



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE BACHARELADO EM ESTATÍSTICA**

GILMAR BARBOSA DA SILVA

**MODELAGEM ESPACIAL DOS CASOS NOTIFICADOS POR COVID-19 NO ESTADO
DA PARAÍBA E POSSÍVEIS ASSOCIAÇÕES COM FATORES SOCIOECONÔMICOS**

CAMPINA GRANDE - PB

2022

GILMAR BARBOSA DA SILVA

MODELAGEM ESPACIAL DOS CASOS NOTIFICADOS POR COVID-19 NO ESTADO DA PARAÍBA E POSSÍVEIS ASSOCIAÇÕES COM FATORES SOCIOECONÔMICOS

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Ricardo Alves de Olinda

CAMPINA GRANDE - PB

2022

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

S586m Silva, Gilmar Barbosa da.
Modelagem espacial dos casos notificados por covid-19 no estado da Paraíba e possíveis associações com fatores socioeconômico [manuscrito] / Gilmar Barbosa da Silva. - 2022.
41 p. : il. colorido.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2022.
"Orientação : Prof. Dr. Ricardo Alves de Olinda, Coordenação do Curso de Estatística - CCT."
1. Saúde pública. 2. Pandemia. 3. Regressão espacial. 4. Paraíba. I. Título

21. ed. CDD 519.5

GILMAR BARBOSA DA SILVA

MODELAGEM ESPACIAL DOS CASOS NOTIFICADOS POR COVID-19 NO ESTADO DA
PARAÍBA E POSSÍVEIS ASSOCIAÇÕES COM FATORES SOCIOECONÔMICOS

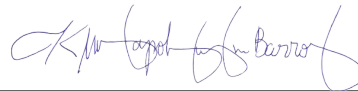
Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Bacharel em Estatística.

Trabalho aprovado em 01 de abril de 2022.

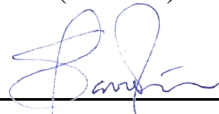
BANCA EXAMINADORA



Prof. Dr. Ricardo Alves de Olinda (Orientador)
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Kleber Napoleão Nunes de Oliveira
Barros
Universidade Federal Rural de Pernambuco
(UFRPE)



Prof. Dr. Silvio Fernando Alves Xavier Júnior
Universidade Estadual da Paraíba (UEPB)

AGRADECIMENTOS

Gostaria de agradecer principalmente a Deus, meus pais e minha filha, que apesar de ser tudo bem recente com ela me fez repensar muito na vida. Obrigado por toda a força e ajuda que foram me dadas ao longo de todos esses anos. Gostaria também de agradecer a mim, por não ter desistido, por ter confiado, por ter acreditado, por ser a pessoas que sou apesar de todas as coisas que aconteceram nesses anos de curso.

Ao Professor e orientador Dr. Ricardo Alves de Olinda pelo conhecimento adquirido, pelo crescimento profissional e principalmente pela paciência.

Ao meu professor de Geografia Valdoíres Cachorrão e a minha professora de Educação Física Zélia Chacon que foram meus professores do Ensino Fundamental. Podem ter certeza que os senhores fazem parte dessa história.

Aos amigos do curso de Estatística, Wylliam Alves, Mateus Peixoto, Beatriz, Rafaella, Mayara, Jiulia, Julyanne, Viviane, Eduardo, Sr. Álvaro, Elizandra, entre outros, por todos os estudos e conhecimentos compartilhados e por serem verdadeiros companheiros de jornada.

Aos amigos do espeto, Sr. Antônio, Jeysianne, Louhanne, Bruno, 'Josimara', Sammara... entre outros amigos de copo.

E por último, não menos importante, gostaria de agradecer a todos os meus professores da graduação. Obrigado por vocês passarem um pouco de seus conhecimentos. Muito obrigado!

*"Não deixe que as pessoas te façam desistir daquilo que você mais quer na vida. Acredite. Lute.
Conquiste. E acima de tudo, seja feliz!"
(Autor Desconhecido)*

*"Sem dados você é apenas mais uma pessoa com uma opinião."
(WILLIAM EDWARDS DEMING)*

RESUMO

No final de dezembro de 2019, na China, surgiu um surto de um novo vírus denominado SARS-CoV-2, que ocasionou uma doença infecto contagiosa causadora da COVID-19. A doença se dispersou rapidamente pelo mundo todo infectando milhares de pessoas, se tornando não só um problema de Saúde Pública mundial, mas causando abalos econômicos, políticos e social. No mês de fevereiro de 2020 o Brasil relata seu primeiro caso de COVID-19 no estado de São Paulo. Chegando a ser confirmado pela primeira vez em solo paraibano (área de estudo) no mês de março de 2020, através de um indivíduo que deu entrada em uma unidade da rede particular com indícios da doença. Após isso, a Paraíba foi o último colocado em casos notificados na região nordeste segundo os dados epidemiológicos do estado. Sendo assim, o trabalho tem como objetivo confirmar a possibilidade de que pode haver uma autocorrelação espacial ou dependência espacial e possíveis fatores socioeconômicos que venham influenciar no aumento dos casos de COVID-19 na região paraibana. Desta forma, aplica-se uma análise exploratória da estatística espacial para Dados de Área, que permite visualizar espacialmente os eventos – neste caso a COVID-19 - em estudo por meio de aglomerações e autocorrelação espacial. Modelos de Regressão Espacial também é uma ferramenta importante para explicar a estrutura de autocorrelação dos casos notificados por COVID-19 e variáveis socioeconômicas. Desta forma, com base nos resultados obtidos, pode-se concluir que, o estudo espacial da COVID-19 nos dez meses do ano de 2020 no estado da Paraíba, mostra que há presença de autocorrelação espacial entre os municípios, formando aglomerações em várias mesorregiões. Desta maneira, o estudo torna-se de total relevância para os órgãos competentes do estado, pois relata as regiões que podem contribuir para o agravamento de problemas na saúde pública

Palavras-chave: Regressão Espacial. Saúde pública. Pandemia.

ABSTRACT

At the end of December 2019, in China, an outbreak of a new virus called *SARS-CoV-2*, a contagious infectious disease that causes COVID-19, emerged. The disease quickly spread around the world, infecting thousands of people, becoming not only a public health problem worldwide, but also causing economic, political and social shocks. In the month of February 2020, Brazil reports its first case of COVID-19 in the state of São Paulo. It was confirmed for the first time on Paraíba soil (study area) in March 2020, through an individual who was admitted to a unit of the private network with evidence of the disease. After that, Paraíba was the last place in cases reported in the northeast region according to the epidemiological data of the state. Therefore, the work aims to confirm the possibility that there may be a spatial autocorrelation or spatial dependence and possible socioeconomic factors that may influence the increase in COVID-19 cases in the Paraíba region. In this way, an exploratory analysis of spatial statistics is applied to Area Data, which allows the spatial visualization of the events - in this case COVID-19 - under study through agglomerations and spatial autocorrelation. Spatial Regression Models is also an important tool to explain the autocorrelation structure of COVID-19 reported chaos and socioeconomic variables. In this way, based on the results obtained, it can be concluded that the spatial study of COVID-19 in the ten months of the year 2020 in the state of Paraíba, shows that there is a presence of spatial autocorrelation between the municipalities, forming agglomerations in several mesoregions. In this way, the study becomes of total relevance for the competent bodies of the state, as it reports the regions that can contribute to the aggravation of problems in public health.

Keywords: Spatial Regression. Public Health; Pandemic.

LISTA DE ILUSTRAÇÕES

Figura 1 – Estado da Paraíba	15
Figura 2 – Matriz de proximidade espacial usando como critério a fronteira entre as áreas.	16
Figura 3 – Áreas e Matriz de proximidade espacial, normalizada pelas linhas.	17
Figura 4 – Diagrama de Espalhamento de Moran.	19
Figura 5 – Boxplot da taxa de incidência dos casos notificados por COVID-19 no estado da Paraíba para o ano de 2020.	25
Figura 6 – Mapa dos quartis da taxa de incidência dos casos notificados por COVID-19 no estado da Paraíba	26
Figura 7 – Espalhamento de Moran para os casos notificados por COVID-19 nos municípios do estado da Paraíba de março a agosto.	28
Figura 8 – Espalhamento de Moran para os casos notificados por COVID-19 nos municípios do estado da Paraíba de setembro a dezembro.	29
Figura 9 – Lisa Map das taxas dos casos notificados por COVID-19 nos municípios do estado da Paraíba de março a junho.	30
Figura 10 – Lisa Map das taxas dos casos notificados por COVID-19 nos municípios do estado da Paraíba de julho a dezembro.	31
Figura 11 – Moran Map das taxas notificados por COVID-19 nos municípios do estado da Paraíba.	31
Figura 12 – Moran Map das taxas notificados por COVID-19 nos municípios do estado da Paraíba.	32
Figura 13 – Mapa dos resíduos para modelos de regressão espacial para a taxa de incidência da COVID-19 no ano de 2020	38

LISTA DE TABELAS

Tabela 1 – Estatística Descritiva para os casos notificados por COVID-19 para os meses de março a julho de 2020.	23
Tabela 2 – Estatística Descritiva para os casos notificados por COVID-19 para os meses de agosto a dezembro de 2020.	23
Tabela 3 – Estatística Descritiva para as taxas de incidência de COVID-19 nos meses de março a julho de 2020.	24
Tabela 4 – Estatística Descritiva para as taxas de incidência de COVID-19 nos meses de agosto a dezembro de 2020.	25
Tabela 5 – Tabela das estimativa do Índice I de <i>Moran</i> para os casos de COVID-19 no estado da Paraíba.	27
Tabela 6 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de abril de 2020.	34
Tabela 7 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de maio de 2020.	34
Tabela 8 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de junho de 2020.	35
Tabela 9 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de julho de 2020.	35
Tabela 10 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de agosto de 2020.	36
Tabela 11 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de outubro de 2020.	36
Tabela 12 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de novembro de 2020.	36
Tabela 13 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de dezembro de 2020.	37

LISTA DE ABREVIATURAS E SIGLAS

RNA	Ácido ribonucléico
PIB	Produto Interno Bruto
OMS	Organização Mundial de Saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
SIDRA	Sistema IBGE de Recuperação Automática
IDHM	Índice de Desenvolvimento Humano Municipal
RDPC	Renda Domiciliar per capita
COB_AB	% Cobertura de Atenção Básica em Saúde
DENSDEMOG2017	Densidade Demográfica
Q1	Primeiro quadrante
Q2	Segundo quadrante
Q3	Terceiro quadrante
Q4	Quarto quadrante
AA	Alto-Alto
AB	Alto-Baixo
BA	Baixo-Alto
BB	Baixo-Baixo
RL	Regressão Linear
SAR	Spatial AutoRegressive Model (Modelo Espacial Autoregressivo Misto)
SEM	Spatial Error Model (Modelo de Erro Espacial)
SDM	Spatial Durbin Model (Modelo Espacial de Durbin)
AIC	Akaike Information Criterion (Critério de Informação de Akaike)
TIC	Taxa de Incidência da COVID-19

LISTA DE SÍMBOLOS

μ	Letra grega minúscula mu
β	Letra grega minúscula beta
ε	Letra grega minúscula epsilon
σ	Letra grega minúscula sigma
ρ	Letra grega minúscula rho
λ	Letra grega minúscula lâmbda
ξ	Letra grega minúscula ksi
Θ	Letra grega maiúscula theta
θ	Letra grega minúscula theta

SUMÁRIO

1	INTRODUÇÃO	13
2	MATERIAL E MÉTODOS ESTATÍSTICOS	15
2.1	Local de Estudo e Material	15
2.2	Metodologia	16
2.2.1	<i>Matriz de Vizinhaça</i>	16
2.2.2	<i>Média Móvel</i>	17
2.2.3	<i>Autocorrelação Espacial</i>	17
2.2.3.1	<i>Autocorrelação Espacial Global (Índice I de Moran Global e a Estatística C de Geary)</i>	17
2.2.3.2	<i>Autocorrelação Espacial Local (I de Moran Local)</i>	18
2.2.4	<i>Gráfico de Espalhamento de Moran</i>	19
2.2.5	<i>Box Map, Lisa Map e Moran Map</i>	20
2.2.6	<i>Modelo de Regressão Linear</i>	20
2.2.7	<i>Modelo Espacial Autorregressivo Misto - SAR</i>	20
2.2.8	<i>Modelo de Erro Espacial - SEM</i>	21
2.2.9	<i>Modelo Espacial de Durbin - SDM</i>	21
2.2.10	<i>Seleção de Modelos</i>	21
3	RESULTADOS E DISCUSSÃO	23
3.1	Análise Descritiva para os Casos Notificados por COVID-19.	23
3.2	Cálculo da Taxa de Indicência dos Casos Notificados por COVID-19.	24
3.3	Análise Descritiva da Taxa de Incidência dos Casos Notificados por COVID-19	24
3.4	Dependência Espacial	27
3.5	Análise de Regressão Espacial	33
3.5.1	<i>Modelagem para os dados Taxa Incidência da COVID-19 nos meses de abril e maio de 2020.</i>	33
3.5.2	<i>Modelagem para os dados Taxa Incidência da COVID-19 nos meses de junho de 2020.</i>	34
3.5.3	<i>Modelagem para os dados Taxa Incidência da COVID-19 nos meses de julho e agosto de 2020.</i>	35
3.5.4	<i>Modelagem para os dados Taxa Incidência da COVID-19 nos meses de outubro, novembro e dezembro de 2020.</i>	36
4	CONCLUSÃO	39

REFERÊNCIAS 40

1 INTRODUÇÃO

O coronavírus denominado SARS-CoV-2, doença que causa COVID-19, foi identificada pela primeira vez na China, em dezembro de 2019 (ZHU et al., 2020). O mesmo é derivado da família coronavírus, que é um amplo grupo de vírus que geram infecções respiratórias. Existem sete tipos de coronavírus identificados que são capazes de causar doenças em humanos (LANA et al., 2020).

A COVID-19 é uma doença infecto contagiosa que atualmente vem causando preocupação na população e governantes em escala global. O vírus composto por ácidos ribonucleico (RNA) responsáveis pela infecção respiratórias em diversos animais, aves e mamíferos (FEHR; PERLMAN, 2015). A pandemia causada pela COVID-19 não só tem impacto na Saúde Pública em proporção mundial, mas também fez com que houvesse um abalo social, político, cultural, histórico e principalmente econômico. No Brasil, segundo Silva, Santos e Soares (2020), o resultado causado pela pandemia do novo coronavírus provocará uma diminuição do Produto Interno Bruto (PIB) brasileiro em 5% no ano de 2020.

De acordo com os dados da Universidade Johns Hopkins, situada em Baltimore, Maryland, Estados Unidos, no ano de 2020 foram registrado mais de 82,5 milhões de pessoas que foram contaminadas pelo vírus em 191 países/regiões diferentes, somando mais de 1,8 milhão de mortos pelo mundo. Usando em comparação com a gripe russa, que ocorreu entre 1889 e 1890, e fez entre um milhão e um milhão e meio de vítimas, o novo coronavírus conseguiu superar esse número e fazer 300 mil vítimas a mais no seu primeiro ano.

Segundo Almeida et al. (2021), os sistema de saúde de todo o mundo vêm indicando, desde a década de 60, óbitos e infecção associado ao coronavírus, mas nenhuma dos casos notificados conseguiu a gravidade da pandemia atual com a descoberta do SARS-CoV-2. Esta doença causa sintomas que variam de uma gripe simples até uma pneumonia mais grave. Tendo como sinais mais frequentes, segundo Singhal (2020), febre (50%) e tosse (38%).

Em 26 de fevereiro de 2020, após quase dois meses depois de a China avisar à Organização Mundial de Saúde (OMS) sobre os primeiros casos de infecção por um novo vírus ainda desconhecido, o Brasil registra seu primeiro caso de COVID-19 no estado de São Paulo. De acordo com Croda e Garcia (2020), no dia 3 de março existiam 488 casos suspeitos notificados, quando foram 2 confirmados e 240 descartados. Os primeiros dois casos confirmados eram de cidadãos do sexo masculino, moradores da cidade de São Paulo, que haviam acabado de voltar de viagem da Itália.

Segundo o site *covid.saude.gov.br*, no ano de 2020, o Brasil registrou cerca 7.675.973 casos confirmados e 194.949 vítimas. Tendo a região Sudeste com a maior quantidade de casos confirmados, sendo 2.688.086. Seguindo das regiões Nordeste que quantificou 1.896.157, Sul com 1.358.823, Norte e Centro-Oeste com 859.150 e 873.757 respectivamente. O ano de 2020 acaba em metade a uma segunda onda da doença, com o descobrimento de uma nova linhagem do coronavírus e um aumento de casos, óbitos e buscas por internações.

No Nordeste, os primeiros casos registrados foram no estado da Bahia e o primeiro óbito em Pernambuco. A região é considerada a segunda do Brasil com o maior número de casos notificados da COVID-19. No estado da Paraíba, segundo Daspett (2020) apud Costa, Tavares e Clemente (2021), a primeira confirmação foi em 18 de março de 2020, quando um indivíduo com indícios clínicos da doença foi se consultar em uma unidade de saúde da rede particular. O estado, localizado na segunda região com mais casos notificados, foram quantificados 166.484 casos em todo o ano de 2020. Ficando atrás de Bahia com 493.400, Ceará 335.091, Pernambuco e Maranhão com 222.166 e 200.938, respectivamente. Em seguida, temos Piauí com 142.672, Rio Grande do Norte 118.083, Sergipe e Alagoas com 112.505 e 104.810 nesta ordem, de acordo com os dados epidemiológicos do estado da Paraíba.

Segundo Silva et al. (2021), para coibir a doença no estado, o Governo da Paraíba contratou profissionais da área de saúde em diversas funções, ampliados leitos de terapias intensivas, enfermaria e também foram construídos 130 leitos com competência de expansão 200 leitos, que foram construídos do zero no estacionamento do Hospital Metropolitano Dom José Maria Pires.

Neste estudo, foram utilizadas técnicas espaciais que é uma área da estatística que nos permite analisar de forma espacial a localização de determinados fenômenos/eventos Andrade et al. (2007). Neste caso, foi utilizado para analisar os casos notificados por COVID-19 no estado da Paraíba no ano de 2020. A estatística espacial contém processos pontuais, superfície contínua e dados de área como seus três principais métodos. No qual, essas técnicas podem ser aplicadas em diversas áreas, como agronomia, geologia, epidemiologia, ecológicos, entre outros.

Esta monografia, tem como objetivo verificar através de modelos espaciais para dados de área a hipótese de que fatores socioeconômicos tenham possíveis associações no aumento ou na diminuição da COVID-19 no estado da Paraíba. Tendo ainda como meta quantificar os casos e verificar a autocorrelação tanto global como local através dos métodos de dados de área.

2 MATERIAL E MÉTODOS ESTATÍSTICOS

2.1 Local de Estudo e Material

Observa-se por meio da Figura 1, o mapa referente ao estado da Paraíba que é um dos nove estados pertencentes a região Nordeste. A Paraíba contém 223 municípios, uma extensão territorial de 56.468,80 km^2 e segundo o Instituto Brasileiro de Geografia e Estatística - (IBGE, 2020)¹, a população estimada para o ano de 2020 era de 4.039.277.

Figura 1 – Estado da Paraíba



Fonte: Produzido pelo autor.

Os dados da COVID-19 utilizados para análise desta monografia estão disponíveis no site: *covid.saude.gov.br*² que por sua vez, foram acessados e baixados pela última vez em 31/05/2021. O site fornece os casos notificados por COVID-19 para todos os dias do ano de 2020 e é utilizado para divulgar boletins que são atualizados diariamente. No decorrer do estudo, percebe-se que foi utilizada a Taxa de Incidência dos casos notificados de COVID-19 sendo explicado no Capítulo 3. Todas as análises estatísticas foram feitas com a linguagem de programação Software R (TEAM et al., 2020) e serão ilustrados na forma de mapas. As variáveis explicativas selecionadas para o ajuste dos modelos de regressão espacial foram retiradas do Sistema IBGE de Recuperação Automática (SIDRA) e são:

- IDHM: Índice de Desenvolvimento Humano Municipal - 2010;
- RDPC: Renda Domiciliar per capita - 2010;
- COB_AB²: % Cobertura de Atenção Básica em Saúde - 2015;
- T_ANALF18M: Taxa de analfabetismo 18 anos ou mais - 2010;
- DENSDEMOG2017: Densidade demográfica - 2017.

¹ <https://www.ibge.gov.br>.

² <https://covid.saude.gov.br>.

2.2 Metodologia

Na estatística espacial há três grande áreas, e uma dessas área é a análise de dados de área, onde os dados são disponibilizados de modo acumulado em áreas que são marcadas por polígonos, como cidades, bairros, etc. Esse método normalmente busca quantificar eventos, como número de óbitos, números de assaltos, números de acidentes, etc. De acordo com Costa (2017), os dados estão mencionados a uma determinado local, que é dividido em n áreas A_i sendo $i = 1, 2, \dots, n$, onde cada área é realizada uma observação aleatória Y_i , em que Y_i é uma conciliação sobre a área A_i , na qual essa mensuração é referente a toda área A_i e não a um certo ponto específico.

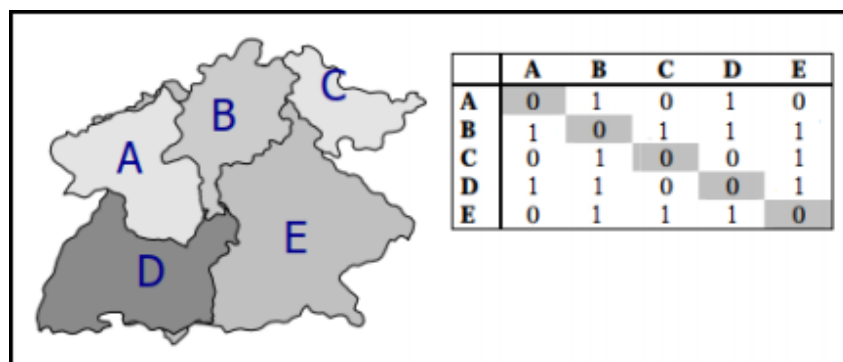
2.2.1 Matriz de Vizinhaça

Os métodos de análise de dados de área inicia-se com a elaboração de uma matriz de vizinhaça, que por sua vez, também é conhecida como matriz de distância ou matriz de proximidade. Onde essa matriz de proximidade indica a relação espacial de cada área com as demais. Essa matriz nos fornece informações sobre o quão forte é a interação entre cidades mais próximas e mais fracas em cidades distantes. De acordo com Costa (2017), a matriz de proximidade é da seguinte maneira:

$$\mathbf{W} = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & W_{n2} & \cdots & W_{nn} \end{bmatrix},$$

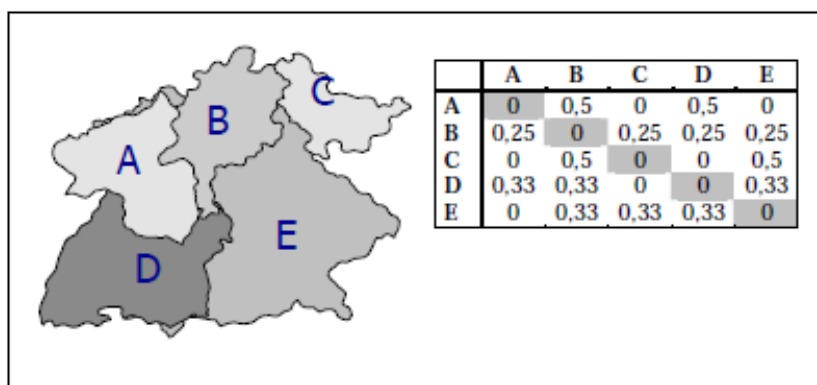
em que n significa o número de observações obtida na matriz \mathbf{W} . E cada elemento W_{ij} (elementos da matriz \mathbf{W}) representa o nível de proximidade entre A_i e A_j . Desta forma, são atribuídos aos elementos da matriz \mathbf{W} valor 1, quando as áreas são consideradas vizinhas, e 0 caso contrário. As Figuras 2 e 3, respectivamente, são modelos básicos da construção de matriz de proximidade, em que os valores 0 e 1 já foram atribuídos seguindo o critério de vizinhos mais próximos.

Figura 2 – Matriz de proximidade espacial usando como critério a fronteira entre as áreas.



Fonte: (CÂMARA et al., 2004)

Figura 3 – Áreas e Matriz de proximidade espacial, normalizada pelas linhas.



Fonte: (CÂMARA et al., 2004)

A Figura 3 nos mostra a necessidade também de normalizar suas linhas, para que a soma dos pesos de cada linha seja igual a 1. E assim, poder calcular a autocorrelação entre esses municípios.

2.2.2 Média Móvel

É possível usar a média móvel para observar a variação de tendência espacial da taxa de pessoas com COVID-19 na Paraíba, pois é a forma mais clara de analisar a variação da tendência espacial dos dados. Que por sua vez, faz com que haja uma diminuição de flutuação em relação aos dados originais (CÂMARA et al., 2004). A expressão é dada por:

$$\hat{\mu}_i = \sum_{i=1}^n \sum_{j=1}^n W_{ij}(y_j - \bar{y}), \quad (2.1)$$

onde W_{ij} são os elementos da matriz de proximidade normalizada, y_j é o valor do atributo em cada área e n é o número de observações na matriz \mathbf{W} .

2.2.3 Autocorrelação Espacial

A dependência espacial que é também conhecida por autocorrelação espacial, como o próprio nome indica, mede a correlação entre duas variáveis da mesma forma que os princípios da correlação estatística. Segundo Griffith (2003), a autocorrelação espacial é a medida de correlação realizada com a mesma variável aleatória, porém medidas em locais distintos no espaço geográfico. A dependência espacial é um conceito muito importante para entender e analisar os eventos espaciais.

2.2.3.1 Autocorrelação Espacial Global (Índice I de Moran Global e a Estatística C de Geary)

A autocorrelação espacial global é um método estatístico utilizado para verificar se há presença ou não de dependência espacial. Há dois métodos que têm sido bastante utilizados para calcular a dependência espacial global em dados de área. São eles o I de Moran Global, que

segundo Almeida, Perobelli e Ferreira (2008), indica à associação linear do vetor de observação da variável estudada em um determinado tempo e os valores da média ponderada de vizinhança. E a Estatística C de Geary, que segundo Lembo(2005) apud Rêgo e Pena (2012) é preferido ao I de Moran quando existe uma pequena quantidade de vizinhanças. O I de Moran, proposto por Moran (1950) pode ser expresso da seguinte maneira:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2.2)$$

e a Estatística C de Geary, proposto por Geary (1954), definido pela expressão:

$$c = \frac{(n-1)}{2 \sum_{i=1}^n \sum_{j \neq 1}^n w_{ij}} \times \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2}. \quad (2.3)$$

Nestes métodos, a análise dos dados é de forma geral, examinando a associação espacial de uma área de forma inteira por meio de um indicador global conhecido por I de Moran Global. Segundo Rêgo e Pena (2012), o I de Moran varia em intervalo de $[-1, 1]$. Onde o valor mais próximo de -1 indica uma autocorrelação espacial negativa (dispersão perfeita), ou seja, o valor da variável de interesse numa determinada área tende a ser diferente dos valores observados em outras áreas (seus vizinhos). Já o valor próximo a 0 significa a ausência de autocorrelação espacial da variável de interesse entre as áreas (um comportamento aleatório). Por último, o valor próximo de 1 caracteriza autocorrelação espacial positiva (associação perfeita), ou seja, o valor da variável de interesse numa determinada área tende a ser igual aos seus vizinhos. Enquanto a Estatística C de Geary varia em intervalo de $[0, 2]$, onde o valor próximo de 0 significa associação perfeita, o valor próximo de 1 indica um comportamento aleatório e o valor próximo de 2 representa dispersão perfeita.

2.2.3.2 Autocorrelação Espacial Local (I de Moran Local)

Ainda que as medidas globais sejam importantes para nos fornecer uma indicação de agrupamento global dos dados, estas medidas precisam ser acrescentadas por medidas locais. Estas avaliam a autocorrelação associada com um ou alguns valores específicos para cada área Fischer e Wang (2011). O I de Moran local pode ser expresso da seguinte maneira:

$$I_i = \frac{(y_i - \bar{y}) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2.4)$$

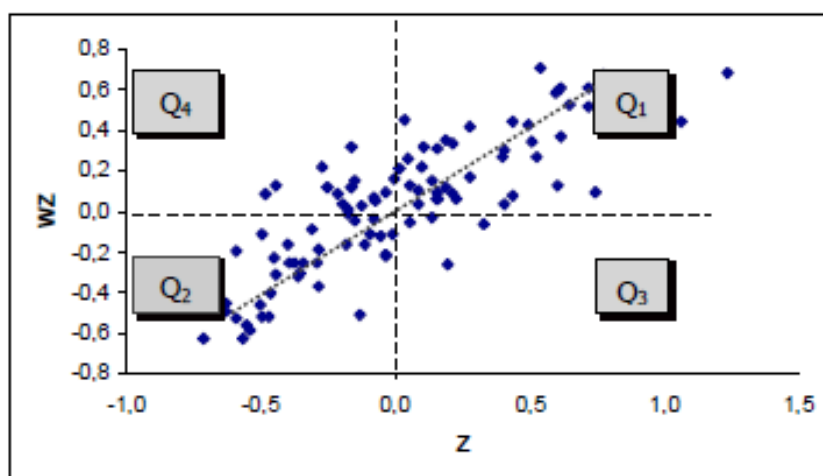
a estatística de teste pode ser interpretada da seguinte modo: valores positivos de I_i significam que existem grupos espaciais com valores parecidos (alto ou baixo) da variável em estudo. Em contrapartida, valores negativos significam que existem grupos espaciais com valores diferentes da variável entre as áreas e seus vizinhos.

2.2.4 Gráfico de Espalhamento de Moran

Essa técnica é uma maneira adicional de visualizar a dependência espacial e também uma forma de verificar o grau de similaridade Câmara et al. (2004). Colocando-se no eixo X o valor do indicador do município W_{ij} normalizado e no eixo Y o valor da estatística de Moran somente entre os vizinhos de W_{ij} . A inclinação da reta é dada pelo Índice I de Moran que é correspondente ao coeficiente de regressão. Dessa forma, o gráfico é dividido em quatro quadrantes (Q1, Q2, Q3 e Q4) que são interpretados da seguinte forma:

- **Q1 (Alto-Alto):** Indica que regiões pertencente a esse quadrante apresentam valores altos e seus vizinhos também apresentam valores acima da média para a variável em estudo (valores positivos, médias positivas).
- **Q2 (Baixo-Baixo):** Mostra que as regiões pertencente a esse quadrante apresentam baixos valores e seus vizinhos também apresentam valores baixos (valores negativos, médias negativas).
- **Q3 (Alto-Baixo):** Terceiro quadrante fornece uma região com valores alto são rodeados por regiões com valores baixos (valores positivos, médias negativas).
- **Q4 (Baixo-Alto) :** Quarto quadrante informa uma determinada região com baixos valores que são cercados por regiões com altos valores (valores negativos, médias positivas).

Figura 4 – Diagrama de Espalhamento de Moran.



Fonte: (CÂMARA et al., 2004)

Note que, as regiões localizadas nos quadrantes Q1 e Q2 indicam pontos de associação espacial positiva, no sentido que um município possui vizinhos com valores iguais. Já as regiões localizadas nos quadrantes Q3 e Q4 indicam pontos de associação espacial negativa, no sentido que há regiões possuem vizinhos com valores distintos.

2.2.5 Box Map, Lisa Map e Moran Map

- **Box Map:** É um mapa em que é extensão do gráfico de espalhamento de Moran, onde os componentes de cada quadrante do gráfico são caracterizados por uma cor particular com os seus respectivos polígonos.
- **LISA Map:** Indica os municípios/regiões que apontam correlações locais significativamente diferentes das demais, onde são especificadas nos grupos, não significativos, significativos a 5%, 1% e 0,1%.
- **Moran Map:** Similar ao LISA Map, apenas os valores significativos são apresentados, porém classificados em quatro categorias, conforme os quadrantes a que pertencem no box map, além dos não significativos.

2.2.6 Modelo de Regressão Linear

A análise de regressão linear verifica o relacionamento entre a variável dependente, também conhecida como variável resposta \mathbf{Y} e uma ou mais variáveis independentes ou explicativas (X_1, \dots, X_p). Esta relação é representada por meio de um modelo matemático, ou seja, por uma equação que associa a variável dependente \mathbf{Y} com as variáveis independentes (X_1, \dots, X_p) (RODRIGUES, 2012). De acordo com Câmara et al. (2004), se houver autocorrelação espacial, o modelo gerado deve adotar uma estrutura espacial, tendo em vista que a dependência entre essas observações possa afetar a competência de explicação do modelo. O modelo de regressão é dado por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2), \quad (2.5)$$

onde que \mathbf{Y} é uma observação da variável dependente, \mathbf{X} é uma observação sobre uma variável explicativa, $\boldsymbol{\beta}$ são os coeficientes de regressão correspondentes e $\boldsymbol{\varepsilon}$ é o erro associado as observações da variável dependente. No qual é necessário que os erros tenham uma distribuição Normal, com média 0 e variância constante σ^2 , e sejam independentes entre si, ou seja, $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$ e $Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$.

2.2.7 Modelo Espacial Autorregressivo Misto - SAR

Segundo Câmara et al. (2004), Modelo Espacial Autorregressivo Misto ("*Spatial Autoregressive – SAR*" ou ainda como "*spatial lag model*") é uma das alternativas para tratar a autocorrelação global em um modelo de regressão. Neste modelo, a informação de seus vizinhos é colocada como variável explicativa como também outras covariáveis. Deste modo, a expressão do modelo é:

$$\mathbf{Y} = \rho \mathbf{W}\mathbf{Y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.6)$$

sendo \mathbf{W} é a matriz de vizinhança, a dependência espacial apresentada pelo produto de $\mathbf{W}\mathbf{Y}$ e ρ o coeficiente espacial autorregressivo. Sendo que, quando ρ é igual a zero, quer dizer que,

não há dependência espacial para a(s) variável(éis), ou seja, as áreas vizinhas não influencia na variável dependente. Desta forma há uma objeção da ideia principal do modelo, ou seja, integrar a autocorrelação espacial como parte do modelo em estudo.

2.2.8 Modelo de Erro Espacial - SEM

Para este modelo, que também é conhecido como Modelo de Erro Espacial (*Spatial Error Model - SEM ou Conditional AutoRegressive – CAR*) tem seu entendimento mais comprometido, devido há não possuir a dependência espacial como variável e sim no erro aleatório do modelo. Sendo a expressa da seguinte forma:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \quad \varepsilon = \lambda \mathbf{W}\varepsilon + \xi, \quad (2.7)$$

em que $\mathbf{W}\varepsilon$ é a componente do erro com efeitos espaciais, λ é o coeficiente autorregressivo e ξ é a componente do erro com variância constante e não correlacionada (CÂMARA et al., 2004).

2.2.9 Modelo Espacial de Durbin - SDM

De acordo com Almeida (2012) apud Tafarelo e Francisco (2018) o modelo também chamado de (*SDM - Spatial Durbin Model*) tem alcance global, que é dado pelo multiplicador da variável resposta defasada espacialmente e também alcance local, dado pelas defasagens das variáveis explicativas. A equação é dada a seguinte forma:

$$\mathbf{Y}_i = \rho \mathbf{W}\mathbf{Y} + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\Theta + \varepsilon, \quad (2.8)$$

onde \mathbf{Y} é o vetor de variáveis respostas, \mathbf{X} é a matriz de dados, onde contém as variáveis explicativas, β é o vetor de parâmetros, \mathbf{W} é a matriz de proximidade espacial e ρ é o parâmetro espacial autorregressivo. Para esse modelo, um acréscimo de variáveis é adicionado, criando um lag espacial das variáveis explicativas, onde seria a multiplicação da matriz de proximidade espacial \mathbf{W} com os parâmetros associados a Θ (COSTA, 2017).

2.2.10 Seleção de Modelos

A escolha do modelo mais adequado é um assunto muito importante. Tendo em vista que, cada um tem suas particularidades em estimação e interpretação. No entanto, sempre há a necessidade de realizar uma análise de resíduo para verificar a qualidade do ajuste de regressão, pois é importante analisar o diagnóstico do modelo e saber se há uma quebra de pressupostos de independência.

Os estimadores e os diagnósticos comuns de regressão não levam em consideração os efeitos espaciais, assim, na regressão espacial as inferências, como por exemplo, as indicações de qualidade de ajuste, baseadas no coeficiente de determinação (R^2) não são confiáveis (RIGHETTO; TACHIBANA, 2015).

Usualmente, essa comparação entre modelos é feita utilizando o Critério de Informação de Akaike (*AIC - Akaike Information Criterion*). A estimativa do valor de AIC para um determinado modelo é dada por:

$$AIC = -2\log L(\hat{\theta}) + 2k, \quad (2.9)$$

em que $\hat{\theta}$ é a função de máxima verossimilhança e k é o número de coeficiente de regressão. Este método informa a distância entre o modelo proposto utilizando os dados e o modelo perfeito para descrever o evento em questão. Logo, quanto menor for o valor entre essas distância, menor o AIC e conseqüentemente, melhor o modelo.

3 RESULTADOS E DISCUSSÃO

3.1 Análise Descritiva para os Casos Notificados por COVID-19.

Inicialmente realizou-se uma análise descritiva dos casos notificados para verificar as medidas de tendências centrais e as medidas de dispersão. Na Tabela 1 é apresentada à análise para os meses de março a julho do ano de 2020 e na Tabela 2 para os meses de agosto a dezembro do mesmo ano. Podemos observar que para os dois primeiros meses (março e abril), 75% dos dados são iguais à zero. Percebe-se também, que ao passar dos meses a média dos casos notificados vão aumentando, tendo um aumento considerável dessa média nos meses de maio a julho e voltando a cair após esse último mês. Sendo esse aumento proporcional ao pico do crescimento semanal que ocorreu nas capitais do nordeste no mesmo período (semanas epidemiológicas 19 a 31)(XIMENES et al., 2021). Nota-se através da Tabelas 1 e 2, que de maio a dezembro temos uma assimetria positiva dos dados e o município de João Pessoa apresenta o maior número de casos notificados por COVID-19 nos dez meses em estudo.

Tabela 1 – Estatística Descritiva para os casos notificados por COVID-19 para os meses de março a julho de 2020.

Estatística	Março	Abril	Maió	Junho	Julho
Mínimo	0,00	0,00	0,00	0,00	0,00
1° Quartil	0,00	0,00	1,00	5,50	11,00
Mediana	0,00	0,00	7,00	24,00	37,00
Média	0,08	3,58	55,37	151,56	160,71
3° Quartil	0,00	0,00	24,50	86,00	107,00
Máximo	12,00	487,00	3460,00	9339,00	8032,00
Variância	0,67	1102,67	69426,13	530633,94	384348,68
Desvio padrão	0,82	33,21	263,49	728,45	619,96

Fonte: Produzido pelo autor.

Tabela 2 – Estatística Descritiva para os casos notificados por COVID-19 para os meses de agosto a dezembro de 2020.

Estatística	Agosto	Setembro	Outubro	Novembro	Dezembro
Mínimo	0,00	0,00	0,00	0,00	0,00
1° Quartil	9,50	6,00	3,00	3,00	12,00
Mediana	26,00	21,00	11,00	10,00	26,00
Média	103,08	69,47	53,31	54,28	95,04
3° Quartil	73,50	56,00	30,50	30,00	69,00
Máximo	5555,00	2807,00	3096,00	4025,00	5066,00
Variância	162741,75	49638,71	49648,36	79215,19	134947,66
Desvio Padrão	403,41	222,80	222,82	281,45	367,35

Fonte: Produzido pelo autor.

Ao observar nas Tabelas 1 e 2, notou-se que há uma dispersão considerável para os casos notificados nos meses apresentados. Desta forma, optou-se por trabalhar com as taxas de incidência, pois nos permite padronizar e fazer com que cidades com menores números de casos venham ter a mesma influência.

3.2 Cálculo da Taxa de Incidência dos Casos Notificados por COVID-19.

A incidência em epidemiologia, trás a ideia de com que força a morbidade incide em uma população num tempo determinado. De acordo com Rouquayrol e Gurgel (2021), o coeficiente de incidência é definido como a divisão entre o número de casos novos de uma determinada doença que incidem numa região num intervalo de tempo definido e o número da população expostos durante no mesmo tempo, multiplicado o resultado por potência de 10, que é base referencial da população. Desta forma, a equação é dada da seguinte maneira:

$$Tax_Incid_COVID19(TIC) = \frac{Casos_notificados}{Pop_Estimada_do_ano} \times 10^n. \quad (3.1)$$

3.3 Análise Descritiva da Taxa de Incidência dos Casos Notificados por COVID-19

Após o cálculo da taxa de incidência dos casos notificados por COVID-19, realizou-se uma nova análise descritiva dos dados com intuito de verificar as medidas de tendências centrais e de dispersão. Como se pode observar nas Tabelas 3 e 4, ao fazer o comparativo entre as médias dos 10 meses (março a dezembro) do ano de 2020, nota-se que, o mês que obtém a maior média de incidência de casos notificados por COVID-19 é o mês de julho. Após o mês de julho, nota-se também que, houve uma diminuição na média da incidência dos casos notificados, vindo a subir novamente no mês de dezembro. Da mesma forma que para os casos notificados mostrados nas Tabelas 1 e 2, porém com uma variabilidade menor. Já para o valor máximo, e o município de Ingá apresenta essa maior taxa de incidência nos meses estudados.

Tabela 3 – Estatística Descritiva para as taxas de incidência de COVID-19 nos meses de março a julho de 2020.

Estatísticas	Março	Abril	Maió	Junho	Julho
Mínimo	0	0	0	0	0
1° Quartil	0	0	2,33	10,31	19,39
Mediana	0	0	9,31	28,86	47,93
Média	0,02	0,39	17,9	50,36	67,73
3° Quartil	0	0	24,12	66,98	93,45
Máximo	3,28	5,96	171,77	278,95	331,79
Variância	0,05	1,04	663,06	3135,76	4339,66
Desvio Padrão	0,22	1,02	25,75	56	65,88

Fonte: Produzido pelo autor.

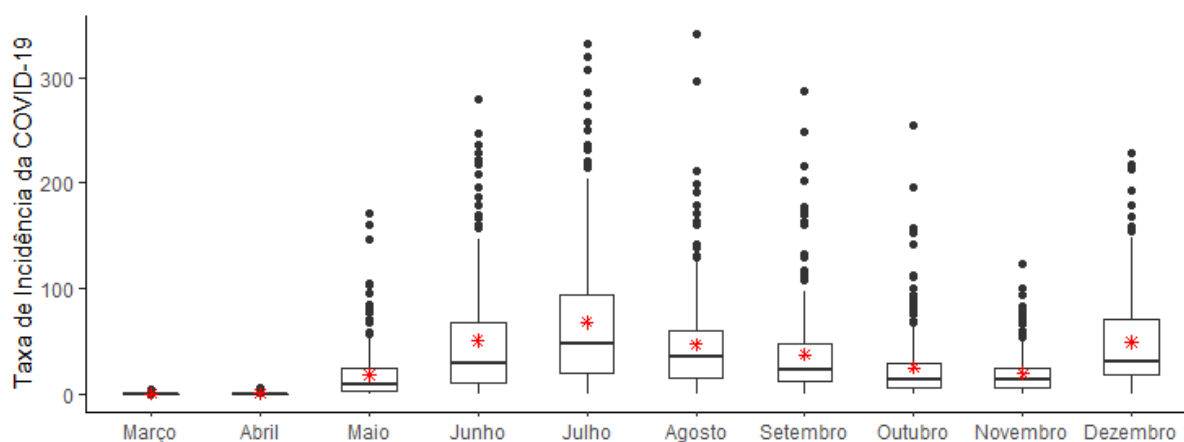
Tabela 4 – Estatística Descritiva para as taxas de incidência de COVID-19 nos meses de agosto a dezembro de 2020.

Estatística	Agosto	Setembro	Outubro	Novembro	Dezembro
Mínimo	0	0	0	0	0
1° Quartil	14,38	11,23	4,81	5,59	17,42
Mediana	35,63	22,75	13,23	13,08	31,03
Média	46,94	37,17	24,66	19,25	48,89
3° Quartil	59,76	47,14	28,69	24,12	69,88
Máximo	341,91	287,39	255,08	123,57	228,99
Variância	2351,17	1958,01	1142,41	406,79	2085,52
Desvio Padrão	48,49	44,25	33,8	20,17	45,67

Fonte: Produzido pelo autor.

A figura 5 corrobora com as Tabelas 3 e 4 apresentadas anteriormente, onde nota-se que há uma assimetria positiva para todos os meses apresentados e vemos que ao longo dos meses a média da incidência aumenta na medida que taxa de incidência aumenta. Percebe-se também que, ainda há uma dispersão entre os dados, mas ao comparar-se com os dados apresentados nas Tabelas 1 e 2, observa-se que há uma diminuição considerável nesta dispersão.

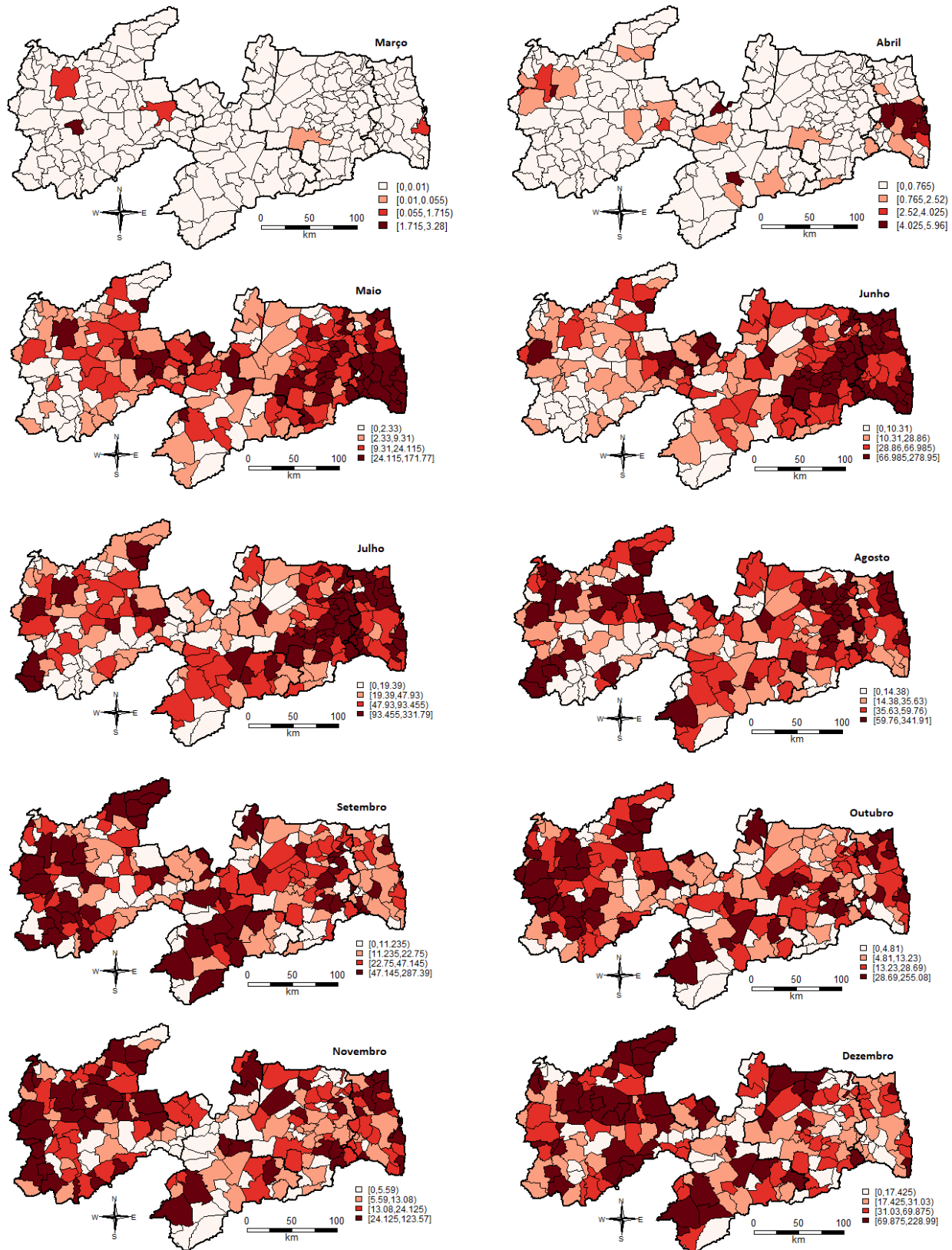
Figura 5 – Boxplot da taxa de incidência dos casos notificados por COVID-19 no estado da Paraíba para o ano de 2020.



Fonte: Produzido pelo autor.

Em seguida, foram utilizados mapas dos quartis para os meses no intuito visualizar as frequências das taxas de incidência da COVID-19 em todos os municípios do estado da Paraíba. Ao observar nas Figura 6, as maiores taxas de incidência, são as que estão apresentadas na tonalidade mais escura. Assim como as áreas com incidência zero estão representadas na tonalidade mais clara. Ainda na Figura 6 observa-se que, nos meses de março e abril são os meses que se iniciam as notificações e os que mais indicam municípios com a tonalidade mais clara. Onde que para esses dois meses mais de 75% dos dados são iguais à zero como mostrado na Tabela 3. De modo geral, nota-se que os meses de maio a agosto os municípios com as maiores taxas foram aqueles que pertencem à mata paraibana e o agreste. E de agosto a dezembro, observa-se um deslocamento da incidência para a mesorregião da Borborema e sertão paraibano.

Figura 6 – Mapa dos quartis da taxa de incidência dos casos notificados por COVID-19 no estado da Paraíba



Fonte: Produzido pelo autor.

3.4 Dependência Espacial

Em um estudo realizado por Maciel, Castro-Silva e Farias (2020) no estado do Ceará para verificar a correlação espacial entre a incidência de COVID-19 e o desenvolvimento humano nos municípios, foi utilizado o Índice I de Moran para verificar a autocorrelação espacial. Constataram que as notificações por COVID-19 no estado do Ceará é espacialmente dependente com um valor de índice de Moran significativo ao nível de 5%. Para esta monografia, obtivemos alguns Índices I de Moran significativos ao nível de 0,1% e 1% para a dependência espacial entre os municípios do estado da Paraíba. Observa-se através da Tabela 5, a existência de autocorrelação espacial positiva para os meses de abril a dezembro. Deste modo, verifica-se uma dependência na propagação da COVID-19 entre os municípios do estado da Paraíba no ano de 2020.

Tabela 5 – Tabela das estimativa do Índice I de Moran para os casos de COVID-19 no estado da Paraíba.

Variáveis	I de Moran	Variância	p-valor
T_Inc_Março	-0,0036	0,00002	0,4217
T_Inc_Abril	0,418	0,001719	2,2e-16***
T_Inc_Maio	0,3513	0,001709	2,2e-16***
T_Inc_Junho	0,4653	0,001788	2,2e-16***
T_Inc_Julho	0,4409	0,001787	2,2e-16***
T_Inc_Agosto	0,2223	0,001732	2,511e-08**
T_Inc_Setembro	0,1459	0,001738	0,0001538***
T_Inc_Outubro	0,1046	0,001697	0,004041*
T_Inc_Novembro	0,1431	0,001769	0,0002247***
T_Inc_Dezembro	0,2319	0,001793	1,174e-08***

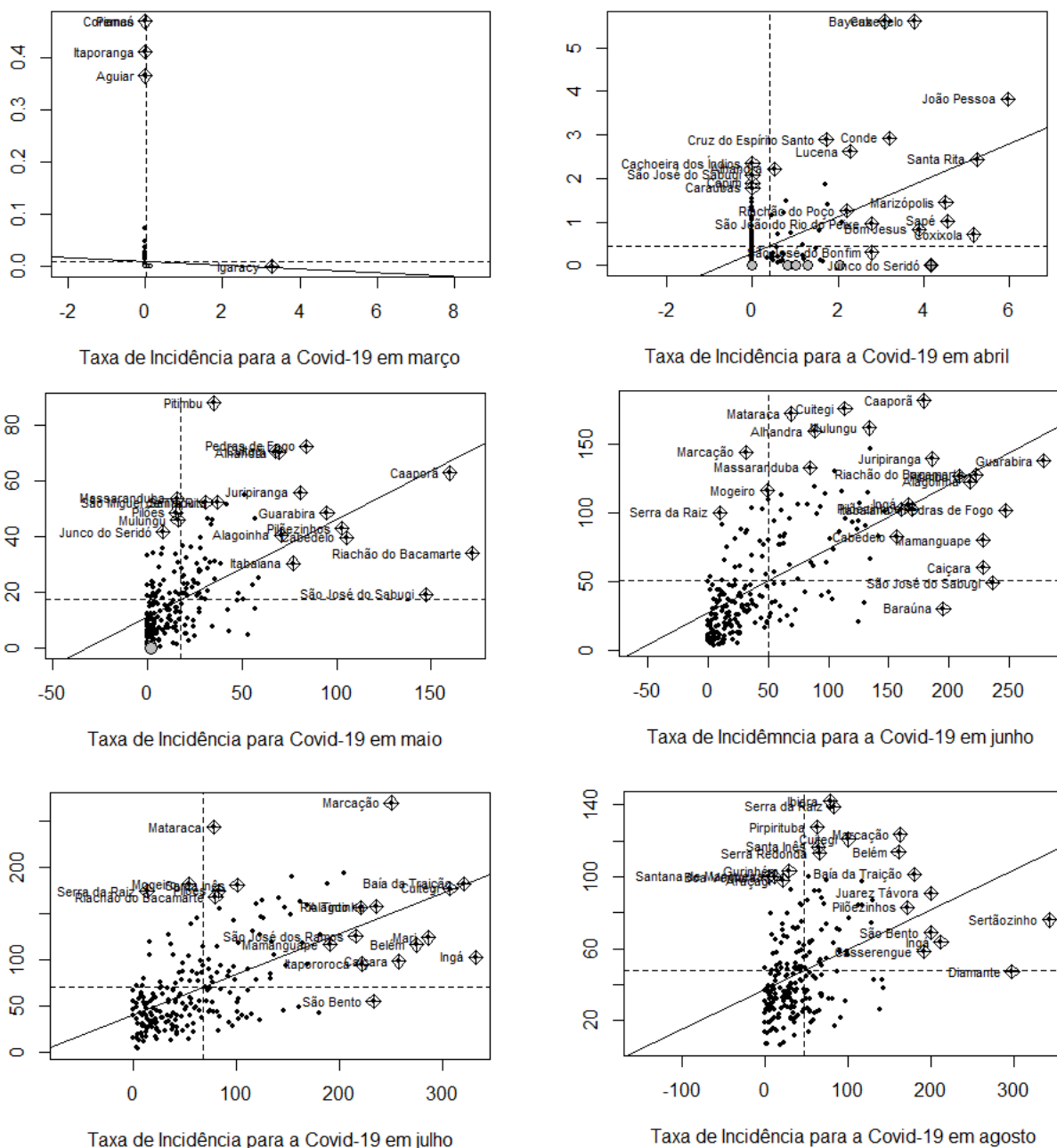
Fonte: Produzido pelo autor.

Após obter o Índice I de Moran global, tem-se valores que indicam uma autocorrelação espacial positiva entre os municípios para os meses de abril a dezembro, dando a entender que os casos de COVID-19, estão de determinado modo relacionados à sua localização e que contribui positivamente para ocorrência de COVID-19. Vale destacar, que o mês de março não apresentou I de Moran global significativo, podendo concluir como independência espacial. Observando o mapa descritivo para o mês de março há uma dispersão na distribuição dos casos de COVID-19, justificando assim, o I de Moran não significativo. No entanto, antes de levar em consideração a ausência de dependência espacial pelo I de Moran global, pode-se utilizar o I de Moran local e o Lisa Map para confirmar essa ausência. Já que o Moran global pode ser influenciado pela proximidade ou distância e o Moran local, leva em consideração os seus vizinhos.

Ao observar o gráfico de espalhamento de Moran, na Figura 7, podemos perceber que há municípios que estão no Q1 e no Q3. Todos os municípios presente nestes quadrantes tem valores positivos, ou seja, esses municípios apresentam valores positivos e influenciam seus vizinhos a terem valores iguais aos seus, e assim, elevando essa taxa de incidência. No ponto de vista clínico, seriam municípios que devem ter uma maior atenção por influenciar seus vizinhos

ao aumento da taxa de COVID-19. Nota-se que no mês de março, há um município presente no Q3 que para esse mês foi o municípios que obteve a maior incidência e tende a influenciar seus vizinhos no aumento dessa incidência.

Figura 7 – Espalhamento de Moran para os casos notificados por COVID-19 nos municípios do estado da Paraíba de março a agosto.



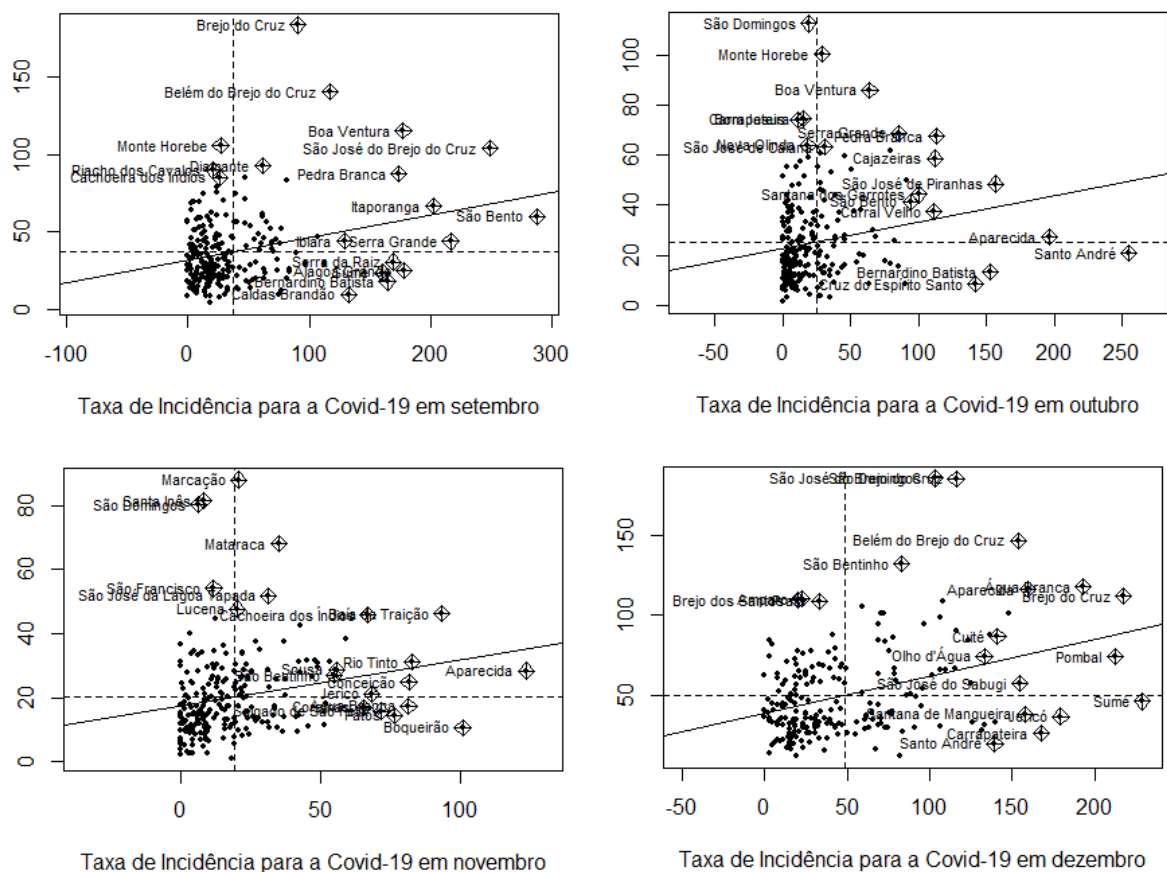
Fonte: Produzido pelo autor.

Ainda observando a Figura 7, pode-se notar que entre os meses de abril a agosto, os municípios presente no Q1 e Q3, ou seja, os municípios que apresentam valores positivos, podendo assim, influenciar seus vizinhos a terem um aumento na taxa de incidência, estão localizados na mata paraibana, como João Pessoa, Cabedelo, Baía da Traição, Lucena, Pitimbu, entre outros. E

no agreste, como Juripiranga, Pedras de fogo, Itabaiana, Alagoinha, etc. Similarmente este fato pode ser observado nas Figura 5, que mostra uma maior concentração de taxas de incidência de COVID-19 nessas mesorregiões entre os meses. O mês de julho, contido nesse intervalo, é o mês entre os 10 meses analisados que apresenta a maior média das taxas de incidência no ano de 2020 para os casos notificados por COVID-19. No entanto, não é o mês que apresenta o maior número de municípios no Q1. Sendo esse, o mês de junho por conta do aumento da média de incidência.

Na Figura 8 temos os meses de setembro, outubro, novembro e dezembro. Onde que, os municípios com as maiores taxas estão localizados na borborema e sertão. Em setembro, os municípios presente no Q1 e Q3 influenciam seus vizinhos, são Boa Ventura, Pedra Branca, Brejo do Cruz, entre outros do sertão do estado. Caldas Brandão e Alagoa Grande no agreste. Já no mês de outubro, Q1 é composto por São José de Piranhas, Boa ventura, entre outros do sertão. E no Q3, Bernadino Batista no sertão e Santo André na Borborema.

Figura 8 – Espalhamento de Moran para os casos notificados por COVID-19 nos municípios do estado da Paraíba de setembro a dezembro.



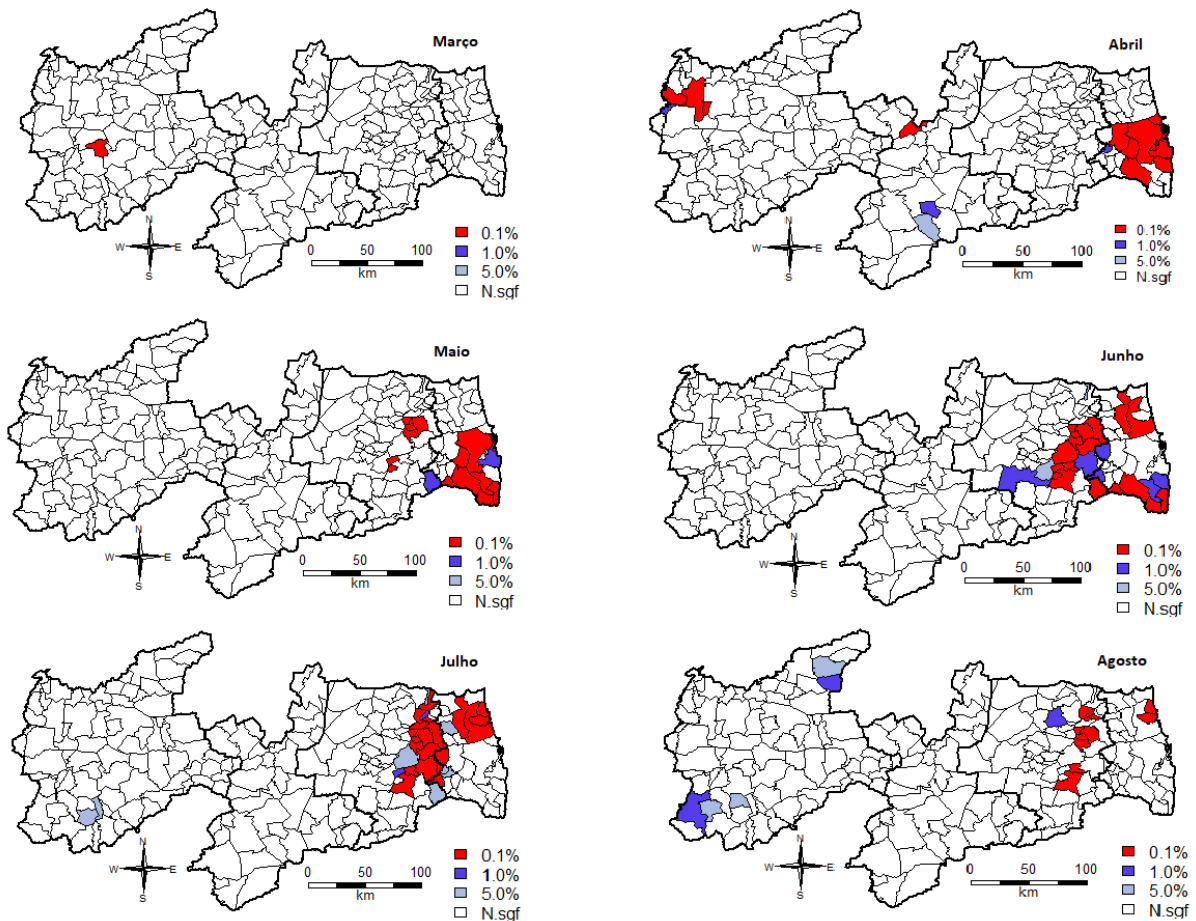
Fonte: Produzido pelo autor.

Em novembro, os municípios que influenciam seus vizinhos para o aumento da taxa alguns deles são, Mataraca, Rio Tinto, Marcação e Baía da Traição na mata paraibana, Cachoeira dos Índios, Conceição e Aparecida, que são situados no sertão do estado. Já para o mês de

dezembro, temos alguns municípios, como São Bentinho, Belém do Brejo do Cruz, Aparecida, Brejo do Cruz, Pombal e Água Branca, que estão no sertão paraibano e do agreste o município de Cuité. Nos meses de novembro e dezembro, percebe-se que grande parte dos municípios que pertencem ao Q1, estão situados no sertão do estado. Corroborando ainda mais com a Figura 5 que mostra que a partir do mês de agosto conseguimos ver uma migração dos casos da região litorânea para o sertão.

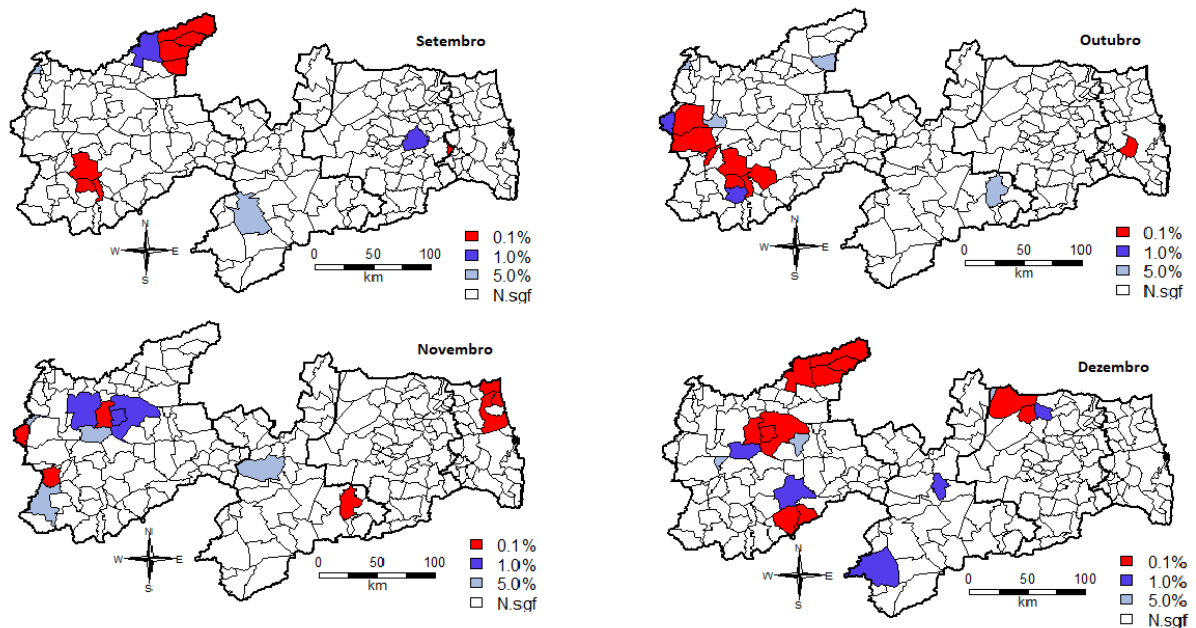
Analisando as imagens nas Figura 9 e 10, contendo o Lisa Map para os meses de março a dezembro de 2020, é possível identificar as áreas que influencia de modo local, com base na significância que varia de não significativo, que são as áreas em branco, significativo a 5% na tonalidade azul claro, 1% áreas em azul escuro e 0,1% que são áreas com dependência local altamente significativa, que sem encontram em vermelho. Tudo isso deduzido pelo p-valor.

Figura 9 – Lisa Map das taxas dos casos notificados por COVID-19 nos municípios do estado da Paraíba de março a junho.



Fonte: Produzido pelo autor.

Figura 10 – Lisa Map das taxas dos casos notificados por COVID-19 nos municípios do estado da Paraíba de julho a dezembro.

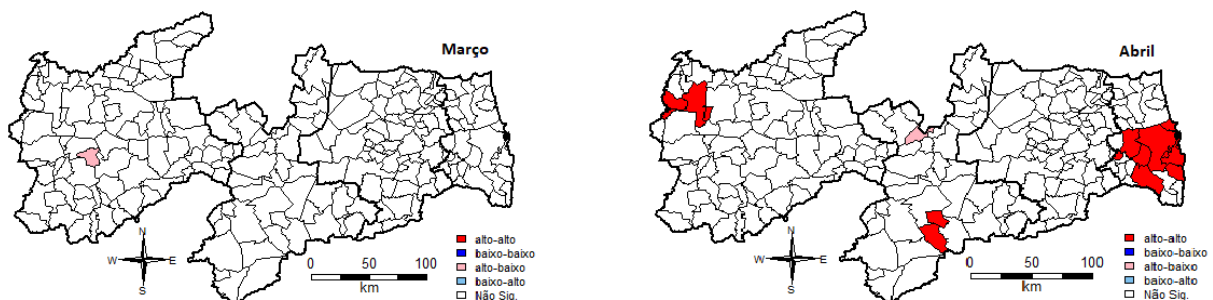


Fonte: Produzido pelo autor.

Ainda com base nas Figuras 9 e 10, nota-se que há uma alteração nas mesorregiões que apresentam autocorrelação local positiva durante os meses do ano de 2020. No entanto, percebe-se que a grande maioria desses municípios que apontam autocorrelação local ao longo dos meses estão localizados na mata e agreste paraibano. Vindo a modificar a partir do mês de setembro, em que a significância local vai migrando para a mesorregião o sertão paraibano.

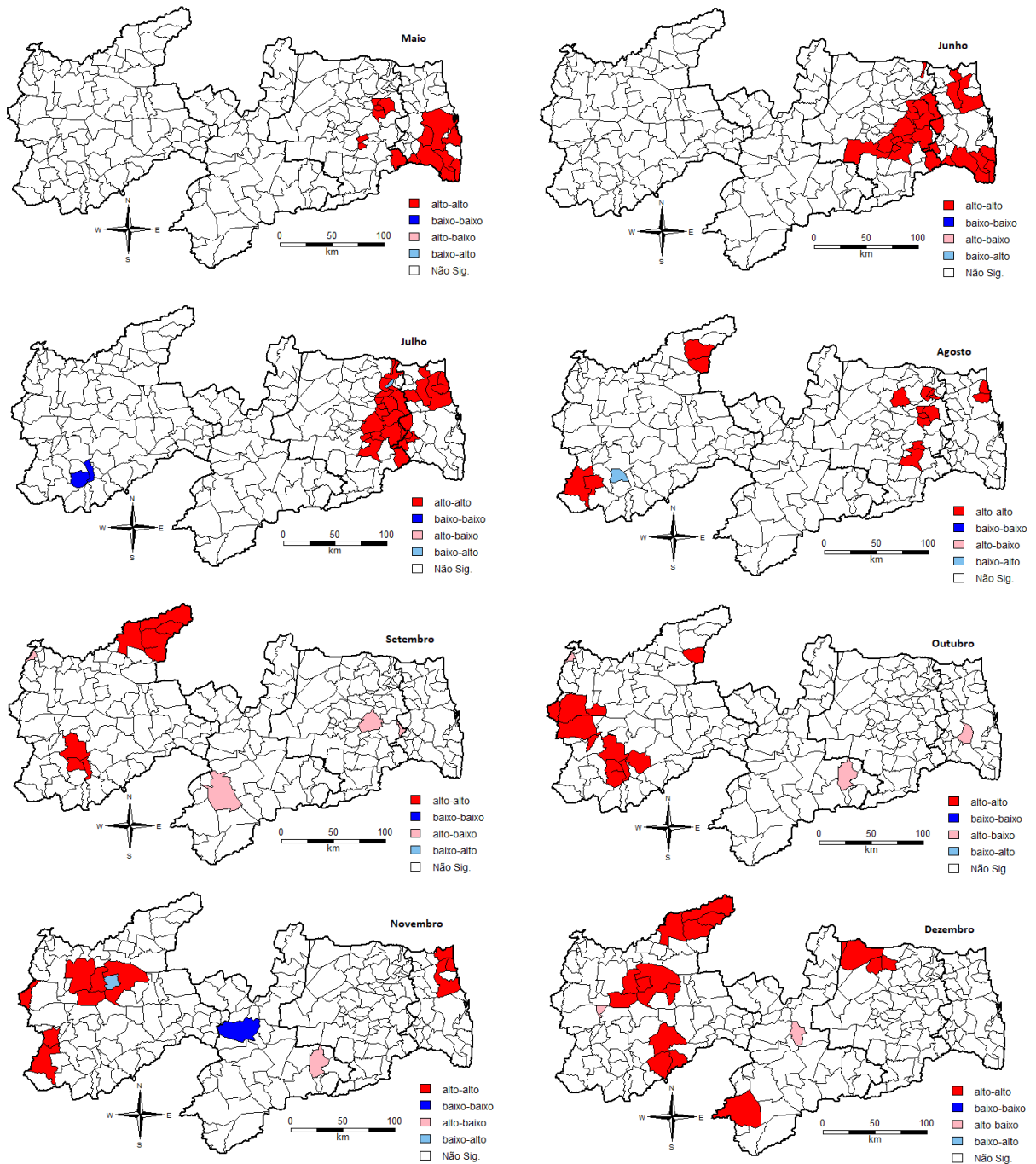
Um estudo proposto por Cavalcante e Abreu (2020) com o objetivo de descrever a distribuição espacial dos primeiros casos e óbitos confirmados de COVID-19 do Rio de Janeiro, constatou que, existem municípios com alta taxa de incidência e está rodeado de municípios que também tem uma alta incidência. Nesta monografia, também podemos observar que o mesmo acontece. Observando as Figuras 11 e 12, há regiões que precisam de um cuidado maior para que se possam controlar o aumento taxa de incidência.

Figura 11 – Moran Map das taxas notificados por COVID-19 nos municípios do estado da Paraíba.



Fonte: Produzido pelo autor.

Figura 12 – Moran Map das taxas notificadas por COVID-19 nos municípios do estado da Paraíba.



Fonte: Produzido pelo autor.

Por meio das Figuras 11 e 12, pode-se notar que o Moran Map evidenciou para os meses de abril a dezembro aglomerações com padrão alto-alto nas mesorregiões da Mata paraibana, Agreste, Borborema e Sertão do estado. No entanto, de abril a julho os municípios que mais apresentam esse padrão alto-alto são os municípios que estão presentes na mata e agreste paraibano. Como, João pessoa, Campina Grande, Santa Rita, Cabedelo, entre outros. No mês de agosto, temos uma maior quantidade de municípios que apresenta este padrão, como Ingá,

Juarez Távora, entre outros presentes no agreste. Já para o sertão, são Conceição e Ibiara. Tendo em vista que, nos dez meses analisados houve também municípios que apresentaram padrão baixo-baixo, como Curral Velho e Pedra Branca no mês de julho e localizados no sertão do estado. E em novembro Taperoá presente na mesorregião da Borborema.

De acordo com um estudo na china proposto por Zheng et al. (2020), indicou que o transporte público se tornou um veículo de disseminação de casos importados de COVID-19 para regiões nas quais ainda não havia casos. Já na Paraíba, devido ao primeiro caso notificado ter sido na capital do estado, podendo-se também haver uma disseminação por meio do transporte já que a cidade tem um volume de migração/emigração de outros municípios do estado por está situada na parte litorânea, ter um grande centro de comércio, universidades e uma busca de emprego considerável. Desta forma, justificando esses *clusters* com padrões alto-alto nas mesorregiões da mata e agreste paraibano no primeiro semestre do ano. Conforme Chen et al. (2020), o estudo realizado na China, que tinha como objetivo estudar a distribuição de casos e da sua emigração populacional de Wuhan evidenciou que a população que emigrou desta cidade foi a fonte de infecção primária para as demais no país.

3.5 Análise de Regressão Espacial

Da mesma maneira que Oliveira (2021) utilizou a modelagem de dados de área para verificar quais modelos propostos conseguem explicar melhor através das variáveis explicativas o número de casos de COVID-19 nos municípios do Estado do Rio de Janeiro, nesta seção, serão apresentados alguns ajustes dos modelos propostos na metodologia, onde tem-se os modelos espaciais - que levam em consideração os efeitos espaciais - e o modelo de regressão linear. Que por sua vez, serão analisados qual o melhor modelo, através do AIC, que esclarece melhor a taxa de incidência para a COVID-19 no estado da Paraíba nos meses em estudo. Para os meses de março e setembro não contém tabelas, pois esses meses não foram significativos.

Como especificado na seção 2.1, para a modelagem, foram utilizados dados referentes ao IDHM (é uma medida resumida do progresso em longo prazo em três dimensões básicas do desenvolvimento humano: renda, educação e saúde, variando de 0 a 1), RDPC (resultado da soma da renda recebida por cada morador, dividido pelo total de moradores do domicílio), COB_{AB}^2 (porcentagem das pessoas que procuram os serviços públicos de saúde), T_ANALF18M (taxa de analfabetismo na população de 18 anos ou mais é um dos indicadores que são utilizados para definir o status de um país na comunidade mundial) e DENSDEMOG2017 (um índice demográfico que permite avaliar a distribuição da população em um dado território ($(Pop_{Abs}/area) * 2017$)).

3.5.1 Modelagem para os dados Taxa Incidência da COVID-19 nos meses de abril e maio de 2020.

Diante dos resultados apresentados nas Tabelas 6 e 7, observa-se que o melhor modelo que se ajustou aos dados para os meses de abril e maio foi o Modelo Espacial Autorregressivo Misto

Tabela 6 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de abril de 2020.

Abril - 2020				
Covariáveis	RL (p-valor)	SAR(p-valor)	SEM(p-valor)	SDM(p-valor)
Intercepto	1,274($\leq 0,05$)	1,033($\leq 0,05$)	1,385($\leq 0,05$)	-
DENSDEMOG2017	0,001($\leq 0,05$)	0,001($\leq 0,05$)	0,001($\leq 0,05$)	-
RDPC	-	-	-	0,0018($\leq 0,05$)
T_ANALF18M	-0,031($\leq 0,05$)	-0,027($\leq 0,05$)	-0,034($\leq 0,05$)	-0,027($\leq 0,05$)
COB_AB ²	-	-	-	-
IDHM	-	-	-	-
ρ	-	0,355($\leq 0,05$)	0,379($\leq 0,05$)	0,386($\leq 0,05$)
AIC	581,41	563,95	566,1	565,78

Tabela 7 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de maio de 2020.

Maio - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	-	-6,841 (0,13)	-	-
DENSDEMOG2017	-	-	-	-
RDPC	0,062($\leq 0,05$)	0,051($\leq 0,05$)	0,058($\leq 0,05$)	0,053($\leq 0,05$)
T_ANALF18M	-	-	-	-
COB_AB ²	-	-	-	-
IDHM	-	-	-	-
ρ	-	0,607($\leq 0,05$)	0,603($\leq 0,05$)	0,608($\leq 0,05$)
AIC	2075,3	2020,5	2021,4	2022,5

(SAR), pois apresenta o menor AIC. Tendo em vista que, ao comparar modelos de regressão tradicional e modelos de regressão que consideram a autorrelação espacial, uma medida como o R^2 não é mais confiável. Desta forma, a equação ajustada para o mês de abril é dada por:

$$\widehat{TIC} = 1,033 + 0,001DENSDEMOG17 - 0,027T_ANALF18M, \quad (3.2)$$

ou seja, à medida que aumentou a unidade na DENSDEMOG17 houve um aumento de 0,001 em média na TIC e à medida que aumentou a unidade na T_ANALF18M, houve uma diminuição em média de 0,027 na TIC. Também sendo significativo a 5%, o mês de maio tem a seguinte equação ajustada:

$$\widehat{TIC} = -6,841 + 0,051RPDC, \quad (3.3)$$

sendo assim, à medida que aumentou a unidade na RDPC houve um acréscimo de 0,051 em média na TIC.

3.5.2 Modelagem para os dados Taxa Incidência da COVID-19 nos meses de junho de 2020.

Por meio da Tabela 8 nota-se que, para o mês de junho, temos que o modelo que obteve o melhor ajuste para os dados foi o Modelo de Erro Espacial (SEM), que por sua vez, teve o menor valor de AIC. A equação para o mês de junho é dada da seguinte forma:

Tabela 8 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de junho de 2020.

Junho - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	47,274(≤ 0,05)	-	12,433(≤ 0,34)	-
DENSDEMOG2017	0,029(≤ 0,05)	-	-	-
RDPC	-	0,107(≤ 0,05)	0,143(≤ 0,05)	0,147(≤ 0,05)
T_ANALF18M	-	-	-	-
COB_AB ²	-	-	-	-
IDHM	-	-	-	-
ρ	-	0,666(≤ 0,05)	0,672(≤ 0,05)	0,669(≤ 0,05)
AIC	2425,6	2337,5	2334,2	2336

$$\widehat{TIC} = 12,433 + 0,143RDPC, \quad (3.4)$$

deste modo, à medida que aumentou a unidade na RDPC houve uma acréscimo em média de 0,143 na TIC.

3.5.3 Modelagem para os dados Taxa Incidência da COVID-19 nos meses de julho e agosto de 2020.

Tabela 9 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de julho de 2020.

Julho - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	-	75,608(≤ 0,05)	-	212,938 (≤ 0,05)
DENSDEMOG2017	-	-	-	-
RDPC	0,14(≤ 0,05)	-	-	-
T_ANALF18M	2,129(≤ 0,05)	-	-	-
COB_AB ²	-	-0,527(≤ 0,05)	-0,473(≤ 0,05)	-
IDHM	-	-	312,7(≤ 0,05)	380,9(≤ 0,05)
ρ	-	0,618(≤ 0,05)	0,662(≤ 0,05)	0,611(≤ 0,05)
AIC	2503,9	2430,8	2424,6	2419,6

Com base nas Tabelas 9 e 10, temos que, para os meses de julho e agosto, o melhor modelo que conseguiu o melhor ajuste para os dados foi o Modelo Espacial de Durbin (SDM) que obteve o menor AIC. No mês de julho temos a seguinte equação:

$$\widehat{TIC} = 212,938 + 380,9IDHM, \quad (3.5)$$

então, à medida que aumentou a unidade no IDHM houve uma aumento de 380,9 em média na TIC. Para o mês de agosto, a equação é dada por:

$$\widehat{TIC} = 107,814 + 290,5IDHM, \quad (3.6)$$

Tabela 10 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de agosto de 2020.

Agosto - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	-258,872($\leq 0,05$)	-217,27($\leq 0,05$)	-	107,814($\leq 0,15$)
DENSDEMOG2017	-	-	-	-
RDPC	-	-	-	-
T_ANALF18M	2,189($\leq 0,05$)	1,444($\leq 0,05$)	-	-
COB_AB ²	-	-	-	-
IDHM	401,7($\leq 0,05$)	337,6($\leq 0,05$)	232,2($\leq 0,05$)	290,5($\leq 0,05$)
ρ	-	0,415($\leq 0,05$)	0,479($\leq 0,05$)	0,446($\leq 0,05$)
AIC	2360,4	2342	2341,3	2338,5

portanto, à medida que aumentou a unidade no IDHM houve um aumento de 290,5 em média na TIC.

3.5.4 Modelagem para os dados Taxa Incidência da COVID-19 nos meses de outubro, novembro e dezembro de 2020.

Tabela 11 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de outubro de 2020.

Outubro - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	59,951($\leq 0,05$)	55,751($\leq 0,05$)	59,609($\leq 0,05$)	-
DENSDEMOG2017	-	-	-	-
RDPC	0,048($\leq 0,05$)	0,045($\leq 0,05$)	0,049($\leq 0,05$)	-
T_ANALF18M	-	-	-	-
COB_AB ²	-0,502($\leq 0,05$)	-0,506($\leq 0,05$)	-0,504($\leq 0,05$)	-0,518($\leq 0,05$)
IDHM	-	-	-	-
ρ	-	0,215($\leq 0,05$)	0,222($\leq 0,05$)	0,225($\leq 0,05$)
AIC	2196,3	2193,1	2192,9	2196,7

Tabela 12 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de novembro de 2020.

Novembro - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	-	-	3,20($\leq 0,47$)	-
DENSDEMOG2017	-	-	-	-
RDPC	0,055($\leq 0,05$)	0,053($\leq 0,05$)	0,057($\leq 0,05$)	0,058($\leq 0,05$)
T_ANALF18M	-	-	-	-
COB_AB ²	-	-	-	-
IDHM	-	-	-	-
ρ	-	0,219($\leq 0,05$)	0,2274($\leq 0,05$)	0,2275($\leq 0,05$)
AIC	1963,1	1958	1957,8	1959,8

Tabela 13 – Tabela dos coeficientes estimados dos modelos de regressão para o mês de dezembro de 2020.

Dezembro - 2020				
Covariáveis	RL (p-valor)	SAR (p-valor)	SEM (p-valor)	SDM (p-valor)
Intercepto	101,122(≤ 0,05)	83,062(≤ 0,05)	-12,310(≤ 0,82)	-
DENSDEMOG2017	-	-	-	-
RDPC	-	-	-	-
T_ANALF18M	-	-	-	-
COB_AB ²	-0,538(≤ 0,05)	-0,557(≤ 0,05)	-0,502(≤ 0,05)	-0,504(0,05)
IDHM	-	-	186,328(≤ 0,05)	218,595(≤ 0,05)
ρ	-	0,400(≤ 0,05)	0,419(≤ 0,05)	0,422(≤ 0,05)
AIC	2336,6	2315,5	2312,6	2315,5

Para os três últimos meses em estudo, obtive-se que o melhor modelo que se adequou aos dados foi o Modelo de Erro Espacial (SEM), pois apresentam o menor valor de AIC. No mês de outubro, temos a equação ajustada da seguinte maneira:

$$\widehat{TIC} = 59,609 + 0,049RDPC - 0,504COB_AB^2, \quad (3.7)$$

assim, à medida que aumenta a unidade RPDC tende à aumentar em média 0,049 na TIC e à medida que acrescenta uma unidade em COB_AB^2 diminui em média 0,504 na TIC. Para o mês de novembro a equação ajustada é dada por:

$$\widehat{TIC} = 3,20 + 0,057RDPC, \quad (3.8)$$

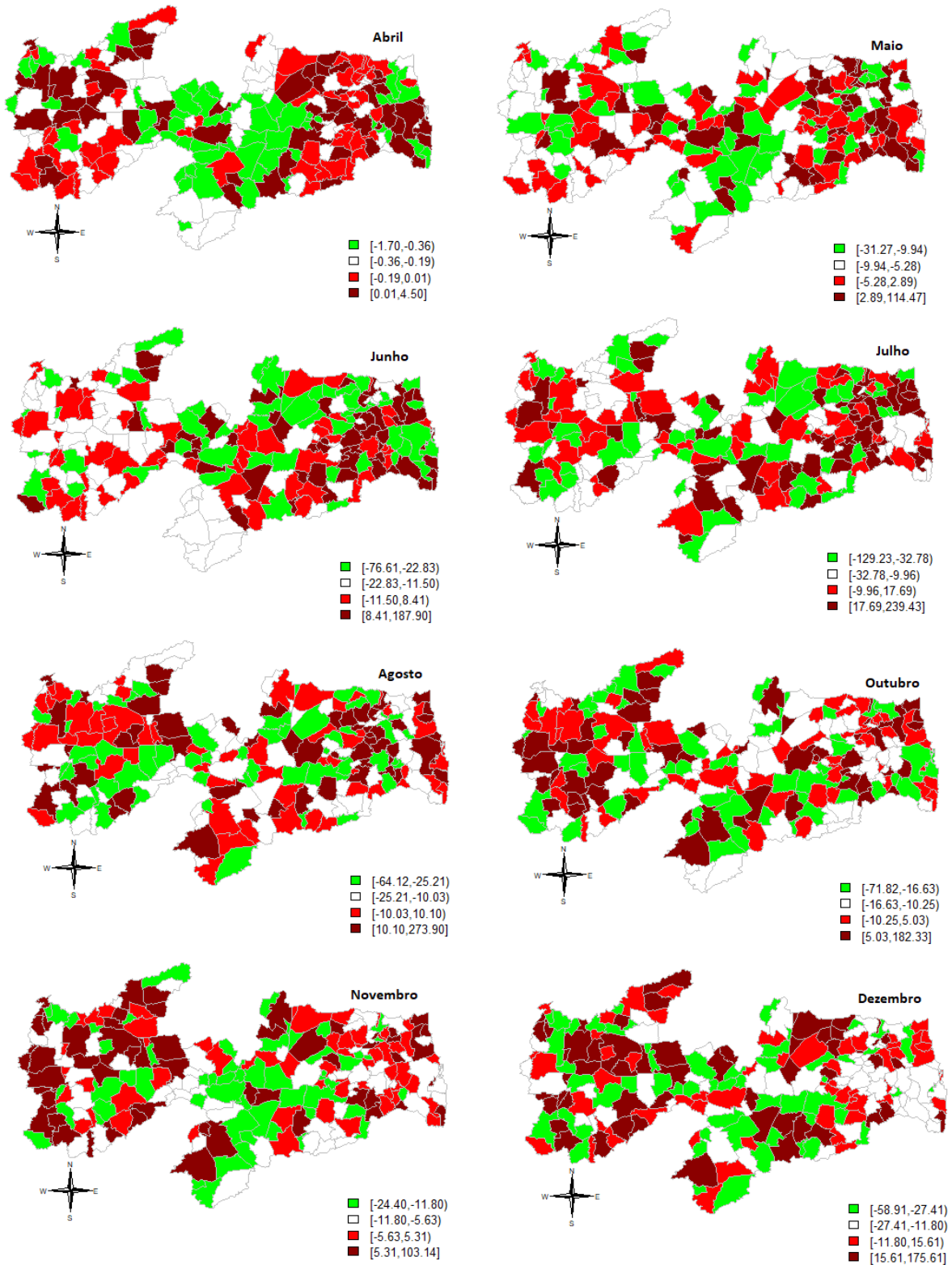
dessa maneira, à medida à medida que aumenta a unidade RPDC aumenta em média 0,057 na TIC. A equação da reta é da seguinte forma:

$$\widehat{TIC} = -12,310 - 0,502COB_AB^2 + 186,328IDHM, \quad (3.9)$$

desta forma, à medida que aumenta a unidade COB_AB^2 há uma diminuição em média de 0,502 na TIC e à medida que aumenta uma unidade do IDHM, tem um acréscimo de 186,328 em média na TIC.

Com base nas Figuras 13 é possível observar os erros dos modelos que foram mais apropriados para os oito meses do ano de 2020 que foram significativos. Destaca-se que no modelo de regressão espacial não há uma limitação de que os valores dos resíduos tenham que permanecer contido no intervalo [-2,2]. Além disso, os mapas de resíduos podem também nos fornecer se há ou não presença de dependência espacial através de valores positivos e negativos alocados próximos uns aos outros. Sendo os menores valores representados pela cor verde, os próximos a zero representados pela cor branca e os maiores valores estão representados pelos tons em vermelho.

Figura 13 – Mapa dos resíduos para modelos de regressão espacial para a taxa de incidência da COVID-19 no ano de 2020 .



Fonte: Produzido pelo autor.

4 CONCLUSÃO

Portanto, diante do que foi visto, podemos afirmar que a análise espacial para dados de área e a modelagem espacial, pode aumentar de forma considerável o entendimento e compreensão de padrões espaciais. Pois no decorrer do estudo, foi encontrada uma autocorrelação espacial para nove meses em dez estados no ano de 2020 indicando uma autocorrelação espacial na taxa de COVID-19 nos municípios do estado da Paraíba, e por meio dos mapas apresentados, pode-se notar que alguns municípios são preocupantes, pois tem uma alta incidência desses casos notificados por COVID-19. A modelagem espacial é de grande importância, visto que para oito entre os dez meses fatores socioeconômicos vem influenciando no aumento e também na diminuição de contágio da doença.

Sendo assim, de total importância para os órgãos dirigente do estado da Paraíba, a obrigação de uma estruturação de programas para o combate a COVID-19 que busquem promover a prevenção da saúde por meio de normas constantes de orientação quanto às medidas de tratamento, autocuidado e segurança, fazendo assim, com que haja uma diminuição nas chances de transmissão da doença para a população residente do estado.

REFERÊNCIAS

- ALMEIDA, E. S. d.; PEROBELLI, F. S.; FERREIRA, P. G. C. Existe convergência espacial da produtividade agrícola no Brasil? *Revista de Economia e Sociologia Rural*, 2008. 18
- ALMEIDA, M. E. et al. O protagonismo extensionista na pandemia em 2020: Inovar e resistir! *Expressa Extensão*, 2021. 13
- ANDRADE, A. L. et al. Introdução à estatística espacial para a saúde pública. 2007. 14
- CÂMARA, G. et al. Análise espacial de áreas. *Análise espacial de dados geográficos. Brasília: EMBRAPA*, p. 157–82, 2004. 16, 17, 19, 20, 21
- CAVALCANTE, J. R.; ABREU, A. d. J. L. d. Covid-19 no município do Rio de Janeiro: análise espacial da ocorrência dos primeiros casos e óbitos confirmados. *Epidemiologia e Serviços de Saúde*, v. 29, 2020. 31
- CHEN, Z.-L. et al. Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China. *Chinese Medical Journal*, Wolters Kluwer Health, v. 133, n. 9, p. 1044, 2020. 33
- COSTA, A. O. Modelagem espacial no mapeamento de doenças transmitidas pelo *Aedes aegypti* no estado da Paraíba. 2017. 16, 21
- COSTA, E. de O.; TAVARES, M. de O.; CLEMENTE, C. M. A COVID-19 na Paraíba: Análise epidemiológica e social. *Diálogos em Saúde*, v. 3, n. 2, 2021. 14
- CRODA, J. H. R.; GARCIA, L. P. Resposta imediata da vigilância em saúde à epidemia da COVID-19. 2020. 13
- FEHR, A. R.; PERLMAN, S. Coronaviruses: an overview of their replication and pathogenesis. *Coronaviruses*, p. 1–23, 2015. 13
- FISCHER, M. M.; WANG, J. *Spatial data analysis: models, methods and techniques*. [S.l.]: Springer Science & Business Media, 2011. 18
- GEARY, R. C. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, v. 5, n. 3, p. 115–146, 1954. 18
- GRIFITH, D. *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization* New York. [S.l.]: Springer-Verlag, 2003. 17
- IBGE. 2020. Disponível em: <<https://www.ibge.gov.br>>. 15
- LANA, R. M. et al. Emergência do novo coronavírus (SARS-CoV-2) e o papel de uma vigilância nacional em saúde oportuna e efetiva. *Cadernos de Saúde Pública*, v. 36, p. e00019620, 2020. 13
- MACIEL, J. A. C.; CASTRO-SILVA, I. I.; FARIAS, M. R. d. F. Análise inicial da correlação espacial entre a incidência de COVID-19 e o desenvolvimento humano nos municípios do estado do Ceará no Brasil. *Revista Brasileira de Epidemiologia*, SciELO Public Health, v. 23, p. e200057, 2020. 27
- MORAN, P. A. Notes on continuous stochastic phenomena. *Biometrika*, JSTOR, v. 37, n. 1/2, p. 17–23, 1950. 18

- OLIVEIRA, G. T. de. Análise espacial dos casos de covid-19 no estado do rio de janeiro. 2021. 33
- RÊGO, C. V.; PENA, M. G. Análise dos modelos de regressão espacial sar, sem e sac. 2012. 18
- RIGHETTO, A. J.; TACHIBANA, V. M. Análise de regressão linear: abordagem tradicional e espacial em um estudo de caso. *Revista Brasileira de Qualidade de Vida*, v. 7, n. 3, p. 169–179, 2015. 21
- RODRIGUES, S. C. A. *Modelo de regressão linear e suas aplicações*. Tese (Doutorado) — Universidade da Beira Interior, 2012. 20
- ROUQUAYROL, M. Z.; GURGEL, M. *Rouquayrol: epidemiologia e saúde*. [S.l.]: Medbook, 2021. 24
- SILVA, A. dos S. et al. Atuação do serviço de controle de infecção hospitalar em hospital público da paraíba no enfrentamento à covid-19: Relato de experiência. *ENVIRONMENTAL SMOKE*, v. 4, n. 1, p. 32–43, 2021. 14
- SILVA, D. S. da C.; SANTOS, M. B. dos; SOARES, M. J. N. Impactos causados pela covid-19: um estudo preliminar. *Revista Brasileira De Educação Ambiental (RevBEA)*, v. 15, n. 4, p. 128–147, 2020. 13
- SINGHAL, T. A review of coronavirus disease-2019 (covid-19). *The indian journal of pediatrics*, v. 87, n. 4, p. 281–286, 2020. 13
- TAFARELO, F.; FRANCISCO, E. de R. Evasão escolar no ensino médio seriado brasileiro através de modelos de regressão espacial: metodologia potencial para aplicação no varejo. In: *CLAV 2018*. [S.l.: s.n.], 2018. 21
- TEAM, R. C. et al. The r project for statistical computing. *Available from:* < *Available from: <https://www.r-project.org/>*>. *Accessed: Mar, 2020*. 15
- XIMENES, R. A. d. A. et al. Covid-19 no nordeste do brasil: entre o lockdown e o relaxamento das medidas de distanciamento social. *Ciência & Saúde Coletiva*, SciELO Public Health, v. 26, p. 1441–1456, 2021. 23
- ZHENG, R. et al. Spatial transmission of covid-19 via public and private transportation in china. *Travel medicine and infectious disease*, v. 23, p. 101626, 2020. 33
- ZHU, N. et al. A novel coronavirus from patients with pneumonia in china. *New England journal of medicine*, 2020. 13