NOTA TÉCNICA 14: EVOLUÇÃO DA EPIDEMIA DO COVID-19 EM JUIZ DE FORA

Fernando A.B. Colugnati[1,2], Mário Círio Nogueira[1,3], Marcel de Toledo Vieira[4,5], Maria Teresa Bustamante-Teixeira[1,2,3], Isabel Cristina Gonçalves Leite[1,2,3,6], Alfredo Chaoubah[2,4,6]

1. Faculdade de Medicina – UFJF
2. PPg Saúde – UFJF
3. Mestrado Profissional em Saúde da Família - UFJF
4. Depto. de Estatística/ ICE – UFJF
5. PPg Economia – UFJF
6. PPg Saúde Coletiva – UFJF

DESTAQUES

- ✓ Esta é a 14ª nota técnica deste grupo responsável pelas análises de dados e modelagem da epidemia de COVID-19 em Juiz de Fora.
- ✓ A nota atual analisa os casos, internações e óbitos por COVID-19 em Juiz de Fora até o dia 16 de janeiro de 2021.
- ✓ Análise dos números de casos e vidas perdidas ao longo das três últimas semanas epidemiológicas confirma que Juiz de Fora e a Macro Sudeste ainda não atendem a dois importantes critérios da OMS, para que possamos considerar a pandemia sob controle no município: redução contínua de pelo menos 50% do número semanal de novos casos por um período de três semanas desde o último pico. Além disso, seria também necessário que o número de vidas perdidas apresentasse uma redução por pelo menos três semanas sucessivas.
- ✓ A distribuição dos casos confirmados, por faixas etárias, indica que 57,5% (em Juiz de Fora) e 68,64% (na Macro Sudeste) são em pessoas com idades entre 20 e 59 anos
- ✓ A análise da distribuição dos óbitos por faixas etárias confirma que a maioria dos pacientes que vieram a falecer, 80,4% (em Juiz de Fora) e 81% (na Macro Sudeste) tinha 60 ou mais anos de idade.
- ✓ A média móvel do número de casos novos dos últimos 7 dias era igual a 110,6 no dia 16 de janeiro, representando aumento de 70,2% em relação a 2 de janeiro (14 dias antes). A média móvel para o número de vidas perdidas era

aproximadamente igual a 5,6 no dia 16 de janeiro, apresentando uma redução de 5,4% em relação ao dia 2 de janeiro.

- ✓ Detectado aumento da taxa de letalidade ao longo das últimas 4 semanas epidemiológicas no município, sendo de 3,94% no dia 16 de janeiro.
- ✓ Outro parâmetro associado a controle de pandemia é o valor do Número de Reprodução Efetivo (Rt) abaixo de 1 por pelo menos 2 semanas. O Rt estimado para Juiz de Fora entre os dias 19 de dezembro e 16 de janeiro apresentou um valor máximo de 1,39 nos dias 8 e 9 de janeiro e um valor mínimo no de 0,49 no dia 1º de janeiro.
- ✓ Dados do Google Mobility indicam redução de ida ao trabalho e a parques no período do feriado de 1º. de janeiro, mas por outro lado às idas às estações de transporte (pontos de ônibus, estações rodoviárias etc.) foram 60% maior em relação ao período de referência no dia 18 de dezembro.
- ✓ A modelagem por séries temporais estima a Função de Resposta ao Impulso apontando que, no período até 13 de janeiro, a cada 84 novos casos suspeitos, esperamos depois de 8 dias um aumento de novos 40 casos suspeitos, cerca de 4 leitos de enfermaria ocupados e 1 leito de UTI.
- ✓ Com base neste modelo, é apresentada uma projeção de leitos ocupados até o dia 27 de janeiro, que pode chegar a 473 pacientes internados entre enfermaria e UTI.

1. Objetivos

Esta é a 14ª nota técnica do grupo responsável pelas análises de dados e modelagem da epidemia da COVID-19 em Juiz de Fora, formado pelos professores supracitados como autores deste documento. Este grupo é parte de uma iniciativa de parceria entre a UFJF e a Prefeitura de Juiz de Fora (PJF), que contempla diversas ações nas mais variadas áreas de conhecimento. A primeira nota técnica foi publicada em 14 de abril com a análise dos dados de notificações até o dia 13 de abril de 2020. A nota atual analisa os dados até o dia 16 de janeiro de 2021.

Este grupo tem como objetivo sistematizar e analisar dados de diversas fontes oficiais sobre a pandemia de COVID-19 no município de Juiz de Fora, fazendo comparações com dados semelhantes do estado de Minas Gerais e do Brasil, quando pertinente. Por meio deste convênio, o acesso aos dados fornecidos diretamente pela Vigilância Epidemiológica e pela PJF tem sido fundamental para um entendimento da situação, sua modelagem e a construção de diferentes cenários possíveis para esta epidemia na cidade. O objetivo maior é auxiliar nos planos de contingenciamento dos leitos, profissionais e equipamentos de saúde no decorrer do crescimento da infecção na cidade.

Cabe ressaltar que os dados analisados por este grupo são referentes e limitados aos seguintes grupos elegíveis para testagem de acordo com a Secretaria de Estado de Saúde de Minas Gerais: a) amostras provenientes de unidades sentinelas de Síndrome Gripal (SG) e Síndrome Respiratória Aguda Grave (SRAG); b) TODOS os casos de SRAG hospitalizados; c) TODOS os óbitos suspeitos; d) profissionais de saúde sintomáticos (neste caso, se disponível, priorizar Teste Rápido e profissionais da assistência direta); e) profissionais de segurança pública sintomáticos (neste caso, se disponível, priorizar Teste Rápido); f) por amostragem representativa (mínimo de 10% dos casos ou 3 coletas), nos surtos de SG em locais fechados (ex: asilos, hospitais, etc); g) público privado de liberdade e adolescentes em cumprimento de medida restritiva ou privativa de liberdade, ambos sintomáticos; h) população indígena aldeada (https://www.saude.mg.gov.br/images/noticias_e_eventos/000_2020/coronavirus-legislacoes/19-05-Atualizacao-Protocolo.pdf). Estão considerados nas análises também os casos testados pela rede privada, desde que tenham sido devidamente notificados.

2. Evolução dos casos e óbitos por COVID-19 confirmados em Juiz de Fora e na Macrorregião de Saúde Sudeste

A evolução dos casos confirmados e dos óbitos causados pelo novo coronavírus (COVID-19) pode ser acompanhada em gráficos disponibilizados na plataforma virtual "JF Salvando Todos" (http://jfsalvandotodos.ufjf.br/). Este site vem sendo desenvolvido por alunos do Curso de Estatística sob a coordenação do terceiro autor desta nota técnica, com a colaboração de outros professores, incluindo os demais autores da presente nota. O principal objetivo da plataforma é permitir o acesso às informações de forma clara e rápida para tomadores de decisão de políticas públicas na área da Saúde e para a população em geral. A plataforma continua sendo aprimorada e permite a visualização de forma amigável de dados de todos os municípios do Brasil e agregados por microrregiões e macrorregiões de saúde, regiões definidas de acordo com os critérios do IBGE, Unidades da Federação e Grandes Regiões.

Dados apresentados por região de saúde permitem a sua utilização para a integração das ações e serviços públicos de saúde no âmbito da pandemia da COVID-19 de forma regionalizada e hierarquizada. Outras novidades são a disponibilização de estimativas de taxas de mobilidade social, do Número de Reprodução Efetivo (Rt), de médias móveis dos números de casos e vidas perdidas, da taxa de letalidade, dos perfis das internações, número de leitos hospitalares disponíveis, e novas abas com links para as notas técnicas do Grupo de Modelagem Epidemiológica da UFJF, para artigos da série 'Descreve, Demonstra e Instrui', e para boletins informativos que são publicados quinzenalmente.

Na análise exploratória apresentada na presente seção voltaremos a levar em consideração alguns dos indicadores para análise do possível controle da pandemia propostos pela Organização Mundial de Saúde. Tais indicadores estão no documento Public health criteria to adjust public health and social measures in the context of COVID-19, publicado em 12 de maio de 2020 e que pode ser acessado em https://www.who.int/publications/i/item/public-health-criteria-to-adjust-public-health-and-social-measures-in-the-context-of-covid-19. São oito indicadores, que foram apresentados nas edições 5 (de 8 de julho de 2020) e 6 (de 22 de julho) do Boletim Informativo COVID-19 da Plataforma "JF Salvando Todos" (http://jfsalvandotodos.ufjf.br/Boletim_Informativo_Edi%C3%A7%C3%A3o_5.pdf e

http://jfsalvandotodos.ufjf.br/Boletim_Informativo_Edi%C3%A7%C3%A3o_6.pdf, respectivamente), sendo eles:

- ✓ Rt abaixo de 1 por 2 semanas;
- ✓ redução contínua de ao menos 50% dos casos suspeitos ou confirmado, por 3 semanas;
- ✓ testes positivos para COVID-19 abaixo de 5% por 2 semanas;
- ✓ proporção de testes positivos para quadros gripais abaixo de 5% por 2 semanas;
- ✓ conhecimento da cadeia de transmissão de pelo menos 80% dos novos casos;
- ✓ queda das mortes por COVID-19 por 3 semanas;
- ✓ queda da morte por pneumonia em faixas etárias onde estava aumentada;
- ✓ redução do número de hospitalizações e internações em UTI por 2 semanas.

No dia 19 de dezembro de 2020 Juiz de Fora tinha 12314 casos confirmados e registrava 437 vidas perdidas de acordo com a Prefeitura de Juiz de Fora. Estes números evoluíram para 15338 casos confirmados e 605 vidas perdidas no dia 16 de janeiro de 2021, representando aumentos de 24,6% e 38,4%, respectivamente, no período de 4 semanas (da 52ª semana epidemiológica de 2020 à 2ª semana epidemiológica de 2021). Na Macrorregião de Saúde Sudeste de MG (Macro Sudeste), neste mesmo período, os aumentos foram de 40,2% e 33,1%, respectivamente, e haviam sido registrados 53658 casos e 1322 vidas perdidas até o dia 16 de janeiro.

A análise dos números de casos e vidas perdidas ao longo das três últimas semanas epidemiológicas confirma que Juiz de Fora e a Macro Sudeste ainda não atendem a dois importantes critérios da OMS, para que possamos considerar a pandemia sob controle no município. De acordo com a OMS, seria necessário que tivesse ocorrido uma redução de pelo menos 50% do número semanal de novos casos por um período de três semanas desde o último pico e que essa redução tivesse ocorrido de forma contínua nestas três semanas. Além disso, seria também necessário que o número de vidas perdidas apresentasse uma redução por pelo menos três semanas sucessivas.

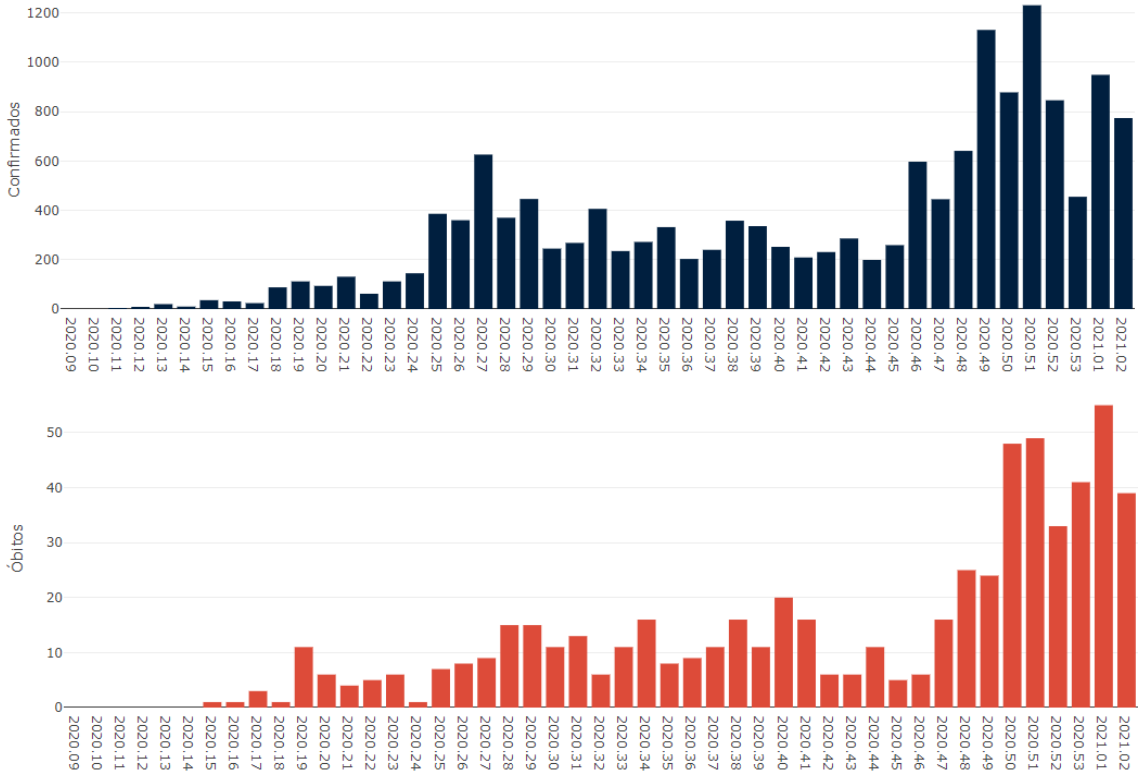

Gráfico 1 - Casos confirmados e vidas perdidas para a Covid-19 por semana epidemiológica em Juiz de Fora. Fonte: jfsalvandotodos.ufjf.br a partir de dados da Prefeitura de Juiz de Fora (PJF).

Os números de casos e óbitos por milhão de habitantes eram de 26.755 e 1055 no dia 16 de janeiro, em Juiz de Fora, e eram de 31.836 e 784 na Macro Sudeste. Estes números eram 39.928 e 988 para o Brasil, e 29.907 e 627 para Minas Gerais.

A distribuição dos casos confirmados, por faixas etárias, indica que 57,5% (em Juiz de Fora) e 68,64% (na Macro Sudeste) são em pessoas com idades entre 20 e 59 anos, ou seja, pertencem majoritariamente à parte da população que é economicamente ativa. Em Juiz de Fora, homens representam 51,2% dos casos e mulheres 48,8%, enquanto que na Macro Sudeste estas proporções são 46,9% e 53,1%, respectivamente.

A análise da distribuição dos óbitos por faixas etárias confirma que a maioria dos pacientes que vieram a falecer, 80,4% (em Juiz de Fora) e 81% (na Macro Sudeste) tinha 60 ou mais anos de idade. Dentre os que perderam suas vidas para a Covid-19, 54% eram homens e 46% eram mulheres em Juiz de Fora, e na Macro Sudeste estas proporções eram 53,8% e 46,2%, respectivamente.

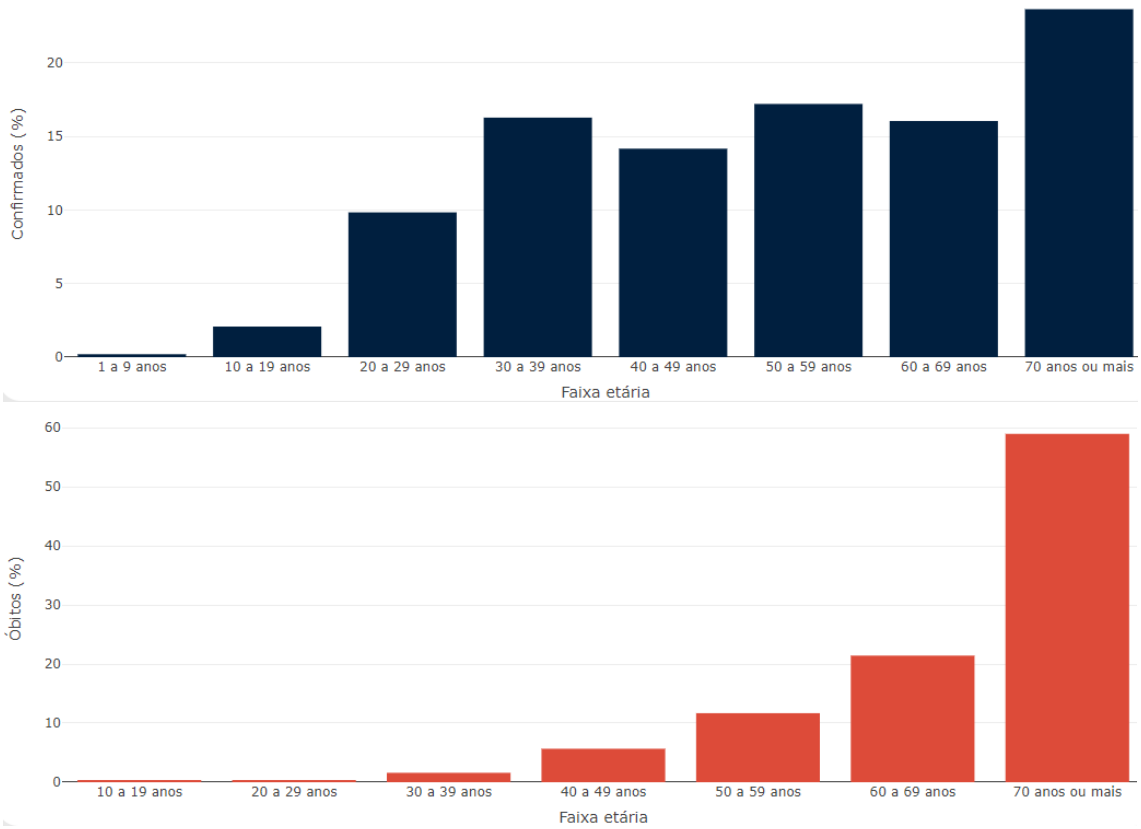

Gráfico 3 – Distribuição dos casos confirmados (acima) e vidas perdidas (abaixo) para a Covid-19 em Juiz de Fora. Fonte: jfsalvandotodos.ufjf.br a partir de dados da PJF.

Em Juiz de Fora, a média móvel do número de casos novos para os últimos 7 dias era igual a 110,6 no dia 16 de janeiro e apresentou um aumento de 70,2% em relação a 2 de janeiro (14 dias antes). A média móvel para o número de vidas perdidas era aproximadamente igual a 5,6 no dia 16 de janeiro, apresentando uma redução de 5,4% em relação ao dia 2 de janeiro. Na Macro Sudeste, considerando o mesmo período, houve aumentos de 63,1% e 28,8%, respectivamente.

O Número de Reprodução Efetivo (Rt) estimado para Juiz de Fora entre os dias 19 de dezembro e 16 de janeiro apresentou um valor máximo de 1,39 nos dias 8 e 9 de janeiro (na Macro Sudeste o máximo foi de 1,35 no dia 10 de janeiro) e um valor mínimo no de 0,49 no dia 1º de janeiro (na Macro Sudeste o mínimo foi de 0,76 no dia 29 de dezembro), e era de 0,91 no dia 16 de janeiro (na Macro Sudeste o Rt era igual a 1,12 nesta data).

De acordo com a Organização Mundial de Saúde, uma das condições para que a pandemia esteja sob controle é de que os valores do Rt sejam menores que 1

persistentemente por pelo menos duas semanas, o que não foi observado em Juiz de Fora e na Macro Sudeste.

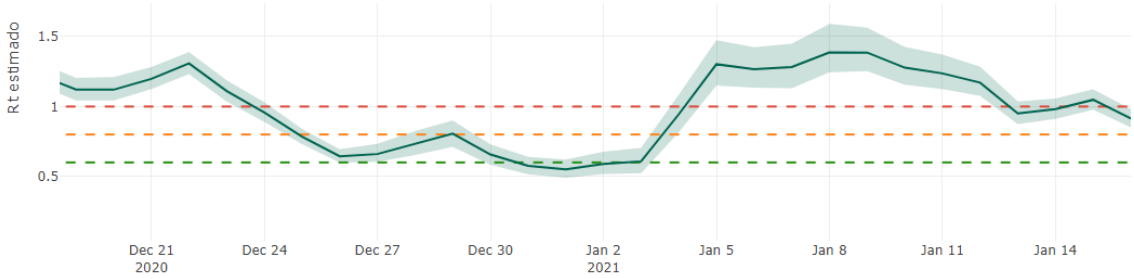

Gráfico 4 – Rt para Juiz de Fora. Fonte: jfsalvandotodos.ufjf.br a partir de dados da PJF.

No dia 16 de janeiro, a taxa de letalidade da Covid-19 era de 3,94% em Juiz de Fora (essa taxa era de 3,55% no dia 19 de dezembro). Percebeu-se um aumento da taxa de letalidade ao longo das últimas 4 semanas epidemiológicas. Para melhor contextualizar, cabe ressaltar que a referida taxa era de 2,46% na Macro Sudeste, 2,1% em Minas Gerais e 2,48% no Brasil no dia 16 de janeiro (esses valores eram de 2,59%, 2,24% e 2,58%, respectivamente, no dia 19 de dezembro).

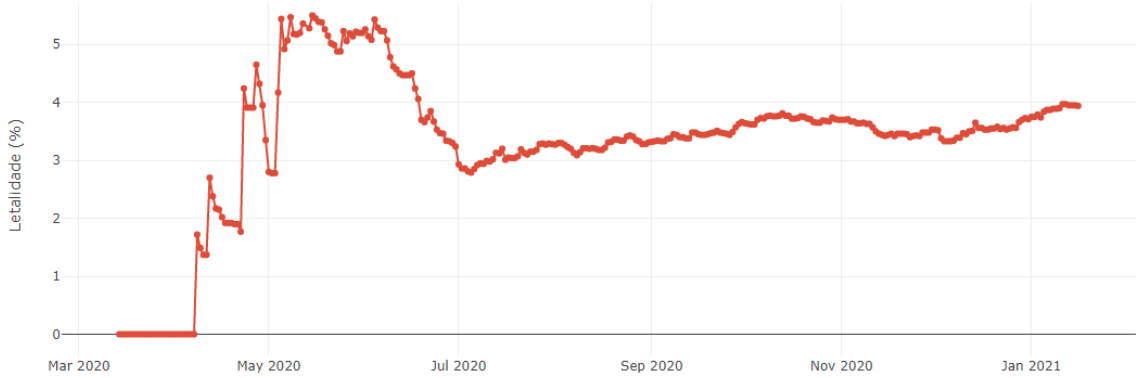

Gráfico 5 – Taxa de Letalidade da Covid-19 para Juiz de Fora. Fonte: jfsalvandotodos.ufjf.br a partir de dados da PJF.

Todas as análises apresentadas acima foram feitas a partir de números oficiais divulgados pela Prefeitura de Juiz de Fora. Sendo assim, não se considerou as possíveis subnotificações de novos casos e de óbitos.

A seguir, reproduzimos a análise dos dados do *Google Mobility* para Juiz de Fora, apresentados na Plataforma JF e que foi apresentada em seu 18º Boletim Informativo (http://jfsalvandotodos.ufjf.br/Boletim_Informativo_Edi%C3%A7%C3%A3o_18.pdf):

> "A partir de dados do *Google Mobility*, apresentados graficamente na Plataforma JF, pode-se estimar que no dia 1º de janeiro (uma 6ª feira, feriado) havia um percentual 23% maior de pessoas em casa em relação ao período de referência anterior ao início da pandemia - esse percentual era de 10% maior no dia 18 de dezembro (uma 6ª feira), indicando **um aumento na adesão ao isolamento social**. A ida aos locais de trabalho apresentava um percentual 77% **menor** que no período de referência – esse percentual indicava uma redução de 5% no dia 18 de dezembro, indicando **um claro efeito do feriado**. Houve uma **redução** na frequência de idas às farmácias e aos pontos de venda de gêneros alimentícios – redução de 77% em relação ao período de referência – havia um aumento de 23% no dia 18 de dezembro em relação ao período de referência. **Reduziu a frequência de idas aos parques** com um percentual 77% menor que no período de referência - essa redução era de 26% no dia 18 de dezembro. Idas aos estabelecimentos de varejo e recreação continuam abaixo do período de referência, com uma redução de 85% - essa redução era de 28% no dia 18 de dezembro sugerindo **menos pessoas foram nestes locais no feriado do dia 1º de janeiro**. Os dados referentes às idas às estações de transporte (pontos de ônibus, estações rodoviárias etc.) apresentaram variação negativa com redução de 13%, - esse percentual era 60% maior em relação ao período de referência no dia 18 de dezembro. Será importante observar a evolução destas taxas em nosso próximo boletim a partir de números não influenciados pelos feriados de final de ano."

3. Modelagem de séries temporais multivariadas para os indicadores de COVID em Juiz de Fora: casos suspeitos, ocupação de leitos de enfermaria e UTI

O Gráfico 6 apresenta a série de indicadores de casos suspeitos e ocupação de leitos no município de Juiz de Fora, referente ao período 30/03/2020 e 13/01/2021

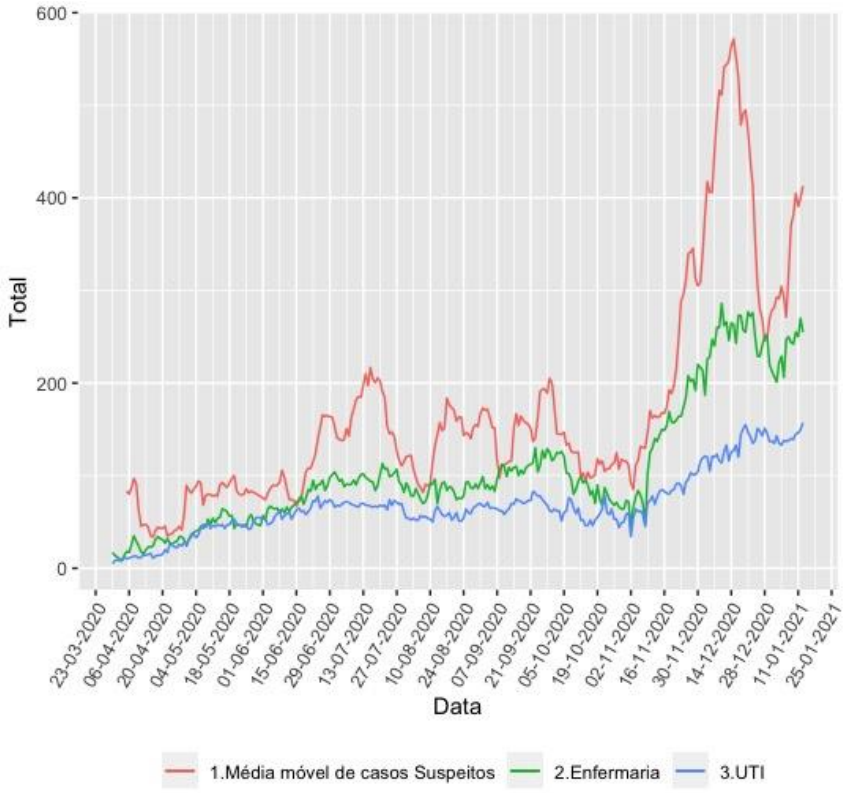

Gráfico 6- Séries dos indicadores de casos suspeitos (média móvel de 7 dias), número de leitos de enfermaria ocupados (Enfermaria) e de UTI (UTI).

Métodos utilizados na análise

Modelos do tipo *Vector Auto-regressive* (VAR) possibilitam a análise de séries temporais multivariadas, permitindo o ajuste de equações simultâneas para estimação dos efeitos de uma série específica nas demais, incluindo nisso a defasagem de tempo, chamada neste relatório de *lag*. No presente contexto, uma série multivariada é um conjunto de indicadores com observações indexadas por uma data, por exemplo, ou qualquer outro índice observado ao longo do tempo. Em resumo, a metodologia aqui considerada é

semelhante à estimação de modelos de regressão simultâneos, um para cada indicador que compõe a série multivariada.

Além disso, por meio das chamadas Funções de Resposta ao Impulso (FRI), podemos entender como e quando um "choque" em determinado indicador impacta nos demais indicadores, o que permite projetar, por exemplo, futuros cenários da ocupação de leitos a partir da observação da série do número de casos suspeitos. O anexo desta nota traz maiores detalhes técnicos sobre o modelo e metodologia considerado. Neste ponto, vale destacar apenas alguns pontos.

Função de Resposta ao Impulso

Sem muitos pormenores, é importante esclarecer que a FRI estima o impacto, em função do tempo *t*, de um "choque" de uma variável explicativa em uma variável resposta. Por choque, neste caso, entende-se um aumento de 1 desvio-padrão da variável explicativa.

A figura 1 abaixo apresenta a FRI para o indicador de casos Suspeitos observados até 13 de janeiro de 2021, impactando em sua própria série e no número de leitos ocupados, de Enfermaria e UTI. Neste caso, o desvio padrão para a série de suspeitos é de 84 casos. Portanto, a cada 84 novos casos suspeitos, esperamos depois de 8 dias um aumento de novos 40 casos suspeitos, cerca de 4 leitos de enfermaria ocupados e 1 leito de UTI.

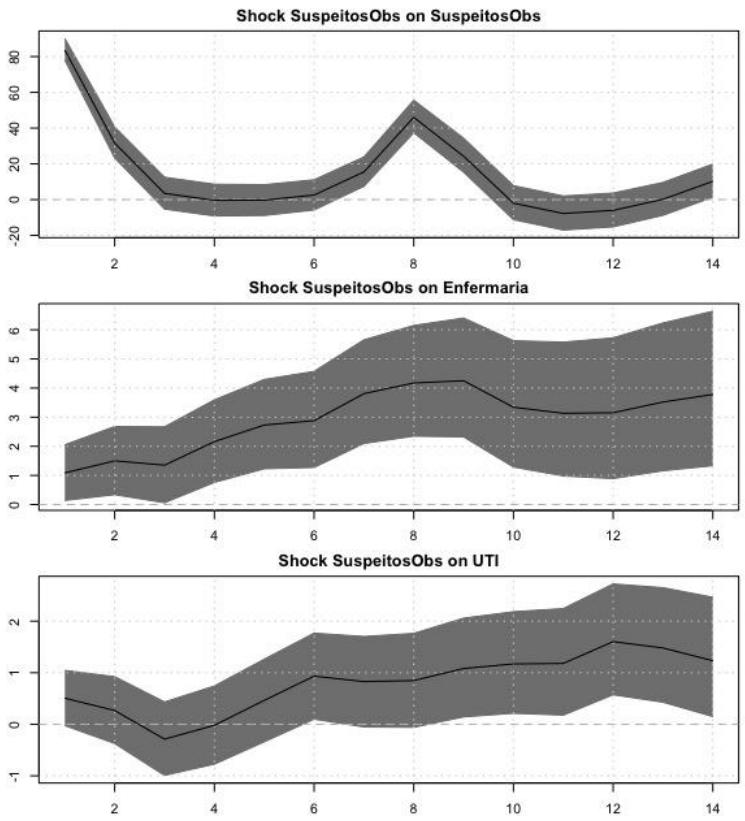

Figura 1- FRI estimada para os dados com série final em 13 de janeiro

As previsões de ocupação de leitos (Enfermaria e UTI) para os próximos 14 dias estão apresentados na figura 2 e quadro 1 a seguir, com respectivos intervalos de confiança de 80%.

Maiores esclarecimentos quanto a essas análises podem ser vistos no anexo técnico na sequência dessa nota.

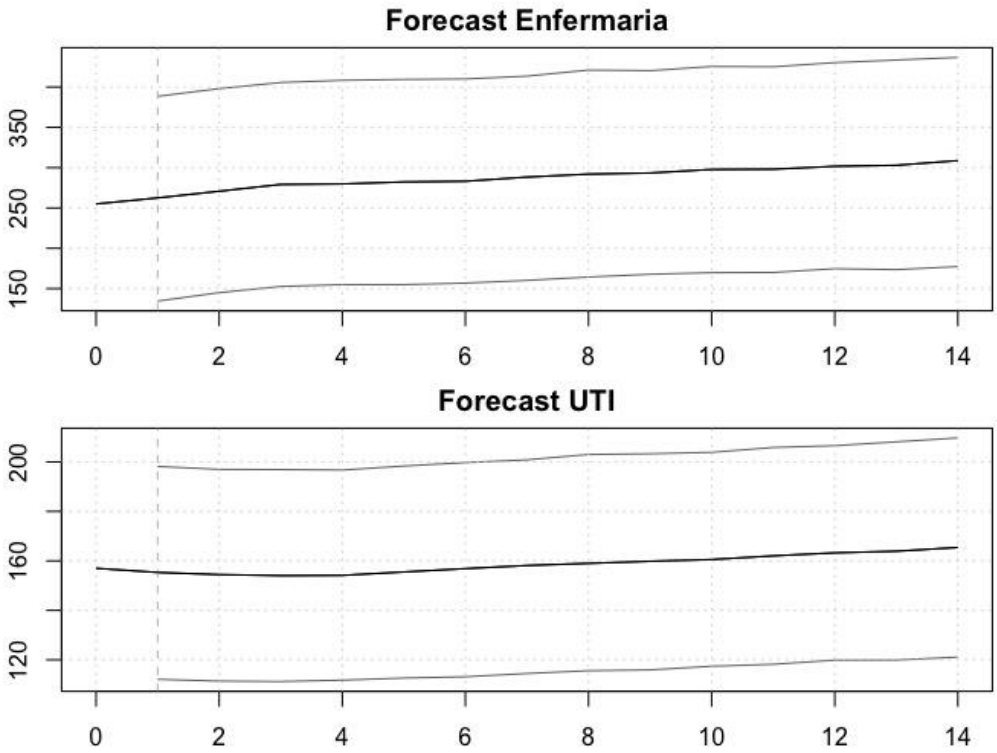

Figura 2- Previsões das séries em horizonte de 14 dias, equivalente a 27 de janeiro.

Quadro 1- Projeções para ocupação de leitos com respectivos intervalos preditivos de 80% de probabilidade

	Enfermaria	UTI	Total (apenas est. pontual)
20/jan/2020	288 [160; 413]	158 [114; 200]	446
27/ jan/ 2020	308 [178; 436]	165 [121; 209]	473

Anexo Técnico

Grupo de Modelagem Epidemiológica da Covid-19 UFJF

09/jan/2020

O Modelo

Foi utilizada a técnica de *Vector Auto-regressive models* (VAR), que possibilita a modelagem de séries temporais multivariadas, sendo possível que se estime equações simultâneas que estimam os efeitos de uma série específica nas demais, incluindo nisso a defazagem de tempo, chamda neste relatório de *lag*. Em resumo, é como se fossem estimados modelos de regressão simultâneos, um para cada indicador que compõe a série multivariada. Entende-se aqui uma série multivariada como um conjunto de indicadores com obserbações indexadas por uma data, por exemplo, ou qualquer outro índice temporal.

Além disso, por meio das chamadas Funções de Resposta ao Impulso (FRI), podemos entender como e quando um "choque" em determinado indicador impacta nos demais indicadores, o que permitirá projetar futuros cenários da ocupação de leitos a partir da observação da série de suspeitos. Segue uma breve introdução sobre o tema.

Para a redação deste anexo, as principais referências são Kuschnig et al. (2020) . Maiores detalhes sobre este tipo de modelagem podem ser encontrados em Stock et al.(1990).

Modelos VAR

Um VAR univariado, também conhecido como Modelo Auto-regressivo (AR) assume a forma $y_t = a_1 y_{t-1} + e_t$. Ou seja, o valor atual observado y depende de seu valor imediatamente anterior (dia anterior, por exempo), onde a_1 denota o coeficiente e o subscrito o *lag* considerado, neste caso 1. Este modelo é chamado de auto-regressivo de ordem 1, AR(1), mas é possível se utilizar ordens maiores de defazagem p, resultando em um modelo AR(p). Os resíduos e_t são assumidos Normais, com média 0 e desvio padrão σ^2.

Um modelo VAR, multivariado, é uma simples extensão do AR(p), e pode ser representado como abaixo, para um caso com 3 séries temporais, ou 3 indicadores.

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix}$$

.

Ou de modo mais compacto

$$y_t = A_1 y_{t-1} + \epsilon_t,$$

onde $y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix}$, $A_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ e $\epsilon_t = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix}$.

Como dito anteriormente, nesta abordagem, tudo depende de tudo, e cada linha das matrizes acima caracterizam um único modelo, ou uma única equação, tendo um indicador como variável resposta e os demais como explicativas, além da própria resposta defasada. Por exemplo, para y_{1t} a equação é dada por $y_{1t} = a_{11} y_{1t-1} + a_{12} y_{2t-1} + a_{13} y_{3t-1} + \epsilon_{1t}$.

A matriz de covariância que contém os ϵ_{it} é simétrica, o que nos leva a concluir que o que está sendo utilizada é a estrutura de covarância, e portanto relações causais não devem ser inferidas por este tipo de modelo. Modelos mais apropriados, como o VAR estruturado, podem trazer inferências causais mais confiáveis.

Função de Resposta ao Impulso

A grosso modo, a FRI estima o impacto, em função do tempo t, de um "choque" de uma variável explicativa em uma variável resposta. Por choque, neste caso, entende-se um aumento de 1 desvio-padrão da variável explicativa, onde este desvio padrão é dado pela raiz de ϵ_{it}.

A figura abaixo apresenta a FRI para o indicador de casos Suspeitos até a data de 3 de janeiro de 2021, impactando em sua própria série e no número de leitos ocupados, de Enfermaria e UTI. Neste caso, o desvio padrão para a série de suspeitos é de 72 casos. Portanto, a cada 72 casos suspeitos, esperamos em 8 dias após um aumento de novos 40 casos suspeitos, cerca de 4 leitos de enfermaria ocupados e 1 leito de UTI.

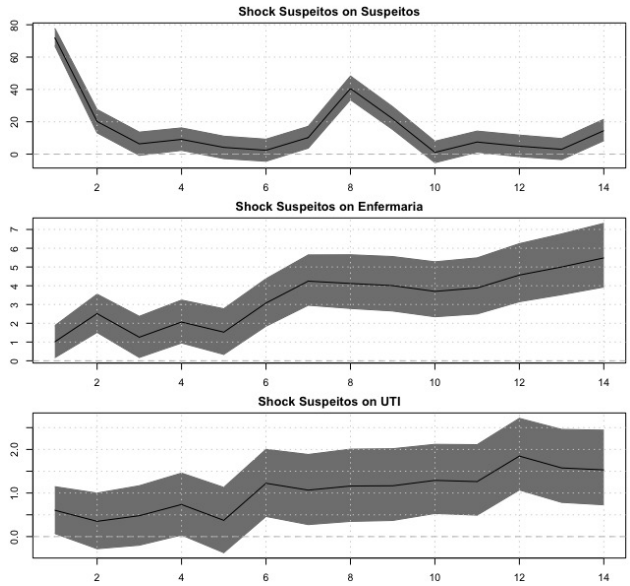

Figura 1- Exemplo de FRI

Estimação Bayesiana

Os modelos VAR possuem forte suposição de estacionaridade das séries envolvidas ao se utilizar métodos de estimação por máxima verossimilhança, sob enfoque frequentista. A grosso modo, as séries devem apresentar uma variabilidade constante ao longo da série, e um mesmo regime.

No entanto, nos chamados BVAR, os VAR estimados a partir de métodos Bayesianos, há um relaxamento desta suposição, principalmente pelo uso de distribuições a priori de determinados parâmetros no modelo (REF Sims). Desta forma, adotou-se esta forma de estimação para as séries da epidemia, que certamente não possuem estacionaridade pela própria característica do fenômeno, com mudanças de regime constantes, que são inclusive o interesse da modelagem.

O pacote BVAR, do R, utilizado nestas análises, considera a abordagem de prióris conjugadas para os parâmetros da chamada família Normal-inverse-Wishart, que inclui diversas formas de ditribuições. Para este caso, o modelo apresentado anteriormente podemos escrever $\boldsymbol{A} = [\boldsymbol{a}_0, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_p]'$ e $\boldsymbol{\beta} = vec(\boldsymbol{A})$. Então, o conjunto de prióris conjugadas é escrito como:

$$\begin{aligned} \boldsymbol{\beta}|\boldsymbol{\Sigma} &\sim \mathcal{N}(\boldsymbol{b}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}), \\ \boldsymbol{\Sigma} &\sim \mathcal{IW}(\boldsymbol{\Psi}, \boldsymbol{d}), \end{aligned} \tag{1}$$

onde $\boldsymbol{b}$, $\boldsymbol{\Omega}$, $\boldsymbol{\Psi}$ e $\boldsymbol{d}$ são funções de dimensão inferior dos hiperparâmetros $\boldsymbol{\gamma}$ da relação

$$p(\boldsymbol{\gamma}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\gamma})p(\boldsymbol{\theta}|\boldsymbol{\gamma})\ p(\boldsymbol{\gamma}), \tag{2}$$

$$p(\boldsymbol{y}|\boldsymbol{\gamma}) = \int p(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\gamma})p(\boldsymbol{\theta}|\boldsymbol{\gamma})d\boldsymbol{\theta}, \tag{3}$$

onde $\boldsymbol{y} = (\boldsymbol{y}_{p+1}, \ldots, \boldsymbol{y}_T)^\top$, os parâmetros autoregressivos e de variância do modelo VAR são denotados por $\boldsymbol{\theta}$.

São utilizados três conjuntos de distribuições a priori para este modelo. Um conjunto refere-se às prióris dos parâmetros própriamente ditos, onde são utilizadas a chamada Priori de Minnesota (*Minnesota Prior*), que considera que cada série individual seguem um processo de passeio aleatório. Matematicamente pode ser caracterizada por seus momentos:

$$\mathbb{E}[(\boldsymbol{A}_s)_{ij}|\boldsymbol{\Sigma}] = \begin{cases} 1 & \text{se } i = j \text{ e } s = 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$cov\left[(\boldsymbol{A}_s)_{ij}, (\boldsymbol{A}_r)_{kl}|\boldsymbol{\Sigma}\right] = \begin{cases} \lambda^2 \frac{1}{s^\alpha} \frac{\boldsymbol{\Sigma}_{ik}}{\psi_j/(d-M-1)} & \text{if } l = j \text{ e } r = s, \\ 0 & \text{caso contrário.} \end{cases}$$

O parâmetro chave λ controla a relação entre a importância da priori em relação aos dados observados, quanto mais próximo de 0 λ for, maior a importância da priori nas estimativas em relação aos dados, e quanto maior este valor, mais espalhará o conjunto de dados.

Os outros dois conjuntos são variáveis indicadoras adicionadas ao modelo para se contornar situações que são em geral as pricipais que levam à quebra das suposições dos modelo VAR. As prióris-de-soma-de-coeficientes e as prióris de raiz-única procuram minimizar os efeitos de componentes determinísticos no modelo, como tendências e sazonalidades. Maiores detalhes em Giannone et al. (2015).

Maiores detalhes sobre as prióris utilizadas serão divulgadas na nota técnica a ser publicada, mas utilizou-se as distribuições padrão do pacote Bvar, do software R.

O BVAR utiliza o algoritmo MCMC para as estmativas, baseadas em amostras das distribuições a posteriori. Nestas análises utilizamos 1 única cadeia com 40.000 simulações sendo 15.000 descartadas, usadas como aquecimento da cadeia (chamadas de *burn*).

Resultados da estimação do modelo de previsão

Análise exploratória para utilização do modelo

Como descrito nas notas técnicas do grupo, e em artigo publicado (Colugnati et al., 2020) o grupo de modelagem utiliza modelos de Poisson segmentados, que captam pontos de mudança de regime nas curvas, para o monitoramento das mesmas. A figura 2 apresenta as estimativas para Suspeitos, Total de leitos de Enfermaria e de UTI, respectivamente. A quebra de maior interesse é a que antecede rápida subida em torno de meados de outubro e início de novembro. O modelo estimou as seguintes datas para esta mudança de regime de rápido crescimento:

- Suspeitos: 20 de outubro de 2020, onde a curva passa a um crescimento de 3,5% diário
- Total de leitos de Enfermaria: 27 de outubro de 2020, com crescimento de 1,7%
- Total de leitos de UTI: 3 de novembro de 2020, com crescimento de 2,3%

Chama atenção a distância de 7 dias entre os pontos de inflexão nas diferentes curvas, que pode evidenciar algum padrão persistente de dependência entre as séries

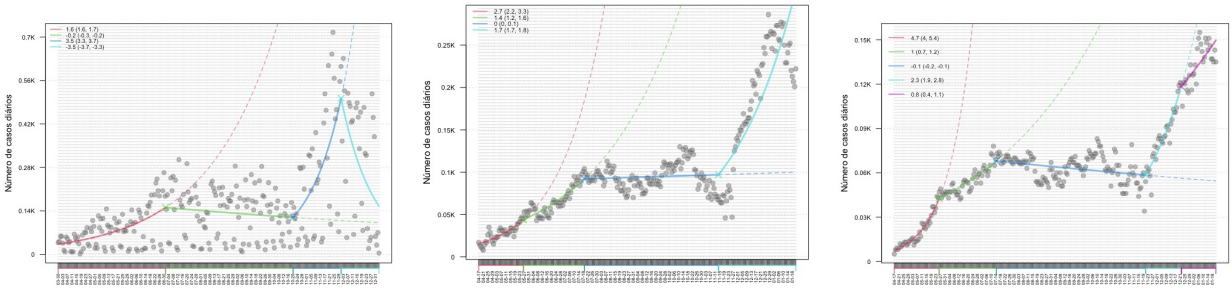

Figura 2- Segmentação das séries de Suspeitos, Enfermaria e UTI, segundo modelo de poisson

É importante avaliar se esta característica é persistente ao longo da série histórica, e entender se os intervalos de tempo entre estes eventos também se manter relativamente constantes. Desta forma, foram ajustados 4 modelos para séries com início em 30 de março de 2020 e final em:

- 31 de outubro de 2020, com desvio padrão estimado para a séria de casos suspeitos de 48 casos
- 30 de novembro de 2020, com desvio padrão estimado para a séria de casos suspeitos de 27 casos
- 28 de dezembro de 2020, com desvio padrão estimado para a séria de casos suspeitos de 48 31 casos
- 3 de janeiro de 2020, com desvio padrão estimado para a séria de casos suspeitos de 72 casos

Estas datas finais são relativas às datas dos relatórios da DVA enviados ao grupo. O objetivo é avaliar se as relações entre os indicadores e os seus valores com *lags* de defazagem, mantém a mesma relação ao longo da série, principalmente desde o início do aumento do número de internações. No corpo deste texto serão apresentadas apenas as FRI tendo Suspeitos como variável de impulso sobre as demais e sobre ela mesma. No anexo técnico, todas as análises e saídas úteis para interpretação dos modelos e seus ajustes, serão disponibilizados.

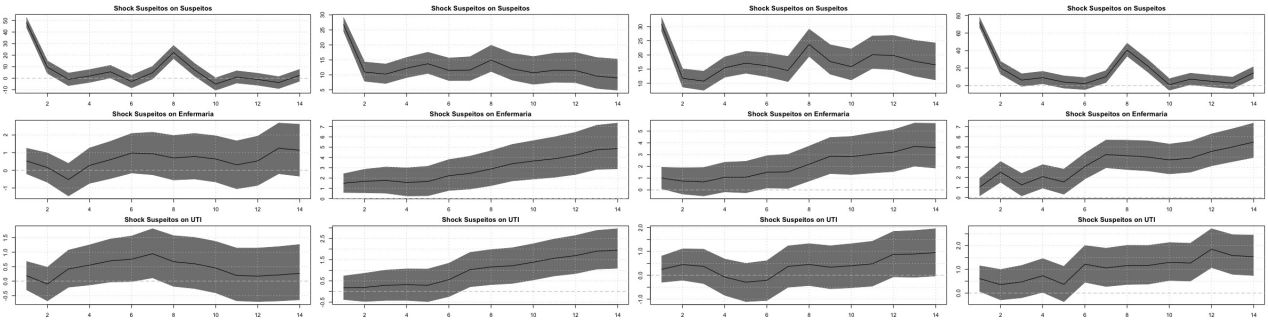

Figura 3- Gráficos das FRI para as séries terminadas em: 1. 31/10/2020, 2. 30/11/2020, 3. 28/12/2020 e 4. 03/01/2021

Podemos observar que existem consistências nas funções impulso estimadas para as 4 séries:

1. O aumento de um dp de casos suspeitos apresenta um pico no oitavo dia após este impulso
2. Este mesmo impulso leva a um impacto relativamente baixo até o sétimo dia tanto para leitos de Enfermaria quanto de UTI, quando o impacto no número de leitos dobra.

Modelo utilizado para as previsões apresentadas

Para o modelo final de previsão, optou-se por um ajuste de *lag* 14, correspondendo a 14 dias ou duas semanas. Esta opção se deu por questões do fenômeno, que como se observa ocorre em picos semanais, tendo sempre 2 semanas como horizonte das previsões. Além disso, por critérios técnicos. Foram ajustados modelos de

ordem 1, 7 e 14. A mediana das distribuições a posteriori da verossimilhança marginal em cada um dos casos foram, -3188,969; -3096.113 e -3026.605 respectivamente, indicando um melhor ajuste para os 14 dias. Além disso, as cadeias simuladas apresentaram melhor comportamento no último caso, de 14 dias de *lag*.

Os gráficos abaixo, Figura 4, apresentam as cadeias de Markov geradas pelo MCMC e as densidades da distribuição a posteriori simuladas para cada parâmetro chave das prióris Minessota, soma-de-coeficientes e raiz única, mostrando boa converência.

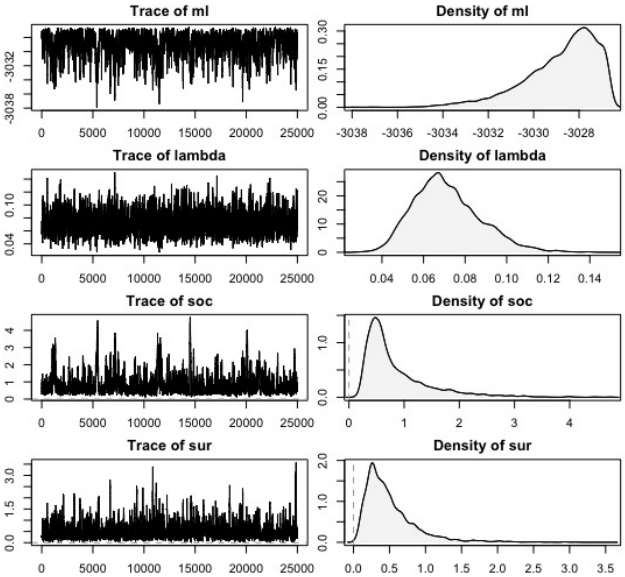

Figura 4- Gráficos diagnósticos do MCMC

A Figura 5 apresenta os resíduos, que realmente evidenciam a heterocedasticidade ao longo da série, o que procurou-se contornar com a estimação bayesiana e as prióris.

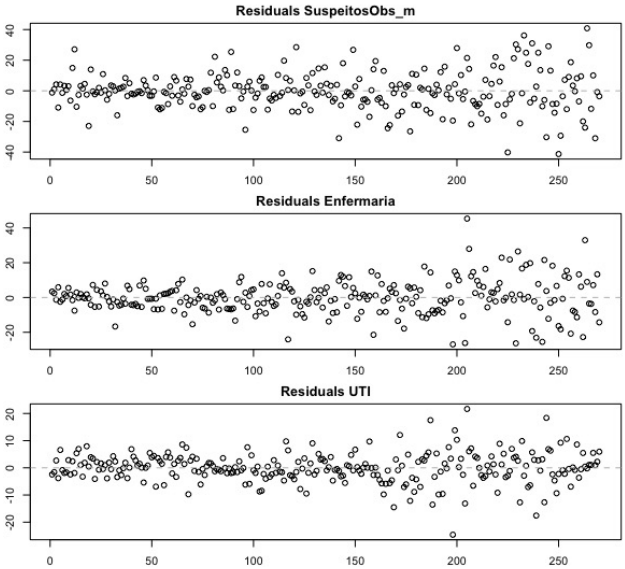

Figura 5- Resíduos gerados pelo modelo BVAR para cada série individual

Referências

Colugnati FAB, Nogueira MC, Vieira MTV, Teixeira MTB, Leite CG, Chaoubah A. Enfrentamento da epidemia da Covid-19 pela análise de dados: relato de uma equipe de pesquisadores. Revista Brasileira de Estatística, 244(78), 162-181

Kuschnig, Nikolas & Vashold, Lukas. (2019). BVAR: Bayesian Vector Autoregressions with Hierarchical Prior Selection in R. 10.13140/RG.2.2.25541.60643.

Stock J, Sims C, Watson M. (1990). Inference in Linear Time Series Models with Some Unit Roots. Econometrica, 58 (1), 113-144.

Giannone D, Lenza M, Primiceri GE (2015). Prior Selection for Vector Autoregressions. Review of Economics and Statistics, 97(2), 436–451