

**UNIVERSIDADE FEDERAL DE OURO PRETO
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO, ADMINISTRAÇÃO E
ECONOMIA – DEPRO**

Samuel Maia Salvador de Assis

**APLICAÇÃO DE MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS PARA
O NÚMERO DE MORTES POR CORONAVÍRUS NO BRASIL**

**Ouro Preto
2020**



MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE FEDERAL DE OURO PRETO
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO, ADMINISTRAÇÃO E ECONOMIA



FOLHA DE APROVAÇÃO

Samuel Maia Salvador de Assis

Aplicação de modelos de previsão em séries temporais para o número de mortes por Coronavírus no Brasil

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheiro de Produção

Aprovada em 07 de janeiro de 2022.

MEMBROS DA BANCA

Prof.º Dr.º **Magno Silvério Campos** - Orientador - Universidade Federal de Ouro Preto

Prof.ª Dr.ª **Maurinice Daniela Rodrigues** – Examinadora Convidada - Universidade Federal de Ouro Preto

Prof.ª **M.ª Samantha Rodrigues de Araújo** - Examinadora Convidada - Universidade Federal de Ouro Preto

Magno Silvério Campos, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOPem 11/01/2022.



Documento assinado eletronicamente por **Magno Silvério Campos**, PROFESSOR DE MAGISTERIO SUPERIOR, em 11/01/2022, às 11:17, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Samantha Rodrigues de Araújo**, PROFESSOR DE MAGISTERIO SUPERIOR, em 11/01/2022, às 11:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Maurinice Daniela Rodrigues**, PROFESSOR DE MAGISTERIO SUPERIOR, em 11/01/2022, às 13:23, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0265521** e o código CRC **B575CF32**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.000237/2022-00

SEI nº 0265521

R. Diogo de Vasconcelos, 122, -
Bairro Pilar Ouro Preto/MG, CEP
35400-000 Telefone: 3135591540
- www.ufop.br

https://sei.ufop.br/sei/controlador.php?acao=documento_imprimir_web&acao_origem=arvore_visualizar&id_documento=295112&infra_sistema=1... 1/1

Resumo

Modelos de previsão de dados são ferramentas de suma utilidade para inúmeras áreas de estudo, sendo uma delas a gestão pública de saúde. No ano de 2020 houve o início da pandemia causada pelo coronavírus(Covid-19), em que sociedade entrou em estado de emergência e concentrou recursos para combater essa triste realidade que continua a assombrando até hoje. Técnicas estatísticas de modo de previsão também foram utilizadas amplamente para acompanhar o andamento da doença, como a sua disseminação e o comportamento dos dados relacionados as mortes causadas pela doença. O modelo mais utilizado e divulgado pelos veículos midiáticos foi o modelo de previsão baseado em Médias Móveis, que é um entre os diversos modelos estatísticos que podemos utilizar para modelar estes dados. Este trabalho realizará a modelagem dos dados de mortes causadas pelo coronavírus utilizando o método de Holt-Winters e também o de Médias Móveis , onde será analisado o resultado e a eficiência de cada um, segundo as métricas estatísticas, qual se mostra mais eficiente. A conclusão foi que o Modelo de Holt-Winters, segundo a métrica de comparação dos Diferença Mínima dos Quadrados(DMQ), se demonstrou mais eficiente em relação ao de Médias Móveis.

Palavras-chave: Coronavírus. Séries Temporais. Médias Móveis. Holt-Winter. Gestão Pública. Modelos de Previsão.

Abstract

Forecast models are extremely useful tools for numerous areas of study, one huge area that use it a lot is the public health management. The year 2020 was marked by the beginning of the coronavirus pandemic (Covid-19), where the whole world entered a state of emergency and concentrated resources to combat this sad reality that continues to haunt us today. For the coronavirus pandemic, statistical techniques of predictive mode were also widely used to monitor the progress of the disease, such as its spread and the behavior of data related to deaths caused by the disease. The most used and disseminated model by media vehicles was the forecast model based on Moving Averages, which is one of the various statistical models that we can use to model these kind of data. This work will carry out the modeling of data on deaths caused by the coronavirus using the Holt-Winters method and also the Moving Averages method, where we will compare the result and efficiency of each one and analyze it, according to the metrics statistics, which is more efficient. The conclusion was that the Holt-Winters Model, according to the efficiency metrics, showed us less errors based on the MSD(Minimum Squares Difference) indicator when compared with de Moving Average model.

Key words: Coronavirus. Time Series. Moving Average. Holt-Winter. Public Administration. Forecast Models.

SUMÁRIO

1. INTRODUÇÃO	7
Formulação do Problema.....	8
Objetivos.....	8
Justificativa e relevância.....	9
2. REVISÃO DE LITERATURA	10
Modelos de Previsão de Séries Temporais	10
Modelo de previsão utilizando médias móveis simples	11
Suavização exponencial sazonal de Holt-Winters.....	11
Média quadrática dos erros (MSD)	12
3. METODOLOGIA	12
4. RESULTADOS E DISCUSSÕES	13
Modelo de médias móveis	13
Modelo de Suavização exponencial sazonal de Holt-Winters.....	16
CONCLUSÕES FINAIS	21
5. REFERÊNCIA BIBLIOGRÁFICA	22

LISTA DE FIGURAS

Figura 1 – Gráficos dos Resultados modelo de Médias Móveis	14
Figura 2 - Gráficos dos resultados dos resíduos do modelo de Médias Móveis	15
Figura 3 - Tabela de Programação do Excel.....	16
Figura 4 - Gráfico dos resultados do modelo de Holt-Winters.....	19
Figura 5 - Resultados dos gráficos de resíduos do modelo de Holt-Winters	20

1. INTRODUÇÃO

O ano de 2020 começou com o surgimento do SARS-CoV2, conhecido popularmente como coronavírus. A aparição deste vírus foi um grande marco para a história, uma vez que suas características possibilitaram uma fácil e altamente rápida contaminação em todo mundo.

O vírus foi descoberto na China e com menos de dois meses ele já espalhou por todo planeta de forma repentina, contaminando dezenas de milhões de pessoas ao redor do mundo e também sendo o responsável pela morte de muitas dessas pessoas. De acordo com a Organização Pan Americana de Saúde (OPAS) e a Organização Mundial de Saúde (OMS), a partir de março de 2021 foi classificada com uma pandemia, onde, até 10 de Novembro de 2021 houveram 259.502.031 casos confirmados e 5.183.003 mortes causas pelo coronavírus.

Devido a esta realidade, países do mundo inteiro se encontraram em estado de alerta. A OMS declarou estado de emergência, avisou todos os países para que concentrassem esforços e recursos da gestão pública e de saúde para cessar essa grande contaminação e tentar diminuir ao máximo as possíveis mortes causadas por ele. Foi visto, em todo o mundo, sistemas de saúde pública como hospitais, postos médicos e afins sendo completamente ocupados por pessoas com sintomas causados pelo vírus além do surgimento de centenas de milhares de mortes causadas por ele. Segundo a Organização Mundial de Saúde (OMS), o objetivo geral do Plano Estratégico de Resposta ao coronavírus é justamente entender a situação vivenciada e lutar contra a disseminação do vírus, suprimindo sua transmissão e prevenindo a doença e a mortalidade associada.

Especificamente no Brasil, a primeira morte causada pelo vírus foi em março de 2020, a partir daí, infelizmente, esse número se expandiu agressivamente ao correr dos dias. Diante dessa situação, assim como no resto do mundo, o coronavírus se tornou o assunto mais falado e discutido na sociedade brasileira e suas instituições, onde governantes, para controlar a contaminação da população, tiveram de tomar decisões importantes sobre o sistema de saúde público e particular. Além disso toda mídia tinha o coronavírus como assunto principal, entregando à população informações sobre as situações dos hospitais, o número de contaminados em todo país, o número de mortes e outros tipos de informações sobre a pandemia.

Estes grandes vetores de comunicação e orientação sobre o vírus, a mídia, o governo e todos seus canais de comunicação, desde o início da pandemia até nos últimos dias buscam informações para trazer ao público sobre as medidas de proteção contra o vírus, os dados sobre contaminações e mortes causadas por ele. Sobre esses últimos dados, a mídia utilizou amplamente para divulga-los um modelo de previsão de dados chamado Médias Móveis para

comunicar a população como está o andamento e a tendência de contaminações e mortes em nosso país, em cada estado e em cada cidade.

Por sua simplicidade, o modelo utilizado para fazer o cálculo foi o mais utilizado para realizar projeções sobre as estatísticas relacionadas ao vírus e, conseqüentemente, mais divulgados para a população brasileira, servindo, durante todo este tempo, como a maior forma de entendimento da população sobre como está o andamento real da situação. Entretanto, este é apenas um dos modelos que podemos usar para prever estes dados, onde existem vários outros modelos que podem ser usados para descrever os dados de variados tipos.

Essas informações não só foram utilizadas para servir de informação para a população no geral, mas também serviu de informação para vários gestores de saúde e gestores públicos tomarem decisões em cima destes dados, vemos a importância de, quando mais assertivo este modelo for, mais assertivas serão também as decisões tomadas em cima destas previsões, e estamos falando sobre decisões que impactarão diretamente a vida das pessoas, podendo prevenir, inclusive, a morte da população.

O tema desse trabalho será focado nesses modelos de previsão, será realizado um estudo com os dados relacionados as mortes pelo coronavírus no Brasil e será feito outro modelo de previsão de séries temporais, analisar seus resultados e comparar a eficiência dos mesmos em relação ao método utilizado pela maioria das vias midiáticas.

Formulação do problema

Analisando estes fatores, o problema do estudo consiste em: o método que está sendo mais utilizado para modelar esses dados, chamado “Médias Móveis” é realmente o melhor método de previsão, estatisticamente falando? Existem outros que possuem uma eficiência maior e que podem melhorar as tomadas de decisão dos gestores de saúde/públicos, além de melhorar a qualidade da comunicação feita pelos veículos midiáticos?

Objetivos

Neste, trabalho, temos como objetivo geral: Aplicar e comparar os métodos de Suavização Exponencial de Holt-Winter e o método de Médias Móveis para os dados de mortes por covid-19 desde o início destes dados no Brasil, ou seja, desde março de 2020 a 20/04/2021.

Para alcançá-lo temos os seguintes objetivos específicos:

- Coletar todos os dados referentes às mortes causadas pelo vírus no Brasil, desde o início da pandemia.

- Realizar um estudo estatístico com outro modelo de previsão além do utilizado hoje pela maioria dos veículos midiáticos.
- Comparar todos os resultados e ver qual tipo de modelo de previsão é o que mais se adequa aos dados propostos.

Justificativa e relevância

A pandemia causada pelo coronavírus está nos primeiros lugares dentre as pandemias que marcaram a história da nossa sociedade, se tornando um assunto de extrema importância para a mesma.

Nos tempos de hoje, temos diversas ferramentas de várias áreas de atuação que nos ajudam de alguma forma a combater um vírus tão devastador como o Covid-19. A estatística é uma delas, uma vez que os conceitos e metodologias dessa área conseguem nos dar visões e previsões do comportamento dos dados relacionados a doença e assim nos guiar para boas tomadas de decisão.

Esse estudo surge para contribuir diretamente para que avaliar essas possíveis análises sobre o comportamento dos dados sobre as mortes das pessoas e servirá como ferramenta para melhores tomadas de decisão. Uma vez que, não é só uma ferramenta de comunicação para a sociedade (de como está/será o comportamento destes dados), mas também servirá de base para todos os níveis de gestão pública tomarem ações assertivas e eficientes no controle da pandemia.

No âmbito da universidade, esse estudo também pode servir como base a comunidade acadêmica também avaliar esse trabalho e tirar suas conclusões sobre outras possibilidades de modelagem estatística quando o assunto é o comportamento dos dados relacionados a uma pandemia, mostrando, então, outras possibilidades e formas de modelagem. Ademais, este estudo pode servir como inspiração para outros estudos de aprofundamento na área, testar outros modelos de previsão, utilizar como base para dissertar sobre a gestão pública e suas ações, dentre outras áreas.

Além disso, estatística, modelagem e tomada de decisão são três palavras que, ao perceber que poderia utilizar ferramentas dentro deste campo de atuação para realizar um trabalho com um assunto tão importante e crucial nos nossos dias, foi, de fato, um grande motivador, tanto entregar isso para a sociedade, quanto aperfeiçoar meus conhecimentos dentro da estatística.

2. REVISÃO DE LITERATURA

Modelos de Previsão de Séries Temporais

Segundo Montgomery(2015) e Farias(2008), séries temporais são séries de dados que se orientam em uma linha cronológica de observações dos respectivos dados. Ou seja, levam essa denominação quando a sequência de dados corre ao longo do tempo.

A análise de séries temporais possui 4 objetivos principais (MORRETIN; TOLOI, 2004):

- Procurar saber quais são as causas geradoras da série temporal;
- Realizar previsões para valores futuros da série, sendo essas previsões tanto de curto ou longo prazo;
- Identificar e modelar o comportamento da série temporal;
- Procurar um nível de periodicidade na série;

Objetivos os quais vão variar dependendo do que se quer alcançar com o estudo a ser feito. Segundo Montgomery(2015), modelos de previsão de séries temporais são de extrema importância para várias áreas, sendo um fator crítico para tomadas de decisão que envolvam o futuro. Além disso, segundo Fialho e Alves(2019), a matemática, com as áreas de modelagem matemática e a previsão estatísticas são enormes aliados da gestão pública de saúde, servindo, assim, como uma ferramenta para nos dar maior confiança e assertividade nas tomadas de decisão. Para combater o coronavírus, foram elaborados modelos de previsão de dados ajudando a interpretar e a conhecer melhor o comportamento da doença.

Os modelos de previsão de séries temporais são descritos da seguinte forma:

$$\hat{y}_{t+1} = f(y_t, y_{t-1}, y_{t-2}, y_{t-3}, \dots)$$

Onde,

- \hat{y}_{t+1} seria o valor ajustado pela previsão no período de tempo t+1;
- y_t é o valor da variável no tempo t;
- y_{t-1} é o valor da variável no tempo t-1;

Além disso, um conceito muito importante a ser comentado são os erros dos modelos de previsão, que, segundo Montgomery(2015), podem ser representados como:

$$e_t = y_t - \hat{y}_t$$

Onde,

- e_t é o erro do modelo naquele determinado período de tempo t;
- y_t é o valor da variável no tempo t;
- \hat{y}_t é o valor ajustado pelo modelo no tempo t.

Um outro ponto muito importante a ser abordado é sobre as classificações de comportamentos de uma série temporal, podendo ser analisados segundo sua estacionariedade e sua sazonalidade.

Segundo Morretin e Toloi(2006, p.16), uma série estacionária “se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo de alguma forma de equilíbrio estável”. As séries que não apresentam este tipo de comportamento, são séries que apresentam tendências, ou seja, os dados seguem um padrão de crescimento/decrescimento ao longo do tempo.

Sobre sazonalidade de uma série, segundo Montgomery(2015), é quando percebe-se um comportamento cíclico dos dados analisados. Consequentemente, se uma série temporal ela se comporta dessa forma, podemos falar que ela possui sazonalidade, de modo contrário, não.

Modelo de previsão utilizando médias móveis simples

Conforme Morretin e Toloi(2006, p.32), “A técnica de média móvel consiste em calcular a média aritmética das r observações mais recentes”. Logo, a previsão de todos valores futuros é dado pela última média móvel calculada. Podemos escrever a equação desta forma:

$$\hat{y}_t = \frac{y_t + y_{t-1} + \dots + y_{t-r+1}}{r}$$

Ainda conforme os autores citados acima, o método das médias móveis tem as seguintes vantagens:

- Simples aplicação;
- Pode ser usado com um número pequeno de dados;
- Permite uma flexibilidade quanto a se determinar o valor de r.

E as seguintes desvantagens:

- Dificuldade em determinar o valor de r.

Suavização exponencial sazonal de Holt-Winters

Conforme Montgomery(2015), o método de Holt-Winters é um método que se inclui em seu cálculo o fator tendência e o fator sazonalidade. Consequentemente, ele é um modelo que leva em conta mais variáveis e isso o deixa um modelo mais robusto e, geralmente, mais assertivo quando os dados possuem este tipo de comportamento.

Ainda segundo a obra citada acima, podemos descrever uma série temporal sazonal desta forma:

- $\hat{y}_{t+n} = E_t + nT_t + S_{t+n-p}$

Sendo:

- $E_t = \alpha(y_{t-})\hat{S}_{t-p} + (1 - \alpha)(E_{t-1} + T_{t-1})$,
- $T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1}$,
- $S_t = \gamma(y_T - E_t) + (1 - \gamma)S_{t-p}$

Onde:

- E_t é o valor ajustado que representa o nível;
- T_t é o valor ajustado que representa a tendência;
- S_T é o valor ajustado que representa a sazonalidade;
- p representa o número de estações consideradas para a série temporal;
- α é a constante de suavização do nível;
- β é a constante de suavização da tendência;
- γ é a constante de suavização da sazonalidade;
- \hat{y}_{t+n} a previsão calculada no tempo t .

Média quadrática dos erros (MSD)

Segundo Ragsdale(2009), ao realizar um modelo de séries temporais, podemos utilizar esta métrica a Erro Quadrático Médio (EQM) para verificar a assertividade do modelo.

Seu cálculo se dá da seguinte forma:

- $$EQM = \frac{1}{n} \sum_{t:1}^n (y_t - \hat{y}_t)^2$$

Onde,

- t representa a unidade de tempo;
- i representa até onde vai esta unidade de tempo;

O EQM, quanto menor ele for, mais assertivo é o modelo será considerado. Ele será a métrica utilizada neste estudo para compararmos e eficiência de um modelo em relação ao outro.

3. METODOLOGIA

Em uma pesquisa científica é necessário entender qual a metodologia a ser utilizada na pesquisa, além disso, é importante deixar claro qual é o tipo de pesquisa a ser utilizado em seu estudo, para facilitar o entendimento do leitor no decorrer do mesmo.(Venanzi e Silva, 2016).

Esta monografia trata-se de uma pesquisa de caráter aplicada, partindo de uma abordagem quantitativa, com o objetivo de descrever os resultados obtidos.

Além da revisão bibliográfica sobre os conceitos abordados neste estudo, a pesquisa se divide em duas grandes etapas: a coleta dos dados relacionados as mortes causadas pelo

coronavírus no Brasil e, posteriormente, utilizar os modelos de previsão e analisar os resultados obtidos.

A coleta dos dados foi feita na base de dados da *Oxford University*, onde eles possuem um site “*Our world in data*”, onde possui os dados sobre a pandemia em todo mundo, incluindo o Brasil.

Após a coleta destes dados, transferimos todos estes dados para uma planilha do Microsoft Excel®. A partir disso, programamos esta planilha para calcular alguns termos matemáticos ligados ao modelo de Holt-Winters(conforme fórmulas apresentadas acima), com o objetivo de encontrar as melhores constantes de suavização para o modelo, baseando no princípio de: as melhores constantes de suavização de um modelo serão as que resultam o menor EQM.

Por fim, foi utilizado o *software* estatístico Minitab®. No Minitab, foram executados dois modelos de previsão de séries temporais: um modelo foi o de médias móveis, o outro modelo foi o de holt-winters. Pelo próprio Minitab®, após rodarmos o modelo, o programa gera o resultado de toda a previsão, do modelo ajustado, e também dos gráficos de análise de assertividade do modelo. O modelo demorou 2 segundos para ser rodado, em um computador com configurações: Processador Intel Core i7-8565U, Memória RAM 8GB e placa de vídeo NVIDIA GeForce MX110 2GB.

Posteriormente aos modelos estarem prontos, realizou-se uma análise visual dos gráficos dos modelos ajustados, valendo como uma análise prévia do comportamento dos modelos em relação aos resultados reais. Além disso, analisa-se os gráficos de resíduos dos modelos, para verificar a assertividade e validade dos modelos. Após isso, analisa-se a métrica que irá, finalmente, comparar a assertividade dos modelos: o EQM. Após analisar estes pontos acima, os dois modelos foram comparados discutiu-se seus resultados.

4. RESULTADOS E DISCUSSÕES

Conforme citado acima, para analisar os dados relacionados as mortes de Covid-19, foram utilizados dois modelos de séries temporais: o modelo de médias móveis e o modelo de holt-winters. Cada um destes tem suas fórmulas de calcular seus respectivos modelos e também alguns parâmetros utilizados em cada um.

Modelo de médias móveis

Utilizando-se o modelo de médias móveis um total de 583 dados, onde o “1º dado” se refere ao dia 17 de março de 2020, dia o qual ocorreu a primeira morte causada pelo vírus, indo até o dia 21 de outubro de 2021.

Como citado acima, a fórmula para o cálculo de cada unidade de um modelo ajustado pelo método de médias móveis é esta:

- $\hat{y}_t = \frac{y_t + y_{t-1} + \dots + y_{t-7+1}}{14}$

Onde, o r utilizado será o intervalo de 14 unidades, já que 14 unidades para o modelo ser igual ao que a mídia está utilizando e propagando por todo o Brasil, a fim de obtermos uma comparação real.

Além disso, como também citado acima, será calculada uma métrica de análise de comparação da eficiência do modelo chamada MSD, que se constrói a partir dessa fórmula:

- $MSD = \frac{1}{583} \sum_{t:1}^{583} (y_t - \hat{y}_t)^2$

Onde, como pode-se ver na fórmula, o valor de “i” é 583, devido ao número de unidades encontrados nos dados.

Ao analisarmos as figuras 1 e 2 que foram gerados pelo Minitab® abaixo, podemos tirar algumas conclusões sobre seus respectivos gráficos:

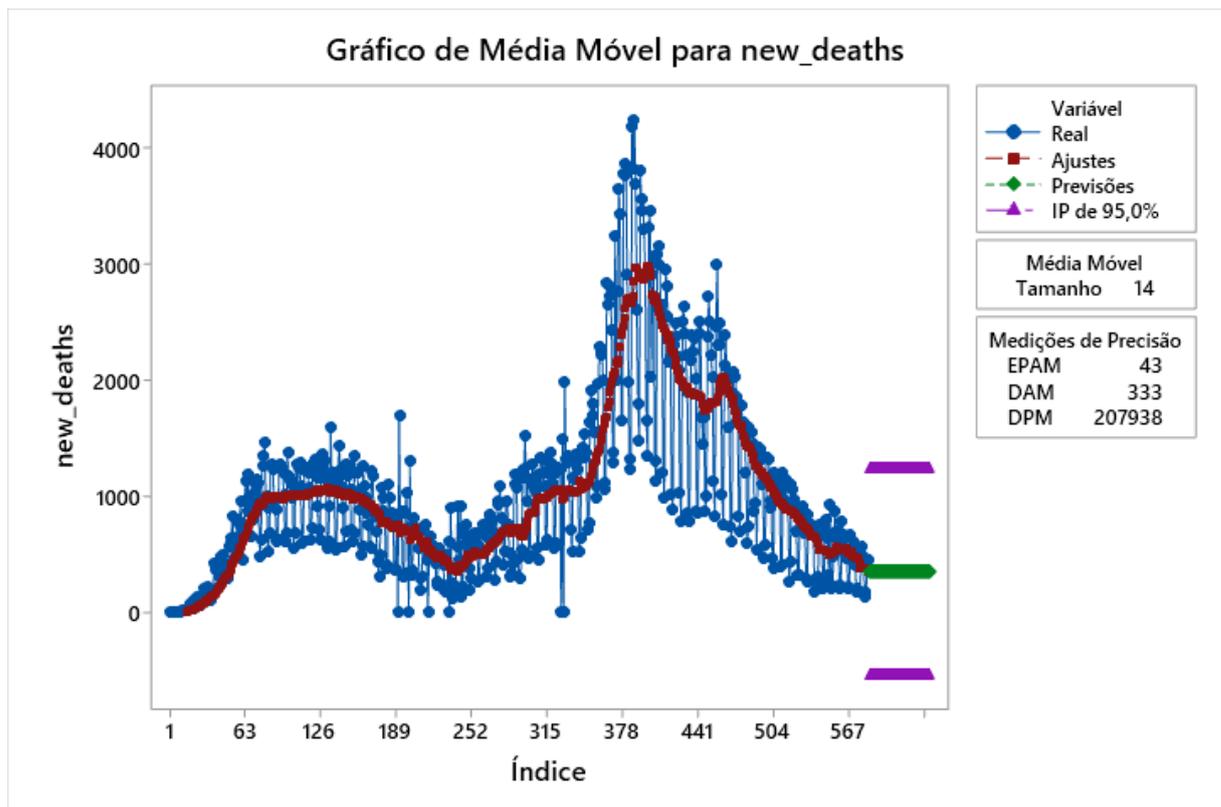


Figura 1 – Gráficos dos Resultados modelo de Médias Móveis – Fonte autoral(2021)

- Analisando visualmente o gráfico, podemos perceber que o modelo de médias móveis acompanha a movimentação “macro” do gráfico, porém sem acompanhar fielmente seus picos e vales, onde a curva do modelo, em vermelho, se mostra muito mais suavizada ao se comparar com a curva dos

dados reais. Podemos analisar que, pelo fato de ser um modelo de 583 unidades e a média móvel estar com um tamanho apenas de 14, numa visão “Macro” do gráfico podemos ver claramente ela acompanhando a sazonalidade do modelo(mesmo que ela não utilize meios matemáticos específicos para isso). Entretanto, justamente, por esses motivos, há essa grande suavização e as vezes o modelo pode fugir um pouco da realidade, podendo prejudicar possíveis tomadas de decisões baseadas nele.

- Ao calcularmos o EQM do modelo, encontramos o resultado de 207.938.

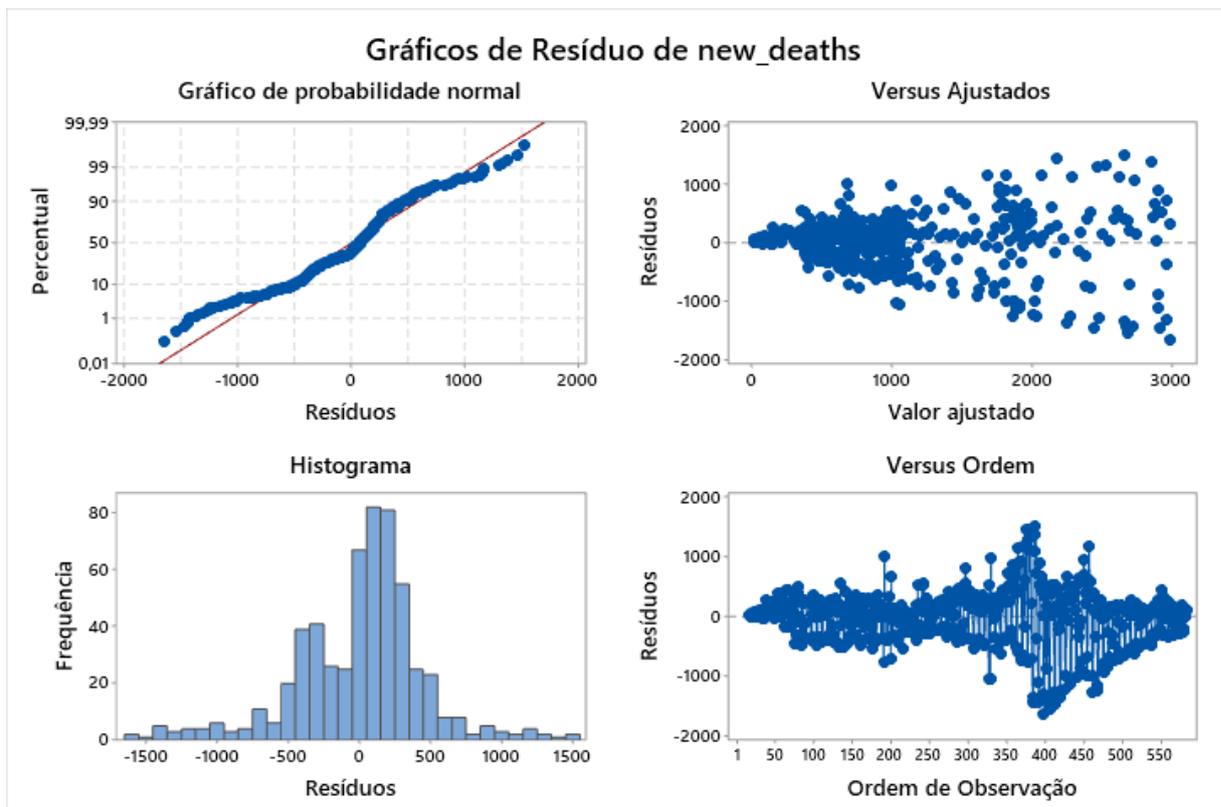


Figura 2 - Gráficos dos resultados dos resíduos do modelo de Médias Móveis – Fonte autoral(2021)

Analisando os gráficos de resíduos gerados pelo modelo, analisando-se as premissas necessárias para a validade do modelo, vemos:

- Segundo o gráfico de probabilidade normal, podemos ver que os pontos acompanham a reta de regressão linear, logo podemos pressupor que os dados residuais seguem uma distribuição normal.
- Olhando para o gráfico de resíduos e valores ajustados, podemos perceber que os dados seguem uma variância não constante, fazendo com que ela aumente de acordo com o crescimento dos valores previstos.

- Analisando-se o histograma, podemos também concluir que os dados são distribuídos simetricamente, aderindo a suposição de serem distribuídos normalmente;
- Sobre o gráfico de resíduos por ordem de observação, podemos também ver que ele está distribuído aleatoriamente variando entre pontos abaixo e acima do gráfico, atendendo a premissa de aleatoriedade dos resíduos para validar o modelo.

Modelo de Suavização exponencial sazonal de Holt-Winters

Utilizando-se o modelo de Holt-Winters, foram utilizados, claramente, os mesmo dados utilizados para o modelo de médias móveis, sendo: um total de 583 dados, onde o “1º dado” se refere ao dia 17 de março de 2020, dia o qual ocorreu a primeira morte causada pelo vírus, indo até o dia 21 de outubro de 2021.

A fórmula matemática para o cálculo do modelo ajustado, como citado acima, segue este formato:

$$\hat{y}_{t+n} = E_t + nT_t + S_{t+n-p}$$

No modelo de Holt Winters, devemos definir previamente quais serão as constantes de suavização que resultarão em um modelo assertivo. Para isso, foi utilizado o Microsoft Excel® para encontrar os valores ótimos para estas respectivas constantes, conforme a tabela contida na figura 3:

t	Mortes	Nível Esperado	Tendência	Fator Sazonal	Previsão	Diferença	α	0,124
1	1	*	*	1,00	*	*	β	0,000
2	2	*	*	1,99	*	*	γ	0,244
3	3	*	*	2,99	*	*	MSD	149.960,54
4	5	*	*	4,98	*	*		
5	4	*	*	3,98	*	*		
6	10	*	*	9,96	*	*		
7	9	*	*	8,96	*	*		
8	12	*	*	11,95	*	*		
9	13	*	*	12,95	*	*		
10	18	*	*	17,93	*	*		
11	15	*	*	14,94	*	*		
12	19	*	*	18,92	*	*		
13	25	*	*	24,90	*	*		
14	23	*	*	22,90	*	*		
15	42	*	*	41,83	*	*		
16	39	*	*	38,84	*	*		
17	84	*	*	83,65	*	*		
18	35	*	*	34,85	*	*		
19	86	*	*	85,64	*	*		
20	41	*	*	40,83	*	*		
21	78	*	*	77,68	*	*		
22	122	*	*	121,49	*	*		
23	133	*	*	132,45	*	*		
24	131	*	*	130,45	*	*		
25	107	*	*	106,55	*	*		
26	67	*	*	66,72	*	*		
27	99	*	*	98,59	*	*		
28	105	*	*	104,56	*	*		
29	204	*	*	203,15	*	*		
30	204	*	*	203,15	*	*		
31	188	*	*	187,22	*	*		
32	217	*	*	216,10	*	*		
33	213	*	*	212,11	*	*		
34	108	*	*	107,55	*	*		

Figura 3 - Tabela de Programação do Excel

Onde:

- $\hat{y}_{t+n} = E_t + nT_t + S_{t+n-p}$;

Sendo:

- $E_t = \alpha(y_{t-})\hat{S}_{t-p} + (1 - \alpha)(E_{t-1} + T_{t-1})$;
- $T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1}$;
- $S_t = \gamma(y_T - E_t) + (1 - \gamma)S_{t-p}$;
- A coluna “t” se refere à cada dia (no intervalo de 0 a 583);
- A coluna “Mortes” se refere a quantidade de mortes causadas pelo coronavírus no determinado dia “t”;
- A coluna “Nível esperado” calcula o valor ajustado que representa o nível, com base na seguinte fórmula: $E_t = \alpha(y_{t-})\hat{S}_{t-p} + (1 - \alpha)(E_{t-1} + T_{t-1})$;
- A coluna “Tendência” calcula o valor ajustado que representa a tendência, com base na seguinte fórmula: $T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1}$;
- A coluna “Fator Sazonal” calcula o valor ajustado que representa a sazonalidade, com base na seguinte fórmula: $S_t = \gamma(y_T - E_t) + (1 - \gamma)S_{t-p}$;
- A célula referente ao cálculo do MSD foi calculada com base na seguinte fórmula: $MSD = \frac{1}{583} \sum_{t:1}^{583} (y_t - \hat{y}_t)^2$.

Para encontrar os valores ótimos de α , γ e δ , foi realizada uma programação dentro do Microsoft Excel baseada no método Simplex, onde configuramos para serem encontrados os seus respectivos valores que resultariam no menor MSD:

- Minimizar $MSD = \frac{1}{n} \sum_{t:1}^n (y_t - \hat{y}_t)^2$

Sujeito a:

- $\alpha, \gamma, \delta < 1$;
- $\alpha, \gamma, \delta > 0$;

Parâmetros do Solver

Definir Objetivo:

Para: Máx. MÍN. Valor de:

Alterando Células Variáveis:

Sujeito às Restrições:

\$P\$4:\$P\$6 <= 1	
\$P\$4:\$P\$6 >= 0	

Tornar Variáveis Irrestritas Não Negativas

Selecionar um Método de Solução:

Método de Solução

Selecione o mecanismo GRG Não Linear para Problemas do Solver suaves e não lineares. Selecione o mecanismo LP Simplex para Problemas do Solver lineares. Selecione o mecanismo Evolutionary para problemas do Solver não suaves.

Onde:

- O “objetivo” da função é alcançar o resultado mínimo do MSD;

Os valores encontrados para essas constantes foram

- $\alpha = 0,124$
- $\gamma = 0,000$
- $\delta = 0,244$

Após encontrarmos os valores ótimos para as constantes, as utilizamos para modelar a previsão:

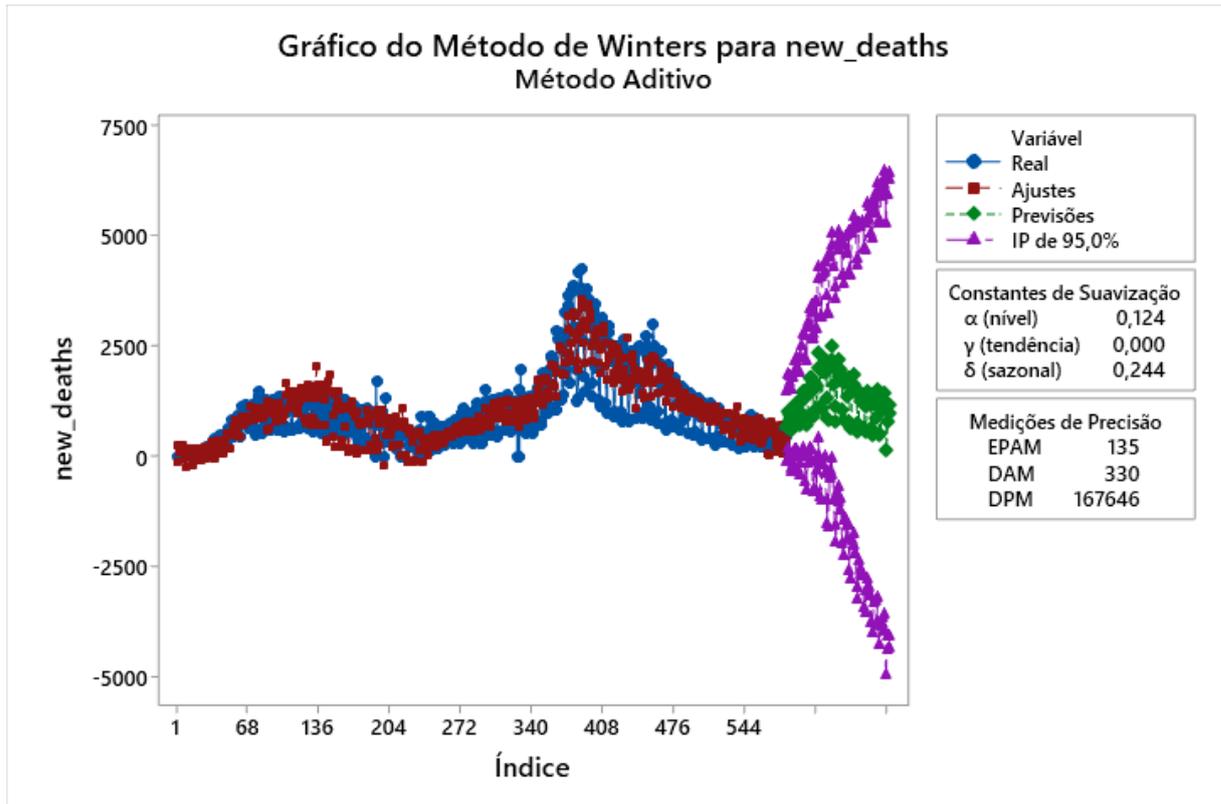


Figura 4 - Gráfico dos resultados do modelo de Holt-Winters

- Analisando visualmente o gráfico gerado pelo modelo de Holt-Winters, podemos analisar que a curva de vermelho (referente ao modelo ajustado) acompanha muito bem a curva dos dados reais, incluindo uma sensibilidade maior aos picos e vales da curva real. Podemos analisar também que o fator sazonal do modelo faz com que ele acompanhe também o aspecto sazonal da variação dos dados. Uma vez que ele é composto por esse aspecto sazonal e conta com uma sensibilidade maior às variações dos dados reais, torna-se um modelo aparentemente muito interessante de se usar para prever os dados em questão.
- Ao calcularmos o MSD do modelo, foi encontrado o valor de 167.646.

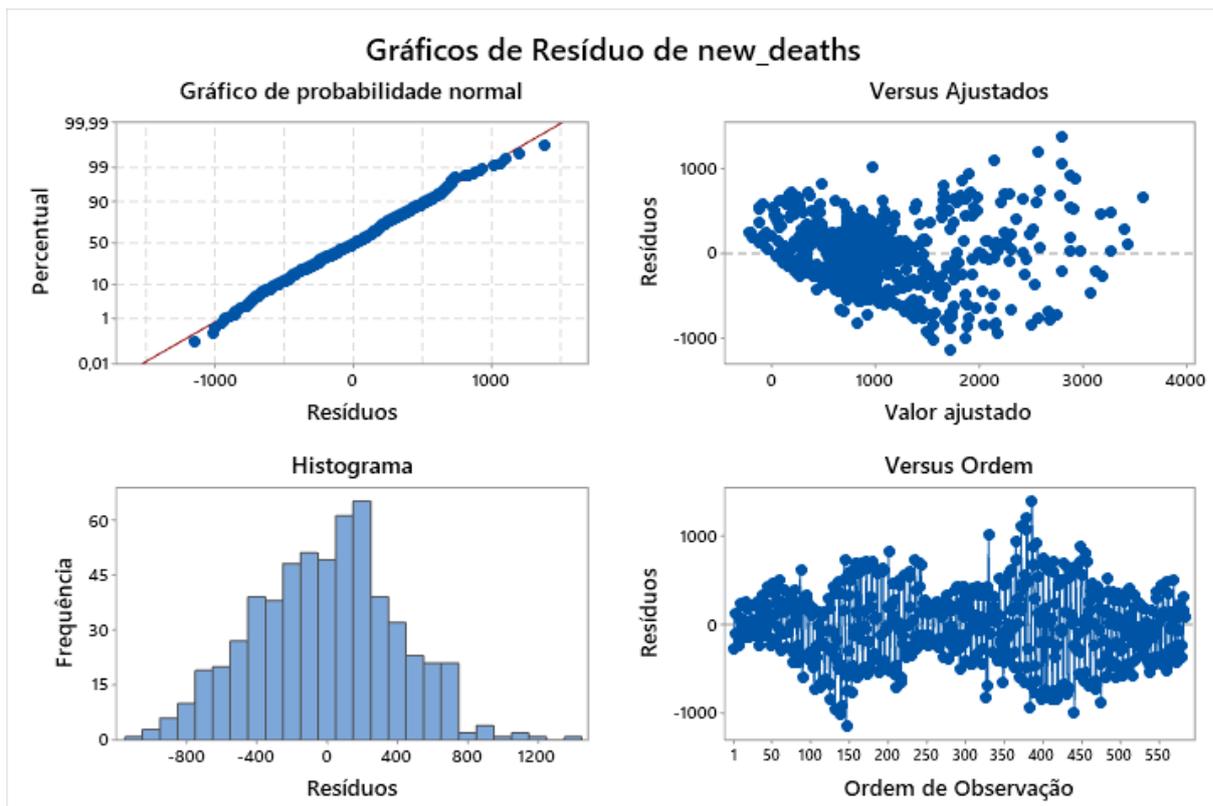


Figura 5 - Resultados dos gráficos de resíduos do modelo de Holt-Winters

Analisando os gráficos de resíduos gerados pelo modelo, analisando-se as premissas necessárias para a validade do modelo, vemos:

- Segundo o gráfico de probabilidade normal, podemos ver que os pontos acompanham a reta de regressão linear, logo podemos pressupor que os dados residuais seguem uma distribuição normal.
- Olhando para o gráfico de resíduos e valores ajustados, podemos perceber que os dados seguem uma variância constante, uma vez que os dados estão bem distribuídos;
- Analisando-se o histograma, podemos também concluir que os dados são distribuídos simetricamente, aderindo a suposição de serem distribuídos normalmente;
- Sobre o gráfico de resíduos por ordem de observação, podemos também ver que ele está distribuído aleatoriamente variando entre pontos abaixo e acima do gráfico, atendendo a premissa requisitada para podermos validar o modelo.

CONCLUSÕES FINAIS

O objetivo geral deste trabalho é construir e ajustar os dois modelos propostos, onde o modelo de Médias Móveis foi o modelo amplamente utilizado pelos veículos de imprensa e o modelo de Holt-Winters foi escolhido pelo fato de ser um modelo que leva mais variáveis em consideração e, possivelmente poderia ser mais fiel aos dados reais.

Como engenheiro de produção, acredito que “desafiar” o método utilizado pelo senso comum com um objetivo de tentar achar soluções mais eficientes é um grande papel que podemos realizar, principalmente dentro de uma grande área da Engenharia da Produção, como a estatística aplicada à gestão pública.

Analisando os resultados e interpretações dos modelos propostos, ficou claro que, segundo a métrica utilizada para comparar a assertividade dos mesmos, o MSD, juntamente com as análises gráficas, podemos concluir que o modelo de Holt-Winters se demonstrou mais efetivo do que o modelo de Médias Móveis.

Primeiramente, porque o MSD deste último é 207.938, enquanto o de Holt-Winters foi 167.646, demonstrando que o modelo “erra” muito menos.

Além disso, segundo as interpretações gráficas de cada um dos modelos, podemos ver que a curva do modelo ajustado de Holt-Winters se aproxima muito mais da curva dos dados reais ao compararmos com o modelo de Médias Móveis. O primeiro tem uma sensibilidade maior aos picos e vales enquanto o segundo se encontra de uma forma bem mais suavizada.

Em termos práticos, estes modelos são utilizados para prever um número tão importante e relevante nos últimos anos como as mortes causadas pela Covid-19, onde servem como base para informar toda a população de um país, além de serem grandes entradas para importantes tomadas de decisão dentro da gestão pública, como elaboração de cronogramas de vacinação, construção de medidas restritivas à população para combater a pandemia e todos seus desdobramentos em cada cidade do país.

Levando em conta todos estes aspectos, conclui-se que o modelo de Holt-Winters se demonstra mais efetivo para ser utilizado como tomadas de decisões tão importantes citadas acima, uma vez que ele representa melhor os dados reais

5. REFERÊNCIA BIBLIOGRÁFICA

VENANZI, D.; SILVA, O. R. **Introdução à engenharia de produção**: conceitos e casos práticos. Rio de Janeiro: Livros Técnicos e Científicos Editora Ltda, 2016.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **“Introduction to time series analysis and forecasting”**.2. ed. Nova Jersey: Wiley, 2015.

MORRETIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Egard Blcuher, 2006.

RAGSDALE, C. T. **Modelagem e Análise de Decisão**. São Paulo: Cengage Learning, 2009.

UNIVERSITY OF OXFORD. **Our world in data**, c2021. Página Brasil. Disponível em: < <https://ourworldindata.org/coronavirus/country/brazil>>. Acesso em: 25 de Out. de 2021.

WORLD HEALTH ORGANIZATION. **Who Coronavírus Dashboard**, c2021. Overview. Disponível em: <<https://ourworldindata.org/coronavirus/country/brazil>> . Acesso em: 10 de Nov. de 2021.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Transmissão do SARS-CoV-2: implicações para as precauções de prevenção de infecção**. Resumo Científico, 2020.

ALVES, C.; FIALHO, R. P. **Atividades de Modelagem Matemática: Saúde Pública**. 1 ed. Belém: Universidade do Estado do Pará, 2019.

FARIAS, A. A. et al. **Introdução à Estatística**. Rio de Janeiro: LTC, 2008.