

Estimation of Reproduction Number for Covid-19 with Statistical Correction of Notifications Delay

Gabriela Pereira Barros, Rafaela Pereira Barros, Everthon de Souza Oliveira, Erivelton Geraldo Nepumuceno e Robson Bruno Dutra Pereira

Abstract—The pandemic of Covid-19 began in Brazil in February 2020. To evaluate the evolution of pandemics some metrics can be estimated, such as the reproduction number, R_t , and the basic reproduction number, R_0 . Due to the delay in the notifications, these estimates may present a bias. Taking the reported data, besides a sample of individuals who reported the day of symptoms onset, it is possible to estimate delay probabilities and to perform a deconvolution to correct the notifications' delay. In this work, it was performed a corrected estimate of R_t . This estimate is done based on the curve of notifications corrected through deconvolution. The approach is applied in three country cities and in the capital of Minas Gerais state. The behavior of R_t concerning the Minas Consciente program was evaluated. It was observed that the corrected R_t was more suitable to measure the effect of the program when compared to the raw R_t . When it was determined a more rigid mobility and activities regime by the program, it was observed a decrease in the median of the variation of the R_t of the cities studied.

Index Terms—Reproduction number, Deconvolution, Notifications delay, Covid-19, Sars-CoV-2.

I. INTRODUÇÃO

A COVID-19 é uma doença causada pelo coronavírus, denominado SARS-CoV-2, que apresenta um espectro clínico variando de infecções assintomáticas a quadros graves. De acordo com a Organização Mundial de Saúde, cerca de 80% dos pacientes com COVID-19 podem ser assintomáticos ou apresentarem poucos sintomas, e aproximadamente 20% dos casos detectados requerem atendimento hospitalar por apresentarem dificuldade respiratória, dos quais aproximadamente 5% podem necessitar de suporte ventilatório [1].

Para o auxílio no controle de pandemias são utilizadas algumas métricas, dentre elas, a taxa de reprodução, R_t , e a taxa básica de reprodução, R_0 , servindo para medir o potencial de crescimento da pandemia. A taxa básica de reprodução, R_0 , é o número de infecções secundárias geradas a partir de um caso

inicial em toda uma população suscetível, sendo um parâmetro do início de uma transmissão. Já a taxa de reprodução, R_t , representa uma quantidade adimensional atualizada no tempo que discrimina se os dados epidemiológicos no instante de tempo t sustentam um crescimento no número de infecções secundárias ($R_t > 1$) [2]. Portanto, R_t é utilizado quando existe um determinado número de pessoas que já se infectaram e possuem imunidade [3].

No decorrer de uma epidemia, é comum a determinação de políticas de intervenção não farmacológicas incluindo restrições de deslocamento, fechamento de escolas e de negócios não essenciais, distanciamento social e isolamento precoce de pacientes confirmados. As intervenções mais eficazes são aquelas que suprimem R_t ao máximo com o menor custo social e econômico [4]. Desde de março de 2020 os municípios de Minas Gerais, assim como de todo o Brasil, têm adotado, à sua maneira, medidas para contenção da pandemia. O Governo de Minas Gerais criou uma proposta para orientar a retomada segura das atividades econômicas nos municípios do estado. O plano “Minas Consciente” sugere a retomada gradual à normalidade através da adoção de um sistema de critérios e protocolos sanitários que garantam a segurança da população [5].

Na modelagem da epidemia, o atraso nas notificações de novos casos pode acarretar subestimativas do status da pandemia. Tais atrasos são decorrentes da baixa capacidade de testes, tempo para diagnóstico laboratorial, além de problemas logísticos e estruturais dos órgãos de saúde pública [6], [7]. A subestimativa de número de casos decorrente do atraso nas notificações afeta a estimativa de R_t e de outros índices de evolução da pandemia, dificultando a mensuração do grau da emergência de saúde pública em curso bem como seu enfrentamento.

Tomando uma amostra pequena mas suficiente de indivíduos que reportaram a data de início dos sintomas, juntamente com a data de notificação, é possível estimar as probabilidades de atraso das notificações. Tomando tais probabilidades juntamente com a curva de indivíduos doentes notificados é possível obter uma aproximação para a curva de indivíduos com manifestação dos sintomas no tempo a partir de deconvolução. Portanto, é possível corrigir o atraso nas notificações e os índices da pandemia, melhorando o controle da emergência de saúde pública.

O presente artigo tem por finalidade propor uma estimativa de R_t considerando a estimativa de indivíduos com manifestação dos sintomas no horizonte de tempo obtida pela deconvolução dos dados notificados de doentes e da probabilidade de atraso nas notificações. O método proposto é aplicado em

Gabriela Pereira Barros is with Center for Innovation in Modelling and Optimization of Systems (CIMOS), Department of Mechanical and Industrial Engineering, Federal University of São João del-Rei, Brazil, e-mail: gabrielapbarros15@aluno.ufsj.edu.br

Rafaela Pereira Barros is with Presidente Antonio Carlos University, Brazil, e-mail: 201-002131@aluno.unipac.br

Everthon de Souza Oliveira is with Department of Electrical Engineering, Federal Center for Technological Education of Minas Gerais, Brazil, e-mail: oliveira@cefetmg.br

Erivelton Geraldo Nepumuceno is with Centre for Ocean Energy Research, Department of Electronic Engineering, Maynooth University, Ireland, e-mail: erivelton.nepumuceno@mu.ie

Robson Bruno Dutra Pereira is with Centre for Innovation in Modelling and Optimization of Systems (CIMOS), Department of Mechanical and Industrial Engineering, Federal University of São João del-Rei, Brazil, e-mail: robsondutra@ufsj.edu.br

três cidades do interior além da capital do estado de Minas Gerais no Brasil. Apresenta-se, também uma estimativa da taxa básica de reprodução, R_0 , para Minas Gerais, considerando os primeiros dados de incidência do estado.

II. TAXA BÁSICA DE REPRODUÇÃO E TAXA DE REPRODUÇÃO

A taxa básica de reprodução, R_0 , pode ser estimada no início da epidemia, considerando os dados de incidência acumulada até quando o crescimento exponencial se sustenta. A taxa de crescimento exponencial pode ser matematicamente descrita como $r = \log(R_0)/s$, onde s é o intervalo serial da doença, isto é, o tempo entre a infecção e o aparecimento dos sintomas. Logo, a forma mais simples de estimar R_0 é fazendo uma regressão do logaritmo da incidência cumulativa em relação ao tempo para obter r e então calcular R_0 , segundo a Equação (1) [8].

$$R_0 = sr + 1 \quad (1)$$

A taxa básica de reprodução é um índice importante, visto que mede a capacidade da doença se espalhar, de forma que $R_0 \geq 1$ indica que a epidemia tende a crescer, enquanto $R_0 < 1$ indica a extinção da epidemia. A taxa básica de reprodução pode ser utilizada para estimar a fração da população que deve ser vacinada, sendo esta igual a $1 - 1/R_0$, de forma que a propagação do patógeno se dissipe e este não tenha a habilidade de reinvidir a população hospedeira [8].

A taxa de reprodução R_t , utilizada para medir a evolução da epidemia de forma dinâmica, pode ser estimada de forma empírica levando em consideração as observações de incidência da doença, conforme a Equação (2), onde I_t é o número de indivíduos doentes no instante t , τ é uma janela de tempo estipulada e $I_{t-\tau+1}$ é o número de indivíduos doentes no instante de tempo $t - \tau + 1$ [9].

$$R_t = \frac{I_t}{I_{t-\tau+1}} \quad (2)$$

Conforme sugere Cori et al. [10], a taxa de reprodução R_t pode ser estimada através da Equação (3), sendo s o intervalo serial médio, com distribuição de probabilidade w_s ; e I_{t-s} o número de indivíduos doentes no instante t que foram infectados a s unidades de tempo atrás e podem estar espalhando a doença. Este método requer os parâmetros da distribuição de probabilidade do intervalo serial w_s sendo recomendada a distribuição gamma. Porém, a abordagem é robusta para outras distribuições discretas positivas, tais como lognormal e weibull. O intervalo serial, necessário à estimativa de R_t , consiste no intervalo de tempo entre o instante da infecção e o momento da manifestação dos sintomas. Este intervalo é estimado, portanto, considerando a ligação da data do início dos sintomas entre pares infectante-infectado, sendo tais ligações difícil de serem estabelecidas. Com a análise de um número muito pequeno de pares pode haver um viés nas estimativas da média e variância do intervalo [11].

$$R_t = \frac{I_t}{\sum_{s=1}^t I_{t-s} w_s} \quad (3)$$

Além do atraso decorrente do aparecimento nos sintomas, contemplado no modelo de R_t de Cori et al. [10], é interessante considerar o atraso nas notificações de casos da doença. A data de notificação geralmente ocorre com um ou mais dias de atraso em relação à data de aparecimento dos sintomas. Por conseguinte, a taxa de reprodução e outras métricas para avaliar a pandemia podem ser subestimadas, de forma que medidas de mitigação podem ser tomadas tardiamente.

Cada indivíduo $i \in \{1, \dots, N\}$ com manifestação dos sintomas no dia M_i é reportado com atraso A_i no dia R_i , podendo o atraso A_i ser descrito por uma distribuição discreta, com probabilidade $p \in \{p_1, p_2, \dots, p_{t-s}\}$, $t \geq s$ onde t é o dia de confirmação do caso e s o de início dos sintomas. A Equação 4 descreve esta relação [12].

$$R_i = M_i + A_i \quad (4)$$

Seja o número de reportados no dia $t \in \{1, \dots, T\}$ igual a $Y_t = \sum_i^N I(R_i = t)$, enquanto o número de pessoas com início de manifestação dos sintomas no dia t igual a $X_t = \sum_i^N I(M_i = t)$, onde $I(x = a) = 1$, se $x = a$. Deseja-se estimar X_t , uma vez que se conhece Y_t e a distribuição de atraso A_i [12].

A esperança do número de reportados Y_t condicionada ao número de indivíduos com manifestação dos sintomas no dia s , X_s , pode ser expressa conforme Equação 5, como uma convolução da série desconhecida do número de indivíduos com manifestação dos sintomas no tempo com a probabilidade de atraso p_{t-s} . Esta quantidade é distribuída por uma Poisson com média análoga, conforme a Equação 6 [12]–[14].

$$E(Y_t|X) = \sum_{s=1}^T X_s \cdot p_{t-s} \quad (5)$$

$$Y_t \sim Poisson \left(\sum_{s=1}^T X_s \cdot p_{t-s} \right) \quad (6)$$

Pode-se estimar o número de indivíduos com início dos sintomas no tempo, X_t , a partir da deconvolução das observações de doentes reportados Y_t considerando as probabilidades de atraso p_{t-s} . Uma forma de fazer este processo é a partir de um estimador robusto de deconvolução de incidência proposto por Miller et al. [12]. Esta abordagem propõe o uso da verossimilhança das observações condicionada ao modelo subjacente dos indivíduos com início dos sintomas no tempo, conforme a Equação 7, com a verossimilhança sendo a Poisson e a média aproximada por uma *spline*, isto é, $Y_t \sim Poisson(\exp(\theta^T B(t)))$, onde $B(t)$ é o valor do modelo de regressão via *B-spline* no instante t . O termo de penalização adicionado, $\lambda \cdot r(\theta)$, garante a suavização da estimativa. Porém, neste trabalho estamos pensando na deconvolução da manifestação dos sintomas e não na incidência em si.

$$L(\theta; y_1, \dots, y_t) = -\ln(P(y_1, \dots, y_t|\theta)) + \lambda \cdot r(\theta) \quad (7)$$

A abordagem proposta por Miller et al. [12] busca aproximar a distribuição à priori de θ , ou mais especificamente um modelo para o instante de manifestação M_n dos sintomas considerando a probabilidade de manifestação dos sintomas

no dia s , $g_\theta \equiv P(M_n = s)$. Tal aproximação pode ser obtida pela maximização da função de verossimilhança, conforme Equação 8.

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta} \sum_i \ln(f(r_i)) \\ &= \arg \max_{\theta} \sum_i \ln \left(\int g_\theta(s) p(r_i|s) ds \right) \end{aligned} \quad (8)$$

Obtendo-se esta estimativa pode-se estimar a curva suavizada de manifestação dos sintomas no tempo via regra de Bayes, conforme Equação 9.

$$g_\theta(s) = \alpha \exp(\theta^T B(s)) \quad (9)$$

Neste trabalho propõe-se a correção da taxa de reprodução, R_t , levando em consideração o atraso nas notificações. Para tal, estimou-se via deconvolução a curva de indivíduos com manifestação dos sintomas a partir da curva de doentes reportados e das probabilidades de atraso nas notificações. O R_t corrigido foi então calculado tomando esta estimativa em detrimento dos dados notificados. O estudo foi aplicado em três cidades do interior - São João del-Rei, Sete Lagoas e Ouro Branco e na capital do estado de Minas Gerais, Belo Horizonte. A taxa básica de reprodução, R_0 , foi estimada para o estado de Minas Gerais.

III. PROCEDIMENTOS METODOLÓGICOS

Os dados para estimativa de R_t para as cidades foram coletados através dos boletins epidemiológicos disponibilizados nos sites das prefeituras a partir do primeiro dia em que foi disponibilizado o número de recuperados, até o último dia do ano de 2020. Já para o cálculo da taxa básica de reprodução, R_0 , foram utilizados os dados disponibilizados por Wesley Cota em seu site [15]. Os dados para estimar o atraso nas notificações foram obtidos no “Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19” do SUS [16] que contém amostras de indivíduos com data de início dos sintomas e a respectiva data de notificação a nível municipal.

Para validar o resultado da deconvolução, tomou-se a amostra de Belo Horizonte de indivíduos que reportaram data de início dos sintomas, além da data de notificação. Procedeu-se com a deconvolução desta amostra afim de avaliar se a curva estimada se aproxima bem da curva amostral de manifestação dos sintomas. Foi realizada uma regressão via *spline* para a curva amostral de indivíduos com manifestação dos sintomas. A deconvolução deve apresentar similaridade à esta aproximação com ajuste competitivo à *spline*. Esta comparação é feita via validação cruzada separando 80% das observações desta amostra para treino e 20% para teste e utilizando-se das métricas de ajuste *RMSE* e *MAE* e R^2 para medir a acuracidade da *spline* aplicada à previsão dos dados de teste. De modo similar, faz-se também o cálculo de tais métricas de ajuste para a curva estimada via deconvolução. Espera-se haver um resultado competitivo da deconvolução em relação à *spline*.

Estimou-se a taxa de reprodução a partir da Equação 3 utilizando-se a curva aproximada de indivíduos com manifestação dos sintomas no tempo, resultante da deconvolução, em

detrimento dos dados de notificações de casos. Tal estimativa corrigida de R_t tem a finalidade de medir o status da pandemia de forma corrigida, viabilizando um melhor controle. Para fins de comparação o R_t bruto também foi estimado pela mesma equação, porém considerando os dados reportados.

Os gráficos de R_t foram plotados com a sobreposição das ondas do programa Minas Consciente. No início do programa constavam quatro ondas: verde, branca, amarela e vermelha. A partir de 1º de Agosto de 2020 constavam apenas verde, amarela e vermelha. Afim de unificar a classificação, as datas de início e término de cada onda foram coletadas considerando apenas as 3 ondas [17]. Até a data final do estudo a capital não aderiu ao programa, por possuir um comitê próprio de enfrentamento à COVID-19. Logo, para Belo Horizonte, foram consideradas as ondas da macrorregião onde a cidade está inserida, que refletem medidas próximas às adotadas pela cidade. Para validar a estimativa de R_t corrigido, calculou-se a mediana da variação do R_t dentro de cada onda para cada cidade estudada, e foi verificado se o R_t corrigido captura melhor o efeito da mudança de onda em relação ao R_t bruto. Espera-se que a mudança para uma onda mais rígida resulte em redução no R_t .

A Tabela I, resume estimativas do intervalo serial da COVID-19 encontradas na literatura. Esta métrica é necessária para o cálculo de R_t e R_0 e foi considerada a estimativa feita a partir de um maior número de pares infectante-infectado [11].

TABELA I
ESTIMATIVAS DO INTERVALO SERIAL

Referência	Pares	Distribuição	Média	Desvio Padrão
Najafi et al. [18]	21	Weibull	5,7	3,9
Li et al. [19]	06	Gama	7,5	3,4
Nishiura et al. [11]	28	Weibull	4,8	2,3
Viego et al. [20]	17	Gama	6,9	5,9
Wang et al. [21]	12	Weibull	3,0	1,9

As análises foram conduzidas na linguagem *R*, sendo utilizados os pacotes: “*fitdistrplus*” [22] para aproximar as distribuições de probabilidade, “*EpiEstim*” [23] para estimativa de R_t e “*incidental*” [24] para deconvolução. As análises e dados foram disponibilizados em <https://github.com/robsonpro/covid19-statistical-correction-delay>.

IV. RESULTADOS

A Figura 1 apresenta a curva de casos acumulados em escala log para o estado de Minas Gerais. Os dados do 5º ao 16º dia, onde se observa um crescimento exponencial da epidemia, foram considerados para estimativa da taxa básica de reprodução, R_0 , utilizando a Equação 1, com r obtido pela inclinação do modelo de regressão linear para o log de casos em função do tempo, resumido na Tabela II, isto é, $r = 0,3983$, enquanto o intervalo serial médio considerado foi $s = 4,8$, tomando a referência com o maior número de pares na estimativa [11], conforme resumido na Tabela I. Logo, a taxa básica de reprodução média obtida para Minas Gerais é de 2,91, com limites de 0,95 de confiança iguais a [2,76; 3,07]. Souza *et al.* [25] obtiveram o mesmo valor médio para São Paulo e Rio de Janeiro. Logo, para que a epidemia perca

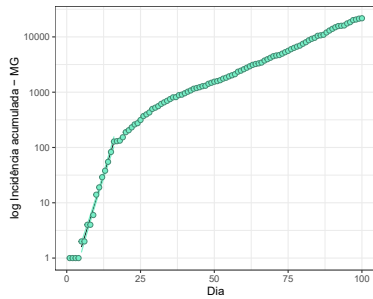


Fig. 1. Casos acumulados em Minas Gerais com eixo y em escala log e regressão loglinear para o período inicial de crescimento exponencial.

força, a proporção média de $1 - 1/R_0 = 65,52\%$ da população do estado precisa ser vacinada.

TABELA II

MODELO LINEAR EM ESCALA LOG PARA CRESCIMENTO EXPONENCIAL DE CASOS

Termo	Coefficiente	Erro-padrão	Valor-t	Pr(> t)
Constante	-1,531	0,1599	-9,57	2,374e-06
log(Dia)	0,3983	0,01447	27,52	9,287e-11
$R_{aj}^2: 0,9857$				

Segundo Shim *et al.* [26] para uma avaliação mais precisa das taxas de crescimento epidêmico, devem ser analisadas as datas de início dos sintomas, e não as datas de notificação. Assim, foi aproximada a probabilidade de atraso das notificações. Na Figura 2 pode-se verificar que as observações de atraso das cidades apresentam um bom ajuste para as distribuições Weibull, Gama e Lognormal. A Tabela III apresenta os parâmetros estimados para tais distribuições. Porém, para estimar os indivíduos com manifestação dos sintomas no tempo via deconvolução foi considerada a distribuição amostral como boa aproximação para o caso populacional.

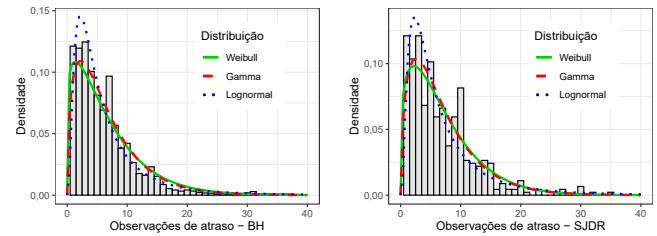
TABELA III

DISTRIBUIÇÕES DE PROBABILIDADE ESTIMADAS PARA O ATRASO NAS NOTIFICAÇÕES

Dist	BH (30425*)	SJDR (456)	OB (265)	SL (799)
Weibull ^a	1,18;7,06	1,26;7,54	1,33;7,62	1,33;7,49
Gamma ^b	1,51;0,23	1,59;0,23	1,78;0,25	1,75;0,26
Lognormal ^c	1,53;0,86	1,59;0,88	1,63;0,82	1,61;0,83

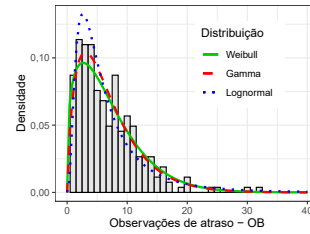
a:forma;escala, b:forma;taxa, c:meanlog;sdlog, *tamanho amostral
BH: Belo Horizonte, SJDR: São João del-Rei, OB: Ouro Branco, SL: Sete Lagoas

A Figura 3 apresenta o resultado da deconvolução para estimar a curva de indivíduos com manifestação dos sintomas no tempo a partir dos dados reportados de doentes e das probabilidades de atraso. Os pontos em '+' vermelho consistem nos dados reportados de doentes. A curva em linha preta contínua consiste na estimativa de indivíduos com manifestação dos sintomas no tempo obtida via deconvolução, podendo ser utilizada para estimar a taxa de reprodução no tempo, viabilizando a obtenção de uma estimativa corrigida de R_t . A linha tracejada azul consiste em uma regressão por *spline* dos dados reportados, enquanto as linhas contínuas

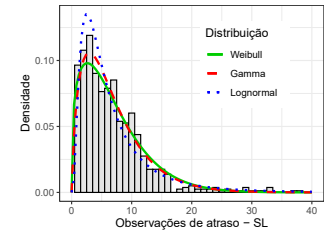


(a) Belo Horizonte

(b) São João Del Rei

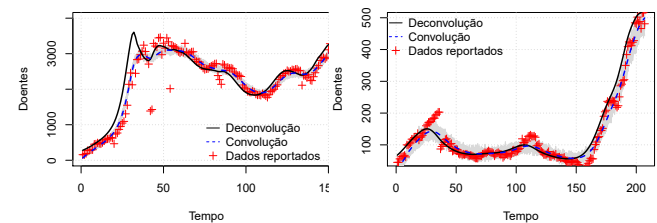


(c) Ouro Branco

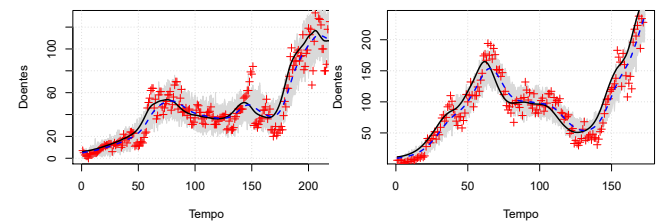


(d) Sete Lagoas

Fig. 2. Histograma e densidades teóricas para os dados de atraso das cidades estudadas.



(a)



(c) Ouro Branco

(d) Sete Lagoas

Fig. 3. Deconvolução dos dados reportados de doentes para as cidades estudadas.

em cinza são resultantes de 100 simulações aleatórias da convolução realizada a partir da curva estimada de indivíduos com manifestação dos sintomas no tempo e das probabilidades de atraso, sendo estas sorteadas aleatoriamente segundo um *bootstrap* não paramétrico, afim de avaliar o erro aleatório associado a tais probabilidades amostrais. As simulações de forma geral abrangem bem os dados reportados e a *spline*. Em alguns momentos percebem-se desvios maiores nas observações, possivelmente devido a presença de dados represados. A diferença observada na dispersão das simulações para as quatro cidades está relacionada principalmente à magnitude da curva de doentes em cada caso, visto que as probabilidades amostrais de atraso são similares para as quatro cidades,

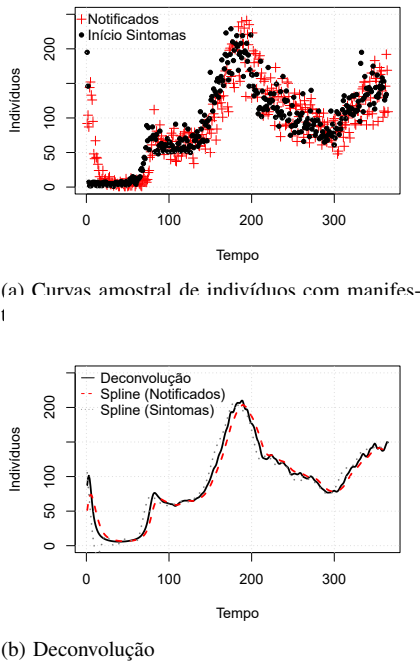


Fig. 4. Amostra de indivíduos que reportaram datas de manifestação dos sintomas para Belo Horizonte.

conforme observado na Figura 2.

Para validar a deconvolução, considerando a amostra de indivíduos que reportaram datas de manifestação dos sintomas além da incidência para a cidade de Belo Horizonte, a Figura 4a expõe graficamente a curva de notificados, em '+' vermelho, juntamente com a respectiva curva dos mesmos indivíduos no instante da manifestação dos sintomas, em pontos pretos, para esta amostra. A Figura 4b ilustra o resultado obtido, sendo a linha preta contínua a estimativa da curva de indivíduos com manifestação dos sintomas no tempo obtida via deconvolução. Esta curva visa aproximar as observações da amostra de indivíduos com manifestação dos sintomas no tempo plotada em pontos pretos na Figura 4a. Já a curva em linha tracejada vermelha da Figura 4b é a regressão por *spline* da curva de notificados em '+' vermelho plotada na Figura 4a, enquanto a curva pontilhada cinza da Figura 4b é uma regressão via *spline* da curva de indivíduos com manifestação dos sintomas no tempo plotada em pontos pretos na Figura 4a. Esta *spline* foi estimada considerando 80% das observações amostrais de indivíduos com manifestação dos sintomas no tempo, sendo as 20% remanescentes separadas para teste e validação, afim de evitar sobreajuste do modelo estimado.

A Tabela IV expõe as estatísticas de ajuste para avaliar a estimativa da curva de indivíduos com manifestação dos sintomas no tempo, obtida via deconvolução, em relação à curva amostral de indivíduos com manifestação dos sintomas no tempo. A *spline* ajustada aos dados de manifestação dos sintomas é referência para esta validação. Comparando-se as métricas *RMSE*, R^2 e *MAE* da regressão via *spline* para os dados de treino e teste, pode-se observar boa proximidade entre as métricas garantindo um bom ajuste às observações e ausência de sobreajuste. Como os valores das métricas para a estimativa obtida via deconvolução apresenta valores

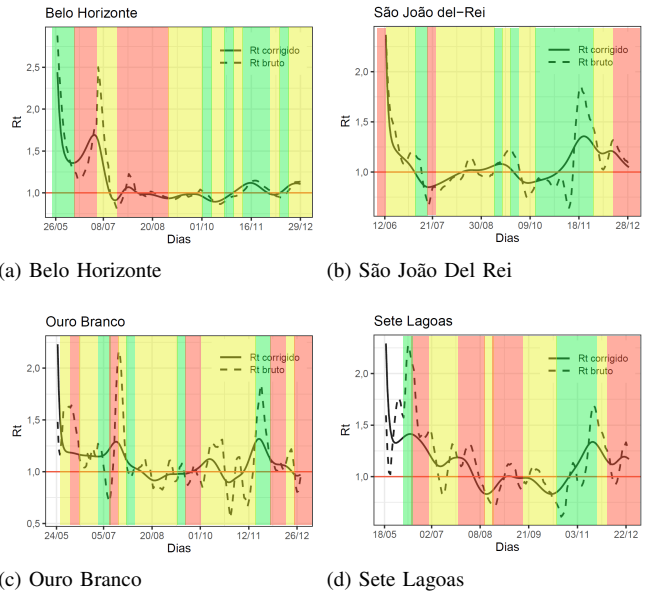


Fig. 5. R_t corrigido e bruto com sobreposição das ondas do programa Minas Consciente.

próximos aos da *spline* aplicada aos dados de teste, pode-se confirmar a eficácia do método para aproximar as observações de indivíduos com manifestação dos sintomas no tempo para toda a população de doentes, conforme feito e plotado na Figura 3.

TABELA IV
MÉTRICAS PARA AVALIAR O AJUSTE DA DECONVOLUÇÃO COMPARATIVAMENTE À *spline* PARA A AMOSTRA DE INDIVÍDUOS COM MANIFESTAÇÃO DOS SINTOMAS NO TEMPO DE BELO HORIZONTE

	<i>RMSE</i>	R^2	<i>MAE</i>
Deconvolução	21.60	0.8567	14.48
<i>Spline</i> (Treino)	18.68	0.8973	13.11
<i>Spline</i> (Validação)	18.09	0.8803	12.21

A Figura 5 apresenta os gráficos com as estimativas de R_t para as cidades estudadas, com a linha contínua trazendo o R_t corrigido considerando a estimativa de indivíduos com manifestação dos sintomas obtidos via deconvolução, enquanto a linha tracejada expressa o R_t bruto estimado considerando os dados de doentes reportados. O sombreado nas cores verde, amarela e vermelha representa as ondas do programa Minas Consciente em cada cidade. A estimativa de R_t corrigida, considerando a deconvolução, pretende medir a evolução da epidemia corrigindo o atraso nas notificações de casos.

Espera-se que a mudança para ondas com mais restrições de circulação no programa Minas Consciente acarrete uma mudança no R_t . Após alguns dias de mudança de onda verde para a amarela ou de onda amarela para a vermelha espera-se uma redução na mobilidade e, conseqüentemente, no R_t . A Figura 6 plota a mediana da variação do R_t em cada onda, sendo que 1, 2 e 3 no eixo horizontal corresponde respectivamente às ondas verde, amarela e vermelha. Os marcadores em círculo vermelho correspondem ao R_t bruto, enquanto os marcadores

em triângulo azul correspondem ao R_t corrigido. A tendência de queda na mediana da variação do R_t com a mudança de onda é observada mais claramente no R_t corrigido. Para as cidades do interior, pode-se constatar que a mediana da variação do R_t corrigido cai com a mudança de onda verde para amarela (1-2) e amarela para a vermelha (2-3). Já para Belo Horizonte não percebe-se um decréscimo na mediana da variação do R_t corrigido na mudança de onda verde para amarela. Esta dificuldade de medir o efeito do Minas Consciente pela taxa de reprodução R_t corrigida para a capital pode ser explicada pela não adesão de Belo Horizonte ao programa. As ondas definidas e plotadas junto com o R_t na Figura 5a são relacionadas à região metropolitana de Belo Horizonte, conforme definido pelo governo do estado. Para o R_t bruto, ou seja, sem correção do atraso nas notificações, para todas as cidades avaliadas, fica difícil se estabelecer uma relação entre o programa e esta métrica sem correção, observando-se para todas as cidades um comportamento confuso da mediana da variação do R_t bruto com a mudança de onda.

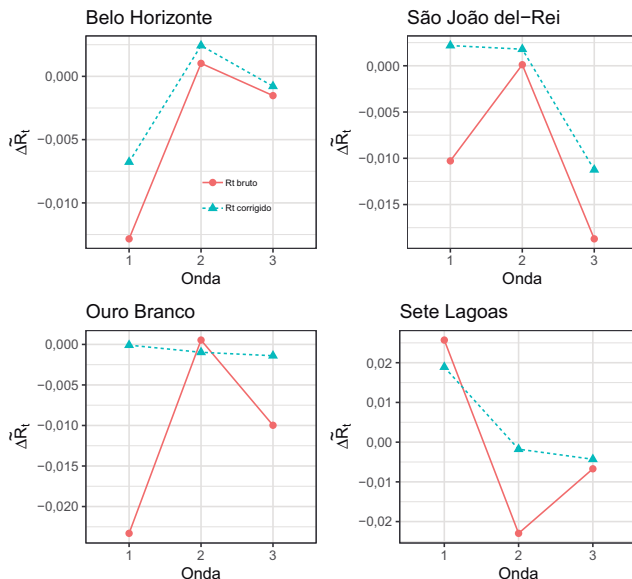


Fig. 6. Relação entre mediana da derivada no tempo do R_t e mudança de ondas do programa Minas Consciente.

V. CONCLUSÕES

Neste trabalho realizou-se uma estimativa corrigida da taxa de reprodução, sendo esta calculada a partir da estimativa de indivíduos com manifestação dos sintomas no tempo obtida via deconvolução. A deconvolução foi realizada considerando os dados reportados de doentes e a probabilidade amostral de atraso nas notificações. A partir destas estimativas calculou-se o R_t corrigido para Belo Horizonte e para três cidades do interior de Minas Gerais. Para fins de comparação, o R_t bruto também foi calculado, a partir dos dados reportados. A taxa básica de reprodução R_0 também foi calculada para o estado de Minas Gerais. Esse dado é importante para medir o potencial de crescimento da pandemia e indicar medidas de enfrentamento ou planejamento vacinal da população.

Constatou-se que a estimativa corrigida de R_t é mais influenciada pelas medidas de restrição de mobilidade se comparada ao R_t estimado com base nos dados de notificações brutas. Este resultado evidencia aspectos importantes no enfrentamento de epidemias a partir de modelos baseados em dados e permite fazer as seguintes conclusões.

- Um sistema de notificações de casos da doença com um atraso aleatório prejudica a tomada de decisão e a avaliação das medidas adotadas e, por isso, é relevante garantir a qualidade das notificações.
- O uso de ferramentas estatísticas pode minimizar o efeito do atraso nas notificações e proporcionar melhorias na modelagem e no monitoramento da epidemia.
- No contexto do problema considerado, as medidas adotadas poderiam equivocadamente ser consideradas ineficientes se desprezados os efeitos do retardo na compilação dos dados diários.

Os resultados apresentados levantam interessantes questões a serem investigadas em trabalhos futuros. Uma delas é considerar essa abordagem para outras epidemias sobre as quais se tenham as informações equivalentes as utilizadas aqui. Além disso, é interessante avaliar a correlação entre a taxa de reprodução corrigida e séries temporais de mobilidade, por exemplo, as obtidos via GPS de smartphones ou via dados reportados de agências de transporte de pessoas.

AGRADECIMENTOS

Os autores agradecem o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERÊNCIAS

- [1] M. da Saúde, "Sobre a doença - coronavírus."
- [2] D. Pasetto, J. C. Lemaître, E. Bertuzzo, M. Gatto, and A. Rinaldo, "Range of reproduction number estimates for covid-19 spread," *Biochemical and biophysical research communications*, vol. 538, pp. 253–258, 2021.
- [3] S. Wang, X. Yang, L. Li, P. Nadler, R. Arcucci, Y. Huang, Z. Teng, and Y. Guo, "A bayesian updating scheme for pandemics: Estimating the infection dynamics of covid-19," *IEEE Computational Intelligence Magazine*, vol. 15, no. 4, pp. 23–33, 2020.
- [4] I. McCulloh, K. Kiernan, and T. Kent, "Improved estimation of daily covid-19 rate from incomplete data," in *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)*, pp. 153–158, IEEE, 2020.
- [5] Governo de Minas Gerais, "Minas consciente - entenda o plano." Acesso em: 11 maio 2021.
- [6] A. J. Sarnaglia, B. Zamprogno, F. A. F. Molinares, L. G. de Godoi, and N. A. J. Monroy, "Correcting notification delay and forecasting of covid-19 data," *Journal of mathematical analysis and applications*, p. 125202, 2021.
- [7] D. A. M. Villela, "How limitations in data of health surveillance impact decision making in the covid-19 pandemic," *Saúde em debate*, pp. 206–218, 2021.
- [8] O. N. Bjornstad, *Epidemics: Models and Data using R*. Switzerland, AG: Springer International Publishing, 2018.
- [9] V. López and M. Čukić, "A dynamical model of sars-cov-2 based on people flow networks," *Safety Science*, vol. 134, p. 105034, 2021.
- [10] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez, "A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics," *American Journal of Epidemiology*, vol. 178, pp. 1505–1512, 09 2013.

- [11] H. Nishiura, N. M. Linton, and A. R. Akhmetzhanov, "Serial interval of novel coronavirus (covid-19) infections," *International Journal of Infectious Diseases*, vol. 93, pp. 284–286, 2020.
- [12] A. C. Miller, L. Hannah, J. Futoma, N. J. Foti, E. B. Fox, A. D'Amour, M. Sandler, R. A. Saurous, and J. Lewnard, "Statistical deconvolution for inference of infection time series," *medRxiv*, 2020.
- [13] E. Goldstein, J. Dushoff, J. Ma, J. B. Plotkin, D. J. Earn, and M. Lipsitch, "Reconstructing influenza incidence by deconvolution of daily mortality time series," *Proceedings of the National Academy of Sciences*, vol. 106, no. 51, pp. 21825–21829, 2009.
- [14] S. Wang, X. Yang, L. Li, P. Nadler, R. Arcucci, Y. Huang, Z. Teng, and Y. Guo, "A bayesian updating scheme for pandemics: estimating the infection dynamics of covid-19," *IEEE Computational Intelligence Magazine*, vol. 15, no. 4, pp. 23–33, 2020.
- [15] Wesley Cota, PhD, "Número de casos confirmados de covid-19 no brasil." Acesso em: 18 maio 2020.
- [16] Governo Federal, "Opendatasus." Acesso em: 05 mar. 2021.
- [17] Governo de Minas Gerais, "Minas consciente - transparência." Acesso em: 19 maio 2021.
- [18] F. Najafi, N. Izadi, S.-S. Hashemi-Nazari, F. Khosravi-Shadmani, R. Nikbakht, and E. Shakiba, "Serial interval and time-varying reproduction number estimation for covid-19 in western iran," *New Microbe and New Infect*, vol. 36, no. 100715, 2020.
- [19] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia," *New England journal of medicine*, 2020.
- [20] V. Viego, M. Geri, J. Castiglia, and E. Jouglard, "Período de incubación e intervalo serial para covid-19 en una cadena de transmisión en bahia blanca (argentina)," *Ciência 'I&' Saúde Coletiva*, vol. 25, no. 9, 2020.
- [21] S. Zhao, D. Gao, Z. Zhuang, M. K. Chong, Y. Cai, J. Ran, P. Cao, K. Wang, Y. Lou, W. Wang, *et al.*, "Estimating the serial interval of the novel coronavirus disease (covid-19): a statistical analysis using the public data in hong kong from january 16 to february 15, 2020," *Frontiers in Physics*, vol. 8, p. 347, 2020.
- [22] M. L. Delignette-Muller and C. Dutang, "fitdistrplus: An R package for fitting distributions," *Journal of Statistical Software*, vol. 64, no. 4, pp. 1–34, 2015.
- [23] A. Cori, *EpiEstim: Estimate Time Varying Reproduction Numbers from Epidemic Curves*, 2019. R package version 2.2-1.
- [24] A. Miller, L. Hannah, N. Foti, and J. Futoma, *incidental: Implements Empirical Bayes Incidence Curves*, 2020. R package version 0.1.
- [25] W. M. de Souza, L. F. Buss, D. da Silva Candido, J.-P. Carrera, S. Li, A. E. Zarebski, R. H. M. Pereira, C. A. Prete, A. A. de Souza-Santos, K. V. Parag, *et al.*, "Epidemiological and clinical characteristics of the covid-19 epidemic in brazil," *Nature Human Behaviour*, vol. 4, no. 8, pp. 856–865, 2020.
- [26] E. Shim, A. Tariq, and G. Chowell, "Spatial variability in reproduction number and doubling time across two waves of the covid-19 pandemic in south korea, february to july, 2020," *International Journal of Infectious Diseases*, vol. 102, pp. 1–9, 2021.



Everthon Souza de Oliveira Graduated in Electrical Engineering from the Federal University of São João Del-Rei (2009) and PhD. in electrical engineering from Federal University of Minas Gerais - UFMG (2021). Everthon is currently a professor at the Federal Center for Technological Education of Minas Gerais. Everthon has experience in the field of Electrical Engineering with an emphasis on Modeling and Control of Non-Linear Dynamic Systems and Fuzzy Systems. Everthon is interested in the following topics: Chaos in Electronic Circuits, UPS Synchronism, Control and Modeling by Fuzzy Logic, Adaptive Control. E-mail: oliveira@cetefmg.br



Erivelton Geraldo Nepomuceno Erivelton Nepomuceno received his BE and PhD in Electrical Engineering from UFSJ (2001) and UFMG (2005) respectively. He was Associate Professor at Federal University of São João del-Rei until 2021, when he was appointed as Assistant Professor at the Centre for Ocean Energy Research and Department of Electronic Engineering at Maynooth University. He was a visiting Research Fellow at Technological Institute of Aeronautics in Brazil (2005), Imperial College London (2013/14), Saint Petersburg Electrotechnical University in Russia (2019) and at City, University of London (2020/21). Dr Nepomuceno is a Senior Member of IEEE and elected Secretary of the IEEE Technical Committee on Nonlinear Circuits and Systems. He has been elected coordinator of the Technical Committee on System Identification and Data Science for Brazilian Association of Automatic Control. He is Deputy EiC of IEEE Latin America Transactions, and he serves as Associate editor for IEEE Transactions on Circuits and Systems II: Express Briefs, Journal of Control, Automation and Electrical Systems, and Mathematical Problems in Engineering. His research interests include analysis, modelling and control of circuits, systems and networks, computer arithmetic, system identification, intelligent systems, chaos, cryptography. Email: erivelton.nepomuceno@mu.ie



Robson Bruno Dutra Pereira Industrial Engineer from Federal University of Ouro Preto (2009) and Ph.D. in Industrial Engineering from Federal University of Itajubá (2017). Robson is Professor at the Federal University of São João del-Rei - UFSJ. Research topics: Process modeling and optimization through the design of experiments, multi-objective optimization, multivariate statistics, statistical process control, measurement systems analysis, statistical learning, and stochastic processes. E-mail: robsondutra@ufsj.edu.br



Gabriela Pereira Barros Industrial Engineer from Federal University of São João del-Rei - UFSJ (2022). Scholarship at Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Published paper: Otimização da Furação Orbital para Obtenção de Furos em Polímero Reforçado por Fibra de Carbono in the 11st Congresso Brasileiro de Engenharia de Fabricação (COBEF), 2021. E-mail: gabrielapbarros15@aluno.ufsj.edu.br



Rafaela Pereira Barros Law student at Presidente Antônio Carlos University – UNIPAC. Currently, she is an intern at a law office and voluntary conciliator at the Justice Court of Minas Gerais – TJMG, located in the Special Civil Court of the District of Barbacena. In the present paper Rafaela contributed to data collect to R_t estimate through the epidemiological bulletins available on the websites of the local public government of the cities studied. E-mail: 201-002131@aluno.unipac.br