



Cadernos

n° 88

Modelos espaço-temporais para avaliação do risco da Covid-19 nos municípios brasileiros

Anderson Castro Soares
de Oliveira

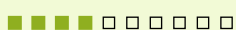
Lia Hanna Martins Morita



Coleção: Covid-19 Fast Track
■■■■■■■■■■

Coleção:

Covid-19 Fast Track



Modelos espaço-temporais para avaliação do risco da Covid-19 nos municípios brasileiros

Autor

Anderson Castro Soares de Oliveira

Coautora

Lia Hanna Martins Morita

Parecerista convidado

Pedro Masson Sesconetto Souza

Este caderno é resultado dos conhecimentos gerados pelas pesquisas realizadas no âmbito do **Programa Cátedras Brasil**, desenvolvido com o objetivo de selecionar projetos de pesquisa que gerem subsídios para o entendimento ou enfrentamento à pandemia de Covid-19 pela Administração Pública. A presente publicação é uma das entregas previstas no Edital nº 69 de 2020.

Conheça os autores



**Anderson Castro Soares
de Oliveira**
Autor



É professor Associado do Departamento de Estatística da Universidade Federal de Mato Grosso. Possui doutorado em Estatística e Experimentação Agropecuária pela Universidade Federal de Lavras. Tem experiência na área de Probabilidade e Estatística, atuando principalmente com: dinâmica de populações, estatística em saúde, mineração de dados e de texto.



**Lia Hanna Martins
Morita**
Coautora



É docente do departamento de estatística da Universidade Federal de Mato Grosso. Possui graduação em Matemática (2002), mestrado em Estatística (2005) e doutorado em Estatística pela Universidade Federal de São Carlos (2017). Tem experiência em Probabilidade e Estatística, com ênfase em análise de sobrevivência e inferência Bayesiana.

Expediente



Escola Nacional de Administração Pública – Enap

Presidente

Diogo Costa

Diretora-Executiva

Rebeca Loureiro de Brito

Diretora de Altos Estudos

Diana Coutinho

Diretor de Educação Executiva

Rodrigo Torres

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretora de Inovação

Bruna Santos

Diretora de Gestão Interna

Alana Regina Biagi Lisboa

Revisão

Renata Mourão

Roberto Araújo

Projeto gráfico

Amanda Soares

Letícia Lopes

Edição eletrônica

Amanda Soares

A Escola Nacional de Administração Pública (Enap) é uma escola de governo vinculada ao Ministério da Economia (ME).

Tem como principal atribuição a formação e o desenvolvimento permanente dos servidores públicos. Atua na oferta de cursos de mestrados profissionais, especialização *lato sensu*, cursos de aperfeiçoamento para carreiras do setor público, educação executiva e educação continuada.

A instituição também estimula a produção e disseminação de conhecimentos sobre administração pública, gestão governamental e políticas públicas, além de promover o desenvolvimento e a aplicação de tecnologias de gestão que aumentem a eficácia e a qualidade permanente dos serviços prestados pelo Estado aos cidadãos. Para tanto, desenvolve pesquisa aplicada e ações de inovação voltadas à melhoria do serviço público.

O público preferencial da Escola são servidores públicos federais, estaduais e municipais. Sediada em Brasília, a Enap é uma escola de governo de abrangência nacional e suas ações incidem sobre o conjunto de todos os servidores públicos, em cada uma das esferas de governo.

O482m Oliveira, Anderson Castro Soares de

Modelos espaço-temporais para avaliação do risco da Covid-19 nos municípios brasileiros / Anderson castro Soares de Oliveira. -- Brasília: Enap, 2021.

83 p. : il. -- (Cadernos Enap, 88)

Inclui bibliografia

ISSN: 0104-7078

1. Avaliação de risco. 2. Saúde Pública. 3. Administração Pública Municipal. 4. Pandemia. I. Título. II. Morita, Lia Hanna Martins.

CDU 35:616-036.21(81)

Bibliotecária: Tatiane de Oliveira Dias – CRB1/2230



Enap, 2021

Este trabalho está sob a Licença Creative Commons – Atribuição: Não Comercial – Compartilha Igual 4.0 Internacional.

As informações e opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista da Escola Nacional de Administração Pública (Enap). É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.



Escola Nacional de Administração Pública (Enap)

Diretoria de Altos Estudos

Coordenação-Geral de Pesquisa

SAIS – Área 2-A – 70610-900 — Brasília-DF, Brasil

CÁTEDRAS FAST-TRACK

Editorial

O ano de 2021 começou com a boa novidade das vacinas, permitindo às pessoas a revisão de suas expectativas quanto ao futuro. Expectativas são baseadas em informações e, para tomar boas decisões, é preciso que os indivíduos estejam bem informados.

A pesquisa científica, por exemplo, é um insumo informacional útil para gestores públicos e privados. Contudo, como ficou evidente neste último ano, a pesquisa leva tempo: seus resultados nem sempre são imediatos. Trata-se de um empreendimento árduo, mas necessário, principalmente quando a pesquisa tem por objetivo auxiliar na formulação de políticas públicas em um período tão atípico como o da pandemia de Covid-19.

É neste contexto que, em 2020, de forma inédita em sua história pela agilidade com a qual foi implementado, a Enap lançou uma chamada pública para seleção de projetos de pesquisas que gerem subsídios para o entendimento ou enfrentamento à pandemia de Covid-19 pela Administração Pública. Ficou conhecido como o edital Cátedras Covid-19 e os dez projetos de pesquisa selecionados foram concluídos até o final do ano de 2020.

O trabalho de Monique Menezes e coautores, sob a ótica das chamadas capacidades estatais, encontra uma heterogeneidade nas políticas públicas adotadas nas capitais brasileiras. A análise de documentos (conteúdo e discurso) mostrou uma articulação entre governos municipais e

estaduais. Além disso, um apanhado de “boas” e “más” práticas no combate à pandemia, por capitais brasileiras, é um interessante subproduto deste trabalho.

O modelo epidemiológico Suscetíveis-Infectedos-Recuperados (SIR) microfundamentado foi usado por Geraldo Sandoval Goés e Luan Borelli para verificar o impacto da pandemia em cinco estados brasileiros: São Paulo, Amazonas, Ceará, Rio de Janeiro e Pernambuco. O objetivo foi comparar dois cenários: um no qual todos os estados seguem uma mesma política de contenção do vírus e outro no qual cada um adota uma política própria, conforme suas particularidades. As evidências das simulações favorecem a adoção de políticas públicas que respeitem as peculiaridades de cada estado.

Por meio de uma extensa base de dados municipais, Janaina Lopes Pereira Peres e coautoras encontram seis *clusters* de municípios espalhados de forma nada trivial pelo território brasileiro. As autoras criaram o termo comorbidade social para designar o “acúmulo de patologias sociais em um determinado território”, o que serviu de base teórica para seu trabalho. A pesquisa mostra evidências de que vários municípios das regiões Norte e Nordeste apresentavam grande quantidade de comorbidades sociais e também um desempenho ruim na pandemia (em termos de casos e óbitos por Covid-19).

Durante os primeiros meses da pandemia, vários governos estaduais buscaram um papel ativo na busca por soluções científicas. Este protagonismo foi detalhadamente estudado pela bolsista Silmary de Jesus Gonçalves Alvim, com um exaustivo e inédito levantamento de legislações com foco em políticas públicas estaduais de Ciência, Tecnologia e de Inovação (CT&I), no qual foram identificadas 118 políticas estaduais voltadas ao combate à Covid-19, sendo 19% delas caracterizadas pela parceria com o governo federal. Estimou-se que estes projetos movimentaram cerca de R\$ 231 milhões em 2020.

O bolsista Anderson Castro Soares de Oliveira e a coautora Lia Hanna Morita utilizaram dados diários do Ministério da Saúde (27 de março a 3 de outubro de 2020) e trabalharam com aproximadamente 1,5 milhão de observações utilizadas em vários modelos espaço-temporais (Poisson, Poisson Hurdle, Poisson Inflado de Zero, Binomial Negativa, Binomial Negativa Hurdle e Binomial Negativa Inflada de Zeros). O painel para a visualização dos resultados é outro subproduto desta pesquisa (<https://lecdufmt.shinyapps.io/COVID/>) e, já na 40ª semana epidemiológica, apontava para uma preocupante situação no estado do Amazonas.

Rafael Mesquita Pereira e seus coautores estudaram o impacto da pandemia no grupo de trabalhadores com mais de 60 anos de idade no período compreendido entre fevereiro e agosto de 2020. Em 18 de junho de 2020, a portaria conjunta do Ministério da Economia, da Saúde e da Agricultura, Pecuária e Abastecimento reconheceu a necessidade de se atribuir tratamento diferenciado geral aos trabalhadores do grupo de risco. O resultado foi uma diminuição nas horas trabalhadas por parte dos funcionários públicos nesta faixa etária, fato não observado no caso dos trabalhadores na mesma faixa etária no setor. Os autores ressaltaram os efeitos diferenciados da legislação que permitiu a funcionários públicos reduzirem suas jornadas de trabalho sem redução de rendimentos, ao passo que os trabalhadores do setor privado, em sua maioria, não puderam evitar o trabalho presencial visando minimizar as perdas em seus rendimentos.

O setor de transporte público municipal é alvo do trabalho de Gabriel Pabst. Segundo o autor, o setor já apresentava um déficit financeiro (projetado pelo autor em R\$ 8 bilhões no final de 2020) e a pandemia gerou um conjunto de medidas restritivas de circulação diminuindo a quantidade de passageiros em circulação. Este é um setor cuja regulação carece de aperfeiçoamentos, como bem discute o pesquisador.

O trabalho infantil é objeto da pesquisa de Thauan Patrello e Ruth Knaak. Por meio de entrevistas envolvendo autoridades do setor no Estado do Espírito Santo,

os autores especulam que mudanças na política educacional seriam importantes para combater o trabalho infantil, indicando o ensino integral como exemplo de política pública interessante. Outra proposta apresentada pelos autores envolve a garantia de algum tipo de auxílio emergencial às famílias das crianças e o reforço às medidas tradicionais de combate ao trabalho infantil.

Nos primeiros meses da pandemia, ainda em 2020, houve um debate importante acerca dos métodos de rastreamento de pessoas contaminadas - o chamado *contact tracing* - a partir de políticas adotadas em alguns países. Ferramentas criadas para o rastreamento levantam questões importantes sobre a privacidade dos dados dos cidadãos. No caso brasileiro, essa discussão encontra eco na nova Lei Geral de Proteção de Dados Pessoais (LGPD). A bolsista Letícia Lobato Anicet Lisboa fez um levantamento das políticas de rastreamento dos Estados Unidos, da União Europeia, da Alemanha, do Reino Unido e de Cingapura. As lições extraídas das políticas públicas desses países podem promover melhorias no arcabouço das políticas de saúde com base em evidências.

O pesquisador Júlio César Andrade de Abreu e coautores estudaram o uso de *business intelligence* (BI) pelos poderes públicos municipais no estado do Rio de Janeiro. Foi apontado que 21 municípios cariocas (aproximadamente 22.8% do total) fizeram uso de alguma ferramenta de BI no acompanhamento da evolução da Covid-19. Este baixo uso é preocupante, não apenas pelo cenário da pandemia no país. Trata-se, afinal, de uma ferramenta de gestão estratégica que pode - e deve - ter seu uso difundido na administração pública, visando informar o cidadão com o máximo de transparência possível dentro do marco da nova LGPD.

A pesquisa é um empreendimento fascinante. Cada projeto de pesquisa aqui resumido abordou o problema da pandemia no território brasileiro sob diferentes óticas teóricas, por meio do uso de abordagens empíricas distintas. A leitura mais atenta do conjunto desses trabalhos mostrará a emergência de alguns consensos e também novas hipóteses a serem pesquisadas.

Sumário

Executivo

Diante da pandemia da Covid-19 e sua interiorização no Brasil, os municípios e sistemas de saúde apresentam desafios e formas diferentes de enfrentar a crise. Não apenas o número de casos difere substancialmente de um município para outro, mas cada localidade tem melhor ou pior condição para lidar com o surto. A variação nos resultados de saúde é motivada por múltiplos e complexos fatores, incluindo o “estoque” de saúde existente na população e a qualidade dos sistemas de saúde e assistência social locais.

No atual contexto de escassez de ferramentas para orientar as ações de vigilância e controle da pandemia, a modelagem estatística pode ser uma boa ferramenta, no intuito de compreender fatores que contribuem para expansão da epidemia e construir indicadores que possam subsidiar ações de enfrentamento. Assim, este trabalho teve por objetivo a utilização de modelos espaço-temporais para o desenvolvimento de indicadores de risco que consideram o avanço da doença e a capacidade do sistema de saúde, no intuito de auxiliar no desenvolvimento de políticas de enfrentamento.

Foram obtidos dados diários do Ministério da Saúde sobre casos e óbitos de Covid-19 de todos os 5570 municípios brasileiros, no período de 27 de março a 03 de outubro de 2020. Esses dados foram agregados por semana epidemiológica, com início no domingo e

término no sábado. Para cada município, foi verificado se o número de casos ou óbitos em uma semana epidemiológica específica possuíam valores negativos. Nesses casos, tais valores foram zerados e subtraídos da semana anterior.

Dessa forma, o banco de dados foi composto de número de casos e óbitos em 27 semanas para cada município, totalizando assim 1.503.980 observações.

Também foram obtidas as informações do Cadastro Nacional de Estabelecimentos de Saúde (CNES) com respeito à infraestrutura de saúde, com as informações referentes a todos os municípios brasileiros no período de fevereiro a setembro de 2020.

As variáveis utilizadas foram:

- número de postos de saúde (NPS);
- número de centros de saúde/unidades básicas (NCS);
- número de leitos clínicos (NLC);
- número de médicos (NM);
- número de enfermeiros e técnicos de enfermagem (NE).

Todas as variáveis foram transformadas em taxas por 10.000 habitantes, da seguinte forma:

$$TPS = \frac{10000NPS}{P}$$

em que P é a população do município.

Em seguida, foram ajustados modelos com e sem a infraestrutura de saúde, considerando diferentes estruturas de dependência temporal e espacial. Além disso, foram testadas seis distribuições de probabilidades, a saber, as distribuições de Poisson, Poisson Hurdle, Poisson inflado de zero (ZIP), binomial negativa, binomial negativa Hurdle e binomial negativa inflada de zeros.

A partir dos modelos definidos para casos e óbitos, foi desenvolvido um painel de visualização com as informações básicas e os indicadores de risco para cada município brasileiro.

Os modelos espaço-temporais desenvolvidos neste projeto contribuem para a interpretação do risco de espalhamento da epidemia de Covid-19. No estudo da evolução dos casos de Covid-19, o modelo espaço-temporal com distribuição de

probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e estrutura de vizinhança do tipo *queen* foi o mais adequado para explicar a evolução de casos nos municípios brasileiros.

Com respeito à evolução dos óbitos de Covid-19, o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e estrutura de vizinhança do tipo *queen* apresentou o melhor ajuste entre os modelos estudados para explicar a distribuição da contagem acumulada de óbitos.

A obtenção de modelos Hurdle, tanto para casos quanto para óbitos, indica que a ocorrência de zeros, ou seja, a falta de casos e mortes, ocorre de forma estrutural, indicando que os valores zero provêm de fonte não aleatória ou não fazem parte do contexto de casos e/ou óbitos.

No estudo da relação entre a estrutura de saúde e a evolução de casos e óbitos, observou-se que as taxas de postos de saúde, centros de saúde, leitos de clínica geral e enfermeiros apresentaram efeitos negativos ao risco de aumento de casos e óbitos por Covid-19, levando à conclusão que o incremento dessas quantidades contribui para a redução do risco.

Esses resultados trazem interpretações importantes, pois, com as estimativas das quantidades *a posteriori*, foi possível estabelecer a distribuição do risco de aumento ou diminuição de casos e óbitos nos municípios.

Foram apresentados gráficos para verificar a boa calibração dos modelos, mapas de visualização dos riscos nos municípios brasileiros e gráficos de tendência temporal ao longo das semanas epidemiológicas, nos modelos para casos e óbitos.

Por fim, foi desenvolvido um painel de visualização para os casos e óbitos decorrentes de Covid-19 no seguinte endereço: <https://lecdufmt.shinyapps.io/COVID/>. O painel de visualização auxiliará os pesquisadores, gestores e profissionais da saúde no entendimento desta epidemia e no desenvolvimento de ações e tomada de decisões de saúde da população.



Clique aqui para baixar
o **Sumário Executivo** separado.
Compartilhe!

Resumo

O avanço da epidemia de Covid-19 no Brasil apresenta desafios aos diferentes estados e municípios. O número de casos e óbitos difere substancialmente entre os locais, assim como as condições de estrutura e saúde. A diversidade da saúde nos locais é influenciada por vários fatores, como o “estoque” de saúde existente na população e a qualidade dos sistemas de saúde e assistência social locais. Entre as ferramentas quantitativas que auxiliam nas decisões de vigilância e controle da pandemia, a modelagem estatística apresenta destaque, pois contribui para o entendimento dos fatores relacionados à expansão da mesma. Este projeto consiste no desenvolvimento de modelos espaço-temporais para criação de indicadores de risco de expansão da epidemia de Covid-19 que contribuam na promoção de ações de atenção à saúde. Foram apresentados vários modelos espaço-temporais com distribuições de probabilidade, estruturas temporais e espaciais distintas, com o objetivo de modelar o comportamento e espalhamento da doença nos municípios brasileiros e interpretar o comportamento de risco de aumento ou diminuição de casos e óbitos. Também foram considerados nos estudos a estrutura de postos de saúde, centros de saúde, leitos de clínica geral, médicos, enfermeiros e técnicos de enfermagem. Por fim, foi desenvolvido um painel de visualização para os casos e óbitos decorrentes de Covid-19. A apresentação desses modelos e do painel de visualização contribui para os pesquisadores, gestores e profissionais da saúde no entendimento desta epidemia e auxilia no desenvolvimento de ações de atenção à saúde.

Palavras-chave: Covid-19, modelos espaço-temporais, saúde, risco

Abstract

The advancement of the Covid-19 epidemic in Brazil presents challenges to different states and municipalities. The number of cases and deaths differs substantially between locations, as do structure and health conditions. Several factors influence the heterogeneity of health situation in the places, such as the existing “stock” of health in the population and the quality of the local health and social assistance systems. Among the quantitative tools that assist in the decisions of surveillance and pandemic control, statistical modeling deserves attention because it contributes to the understanding of the factors related to its expansion. This project consists of the development of Spatio-temporal models for the creation of risk indicators for the evolution of the Covid-19 epidemic in the advancement of health care actions. Several Spatio-temporal models with different probability distributions, temporal and spatial structures were developed to model the behavior and spread of the disease in Brazilian municipalities and understanding the risk behavior of increasing or decreasing cases and deaths. The structure of health posts, health centers, general practice beds, doctors, nurses, and nursing technicians were also considered in the studies. Finally, a visualization panel was developed for cases and deaths resulting from Covid-19. The presentation of these models and the visualization panel may give support to researchers, managers, and health professionals to understand this epidemic and assist in the development of health care actions.

Key-words: Covid-19, spatio-temporal models, health, risk

Sumário

1.

Introdução

Pág. 20

2.

Revisão bibliográfica

Pág. 23

3.

Metodologia

Pág. 37

4.

Resultados e discussão

Pág. 45

5.

Conclusões

Pág. 71



Lista de quadros

Quadro 1 – Tipos de arquivos disponíveis nos microdados do CNES

Pág. 26

Lista de tabelas

Tabela 1 – Valores de DIC e LPML dos modelos espaço-temporais para casos com diferentes modelos de probabilidade e estruturas temporais, e com vizinhança do tipo *queen*

Pág. 47

Tabela 2 – Valores de DIC e LPML dos modelos espaço-temporais para casos com distribuição binomial negativa Hurdle, efeito temporal de 2ª ordem e diferentes estruturas de vizinhança

Pág. 48

Tabela 3 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

Pág. 49

Tabela 4 – Valores de DIC e LPML – nos modelos para óbitos com vizinhança do tipo *queen*

Pág. 54

Tabela 5 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*

Pág. 55

Tabela 6 – Valores de DIC e LPML – modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 2ª ordem e vizinhança tipo *queen*, com e sem estrutura de saúde

Pág. 64

Tabela 7 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 2ª ordem e vizinhança tipo *queen*, com estrutura de saúde

Pág. 64

Tabela 8 – Valores de DIC e LPML – modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança tipo *queen*, com e sem estrutura de saúde

Pág. 65

Tabela 9 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança tipo *queen*, com estrutura de saúde

Pág. 65

Lista de figuras

Figura 1 – Gráfico de histograma da transformação integral de probabilidade (PIT) do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

Pág. 49

Figura 2 – Médias (a) e probabilidades (b) *a posteriori* dos efeitos espaciais (ζ_i) para cada município brasileiro, considerando o modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

Pág. 50

Figura 3 – Tendência temporal da média *a posteriori* do risco de aumentos de casos, considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança tipo *queen*

Pág. 51

Figura 4 – Risco de aumento de Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (27 de setembro a 03 de outubro de 2020), considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

Pág. 52

Figura 5 – Prevalência de casos de Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)

Pág. 53

Figura 6 – Gráfico de histograma da transformação integral de probabilidade (PIT) do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*
Pág. 55

Figura 7 – Médias (a) e probabilidades (b) *a posteriori* dos efeitos espaciais (ζ_i) para cada município brasileiro, considerando o modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança do tipo *queen*
Pág. 56

Figura 8 – Tendência temporal da média *a posteriori* do risco de aumentos de óbitos, considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança tipo *queen*
Pág. 57

Figura 9 – Risco de aumento de óbitos pela Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (27 de setembro a 03 de outubro de 2020), considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*
Pág. 58

Figura 10 – Taxa de mortalidade para Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)
Pág. 59

Figura 11 – Taxas de postos de saúde por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)
Pág. 60

Figura 12 – Taxas de centros de saúde por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)

Pág. 60

Figura 13 – Taxas de leitos em clínica geral por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)

Pág. 61

Figura 14 – Taxas de médicos por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)

Pág. 62

Figura 15 – Taxas de enfermeiros e técnicos de enfermagem por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)

Pág. 63

Figura 16 – Informações da aba Brasil no painel de visualização

Pág. 66

Figura 17 – Informações da aba Estados no painel de visualização

Pág. 67

Figura 18 – Informações da aba *Municípios* no painel de visualização

Pág. 67

Figura 19 – Informações da aba *Risco* no painel de visualização

Pág. 68

1.

Introdução



1. Introdução

Em 31 de dezembro de 2019, a Organização Mundial da Saúde (OMS) foi alertada sobre vários casos de pneumonia na cidade de Wuhan, província de Hubei, na República Popular da China. A nova doença foi denominada de Covid-19 e seu agente etiológico foi identificado como Coronavírus da Síndrome Respiratória Aguda Grave 2 (SARS-CoV-2), que é um novo tipo de coronavírus da família betacoronavírus. Após o primeiro caso detectado na China, a doença continuou a se espalhar globalmente, com casos confirmados em todos os continentes do mundo. Em 12 de março de 2020, a Organização Mundial da Saúde (OMS) declarou a Covid-19 como doença pandêmica, quando mais de 20 mil casos e quase mil mortes foram registradas na região europeia – o epicentro dessa pandemia, de acordo com o Comitê Permanente da Europa (HUI *et al.*, 2020; WORLD HEALTH ORGANIZATION, 2020).

No Brasil, o primeiro caso confirmado de Covid-19 ocorreu em 25 de fevereiro de 2020 e a primeira morte no dia 17 de março de 2020, ambos na cidade de São Paulo. Ao longo dos meses, a doença continuou se espalhando e, no dia 03 de outubro, já tínhamos mais de 4.906.833 casos e 145.987 mortes, em 5.561 dos 5.570 municípios brasileiros (BRASIL, 2020b).

A epidemia de Covid-19 trouxe um grande desafio, a todos os países do mundo, no delineamento de políticas públicas, devido principalmente à velocidade assustadora com a qual essa doença se espalha pelas comunidades, ao colapso nos sistemas de saúde e à redução da atividade econômica (GOSTIN; FRIEDMAN; WETTER, 2020; WHITWORTH, 2020).

No caso do Brasil, preservando as devidas proporções, a Covid-19 não é o primeiro surto significativo de saúde. Ano após ano, temos enfrentado uma gama de doenças, tais como dengue, zika, chikungunya, influenza. Algumas características particulares da população brasileira trazem grandes dificuldades ao enfrentamento de grandes surtos, entre as quais podemos destacar o tamanho territorial e populacional, a heterogeneidade intrínseca ao seu extenso território e a presença de subnotificação de casos dessas doenças (DE OLIVEIRA; LOSCHI; ASSUNÇÃO, 2017; STONER; ECONOMOU; SILVA, 2019).

Diante da pandemia da Covid-19 e sua interiorização no Brasil, os municípios e sistemas de saúde têm desafios e formas diferentes de enfrentar a crise. Não apenas o número de casos difere substancialmente de um município para outro, mas cada localidade tem melhor ou pior condição para lidar com o surto. A variação nos resultados de saúde será motivada por múltiplos e complexos fatores, incluindo o estoque de saúde existente na população e a qualidade dos sistemas de saúde e assistência social locais (ANDRADE *et al.*, 2020; SANTANA *et al.*, 2020).

No atual contexto de escassez de ferramentas para orientar as ações de vigilância e controle da pandemia, a modelagem estatística pode ser uma boa ferramenta, no intuito de compreender fatores que contribuem para a expansão da epidemia e construir indicadores que possam subsidiar ações de enfrentamento.

Estudos de padrões espaço-temporais podem ajudar a compreender os mecanismos de disseminação da doença na população e identificar fatores associados à distribuição geográfica heterogênea (MOHEBBI; WOLFE; FORBES, 2014; ELLIOTT, 2000; ADIN; GOICOA; UGARTE, 2019).

Nesse contexto, este trabalho teve por objetivo a utilização de modelos espaço-temporais para o desenvolvimento de indicadores de risco que consideram o avanço da doença e a capacidade do sistema de saúde, no intuito de auxiliar no desenvolvimento de políticas de enfrentamento.

2.

Revisão bibliográfica



2. Revisão bibliográfica

2.1 Sistemas de Informação em Saúde

O Brasil dispõe de uma ampla rede de Sistemas de Informação em Saúde (SIS), sendo que sua maior parte encontra-se disponível para o acesso público. É administrado pelo Ministério da Saúde (MS) através do Departamento de Informática do Sistema Único de Saúde (DATASUS – <http://www.datasus.gov.br>). A apropriação dessas informações pelos gestores e pela sociedade organizada para subsidiar suas atividades deve considerar seus pontos fortes e suas limitações, sendo resultado de avaliações regulares e sistemáticas dos dados disponibilizados (LIMA *et al.*, 2009).

A importância do SIS reside na capacidade da gestão do Sistema Único de Saúde (SUS) na difusão de saberes, o que nos remete ao desenvolvimento da inteligência da informação em saúde, por meio da coleta de dados organizados e transformados em conhecimento para oferecer subsídios ao planejamento, de acordo com as necessidades da população, em cumprimento ao proposto pela Reforma Sanitária Brasileira (SILVA, 2020).

Com a emergência da pandemia do Coronavírus, foi necessário disponibilizar um SIS para notificação de casos e óbitos. O SUS já dispunha do Sistema de Vigilância Sentinela de Síndromes Gripais e do Sistema de Informação das Síndromes Respiratórias Agudas Graves, os quais foram mantidos, e implantou uma nova ferramenta de vigilância, o e-SUS VE, no qual situou um módulo específico para registrar as notificações de Covid-19 (BRASIL, 2020c).

Assim foi criado o Portal da Covid-19 (<https://covid.saude.gov.br>), em que diariamente o Ministério da Saúde, por meio da Secretaria de Vigilância em Saúde (SVS/MS) divulga dados consolidados sobre a situação epidemiológica da Covid-19 no Brasil. O processo de atualização dos dados sobre casos e óbitos confirmados por Covid-19 no Brasil é realizado através do repasse das informações oficiais pelas Secretarias Estaduais de Saúde das 27 unidades federativas brasileiras. Os dados fornecidos pelos estados são consolidados e disponibilizados publicamente todos os dias, em torno das 19h (BRASIL, 2020b).

2.1.1 Cadastro Nacional de Estabelecimentos de Saúde

O Cadastro Nacional de Estabelecimentos de Saúde (CNES) é o sistema de informação oficial de cadastramento de informações de todos os estabelecimentos de saúde no Brasil, vinculado ao Ministério da Saúde (MS) no tocante à realidade da capacidade instalada e mão de obra assistencial de saúde no Brasil nos estabelecimentos de saúde públicos ou privados, podendo ou não serem conveniados ao SUS (BRASIL, 2020a).

A criação do CNES se deu pela Portaria SAS/MS nº 376, de 03 de outubro de 2000. Após um período de consulta pública, no mesmo ano foi publicada a Portaria SAS/MS nº 511/2000, que passou então a servir de norma para o processo de cadastramento no sistema em todo território nacional (BRASIL, 2020a).

O CNES possui as seguintes finalidades:

- cadastrar e atualizar as informações sobre os estabelecimentos de saúde e suas dimensões, como recursos físicos, trabalhadores e serviços;
- disponibilizar informações dos estabelecimentos de saúde para os outros sistemas de informação;
- ofertar para a sociedade as informações sobre a disponibilidade de serviços nos territórios, formas de acesso e funcionamento;
- fornecer informações que apoiem as tomadas de decisões, o planejamento, a programação e o conhecimento pelos gestores, pesquisadores,

trabalhadores e a sociedade em geral acerca da organização, existência e disponibilidade dos serviços, da força de trabalho e da capacidade instalada nos estabelecimentos de saúde e territórios.

Os dados do CNES são importantes para a área de planejamento, controle e avaliação em saúde, pois são a principal fonte de informação da rede física assistencial no país, principalmente pela quantidade e variedade de informações sobre os serviços oferecidos, com localização, características de recursos humanos, habilitações, infraestrutura, entre outras. A estrutura de dados do CNES fornece suporte para todos os sistemas de informação que utilizam dados relativos a estabelecimentos de saúde (PELISSARI, 2019).

De acordo com Saldanha, Bastos e Barcellos (2019), a disseminação desses dados do CNES é feita no formato de microdados, contendo 14 tipos de arquivos distintos, conforme Quadro 1.

Quadro 1 – Tipos de arquivos disponíveis nos microdados do CNES

Arquivo	Sigla	Cobertura	Periodicidade
Leitos	CNES-LT	Outubro/2005 - atual	Mensal
Estabelecimentos	CNES-ST	Agosto/2005 - atual	Mensal
Dados complementares	CNES-EQ	Agosto/2005 - atual	Mensal
Equipamentos	CNES-EQ	Agosto/2005 - atual	Mensal
Serviços Especializados	CNES-SR	Agosto/2005 - atual	Mensal
Habilitações	CNES-HB	Março/2007 - atual	Mensal
Profissionais	CNES-PF	Agosto/2005 - atual	Mensal
Equipes	CNES-EP	Abril/2007 - atual	Mensal
Regras Contratuais	CNES-RC	Março/2007 - atual	Mensal
Incentivos	CNES-IN	Novembro/2007 - atual	Mensal
Estabelecimentos de Ensino	CNES-EE	Março/2007 - atual	Mensal
Estabelecimentos Filantrópicos	CNES-EF	Março/2007 - atual	Mensal
Gestão e Metas	CNES-GM	Junho/2007 - atual	Mensal

Fonte: adaptado de Saldanha, Bastos e Barcellos (2019).

2.2 Mapeamento de doenças

O mapeamento de doenças tem uma longa história em epidemiologia e tem por objetivo estudar o padrão espacial e espaço-temporal do risco de uma doença em uma determinada região geográfica, de modo que as áreas de alto risco possam ser identificadas. O mapeamento de doenças utiliza as estimativas de prevalência, incidência ou taxa de mortalidade padronizadas de uma determinada doença, obtidas a partir de contagens particulares de cada área geográfica, oferecendo estimativas do risco relativo nessas áreas (ADIN; GOICOA; UGARTE, 2019; ELLIOTT, 2000; MOHEBBI; WOLFE; FORBES, 2014).

Devido à grande variabilidade inerente às medidas clássicas de estimação, como taxas padronizadas de mortalidade e/ou incidência, torna-se necessário o uso de modelos para suavizar os riscos da utilização das informações dos vizinhos espaciais e temporais. As pesquisas sobre o mapeamento espacial e espaço-temporal de doenças no contexto de uma estrutura bayesiana, através dos modelos lineares mistos generalizados (GLMM), têm ganhado destaque nos últimos anos. Duas abordagens têm sido utilizadas amplamente para o ajuste e inferência do modelo, a saber, métodos bayesianos empíricos e métodos totalmente bayesianos. O crescimento do uso dessas técnicas se deve ao grande avanço nos recursos computacionais disponíveis e também à necessidade de se obter estimativas mais suavizadas, que reflitam de forma mais adequada a característica subjacente do evento de interesse, sem a influência da alta variabilidade observada nos dados brutos (ADIN; GOICOA; UGARTE, 2019).

Os métodos bayesianos empíricos utilizam a quase-verossimilhança penalizada para realizar inferências (BRESLOW; CLAYTON, 1993). Os métodos totalmente bayesianos geralmente utilizam técnicas de Monte Carlo com cadeias de Markov (MCMC). Além dos métodos MCMC, uma nova estratégia baseada em aproximações de Laplace aninhadas integradas (INLA) foi derivada recentemente para estimar as quantidades de interesse *a posteriori* (RUE; MARTINO; CHOPIN, 2009).

2.3 Modelos espaço-temporais

Considerando o problema em que desejamos estudar as n áreas de uma região geográfica com número de casos Y_i de uma determinada doença, em geral a estimação é através do modelo de Poisson.

Seja E_i o número de casos esperados e θ_i o risco relativo na i -ésima área de estudo (BESAG; YORK; MOLLÍÉ, 1991; ELLIOTT, 2000; MOHEBBI; WOLFE; FORBES, 2014), então temos:

$$Y_i | \theta_i \sim \text{Poisson}(E_i \theta_i)$$

Dessa forma, o risco relativo pode ser estimado utilizando um preditor linear logarítmico

$$\eta_i = \log(\theta_i) = \alpha + \mu_i$$

em que μ_i descreve uma componente aleatória espacialmente estruturada.

Suponha agora que, para cada área i , os dados estão disponíveis em diferentes períodos de tempo rotulados por $t = 1, \dots, T$. Assim, temos

$$Y_{i,t} | \theta_{i,t} \sim \text{Poisson}(E_{i,t} \theta_{i,t})$$

Nessa abordagem, o risco relativo também pode ser estimado utilizando um preditor linear logarítmico

$$\eta_{i,t} = \log(\theta_{i,t}) = \alpha + \mu_i + \gamma_t$$

em que γ_t descreve uma componente aleatória temporal estruturada.

Os modelos espaço-temporais que incluem os efeitos de área e tempo de forma aditiva podem ser muito restritivos na prática, então os termos de interação são geralmente adicionados ao modelo (KNORR-HELD, 2000), como vemos abaixo:

$$\eta_{i,t} = \log(\theta_{i,t}) = \alpha + \mu_i + \gamma_t + \delta_{it} \tag{1}$$

2.3.1 Modelos Inflados

A utilização do modelo de Poisson assume a igualdade nos valores da média e variância. Entretanto, essa suposição nem sempre é satisfeita, principalmente porque em algumas áreas pode não ocorrer a doença estudada, podendo resultar em excesso de zeros. Nesses casos, pode haver superdispersão dos dados, sendo a variância maior do que a média. Assim, podem ser utilizados outros modelos, tais como o binomial negativo, os modelos inflados de zeros e os modelos Hurdle (HINDE; DEMÉTRIO, 1998; HU; PAVLICOVA; NUNES, 2011; MOHEBBI; WOLFE; FORBES, 2014).

O modelo binomial negativo (BN) incorpora um parâmetro que modela a variação adicional inerente aos dados. Assim, temos:

$$Y_i | \theta_i \sim \text{BN}(E_i \theta_i, \tau)$$

em que τ é um parâmetro de escala e representa a superdispersão.

O modelo BN pode ser utilizado na presença de superdispersão, entretanto esse modelo não se caracteriza como a melhor alternativa para as situações em que essa superdispersão é causada pelo excesso de zeros. Nesse contexto, podem ser utilizados os modelos inflados de zeros (ZI), que consistem na mistura entre a distribuição de Bernoulli e as distribuições Poisson ou BN em sua formulação (CAMERON; TRIVEDI, 2013; GHOSH; MUKHOPADHYAY; LU, 2006).

Nos modelos inflados de zeros, presume-se que dois tipos de zeros podem ocorrer provenientes de dois processos: um processo que gera apenas zeros e outro processo que gera zeros e não-zeros. Esses modelos podem ser descritos como a mistura de uma distribuição de Bernoulli em zero com probabilidade π_i e uma distribuição de Poisson (ou, alternativamente, BN) com probabilidade $(1 - \pi_i)$ (CAMERON; TRIVEDI, 2013; GHOSH; MUKHOPADHYAY; LU, 2006). Assim, podemos descrever os seguintes modelos:

- Modelo Poisson inflado de zero (ZIP)

$$Y_{i,t} | \theta_{i,t}, \pi_{i,t} \sim \text{Poisson}(E_{i,t} \theta_{i,t}, \pi_{i,t})$$

$$P[Y_{i,t} = y] = \begin{cases} \pi_{i,t} + (1 - \pi_{i,t}) \text{Poisson}(E_{i,t} \theta_{i,t}) & \text{se } y = 0 \\ (1 - \pi_{i,t}) \text{Poisson}(E_{i,t} \theta_{i,t}) & \text{se } y > 0 \end{cases}$$

- Modelo binomial negativo inflado de zero (ZIBN)

$$Y_{i,t} | \theta_{i,t}, \pi_{i,t} \sim \text{BN}(E_i \theta_i, \tau, \pi_{i,t})$$

$$P[Y_{i,t} = y] = \begin{cases} \pi_{i,t} + (1 - \pi_{i,t}) \text{BN}(E_i \theta_i, \tau) & \text{se } y = 0 \\ (1 - \pi_{i,t}) \text{BN}(E_i \theta_i, \tau) & \text{se } y > 0 \end{cases}$$

Os modelos Hurdle também são definidos como mistura de uma distribuição Bernoulli e as distribuições Poisson ou BN, mas assume-se que todos os valores iguais a zero provêm de uma fonte estrutural, enquanto que os valores diferentes de zero provêm de “amostragem” e são definidos por uma distribuição de Poisson ou BN truncada (HU; PAVLICOVA; NUNES, 2011). Assim, temos

- Modelo Hurdle Poisson (PH)

$$Y_{i,t} | \theta_{i,t}, \pi_{i,t} \sim \text{Poisson}(E_{i,t} \theta_{i,t}, \pi_{i,t})$$

$$P[Y_{i,t} = y] = \begin{cases} \pi_{i,t} & \text{se } y = 0 \\ (1 - \pi_{i,t}) \text{Poisson}(E_{i,t} \theta_{i,t}) & \text{se } y > 0 \end{cases}$$

- Modelo Hurdle binomial negativo (NBH)

$$Y_{i,t} | \theta_{i,t}, \pi_{i,t} \sim \text{BN}(E_i \theta_i, \tau, \pi_{i,t})$$

$$P[Y_{i,t} = y] = \begin{cases} \pi_{i,t} & \text{se } y = 0 \\ (1 - \pi_{i,t}) \text{BN}(E_{i,t} \theta_{i,t}) & \text{se } y > 0 \end{cases}$$

2.3.2 Definição da estrutura espacial

No modelo espaço-temporal definido na equação 1, a componente aleatória espacialmente estruturada μ_i representa o efeito do espaço sobre a variável de interesse. Esse efeito espacial é definido como autocorrelação espacial, que indica quanto uma variável associada a uma localização se assemelha mais aos valores de suas observações vizinhas. No contexto de análise de áreas, é necessário definir a estrutura de vizinhança, que verifica a dependência espacial a partir da comparação entre os valores de uma amostra e de seus vizinhos (CÂMARA *et al.*, 2001; CRESSIE, 2015; ELLIOTT, 2000).

Uma área pode ser considerada vizinha a outra por diferentes critérios, como, por exemplo (BIVAND; PEBESMA; GOMEZ-RUBIO, 2013):

- vizinhança do tipo *queen* – as áreas vizinhas são aquelas que possuem fronteira entre si;
- vizinhança por distância – as áreas vizinhas são aquelas que compreendem uma determinada distância umas das outras.

A estrutura de vizinhança é expressa pela matriz de proximidade espacial ou matriz de vizinhança, através de uma matriz quadrada e não estocástica representada por $W_{n \times n}$, em que as linhas e as colunas apresentam todos os atributos, e a intersecção de cada elemento tem um peso w_{ij} , que representa uma medida de proximidade entre A_i e A_j , com $i = 1, \dots, m$ sendo o local considerado e $j = 1, \dots, k$ sendo o local vizinho. Os pesos podem assumir valores $w_{ij} = 1$, quando as áreas A_i e A_j são vizinhas, e $w_{ij} = 0$, no caso contrário (CÂMARA *et al.*, 2001; CRESSIE, 2015).

Modelos autorregressivos condicionais (CAR) propostos por Besag (1974) são frequentemente utilizados para descrever a variação espacial de dados de área, com o propósito de modelar os efeitos aleatórios, sendo definidos por:

$$Y_i | Y_{j \neq i} \sim \text{Normal} \left(\mu_i + \rho \sum_{j \neq i} w_{ij} (Y_j - \mu_j), \frac{\sigma^2}{m_i} \right),$$

em que, $Y_{j \neq i} = (Y_1, \dots, Y_{i-1}, Y_{i+1}, \dots, Y_n)$ é o vetor de observações, μ_i é o valor esperado em i , ρ é um parâmetro de correlação espacial que determina a dependência do efeito da vizinhança, w_{ij} 's são os pesos da matriz de vizinhança, $\sigma^2 > 0$ é o termo da variância, m_i é o número de regiões vizinhas.

2.3.3 Definição da estrutura temporal

No modelo espaço-temporal definido na equação 1, é necessário definir a componente aleatória temporal γ_t . Essa componente representa a autocorrelação temporal, ou seja, indica como as observações se relacionam no tempo, ou a intensidade em que um valor passado pode influenciar em um valor presente (CHATFIELD, 2000; ELLIOTT, 2000; MORETTIN; TOLOI, 2018).

A autocorrelação temporal pode ocorrer de diferentes formas e é definida por ordem ou *lag*. Sendo assim, a autocorrelação de primeira ordem indica que um observação está relacionada com a observação imediatamente anterior, a autocorrelação de segunda ordem caracteriza que uma observação está correlacionada com as observações a 2 unidades de tempo no passado e assim por diante (CHATFIELD, 2000; MORETTIN; TOLOI, 2018).

Os efeitos de tempo podem ser modelados por passeios aleatórios de primeira (RW1) ou de segunda ordem (RW2). Um passeio aleatório RW1 indica que os valores de uma determinada área dependem dos valores observados nessa área no tempo anterior. Por sua vez, um passeio aleatório do tipo RW2 indica que os valores dessa área dependem de seus valores observados em até dois tempos anteriores. Os passeios aleatórios são definidos pelas diferenças com adição de um efeito residual (CHATFIELD, 2000; DEDEFO *et al.*, 2018; KNORR-HELD, 2000).

2.4 Inferência Bayesiana

A inferência estatística tem por finalidade apontar as generalizações sobre uma população (conjunto de todas as observações possíveis) a partir de uma amostra. Para tanto, duas abordagens podem ser utilizadas: a bayesiana e a frequentista. Na abordagem frequentista, um parâmetro de interesse θ é considerado

desconhecido, porém é definido como uma quantidade fixa. Já na abordagem bayesiana, este é avaliado como uma quantidade desconhecida cuja variação pode ser descrita por uma distribuição de probabilidade subjetiva, denominada distribuição *a priori*, a qual é atualizada com a informação contida na amostra, dando origem à distribuição *a posteriori* (PAULINO *et al.*, 2018).

A inferência bayesiana é baseada no Teorema de Bayes:

$$h(\theta|y) = \frac{f(y|\theta)h(\theta)}{\int_{\theta} f(\theta)f(y|\theta)d\theta}$$

$$h(\theta|y) \propto f(y|\theta)h(\theta),$$

em que $h(\theta)$ é a distribuição *a priori*, a distribuição amostral $f(y|\theta)h(\theta)$ é a função de verossimilhança (distribuição dos dados) e $h(\theta|y)$ é a distribuição *a posteriori*.

Toda inferência sobre θ é feita a partir da distribuição *a posteriori*, que detém toda informação probabilística desse parâmetro. A *posteriori* é utilizada para obter estimativas pontuais, bem como estimativas intervalares, conhecidas como intervalos de credibilidade (PAULINO *et al.*, 2018).

As estimativas pontuais mais comumente utilizadas na análise bayesiana são a moda *a posteriori*, a média *a posteriori*, o quantil α *a posteriori* e a variância *a posteriori*, conforme definidas nas expressões a seguir (PAULINO *et al.*, 2018)):

- Moda *a posteriori* $\hat{\theta}$ com:

$$h(\hat{\theta}|x) = \max\{h(\theta|x)\}_{\theta \in \theta}$$

- Média *a posteriori* $\hat{\theta} = E(\theta|x)$ tal que:

$$E(\theta|x) = \int \theta h(\theta|x)d\theta, \theta \in \theta$$

- Quantil α *a posteriori* $Q(\alpha)$ devido a:

$$Q(\alpha) = \left\{ \theta \in \theta : \int_{-\infty}^{\theta} h(\theta|x)d\theta = \alpha \right\}, \alpha \in (0,1);$$

observa-se que quando $\alpha=0,5$ obtém-se a mediana *a posteriori*.

- Variância *a posteriori* $\hat{\theta} = Var(\theta|x)$ tal que:

$$Var(\theta|x) = \int (\theta - E(\theta|x))^2 h(\theta|x)d\theta, \theta \in \theta;$$

a raiz quadrada da variância equivale ao desvio padrão *a posteriori*.

No contexto bayesiano, a estimação intervalar é feita por meio do intervalo de credibilidade $100(1 - \alpha)\%$,

$$(A, B) = \int_A^B h(\theta|x)d\theta = 1 - \alpha, \quad \alpha \in (0,1);$$

sendo assim para um mesmo nível de credibilidade, existe uma infinidade de intervalos distintos. A região com menor amplitude para um mesmo nível de credibilidade é chamada de intervalo de máxima densidade *a posteriori* (HPD), do inglês *Highest posterior density*, e consiste em tomar os valores de θ com maior densidade *a posteriori* (BLANGIARDO; CAMELETTI, 2015).

A distribuição *a posteriori* nem sempre possui forma fechada e as integrais necessárias para a obtenção das distribuições marginais e avaliação dos estimadores bayesianos não possuem solução analítica. Desse modo, os métodos numéricos são utilizados para a obtenção da distribuição *a posteriori*, bem como dos estimadores. Existem diversos algoritmos para inferência bayesiana, dentre os quais se destacam os métodos de Monte Carlo via cadeias de Markov (MCMC), que geram valores através de uma distribuição condicional *a posteriori* para cada parâmetro (PAULINO *et al.*, 2018).

Os métodos MCMC, apesar de serem amplamente utilizados, apresentam um tempo computacional muito elevado e, em alguns casos, não atingem convergência. Como alternativa aos métodos MCMC, Rue, Martino e Chopin (2009) propuseram a aproximação de Laplace aninhada integrada (INLA), que propicia maior eficiência computacional e pode ser utilizada no contexto de modelos gaussianos latentes.

2.4.1 Aproximação INLA

Um dos pilares do INLA são os modelos gaussianos latentes (MGL), que são formulados de forma similar aos modelos lineares generalizados (MLG). Em um MGL, a variável resposta pertence a uma família de distribuições, em que algum parâmetro μ_i é ligado a um preditor aditivo estruturado η_i por uma função de ligação $g(\cdot)$, em que $g(\mu_i) = \eta_i$ (BLANGIARDO; CAMELETTI, 2015).

O preditor linear contém todas as covariáveis e pode ser escrito como

$$\eta_i = \alpha + \sum_{k=1}^m \beta_k x_{ki} + \sum_{j=1}^n f^{(j)}(z_{ji} + \epsilon_i),$$

em que :

- α é um escalar que representa o intercepto;
- β_k representam o efeito linear das variáveis explicativas x_k ;
- $f^{(j)}(\cdot)$ são efeitos não lineares das covariáveis e podem assumir diferentes formas.

De acordo com Natário (2013), para todos os parâmetros do preditor (as variáveis latentes α , β_k e $f^{(j)}$), podem ser atribuídas distribuições *a priori* gaussianas, assumindo-se ainda que são condicionalmente independentes dois a dois, dados os restantes.

No contexto de modelos espaço-temporais, as funções $f^{(j)}(\cdot)$ podem assumir as seguintes formas (NATÁRIO, 2013):

- efeitos espaciais estruturados $f^{(1)}(\cdot) \sim CAR$;
- efeitos espaciais não estruturados $f^{(2)}(\cdot) \sim Normal(0, \sigma_{f^{(2)}}^2)$;
- efeitos temporais $f^{(3)}(\cdot) \sim AR(\phi, \sigma_{f^{(3)}}^2)$;
- efeitos temporais não estruturados $f^{(4)}(\cdot) \sim Normal(0, \sigma_{f^{(4)}}^2)$.

De acordo com Rue, Martino e Chopin (2009), os MGLs podem ser definido em três estágios:

- Equação de observações

$$y|x, \theta_1 \sim h(y|x, \theta_1)$$

- Campo gaussiano latente

$$x|\theta_2, y \sim Normal(\mu(\theta_2), Q^{-1}(\theta_2))$$

- Hiperparâmetros

$$\theta \sim h(\theta)$$

em que $Q^{-1}(\theta_2)$ é a matriz de precisão do processo gaussiano do vetor x .

O objetivo da abordagem INLA é estimar as distribuições marginais *a posteriori*

$h(x_i|y)$ e $h(\theta_i|y)$:

$$h(x_i|y) = \int h(x_i|\theta, y)h(\theta|y)d\theta$$

$$h(\theta_j|y) = \int h(\theta|y)d\theta_{-j}$$

sendo assim, utilizam-se as aproximações de Laplace para obter as distribuições aproximadas.

A aproximação INLA de $h(x_i | \mathbf{y})$, é feita da seguinte forma:

1. Primeiramente, é realizada uma aproximação para a *posteriori* conjunta:

$$h(\theta | \mathbf{y}) \approx \frac{h(x, \theta, \mathbf{y})}{\tilde{h}(x | \mathbf{y}), \theta}$$

em que $\tilde{h}(x | \mathbf{y}), \theta$ é a aproximação gaussiana para $h(x | \mathbf{y}), \theta$, que corresponde à distribuição teórica da moda.

2. A seguir, é realizada a aproximação de Laplace de $h(x_i | \mathbf{y}), \theta$ para valores fixados de θ
3. Por fim, combinam-se os dois passos anteriores usando integração numérica:

$$\tilde{h}(x_i | \mathbf{y}) = \sum_m \tilde{h}(x_i | \theta_m, \mathbf{y}) \tilde{h}(\theta_m | \mathbf{y}) \Delta_m.$$

2.4.2 Seleção de modelos

Os métodos de seleção de modelos servem para avaliar qual dos modelos tem menor erro na estimação das observações. Assim, o intuito é encontrar o modelo que melhor se ajusta aos dados disponíveis e comparar vários modelos distintos. Em geral são utilizadas metodologias baseadas na distribuição preditiva e na função *deviance*, que são sucintamente apresentadas a seguir (BLANGIARDO; CAMELETTI, 2015).

De acordo com Blangiardo e Cameletti (2015), os métodos baseados na distribuição preditiva consistem em dividir a amostra de observações \mathbf{y} em dois grupos:

- y_f – utilizado para o ajuste do modelo e para estimar a distribuição *a posteriori* dos parâmetros;
- y_c – utilizado para realizar crítica do modelo, ou seja, verificar a sua plausibilidade.

A avaliação da performance preditiva do modelo pode ser feita por meio da validação cruzada *Leave-One-Out*, em que $y_f = y_{-i}$ e $y_c = y_i$. Assim, dois índices podem ser utilizados: a ordenada preditiva condicional (*conditional predictive ordinate – CPO*) e a transformação integral de probabilidade (*probability integral transform – PIT*) (BLANGIARDO; CAMELETTI, 2015).

A CPO (PETTIT, 1990) é a probabilidade *a posteriori* de verificar uma observação quando o modelo é ajustado usando todos os dados, exceto y_i :

$$CPO_i = p(y_i|y_f)$$

em que os valores altos indicam um melhor ajuste do modelo aos dados, enquanto valores baixos indicam um ajuste inadequado do modelo a essa observação e possivelmente a presença de *outliers*. Uma medida que resume a CPO é

$$LPML = - \sum_i^n \log(CPO_i)$$

com valores menores indicando um melhor ajuste do modelo.

A quantidade PIT (MARSHALL; SPIEGELHALTER, 2003) mede, para cada observação, a probabilidade de um novo valor ser inferior ao valor real observado:

$$PIT_i = p(y \leq y_i|y_f)$$

Se os dados forem discretos, a PIT ajustada é calculada como

$$PIT_i^{adj} = PIT_i - 0.5CPO_i$$

Assim, se o modelo se ajusta bem aos dados, a distribuição dos diferentes valores deve ser próxima a uma distribuição uniforme entre 0 e 1.

O critério de informação do desvio (DIC) proposto por Spiegelhalter *et al.* (2002) é um critério popular para escolha de modelos bayesianos semelhante ao critério de Akaike (AIC) (SAKAMOTO; ISHIGURO; KITAGAWA, 1986). Esse critério leva em consideração a adequação do ajuste e um termo de penalização que denota a complexidade do modelo por meio do número efetivo de parâmetros estimados, conforme indica a fórmula abaixo:

$$DIC = D(\underline{\theta}) + 2p_D$$

em que

- $D(\underline{\theta})$ é a média *a posteriori* da *deviance* do modelo;
- p_D é o número efetivo de parâmetros dada por:

$$p_D = \underline{D} - D(\underline{\theta}),$$

Dado um conjunto de modelos candidatos para os dados, o modelo preferido é aquele com o menor valor de DIC.

3.

Metodologia





3. Metodologia

3.1 Modelos para casos e óbitos

3.1.1 Dados

Os dados são referentes às coletas diárias, realizadas pelo Ministério da Saúde, de casos e óbitos de Covid-19 de todos os 5.570 municípios brasileiros, no período de 27 de março a 03 de outubro de 2020.

Inicialmente os dados foram agregados por semana epidemiológica, com início no domingo e término no sábado. Para cada município, foi verificado se o número de casos ou óbitos em uma semana epidemiológica específica possuíam valores negativos. Nesses casos, tais valores foram zerados e subtraídos da semana anterior.

Assim, o banco de dados foi composto de número de casos e óbitos em 27 semanas para cada município, totalizando assim 1.503.980 observações.

3.1.2 Modelagem

Os modelos espaço-temporais foram ajustados aos dados utilizando a metodologia INLA, em que foram testadas diferentes distribuições de probabilidade com a variável

resposta sendo número de casos e óbitos de Covid-19. Para a seleção do modelo que melhor se adequa às variáveis de interesse, foram utilizados o critério DIC e LMPL e o histograma das transformações integrais de probabilidade.

As distribuições de probabilidade testadas foram:

- Poisson

$$Y_{i,t}|\theta_i \sim \text{Poisson}(E_{i,t}\theta_{i,t})$$

- Poisson Hurdle (HP)

$$Y_{i,t}|\theta_{i,t}, \pi_{i,t} \sim \text{HP}(E_{i,t}\theta_{i,t}, \pi_{i,t})$$

- Poisson inflado de zero (ZIP)

$$Y_{i,t}|\theta_{i,t}, \pi_{i,t} \sim \text{ZIP}(E_{i,t}\theta_{i,t}, \pi_{i,t})$$

- Binomial negativo (BN)

$$Y_i|\theta_i \sim \text{BN}(E_i\theta_i, \tau)$$

- Binomial negativo Hurdle (HBN)

$$Y_{i,t}|\theta_{i,t}, \pi_{i,t} \sim \text{HBN}(E_i\theta_i, \tau, \pi_{i,t})$$

- Binomial negativo inflado de zero (ZIBN)

$$Y_{i,t}|\theta_{i,t}, \pi_{i,t} \sim \text{ZIBN}(E_i\theta_i, \tau, \pi_{i,t})$$

O número esperado de casos e óbitos foi dado pelos valores da semana anterior:

$$E_{i,t} = Y_{i,t-1} + 1$$

Assim, θ_i representa o risco de novos casos ou óbitos, e o preditor linear do modelo não paramétrico aditivo espaço-temporal foi definido como (KNORR-HELD, 2000)

$$\eta_{i,t} = \log(\theta_{i,t}) = \alpha + \mu_i + \nu_i + \gamma_t + \phi_t + \delta_{it},$$

com $i = 1, \dots, 5570$ representando cada município brasileiro e $t = 2, \dots, 27$ representando as semanas epidemiológicas no período de 27 de março a 03 de outubro de 2020.

O efeito aleatório referente a cada município é representado pelo modelo de BYM (BESAG; YORK; MOLLIE, 1991) e possui duas componentes:

- μ_i efeito estruturado no espaço;
- ν_i efeito não estruturado no espaço.

Para o efeito estruturado espacialmente μ_i , aplica-se uma estrutura espacial autorregressiva condicional intrínseca (iCAR)

$$\mu_i | \mu_{j \neq i} \sim \text{Normal} \left(\frac{1}{m_i} \sum_j w_{ij} \mu_j, \frac{1}{m_i \tau_\mu} \right),$$

em que m_i é o número de regiões vizinhas, w_{ij} assumem valores iguais a 1 se as áreas i e j são vizinhas e valores nulos caso contrário. A distribuição *a priori* do hiperparâmetro de precisão τ_μ é definida para $\log(\tau_\mu) \sim \text{loggamma}(1,5 \times 10^{-5})$.

Foram testadas diferentes estruturas de vizinhança:

- Fronteira – tipo *queen*.
- Por distância:
 - até 50km;
 - até 100km;
 - até 150km;
 - até 200km;
 - até 250km.

Para o efeito espacial não estruturado, assume-se $\nu_i \sim \text{Normal}(0, \frac{1}{\tau_\nu})$, e a distribuição *a priori* do hiperparâmetro de precisão τ_ν sendo definida para $\log(\tau_\nu) \sim \text{loggamma}(1,5 \times 10^{-5})$.

O efeito aleatório referente às semanas epidemiológicas possui duas componentes:

- γ_t efeito estruturado no tempo;
- ϕ_t efeito não estruturado no tempo.

Para o efeito estruturado no tempo γ_t foram testadas duas estratégias de modelagem com passeios aleatórios:

- passeio aleatório de primeira ordem RW1 – indicando dependência de uma semana:

$$\gamma_t | \gamma_{t-1} \sim \text{Normal} \left(\gamma_{t-1}, \frac{1}{\tau_\gamma} \right);$$

- passeio aleatório de segunda ordem RW2 – indicando dependência de duas semanas:

$$\gamma_t | \gamma_{t-1}, \gamma_{t-2} \sim \text{Normal} \left(2\gamma_{t-1} + \gamma_{t-2}, \frac{1}{\tau_\gamma} \right),$$

onde a distribuição *a priori* do hiperparâmetro de precisão τ_γ é definida para $\log(\tau_\gamma) \sim \text{loggamma}(1,5 \times 10^{-5})$.

O parâmetro não estruturado no tempo assume $\phi_t \sim \text{Normal}(0; \tau_\phi)$, com a *a priori* do hiperparâmetro de precisão τ_ϕ definida para $\log(\tau_\phi) \sim \text{loggamma}(1,5 \times 10^{-5})$.

O efeito aleatório δ_{it} representa interação entre espaço e tempo, e assume $\delta_{it} \sim \text{Normal}(0; \tau_\delta)$, com a *a priori* do hiperparâmetro de precisão τ_δ sendo definida para $\log(\tau_\delta) \sim \text{loggamma}(1,5 \times 10^{-5})$.

3.2 Modelos com estrutura de Saúde

3.2.1 Dados do CNES

Os dados do CNES foram obtidos com respeito à infraestrutura de saúde, utilizando o pacote *microdatasus* (SALDANHA; BASTOS; BARCELLOS, 2019) do software R (R Core Team, 2020), com as informações referentes a todos os municípios brasileiros no período de fevereiro a setembro de 2020.

Assim, foram utilizados os dados dos estabelecimentos (CNES-ST), sendo considerados apenas a quantidade de postos de saúde e centros de saúde ou unidades básicas de saúde (BRASIL, 2020a). Cada estabelecimento no CNES detém um código como descrito abaixo:

- TP_UNID01 – posto de saúde: unidade destinada à prestação de assistência a uma determinada população, de forma programada ou não, por profissional de nível médio, com a presença intermitente ou não do profissional médico.
- TP_UNID02 – centro de saúde/unidade básica de saúde: unidade para realização de atendimentos de atenção básica e integral a uma população, de forma programada ou não, nas especialidades básicas, podendo oferecer

assistência odontológica e de outros profissionais de nível superior. A assistência deve ser permanente e prestada por médico generalista ou especialista nessas áreas, podendo ou não oferecer: serviços auxiliares de diagnose e terapia (SADT) e pronto atendimento de 24 horas.

Também foram utilizados os dados referentes aos leitos (CNES-LT), sendo considerados leitos clínicos e de UTI Covid, de acordo com tabela de *Habilitações e Leitos* (BRASIL, 2020a).

- CODLEITO33 – leitos clínica geral: leito de internação hospitalar destinado a acomodar pacientes de qualquer especialidade clínica.

Para determinar a força de trabalho em cada município, foram utilizados os dados de profissionais (CNES-PF) e obtidas as quantidades totais de médicos, enfermeiros e técnicos de enfermagem (BRASIL, 2020a). Dentro desse banco de dados, cada profissional é identificado pelo seu código na Classificação Brasileira de Ocupações (CBO):

- Médicos: foram considerados os profissionais com CBO iniciado com 2231.
- Enfermeiros: foram considerados os profissionais com CBO iniciado com 2235;
- Técnicos de enfermagem: foram considerados os profissionais com CBO iniciado com 3222.

Dessa forma, foram consideradas as seguintes variáveis da estrutura de saúde e força de trabalho:

- número de postos de saúde (NPS);
- número de centros de saúde/unidade básica (NCS);
- número de leitos clínicos (NLC);
- número de médicos (NM);
- número de enfermeiros e técnicos de enfermagem (NE).

Todas as variáveis obtidas no CNES foram transformadas em taxas por 10.000 habitantes, da seguinte forma:

$$TPS = \frac{10000NPS}{P}$$

em que P é a população do município.

Como os dados obtidos no CNES são mensais e os dados de casos e óbitos são semanais, os valores das taxas de estrutura de saúde e força de trabalho foram transformados em semanais utilizando a seguinte estratégia:

- Para cada semana epidemiológica, foram considerados os valores das taxas do mês anterior. Por exemplo, para a semana 15, que compreende o período de 05 de abril a 11 de abril, foram utilizadas as taxas de estrutura de saúde e força de trabalho referentes ao mês de março.
- Quando uma semana epidemiológica continha dias provenientes de dois meses distintos, foi calculada uma média ponderada. Por exemplo, para a semana 40, que compreende o período de 27 de setembro a 03 de outubro, foram utilizadas as taxas de estrutura de saúde e força de trabalho referentes aos meses de agosto e setembro:

$$\frac{4T_{agosto} + 3T_{setembro}}{7}$$

3.2.2 Modelagem

Os modelos espaço-temporais foram ajustados considerando as variáveis respostas números de casos e óbitos, e as variáveis explicativas sendo as taxas de estrutura de saúde e força de trabalho, com o preditor linear não paramétrico aditivo espaço-temporal definido como (KNORR-HELD, 2000):

$$\eta_{i,t} = \log(\theta_{i,t}) = \alpha + \sum_i \beta_i X_i + \mu_i + \nu_i + \gamma_t + \phi_t + \delta_{it}$$

onde $i = 1, \dots, 5570$ representam cada município brasileiro e $t = 2, \dots, 27$ representam as semanas epidemiológicas compreendidas no período de 27 de março a 03 de outubro de 2020. As quantidades X_i 's representam os efeitos das taxas de estrutura de saúde e força de trabalho.

A demais componentes utilizadas foram idênticas àquelas utilizadas na estratégia de modelagem anterior.

3.3 Análises estatísticas

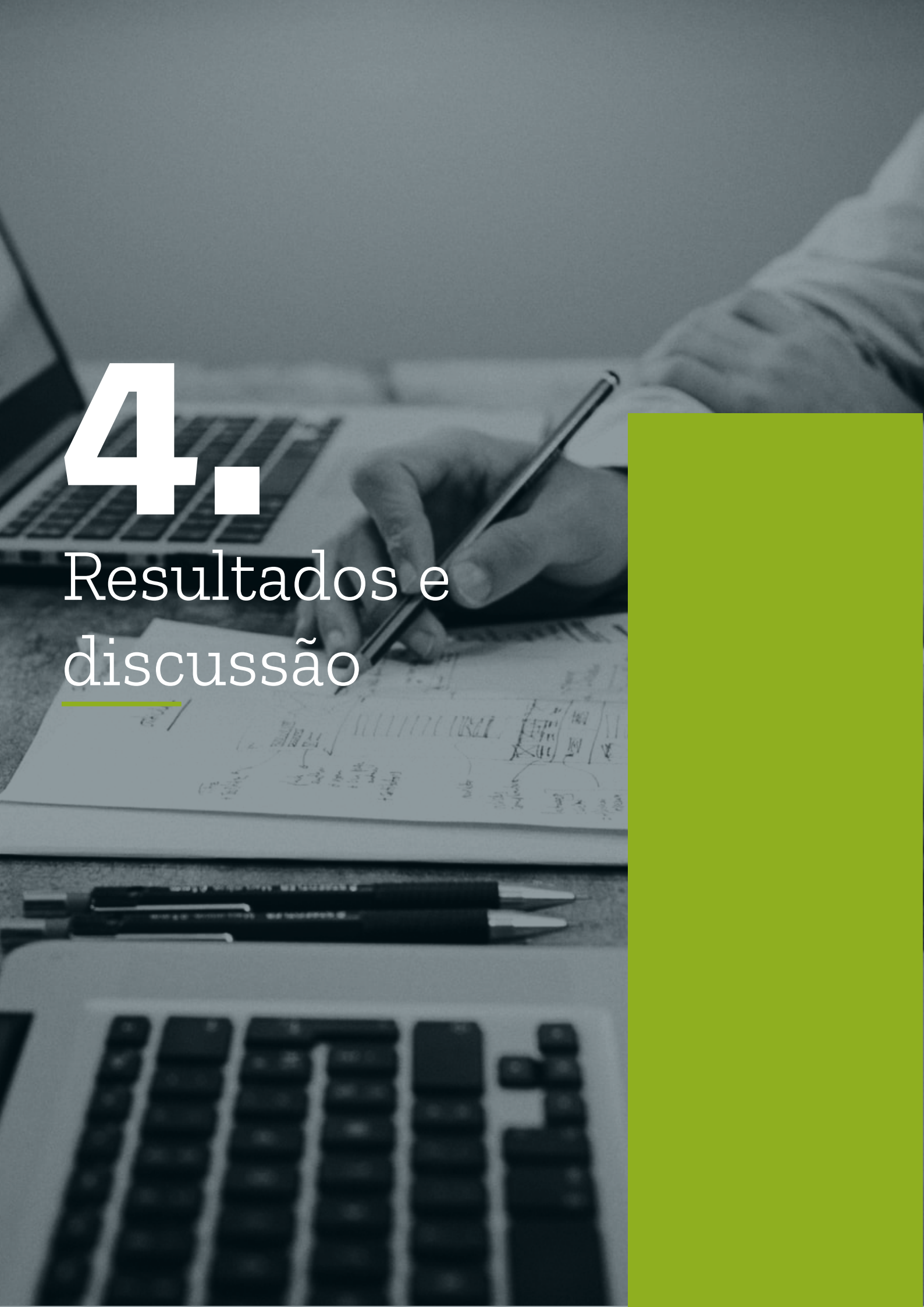
Todas as análises estatísticas foram implementadas no software R (R CORE TEAM, 2020). Para a aplicação da metodologia INLA, foi utilizado o pacote *INLA* (RUE; MARTINO; CHOPIN, 2009; LINDGREN; RUE, 2015). Para obtenção do mapa do Brasil, foi utilizado o pacote *brazilmaps* (PRADO SIQUEIRA, 2020) e a definição da estrutura espacial de vizinhança foi feita por meio do pacote *spdep* (BIVAND; PEBESMA; GOMEZ-RUBIO, 2013).

3.4 Painel de visualização

A partir dos modelos definidos para casos e óbitos, foi desenvolvido um painel de visualização com as informações básicas e os indicadores de risco para cada município brasileiro. Esse painel foi construído utilizando o pacote *shiny* (CHANG *et al.*, 2020) do R Core Team (2020), que cria ferramentas interativas, permitindo que possamos compartilhar os resultados das análises em aplicativos publicáveis na internet, integralmente e de forma gratuita. As funcionalidades disponíveis para o desenvolvedor permitem a criação de aplicativos úteis, fáceis de usar e esteticamente agradáveis, além de amigáveis para o usuário. Outra funcionalidade do *shiny* é o desenvolvimento de *dashboards*. Além disso, é possível utilizar outras linguagens além de R, como HTML e JavaScript, para configurar a interface desse aplicativo.

4.

Resultados e discussão





4. Resultados e discussão

4.1 Modelos para casos

Na Tabela 1, são apresentados os valores do critério de informação de *deviance* (DIC), bem como os valores da pseudo log-verossimilhança marginal (LPML) dos modelos espaço-temporais para casos com diferentes distribuições de probabilidade e estruturas temporais, considerando a vizinhança do tipo *queen*, em que se verifica que o melhor modelo foi aquele com distribuição de probabilidade binomial negativa Hurdle e efeito temporal de 2ª ordem (BNH2), uma vez que este apresenta os menores valores de $DIC = 817846,6$ e $LPML = 410554,4$.

Tabela 1 – Valores de DIC e LPML dos modelos espaço-temporais para casos com diferentes modelos de probabilidade e estruturas temporais, e com vizinhança do tipo *queen*

Modelo	Efeito Temporal	DIC	LPML
Poisson	Sem efeito	2014436,2	968956,4
	Efeito de 1ª ordem	2014436,1	968936,5
	Efeito de 2ª ordem	2014435,7	968936,2
Poisson inflado de zero	Sem efeito	2001007,0	961921,0
	Efeito de 1ª ordem	2001006,7	961845,4
	Efeito de 2ª ordem	2001005,1	961842,8
Poisson Hurdle	Sem efeito	1956493,7	939734,7
	Efeito de 1ª ordem	1956493,4	939732,2
	Efeito de 2ª ordem	1956493,1	939681,7
Binomial negativa	Sem efeito	883350,0	443300,0
	Efeito de 1ª ordem	875654,3	439598,5
	Efeito de 2ª ordem	875652,1	439596,9
Binomial negativa inflado de zero	Sem efeito	838829,7	421358,2
	Efeito de 1ª ordem	838824,9	421281,5
	Efeito de 2ª ordem	838824,7	421280,2
Binomial negativa Hurdle	Sem efeito	817847,7	410560,1
	Efeito de 1ª ordem	817847,1	410559,7
	Efeito de 2ª ordem	817846,6	410554,4

Fonte: elaboração própria.

Foram testados diferentes modelos espaço-temporais com estruturas de vizinhança espaciais distintas. Entre as estruturas de vizinhança exploradas, o modelo BNH2 apresentou a melhor performance. Sendo assim, a estrutura de tipo *queen* apresentou melhor desempenho entre as estruturas de vizinhança apresentadas (Tabela 2).

Tabela 2 – Valores de DIC e LPML dos modelos espaço-temporais para casos com distribuição binomial negativa Hurdle, efeito temporal de 2ª ordem e diferentes estruturas de vizinhança

Tipo de Vizinhança	DIC	LPML
<i>Queen</i>	817846,6	410554,4
Até 50km	886199,1	443635,7
Até 100km	878939,5	440046,4
Até 150km	821086,7	411022,3
Até 200km	821086,4	411024,7
Até 250km	821089,2	411021,3

Fonte: elaboração própria.

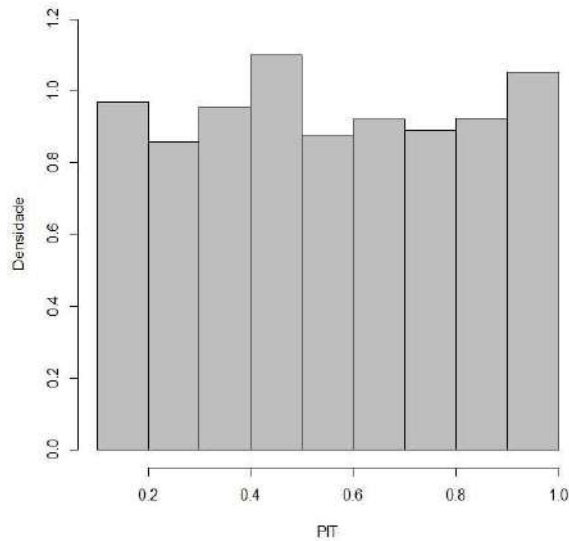
Dos resultados das Tabelas 1 e 2, conclui-se que o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen* adere melhor ao risco de aumento de casos de Covid-19. Assim, a ocorrência de zeros é estrutural, indicando que os valores provêm de fonte não aleatória ou não fazem parte do contexto de casos. O efeito temporal indica que o risco depende dos valores observados em até duas semanas anteriores e, para um determinado município, esse risco depende dos valores observados em seus municípios vizinhos com fronteira geográfica.

Para verificar a boa calibração do modelo espaço-temporal BNH2 com vizinhança do tipo *queen*, foi construído o gráfico de histograma da transformação integral de probabilidade (Figura 1), em que é possível observar um comportamento próximo da distribuição uniforme, indicando que o modelo está bem calibrado. Assim, apresenta-se um resumo dos efeitos do modelo na Tabela 3, e, a partir do intervalo de credibilidade de 95%, verifica-se que todos os parâmetros e hiperparâmetros desse modelo são significativos.

A Figura 2) a) apresenta a média *a posteriori* do efeito espacial para cada município brasileiro. Observa-se que os municípios das regiões Centro-Oeste e Norte apresentam os efeitos espaciais mais elevados.

Na Figura 2) b), é apresentada a probabilidade *a posteriori* de que o efeito espacial seja maior do que 1, que auxilia no mapeamento das incertezas associadas às médias *a posteriori* da Figura 2. Assim, verifica-se que os municípios das Regiões Centro-Oeste e Norte apresentaram probabilidades altas de os efeitos espaciais serem maiores do que 1, indicando que, dentro da sua vizinhança, estes têm contribuição forte para a determinação do risco de aumento de casos da Covid-19.

Figura 1 – Gráfico de histograma da transformação integral de probabilidade (PIT) do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo queen



Fonte: elaboração própria.

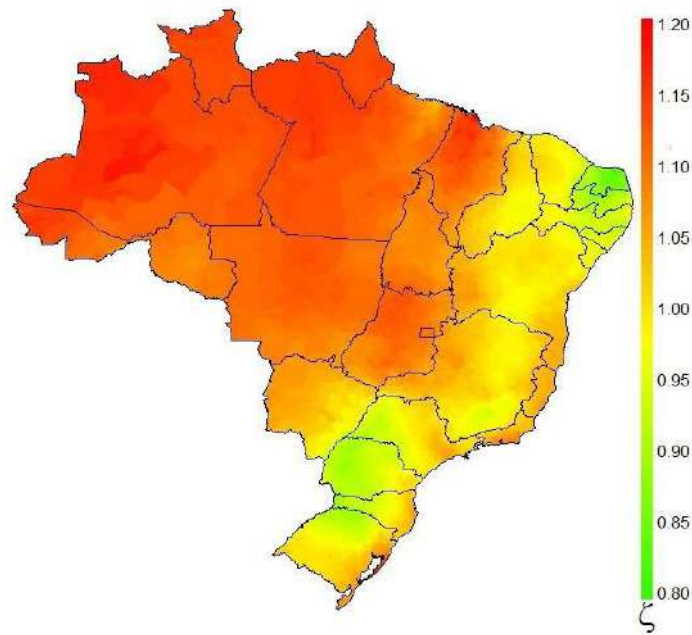
Tabela 3 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo queen

Parâmetro	Média	Moda	Desvio-padrão	Intervalo de Credibilidade (95%)	
	<i>a posteriori</i>	<i>a posteriori</i>	<i>a posteriori</i>	L.I.	L.S.
Intercepto	-0,020	-0,020	0,002	-0,007	-0,024
Hiperparâmetros					
θ	64,141	58,780	11,610	46,637	97,319
π	0,350	0,350	0,001	0,348	0,351
τ_{ν}	9249,027	7862,923	3509,741	3769,043	17310,657
τ_{μ}	279,533	258,320	29,029	237,006	349,446
τ_{γ}	83,250	72,639	25,097	45,095	142,591
τ_{ϕ}	13477,069	3483,212	13741,598	1234,779	50492,615
τ_{δ}	1,426	1,430	0,011	1,396	1,453

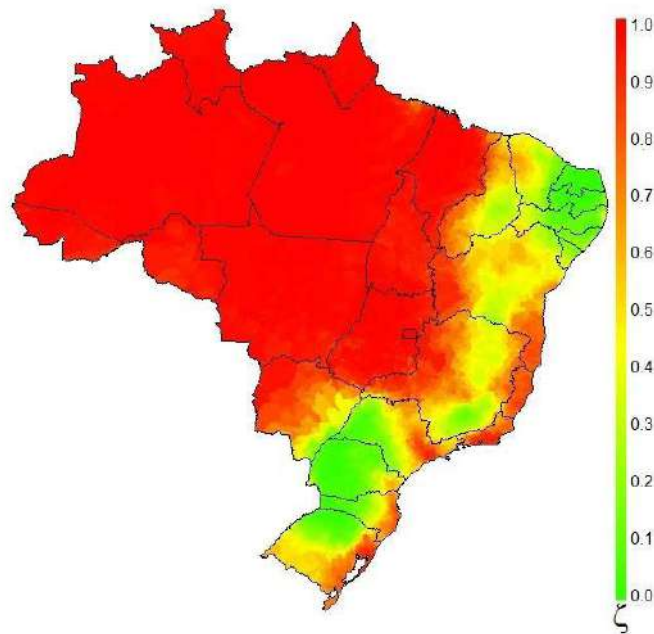
Fonte: elaboração própria.

Figura 2 – Médias (a) e probabilidades (b) *a posteriori* dos efeitos espaciais (ζ_i) para cada município brasileiro, considerando o modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

a) médias



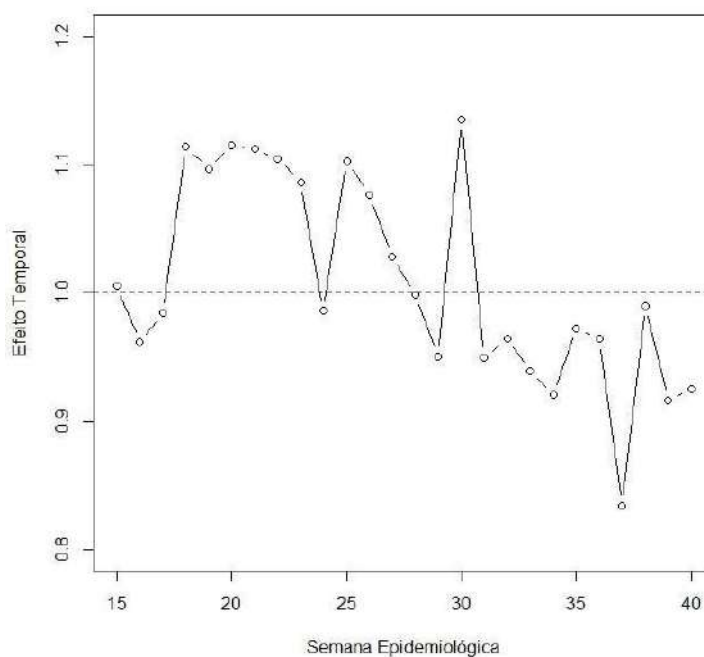
b) probabilidades



Fonte: elaboração própria.

A média *a posteriori* da tendência temporal para o período compreendido entre a 15ª e a 40ª semanas epidemiológicas é apresentada na Figura 3, segundo o modelo espaço-temporal com distribuição de probabilidade BNH2 e estrutura de vizinhança do tipo *queen*. Em geral, houve um efeito positivo do tempo até a 30ª (período de 19 de julho a 25 de julho de 2020), indicando um crescimento dos casos de Covid-19. E, após essa semana, o efeito médio foi inferior a 1, indicando uma diminuição do risco de aumento de casos.

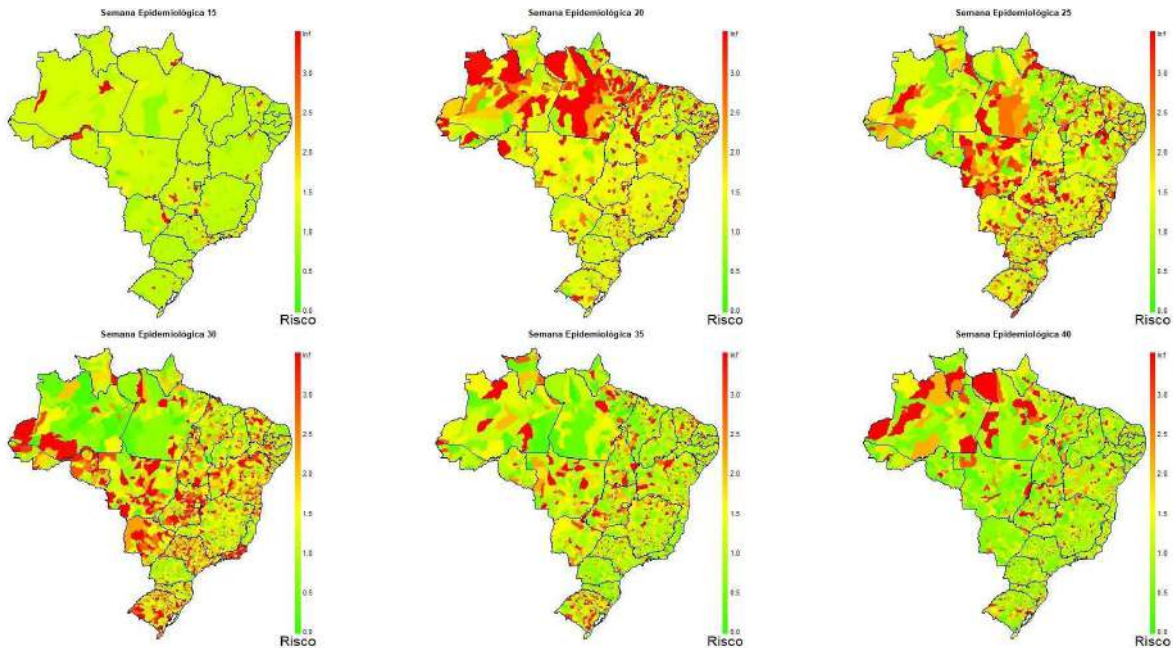
Figura 3 - Tendência temporal da média *a posteriori* do risco de aumentos de casos, considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança tipo *queen*



Fonte: elaboração própria.

Na Figura 4, são apresentados os valores de risco de aumento da Covid-19 para algumas semanas epidemiológicas. Os municípios com coloração mais próxima de verde são aqueles que apresentam risco abaixo de 1, indicando que não há risco de aumento de casos. Para melhor interpretação do risco de aumento da doença, a Figura 5 apresenta os valores das prevalências de casos de Covid-19.

Figura 4 – Risco de aumento de Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (27 de setembro a 03 de outubro de 2020), considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança do tipo *queen*

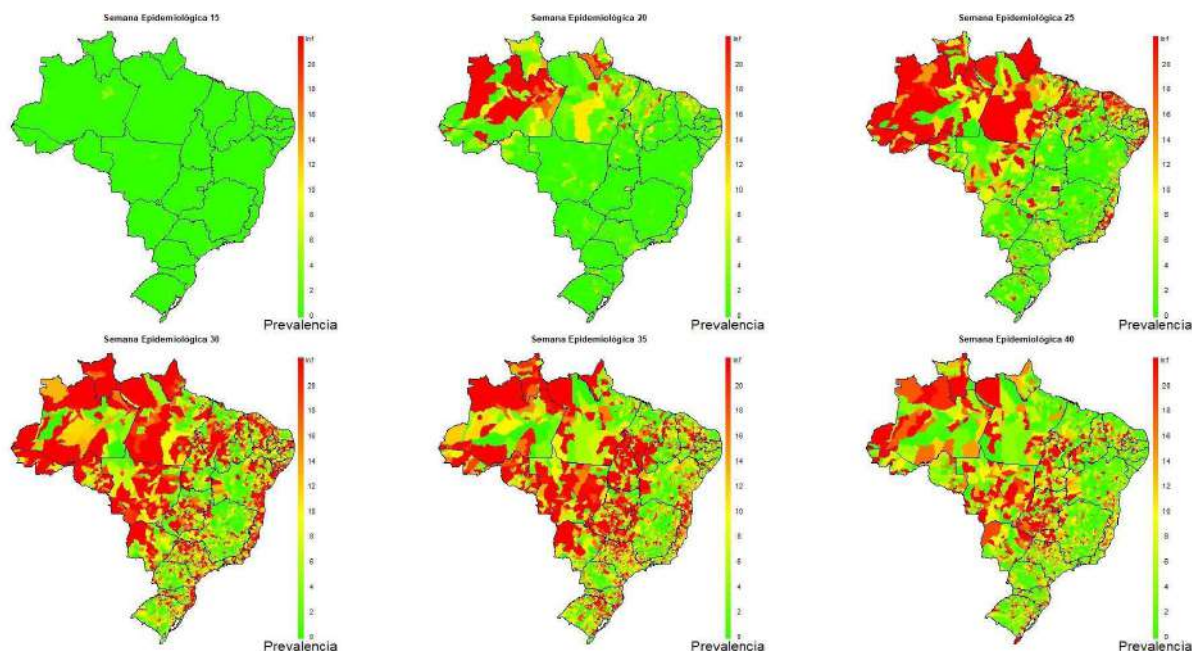


Fonte: elaboração própria.

Assim, por meio das Figuras 5 e 6, observa-se que:

- Na 15ª semana epidemiológica (período de 05 de abril a 11 de abril de 2020), muitos municípios apresentavam prevalência baixa (cor verde), e risco médio (cor amarela).
- Na 25ª semana epidemiológica (período de 14 de junho a 20 de junho de 2020), os municípios da região Norte apresentavam prevalência alta (cor vermelha) e risco médio ou alto (cores amarela e vermelha).
- Na 40ª semana epidemiológica (período de 27 de setembro a 03 de outubro de 2020), a maioria municípios apresentavam prevalência média ou alta (cores amarela e vermelha) e risco baixo ou médio (cores verde e amarela).

Figura 5 – Prevalência de casos de Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



Fonte: elaboração própria.

4.2 Modelos para óbitos

Os valores de DIC e LPML dos modelos espaço-temporais para casos com diferentes distribuições de probabilidade e estruturas temporais, considerando a vizinhança do tipo *queen*, são apresentados na Tabela 4, em que se verifica que o melhor modelo ajustado corresponde a distribuição de probabilidade binomial negativa Hurdle e efeito temporal de 1ª ordem (BNH1), pois esse modelo apresenta os menores valores de $DIC = 973843,3$ e $LPML = 487029,8$.

Assim, os óbitos de Covid-19 também são representados por um modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, porém com efeito temporal de 1ª ordem e vizinhança do tipo *queen*. Dessa forma, observa-se que a não ocorrência de óbitos aparece de forma estrutural e que o risco do aumento de óbitos depende dos valores observados na semana anterior e nos municípios que fazem fronteira.

Na Figura 6, é apresentado o gráfico de histograma da transformação integral de probabilidade (PIT) do modelo espaço-temporal para óbitos com distribuição de probabilidade BNH1 e vizinhança do tipo *queen*, em que é possível observar

um comportamento próximo da distribuição uniforme, mostrando uma boa calibração do modelo. Assim, apresenta-se um resumo dos efeitos do modelo na Tabela 5 e, a partir do intervalo de credibilidade de 95%, verifica-se que todos os parâmetros e hiperparâmetros são significativos.

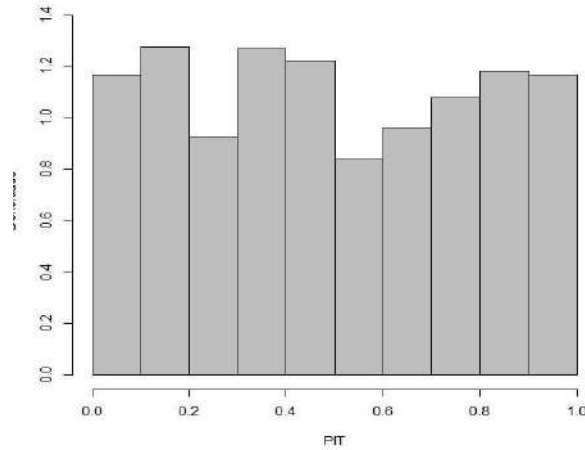
Tabela 4 – Valores de DIC e LPML – nos modelos para óbitos com vizinhança do tipo *queen*

Modelo	Efeito Temporal	DIC	LPML
Poisson	Sem efeito	1524373,7	773002,2
	Efeito de 1ª ordem	1524372,0	772995,7
	Efeito de 2ª ordem	1524379,6	772990,3
Poisson inflado de zero	Sem efeito	1406076,1	717519,4
	Efeito de 1ª ordem	1415007,9	729321,2
	Efeito de 2ª ordem	1405161,4	716474,6
Poisson Hurdle	Sem efeito	1322319,1	667232,9
	Efeito de 1ª ordem	1322317,7	667228,9
	Efeito de 2ª ordem	1322316,8	667224,7
Binomial negativa	Sem efeito	1198656,7	599298,5
	Efeito de 1ª ordem	1198663,5	599299,8
	Efeito de 2ª ordem	1198664,0	599299,9
Binomial negativa inflado de zero	Sem efeito	1198673,6	599305,9
	Efeito de 1ª ordem	1198680,0	599309,2
	Efeito de 2ª ordem	1198690,6	599313,2
Binomial negativa Hurdle	Sem efeito	1345068,9	684174,5
	Efeito de 1ª ordem	973843,3	487029,8
	Efeito de 2ª ordem	973903,9	487055,2

Fonte: elaboração própria.

As médias *a posteriori* do efeito espacial para cada município brasileiro, bem como suas probabilidades *a posteriori* de esses valores serem maiores do que 1, é apresentada nas Figuras 7)a) e 7)b). Assim, verifica-se que os municípios das regiões Nordeste e Sul apresentaram probabilidades baixas de terem os efeitos espaciais maiores do que 1, indicando que, dentro da sua vizinhança, esses valores não apresentam contribuição para a determinação do risco de aumento de óbitos da Covid-19.

Figura 6 – Gráfico de histograma da transformação integral de probabilidade (PIT) do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*



Fonte: elaboração própria.

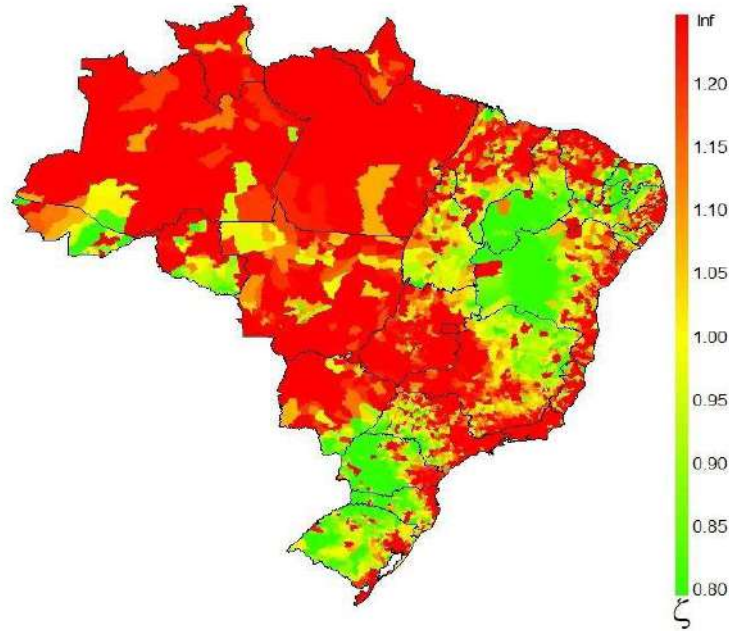
Tabela 5 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*

Parâmetro	Média	Moda	Desvio-padrão	Intervalo de Credibilidade (95%)	
	<i>a posteriori</i>	<i>a posteriori</i>	<i>a posteriori</i>	L.I.	L.S.
intercepto	-0,949	-0,944	0,021	-0,993	-0,911
Hiperparâmetros					
θ	5,103	5,021	0,258	4,652	5,662
π	0,801	0,801	0,001	0,799	0,804
τ_ν	8,342	7,867	0,813	7,082	10,234
τ_μ	3,964	3,827	0,410	3,262	4,871
τ_γ	3624,343	3132,650	10075,394	1075,600	31732,765
τ_ϕ	1172,095	390,274	1016,023	136,399	3831,026
τ_δ	3,617	3,584	0,147	3,350	3,928

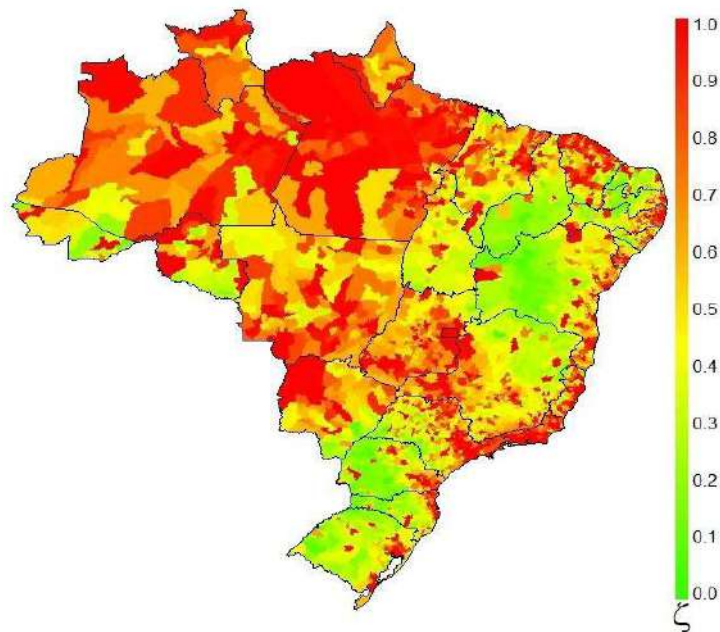
Fonte: elaboração própria.

Figura 7 – Médias (a) e probabilidades (b) *a posteriori* dos efeitos espaciais (ζ_i) para cada município brasileiro, considerando o modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança do tipo *queen*

a) médias



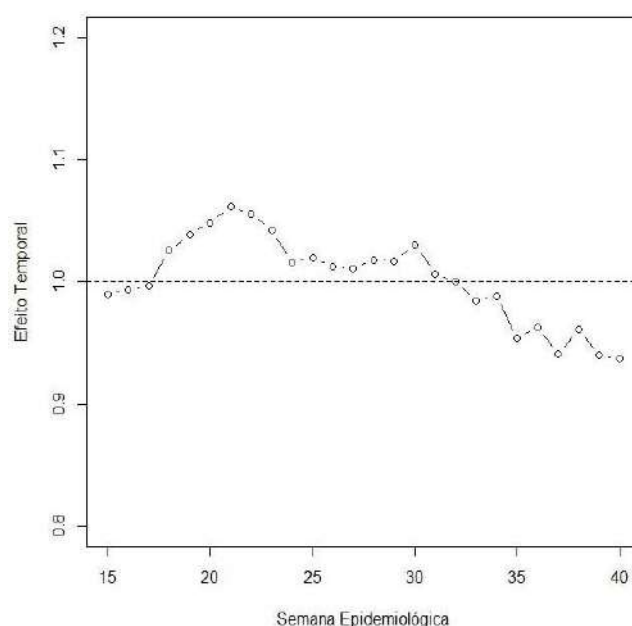
a) probabilidades



Fonte: elaboração própria.

Na Figura 8, é apresentada a média *a posteriori* da tendência temporal para o período compreendido entre a 15ª e a 40ª semanas epidemiológicas. Em geral, houve um efeito positivo do tempo até a 30ª semana (período de 19 de julho a 25 de julho de 2020), indicando um crescimento dos óbitos de Covid-19. E, após essa semana, o efeito médio foi inferior a 1, indicando uma diminuição do risco de aumento de mortes.

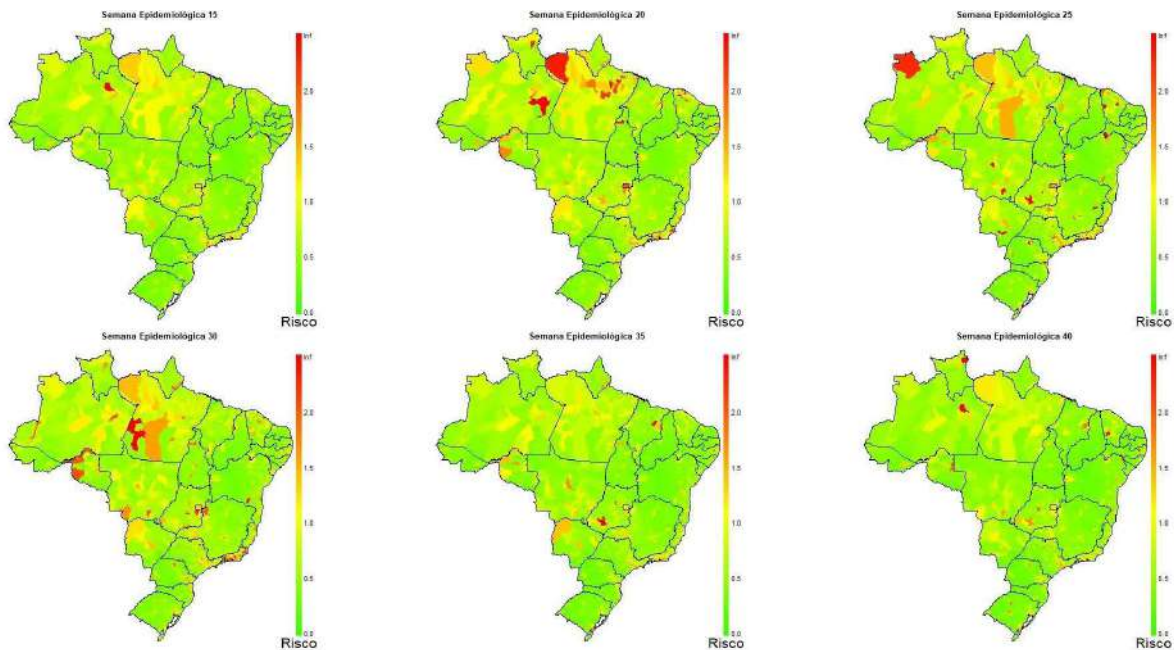
Figura 8 - Tendência temporal da média *a posteriori* do risco de aumentos de óbitos, considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança tipo *queen*



Fonte: elaboração própria.

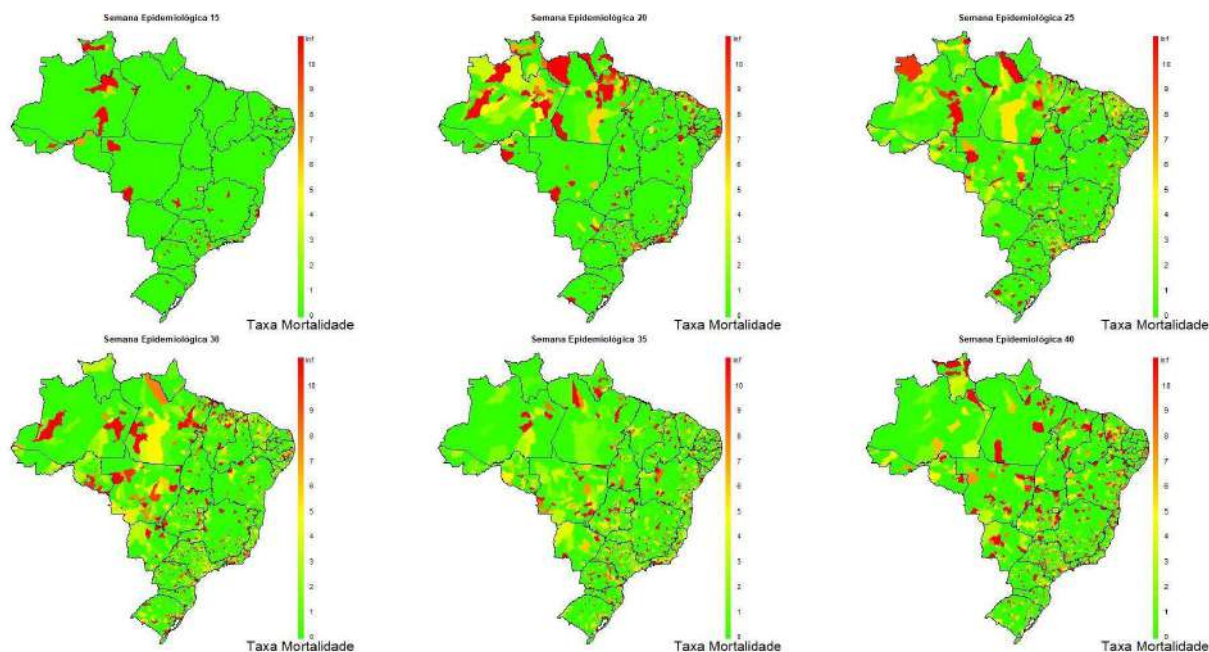
Os valores de risco de aumento de óbitos pela Covid-19 são apresentados para algumas semanas epidemiológicas na Figura 9, e a taxa de mortalidade na Figura 10. Para análise do aumento de óbitos, os municípios com coloração muito próxima de verde são aqueles que apresentam risco abaixo de 1, indicando que não há risco de aumento de casos.

Figura 9 – Risco de aumento de óbitos pela Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (27 de setembro a 03 de outubro de 2020), considerando o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e vizinhança do tipo *queen*



Fonte: elaboração própria.

Figura 10 – Taxa de mortalidade para Covid-19 na 15ª semana (período de 05 de abril a 11 de abril de 2020), 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



Fonte: elaboração própria.

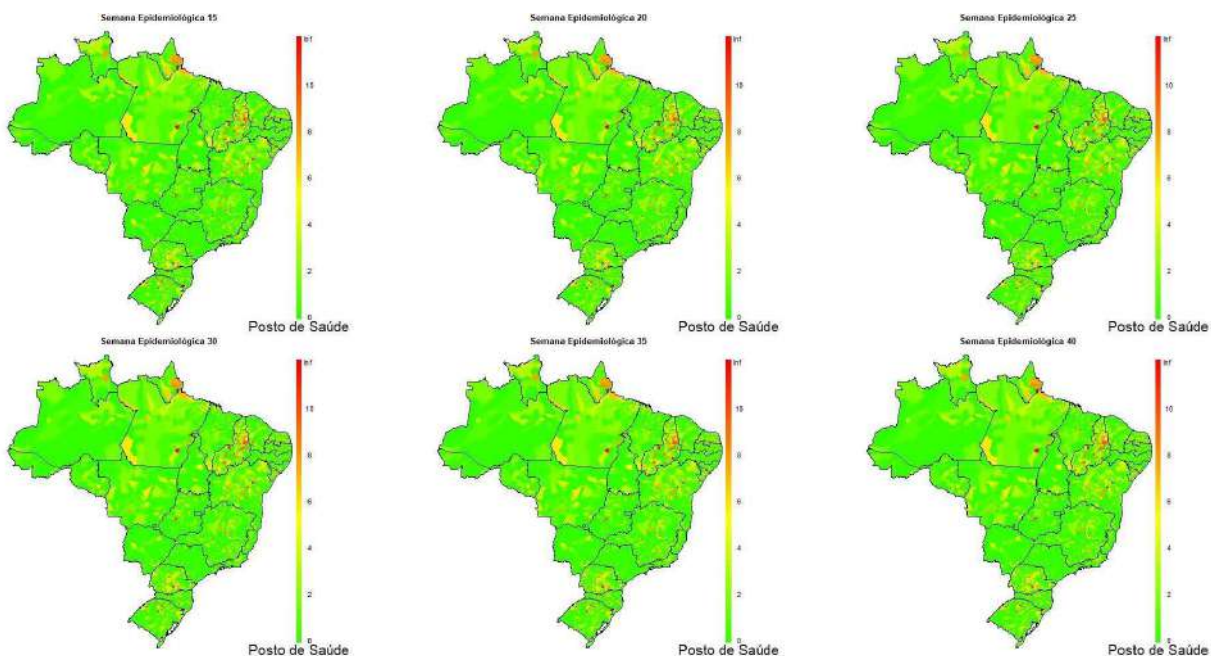
4.3 Modelos com estrutura de saúde

Nas Figuras 11 a 15, são apresentadas as taxas por 10.000 habitantes de postos de saúde, centros de saúde, leitos de clínica geral, médicos, enfermeiros e técnicos de enfermagem. Observa-se que, para todos os indicadores de estrutura de saúde, não houve mudanças ao longo do tempo.

Em relação aos postos de saúde (Figura 11), verifica-se que a maioria dos municípios tem menos de 2 postos por 10.000 habitantes. Quanto aos centros de saúde, com exceção da região Norte, a maioria dos municípios tem mais que 3 centros (Figura 12). Os leitos de clínica geral (Figura 13) tem comportamento similar ao de postos de saúde; a maioria dos municípios tem menos de 2 leitos por 10.000 habitantes.

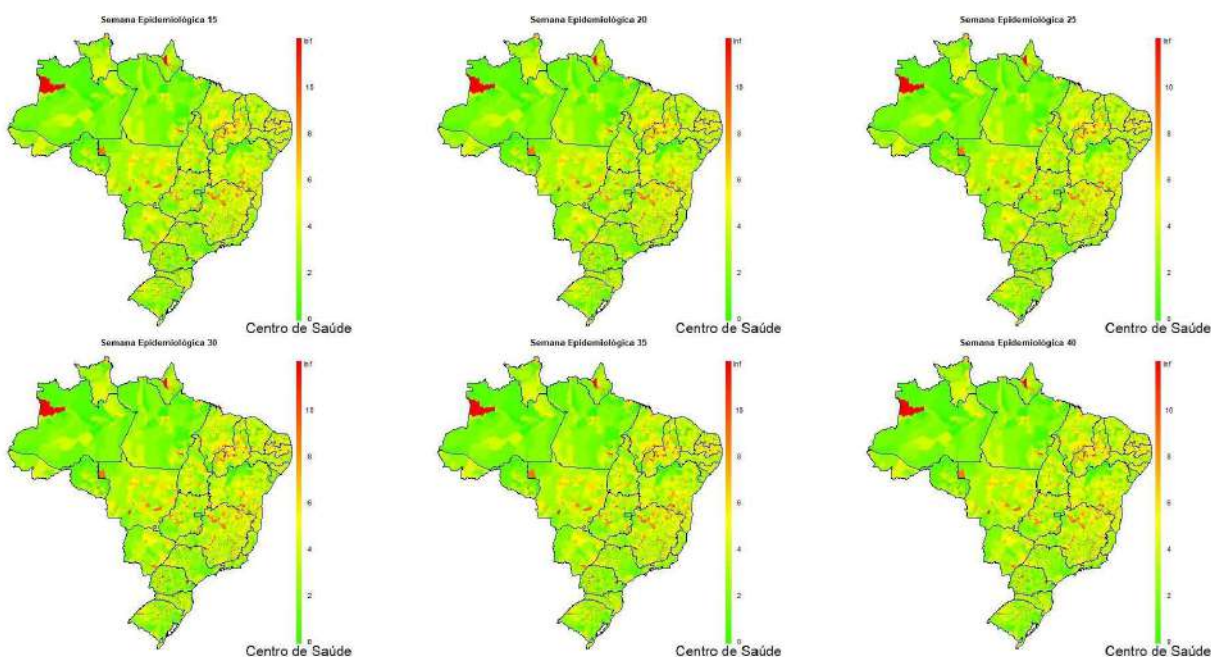
Para os profissionais de saúde, observa-se que o número de médicos por 10.000 habitantes é inferior a 1 para a maioria dos municípios brasileiros (Figura 14). Em relação aos enfermeiros e técnicos de enfermagem (Figura 15), observa-se que a maioria dos municípios tem entre 30 a 60 profissionais de enfermagem por 10.000 habitantes.

Figura 11 – Taxas de postos de saúde por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



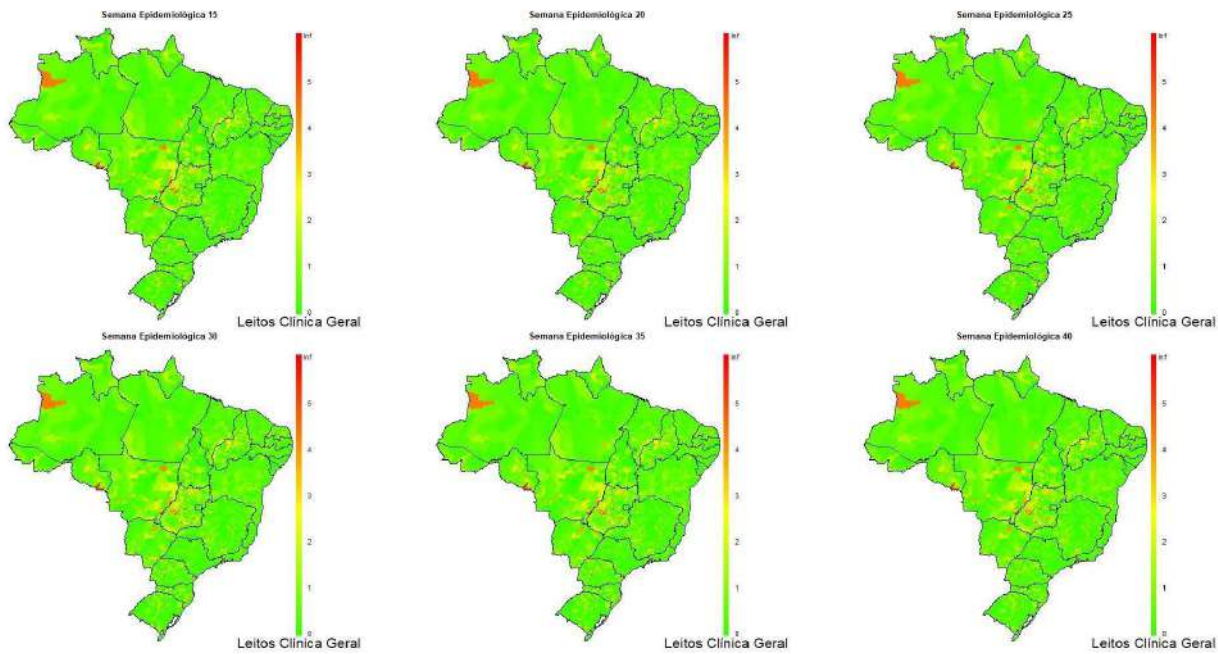
Fonte: elaboração própria.

Figura 12 – Taxas de centros de saúde por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



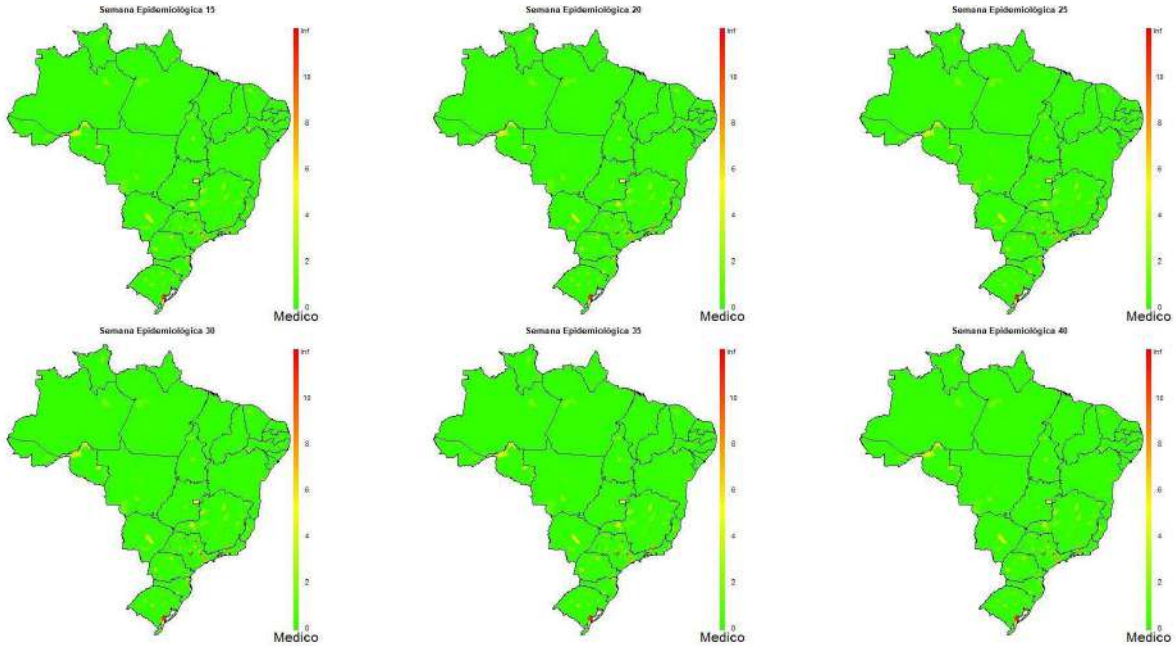
Fonte: elaboração própria.

Figura 13 – Taxas de leitos em clínica geral por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



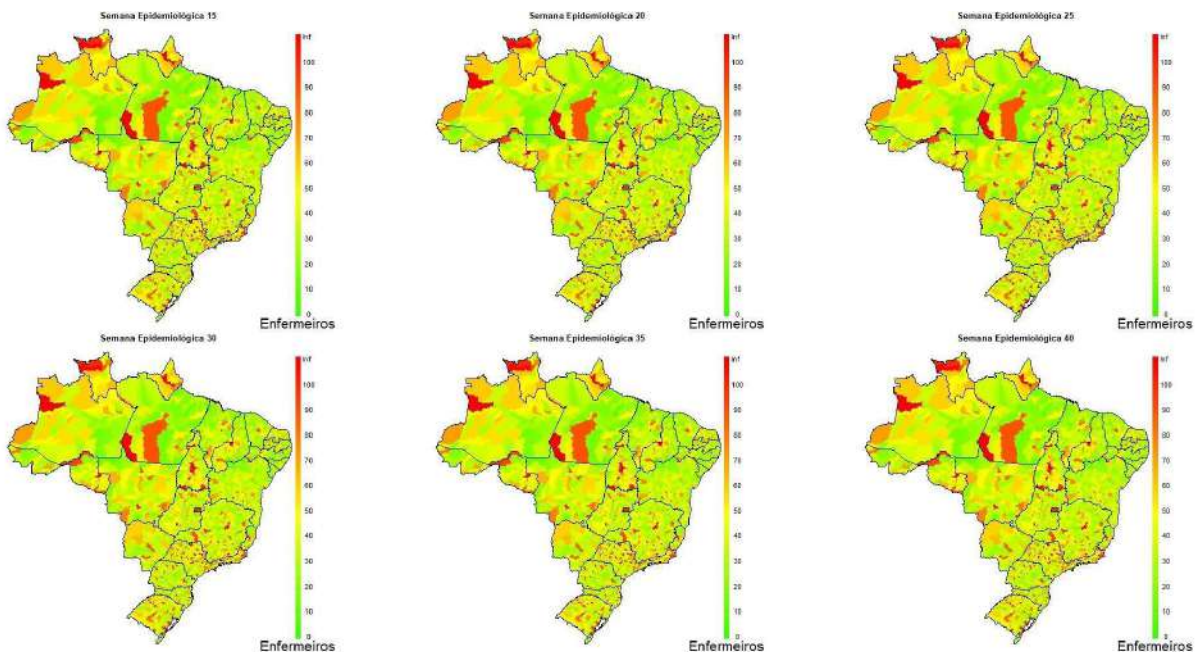
Fonte: elaboração própria.

Figura 14 – Taxas de médicos por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



Fonte: elaboração própria.

Figura 15 – Taxas de enfermeiros e técnicos de enfermagem por 10.000 habitantes na 20ª semana (período de 10 de maio a 16 de maio de 2020), 25ª semana (período de 14 de junho a 20 de junho de 2020), 30ª semana (período de 19 de julho a 25 de julho de 2020), 35ª semana (período de 23 de agosto a 29 de agosto de 2020) e 40ª semana (período de 27 de setembro a 03 de outubro de 2020)



Fonte: elaboração própria.

4.4 Modelos para casos

Na Tabela 6, são apresentados os valores DIC e LPML dos modelos espaço-temporais BNH2 com vizinhança tipo *queen*, com e sem estrutura de saúde, em que se verifica que, adicionando a estrutura de saúde, há melhora na performance do modelo, com redução dos valores de $DIC = 815173,8$ e $LPML = 409444,5$. Um resumo do modelo com estrutura de saúde é descrito na Tabela 7, em que se verifica que todos os parâmetros e hiperparâmetros são significativos.

Tabela 6 – Valores de DIC e LPML – modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 2ª ordem e vizinhança tipo queen, com e sem estrutura de saúde

Modelo	DIC	LPML
Sem Estrutura de Saúde	817846,6	410554,4
Com Estrutura de Saúde	815173,8	409444,5

Fonte: elaboração própria.

Tabela 7 – Resumos *a posteriori* dos efeitos do modelo espaço-temporal para casos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 2ª ordem e vizinhança tipo queen, com estrutura de saúde

Parâmetro	Média	Moda	Desvio-padrão	Intervalo de Credibilidade (95%)	
	<i>a posteriori</i>	<i>a posteriori</i>	<i>a posteriori</i>	L.I.	L.S.
Intercepto	0,2598	0,2598	0,0104	0,2393	0,2802
Posto de saúde	-0,0307	-0,0307	0,0017	-0,0341	-0,0272
Centro de saúde	-0,0678	-0,0678	0,0019	-0,0715	-0,0641
Leitos clínica geral	-0,0515	-0,0516	0,0054	-0,0621	-0,0410
Médico	0,0048	0,0048	0,0003	0,0041	0,0054
Enfermeiros	-0,0003	-0,0003	0,0002	-0,0007	-0,0002
Hiperparâmetros					
θ	0,9258	0,9255	0,0052	0,9158	0,9362
π	0,0012	0,0001	0,0001	0,0001	0,0005
τ_ν	2681,2881	668,6498	4359,2953	318,3864	12640,0525
τ_μ	13,9242	14,0675	0,3579	13,1581	14,5532
τ_γ	20,6948	19,0532	6,8866	8,9188	35,2002
τ_ϕ	8079,1785	2075,9309	9050,4467	776,5827	31566,3337
τ_δ	1,1759	1,1777	0,0081	1,1559	1,1946

Fonte: elaboração própria.

A partir dos resultados da Tabela 7, observa-se que as taxas de postos de saúde, centros de saúde, leitos de clínica geral e enfermeiros apresentam efeitos negativos ao risco. Dessa forma, com o aumento dessas variáveis, tem-se uma redução do risco de aumento de casos. As taxas de médicos apresentaram efeitos positivos, ou seja, o incremento dessa variável também corrobora para o incremento do risco de aumento de casos.

4.4.1 Modelos para óbitos

Os valores DIC e LPML dos modelos espaço-temporais BNH1 com vizinhança tipo *queen*, com e sem estrutura de saúde, são apresentados na Tabela 8. Assim, verifica-se que a incorporação da estrutura de saúde melhora a performance do modelo. Um resumo do modelo com estrutura de saúde é apresentado na Tabela 9, em que se verifica que o parâmetro para taxa de médicos é não significativo, pois seu intervalo de credibilidade contém o zero.

Tabela 8 – Valores de DIC e LPML – modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança tipo *queen*, com e sem estrutura de saúde

Modelo	DIC	LPML
Sem estrutura de saúde	973843,3	487029,8
Com estrutura de saúde	910927,7	455577,6

Fonte: elaboração própria.

Tabela 9 – Resumos a *posteriori* dos efeitos do modelo espaço-temporal para óbitos com distribuição de probabilidade binomial negativa Hurdle com efeito temporal de 1ª ordem e vizinhança tipo *queen*, com estrutura de saúde

Parâmetro	Média	Moda	Desvio Padrão	Intervalo de Credibilidade	
	<i>a posteriori</i>	<i>a posteriori</i>	<i>a posteriori</i>		
Intercepto	2,8231	2,8231	0,0135	2,7966	2,8495
Posto de saúde	-0,1379	-0,1379	0,0023	-0,1424	-0,1333
Centro de saúde	-0,2756	-0,2756	0,0026	-0,2807	-0,2706
Leitos clínica geral	-0,2979	-0,2979	0,0079	-0,3134	-0,2823
Médico	0,0439	0,0437	0,0069	0,0305	0,0577
Enfermeiros	0,0145	0,0145	0,0003	0,0139	0,0151
Hiperparâmetros					
θ	0,3544	0,3544	0,0016	0,3514	0,3575
π	0,0014	0,0008	0,0009	0,0003	0,0038
τ_v	99,0403	91,8495	25,2439	57,2278	155,9272

τ_{μ}	82,9575	71,4282	25,4711	45,0313	143,7906
τ_{γ}	7,5903	7,0265	2,0889	4,1146	12,2378
τ_{ϕ}	8079,1785	2075,9309	9050,4467	776,5827	31566,3337
τ_{δ}	1,1759	1,1777	0,0081	1,1559	1,1946

Fonte: elaboração própria.

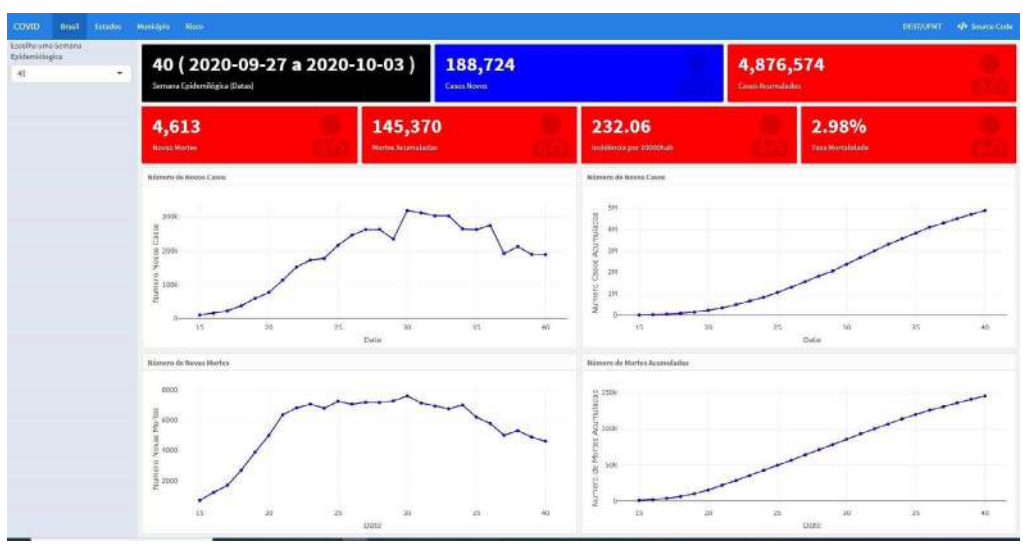
A partir dos resultados da Tabela 9, observa-se que as taxas de postos de saúde, centros de saúde, leitos de clínica geral e enfermeiros apresentam efeitos negativos ao risco. Dessa forma, com o aumento dessas variáveis, tem-se uma redução do risco de aumento de mortes.

4.5 Painel de visualização

Como fruto deste projeto, foi desenvolvido um painel de visualização para os casos e óbitos decorrentes de Covid-19, disponível no seguinte endereço: <<https://lecdufmt.shinyapps.io/Covid/>>. Esse painel é dividido em quatro abas: Brasil, Estados, Municípios e Risco.

Na aba *Brasil*, é apresentada a evolução da doença no Brasil, onde é possível acessar o número de casos e óbitos novos, bem como os seus valores acumulados (Figura 16). No canto superior esquerdo, também é possível aplicar um filtro por semana epidemiológica.

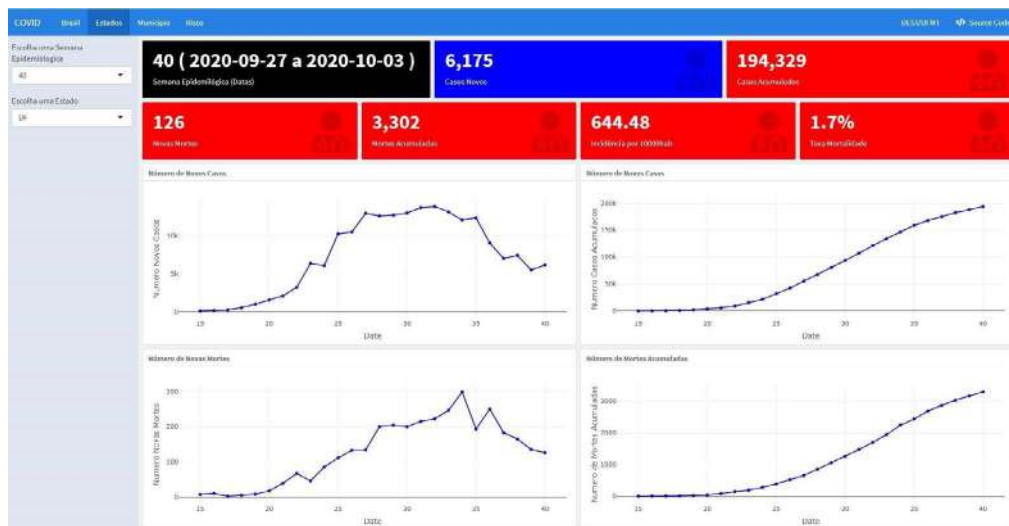
Figura 16 – Informações da aba *Brasil* no painel de visualização



Fonte: elaboração própria.

Na Figura 17, é apresentada a aba *Estados*. No canto superior esquerdo, além do filtro por semana epidemiológica, também é possível filtrar o estado e, de acordo com este, é apresentada a evolução da doença.

Figura 17 – Informações da aba *Estados* no painel de visualização



Fonte: elaboração própria.

A aba *Municípios* apresenta a evolução da doença para cada município brasileiro (Figura 18). Para obter as informações da evolução da doença em um dado município, deve-se escolher o estado na parte superior e em seguida, no canto superior esquerdo, escolher a semana epidemiológica e o município.

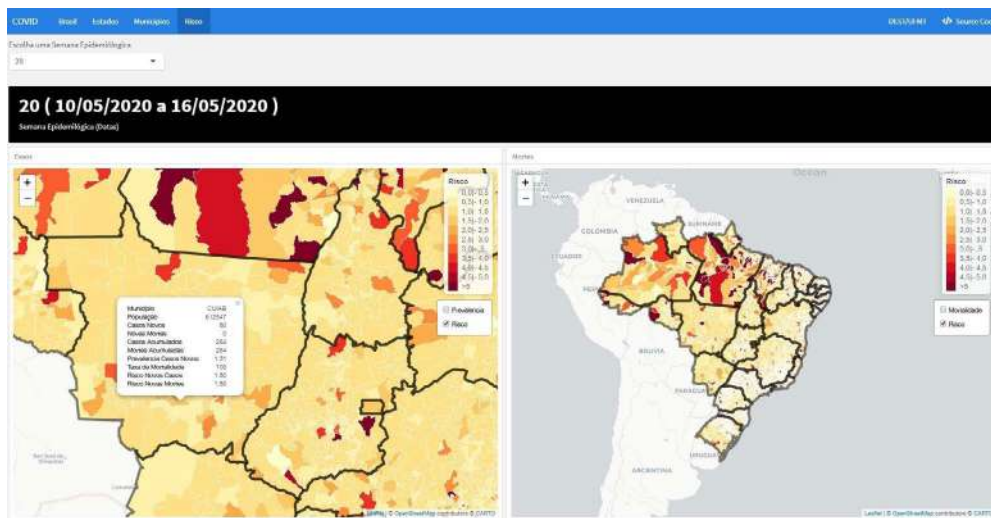
Figura 18 – Informações da aba *Municípios* no painel de visualização



Fonte: elaboração própria.

Na aba *Risco*, é apresentado o risco para cada município brasileiro (Figura 19). Essas informações foram obtidas a partir dos modelos espaço-temporais apresentados neste trabalho. É possível aplicar o filtro por semana epidemiológica e assim são apresentados dois mapas: um mapa que descreve os casos e outro mapa que descreve os óbitos. Dentro do mapa de casos, é possível escolher uma camada com o risco de aumento de casos ou prevalência de casos. No mapa de óbitos, verifica-se uma camada com o risco de aumento de óbitos e as taxas de mortalidade. Também é possível aumentar ou diminuir o *zoom* no mapa e, ao clicar em um município, são acessadas as suas informações básicas: nome, população, número de casos novos, número de novos óbitos, número de casos acumulados, número de óbitos acumulados, prevalência de casos novos, taxa de mortalidade, risco de aumento de casos e risco de aumento de óbitos.

Figura 19 – Informações da aba *Risco* no painel de visualização



Fonte: elaboração própria.

4.6 Discussão

Os casos e óbitos para Covid-19 apresentaram modelos semelhantes, pois ambos tiveram melhor performance com vizinhança do tipo *queen*, refletindo que a dinâmica de espalhamento da doença entre os municípios está relacionada com a estrutura de fronteira. Em relação à dependência temporal, observou-se que os casos apresentam dependência de 2 semanas e os óbitos de 1 semana. Dias *et al.* (2020) relatam que as pessoas infectadas pelo novo coronavírus podem apresentar sintomas entre 4 e 5 dias após a contaminação, mas os sintomas podem variar entre 1 e 14 dias para aparecerem. Em relação ao agravamento e possível morte, em geral ocorrem após 10 dias desde o início dos sintomas.

Em relação à presença de zeros, observou-se que, no comportamento da epidemia tanto para casos como para óbitos, esses valores nulos ocorrem de forma estrutural, com melhor performance dos modelos Hurdle. Entretanto essa suposição sobre os mecanismos de geração de zeros não é simples de ser validada. Ainda assim, os dados com presença de zeros não devem ser ignorados e excluídos da análise, pois podem fornecer informações importantes sobre o processo gerador da doença. O fato de haver uma grande proporção de zeros pode indicar características importantes da doença ou condição em estudo, incluindo sua prevalência na população, detectabilidade (ou taxa de captura) e conscientização da população em relação à doença. Costa *et al.* (2018) estudaram o processo de difusão do vírus influenza A (H1N1) no Paraná, em que as unidades de análise espacial e temporal foram o município de residência dos casos e as semanas epidemiológicas do ano de 2009, respectivamente. Nesse estudo, os autores também desenvolveram um modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e vizinhança tipo *queen*. Gayawan *et al.* (2020) estudaram a dinâmica espaço-temporal da Covid-19 nos primeiros 62 dias de existência no continente africano, em que o modelo Poisson Hurdle apresentou uma boa performance para explicação do espalhamento da doença.

Assim, o modelo espaço-temporal binomial negativo Hurdle fornece uma estrutura flexível para modelar dados epidemiológicos com excesso de zeros. Particularmente em relação à Covid-19, esse modelo apresentou bom desempenho, permitindo levantar suposições sobre o processo gerador espacial e temporal.

A inclusão das variáveis de estrutura de saúde contribuiu para aperfeiçoamento do modelo. Cabe salientar que, no Brasil, a estrutura de saúde é bem diversa e, devido à sua grande dimensão territorial, esta apresenta assimetrias marcantes (SANTOS, 2020). As variáveis de estrutura de saúde utilizadas neste trabalho evidenciaram que essas assimetrias podem contribuir para uma maior difusão da doença. Assim, o enfrentamento da pandemia no país pressupõe mudanças substanciais na forma como os cuidados de saúde são prestados, bem como a reorganização de toda a rede assistencial.

No que se refere à força de trabalho, a Organização Mundial da Saúde (WORLD HEALTH ORGANIZATION, 2017) definiu uma agenda política para garantir uma força de trabalho que seja adequada ao propósito de atingir as metas dos *Objetivos de Desenvolvimento Sustentáveis*. Esse documento de base analisa as implicações quantitativas e os requisitos para a sua implementação, preconizando como ideal

o índice de 4,45 profissionais entre médicos, enfermeiros e parteiras por 1.000 habitantes, que é um índice bem superior ao praticado no Brasil. Assim, a força de trabalho disponível apresenta impacto na forma de enfrentamento da pandemia.

No contexto da pandemia da Covid-19, o excesso de informações disponíveis torna difícil encontrar aquelas que auxiliam na tomada de decisão por gestores e profissionais da saúde. Assim, o painel desenvolvido visa contribuir para a difusão de informação e a visualização de tendências, não somente para gestores, mas também para a população. De acordo com Everts (2020), os painéis de visualização são o artefato cultural mais impressionante da atual pandemia de coronavírus, e estes podem auxiliar não somente na compreensão da doença, mas também no planejamento de preparação para o enfrentamento à pandemia.



5.

Conclusões



5. Conclusões

Os modelos espaço-temporais desenvolvidos neste projeto contribuem para a interpretação do risco de espalhamento da epidemia de Covid-19. No estudo da evolução dos casos de Covid-19, o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 2ª ordem e estrutura de vizinhança do tipo *queen* foi o mais adequado para explicar a evolução de casos nos municípios brasileiros.

Com respeito à evolução dos óbitos de Covid-19, o modelo espaço-temporal com distribuição de probabilidade binomial negativa Hurdle, efeito temporal de 1ª ordem e estrutura de vizinhança do tipo *queen* apresentou o melhor ajuste entre os modelos estudados para explicar a distribuição da contagem acumulada de óbitos.

No estudo da relação entre estrutura de saúde com a evolução de casos e óbitos, observou-se que as taxas de postos de saúde, centros de saúde, leitos de clínica geral e enfermeiros apresentaram efeitos negativos ao risco de aumento de Covid-19, levando à conclusão de que o incremento dessas quantidades contribui para a redução do risco de aumento de casos de Covid-19.

Esses resultados trazem interpretações importantes, pois, com as estimativas das quantidades *a posteriori*, foi possível estabelecer a distribuição do risco de aumento ou diminuição de casos e óbitos nos municípios.

Foram apresentados gráficos para verificar a boa calibração dos modelos, mapas de visualização dos riscos nos municípios brasileiros e gráficos de tendência temporal ao longo das semanas epidemiológicas, nos modelos para casos e óbitos.

Por fim, foi desenvolvido um painel de visualização para os casos e óbitos decorrentes de Covid-19 no seguinte endereço: <<https://lecdufmt.shinyapps.io/Covid/>>. O painel de visualização auxiliará os pesquisadores, gestores e profissionais da saúde no entendimento desta epidemia e no desenvolvimento de ações e tomada de decisões de saúde da população.

REFERÊNCIAS

- ADIN, A.; GOICOA, T.; UGARTE, M. D. Online relative risks/rates estimation in spatial and spatio-temporal disease mapping. *Computer Methods and Programs in Biomedicine*, v. 172, p. 103-116, 2019.
- ANDRADE, L. A. et al. Surveillance of the first cases of Covid-19 in Sergipe using a prospective spatiotemporal analysis: the spatial dispersion and its public health implications. *Revista da Sociedade Brasileira de Medicina Tropical*, v. 53, p. 1–5, 2020.
- BESAG, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 36, n. 2, p. 192-236, 1974.
- BESAG, J.; YORK, J.; MOLLIÉ, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, v. 43, p. 1-20, 1991.
- BIVAND, R. S.; PEBESMA, E.; GOMEZ-RUBIO, V. *Applied spatial data analysis with R*. 2. ed. NY: Springer, 2013. Disponível em: <<http://www.asdar-book.org/>>.
- BLANGIARDO, M.; CAMELETTI, M. *Spatial and spatio-temporal Bayesian models with R – INLA*. [S.l.]: Wiley, 2015.
- BRASIL. Ministério da Saúde. *Cadastro Nacional de Estabelecimentos de Saúde*. Brasília, DF: Ministério da Saúde 2020a.
- Disponível em: <<http://cnes.datasus.gov.br/>>. Acesso em: 10/09/2020.
- BRASIL. Ministério da Saúde. *Painel da Situação Epidemiológica da Covid-19 no Brasil*. Brasília, DF: Ministério da Saúde, 2020b. Disponível em: <<https://covid.saude.gov.br>>.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. *Guia de vigilância epidemiológica: emergência de saúde pública de importância nacional pela doença pelo coronavírus 2019*. Brasília, DF: Ministério da Saúde, 2020c. Disponível em: <<https://coronavirus.saude.gov.br/guia-de-vigilancia-epidemiologica-covid-19>>. Acesso em: 03/10/2020.
- BRESLOW, N. E.; CLAYTON, D. G. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, Taylor & Francis, v. 88, n. 421, p. 9-25, 1993.
- CÂMARA, G. et al. *Análise espacial de dados geográficos*. São José dos Campos: INPE, 2001. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro>>.
- CAMERON, A.; TRIVEDI, P. *Regression analysis of count data*. [S.l.]: Cambridge University Press, 2013. (Econometric Society Monographs).
- CHANG, W. et al. *Shiny: web application framework for R*. [S.l.], 2020. R package version 1.5.0. Disponível em: <<https://CRAN.R-project.org/package=shiny>>.

CHATFIELD, C. *Time-series forecasting*. [S.l.]: CRC Press, 2000.

COSTA, A. C. C. da *et al.* Spatiotemporal diffusion of influenza a (h1n1): starting point and risk factors. *PLOS ONE*, Public Library of Science, v. 13, n. 9, p. 1-20, set. 2018.

CRESSIE, N. *Statistics for spatial data*. [S.l.]: Wiley, 2015. (Wiley Series in Probability and Statistics).

DE OLIVEIRA, G. L.; LOSCHI, R. H.; ASSUNÇÃO, R. M. A random-censoring poisson model for underreported data. *Statistics in Medicine*, Wiley Online Library, v. 36, n. 30, p. 4873-4892, 2017.

DEDEFO, M. *et al.* Spatiotemporal mapping and detection of mortality cluster due to cardiovascular disease with bayesian hierarchical framework using integrated nested laplace approximation: a discussion of suitable statistic applications in kersa, oromia, ethiopia. *Geospatial Health*, v. 13, n. 2, p. 336—334, nov. 2018.

DIAS, V. M. de C. H. *et al.* Orientações sobre diagnóstico, tratamento e isolamento de pacientes com covid-19. *Journal Infection Control*, v. 9, n. 2, p. 56-75, abr. 2020.

ELLIOTT, P. *Spatial epidemiology: methods and applications*. [S.l.]: Oxford University Press, 2000. (Oxford medical publications).

EVERTS, J. The dashboard pandemic. *Dialogues in Human Geography*, v. 10, n. 2, p. 260-264, 2020.

GAYAWAN, E. *et al.* The spatio-temporal epidemic dynamics of covid-19 outbreak in africa. *medRxiv*, Cold Spring Harbor Laboratory Press, v. 148, n. e212., p.1-11, 2020.

GHOSH, S. K.; MUKHOPADHYAY, P.; LU, J.-C. Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference*, v. 136, n. 4, p. 1360-1375, 2006. ISSN 0378–3758.

GOSTIN, L. O.; FRIEDMAN, E. A.; WETTER, S. A. Responding to Covid-19: how to navigate a public health emergency legally and ethically. *Hastings Center Report*, v. 50, n. 2, p. 8-12, 2020.

HINDE, J.; DEMÉTRIO, C. G. Overdispersion: models and estimation. *Computational Statistics & Data Analysis*, v. 27, n. 2, p. 151-170, 1998.

HU, M.-C.; PAVLICOVA, M.; NUNES, E. V. Zero-inflated and Hurdle models of count data with extra zeros: examples from an hiv-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse*, Taylor & Francis, v. 37, n. 5, p. 367-375, 2011.

HUI, D. S. *et al.* The continuing 2019-ncov epidemic threat of novel coronaviruses to global health – the latest 2019 novel coronavirus outbreak in Wuhan, China. *International Journal of Infectious Disease*, v. 91, p. 264-266, 2020.

KNORR-HELD, L. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, v. 19, n. 17-18, p. 2555-2567, 2000.

LIMA, C. R. de A. *et al.* Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. *Cadernos de Saúde Pública*, v. 25, p. 2095-2109, out. 2009.

LINDGREN, F.; RUE, H. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, v. 63, n. 19, p. 1-25, 2015.

LITTLE, Roderick J. A.; RUBIN, Donald B. *Statistical analysis with missing data*. 2. ed. New York: Wiley & Sonc, Inc, 2002.

MARSHALL, E. C.; SPIEGELHALTER, D. J. Approximate cross-validatory predictive checks in disease mapping models. *Statistics in Medicine*, v. 22, n. 10, p. 1649-1660, 2003.

MOHEBBI, M.; WOLFE, R.; FORBES, A. Disease mapping and regression with count data in the presence of overdispersion and spatial autocorrelation: a Bayesian model averaging approach. *International Journal of Environmental Research and Public Health*, v. 11, n. 1, p. 883-902, 2014.

MORETTIN, P.; TOLOI, C. *Análise de séries temporais: modelos lineares univariados*. [S.l.]: Blucher, 2018.

NATÁRIO, I. Métodos computacionais:INLA, integrated nested laplace approximation. *Boletim Sociedade Portuguesa de Estatística*, n. 2, p. 52-57, 2013.

PAULINO, C. D. *et al.* *Estatística Bayesiana*. 2. ed. Lisboa: Fundação Calouste Gulbenkian, 2018.

PELISSARI, M. R. CNES como instrumento de gestão e sua importância no planejamento das ações em saúde. *Revista Saúde Pública do Paraná*, v. 2, n. 1, p. 160-166, 2019.

PETTIT, L. I. The conditional predictive ordinate for the normal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 52, n. 1, p. 175-184, 1990.

PRADO SIQUEIRA, R. *Brazilmaps*: Brazilian maps from different geographic levels. [S.l.], 2020. R package version 0.1.0. Disponível em: <<http://github.com/rpradosiqueira/brazilmaps>>.

R CORE TEAM. *R: A language and environment for statistical computing*. Vienna, Austria, 2020. Disponível em: <<https://www.R-project.org/>>.

RUE, H.; MARTINO, S.; CHOPIN, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 71, n. 2, p. 319-392, 2009.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. *Akaike information criterion statistics*. 1. ed. [S.l.]: Springer Netherlands, 1986. v. 1.

SALDANHA, R. D. F.; BASTOS, R. R.; BARCELLOS, C. *Microdatasus*: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS).

Cadernos de Saúde Pública, v. 35, n. 9, p. 86-90, 2019.

SANTANA, G. B. de A. *et al.* Spatiotemporal evolution of coronavirus disease 2019 mortality in Brazil in 2020. *Revista da Sociedade Brasileira de Medicina Tropical*, v. 53, p1--3, 2020.

SANTOS, J. A. F. Covid-19, causas fundamentais, classe social e território. *Trabalho, Educação e Saúde*, v. 18, p. 1--7, 2020. ISSN 1981-7746.

SILVA, M. V. S. da. Sistemas de informação em saúde em tempos de Covid-19. *Cadernos ESP*, v. 14, n. 1, p. 86-90, 2020.

SPIEGELHALTER, D. J. *et al.* Bayesian measures of model complexity and fit. *Journal of the Royal*

Statistical Society: Series B (Statistical Methodology), v. 64, n. 4, p. 583-639, 2002.

STONER, O.; ECONOMOU, T.; SILVA, G. D. M. da. A hierarchical framework for correcting under-reporting in count data. *Journal of the American Statistical Association*, Taylor & Francis, v. 114, n. 528, p. 1481–1492, 2019.

WHITWORTH, J. Covid-19: a fast evolving pandemic. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, v. 114, n. 4, p. 241-248, mar. 2020.

WORLD HEALTH ORGANIZATION. *World health statistics 2017: monitoring health for the SDGs, sustainable development goals*. [S.l.]: World Health Organization, 2017. 116 p.

WORLD HEALTH ORGANIZATION. *WHO announces Covid-19 outbreak a pandemic*. 2020. Disponível em: <http:

[//www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic](http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic)>. Acesso em: 12/08/2020.

ANEXO I: TUTORIAL PREDIÇÃO

A inferência bayesiana trata as observações ausentes como qualquer outro parâmetro no modelo (LITTLE; RUBIN, 2002). Quando as observações ausentes estão na variável de resposta, o INLA calcula automaticamente a distribuição preditiva do preditor linear correspondente e os valores ajustados.

O INLA pode ajustar modelos com observações ausentes na resposta computando sua distribuição preditiva. Para isso, deve ser atribuído **NA** à variável resposta. Assim, são obtidas as previsões da distribuição preditiva do modelo.

Assim, para fazer a predições com INLA no R, podem-se utilizar os códigos abaixo:

1. Pacotes necessários para estimar o modelo e verificar as predições:

```
require(rgdal) #pacote para ler os mapas
```

```
require(spdep)
```

```
require(dplyr)
```

```
require(INLA)
```

```
require(brazilmaps)
```

2. Obter a estrutura espacial:

```
BR=get_brmap(geo = "City",class = "SpatialPolygonsDataFrame")
```

```
BR1=get_brmap(geo = "State",class = "SpatialPolygonsDataFrame")
```

```
CRS.new <-CRS("+proj=longlat+ellps=GRS80+no_defs  
+towgs84=0,0,0,0,0,0,0")
```

```
BR <-spTransform(BR, CRS.new)
```

```
BR1 <-spTransform(BR1, CRS.new)
```

```
viz=poly2nb(BR)
```

```
nb2INLA("adj.mat", viz)
```

```
g = inla.read.graph("adj.mat")
```

```
summary(g)
```

3. Obter o banco de dados:

```

dados=read.csv("final.csv",sep = ";",dec=",")
dados$E=0
s1=unique(dados$semanaEpi)
n=length(s1)
for(i in 2:n){
  w=which(dados$semanaEpi==s1[i])
  w1=which(dados$semanaEpi==s1[i-1])
  aux1=dados[w1,]
  dados$E[w]=aux1$casos
}
dados$E=dados$E+1

```

4. Estabelecer as predições. Neste exemplo serão utilizadas 5 predições para cada município:

```

NPred=5 #numero de previsões a serem obtidas
dados1=dados%>%
  filter(semanaEpi>14)%>%
  arrange(semanaEpi)
nm=length(unique(dados1$codmun))
dadospred=NULL
semanaEpi=max(unique(dados1$semanaEpi))
for(i in 1:NPred){
  aux1=dados1[1:nm,]
  semanaEpi=semanaEpi+1
  aux1$semanaEpi=semanaEpi
  dadospred=rbind(dadospred,aux1)
}

```

```
dadospred$casos=NA
```

```
dadospred$E=NA
```

5. Organizar o banco de dados para estimar o modelo:

```
dados1=rbind(dados1,dadospred)
```

```
dados1=dados1%>%
```

```
  arrange(codmun,semanaEpi)
```

```
  s=unique(dados1$semanaEpi)
```

```
  n=length(s)
```

```
  ID=seq(1,nm)
```

```
  ID=sort(rep(ID,n))
```

```
  dados1$Sem=rep(1:n,nm)
```

```
  dados1$ID=ID
```

```
  dados2=data.frame(ID=as.numeric(ID),Sem=as.numeric(dados1$Sem),casos=as.numeric(dados1$casos),E=as.numeric(dados1$E),pop=as.numeric(dados1$populacao))
```

```
  dados2$Sem1=dados2$Sem
```

```
  dados2$ID1=dados2$ID
```

```
  dados2$ID.Sem=seq(1,nrow(dados2))
```

6. Estimar o modelo:

```
modelCASO <- inla(casos ~ 1+
```

```
  f(ID, model = "bym", graph = g)+
```

```
  f(Sem, model = "iid") +
```

```
  f(Sem1, model = "rw2") +
```

```
  f(ID.Sem, model = "iid"),
```

```
  data = dados2, E=E,family = "zeroinflatednbinomial1",
```

```
  control.predictor = list(compute = TRUE, link = 1),
```

```
  control.compute = list(dic=TRUE,cpo=TRUE,config = TRUE))
```


7. Obter as previsões:

```
Pred=exp(modelCASO$summary.linear.predictor$mode)
```

8. Avaliar o erro percentual médio (MAPE) nos dados que não foram oriundos de previsão:

```
dados3=cbind(dados2,Pred)
```

```
dados3=dados3%>%
```

```
  filter(Sem<=26)
```

```
CasosP=dados3$Pred*dados3$E
```

```
Casos=dados3$casos+1
```

```
MAPE=sum((CasosP-Casos)/(Casos))
```

9. Quando obter os dados reais do intervalo de previsão, obter o MAPE:

```
dados5=read.csv("final1.csv",sep = ";",dec=",")
```

```
dados5$E=0
```

```
s1=unique(dados5$semanaEpi)
```

```
s1=c(max(dados$semanaEpi),s1)
```

```
n=length(s1)
```

```
for(i in 2:n){
```

```
  w=which(dados5$semanaEpi==s1[i])
```

```
  if(i==2){
```

```
    w1=which(dados$semanaEpi==s1[i-1])
```

```
    aux1=dados[w1,]
```

```
  }
```

```
  if(i>2){
```

```
    w1=which(dados5$semanaEpi==s1[i-1])
```

```
    aux1=dados5[w1,]
```

```
  }
```

```

dados5 = dados5 + w * casos
}

dados5 = dados5 + 1

dados5 = dados5 %>%
  arrange(codmun, semanaEpi)

dados6 = dados3 %>%
  filter(Sem > 26)

CasosP = dados6$Pred * dados5$E

Casos = dados5$casos + 1

MAPE = sum((CasosP - Casos) / (Casos))

```

10. Comparar o MAPE da etapa 8 e 9, se a diferença entre os dois for superior a 30%, indicativo de modelo descalibrado. Deve-se estimar o modelo novamente.



Cadernos

Caderno nº 88

Coleção: Covid-19 Fast Track

