



CÂMARA DOS DEPUTADOS  
CENTRO DE FORMAÇÃO, TREINAMENTO E APERFEIÇOAMENTO  
PROGRAMA DE PÓS-GRADUAÇÃO  
MESTRADO PROFISSIONAL EM PODER LEGISLATIVO

**Maurício de Moraes Rêgo Soares**

**ANÁLISE COMPARADA DOS MODELOS DE MEDIÇÃO DA GEOGRAFIA  
ELEITORAL: estudo da efetividade por meio de abordagem quantitativa**

Brasília  
2018

**Maurício de Moraes Rêgo Soares**

**ANÁLISE COMPARADA DOS MODELOS DE MEDIÇÃO DA GEOGRAFIA  
ELEITORAL: estudo da efetividade por meio de abordagem quantitativa**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do grau de Mestre no Curso de Mestrado Profissional do Programa de Pós-Graduação do Centro de Formação, Treinamento e Aperfeiçoamento da Câmara dos Deputados/Cefor.

Orientador Prof. Dr.: Fabiano Peruzzo Schwartz

Área de Concentração: Poder Legislativo

Linha de Pesquisa: Processos Políticos do Poder Legislativo

Brasília  
2018

## Autorização

Autorizo a divulgação do texto completo no sítio da Câmara dos Deputados e a reprodução total ou parcial, exclusivamente, para fins acadêmicos e científicos.

Assinatura: \_\_\_\_\_

Data: \_\_/\_\_/\_\_

Soares, Maurício de Moraes Rêgo.

Análise comparada dos modelos de medição da geografia eleitoral  
[manuscrito] : estudo da efetividade por meio de abordagem quantitativa /  
Maurício de Moraes Rêgo Soares. -- 2018.

99 f.

Orientador: Fabiano Peruzzo Schwartz.

Impresso por computador.

Dissertação (mestrado) -- Câmara dos Deputados, Centro de Formação,  
Treinamento e Aperfeiçoamento (Cefor), 2018.

1. Sistema eleitoral, Brasil. 2. Voto, distribuição, Brasil. I. Título.

CDU 342.8(81)

Bibliotecária: Fabyola Lima Madeira – CRB1: 2109



## FOLHA DE APROVAÇÃO

**Título:** ANÁLISE COMPARADA DOS MODELOS DE MEDIÇÃO DA GEOGRAFIA ELEITORAL: UMA ABORDAGEM QUANTITATIVA.

**Autor (a):** Maurício de Moraes Rego Soares

**Área de concentração:** Poder Legislativo

**Linha de pesquisa:** Gestão Pública no Poder Legislativo

Dissertação submetida à Comissão Examinadora designada pela Coordenação do Programa de Pós-graduação do Centro de Formação, Aperfeiçoamento e Treinamento da Câmara dos Deputados como requisito parcial para obtenção do título de **Mestre** em Poder Legislativo.

Dissertação aprovada em 10 de dezembro de 2018.

---

**Prof. Dr.<sup>a</sup> Prof. Dr. Fabiano Peruzzo Schwartz**  
Presidente da Banca - Câmara dos Deputados

---

**Prof. Dr. João Luiz Pereira Marciano**  
Membro - Câmara dos Deputados

---

**Prof. Dr. David Verge Fleischer**  
Membro - UnB

Dedico este trabalho aos meus pais Elita e Wanderley (*in memoriam*),  
pelo incentivo aos estudos.

## **Agradecimentos**

Ao meu filho Iran pela companhia.  
A minha esposa Gláucia pela companhia e  
pelo apoio, além da inestimável contribuição com a revisão deste trabalho.  
Ao meu orientador Fabiano por me nortear nos caminhos que percorri nessa jornada.

## Resumo

Este trabalho se dedica a avaliar os modelos de mensuração da geografia eleitoral usados na literatura da Ciência Política brasileira. Para isso, foi feita uma análise comparativa dos três índices mais utilizados para aferir a concentração eleitoral das votações parlamentares, conhecidos por N, G e I de Moran, juntamente com um novo indicador, o índice M. O novo indicador foi proposto por este estudo com o objetivo de se constituir em alternativa de medição que pese proporcionalmente as áreas (municípios) testadas em face dos votos de todo o estado. O estudo se fundamenta em pesquisa de caráter exploratório-explicativa, com abordagem quantitativa aplicada a dados eleitorais das votações dos 513 candidatos eleitos para deputado federal nos 26 estados e no Distrito Federal, nas eleições de 2014. Os índices N, G, Moran e M foram testados usando a agregação espacial do município, sendo que o índice N também foi testado empregando a agregação das zonas eleitorais, e o índice M, utilizando um mapa de agregação espacial misto, combinando municípios e zonas eleitorais de grandes municípios. A partir desses dados, foi possível compor um quadro comparativo com os coeficientes dos indicadores de todos os estados. Os resultados mostraram diferença entre os indicadores. O índice N por município foi aquele que apresentou menor capacidade de traduzir a distribuição eleitoral dos candidatos. Os índices G e N por zona eleitoral conseguiram espelhar melhor a realidade das votações em seus coeficientes, especialmente o índice N por zona eleitoral que, por utilizar uma agregação eleitoral com menor variação de tamanho, obteve melhores resultados do que aqueles apresentados pelo N por município. Os coeficientes calculados pelo I de Moran não foram significativos. Já o índice M, em comparação aos demais indicadores, apresentou os resultados mais satisfatórios para representar as votações dos deputados eleitos, em especial, os cálculos feitos usando a agregação mista de votos. Por fim, o índice M foi testado utilizando as seções de votação como base de agregação eleitoral em pequenos estados. Essa agregação eleitoral foi mais eficiente em refletir a distribuição eleitoral quando comparada com as demais agregações em estados de mesma magnitude.

**Palavras-Chave:** Poder Legislativo. Geografia Eleitoral. Índice de Concentração Eleitoral. Densidade Eleitoral. Índice M.

## **Abstract**

This research evaluates most used metric models for electoral geography in the literature of Brazilian Political Science. In order to accomplish that, it was developed a comparative analysis of the two most used indices to measure the electoral concentration of votes, called N and G, together with a new indicator, the M index. The new indicator was proposed by this study as an alternative metric which consider the proportional weight of the areas (counties) tested in comparison of votes of the entire state. The study is based on an exploratory research, with quantitative approach and applied to the voting of 513 candidates elected to Congress in 26 States and the Federal District, in 2014 elections. N, G and M indices were tested by using the spatial aggregation of counties. The N index was also tested by using the aggregation of congressional district, and M index was tested by using a mixed aggregation of counties and congressional district when in larger counties. These data made it possible to build a comparative table with the coefficients of the indicators in those states. The results showed a difference between the indicators. The N index by counties showed the least capacity to translate the voting distribution. The G index and N index by congressional district were able to reflect a better reality of polls by their coefficients. The N index by congressional district produced better results by using an electoral aggregation with smaller variation of size than those presented by N by county. The coefficients calculated by index I of Moran have not achieved significant results. The M index, in comparison to the other indicators, presented better results in representing the votes of the deputies elected, specially when it was used a mixed spatial aggregation of votes. Finally, the index M was tested using the congressional district as a base of spatial aggregation of votes in smaller states. Such a spatial aggregation of votes was more efficient in reflecting the electoral distribution than other spatial aggregations in smaller states.

**Key-words:** Legislative Power. Electoral Geography. Electoral Concentration Index. Electoral Density. M Index.



## LISTA DE TABELAS

Tabela 1 – Número de áreas e médias dos índices por estado.....	62
Tabela 2 – Médias por região do número de áreas por agregação mista, MZE, MZEn e percentual da maior área de agregação de cada estado .....	66
Tabela 3 – Densidades das votações por estado .....	67
Tabela 4 – Densidades das votações por região .....	68
Tabela 5 – Índices dos deputados federais eleitos por Minas Gerais em 2014 .....	69
Tabela 6 – Índices dos deputados federais eleitos pelo Rio de Janeiro em 2014.....	73
Tabela 7 – Índices dos deputados federais eleitos pelo Acre em 2014 .....	75
Tabela 8 – Índices dos deputados federais eleitos pelo Rio Grande do Sul em 2014 .....	77
Tabela 9 – Índices dos deputados federais eleitos por Pernambuco em 2014.....	79
Tabela 10 – Comparação dos índices de cinco candidatos de estados diferentes .....	80
Tabela 11 – Número e maior área e médias M com a agregação mista .....	83
Tabela 12 – Número e maior área e médias M com a agregação por seção eleitoral.....	83

## LISTA DE QUADROS

Quadro 1 – Modelo Proposto por Ames (2003) .....	26
Quadro 2 – Índices de medição de concentração/dispersão da geografia eleitoral .....	39
Quadro 3 – Comparação dos dados eleitorais de Minas Gerais e Rio de Janeiro .....	43
Quadro 4 – Votos por Município dos Candidatos Selecionados para Análise a partir de Simulação .....	47
Quadro 5 – Índices de Concentração para os Candidatos Selecionados da Simulação.....	48
Quadro 6 – Votos por Município dos Candidatos Selecionados para Simulação .....	51
Quadro 7 – Síntese dos problemas dos índices de concentração eleitoral.....	53
Quadro 8 – Índices de medição da geografia eleitoral .....	57

## LISTA DE FIGURAS

Figura 1 – Padrão espacial das variáveis .....	32
Figura 2 – Matriz de ponderação, contiguidade queen de primeira ordem .....	33
Figura 3 – Diagrama de Dispersão de Moran.....	36
Figura 4 – Mapa de cluster .....	37
Figura 5 – Mapa de significância .....	38

## SUMÁRIO

1. INTRODUÇÃO.....	12
1.1 OBJETIVOS.....	17
1.2 JUSTIFICATIVA.....	18
2. OS MODELOS DE MEDIÇÃO DA GEOGRAFIA ELEITORAL.....	19
2.1 Índice de Rae .....	22
2.2 Índice N .....	23
2.3 Índice G .....	28
2.4 Índice I de Moran.....	31
2.5 Síntese dos índices de correlação espacial.....	39
3 AS FRAGILIDADES DOS ÍNDICES DE MEDIÇÃO DA GEOGRAFIA ELEITORAL ..	40
3.1 Comparação entre os estados do Rio de Janeiro e Minas Gerais – estudo exploratório. 42	
3.2 Comparação entre agregações espaciais .....	44
3.3 Vantagens e desvantagens do índice G .....	46
3.4 Síntese do problema dos índices de concentração eleitoral.....	52
3.5 A proposta de um novo indicador: o índice M.....	53
4 METODOLOGIA.....	58
5 RESULTADOS E DISCUSSÕES SOBRE OS ÍNDICES.....	61
5.1 Comparação dos índices em todos os estados.....	61
5.2 Comparações de índices de votações do mesmo estado .....	68
5.3 Comparação de votações de estados diferentes .....	79
5.4 Estados com agregação eleitoral por seção de votação.....	82
6 CONCLUSÃO.....	85
REFERÊNCIAS .....	87
APÊNDICE – CÓDIGO R.....	91

## 1. INTRODUÇÃO

Na Ciência Política, uma forma de se entender o comportamento parlamentar é por meio do perfil espacial da votação recebida pelo candidato em uma eleição. Isto porque, se partirmos da premissa que um dos principais ou o principal objetivo de um político é a sua manutenção no poder, conclui-se que, para alcançar tal intento, buscar a reeleição é nada mais que uma meta natural do mandato. Dessa forma, faz sentido que o eleito procure atender aos anseios de seus eleitores, visando à recompensa futura da reeleição. Esse é o raciocínio que sustenta o conceito de conexão eleitoral, que foi desenvolvido por David Mayhew (1974) em seus estudos sobre o Congresso americano. Simões define bem essa teoria no trecho:

[...] os parlamentares são atores políticos racionais e que entre seus principais interesses está a manutenção de sua carreira política, ou seja, a (re) eleição está no centro de suas preocupações. Sendo a reeleição o principal objetivo dos parlamentares, o que mais lhes interessa é garantir uma boa relação frente a seus eleitores. E de que maneira eles conseguem cumprir esse objetivo? Moldando os recursos e as estruturas institucionais internas do Congresso com o objetivo de viabilizar o seu uso estratégico. Significa dizer que a chave explicativa para a produção e organização legislativas se encontra no “momento eleitoral”, na forma como os representantes se elegem. Daí se estabelece o que Mayhew chama de *a lógica da conexão eleitoral*. (SIMÕES, 2016, p.2)

Assim, fica nítida a importância de se conhecer onde estão localizados os votos dos candidatos eleitos, para se entender um vetor fundamental de incentivo à atuação parlamentar. Esse é o objeto de estudo da geografia eleitoral, que é o tema desta pesquisa e tem sido abordado no Brasil a partir de reflexões que fazem comparações com o sistema político americano. Nos Estados Unidos a conexão eleitoral é clara: devido ao seu sistema eleitoral ser majoritário, o político tem todos os incentivos para buscar recursos para suas bases eleitorais, que são geograficamente definidas.

No Brasil, por adotarmos o sistema proporcional, em princípio a arena eleitoral geraria incentivos universalistas aos parlamentares, diferindo do sistema americano. Porém, nosso sistema multipartidário, de lista aberta e com grandes distritos, torna as avaliações mais complexas. Devido a essas especificidades, verifica-se que aqui são possíveis tanto estratégias eleitorais típicas de sistemas proporcionais que visem atingir eleitores espalhados pelo distrito eleitoral - no caso brasileiro representado pelo estado -, quanto estratégias que se concentrem em áreas geograficamente limitadas do eleitorado para obtenção do voto, típicas de sistemas distritais. Para descobrir que tipo de conexão eleitoral é predominante em nosso sistema, desenvolveram-se modelos que pretendem medir o nível de concentração espacial das

votações dos candidatos para pleitos legislativos. Nesse sentido, quanto maior for a concentração dos votos recebidos pelo parlamentar, maior será o incentivo que ele terá para atuar em favor dessas localidades, já que foram importantes para sua eleição.

Esse tipo de conexão se assemelha àquela existente no sistema eleitoral majoritário americano, que tem como consequência o modelo batizado de distributivista. A teoria distributivista, conhecida como política de *pork barrel*, diz que o deputado orienta sua atuação parlamentar de forma a carrear recursos do Estado em proveito de sua base, no intuito de auferir dividendos eleitorais; no Brasil, essa teoria é identificada com o paroquialismo político. Nesta perspectiva, Ames (1995) sustenta que a transferência de recursos públicos do orçamento em prol de regiões específicas do eleitorado é um fator que impulsiona o sucesso eleitoral do candidato.

No caso contrário, ao se perceber que o eleito teve sua votação de forma dispersa, deve-se concluir que, apesar das especificidades de nosso sistema eleitoral, este consegue manter a característica fundamental da proporcionalidade, que consiste em representar frações de opinião relevantes de um distrito. A ação parlamentar neste tipo de conexão eleitoral tende a visar o interesse universal dos eleitores.

Por conta da citada atipia do sistema político eleitoral brasileiro, não é possível afirmar a priori qual tipo de conexão eleitoral está mais presente na nossa realidade política, se de cunho paroquial ou universal. Para compreender melhor o comportamento parlamentar no Brasil cabe, então, descobrir qual dos dois perfis de distribuição geográfica das votações está mais presente em nossos pleitos para o Legislativo: se aquele que busca votos dispersos por todo o distrito, característico da proporcionalidade; ou se aquele que busca concentrar a votação em áreas determinadas, formando distritos informais, condizentes com sistemas majoritários.

Seguindo o caminho aberto por precursores nos estudos da geografia eleitoral, Barry Ames (2003) desenvolveu uma elaborada metodologia para mensurar o nível de dominância eleitoral de um candidato - ao qual combinou esse vetor de análise com outro já existente e que mede a concentração espacial dos votos recebidos pelos candidatos -, usando uma medida de estatística espacial chamada de I de Moran. Assim, o autor utilizou dois recortes sobrepostos de análise das votações: um vertical, que se refere à dominância e calcula o percentual de votos que o candidato recebeu dos votos totais de um município, variando assim seu perfil de dominante a compartilhado; e outro vetor horizontal, que se refere à concentração e calcula a agregação espacial dos votos recebidos pelos candidatos nos municípios em que ele obteve melhor desempenho, criando deste modo um eixo que varia de

concentrado a disperso.

Carvalho (2003) adota a mesma linha de análise, conjugando dois vetores geográficos de leitura das votações: um vertical com outro horizontal, com a diferença que ele usa o índice de fragmentação de Rae<sup>1</sup> para aferir o vetor horizontal de distribuição espacial dos votos. Tal método de observação da geografia eleitoral tornou-se predominante entre os autores que analisam o perfil das votações dos candidatos e dele derivam as conclusões acerca do tema na Ciência Política brasileira contemporânea. Falando sobre os estudos de geografia do voto, Santos (2015, p.7) afirma que Ames e Carvalho são os “[...] que possuem os trabalhos mais comentados da atualidade.”

Nos estudos sobre o tema, observa-se que o município é a unidade fundamental de análise, tanto para o cálculo da dominância, quanto para o cálculo da concentração eleitoral. Tal adoção parece coerente quando se trata da dominância, pois nesse caso, que se refere ao recorte vertical das votações, se deseja aferir o quanto determinado político controla ou partilha o eleitorado de uma determinada área, assim não há problema de se adotar o município para fazer essa avaliação, uma vez que o índice pondera o peso do município frente ao percentual de votos total do candidato.

Já no caso do recorte horizontal, que mede a distribuição espacial dos votos e varia em uma escala de concentrado a fragmentado, a tradicional abordagem a partir do município não parece ser a mais adequada. Isso porque há grande variação no tamanho dos municípios; na proporção dos seus votos em relação ao tamanho dos distritos (estados); e no número de votos necessários para eleição de um deputado, que depende da magnitude do distrito (estado), que não é levada em consideração na composição do índice de concentração horizontal dos votos.

Gonçalves (2015) fez um levantamento dos trabalhos que utilizaram a espacialidade como ferramenta de análise das eleições presidenciais e percebeu a preponderância da utilização do município como unidade espacial dos estudos e constatou o problema da variabilidade de tamanho dos mesmos, nas suas palavras:

O município é a unidade espacial de análise mais utilizada pelos trabalhos, presente em 11 deles, muito provavelmente pela facilidade de encontrar dados sociais georeferenciados a essa unidade. Embora os municípios sejam unidades estáveis e reconhecidas, as análises não levam em conta a população diferenciada de cada unidade, super e sub-representando os municípios, [...] São Paulo-SP com quase 9 milhões de eleitores fica em pé de igualdade com Araguinha-MT de menos de 900 eleitores, e o desvio padrão ( $\sigma = 154.739$ ) mostra que a variação de eleitores em relação a média é alta, unidades espaciais extremamente heterogêneas quanto ao número de eleitores. (GONÇALVES, 2015, p.376)

---

<sup>1</sup> Ver item 2.1

Dessa forma, não parece adequada a métrica mais usada no Brasil atualmente para se medir a concentração espacial dos votos dos candidatos ao Legislativo. Ela não contempla a variabilidade da magnitude dos municípios que formam os distritos (estados), estes mais amplos e também de diversos tamanhos e que, dependendo da dimensão, podem eleger vários deputados com seus votos ou contribuir com uma pequena fração necessária para eleger apenas um.

Como exemplo dessa aparente contradição, podemos comparar as cidades do Rio de Janeiro e de Belo Horizonte, ambas capitais, mas cada uma com participação percentual de eleitores bem diferente nos seus estados. A primeira tem cerca de 37% do eleitorado do estado, já a segunda conta com cerca de 11% dos votos de Minas Gerais. Fica claro, então, que concentrações semelhantes de votações recebidas nessas cidades terão significados distintos, ou seja, um candidato votado no Rio, com a mesma concentração eleitoral que outro em Belo Horizonte, terá maior chance de ser eleito, já que a primeira cidade tem eleitorado proporcionalmente maior no seu estado que a segunda. Apesar disso, o índice comumente usado para medir a concentração-dispersão eleitoral não traduz essas nuances da demografia dos votos.

Como consequência de tal inconsistência presente na literatura, mas não percebida pela maioria dos estudiosos do tema, pode-se citar o caso de Carvalho (2003), que no detalhado estudo que realizou sobre a distribuição da geografia eleitoral brasileira, notou que muitos dos deputados federais eleitos nos anos de 1994 e 1998 para a Câmara dos Deputados, tiveram perfil de votação com alto grau de concentração, tendo recebido seus votos em grandes colégios eleitorais, como de capitais e cidades de regiões metropolitanas. Diferente do que supõe a teoria para candidatos com votações concentradas, nesses casos os eleitos não possuíam perfil paroquial.

Isso se explica porque o índice utilizado para medir a concentração espacial das votações não pondera a magnitude dos municípios analisados, porém o autor toma tais casos como exceções, sem observar qualquer problema no método de mensuração empregado. Interpretando tal achado de outro modo, o presente trabalho levanta a hipótese de que a ineficiência do índice largamente usado pela literatura é a causa desse tipo de classificação equivocada.

Outra questão relevante que se coloca na avaliação da geografia eleitoral é o da contiguidade. Os métodos amplamente usados para aferir a concentração dos votos não levam em conta a adjacência dos municípios testados. Na maior parte dos estudos, diante de índices

de concentração altos, supõe-se que as áreas onde os votos foram dados estão espacialmente próximas. Porém, apesar de ser uma consequência intuitiva, essa inferência não é demonstrada empiricamente, pois os índices usados não testam essa característica das votações. Carvalho pressupõe a contiguidade como inferência natural derivada da concentração, no trecho:

Vale assinalar, ainda, que o cenário espacial, ou seja, a contiguidade ou não-contiguidade dos municípios, se vê indiretamente contemplada. É de se esperar a correlação positiva entre concentração e contiguidade espacial, por um lado, e fragmentação e descontiguidade espacial, por outro – sobretudo nas faixas extremas. (CARVALHO, 2003, p.96)

Pioneiro também nesse aspecto, Ames (2003) utiliza o índice de Moran, que considera no seu cálculo a contiguidade das áreas pesquisadas, para avaliar votações dos candidatos à Câmara Federal; contudo, o autor não especifica como chegou aos seus resultados, nem detalha como o índice funciona. Nesse sentido, trabalhos mais recentes também utilizam o I de Moran, como os de Marzagão (2013), que analisou a base geográfica de Dilma Roussef na eleição presidencial de 2010, e de Nicolau e Terron (2012), que procurou explorar a configuração social e espacial dos candidatos no segundo turno nas eleições para prefeito do Rio de Janeiro em 2008. Entretanto, a maioria desses estudos trata de eleições majoritárias para o Executivo. Já o estudo de Maciel (2014), que analisa os padrões espaciais de voto, bases sociais e políticas dos deputados federais brasileiros entre 1994 e 2010, constitui-se uma exceção a essa regra.

Nessa perspectiva, a partir da identificação das fragilidades dos modelos vigentes de cálculo da concentração eleitoral no Brasil, cabe avançar na análise de como aperfeiçoar os métodos já existentes ou desenvolver um novo método mais condizente com a realidade do sistema eleitoral brasileiro, que deve considerar a ponderação da magnitude dos colégios eleitorais (municípios), além de ter em vista a proximidade espacial do eleitorado, para ser um tradutor mais fidedigno dos perfis de votação dos candidatos no Brasil. Dessa forma, é possível verificar se esses perfis estão mais próximos da lógica distributivista paroquial ou se têm contornos típicos da proporcionalidade, e em que medida há presença de um ou de outro em nosso sistema político.

Para tentar elucidar essas dúvidas, este trabalho se fundamenta em pesquisa de caráter exploratório-explicativa, com abordagem quantitativa aplicada a dados eleitorais em eleições proporcionais para a Câmara dos Deputados no ano de 2014. As dúvidas suscitadas pelas aparentes deficiências nos indicadores que medem a distribuição espacial das votações são o cerne do problema de pesquisa deste estudo: **Em que medida os índices propostos para**



**mensurar a concentração/dispersão das votações refletem a geografia eleitoral no Brasil? Quais são as suas virtudes e limitações? Há uma forma mais adequada para tais aferições?**

A hipótese que se pretende averiguar é a de que os índices mais largamente utilizados pela literatura para medir a concentração eleitoral no Brasil, em especial o índice N, não refletem a realidade das votações, pois não ponderam o peso relativo dos municípios no total de votos de um estado. A teoria de base utilizada é a da Geografia Eleitoral.

## **1.1 OBJETIVOS**

O objetivo principal deste trabalho é avaliar a efetividade dos indicadores de medição da geografia eleitoral quanto ao poder explicativo que os mesmos têm em relação à observação empírica do processo eleitoral.

Como objetivos específicos, destacam-se:

- comparar os índices de concentração eleitoral aplicados nas votações para deputado federal, apresentando os prós e contras de cada um deles;
- propor um novo indicador, o índice M, baseado em um novo conceito de densidade eleitoral, que seja capaz de medir a geografia eleitoral de estados com diferentes demografias;
- aplicar e avaliar o índice M em todos estados e no Distrito Federal;
- compor um novo mapa de agregação eleitoral, mesclando municípios e zonas eleitorais de grandes municípios para serem utilizados como base de aplicação para o índice M;
- avaliar o índice M tendo como áreas de agregação eleitoral em pequenos estados as seções de votação;
- construir base de dados com vistas ao estudo da efetividade dos índices e disponibilizar, em repositório público, o conteúdo e scripts de programação utilizados na geração.

## 1.2 JUSTIFICATIVA

Este estudo pretende contribuir com a produção de conhecimento sobre geografia eleitoral na Ciência Política brasileira a partir de uma perspectiva crítica, que visa ampliar o conhecimento sobre os métodos existentes de medição da concentração/dispersão eleitoral. Conhecer a distribuição eleitoral de um político – se ele recebeu votos distribuídos por todo estado ou se seu eleitorado está concentrado em poucas cidades – é o ponto essencial das discussões sobre o tema da conexão eleitoral e impacta diretamente na avaliação se aquele candidato eleito terá perfil paroquial ou se terá preocupações mais universais no exercício do mandato.

Por isso, é importante que um indicador que se proponha a medir a geografia eleitoral seja o mais representativo possível da realidade das votações, cabendo, nesse contexto, conhecer a fundo as principais potencialidades e limitações do indicador. Além disso, o debate em torno de uma reforma política se vê historicamente influenciado por estudos acadêmicos relacionados às características do sistema eleitoral e o tipo de representação que ele produz.

Segundo o professor Gláucio Soares (2001, p.294), o modelo de representação política adotado pelo país não reflete adequadamente a distribuição espacial da população, nem a estrutura de poder e de dominação, nem a estrutura de propriedades dos bens de produção, nem a estrutura de classes. Fleischer (1976), em seu estudo sobre a concentração geográfica dos deputados eleitos em Minas Gerais nas eleições de 1966, 1970 e 1974, dentre outras conclusões, afirma que:

Para os deputados eleitos nos últimos três pleitos, observa-se [...] que um sistema eleitoral seria mais prejudicial para os federais que os estaduais. Fato que leva alguns estudiosos do assunto a sugerir que se adote o sistema alemão (deputados gerais e distritais) para eleições federais, e o puro para as estaduais. (FLEISCHER, 1976, p. 25)

Nesse sentido, fica claro que uma parte significativa das questões que permeiam a análise das distorções do sistema eleitoral brasileiro é influenciada diretamente pelos resultados produzidos em estudos sobre a geografia eleitoral. Tal fato constitui a base que justifica os propósitos do presente trabalho.

## 2. OS MODELOS DE MEDIÇÃO DA GEOGRAFIA ELEITORAL

O tema da geografia eleitoral surge na Ciência Política brasileira seguindo as reflexões trazidas por estudos sobre o Congresso americano. Pereira e Rennó (2001) analisaram as chances de reeleição dos deputados federais no pleito de 1998 e concluíram que, apesar de coexistirem tanto estratégias políticas visando o posicionamento de opinião em relação a questões nacionais, quanto estratégias que colocam como prioritário o interesse local, as últimas “prevalecem na arena eleitoral porque as demandas locais parecem ter impacto mais forte no sucesso eleitoral.” Em oposição a essa visão, Samuels (2002) argumenta a partir de outra perspectiva. Para o autor, quando o deputado leva recursos públicos do orçamento para uma localidade, na realidade ele está visando captar financiamento privado para sua campanha de reeleição, em vez de almejar o reconhecimento do eleitor por ter carregado esses benefícios para a sua região. Ambos os estudos quantificaram a distribuição eleitoral dos candidatos para chegar às suas conclusões.

Carvalho (2003) faz um extenso trabalho de pesquisa, conjugando uma série de levantamentos referente à atuação parlamentar com a aferição da distribuição eleitoral do deputado. O autor verificou que cerca de metade dos parlamentares foram eleitos com votação concentrada/dominante e dispersa/dominante, evidenciando um perfil mais condizente com o modelo distributivista em que o deputado pauta sua atuação na obtenção de recursos para o reduto eleitoral que o elegeu.

Tais trabalhos estão atrelados à origem do institucionalismo da escolha racional e, nessa perspectiva teórica, partem da ideia de que, dadas as regras do jogo, os atores políticos se comportam de forma instrumental, a fim de maximizar seus ganhos mediante às condições estabelecidas e levando em conta também o comportamento previsível dos demais atores. Suas preferências são prévias e visam apenas ao autointeresse, mas são condicionadas pelas normas institucionais vigentes.

A abordagem racionalista na Ciência Política é advinda da Teoria Econômica, que preceitua, na mesma linha, que os indivíduos determinam suas escolhas por uma lógica puramente focada na relação risco/retorno de suas ações. Nessa visão, é coerente supor previamente os interesses dos indivíduos, pois eles estão baseados numa lógica universal, onde as especificidades das diversas culturas e sociedades não têm maior influência. De forma racional, os atores calculam sua estratégia sob a perspectiva utilitarista, baseados na expectativa previsível do comportamento dos demais atores.

Outro aspecto importante da teoria da escolha racional é o papel que as instituições cumprem, gerando balizas para os atores políticos. As instituições criam um conjunto de incentivos que possibilita o atendimento de interesses individuais de maneira coordenada num arranjo comum entre os participantes. A abordagem feita por David Mayhew (1974) sobre o Congresso norte-americano é um bom exemplo disso. O autor americano vê que, de um lado, o sistema eleitoral majoritário gera incentivos distributivistas para os deputados americanos. Como o parlamentar é eleito em distritos uninominais – aqueles em que o eleitorado elege apenas um representante –, sua preocupação durante o mandato seria o atendimento de demandas específicas da sua base eleitoral, dando pouca atenção a questões de âmbito universal. Dessa forma sua atuação seria orientada a buscar a maximização das suas chances de reeleição.

Essa dispersão de interesses poderia gerar dificuldade de governabilidade, que de fato não se verificava. Isso porque, nessa visão, a organização do processo legislativo do Congresso, baseado no papel central ocupado pelas comissões, viabilizaria a aprovação das matérias dos deputados, em uma espécie de troca-troca de aprovações, tornando um jogo que inicialmente seria concorrencial de todos contra todos pela aprovação das suas matérias, por um jogo colaborativo em que todos cooperam.

Partindo desse modo de abordar a realidade política e institucional, os estudos iniciais sobre a geografia eleitoral brasileira procuraram entender a dinâmica da atuação parlamentar em nosso Legislativo, tendo como vetor explicativo fundamental a demografia das votações. Se na realidade americana a conexão eleitoral é mais simples, pois o deputado é eleito por meio do sistema majoritário em distritos uninominais, no Brasil, o sistema proporcional de lista aberta que elege deputados em distritos plurinominais torna nosso sistema incomum e mais difícil de ser interpretado, o que motivou os estudiosos a aprofundar o conhecimento do tema.

Um dos pioneiros na investigação de nossa geografia eleitoral foi David Fleischer (1976), que se debruçou sobre a concentração eleitoral em Minas Gerais nos pleitos para deputados federais e estaduais de 1966, 1970 e 1974. Seu intuito foi conhecer a distribuição geográfica das votações e suas consequências para a representação regional e partidária, além de testar se já haveria na época uma suposta distritalização das votações dos deputados e descobrir quais seriam as consequências, caso se adotasse o sistema majoritário. Nessa área de estudo, Fleischer e Ribeiro (1983) publicaram ainda um trabalho que simulou as probabilidades de êxito eleitoral dos candidatos e, de forma geral, as consequências, caso fosse adotado o sistema majoritário na eleição para Câmara dos Deputados em 1978.

No primeiro estudo, Fleischer adotou as votações dos deputados nas zonas eleitorais como unidade de análise: as zonas que obtivessem mais de 1% do total de votos recebidos pelo candidato passariam a figurar no mapa. A partir da composição do mapa, foram analisados os aglomerados de áreas mais ou menos contíguas em que se distribuía a maioria dos votos recebidos pelo candidato, formando assim redutos ou núcleos eleitorais (*clusters*). Em sequência, calculou-se uma série de valores: o número de zonas eleitorais com mais de 1% dos votos; a quantidade de redutos observados; a quantidade de zonas do maior reduto; e o percentual dos votos totais do maior reduto. Baseado nesses dados, foi avaliado o nível de concentração-dispersão eleitoral dos candidatos.

Entre outros achados, Fleischer concluiu que, diferente do que se supunha, os candidatos da Arena<sup>2</sup> não seriam beneficiados em detrimento dos do MDB<sup>3</sup> no caso de se adotar o sistema distrital, ao contrário: “Nota-se, também, uma ligeira tendência do MDB de ter uma maior proporção dos seus deputados entre as fileiras dos ‘distritáveis’ do que a Arena, principalmente no plano estadual.” (FLEISCHER, 1976, p.25). Em vista dos resultados da pesquisa o autor aponta em suas conclusões um argumento em favor do sistema distrital:

Constatou-se que o atual sistema de eleições proporcionais não significa uma representação política fiel à distribuição geográfica da população mineira, deixando algumas regiões sub-representadas e outras super-representadas. Fenômeno este que não ocorreria com o sistema distrital. (FLEISCHER, 1976, p.33)

Já o segundo trabalho analisou as votações recebidas por todos os candidatos nos estados de Pernambuco, Sergipe, Bahia, Espírito Santo, Rio de Janeiro e Santa Catarina, caso o sistema de escolha fosse distrital. O estudo constatou que, se esse fosse o nosso sistema eleitoral, a distribuição de candidatos distritáveis se daria de forma desigual dentre os estados testados, tendo Bahia e Pernambuco menos candidatos com perfis de eleitorado distritáveis do que não-distritáveis. Já nos demais estados a situação seria inversa. Outra conclusão dos autores corrobora os resultados do trabalho anterior: os candidatos do MDB seriam beneficiados, pois seus perfis eleitorais se apresentavam mais distritáveis que dos candidatos da Arena. Nas palavras do autor:

O MDB teria, assim, mais eleitos distritáveis que a ARENA e, mais ainda, os possíveis adeptos do sistema distrital formariam maioria dentro do MDB e minoria dentro da ARENA. [...] a distribuição natural da votação do MDB coadunar-se-ia melhor com o sistema distrital. (FLEISCHER; RIBERIRO, 1983, p.232)

---

<sup>2</sup> Aliança Renovadora Nacional, partido que dava sustentação política à ditadura militar.

<sup>3</sup> Movimento Democrático Brasileiro, partido de oposição à ditadura militar.

## 2.1 Índice de Rae

Outros autores seguiram o caminho aberto por Fleischer, como Indjaian (1981), que fez uma “Análise preliminar da bancada federal paulista eleita pelo MDB em 1978”, utilizando o índice de fragmentação de Rae, para mensurar o grau de concentração/dispersão eleitoral. Segundo Nicolau (2005, p. 599), tal índice foi proposto por Douglas Rae, cientista político norte-americano, que com o intuito de medir “[...] a concentração/dispersão de votos em uma eleição, ou das cadeiras no Legislativo [...]” sob o prisma da dispersão, trouxe da economia o Índice de concentração de Herfindal-Hirschman (HH) invertendo-o, ao subtrair-lo de 1, onde,  $Rae = 1 - HH$ . Jairo Nicolau explica que:

Nos anos 1940, um índice foi criado por economistas com o intuito de mensurar a concentração de firmas em um determinado setor. Imagine-se, por exemplo, a distribuição de empresas no mercado de automóveis. Como poderemos mensurar o padrão de concentração/dispersão, levando em conta a fatia de mercado que cada empresa controla? O índice de concentração de Herfindal-Hirschman (HH), como ficou conhecido em homenagem aos seus autores, é calculado de maneira bastante simples: Índice de concentração de Herfindal-Hirschman (HH):  $HH = \sum_{i=1}^n p_i^2$  (onde  $p_i$  é a proporção do mercado controlado pela empresa  $i$ ). O índice varia de 0 a 1; quanto mais próximo de 1, maior a concentração. Imagine-se, por exemplo, o mercado de refrigerantes, com cinco firmas, cada uma delas com a seguinte fatia: 40%, 30%, 15%, 10% e 5%. O índice HH seria calculado da seguinte forma:  $(0,40 \times 0,40) + (0,30 \times 0,30) + (0,15 \times 0,15) + (0,10 \times 0,10) + (0,05 \times 0,05)$ . O resultado é igual a 0,29. (NICOLAU, 2005, p.598)

Para o cálculo do índice de Rae, teríamos a equação:

$$Rae = 1 - \sum_{i=1}^n \left( \frac{Vm}{Ve} \right)^2 \quad (2.1)$$

Sendo:

$n$ , o número de municípios;

$Vm$ , os votos do deputado no município;

$Ve$ , os votos do deputado em todo o estado.

Dito de outra forma, a apuração do índice se expressa subtraindo 1 da soma dos quadrados das divisões entre os votos de um candidato em cada município pelos votos recebidos em todo o estado. Assim, o índice de Rae varia em um intervalo de zero (máxima dispersão) a 1 (máxima concentração). Esse indicador tem por base de cálculo os percentuais das votações do candidato em cada município pelos votos totais que ele obtém em todo o estado. No índice de Rae, quanto menor for o número de municípios onde o candidato recebe suas maiores votações maior será o coeficiente de concentração apurado, isto é, o resultado do

indicador está associado ao número de municípios onde foi distribuída a votação do candidato, sem considerar a magnitude desses municípios.

Em relação ao trabalho de Indjaian, Kinzo, Borin e Martins Jr. dizem que, calculando os percentuais obtidos pelos candidatos nos dez municípios mais votados,

Os resultados indicaram elevada concentração espacial dos votos recebidos pelos deputados eleitos, independentemente do total de votos obtidos, da corrente político-ideológica a que se filiavam ou da experiência política anterior. (KINZO; BORIN; MARTINS JR., 2003, p. 47, nota nº 7)

## 2.2 Índice N

Na mesma área de estudo, Dias (1991) utilizou o índice N de Laakso e Taagepera, que normalmente é usado para mensurar o número efetivo de partidos (NEP), para medir o nível de competitividade nas eleições para deputados estadual e federal nas circunscrições (zonas eleitorais da capital e municípios do interior do estado) do Rio de Janeiro, em 1982 e 1986, e no Paraná, nos pleitos de 1982 e 1990. Aqui, o índice N foi usado para medir o número efetivo de candidatos concorrentes em uma circunscrição.

Desse modo, quanto maior o número de candidatos efetivos, maior será a competitividade e fragmentação eleitoral nessas áreas e, inversamente, quanto menor o número de candidatos, menor será a competitividade e fragmentação eleitoral nas respectivas localidades. Seguindo esse raciocínio, quanto menor a competitividade eleitoral maior será a tendência a distritalização. O autor conclui que:

No que se refere à competição política no Estado, quando avaliada pelo prisma da concentração/dispersão do voto [...] detectamos alguns aspectos importantes. Sob o ponto de vista das representações eleitas, percebemos não soa manifestação clara, no plano federal e estadual, em ambos os estados, do fenômeno da distritalização do voto, como os traços gerais de seu processo de transformação. (DIAS, 1991, p.80)

Assim como o indicador de Rae, o índice de Laakso e Taagepera também advém de uma operacionalização do índice HH; o valor é obtido pela divisão de 1 por HH, ou  $N = 1/HH$ , expresso na equação:

$$N = \frac{1}{\sum_{i=1}^n \left(\frac{Vm}{Ve}\right)^2} \quad (2.2)$$

Em sua proposta original, a vantagem desse índice em relação ao índice de Rae seria a de supostamente resultar no número efetivo de partidos. Quando aplicado a cálculos de

concentração eleitoral, o índice de Laakso e Taagepera traduziria o número efetivo de municípios onde o candidato foi votado, ou seja, seu coeficiente varia de 1 (máxima concentração), diminuindo a concentração eleitoral conforme aumenta o valor do coeficiente, não tendo limite máximo de dispersão.

Do mesmo modo que o índice de Rae, o índice N tem seu resultado obtido a partir do número de municípios em que o candidato concentra sua votação. Quanto maior o número de municípios onde a votação está distribuída, mais disperso será o coeficiente resultante; do contrário, quanto menor for o número de municípios onde o político foi votado, mais concentrado será o resultado do indicador.

Ainda que seja uma importante referência nos estudos de geografia eleitoral, a pesquisa de Dias comete uma incorreção. Apesar de ser uma das primeiras pesquisas a usar o índice N, o autor diz textualmente ter usado o índice de Rae, em suas palavras “[...] utilizaremos a seguir o conhecido índice de fragmentação de Rae, ligeiramente modificado.” (DIAS, 1991, p. 61). Tal equívoco do autor explica uma série de citações erradas posteriores na literatura, atribuindo-se ao índice Rae, quando na verdade era usado o índice N. Verifica-se o erro em três exemplos hipotéticos apresentados pelo autor: neles, os índices resultam nos valores 4,4, 6,5 e 11,1 - impossíveis na escala Rae, que varia de 0 a 1 -, no entanto, condizentes com a variação do índice de Laakso e Taagepera.

Seguindo a mesma linha de pesquisa, Kinzo utiliza o índice N como ferramenta para “[...] identificar o número de competidores que efetivamente disputaram os votos nos distritos paulistanos [...]” (2003, p. 53). O autor chega à conclusão de que a competição eleitoral foi alta nos pleitos de 1992, 1996 e 2000 para a Câmara Municipal paulistana, afirmando que:

[...] é surpreendente o nível de votação individual. Mesmo entre os distritos menores inexitem casos em que haja um único competidor dominando a disputa; ao contrário, considerando-se as três eleições, o N mínimo encontrado foi de 3,5. (KINZO; BORIN; MARTINS JR., 2003, p.54)

Percebe-se assim que, tanto o índice N quanto o índice de Rae, podem ser usados de diferentes modos, desde aferir a concentração partidária de um Parlamento, além da competitividade de um distrito, até, como veremos a seguir, avaliar a concentração/dispersão eleitoral da votação recebida pelos candidatos, sendo o último caso o foco deste trabalho.

Estabelecendo um marco nas pesquisas sobre a geografia eleitoral, Barry Ames (2003) propõe uma taxonomia mais sofisticada de padrões espaciais das votações, partindo de dois eixos de análise: um horizontal e outro vertical. O primeiro eixo corresponde exatamente a já citada dicotomia concentração/fragmentação, de acordo com Ames (1995, p. 63) “[...] a distribuição espacial dos municípios em que o candidato se sai bem.” Para operacionalizar



essa mensuração, Ames parte novamente do cálculo da

[...] percentagem que coube ao candidato  $i$  da totalidade dos votos dados em cada município, mas essa dimensão utiliza uma medida estatística denominada Moran I, que calcula a distribuição espacial dos municípios em que o candidato tem bom desempenho. (AMES, 2003, p.65)

No texto não fica claro como Ames compôs esse cálculo, nem o critério que determinaria os municípios em que o candidato tem bom desempenho. São apresentados alguns exemplos-modelo com mapas mostrando a distribuição de alguns deputados, mas sem explicar de forma objetiva como se chegou a esses resultados. Câmara, Monteiro, Fucks e Carvalho (2004) e Cardoso (2004) explicitam melhor como se chega a Moran I:

A dependência espacial pode ser medida de diferentes formas. O índice de Moran (I) é a estatística mais difundida e mede a autocorrelação espacial a partir do produto dos desvios em relação à média. Este índice é uma medida global da autocorrelação espacial, pois indica o grau de associação espacial presente no conjunto de dados. (CARDOSO, 2004, p. 2)

De uma forma geral, o índice de Moran presta-se a um teste cuja hipótese nula é de independência espacial; neste caso, seu valor seria zero. Valores positivos (entre 0 e +1) indicam para correlação direta e negativos, (entre 0 e - 1) correlação inversa. (CÂMARA; MONTEIRO; FUCKS e CARVALHO, 2004, p.14)

O segundo vetor de análise utilizado por Ames para determinar o perfil eleitoral de um candidato varia de dominante a compartilhado, explicado pelo autor no trecho: “Definimos dominação municipal de cada candidato como a cota do candidato dos votos totais dados aos membros de todos partidos.” (AMES, 1995, p. 63). Sendo “[...] a dominância média de cada candidato em todos os municípios do estado, ponderada pela percentagem da totalidade dos votos do candidato para a qual cada município contribui.” (AMES, 2003, p. 65). Desta forma, a dominância se expressa na equação:

$$D = \sum_{i=1}^n \left( \frac{Vm}{Ve} \right) \left( \frac{Vm}{VM} \right) \quad (2.3)$$

Sendo:

$n$ , o número de municípios;

$Vm$ , os votos do deputado no município;

$Ve$ , os votos do deputado em todo o estado;

$VM$ , os votos totais do município.

Assim, o índice de dominância aferiria o percentual de votos de um candidato em cada

um dos municípios de um estado ponderado pelo peso dos votos desses municípios no total de votos do estado. Com a dimensão da dominância eleitoral, Ames adiciona à literatura mais uma ferramenta de avaliação da distribuição espacial dos sufrágios.

A partir da combinação desses dois eixos, então, é possível produzir quatro padrões espaciais: concentrado-dominante, concentrado-compartilhado, disperso-compartilhado e disperso-dominante. O autor compôs uma matriz que categorizou de forma mais precisa a variabilidade dos tipos de votação recebidos pelos candidatos ao Legislativo (vide Quadro 1). É importante destacar que o autor não explicita a partir de quais critérios classificou o que seria uma votação dispersa ou concentrada, nem distingue a dicotomia dominante-compartilhada.

Quadro 1 – Modelo Proposto por Ames (2003)

Padrões estaduais de distribuição espacial % do total de votos em municípios principais			
		Baixa	Alta
Distribuição espacial de municípios principais	Dispersa	Dispersa-compartilhada	Dispersa-dominante
	Contígua	Concentrada-compartilhada	Concentrada-dominante

FONTE: Ames (2003), p. 65

No estudo de Ames (1995) fica implícita a conclusão de que, para se eleger no Brasil, são possíveis diferentes estratégias, tanto aquelas em que os candidatos procuram ter votações concentradas em regiões limitadas, típicas da política paroquial, quanto aquelas em que os candidatos eleitos têm seu eleitorado distribuído por vários municípios do estado, comumente associadas à atuação de políticas universais. Ambas as configurações de base eleitoral estariam disponíveis aos políticos que se candidatassem. Apesar disso, o autor foca suas conclusões em relação à predominância do voto paroquial, característica de redutos clientelistas. Isso se percebe nos trechos:

Os deputados com distribuições dominantes de votos conseguem resistir mais a oscilações partidárias do que aqueles com distribuições compartilhadas. Mas a maioria dos deputados ganha pouco concentrando suas distribuições de votos ou fazendo apelos ideológicos ou a grupos, e os padrões de carreiras não têm um efeito amplo sobre a sorte em eleições. (AMES, 1995, p. 63)

O sistema eleitoral brasileiro induz os deputados a prover benefícios para clientelas eleitorais. (AMES, 2003, p. 130)

Em seu livro “E no início eram as bases”, Carvalho (2003), com pequenas diferenças

metodológicas, segue a mesma linha iniciada por Ames. Nesse trabalho, tendo por base os padrões espaciais propostos por Ames, ele avança em análises estatísticas das votações recebidas pelos deputados eleitos, incluindo nessa análise o nível de competição eleitoral dos municípios; as diferenças nas dicotomias urbano-rural e capital-interior; e a distribuição partidária dos eleitos, além de outras variáveis de cunho legislativo, como projetos de lei e emendas orçamentárias.

Carvalho (2003) adota o mesmo modelo proposto por Ames no Quadro 1; ele utilizou também dois vetores de análise da demografia eleitoral. Seu índice de dominância é o mesmo criado por Ames. Já para avaliar a concentração eleitoral ele usa o índice de fragmentação de Rae, que “[...] usualmente, se adota para determinar o número efetivo de partidos políticos presentes nos sistemas políticos.” (CARVALHO, 2003, p. 95), e também é utilizado para medir o nível de competitividade eleitoral em um município.

Carvalho (2003) categoriza o índice em quatro faixas de concentração: concentração alta, em que o índice varia de 1,0 a 4,0; concentração média, que varia de 4,0 a 7,0; dispersão média, que varia de 7,0 a 11,0; e dispersão alta, para índice maior que 11,0. Essas categorias foram determinadas de forma arbitrária, sem se explicar a metodologia da adoção de tais faixas de concentração/fragmentação eleitoral: se havia algum significado objetivo na definição desses valores ou se foi apenas uma convenção determinada pela conveniência. Ainda assim, foi a partir desses parâmetros que se desenvolveu todo o restante do livro e se analisou a dicotomia concentração/fragmentação na geografia eleitoral brasileira.

Vale assinalar que Carvalho (2003) comete o mesmo equívoco de Dias (1991): apesar de se referir ao índice de Rae, constata-se pelos valores medidos que, na realidade, ele utiliza o índice N de Laakso e Taagepera. Talvez isso se explique por Dias ter sido uma de suas referências.

Na obra referência de Carvalho (2003), o autor analisou os dados das votações para deputado federal nos anos de 1994 e 1998; ele concluiu que não há preponderância de eleitos com votação concentrada, defendendo a tese de que, em nosso sistema, é possível haver várias estratégias eleitorais.

Verificamos, do ponto de vista nacional, o equilíbrio entre os deputados eleitos com votação concentrada (46%) e os deputados de votação fragmentada (54%). Nosso sistema proporcional – em sua operacionalização – parecia contemplar de forma praticamente equânime tanto a fragmentação como a concentração de votos. (CARVALHO, 2003, p. 97)

A ampla pesquisa empreendida por Carvalho (2003), que de forma aprofundada demonstrou a relação entre a demografia do voto com a atuação parlamentar, tornou-se uma

das principais referências no campo dos estudos sobre conexão eleitoral no Brasil. Vários outros trabalhos seguiram nessa direção interpretativa da geografia eleitoral, como o de Corrêa (2012, p. 2), que analisando os “[...] padrões espaciais de votação dos deputados estaduais eleitos em 2006 na RMRJ<sup>4</sup> [...]”, chega à conclusão de que mesmo quando a concentração não está aliada à dominância eleitoral se verifica um viés localista na votação do parlamentar:

Diferentemente da formulação teórica dada por Ames (2003) para os incentivos ao comportamento parlamentar oriundos do padrão espacial de votação concentrado-dominante no nível dos municípios, consideramos que a concentração somada ao compartilhamento de votação no interior da RMRJ, é indicativa de relação estrita com o território. (CORRÊA, 2012, p. 17)

Pode-se citar, dentre outros trabalhos acadêmicos na área da geografia eleitoral, os de Vieira (2012), Borges, Paula e Silva (2016) e Saugo (2007), que apesar de fazerem diferentes abordagens para estudar o sistema político regional e nacional, se basearam na classificação de demografia eleitoral construída e aperfeiçoada por Ames e Carvalho, que é balizada no duplo recorte vertical/horizontal como indicativo dos níveis de dominância e concentração das votações. Nessas pesquisas o município se destaca como unidade básica eleitoral de análise.

### 2.3 Índice G

Já em “A Concentração Eleitoral nas Eleições Paulistas: Medidas e Aplicações”, de Avelino, Biderman e Silva (2011), surge uma reflexão mais complexa relativa à conexão eleitoral. Nesse estudo, os autores propõem uma nova forma para mensurar a concentração eleitoral, por meio de um novo e único índice, que pretende unir as dimensões horizontal e vertical de distribuição do voto. Os propositores batizaram-lhe de índice G. A nova métrica veio da

[...] adaptação de um índice amplamente utilizado na literatura de economia regional para identificar o grau de concentração geográfica dos setores. Este índice, geralmente denominado por G (que será denominação adotada neste artigo), foi utilizado pela primeira vez na literatura de ciências regionais por Florence (1947). (AVELINO; BIDERMAN; SILVA, 2011, p.324)

O índice proposto por Avelino, Biderman e Silva (2011) é expresso na equação:

---

<sup>4</sup> Região Metropolitana do Rio de Janeiro

$$G = \sum_{i=1}^n \left( \frac{Vm}{Ve} - \frac{VM}{V} \right)^2 \quad (2.4)$$

Sendo:

$n$ , o número de municípios;

$Vm$ , os votos do deputado no município;

$Ve$ , os votos do deputado em todo o estado;

$VM$ , os votos totais do município;

$V$ , todos votos para deputado no estado.

O índice  $G$ , assim como os outros índices de concentração, calcula a proporção de votos do deputado no município pelos seus votos recebidos em todo estado, mas, ao invés de imediatamente elevar esse valor ao quadrado como nos índices anteriormente citados, primeiro se deduz daquele valor a proporção de votos totais do município pelos votos totais para deputado no estado. O índice  $G$  leva em conta, para sua operacionalização, o peso do eleitorado do município em todo o estado, preenchendo à primeira vista uma lacuna que observamos nos índices mais usados na literatura. Na explicação dos autores:

[...] o índice deduz dos votos observados para um dado deputado a votação que seria esperada para esse mesmo deputado caso os votos fossem distribuídos de forma aleatória entre os municípios. Uma maneira intuitiva de entender este índice é observando que se um deputado tivesse os seus votos distribuídos exatamente de acordo com a distribuição dos eleitores no estado seu índice seria zero. (AVELINO; BIDERMAN; SILVA, 2011, p. 324)

O índice  $G$  mede a dicotomia dispersão/concentração eleitoral em uma escala que parte do zero, que seria a máxima dispersão eleitoral, e varia positivamente, indicando maior concentração eleitoral de uma votação. Na composição desse indicador, os autores adotaram o conceito de dispersão com o sentido de proporcionalidade. Essa variação do índice é explicitada no trecho:

O índice  $G$ , amplamente utilizado em trabalhos de economia regional, possui um contrafactual claro: o valor será zero quando a dispersão for total. O critério de dispersão total também é bastante claro: se a proporção de votos de um candidato em uma determinada região é idêntica à proporção dessa região no colégio eleitoral, o candidato não pode ser considerado concentrado nessa região. Para o caso da análise eleitoral,  $G$  será zero se a distribuição dos votos de um candidato for igual à distribuição dos eleitores no estado. (AVELINO; BIDERMAN; SILVA, 2011, p. 342)

Em acordo com a crítica proposta neste trabalho, os autores analisaram as votações dos candidatos a deputado federal em São Paulo, no período entre 1944 e 2010, testando outras dimensões de agregação geográfica, além do município, e usando unidades espaciais

mais amplas, baseadas nas micro e mesorregiões geográficas definidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE); e nas Regiões de Governo (RGs) e Regiões Administrativas (RAs) estabelecidas pelo estado de São Paulo. As comparações não demonstraram diferenças nos diversos escopos de agregação analisados:

O mais relevante nesses resultados é a pequena diferença entre os índices nas diversas agregações. Os índices pra RGs ajustadas é, no máximo, 55% acima do índice para municípios, e cada RG tem bem mais de 1,5 municípios. (AVELINO; BIDERMAN; SILVA, 2011, p. 337)

Preliminarmente, nota-se que o índice G propõe uma métrica mais sofisticada de aferição da distribuição espacial dos votos recebidos pelos candidatos, constituindo-se, aparentemente, numa evolução metodológica em relação às proposições anteriores, talvez por isso Gonçalves (2016, p.72) afirme que ele “[...] é o índice mais utilizado pelas pesquisas recentes”. Percebe-se que os autores tentaram, com a formulação desse novo cálculo, dirimir as deficiências dos demais índices em uso na Ciência Política brasileira. Cabe avaliar empiricamente se esse índice realmente é um tradutor eficiente das diversas distribuições espaciais das votações e equaciona as deficiências dos índices citados anteriormente.

Por fim, deve-se citar o recente estudo de Gonçalves (2016): “Onde agrego os votos? Contribuições à geografia eleitoral aplicada a problemas político-eleitorais brasileiros”, que buscou entender os métodos de análise de dados eleitorais com estruturação espacial, fazendo um apanhado geral da literatura relacionada à geografia das votações no Brasil. Nele, o autor percebeu que ainda são poucos os trabalhos que se valem dos métodos geoespaciais de análise e estatística espacial, ainda que

[...] a complexidade das técnicas analíticas espaciais venha sendo diminuída pelo aumento de dados georeferenciados e softwares estatísticos de simples acesso, os dados mostram que a produção no campo é reduzida. (GONÇALVES, 2016, p.64)

Como vimos até aqui na revisão de literatura sobre a geografia eleitoral da Ciência Política, as formas de mensuração se limitam a dados eleitorais, sem levar em conta a distribuição espacial dos votos. Mesmo quando se chega a resultados de votações concentradas em alguns municípios, não é possível afirmar categoricamente que essas cidades estão próximas ou contíguas, infere-se a distribuição espacial concentrada, sem se demonstrar empiricamente. A exceção a essa regra é o trabalho de Ames, que usa o índice I de Moran, apesar de ele também não explicitar claramente sua operacionalização e como chegou aos seus resultados.

Gonçalves (2016) discorre sobre as possibilidades e limitações das ferramentas de

estatística espacial no emprego dos estudos de geografia eleitoral. Uma questão que ele coloca é a da agregação de dados espaciais da área que se pretende testar e que pode variar dependendo da escala escolhida: qual será a escala geográfica escolhida para a seleção dos dados a serem analisados? Unidades de agregação muito amplas podem esconder informações importantes, assim, sempre que possível, é preferível o uso de agregações menores, mais detalhadas e homogêneas para se evitar distorções nos resultados.

Em linha com essa preocupação metodológica, que na geografia eleitoral se reflete no problema da heterogeneidade dos municípios como unidades de pesquisa, o estudo de Gonçalves (2016) utilizou o índice G para comparar diferentes amplitudes de unidades geográficas, sendo elas: Mesorregiões, Microrregiões, Zonas Eleitorais e Municípios na eleição de 2014 para a Câmara dos Deputados no estado do Paraná. A partir dessas agregações espaciais o trabalho analisou também a distribuição de votos do PT, em todo o país, no segundo turno do pleito presidencial no mesmo ano. Em ambos os casos e da mesma forma que Avelino, Biderman e Silva (2011), Gonçalves (2016) não viu diferenças relevantes nos diversos níveis de agregação testados:

Os testes do índice G, que quantifica a distribuição dos votos pelas unidades, não mostraram diferenças significativas entre as médias dessas agregações. A agregação dos dados influenciou pouco os resultados. (GONÇALVES, 2016, p.93)

## 2.4 Índice I de Moran

Um conceito trazido da geografia e que é base para as técnicas de dependência espacial é a ideia de que "Tudo está relacionado com todo o resto, mas coisas próximas estão mais relacionadas do que coisas distantes.", também conhecido como Lei de Tobler (TOBLER, 1970, tradução nossa). Quando, a partir de uma variável, se mede fenômenos e se percebe que eles estão mais relacionados quanto maior for a proximidade, pode-se dizer que eles são concentrados espacialmente. Quando a relação é inversa, ou seja, quanto mais próximos menos relacionados, pode-se concluir que a distribuição espacial é dispersa. Quando não é possível estabelecer uma relação entre o fenômeno mensurado e a proximidade, então temos um padrão aleatório da espacialidade. A Figura 1 exemplifica essas distribuições espaciais.

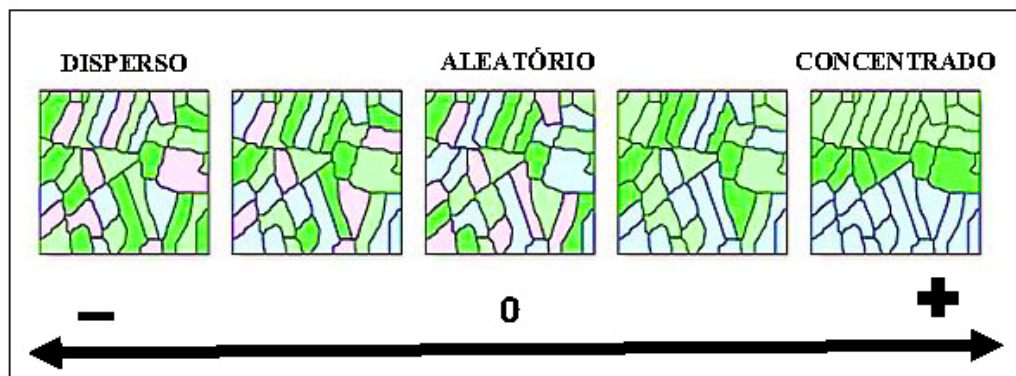
A partir desse *continuum* concentrado-aleatório-disperso é possível estabelecer se há autocorrelação espacial de um atributo e em que grau isso ocorre. Para realizar essa medição foram desenvolvidos testes estatísticos, dentre eles se destaca o I de Moran (1950), usado por Ames (2003) em seu trabalho, também chamado de Moran Global. Seu coeficiente varia de -1

a +1, sendo o +1 a máxima concentração; o -1, a máxima dispersão; e o zero, a distribuição aleatória de uma variável. Assim, estatisticamente, a hipótese nula ( $H_0$ ) seria a aleatoriedade, e  $H_1$  é a hipótese que corresponde à situação onde há dependência espacial nos dados analisados e quando o coeficiente do índice I se distancia do zero. Na explicação de Gonçalves:

O Índice de Moran indica a correlação entre o valor de uma unidade espacial e a média ponderada dos valores das unidades espaciais vizinhas. Quando existe autocorrelação positiva, o índice assume um valor positivo e isso significa que as áreas vizinhas entre si apresentam valores semelhantes para a variável analisada. Analogamente, um valor negativo indica a autocorrelação negativa, o que caracteriza dissimilaridade entre os valores das áreas adjacentes. (GONÇALVES, 2016, p.38)

Em vista disso, para se aferir a autocorrelação, resalta-se a importância de se qualificar os parâmetros da espacialidade do atributo a ser medido: com qual critério é possível distinguir se uma área encontra-se perto ou longe de outra? Para responder a essa questão, é preciso definir uma matriz de vizinhança. A matriz de vizinhança também é chamada de outras formas: matriz de peso espacial, de proximidade ou de ponderação. É assim definida por Gonçalves (2016, p.43): “Denominada  $W$ , a matriz corresponde ao grau de interação espacial entre as unidades analisadas. Simplificando,  $W$  define o que é ‘vizinho’ e o que não é, ou ‘quanto’ o vizinho é próximo.”

Figura 1 – Padrão espacial das variáveis



Fonte: Gonçalves, 2016, p.38

A matriz de peso espacial é uma matriz quadrada, ou seja, com número igual de linhas e colunas, que representam as áreas a serem testadas, onde as intersecções da tabela representam o cruzamento de duas localidades. Se for determinada a vizinhança, atribui-se o valor 1 (um) ao campo correspondente, quando não há vizinhança entre os elementos o valor do campo é 0 (zero). Por convenção, a diagonal central da matriz é igual a zero, pois esses



campos representam a intersecção entre as mesmas áreas de observação.

Existem diferentes critérios de vizinhança: o mais simples é o de contiguidade, referente às localidades que têm algum limite comum; o mais usual é a contiguidade de primeira ordem, que é aquela verificada entre vizinhos, já as de segunda ordem compreendem os vizinhos dos vizinhos e assim sucessivamente. Há três formas de qualificar a vizinhança de contiguidade: torre (*rook*), bispo (*bishop*) e rainha (*queen*), na explicação de Terron:

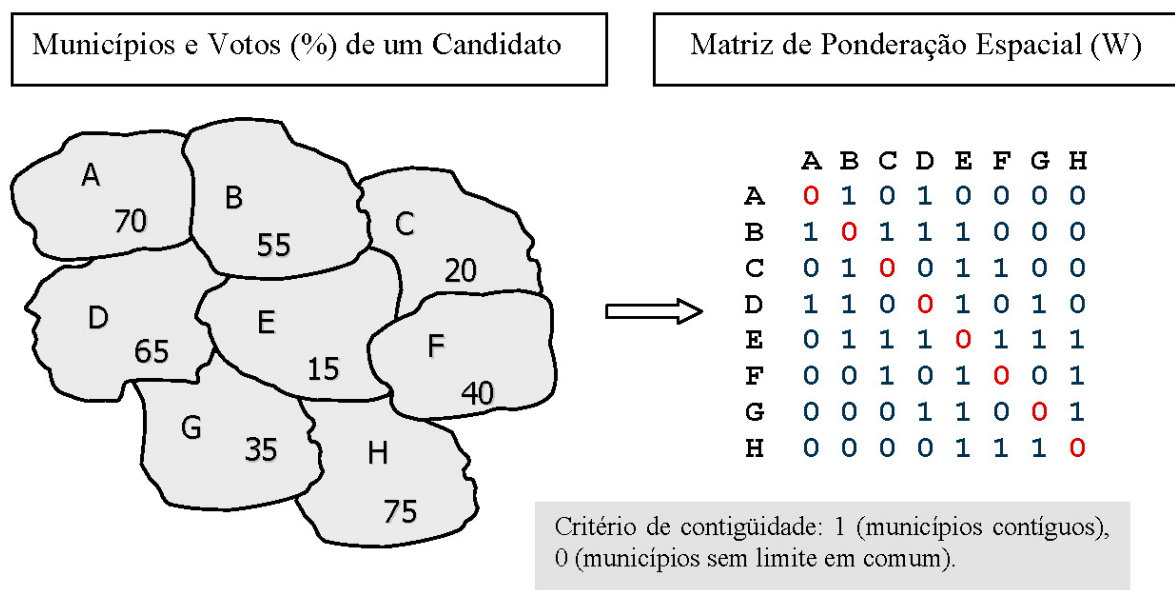
Os nomes fazem alusão a direção dos movimentos permitidos às respectivas peças no jogo de xadrez. No tipo torre, as unidades contíguas são aquelas que compartilham um lado em comum. No tipo bispo as unidades são contíguas se compartilham um vértice em comum. O tipo rainha incorpora os dois critérios, torre e bispo, sendo que qualquer objeto que compartilhe uma borda ou um vértice em comum é definido como vizinho. A contiguidade também pode ser determinada por círculos concêntricos de proximidade. (TERRON, 2009, p. 63)

Há ainda a vizinhança do tipo euclidiana que, em vez de variar os valores de forma binária, se distribuem em uma escala contínua baseada na proporção inversa à distância entre duas áreas de observação. Como assevera Terron, há ainda

[...] outros critérios concebidos para o espaço Euclidiano, baseados em distâncias, como o critério k-nearest. Nesse caso considera-se como vizinhas apenas as k unidades mais próximas, segundo a distância euclidiana entre as unidades. (TERRON, 2009, p. 64)

A seguir, temos a Figura 2, representando uma matriz de ponderação espacial de contiguidade de primeira ordem do tipo *queen*.

Figura 2 – Matriz de ponderação, contiguidade queen de primeira ordem



A partir da definição da matriz de ponderação é possível realizar o cálculo de Moran I, expresso em:

$$I = \frac{n}{W} \left( \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} \right) \quad (2.5)$$

Sendo:

$n$ , o número de observações;

$w_{ij}$ , o elemento na matriz de vizinhança para o par  $i$  e  $j$ ;

$W$ , a soma dos ponderadores da matriz, ou  $\sum_i \sum_j w_{ij}$ ;

$z_i$  e  $z_j$ , os desvios em relação à média ( $z_i - \bar{z}$ ), ( $z_j - \bar{z}$ ),  $\bar{z}$  é a média da variável.

Em linhas gerais, pode-se dizer que o índice Moran I é o resultado do produto cruzado de todas as observações de uma variável em seus polígonos, subtraídas pela média geral dessa variável no conjunto de polígonos (desvios), multiplicadas pelos desvios de sua respectiva vizinhança (de acordo com a matriz de ponderação). Por isso, se diz que:

[...] o Índice de Moran Global (I), que é a versão espacial do coeficiente de correlação linear de Pearson (Costa Neto, 2002) e que representa o coeficiente de correlação para o relacionamento entre os valores de uma variável espacial (atributo) e o valor médio desta variável. (LUZARDO; CASTAÑEDA FILHO; RUBIM, 2017, p.163)

Outra ferramenta de avaliação da dependência espacial é o Moran Local, desenvolvido por Luc Anselin (1995), também conhecido como *Local Indicators of Spatial Association* (Lisa). O Moran Local quantifica a associação espacial de cada área de um conjunto de amostras, além de identificar *clusters* e *outliers* no agrupamento de polígonos espaciais testados.

Enquanto o índice Moran Global atribui um valor de interdependência espacial entre todas as áreas testadas, o Moran Local calcula individualmente a autocorrelação das áreas em relação ao seu entorno, gerando para cada polígono um valor de índice. Com ele, é possível identificar diferentes padrões de dependência espacial das regiões de um mapa, viabilizando uma compreensão mais minuciosa da espacialidade estudada. É o que afirma Luzardo, Castañeda Filho e Rubim:

[...] quando se lida com um grande número de áreas em uma região, é muito provável que ocorram diferentes regimes de associação espacial e que apareçam máximos locais de autocorrelação onde a dependência espacial é ainda mais pronunciada. (LUZARDO; CASTAÑEDA FILHO; RUBIM, 2017, p.166)

Em certo sentido, pode-se dizer que o Lisa é um Moran Global decomposto, ou que o Moran Global é a síntese de todos os Moran Locais de um mapa. O Moran Local é calculado a partir da subtração do valor da variável de um polígono pela média das áreas adjacentes dividido pela sua variância. Sua formalização é dada por:

$$I_i = \frac{z_i}{S_i^2} \sum_j w_{ij} z_j \quad (2.6)$$

Sendo:

$$S_i^2 = \frac{\sum_j z_j^2}{n} \quad (2.7)$$

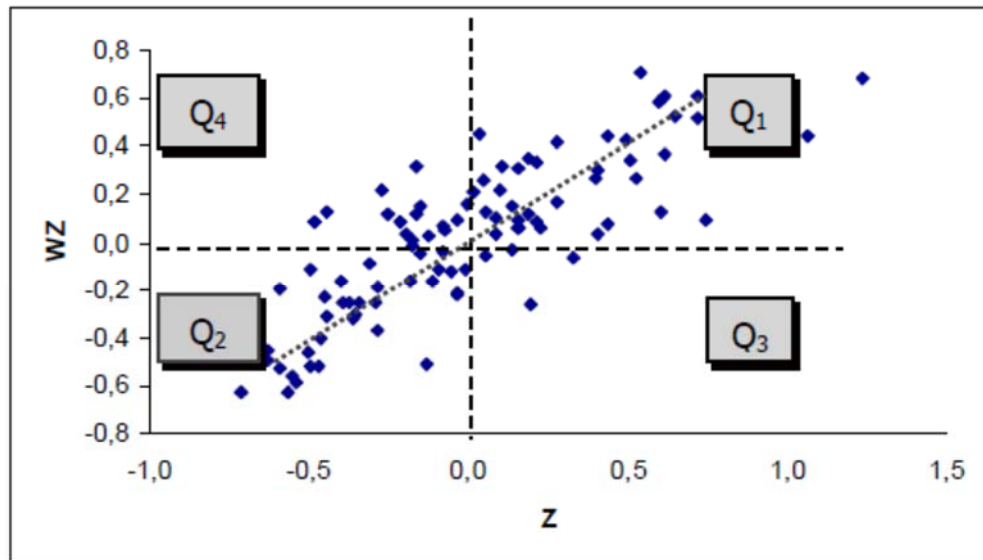
Ao medir a dependência espacial, tanto a Global como a Local, deve-se avaliar se os resultados são significativos estatisticamente. Para realizar esse diagnóstico o mais comum é o uso do método de Monte Carlo, que consiste na repetição de simulações sucessivas para calcular a probabilidade do resultado ser fruto ou não da aleatoriedade. Assim, temos um teste de pseudosignificância, que gera  $n$  permutações, distintas entre si, das variáveis associadas às áreas pesquisadas. Câmara, Monteiro, Fucks e Carvalho explicam como funciona o teste de pseudosignificância:

[...] são geradas diferentes permutações dos valores de atributos associados às regiões; cada permutação produz um novo arranjo espacial, onde os valores estão redistribuídos entre as áreas. Como apenas um dos arranjos corresponde à situação observada, pode-se construir uma distribuição empírica de  $I$ [...] Se o valor do índice  $I$  medido originalmente corresponder a um “extremo” da distribuição simulada, então trata-se de valor com significância estatística. (CÂMARA; MONTEIRO; FUCKS e CARVALHO, 2004. p.14)

O índice de  $I$  de Moran pode ser visualizado graficamente por meio do seu Diagrama de Dispersão (*Moran scatter plot*), também conhecido como Diagrama de Espalhamento. Concebido por Anselin (1995), o Diagrama de Espalhamento dispõe pontos em um gráfico bidimensional, tendo no eixo  $z$  (horizontal) os valores normalizados (subtraídos da média e divididos pelo desvio padrão) de uma variável; já no eixo  $wz$  (vertical) estão as médias dos vizinhos da variável normalizada. A Figura 3 exemplifica um Diagrama de Dispersão e Câmara, Monteiro, Fucks e Carvalho (2004) explicam sobre os quatro quadrantes que se formam:

Q1 (valores positivos, médias positivas) e Q2 (valores negativos, médias negativas): indicam pontos de associação espacial positiva, no sentido que uma localização possui vizinhos com valores semelhantes. Q3 (valores positivos, médias negativas) e Q4 (valores negativos, médias positivas): indicam pontos de associação espacial negativa, no sentido que uma localização possui vizinhos com valores distintos. (CÂMARA; MONTEIRO; FUCKS e CARVALHO, 2004, p. 20)

Figura 3 – Diagrama de Dispersão de Moran



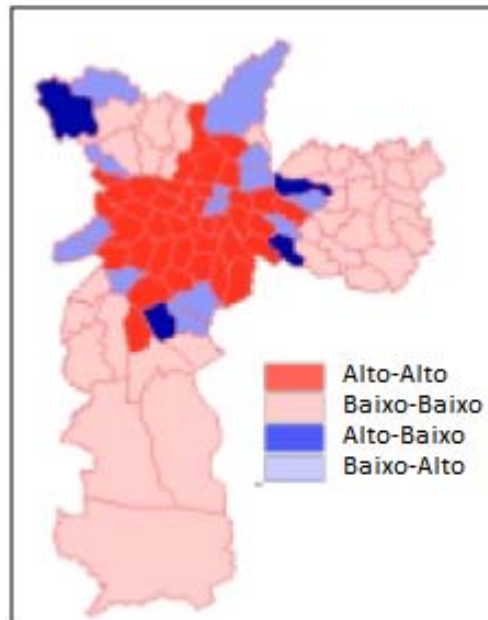
Fonte: Câmara, Monteiro, Fucks e Carvalho (2004, p.20)

Por meio de uma interpretação visual simples do Diagrama de Dispersão é possível perceber qual é a autocorrelação espacial das áreas testadas: a posição de cada ponto traduz o comportamento de um atributo em um polígono e no seu entorno. Assim, quanto maior for o número de áreas plotadas nos quadrantes Q1 e Q2, maior será a autocorrelação resultante dos dados analisados - no caso, demonstrando concentração espacial -, resultando em uma reta crescente e pronunciada. Se houver maior número de pontos nos quadrantes Q3 e Q4, teremos uma reta decrescente, o que significa uma correlação significativa, mas nesse caso negativa, indicando dispersão espacial da variável. Já quando houver uma maior concentração de dados na intersecção dos dois eixos do diagrama, a interpretação imediata decorrente dessa configuração é que há pouca ou nenhuma correlação espacial nos dados, ou seja, a disposição da variável no espaço é aleatória, nessa situação a reta será paralela ao eixo horizontal. Desse modo, vê-se que a reta que sintetiza os pontos plotados no diagrama é a representação gráfica do índice Moran. Nesse sentido, é pertinente afirmar que a reta equivale à regressão linear.

Outra representação visual da estatística de Moran é o mapa de *cluster*, técnica também desenvolvida por Anselin (1995), que se destina a demonstrar diretamente na geografia da área estudada as relações espaciais nele existentes. Para isso os dados calculados de cada polígono, que representavam pontos no diagrama de dispersão, agora são referências para o mapa de *cluster*. O mapa comunica graficamente *clusters* e *outliers* no conjunto de áreas testadas, explicitando visualmente a dependência espacial local (Lisa).

A diferenciação da cor dos polígonos no mapa se dá de acordo com o quadrante que ele pertence no diagrama. É possível distinguir quatro cores que identificam áreas onde o valor da variável foi alto e o da vizinhança também (Alto-Alto); regiões onde o valor da variável foi baixo e o do seu entorno da mesma forma (Baixo-Baixo); e zonas de transição, que seriam aquelas onde o atributo correspondente a um polígono é alto, mas ele está situado em uma vizinhança com valor de atributo baixo (Alto-Baixo), e aqueles polígonos com valores baixos em meio a uma vizinhança com variável alta (Baixo-Alto). Na Figura 4, há o exemplo dos dados referentes ao diagrama da Figura 3 representados em mapa de cluster.

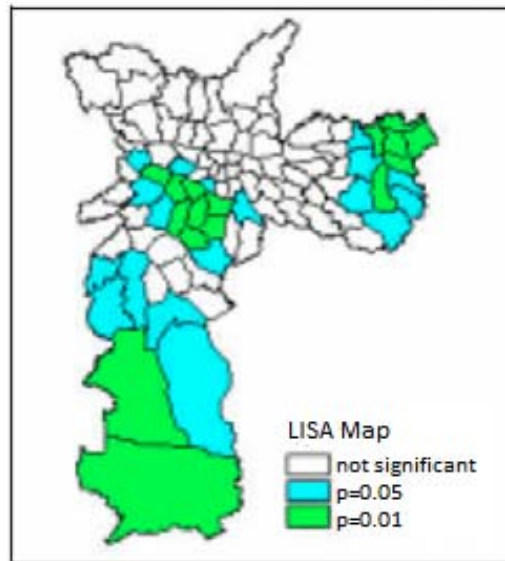
Figura 4 – Mapa de cluster



Fonte: Câmara, Monteiro, Fucks e Carvalho (2004, p.21)

No entanto, não se deve fazer a interpretação imediata desse mapa, há que se levar em conta, antes, a significância estatística dos valores associados individualmente aos polígonos. Nesse sentido, por meio do teste de pseudosignificância é possível compor também um mapa com os níveis de significância estatística calculados para cada observação, como se pode ver na Figura 5.

Figura 5 – Mapa de significância



Fonte: Câmara, Monteiro, Fucks e Carvalho (2004, p.22)

Hoje, com os softwares de estatística espacial atuais, como o GeoDA (2017), já é possível compor mapas unificados que integrem os exemplos das Figuras 4 e 5 - os mapas de *cluster* e de significância, respectivamente. Nesse caso, os mapas de dispersão não têm todas as suas áreas colorizadas de acordo com a localização no respectivo quadrante do Digrama de Dispersão, apenas os polígonos que tiverem significância estatística serão destacados no mapa.

Ao contrário dos outros indicadores, o índice Moran I, trazido para a Ciência Política brasileira por Ames (2003), tem a vantagem de levar em conta a distribuição espacial dos votos em sua apuração. Moran, quando mede a autocorrelação espacial dos votos, aborda questão fundamental na geografia eleitoral - a contiguidade espacial das áreas analisadas. Já o índice Moran Global gera um coeficiente que leva em conta a dependência espacial dos votos. No entanto, do mesmo modo que a maioria dos outros índices, sua formulação não pondera a magnitude das áreas (municípios) analisadas, causando distorção nos resultados, já que seu cálculo coloca em pé de igualdade municípios com poucas centenas de eleitores e outros com alguns milhões, sem pesar tal disparidade.

## 2.5 Síntese dos índices de correlação espacial

Com intuito de facilitar a visualização dos índices de concentração/dispersão eleitoral existentes na literatura e de estabelecer parâmetros de comparação entre eles, montamos o Quadro 2.

Quadro 2 – Índices de medição de concentração/dispersão da geografia eleitoral

Índice	Rae	Índice N	Índice G	Moran I
Autor	Douglas Rae	Laakso e Taagepera	Avelino, Biderman e Silva	Patrick Moran
Quem usa	Indjaian	Dias / Kinzo, Borin e Martins Jr. / Carvalho	Avelino, Biderman e Silva / Gonçalves	Ames / Marzagão / Nicolau e Terron
Cálculo	$1 - \sum_{i=1}^n \left(\frac{Vm}{Ve}\right)^2$	$1 / \sum_{i=1}^n \left(\frac{Vm}{Ve}\right)^2$	$\sum_{i=1}^n \left(\frac{Vm}{Ve} - \frac{VM}{V}\right)^2$	$\frac{n}{W} \left(\frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2}\right)$
Variação	De 0 (disperso) a 1 (concentrado)	De 1 (concentrado) a >100 (disperso)	De 0 (disperso) a >100 (concentrado)	De -1 (disperso) a +1 (concentrado), o 0 (aleatório)
Conceito	A concentração/dispersão eleitoral varia de acordo com o número de municípios onde a votação se distribui	A concentração/dispersão eleitoral varia de acordo com o número de municípios onde a votação se distribui	Dispersão eleitoral entendida como proporcionalidade eleitoral	A concentração/dispersão eleitoral entendida como o grau de proximidade espacial das áreas em que os votos são dados

Fonte: elaboração própria

$Vm$  = Votos do deputado no município

$Ve$  = Votos do deputado no estado

$VM$  = Todos os votos para deputado no município

$V$  = Todos os votos para deputado no estado

$n$  = o número de observações;

$w_{ij}$  = o elemento na matriz de vizinhança para o par  $i$  e  $j$ ;

$W$  = a soma dos ponderadores da matriz, ou  $\sum_i \sum_j w_{ij}$ ;

$z_i$  e  $z_j$ , os desvios em relação à média  $(z_i - \bar{z})$ ,  $(z_j - \bar{z})$ ,  $\bar{z}$  é a média da variável.

### **3 AS FRAGILIDADES DOS ÍNDICES DE MEDIÇÃO DA GEOGRAFIA ELEITORAL**

Diante do panorama exposto no capítulo anterior, convém aprofundar a análise sobre os indicadores que medem a concentração das votações no Brasil, com o objetivo de ampliar a visão sobre essas ferramentas, apontando virtudes e deficiências, quando for o caso. Em face dessa necessidade, buscou-se sintetizar neste tópico as questões que envolvem o tema.

No Brasil, o estudo da geografia eleitoral relacionada a candidatos, em eleições proporcionais para o Legislativo, se baseia na ideia central de que o candidato com votação concentrada espacialmente tem seu comportamento parlamentar fortemente influenciado pelo vetor do paroquialismo político. No caso contrário, quando a votação for dispersa, seu mandato terá preocupações mais universalistas. Para descobrir qual das duas características está mais presente na nossa política é necessário conhecer o perfil das votações dos candidatos.

Quando se analisa individualmente uma votação, não parece ser difícil classificá-la como concentrada ou dispersa: mesmo de forma subjetiva, é possível identificar se o candidato focalizou seus esforços em uma determinada área ou se conseguiu votos em todo estado. Parece evidente que o candidato conhece, mesmo que não tenha dados mais detalhados, quem é seu eleitor ou onde possui força política. Porém, quando se quer atribuir um perfil eleitoral a vários candidatos/eleitos, é preciso uma avaliação mais objetiva para fazer tal classificação, assim, é possível conhecer um incentivo de atuação política importante dos eleitos para o Parlamento.

Um primeiro problema detectado na tradicional forma de avaliar a geografia dos votos é tomar o município como unidade básica dos dados, pois há grande variação na dimensão do eleitorado dos municípios. Como colocar em pé de igualdade áreas com mil votos e outras com um milhão sem ponderar essa disparidade?

Apesar disso, é importante lembrar que o município é usado como base na maioria dos estudos de geografia eleitoral. Os índices de Rae, N, G e mesmo o I de Moran, que são os indicadores mais utilizados para medir a distribuição eleitoral no Brasil, se baseiam na agregação dos votos dos municípios para realizar seus cálculos. Segundo Gonçalves (2016, p. 64) “O município como unidade espacial de análise é o mais utilizado, corresponde a 60% do total das escalas utilizadas.” Isso se deve à facilidade de acesso aos dados eleitorais por município e à compreensão de que a política acontece em nível local, em contato com as lideranças regionais como vereadores e prefeitos. Partindo desse ponto de vista, Carvalho



argumenta a favor da adoção do município, no trecho:

[...] a unidade eleitoral de fato no Brasil tem por sede o Município. Ora quando defendemos que o Município tem este estatuto, estamos a sugerir que é ali onde de fato se trava a competição por votos entre os postulantes a mandatos eletivos na Câmara dos Deputados. (CARVALHO, 2003. p.68)

Entretanto, a despeito do embate político ser travado em nível municipal, há de se ter em conta que poucos são os municípios que, em virtude da sua variabilidade de tamanho, contêm todos os votos necessários para um deputado se eleger: mesmo os políticos com perfil paroquial possuem estratégias que miram mais de um município para compor sua base eleitoral. Como bem observaram Avelino, Biderman e Silva:

Em termos empíricos, os dados do último Censo do IBGE indicam que cerca de 70% dos municípios brasileiros tem até 20 mil habitantes, um número que representaria em média de 14 mil eleitores. Dada essa distribuição dos eleitores são poucos os municípios com eleitores suficientes para garantir a vitória em uma eleição estadual. Desta forma, não é surpreendente que os candidatos visem regiões, agrupamento de municípios, ao invés de municípios individuais. (AVELINO; BIDERMAN; SILVA, 2016. p. 1108)

A grande diferença na composição dos estados é outro fator que não favorece a utilização do município como unidade de teste, sem que se considere seu peso relativo no total do eleitorado. Há estados com centenas de municípios pequenos e outros com duas ou três dezenas deles, geralmente com uma capital com grande peso sobre o total dos votos; em outras palavras, estados grandes são analisados com o mesmo parâmetro de estados pequenos, sem que se leve em conta o perfil eleitoral totalmente diverso.

Essa variabilidade de demografias dos estados - combinada com a utilização dos municípios, que também são heterogêneos em sua conformação, como unidades de agregação de votos - ocasiona um segundo problema: os coeficientes resultantes desses cálculos têm dificuldade para comparar as votações dos diversos estados, pois resultam em valores condizentes com cada uma dessas demografias eleitorais díspares, sem equalizar tais diferenças. Por exemplo: quando um candidato no Acre e outro em Minas Gerais obtêm o mesmo coeficiente de um determinado indicador, esses valores não podem ser equiparados, pois não traduzem concentrações eleitorais equivalentes, já que os resultados são determinados por duas demografias eleitorais muito distintas. Nos valores calculados para o Acre, que possui uma capital com quase metade da votação do estado e apenas vinte e dois municípios, a votação de Rio Branco terá tratamento similar à pequena cidade de Santa Rosa do Purus, que tem pouco impacto eleitoral, menos de 1% do eleitorado estadual. Ambos os municípios terão o mesmo peso no cálculo dos indicadores.

Já em Minas Gerais, que distribui seu eleitorado por mais de oitocentos municípios e a capital tem apenas 11% da votação do estado, os indicadores tendem a captar melhor as características das votações, pois os cálculos são feitos tendo por base dados mais detalhados, onde um único município não pesa tanto para o cômputo final do cálculo. Apesar dessa inconsistência, é comum a utilização dos índices para comparar dados de estados com configurações eleitorais muito diferentes.

O número de vagas em disputa em cada unidade da federação é outra consequência direta da diversidade de demografias estaduais, que influi nas interpretações dos indicadores e dificulta uma avaliação comparativa dos resultados quando confrontados em mais de um estado. A concentração/dispersão das votações está atrelada ao número de eleitos em um estado, pois quanto menos cadeiras estão em jogo, maior será a necessidade do político dispersar seus votos por um percentual maior do eleitorado estadual. Assim como, quanto maior o número de vagas, menor a necessidade de dispersão eleitoral do candidato.

Enquanto o Acre elege oito, Minas Gerais elege cinquenta e três deputados para a Câmara dos Deputados; nesses dois casos, se o eleitorado for dividido pelo número de vagas, o deputado acreano representaria 12,5% dos votos do estado. Por sua vez, em Minas, esse percentual é de 1,88%, ou seja, duas necessidades eleitorais bem desiguais. Portanto, a disparidade de vagas em disputa também deve ser levada em conta quando se analisa comparativamente dados eleitorais de mais de um estado, isso se há a intenção de usar uma métrica de avaliação da geografia eleitoral que possa ser também um meio de comparação das votações nos diferentes estados.

### **3.1 Comparação entre os estados do Rio de Janeiro e Minas Gerais – estudo exploratório**

Para exemplificar a diferença de configuração eleitoral entre dois estados, mesmo quando ambos tenham dimensão eleitoral equivalente, o presente trabalho desenvolveu estudo exploratório em que foi estabelecida uma comparação dos dados eleitorais dos estados do Rio de Janeiro e de Minas Gerais, elencando o número de municípios, de zonas eleitorais, a média de votos e o desvio padrão dessas duas agregações espaciais, conforme Quadro 3. Os dois estados têm eleitorados com ordens de grandeza próxima, Minas Gerais com cerca de nove milhões de eleitores e o Rio de Janeiro com cerca de sete milhões, mas esses votantes se distribuem por uma quantidade bem diferente de municípios, 853 e 92, respectivamente, mais de 1.000% de diferença entre as duas áreas. Já quando se analisa o número de zonas eleitorais, observa-se que a diferença não é tão significativa, Minas com 351 zonas eleitorais e o Rio

com 249, diferença de 40%.

Também foi calculado o número efetivo de municípios (NEM) e o número efetivo de zonas eleitorais (NEZE)<sup>5</sup>. Ambos os indicadores são variações do índice N que avaliam a distribuição espacial média dos eleitores de um estado, tomando por base de agregação os municípios e as zonas eleitorais, ao invés de testarem a concentração dos votos de um candidato. Assim, quanto menor for o valor calculado para um estado, maior será a concentração espacial dos eleitores nesse estado, e de modo inverso, quanto maior for o índice apurado, menos concentrado será o eleitorado do estado.

Os resultados encontrados mostram maior proximidade dos índices quando a agregação escolhida se dá pela zona eleitoral (243,72 e 206,99), uma diferença de 17% entre Minas e Rio, sendo bem mais discrepantes quando a agregação municipal é considerada (52,61 e 6,49), com disparidade de 710% entre os dois estados. Do mesmo modo, quando se analisam os valores da média da magnitude dos eleitorados por município e zona eleitoral no Quadro 3, observa-se que os resultados das zonas eleitorais apresentam menor variação do que os dos municípios na comparação entre os dois estados. O Rio apresenta média de 76.946 eleitores por município, enquanto Minas Gerais tem 10.890, ou seja, valores bem distantes nos dois estados. Quando a análise é por zona eleitoral, os resultados são bem mais próximos que os dos municípios: o Rio tem média de 28.429, enquanto Minas apresenta 26.466 de média.

Quadro 3 – Comparação dos dados eleitorais de Minas Gerais e Rio de Janeiro

	<b>RIO DE JANEIRO</b>	<b>MINAS GERAIS</b>
<b>Votos para Deputado Federal em 2014</b>	7.079.046	9.289.851
<b>Municípios</b>	92	853
<b>Zonas Eleitorais</b>	249	351
<b>Índice NEM</b>	6,49	52,61
<b>Índice NEZE</b>	206,99	243,72
<b>Média de Votos (Mun.)</b>	76.946	10.890
<b>Coefficiente de Variação (Mun.)</b>	364%	390%
<b>Média de Votos (ZE)</b>	28.429	26.466
<b>Coefficiente de Variação (ZE)</b>	45%	66%

Fonte: elaboração própria

<sup>5</sup> Tomando por base os votos nominais para deputado federal em cada estado, apurou-se uma espécie de número efetivo de municípios (NEM) e número efetivo de zonas eleitorais (NEZE), medindo a concentração/dispersão média dos votos nestas duas agregações dos dois estados, por meio do cálculo  $N = 1/\sum_{i=1}^n (VA/VE)^2$ , sendo VA, todos os votos para deputado na agregação espacial calculada (município e zona eleitoral); e VE, todos os votos para deputado no estado.

Por fim, ao comparar os coeficientes de variação dos eleitorados dos dois estados verifica-se que os percentuais são próximos, tanto os valores calculados por município (364% e 390%), como os percentuais aferidos por zona eleitoral (45% e 66%). No entanto, o que ressalta nesses resultados é a grande variância dos municípios quando se compara com a apuração feita por zona eleitoral, na qual os coeficientes são bem menores, indicando haver menor variabilidade nessa agregação espacial.

De forma geral, esses testes preliminares com o índice N – aqui usado para medir a concentração média de dois estados cujas magnitudes são semelhantes, mas com distribuições demográficas diferentes – indicam que, no caso das zonas eleitorais, as diferenças do índice são minimizadas para os dois estados. Isso é explicado pelo fato dessa dimensão ter menor variação na distribuição dos seus eleitorados quando comparada a dos municípios. Nesse sentido, a zona eleitoral se destaca como uma alternativa interessante de unidade de agregação espacial a ser adotada.

Além disso, pode-se dizer que os índices N e Rae, que são os mais tradicionais na geografia eleitoral, não ponderam a variabilidade dos municípios, que são suas unidades básicas de avaliação. O cálculo de ambos os indicadores é fruto da soma dos quadrados da divisão do número de votos do candidato em cada um dos municípios pelo seu total de votos no estado, porém, não leva em conta o tamanho relativo do eleitorado dessas áreas no total do estado. Então, talvez adotar a zona eleitoral como unidade de agregação espacial seja um caminho alternativo para os dois índices citados a fim de dirimir essa deficiência.

### **3.2 Comparação entre agregações espaciais**

Seguindo essa perspectiva, Gonçalves (2016), em recente trabalho, comparou quatro unidades de agregação espacial dos votos recebidos pelos deputados federais no Paraná em 2014. Para isso, ele utilizou o índice G para medir as votações dos trinta deputados eleitos. A partir da estatística descritiva dos dados consolidados do indicador, o autor afirma que não existe diferença significativa entre as escalas testadas, ou seja, entre as mesorregiões, microrregiões, zonais eleitorais e os municípios do estado. Isso fica claro no trecho: “Da perspectiva metodológica indica que os resultados não foram influenciados pela unidade espacial de análise, ou foram pouco influenciados e a troca da unidade não mudaria a interpretação da realidade.” (GONÇALVES, 2016. p.76)

Notadamente, na visão do autor, verifica-se pouca diferença quando se compara especificamente os indicadores das duas menores agregações espaciais. Para Gonçalves

(2016, p. 78), “O resultado mais significativo é a correlação entre os índices testados em Municípios e Zonas Eleitorais.”

Tal resultado, que indica não haver diferença quando são utilizados os municípios ou as zonas eleitorais, talvez se explique por dois fatores: o primeiro é que o indicador utilizado foi o índice G, que em sua formulação contempla o peso relativo dos municípios no total de votos do estado; o outro fator é que o Paraná possui uma demografia mais estável quando observamos seus municípios. Comparando-o a outros estados, o tamanho do eleitorado das cidades não varia tanto e o peso da capital no eleitorado total do estado não é tão grande como no Rio de Janeiro, por exemplo. Por apresentarem essa configuração mais uniforme, os municípios paranaenses não estariam em desvantagem em relação às zonas eleitorais como unidades espaciais de agregação de votos, corroborando, nesse caso, a afirmação do autor de que: “[...] faz pouco sentido utilizar as zonas eleitorais pelo esforço de criação das bases geográficas e falta de dados georreferenciados à essa unidade, já que os resultados são equivalentes.” (GONÇALVES, 2016. p.95).

Assim, deduz-se que o resultado desse trabalho, que não encontrou diferença entre o município e a zona eleitoral para testar a espacialidade, é explicado por ter empregado um indicador que já leva em conta o peso relativo dos municípios. Por isso a zona eleitoral, que é uma dimensão mais estável, não se mostrou mais eficiente na comparação. Possivelmente, o teste faria mais sentido se, ao contrário de como foi realizado, tivesse utilizado indicadores que não contemplam o peso relativo dos municípios, como o índice N e o de Rae.

No artigo em que propõem o índice G, assim como Gonçalves (2016), Avelino, Biderman e Silva (2011) testaram outras agregações espaciais, além do município, e chegaram a conclusão semelhante a do autor paranaense, verificando pouca diferença nas agregações testadas. Os autores testaram os resultados do índice G tendo como base municípios, micro e mesorregiões, regiões de governo e regiões administrativas. Nessa comparação, o município se constitui na menor unidade, sendo as demais agregações compostas por ela.

Isso explica o fato dos resultados não espelharem diferenças significativas, pois as agregações mais abrangentes, apesar de se constituírem em unidades mais homogêneas que os municípios, representam apenas uma síntese dos dados contidos neles. Em tese, uma comparação entre município e zona eleitoral poderá produzir melhores resultados, pois a zona eleitoral, além de ser uma unidade espacial mais estável, também possui um universo de amostras mais numeroso daqueles apresentados por outras agregações, como nos casos das micro e das mesorregiões, e próximo àquele apresentado pelo município.

Baseado no fato de o Paraná ter uma demografia mais homogênea, como se vê no estudo de Gonçalves (2016), e na comparação que fizemos entre os estados do Rio e de Minas, pode-se aventar um segundo problema quando se pretende mensurar a geografia eleitoral: os estados têm diferentes configurações demográficas. Isso possivelmente implica a adoção de um modo de análise das votações que leve em conta o perfil demográfico diferenciado de cada estado ou que se empregue uma base de agregação eleitoral que equalize as discrepâncias existentes entre as diversas demografias estaduais.

À primeira vista, nota-se que existem estados com grande número de eleitores e uma capital com grande peso no total dos votos, como no caso do Rio Janeiro. De outro lado, vê-se o caso de Minas Gerais e Paraná, que possuem eleitorados dispersos por um grande número de municípios, havendo em Minas 852 e no Paraná, 399 municípios. Assim, destacam-se dois tipos diferentes de perfis demográficos e cabe investigar se, além destes, existem outros tipos de perfis na geografia eleitoral brasileira, a fim de descobrir qual a metodologia de medição mais adequada a ser adotada.

### **3.3 Vantagens e desvantagens do índice G**

Em vista do estudo de Gonçalves (2016), pode-se dizer que o índice G, por ponderar o peso dos municípios em seu cálculo, a princípio parece ser uma metodologia mais adequada para medir a concentração eleitoral, pois o indicador estima com maior precisão a distribuição eleitoral das votações. Além disso, cabe uma análise sobre o método de avaliação da geografia eleitoral, que combina recorte (dominância) e recorte horizontal (concentração), e que foi proposto por Ames (2003), sendo seguido por vários outros autores da Ciência Política.

Quando o cálculo de concentração eleitoral é eficiente em sua medição, talvez não haja a necessidade de um segundo vetor de análise. Isso porque diminui a relevância de saber se porventura um candidato dominou eleitoralmente um município ou não, tendo tido ou não grande percentual dos votos de uma cidade. O mais importante é saber qual é o tamanho do eleitorado em que ele recebeu sua votação, ou qual é a dimensão do eleitorado em que ele diluiu seus votos.

Por exemplo, o candidato X concentra todos os seus eleitores em um município, sendo que essa votação corresponde a 60% dos votos da cidade A, e o concorrente Y divide todos os seus eleitores em três municípios menores, em média um terço do tamanho da cidade A. Caso o candidato Y tenha recebido 20% dos votos nas três cidades, pode-se afirmar que a

concentração eleitoral de ambos os candidatos é equivalente. No entanto, o candidato X obteve grande dominância no município em que foi votado, o que não ocorreu com candidato Y. Portanto, é possível haver candidatos com a mesma concentração eleitoral, mas que tenham dominâncias bastante diferentes. Neste sentido, a simplificação da mensuração em apenas um coeficiente, em princípio, pode representar mais uma vantagem do que desvantagem desse índice em relação aos demais.

O Quadro 4 apresenta uma votação hipotética, que os autores do índice G elaboraram para representar um universo de eleitores que se distribuem por dez municípios e dez candidatos (AVELINO; BIDERMAN; SILVA, 2011, p.326). Nela, pode-se perceber a ponderação do peso dos municípios em relação ao total de votos que o índice G proporciona. Os candidatos 2 e 4 possuem a mesma votação total e a mesma distribuição de votos em três municípios nos quais possuem eleitores. Contudo, esses votos estão localizados em municípios cujos tamanhos populacionais são bem diferentes e, em razão disso, resultam em coeficientes G diferentes.

Quadro 4 – Votos por Município dos Candidatos Selecionados para Análise a partir de Simulação

Município	Candidato										Tamanho do Município
	1	2	3	4	5	6	7	8	9	10	
A	80.000	0	0	0	26.593	8.000	0	0	0	2.659	10.000.000
B	0	37.742	0	0	17.826	0	3.774	0	0	1.783	6.703.200
C	0	25.299	0	0	11.949	0	2.530	0	0	1.195	4.493.290
D	0	16.959	0	0	8.010	0	1.696	0	0	801	3.011.942
E	0	0	0	0	5.369	0	0	0	0	537	2.018.965
F	0	0	0	0	3.599	0	0	0	0	360	1.353.353
G	0	0	0	0	2.412	0	0	0	0	241	907.180
H	0	0	0	0	1.617	0	0	0	0	162	608.101
I	0	0	80.000	0	1.084	0	0	8.000	0	108	407.622
J	0	0	0	37.742	727	0	0	0	3.774	73	273.237
K	0	0	0	25.299	487	0	0	0	2.530	49	183.156
L	0	0	0	16.959	326	0	0	0	1.696	33	122.773
<b>Total</b>	<b>80.000</b>	<b>80.000</b>	<b>80.000</b>	<b>80.000</b>	<b>80.000</b>	<b>8.000</b>	<b>8.000</b>	<b>8.000</b>	<b>8.000</b>	<b>8.000</b>	<b>30.082.819</b>

Fonte: Avelino, Biderman e Silva, 2011, p.326

O candidato 2 recebeu seus votos em municípios com um total de cerca de quatorze milhões de eleitores, enquanto que o candidato 4 teve sua votação num universo de eleitores que não chegou a seiscentos mil. Portanto, é possível notar que o índice G qualifica o peso relativo das áreas testadas em face do conjunto de votos de um estado. Comparando com outros indicadores, vale citar que, no caso de se aplicar o índice N para medir a concentração/fragmentação de votos dos candidatos 2 e 4, os coeficientes calculados seriam

os mesmos, não distinguindo uma área da outra.

Porém, há que se fazer uma ressalva a esse indicador. Avelino, Biderman e Silva, quando propuseram o índice G, perceberam um problema na simulação que fizeram para compará-lo a outros índices. Os candidatos 1 e 4 obtiveram valores equivalentes, apesar dos seus votos estarem distribuídos de forma desigual, conforme pode ser ver no Quadro 4, que simula a distribuição das votações de doze candidatos por dez municípios, e no Quadro 5, que mostra os resultados dos índices de concentração de acordo com o quadro anterior. Os autores observaram essa contradição no trecho:

Os únicos que o índice G tem dificuldade de distinguir são os candidatos 1 e 4 e seus correspondentes “menores”, 6 e 9. Mas esta dificuldade é legítima. Seria mais concentrado um candidato com todos os votos em um grande município, ou um candidato com votos divididos em três pequenos municípios? Não há como afirmar, a priori, qual seria o resultado esperado. (AVELINO; BIDERMAN; SILVA, 2011, p. 328)

Quadro 5 – Índices de Concentração para os Candidatos Selecionados da Simulação

Candidato	Índice			
	Ames (%)	Samuels (%)	R & P (%)	G (%)
1	0,80	0,80	100,00	53,58
2	0,56	0,56	47,18	22,10
3	19,63	19,63	100,00	117,36
4	13,81	13,81	47,18	55,40
5	0,27	0,27	33,24	0,00
6	0,08	0,08	100,00	53,58
7	0,06	0,06	47,18	22,10
8	1,96	1,96	100,00	117,36
9	1,38	1,38	47,18	55,40
10	0,03	0,03	33,24	0,00

Fonte: Avelino, Biderman e Silva, 2011, p.327

A simulação realizada retorna os valores 53,58 para o índice G do candidato 1 e 55,40 para o candidato 4. Entretanto o primeiro concentrou toda sua votação no município A, que tem dez milhões de votantes, enquanto o candidato 4 obteve o mesmo total de oitenta mil votos, nos municípios J, K e L, que somados têm 579.166 eleitores. Ou seja, o candidato 4 recebeu a mesma votação que o candidato 1, contudo, isso aconteceu em um universo bem menor de eleitores, quase vinte vezes menor que o do outro candidato. Tais números respondem a dúvida dos autores de qual candidato teria votação mais concentrada: aquele “com todos os votos em um grande município, ou um candidato com votos divididos em três pequenos municípios?” (AVELINO; BIDERMAN; SILVA, 2011, p. 328). A resposta a esse



questionamento é aquele que recebeu sua votação em um colégio de eleitores mais restrito e, no exemplo apresentado, seria o segundo candidato. Os resultados apresentados pelo indicador apontam uma divergência entre as votações dos dois candidatos, já que, apesar dos valores de ambos serem tão próximos, suas votações se concentraram em eleitorados com tamanhos muito diferentes.

Assim, verifica-se uma contradição nos resultados da simulação com o índice G, evidenciando uma possível fragilidade desse método de medição. O indicador consiste na soma dos quadrados das proporções dos votos recebidas por um candidato em cada município ( $V_m$ ) pela votação total que ele recebeu no estado ( $V_e$ ), subtraídos pela proporção de votos totais destes municípios ( $VM$ ) no eleitorado total do estado ( $V$ ), expresso pela Equação 1.4. Nessa equação, a divisão de ( $VM$ ) por ( $V$ ) representa o valor esperado de um município, ou seja, o percentual de votos de um município dos votos totais do estado. Quanto menor for a diferença entre o percentual de votos recebidos por um candidato em relação aos seus votos totais no estado e o valor esperado, mais proporcional será sua votação em relação à distribuição eleitoral do estado.

Ao observar mais detidamente seu cálculo, vê-se que o indicador, apesar de considerar o peso relativo dos municípios no estado, apresenta dificuldade de mensurar quando a votação obtida em uma área se distancia do seu valor esperado. Isso porque, quando se eleva ao quadrado a subtração da proporção de votos dados no candidato pela votação esperada para o município, o indicador uniformiza as diferenças em relação à média. Sendo positivos ou negativos, os desvios acabam, em ambos os casos, resultando em incremento positivo no cômputo do índice. Os autores citam essa característica no trecho a seguir: “O fator quadrático permite que desvios abaixo ou acima da proporção do eleitorado sejam contabilizados da mesma forma.” (AVELINO; BIDERMAN; SILVA, 2011, p.325).

Em consequência disso, o indicador sofre uma distorção, pois quanto mais distante a votação for da média esperada para um município, por causa da função quadrática, maior será o aumento no índice em favor de um coeficiente mais concentrado, mesmo que o desvio aferido seja negativo para a localidade testada. Nesse caso, o indicador poderá gerar distorções indesejadas, como sobrepesar a importância da diferença em relação à votação esperada, em detrimento da ponderação do peso municipal para o total de votos do estado.

Isso acontece porque os autores, quando compuseram o índice, fundamentaram a formulação do cálculo na ideia de que a proporcionalidade se confunde com dispersão. Observa-se esse raciocínio dos autores nos trechos a seguir:

Uma maneira intuitiva de entender este índice é observando que se um deputado tivesse os seus votos distribuídos exatamente de acordo com a distribuição dos eleitores no estado seu índice seria zero.

Ou seja, se um candidato dividisse todos os seus recursos (entendidos de maneira ampla) de forma proporcional ao eleitorado em cada região, seus votos, em média, estariam distribuídos de maneira proporcional ao eleitorado. Se, por outro lado, o candidato opta por concentrar seus recursos em apenas alguns municípios, teria uma participação mais do que proporcional nesses municípios, aumentando o valor do índice. (AVELINO; BIDERMAN; SILVA, 2011, p.324)

No entanto, apesar de serem conceitos próximos, proporcionalidade e dispersão não são equivalentes. É possível haver uma distribuição de votos dispersa, mas que não seja considerada proporcional, e o inverso também ocorre. Por exemplo: um político que receba dez votos em cada um dos noventa e dois municípios do Rio de Janeiro terá sua votação dispersa por todo estado fluminense. Entretanto, seu eleitorado não será proporcional, uma vez que as dezenas de municípios do estado possuem uma grande variação no número de eleitores. Se tivessem em cada um deles a mesma população habilitada para o pleito, aí sim, a distribuição eleitoral desse político seria proporcional. O caso contrário pode ser exemplificado pelo candidato 4 da simulação feita no Quadro 5: sua votação se distribui proporcionalmente pelos municípios J, K e L; seus votos correspondem a 13,81% do total de votos dessas áreas, porém, não se pode dizer que os votos estão dispersos pelo estado, pois a votação está concentrada numa fração dos eleitores do estado.

Para ilustrar possíveis distorções advindas dessa característica do indicador, o Quadro 6 apresenta testes baseados nos colégios eleitorais propostos na simulação feita por Avelino, Biderman e Silva (2011). No total, foram testadas onze distribuições diferentes de votações, mas todas com o mesmo total de 30.000 votos. O hipotético candidato 1 teria todo o seu eleitorado em uma área, o município D. Os demais candidatos teriam a mesma quantidade de votos do candidato 1, distribuída por duas a seis localidades cada, esta sequência também identifica a numeração dos candidatos. Os candidatos 7 a 11 repetem a dispersão eleitoral dos candidatos 2 a 6 nos mesmos municípios (D, E, F, G, H, I, J, K e L), sendo que os candidatos 2 a 6 têm o total de 30.000 votos distribuídos igualmente pelos municípios testados. Já os candidatos 7 a 11 têm seus votos distribuídos proporcionalmente nos municípios (D a L).

A linha “% do D” mostra o percentual do tamanho dos municípios de cada candidato comparado aos votos do candidato 1, que teve todos seus eleitores no município D. O candidato 2 teve seus votos nas áreas E e G, esses municípios somados representam 97,15% do candidato 1; já o candidato 3 recebeu sufrágio em três municípios, em um total de

100,75% dos votos do primeiro candidato testado, e assim sucessivamente. Os percentuais mostram equivalência entre as votações avaliadas. A primeira coluna contém a identificação do município e o número de votos esperado para ele, expressando a proporcionalidade de cada área no total de votos do estado simulado.

Quadro 6 – Votos por Município dos Candidatos Selecionados para Simulação

Município / Votos esperados	Em um	Divididos igualmente nas áreas					Proporcionais ao eleitorado					Tamanho do Município
	1	2	3	4	5	6	7	8	9	10	11	
A 9.972												10.000.000
B 6.685												6.703.200
C 4.481												4.493.290
D 3.004	30.000											3.011.942
E 2.013		15.000	10.000				20.699	19.959				2.018.965
F 1.350				7.500	6.000	5.000			13.572	13.651	13.771	1.353.353
G 905		15.000		7.500	6.000		9.301		9.098	9.151		907.180
H 606			10.000	7.500		5.000		6.012	6.098		6.188	608.101
I 406			10.000		6.000	5.000		4.030		4.112	4.148	407.622
J 272						5.000					2.780	273.237
K 183					6.000	5.000				1.848	1.864	183.156
L 122				7.500	6.000	5.000			1.231	1.238	1.249	122.773
<b>Total</b>	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	<b>30.082.819</b>
<b>Índice G</b>	<b>100,04</b>	<b>60,34</b>	<b>46,67</b>	<b>40,09</b>	<b>36,11</b>	<b>33,46</b>	<b>66,15</b>	<b>60,04</b>	<b>47,28</b>	<b>46,09</b>	<b>43,10</b>	
<b>% do D</b>	-	97,15	100,76	99,32	98,74	97,89	97,15	100,76	99,32	98,74	97,89	

Fonte: elaboração própria

Na simulação do Quadro 6 o candidato 1 concentra todos os seus votos no município D, que possui um colégio com cerca de três milhões de votantes; seu índice G é de 100,04, a mais alta concentração entre todas as votações analisadas. Ao observar os candidatos, verifica-se que todos têm o mesmo número de votos, que se distribuem por áreas que possuem no total o mesmo universo de eleitores do candidato 1, porém seus índices G retornam resultados bem diferentes, que variam de 43,10 à 60,34, uma faixa de valores que aponta uma fragmentação maior de votos segundo o indicador. No entanto, não parece coerente que o candidato 3, por exemplo, detenha votação mais dispersa que o candidato 1, já que seus parâmetros globais são os mesmos. Esses resultados contraditórios suscitam a dúvida: por que trinta mil eleitores de um candidato dispersos em um universo de três milhões de votos podem resultar em um índice de concentração diferente de outro candidato com as mesmas características?

O indicador tende progressivamente à fragmentação quando aumenta o número de municípios em que os votos são distribuídos, como acontece com os coeficientes dos candidatos 2 a 6 e de 7 a 11, resultando em um valor coerente, pois aumenta a dispersão eleitoral dessas votações. No entanto, votações com um número grande de eleitores agrupados nos municípios menores, como o caso do candidato 6, que tem metade de seus votos nas três menores áreas, resultam em um coeficiente mais baixo de concentração que os demais, o que denota contradição no cálculo do indicador. A explicação para essas incongruências é a formulação do índice, pois adiciona concentração ao seu coeficiente quando os valores se desviam dos votos esperados para um município, baseados na premissa que confunde proporcionalidade com dispersão. É por isso que, apesar das onze votações simuladas estarem distribuídas no mesmo universo de eleitores, os índices calculados para esses candidatos são diferentes.

### **3.4 Síntese do problema dos índices de concentração eleitoral**

A falta de ponderação do peso relativo dos municípios no eleitorado total dos estados emerge como um dos problemas na forma como se mede a geografia eleitoral. Isso pode gerar distorções nos resultados desses indicadores, que pretendem refletir a concentração/dispersão dos votos. Há dois caminhos a serem tomados para contornar esse problema: o primeiro caminho é utilizar outro tipo de agregação espacial, para isso a zona eleitoral parece se constituir em uma alternativa viável, pois são áreas mais estáveis que os municípios, ainda que o estudo de Gonçalves (2016) tenha chegado a um resultado em sentido oposto. Pode-se partir também para uma solução alternativa, que seria compor uma agregação espacial combinando municípios e zonas eleitorais de grandes municípios, tendo assim, uma base territorial mais detalhada para análise. Se necessário, há a possibilidade de se desagregar mais ainda a votação, analisando os votos por meio das seções eleitorais.

Este estudo testará essas possibilidades: tanto a zona eleitoral como base de agregação espacial, quanto uma base de agregação mista, composta de municípios e zonas eleitorais, podendo chegar à desagregação por seção de votação nos estados em que a maior área da agregação mista represente mais de 10% dos votos e/ou o total de áreas da agregação mista seja inferior a setenta. Outro caminho é adotar um índice que leve em conta o peso dos municípios na operacionalização do seu cálculo, para isso há o índice G e este estudo propõe o índice M, que será apresentado na próxima seção, como ferramenta alternativa de avaliação da concentração eleitoral.

Um segundo problema que se destaca é o da heterogeneidade de configuração demográfica dos 26 estados brasileiros e do Distrito Federal. Os estados têm composição muito diversa, dificultando a adoção de apenas uma metodologia para se aferir a distribuição eleitoral, que seja capaz de comparar as votações desses colégios de modo eficiente, uniformizando as diferenças existentes entre eles. Talvez faça sentido utilizar um método de cálculo, mas com adaptações de acordo com a configuração de distribuição do eleitorado de cada estado.

Por fim, há a questão da contiguidade espacial das votações, que é um pressuposto teórico para se determinar efetivamente se há ou não concentração eleitoral de um candidato. Só faz sentido afirmar que um eleitorado é geograficamente concentrado se for demonstrada a proximidade das áreas em que esses votos foram recebidos, no entanto, apenas o índice de Moran considera a vizinhança espacial em sua formulação, tornando-se assim ferramenta obrigatória a quem queira ponderar a contiguidade em qualquer metodologia de medição.

Quadro 7 – Síntese dos problemas dos índices de concentração eleitoral

<b>Problema</b>	<b>Motivo</b>	<b>Possíveis soluções</b>
Os índices não ponderam o peso relativo da unidade de agregação espacial utilizada para medição	Municípios são muito heterogêneos	<ul style="list-style-type: none"> <li>- Utilizar as zonas eleitorais como unidade de agregação</li> <li>- Utilizar indicador que pondere o peso relativo dos municípios</li> <li>- Utilizar áreas de agregação mista, composta por municípios e zonas eleitorais</li> </ul>
Impossibilidade de comparar índices de estados diferentes	<ul style="list-style-type: none"> <li>- Municípios são muito heterogêneos</li> <li>- A grande variabilidade demográfica dos estados condiciona os resultados</li> <li>- Diversidade no número de deputados eleitos por cada estado</li> </ul>	<ul style="list-style-type: none"> <li>- Utilizar áreas de agregação mista, composta por municípios e zonas eleitorais</li> <li>- Normalizar índice proposto, possibilitando a comparação entre os valores de diferentes estados</li> </ul>

Fonte: elaboração própria

### 3.5 A proposta de um novo indicador: o índice M

Em vista das limitações apontadas nos índices utilizados para medir nossa geografia eleitoral, abre-se espaço para uma nova abordagem que leve em conta fatores como o peso

relativo dos municípios no total de votos do estado, as várias configurações de distribuição eleitoral dos estados e a contiguidade das áreas analisadas. A ideia principal que sustenta esse novo cálculo é a seguinte: para ser eleito um candidato precisa de votos; quanto maior for a quantidade de votos, menor será o risco de não se eleger, independente de qual votação receba sua legenda.

A chance de vitória aumenta na proporção dos votos recebidos, já que estes possibilitam melhorar sua posição na lista de candidatos da coligação. Nesse sentido, é essencial o entendimento do conceito de concentração/dispersão eleitoral para descobrir se uma distribuição de votos é concentrada ou dispersa: considera-se um eleitorado concentrado ou disperso em função da quantidade de votos de um candidato pela votação total nas áreas onde esses votos foram dados – no caso, os municípios – em face ao percentual que as áreas representam no eleitorado total do estado.

A concentração/dispersão eleitoral é uma relação entre quatro fatores: i) os votos do candidato nos colégios eleitorais; ii) a magnitude desses colégios eleitorais; iii) a proporção que esses colégios representam nos votos totais do estado; iv) e o percentual de votos correspondente a um deputado na divisão de cadeiras em disputa no estado. A ideia que sustenta a proposta do novo indicador está baseada no peso dos colégios eleitorais em que a votação majoritária de um candidato é recebida. Diferente do que ocorre com o indicador proposto, o índice N considera o número de municípios em que a votação foi dada, sem levar em conta o peso eleitoral dessas áreas em relação ao eleitorado total do estado.

Essa nova proposta metodológica de medição da geografia eleitoral é composta por duas vertentes: um indicador para mensurar a concentração/dispersão dos votos recebidos por candidatos em eleições proporcionais em face do percentual de votos; e a normalização desse indicador, por meio da sua divisão pelo percentual que corresponde uma cadeira em disputa no estado.

A partir disso, para realizar o cálculo, identifica-se onde o candidato recebeu a **maioria** da sua votação (por isso chamaremos de índice M), ou seja, ordena-se de forma decrescente os municípios em que o candidato recebeu mais votos. Serão selecionados para análise aquelas cidades nas quais, somadas cumulativamente, os votos atingiram cinquenta por cento mais um dos votos totais dados ao candidato.

Tem-se, então, uma lista de municípios onde o candidato teve a maioria dos seus votos. Supõe-se, a partir desses dados, que o político concentrou seu esforço de campanha nessas localidades e que elas são sua base eleitoral. A partir da lista de municípios, calcula-se o percentual que os votos dessas localidades representam no total de votos do estado. Por fim,

esse valor é dividido pelo percentual que corresponde à fatia eleitoral de um deputado no estado. Essa última operação do cálculo busca normalizar o indicador para que ele possa comparar estados com diferentes distribuições espaciais de votos. A formalização do índice se dá pela Equação 3.8:

$$M = \frac{C}{V} \sum_i VM_i \quad (3.8)$$

Onde:

$VM_i$  = Todos os votos para deputado no município  $i$

$V$  = Todos os votos para deputado no estado

$C$  = número de cadeiras em disputa no estado

$i$  = municípios em que o candidato recebeu 50% dos seus votos

O índice  $M$  resulta em um coeficiente que variou de 0,11 a 28,63, nos cálculos realizados com os deputados eleitos para a Câmara dos Deputados, e reflete quantas vezes o percentual do eleitorado, onde o candidato recebeu a maioria da sua votação, equivale em relação à fração eleitoral correspondente a uma vaga em disputa no estado. Essa última operação, para apurar o valor do índice, busca equalizar as diferenças das demografias estaduais por meio da parcela dos votos equivalente a um distrito eleitoral, típico do sistema majoritário. Quanto maior o índice  $M$ , mais dispersa é a votação do candidato, ou seja, os votos recebidos pelo candidato estão distribuídos em um percentual maior do eleitorado; de modo oposto, quanto menor o valor de  $M$ , mais concentrado é o eleitorado do candidato, ou seja, os votos recebidos pelo candidato estão distribuídos em um percentual menor do eleitorado. Nesse sentido, uma interpretação possível é a de que quanto menor o indicador, mais próximo da lógica distrital seria a votação do candidato, e quanto maior o valor do indicador mais coerente com o sistema proporcional.

Essa nova fórmula de cálculo da distribuição dos votos pondera o peso que cada município tem no estado, o que representa uma inovação em relação aos índices mais usados na literatura, uma vez que a grande maioria dos indicadores que medem a geografia eleitoral não leva em conta essa ponderação, gerando distorções em seus resultados. Ou seja, diferentemente dos demais índices que medem a concentração/dispersão eleitoral, onde o número de áreas (municípios) é o ponto central do cálculo de ponderação eleitoral, o novo indicador visa avaliar a distribuição de votos em face do peso eleitoral dos colégios em que esses votos foram dados. Por isso, cabe conceituar de modo distinto o que o novo índice

almeja medir dos demais indicadores: o índice M mede a densidade eleitoral das votações. No escopo deste trabalho, entende-se **densidade eleitoral** como a distribuição da votação na proporção da dimensão do eleitorado dessas áreas.

Essa conceituação se distingue da aceção corrente na qual a densidade eleitoral é o peso dos votos em face da dimensão geográfica de uma determinada área. Segundo Costa (2007, p.181) “[...] a densidade eleitoral é a relação entre o total de eleitores inscritos, dividido pela extensão territorial.” Portanto, nesse sentido, a densidade eleitoral advém da ideia de densidade demográfica, possuindo outro significado daquele proposto neste trabalho.

Baseado nos cálculos apurados pelo índice M, as votações podem ser classificadas em uma escala que se divide em quatro graus de densidade eleitoral: baixa, média-baixa, média-alta e alta. Para se ajustar os valores limítrofes correspondentes a essas faixas de densidade eleitoral, este estudo se fundamentou na premissa de que o indicador representa o número de vezes em que a maioria da votação de um candidato equivale ao percentual do eleitorado necessário para se obter uma cadeira em disputa. Partindo da observação empírica dos resultados, constatou-se que o intervalo de 0 a 1,5 pode ser usado para a alta densidade; 1,5 a 4,5, para a média-alta densidade; 4,5 a 8, para a média-baixa densidade; e acima de 8, para a baixa densidade.

A partir dos dados da simulação proposta por Avelino, Biderman e Silva (2011), pode-se comparar os candidatos 1 e 4 do Quadro 4 utilizando o índice M. Os dois candidatos obtêm, respectivamente, os percentuais de 33% e 13% ao serem aferidos pelo novo indicador, (coeficiente não ponderado pelo número de vagas em disputa). Nesse teste, não resta dúvida de que o candidato 4 obteve votação mais concentrada do que o candidato 1, visto que o cálculo do seu índice resultou em valor mais próximo de zero. Apesar de, em tese, o cálculo de Avelino, Biderman e Silva (2011) contemplar o peso relativo dos municípios, a simulação apresentada pelos autores demonstrou resultado contraditório, o que não ocorreu com o cálculo do índice M.

Da mesma forma que o índice G, este indicador também unifica os vetores horizontal e vertical da geografia eleitoral em um só coeficiente, prescindindo do cálculo de dominância, visto que um indicador de concentração efetivo é suficiente para mensurar a geografia eleitoral. Outro ponto a favor do índice M é a simplicidade da sua formulação, pois seu coeficiente traduz uma metodologia de cálculo mais simplificada e de melhor visualização do fenômeno eleitoral, além de estabelecer um critério único a ser empregado nos diversos estados brasileiros. Pois, o novo indicador, por normalizar seu coeficiente dividindo-o pelo percentual correspondente a uma cadeira em disputa no estado, ou, dito de outra forma,



multiplicando-o pelas cadeiras em disputa, se caracteriza por servir tanto para medições de votações de candidatos individualmente, quanto para fazer comparações globais de estados e regiões.

O Quadro 8 apresenta um resumo das vantagens e desvantagens dos índices que medem a concentração eleitoral no Brasil.

Quadro 8 – Índices de medição da geografia eleitoral

<b>Índice</b>	<b>N/Rae</b>	<b>Índice G</b>	<b>Moran I</b>	<b>Índice M</b>
<b>Vantagens</b>	-	Pondera a dimensão dos municípios  Unifica os índices	Pondera a contiguidade espacial	Pondera a dimensão dos municípios  Unifica os índices  Cálculo e interpretação simplificados  Quando normalizado, padroniza a métrica eleitoral, viabilizando comparações entre diferentes estados
<b>Desvantagens</b>	Não pondera a dimensão dos municípios  Não pondera a contiguidade espacial  Não é possível comparar os coeficientes de estados diferentes	Não pondera a contiguidade espacial  Os desvios em relação à média esperada nos municípios superestimam a concentração  Não é possível comparar os coeficientes de estados diferentes	Não pondera a dimensão dos municípios  Não é possível comparar os coeficientes de estados diferentes	Não pondera a contiguidade espacial

Fonte: elaboração própria

## 4 METODOLOGIA

Para testar a hipótese central deste trabalho - a de que os índices mais largamente utilizados pela literatura para medir a concentração eleitoral no Brasil, em especial o índice N, não refletem a realidade das votações, pois não ponderam o peso relativo dos municípios no total de votos de um estado -, foram utilizados os dados eleitorais do pleito de 2014 referentes aos deputados federais eleitos em todos os estados. Essas informações foram extraídas do Repositório de Dados Eleitorais (BRASIL, 2018) na internet, que possui informação de todos os pleitos realizados desde 1945. Para se aferir o índice I de Moran dos candidatos foi preciso baixar os *shapefiles*<sup>6</sup> de todos os estados brasileiros e do Distrito Federal no Portal de Mapas (IBGE, 2018). Os arquivos de extensão “txt”<sup>7</sup> com as informações das votações e os *shapefiles* com os dados geoespaciais dos estados foram processados no ambiente de programação RStudio (2017), resultando em uma planilha com os índices calculados para a pesquisa.

O primeiro passo do trabalho foi preparar os dados disponibilizados. Há uma série de possibilidades de como se extrair as informações desejadas no Repositório de dados do TSE. Das eleições mais recentes é possível filtrar os dados de diversas formas: pelo boletim de urna; por meio da votação nominal por município e zona eleitoral; pela votação por seção eleitoral; dentre outras formas.

Foi feito o download dos arquivos compactados da votação nominal por município e zona eleitoral, que contêm os votos de todos os estados; posteriormente foram baixados e descompactados também os arquivos com informações das votações por seção eleitoral dos estados do Acre, Amapá, Rondônia, Roraima, Tocantins e do Distrito Federal. Em ambos os casos foi realizada uma série de procedimentos para organizar e limpar o arquivo só com as informações necessárias para o estudo. Para isso, inicialmente foram selecionadas, dos dados brutos, as variáveis necessárias para os cálculos: número da seção eleitoral; nome do município; número da zona eleitoral; nome do candidato, cargo em disputa, situação do candidato e quantidade de votos.

Os *shapefiles* necessários para calcular o I de Moran extraídos do Portal de Mapas do IBGE também necessitaram de tratamento. Os arquivos usados tiveram como referência os dados geoespaciais do ano de 2014 para facilitar o cruzamento de informações com os dados

---

<sup>6</sup> O Esri Shapefile ou simplesmente shapefile é um formato popular de arquivo contendo dados geoespaciais em forma de vetor usado por Sistemas de Informações Geográficas também conhecidos como SIG.

<sup>7</sup> Arquivo TXT é um arquivo de texto no formato ASCII, sem formatação.

das votações fornecidos pelo TSE. A principal dificuldade enfrentada foi uniformizar a nomenclatura dos municípios nas duas bases de dados para que os dados eleitorais e geoespaciais dos municípios pudessem ser processados de forma conjunta. Outra dificuldade encontrada foi equalizar a base de dados do estado do Rio Grande do Sul, isso porque a Lagoa dos Patos figura na relação de áreas do IBGE, no entanto, não consta na relação do TSE, por ser uma área que não possui seções de votação. A solução encontrada foi excluir a Lagoa dos Patos do *shapefile* do estado.

Após a preparação dos dados, os testes foram realizados em etapas. A primeira etapa se consistiu no cálculo, para todos os deputados eleitos nos vinte e seis estados brasileiros e no Distrito Federal<sup>8</sup>, de várias combinações de índices com agregações eleitorais: N por município; N por zona eleitoral; G, I de Moran e M por município (M), M por agregação mista, ou seja, por município e zona eleitoral (MZE); e M por agregação mista normalizado (MZEn), que é o MZE multiplicado pelas cadeiras em disputa no estado. A partir desses resultados foi possível compor as Tabelas 1 e 2 com as médias dos índices por estado. Essa etapa serviu para a escolha dos estados que tiveram votações comparadas, baseando-se no critério de realização dos testes em unidades da federação que representassem tipos diferentes de demografia eleitoral, em diferentes regiões do país, com diferentes médias de densidades, além de servir para uma primeira avaliação do potencial explicativo do índice M em relação aos demais.

Uma vez escolhidos os estados e de posse dos resultados dos indicadores, na segunda etapa foram comparados sempre dois candidatos com valores próximos do índice N e que nos demais indicadores os resultados se mostrassem contraditórios. Essa comparação pretende mostrar possíveis inconsistências dos indicadores por meio da análise detalhada da votação dos candidatos.

Nessa etapa, foi feita ainda uma comparação entre cinco candidatos eleitos em cada um dos cinco estados escolhidos: Minas Gerais, Acre, Rio de Janeiro, Rio Grande do Sul e Pernambuco. Os cinco deputados foram selecionados por terem coeficientes MZEn bem próximos e com a finalidade de verificar a efetividade do indicador por meio da identificação das semelhanças e possíveis discrepâncias na distribuição das votações dos candidatos. O MZEn foi usado nas análises de candidatos de estados diferentes, pois apresenta maior sensibilidade para esse tipo de comparação. No caso das comparações entre votações de candidatos do mesmo estado, o índice normalizado não foi necessário, ainda que possa ser utilizado.

---

<sup>8</sup> Por não se subdividir em municípios, o Distrito Federal foi a única unidade federativa que não se calculou os índices N, G, M e I de Moran por município.

A terceira etapa testou a agregação por seção eleitoral nos estados com menos de setenta áreas (municípios + zonas eleitorais) e/ou que têm sua maior área com mais de 10% dos votos totais do estado. Esses critérios foram utilizados, pois se observou que, quando os estados têm ao menos uma dessas características, a agregação mista (por municípios e zonas eleitorais) parece não ser suficiente para medir as votações. Comparou-se os resultados das médias dos coeficientes de M para esses estados, por agregação mista e por seção eleitoral, com o intuito de verificar se há a necessidade ou não de se detalhar mais a unidade eleitoral, por meio da adoção das seções de votação, para agregar os votos.

## 5 RESULTADOS E DISCUSSÕES SOBRE OS ÍNDICES

As tabelas de comparação apresentadas neste capítulo foram estruturadas de forma a apresentar as diversas formas de cálculo - em especial do índice M proposto neste trabalho -, com o propósito de representar a sequência dos passos percorridos nesta pesquisa na busca do modelo mais apropriado para se aferir a geografia eleitoral. Alguns dos critérios utilizados para o cálculo dos indicadores - como o percentual dos votos do candidato, que representa a maioria da sua votação para o índice M, apurado tanto pela agregação municipal, quanto pela agregação combinada, assim como o teste de significância do I de Moran - não estão representados nas tabelas, com vistas a não prejudicar a visualização das informações relevantes. Contudo, é possível identificar o tratamento desses elementos nos *scripts* de programação disponíveis no Anexo A.

### 5.1 Comparação dos índices em todos os estados

Os testes dos índices de concentração eleitoral de todos os estados revelou como a demografia diferenciada de cada unidade federativa pode impactar a medição realizada pelos indicadores. Analisando a Tabela 1, que agrega as médias dos índices nos estados, observa-se que o índice N por município segue na mesma direção do número de municípios que o estado possui, ou seja, quanto menor o número de municípios menor é o valor do índice, indicando maior concentração eleitoral. Isso se explica pela ausência de ponderação proporcional de cada município para o eleitorado total do estado, que é uma característica desse cálculo, assim o número de municípios é determinante no resultado do indicador. O coeficiente de concentração aumenta mais ainda quando existe um município com grande peso, como nos casos de capitais populosas, como o Rio de Janeiro, que possui 37,35% dos votos do estado, e Rio Branco que representa 49,24% do eleitorado acreano. Não por acaso, os dois estados, juntamente com os de Roraima, Amazonas e Amapá, são aqueles que apresentam a maior concentração média. Essas unidades da federação possuem poucos municípios e/ou uma capital com grande peso em relação à população total do estado. Corroborando essa ideia, os estados com maior número de municípios também são aqueles com os maiores e mais dispersos coeficientes de N por município, são eles: Bahia, Minas Gerais e Rio Grande do Sul, que obtiveram as médias de índice N por município 26,34; 27,13 e 32,71 respectivamente.

Tabela 1 – Quantitativo de áreas e médias dos índices por estado

Estado	Votos	áreas (M)	áreas (ZE)	áreas (MZE)	Vagas	maior MZE%	média N (M)	média N (ZE)	média G (M)	média M (M)%	média MZE%	média MZEn	maior MZEn	menor MZEn	média Moran
DF	1.362.160	1	21	21	8	6,7	NA	17,45	NA	NA	42,1	3,37	3,75	2,85	NA
GO	2.824.329	246	130	263	17	4,3	17,84	30,94	7,2	31,9	19,3	3,30	6,13	1,00	0,09
MS	1.174.221	79	54	88	8	6,5	11,66	28,76	3,6	47,1	40,2	3,23	4,38	1,12	0,00
MT	1.334.861	141	60	152	8	4,9	18,12	24,01	5,1	36,8	26,6	2,12	3,03	0,78	0,14
AL	1.283.120	102	55	106	9	9,4	18,42	22,42	6,6	37,8	30,8	2,76	3,96	1,15	0,04
BA	5.982.371	417	205	447	39	1,3	26,34	35,98	5,8	20,6	7,8	3,02	11,46	0,62	0,12
CE	4.002.492	184	123	201	22	2,9	15,77	27,06	7,7	29,1	13,7	3,00	5,84	0,84	0,13
MA	2.892.959	217	111	237	18	3	24,14	28,44	5,8	21,7	11,4	2,06	3,44	0,86	0,10
PB	1.773.112	223	77	233	12	4,5	22,27	26,09	3,7	29,8	22,1	2,66	4,37	0,97	0,14
PE	4.129.147	185	151	212	25	2,3	14,64	26,71	7,5	22,4	10,6	2,64	7,24	0,77	0,26
PI	1.587.477	224	98	232	10	6,1	25,79	34,22	4,9	37,1	32,7	3,28	4,51	1,24	0,06
RN	1.451.341	167	69	172	8	6,3	20,93	27,71	2,8	36,8	33,4	2,68	3,41	1,26	0,16
SE	974.311	75	36	78	8	7,6	17,49	20,10	2,9	42,6	35,6	2,86	3,95	1,00	0,13
AC	368.332	22	10	24	8	18,6	3,47	5,13	8,7	46,4	42,9	3,43	4,71	0,94	-0,01
AP	368.061	16	13	17	8	33,8	2,96	4,26	9,2	58,6	52,2	4,19	5,52	1,24	-0,05
AM	1.560.085	62	70	74	8	7,1	5,83	27,65	5,3	60,6	42,2	3,37	4,76	1,40	-0,01
PA	3.496.764	144	104	158	17	3,4	20,10	28,56	7,9	30,3	17	2,88	5,53	0,59	0,18
RO	740.924	52	35	64	8	4,7	9,19	13,92	14,8	33,2	17,2	1,37	3,23	0,61	0,03
RR	225.631	15	8	16	8	44,5	3,02	3,95	4	65,1	58,3	4,67	5,13	3,56	-0,13
TO	670.894	139	35	140	8	16,6	19,62	12,98	4,8	36,1	36	2,89	4,19	0,99	0,03
ES	1.665.277	78	58	85	10	5,3	13,32	14,88	15	23,1	16,6	1,67	4,18	0,44	0,24
MG	9.273.472	853	351	905	53	0,8	27,13	39,92	8,9	14,6	6,8	3,58	12,39	0,75	0,17
RJ	7.079.046	92	249	262	46	0,9	4,83	74,22	17	32,1	14,6	6,71	18,83	0,37	0,10
SP	19.072.417	645	425	792	70	0,7	12,34	74,16	14,7	25,3	11,6	8,07	28,64	0,12	0,13
PR	5.275.880	399	206	443	30	2,1	20,76	30,17	12,3	16,7	9,4	2,82	7,33	0,44	0,18
RS	5.501.353	497	173	527	31	1,4	32,71	38,64	5,3	22,7	14,7	4,54	9,92	0,87	0,20
SC	3.135.381	295	105	316	16	1,9	23,47	23,94	8	18,9	12,4	1,99	4,77	0,35	0,22

Fonte: elaboração própria

Da mesma forma que o N por município, o índice N por zona eleitoral varia de acordo com o número de zonas eleitorais do estado nos seus valores extremos: os estados que possuem menos zonas, Roraima, Acre e Amapá são aqueles com os coeficientes médios mais concentrados; e São Paulo, com o maior número de zonas eleitorais (425), apresenta a segunda mais alta dispersão dos votos. Apesar disso, no geral, o índice N por zona eleitoral demonstra outro comportamento, já que sua base de agregação espacial é diferente ao utilizar a zona eleitoral, que é mais homogênea que o município, para calcular a dispersão das votações, gerando resultados mais consistentes.

Nesse indicador, o Rio de Janeiro, que possui 249 zonas, apresenta uma fragmentação de 74,22, que é ainda maior que a de São Paulo (74,16), evidenciando como essa agregação espacial pode amenizar as distorções geradas em estados que possuem cidades com grande peso relativo sobre o total de votos, por isso seu coeficiente é tão discrepante em relação ao índice N convencional, que é 4,83. Outro fenômeno interessante ao se comparar os cálculos de N por município e zona eleitoral é o aumento das médias dos coeficientes de todos estados, com exceção do Tocantins, que teve uma diminuição de 19,62 pra 12,98. Isso também é consequência da agregação eleitoral composta por zonas eleitorais, que propicia uma base de votação mais uniforme, apesar de conter uma quantidade menor de áreas que a agregação formada por municípios.

As médias do índice G para os estados revelam uma faixa de variação muito pequena, de 2,8 a 17,0. Se considerarmos que o indicador pode variar de zero a mais de 100, essas médias revelam que a dispersão eleitoral dos estados é alta, pois os valores estão próximos do zero. Minas Gerais e o Acre obtiveram médias quase idênticas, 8,9 e 8,7, respectivamente. Tal fato chama atenção, visto que as demografias eleitorais desses estados são muito díspares. Minas possui mais de oitocentos municípios, enquanto o Acre apenas vinte e dois. Ao analisar também os estados com as médias mais baixas do índice, da mesma forma que na comparação anterior, não é possível detectar algum padrão demográfico que influencie os resultados. Por um lado é verdade que, entre as médias com os valores mais baixos do índice G, indicando maior dispersão, estão estados pequenos com capitais que representam uma fatia grande do total de votos, como o Rio Grande do Norte, Sergipe, Paraíba e Mato Grosso do Sul. Por outro lado, também é fato que entre as maiores médias do indicador estão estados com essas mesmas características, como Rondônia e Espírito Santo; e estados grandes, como Rio de Janeiro e São Paulo. Portanto, não é possível identificar padrões demográficos que condicionem os coeficientes de G para os estados.

As médias estaduais do índice I de Moran, assim como os do índice G, apresentam

pouca variação. O estado com o valor mais baixo é Roraima (-0,13), e com o maior coeficiente é Pernambuco (0,26). Nesse indicador, os valores negativos indicam dispersão e os positivos indicam concentração das votações; no entanto, o intervalo das médias dos estados encontra-se próximo de zero, o que revela aleatoriedade espacial das votações. Como os valores são médios, esse pequeno intervalo poderia esconder grande oscilação nas medições individuais das votações, mas não é o que se verifica. Dos 505 deputados calculados, apenas 77 tiveram votações fora do intervalo, setenta com I de Moran maior que 0,26 e sete abaixo de -0,13. Isso demonstra que as medições giraram realmente em torno do zero, indicando aleatoriedade.

Cabe assinalar que os valores negativos e os positivos mais próximos do zero foram de estados pequenos, que têm poucos municípios, como Acre (-0,01), Amapá (-0,05) e Rondônia (0,03). Isso sinaliza que o indicador tem mais dificuldade de detectar alguma dependência espacial quando a amostra é menor. Essa conclusão é confirmada pelos resultados do teste de pseudosignificância, que serve para aferir a validade do I de Moran. Dos nove estados medidos, que elegem oito ou nove deputados para Câmara, apenas o Mato Grosso teve mais de 25% das votações validadas pelo teste de pseudosignificância, demonstrando assim a dificuldade do indicador em medir estados com poucas áreas.

Os primeiros testes realizados com o índice M foram feitos utilizando a agregação por município. A partir do cálculo das médias do indicador em três estados - Minas Gerais, Rio e Acre -, seus respectivos coeficientes (15,09; 32,20 e 46,51) se aproximaram dos percentuais do eleitorado das suas capitais, que são as maiores cidades dos estados: Belo Horizonte com 11%; Rio de Janeiro com 37%; e Rio Branco com 49%. Esses resultados sugeriram uma ponderação: unidades de análise, que têm grande impacto em um estado, devem ser decompostas para aumentar a eficiência do indicador. Baseado nessa ponderação, além de testar o índice M por município, foi medido também o índice M usando a agregação mista - aquela que combina municípios e zonas eleitorais - nos vinte e seis estados brasileiros e no Distrito Federal. Nesse último, só foi testada a segunda agregação, pois essa unidade federativa não é dividida em municípios

Dos resultados aferidos pelo índice M por agregação mista, destaca-se que em apenas 24 candidatos, de um total de 513 de todos os estados, o indicador superou o valor de 0,5. Isso ocorre porque o índice M, por apurar o tamanho do eleitorado nas áreas em que o candidato recebeu mais votos até o montante de 50% da sua votação total mais um, tende a variar entre 0 e 0,5 na maioria das vezes. Esse número somente excederá 0,5 se a configuração demográfica do estado for concentrada em poucos ou em apenas um município com grande



percentual de votos sobre o eleitorado total do estado.

As médias dos estados indicaram diminuição dos coeficientes do indicador quando foi utilizada a agregação mista em vez da agregação convencional, a qual se baseia nos votos por município. Com exceção do Tocantins, que manteve o índice, os demais estados tiveram redução dos coeficientes, chegando a diminuir de 20,6 para 7,8, como no caso da Bahia. Isso se deve ao desmembramento das votações dos grandes municípios em zonas eleitorais. Na agregação municipal, os votos recebidos nas grandes cidades distorciam os resultados, pois aumentavam muito o impacto sobre o indicador calculado; na média, a redução foi de 32% nos valores dos índices. Portanto, a utilização da agregação mista trouxe uma visão mais coerente com a distribuição dos votos dos candidatos: na média, observa-se que as votações são mais concentradas do que se supunha.

Apesar das votações terem sido decompostas em agregações mistas, os resultados revelaram, ainda assim, que as médias dos estados com menos áreas foram aquelas com os maiores índices M, o que recomenda uma atenção maior ao fenômeno. Os estados com os maiores índices M foram Roraima, Amapá, Acre e Amazonas, com 58,3; 52,2; 42,9; e 42,2 respectivamente. Esses mesmos estados são aqueles com menor número de áreas de agregação, 16, 17, 24 e 64, nessa ordem. Ao analisar as áreas de agregação desses estados, verifica-se que, mesmo depois de eles terem sido decompostos, persistem áreas que representam um percentual alto do eleitorado do estado. Um exemplo é Roraima, cuja maior área representa 45% do eleitorado do estado. Esses resultados sugerem que se teste uma agregação mais detalhada para estados com poucas áreas de agregação e/ou que a maior área de agregação represente um percentual alto dos votos do estado, o que será feito na seção 5.4.

A solução encontrada para viabilizar a utilização do índice M na comparação de votações de diferentes estados foi normalizar o indicador, multiplicando-o pelo número de cadeiras em disputa no estado. Essa última etapa do cálculo permite que se tenha uma visão global das votações. Com essa operação, os estados com maiores coeficientes passam a ser Rio de Janeiro (6,71) e São Paulo (8,07), e os de menores índices são Rondônia (1,37) e Espírito Santo (1,67). Ou seja, há uma mudança na posição relativa dos estados quando se comparam suas votações.

Na Tabela 1 pode-se ver o maior e o menor MZEn dos estados. A partir dessa observação, é possível identificar que grandes estados - com numerosas áreas de agregação de votos, como Rio de Janeiro, São Paulo e Minas Gerais - possuem o maior intervalo de coeficientes MZEn, ao mesmo tempo que no Distrito Federal e em Roraima, estados com poucas áreas, 21 e 16 respectivamente, se verifica um intervalo menor do índice. Conclui-se, a

partir dessa observação, que os estados maiores, onde há maior número de vagas em disputa, são aqueles que possibilitam mais tipos diferentes de estratégias eleitorais, tanto de votos mais concentrados, típicos da lógica distrital, como votações mais dispersas, condizentes com o sistema proporcional.

A Tabela 2 mostra que quanto maior o número de áreas agregadas, menor é o peso da maior área de uma região, e o inverso também é verdadeiro. Essa lógica se reflete também nos MZEs. Sul e Sudeste, que são as regiões com maior número de áreas e de menor tamanho, são as regiões onde o MZE são mais baixos, 12,17 e 12,40, respectivamente. Do mesmo modo o Norte segue essa lógica: sendo a região com menor número de áreas, é também aquela com a maior média de MZE, 37,97. Outro ponto que se destaca na observação da Tabela 2 é o efeito da normalização no cálculo do índice M. O Sudeste, que é a região com a segunda maior densidade de acordo com o MZE, é ultrapassado pela região Norte quando se apura o valor de MZEn. Isso ocorre porque a normalização do índice M considera o percentual de cada cadeira em disputa para o cálculo do indicador.

Tabela 2 – Médias por região do número de áreas por agregação mista, MZE, MZEn e percentual da maior área de agregação de cada estado

<b>Regiões</b>	<b>áreas (MZE)</b>	<b>maior área (%)</b>	<b>média (MZE) (%)</b>	<b>média (MZEn)</b>
CENTRO-OESTE	168	5,23	28,7	2,88
NORDESTE	213	4,82	22,01	2,77
NORTE	70	18,39	37,97	3,26
SUDESTE	511	1,92	12,4	5,01
SUL	429	1,8	12,17	3,12

Fonte: elaboração própria

Nesse estudo, foi adotado o termo densidade eleitoral para propor faixas de classificação da distribuição das votações. A partir da observação dos resultados de MZEn de todos os deputados federais eleitos em 2014, foi possível compor uma escala de densidade que se divide em quatro faixas: alta densidade, para candidatos que obtiverem MZEn de 0 a 1,5; média-alta densidade, para aqueles com MZEn de 1,5 a 4,5; média-baixa densidade para os que tiverem MZEn de 4,5 a 8,0; e baixa densidade para os candidatos com MZEn maior que 8,0.

Na Tabela 3, os candidatos eleitos em 2014 se caracterizam pelo predomínio de densidades mais elevadas. Dentre os 513 deputados, 59 têm baixa densidade; 80, média-baixa; 248, média-alta; e 126, alta. Nessa tabela, é possível perceber que há somente cinco estados que têm votações de baixa densidade: Bahia, Minas, Rio, São Paulo e Rio Grande do Sul. Os quatro primeiros são aqueles estados que estão entre os cinco que mais têm votações de alta

densidade. Isso confirma a percepção de que estados maiores possibilitam uma maior diversidade de estratégias eleitorais.

Tabela 3 – Densidades das votações por estado

Estado	Alta	Média-Alta	Média-Baixa	Baixa
AC	1	6	1	0
AL	2	7	0	0
AP	1	2	5	0
AM	1	6	1	0
BA	12	21	3	3
CE	5	12	5	0
ES	7	3	0	0
DF	0	8	0	0
GO	4	8	5	0
MA	5	13	0	0
MG	11	26	11	5
MS	1	7	0	0
MT	2	6	0	0
PA	3	12	2	0
PB	1	11	0	0
PE	11	9	5	0
PI	1	8	1	0
PR	14	9	7	0
RJ	10	10	8	18
RN	1	7	0	0
RO	6	2	0	0
RR	0	2	6	0
RS	4	12	12	3
SC	5	10	1	0
SE	1	7	0	0
SP	16	17	7	30
TO	1	7	0	0
	126	248	80	59

Fonte: elaboração própria

A Tabela 4 mostra o número de votações de acordo com suas densidades e regiões. A quase totalidade dos candidatos eleitos com eleitorados de baixa densidade é proveniente da região Sudeste. As regiões Centro-Oeste e Norte não apresentam eleitorado com baixa densidade. A densidade de quase metade das votações (248) se caracterizou como média-alta; a região que mais contribui para essa densidade foi a região Nordeste, com 95 votações. Entre as altas densidades, o Nordeste também se destacou com 39 das 126; somente o Sudeste teve mais deputados com eleitorados de alta densidade. O Sul, o Norte e o Centro-Oeste distribuíram suas densidades de forma semelhante: maior número de votações médias-altas; quase a mesma

quantidade de votações alta e média-baixa; e quase nenhuma votação de baixa densidade.

Tabela 4 – Densidades das votações por região

Região	Alta	Média-Alta	Média-Baixa	Baixa
CENTRO-OESTE	7	29	5	0
NORDESTE	39	95	14	3
NORTE	13	37	15	0
SUDESTE	44	56	26	53
SUL	23	31	20	3
	126	248	80	59

Fonte: elaboração própria

## 5.2 Comparações de índices de votações do mesmo estado

Analisando caso a caso, as comparações realizadas, por meio da verificação do número e do tamanho dos municípios onde o candidato foi votado, é possível identificar inconsistências nos índices testados. Os resultados dos cinquenta e três deputados mineiros eleitos, quarenta e seis fluminenses, oito acreanos, trinta e um gaúchos e vinte e cinco pernambucanos estão representados, respectivamente, nas Tabelas 5, 6, 7, 8 e 9 em ordem crescente de valores do índice N por município. Ao analisar os dados dos candidatos eleitos em Minas Gerais (vide Tabela 5), percebe-se algumas discrepâncias quando se comparam os resultados. Os deputados mineiros Laudívio Carvalho e Tenente Lúcio obtiveram valores do índice N bem próximos, 3,60 e 3,86, respectivamente, apontando alta densidade eleitoral para os dois candidatos. No entanto, os coeficientes calculados com os outros indicadores demonstram contradição nos resultados. O N por zona eleitoral retorna valores bem diferentes, 41,41 e 16,48, indicando uma votação bem mais concentrada para o Tenente Lúcio, coerente com o resultado do índice G, que retornou os valores de 16,71 para Laudívio Carvalho e 24,03 para o Tenente Lúcio. É importante lembrar que a interpretação do indicador G deve ser feita inversamente em relação às demais: quanto menor for o coeficiente, mais fragmentado será o eleitorado. Assim, o valor de G é coerente com o cálculo do N por zona eleitoral, mas diverge do índice N por município. Já os resultados do índice M por município (M) e por agregação mista (MZE) para os deputados também espelham divergência quanto às duas votações. Tenente Lúcio recebeu mais da metade de seus votos em apenas 3,28% do eleitorado total do estado nas duas agregações calculadas (M) e (MZE). No caso de Laudívio, a maioria dos seus votos veio de municípios que perfazem 11,68% (M) do total do eleitorado e quase o mesmo percentual, 10,26%, quando se utiliza a agregação mista.

Tabela 5 – Índices dos deputados federais eleitos por Minas Gerais em 2014

Candidato Eleito	Votos	N (M)	N (ZE)	G (M)	M (M)%	MZE%	MZEn	Moran
ODELMO LEÃO	179.652	1,36	6,62	69,52	3,28	2,08	1,10	0,01
MARCELO ALVARO	60.384	1,82	5,25	39,02	11,68	1,42	0,75	0,05
MARGARIDA SALOMÃO	78.973	2,16	12,32	43,96	2,43	2,10	1,11	0,03
DELEGADO EDSON	49.391	2,87	32,34	22,40	11,68	7,56	4,01	0,10
PATRUS ANANIAS	147.175	3,28	48,31	18,75	11,68	10,37	5,50	0,08
LAUDÍVIO CARVALHO	78.762	3,60	41,41	16,71	11,68	10,26	5,44	0,18
TENENTE LÚCIO	67.459	3,86	16,48	24,03	3,28	3,28	1,74	0,08
MARCELO ARO	87.113	3,99	39,83	14,66	13,38	9,56	5,07	0,08
JULIO DELGADO	86.245	4,18	20,33	22,50	2,50	2,50	1,33	0,03
JÔ MORAES	67.650	4,28	42,01	13,48	14,44	9,79	5,19	0,21
MARCOS MONTES	116.175	4,61	14,27	21,89	1,82	1,82	0,97	0,06
LUIS TIBE	114.948	5,77	45,88	9,09	14,74	9,26	4,91	0,10
ADELMO LEAO	57.921	8,03	25,48	11,19	16,44	4,13	2,19	0,09
WELITON PRADO	186.098	8,19	32,75	7,99	17,72	6,77	3,59	0,10
MIGUEL CORREA	93.450	8,65	25,54	5,78	14,23	4,78	2,53	0,32
DÂMINA PEREIRA	52.679	11,99	10,40	9,21	14,20	2,56	1,36	0,07
LEONARDO QUINTÃO	118.470	12,37	67,92	3,16	18,60	14,32	7,59	0,18
LINCOLN PORTELA	98.834	13,69	79,67	2,31	21,52	16,09	8,53	0,17
JAIMINHO MARTINS	158.907	15,01	17,14	7,59	2,02	2,02	1,07	0,26
GEORGE HILTON	146.792	15,28	105,89	1,66	25,79	21,69	11,49	0,23
SUB-TENENTE GONZAGA	93.997	15,67	117,17	1,63	27,84	23,38	12,39	0,12
STEFANO AGUIAR	144.153	16,62	92,65	1,55	23,29	18,72	9,92	0,31
EROS BIONDINI	179.073	17,31	101,08	1,54	25,49	19,93	10,57	0,08
LEONARDO MONTEIRO	115.336	17,65	25,48	6,58	6,61	3,57	1,89	0,10
NEWTON CARDOSO JR	128.489	19,66	35,21	4,89	18,94	5,42	2,87	0,09
FABINHO RAMALHO	83.567	20,83	13,77	5,67	13,23	1,64	0,87	0,32
AELTON FREITAS	91.103	21,12	29,40	5,42	14,95	3,87	2,05	0,14
TONINHO PINHEIRO	148.239	22,26	26,16	4,09	18,93	5,35	2,84	0,31
DIEGO ANDRADE	114.240	23,26	24,46	3,26	16,42	4,31	2,28	0,21
CAIO NARCIO	101.040	24,95	24,07	2,99	16,82	4,01	2,12	0,28
RAQUEL MUNIZ	96.073	26,84	28,84	4,28	16,94	4,54	2,41	0,19
MARIO HERINGER	90.738	27,09	32,83	1,91	19,56	6,75	3,58	0,15
CARLOS MELLES	107.906	27,20	23,57	4,92	2,56	2,56	1,36	0,28
BONIFÁCIO ANDRADA	83.628	29,07	27,78	3,55	17,13	3,38	1,79	0,12
MISAEEL VARELLA	258.363	30,88	18,36	4,55	3,64	3,55	1,88	0,28
BRUNNY	45.381	30,92	41,75	3,32	15,89	4,50	2,38	0,10
DOMINGOS SÁVIO	143.901	31,26	34,27	3,11	14,98	4,15	2,20	0,22
BILAC PINTO	123.377	31,65	20,22	4,37	2,16	2,16	1,15	0,38
GABRIEL GUIMARÃES	200.014	33,40	51,02	1,53	21,30	10,39	5,51	0,16
MARCUS PESTANA	131.687	35,97	59,39	2,45	17,11	7,03	3,72	0,07
LUIZ FERNANDO	117.542	40,75	31,33	3,36	19,03	4,57	2,42	0,23
MAURO LOPES	129.795	45,92	30,84	3,20	14,35	2,78	1,47	0,25
RODRIGO PACHECO	92.743	49,30	49,72	1,89	17,87	4,68	2,48	0,22
ODAIR CUNHA	201.782	51,61	37,31	3,23	5,35	5,35	2,83	0,32
EDUARDO BARBOSA	130.453	52,88	40,88	2,36	22,25	11,46	6,07	0,11
DIMAS FABIANO	129.096	53,93	35,45	2,69	17,63	4,04	2,14	0,34
RENZO BRAZ	109.510	58,61	35,34	3,01	5,89	3,51	1,86	0,19
RODRIGO DE CASTRO	292.848	58,65	67,29	1,14	17,27	9,62	5,10	0,08
SARAIVA FELIPE	111.317	64,52	39,09	2,61	16,23	3,04	1,61	0,24
PADRE JOÃO	112.722	65,91	57,48	1,83	28,71	8,99	4,76	0,27
PAULO ABI-ACKEL	104.849	66,16	46,62	1,86	14,88	3,71	1,97	0,18
ZÉ SILVA	109.925	88,22	47,95	2,14	18,86	4,55	2,41	0,29
REGINALDO LOPES	310.226	102,80	78,93	1,94	26,71	10,39	5,51	0,16
<b>Média de MG</b>	<b>122.268</b>	<b>27,13</b>	<b>39,92</b>	<b>8,83</b>	<b>14,56</b>	<b>6,76</b>	<b>3,58</b>	<b>0,17</b>

Fonte: elaboração própria

O primeiro candidato recebeu a maioria de seus votos em um eleitorado com quase um quarto do tamanho do segundo. Por sua vez o I de Moran indica, em sentido oposto aos outros índices, uma maior concentração na votação de Laudívio (0,18) do que no eleitorado de Tenente Lúcio (0,08). No entanto, ambos os coeficientes são pouco representativos, pois estão próximos da aleatoriedade representada pelo zero do indicador.

Analisando em detalhe a votação dos dois candidatos, é possível ver que Laudívio Carvalho conseguiu 39.835 votos, ou 50,57% da sua votação, em Belo Horizonte, o que representa 11,68% do eleitorado de Minas e corresponde ao coeficiente calculado pelo índice M por município. Quando a agregação é a combinada (MZE), o resultado é parecido, a maioria dos seus votos (51,86%) se concentra em dezesseis zonas eleitorais de Belo Horizonte e da Grande BH (Contagem, Betim e Santa Luzia). Tenente Lúcio recebeu 33.830 votos (50,15% da sua votação global) em Uberlândia, município que possui 3,28% dos votos mineiros. Ou seja, ambos os candidatos receberam a maioria dos seus sufrágios em um só município, entretanto o primeiro tem seu reduto na capital do estado, já o segundo tem sua base localizada em um município bem menor. Portanto, dificilmente se poderia afirmar que eles têm o mesmo perfil de votação, como indica o índice N por município. Os demais indicadores testados mostraram diferença entre os dois candidatos, mas revelaram coeficientes bem mais dispersos do que aqueles encontrados pelos índices N e M por município e MZE, com exceção do I de Moran, que mostrou coeficientes pouco representativos e distantes do resultado de todos os outros indicadores.

A média do índice N para os candidatos mineiros foi de 27,13, por conseguinte Mario Heringer e Carlos Melles, que apresentaram 27,09 e 27,20 respectivamente para este indicador, obtiveram concentração média dentre os candidatos eleitos por Minas. No caso do índice N por zona eleitoral, os resultados foram 32,83 e 23,57, coeficientes não tão próximos entre os dois deputados e um pouco mais distantes da média do índice para o estado, que foi de 39,92. Ainda assim, indicam certa proximidade com a concentração média ao se comparar com os demais eleitos. Já no caso do índice G, apesar dos valores de Mario Heringer e Carlos Melles apresentarem aproximação, 1,91 e 4,92, respectivamente, ambos os coeficientes significam alta fragmentação, próxima da fragmentação máxima que uma votação pode obter.

Quando se testa o índice M por município, as votações resultam bem discrepantes: a de Carlos Melles (2,56) bem mais concentrada que a de Mario Heringer (19,56), a última mais próxima da média do indicador para o estado: 14,56. Ou seja, o resultado de M por município difere dos demais indicadores. Quando se afere o índice M por agregação mista (MZE), a disparidade entre os dois candidatos não se mostra tão grande: Carlos Melles tem os mesmos

2,56 da agregação municipal, e Mario Heringer tem 6,75. O índice de Moran também apresenta uma votação mais concentrada para Melles (0,15) do que para Heringer (0,28).

Ao analisar detalhadamente a distribuição eleitoral dos dois políticos, constata-se que, por agregação municipal, Mario Heringer distribuiu a maior parte da sua votação por onze municípios, dentre eles a capital Belo Horizonte, onde recebeu a maior votação, 10.571 votos, e cidades médias como Contagem, Governador Valadares e Juiz de Fora. Mas quando se examinam os dados por agregação mista, percebe-se que a maioria dos votos de Heringer é mais concentrada. Apesar de se distribuir por vinte seis zonas eleitorais de vinte e três municípios - sendo Araxá, com 8.349, e Manhuaçu, com 5.978, as cidades em que o candidato foi mais votado -, verifica-se que essas cidades representam um eleitorado mais restrito (com menor percentual do estado) do que aquele encontrado quando se mede o índice M por agregação municipal.

Já no caso de Carlos Melles, apesar da maioria de seus eleitores estar disposta em quatorze municípios de Minas Gerais, o tamanho dessas cidades era bem menor do que as de Heringer. Basta dizer que o município em que foi mais votado foi o de São Sebastião do Paraíso, pequena cidade que não chega a ter nem 0,5% do eleitorado mineiro, onde recebeu 15.319 de um total de 30.968 votos dados a deputados federais na cidade, cerca de metade do eleitorado.

Diante de tais números é difícil dizer que o eleitorado dos dois candidatos tem perfil espacial parecido. Nessa avaliação, excetuando os índices M por município, o MZE e o I de Moran que refletiram melhor as votações, os demais indicadores pecam em seus cálculos, pois colocam em patamar próximo as duas votações. Os índices N por município e N por zona eleitoral não refletem bem o eleitorado de Carlos Melles, pois sua votação é típica de eleitorado concentrado e os indicadores retornaram valores médios. Já o índice G indica alta fragmentação para os dois candidatos, muito distante da realidade do eleitorado de ambos os candidatos. Essa última discrepância parece ser a mais grave nessa comparação.

Na comparação entre as agregações testadas pelo índice M, a agregação mista (MZE) permitiu melhor observação da votação de Mario Heringer, pois, ao decompor a votação de Belo Horizonte em zonas eleitorais, percebe-se que a maior parte dos votos do candidato se distribui por pequenas cidades mineiras, indicando uma densidade eleitoral mais alta. De toda forma, nas duas comparações realizadas em Minas Gerais, não se observou grande disparidade nos resultados das duas agregações espaciais; isso se explica pela homogeneidade da demografia mineira.

A Tabela 6 apresenta os resultados das votações dos deputados federais eleitos no Rio

de Janeiro no pleito de 2014. Nessas votações, também foi possível apontar distorções entre os indicadores testados. Dentre tantas outras possíveis comparações em que se verificou discrepâncias, destacamos as votações dos deputados eleitos Alessandro Molon e Fernando Jordão, que obtiveram 2,01 e 2,06 de índice N, respectivamente. Portanto, concentrações eleitorais bastante próximas segundo esse cálculo. Entretanto, quando se aplica os outros quatro indicadores os resultados são diferentes.

O índice N por zona eleitoral retorna os valores de 129,34 para o primeiro candidato e 4,1 para o segundo, isto é, concentrações eleitorais bem diferentes: a de Molon se revelou bastante dispersa e a de Jordão extremamente concentrada. Os outros indicadores seguem a interpretação dos resultados calculados a partir da zona eleitoral. Quando se aplica o índice G, Molon obtém 11,05 e Jordão 58,74, enquanto que os valores medidos pelo M por município são 37,35 e 1,10, e pelo M por agregação mista são 20,43 e 1,10, respectivamente, para os dois candidatos. Já o cálculo do índice de Moran retorna valores bem próximos para os dois candidatos, 0,03 para Molon e 0,05 para Jordão, coeficientes bem parecidos e pouco representativos, pois estão próximos do zero do indicador.

Ao verificar as votações dos candidatos, observa-se que Alessandro Molon recebeu 70,06% dos seus 87.003 votos no município do Rio de Janeiro. Enquanto isso, Fernando Jordão também obteve 69,21% da sua votação em apenas uma cidade, no caso Angra dos Reis; no entanto esse município representa somente 1,10% do eleitorado total do estado. É interessante destacar essa comparação, pois os dois candidatos apresentaram quase três quartos dos seus votos em apenas um município, porém Jordão teve seu eleitorado concentrado em uma localidade mais de trinta vezes menor que a de Molon. Logo, seria muito complicado afirmar que as dispersões das duas votações são equivalentes, evidenciando fragilidade no cálculo do índice N nessa comparação.

Contudo, quando se desagregam os votos da cidade do Rio de Janeiro em zonas eleitorais é possível ter uma percepção mais detalhada da votação de Molon. Apesar de Molon concentrar sua votação do Rio de Janeiro, a maioria dos seus votos (50,24%) foi dada em 46 das 132 zonas eleitorais da capital, e em mais três zonas eleitorais de municípios do Grande Rio (Mesquita, Niterói e Nova Iguaçu). Portanto uma área mais restrita do que aquela encontrada quando a agregação de votos é municipal. Isso demonstra que a disparidade entre os dois candidatos não é tão acentuada, particularidade que apenas a medição realizada utilizando a agregação mista conseguiu detectar.



Tabela 6 – Índices dos deputados federais eleitos pelo Rio de Janeiro em 2014

Candidato Eleito	Votos	N (M)	N (ZE)	G (M)	M (M)%	MZE%	MZEn	Moran
PEDRO PAULO	162.403	1,35	89,81	24,49	37,35	16,30	7,50	-0,01
OTAVIO LEITE	106.398	1,41	87,84	22,49	37,35	15,35	7,06	0,00
DR JOÃO	65.624	1,66	8,14	65,67	2,98	1,70	0,78	0,06
LUIZ CARLOS DO CHAPEU	33.221	1,68	38,66	16,50	37,35	8,15	3,75	0,02
CRISTIANE BRASIL	81.817	1,71	99,94	16,11	37,35	16,50	7,59	-0,01
MIRO TEIXEIRA	58.409	1,74	78,35	15,34	37,35	10,80	4,97	0,00
RODRIGO MAIA	53.167	1,81	119,76	14,09	37,35	21,82	10,04	0,01
CHICO ALENCAR	195.964	1,90	107,58	12,64	37,35	15,88	7,31	0,01
ALESSANDRO MOLON	87.003	2,01	129,34	11,05	37,35	20,43	9,40	0,03
FERNANDO JORDÃO	47.188	2,06	4,10	58,74	1,10	1,10	0,50	0,05
JEAN WYLLYS	144.770	2,11	119,90	10,06	37,35	17,53	8,06	0,01
WASHINGTON REIS	103.190	2,29	15,90	47,23	5,14	3,72	1,71	0,07
ALEXANDRE VALLE	26.526	2,36	2,45	47,23	0,81	0,81	0,37	0,09
JANDIRA FEGHALI	68.531	2,55	165,75	6,21	37,35	30,20	13,89	0,03
ROBERTO SALES	124.087	2,73	101,17	5,68	37,35	20,05	9,22	-0,01
DELEY	48.874	2,86	9,40	42,76	2,08	2,08	0,96	0,15
INDIO DA COSTA	91.523	2,93	127,47	4,58	37,35	23,66	10,88	0,02
FELIPE BORNIER	105.517	2,94	19,98	38,46	4,71	4,02	1,85	0,07
JAIR BOLSONARO	464.572	3,05	174,12	3,75	37,35	33,75	15,52	0,05
SIMÃO SESSIM	58.825	3,08	12,16	37,76	1,24	1,24	0,57	0,17
CELSO PANSERA	58.534	3,18	23,24	26,41	5,14	4,20	1,93	0,18
EZEQUIEL TEIXEIRA	35.701	3,34	89,27	3,29	37,35	17,01	7,82	0,15
ALTINEU CORTES	40.593	3,70	9,62	33,48	6,64	2,63	1,21	0,28
GLAUBER	82.236	4,12	11,07	32,77	38,57	1,39	0,64	0,06
FABIANO HORTA	37.989	4,26	5,90	18,41	38,23	3,23	1,48	0,16
SÓSTENES CAVALCANTE	104.697	4,57	156,53	0,80	42,49	34,78	16,00	0,11
AROLDE DE OLIVEIRA	55.380	4,60	145,05	0,95	42,53	30,58	14,07	0,11
CABO DACIOLO	49.831	4,70	161,55	0,79	45,73	31,64	14,55	0,06
MARCOS SOARES	44.440	4,76	157,43	0,64	42,49	35,81	16,47	0,11
MARCELO MATOS	47.370	5,01	27,44	20,80	40,33	4,26	1,96	0,18
PAULO FEIJÓ	48.058	5,44	14,11	29,45	3,49	2,67	1,23	0,35
SERGIO ZVEITER	57.587	6,04	71,31	5,57	40,52	12,96	5,96	0,04
CHICO D'ANGELO	52.809	6,10	40,74	20,96	11,54	5,90	2,72	0,18
AUREO	58.117	6,12	39,64	20,06	45,66	7,56	3,48	0,09
ROSANGELA GOMES	101.686	6,33	87,35	2,31	47,20	19,94	9,17	0,34
CLARISSA GAROTINHO	335.061	6,89	102,32	1,78	45,26	22,85	10,51	0,05
ALEXANDRE SERFIOTIS	48.879	7,03	9,74	25,56	2,17	1,80	0,83	0,63
LEONARDO PICCIANI	180.741	7,54	55,31	2,07	40,28	16,44	7,56	0,03
BENEDITA DA SILVA	48.163	8,10	110,83	0,88	50,59	28,21	12,98	0,12
FRANCISCO FLORIANO	47.157	8,46	181,06	0,35	55,58	40,94	18,83	0,11
MARCO ANTÔNIO CABRAL	119.584	10,19	95,07	1,60	49,36	19,13	8,80	-0,01
JÚLIO LOPES	96.796	10,20	103,83	1,56	54,77	18,08	8,32	-0,02
EDUARDO CUNHA	232.708	10,42	77,68	2,63	51,30	19,59	9,01	0,17
HUGO LEAL	85.449	10,87	54,53	8,95	45,46	9,45	4,35	0,01
SORAYA SANTOS	48.204	12,41	30,02	12,90	49,75	8,00	3,68	0,37
LUIZ SÉRGIO	48.903	13,38	41,73	7,03	45,20	6,41	2,95	0,16
<b>Média de RJ</b>	<b>93.354</b>	<b>4,83</b>	<b>74,22</b>	<b>17,02</b>	<b>32,20</b>	<b>14,58</b>	<b>6,71</b>	<b>0,10</b>

Fonte: elaboração própria

Na segunda comparação entre deputados fluminenses, os candidatos Marcelo Matos e Paulo Feijó obtiveram índices N por município parecidos, 5,01 e 5,44, respectivamente; ambos os coeficientes se situam perto da média do indicador para o estado, que foi de 4,83. Já quando se aplica os demais indicadores, os resultados se revelam de outra forma. Os valores calculados de N por zona eleitoral foram 27,44 e 14,11 para os dois deputados, sugerindo concentrações um pouco maiores, pois a média do estado para esse índice é de 74,22. Os resultados do índice G seguiram na mesma linha: concentrações um pouco mais altas que a média de 17,02, sendo que os coeficientes resultantes para Matos e Feijó foram 20,80 e 29,45, respectivamente. O I de Moran calculado também indica maior concentração eleitoral para Feijó, pois o candidato obteve 0,35, enquanto o coeficiente de Matos foi de 0,18. Do mesmo modo que o índice N por zona eleitoral, os coeficientes de G e Moran indicam que Matos tem sua base eleitoral mais dispersa que Feijó. Nessa comparação, N foi o índice que indicou dispersão de votos mais semelhantes para os dois candidatos; os demais indicadores apontaram resultados diferentes, sendo a votação de Feijó mais concentrada do que a do seu oponente. No entanto, o índice que demonstrou maior disparidade entre os dois candidatos foi o M por município.

Com uma discrepância mais acentuada que os coeficientes anteriores, o índice M por município resultou em 40,33 para Marcelo Matos e 3,49 para Paulo Feijó. Marcelo Matos obteve 61,88% dos seus votos em três municípios da Região Metropolitana do Rio: Rio de Janeiro, Nova Iguaçu e São João de Meriti. No último município, recebeu sua maior votação, 19.735 votos. Os 47.370 votos de Matos se concentraram em grandes cidades do estado, mostrando que seu eleitorado, apesar de se concentrar em apenas três municípios, se localizam em cidades que abrigam grandes colégios eleitorais. Já Paulo Feijó concentra 65,92% dos seus 48.058 votos - quase dois terços do seu eleitorado - em somente três municípios do norte fluminense, demonstrando assim ter uma votação significativamente mais concentrada que o outro candidato.

Quando se analisam os votos por agregação mista, o resultado de Marcelo Matos é bastante diferente daquele calculado pela agregação por município. Nessa agregação, o índice M resulta em uma densidade bem mais alta para Matos, 4,26. A distribuição dos seus votos demonstra essa maior densidade; 51,48% da sua votação estão localizados em apenas dez áreas, seis zonas eleitorais de São João de Meriti, uma de Teresópolis e Nova Iguaçu e nas cidades de Engenheiro Paulo de Frontin e Guapimirim. Já Paulo Feijó apresenta pequena variação quando se usa a agregação mista, 2,67. Nessa agregação, a maioria dos seus votos (54,13%) é dada em quatro zonas eleitorais de Campos dos Goytacazes e nos pequenos

municípios de São Fidélis e São Francisco do Itabapoana. Nessa comparação, o índice M por agregação mista (MZE) conseguiu retratar com mais detalhe e precisão a votação de Marcelo Matos, que demonstrou ter uma densidade eleitoral alta, parecida com a votação e Paulo Feijó, assim como indicou o índice N por município.

A Tabela 7 apresenta os índices das votações dos oito candidatos eleitos pelo Acre em 2014. O perfil do eleitorado dos deputados, segundo os indicadores testados, se revelou semelhante, com exceção dos votos de Jéssica Sales que obteve 3,87 pelo cálculo do índice N por município, valor próximo do resultado calculado para este indicador do deputado Léo do PT, que obteve 4,36. Quando se testam essas votações com os outros indicadores, as disparidades aparecem.

O índice N por zona eleitoral resultou em 1,56 para Jéssica Sales e 7,49 para Léo do PT. Tais valores demonstram maior concentração eleitoral da primeira candidata em relação ao segundo candidato testado. Da mesma forma, o índice G calculado apresenta em seus coeficientes a disparidade entre os dois eleitos, Sales com 33,36 e Léo do PT com 1,14; o segundo deputado revelou uma distribuição de votos bem próxima da dispersão máxima do indicador. O I de Moran calculado para Sales é de 0,22, apesar de não ser um coeficiente de concentração expressivo, o índice da candidata é o único com significância estatística no estado. Para o mesmo indicador, Léo do PT obteve -0,03, coeficiente que indica aleatoriedade dos dados. Além disso, o valor não passou no teste realizado de pseudosignificância. Os cálculos dos índices M por agregação municipal e por agregação mista resultaram nos mesmos valores para os dois deputados, 11,70 para Jéssica Sales e 53,16 para Léo do PT. Esses dois cálculos mostraram que a candidata obteve votação bem mais concentrada que os demais eleitos no estado, do mesmo modo que os indicadores N por zona eleitoral e G.

Tabela 7 – Índices dos deputados federais eleitos pelo Acre em 2014

<b>Candidato Eleito</b>	<b>Votos</b>	<b>N (M)</b>	<b>N (ZE)</b>	<b>G (M)</b>	<b>M (M)%</b>	<b>MZE%</b>	<b>MZEn</b>	<b>Moran</b>
ALAN RICK	17.903	1,45	3,98	12,33	49,24	34,53	2,76	-0,07
ANGELIM	39.844	1,51	4,16	11,32	49,24	34,53	2,76	-0,08
MAJOR ROCHA	23.466	2,38	5,61	3,33	49,24	49,24	3,94	-0,02
FLAVIANO MELO	18.372	2,92	5,58	1,60	49,24	49,24	3,94	0,00
JÉSSICA SALES	20.339	3,87	1,56	33,36	11,70	11,70	0,94	0,22
LÉO DO PT	20.876	4,36	7,49	1,14	53,16	53,16	4,25	-0,03
SIBÁ MACHADO	18.395	4,64	7,16	1,67	51,37	51,37	4,11	0,01
CÉSAR MESSIAS	26.448	6,64	5,50	4,68	58,84	58,84	4,71	-0,14
<b>Média de AC</b>	<b>23.205</b>	<b>3,47</b>	<b>5,13</b>	<b>8,68</b>	<b>46,51</b>	<b>42,83</b>	<b>3,43</b>	<b>-0,01</b>

Fonte: elaboração própria

Analisando mais detidamente a distribuição dos votos dos dois políticos, é possível observar que Léo do PT obteve a maioria dos seus 20.876 votos (53,76%) em apenas dois municípios: a capital Rio Branco, que detém metade do eleitorado do estado (49,24%) e Tarauacá, município que abrange 3,97% dos votos do Acre. Enquanto isso, Jéssica Sales recebeu dois terços da sua votação em apenas três municípios acreanos, sendo que quatro de cada dez eleitores seus estão situados na cidade de Cruzeiro do Sul. Como se pode ver na comparação com Léo do PT, sua votação possui perfil mais concentrado do que o deputado petista, pois, apesar de ambas as votações se concentrarem em poucos municípios, o eleitorado de Jéssica Sales encontra-se em áreas menos populosas que Léo do PT.

Nesta comparação, nota-se que o índice N destoa dos demais cálculos. N é o único indicador que não detecta o eleitorado mais concentrado da deputada acreana. Outra observação que se destaca nesse cotejo é como a utilização da agregação mista para o cômputo do índice M não alterou os resultados. Isso porque o Acre é um estado muito pequeno, constituído por vinte e dois municípios, que quando são decompostos só adicionam duas áreas ao total de áreas formada na agregação mista, compondo assim uma agregação com apenas 24 áreas. Essa é uma estratificação muito pequena dos dados eleitorais, indicando a necessidade de se segmentar ainda mais os dados das votações em casos de estados muito pequenos, esse ponto será tratado mais adiante, na seção 5.4.

Os índices dos candidatos eleitos pelo Rio Grande do Sul estão na Tabela 8. Nela, pode-se ver que Onyx Lorenzoni e Perondi obtiveram índices N por município, muito parecidos, 12,01 e 12,64, respectivamente. O I de Moran calcula 0,13 para Lorenzoni e 0,08 para Perondi, números pouco expressivos de concentração, pois, apesar de serem positivos, estão muito próximos do zero. Os demais indicadores mostraram diferença entre os dois candidatos, tendo Perondi uma votação bem mais concentrada que o outro concorrente. Onyx Lorenzoni e Perondi obtiveram para o índice N por zona eleitoral 61,41 e 9,00, respectivamente; para o índice G, 2,61 e 8,68; para o índice M por município, 29,72 e 2,80; e para o índice M por zona eleitoral, 23,99 e os mesmos 2,80 do cálculo anterior.

Quando se examina as duas votações, é possível notar as diferenças de composição dos eleitorados. Onyx Lorenzoni recebeu 51,10% dos seus votos em nove municípios, sendo cinco deles da Região Metropolitana de Porto Alegre, incluindo a Capital, Caxias do Sul e Pelotas, segunda e terceira maiores cidades do estado e as cidades médias de Santa Maria e Passo Fundo. Ou seja, a votação se divide em várias cidades grandes e médias, caracterizando um eleitorado com baixa densidade. Ao explorar as votações por agregação mista a percepção é semelhante. Os votos de Lorenzoni se distribuem por dez zonas eleitorais de Porto Alegre,

mais quinze áreas de cidades já citadas anteriormente, além de Esteio, Encruzilhada do Sul, Rosário do Sul e Novo Hamburgo. Já Perondi concentra a maioria dos seus votos (51,40%) nas pequenas cidades de Ijuí, Santo Ângelo, Panambi, Cruz Alta e Três Passos, configurando uma típica votação com alta densidade.

Tabela 8 – Índices dos deputados federais eleitos pelo Rio Grande do Sul em 2014

Candidato Eleito	Votos	N (M)	N(ZE)	G (M)	M (M)%	MZE%	MZEn	Moran
PEPE VARGAS	109.469	2,78	7,99	32,48	3,87	3,87	1,20	0,04
JOÃO DERLY	106.991	5,56	42,13	8,94	16,69	16,19	5,02	0,14
MARIA DO ROSÁRIO	127.919	6,83	46,73	6,89	20,67	16,98	5,26	0,11
JOSE STEDILE	60.523	7,47	14,57	12,16	16,69	5,56	1,72	0,18
JOSÉ OTAVIO GERMANO	81.503	8,59	8,24	12,21	15,11	3,73	1,16	0,03
ONYX LORENZONI	148.302	12,01	61,41	2,61	29,72	23,99	7,44	0,13
PERONDI	109.864	12,64	9,00	8,68	2,80	2,80	0,87	0,08
GIOVANI FELTES	151.406	13,35	17,01	7,56	6,50	5,68	1,76	0,38
HENRIQUE FONTANA	128.981	13,37	41,45	3,16	15,43	12,55	3,89	0,21
SÉRGIO MORAES	115.155	14,28	10,97	7,99	3,09	3,09	0,96	0,36
MARCO MAIA	133.639	14,32	39,24	5,09	22,39	14,96	4,64	0,21
CARLOS GOMES	92.323	14,96	69,54	1,53	27,68	25,53	7,92	0,29
AFONSO HAMM	132.202	15,01	21,34	6,84	17,56	6,76	2,10	0,16
PAULO PIMENTA	140.868	15,58	34,55	5,30	21,44	13,38	4,15	0,10
NELSON MARCHEZAN JUNIOR	119.375	16,84	56,06	1,83	26,15	22,40	6,94	0,15
OSMAR TERRA	120.755	17,64	14,78	4,94	18,61	6,15	1,91	0,16
DANREI DE DEUS GOLEIRO	158.973	19,85	87,22	0,92	34,93	30,79	9,54	0,24
AFONSO MOTTA	90.917	23,05	32,14	2,62	22,26	17,57	5,45	0,12
RENATO MOLLING	102.770	24,69	15,75	4,80	7,26	6,39	1,98	0,57
HEITOR SCHUCH	101.243	28,52	27,00	4,15	25,58	8,58	2,66	0,21
RONALDO NOGUEIRA	77.017	30,53	40,53	1,49	38,08	32,01	9,92	0,17
BUSATO	130.807	30,93	58,12	2,67	30,02	20,59	6,38	0,18
MÁRCIO BIOLCHI	119.190	36,62	30,06	3,03	29,44	15,22	4,72	0,09
BOHN GASS	100.841	45,46	32,14	2,23	22,50	11,18	3,47	0,20
LUIZ CARLOS HEINZE	162.462	58,35	41,75	2,67	23,68	12,41	3,85	0,33
ALCEU MOREIRA	152.421	69,21	38,76	2,43	28,99	14,70	4,56	0,41
POMPEO DE MATTOS	91.849	69,31	69,38	0,88	37,85	28,48	8,83	0,16
GIOVANI CHERINI	115.294	71,45	45,84	2,09	25,11	14,15	4,39	0,21
MARCON	116.178	85,56	73,89	1,04	37,51	23,95	7,42	0,31
JERÔNIMO GOERGEN	115.173	98,51	57,30	1,61	38,92	21,02	6,52	0,17
COVATTI FILHO	115.131	130,87	52,99	1,85	35,20	13,56	4,20	0,23
<b>Média de RS</b>	<b>117.082</b>	<b>32,71</b>	<b>38,64</b>	<b>5,25</b>	<b>22,64</b>	<b>14,65</b>	<b>4,54</b>	<b>0,20</b>

Fonte: elaboração própria

Nessa comparação com votações de candidatos gaúchos, os índices N e I de Moran não conseguiram refletir em seus coeficientes as distribuições eleitorais analisadas. Os demais indicadores - N por zona eleitoral, G, M por município e por agregação mista - retrataram com maior acerto as votações dos dois candidatos. Nem mesmo a agregação mista representou grande ganho nas medições do índice M, isso porque a demografia eleitoral do Rio Grande do Sul é bastante homogênea. Porto Alegre, por exemplo, não excede em 11% do total de votos do estado.

A Tabela 9 apresenta os índices dos deputados eleitos por Pernambuco em 2014. Sebastião Oliveira e Bruno Araújo alcançaram os coeficientes de 23,61 e 24,03, respectivamente, para o índice N por município, valores que representam a segunda e a terceira votações mais dispersas do estado para o indicador. O índice N apurado por zona eleitoral também resulta em valores próximos com eleitorados dispersos dos dois candidatos, em que Oliveira obtém uma votação um pouco mais concentrada (20,64) que a de Araújo (24,30).

O índice G indica votações ainda mais parecidas para os dois candidatos, com Oliveira (7,07) aparecendo ligeiramente mais concentrado que Araújo (6,81). Os resultados do índice de Moran expressam votações na mesma linha dos indicadores anteriores: Bruno Araújo (0,28) com valor próximo, mas pouco mais concentrado que Sebastião Oliveira (0,23). O índice M por município indica resultados contraditórios em relação aos cálculos anteriores. Sebastião Oliveira obtém 23,97 e Bruno Araújo, 6,43. Nesse índice, Araújo teria uma votação concentrada, diferente da votação do outro candidato e do que indica os outros índices. Já o M por agregação mista retorna o mesmo índice 6,43 para Araújo, mas para Oliveira o coeficiente calculado é de 5,18, portanto eleitorado mais concentrado ainda que o do concorrente.

Quando se analisa a votação de Bruno Araújo, vê-se que ele recebeu a maioria (53,55%) dos seus votos em dez pequenos municípios de Pernambuco. O maior deles, Camaragibe, possui 1,8% do eleitorado do estado, denotando assim possuir uma alta densidade eleitoral. Já Sebastião Oliveira distribui a maioria dos seus votos (50,76%) em 13 municípios do estado, incluindo Recife, mas quando o recorte eleitoral é feito pela agregação combinada (MZE) a Capital não faz parte das cidades em que o deputado recebeu a maioria da sua votação, sendo substituída nessa relação por Triunfo, cidade com 0,16% do eleitorado do estado. Por isso, a densidade calculada para Oliveira, por meio do índice M, aumenta muito quando a agregação escolhida é a que combina municípios e zonas eleitorais.

Tabela 9 – Índices dos deputados federais eleitos por Pernambuco em 2014

<b>Candidato Eleito</b>	<b>Votos</b>	<b>N (M)</b>	<b>N (ZE)</b>	<b>G (M)</b>	<b>M (M)%</b>	<b>MZE%</b>	<b>MZEn</b>	<b>Moran</b>
FELIPE CARRERAS	187.348	3,26	29,38	12,80	18,96	16,99	4,25	0,17
WOLNEY QUEIROZ	86.739	3,39	8,55	27,89	3,08	3,08	0,77	0,01
DANIEL COELHO	138.825	4,19	29,58	8,26	22,15	17,72	4,43	0,37
ADALBERTO CAVALCANTI	99.912	5,09	9,14	21,21	3,67	3,67	0,92	0,27
BETINHO	97.269	5,34	12,58	14,88	8,87	6,63	1,66	0,33
LUCIANA SANTOS	85.053	6,48	18,91	9,66	23,76	7,65	1,91	0,37
ANDERSON FERREIRA	150.565	6,50	36,77	3,54	30,22	21,53	5,38	0,40
JARBAS	227.470	6,94	39,99	3,54	26,84	24,04	6,01	0,26
MENDONÇA FILHO	88.250	8,28	24,17	3,85	30,98	20,77	5,19	0,20
FERNANDO FILHO	112.684	8,34	16,64	11,62	23,11	4,83	1,21	0,16
<b>KAIO MANIÇOBA</b>	<b>28.585</b>	<b>12,46</b>	<b>10,69</b>	<b>6,45</b>	<b>20,90</b>	<b>3,83</b>	<b>0,96</b>	<b>0,26</b>
PASTOR EURICO	233.762	15,14	58,14	0,55	39,07	28,95	7,24	0,48
DANILO CABRAL	113.588	15,63	24,83	2,82	25,81	6,27	1,57	0,21
TADEU ALENCAR	102.669	16,51	40,34	1,51	32,79	15,08	3,77	0,24
EDUARDO DA FONTE	283.567	17,76	48,02	1,11	38,04	24,30	6,07	0,54
ZECA CAVALCANTI	97.057	17,83	16,65	8,50	3,33	3,33	0,83	0,39
JOAO FERNANDO COUTINHO	120.059	19,04	39,40	2,67	32,00	10,37	2,59	0,26
RICARDO TEOBALDO	92.262	19,06	21,40	7,98	21,97	3,28	0,82	0,25
GONZAGA PATRIOTA	101.452	20,12	27,26	4,99	24,15	5,74	1,44	0,13
ANDRE DE PAULA	100.875	22,98	30,77	4,15	22,64	5,68	1,42	0,10
JORGE CÔRTE REAL	86.023	23,10	28,64	4,25	24,60	8,26	2,07	0,13
MARINALDO ROSENDO	97.380	23,40	22,54	5,16	23,81	5,76	1,44	0,41
<b>SEBASTIAO OLIVEIRA</b>	<b>115.926</b>	<b>23,61</b>	<b>20,64</b>	<b>7,07</b>	<b>23,97</b>	<b>5,18</b>	<b>1,30</b>	<b>0,23</b>
<b>BRUNO ARAÚJO</b>	<b>131.768</b>	<b>24,03</b>	<b>24,30</b>	<b>6,81</b>	<b>6,43</b>	<b>6,43</b>	<b>1,61</b>	<b>0,28</b>
SILVIO COSTA	103.461	37,40	28,48	4,48	28,28	4,04	1,01	0,17
<b>Média de PE</b>	<b>123.302</b>	<b>14,64</b>	<b>26,71</b>	<b>7,43</b>	<b>22,38</b>	<b>10,54</b>	<b>2,63</b>	<b>0,26</b>

Fonte: elaboração própria

Na comparação entre Araújo e Oliveira, o índice M por agregação mista é o único cálculo que consegue traduzir a alta densidade eleitoral de ambos os candidatos, superando o cálculo feito com a agregação municipal, onde o peso do eleitorado de Recife desequilibrava o cômputo do indicador. Os demais índices indicaram equilíbrio entre a votação dos dois candidatos, mas apontaram alta dispersão do eleitorado, resultado não condizente com a análise das votações desses deputados.

### 5.3 Comparação de votações de estados diferentes

Um problema importante dos índices que medem a geografia eleitoral brasileira é que eles têm dificuldade em comparar votações de diferentes estados, pois nossa federação é

composta por unidades de demografia bastante variada, o que influencia muito os cálculos dos indicadores. O índice M propõe uma forma de padronizar essas disparidades regionais por meio do número de cadeiras em disputa em cada estado. A etapa final do cálculo do índice M é, portanto, a multiplicação do número de cadeiras pelos percentuais do eleitorado onde o candidato recebeu a maioria dos seus votos. Dessa forma, as diferentes necessidades que os candidatos têm de conquista do eleitorado em cada estado seriam equalizadas. Por exemplo: em Minas Gerais, um candidato concorre a uma das 53 cadeiras em disputa a deputado federal, portanto, o percentual do eleitorado que necessita conquistar é proporcional ao número de vagas, no caso  $1 / 53 = 1,88\%$ . Já um candidato no Acre, que concorre por uma das 8 cadeiras em disputa, representará  $1 / 8 = 12,5\%$  dos eleitores e terá que conquistar uma fatia bem maior do eleitorado do seu estado.

Na Tabela 10, estão os índices dos candidatos escolhidos para a comparação, são eles: Jéssica Sales do Acre, Deley do Rio de Janeiro, Sérgio Moraes do Rio Grande do Sul, Kaio Maniçoba de Pernambuco e Marcos Montes de Minas Gerais. Os cinco deputados escolhidos obtiveram índices MZEn bem parecidos, que variaram de 0,94 a 0,96. Os coeficientes MZEn são o resultado da divisão entre o percentual do eleitorado do deputado que lhe deu  $50\% + 1$  dos votos, e o percentual de eleitores necessário para se conquistar uma cadeira no estado. Pode-se então concluir que os deputados escolhidos possuem alta densidade eleitoral, pois seus coeficientes são menores que 1.

Tabela 10 – Comparação dos índices de cinco candidatos de estados diferentes

UF	Candidato Eleito	Votos	N (M)	N (ZE)	G (M)	M (M)%	MZE %	MZEn	Moran
AC	JÉSSICA SALES	20.339	3,87	1,56	33,36	11,7	11,7	0,94	0,22
RJ	DELEY	48.874	2,86	9,4	42,76	2,08	2,08	0,96	0,15
RS	SÉRGIO MORAES	115.155	14,28	10,97	7,99	3,09	3,09	0,96	0,36
PE	KAIO MANIÇOBA	28.585	12,46	10,69	6,45	20,9	3,83	0,96	0,26
MG	MARCOS MONTES	116.175	4,61	14,27	21,89	1,82	1,82	0,97	0,06

Fonte: elaboração própria

Os índices N por município e por zona eleitoral medidos para os cinco candidatos apresentam considerável variação: o N por município variou de 2,86 para Deley, o candidato mais concentrado, a 14,28 para Sérgio Moraes, sendo o deputado gaúcho o mais disperso nesse cálculo. Já o N por zona eleitoral resultou em coeficientes de 1,56 a 14,27, tendo Jéssica Sales a votação mais concentrada e Marcos Montes, o menos concentrado nessa comparação. Há também grande variação quando se observam os valores dos índices G e o I de Moran: no primeiro indicador, os candidatos obtiveram coeficientes que variaram de 6,45 para o



pernambucano Kaio Maniçoba (mais disperso) a 42,76 para o deputado fluminense Deley, o mais concentrado. O resultado do I de Moran mostra variação de 0,06 do deputado Marcos Montes, o mais disperso, a 0,36 do gaúcho Sérgio Moraes, o mais concentrado.

Os cálculos dos índices M por município e por agregação mista trouxeram praticamente os mesmos números, variando de 1,82 para Marcos Montes, com a densidade mais alta, a 11,70 de Jéssica Sales, com a menor densidade. Apenas um dos cinco deputados teve um coeficiente diferente nas duas agregações: Kaio Maniçoba computou 20,9 para a agregação municipal e 3,83 para agregação mista. Essa igualdade dos resultados se explica porque os candidatos escolhidos têm alta densidade eleitoral e seus votos se distribuem por poucas áreas e/ou por pequenos colégios eleitorais.

A análise comparativa dos cinco candidatos traz alguns resultados interessantes. Marcos Montes é o mais disperso para o índice N e I de Moran e o mais concentrado para os coeficientes M apurados. Sérgio Moraes é o mais concentrado para o I de Moran e o mais disperso para o índice N por município. Jéssica Sales tem a votação mais dispersa para os cálculos realizados pelo índice M e, ao mesmo tempo, o voto mais concentrado para o índice N por zona eleitoral. Nesses resultados, as divergências mais esperadas são aquelas que opõem os cálculos de M aos demais indicadores, pois o índice M baseia seu cálculo na proporção do eleitorado em que o candidato recebeu determinada votação, enquanto que os demais índices fundamentam seu cômputo no número de municípios em que os votos foram dados.

Assim, cabe uma observação mais pormenorizada da distribuição eleitoral dos cinco candidatos escolhidos para verificar se o índice M, normalizado pelo número de cadeiras em disputa, é capaz de estabelecer uma métrica coerente que consiga comparar as votações de candidatos de estados com distintas demografias.

Sérgio Moraes, o deputado eleito no Rio Grande do Sul, teve a maioria da sua votação distribuída por sete municípios gaúchos, o maior deles foi Vale do Sol, cidade que possui 3,08% dos votantes do estado. Jéssica Sales concentra quase metade de seus votos (47,06%) no município de Cruzeiro do Sul, que representa 9,59% dos votos do Acre. Apesar de Cruzeiro do Sul ser o segundo maior eleitorado do estado, o que aparentemente distanciaria essa das outras votações, quando se normaliza o índice M, verifica-se que o percentual dos votos que um candidato necessita conquistar para se eleger no Acre é bem maior do que nos outros estados da comparação. Isso porque, elegem-se somente oito deputados no estado para a Câmara dos Deputados. Marcos Montes, eleito por Minas Gerais, recebeu a maior parte dos seus votos (53,05%) em apenas duas cidades do estado, Uberaba e Araxá. A primeira cidade

foi a que o candidato recebeu mais votos, 52.735 de um total de 116.175. Apesar de ser a oitava população do estado, juntamente com Araxá representam somente 1,82% do eleitorado total de Minas Gerais.

Deley, eleito para a Câmara dos Deputados pelo Rio de Janeiro, concentrou 56.615 do seu eleitorado em apenas um município do sul fluminense: Volta Redonda. A cidade representa somente 2,07% dos votos do estado. Já Kaio Maniçoba, que foi eleito deputado federal por Pernambuco, teve sua votação localizada em cinco municípios do estado, incluindo a capital. Ao decompor os votos de Recife, a agregação mista mostra que o candidato não teve grande votação nas zonas eleitorais da capital, portanto, a maioria do seu eleitorado (51,19%) estava concentrada em sete pequenas cidades de Pernambuco, a maior delas é São Lourenço da Mata, que representa 1,21% dos votos do estado.

A análise minuciosa das votações dos cinco candidatos mostrou que, a despeito de tais eleitorados terem se distribuído por estados com demografias distintas, o índice MZEn, normalizado pelo número de cadeiras por estado, conseguiu categorizar de forma coerente as votações dos candidatos analisados.

#### **5.4 Estados com agregação eleitoral por seção de votação**

Ao analisar a agregação mista, que combina municípios e zonas eleitorais, observa-se que alguns estados permanecem com áreas que representam um percentual alto do eleitorado do estado, por isso surgiu a necessidade de se desagregar ainda mais os votos desses estados. Optou-se por utilizar as seções eleitorais como unidade de agregação de votos. Os critérios utilizados para selecionar os estados que tiveram seus eleitorados agregados por seções de votação foram (1) ter menos de setenta áreas e/ou (2) ter a maior área superior a 10% dos votos do estado. Identificou-se que, nos estados que se enquadram em pelo menos um dos dois critérios, a agregação mista se torna insuficiente para medir a densidade das votações. Os estados selecionados por esses critérios foram Distrito Federal, Acre, Amapá, Rondônia, Roraima e Tocantins.

Nas Tabelas 11 e 12, pode-se comparar a alteração dos resultados com a nova agregação. Destaca-se o aumento do número de áreas em que os votos são analisados: na agregação mista, o estado com mais áreas é Tocantins, com 140; na agregação por seção de votação, o número de áreas salta para, no mínimo, 1.059 para Roraima, indo até 6.452, para o Distrito Federal. Também decresce vertiginosamente o tamanho das maiores áreas: na agregação mista o percentual dos votos do estado variava de 6,70%, no Distrito Federal, a

44,50%, em Roraima; já na agregação por seção, esse tamanho reduz para cerca de 0,1% do eleitorado nos seis estados pesquisados.

Tabela 11 – Número e maior área e médias M com a agregação mista

<b>Estado</b>	<b>áreas (MZE)</b>	<b>maior área (%)</b>	<b>média (MZN)</b>	<b>média (MZEn)</b>
DF	21	6,7	42,1	3,37
AC	24	18,6	42,9	3,43
AP	17	33,8	52,2	4,19
RO	64	4,7	17,2	1,37
RR	16	44,5	58,3	4,67
TO	140	16,6	36	2,89

Fonte: elaboração própria

A partir desses dados mais detalhados, os cálculos das médias dos seis estados sofreram alteração; na maioria deles houve grande redução nos coeficientes, o que traduz um aumento considerável da densidade das votações. O estado do Tocantins teve a maior redução em seus coeficientes, de 2,89 para 1,23, representando 57,34% a menos na média do índice M calculado.

Tabela 12 – Número e maior área e médias M com a agregação por seção eleitoral

<b>Estado</b>	<b>áreas (SEC)</b>	<b>maior área (%)</b>	<b>média (MSEC)</b>	<b>média (MSECn)</b>
DF	6.452	0,02	31,22	2,52
AC	1.635	0,1	24,11	1,94
AP	1.461	0,11	23,89	1,91
RO	3.629	0,04	11,89	0,94
RR	1.059	0,15	28,22	2,26
TO	3.514	0,04	15,33	1,23

Fonte: elaboração própria

Já o Distrito Federal foi o estado que reduziu menos a sua média, de 3,37 para 2,52, uma redução de 25,20% no seu coeficiente. Rondônia, junto com o Distrito Federal, foram os dois estados que tiveram menores reduções dos seus coeficientes de M, por isso não se mostrou válida a desagregação desses dois estados no nível de seções eleitorais. Nesses dois estados, a dificuldade na manipulação dos dados não justifica o ganho observado nos resultados, uma vez que a agregação mista mostrou-se satisfatória para aferir a densidade das votações. Já nos outros quatro estados - Acre, Amapá, Roraima e Tocantins -, a diferença nos resultados justificou o uso de uma agregação mais detalhada, que foi o caso das seções de

votação. Tocantins foi o caso mais extremo de alteração, sua densidade média passou de média-alta (2,89) para alta (1,23).

Os resultados indicaram que os quatro estados onde se mostrou válido o uso da agregação por seções eleitorais foram exatamente aqueles escolhidos pelo critério de seleção dos estados que tinham áreas de agregação com percentual superior a 10% do eleitorado total do estado, pois o peso relativo dessas áreas distorce o cálculo do indicador. Já o critério que selecionava os estados por terem menos de setenta áreas de agregação, não se mostrou válido, uma vez que, nos estados com poucas áreas, mas que possuíam a maior área com menos de 10% dos votos, a agregação mista se mostrou satisfatória para medir a densidade das votações.

## 6 CONCLUSÃO

Esta pesquisa, de caráter exploratório-explicativo e com abordagem quantitativa aplicada a dados eleitorais, pretendeu esmiuçar os índices que medem a geografia eleitoral brasileira, revelando suas vantagens e desvantagens, com o intuito de elucidar o problema de pesquisa ora enunciado: **Em que medida os índices propostos para mensurar a concentração/dispersão das votações refletem a geografia eleitoral no Brasil?** Os achados sobre os indicadores mais usados confirmam a hipótese de que esses índices não refletem a realidade das votações de forma satisfatória, pois não ponderam o peso relativo dos municípios no total de votos de um estado.

Em síntese, ao avaliar a efetividade dos índices de concentração eleitoral, que é o objetivo principal deste trabalho, pode-se dizer que os testes indicaram o índice N convencional, que utiliza o município como base de cálculo e é o mais utilizado nas medições, como sendo o mais inconsistente nas comparações realizadas. A ausência de ponderação do peso relativo dos municípios é uma fragilidade que prejudica seus resultados, pois superestima pequenos municípios e subdimensiona grandes municípios. O índice N por zona eleitoral mostrou-se mais eficiente em refletir a demografia das votações do que o índice N por município, pois utiliza uma unidade de agregação espacial mais estável, a zona eleitoral.

Já o índice G mostrou eficiência mediana, situando-se entre os dois indicadores anteriores, uma vez que seu cálculo supervaloriza os desvios dos valores esperados para cada município. O I de Moran, apesar de ser o único dos indicadores que contempla a contiguidade espacial em sua formulação, padece do mesmo problema que o índice N: seu cálculo não leva em conta o peso relativo do eleitorado das áreas em que as votações ocorrem. Talvez, por isso, seus resultados se mostraram pouco significativos.

De acordo com os testes realizados e comparando-o aos demais indicadores, pode-se afirmar que o índice M<sub>z</sub> proposto, como um dos objetivos específicos desta pesquisa, refletiu o perfil eleitoral dos candidatos de forma mais realista, traduzindo-se numa ferramenta mais confiável para retratar a mensuração da geografia eleitoral. Com esses testes foram alcançados dois outros objetivos específicos: o de apresentar os prós e contras dos índices de concentração eleitoral aplicados nas votações para deputado federal; e o de aplicar e avaliar o índice M em todos estados e no Distrito Federal.

O cálculo do índice M usando os municípios como áreas de agregação eleitoral trouxe bons resultados, contudo, a utilização de zonas de agregação mista, que combina municípios e zonas eleitorais e que foi desenvolvida para este estudo, obteve resultados ainda mais

satisfatórios na medição da distribuição eleitoral das votações, pois conseguiu retratar com maior precisão a densidade das votações, minimizando a distorção no indicador causada pelos grandes municípios. A utilização das seções eleitorais em pequenos estados, cujas maiores áreas de agregação representam mais de 10% do total de votos do estado, se mostrou mais eficiente para medir a geografia eleitoral do que as demais agregações testadas. Os testes realizados com essas duas formas de agregação espacial, alternativas à forma tradicional que utiliza municípios, permitiram que se atingissem mais dois objetivos específicos do estudo: a composição de um novo mapa de agregação eleitoral, mesclando municípios e zonas eleitorais de grandes municípios; e a avaliação do índice M, tendo as seções de votação como áreas de agregação eleitoral em pequenos estados.

A normalização do índice M, por meio da multiplicação do seu percentual pelo número de vagas em disputa no estado, permite que se comparem votações de estados com diferentes demografias, superando a dificuldade presente nos índices mais usados na Ciência Política brasileira e aperfeiçoando a formulação do índice M.

Ainda que os resultados dos testes com o índice M tenham sido positivos, é válido que se reproduza essa metodologia de cálculo da geografia eleitoral em outras votações, futuras ou mesmo já ocorridas, inclusive em pleitos de legislativos estaduais e municipais, para que se confirme sua efetividade. Para auxiliar nessa tarefa e cumprindo o último objetivo específico, esta pesquisa disponibilizou os scripts de programação utilizados no Anexo A, assim como a base de dados que foi produzida em repositório público<sup>9</sup>.

Por fim, cabe ressaltar que a metodologia de cálculo proposta neste estudo pode ser aperfeiçoada. Como a maioria dos indicadores, o índice M também não contempla a contiguidade espacial das áreas de votação. Por isso, em trabalhos futuros, seria oportuno que se tentasse conciliar ao índice M algum método que avaliasse a contiguidade espacial das áreas, como o índice I de Moran.

---

<sup>9</sup> <https://github.com/Cefor/geografiaeleitoral>

## REFERÊNCIAS

AMES, Barry. Electoral rules, constituency pressures, and pork barrel: bases of voting in the Brazilian Congress. **The Journal of Politics**, [s. l.], v. 57, n. 2, p. 324-343, May 1995.

\_\_\_\_\_. **Os entraves da democracia no Brasil**. Rio de Janeiro: Ed. FGV, 2003.

ANSELIN, Luc. Local Indicators of Spatial Association – LISA. **Geographical Analysis**, [s. l.], v. 27, n. 2, p. 93-115, 1995.

AVELINO, George; BIDERMAN, Ciro; SILVA, Glauco Peres da. A concentração eleitoral nas eleições paulistas: medidas e aplicações. **DADOS – Revista de Ciências Sociais**, Rio de Janeiro, v. 54, n. 2, p. 319-347, 2011. Disponível em: <http://www.scielo.br/pdf/dados/v54n2/v54n2a04.pdf>. Acesso em: 13 mar. 2019.

\_\_\_\_\_. A concentração eleitoral no Brasil (1994-2014). **DADOS – Revista de Ciências Sociais**, Rio de Janeiro, v. 59, n. 4, p. 1.091-1.125, 2016. Disponível em: <http://www.scielo.br/pdf/dados/v59n4/0011-5258-dados-59-4-1091.pdf>. Acesso em: 13 mar. 2019.

BORGES, André; PAULA, Carolina de; SILVA, Adriano da Nóbrega. Eleições legislativas e geografia do voto em contexto de preponderância do Executivo. **Revista de Sociologia e Política**, Curitiba, v. 24, n. 58, p. 31-58, jun. 2016. Disponível em: <http://www.scielo.br/pdf/rsocp/v24n58/0104-4478-rsocp-24-58-0031.pdf>. Acesso em: 13 mar. 2019.

BRASIL. Tribunal Superior Eleitoral. **Repositório de dados eleitorais**. 2018. Disponível em: <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1/repositorio-de-dados-eleitorais>. Acesso em: 22 mar. 2019.

CÂMARA, Gilberto; MONTEIRO, Antônio Miguel; FUCKS, Suzana Druck; CARVALHO, Marília Sá. Análise espacial e geoprocessamento. In: DRUCK, Suzana *et al.* (ed.). **Análise espacial de dados geográficos**. Brasília, DF: Embrapa, 2004. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/cap1-intro.pdf>. Acesso em: 13 mar. 2019.

CARDOSO, Carlos Eduardo de Paiva. **Mobilidade em São Paulo**: a importância dos fatores socioeconômicos. [S. l.: s. n.], 2004. Disponível em: <http://www.sinaldetransito.com.br/artigos/mobilidadesp.pdf>. Acesso em: 22 jul. 2018.

CARVALHO, Nelson Rojas de. **E no início eram as bases**: geografia política do voto e comportamento legislativo no Brasil. Rio de Janeiro: Revan, 2003.

CORRÊA, Filipe Souza. Conexões territoriais: uma proposta de qualificação das conexões eleitorais na Região Metropolitana do Rio de Janeiro. In: CONGRESO LATINOAMERICANO DE CIENCIA POLÍTICA, 6., 2012, Quito. **Anales [...]**. Quito: Alacip, 2012.

COSTA, Homero de Oliveira. **Democracia e representação política no Brasil**: uma análise das eleições presidenciais (1989-2002). Porto Alegre: Sulina, 2007.

DIAS, José Luciano de Mattos. **O jogo e os jogadores: legislação eleitoral e padrões de competição política.** 1991. Dissertação (Mestrado em Ciência Política) – Instituto Universitário de Pesquisas do Rio de Janeiro, Rio de Janeiro, 1991.

FLEISCHER, David V. Concentração e dispersão eleitoral: um estudo da distribuição geográfica do voto em Minas Gerais (1966-1974). **Revista de Ciência Política**, Rio de Janeiro, v. 19, n. 3, p. 15-36, jul./set. 1976. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/rcp/article/view/59689/58032>. Acesso em: 13 mar. 2019.

FLEISCHER, David V.; RIBEIRO, Sérgio de Otero. Simulações de divisões distritais dos estados brasileiros para as eleições federais de 1978. **Revista de Informação Legislativa**, Brasília, DF, ano 20, n. 78, p. 205-232, abr./jun. 1983. Disponível em: <http://www2.senado.leg.br/bdsf/bitstream/handle/id/181444/000399848.pdf?sequence=3>. Acesso em: 13 mar. 2019.

GEODA. **Download GeoDa Software.** Version 1.12. [S. l.]: GeoDa, 2017. Disponível em: <http://geodacenter.github.io/download.html>. Acesso em: 22 jul. 2018.

GONÇALVES, Ricardo Dantas. Eleições mapeadas: como se infere sobre a dimensão geográfica das eleições presidenciais brasileiras? **Revista Eletrônica de Ciência Política**, [s. l.], v. 6, n. 2, p. 364-381, 2015. Disponível em: <https://revistas.ufpr.br/politica/article/view/42646/26899>. Acesso em: 13 mar. 2019.

\_\_\_\_\_. **Onde agrego os votos?** Contribuições à geografia eleitoral aplicada a problemas político-eleitorais brasileiros. 2016. Dissertação (Mestrado em Ciência Política) – Universidade Federal do Paraná, Curitiba, 2016. Disponível em: <https://acervodigital.ufpr.br/bitstream/handle/1884/55225/R%20-%20D%20-%20RICARDO%20DANTAS%20GONCALVES.pdf?sequence=1&isAllowed=y>. Acesso em: 13 mar. 2019.

INDJAIAN, Maria Lúcia. **Análise preliminar da bancada federal paulista eleita pelo MDB em 1978.** São Paulo: [s. n.], 1981. Relatório parcial de pesquisa apresentado à FAPESP. Mimeografado.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Portal de mapas.** 2018. Disponível em: <https://portaldemapas.ibge.gov.br/portal.php#homepage>. Acesso em: 22 mar. 2019.

KINZO, Maria D'Alva; BORIN, Ivan; MARTINS JR., José Paulo. Padrões de competição eleitoral na disputa para a Câmara paulistana: 1992-2000. **Novos Estudos Cebrap**, São Paulo, n. 65, p. 45-56, mar. 2003. Disponível em: [http://bibliotecadigital.tse.jus.br/xmlui/bitstream/handle/bdtse/5157/2003\\_kinzo\\_padroes\\_competicao\\_eleitoral.pdf?sequence=1&isAllowed=y](http://bibliotecadigital.tse.jus.br/xmlui/bitstream/handle/bdtse/5157/2003_kinzo_padroes_competicao_eleitoral.pdf?sequence=1&isAllowed=y). Acesso em: 13 mar. 2019.

LUZARDO, Antonio José Rocha; CASTAÑEDA FILHO, Rafael March; RUBIM, Igor Brum. Análise espacial exploratória com emprego do índice de Moran. **Revista GEOgraphia**, Niterói, v. 19, n. 40, p. 161-179, maio/ago. 2017. Disponível em: <http://periodicos.uff.br/geographia/article/view/13807/9007>. Acesso em: 13 mar. 2019.



MACIEL, Natalia. Padrões espaciais de voto, bases sociais e políticas dos deputados federais brasileiros: um estudo comparado entre PT, PSDB, PMDB e PFL/DEM. *In: ENCONTRO ANUAL DA ANPOCS*, 38., 2014, Caxambú. **Anais [...]**. Caxambú: Anpocs, 2014. p. 1-27. Disponível em: <https://www.anpocs.com/index.php/papers-38-encontro/spg-1/spg13-1/9287-padroes-espaciais-de-voto-bases-sociais-e-politicas-dos-deputados-federais-brasileiros-um-estudo-comparado-entre-pt-psdb-pmdb-e-pfl-dem/file>. Acesso em: 13 mar. 2019.

MARZAGÃO, Thiago. A dimensão geográfica das eleições brasileiras. **Opinião Pública**, Campinas, v. 19, n. 2, p. 270-290, nov. 2013. Disponível em: <http://www.scielo.br/pdf/op/v19n2/v19n2a02.pdf>. Acesso em: 13 mar. 2019.

MAYHEW, David R. **Congress: the electoral connection**. New Haven: Yale University Press, 1974.

MORAN, P. A. P. Notes on continuous stochastic phenomena. **Biometrika**, [s. l.], v. 37, n. 1/2, p. 17-23, 1950.

NICOLAU, Jairo. Partidos na República de 1946: uma réplica metodológica. **Dados – Revista de Ciências Sociais**, Rio de Janeiro, v. 48, n. 3, p. 589-609, 2005. Disponível em: <http://www.scielo.br/pdf/dados/v48n3/a05v48n3.pdf>. Acesso em: 13 mar. 2019.

NICOLAU, Jairo; TERRON, Sonia. Uma cidade partida? As eleições para prefeito do Rio de Janeiro em 2008. *In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE CIÊNCIA POLÍTICA*, 8., 2012, Gramado. **Anais [...]**. Gramado: ABCP, 2012.

PEREIRA, Carlos; RENNÓ, Lucio. O que é que o reeleito tem? Dinâmicas político-institucionais locais e nacionais para a Câmara dos Deputados. **DADOS – Revista de Ciências Sociais**, Rio de Janeiro, v. 44, n. 2, p. 133-172, 2001. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0011-52582001000200004&lng=en&nrm=iso&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582001000200004&lng=en&nrm=iso&tlng=pt). Acesso em: 15 maio 2018.

RSTUDIO. **Choose your version of RStudio**. Version 1.0.153. Boston, MA: RStudio, Inc, 2017. Disponível em: <https://www.rstudio.com/products/rstudio/download/>. Acesso em: 22 jul. 2018.

SAMUELS, David J. Pork barreling is not credit claiming or advertising: campaign finance and the sources of personal vote in Brazil. **The Journal of Politics**, [s. l.], v. 64, n. 3, p. 845-863, 2002.

SANTOS, Diego A. dos. “Geografia do voto”: breves apontamentos. *In: SEMINÁRIO INTERNACIONAL DE CIÊNCIA POLÍTICA*, 1., 2015, Porto Alegre. **Anais [...]**. Porto Alegre: UFRGS, 2015. Disponível em: <https://www.ufrgs.br/sicp/wp-content/uploads/2015/09/DiegoSantos.pdf>. Acesso em: 13 mar. 2019.

SAUGO, Josiana Fatima. **Geografia do voto e conexão eleitoral no Rio Grande do Sul: 1994-2006**. 2007. Dissertação (Mestrado em Ciência Política) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/12516/000627205.pdf?sequence=1>. Acesso em: 13 mar. 2019.

SIMÕES, Jessica Kamille Cruz. Conexão eleitoral e comportamento parlamentar no estado do Espírito Santo. *In: SEMINÁRIO DE CIÊNCIAS SOCIAIS*, 1., 2016, Vitória. **Anais** [...]. Vitória: UFES, 2016. p. 1-15. Disponível em: <https://eventos.ufes.br/scs/SCS2016/paper/viewFile/1930/18>. Acesso em: 13 mar. 2019.

SOARES, Gláucio Ary Dillon. **A democracia interrompida**. Rio de Janeiro: Ed. FGV, 2001.

TERRON, Sonia Luiza. **A composição de territórios eleitorais no Brasil**: uma análise das votações de Lula (1989-2006). 2009. Tese (Doutorado em Ciência Política) – Instituto Universitário de Pesquisas do Rio de Janeiro, Universidade Cândido Mendes, Rio de Janeiro, 2009. v. 1. Disponível em: <http://bibliotecadigital.tse.jus.br/xmlui/handle/bdtse/4062>. Acesso em: 13 mar. 2019.

TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. **Economic Geography**, [s. l.], v. 46, p. 234-240, Jun. 1970. Disponível em: [https://www.jstor.org/stable/pdf/143141.pdf?refreqid=excelsior%3Ac529766f408f528030ea904a74d546e7&seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/pdf/143141.pdf?refreqid=excelsior%3Ac529766f408f528030ea904a74d546e7&seq=1#page_scan_tab_contents). Acesso em: 22 jul. 2018.

VIEIRA, Márcia Paula Chaves. **Poder Legislativo no Ceará**: geografia do voto e ação política na Assembleia Legislativa. 2012. Dissertação (Mestrado em Sociologia) – Universidade Federal do Ceará, Fortaleza, 2012. Disponível em: <http://www.repositorio.ufc.br/bitstream/riufc/6416/1/2012-DIS-MPCVIEIRA.pdf>. Acesso em: 13 mar. 2019.

## APÊNDICE – CÓDIGO R

```

# limpa base de dados
rm(list = ls(all = TRUE))

# Baixar bibliotecas
library(stringr)
library(rgdal)
library(spdep)
set.ZeroPolicyOption(TRUE)

# siglas dos estados
siglas <- c("AC", "AL", "AP", "AM", "BA", "CE", "ES", "DF",
           "GO", "MA", "MG", "MS", "MT", "PA", "PB", "PE",
           "PI", "PR", "RJ", "RN", "RO", "RR", "RS", "SC",
           "SE", "SP", "TO")

# Data Frame regiões
RG <- data.frame(siglas, "regioes")
names(RG) <- c("UF", "regioes")
RG$regioes <- "NORDESTE"
RG$regioes[c(1, 3, 4, 14, 21, 22, 27)] <- "NORTE"
RG$regioes[c(8, 9, 12, 13)] <- "CENTRO-OESTE"
RG$regioes[c(7, 11, 19, 26)] <- "SUDESTE"
RG$regioes[c(18, 23, 24)] <- "SUL"

# Criando vetores
nome_arq <- siglas
nome_grv <- siglas
nome_shp <- siglas
nome_nb <- siglas
compara <- data.frame(estado="", votos="", areas_m="", areas_z="", areas_mz="",
                     vagas="", media_n_mun="", media_n_ze="", media_g_mun="",
                     media_m="", media_mze="", media_mzen="", maior_mzen="",
                     menor_mzen="", maior_area="", media_moran="", media_moranP="")
compara <- compara[-1,]
BR <- data.frame(UF="", Candidato_Eleito="", Votos="", N_mun="", N_ZE="", G_Mun="",
                 M_Mun="", M_MZE="", MZEN="", Moran_I="", Moran_P="", stringsAsFactors=FALSE)
BR <- BR[-1,]

# Baixa e seleciona os dados de cada estado
for(x in 1: length(siglas)){
  nome_arq[x] <- sub("UF", siglas[x], "votacao_candidato_munzona_2014_UF.txt")
  EST <- read.table(nome_arq[x], sep=";", header = FALSE)
  EST <- EST[,c(8,9,10,11,12,15,16,22,24,28,29)]

# Ajustando nomes de municípios aos dos Shapefiles
# Rio Grande do Norte
if (siglas[x]=="RN") {
  EST[,2] <- sub("ASSÚ", "AÇU (ASSÚ)", EST[,2])
  EST[,2] <- sub("CAMPO GRANDE", "AUGUSTO SEVERO (CAMPO GRANDE)", EST[,2])
  EST[,2] <- sub("BOA SAÚDE", "JANUÁRIO CICCÓ (BOA SAÚDE)", EST[,2])
}

# Minas Gerais

```

```

if (siglas[x]=="MG") {
  EST[,2] <- sub("PASSA VINTE","PASSA-VINTE", EST[,2])}

# Rio Grande do Sul
if (siglas[x]=="RS") {
  EST[,2] <- sub("SANTANA DO LIVRAMENTO","SANT'ANA DO LIVRAMENTO", EST[,2])}

# Preparando dados p/ agregação mista e nomeando variáveis
EST[,3] <- as.factor(EST[,3])
EST[,3] <- str_pad(EST[,3], width = 3, side = "left", pad = "0")
EST$V11 <- paste(EST[,2],EST[,3])
EST$V11 <- as.factor(EST$V11)
names(EST)[1:11] <- c("CD_MUNICIPIO","NM_MUNICIPIO","NR_ZONA", "NM_MUN_ZE",
  "NR_CANDIDATO","NM_URNA_CANDIDATO", "DS_CARGO",
  "DS_SIT_CAND_TOT","SG_PARTIDO", "COLIGACAO","QT_VOTOS")

# Seleciona os votos p/ deputado federal dos candidatos eleitos
EST <- subset(EST,DS_CARGO=="DEPUTADO FEDERAL")
eleito_EST <- subset(EST,DS_SIT_CAND_TOT=="ELEITO POR
QP"|DS_SIT_CAND_TOT=="ELEITO POR MÉDIA")
eleito_tot_EST <- aggregate(QT_VOTOS ~ NM_URNA_CANDIDATO + NR_CANDIDATO, data =
eleito_EST, sum)
names(eleito_tot_EST) <- c("Candidato_Eleito", "Número", "Votos")

# Seleciona os Votos por Municipio
eleito_EST1 <- aggregate(QT_VOTOS ~ NM_MUNICIPIO + NM_URNA_CANDIDATO +
NR_CANDIDATO, data = eleito_EST, sum)
VPMUN <- aggregate(QT_VOTOS ~ NM_MUNICIPIO, data = EST, sum)
VPMUN$QT_VOTOS_MUN <- VPMUN$QT_VOTOS/sum(VPMUN$QT_VOTOS)
for(i in 1: length(eleito_EST1$QT_VOTOS)) {
  eleito_EST1$QT_VOTOS_MUN[i] <- subset(VPMUN$QT_VOTOS_MUN,
  eleito_EST1$NM_MUNICIPIO[i]==VPMUN$NM_MUNICIPIO)
}

### Calcula o índice N de Todos os eleitos nos estados (por Municipio)
for(i in 1:length(eleito_tot_EST$Votos)) {
  eleito_tot_EST$N_Mun[i] <- 1/ sum((subset(eleito_EST1$QT_VOTOS,
  eleito_tot_EST$Candidato[i]==eleito_EST1$NM_URNA_CANDIDATO)/eleito_tot_EST$Votos[i])
^2)
}

# Seleciona os Votos por zona eleitoral (ZE)
eleito_EST2 <- aggregate(QT_VOTOS ~ NR_ZONA + NM_URNA_CANDIDATO +
NR_CANDIDATO, data = eleito_EST, sum)
VPZE <- aggregate(QT_VOTOS ~ NR_ZONA, data = EST, sum)
VPZE$QT_VOTOS_MUN <- VPZE$QT_VOTOS/sum(VPZE$QT_VOTOS)
for(i in 1: length(eleito_EST2$QT_VOTOS)) {
  eleito_EST2$QT_VOTOS_MUN[i] <- subset(VPZE$QT_VOTOS_MUN,
  eleito_EST2$NR_ZONA[i]==VPZE$NR_ZONA)
}

### Calcula o índice N de Todos os eleitos nos estados (por ZE)
for(i in 1:length(eleito_tot_EST$Votos)) {
  eleito_tot_EST$N_ZE[i] <- 1/sum((subset(eleito_EST2$QT_VOTOS,

```

```

eleito_tot_EST$Candidato[i]==eleito_EST2$NM_URNA_CANDIDATO)/eleito_tot_EST$Votos[i]^
2)
}

```

```

### Calcula o Índice G de Todos os eleitos nos estados

```

```

for(i in 1:length(eleito_tot_EST$Votos)) {
  eleito_tot_EST$G_Mun[i] <- sum((subset(eleito_EST1$QT_VOTOS,
  eleito_tot_EST$Candidato[i]==eleito_EST1$NM_URNA_CANDIDATO)/eleito_tot_EST$Votos[i]
  -subset(eleito_EST1$QT_VOTOS_MUN,
  eleito_tot_EST$Candidato[i]==eleito_EST1$NM_URNA_CANDIDATO))^2)
}

```

```

# Prepara os dados p/ calcular o índice M

```

```

eleito_EST1 <- eleito_EST1[order(eleito_EST1$NM_URNA_CANDIDATO,-
eleito_EST1$QT_VOTOS),]
eleito_EST1$VOTOS_ACUM <- cumsum(eleito_EST1$QT_VOTOS_MUN)
eleito_EST1$VOTOS_ACUM <- eleito_EST1$VOTOS_ACUM -
floor(eleito_EST1$VOTOS_ACUM)
eleito_EST1$VOTOS_ACUM[eleito_EST1$VOTOS_ACUM==0] <- 1
for(i in 1: length(eleito_EST1$QT_VOTOS)) {
  eleito_EST1$PER_VOTOS_MUN[i] <- eleito_EST1$QT_VOTOS
[i]/(subset(eleito_tot_EST$Votos,
  eleito_EST1$NM_URNA_CANDIDATO[i]==eleito_tot_EST$Candidato_Eleito))
}
eleito_EST1$PER_VOTOS_ACUM <- cumsum(eleito_EST1$PER_VOTOS_MUN)
eleito_EST1$PER_VOTOS_ACUM <- eleito_EST1$PER_VOTOS_ACUM -
floor(eleito_EST1$PER_VOTOS_ACUM)
eleito_EST1$PER_VOTOS_ACUM[eleito_EST1$PER_VOTOS_ACUM==0] <- 1

```

```

### Cálculo do índice M de todos os eleitos nos estados

```

```

PERCE <- 0.5
for(i in 1:length(eleito_tot_EST$Votos)) {
  eleito_tot_EST$M_Mun[i] <- min(subset(eleito_EST1$VOTOS_ACUM,
  eleito_tot_EST$Candidato[i]==eleito_EST1$NM_URNA_CANDIDATO &
  eleito_EST1$PER_VOTOS_ACUM > PERCE))
}
for(i in 1:length(eleito_tot_EST$Votos)) {
  eleito_tot_EST$Percentual_M[i] <- min(subset(eleito_EST1$PER_VOTOS_ACUM,
  eleito_tot_EST$Candidato[i]==eleito_EST1$NM_URNA_CANDIDATO &
  eleito_EST1$PER_VOTOS_ACUM > PERCE))
}

```

```

# Organiza os dados por Agregação mista (Município+ZE)

```

```

VPMUN_ZE <- aggregate(QT_VOTOS ~ NM_MUN_ZE + NM_MUNICIPIO, data = EST, sum)
VPMUN_ZE$AREA <- 1
votacao <- aggregate(AREA ~ NM_MUNICIPIO, data = VPMUN_ZE, sum)
for(i in 1: length(VPMUN_ZE$QT_VOTOS)) {
  VPMUN_ZE$AREA[i] <-
subset(votacao$AREA,VPMUN_ZE$NM_MUNICIPIO[i]==votacao$NM_MUNICIPIO)
}
VPMUN_ZE$QT_VOTOS_MUN_ZE <-
VPMUN_ZE$QT_VOTOS/sum(VPMUN_ZE$QT_VOTOS)
for(i in 1: length(VPMUN_ZE$QT_VOTOS)) {

```

```

VPMUN_ZE$NM_AREA[i] <-
if(VPMUN_ZE$AREA[i]==1){str_sub(VPMUN_ZE$NM_MUN_ZE[i],end = -4)}
else {str_sub(VPMUN_ZE$NM_MUN_ZE[i])}
}
VPMUN_ZE <- VPMUN_ZE[, c(6, 1, 3, 5)]

# Prepara os dados p/ calcular o indice por (Mun+ZE)
eleito_EST3 <- aggregate(QT_VOTOS ~ NM_MUN_ZE + NM_MUNICIPIO +
NM_URNA_CANDIDATO
+ NR_CANDIDATO, data = eleito_EST, sum)
for(i in 1: length(eleito_EST3$QT_VOTOS)) {
eleito_EST3$QT_VOTOS_MUN_ZE[i] <- subset(VPMUN_ZE$QT_VOTOS_MUN_ZE,
eleito_EST3$NM_MUN_ZE[i]==VPMUN_ZE$NM_MUN_ZE)
}
VPMUN_ZE$NM_MUN_ZE <- NULL

# Prepara os dados p/ calcular o indice por (Mun+ZE) (2)
eleito_EST3 <- eleito_EST3[order(eleito_EST3$NM_URNA_CANDIDATO,-
eleito_EST3$QT_VOTOS),]
eleito_EST3$VOTOS_ACUM <- cumsum(eleito_EST3$QT_VOTOS_MUN_ZE)
eleito_EST3$VOTOS_ACUM <- eleito_EST3$VOTOS_ACUM -
floor(eleito_EST3$VOTOS_ACUM)
eleito_EST3$VOTOS_ACUM[eleito_EST3$VOTOS_ACUM==0] <- 1
for(i in 1: length(eleito_EST3$QT_VOTOS)) {
eleito_EST3$PER_VOTOS_MUN_ZE[i] <- eleito_EST3$QT_VOTOS
[i]/(subset(eleito_tot_EST$Votos,
eleito_EST3$NM_URNA_CANDIDATO[i]==eleito_tot_EST$Candidato_Eleito))
}
eleito_EST3$PER_VOTOS_ACUM <- cumsum(eleito_EST3$PER_VOTOS_MUN_ZE)
eleito_EST3$PER_VOTOS_ACUM <- eleito_EST3$PER_VOTOS_ACUM -
floor(eleito_EST3$PER_VOTOS_ACUM)
eleito_EST3$PER_VOTOS_ACUM[eleito_EST3$PER_VOTOS_ACUM==0] <- 1

#### Cálculo do índice M (por Mun+ZE) de todos os eleitos nos estados
PERCE <- 0.5
for(i in 1:length(eleito_tot_EST$Votos)) {
eleito_tot_EST$M_MZE[i] <- min(subset(eleito_EST3$VOTOS_ACUM,
eleito_tot_EST$Candidato[i]==eleito_EST3$NM_URNA_CANDIDATO &
eleito_EST3$PER_VOTOS_ACUM > PERCE ))
}
for(i in 1:length(eleito_tot_EST$Votos)) {
eleito_tot_EST$Percentual_M_ZE[i] <- min(subset(eleito_EST3$PER_VOTOS_ACUM,
eleito_tot_EST$Candidato[i]==eleito_EST3$NM_URNA_CANDIDATO &
eleito_EST3$PER_VOTOS_ACUM > PERCE ))
}

#### Cálculo do MZEN
eleito_tot_EST$MZEN <- eleito_tot_EST$M_MZE/(1/(length(eleito_tot_EST$Candidato_Eleito)))

## Prepara os dados para o cálculo do I de Moran
# Importa e ordena Dados dos Shapefiles
nome_shp[x] <- sub("UF", siglas[x], "UF_municipios")
shape <- readOGR(nome_shp[x])
shape$NM_MUNICIP <- iconv(shape$NM_MUNICIP, from="UTF-8", to="latin1//TRANSLIT")

```

```

# Retira Lagoa dos Patos e Lagoa Mirim do RS
if (siglas[x]=="RS") {shape <- shape[-c(229,231),]}

#### Calcula Matriz de Ponderação dos Shapefiles
shape <- shape[order(shape$NM_MUNICIP),]

# Exclui o DF do cálculo de Moran I
if (siglas[x]=="DF") { eleito_tot_EST[c(4,6,7,8)] <- NA } else {
queen_R_nb <- poly2nb(shape)
}

# ordena dados p/ o cálculo
votacao <- votacao[order(votacao$NM_MUNICIPIO),]
eleito_EST1 <- eleito_EST1[order(eleito_EST1$NM_MUNICIPIO),]

#### Calcula I de Moran
for(z in 1:length(eleito_tot_EST$Votos)) {
  cand <-
  eleito_EST1$QT_VOTOS[eleito_EST1$NM_URNA_CANDIDATO==eleito_tot_EST$Candidato_El
eito[z]]
  votacao <- cbind(votacao,cand)
  vt <- as.data.frame(votacao[z+2])
  if (siglas[x]=="DF") { eleito_tot_EST$Moran_I[z] <- NA
    eleito_tot_EST$Moran_P[z] <- NA } else {
    eleito_tot_EST$Moran_I[z] <- as.numeric(moran.mc(vt$cand, nb2listw(queen_R_nb), nsim=99)[1])
    eleito_tot_EST$Moran_P[z] <- as.numeric(moran.mc(vt$cand, nb2listw(queen_R_nb), nsim=99)[3])
  }}

# Prepara e grava as Matrizes de votos dos deputados por Estado/Município
dep <- as.character(eleito_tot_EST$Candidato_Eleito)
names(votacao)[3:(length(dep)+2)] <- c(dep[1:length(dep)])
votacao <- cbind(votacao, shape$NM_MUNICIP)[,c(1,(length(dep)+3),2:(length(dep)+2))]
nome_nb[x] <- nome_nb[x] <- sub("UF", siglas[x], "Municipios_UF.csv")
write.table(votacao,file = nome_nb[x],row.names = FALSE, sep = ";", dec = ",")

# Armazena Matriz com todos os índices de todos os eleitos do país
BR <- rbind(BR,cbind(siglas[x], eleito_tot_EST[,c(8,10)]))

# Ordena os candidatos pelo N (por Mun)
eleito_tot_EST <- eleito_tot_EST[order(eleito_tot_EST$N_Mun),]

# Insere as médias do estado e número de Moran P(value) < 0.05
media <- as.list(replicate(11,paste("Média de",siglas[x])))
media[2] <- ""
media[3:12] <- lapply(eleito_tot_EST[3:12], mean)
if (is.na(min(eleito_tot_EST$Moran_P))) { media[13] <- NA } else {
media[13] <- as.numeric(length(eleito_tot_EST$Moran_P[eleito_tot_EST$Moran_P<0.05]))
}
media <- as.data.frame(media)
names(media) <- colnames(eleito_tot_EST)
eleito_tot_EST <- rbind(eleito_tot_EST, media)

#Acerta os índices e arredonda médias do estado
eleito_tot_EST[,c(6,7)] <- eleito_tot_EST[,c(6,7)]*100

```

```

#eleito_tot_EST[4:12] <- round(eleito_tot_EST[,4:12],2)
eleito_tot_EST[3] <- round(eleito_tot_EST[,3],0)

#Grava planilhas dos estados com todos os índices dos candidatos
nome_grv[x] <- nome_grv[x] <- sub("UF", siglas[x], "Indices_UF.csv")
write.table(eleito_tot_EST, file = nome_grv[x], row.names = FALSE, sep = ";", dec = ",")
nome_grv[x] <- nome_grv[x] <- sub("UF", siglas[x], "Cidades_UF.csv")
write.table(VPMUN, file = nome_grv[x], row.names = FALSE, sep = ";", dec = ",")

# Monta e Grava a planilha de comparação entre os estados
linha <- data.frame(estado=siglas[x],votos=sum(VPMUN_ZE$QT_VOTOS),
areas_m=length(VPMUN$NM_MUNICIPIO),
areas_z=length(VPZE$NR_ZONA), areas_mz=length(VPMUN_ZE$NM_AREA),
vagas=length(eleito_tot_EST$Candidato_Eleito)-1,
media_n_mun=mean(eleito_tot_EST$N_Mun),
media_n_ze=mean(eleito_tot_EST$N_ZE), media_g_mun=mean(eleito_tot_EST$G_Mun),
media_m=mean(eleito_tot_EST$M_Mun), media_mze=mean(eleito_tot_EST$M_MZE),
media_mzen=mean(eleito_tot_EST$MZEN), maior_mzen=max(eleito_tot_EST$MZEN),
menor_mzen=min(eleito_tot_EST$MZEN),
maior_area=max(VPMUN_ZE$QT_VOTOS_MUN_ZE),
media_moran=mean(eleito_tot_EST$Moran_I), N_moranP=as.numeric(media[13]))
linha[2:15] <- round(linha[,2:15],3)
compara <- rbind(compara, linha)
}
compara <- merge(compara, RG, by.x="estado", by.y="UF")
write.table(compara, file = "Comparacao_estados.csv", row.names = FALSE, sep = ";", dec = ",")

# Tabela com médias de MZE e MZEn das regiões
regioes <- aggregate(. ~ regioes, data= compara, mean)[,c(1,6,16,12,13)]
regioes[,c(3,4)] <- regioes[,c(3,4)]*100
regioes[,3:5] <- round(regioes[,3:5],2)
regioes[,2] <- round(regioes[,2],0)
write.table(regioes, file = "Regioes_medias.csv", row.names = FALSE, sep = ";", dec = ",")

# Grava a planilha com os índices de todos os deputados eleitos em todo o Brasil
names(BR)[1] <- "UF"
BR$Densidade <- ("Alta")
BR$Densidade[BR$MZEN>1.5 & BR$MZEN<=4.5] <- "Média-Alta"
BR$Densidade[BR$MZEN>4.5 & BR$MZEN<=8] <- "Média-Baixa"
BR$Densidade[BR$MZEN>8] <- "Baixa"
BR <- merge(BR, RG)[, c(1,14,2:13)]
write.table(BR, file = "Brasil_indices.csv", row.names = FALSE, sep = ";", dec = ",")

# Tabelas de incidência de densidade por estado e região
DE <- table(BR$UF, as.factor(BR$Densidade)) [,c(1,3,4,2)]
DE <- rbind(DE, colSums(DE))
write.table(DE, file = "Densidade_estados.csv", sep = ";", dec = ",")
DR <- table(BR$regioes, as.factor(BR$Densidade)) [,c(1,3,4,2)]
DR <- rbind(DR, colSums(DR))
write.table(DR, file = "Densidade_regioes.csv", sep = ";", dec = ",")

# Comparação entre deputados de 5 estados
depMZEN <- BR[order(BR$MZEN),]
depMZEN <- depMZEN [c(45:47,49,50),c(1,2,4:11)]
depMZEN[,6:8] <- depMZEN[,6:8] * 100

```



```

depMZEN[,4:10] <- round(depMZEN[,4:10],2)
write.table(depMZEN,file = "Cinco_deputados.csv", row.names = FALSE, sep = ";", dec = ",")

##### Cálculo por Seção eleitoral
# Monta e Grava a planilha de comparação entre os estados
compara1 <- data.frame(estado="", areas_sec="", media_SEC="",
                      media_SEcN="", maior_sec="")
compara1 <- compara1[-1,]

# Critério de Seleção de estados p/ avaliar por seção de votação
siglas <- subset(compara$estado, compara$areas_mz<70 | compara$maior_area>0.1)

# Baixa e seleciona os dados por Seção Eleitoral dos estados
for(x in 1: length(siglas)){
  nome_arq[x] <- sub("UF", siglas[x],"votacao_secao_2014_UF.txt")
  eleito_SECAO <- read.table(nome_arq[x],sep=";", header = FALSE)
  eleito_SECAO <- eleito_SECAO[,c(9:11,13:15)]
  eleito_SECAO <- eleito_SECAO[str_length(eleito_SECAO$V14)==4,]
  names(eleito_SECAO) <- c("NM_MUNICIPIO","NR_ZONA","NM_SECAO",
                        "DS_CARGO","NR_CANDIDATO","QT_VOTOS")

# Baixa e seleciona os dados por Zona Eleitoral dos estados
nome_arq[x] <- sub("UF", siglas[x],"votacao_candidato_munzona_2014_UF.txt")
EST_SEC <- read.table(nome_arq[x],sep=";", header = FALSE)
EST_SEC <- EST_SEC[,c(8,9,10,11,12,15,16,22,24,28,29)]
EST_SEC <- subset(EST_SEC,V16=="DEPUTADO FEDERAL")
names(EST_SEC)[1:11] <- c("CD_MUNICIPIO", "NM_MUNICIPIO", "NR_ZONA",
                        "NM_MUN_ZE",
                        "NR_CANDIDATO", "NM_CANDIDATO", "DS_CARGO",
                        "DS_SIT_CAND_TOT", "SG_PARTIDO", "COLIGACAO", "QT_VOTOS")

# Seleciona os votos p/ deputado federal dos candidatos eleitos por seção
eleito_tot_EST1 <- aggregate(QT_VOTOS ~ NM_CANDIDATO + NR_CANDIDATO +
                            DS_SIT_CAND_TOT,
                            data = EST_SEC, sum)
#names(eleito_tot_EST1)[2] <- "Número"
eleito_SECAO <- merge(eleito_SECAO, eleito_tot_EST1[
,c("NR_CANDIDATO","NM_CANDIDATO",
   "DS_SIT_CAND_TOT")], by="NR_CANDIDATO") [, c(7,1:6,8)]
eleito_SECAO[,4] <- str_pad(eleito_SECAO[,4], width = 3, side = "left", pad = "0")
eleito_SECAO[,5] <- str_pad(eleito_SECAO[,5], width = 3, side = "left", pad = "0")
eleito_SECAO$ZONA_SECAO <- as.factor(paste(eleito_SECAO[,4],"(",eleito_SECAO[,5],")"))
eleito_SECAO <- eleito_SECAO[, c(1:5,9,6:8)]
eleito_EST4 <- subset(eleito_SECAO,DS_SIT_CAND_TOT=="ELEITO POR
QP"|DS_SIT_CAND_TOT=="ELEITO POR MÉDIA")
eleito_EST4$DS_SIT_CAND_TOT <- NULL
eleito_EST4 <- eleito_EST4[order(eleito_EST4$NM_CANDIDATO,eleito_EST4$NM_MUNICIPIO,
                               eleito_EST4$NR_ZONA,eleito_EST4$NM_SECAO),]

# Votos por Seção Eleitoral
eleito_EST4 <- aggregate(QT_VOTOS ~ ZONA_SECAO + NM_CANDIDATO, data = eleito_EST4,
sum)
VP_SEC <- aggregate(QT_VOTOS ~ ZONA_SECAO, data = eleito_SECAO, sum)

```

```

VP_SEC$QT_VOTOS_SEC <- VP_SEC$QT_VOTOS/sum(VP_SEC$QT_VOTOS)
for(j in 1: length(eleito_EST4$QT_VOTOS)) {
  eleito_EST4$QT_VOTOS_SEC[j] <- subset(VP_SEC$QT_VOTOS_SEC,
eleito_EST4$ZONA_SECAO[j]
                                     ==VP_SEC$ZONA_SECAO)
}

# Prepara p/ indice M
eleito_EST4 <- eleito_EST4[order(eleito_EST4$NM_CANDIDATO,-eleito_EST4$QT_VOTOS),]
eleito_EST4$VOTOS_ACUM[1] <- eleito_EST4$QT_VOTOS_SEC[1]
for(i in 2: length(eleito_EST4$QT_VOTOS)) {
  if (eleito_EST4$NM_CANDIDATO[i]==eleito_EST4$NM_CANDIDATO[i-1]) {
    eleito_EST4$VOTOS_ACUM[i] <-
eleito_EST4$QT_VOTOS_SEC[i]+eleito_EST4$VOTOS_ACUM[i-1]
  } else { eleito_EST4$VOTOS_ACUM[i] <- eleito_EST4$QT_VOTOS_SEC[i]}
}

eleito_tot_EST1 <- subset(eleito_tot_EST1,DS_SIT_CAND_TOT=="ELEITO POR
QP"|DS_SIT_CAND_TOT=="ELEITO POR MÉDIA")
for(i in 1: length(eleito_EST4$QT_VOTOS)) {
  eleito_EST4$PER_VOTOS_SEC[i] <- eleito_EST4$QT_VOTOS [i] /
(subset(eleito_tot_EST1$QT_VOTOS, eleito_EST4$NM_CANDIDATO[i]
       ==eleito_tot_EST1$NM_CANDIDATO))
}
eleito_EST4$PER_VOTOS_ACUM <- cumsum(eleito_EST4$PER_VOTOS_SEC)
eleito_EST4$PER_VOTOS_ACUM <- eleito_EST4$PER_VOTOS_ACUM -
floor(eleito_EST4$PER_VOTOS_ACUM)
eleito_EST4$PER_VOTOS_ACUM[eleito_EST4$PER_VOTOS_ACUM==0] <- 1

# Calcula o índice M (por Seção Eleitoral) de todos os eleitos
PERCE <- 0.5
for(i in 1:length(eleito_tot_EST1$QT_VOTOS)) {
  eleito_tot_EST1$M_por_Secao[i] <- round(min(subset(eleito_EST4$VOTOS_ACUM,
eleito_tot_EST1$NM_CANDIDATO[i]==eleito_EST4$NM_CANDIDATO &
eleito_EST4$PER_VOTOS_ACUM > PERCE ),5)
}
for(i in 1:length(eleito_tot_EST1$QT_VOTOS)) {
  eleito_tot_EST1$Percentual_M[i] <- min(subset(eleito_EST4$PER_VOTOS_ACUM,
eleito_tot_EST1$NM_CANDIDATO[i]==eleito_EST4$NM_CANDIDATO &
eleito_EST4$PER_VOTOS_ACUM > PERCE ))
}

## Cálculo do MZEN (por Seção)
eleito_tot_EST1$MZEN <- eleito_tot_EST1$M_por_Secao /
(1/(length(eleito_tot_EST1$NM_CANDIDATO)))

# Inse as médias do estado e número de MoranP value < 0.06
media1 <- as.list(replicate(7,paste("Média de",siglas[x])))
media1[4:7] <- lapply(eleito_tot_EST1[4:7], mean)
media1 <- as.data.frame(media1)
media1[,2:3] <- NA
names(media1) <- colnames(eleito_tot_EST1)
media1[4] <- floor(media1$QT_VOTOS)
eleito_tot_EST1 <- rbind(eleito_tot_EST1, media1)

```

```

#Arredonda médias do estado
eleito_tot_EST1[5:7] <- round(eleito_tot_EST1[,5:7],2)

#Grava planilhas dos estados
nome_grv[x] <- nome_grv[x] <- sub("UF", siglas[x], "Indices_UF_Secoos.csv")
write.table(eleito_tot_EST1, file = nome_grv[x], row.names = FALSE, sep = ";", dec = ",")

# Monta Planilha de comparação entre os estados por seção
linha <- data.frame(estado=siglas[x], areas=length(VP_SEC$ZONA_SECAO),
                    media_SEC=mean(eleito_tot_EST1$M_por_Secao),
                    media_SECn=mean(eleito_tot_EST1$MZEN),
                    maior_sec=max(VP_SEC$QT_VOTOS_SEC))
linha[3:5] <- round(linha[,3:5],4)
compara1 <- rbind(compara1, linha)
}

# Compara e grava as tabelas de estados por Seções e agregação mista
agmista <- compara[c(1,14,15,18:20),][,c(1,5,15,11,12)]
agmista[,c(3,4)] <- agmista[,c(3,4)]*100
agsec <- compara1[,c(1,2,5,3,4)]
agsec[,c(3,4)] <- agsec[,c(3,4)]*100
write.table(agmista, "Agreg_mista.csv", row.names = FALSE, sep = ";", dec = ",")
write.table(agsec, "Agreg_secao.csv", row.names = FALSE, sep = ";", dec = ",")

```