

Santos, ALP, Figueiredo, MPS, Ferreira, TAE, Gomes-Silva, F, Moreira, GR, Silva, JE & Freitas, JR (2020). Analysis and forecasting of the evolution of COVID-19 death numbers in the state of Pernambuco and Ceará using regression models. *Research, Society and Development*, 9(7): 1-24, e602974551.

**Análise e previsão da evolução do número de óbitos por COVID-19 do estado de Pernambuco e Ceará utilizando modelos de regressão**  
**Analysis and forecasting of the evolution of COVID-19 death numbers in the state of Pernambuco and Ceará using regression models**  
**Análisis y pronóstico de la evolución del número de muertes de COVID-19 en el estado de Pernambuco y Ceará utilizando modelos de regresión**

Recebido: 15/05/2020 | Revisado: 18/05/2020 | Aceito: 19/05/2020 | Publicado: 27/05/2020

**André Luiz Pinto dos Santos**

ORCID: <https://orcid.org/0000-0002-7703-2102>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [andredefensor@hotmail.com](mailto:andredefensor@hotmail.com)

**Marcela Portela Santos de Figueiredo**

ORCID: <https://orcid.org/0000-0002-5042-1247>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [portela.marcela.producao@gmail.com](mailto:portela.marcela.producao@gmail.com)

**Tiago Alessandro Espínola Ferreira**

ORCID: <https://orcid.org/0000-0002-2131-9825>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [taef.first@gmail.com](mailto:taef.first@gmail.com)

**Frank Gomes-Silva**

ORCID: <https://orcid.org/0000-0002-3481-3099>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [franksinatrags@gmail.com](mailto:franksinatrags@gmail.com)

**Guilherme Rocha Moreira**

ORCID: <https://orcid.org/0000-0001-6344-1151>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [guirocham@gmail.com](mailto:guirocham@gmail.com)

**José Eduardo Silva**

ORCID: <https://orcid.org/0000-0002-8483-0305>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [profeduardosilva3@gmail.com](mailto:profeduardosilva3@gmail.com)

**Jucarlos Rufino de Freitas**

ORCID: <https://orcid.org/0000-0002-3497-4263>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [jucarlos123@hotmail.com](mailto:jucarlos123@hotmail.com)

## Resumo

Objetivou-se ajustar diferentes modelos não lineares na análise a dados de óbitos por COVID-19 nos estado de Pernambuco e Ceará e fazer extrapolações do número de óbitos por meio de previsões. Neste relatório, analisou-se os dados oficiais epidêmicos disponibilizados pelo Ministério da saúde do Brasil (MS), referindo-se ao período 25/03/2020 a 11/05/2020 para o estado de pernambuco - PE e no estado do Ceará no período de 26/03/2020 a 11/05/2020 do número de óbitos confirmados por COVID-19. Para a comparação entre os modelos empregaram-se o coeficiente de determinação ajustado ( $R_{\alpha_j}^2$ ), quadrado médio dos resíduos (QMR) e critério de informação de Akaike (AIC). Todos os modelos tiveram bons ajustes, com valores de aproximadamente 99%. A verificação dos pressupostos dos resíduos foi realizada por meio de análises gráficas, e os pressupostos foram atendidos. Calculou-se o número acumulado de mortes no período de 12/05/2020 a 10/10/2020 para o estado de Pernambuco e 12/05/2020 a 11/10/2020 para o estado do Ceará, além da extrapolação da taxa de crescimento absoluto (TCA) para os respectivos intervalos. As análises indicaram que os pontos de inflexões de todos os modelos ocorreram dentro do período de 200 dias após o início da pandemia. Entretanto, não é possível ainda fazer projeções seguras de quando os números de casos confirmados de óbitos minimizarão. Independentemente da possível incerteza da previsão dos modelos, as observações indicam que os próximos dias podem ser críticos para determinar o crescimento futuro dos casos de óbitos.

**Palavras-chave:** Coronavírus; Previsão; Modelagem Epidêmica; Pandemia.

## Abstract

The purpose was defined to adjust different non-linear models in the analysis to death data by

COVID-19 in Pernambuco and Ceará and to extrapolate the deaths numbers through forecasts. In this report, we analyze the official epidemic data available by the Ministry of Health of Brazil (MS), referring to the period of 25/03/2020 to 11/05/2020 for Pernambuco - PE and in the period of 26/03/2020 to 11/05/2020 for Ceará, of the deaths numbers, COVID-19 confirmed. For the comparison between the models, the adjusted coefficient of determination ( $R_{aj}^2$ ), residual mean squares (RMS), and Akaike information criterion (AIC) were used. All models had good adjustments, with values of approximately 99%. The verification of the assumptions of the residues was carried out through graphic analyzes, and the assumptions were met. The cumulative deaths' numbers in the period from 12/05/2020 to 10/10/2020 was calculated for Pernambuco and 12/05/2020 to 11/10/2020 for Ceará, in addition to the extrapolation of the absolute growth rate (AGR) for the respective intervals. The analyzes indicated that the inflection points of all models occurred within 200 days after the start of the pandemic. However, it is not yet possible to make reliable projections of when the numbers of confirmed deaths will minimize. Regardless of the possible uncertainty of the models' prediction, our observations indicate that the next few days may be critical in determining the future growth of death cases.

**Keywords:** Coronavirus; Forecast; Epidemic Modeling; Pandemic.

### Resumen

El propósito se definió para ajustar diferentes modelos no lineales en el análisis a los datos de muerte por COVID-19 en Pernambuco y Ceará y extrapolar los números de muertes a través de pronósticos. En este informe, analizamos los datos de epidemia oficiales disponibles por el Ministerio de Salud de Brasil (MS), en referencia al período del 25/03/2020 a 11/05/2020 para Pernambuco - PE y en el período del 26/03/2020 a 11/05/2020 para Ceará, de las cifras de muertes, confirmó COVID-19. Para la comparación entre los modelos, se utilizaron el coeficiente de determinación ajustado ( $R_{aj}^2$ ), los cuadrados medios residuales (RMS) y el criterio de información de Akaike (AIC). Todos los modelos tuvieron buenos ajustes, con valores de aproximadamente el 99%. La verificación de los supuestos de los residuos se llevó a cabo mediante análisis gráficos y se cumplieron los supuestos. Los números de muertes acumuladas en el período del 12/05/2020 a 10/10/2020 se calcularon para Pernambuco y 12/05/2020 a 11/10/2020 para Ceará, además de la extrapolación de la tasa de crecimiento absoluta (AGR) para los respectivos intervalos. Los análisis indicaron que los puntos de inflexión de todos los modelos ocurrieron dentro de los 200 días posteriores al inicio de la

pandemia. Sin embargo, aún no es posible hacer proyecciones confiables de cuándo se minimizará el número de muertes confirmadas. Independientemente de la posible incertidumbre de la predicción de los modelos, nuestras observaciones indican que los próximos días pueden ser críticos para determinar el crecimiento futuro de los casos de muerte.

**Palabras clave:** Coronavirus; Pronóstico; Modelado epidémico; Pandemia.

## 1. Introdução

No final de dezembro de 2019, na China, diversas unidades de saúde relataram aglomerados de pacientes com pneumonia de causa desconhecida, que estavam epidemiologicamente ligados a um mercado atacadista de frutos do mar em Wuhan, província de Hubei, China. Em 31 de dezembro, o Centro Chinês de Controle e Prevenção de doenças enviou uma equipe de resposta rápida para acompanhar a província de Hubei e Wuhan, e pesquisadores relataram os resultados desta investigação, identificando a fonte dos aglomerados de pneumonia e descreveram um novo coronavírus 2019-nCoV, detectado em pacientes com pneumonia (Na Zhu et al., 2020).

Mesmo com este cenário, o Aeroporto de Wuhan continuou funcionando com voos para 113 destinos em 22 países, além de 78 voos domésticos. Casos internacionais importados começaram a ocorrer especialmente na Ásia (Rodríguez-Morales, MacGregor, Kanagarajah, Patel, & Schlagenhauf, 2020a), mas também se espalhou, atingindo países em todas as partes do globo (Fanelli & Piazza, 2020; Rodríguez-Morales et al., 2020a).

Este vírus, também chamado de COVID-19 pela World Health Organization (WHO) (World Health Organization, 2020) é causado por um betacoronavírus oficialmente denominado SARS-Cov-2 (Bedford et al., 2020), que afeta o trato respiratório inferior e se manifesta como pneumonia em humanos. Apesar de tomadas diversas atitudes pelas autoridades governamentais, tais como contenção e quarentena, a incidência de COVID-19 continua a aumentar no mundo (Sohrabi et al., 2020).

Em 25 de abril de 2020, o ministro da saúde do Brasil, confirmou o primeiro caso de COVID-19 no país, um homem que havia viajado para a Itália. Este foi o primeiro caso de coronavírus na América do Sul (Rodríguez-Morales et al, 2020b).

Em 12 de março de 2020, os primeiros casos de covid-19 foram confirmados pelo estado de Pernambuco, um casal de idosos que veio da Itália. Pernambuco foi o terceiro estado do Nordeste a registrar casos de coronavírus (Carneiro, 2020). Na data de onze de maio

de dois mil e vinte, o estado possuía 13768 casos confirmados de coronavírus, assim como 1087 óbitos, 61 dias após o primeiro caso confirmado de covid-19 no estado. Já o estado do Ceará teve seu primeiro registro no dia 15 de março e em onze de maio de dois mil e vinte possuía 17599 casos confirmados e 1189 óbitos (MS,2020a).

Segundo o boletim epidemiológico número 15 do Ministério da Saúde, Pernambuco e Ceará estão entre os quatro primeiros estados do Brasil com maiores incidências de óbitos no país (MS, 2020b). O Ministério da Saúde sugere que as decisões sobre a adoção ou flexibilização de medidas restritivas devam ser realizadas pela autoridade local, assim como a avaliação de riscos à nível municipal, estadual e regional (MS, 2020c). Dessa forma, estudos que contemplem os dados estaduais poderiam auxiliar na tomada de decisão dos governantes, por refletirem melhor as características da evolução da pandemia a nível local.

Na literatura, são encontrados diversos modelos matemáticos não lineares e lineares que podem ser utilizados para descrever dados epidemiológicos. Ghosal, Sengupta, Majumder e Sinha (2020) fizeram a predição de número de mortos devido à Covid-19 utilizando regressão linear em dados da Índia para um intervalo de seis semanas. Prata, Rodrigues e Bremejo (2020) utilizaram regressão polinomial linear para modelar os dados de 27 capitais brasileiras provenientes do Ministério da Saúde, e prever casos confirmados de COVID-19 no Brasil, além de analisarem a relação entre temperatura e os casos.

Já Santiago et al. (2020) buscaram ajustar diversos modelos de regressão não linear com dados de mortalidade e casos confirmados da COVID-19 no Brasil, Itália e no mundo. Os modelos estudados pelos pesquisadores foram adequados para modelar os números de casos confirmados e de mortes nos países selecionados, assim como mundialmente, porém não puderam prever quando os casos de COVID-19 diminuirão.

De acordo com Yang, Zhang, Peng, Zhuge e Hong (2020), as funções sigmóides: Hill, Logistic, Gompertz e Richards são mais adequadas para a previsão de epidemias, enquanto as funções lineares, quadráticas, cúbicas e exponenciais não podem capturar adequadamente a curva típica em forma de S. No entanto, ainda existem poucos estudos comparando modelos não lineares para ajustes de curvas para os casos de óbitos.

O crescimento de mortes no estado de Pernambuco e Ceará pode ser analisado de uma forma prática e eficiente por meio do estudo de curvas de crescimento, as quais descrevem uma relação funcional entre mortes e tempo em dias.

Assim, realizou-se este trabalho com objetivo de ajustar diferentes modelos matemáticos de regressão não linear a dados de óbitos por COVID-19 nos estados de Pernambuco-PE e Ceará-CE no período dos dados observados e fazer extrapolações do

número de óbitos por meio de previsões.

## 2. Metodologia

### *Dados*

Neste trabalho foram utilizados duas bases de dados do número acumulado de mortes por COVID-19 no estado de Pernambuco de vinte e cinco de março de dois mil e vinte até onze de maio de dois mil e vinte e no estado do Ceará de vinte e seis de março de dois mil e vinte até onze de maio de dois mil e vinte. Os dados analisados neste trabalho foram obtidos no site do Ministério da Saúde do Brasil (MS, 2020a), instituição que publica diariamente estatísticas sobre a pandemia de COVID-19.

### *Modelos Utilizados*

Recentemente, Santos et al. (2018) apresentaram alguns modelos como subcasos do que eles denominaram de método gerador de modelos de crescimento e/ou decrescimento, obtidos a partir de equações diferenciais. Assim, foram ajustados os modelos apresentados na Tabela 1 à variável número de óbitos por COVID-19 no estado de Pernambuco e Ceará.

**Tabela 1.** Descrição dos modelos estatísticos avaliados.<sup>(1)</sup>

Modelos	Parâmetros	Equações
Gompertz (Laird, 1965)	3	$W(t) = \alpha e^{-\beta e^{-kt}} + \varepsilon$
Logístico (Nelder, 1961)	3	$W(t) = \alpha \{1 + \beta e^{-kt}\}^{-1} + \varepsilon$
Von Bertalanffy (Von Bertalanffy, 1957)	3	$W(t) = \alpha \{1 - \beta e^{-kt}\}^3 + \varepsilon$
Santos (Santos et al., 2018)	4	$W(t) = \alpha \left(1 + \beta e^{\gamma e^{kt}}\right) + \varepsilon$

<sup>(1)</sup> Na Tabela 1 tem-se que:  $W(t)$  o número de óbitos confirmados da COVID-19 no tempo  $t$ ;  $t$  é tempo;  $k$  indica a velocidade de crescimento da curva para atingir o número de mortes assintótico;  $\alpha$  é o número de mortos por COVID-19 quando  $t$  tende a infinito;  $\beta$  e  $\gamma$  são parâmetros de forma;  $e$  é a exponencial; e  $\varepsilon$  é o erro aleatório associado seguindo distribuição normal com média zero e variância constante.

### ***Estimação dos parâmetros dos modelos não-lineares***

Em seguida, as estimativas dos parâmetros de crescimento dos diferentes modelos não lineares foram obtidas por meio do método de mínimos quadrados, usando o processo iterativo de Gauss Newton por meio da função Nonlinear Least Squares (nls) do pacote *Stats*, da ferramenta computacional R. Para realização das análises estatísticas utilizou-se o software gratuito, R versão 3.5.1 (R DEVELOPMENT CORE TEAM, 2018) e para análises gráficas e de previsão foi utilizada a linguagem de programação Python 3.6 utilizando a biblioteca *matplotlib* no ambiente do *Google Colab*. Foram criados algoritmos próprios para a previsão.

### ***Avaliadores da qualidade de ajuste***

Posteriormente, os critérios adotados para verificar a qualidade do ajuste dos modelos matemáticos foram: coeficiente de determinação ajustado ( $R_{aj}^2$ ), quadrado médio do resíduo (QMR) e critério de informação de Akaike (AIC) conforme Tabela 2.

**Tabela 2.** Descrição matemática dos critérios de seleção.

Equações
$R_{aj}^2 = R^2 - (p - 1/n - p)(1 - R^2)$
$QMR = \sum_{i=1}^n  y_i - \hat{y}_i /n$
$AIC = n + n \cdot \log(2\pi) + n \cdot \log(SQ_{res}/n) + 2(p + 1)$

Os termos na Tabela 2, são descritos da seguinte maneira, a saber:  $SQ_{res}$  é a soma de quadrados dos resíduos definida por  $SQ_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ;  $y_i$  é o número de óbitos observado;  $\hat{y}_i$  o número de óbitos estimado (previsão) de  $y_i$ ;  $n$  o número de observações e  $p$  o número de parâmetros livres do modelo. Assim, o melhor modelo ajustado é aquele que apresentar o maior valor para  $R_{aj}^2$ , e os menores valores para QMR, DMA, AIC e BIC (Santos et. al., 2019).



### 3. Resultados e Discussão

As estimativas dos parâmetros, os critérios utilizados e intervalos de confiança para cada parâmetro dos quatro modelos avaliados são apresentados na Tabela 3. Por meio desta tabela, é possível perceber que os modelos Santos e Logístico foram os que apresentaram menores amplitudes no intervalo para os parâmetros estimados.

O parâmetro  $k$ , representa a taxa de declínio na taxa de crescimento relativo, *i.e.*, quanto maior o valor de  $k$ , menor é o crescimento do número de mortes. Na estimativa do parâmetro  $k$ , pôde-se perceber que a estimativa do modelo Logístico (0,11) foi superior aos modelos Santos (0,08), Gompertz (0,03) e Von Bertalanffy (0,007). Para o estado do Ceará, a estimativa do modelo Logístico (0,11) foi superior aos modelos Santos (0,09) e Gompertz (0,02). Assim, é possível inferir que os modelos com menor taxa de declínio, está associado ao maior número de mortes.

Outro parâmetro importante é o  $\alpha$ , que representa o número de pessoas mortas por COVID-19 no limite assintótico, ou seja, não é número máximo de mortes, e sim o número de óbitos médio quando a pandemia atingir a maturidade livre das variações sazonais. Quando se comparam às estimativas dos óbitos ( $\alpha$ ) obtidos pelos quatros modelos para o estado de Pernambuco, pode-se verificar que as estimativas do parâmetro  $\alpha$ , foram distintas apresentando-se maior discrepância no modelo de Von Bertalanffy e Gompertz em relação aos outros dois modelos, ou seja, superestimando o número de óbitos predito. O modelo Von Bertalanffy emprega um número de óbitos assintótico de 48340 seguido pelo modelo Gompertz (4391), Logístico (1624) e Santos (1612,63). Porém, quando o modelo Von Bertalanffy é efetivamente empregado para prever o número de óbitos aos 48 dias, prediz o número de óbitos com maior precisão em torno de 1080 óbitos, seguido do modelo Santos (1076), Gompertz (1071) e Logístico (1056).

Já para o estado do Ceará, o número de mortes predito pelo modelo Gompertz emprega um número de óbitos assintótico de (43200), indicando que superestimou o número de óbitos, seguido pelos modelos Logístico (2675) e Santos (2196). No entanto, quando os modelos são efetivamente empregados para calcular o número de óbitos aos 47 dias, em 11/05/2020, com o valor observado de 1189 óbitos. O modelo Logístico prediz o número de óbitos em torno de 1191, sendo este mais preciso, seguido do modelo Santos (1196) e Gompertz (1204). A curva do Von Bertalanffy não obteve convergência com os dados do



Ceará.

Com relação à qualidade do ajuste do modelo, que se baseou no  $R_{aj}^2$ , todos os modelos são adequados para inferir o número de óbitos em estudo no intervalo de dias dos dados observados, com o coeficiente de determinação ajustado superior a 99,59% Figuras 1 e 2. Porém, foi observado na Tabela 3 que o melhor ajuste foi do modelo Von Bertalanffy para o estado de Pernambuco seguido dos modelos Gompertz, Santos e Logístico. Para o estado do Ceará, o modelo mais ajustado para prever o número de óbitos foi o de Santos, seguido do modelo Gompertz e Logístico para o intervalo de dados observados.

Assim, o modelo Von Bertalanffy e Santos foram os melhores modelos para o estado de Pernambuco e Ceará, respectivamente, por apresentarem os menores valores para AIC e QMR e os maiores valores para o  $R_{aj}^2$ . Porém, nenhum modelo é confiável, no estágio inicial de uma epidemia, quando os dados são limitados.

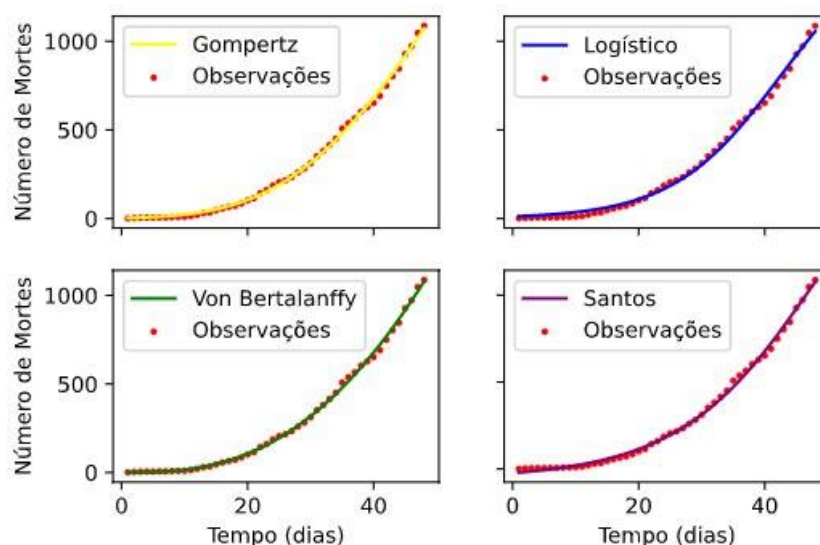
Portanto, não há como saber qual modelo é superior ao outro. Os modelos podem subestimar ou superestimar a situação real de uma epidemia de maneira imprevisível (Yang et al., 2020).

**Tabela 3.** Estimativas dos parâmetros, intervalo de confiança ( $IC_{95\%}$ ) critérios de seleção e modelos ajustados aos dados de óbitos por COVID-19 no estado de PE e Ceará.

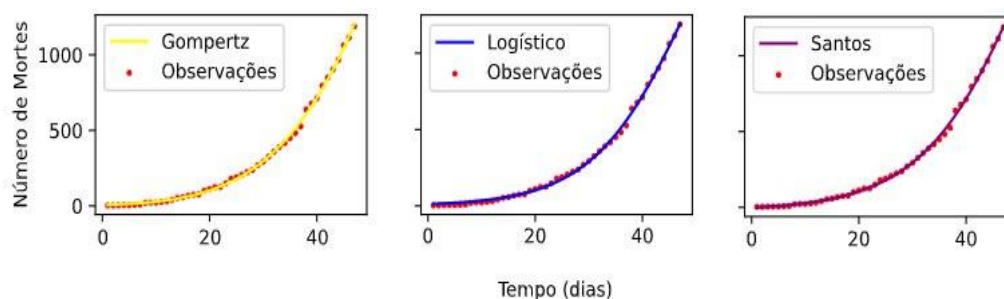
Modelos	Parâmetros		Critérios			
	$\alpha$	$k$	$R_{aj}^2$	AIC	QMR	
PE	Santos	1612,63* IC (1384,25;2084,43)	0,08* IC (0,07;0,09)	99,77%	406,42	246,65
	Logístico	1624,46* IC (1458,45;1857,49)	0,11* IC (0,10;0,12)	99,60%	430,94	418,99
	Gompertz	4391* IC (3507,54;5785,01)	0,03* IC (0,03;0,04)	99,81%	390,04	178,73
	Von Bertalanffy	48340* IC (22491,88;226568,10)	0,007* IC (0,004;0,01)	99,87%	375,37	131,66
CE	Santos	2196* IC (1839,35;2922,12)	0,09* IC (0,08;0,1)	99,89%	367,13	127,69
	Logístico	2675* IC (2347,10;3145,42)	0,11* IC (0,10;0,12)	99,85%	379,96	171,09
	Gompertz	43200* IC (22181,01;109793,5)	0,02* IC (0,01;0,02)	99,86%	374,03	150,78

\*Significativo a 95%

**Figura 1.** Estimativa do número de mortos em função dos dias, ajustado pelos modelos Santos, Von Bertalanffy, Gompertz e Logístico para o estado e Pernambuco.



**Figura 2.** Estimativa do número de mortos em função dos dias, ajustado pelos modelos Santos, Gompertz e Logístico para o estado do Ceará.

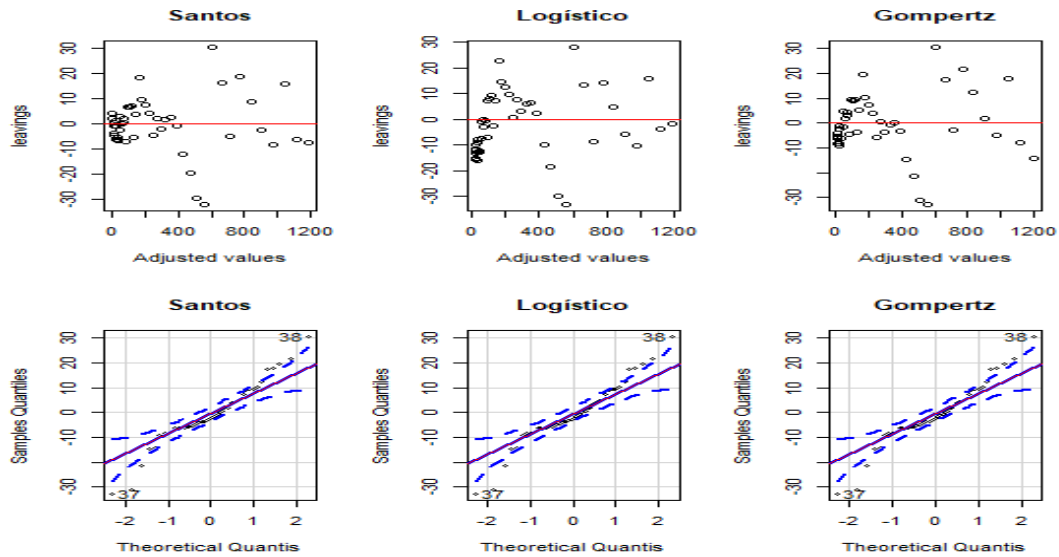


É fundamental a constatação das pressuposições para os modelos de regressão não lineares, pois, uma vez que não forem atendidas, considera-se que o modelo não é adequado e este desvio deve ser considerado ou revisto no modelo (Santos et al., 2019). Deste modo, além de checar a adequação do ajuste pelas Figuras 1 e 2, é importante fazer uma análise dos resíduos para verificar os pressupostos do modelo.

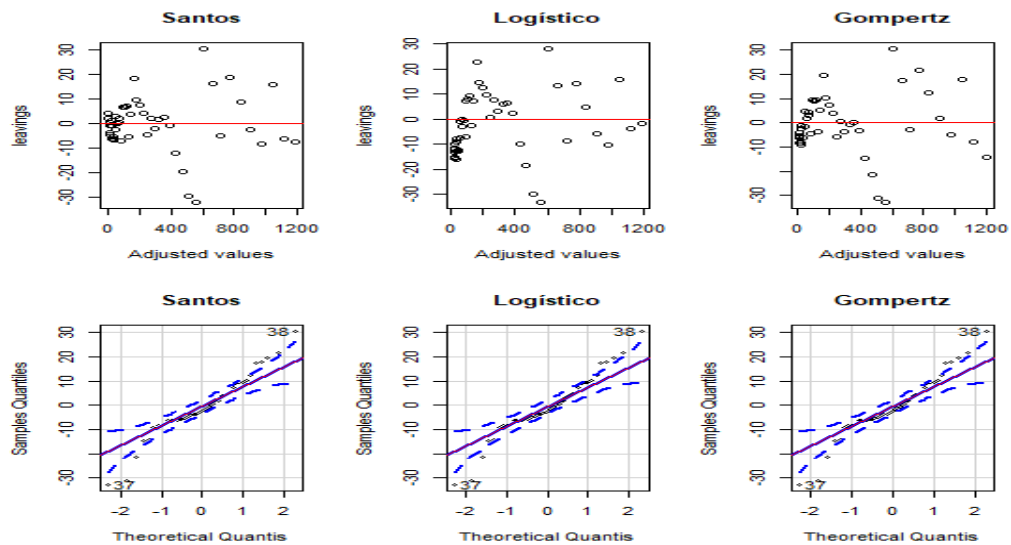
Para avaliar a qualidade do ajuste pela análise de resíduos, utilizou-se o diagrama de dispersão dos resíduos em função dos valores ajustados e o gráfico normal de probabilidade quantil-quantil com envelope dos resíduos apresentados na Figura 2 para o estado de Pernambuco – PE e Figura 3 para o Ceará – CE para os resíduos projetados do ajuste dos modelos aos dados de óbitos. Pelos gráficos de diagnóstico dos resíduos, não se tem motivos

para negar que os pressupostos do modelo foram atendidos, ou seja, indicam evidências de bons ajustes dos modelos.

**Figura 3.** Diagrama de dispersão e normalidade dos modelos estatísticos por meio dos resíduos para o número de óbitos do estado de Pernambuco - PE.



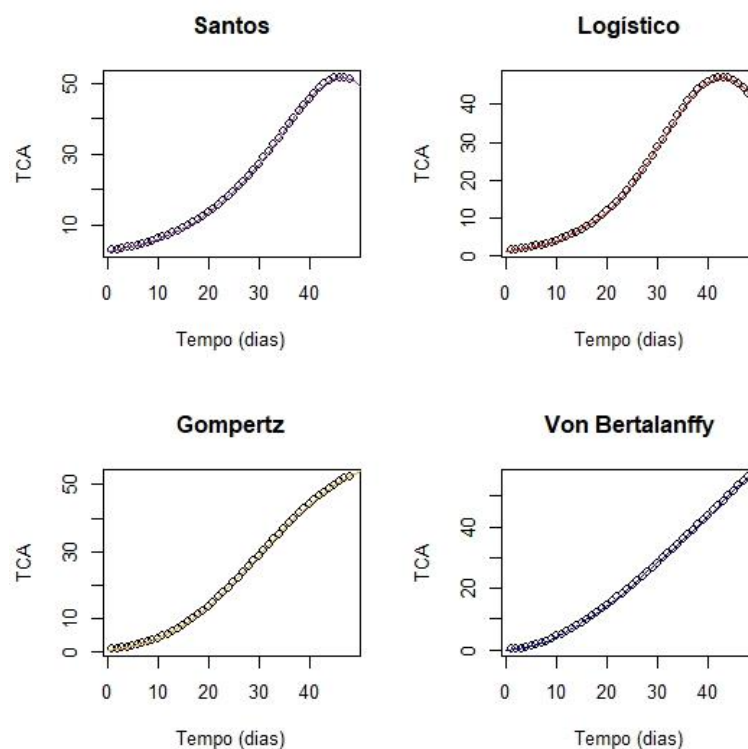
**Figura 4.** Diagrama de dispersão e normalidade dos modelos estatísticos por meio dos resíduos para o número de óbitos do estado de Pernambuco - PE.



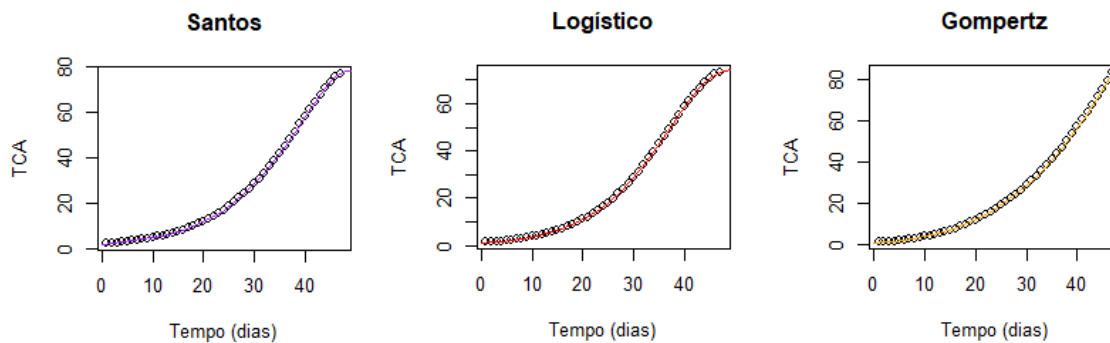
A taxa de crescimento absoluto (TCA) obtida pela primeira, derivada dos modelos em função do tempo, é apresentada nas Figura 5 e 6. Essas figuras apresentam a TCA do total de óbitos do estado de PE e (CE) para os dados observados de vinte e cinco de março de dois mil e vinte a onze de maio de dois mil e vinte e para vinte e seis de março de dois mil e vinte a (onze de maio de dois mil e vinte), respectivamente.

A evolução da taxa do modelo de Santos indica um decrescimento a partir do quadragésimo sétimo dia da pandemia e o modelo Logístico no quadragésimo terceiro dia. Já a taxa dos modelos Gompertz e Von Bertalanffy ainda estão em pleno crescimento durante todos os dias observados para o estado de Pernambuco. Para o estado do Ceará, a evolução da taxa do modelo Santos e logístico apresenta sinais de decrescimento a partir do quadragésimo sétimo dia da pandemia, enquanto a taxa do modelo Gompertz continua a crescer. Isso leva a crer que os estados analisados ainda estão em um estágio inicial do surto, *i.e.*, o número de casos ainda não está atingindo o ponto de inflexão.

**Figura 5.** Taxa de crescimento absoluto do número de mortes total do estado de Pernambuco devido a COVID-19.



**Figura 6.** Taxa de crescimento absoluto do número de mortes total do estado de Ceará devido a COVID-19.



Em seguida, calculou-se a previsão dos modelos para o dia seguinte e para um horizonte maior de dias. A Tabela 4 apresenta os resultados para o estado e Pernambuco.

Pode-se evidenciar a partir da Tabela 4 que os modelos predizem valores próximos para os primeiros dias do horizonte de previsão, e que essa diferença vai aumentando ao decorrer do tempo. A partir do décimo dia de previsão, existe uma diferença de 332 pessoas mortas, entre o menor valor previsto neste dia, pelo modelo logístico (1390), e o maior valor, para o modelo Von Bertalanffy (1722). Essa discrepância entre o menor valor previsto e o maior valor segue crescente até o último dia de previsão. Segundo Wu, Darcet, Wang e Sornette (2020), com mais dados, os cenários se tornam mais precisos, especialmente ao usar dados realizados após o pico.

Felix e Fontenele (2020) estudaram a evolução da epidemia COVID-19, no estado do Ceará, por meio de uma avaliação gráfica e chegaram à conclusão de que a epidemia ainda está longe de ser controlada no Ceará, sendo que existe uma pequena credibilidade em previsões a partir dos dados. Além do mais, estão enviesados ou imprecisos de alguma forma.

**Tabela 4.** Previsão do número de mortos por COVID-19 no estado de PE para um horizonte longo.

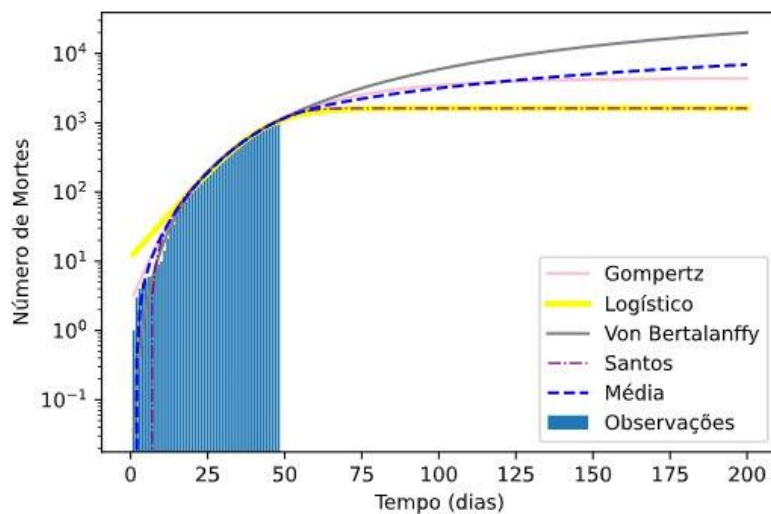
Data	Número total de óbitos previstos			
	Santos	Logística	Gompertz	Von Bertalanffy
12/05	1126	1098	1124	1136
13/05	1176	1138	1177	1195
14/05	1224	1177	1231	1255
15/05	1271	1213	1286	1317
16/05	1315	1248	1341	1381
21/05	1492	1390	1620	1722
26/05	1582	1484	1900	2102
31/05	1608	1542	2171	2518
02/07	1612	1622	3483	5924

Pode-se ver na Figura 7 que o modelo Santos e Logístico convergem para o limite assintótico mais rapidamente que os modelos de Gompertz e Von Bertalanffy, no horizonte de dias especificado vide Tabela 4. A linha tracejada representa o comportamento médio (CM) dos quatro modelos avaliados, que foi calculado por meio da Equação 1, com valores de  $i=1, 2, 3, \dots, 100$ :

$$CM[i] = \frac{(Logístico[i] + Santos[i] + Von\ Bertalanffy[i] + Gompertz[i])}{4} \quad (1)$$

O modelo Gompertz se ajusta melhor à CM que os outros três modelos. Ainda se percebe na Figura 7, que o modelo de Von Bertalanffy está longe de chegar próximo à estabilização. Segundo Santiago et al. (2020) espera-se que os erros de estimativas diminuam à medida que mais dados são obtidos e que traduzam os números mais claro do cenário da epidemia ao longo da evolução da doença.

**Figura 7.** Extrapolação do número total de óbitos por COVID para o estado de PE até 10/10/2020 em escala logarítmica.



Já para a Tabela 5, apresenta as previsões para o estado do Ceará. Assim, como na tabela anterior que, nos primeiros dias de previsão do número de mortos devido por COVID-19 dos três modelos possuem valores próximos, e só começa a divergir a partir do decimo dia e segue ampliando até o último dia de previsão, conforme a Figura 8.

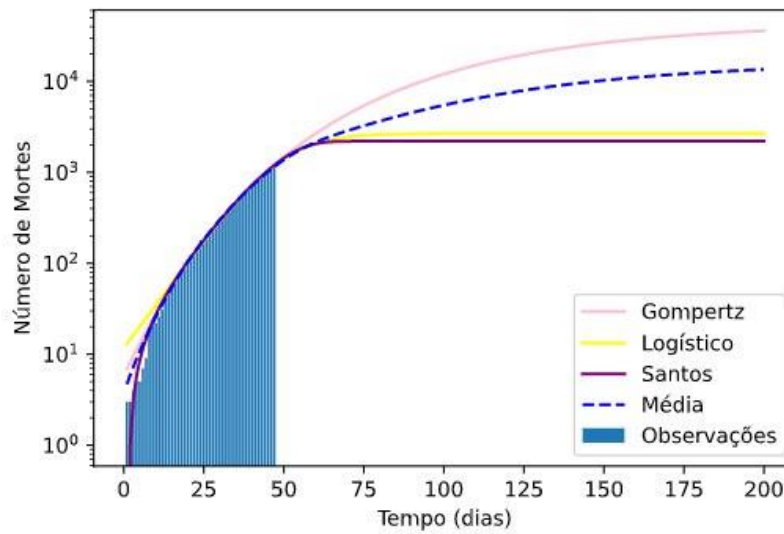


**Tabela 5.** Previsão do número de mortos por COVID-19 no estado de CE para um horizonte longo.

Data	Número de mortos previstos		
	Santos	Logística	Gompertz
12/05	1274	1264	1289
13/05	1352	1338	1379
14/05	1431	1412	1474
15/05	1510	1486	1572
16/05	1587	1559	1676
21/05	1923	1896	2264
26/05	2121	2164	2973
31/05	2186	2355	3808
03/07	2195	2665	12013

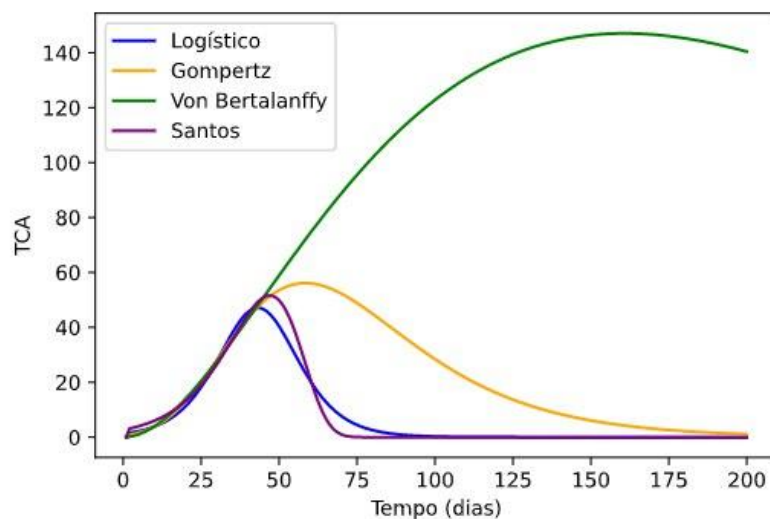
Na Figura 8, o modelo Gompertz tem um crescimento acelerado em relação aos demais modelos. O modelo Santos e Logístico possuem comportamento assintótico semelhante. Os modelos Gompertz, Santos e Logístico possuem um afastamento da média significativo a partir do dia 60. Wu et al. (2020) identificaram que os modelos do tipo logística sempre subestimam o número total de casos infectados no estágio inicial, enquanto o modelo Gompertz faz uma superestimação em todos os casos estudados.

**Figura 8.** Extrapolação do número total de óbitos para o estado de CE por COVID-19 até 11/10/2020 em escala logarítmica.



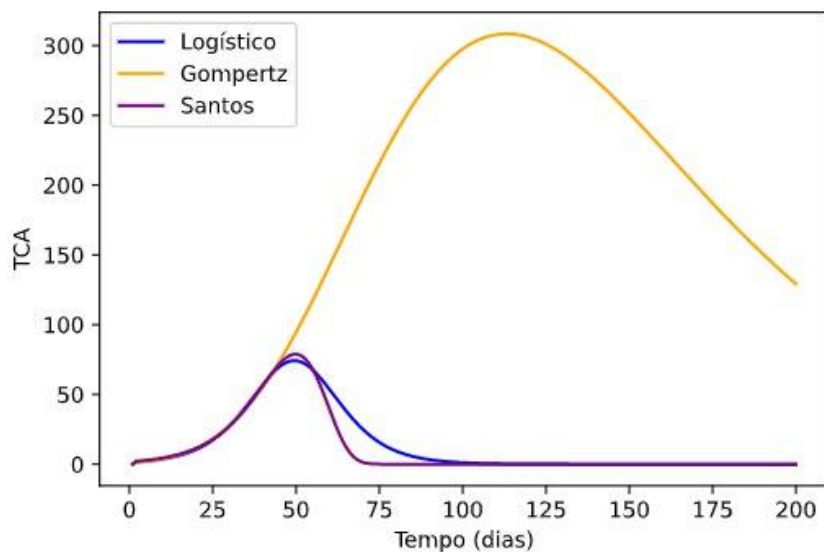
A Figura 9 representa uma extrapolação da TCA do número total de mortes por COVID-19 do estado de PE. O modelo Gompertz indica um decrescimento da taxa de óbitos a partir do dia 56, ou seja, dezoito de maio de dois mil e vinte. O modelo de Von Bertalanffy ainda apresenta um crescimento na taxa no período analisado, atingindo um pico no dia 150, em vinte e um de agosto de dois mil e vinte. Entretanto o modelo Santos e Logístico indicaram um decrescimento na faixa de dados observados analisados vide Figura 5.

**Figura 9.** TCA para o número de mortes extrapoladas para PE.



A Figura 10 apresenta a simulação da TCA do Ceará. A TCA do modelo logístico e Santos apresenta um decréscimo a partir do dia 47 e 48, ou seja, 11/05/2020 e 12/05/2020, respectivamente. A taxa de crescimento do modelo Gompertz continua a crescer no intervalo considerado, atingindo um pico em 111 dias após o início da pandemia, em 14/07/2020.

**Figura 10.** TCA para o número de mortes extrapoladas para CE.



O ponto de inflexão da curva de número de óbitos é o ponto crítico da taxa de mortes, ou seja, quando a derivada da taxa de mortes é igual a zero. Foi calculado o ponto de inflexão para cada modelo, por análise numérica Tabela 6.

**Tabela 6.** Pontos de Inflexão para os diferentes modelos para o estado de Pernambuco e Ceará.

Estado	Modelo	Ponto de Inflexão
PE	Logístico	04/05/2020
	Santos	08/05/2020
	Gompertz	19/05/2020
	Von Bertalanffy	21/08/2020
CE	Logístico	11/05/2020
	Santos	12/05/2020
	Gompertz	14/07/2020

Os resultados encontrados estão resumidos na Tabela 6, corroborando com os achados discutidos anteriormente.

#### 4. Considerações Finais

Para o estado de Pernambuco, no intervalo de dados observados, todos os modelos estudados foram bem ajustados conforme  $R_{aj}^2$ , com prevalência do modelo de Von Bertalanffy que obteve o melhor ajuste por meio dos critérios adotados. Todavia, este superestimou o parâmetro  $\alpha$  de mortes no limite assintótico, prevendo um número exagerado de óbitos em um horizonte de tempo longo.

Já os modelos Santos e Logístico subestimaram esse mesmo parâmetro, por apresentarem um número de mortes do limite assintótico baixo em relação ao número total de óbitos no último dia de observação onze de maio de dois mil e vinte, indicando que poucas mortes irão ocorrer em um limite de 100 dias, o que não aparenta ser uma tendência dos últimos dias observados. O modelo Gompertz apresentou um comportamento próximo à média dos modelos estudados para o estado de Pernambuco.

Em relação ao estado do Ceará, o modelo de Von Bertalanffy não conseguiu se ajustar aos dados. Os três modelos restantes obtiveram bons ajustes segundo o  $R_{aj}^2$ , porém o modelo de Gompertz superestimou o parâmetro  $\alpha$  de número de mortes no limite assintótico, indicando parecer ser inadequado para descrever o número de óbitos por COVID-19, especialmente na fase final de estabilidade. O modelo Santos e Logístico apresentaram uma tendência semelhante, porém subestimando o valor do parâmetro  $\alpha$ .

O modelo Logístico e Santos apresentaram pontos de inflexão dentro do intervalo de dados observados. Já os modelos Gompertz e Von Bertalanffy dentro do período de dados extrapolados, o que indica que o crescimento da taxa do número de óbitos por COVID-19 no estado de PE irá diminuir nesse intervalo de tempo, todavia ainda necessita de mais estudos sobre o tema.

Já para o estado do Ceará, o modelo Logístico apresentou o ponto de inflexão no período de dados observados, enquanto que o modelo Santos e Gompertz foram no período dos dados extrapolados, o que sugere que o crescimento da taxa do número de óbitos por COVID-19, no estado de CE, irá diminuir nesse intervalo de tempo, embora ainda exija mais evidências.

As previsões dos dez primeiros dias após a última observação podem apresentar resultados mais confiáveis em relação à previsão do número total de mortos nesses respectivos dias. Horizontes mais longos, possuem uma incerteza maior para o número total de mortos. Porém, esses resultados obtidos ainda não permitem fazer projeções seguras de quando os números de óbitos por COVID-19 no estado de Pernambuco e Ceará diminuirão.

No entanto, essas estimativas podem ajudar as autoridades a monitorar a epidemia e auxiliar na tomada de decisão sobre a implementação de estratégias de isolamento social para gerenciar a pandemia do COVID-19. Não se pode relaxar no confinamento. Parecem uma conclusão lógica deste cenário epidemiológico a continuidade, prevenção, e até mesmo o fortalecimento das medidas de distanciamento social.

## Referências

Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., Kobinger, G., Lane, H. C. F., Memish, Z., Oh, M., Sall, A. A., Schuchat, A., Ungchusak, K., & Lothar H Wieler.

(2020). COVID-19: towards controlling of a pandemic. *The Lancet*, 395(10229), 1015-1018.  
doi: 10.1016/s0140-6736(20)30673-5

Carneiro, G. (2020). Portal Folha PE. Recuperado em 12 de Março de 2020 de:  
<https://www.folhape.com.br/noticias/noticias/coronavirus/2020/03/12/NWS,133219,70,1668,NOTICIAS,2190-CASAL-RECIFE-SAO-PRIMEIROS-CASOS-CORONAVIRUS-PERNAMBUCO.aspx>

Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*, 134, 109761.  
<https://doi.org/10.1016/j.chaos.2020.109761>

Felix, F. H., & Fontenele, J. (2020). Avaliação gráfica simplificada da evolução da epidemia COVID-19 no estado do Ceará.

Ghosal, S., Sengupta, S., Majumder, M., & Sinha, B. (2020). Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020). *Diabetes & metabolic syndrome*, 14(4), 311–315. Advance online publication. <https://doi.org/10.1016/j.dsx.2020.03.017>

Laird, A. K. (1965). Dynamics of relative growth. *Growth*, 29(9), 249-263.

Ministério da Saúde [MS] (2020a). Painel Coronavírus. Recuperado em 12 de março de 2020 de: <https://covid.saude.gov.br/>

Ministério da Saúde [MS] (2020b). 15º boletim epidemiológico especial COE- COVID19. Recuperado em 12 de maio de 2020 de:  
<https://portalarquivos.saude.gov.br/images/pdf/2020/May/09/2020-05-06-BEE15-Boletim-do-COE.pdf>

Ministério da Saúde [MS] (2020c). Recuperado em 12 de maio de 2020 de:  
<https://www.saude.gov.br/images/pdf/2020/May/11/11.05.2020.Apresenta%C3%A7%C3%A3o%20Diretrizes%20de%20distanciamento%20FINAL.pdf>

Na Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., & Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*. 382, 727-733. doi: 10.1056/NEJMoa2001017

Nelder, J. A. (1961). The fitting of a generalization of the logistic curve. *Biometrics*, 17(1), 89-110. doi: <https://www.jstor.org/stable/2527498>

Prata, D. N., Rodrigues, W., & Bermejo, P. H. (2020). Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil. *Science of the Total Environment*, 729, 138862. <https://doi.org/10.1016/j.scitotenv.2020.138862>

Portal G1. (2020). Recuperado em 08 de Maio de 2020 de:  
<https://g1.globo.com/pe/peernambuco/noticia/2020/05/08/com-mais-82-mortes-e-763-casos-de-covid-19-pe-chega-a-927-obitos-e-11587-pacientes-com-o-novo-coronavirus.ghtml>

Rodríguez-Morales, A. J., MacGregor, K., Kanagarajah, S., Patel, D., & Schlagenhauf, P. (2020a). Going global - Travel and the 2019 novel coronavirus. *Travel medicine and infectious disease*, 33, 101578. <https://doi.org/10.1016/j.tmaid.2020.101578>

Rodríguez-Morales, A. J., Gallego, V., Escalera-Antezana, J. P., Méndez, C. A., Zambrano, L. I., Franco-Paredes, C., Suárez, J. A., Rodriguez-Enciso, H. D., Balbin-Ramon, G. J., Savio-Larriera, E., Risquez, A., & Cimerman, S. (2020b). COVID-19 in Latin America: The implications of the first confirmed case in Brazil. *Travel medicine and infectious disease*, 101613. Advance online publication. <https://doi.org/10.1016/j.tmaid.2020.101613>

Santiago, E. J. P., da Silva Freire, A. K., Cunha Filho, M., Moreira, G. R., de Almeida Ferreira, D. S., & Cunha, A. L. X. (2020). Non-linear models applicable to mortality and cases of COVID-19 in Brazil, Italy and the world. *Research, Society and Development*, 9(6), 117963561. doi: <http://dx.doi.org/10.33448/rsd-v9i6.3561>

Santos, A., Moreira, G. R., Gomes-Silva, F., Brito, C., da Costa, M., Pereira, L., Maurício, R. M., Azevêdo, J., Pereira, J. M., Ferreira, A. L., & Filho, M. C. (2019). Generation of models



from existing models composition: An application to agrarian sciences. *PloS one*, 14(12), e0214778. <https://doi.org/10.1371/journal.pone.0214778>

Santos, A. L. P., Moreira, G. R., de Brito, C. C. R., Gomes-Silva, F., da Costa, M. L. L., Pimentel, P. G., Cunha Filho, M., & Mizubuti, I. Y. (2018). Method to generate growth and degrowth models obtained from differential equations applied to agrarian sciences. *Semina: Ciências Agrárias*, 39(6), 2659-2672. doi: <http://dx.doi.org/10.5433/1679-0359.2018v39n6p2659>

Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Losifidis, C., & Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*. 76, 71-76. <https://doi.org/10.1016/j.ijisu.2020.02.034>

Von Bertalanffy, L. (1957). Quantitative laws in metabolism and growth. *The quarterly review of biology*, 32(3), 217-231. doi: <https://doi.org/10.1086/401873>

World Health Organization, WHO Director-General's Remarks at the Media Briefing on 2019-nCoV on 11 February 2020, (2020). <https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020>

Wu, K., Darcet, D., Wang, Q., & Sornette, D. (2020). Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. arXiv preprint arXiv:2003.05681.

Yang, W., Zhang, D., Peng, L., Zhuge, C., & Hong, L. (2020). Rational evaluation of various epidemic models based on the COVID-19 data of China. *arXiv preprint arXiv:2003.05666*.

**Porcentagem de contribuição de cada autor no manuscrito**

André Luiz Pinto dos Santos – 20%

Marcela Portela Santos de Figueiredo – 20%

Tiago Alessandro Espínola Ferreira – 15%

Frank Gomes-Silva – 15%

Guilherme Rocha Moreira – 10%

José Eduardo da Silva – 10%

Jucarlos Rufino de Freitas – 10%