Geografia Física

Orientador: Prof. Dr. Alfredo Pereira de Queiroz Filho

São Paulo

2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

J96a Justiniano, Eduardo
Análise da eficiência de zoneamentos do
desempenho de alunos no Exame Nacional do Ensino
Médio (2012-2018) / Eduardo Justiniano ; orientador
Alfredo Queiroz Filho. - São Paulo, 2020.
114 f.

Dissertação (Mestrado)- Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade de São
Paulo. Departamento de Geografia. Área de
concentração: Geografia Física.

1. Zoneamento. 2. Enem. 3. MAUP. I. Queiroz
Filho, Alfredo, orient. II. Título.

ENTREGA DO EXEMPLAR CORRIGIDO DA DISSERTAÇÃO/TESE**Termo de Ciência e Concordância do (a) orientador (a)**

Nome do (a) aluno (a): Eduardo Felix Justiniano

Data da defesa: 28/08/2020

Nome do Prof. (a) orientador (a): Prof. Dr. Alfredo Pereira de Queiroz Filho

Nos termos da legislação vigente, declaro **ESTAR CIENTE** do conteúdo deste **EXEMPLAR CORRIGIDO** elaborado em atenção às sugestões dos membros da comissão Julgadora na sessão de defesa do trabalho, manifestando-me **plenamente favorável** ao seu encaminhamento e publicação no **Portal Digital de Teses da USP**.

São Paulo, 14/10/2020.



(Assinatura do (a) orientador (a))

Nome: JUSTINIANO, Eduardo Felix

Título: Análise da eficiência de zoneamentos do desempenho de alunos no Exame Nacional do Ensino Médio (2012-2018)

Dissertação apresentada à Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Aprovado em: 28 ago. 2020.

Banca Examinadora

Prof. Dra. Flávia da Fonseca Feitosa

Instituição: UFABC

Julgamento: _____

Prof. Dra. Camila Fernanda Bassetto Sampaio

Instituição: UNESP

Julgamento: _____

Prof. Dra. Lígia Vizeu Barrozo

Instituição: FFLCH - USP

Julgamento: _____

Dedicatória

*Aos alunos da Escola Municipal Francisca Amélia de Azevedo
do município de Pedra Lavrada, no sertão da Paraíba,
que assistiam aula numa sala sem luz, sem livro didático e sem merenda.*

AGRADECIMENTOS

É difícil escrever esta última folha e não lembrar das pessoas que, desde o primeiro dia na Geografia, ajudaram em minha formação como geógrafo e na produção desta pesquisa de pós-graduação. A ajuda não foi só acadêmica, mas também econômica e afetiva

Agradeço aos professores da graduação e da pós-graduação que muito contribuíram para a minha formação.

Agradeço ao Professor Alfredo pela paciência e pela orientação no desenvolvimento deste trabalho.

Agradeço à minha companheira Noelí e ao nosso filho socioafetivo Ítallo pelo incentivo e compreensão pelas minhas ausências durante os trabalhos de pesquisa.

Agradeço ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira pela disponibilidade de dados, sem os quais não seria possível a pesquisa.

Agradeço à Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes, pelo financiamento da pesquisa.

RESUMO

JUSTINIANO, Eduardo Felix. **Análise da eficiência de zoneamentos do desempenho de alunos no Exame Nacional do Ensino Médio (2012-2018)**. 2020. 114 f. Dissertação (Mestrado em Ciências) – Departamento de Geografia, Universidade de São Paulo, São Paulo, 2020.

A pesquisa analisou a representação espacial do desempenho dos concluintes do ensino médio no Exame Nacional do Ensino Médio (ENEM) de 2012 a 2018. Foram analisados os dados de 7,4 milhões de alunos de até 18 anos que se inscreveram e pontuaram em todas as provas do teste educacional. A proficiência dos alunos, considerando todas as áreas do conhecimento avaliadas pelo ENEM, foi utilizado como parâmetro de desempenho. Após constatar a dificuldade de representação e análise conjunta dos alunos das esferas privada e pública, em função das diferenças da renda familiar e da menor distribuição espacial das escolas particulares, optou-se por reduzir o universo de análise. Assim, foram agrupados os dados de 5,6 milhões de concluintes da esfera pública em diferentes níveis de agregação, de municípios a regiões. A análise da eficiência dos zoneamentos do desempenho no exame, em diferentes faixas de renda familiar per capita, foi realizada a partir da metodologia de Openshaw (1984), adaptada para o objetivo desta pesquisa. Demonstrou-se que a participação no ENEM por unidade de área variou, comprometendo a utilização do exame como parâmetro de avaliação da qualidade de ensino válido para todo o Brasil. E, que o zoneamento por microrregiões foi mais eficiente para representar o desempenho no exame e sua evolução temporal. Diferentemente do esperado, os melhores rendimentos médios de alunos não ocorreram, predominantemente, nas áreas maior produto interno bruto (PIB). Apesar de os alunos de menor renda terem sido menos proficientes, a evolução do desempenho deles foi superior ao dos de maior renda, principalmente em áreas localizadas nas regiões Norte, Nordeste e Centro-Oeste. Os mapas produzidos nesta pesquisa indicaram não só as áreas de melhor rendimento no ENEM, mas também as áreas onde o poder público deve atuar para prover melhores condições para que os concluintes do ensino médio obtenham a proficiência adequada para esta importante etapa escolar.

Palavras-chave: Zoneamento, ENEM, MAUP.

ABSTRACT

JUSTINIANO, Eduardo Felix. **Análise da eficiência de zoneamentos do desempenho de alunos no Exame Nacional do Ensino Médio (2012-2018)**. 2020. 114 f. Dissertação de Mestrado – Departamento de Geografia, Universidade de São Paulo, São Paulo, 2020.

This research analyzed the spatial representation of the performance of high school graduate in the Brazilian National High School Examination (ENEM) from 2012 to 2018. Data was compiled from 7.4 million students up to 18 years of age who enrolled and scored in all educational tests. The proficiency of students, considering all areas of knowledge assessed by ENEM, was used as a performance parameter. It was decided to reduce the universe of analysis after realizing the difficulty of proper representation and joint analysis of students from the private and public schools, due to differences in family income and the lower spatial distribution of private schools. Thus, the data of 5.6 million graduates from public schools was grouped at different levels of aggregation, from municipalities to regions. The analysis of the zoning efficiency of the performance in the exam, in different ranges of family income per capita, was carried out using the methodology of Openshaw (1984), adapted for the purpose of this research. It was shown that the participation in ENEM per unit of area varied, compromising the use of the exam as a parameter for evaluating the quality of teaching throughout Brazil. It also showed that microregions zoning was more efficient to represent the performance in the exam and its temporal evolution. Contrary to expectations, the best average student performance did not predominantly occur in the areas of higher gross domestic product (GDP). Despite the fact that lower-income students were less proficient, their performance evaluation was better than that of higher income students, especially in areas located in the North, Northeast and Midwest regions. The maps produced in this research indicated not only the areas with the best performance in ENEM, but also the areas where the government should act to provide better conditions for high school graduates to obtain adequate proficiency for this important school stage.

Keywords: Zoning, ENEM, MAUP.

LISTA DE FIGURAS

Figura 1 – Fluxograma de processamento de dados	19
Figura 2 – Fluxograma dos procedimentos para testar os zoneamentos	21
Figura 3 – Alunos concluintes da rede particular de acordo com a participação no ENEM	45
Figura 4 – Alunos concluintes da rede pública de acordo com a participação no ENEM	45
Figura 5 – Pontuações médias dos alunos concluintes de até 18 anos do ensino médio da rede particular por área do conhecimento	46
Figura 6 – Pontuações médias dos alunos concluintes do ensino médio de até 18 anos da rede pública por área do conhecimento	47
Figura 7 – Pontuações médias no ENEM dos concluintes do ensino médio de até 18 anos	47
Figura 8 – Histograma da pontuação média dos concluintes do ensino médio de até 18 anos (2012-2018)	48
Figura 9 – Variação percentual das proficiências obtidas pelos alunos concluintes do ensino médio de até 18 anos da rede particular	50
Figura 10 – Variação percentual das proficiências obtidas pelos alunos concluintes do ensino médio de até 18 anos da rede pública	51
Figura 11 – Distribuição do número de alunos concluintes do ensino médio pontuados no ENEM por faixa de renda familiar <i>per capita</i> e por rede de ensino, dado que a renda familiar não foi superior a 20 salários mínimos (2012-2018)	52
Figura 12 – Municípios com escolas de ensino médio da rede particular (2012-2018)	53

Figura 13 – Evolução das proficiências médias dos alunos concluintes do ensino médio da rede pública por faixa de renda familiar <i>per capita</i> (2012-2018).....	55
Figura 14 – Gráficos de dispersão para as variáveis proficiência no ENEM dos alunos concluintes do ensino médio de até 18 anos e renda <i>per capita</i> para o zoneamento por regiões (2012-2018).....	56
Figura 15 – Diagramas de caixa das proficiências médias por mesorregião dos concluintes do ensino médio de até 18 anos.....	58
Figura 16 – Diagramas de caixa das proficiências médias por mesorregião.	59
Figura 17 – Proficiências médias no ENEM por Região (2012-2018).	65
Figura 18 – Proficiências médias no ENEM por UF (2012-2018).....	66
Figura 19 – Proficiências médias no ENEM por mesorregião (2012-2018).....	67
Figura 20 – Proficiências médias no ENEM por microrregião (2012-2018).....	68
Figura 21 – Proficiências médias no ENEM por município (2012-2018)	69
Figura 22 – Evolução temporal da proficiência média por município (2012-2018)....	72
Figura 23 – Evolução temporal da proficiência média por microrregião (2012-2018)	73
Figura 24 – Evolução temporal da proficiência média por mesorregião (2012-2018)	74
Figura 25 – Regiões metropolitanas e municípios onde os alunos concluintes do EM da rede pública de até 18 anos obtiveram proficiência média acima de quatro no ENEM (2012-2018).....	77
Figura 26 – Mapas de <i>clusters</i> e diagramas de espalhamento LISA das variáveis renda <i>per capita</i> média e proficiência média dos cinco zoneamentos avaliados.....	79

LISTA DE TABELAS

Tabela 1 – Número de municípios com escolas que ofereceram o último ano do ensino médio regular, dado que tiveram concluintes de até 18 anos que prestaram o ENEM	24
Tabela 2 – Alunos concluintes do ensino médio no Brasil (1.000 alunos).....	24
Tabela 3 – Alunos concluintes do ensino médio inscritos no ENEM no Brasil (1.000 alunos)	25
Tabela 4 – Alunos concluintes do ensino médio que pontuaram em todas as provas do ENEM (1.000 alunos)	26
Tabela 5 – Alunos concluintes do ensino médio de até 18 anos no Brasil no ENEM (2012-2018) que pontuaram em todas as provas e na redação por faixa de renda familiar e por rede de ensino (1.000 alunos).....	28
Tabela 6 – Alunos concluintes do ensino médio de até 18 anos que pontuaram em todas as provas do ENEM de 2012 a 2018 por faixa de renda familiar <i>per capita</i> , dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).....	29
Tabela 7 – Alunos concluintes do ensino médio da rede pública que pontuaram em todas as provas do ENEM, de 2012 a 2018, por faixas de renda <i>per capita</i> , dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).....	30
Tabela 8 – Alunos concluintes do ensino médio de até 18 anos da rede particular que pontuaram em todas as provas do ENEM, de 2012 a 2018, por faixas de renda <i>per capita</i> , dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).....	31
Tabela 9 – Equivalência entre dólar e real e condição econômica.....	31

Tabela 10 – Pontuação média dos concluintes do ensino médio da rede particular.....	33
Tabela 11 – Pontuação média dos alunos concluintes do ensino médio da rede pública	33
Tabela 12 – Pontuação média no ENEM dos alunos concluintes do ensino médio no Brasil, por rede de ensino (2012-2018)	34
Tabela 13 – Resumo estatístico das pontuações dos alunos concluintes do ensino médio de até 18 anos no ENEM (2012-2018).....	35
Tabela 14 – Percentual de concluintes de até 18 anos da rede particular por número de proficiências obtidas no ENEM (2012-2018)	36
Tabela 15 – Percentual de alunos concluintes de até 18 anos da rede pública por número de proficiências obtidas no ENEM (2012-2018)	36
Tabela 16 – Número de alunos concluintes do ensino médio da rede pública que pontuaram em todas as provas do ENEM de 2012 a 2018, segundo declaração de cor/raça (1.000 alunos)	37
Tabela 17 – Percentual de alunos concluintes do ensino médio da rede pública no Brasil que pontuaram em todas as provas no ENEM, segundo declaração de cor/raça	38
Tabela 18 – Correlações individuais com a variável pontuação dos alunos concluintes do ensino médio de até 18 anos da rede pública (2012-2018).	41
Tabela 19 – Variáveis correlacionadas com a pontuação no ENEM dos alunos concluintes da rede pública de até 18 anos e variáveis agrupadas por escolas (2012-2018).....	41
Tabela 20 – Proficiências médias dos concluintes do ensino médio da rede pública de até 18 anos por faixa de renda <i>per capita</i> e por ano	42

Tabela 21 – Correlações ecológicas e individual entre proficiência no ENEM dos alunos concluintes da rede pública de até 18 anos e renda familiar <i>per capita</i> (2012-2018).....	43
Tabela 22 – Número zonas com escolas que ofereceram o último ano do ensino médio para alunos concluintes do ensino médio de até 18 anos (2012-2018).....	52
Tabela 23 – Mesorregiões com valores discrepantes, abaixo dos limites inferiores do bigode.....	58
Tabela 24 – Microrregiões com valores discrepantes, acima dos limites superiores do bigode.....	59
Tabela 25 – Microrregiões com valores discrepantes, abaixo dos limites inferiores do bigode.....	59
Tabela 26 – Municípios com valores discrepantes da proficiência média (2012-2018).	60
Tabela 27 – Zonas aprovadas para análise longitudinal da proficiência média de 2012 a 2018	71
Tabela 28 – Principais regiões metropolitanas brasileiras (2017)	76
Tabela 29 – Valores limites das variáveis dos mapas LISA	78
Tabela 30 – Teste com valores aleatórios não ordenados.....	82
Tabela 31 – Teste com valores aleatórios ordenados.....	82
Tabela 32 – Correlações ecológicas entre a médias da renda <i>per capita</i> e proficiência no Enem dos alunos concluintes de escolas públicas que pontuaram em todas as provas e na redação, de 2012 a 2018, por UF e por região.....	83

Tabela 33 – Proficiências médias por escola obtidas pelos alunos concluintes do ensino médio no ENEM, no município de Ipanguaçu, RN (2012-2018)	86
Tabela 34 – Municípios com maior e menor proficiência média dos alunos concluintes do ensino médio de 16 a 18 anos no ENEM (2012-2018).....	86
Tabela 35 – Proficiência média por município dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM da microrregião de Boa Vista (2012-2018)	87
Tabela 36 – Proficiência média por microrregião dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM da mesorregião Norte de Roraima (2012-2018).....	87
Tabela 37 – Proficiência média por mesorregião dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM do Estado de Roraima (2012-2018).	88

LISTA DE QUADROS

Quadro 1 – Etapas de ensino constantes nos censos escolares (2012-2018).....	17
Quadro 2 – Procedimentos para testar zoneamentos	20
Quadro 3 – Dados utilizados	27
Quadro 4 – Grau de instrução dos pais segundo os microdados do ENEM	39
Quadro 5 – Testes da etapa 2 dos procedimentos propostos por Openshaw (1984).....	70

LISTA DE ABREVIATURAS E SIGLAS

BET	Bases e Estruturas Territoriais
CH	Ciências humanas e suas tecnologias
CN	Ciências da natureza e suas tecnologias
EJA	Educação de Jovens e Adultos
ENEM	Exame Nacional do Ensino Médio
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LC	Linguagem e códigos e suas tecnologias
LISA	Índice local de associação espacial (<i>Local Indicators of Spatial Association</i>)
MAUP	Problema da unidade de área modificável (<i>Modifiable Areal Unit Problem</i>)
MT	Matemática e suas tecnologias
PIB	Produto interno bruto
RID	Região Integrada de Desenvolvimento
RED	Redação
RM	Região Metropolitana
SARESP	Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo
SIG	Sistema de informação geográfica
Sisu	Sistema de Seleção Unificada
SM	salário mínimo
TCT	teoria clássica dos testes
TRI	teoria de resposta ao item

SUMÁRIO

1. Introdução	1
1.1. Objetivo	2
1.2. Objetivos específicos.....	3
2. Embasamento teórico	4
2.1. Correlações com o desempenho educacional.....	4
2.2. ENEM	7
2.3. Zoneamento	9
2.4. MAUP	11
3. Metodologia e procedimentos	16
3.1. Obtenção e processamento de dados.....	16
3.2. Avaliação de zoneamentos.....	20
4. Análise descritiva e evolutiva dos dados.....	23
4.1. Universo dos concluintes do ensino médio no ENEM	24
4.2. Variáveis socioeconômicas	27
4.3. Desempenho dos concluintes do ensino médio	32
4.4. Correlações	40
5. Resultados	44
5.1. Participação no ENEM	44
5.2. Desempenho no ENEM.....	46
5.3. Avaliação conjunta dos alunos da rede particular e pública	51
5.4. Variáveis associadas ao desempenho	54
5.5. Valores discrepantes	57
6. Discussão	62
6.1. Desempenho no ENEM.....	62
6.2. Avaliação de zoneamentos.....	63

6.2.1. Evolução temporal das proficiências.....	71
6.2.2. Áreas de melhor desempenho	75
6.3. Autocorrelação espacial	78
6.4. Interpretando o MAUP	80
6.4.1. Falácias.....	85
7. Conclusões	89
8. Referências bibliográficas	92

1. INTRODUÇÃO

O ENEM é um dos maiores testes educacionais do mundo em relação ao número de candidatos (TRAVITZKI, 2017). Os microdados do ENEM contêm dados do desempenho escolar e das características dos alunos, de suas famílias e residências. Informações relativas a características das escolas e da quantidade de alunos podem ser obtidas nos microdados dos censos escolares. Esses dados possibilitam análises diversas de uma ou mais variáveis que podem estar correlacionadas com o desempenho de alunos do ensino médio no teste educacional.

A utilização do coeficiente de correlação para avaliar a associação entre o desempenho de alunos ou seu quociente de inteligência e outras variáveis originou-se há mais de cem anos, a partir dos trabalhos de Francis Galton (1822-1911) e Karl Pearson (1857-1936). Um dos primeiros trabalhos que empregavam esse coeficiente em estudos de educação foi desenvolvido por Burt (1911).

Diversos estudos passaram a utilizar a correlação entre o desempenho escolar e as características socioeconômicas dos alunos ou das famílias, contextualizados temporal e espacialmente, a exemplo de Duff e Thomson (1923), Jones e Carr-Saunders (1927), Collins e Douglass (1937), Fleming (1943) e Bankston III e Caldas (1998).

Quase duzentos estudos que consideravam a correlação entre variáveis socioeconômicas e desempenho acadêmico foram avaliados por White (1982). Neles, os valores de correlações foram fortes em dados agrupados e fracos quando a unidade de análise foi o indivíduo.

Nos zoneamentos, os agrupamentos incorporaram a contiguidade espacial nas zonas (OPENSHAW, 1984). Nelas podem ocorrer inferências inadequadas de resultados obtidos de dados agrupados em outros níveis de agrupamentos ou em indivíduos.

A correlação de variáveis em zonas que podem ser modificáveis é uma discussão antiga da Geografia que se iniciou com Henry Sheldon. Em 1931, Sheldon observou a modificação de valor da correlação entre variáveis na medida em que a dimensão e o número de áreas analisados eram alterados. Também observou que quanto maior a dimensão das áreas e menor o número de áreas, maior era a

correlação. Isso passou a ser reconhecido como o problema das unidades de área modificáveis (*Modifiable Areal Unit Problem – MAUP*). Publicações que tratam desse tema, como as de Gehlke e Biehl (1934), de Neprash (1934), de Yule e Kendall (1950), de Robinson (1950), de Alker (1969), de Clark e Avery (1976), de Openshaw (1984) e de Jelinski e Wu (1996), orientam e demonstram como utilizar a estatística na interpretação dos dados agrupados em zonas.

Nesta pesquisa, a utilização de dados individuais de alunos e agrupados em zonas possibilitou não só estabelecer correlações entre variáveis socioeconômicas e desempenho no ENEM, mas também verificar problemas referentes às falácias descritas por Robinson (1950) e por Alker (1969), que ocorrem quando os resultados médios ou das correlações entre variáveis de um grupo de indivíduos inseridos numa zona são atribuídos a agrupamentos maiores, aos subgrupos ou ao indivíduo componente desse grupo.

A discussão que envolve o MAUP poderia auxiliar na escolha do zoneamento mais apropriado para representar o desempenho médio dos alunos de escolas públicas e sua correlação com a renda familiar *per capita* para diversos zoneamentos brasileiros.

No Brasil foram poucos os trabalhos que espacializaram a relação entre o rendimento escolar e as características socioeconômicas, a exemplo de Torres et al. (2003) e de Cunha et al. (2009). Porém, até onde pude verificar não há pesquisas no país que avaliem os zoneamentos para a representação espacial do rendimento de alunos no ENEM. Esta pesquisa pretende preencher essa lacuna temática.

1.1. Objetivo

O objetivo deste trabalho é identificar os zoneamentos que melhor representem o desempenho no ENEM dos alunos concluintes do ensino médio de escolas públicas, correlacionado com a renda familiar, no Brasil, de 2012 a 2018, orientando-se pela questão de que os melhores desempenhos ocorreram nas áreas de maior produto interno bruto (PIB).

1.2. Objetivos específicos

São objetivos específicos:

- Adquirir e analisar dados sobre escolas, alunos e ENEM.
- Quantificar a participação dos concluintes do ensino médio no ENEM.
- Identificar as variáveis socioeconômicas e das escolas que mais se correlacionam com o desempenho dos alunos no ENEM.
- Verificar se os alunos da rede particular de ensino podem ser analisados conjuntamente com os alunos da rede pública.
- Estabelecer uma metodologia de avaliação de zoneamento.
- Identificar as áreas onde ocorrem os melhores desempenhos no ENEM.
- Identificar resultados discrepantes do rendimento dos alunos.
- Compreender as falácias que podem ocorrer na interpretação de dados agrupados em zonas.

2. EMBASAMENTO TEÓRICO

2.1. Correlações com o desempenho educacional

O estudo pioneiro sobre a força da correlação entre variáveis foi desenvolvido por Galton (1888) ao analisar a relação entre o tamanho de cúbito e a estatura de homens adultos. Conceituou a correlação como uma função que descreve uma relação entre variáveis. Para definir o grau de correlação, esse autor utilizou termos como correlacionado, correlação próxima, moderadamente próxima e não correlacionado.

Pearson (1896) descreveu que os órgãos de um mesmo indivíduo são correlatos quando a média encontrada numa série do tamanho de um órgão for uma função do tamanho de outro órgão, e que se a média do tamanho de um órgão independe do tamanho do outro órgão, eles não são correlatos. A correlação foi definida como qualquer constante ou série de constantes que determinam essa função.

Após o estabelecimento do coeficiente de correlação por Pearson (1896), houve trabalhos que utilizaram estudo de correlações entre desempenho de alunos e diversas variáveis. Burt (1911) foi um dos pioneiros na utilização da correlação em análises de desempenho e de inteligência de estudantes, analisando criticamente os dados coletados por pessoas não treinadas em métodos científicos.

Duff e Thomson (1923) procuraram ranquear a ocupação dos pais com a média do quociente de inteligência dos alunos e entender a distribuição espacial, por distrito, dos alunos do ensino elementar de acordo com o quociente de inteligência. Os alunos de maiores habilidades localizavam-se em locais próximos e afastados das cidades e nas localidades de distância intermediária foram encontrados poucos casos de alta habilidade.

Para avaliar a influência da ocupação dos pais, da propriedade da residência, da necessidade de assistência social e da posse de telefone residencial sobre o desempenho escolar de alunos, Collins e Douglass (1937) trabalharam com três grupos com quociente de inteligência média aproximadamente iguais, separados por

desempenho escolar. Um *status* socioeconômico menor e os lares mais desfavoráveis contribuíram para um desempenho escolar insatisfatório.

Mesmo os estudos que procuraram avaliar o desempenho escolar desvinculado das influências da família, da classe social e da vizinhança, como o desenvolvido por Jones e Carr-Saunders (1927) acabaram por valorizar a influência deles. Esses autores analisaram a influência de fatores extraescolares ao comparar a variabilidade de desempenho de crianças em testes educacionais, de acordo com o tempo de residência em orfanato. Pareceu haver uma diminuição na variabilidade no resultado dos testes de inteligência conforme o período de residência se estende e, possivelmente, conforme a idade aumenta. Supomos que esse estudo evidencia a ocorrência da diminuição das diferenças entre os desempenhos escolares dessas crianças ao serem isoladas de seus lugares, de suas famílias e da diferenciação econômica, social e alimentar.

Para Fleming (1943), o desempenho escolar médio de alunos apresentou relação direta com o grau de prosperidade do grupo social em que eles estavam inseridos. Porém, as pontuações mais elevadas pareciam ser de alunos provenientes de lares com níveis sociais médios.

Para a correlação com o desempenho dos alunos, White (1982) identificou diversas variáveis utilizadas em pesquisas que mensuram o *status* socioeconômico dos alunos, como renda familiar, tamanho da família, instrução dos pais, ocupação, etnia, atividades culturais e intelectuais, salário de professores, proporção aluno-professor e outros. A ampla variação na força da relação entre as variáveis encontradas em diferentes estudos explicava-se pela importância da unidade de análise utilizada no cálculo do coeficiente de correlação. A força dessa relação em unidades agregadas não pôde e não deveria ser generalizada para situações em que o aluno for a unidade de análise. A dificuldade de fazer inferências a partir dos resultados obtidos de dados agrupados em outros níveis de agrupamentos também foi abordada por Robinson (1950), Alker (1969) e Clark e Avery (1976).

Para Patterson, Kupersmidt e Vaden (1990), a renda e o gênero foram os melhores preditores da conduta das crianças enquanto a renda e as características de residência seriam melhores preditores para o desempenho escolar.

Ao correlacionar variáveis de desempenho escolar com características socioeconômicas das famílias e as características de vizinhança, Bankson III e Caldas

(1998) pontuaram que, provavelmente, os alunos levam consigo os problemas de seu bairro e de suas famílias para a escola, o que poderia estar relacionado com a modelagem de ambientes sociais dentro da escola.

Na Região Metropolitana de São Paulo, Torres et al. (2003) relacionaram os resultados nas provas de língua portuguesa, obtidos por meio do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP), com um índice de privação que levou em consideração a escolaridade do responsável pelo domicílio, a renda, as condições de moradia, a idade dos responsáveis pelo domicílio, a estrutura familiar e a proporção de mulheres responsáveis pelo domicílio com educação primária ou menor.

Cunha et al. (2009) abordaram a relação entre os resultados obtidos pelo SARESP e as condições socioeconômicas e de moradia e as estruturas das escolas, obtendo uma alta correlação espacial entre as escolas em condições precárias e a pobreza em áreas do município de Campinas, indicando a desigualdade socioespacial no município. As escolas situadas em áreas periféricas e com menor infraestrutura têm alunos com menor rendimento escolar do que as escolas localizadas em áreas centrais e mais bem equipadas. Os autores evidenciaram, porém, que apesar da influência da renda e do nível educacional dos pais, o efeito da localização geográfica da escola permaneceu significativo para os testes de matemática.

Embora tenha ocorrido variação de resultados entre muitos países, o *status* de ocupação dos pais apresentou forte influência no desenvolvimento do aluno e a educação das mães também foi importante. Contudo, houve indicativos do aumento da importância do *status* de ocupação e da educação das mães no desempenho escolar de 1983 a 2000 (MARKS, 2008). A variação de resultados entre diferentes países também foi observada por Contreras et al. (2015) ao abordar a correlação entre desigualdade social, *bullying* e desempenho escolar.

Apesar das correlações, há estudos que procuraram relativizar os resultados dos testes educacionais e a qualidade da educação escolar. Meyer (1997) considerou que os resultados dos testes educacionais como indicador escolar foram falhos porque o desempenho dos alunos pode estar contaminado pela mobilidade deles entre escolas, por fatores extraescolares como família, características da comunidade e experiências anteriores dos estudantes. Para Wodtke et al. (2011), a exposição contínua dos alunos a bairros desfavorecidos caracterizados pela alta pobreza,

desemprego e recebimento de benefício social, por famílias numerosas, chefiadas por mulheres e por poucos adultos bem graduados (ensino médio e superior) durante todo o curso de vida da infância teve um impacto devastador sobre as chances de formação no ensino médio. Thieme et al. (2016) consideraram que há variáveis contextuais que estão além do controle da escola, das quais as características socioeconômicas dos alunos foram o exemplo mais notável. Tekwe et al. (2004) questionaram se as escolas podem ser responsabilizadas pelos efeitos significativos de fatores sociodemográficos nos resultados obtidos pelos seus alunos em testes educacionais.

2.2. ENEM

O ENEM, organizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), foi instituído em 1998 com objetivos de possibilitar a autoavaliação do cidadão, de criar uma referência nacional dos egressos do ensino médio, de fornecer subsídios para o acesso à educação superior e de constituir-se como acesso a cursos profissionalizantes pós-ensino médio (BRASIL, 1998). O ENEM também tem como objetivo a avaliação do desempenho acadêmico dos ingressantes nos cursos de graduação (BRASIL, 2009). A inscrição para o exame é gratuita para alunos concluintes das escolas públicas e para pessoas de baixa renda (BRASIL, 2013).

A aceitação do ENEM tanto pelos estudantes quanto pelas instituições de ensino não ocorreu de forma plena desde a sua instituição. Em 1998, havia 1.330.150 alunos concluintes do ensino médio,¹ e o número de participantes do ENEM foi 115.575,² o que corresponde a 8,7% dos alunos concluintes. O número de

1 Fonte: <http://download.inep.gov.br/download/censo/1998/basica/censo-miolo-98.pdf>. Acesso em: 2 jan. 2019.

2 Fonte: http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/494207. Acesso em: 2 jan. 2019.

participantes evoluiu até 2014, com mais de 8,7 milhões de inscritos, porém 2,5 milhões abstiveram-se da prova.³

Em relação ao número de estudantes ausentes das provas, Travitzki, Calero e Boto (2014) consideraram que a classificação das escolas no ENEM abarcou uma amostragem distorcida da totalidade do ensino médio no Brasil, provavelmente por estar relacionada com o caráter voluntário do exame.

Da Silva, Lamb e Barbosa (2016) propuseram que os testes educacionais organizados pelas diferentes instituições dificultaram a mobilidade de estudantes no Brasil e que, após a implantação do ENEM como um teste válido para todo o país, ocorreria um aumento da migração de estudantes de um estado para outro.

Esse teste educacional pôde ser utilizado para medição da qualidade do Ensino Médio, implementação de políticas públicas, aperfeiçoamento dos currículos do Ensino Médio e desenvolvimento de estudos e indicadores sobre a educação brasileira (BRASIL, 2011).

Segundo Soares (2005, p. 84), era comum a avaliação de “determinados constructos teóricos, que se caracterizam por serem variáveis latentes (não diretamente observadas) como por exemplo, a habilidade cognitiva de um aluno em língua portuguesa”. Esses constructos seriam “abstrações teóricas definidas dentro de uma rede de relações produzidas a partir de vários outros constructos” (SOARES, 2005, p. 84) que não são medidos diretamente, mas por meio de testes, questionários e outros que se constituem de itens como questões ou perguntas específicas, estruturados de acordo com um modelo. As técnicas para obter a medida do constructo dividiam-se principalmente em duas: a primeira baseava-se na teoria clássica dos testes (TCT), em que “o modelo para a construção da escala é baseado no instrumento como um todo”, e a segunda, conhecida como teoria de resposta ao item (TRI), englobaria “modelos específicos e construídos para cada item do instrumento, e a construção da escala consideraria todos esses modelos individuais” (SOARES, 2005, p. 84). O questionário do ENEM até 2008 baseava-se no TCT, e a partir de 2009 o ENEM começou a utilizar o TRI na avaliação.

3 Fonte: INEP: http://download.inep.gov.br/imprensa/2016/dados_de_abstencao_e_eliminacao_enem_2016.pdf. Acesso em: 21 maio 2018.

Para avaliar as competências e habilidades do aluno, as provas objetivas do ENEM ocorriam em dois dias e estavam divididas em quatro áreas do conhecimento: linguagens, códigos e suas tecnologias (LC); matemática e suas tecnologias (MT); ciências humanas e suas tecnologias (CH); ciências da natureza e suas tecnologias (CN), além da redação (RED). A partir de 2009, as provas do ENEM passaram a ter como base a teoria de resposta ao item (TRI) para mensurar o conhecimento dos inscritos e para possibilitar o acompanhamento e o desenvolvimento dos participantes ao longo dos anos (BRASIL, 2011). Mas essa implementação não foi feita de forma transparente e não possibilitou análises independentes pela comunidade científica (DA SILVA; LAMB; BARBOSA, 2016).

Travitzki (2017) considerou a evolução e a consistência da avaliação nas edições ao estudar a TRI nas edições 2009 e 2011 do ENEM, porém alertou para as melhorias a serem feitas e para a falta de clareza do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) sobre a constância dos itens da prova e sobre os parâmetros e procedimentos para garantir a qualidade dos itens e da prova.

2.3. Zoneamento

O Brasil pode ser dividido em vários níveis geográficos: regiões, unidades da Federação, mesorregiões, microrregiões, municípios, distritos, subdistritos e setor censitário, além de regiões metropolitanas, bairros, favelas, comunidades, subprefeituras e outros que se configuram como recorte de interesse para fins de produção e veiculação de informações estatísticas e geográficas. O Instituto Brasileiro de Geografia e Estatística sistematizou e classificou esses recortes na constituição das Bases de Estruturas Territoriais (BET), que

são constituídas pelas informações gráficas e alfanuméricas que definem recortes territoriais, legais ou institucionais, em termos de localização, forma, dimensão, dados histórico-legais e outros atributos. Possibilitam a produção e veiculação de informações estatísticas e geocientíficas para os recortes territoriais de interesse, vinculando elementos do meio físico a fenômenos socioeconômicos e de ocupação do território, entre outros.

Para fins de sistematização, podem ser classificadas em:

- base de estruturas político-administrativas (país, estados e municípios);
- base de estruturas geográficas (regiões, mesorregiões e microrregiões geográficas);
- base de estruturas da administração municipal (distritos, subdistritos, regiões, zonas de planejamento, zonas urbanas e rurais, bairros e similares e outras;
- base de estruturas de interesse específico (bacias hidrográficas, corredores industriais e agrícolas, regiões metropolitanas, favelas ou similares, áreas de proteção ambiental e de comunidades indígenas, e outras); e
- base de estruturas para censos e pesquisas (setores censitários e unidades de investigação – domicílios, estabelecimentos etc.). (IBGE, 1994, p. 19-20)

As mesorregiões são divisões das unidades da federação que apresentaram

formas de organização do espaço geográfico definidas pelas seguintes dimensões: o processo social como determinante, o quadro natural como condicionante e a rede de comunicação e de lugares como elemento de articulação espacial. Estas três dimensões possibilitam que o espaço delimitado como mesorregião tenha uma identidade regional. Esta identidade é uma realidade construída ao longo do tempo pela sociedade que aí se formou. (IBGE, 1990, p.8)

As microrregiões são subdivisões da mesorregião que apresentaram especificidades quanto à organização do espaço.

Essas especificidades não significam uniformidade de atributos nem conferem às microrregiões autossuficiência e tampouco o caráter de serem únicas devido a sua articulação a espaços maiores [...]. Essas especificidades referem-se à estrutura de produção agropecuária, industrial, extrativismo mineral ou pesca. Essas estruturas de produção diferenciadas podem resultar da presença de elementos do quadro natural ou de relações sociais e econômicas particulares, a exemplo, respectivamente, das serras

úmidas nas áreas sertanejas ou à presença dominante da mão de obra numa área de estrutura social capitalista. (IBGE, 1990, p.8)

A atual Divisão Regional do Brasil em Norte, Nordeste, Centro-Oeste, Sudeste e Sul foi definida pelo decreto nº 67.647 de 23 de novembro de 1970 para fins estatísticos.⁴ Os limites das divisões regionais foram alterados a partir do desmembramento do estado de Goiás. A parte norte passou a ser denominada Tocantins e foi integrada à Região Norte, enquanto a parte sul manteve o nome, na Região Centro-Oeste. Esse estado foi instituído pelo Ato das Disposições Constitucionais Transitórias da Constituição Federal de 1.988.⁵

Os estados e municípios são criados a partir de desmembramentos de áreas, definidos por leis, a exemplo do estado de Mato Grosso do Sul que foi estabelecido pela Lei Complementar nº 31, de 11 de outubro de 1977.⁶

2.4. MAUP

Na Geografia, o estudo da correlação entre variáveis em unidades de áreas iniciou-se na década de 1930. Gehlke e Biehl (1934) estudaram as correlações entre criminalidade de jovens do sexo masculino e os valores mensais de locação e, também, entre o valor dos produtos agrícolas e o número de fazendeiros. Demonstraram que a variação nos valores do coeficiente de correlação pode estar condicionada ao tamanho das unidades de área utilizadas e que quanto maior a área e o agrupamento, maior seria a correlação entre as variáveis. Os autores creditam esse achado a Henry Sheldon que, em 1931, demonstrou o aumento de valor do coeficiente de correlação, na medida em que as dimensões das áreas analisadas

4 Fonte: <https://www2.camara.leg.br/legin/fed/decret/1970-1979/decreto-67647-23-novembro-1970-409148-publicacaooriginal-1-pe.html>. Acesso em: 27 fev. 2020.

5 Fonte: http://www.planalto.gov.br/ccivil_03/constituicao/ConstituicaoCompilado.htm. Acesso em: 27 fev. 2020.

6 Fonte: <https://legis.senado.leg.br/norma/540736/publicacao/15809214>. Acesso em: 27 fev. 2020.

aumentam de um setor censitário para vários setores, diminuindo-se o número de áreas.

O coeficiente de correlação não seria apenas uma medida do grau de relacionamento entre variáveis, mas uma medida do grau de correspondência em suas respectivas distribuições geográficas e, para uma área heterogênea, a correlação pode ser alta numa parte e ausente noutra parte (NEPRASH, 1934).

Yule e Kendall (1950) estudaram a correlação entre a produtividade de batata e a de trigo e questionaram se os resultados encontrados foram “reais” ou “ilusórios” porque, do ponto de vista agrícola, era impossível ou impraticável plantar batata e trigo simultaneamente no mesmo lugar. Os autores estabeleceram as correlações em unidades de área modificáveis da produção agrícola de 1936 em 48 condados da Inglaterra e a agregação deles em 24, doze, seis e três zonas. Os valores de correlação obtidos foram 0,2189, 0,2963, 0,5757, 0,7649 e 0,9902, respectivamente. Os autores também pontuaram que as observações ao longo do tempo, correlacionando variáveis em bases mensais, trimestrais, anuais ou trienais também são uma escolha modificável.

Robinson (1950) estabeleceu uma diferenciação entre correlação ecológica e correlação individual. Na correlação individual, o objeto estatístico é indivisível e as variáveis seriam propriedades descritivas de indivíduos; o valor era o da correlação de Pearson entre duas variáveis de todos os indivíduos, desconsiderando-se o posicionamento geográfico. Na correlação ecológica, o objeto estatístico seria um grupo de pessoas e as variáveis poderiam ser expressas em porcentagens, medianas e, principalmente, médias. O valor da correlação ecológica depende do número de zonas, não seria normalmente calculado com ponderação e certamente não seria igual ao valor da correlação individual correspondente. As correlações ecológicas não poderiam ser usadas como substitutas das correlações individuais.

Alker (1969) descreveu como falácia ecológica a atribuição de valores das correlações (ou das covariâncias) ecológicas aos indivíduos; na direção inversa, definiu a falácia individualista como a atribuição de valores das correlações individuais para o zoneamento. Além dessas duas, as falácias poderiam ocorrer na atribuição de valores de covariância ou de correlação de um indivíduo para uma zona (falácia universal), da zona para o indivíduo (falácia seletiva) e de uma zona para o

zoneamento ou grupo de zonas e vice-versa (falácia entre escalas).⁷ Além disso, a falácia ocorreria por fatores contextuais, quando o contexto ou a estrutura social altera a força ou a forma das relações causais ou estatísticas.

Muitas análises estatísticas nas ciências sociais envolveram unidades observacionais que seriam agregados de unidades menores e, em alguns casos, os dados disponíveis para pesquisa encontram-se somente na forma agregada. Os cientistas sociais, sobretudo os geógrafos, trabalhariam com dados agregados como substitutos dos dados individuais. A principal desvantagem de utilizar dados agregados seria a dificuldade de fazer inferências válidas em vários níveis baseadas no resultado num único nível (CLARK; AVERY, 1976).

Para Openshaw e Taylor (1979), o problema de agregação para o geógrafo pareceu ser muito mais simples, pois nenhuma teoria especial de ligação entre escalas era essencial e o necessário seria um acordo sobre o que constitui os objetos da investigação geográfica.

O MAUP seria composto por dois problemas distintos, mas muito relacionados: o primeiro problema seria a escala, em que a variação nos resultados poderia ser frequentemente obtida quando os dados de um conjunto de unidades de área passariam a ser progressivamente agregados para análise em menos unidades, porém maiores. O segundo problema seria a agregação que se evidencia quando há variações de resultados ao utilizar unidades de área alternativas de análise, mantendo-se o número de unidades (OPENSHAW, 1984).

Enquanto os dados censitários eram coletados essencialmente em unidades não modificáveis (residências e pessoas), os resultados dessas coletas seriam agrupados e reportados em distritos, setores e municípios, e os valores de qualquer pesquisa baseadas nessas áreas não teriam validade independente das unidades de áreas analisadas (OPENSHAW, 1984).

Openshaw (1984) definiu o zoneamento como apenas um caso especial de um sistema de agrupamento que incorpora uma restrição de contiguidade. Os zoneamentos atuariam como um detector de padrões espaciais, e a detecção dos padrões dependeria do desenho do zoneamento. As zonas constituíram-se em

7 O termo utilizado pelo autor em inglês é *cross-level fallacy*.

objetos ou áreas geográficas individualizadas que seriam a unidade básica para a observação e para a medição de fenômenos espaciais. Em vez de aceitar qualquer resultado que um zoneamento produz, deve-se especificar qual o resultado esperado a ser alcançado sem violar suposições estatísticas ou fatores geográficos.

A agregação de área diminuiria os valores da variância da população e, como consequência, a heterogeneidade espacial seria perdida ou distorcida (JELINSKI; WU, 1996).

Dependendo do número de áreas, “a natureza dos processos envolvidos é tal que é muito provável a existência de diferentes regimes de correlação espacial em diferentes sub-regiões”. Estes regimes podem ser evidenciados na utilização dos “indicadores locais de autocorrelação espacial e o mapa de espalhamento de Moran” (CÂMARA et al., 2004, p. 11).

Para Anselin (1995), os indicadores locais de associação espacial (LISA – *Local Indicators of Spatial Association*) seriam qualquer estatística que satisfaz duas condições:

- o LISA de cada observação fornece um indicativo da extensão de agrupamentos espaciais significativos com valores semelhantes em torno dessa observação;
- o somatório de LISA para todas as observações é proporcional a um indicador global de associação espacial.

Câmara et al. (2004) definiram que os agrupamentos locais (*clusters*), comumente nomeados *hot spots*, seriam as localizações ou grupo de localizações contíguas nas quais o LISA é significativo.

O diagrama de espalhamento de Moran seria uma “maneira adicional de visualizar a dependência espacial. Construído com base nos valores normalizados [...], permite analisar o comportamento da variabilidade espacial” (CÂMARA et al., 2004, p. 19-20). Pode ser composto por um gráfico bidimensional, dividido em quatro quadrantes que mostram valores positivos e médias positivas, valores negativos e médias negativas, valores positivos e médias negativas e valores negativos e médias positivas (CÂMARA et al., 2004). O tradicional diagrama de espalhamento univariado de Moran poderia ser adaptado para representar também a autocorrelação bivariada – nesse caso, a correlação entre uma variável num local e uma variável diferente nos

locais vizinhos (ANSELIN; SYABRI; KHO, 2006; ANSELIN; SYABRI; SMIRNOV, 2002).

Manley (2014) considerou que a autocorrelação e a correlação cruzada foram muito importantes para entender o grau e a gravidade do MAUP. Além disso, pontuou que é improvável que a solução para esse problema seja encontrada, porque fazê-lo seria negar a espacialidade inerente dos dados que estão sob investigação.

3. METODOLOGIA E PROCEDIMENTOS

Esta pesquisa pode ser classificada como descritiva ou não experimental. Uma pesquisa que, segundo Köche (1997, p.124), “estuda as relações entre duas ou mais variáveis de um dado fenômeno sem manipulá-las”. Também é evolutiva por levar em consideração a alteração de valores de 2012 a 2018.

A medida da relação entre variáveis utilizada foi a da correlação, desenvolvida por Pearson (1896). Na interpretação espacial dessa medida, utilizamos observações e definições de pesquisadores sobre as alterações do valor da correlação relacionados às variações de escala (GEHLKE; BIEHL, 1934; OPENSHAW, 1984), ao posicionamento geográfico (NEPRASH, 1934) e à diminuição da variância na agregação (JELINSKI; WU, 1996).

Para compreender e exemplificar os problemas de inferência das medidas estatísticas em dados agrupados sobre indivíduos e em outros níveis de agrupamentos, empregamos os conceitos de falácia, descritos por Robinson (1950) e Alker (1969).

Os procedimentos envolveram a aquisição de dados de concluintes do ensino médio, a análise descritiva e evolutiva desses dados e a avaliação dos zoneamentos.

3.1. Obtenção e processamento de dados

Utilizamos microdados secundários e oficiais do ENEM e dos censos escolares de 2012 a 2018, disponibilizados gratuitamente no sítio eletrônico do INEP.⁸ Esses microdados têm mais de 350 milhões de linhas, que abrangem informações sobre os candidatos e sobre as escolas e alunos desde a pré-escola até os cursos técnicos subsequentes ao ensino médio.

8 Fonte: <http://inep.gov.br/microdados>. Acesso em: 11 fev. 2020.

A partir de uma solicitação enviada pelo Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC), recebemos dados das escolas contendo a localização delas por endereço físico e por zonas (informação).⁹

Dos microdados do ENEM, selecionamos os inscritos que declararam estar matriculados no último ano letivo do ensino médio e ter de 16 a 18 anos para não contemplar os alunos que não estavam em idade escolar ideal para o último ano do ensino médio.

Dos microdados dos censos escolares, tabelas de matrículas e de escolas,¹⁰ selecionamos as instituições de ensino que ofereceram o ensino médio regular e os alunos de 16 a 18 anos do último ano dos cursos regulares propedêutico, normal e técnico. Também selecionamos campos relacionados com características das escolas que pudessem interferir no desempenho dos alunos, como esfera administrativa, números de funcionários, de equipamentos multimídia e de computadores.

Da tabela de matrículas dos microdados dos censos escolares, selecionamos os estudantes de 16 a 18 das etapas finais do ensino médio (Quadro 1).

Quadro 1 – Etapas de ensino constantes nos censos escolares (2012-2018)

Código	Descrição
27	Ensino Médio - 3ª Série
28	Ensino Médio - 4ª Série
32	Ensino Médio - Integrado 3ª Série
33	Ensino Médio - Integrado 4ª Série
37	Ensino Médio - Normal/Magistério 3ª Série
38	Ensino Médio - Normal/Magistério 4ª Série

Fonte: INEP (2016a, 2016b, 2016c, 2016d, 2016e, 2018a, 2019a).

A concatenação dos microdados dos censos escolares e das edições do ENEM de 2012 a 2018 permitiu eliminar da análise os inscritos que, erroneamente,

9 Mensagem recebida por e.justiniano@usp.br em 11 fev. 2020.

10 Para cada ano há cinco tabelas de matrículas e uma de escola. São 35 tabelas de matrículas e sete de escolas, no total.

declararam ser concluintes de escolas que não ofereceram o ensino médio, no ano de inscrição para o exame, possibilitando conhecer o universo dos alunos concluintes do ensino médio, quantos deles se inscreveram e quantos pontuaram no ENEM.

Nessa concatenação levamos em consideração o ano, porque ocorreram casos de uma mesma escola mudar de esfera administrativa e de escolas que encerraram, interromperam ou iniciaram atividades no decorrer do período avaliado.

Na análise dos microdados, utilizamos os *softwares* SPSS, Tableau e os aplicativos MsOffice Access e Excel para recortar e concatenar tabelas, para obter medidas de tendência central e de dispersão, para produzir gráficos e diagramas, para calcular correlações e para produzir tabelas para o sistema de informação geográfica (SIG) (Figura 1).

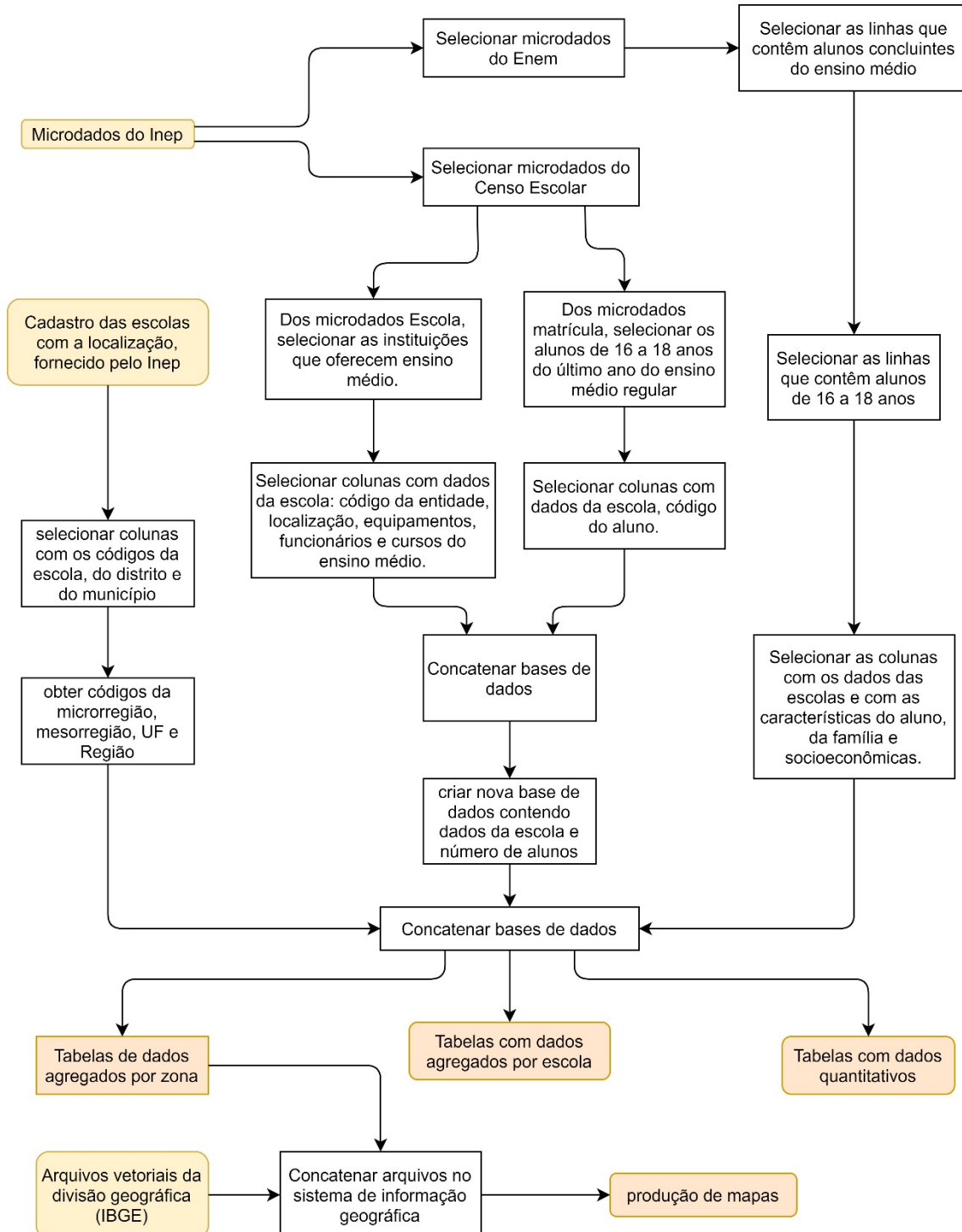
Na análise, também identificamos valores discrepantes por meio de diagramas de caixa, ou “caixas-de-bigodes”, que se baseiam em quartis (Q). Valores acima de $Q_3 + 1,5 \cdot (Q_3 - Q_1)$ e abaixo de $Q_1 - 1,5 \cdot (Q_3 - Q_1)$ são considerados valores atípicos ou discrepantes (BUSSAB; MORETTIN, 2010).

No *software* ArcMap, importamos arquivos vetoriais (*shapefiles*) das malhas territoriais, adquiridos no sítio eletrônico do IBGE.¹¹ Esses arquivos foram concatenados com as tabelas para a produção dos mapas utilizados na avaliação de zoneamentos.

Disponibilizamos dados agrupados e mapas dinâmicos para *download* em <https://public.tableau.com/profile/eduardo.justiniano#!/>.

11 Fonte: <https://portaldemapas.ibge.gov.br/portal.php#homepage>. Acesso em: 27 set. 2019.

Figura 1 – Fluxograma de processamento de dados



Fonte: elaboração do autor.

3.2. Avaliação de zoneamentos

Para avaliar os zoneamentos, empregamos os procedimentos propostos por Openshaw (1984), porém de forma simplificada e adaptada para o propósito desta pesquisa (Quadro 2).

Quadro 2 – Procedimentos para testar zoneamentos

Etapa	Procedimentos
1	Definir o objetivo do teste de forma explícita, especulando qual o resultado esperado, tomando como base o conhecimento prévio. Esse resultado pode ser expresso por uma pergunta ou hipótese e estabelecido como uma função objetivo.
2	Obter o resultado esperado identificando os zoneamentos que satisfaça a função objetivo.
3	Interpretar os resultados no sentido estatístico e verificar se os resultados esperados foram alcançados com grau de erro tolerável. Se não, rejeitar ou alterar a hipótese ou a pergunta e retomar a etapa 1.
4	Verificar se o zoneamento torna visível a interação entre os dados agregados.

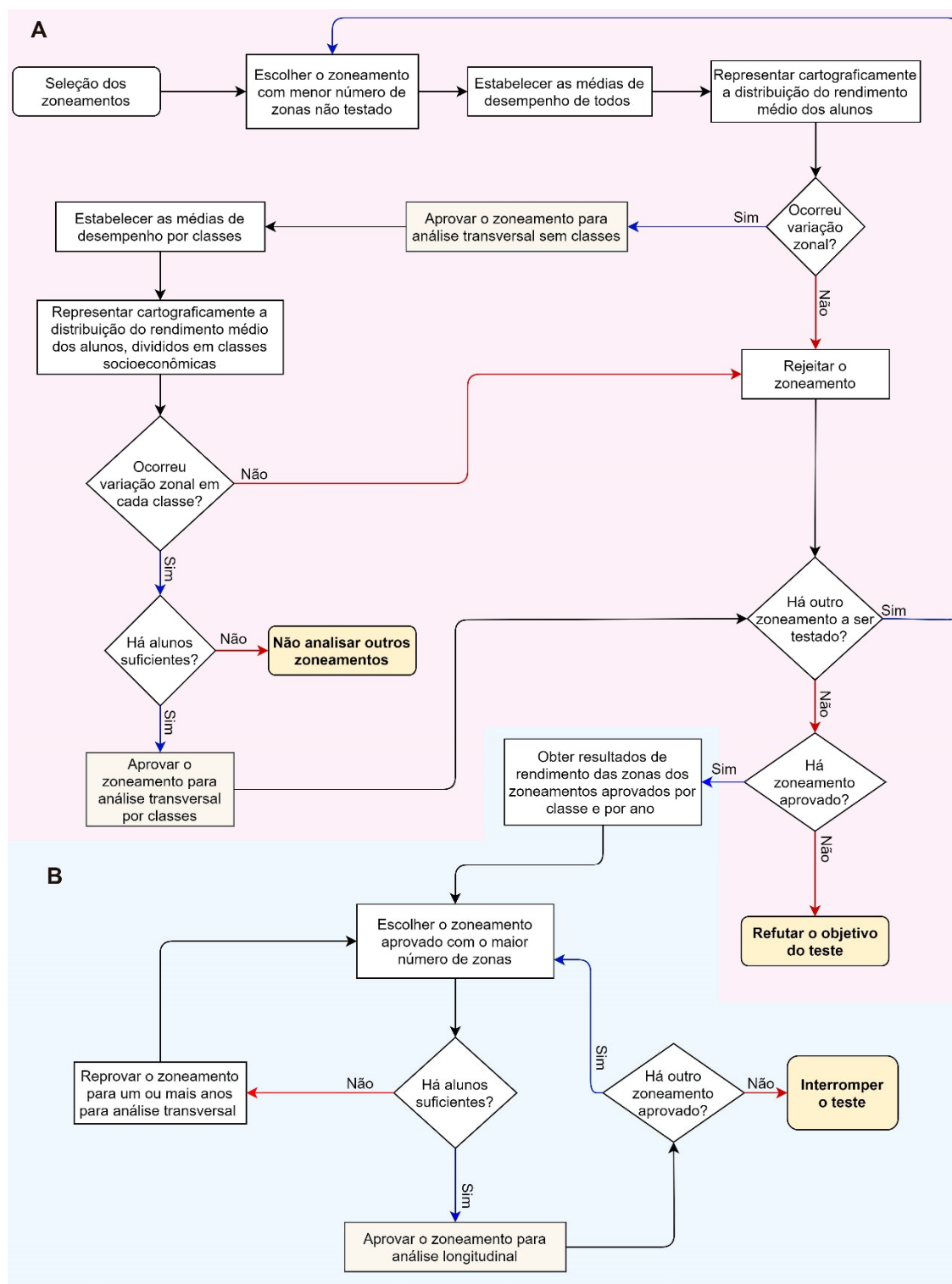
Fonte: adaptado de Openshaw (1984).

O objetivo da avaliação foi identificar zoneamentos eficientes para a representação espacial do desempenho de alunos concluintes do ensino médio no ENEM. Definimos como função objetivo a representação da diferenciação zonal dos resultados obtidos no ENEM no período 2012 a 2018, considerando a influência da variável que apresentou maior correlação com o desempenho médio dos alunos. Esperamos um desempenho melhor dos alunos no ENEM nas zonas que abrangem as dez áreas metropolitanas de maior produto interno bruto.

A diferenciação zonal foi verificada considerando todos os alunos e os alunos por faixa de renda. Na avaliação, organizamos os zoneamentos por ordem de número de zonas e procuramos constatar a variação espacial do desempenho no ENEM. A interpretação dos resultados no sentido estatístico, dos zoneamentos que otimizem a

função objetivo, levou em consideração o número mínimo de alunos por zona, por faixa de renda (Figura 2A)

Figura 2 – Fluxograma dos procedimentos para testar os zoneamentos



Fonte: elaboração do autor.

Para analisar a eficiência dos zoneamentos para representar a variação do desempenho dos concluintes de 2012 a 2018, executamos novamente os procedimentos do Quadro 2. Porém, modificamos a função objetivo para identificar as zonas que tiveram um número mínimo de alunos por ano e faixa de renda e baixo coeficiente de variação dos resultados médios anuais (Figura 2B). Os zoneamentos aprovados nessa segunda análise seriam eficientes também para representar a evolução temporal do rendimento dos alunos, no período observado.

Após a definição dos zoneamentos mais eficientes para representar o desempenho de alunos no ENEM de 2012 a 2018, utilizamos o *software* GeoDa para demonstrar como as variáveis estavam correlacionadas no espaço, por meio do índice global de Moran I e do mapa de espalhamento de Moran.

4. ANÁLISE DESCRITIVA E EVOLUTIVA DOS DADOS

Nesta análise descritiva e evolutiva dos dados, procuramos conhecer o universo dos alunos concluintes do ensino médio por rede de ensino e algumas das características das escolas constantes nos censos escolares de 2012 a 2018. Quantificamos os alunos que se inscreveram e pontuaram no ENEM. Também levantamos suas características socioeconômicas e seus desempenhos no teste educacional que puderam ser correlacionadas.

Os microdados das diversas edições do ENEM possuem muitas variáveis socioeconômicas ou relacionadas com o padrão de consumo dos alunos e de suas famílias. Porém, no período observado, muitas dessas variáveis não puderam ser utilizadas na correlação, porque a pergunta não era a mesma para todos os anos ou porque o item possuído pelas famílias não teria o mesmo valor ao longo do tempo. Em INEP (2013), por exemplo, uma das perguntas do questionário era sobre a quantidade de empregadas mensalistas na casa do inscrito, e em INEP (2019a), a pergunta era sobre quantos dias por semana havia empregado doméstico trabalhando. Outro exemplo pode ser encontrado em INEP (2012): nesse ano, a primeira pergunta do questionário socioeconômico do ENEM 2011 era sobre quantas pessoas moram com você, enquanto nas edições do ENEM de 2012 a 2018, a pergunta era sobre quantas pessoas moram em sua casa/residência (incluindo você). Essa diferença na pergunta inviabilizou incorporar os dados do ENEM 2011 na análise.

Nesse período, 5.561 municípios, 99,8% do total de municípios brasileiros, tiveram escolas que ofereceram o último ano do ensino médio regular, para concluintes de até 18 anos que prestaram o ENEM. Anualmente, o número de municípios com escolas públicas que atenderam esse público oscilou, enquanto o número de municípios com escolas particulares apresentou queda de 2015 a 2018 (Tabela 1).

Tabela 1 – Número de municípios com escolas que ofereceram o último ano do ensino médio regular, dado que tiveram concluintes de até 18 anos que prestaram o ENEM

Ano	Rede de ensino		
	Particular	Pública	Todas
2012	1.469	5.545	5.545
2013	1.481	5.550	5.550
2014	1.495	5.557	5.557
2015	1.484	5.556	5.556
2016	1.477	5.559	5.559
2017	1.459	5.531	5.531
2018	1.449	5.546	5.546

Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

4.1. Universo dos concluintes do ensino médio no ENEM

Nas tabelas de matrícula dos censos escolares contabilizamos o número de alunos concluintes do ensino médio pelo código do aluno (Tabela 2). Ocorreram casos de um mesmo aluno ter mais de uma matrícula, porém em escolas diferentes. Nesses casos, consideramos somente uma das matrículas.

Tabela 2 – Alunos concluintes do ensino médio no Brasil (1.000 alunos)

Ano	Dependência administrativa				Total
	Federal	Estadual	Municipal	Privada	
2012	25,2	1.310,6	15,0	295,1	1.645,8
2013	30,5	1.325,7	13,9	299,8	1.669,8
2014	33,4	1.365,2	12,9	302,9	1.714,3

continua

conclusão

Ano	Dependência administrativa				Total
	Federal	Estadual	Municipal	Privada	
2015	32,9	1.353,9	12,0	295,9	1.694,7
2016	35,4	1.435,1	12,2	302,3	1.784,9
2017	39,8	1.393,2	10,9	293,2	1.737,1
2018	45,1	1.379,0	10,3	266,8	1.701,3
Total	242,5	9.562,8	84,1	2.060,29	11.949,7

Fonte: INEP (2016a, 2016b, 2016c, 2016d, 2016e, 2018b, 2019b).

Somente 72,8% dos concluintes do ensino médio se inscreveram no ENEM. Proporcionalmente, os alunos da rede particular de ensino participaram mais na inscrição (Tabela 3).

Tabela 3 – Alunos concluintes do ensino médio inscritos no ENEM no Brasil (1.000 alunos)

Ano	Dependência administrativa				Total
	Federal	Estadual	Municipal	Privada	
2012	18,6	840,1	9,4	265,0	1.133,0
2013	20,7	897,1	9,4	276,3	1.203,4
2014	23,2	940,6	9,1	281,0	1.253,9
2015	24,8	950,6	8,8	272,5	1.256,7
2016	28,1	1.073,2	9,3	284,0	1.394,6
2017	32,7	1.025,2	8,5	268,4	1.334,8
2018	31,7	850,9	7,3	228,0	1.117,8
Total	179,9	6.577,6	61,8	1.875,1	8.694,3

Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

Dos alunos concluintes que se inscreveram no ENEM, 84,9% pontuaram em todas as provas e na redação (Tabela 4). Eles corresponderam a 61,8% dos concluintes do ensino médio.

Tabela 4 – Alunos concluintes do ensino médio que pontuaram em todas as provas do ENEM (1.000 alunos)

Ano	Dependência administrativa				Total
	Federal	Estadual	Municipal	Privada	
2012	16,5	687,3	7,9	249,0	960,7
2013	19,3	756,4	8,2	261,5	1.045,4
2014	21,2	774,7	7,9	265,4	1.070,0
2015	22,8	831,6	7,9	262,0	1.124,8
2016	25,6	884,6	8,1	269,1	1.188,1
2017	29,2	785,9	7,0	251,3	1.074,2
2018	30,1	666,8	6,0	216,6	920,5
Total	164,6	5.387,4	53,0	1.775,0	7.383,7

Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

Os alunos das escolas de dependências administrativas municipal e federal corresponderam a uma pequena parcela de todos os inscritos. Por isso, optamos por agrupá-los na rede pública de ensino, junto com os alunos das escolas estaduais. Para padronizar, optamos por denominar alunos da rede particular os estudantes de escolas da dependência administrativa privada.

4.2. Variáveis socioeconômicas

Dos microdados do ENEM de 2012 a 2018, selecionamos as colunas que contêm informações que possam ser comparadas ao longo do tempo e que foram importantes para o desenvolvimento desta pesquisa (Quadro 3).

Quadro 3 – Dados utilizados

Campo	Utilização
Número de inscrição	Verificação de duplicidade e concatenação de tabelas.
Ano do ENEM	Separação dos dados para a análise temporal dos resultados de 2012 a 2018.
Idade	Recorte dos dados da pesquisa.
<i>Status</i> de conclusão	Seleção dos alunos do último ano do ensino médio.
Código da escola	Vinculação do aluno com a escola georreferenciada, possibilitando análise espacial.
Código do município da escola	Recorte dos dados para contemplar somente os alunos de escolas localizadas no município de São Paulo.
Dependência administrativa	Recorte dos dados e comparação dos resultados entre os alunos de diferentes dependências administrativas.
Presença nas provas	Quantificação dos alunos inscritos e dos que pontuaram em todas as provas.
Notas obtidas	Cálculo da pontuação média dos alunos por escola e por zona.
<i>Status</i> da redação	Obtenção do percentual de alunos que pontuaram na redação.
Nota da redação	Cálculo da pontuação média dos alunos por escola e por zona.
Moradores	Obtenção da correlação entre número de moradores e nota obtida.

continua

conclusão

Campo	Utilização
Renda familiar	Obtenção da correlação entre a renda familiar e nota.
Estudo dos pais	Obtenção da correlação entre a formação escolar dos pais e nota.

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A renda familiar dos inscritos, de 2012 a 2018, está dividida em 17 faixas (Tabela 5). A diferença entre os números totais de alunos para com os da Tabela 4 explica-se pela ocorrência de valores nulos. Na faixa de renda acima de 20 salários mínimos, encontraram-se 0,11% dos alunos da rede pública e 6,85% dos alunos da rede particular.

Tabela 5 – Alunos concluintes do ensino médio de até 18 anos no Brasil no ENEM (2012-2018) que pontuaram em todas as provas e na redação por faixa de renda familiar e por rede de ensino (1.000 alunos)

Renda familiar (SM)	Rede de ensino		Total
	Pública	Particular	
Sem renda	138,6	8,2	146,8
Acima de 0 até 1	1.574,1	74,7	1.648,8
Acima de 1 até 1,5	1.321,0	135,8	1.456,8
Acima de 1,5 até 2	745,9	128,0	873,8
Acima de 2 até 2,5	514,5	118,8	633,4
Acima de 2,5 até 3	382,0	128,2	510,3
Acima de 3 até 4	372,8	172,2	545,0
Acima de 4 até 5	216,5	163,0	379,4
Acima de 5 até 6	130,8	143,1	273,9
Acima de 6 até 7	72,2	114,6	186,7
Acima de 7 até 8	42,5	84,0	126,5
Acima de 8 até 9	26,7	69,7	96,4

continua

conclusão

Renda familiar (SM)	Rede de ensino		Total
	Pública	Particular	
Acima de 9 até 10	19,2	64,4	83,6
Acima de 10 até 12	21,0	91,5	112,5
Acima de 12 até 15	13,2	82,8	95,9
Acima de 15 até 20	7,7	74,5	82,2
Acima de 20	6,2	121,5	127,8
Total	5.604,8	1.775,0	7.379,8

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Para definir a renda familiar *per capita*, optamos pelo valor médio de cada faixa de renda e dividimos pelo número de habitantes (Tabela 6). Excluimos os concluintes de renda familiar acima de 20 salários mínimo pela impossibilidade de estabelecer valor médio para essa faixa de renda.

Tabela 6 – Alunos concluintes do ensino médio de até 18 anos que pontuaram em todas as provas do ENEM de 2012 a 2018 por faixa de renda familiar *per capita*, dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).

Renda familiar <i>per capita</i> (SM)	Rede de ensino		Total
	Pública	Particular	
Sem renda	138,6	8,2	146,8
Acima de 0 até 0,2	1.509,6	67,2	1.576,8
Acima de 0,2 até 0,4	1.291,1	124,9	1.416,1
Acima de 0,4 até 0,6	1.226,8	227,6	1.454,4
Acima de 0,6 até 0,8	482,0	142,3	624,3
Acima de 0,8 até 1,0	350,5	156,7	507,2
Acima de 1,0 até 1,2	236,3	160,4	396,7
Acima de 1,2 até 1,4	97,4	98,8	196,2

continua

conclusão

Renda familiar <i>per capita</i> (SM)	Rede de ensino		Total
	Pública	Particular	
Acima de 1,4 até 1,6	56,9	65,3	122,2
Acima de 1,6 até 1,8	52,2	76,7	128,9
Acima de 1,8 até 2,0	50,3	92,1	142,4
Acima de 2,0 até 2,2	30,7	78,6	109,3
Acima de 2,2 até 2,4	16,8	44,6	61,4
Acima de 2,4 até 2,6	8,0	22,7	30,7
Acima de 2,6 até 2,8	17,0	66,9	83,9
Acima de 2,8 até 3,0	5,4	20,7	26,0
Acima de 3,0	29,0	199,8	228,8
Total	5.598,6	1.653,4	7.252,0

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Optamos por reagrupar os concluintes da rede pública em quatro faixas de renda familiar *per capita* (Tabela 7), a partir do agrupamento das faixas da Tabela 6.

Tabela 7 – Alunos concluintes do ensino médio da rede pública que pontuaram em todas as provas do ENEM, de 2012 a 2018, por faixas de renda *per capita*, dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).

Renda <i>per capita</i> (SM)	Alunos	Frequência relativa (%)
até 0,2	1.648,2	29,4
Acima de 0,2 até 0,4	1.291,1	23,1
Acima de 0,4 até 0,8	1.708,8	30,5
Acima de 0,8	950,4	17,0

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Essa reclassificação em quatro faixas de renda familiar *per capita* não se mostrou adequada para os concluintes da rede particular, pois a maior parte deles estava na faixa superior (Tabela 8). Além disso, 12,2% desses concluintes estavam nas faixas até 0,4 SM, enquanto a participação dos alunos da rede pública nessas faixas era 52,5%.

Tabela 8 – Alunos concluintes do ensino médio de até 18 anos da rede particular que pontuaram em todas as provas do ENEM, de 2012 a 2018, por faixas de renda *per capita*, dado que a renda familiar não foi superior a 20 salários mínimos (1.000 alunos).

Renda <i>per capita</i> (SM)	Alunos	Frequência relativa (%)
até 0,2	75,4	4,6
Acima de 0,2 até 0,4	124,9	7,6
Acima de 0,4 até 0,8	369,9	22,4
Acima de 0,8	1.083,3	65,5

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Consideramos que o limite superior de renda *per capita* diário para ser considerado pobre foi US\$ 5,50, enquanto US\$ 1,90 foi o limite para ser extremamente pobre (Tabela 9). Dessa forma, os concluintes de renda familiar *per capita* de até 0,2 SM seriam extremamente pobres. Os de renda entre 0,2 e 0,4 seriam pobres ou extremamente pobres, de 0,4 a 0,8 foram pobres, e os de renda superior a esse limite foram não pobres.

Tabela 9 – Equivalência entre dólar e real e condição econômica

Ano	Cotação no primeiro dia útil de fevereiro			Limites econômicos em SM	
	SM (R\$)	dólar	SM (US\$)	Extrema pobreza	Pobreza
2012	622,00	1,74	358,09	0,2	0,5
2013	678,00	1,98	341,77	0,2	0,5
2014	724,00	2,41	300,61	0,2	0,5

continua

Ano	Cotação no primeiro dia útil de fevereiro			Limites econômicos em SM	
	SM (R\$)	dólar	SM (US\$)	Extrema pobreza	Pobreza
2015	788,00	2,69	293,07	0,2	0,6
2016	880,00	4,00	220,12	0,3	0,8
2017	937,00	3,15	297,72	0,2	0,6
2018	954,00	3,17	300,95	0,2	0,5

Fonte: Ipea;¹² World Bank.¹³

4.3. Desempenho dos concluintes do ensino médio

A avaliação das competências e das habilidades dos alunos foi mensurada por provas objetivas de cada área do conhecimento (linguagem e códigos e suas tecnologias (LC), matemática e suas tecnologias (MT), ciências humanas e suas tecnologias (CH) e ciências da natureza e suas tecnologias (CN), já mencionadas) e pela redação, cujas pontuações variam de 0 a 1.000.

A comparação da pontuação média anual dos alunos concluintes do ensino médio que pontuaram em todas as provas e na redação ao longo do tempo pode ser feita por meio dos resultados de cada prova ou pelo conjunto de todas as provas, podendo ser pela média ou pela somatória das pontuações das provas.

Por um lado, os alunos concluintes da rede particular obtiveram pontuações médias sempre superiores a 500 pontos por ano e por prova (Tabela 10).

¹² Fonte: <http://www.ipeadata.gov.br/>. Acesso em: 10 fev. 2020.

¹³ Disponível em: <http://iresearch.worldbank.org/PovcalNet/methodology.aspx>.

Tabela 10 – Pontuação média dos concluintes do ensino médio da rede particular

Ano	CN	CH	LC	MT	RED	
2012	586,8	544,1	547,0	619,6	606,9	Médias
2013	584,8	542,1	549,4	605,4	619,6	
2014	603,1	552,5	562,7	573,2	625,8	
2015	610,8	545,0	560,4	569,7	652,2	
2016	595,6	552,0	572,4	583,9	664,6	
2017	583,9	570,2	558,3	609,9	668,4	
2018	625,6	559,5	581,5	624,6	675,7	
Amplitude	41,7	28,0	34,5	54,9	68,8	Medidas de dispersão
Média	598,7	552,2	561,7	598,1	644,7	
Desvio padrão	14,4	9,2	11,3	20,7	25,1	
Coeficiente de variação (x100)	2,4	1,7	2,0	3,5	3,9	

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Por outro lado, 45,7% das pontuações médias de cada prova por ano foi inferior a 500 pontos para os concluintes da rede pública (Tabela 11).

Tabela 11 – Pontuação média dos alunos concluintes do ensino médio da rede pública

Ano	CN	CH	LC	MT	RED	-
2012	509,6	458,7	484,2	495,6	497,1	Médias
2013	503,2	461,3	482,0	501,5	511,7	
2014	541,2	477,0	508,4	466,7	496,5	
2015	548,1	466,0	497,8	459,6	533,9	
2016	527,9	468,8	516,6	476,9	539,0	
2017	509,8	496,7	502,5	507,5	551,5	
2018	557,7	480,9	518,4	522,2	509,9	

continua

conclusão

Ano	CN	CH	LC	MT	RED	-
Amplitude	54,5	38,0	36,4	62,6	54,9	Medidas de dispersão
Média	528,2	472,8	501,4	490,0	519,9	
Desvio padrão	19,8	12,2	13,4	21,2	20,0	
Coeficiente de variação (x100)	3,7	2,6	2,7	4,3	3,8	

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A variável pontuação média de todas as provas apresentou menor variação anual. O valor dos coeficientes de variação por rede de ensino foi menor do que os encontrados na Tabela 10 e na Tabela 11, exceto para CN para os concluintes da rede particular (Tabela 12).

Tabela 12 – Pontuação média no ENEM dos alunos concluintes do ensino médio no Brasil, por rede de ensino (2012-2018)

Ano	Rede de ensino			-
	Pública	Particular	Todas	
2012	489,0	580,9	512,8	Médias
2013	491,9	580,3	514,0	
2014	498,0	583,4	519,2	
2015	501,1	587,6	521,3	
2016	505,8	593,7	525,8	
2017	513,6	598,1	533,4	
2018	517,8	613,4	540,3	
Amplitude	28,8	33,1	27,5	
Média	502,5	591,1	523,8	
Desvio padrão	9,9	11,0	9,4	
Coeficiente de variação (x100)	2,0	1,9	1,8	

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A pontuação média de todos os concluintes do ensino médio de até 18 anos, sem agregação por ano, foi superior à mediana, indicando que a maioria dos alunos obteve pontuação inferior à média, de 2012 a 2018. Os valores da amplitude, do desvio padrão e o coeficiente de variação foram elevados (Tabela 13).

Tabela 13 – Resumo estatístico das pontuações dos alunos concluintes do ensino médio de até 18 anos no ENEM (2012-2018)

Medida	Pública	Particular	Todas
Média	502,5	590,6	523,7
Mediana	497,0	589,7	514,2
Amplitude	567,5	562,5	586,6
Desvio padrão	65,2	78,6	78,3
Coeficiente de variação (x100)	13,0	13,3	14,9

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Para comprovar a proficiência, o inscrito deveria obter no mínimo 450 pontos em cada área do conhecimento e 500 pontos na redação (INEP, 2014). Classificamos os alunos pelo número de proficiências alcançadas, variando de 0 a 5. Consideramos proficiente o inscrito que obteve pelo menos a pontuação mínima em todas as provas. Para ser considerado parcialmente proficiente, o inscrito alcançou a pontuação mínima em uma a quatro provas. Não proficiente foi o inscrito que não alcançou a pontuação mínima em nenhuma prova.

Apesar de ter ocorrido um aumento contínuo das pontuações médias (Tabela 12), não houve uma melhoria constante na proficiência dos alunos da rede particular de ensino (Tabela 14).

Tabela 14 – Percentual de concluintes de até 18 anos da rede particular por número de proficiências obtidas no ENEM (2012-2018)

Ano	Não proficiente	Parcialmente proficiente				Proficiente
	0	1	2	3	4	5
2012	0,9	2,1	3,9	7,5	19,9	65,7
2013	0,7	1,9	3,8	7,6	19,6	66,4
2014	0,4	1,2	3,4	8,5	21,2	65,4
2015	0,3	1,1	3,1	8,4	20,7	66,4
2016	0,2	0,9	2,6	7,4	19,0	69,9
2017	0,2	0,7	2,0	5,1	14,8	77,2
2018	0,1	0,6	2,0	5,9	17,8	73,6

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Os concluintes da rede pública mostraram-se menos proficientes que os da rede particular. Também não ocorreu melhoria constante da proficiência plena ou parcial desses alunos no ENEM (Tabela 15).

Tabela 15 – Percentual de alunos concluintes de até 18 anos da rede pública por número de proficiências obtidas no ENEM (2012-2018)

Ano	Não proficiente	Parcialmente proficiente				Proficiente
	0	1	2	3	4	5
2012	7,3	11,5	14,8	18,0	22,8	25,6
2013	5,6	11,2	15,0	18,6	23,6	26,1
2014	2,7	7,6	15,1	22,3	26,6	25,8
2015	2,5	7,9	14,8	23,1	26,8	24,9
2016	2,2	7,0	14,2	22,8	26,9	26,9

continua

Ano	conclusão					
	Não proficiente	Parcialmente proficiente				Proficiente
	0	1	2	3	4	5
2017	2,0	6,2	11,3	17,1	25,4	38,0
2018	1,1	5,2	12,1	20,1	27,9	33,5

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A maior parte dos alunos concluintes da rede pública de até 18 anos que pontuaram em todas as provas declarou ser de cor/raça branca ou parda. A variação do número de alunos de cada cor/raça também está relacionada com a quantidade de inscritos em cada ano (Tabela 16).

Tabela 16 – Número de alunos concluintes do ensino médio da rede pública que pontuaram em todas as provas do ENEM de 2012 a 2018, segundo declaração de cor/raça (1.000 alunos)

Ano	Cor / Raça					
	Não declarada	Branca	Preta	Parda	Amarela	Indígena
2012	10,7	320,6	69,6	291,9	14,3	4,4
2013	10,7	334,7	81,9	335,7	16,0	4,9
2014	9,7	336,7	84,5	352,6	15,5	4,7
2015	9,8	335,3	97,7	395,3	17,4	4,6
2016	13,3	349,7	107,3	423,4	19,2	5,4
2017	14,4	302,2	95,7	386,3	18,4	5,1
2018	12,5	250,3	81,5	338,5	15,5	4,6

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A partir de 2012, ser estudante de escola pública, preto, pardo ou indígena e de baixa renda, dependendo da unidade da Federação, facilitou o acesso à universidade pública. Talvez por esse motivo, ocorreu um aumento percentual do número de inscritos de escolas públicas que se declararam pretos e pardos e uma diminuição da participação de alunos que declararam ser brancos (Tabela 17). Por esse motivo, optamos por não considerar a variável cor/raça nas análises.

Tabela 17 – Percentual de alunos concluintes do ensino médio da rede pública no Brasil que pontuaram em todas as provas no ENEM, segundo declaração de cor/raça

Ano	Cor / Raça					
	Não declarada	Branca	Preta	Parda	Amarela	Indígena
2012	1,5	45,1	9,8	41,0	2,0	0,6
2013	1,4	42,7	10,4	42,8	2,0	0,6
2014	1,2	41,9	10,5	43,9	1,9	0,6
2015	1,1	39,0	11,4	46,0	2,0	0,5
2016	1,4	38,1	11,7	46,1	2,1	0,6
2017	1,8	36,8	11,6	47,0	2,2	0,6
2018	1,8	35,6	11,6	48,2	2,2	0,7

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O grau de instrução dos pais foi dividido em nove categorias nos microdados das edições do ENEM de 2012 a 2014 e em oito categorias nas edições de 2015 a 2018. As principais diferenças estavam relacionadas com o ensino médio e o curso superior. Além disso, as faixas de estudos mostraram-se confusas. Apesar desses problemas, atribuímos um valor médio de anos de estudo para cada classe para possibilitar a correlação entre as variáveis pontuação e grau de instrução dos pais, considerando que a duração do ensino fundamental foi nove anos e que a duração do curso superior era de quatro anos (Quadro 4).

Quadro 4 – Grau de instrução dos pais segundo os microdados do ENEM

2012 a 2014	2015 a 2018	Anos de estudo
Não estudou.	Nunca estudou.	0
Da 1ª à 4ª série do ensino fundamental (antigo primário).	Não completou a 4ª série/5º ano do ensino fundamental.	2,5
Da 5ª à 8ª série do ensino fundamental (antigo ginásio).	Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano do ensino fundamental.	7
Ensino médio (antigo 2º grau) incompleto.	Completou o ensino fundamental, mas não completou o ensino médio.	10,5
Ensino médio (antigo 2º grau).	-	12
Ensino superior incompleto.	Completou o ensino médio, mas não completou a faculdade.	14
Ensino Superior.	-	16
-	Completou a faculdade, mas não completou a pós-graduação.	17,5
Pós-graduação.	Completou a pós-graduação.	19
Não sabe.	Não sabe.	Nulo

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Nos microdados do ENEM, o número de habitantes na residência não está dividido em categorias, possibilitando correlacioná-lo com a pontuação dos alunos concluintes sem agrupamento.

4.4. Correlações

A quantificação do grau de associação entre duas variáveis é feita pelo coeficiente de correlação que descreve por um único número a associação ou dependência entre duas variáveis (BUSSAB; MORETTIN, 2010). A identificação do grau de dependência entre variáveis pode auxiliar na formulação de hipóteses sobre a dependência espacial entre fenômenos geográficos (FERREIRA, 2014).

O coeficiente de correlação é dado pela fórmula

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right] \left[\sum_{i=1}^n (y_i - \bar{y})^2 \right]}}$$

onde \bar{x} e \bar{y} são médias aritméticas das variáveis x e y , respectivamente, e o coeficiente de correlação (r) varia de 1 a -1. Acima de zero, quanto maior for o valor de r , maior será a correlação positiva entre as variáveis; abaixo de zero, quanto menor for o valor de r , maior será a correlação negativa. Os valores próximos de 0 indicam a não correlação entre as variáveis.

Para os alunos concluintes do ensino médio de 16 a 18 anos da rede pública determinamos os coeficientes de correlação individual, por escola e por zonas, dos alunos cuja renda familiar declarada não foi superior a 20 SM.

Observamos uma forte correlação individual positiva entre a pontuação e a proficiência (Tabela 18). Apesar da baixa diferença encontrada nos valores de correlação individual entre pontuação e as variáveis renda familiar *per capita* e renda familiar, optamos por utilizar a primeira nas avaliações dos zoneamentos.

Tabela 18 – Correlações individuais com a variável pontuação dos alunos concluintes do ensino médio de até 18 anos da rede pública (2012-2018).

Variável	Proficiência
Proficiências alcançadas	0,836201
Renda <i>per capita</i>	0,284918
Renda familiar	0,280623
Anos de estudo - Pai	0,249858
Anos de estudo - Mãe	0,228373
Moradores	-0,103530

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Agrupamos alunos por escola e estabelecemos valores médios das variáveis da Tabela 18. Também por escola, obtivemos a participação dos alunos no ENEM, o percentual de alunos por sexo, a quantidade de alunos e os números de computadores, funcionários e equipamentos multimídia. Posteriormente, correlacionamos essas variáveis com a pontuação média (Tabela 19).

Tabela 19 – Variáveis correlacionadas com a pontuação no ENEM dos alunos concluintes da rede pública de até 18 anos e variáveis agrupadas por escolas (2012-2018)

Variável	Correlação
Proficiência média	0,950339
Renda <i>per capita</i> média	0,602034
Renda familiar média	0,584255
Média de moradores por família	-0,410568
Média dos anos de estudo dos pais	0,493646
Média dos anos de estudo das mães	0,435501
Participação no exame	0,394337
Número de alunos concluintes da escola	0,152946

continua

conclusão

Variável	Correlação
Porcentagem de alunos homens na escola	0,079376
Porcentagem de alunas mulheres na escola	-0,079376
Número de computadores para alunos	0,443638
Número de funcionários da escola	0,272662
Número de equipamentos multimídia	0,372857

Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

Ser ou não proficiente em uma ou mais áreas do conhecimento foi um parâmetro estabelecido pelo MEC, com base no desempenho no ENEM. Em razão desse estabelecimento e da alta correlação com a pontuação, optamos por utilizar a proficiência média na avaliação dos zoneamentos.

Os resultados médios das proficiências por faixa de renda familiar *per capita* evidenciaram a correlação entre essas variáveis em todos os anos, pois em todos os anos a média dos alunos aumentou da faixa de renda menor para a maior (Tabela 20).

Tabela 20 – Proficiências médias dos concluintes do ensino médio da rede pública de até 18 anos por faixa de renda *per capita* e por ano

Ano	Até 0,2	Acima de 0,2 até 0,4	Acima de 0,4 até 0,8	Acima de 0,8
2012	2,57	3,03	3,37	3,77
2013	2,70	3,15	3,47	3,85
2015	2,91	3,26	3,56	3,91
2014	2,96	3,34	3,61	3,94
2016	3,04	3,36	3,64	3,98
2017	3,30	3,66	3,92	4,24
2018	3,29	3,62	3,88	4,20

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Avaliamos a eficiência de cinco zoneamentos: regiões, unidades federativas, mesorregiões, microrregiões e municípios. As correlações ecológicas entre as variáveis proficiência e renda familiar *per capita* foram fortes e positivas em todos os zoneamentos e, de maneira geral, ocorreu um aumento da força da correlação entre essas variáveis na medida em que o número de zonas diminui. As correlações individuais em cada zona foram mais fracas do que as correlações ecológicas, podendo ser inexistentes ou negativas (Tabela 21).

Tabela 21 – Correlações ecológicas e individual entre proficiência no ENEM dos alunos concluintes da rede pública de até 18 anos e renda familiar *per capita* (2012-2018)

Zoneamento	Zonas	Zonas com escolas	Correlação ecológica	Correlação individual por zona
Regiões	5	5	0,933589	0,244213 a 0,277545
UF	27	27	0,814791	0,200971 a 0,355581
Mesorregiões	137	137	0,776775	0,091754 a 0,391367
Microrregiões	558	558	0,761100	0,046786 a 0,454123
Municípios	5.570	5.561	0,662224	-0,304343 a 0,533128

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

5. RESULTADOS

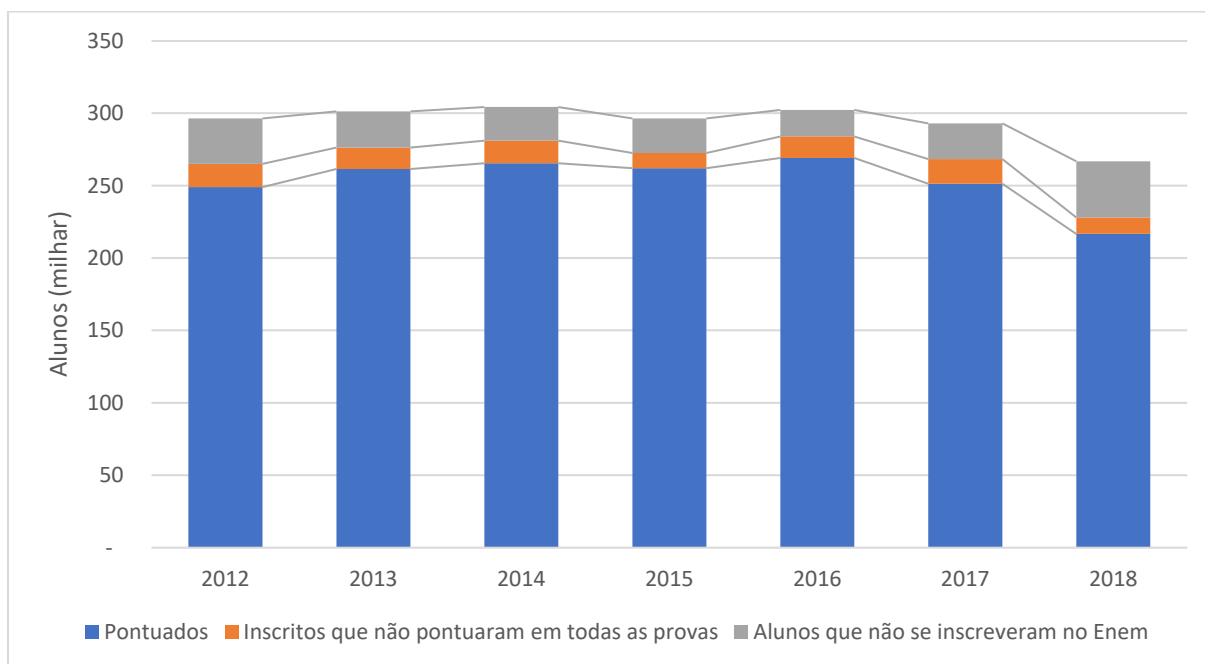
A análise descritiva e evolutiva dos dados possibilitou quantificar a participação dos alunos concluintes do ensino médio de até 18 anos no ENEM. Também forneceu parâmetros para a escolha da medida de desempenho dos concluintes no ENEM para verificar se os alunos da rede pública podem ser analisados em conjunto com os da rede particular e para identificar as variáveis que mais se correlacionam com esse desempenho no teste educacional.

5.1. Participação no ENEM

Ocorreu uma baixa variação do número de concluintes do ensino médio de até 18 anos da rede particular de 2012 a 2016. Em 2017 e 2018 diminuíram o número de concluintes e a participação deles no exame. No entanto, a maioria dos estudantes da rede particular participou no ENEM (Figura 3).

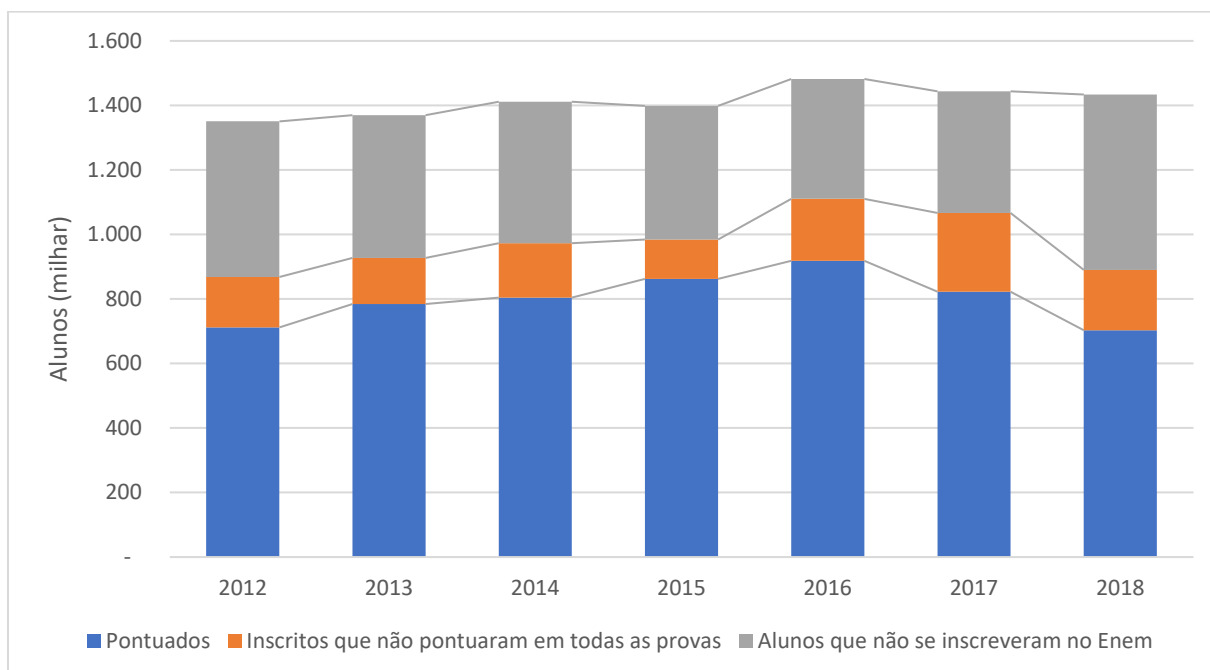
De 2012 a 2017, a maioria dos concluintes do ensino médio de até 18 anos da rede pública se inscreveu e pontuou em todas as provas do ENEM. Em 2018, a quantidade de não inscritos e dos que faltaram ou foram eliminados na prova foi superior ao número dos que pontuaram em todas as provas e na redação (Figura 4).

Figura 3 – Alunos concluintes da rede particular de acordo com a participação no ENEM



Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

Figura 4 – Alunos concluintes da rede pública de acordo com a participação no ENEM



Fonte: INEP (2013, 2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e, 2017a, 2017b, 2018a, 2018b, 2019a, 2019b).

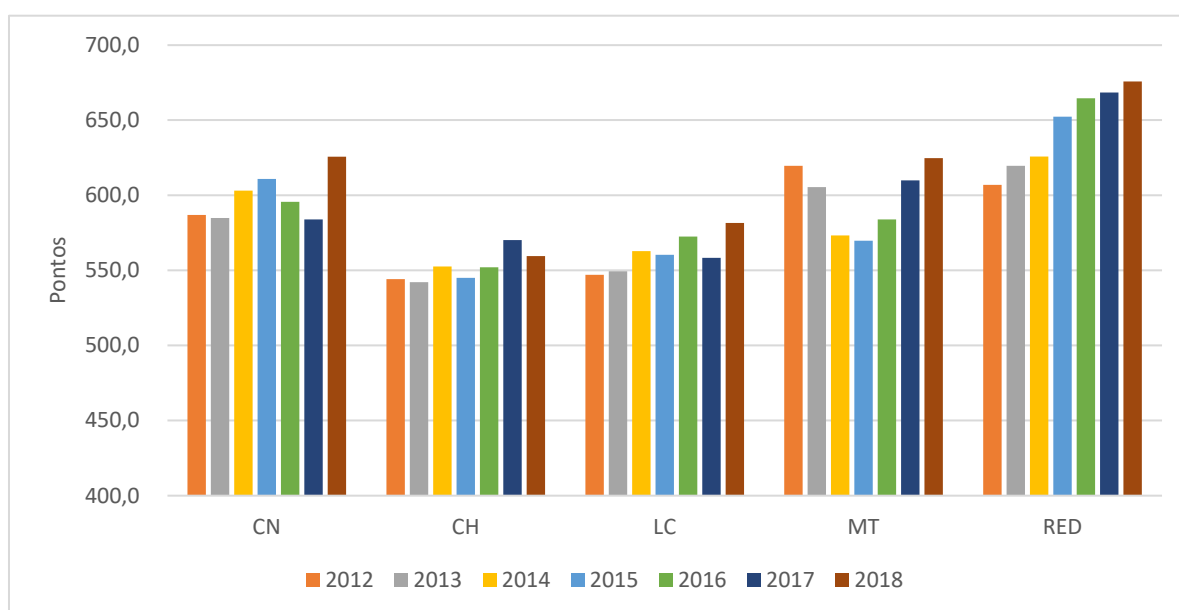
Dos concluintes do ensino médio de até 18 anos da rede pública de 2012 a 2018, 56,7% pontuaram no ENEM, enquanto a participação dos alunos concluintes da

rede particular foi 83,2%. Apesar disso, 75,9% dos concluintes do ensino médio de até 18 anos que prestaram o ENEM eram da rede pública.

5.2. Desempenho no ENEM

As pontuações médias anuais de cada área do conhecimento não se mostraram uniformes para os concluintes da rede particular, principalmente nas avaliações de matemática e da redação (Figura 5). Essa variação da pontuação indicaria a dificuldade de comparação de resultados de um ano para o outro.

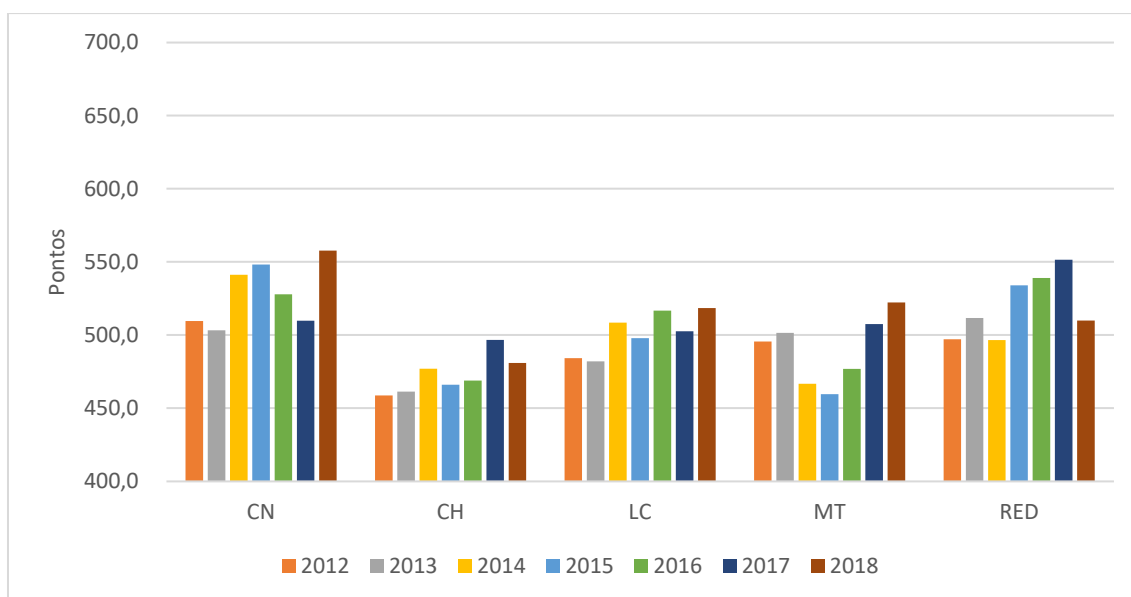
Figura 5 – Pontuações médias dos alunos concluintes de até 18 anos do ensino médio da rede particular por área do conhecimento



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

As pontuações médias anuais por área do conhecimento dos concluintes da rede pública foram ainda mais irregulares (Figura 6). Essa maior variabilidade pode ser confirmada pela diferença de valores dos coeficientes de variação para cada área do conhecimento (Tabela 11 e Tabela 12).

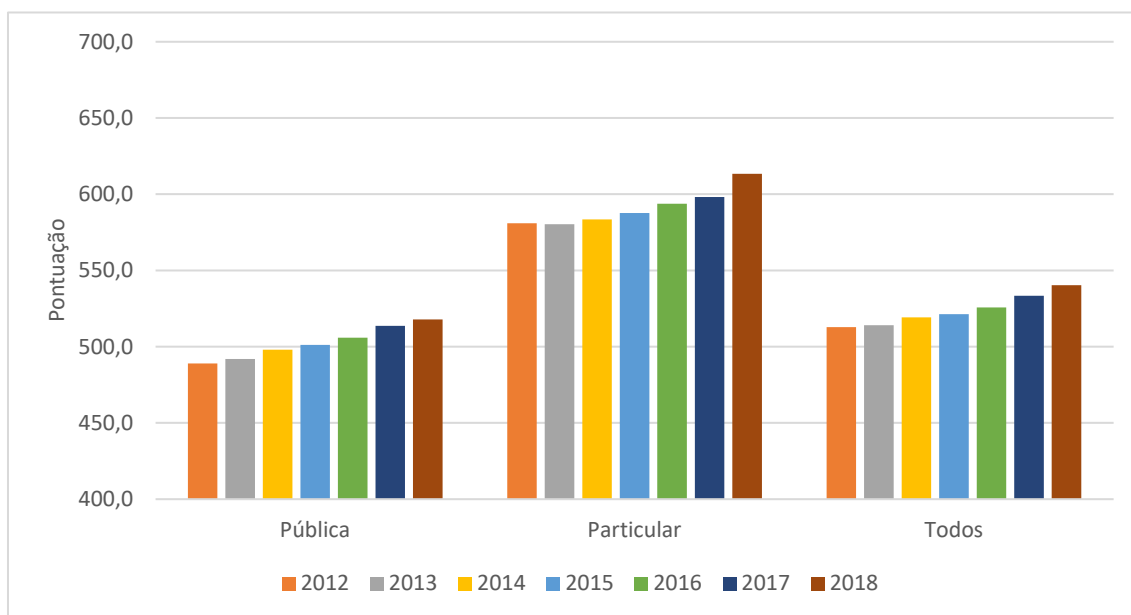
Figura 6 – Pontuações médias dos alunos concluintes do ensino médio de até 18 anos da rede pública por área do conhecimento



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A variação das pontuações médias anuais de todas as áreas do conhecimento foi menor, apesar de ter ocorrido um aumento das pontuações de 2012 a 2018 (Figura 7). Assim, essa medida seria mais apropriada para a comparação de resultados de um ano para o outro.

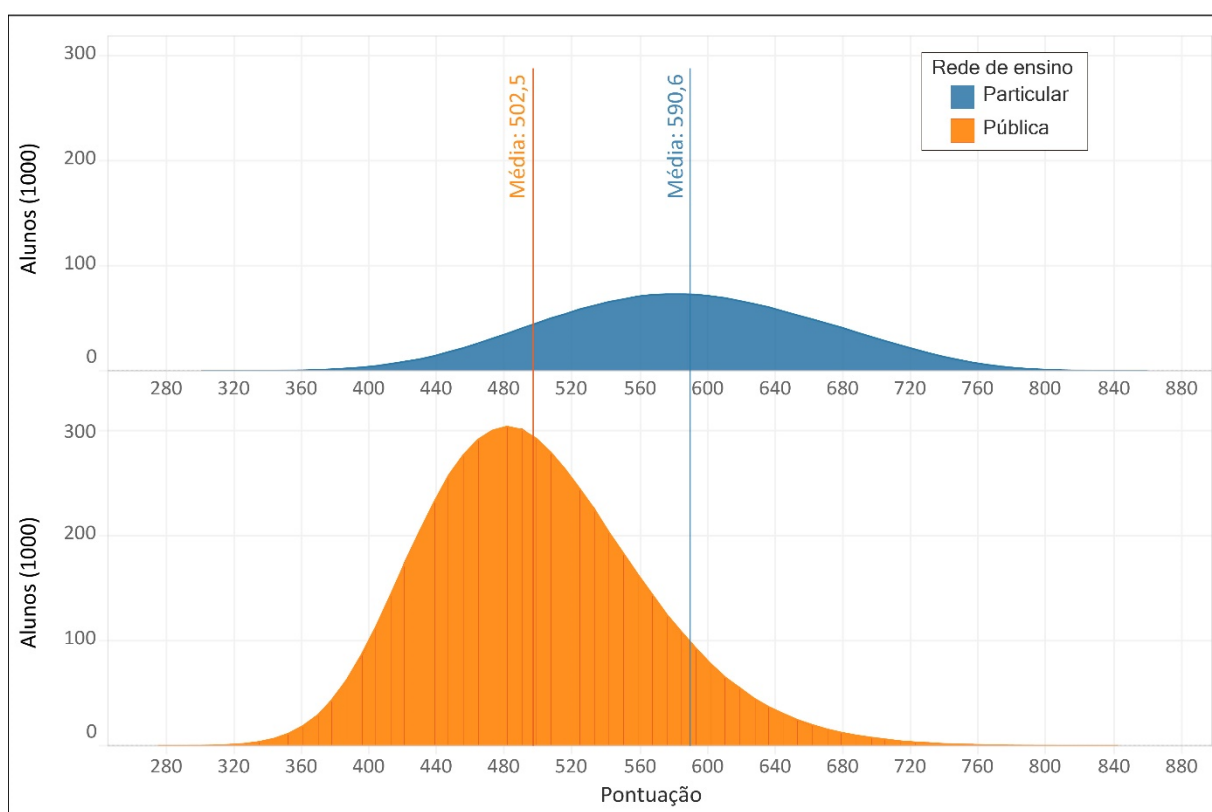
Figura 7 – Pontuações médias no ENEM dos concluintes do ensino médio de até 18 anos



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

As altas amplitudes das pontuações médias por rede de ensino, demonstradas pelas caudas longas dos gráficos de distribuição de frequência (Figura 8), indicaram a ocorrência de concluintes da rede particular com pontuações inferiores à média dos concluintes da rede pública. Também indicou que parte dos concluintes da rede pública obteve pontuação superior à pontuação média dos alunos concluintes da rede particular.

Figura 8 – Histograma da pontuação média dos concluintes do ensino médio de até 18 anos (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Um dos objetivos do ENEM é fornecer subsídios para o acesso à educação superior (BRASIL, 1998). Nesse caso, a pontuação suficiente no exame para cumprir esse objetivo pode não ser homogêneo para todo o Brasil.

Pelo Sistema de Seleção Unificada (Sisu), uma pontuação pode garantir o ingresso em um curso superior de uma instituição pública de ensino superior e não ser suficiente para ingressar em outra instituição. Por exemplo, a nota mínima para o ingresso num curso de Geografia variou em 2018 de 528,74 pontos (Universidade

Federal de Mato Grosso) a 747,86 pontos (Instituto Federal de Educação, Ciência e Tecnologia do Ceará).¹⁴

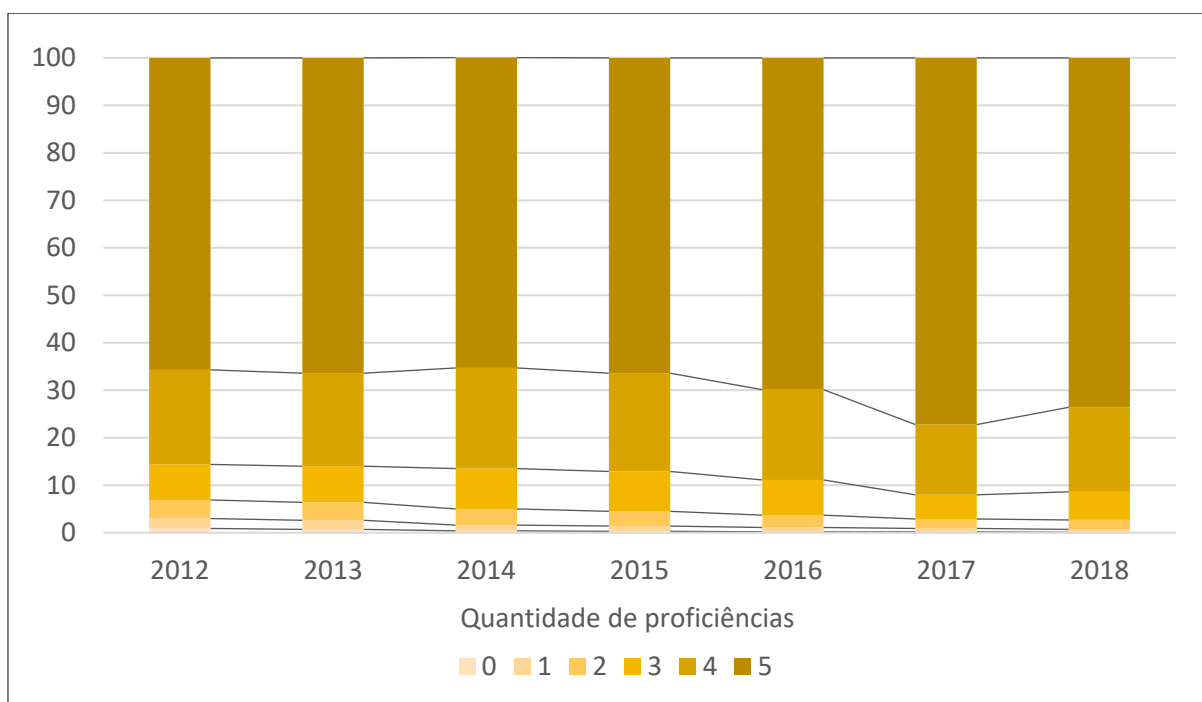
Outro objetivo do ENEM é a criação de uma referência nacional dos alunos concluintes do ensino médio. Optamos por utilizar a proficiência dos concluintes, mensurada por meio dos resultados no teste educacional, como uma referência para comparar espacial e temporalmente o desempenho dos concluintes do ensino médio.

No exame são avaliadas quatro áreas do conhecimento (CH, CN, MT e LC) já mencionadas e a redação, que faz parte da avaliação de LC. Para comprovar a proficiência, o inscrito deveria obter no mínimo 450 pontos em cada área do conhecimento e 500 pontos na redação (INEP, 2014). Como já explicamos em relação à proficiência, ela ocorre quando o aluno obtém a pontuação mínima em todas as provas, e a proficiência parcial, quando o aluno obtém a pontuação mínima em 1, 2, 3 ou 4 provas. Dessa forma, classificamos cada aluno por número de proficiências alcançadas, variando de 0 a 5.

Apesar do aumento constante da pontuação média dos concluintes do ensino médio de até 18 anos, de 2012 a 2018 (Figura 7), a evolução da proficiência não ocorreu da mesma forma. Houve uma diminuição do percentual de alunos da rede particular que demonstraram ser proficientes em até três provas. No entanto, variou positiva e negativamente o percentual de alunos plenamente proficientes e parcialmente proficientes que obtiveram a pontuação mínima em quatro provas (Figura 9).

14 Simulador Sisu: http://portalbi.mec.gov.br/portal-bi/relatorios/simulador_sisu/painel_direto.php. Acesso em: 31 dez. 2019.

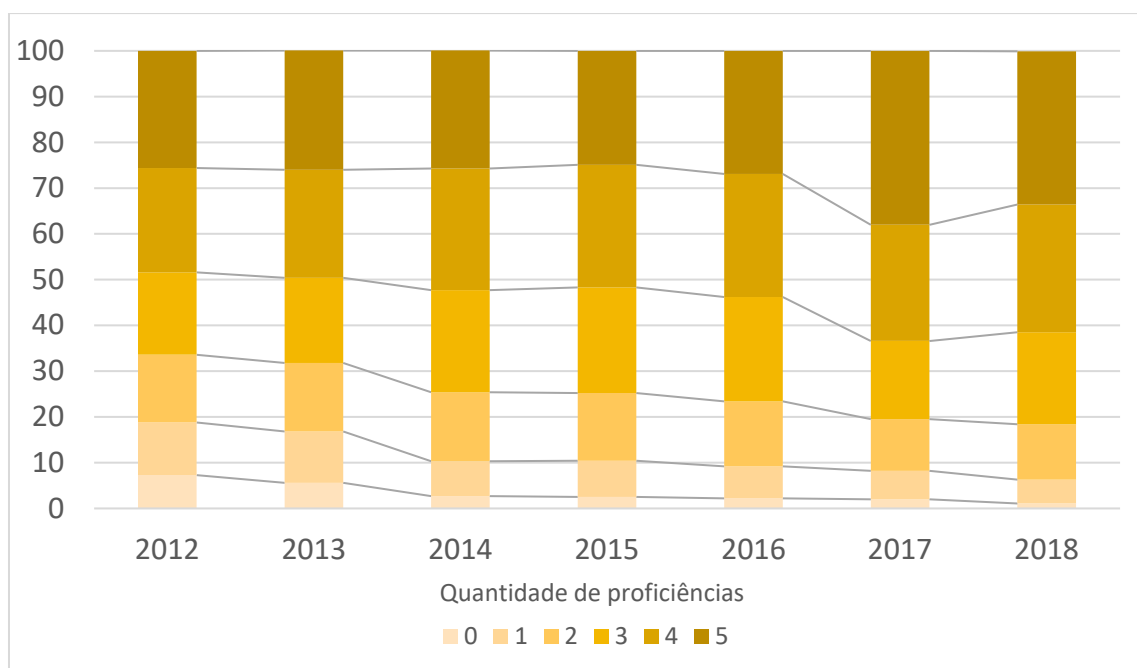
Figura 9 – Variação percentual das proficiências obtidas pelos alunos concluintes do ensino médio de até 18 anos da rede particular



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Ocorreu uma constante diminuição do percentual de alunos da rede pública não proficientes e dos que obtiveram uma proficiência. Porém, isso não resultou no aumento contínuo dos plenamente proficientes. Variaram positiva e negativamente os percentuais de alunos parcialmente proficientes que obtiveram nota mínima em três e quatro provas e dos plenamente suficientes (Figura 10).

Figura 10 – Variação percentual das proficiências obtidas pelos alunos concluintes do ensino médio de até 18 anos da rede pública

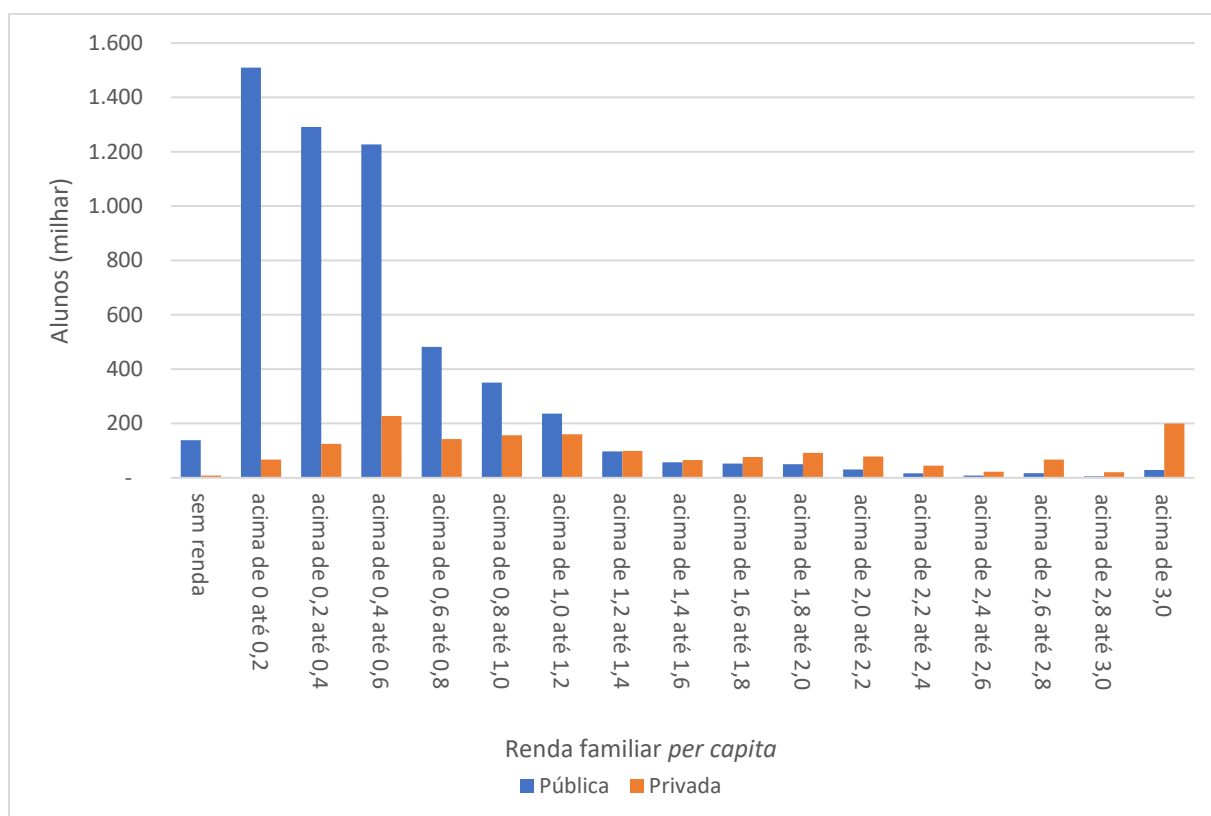


Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

5.3. Avaliação conjunta dos alunos da rede particular e pública

A distribuição do número de concluintes do ensino médio de até 18 anos por renda *per capita* no período de 2012 a 2018 não foi normal. A distribuição para os alunos da rede pública apresentou assimetria positiva; para os alunos da rede particular a distribuição foi assimétrica. Cerca de 75% dos alunos da escola pública apresentaram renda inferior a 0,6 salários mínimos. Além disso, muitos alunos da rede particular encontraram-se na faixa superior de renda, indicando que a análise conjunta dos concluintes de ambas as redes não seria adequada, para esta segmentação de renda (Figura 11).

Figura 11 – Distribuição do número de alunos concluintes do ensino médio pontuados no ENEM por faixa de renda familiar *per capita* e por rede de ensino, dado que a renda familiar não foi superior a 20 salários mínimos (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

As escolas da rede pública de ensino médio estavam presentes na maioria dos municípios e distritos brasileiros, enquanto as da rede particular estavam ausentes na maioria dos municípios e distritos (Tabela 22).

Tabela 22 – Número zonas com escolas que ofereceram o último ano do ensino médio para alunos concluintes do ensino médio de até 18 anos (2012-2018)

Zoneamento	Zonas	Zonas com escolas (%)	
		Rede particular	Rede pública
Regiões	5	100,0	100,0
UF	27	100,0	100,0
Mesorregiões	137	97,1	100,0
Microrregiões	558	90,7	100,0

continua

conclusão

Zoneamento	Zonas	Zonas com escolas (%)	
		Rede particular	Rede pública
Municípios	5570	29,8	99,8
Distritos	10283	17,8	65,3

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Além disso, os municípios que contiveram escolas particulares de ensino médio, cujos alunos pontuaram no ENEM de 2012 a 2018, estavam dispersos em algumas áreas e concentrados em outras, o que dificultaria a comparação de rendimento entre os resultados por rede de ensino (Figura 12). Assim, optamos por analisar e representar espacialmente somente o desempenho no ENEM dos estudantes da rede pública de ensino.

Figura 12 – Municípios com escolas de ensino médio da rede particular (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

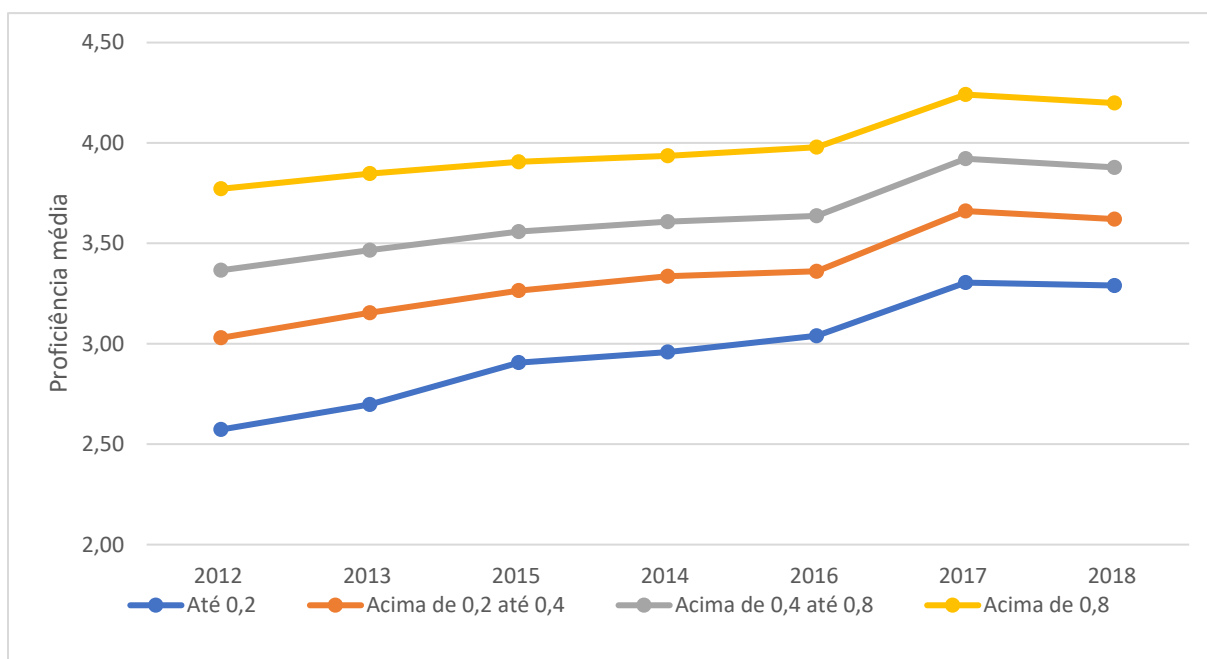
5.4. Variáveis associadas ao desempenho

Para os concluintes de até 18 anos da rede pública de 2012 a 2018, a correlação individual entre a pontuação no ENEM e o número de proficiências obtidas nas áreas do conhecimento foi alta: 0,836. Ao agregar essas variáveis de desempenho pelas médias obtidas por escola, o valor dessa correlação aumentou para 0,950, passando a ser muito forte. Na agregação por zonas, o valor da correlação tornou-se maior, variando de 0,976 (municípios) a 0,987 (regiões). Isso reforçou a certeza da opção pela utilização da proficiência para a representação do desempenho dos alunos concluintes no exame.

A variável socioeconômica que apresentou maior correlação ecológica com a proficiência foi a renda familiar *per capita*. O valor da correlação na agregação por zonas variou de moderadamente positiva (0,662 no zoneamento municípios) a altamente positiva (0,934 no zoneamento regiões). Consideramos que a variável explicativa foi a renda familiar *per capita* média, e a variável resposta, a proficiência média.

Apesar de terem ocorrido variações percentuais do número de concluintes do ensino médio de até 18 anos nas diferentes proficiências obtidas (Tabela 15), a proficiência média aumentou em todas as faixas de renda familiar *per capita* de 2012 a 2017. Em 2018, houve uma redução da proficiência média, em todas as faixas. A proficiência média anual mais elevada para os alunos concluintes de maior renda foi um indicativo da alta correlação entre essas duas variáveis (Figura 13).

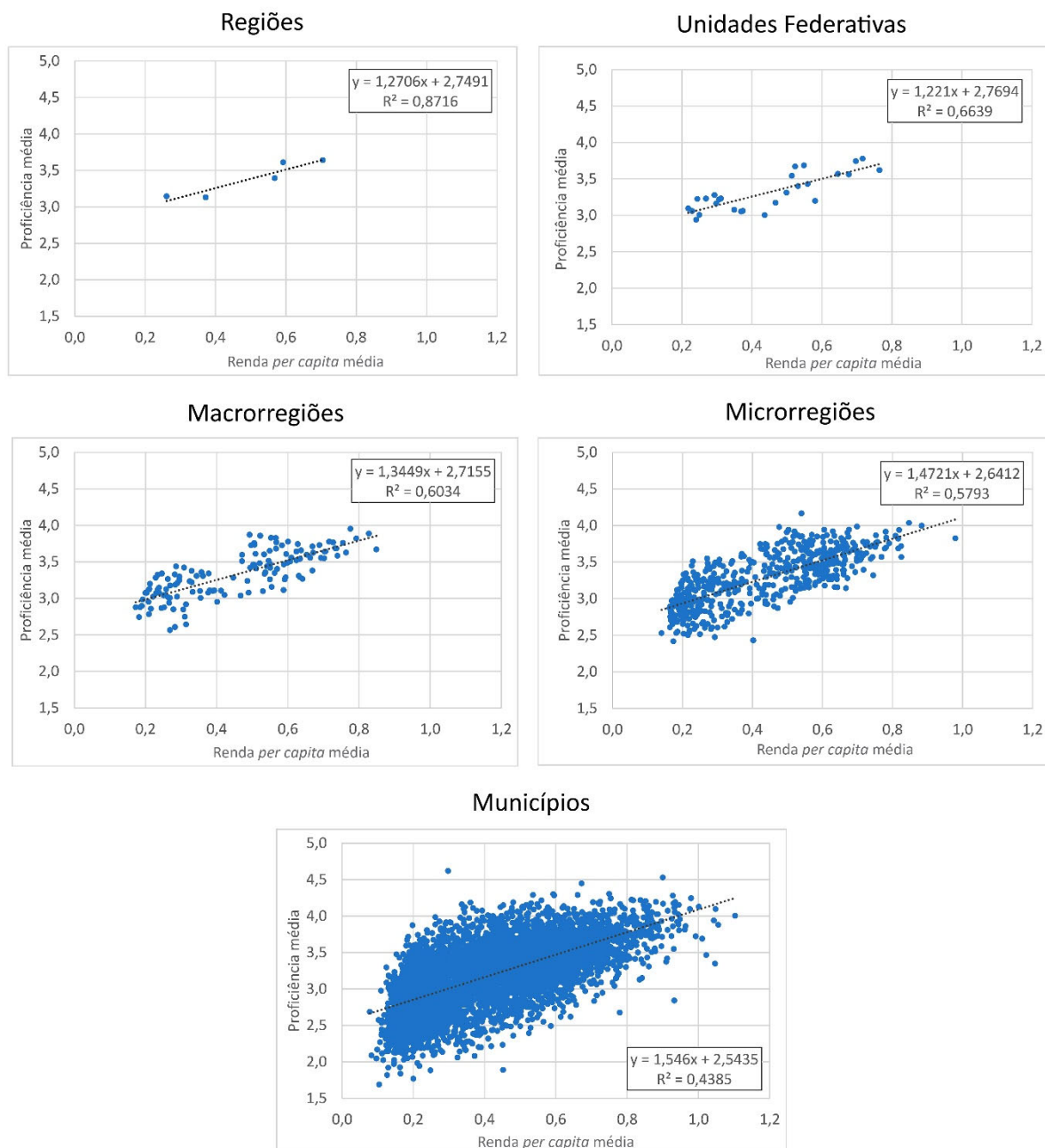
Figura 13 – Evolução das proficiências médias dos alunos concluintes do ensino médio da rede pública por faixa de renda familiar *per capita* (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Embora a amplitude da proficiência e da renda familiar *per capita* tenha diminuído na medida em que se aumentou o nível de agregação, analisamos os zoneamentos do menor para o maior número de zonas (Figura 14).

Figura 14 – Gráficos de dispersão para as variáveis proficiência no ENEM dos alunos concluintes do ensino médio de até 18 anos e renda *per capita* para o zoneamento por regiões (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Ao desprezar as casas decimais da proficiência média, não haveria diferenciação zonal desta variável no zoneamento regiões, pois os valores variaram de 3,1 a 3,6. Contudo, haveria variação zonal para a variável renda familiar *per capita*, nos intervalos estabelecidos na Tabela 8, contemplando duas faixas de renda.

No zoneamento por UF, a proficiência média também apresentou baixa amplitude, variando de 2,9 a 3,8. Em somente uma unidade da federação, Maranhão, a proficiência média esteve abaixo de três, resultando numa baixa diferenciação zonal, relacionada a esta variável. Neste zoneamento, as rendas familiares *per capita* média continuaram contempladas em duas faixas de renda.

No zoneamento Mesorregiões, a proficiência média variou de 2,6 (Norte do Amapá, AP) a 3,9 (Centro Ocidental Rio-Grandense, RS). A partir deste nível de agregação, todas as faixas de renda familiar *per capita* passaram a ser representadas. Ocorreu variação zonal das proficiências médias. Porém, estas estavam distribuídas em duas faixas: acima de dois até três e acima de três até quatro.

No zoneamento Microrregiões, a proficiência média variou de 2,4 (Rosário Oeste, MT) a 4,2 (Nova Friburgo, RJ). Neste nível, três faixas de proficiência média estavam contempladas.

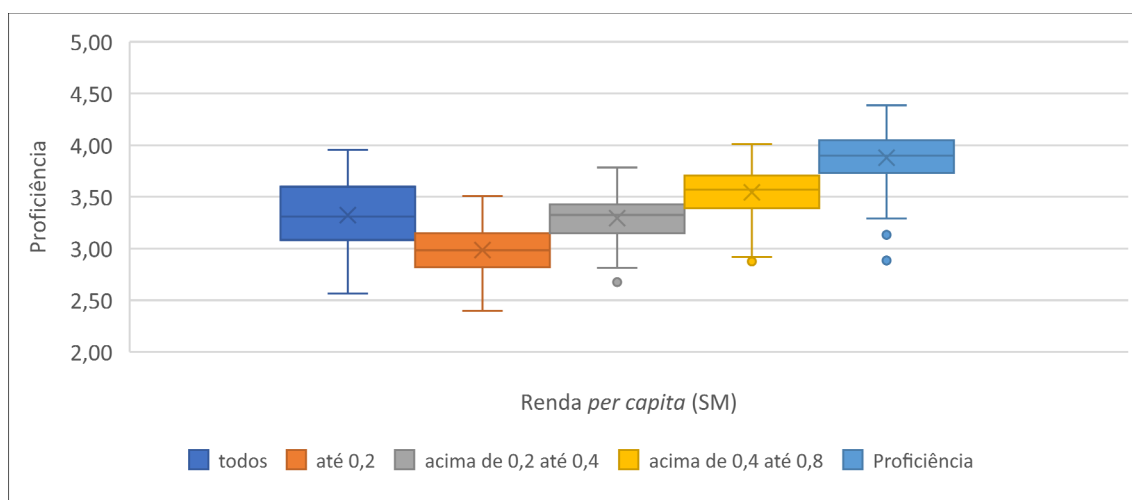
No zoneamento Municípios, a proficiência média variou de 1,7 (Jatobá, MA) a 4,6 (Cocal dos Alves, PI). Quatro faixas de proficiência média estabelecidos passaram a ser representadas.

5.5. Valores discrepantes

Encontramos valores discrepantes em três zoneamentos: Mesorregiões, Microrregiões e Municípios.

Nas mesorregiões, os resultados discrepantes que estão representados abaixo do limite inferior do bigode do diagrama de caixa (Figura 15) se referem aos concluintes de renda familiar *per capita* acima de 0,2 SM e às zonas localizadas nas regiões Norte, Nordeste e Centro-Oeste (Tabela 23).

Figura 15 – Diagramas de caixa das proficiências médias por mesorregião dos concluintes do ensino médio de até 18 anos.



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

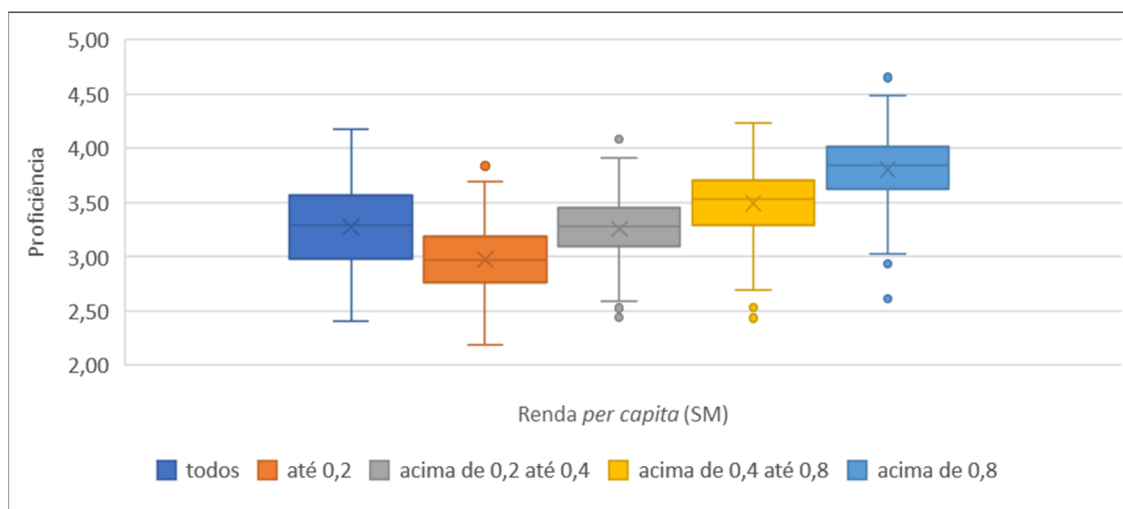
Tabela 23 – Mesorregiões com valores discrepantes, abaixo dos limites inferiores do bigode.

Renda familiar <i>per capita</i> (SM)	Mesorregião	UF	Proficiência
acima de 0,2 até 0,4	Sul de Roraima	RR	2,70
	Sudoeste Amazonense	AM	2,69
	Norte do Amapá	AP	2,67
acima de 0,4 até 0,8	Sul de Roraima	RR	2,87
acima de 0,8	Sul de Roraima	RR	3,13
	Norte do Amapá	AP	2,88

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Nas microrregiões, encontramos resultados discrepantes na segmentação por faixa de renda familiar *per capita*, com valores acima e abaixo dos limites do bigode (Figura 16).

Figura 16 – Diagramas de caixa das proficiências médias por mesorregião.



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Os resultados discrepantes acima do limite superior do bigode correspondem às microrregiões inseridas na Região Sudeste (Tabela 24).

Tabela 24 – Microrregiões com valores discrepantes, acima dos limites superiores do bigode.

Renda familiar <i>per capita</i> (SM)	Microrregião	UF	Proficiência
Até 0,2	Nova Friburgo	RJ	3,82
Acima de 0,2 até 0,4	Nova Friburgo	RJ	4,08
Acima de 0,8	Viçosa	MG	4,65

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

As microrregiões com resultados discrepantes, abaixo do limite inferior do bigode, estão localizadas nas regiões Norte, Nordeste e Centro-Oeste (Tabela 25).

Tabela 25 – Microrregiões com valores discrepantes, abaixo dos limites inferiores do bigode.

Renda familiar <i>per capita</i> (SM)	Microrregião	UF	Proficiência
Acima de 0,2 até 0,4	Juruá	AM	2,57
	Traipu	AL	2,56

continua

conclusão

Renda familiar <i>per capita</i> (SM)	Microrregião	UF	Proficiência
	Amapá	AP	2,53
	Rosário Oeste	MT	2,44
Acima de 0,4 até 0,8	Traipu	AL	2,57
	Rosário Oeste	MT	2,53
	Nordeste de Roraima	RR	2,49
Acima de 0,8	Jalapão	TO	2,93
	Amapá	AP	2,64
	Traipu	AL	2,58

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Há 4 municípios com valores discrepantes para as proficiências médias de todos os alunos e dois destes possuem uma única escola pública de ensino médio. Na segmentação por faixa de renda, encontramos 186 dados discrepantes e 121 destes referem-se aos municípios que possuem uma escola pública de ensino médio (Tabela 26).

Tabela 26 – Municípios com valores discrepantes da proficiência média (2012-2018).

Renda familiar <i>per capita</i> (SM)	Limites do bigode		Municípios com resultados discrepantes	
	Superior	Inferior	todos	Com uma escola
Todos	4,501	1,860	4	2
Até 0,2	4,069	1,792	24	15
Acima de 0,2 até 0,4	4,274	2,104	50	41
Acima de 0,4 até 0,8	4,510	2,316	51	33
Acima de 0,8	4,808	2,654	61	32

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O município que obteve a maior proficiência média dos concluintes do ensino médio de até 18 anos de 2012 a 2018 foi Cocal dos Alves (PI). O valor obtido pelos alunos foi 4,619 e corresponde ao desempenho da EE Augustinho Brandão. Num ranqueamento por escola, esta ocuparia a posição 379, dentre as unidades escolares com no mínimo dez concluintes que pontuaram no ENEM. Desta forma, a análise de valores discrepantes em municípios levaria a comparar municípios com escolas.

Somente uma microrregião conteve uma única escola: Fernando de Noronha. Porém, os resultados médios anuais de seus alunos nas diferentes faixas de renda não foram discrepantes.

6. DISCUSSÃO

Foi possível analisar dados de desempenho somente dos estudantes concluintes de até 18 anos no ENEM que, obviamente, tiveram oportunidades para completar o ensino médio em idade escolar ideal. Estes dados possibilitaram avaliar zoneamentos, identificar as áreas de melhor desempenho no exame e visualizar a dependência espacial das variáveis renda familiar *per capita* e da proficiência média no ENEM. Além disto, agrupar dados individuais em diferentes escalas proporcionou interpretar o MAUP e identificar falácias nos zoneamentos abordados nesta pesquisa.

6.1. Desempenho no ENEM

Na RMSP, segundo Torres et al. (2003, p. 110), o padrão radial-concêntrico (centro-periferia) da “segregação parece ser muito mais forte na educação que no item renda”, pois “é possível encontrar lugares com níveis mais altos de renda e distantes do centro – embora não com níveis muito altos de escolaridade”. Neste caso, o posicionamento geográfico apresentaria uma influência maior no rendimento escolar que a renda.

A segregação espacial também foi importante para Cunha et al. (2009) na análise das escolas e da pobreza no município de Campinas para compreender o desempenho de alunos. A localização geográfica foi significativa no desempenho no SARESP, pois os alunos das escolas localizadas nas áreas mais centrais e mais ricas tiveram desempenho superior aos alunos de escolas de áreas mais pobres.

Nesta pesquisa ao analisar a questão em outra escala, considerando o Brasil como um todo, o desempenho dos concluintes do ensino médio no ENEM pareceu não seguir esta lógica. Os melhores desempenhos médios no ENEM não se encontraram, em sua maioria, nas regiões metropolitanas de maior PIB.

Isto não significa contradição com estes estudos porque o resultado médio de uma cidade como São Paulo elimina a variância interna do município. Neste, a proficiência média dos concluintes de até 18 anos de 2012 a 2018 foi 3,6. Na

agregação de dados em subprefeituras, a proficiência média variou de 3,2 a 4,1, sendo os valores mais baixos em áreas periféricas e os valores mais altos nas áreas centrais do município de São Paulo, corroborando Cunha et al. (2009)

De maneira geral, os concluintes das escolas públicas de até 18 anos de renda familiar *per capita* acima de 0,8 SM apresentaram melhores resultados médios e os menores aumentos da proficiência nas zonas inseridas nas regiões Sul, Sudeste e Nordeste do Brasil, enquanto os que renda inferior a 0,2 SM apresentaram os piores rendimentos e os melhores aumentos da proficiência na maior parte do país, principalmente nas regiões Norte, Nordeste e Centro Oeste.

Apesar da gratuidade de inscrição para os concluintes da rede pública, apenas 57% dos tinham até 18 anos se inscreveram e pontuaram no ENEM de 2012 a 2018. Porém, a participação no exame em unidades agregadas não é homogênea. Por exemplo, em 291 municípios mais de 80% dos concluintes pontuaram no exame enquanto em 20 municípios, menos de 20% dos concluintes pontuaram. Mesmo quando se compara capitais, a variação na participação no ENEM é grande, de 31,1% (Maceió, AL) a 66,3% (Belém, PA). Desta forma, torna-se comprometida a utilização dos dados de desempenho agregados para mensurar a qualidade de ensino, principalmente na comparação de resultados entre unidades agregadas, escolas ou zonas. Isto corrobora com a afirmação de Travitzki et al. (2014) sobre a classificação das escolas no ENEM abarcar uma amostragem distorcida da totalidade do ensino médio no Brasil, provavelmente relacionado com o caráter voluntário do exame.

O ENEM, baseado na TRI, possibilitaria a comparação de resultados entre anos diferentes (DA SILVA; LAMB; BARBOSA, 2016; MINISTÉRIO DA EDUCAÇÃO, 2011). Em função da variabilidade anual dos resultados de cada prova (Figura 5 e Figura 6), esta comparação de resultados foi mais eficiente ao considerar o desempenho de todas as provas (Figura 7).

6.2. Avaliação de zoneamentos

Os procedimentos que utilizamos para avaliar os zoneamentos foi proposto por Openshaw (1984) como uma solução geográfica para o MAUP, na avaliação dos

zoneamentos que seriam gerados automaticamente por processos computacionais. A adaptação destes procedimentos possibilitou estabelecer parâmetros para avaliação dos zoneamentos da BET, para a representação espacial do desempenho de alunos no ENEM.

Para testar, seguindo os procedimentos da Figura 2A, produzimos cinco conjuntos de mapas, um para cada zoneamento. Em cada conjunto, o mapa maior representa a distribuição da proficiência média de todos os alunos, enquanto os menores representam a distribuição da proficiência média de alunos por faixas de renda familiar *per capita*.

Dividimos as proficiências médias em cinco faixas: até um, acima de um até dois, acima de dois até três, acima de três até quatro e acima de quatro até cinco.

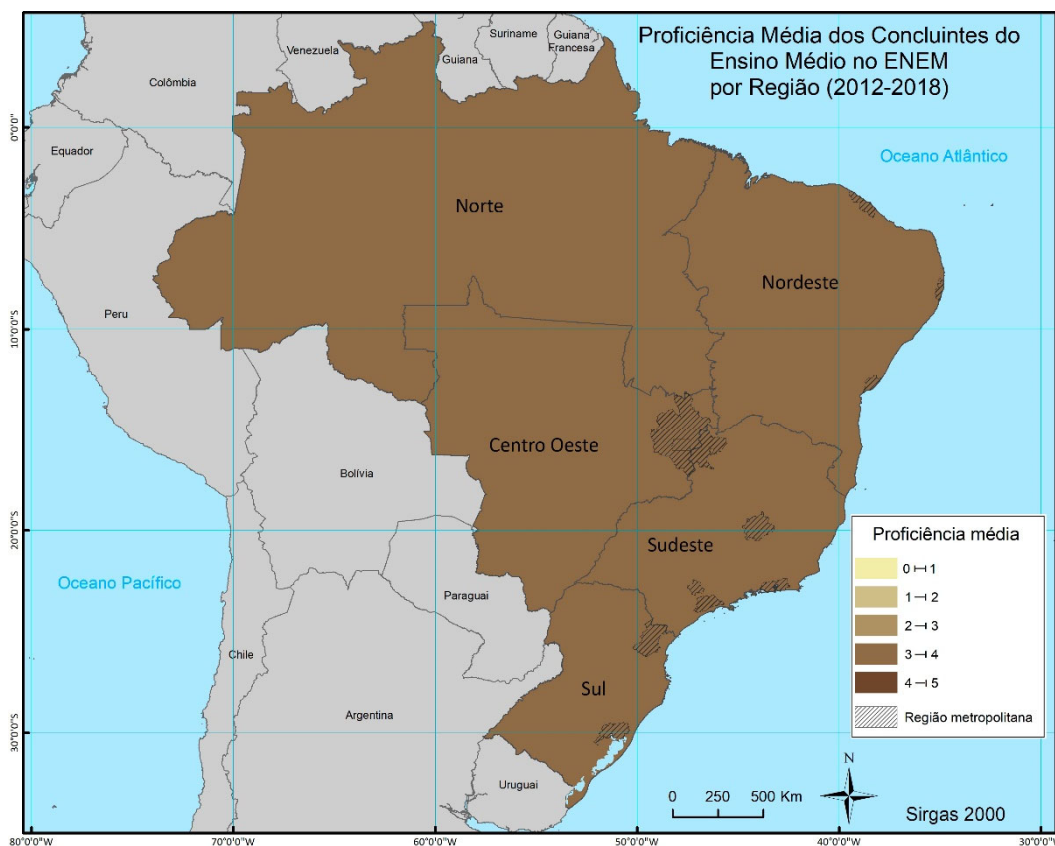
Neste estudo, a função objetivo foi a diferenciação zonal da proficiência média dos alunos concluintes do ensino médio, em diferentes faixas de renda familiar *per capita*. Consideramos como eficientes os zoneamentos que satisfizeram esta função.

O organograma de avaliação (Figura 2A) previu a ordem de avaliação dos zoneamentos do menor número de zonas para o maior e indicou um momento de interrupção das avaliações. A constatação de que não houve escolas de ensino médio em todos os municípios com certo número mínimo de concluintes (Tabela 22) indicou que não foi necessário avaliar zoneamentos com menor nível de agregação como os distritos.

Na análise dos mapas, procedemos a Etapa 2 dos procedimentos propostos por Openshaw (1984) (Quadro 2). Esperamos pela variação zonal das proficiências médias em todas as faixas de renda familiar *per capita*, obtendo-se as melhores médias nas zonas onde se localizam as áreas metropolitanas de maior PIB.

O zoneamento regiões não apresentou variação espacial das proficiências médias para todos os alunos. Por esse motivo reprovamos esse zoneamento e julgamos não ser necessário representar a distribuição espacial das proficiências por faixa de renda (Figura 17).

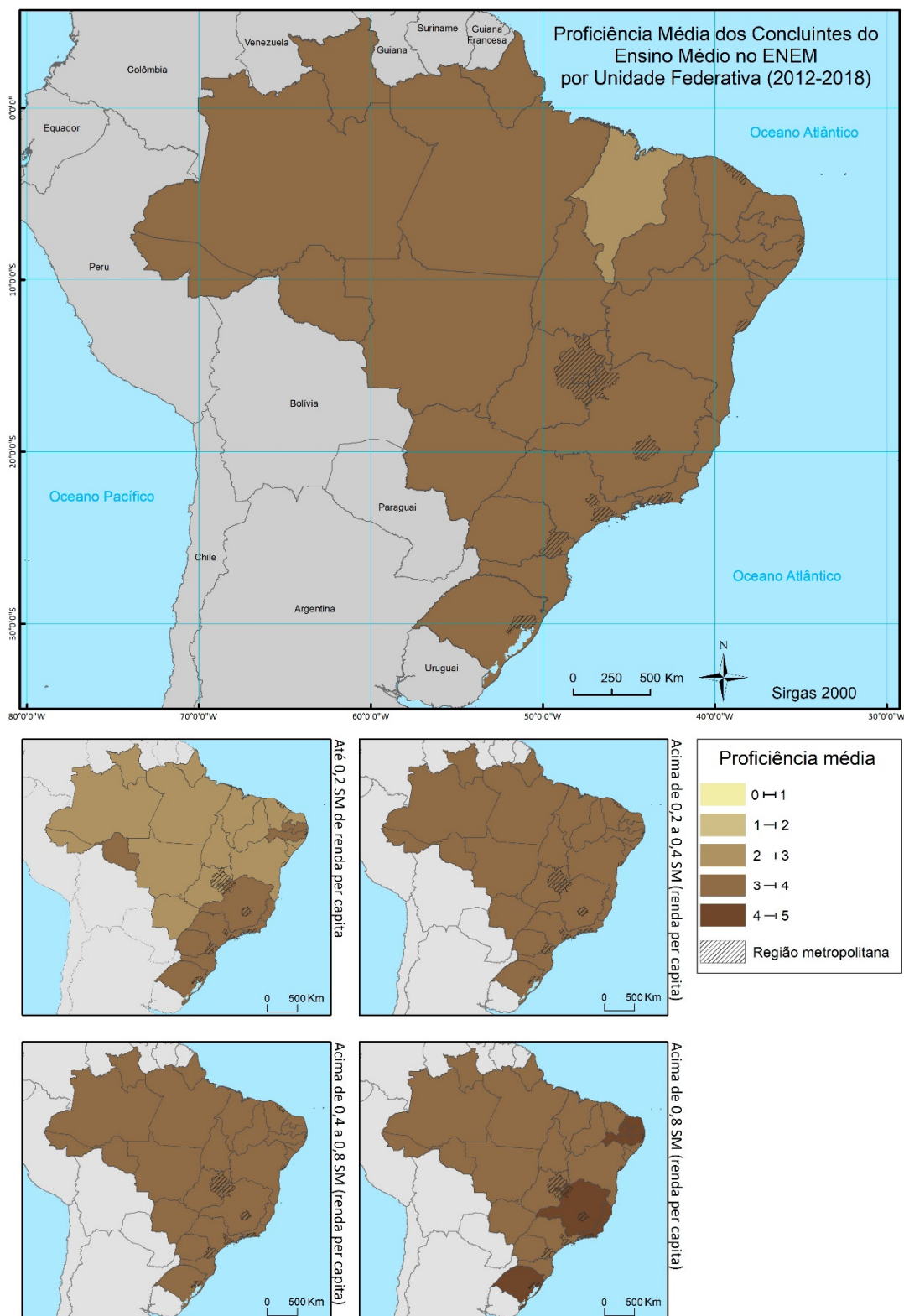
Figura 17 – Proficiências médias no ENEM por Região (2012-2018).



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

No zoneamento unidades federativas, ocorreu variação espacial das proficiências médias para todos os alunos. Porém, a variação foi resultado do menor desempenho dos alunos concluintes do Maranhão. Reprovamos esse zoneamento ao não observar variações espaciais para as faixas intermediárias de renda familiar *per capita* (Figura 18).

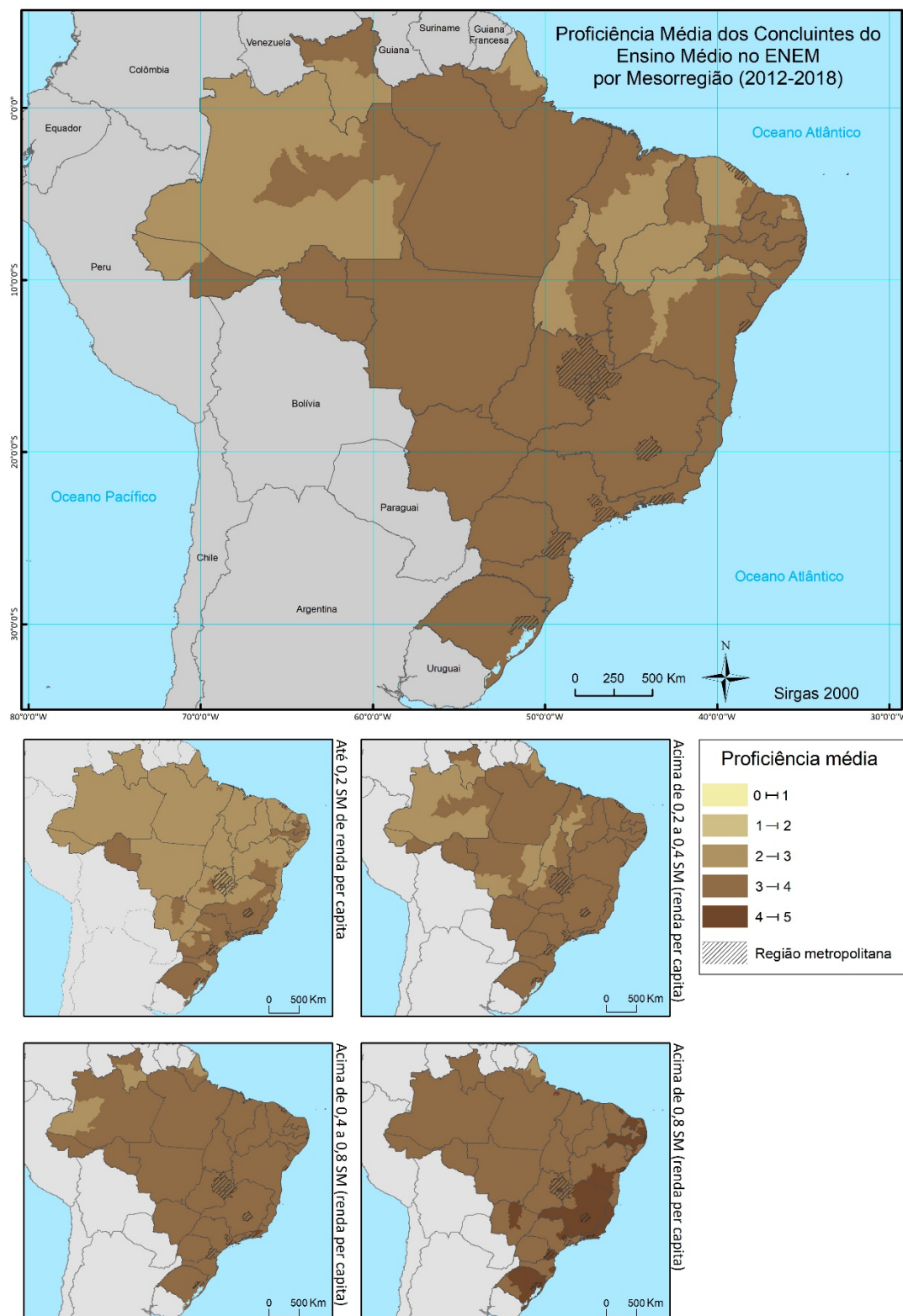
Figura 18 – Proficiências médias no ENEM por UF (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Nas mesorregiões, ocorreu a variação espacial das proficiências médias para todos os alunos e para cada faixa de renda familiar *per capita* (Figura 19).

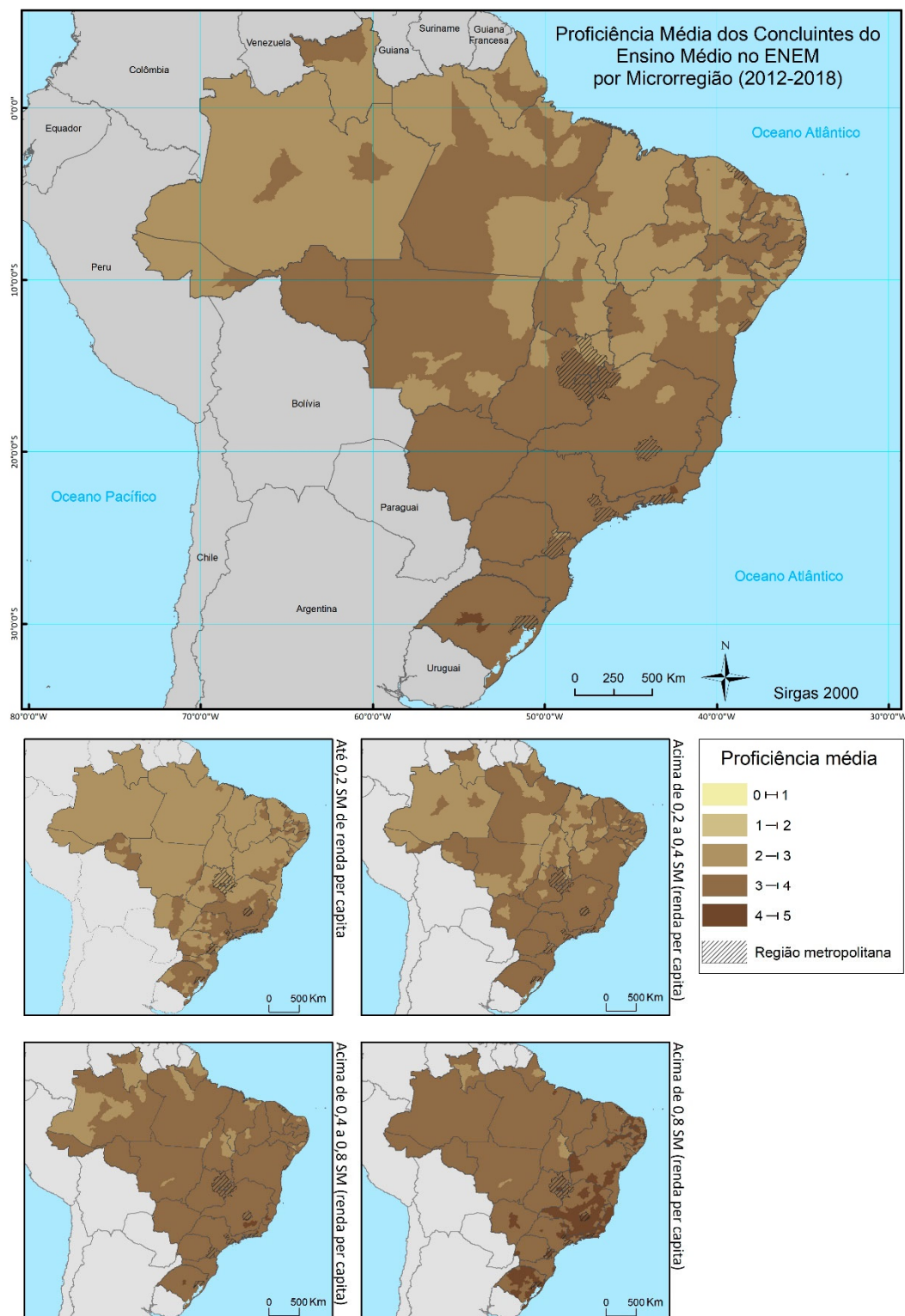
Figura 19 – Proficiências médias no ENEM por mesorregião (2012-2018).



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Nas microrregiões, ocorreu a variação espacial das proficiências médias para todos os alunos e para cada faixa de renda familiar *per capita* (Figura 20).

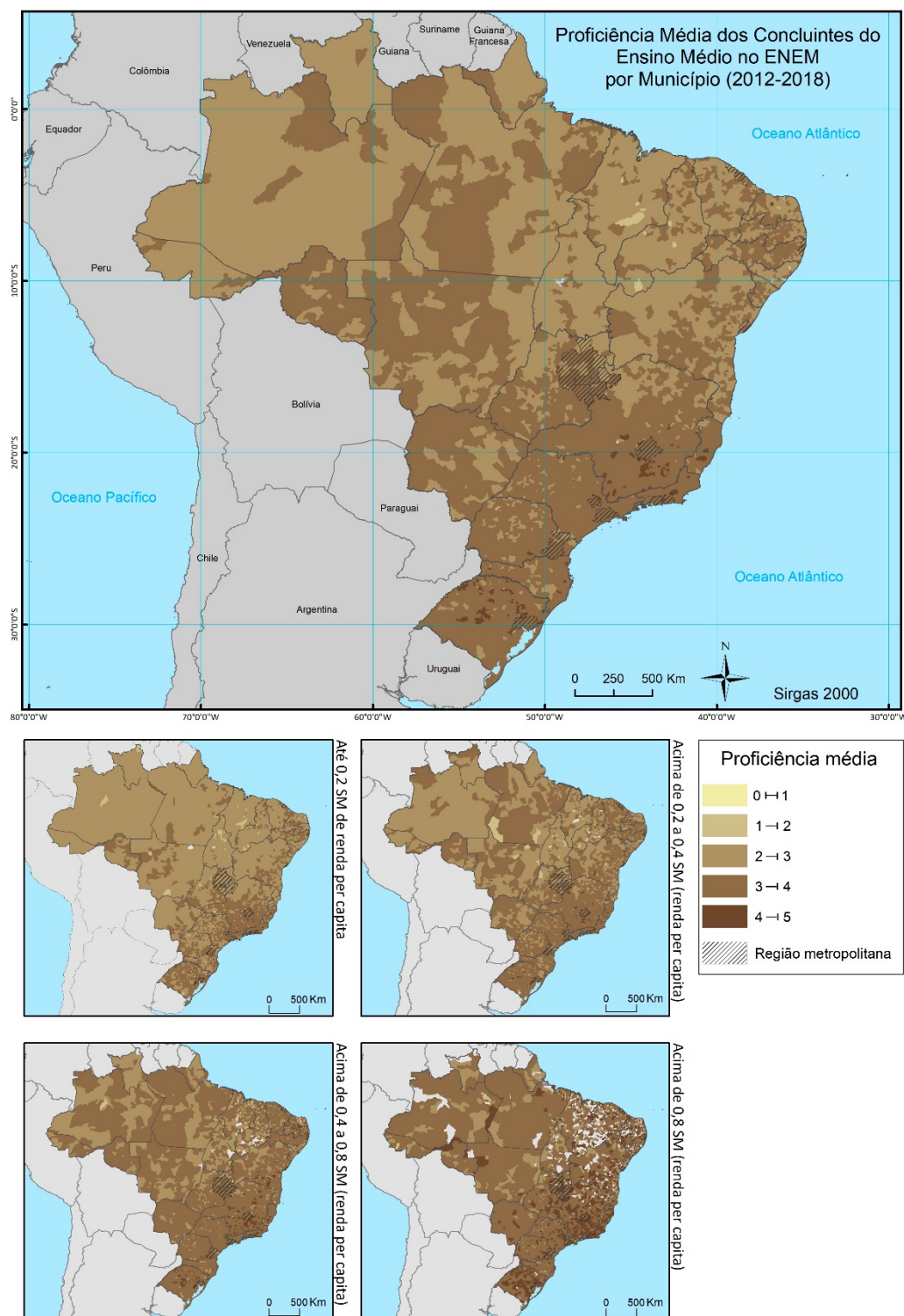
Figura 20 – Proficiências médias no ENEM por microrregião (2012-2018).



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Dentre os zoneamentos avaliados, o zoneamento municípios apresentou maior heterogeneidade espacial das proficiências médias para todos os alunos e para cada faixa de renda familiar *per capita* (Figura 21).

Figura 21 – Proficiências médias no ENEM por município (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A etapa 3 dos procedimentos propostos por Openshaw (1984) é a interpretação dos resultados no sentido estatístico e a verificação de se os resultados esperados foram alcançados com grau de erro tolerável.

De 2012 a 2018, o município com menor número de alunos concluintes do ensino médio de até 18 anos foi Uiramutã, RR, com 21 alunos. Na segmentação por faixa de renda, 1519 municípios tiveram menos de dez alunos em uma ou mais faixas de renda familiar *per capita*. Dessa forma, o zoneamento municípios foi aprovado para representar a proficiência média de todos os alunos, mas não foi aprovado para a representação por faixa de renda familiar *per capita*.

Nesse período, somente três microrregiões tiveram menos de dez alunos: Fernando de Noronha, PE, com oito alunos na faixa acima de 0,2 a 0,4 SM; Japurá, AM, com cinco alunos, e Umbuzeiro, PB, com nove alunos na faixa acima de 0,8 SM. Apesar disso, aprovamos o zoneamento microrregiões para a representação da proficiência média de todos os alunos e por faixa de renda familiar *per capita* (Quadro 5).

O zoneamento mesorregiões foi aprovado porque teve em todas as zonas um número suficiente de alunos para a representação com e sem segmentação por faixa de renda.

Quadro 5 – Testes da etapa 2 dos procedimentos propostos por Openshaw (1984)

Zoneamento	Variação zonal da proficiência	Variação zonal da proficiência de alunos nas diferentes classes de renda	Dez ou mais alunos por zona e faixa de renda
Regiões	Não	-	-
UF	Sim	Não	-
Mesorregiões	Sim	Sim	Sim
Microrregiões	Sim	Sim	Não
Municípios	Sim	Sim	Não

Fonte: elaboração do autor.

6.2.1. Evolução temporal das proficiências

Para calcular a evolução temporal das proficiências em cada zona, subtraímos o valor médio do primeiro ano (2012) do valor médio do último ano (2018). Dessa forma, o valor representado no mapa é a variação da proficiência.

Houve casos de municípios que começaram a ter alunos concluintes que pontuaram no ENEM após 2012, em decorrência de inauguração de escola ou do início do oferecimento do último ano letivo do ensino médio. Nesses casos, consideramos o valor médio do primeiro ano em que observamos os resultados dos concluintes no ENEM.

Para os procedimentos previstos na Figura 2B, representamos as zonas nas quais houve um mínimo de dez concluintes por ano que pontuaram no ENEM. Também representamos as zonas que apresentaram de cinco a nove concluintes e coeficiente de variação da proficiência média anual inferior a 0,2 (Tabela 27), eliminando da análise as áreas com poucos alunos que apresentaram resultados anômalos em um ou mais anos.

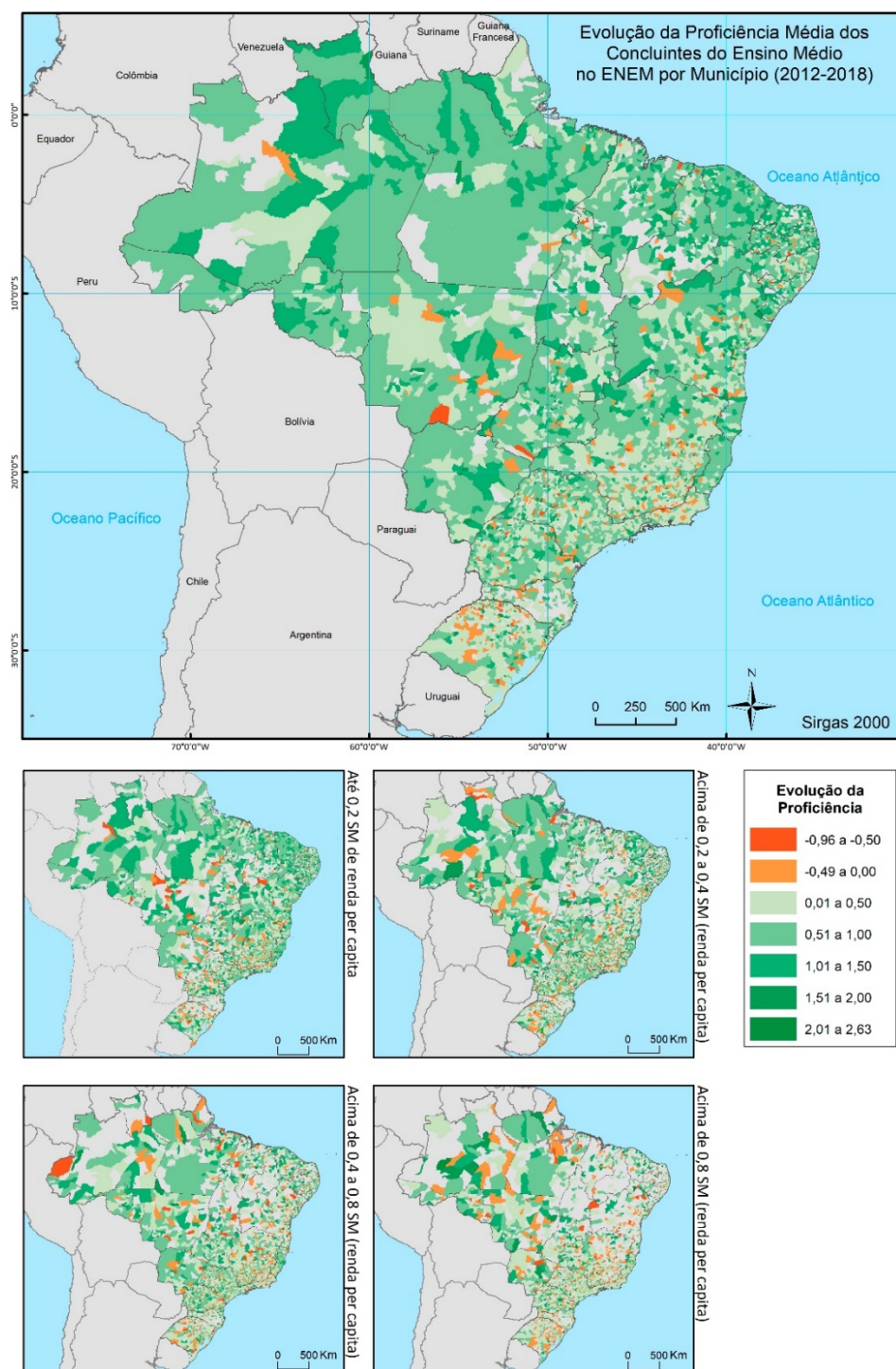
Tabela 27 – Zonas aprovadas para análise longitudinal da proficiência média de 2012 a 2018

Zoneamento	Zonas com escolas	Percentual de zonas aprovadas por faixa de renda familiar <i>per capita</i> .				
		Todos	até 0,2	Acima de 0,2 até 0,4	Acima de 0,4 até 0,8	Acima de 0,8
Regiões	5	100,0	100,0	100,0	100,0	100,0
UF	27	100,0	100,0	100,0	100,0	100,0
Mesorregiões	137	100,0	100,0	100,0	100,0	100,0
Microrregiões	558	99,6	99,1	98,9	98,4	88,0
Municípios	5.561	89,7	56,3	44,0	46,7	26,6

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O zoneamento municípios não foi aprovado para representar a evolução temporal da proficiência por ter muitas zonas com poucos ou nenhum aluno em 2012 e 2018. No mapa, essas zonas não foram representadas em cores, indicando a não ocorrência de alunos em número suficiente para análise temporal (Figura 22).

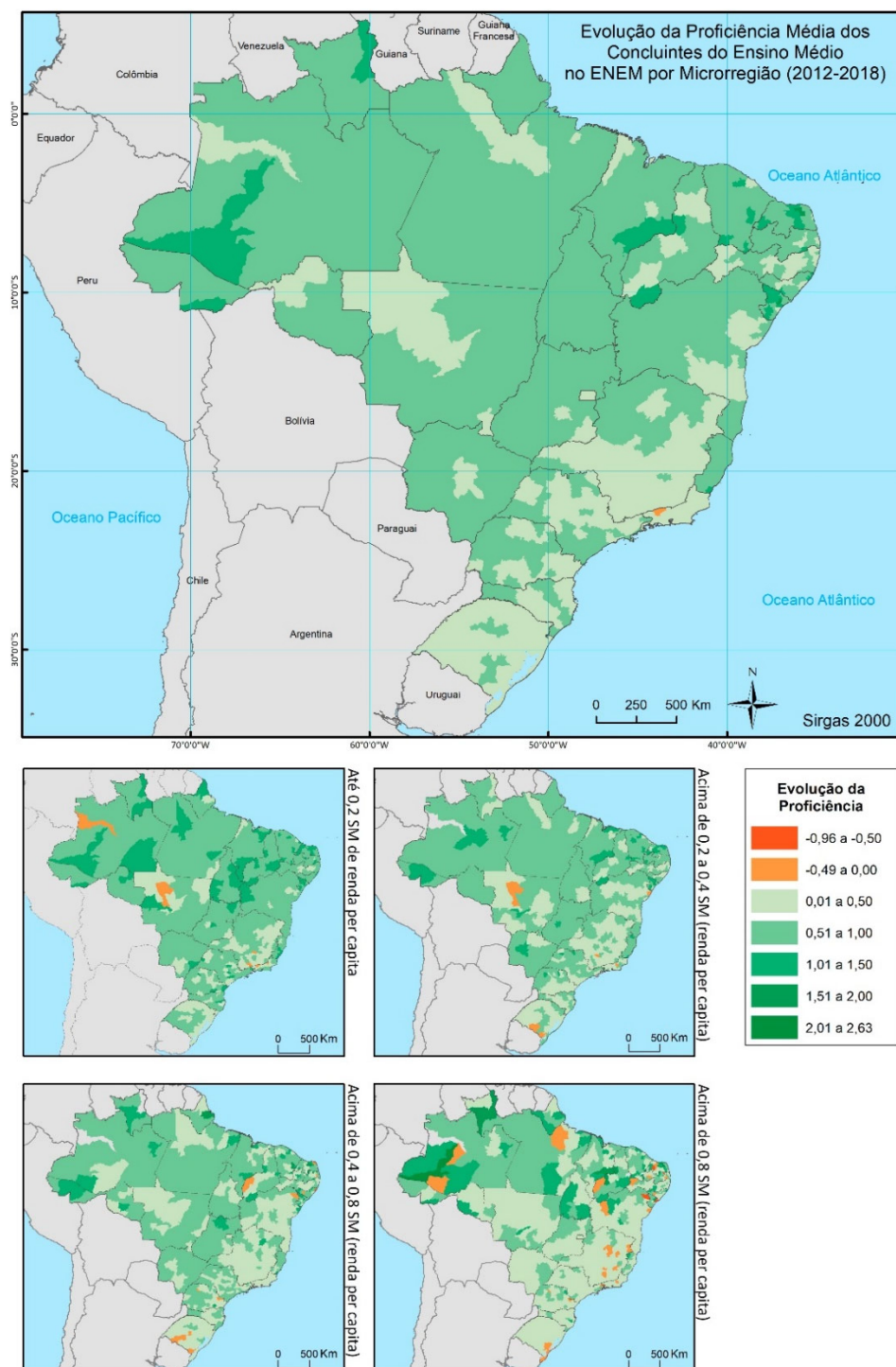
Figura 22 – Evolução temporal da proficiência média por município (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O zoneamento microrregiões mostrou-se eficiente para a representação da evolução temporal da proficiência média dos concluintes do ensino médio por terem, na maioria das zonas, alunos suficientes por faixa de renda e por ano (Figura 23).

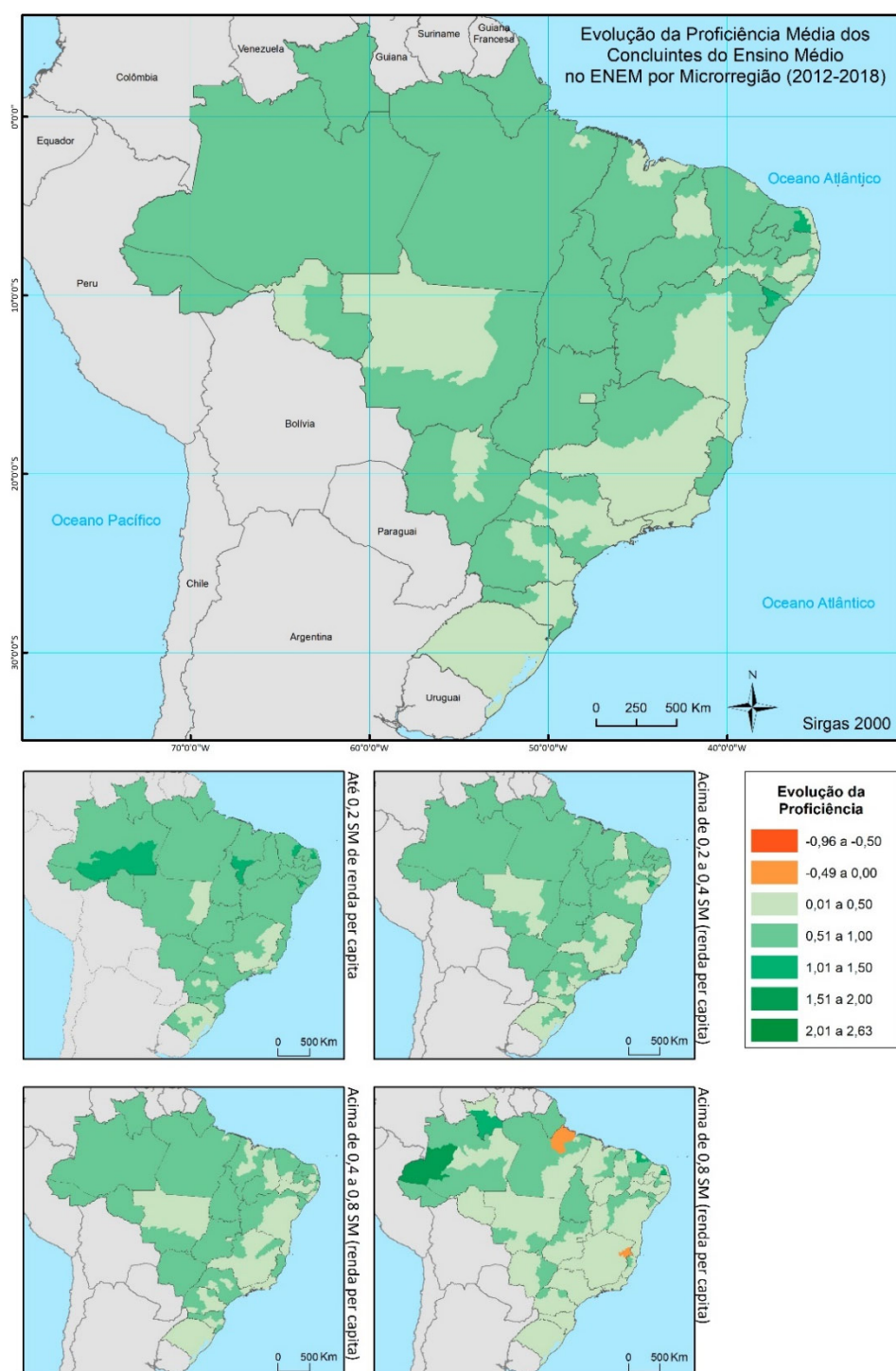
Figura 23 – Evolução temporal da proficiência média por microrregião (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O zoneamento mesorregiões mostrou-se eficiente para a representação da evolução temporal da proficiência média dos alunos concluintes do ensino médio por possuírem, em todas as zonas, alunos suficientes por faixa de renda e por ano (Figura 24).

Figura 24 – Evolução temporal da proficiência média por mesorregião (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

De maneira geral, observamos nos estados do Norte, Nordeste e Centro-Oeste maiores aumentos da proficiência média, superior a 0,5, principalmente para os alunos de renda familiar *per capita* de até 0,2 SM e um menor aumento da proficiência no Sul e Sudeste. Ocorreu diminuição da proficiência média dos alunos de 2012 a 2018 em poucas zonas (Figura 23). As correlações ecológicas entre renda familiar *per capita* média e evolução média da proficiência para os zoneamentos mesorregiões e microrregiões foram -0,678585 e -0,601166, respectivamente.

6.2.2. Áreas de melhor desempenho

O resultado esperado era o melhor rendimento dos concluintes do ensino médio nas zonas que contiveram ou que foram contidas pelas dez regiões metropolitanas de maior PIB (Tabela 28).

Tabela 28 – Principais regiões metropolitanas brasileiras (2017)

UF	Região metropolitana	PIB 2017 (R\$ 1.000,00)	Municípios
SP	RM São Paulo	1.140.570.430	39
RJ	RM Rio de Janeiro	503.827.697	21
DF	Região Integrada de Desenvolvimento do Distrito Federal e Entorno ¹⁵	266.730.019	22
MG	RM Belo Horizonte	203.908.869	50
SP	RM Campinas	194.669.715	20
RS	RM Porto Alegre	180.402.400	34
PR	RM Curitiba	156.042.029	29
BA	RM Salvador	118.217.726	13
PE	RM Recife	105.513.589	14
CE	RM Fortaleza	93.130.076	19

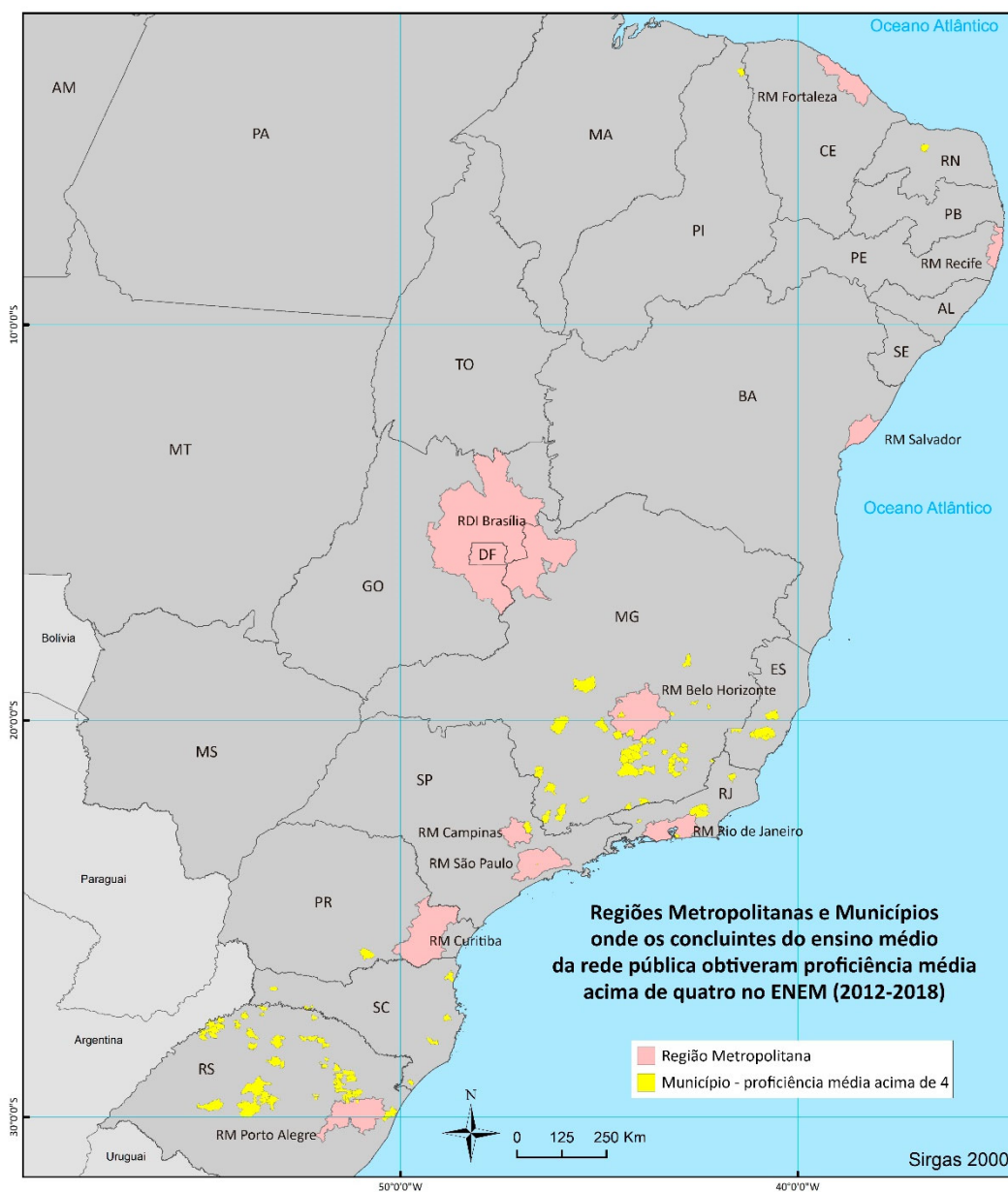
Fonte: IBGE.¹⁶

Para avaliar esse resultado, dispusemos num mapa as regiões metropolitanas e os 120 municípios cujos alunos, em média, obtiveram proficiência acima de quatro (Figura 25). Apenas oito desses municípios localizam-se nas principais metrópoles: Bonfim, Florestal e Itaguara (RM Belo Horizonte); Dois Irmãos, São Sebastião do Caí e Feliz (RM Porto Alegre); São Caetano do Sul (RM São Paulo); Niterói (RM Rio de Janeiro). Observamos que a maioria dos municípios de melhor desempenho no ENEM não se localiza nas áreas de maior PIB.

15 Ampliada pela Lei Complementar n° 163, de 14 de junho de 2018. Disponível em http://www.planalto.gov.br/ccivil_03/leis/lcp/Lcp163.htm. Acesso em: 25 mai. 2020.

16 Fonte: ftp://ftp.ibge.gov.br/Pib_Municipios/2017/base/base_de_dados_2010_2017_xls.zip. Acesso em: 5 jan. 2020.

Figura 25 – Regiões metropolitanas e municípios onde os alunos concluintes do EM da rede pública de até 18 anos obtiveram proficiência média acima de quatro no ENEM (2012-2018)



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a); IBGE¹⁷.

17 Fonte: ftp://ftp.ibge.gov.br/Pib_Municipios/2017/base/base_de_dados_2010_2017_xls.zip. Acesso em: 5 jan. 2020

6.3. Autocorrelação espacial

Os mapas produzidos para os testes (p. 65-69) tiveram por objetivo auxiliar a avaliação da eficiência dos zoneamentos. Embora todos os mapas apresentem uma ordem visual que permitiu a comparação da proficiência média dos alunos entre zonas, não é facilmente perceptível a influência da renda no desempenho dos alunos, pois é necessário a observação conjunta dos quatro mapas menores.

Anselin, Syabri e Kho (2006) apresentaram as possibilidades de trabalhar e visualizar dados estatísticos como a autocorrelação espacial e o LISA, por meio do *software* Geoda.

Utilizamos este *software* para produzir mapas de *cluster* LISA com significância de 0,05 dos zoneamentos avaliados. Os valores considerados altos e baixos não foram os mesmos para cada zoneamento (Tabela 29). Além disso, a diferença entre limites baixo e alto para a variável renda foi muito pequena, principalmente para o zoneamento municípios, onde essa diferença foi inferior a 0,01 SM.

Tabela 29 – Valores limites das variáveis dos mapas LISA

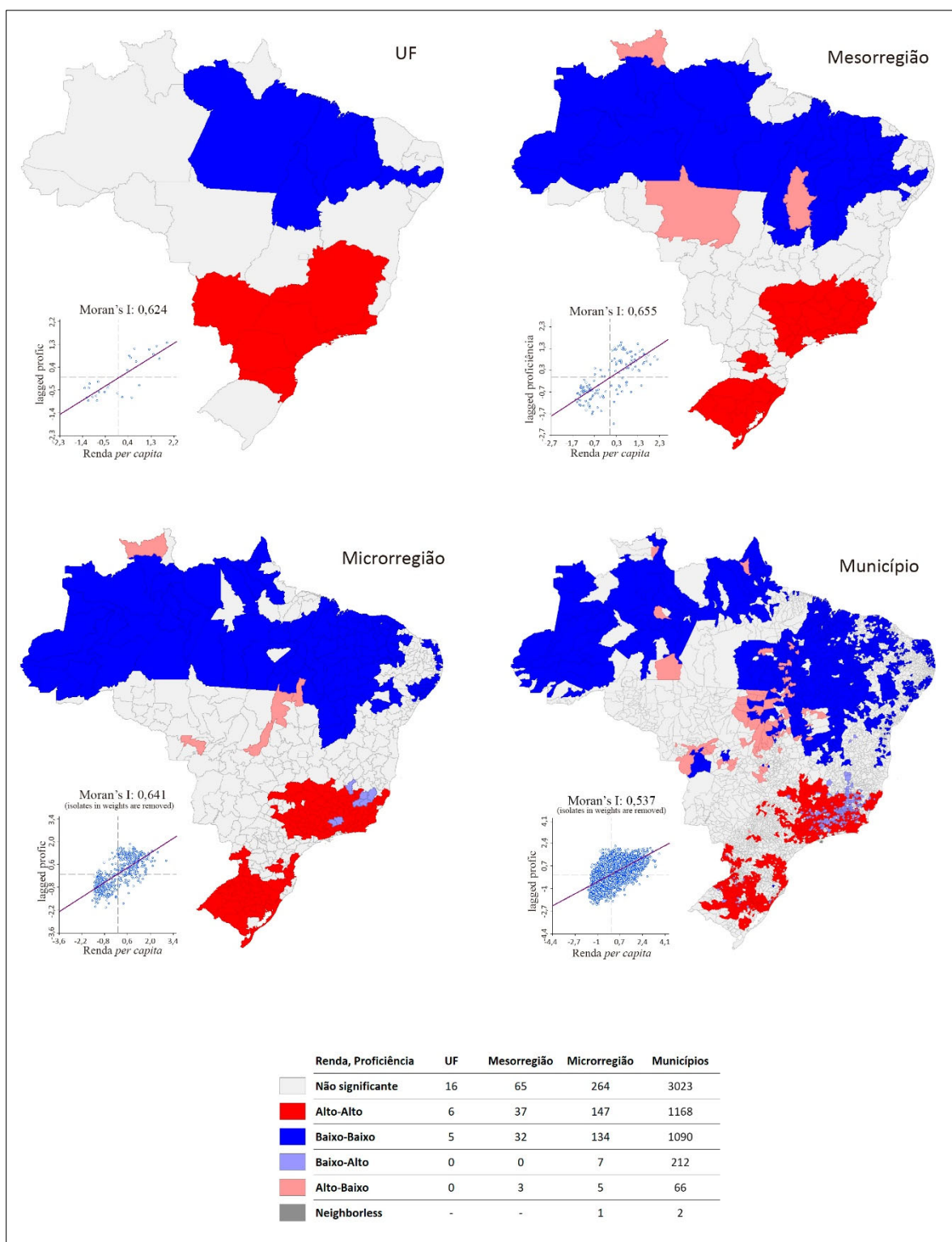
Zoneamento	Menor valor alto		Maior valor baixo	
	Renda	Proficiência	Renda	Proficiência
UF	0,52	3,43	0,44	3,00
Mesorregiões	0,47	3,74	0,42	3,27
Microrregiões	0,43	3,45	0,41	2,76
Municípios	0,41	3,56	0,41	3,36

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Os mapas LISA dos três zoneamentos aprovados (Figura 26) possibilitou visualizar *clusters* de

- maior renda e proficiência médias, no Sudeste e Sul do Brasil;
- menor renda e proficiências médias, no Norte e Nordeste do Brasil;
- *outliers*.

Figura 26 – Mapas de *clusters* e diagramas de espalhamento LISA das variáveis renda *per capita* média e proficiência média dos cinco zoneamentos avaliados



Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Convém ressaltar que, nesse caso, os *outliers* referem-se a resultados em que a ocorrência de valores alto-baixo e baixo-alto foram significativos. Não são os mesmos *outliers* da proficiência média, baseados em quantis que referenciamos como valores discrepantes.

O mapa LISA do zoneamento UF que fora reprovado evidenciou um problema deste. A distribuição dos *clusters* não foi coerente com a distribuição dos *clusters* nos zoneamentos aprovados. Essa constatação reforçou o argumento de Openshaw (1984) de que não podemos simplesmente aceitar qualquer resultado que um zoneamento produz.

6.4. Interpretando o MAUP

A variação dos valores de correlação e da média nos zoneamentos avaliados estaria relacionada com o problema de escala, descrito por Openshaw (1984), pois, dos municípios para as regiões, os conjuntos de unidades de área são progressivamente agregados em menos unidades, porém maiores. Na agregação de áreas, do município para a região, ocorreu a diminuição da diferenciação zonal da proficiência média corroborando a proposição de Jelinski e Wu (1996).

As microrregiões são compostas por municípios que compartilham certas especificidades relacionadas à estrutura de produção que podem ser resultado do ambiente natural e das relações sociais e econômicas (IBGE, 1994). Dessa forma, os municípios próximos que foram agregados em uma zona estariam mais relacionados que os municípios mais distantes e que foram agregados em outra zona. As mesorregiões, por sua vez, são definidas pelo processo social, pelo quadro natural e pela rede de comunicação e de lugares (IBGE, 1994). As microrregiões de uma mesorregião seriam mais relacionadas do que as microrregiões de outra mesorregião. Nesse caso, a característica de agregação da BET é compatível com a primeira lei da geografia, proposta por Tobler (1970).

Em cada uma dessas zonas e nas outras avaliadas nesta pesquisa, calculamos o coeficiente de correlação linear de Pearson entre as variáveis proficiência e renda familiar *per capita*. Esse coeficiente de correlação linear de Pearson expressa a dependência entre variáveis contínuas e baseia-se na regressão linear entre série de dados, dispostos em pares (FERREIRA, 2014). Podemos calculá-lo considerando todos os concluintes (correlação individual). Nesse contexto, o valor do coeficiente independe da localização de cada indivíduo. Ao agrupar os concluintes em zonas, o posicionamento geográfico assume importância. Mudar o indivíduo de uma zona para outra alteraria o valor do coeficiente de correlação do zoneamento (correlação ecológica). Segundo Robinson (1950), o cálculo dessa correlação em dados agrupados em zonas considera os valores das variáveis expressos em porcentagens, medianas ou médias. Nesta pesquisa, utilizamos os valores médios da proficiência e da renda familiar *per capita* para calcular a correlação ecológica de cada zoneamento.

A agregação diminuiu a amplitude das médias da renda familiar *per capita* e da proficiência (Figura 14) enquanto resultou no aumento dos valores das correlações.

Para entender esse aumento, realizamos alguns ensaios a partir da seguinte afirmação: o aumento do valor da correlação não ocorre necessariamente somente com a agregação de dados. Para provar essa afirmação, realizamos dez testes com os seguintes procedimentos:

1. Atribuímos valores aleatórios de variáveis dispostos em duas colunas e nove linhas.
2. Calculamos a correlação individual entre os pares de números aleatórios.
3. Agregamos os pares de números aleatórios em três grupos com três componentes, cujos valores foram a média desses componentes.
4. Calculamos a correlação entre os pares de valores de cada grupo.
5. Verificamos a variação dos valores de correlação.
6. Ordenamos as variáveis aleatórias.
7. Repetimos os procedimentos 1 a 5.

Não observamos o aumento da correlação entre variáveis ao agregar dados aleatórios não ordenados em todos os casos, como o exemplificado na Tabela 30.

Tabela 30 – Teste com valores aleatórios não ordenados

Números aleatórios		Média	
Variável 1	Variável 2	Variável 1	Variável 2
9,9	0,0	9,1	5,9
9,6	8,2		
7,7	9,4		
7,3	4,3	6,3	5,5
2,8	9,9		
8,9	2,3		
5,7	10,0	7,2	7,7
8,5	4,3		
7,5	8,8		
Correlação individual	-0,633361	Correlação grupal	-0,027355

Fonte: elaboração do autor.

Porém, ao ordenar as variáveis aleatórias, ocorreu o aumento da força da correlação na agregação de dados, em todos os testes, como o exemplificado na Tabela 31.

Tabela 31 – Teste com valores aleatórios ordenados

Números aleatórios ordenados		Média	
Variável 1	Variável 2	Variável 1	Variável 2
2,8	0,0	5,2	2,2
5,7	2,3		
7,3	4,3		
7,5	4,3	7,9	7,1

continua

Números aleatórios ordenados		Média	
7,7	8,2		
8,5	8,8		
8,9	9,4	9,5	9,8
9,6	9,9		
9,9	10,0		
Correlação individual	0,937040	Correlação grupal	0,999766

Fonte: elaboração do autor.

No agrupamento das UF em regiões (Tabela 32), os estados das regiões Norte e Nordeste possuem renda *per capita* e proficiências médias inferiores às dos alunos das outras regiões. Na agregação, os valores da correlação ecológica ampliaram-se, de maneira análoga aos testes com valores aleatórios ordenados. Diante disso, podemos considerar que, na agregação, a localização das zonas interferiu no aumento do valor da correlação ecológica.

Tabela 32 – Correlações ecológicas entre a médias da renda *per capita* e proficiência no Enem dos alunos concluintes de escolas públicas que pontuaram em todas as provas e na redação, de 2012 a 2018, por UF e por região

Estado	Média (UF)		Região	Média (Região)	
	Renda <i>per capita</i>	Proficiência		Renda <i>per capita</i>	Proficiência
Alagoas	0,25	3,01	Nordeste	0,26	3,15
Bahia	0,31	3,22			
Ceará	0,22	3,09			
Maranhão	0,24	2,94			
Paraíba	0,24	3,23			
Pernambuco	0,29	3,28			
Piauí	0,23	3,06			

continua

Estado	Média (UF)		Região	Média (Região)	
	Renda <i>per capita</i>	Proficiência		Renda <i>per capita</i>	Proficiência
Rio Grande do Norte	0,31	3,23			
Sergipe	0,27	3,23			
Acre	0,37	3,06	Norte	0,37	3,13
Amapá	0,35	3,08			
Amazonas	0,37	3,05			
Pará	0,30	3,16			
Rondônia	0,53	3,40			
Roraima	0,47	3,17			
Tocantins	0,44	3,00			
Distrito Federal	0,72	3,78			
Goiás	0,50	3,31			
Mato Grosso	0,58	3,20			
Mato Grosso do Sul	0,56	3,43			
Espírito Santo	0,51	3,54	Sudeste	0,59	3,61
Minas Gerais	0,52	3,67			
Rio de Janeiro	0,55	3,69			
São Paulo	0,65	3,57			
Paraná	0,68	3,56	Sul	0,70	3,64
Rio Grande do Sul	0,70	3,74			
Santa Catarina	0,76	3,62			
Correlação ecológica por UF	0,811960		Correlação ecológica por região	0,941846	

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

O MAUP evidencia-se na comparação entre zoneamentos. Esse problema diz respeito à variação dos valores das medidas estatísticas quando ocorre a agregação espacial dos dados. O entendimento desse problema pode auxiliar na interpretação e evitar conclusões equivocadas relativas às falácias descritas por Robinson (1950) e Alker (1969).

A agregação aleatória ordenada (Tabela 31) e o agrupamento de UF em regiões (Tabela 32) evidenciaram o argumento de Clark e Avery (1976) de que a agregação, aparentemente, tem o efeito de agrupar valores altos das variáveis e, em outros grupos, valores mais baixos das variáveis, diminuindo a heterogeneidade entre os pares de variáveis e resultando na diminuição da autocorrelação espacial, quando regiões maiores são formadas. Isso explicaria a ausência desta autocorrelação no zoneamento regiões (Figura 26).

6.4.1. Falácias

A falácia ecológica, descrita por Robinson (1950), refere-se a uma incorreta suposição de que a relação entre variáveis observadas na forma agregada ou ecológicas é a mesma que a dos dados individuais (JARGOWSKY, 2005). Procuramos exemplificar não só a falácia ecológica, mas também as de outros tipos descritos por Alker (1969).

O desempenho médio elevado dos alunos agregados em uma zona não necessariamente representou um bom resultado médio de cada aluno (falácia ecológica) ou de cada subzona (falácia de escala). Em Ipanguaçu, RN, por exemplo, a proficiência média dos alunos concluintes do ensino médio no ENEM foi 4,3. Porém o desempenho elevado derivou, principalmente, das proficiências que obtiveram os alunos de uma das escolas desse município (Tabela 33).

Tabela 33 – Proficiências médias por escola obtidas pelos alunos concluintes do ensino médio no ENEM, no município de Ipanguaçu, RN (2012-2018)

Escola	alunos	Proficiência		
		Média	Mínima	Máxima
EE João Francisco da Costa	14	2,29	0	5
EE Manoel de Melo Montenegro	215	2,83	0	5
EE Maria da Glória de A. Luna	21	1,90	0	3
IF de Educação, Ciência e Tecnologia do RN – Campus Ipanguaçu	478	4,68	2	5

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Com resultados médios extremos, nos municípios de Cocal dos Alves, PI, e de Jatobá, MA, há alunos que se mostraram plenamente proficientes no ENEM e alunos com fraco desempenho. Não podemos afirmar que não houve alunos que obtiveram proficiência plena no município de pior rendimento no ENEM, assim como não podemos afirmar que o município de melhor rendimento médio no Brasil não teve alunos de fraco desempenho (falácia ecológica) (Tabela 34).

Tabela 34 – Municípios com maior e menor proficiência média dos alunos concluintes do ensino médio de 16 a 18 anos no ENEM (2012-2018)

Município	Escolas	Alunos	Proficiência		
			Média	Mínima	Máxima
Cocal dos Alves, PI	1	105	4,6	2	5
Jatobá, MA	3	154	1,7	0	5

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

Os melhores exemplos das falácias de escala podem ser encontrados no estado de Roraima, onde 15.255 alunos concluintes do ensino médio pontuaram em todas as provas e na redação de 2012 a 2018 e obtiveram 3,17 de proficiência média.

A microrregião de Boa Vista é composta por quatro municípios e a proficiência média superior a três não poderia ser atribuída a três municípios (Tabela 35).

Tabela 35 – Proficiência média por município dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM da microrregião de Boa Vista (2012-2018)

Município	Alunos	Proficiência média
Alto Alegre	401	2,57
Amajari	133	2,87
Pacaraima	193	2,97
Boa Vista	12.281	3,30

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A mesorregião Norte de Roraima apresenta proficiência média de 3,23 e é composta por duas microrregiões (Tabela 36). Não podemos atribuir o resultado médio da mesorregião à microrregião Nordeste de Roraima.

Tabela 36 – Proficiência média por microrregião dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM da mesorregião Norte de Roraima (2012-2018)

Microrregião	Alunos	Proficiência
Boa Vista	13.008	3,26
Nordeste de Roraima	564	2,59

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

No estado de Roraima, a proficiência média foi 3,17. Esse resultado é próximo da proficiência média da mesorregião Norte de Roraima e bastante diferente da proficiência média da mesorregião Sul de Roraima, de forma que não podemos atribuir a esta o resultado médio superior a três proficiências dos concluintes desse estado federativo (Tabela 37).

Tabela 37 – Proficiência média por mesorregião dos alunos concluintes do ensino médio de 16 a 18 anos que pontuaram em todas as provas do ENEM do Estado de Roraima (2012-2018).

Mesorregião	Alunos	Proficiência
Norte de Roraima	13.572	3,23
Sul de Roraima	1.683	2,64

Fonte: INEP (2013, 2015a, 2015b, 2017a, 2017b, 2018a, 2019a).

A falácia de escala ocorreu de forma mais intensa no estado de Roraima em todos os níveis de agregação porque 80,5% dos 15.255 concluintes do ensino médio das escolas públicas de até 18 anos estudaram nas unidades escolares de Boa Vista, capital do estado.

7. CONCLUSÕES

Este trabalho objetivou identificar o zoneamento mais eficiente para representar os alunos concluintes do ensino médio no ENEM. A metodologia empregada possibilitou reconhecer os zoneamentos eficientes e os não eficientes para a representação espacial do desempenho no ENEM, correlacionado com a renda familiar *per capita*. Sem esse parâmetro, dificilmente poderíamos afirmar que o zoneamento UF é ineficiente, ainda que possamos verificar a incoerência desse zoneamento em relação aos três aprovados na comparação da autocorrelação espacial bivariada (Figura 26).

De 2012 a 2018, os zoneamentos municípios, microrregiões e mesorregiões foram aprovados para a representação espacial do desempenho dos alunos concluintes do ensino médio de 16 a 18 anos correlacionado com a renda familiar *per capita* no ENEM. Porém, o zoneamento municípios não foi aprovado para representar a evolução temporal, por conter zonas com nenhum ou poucos alunos em um ou mais anos. O zoneamento microrregiões foi o mais eficiente para representar o desempenho e sua evolução temporal no período avaliado, considerando-se a efetividade da aferição dos resultados de todas as provas no ENEM por meio da TRI.

De modo geral, por um lado, as proficiências médias nas microrregiões são menos elevadas nas regiões Norte e Nordeste do Brasil. Na classe de renda mais baixa (até 0,2 SM), os melhores desempenhos médios ocorreram principalmente nas zonas das regiões Sul e Sudeste. Na classe acima de 0,8 SM, os estados que apresentaram mais zonas com desempenho maior no ENEM foram Minas Gerais, Rio Grande do Sul, Rio de Janeiro e Espírito Santo. Por outro lado, os maiores aumentos da proficiência média no ENEM ocorreram, principalmente, nos estados do Norte e Nordeste e para os concluintes de menor renda familiar *per capita*.

A sobreposição das áreas metropolitanas de maior PIB, nos zoneamentos aprovados com menor número de zonas, microrregiões e mesorregiões, indicou que essas áreas metropolitanas se localizam, majoritariamente, nas zonas de maior desempenho médio no ENEM, para as faixas de proficiências utilizadas. Porém, ao analisar esse desempenho por faixa de renda familiar *per capita*, essa localização predominante das áreas metropolitanas em zonas de melhor desempenho no ENEM

não ocorreu. No zoneamento municípios, a maioria das zonas de melhor desempenho não se localizaram em regiões metropolitanas de maior PIB.

Por um lado, considerando o coeficiente de variação das pontuações médias anuais, de 2012 a 2018, dos concluintes do ensino médio de até 18 anos da rede pública de ensino, as pontuações médias anuais por área do conhecimento ou pela redação dificultaria o acompanhamento de resultado dos participantes no exame ao longo do tempo. Por outro lado, a somatória ou a média das pontuações de todas as provas mostrou-se mais apropriada para esse acompanhamento temporal.

No período estudado, os alunos das escolas particulares participaram mais e obtiveram melhores rendimentos médios tanto em proficiência quanto em pontuação do que os alunos das escolas públicas. As escolas particulares estavam presentes em 29,4% dos municípios brasileiros, enquanto as públicas, em 99,8%. A distribuição da renda familiar era diversa para essas esferas administrativas. Essas características dificultaram a avaliação conjunta de todos os alunos.

Apesar de a pontuação média no ENEM apresentar correlação individual e ecológica forte e positiva com a proficiência, a evolução temporal dessas variáveis não ocorre da mesma forma. A pontuação média aumentou de 2012 a 2018 (Figura 7), enquanto a proficiência plena se manteve estável de 2012 a 2015, aumentou em 2016 e 2017 e diminuiu em 2018 (Figura 10).

Nesse último ano, também ocorreu uma diminuição do número de alunos concluintes do ensino médio. A partir de 2017, diminuiu o número de inscritos no ENEM. Por último, a partir de 2016 houve uma redução dos alunos concluintes que pontuaram em todas as provas e na redação. Apesar desses números nos últimos anos, a escala temporal utilizada nesta pesquisa não foi suficiente para prever tendências. Seria necessário analisar os anos subsequentes para compreender melhor a variação do número de alunos inscritos e seu desempenho.

A certificação da proficiência plena ou parcial foi determinada pelo INEP para os alunos acima de 18 anos que não concluíram o ensino médio, com base nos resultados obtidos nas provas objetivas e na redação no ENEM. Nessa pesquisa, verificamos que a maioria dos concluintes do ensino médio de até 18 anos não obteve proficiência plena, a qual seria pré-requisito para a obtenção do certificado do ensino médio pelos avaliados mais velhos no exame. Dessa forma, os mapas na Figura 19, Figura 20 e Figura 21 evidenciaram não só as áreas nas quais os concluintes do

ensino médio de até 18 obtiveram as maiores proficiências médias, mas também as áreas onde o poder público mais falhou em prover as condições necessárias para que os alunos concluintes do ensino médio obtivessem a proficiência adequada para essa etapa escolar.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ALKER, H. R. A Typology of Ecological Fallacies. In: DOGAN, M.; ROKKAN, S. (Eds.). **Quantitative Ecological Analysis in the Social Sciences**. Cambridge: MIT Press, 1969. p. 69-86.

ANSELIN, L. Local Indicators of Spatial Association – LISA. **Geographical Analysis**, v. 27, n. 2, p. 93-115, 1995.

ANSELIN, L.; SYABRI, I.; KHO, Y. GeoDa: An Introduction to Spatial Data Analysis. **Geographical Analysis**, v. 38, n. 1, p. 5-22, 2006.

ANSELIN, L.; SYABRI, I.; SMIRNOV, O. Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows. **Urbana**, v. 51, p. 8, 2002.

BANKSTON III, C. L.; CALDAS, S. J. Family Structure, Schoolmates, and Racial Inequalities in School Achievement. **Journal of Marriage and the Family**, v. 60, n. 3, p. 715-723, ago. 1998.

BRASIL. Ministério da Educação. Portaria n° 438 de 28 de maio de 1998. **Diário Oficial da União**, Brasília, 1º maio 1998. Seção 1, p. 5.

_____. Portaria n° 462 de 27 de maio de 2009. **Diário Oficial da União**, Brasília, 28 maio 2009. Seção 1, p. 54.

_____. Ministério da Educação. Nota Técnica. Assunto: Procedimento de cálculo das notas do ENEM. p. 1-5, 2011. Disponível em: <http://download.inep.gov.br/educacao_basica/enem/nota_tecnica/2011/nota_tecnica_procedimento_de_calculo_das_notas_enem_2.pdf>. Acesso em: 7 maio 2020.

_____. Lei n° 12.799 de 10 de abril de 2013. Dispõe sobre a isenção de pagamento de taxas para inscrição em processos seletivos de ingresso nos cursos das instituições federais de educação superior. **Diário Oficial da União**, Brasília, 10 abr. 2013. Seção 1, p. 1.

BURT, C. Experimental Tests of Higher Mental Processes and their Relation to General Intelligence. **Journal of Experimental Pedagogy**, v. 1, p. 93-112, 1911.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 6.ed. São Paulo: Saraiva, 2010.

CÂMARA, G. et al. Análise espacial de áreas. In: DRUCK, S. et al. (Eds.). **Análise espacial de dados geográficos**. Brasília, DF: Embrapa, 2004. p. 107-150.

CLARK, W. A. V.; AVERY, K. L. The Effects of Data Aggregation in Statistical Analysis. **Geographical Analysis**, v. 8, n. 4, p. 428-438, 1976.

COLLINS, J. H.; DOUGLASS, H. R. The Socio-Economic Status of the Home as a Factor in Success in the Junior High School. **The Elementary School Journal**, v. 38, n. 2, p. 107-113, out. 1937.

CONTRERAS, D. et al. Income Inequality or Performance Gap? A Multilevel Study of School Violence in 52 Countries. **Journal of Adolescent Health**, v. 57, n. 5, p. 545-552, 2015.

CUNHA, J. M. P. da et al. Social Segregation and Academic Achievement in State-run Elementary Schools in the Municipality of Campinas, Brazil. **Geoforum**, v. 40, n. 5, p. 873-883, 2009.

DA SILVA, R.; LAMB, L. C.; BARBOSA, M. C. Universality, Correlations, and Rankings in the Brazilian Universities National Admission Examinations. **Physica A: Statistical Mechanics and its Applications**, v. 457, p. 295-306, 2016.

DUFF, J. F.; THOMSON, G. H. The Social and Geographical Distribution of Intelligence in Northumberland. **British Journal of Psychology**, General Section, v. 14, n. 2, p. 192-198, out. 1923.

FERREIRA, M. C. **Iniciação à análise geoespacial: teoria, técnicas e exemplos para geoprocessamento**. São Paulo: Editora Unesp, 2014.

FLEMING, C. M. Socio-Economic Level and Test Performance. **British Journal of Educational Psychology**, v. 13, n. 2, p. 74-82, 1943.

GALTON, F. Co-Relations and Their Measurement, Chiefly from Anthropometric Data. **Proceedings of the Royal Society of London**, v. 45, p. 135-145, 1888.

GEHLKE, C. E.; BIEHL, K. Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material. **Journal of the American Statistical Association**, v. 29, n. 185, p. 169, 1934.

IBGE. **Divisão do Brasil em mesorregiões e microrregiões geográficas**. Departamento de Geografia. Rio de Janeiro, 1990. Disponível em: <https://servicodados.ibge.gov.br/Download/Download.ashx?http=1&u=biblioteca.ibge.gov.br/visualizacao/livros/liv2269_1.pdf>. Acesso em: 12 dez. 2019.

_____. **Geoprocessamento no IBGE**. Rio de Janeiro, 1994. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv24556.pdf>>. Acesso em: 2 set. 2018.

INEP. Microdados do ENEM 2011. Brasília, 2012. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 1º mar. 2017.

_____. Microdados do ENEM 2012. Brasília, 2013. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 1º mar. 2017.

INEP. Portaria nº 179, de 28 de abril de 2014. **Diário Oficial da União**, Brasília, 29 abr. 2014. Seção 1, p. 40.

_____. Microdados do ENEM 2013. Brasília, 2015a. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 7 jun. 2017.

_____. Microdados do ENEM 2014. Brasília, 2015b. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 7 jun. 2017.

_____. Microdados do Censo da Educação Básica 2012. Brasília, 2016a. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 21 maio. 2018.

_____. Microdados do Censo da Educação Básica 2013. Brasília, 2016b. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 21 maio. 2018.

_____. Microdados do Censo da Educação Básica 2014. Brasília, 2016c. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 19 maio. 2018.

_____. Microdados do Censo da Educação Básica 2015. Brasília, 2016d. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 21 maio. 2018.

_____. Microdados do Censo da Educação Básica 2016. Brasília, 2016e. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 21 maio. 2018.

_____. Microdados do ENEM 2015. Brasília, 2017a. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 29 maio. 2017.

_____. Microdados do ENEM 2016. Brasília, 2017b. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 10 maio. 2018.

_____. Microdados do ENEM 2017. Brasília, 2018a. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 21 maio. 2018.

_____. Microdados do Censo da Educação Básica 2017. Brasília, 2018b. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 26 maio. 2018.

_____. Microdados do ENEM 2018. Brasília, 2019a. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 15 jul. 2019.

_____. Microdados do Censo da Educação Básica 2018. Brasília, 2019b. Disponível em: <<http://inep.gov.br/microdados>>. Acesso em: 15 jul. 2019.

JARGOWSKY, P. A. Ecological Fallacy. **Encyclopedia of Social Measurement**, v. 1, p. 715-722, 2005.

JELINSKI, D. E.; WU, J. The Modifiable Areal Unit Problem and Implications for Landscape Ecology. **Landscape Ecology**, v. 11, n. 3, p. 129-140, jun. 1996.

JONES, D. C.; CARR-SAUNDERS, A. M. The Relation between Intelligence and Social Status among Orphan Children. **British Journal of Psychology, General Section**, [S. I.], v. 17, n. 4, p. 343-364, 1927.

KÖCHE, J. C. **Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa**. 23.ed. Petrópolis: Vozes, 1997.

MANLEY, D. Scale, Aggregation, and the Modifiable Areal Unit Problem. In: MANFRED M. FISCHER; NIJKAMP, P. (Eds.). **Handbook of Regional Science**. Berlin: Springer, 2014. p. 1157-1171.

MARKS, G. N. Are Father's or Mother's Socioeconomic Characteristics more Important Influences on Student Performance? Recent International Evidence. **Social Indicators Research**, v. 85, n. 2, p. 293-309, 2008.

MEYER, R. H. Value-Added Indicators of School Performance: A Primer. **Economics of Education Review**, v. 16, n. 3, p. 283-301, 1997.

NEPRASH, J. A. Some Problems in the Correlation of Spatially Distributed Variables. **Journal of the American Statistical Association**, v. 29, n. 185, p. 167-168, 1934.

OPENSHAW, S. The Modifiable Areal Unit Problem. **Concepts and Techniques in Modern Geography**, v. 38, p. 169-174, 1984.

OPENSHAW, S.; TAYLOR, P. A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem. **Statistical Methods in the Spatial Sciences**, London, p. 127-144, 1979.

PATTERSON, C. J.; KUPERSMIDT, J. B.; VADEN, N. A. Income Level, Gender, Ethnicity, and Household Composition as Predictors of Children's School-Based Competence. **Child Development**, v. 61, n. 2, p. 485-494, 1990.

PEARSON, K. VII. Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia. **Philosophical Transactions of the Royal Society of London**. Series A, Containing Papers of a Mathematical or Physical Character, v. 187, p. 253-318, 31 dez. 1896.

ROBINSON, W. S. Ecological Correlations and the Behavior of Individuals. **American Sociological Review**, v. 15, n. 3, p. 351-357, jun. 1950.

SOARES, T. M. Utilização da teoria da resposta ao item na produção de indicadores sócioeconômicos. **Pesquisa Operacional**, v. 25, n. 1, p. 83-112, 2005.

TEKWE, C. D. et al. An Empirical Comparison of Statistical Models for Value-Added Assessment of School Performance. **Journal of Educational and Behavioral Statistics**, v. 29, n. 1, p. 11-36, 23 mar. 2004.

THIEME, C. et al. Value Added, Educational Accountability Approaches and their Effects on Schools' Rankings: Evidence from Chile. **European Journal of Operational Research**, v. 253, n. 2, p. 456-471, 2016.

TOBLER, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. **Economic Geography**, v. 46, n. 332, p. 234-240, 14 jun. 1970.

TORRES, H. da G. et al. Pobreza e espaço: padrões de segregação em São Paulo. **Estudos Avançados**, v. 17, n. 47, p. 97-128, abr. 2003.

TRAVITZKI, R. Avaliação da qualidade do ENEM 2009 e 2011 com técnicas psicométricas. **Estudos em Avaliação Educacional**, v. 28, n. 67, p. 256-288, 2017.

TRAVITZKI, R.; CALERO, J.; BOTO, C. ¿Qué información proporciona el examen nacional de enseñanza media (Enem) a la sociedad Brasileña? **Cepal Review**, n. 113, p. 163-181, 2014.

WHITE, K. R. The Relation between Socioeconomic Status and Academic Achievement. **Psychological Bulletin**, v. 91, n. 3, p. 461-481, 1982.

WODTKE, G. T.; HARDING, D. J.; ELWERT, F. Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation. **American Sociological Review**, v. 76, n. 5, p. 713-736, 2011.

YULE, G. U.; KENDALL, M. G. **An Introduction to the Theory of Statistics**. 5.ed. London: Charles Griffin, 1950.