

APLICAÇÃO DE MÉTODOS NUMÉRICOS COM O EXCEL PARA A ANÁLISE E MODELAGEM MATEMÁTICA DA CURVA DE CONTAMINAÇÃO DA COVID-19 NOS ESTADOS DO BRASIL

Alan Soares de Lima¹

Anderson Augusto de Oliveira Moura²

Juliano Pedro da Silva³

Maria Luiza Vieira Reis⁴

Sidicley Borba dos Santos⁵

Engenharia Mecatrônica



ISSN IMPRESSO 1980-1777

ISSN ELETRÔNICO 2316-3135

RESUMO

Atualmente estamos vivenciando a triste realidade dos impactos gerados pela pandemia do novo Coronavírus (COVID-19) e dentre os principais efeitos causados por essa pandemia estão a crise econômica e a crise nos sistemas de saúde em vários países ao redor do mundo. Diante dessa situação surge a importância da utilização de ferramentas de análise de dados para acompanhar o crescimento do número de casos confirmados pela COVID-19, e ajudar os governos na elaboração de políticas públicas que possam mitigar os problemas gerados pela pandemia. Este trabalho tem como objetivo desenvolver um método numérico para a geração de modelos matemáticos que possam ser usados para representar curvas de tendências, e realizar estimativas da previsão do número de casos confirmados da COVID-19 em vários estados do Brasil. Utilizando-se de técnicas matemáticas de interpolação polinomial e eliminação gaussiana e de técnicas de programação em VBA, no Excel, foi desenvolvido um programa computacional que possibilita ao usuário, por meio de uma interface interativa, analisar o crescimento do número de casos confirmados da COVID-19 no estado de Alagoas e realizar estudos comparativos da situação com outros estados do Brasil. Os resultados obtidos com a análise da curva de tendência dos dados da COVID-19 coletados em Alagoas e em outros estados comprovaram eficácia do método numérico proposto, constituindo-se assim de um método útil para a análise de cenários de crise.

PALAVRAS-CHAVE

Análise de Dados. COVID-19. Interpolação Polinomial. Curvas de Tendência. Programação em VBA. Excel.

ABSTRACT

We are currently experiencing the sad reality of the impacts generated by the pandemic of the new Coronavirus (COVID-19) and among the main effects caused by this pandemic are the economic crisis and the crisis in health systems and several countries around the world. Given this situation arises the importance of using data analysis tools to monitor the growth in the number of cases confirmed by COVID-19, and to assist governments in the development of public policies that can mitigate the problems generated by the pandemic. This work aims to develop a numerical method for the generation of mathematical models that can be used to represent trend curves, and to estimate the prediction of the number of confirmed cases of COVID-19 in several states of Brazil. Using mathematical techniques of polynomial interpolation and Gaussian elimination and programming techniques in VBA, Excel, a computational program was developed that allows the user, through an interactive interface, to analyze the growth of the number of confirmed cases of COVID-19 in the state of Alagoas and to conduct comparative studies of the situation with other states of Brazil. The results obtained with the analysis of the trend curve of covid-19 data collected in Alagoas and other states proved efficacy of the proposed numerical method, thus constituting a useful method for the analysis of crisis scenarios.

KEYWORDS

Data analysis, COVID-19, Polynomial interpolation, Trend curves, VBA programming, Excel.

1 INTRODUÇÃO

Em meados do século XVII na Europa havia uma crença comum e generalizada de que todos os cisnes eram da cor branca por natureza, entretanto a descoberta da existência de cisnes negros na Austrália causou um grande impacto na sociedade europeia da época, pois esse evento não era previsto e ao menos citado nas estatísticas da época. Desde então o termo “Cisne negro” passou a ser usado para descrever eventos improváveis e transformadores, mas possíveis de serem previstos.

A pandemia da COVID-19 tem se figurado como o grande cisne negro de 2020, que no Brasil tem causado uma série de problemas no âmbito social e econômico de muitos estados. Até o dia 18/05/2020 foram contabilizados no Brasil 254.220 casos confirmados e 16.792 mortes pela doença. A Tabela 1 revela que as regiões mais afeta-

das pela pandemia foram a Sudeste e Nordeste que juntas somam aproximadamente 74% do total de casos confirmados pela doença no país.

Tabela 1 – Casos confirmados da COVID-19 por região do Brasil

NÚMERO DE CASOS CONFIRMADOS DA COVID-19 POR REGIÃO DO BRASIL	
REGIÃO	CASOS CONFIRMADOS
Centro - Oeste	7.918
Sul	11.270
Norte	47.319
Nordeste	86.130
Sudeste	101.583

Fonte: Secretarias Estaduais de Saúde, Brasil (2020).

Em contrapartida a esses números os governos estaduais têm adotado algumas medidas de distanciamento social para tentar desacelerar seu crescimento e evitar o colapso nos sistemas de saúde (Hospitais e Unidades de emergência). No enfrentamento de cenários de crise é de fundamental importância que as ações governamentais de cada localidade sejam pautadas na análise prévia de dados da situação, e para que isso seja possível faz-se necessário colocar a disponibilidade das autoridades ferramentas de análise de dados que permitam analisar a situação de uma forma simples e eficiente.

Cientistas em todo mundo tem trabalhado no desenvolvimento de soluções de análise de dados para auxiliar os governos no planejamento de ações mitigadoras e conscientizar a população do novo cenário. Por meio de bancos de dados disponíveis sobre a situação da COVID-19 em uma localidade e com uso da técnica matemática de interpolação polinomial foi possível a criação de poderosos algoritmos numéricos que possibilitam a geração de modelos matemáticos que podem ser utilizados como modelos de previsão de casos da COVID-19.

Dentre os *softwares* disponíveis atualmente para o desenvolvimento de soluções em análise de dados da COVID-19, o Excel oferece uma série de recursos poderosos a saber, cálculos, plotagem de gráficos, programação e integração com outras plataformas. O desenvolvimento de métodos numéricos com o Excel é objeto chave deste trabalho, com o objetivo de prover a comunidade acadêmica conhecimentos técnicos e práticos que possam ser utilizados para desenvolver soluções de análise de dados no *software*, para avaliar e fazer previsões sobre os possíveis impactos da pandemia da COVID-19 nos estados do Brasil.

2 METODOLOGIA

A análise de dados é processo de transformação de números em informação, em que os dados coletados são utilizados para determinar padrões e tendências que

podem suportar ou rejeitar teorias sobre causa e efeito. Neste trabalho a análise dos dados da COVID19, nos estados do Brasil, foi dividida em 3 passos fundamentais: Coleta e processamento de dados, Modelagem matemática e o desenvolvimento de uma interface gráfica para a apresentação de resultados.

2.1 COLETA E PROCESSAMENTO DOS DADOS

Todos os dados necessários para a composição deste trabalho foram coletados a partir do portal digital Coronavírus Brasil (Veículo oficial de comunicação sobre a situação epidemiológica da COVID-19 no Brasil) durante o período de 18/03/2020 a 18/05/2020. Nesse período foi analisado a evolução diária da pandemia da COVID-19 no Brasil e realizada a filtragem do número de casos confirmados pela doença em cada estado do país com o objetivo de criar uma base de dados no Excel.

Para este trabalho serão apresentados e manipulados apenas os dados referentes a situação da pandemia da COVID-19 no estado de Alagoas, uma vez que a metodologia aplicada será a mesma para os outros estados. Na Tabela 2, é demonstrada a correlação entre data de ocorrência e o número de casos confirmados e acumulados em Alagoas.

Tabela 2 – Evolução diária de casos confirmados da COVID-19 em Alagoas

SITUAÇÃO EPIDEMIOLÓGICA DE ALAGOAS					
DATA	Nº DE PESSOAS INFECTADAS	DATA	Nº DE PESSOAS INFECTADAS	DATA	Nº DE PESSOAS INFECTADAS
18-mar	1	8-abr	37	29-abr	957
19-mar	4	9-abr	37	30-abr	1044
20-mar	5	10-abr	45	1-mai	1226
21-mar	7	11-abr	48	2-mai	1372
22-mar	7	12-abr	48	3-mai	1441
23-mar	7	13-abr	50	4-mai	1538
24-mar	9	14-abr	72	5-mai	1606
25-mar	11	15-abr	82	6-mai	1703
26-mar	11	16-abr	89	7-mai	1867
27-mar	11	17-abr	110	8-mai	2033
28-mar	14	18-abr	132	9-mai	2172
29-mar	17	19-abr	159	10-mai	2258
30-mar	17	20-abr	171	11-mai	2343
31-mar	18	21-abr	210	12-mai	2580
1-abr	18	22-abr	243	13-mai	2761
2-abr	18	23-abr	324	14-mai	2972
3-abr	22	24-abr	413	15-mai	3212
4-abr	23	25-abr	501	16-mai	3593
5-abr	28	26-abr	554	17-mai	3816
6-abr	31	27-abr	643	18-mai	4031
7-abr	32	28-abr	777		

Fonte: Secretaria de Estado da Saúde (SESAU), Alagoas (2020).

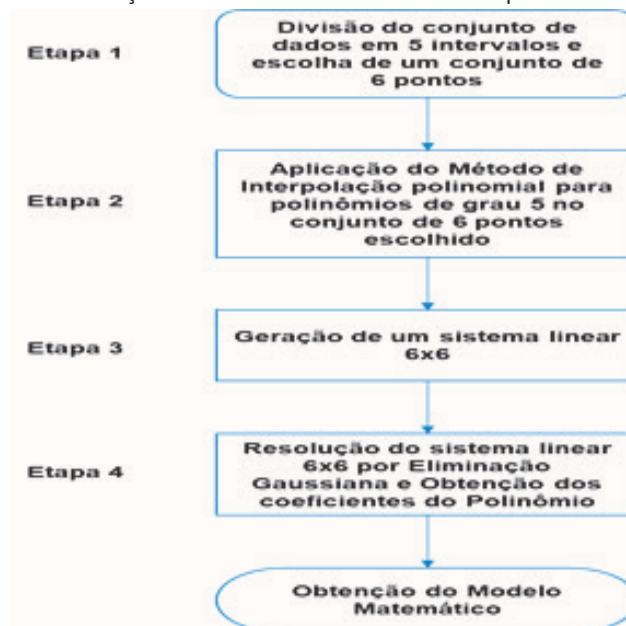
Os dados da Tabela 2 revelam que existe uma relação não linear entre as variáveis dependentes: data de ocorrência e número de pessoas infectadas. Essa relação é observada na Figura 1.

Figura 1 – Evolução de casos confirmados da COVID-19 em Alagoas

Fonte: Autores (2020).

2.2 CRIAÇÃO DO MODELO MATEMÁTICO PARA O ESTADO DE ALAGOAS UTILIZANDO O MÉTODO DA INTERPOLAÇÃO POLINOMIAL DE LAGRANGE NOS DADOS COLETADOS

O processo de criação do modelo matemático para descrever o avanço da COVID-19 no estado de Alagoas durante o período analisado, foi desenvolvido de acordo com as etapas descritas no fluxograma da Figura 2.

Figura 2 – Processo de criação do modelo matemático para o estado de Alagoas

Fonte: Autores (2020).

Com o objetivo de determinar uma equação que descreva o número de pessoas infectadas em função da data de ocorrência, na Etapa 1 foi feita a divisão dos dados da Tabela 2 em cinco intervalos, e escolhido o conjunto de pontos¹ demonstrados na Tabela 3.

Tabela 3 – Conjunto de pontos escolhidos para a aplicação da interpolação polinomial

X	Y
1	1
5	7
20	31
35	210
50	1.703
60	3.593

Fonte: Autores (2020).

O método da interpolação polinomial de Lagrange consiste em uma operação matemática que têm por objetivo determinar um polinômio que melhor represente um conjunto de dados conhecido. Esse método foi desenvolvido por Joseph Louis Lagrange, matemático italiano, e é essencial para a aferição de análise de dados que apresentem um comportamento não linear entre as variáveis dependentes, gerando como resultado da operação um conjunto de sistemas lineares que podem ser resolvidos via eliminação gaussiana, método desenvolvido por Carl Friedrich Gauss (Matemático e Físico alemão). A eliminação gaussiana consiste em manipular os sistemas lineares por meio de determinadas operações elementares, transformando a matriz estendida desses sistemas em uma matriz triangular e obtendo a solução via substituição regressiva.

Na Etapa 2, aplicando o método de interpolação polinomial no conjunto de pontos definido foi feita a aproximação da equação desejada para um polinômio de grau 5, como mostrado na Equação (2.1).

$$y = a_1 + a_2x + a_3x^2 + a_4x^3 + a_5x^4 + a_6x^5 \quad (2.1)$$

Para obter os coeficientes $a_1, a_2, a_3, a_4, a_5, a_6$ que definem o polinômio foi aplicado o método de interpolação polinomial no conjunto de pontos definido, gerando como resultado na Etapa 3 um sistema linear 6x6 demonstrado na Figura 3.

Figura 3 – Sistema Linear 6x6 gerado

$$\begin{cases} a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = 1 & (L1) \\ a_1 + 5a_2 + 25a_3 + 125a_4 + 625a_5 + 3125a_6 = 7 & (L2) \\ a_1 + 20a_2 + 400a_3 + 8000a_4 + 160000a_5 + 3200000a_6 = 31 & (L3) \\ a_1 + 35a_2 + 1225a_3 + 42875a_4 + 1500625a_5 + 52521875a_6 = 210 & (L4) \\ a_1 + 50a_2 + 2500a_3 + 125000a_4 + 6250000a_5 + 312500000a_6 = 1703 & (L5) \\ a_1 + 60a_2 + 3600a_3 + 216000a_4 + 12960000a_5 + 777600000a_6 = 3593 & (L6) \end{cases}$$

Fonte: Autores (2020).

¹ X – Representa o número de dias transcorridos a contar do dia 18/03/2020.

Y – Representa o número de pessoas infectadas a contar do dia 18/03/2020.

Na Etapa 4 para a resolução do sistema linear 6x6 no Excel, foi extraído a matriz estendida dos coeficientes e aplicado o método da Eliminação Gaussiana que consiste em realizar uma transformação triangular da matriz do sistema linear, permitindo obter de forma individual os valores dos coeficientes. O processo de transformação triangular do sistema linear é demonstrado na Figura 4 que reúne o conjunto de equações que foram desenvolvidas.

Figura 4 – Transformação Triangular do Sistema Linear 6x6

	1	1	1	1	1	1	1
	0	4	24	124	624	3124	6
$L_{n+1} = L_{n+1} - L_1$	0	19	399	7999	159999	3199999	30
	0	34	1224	42874	1500624	52521874	209
	0	49	2499	124999	6249999	312499999	1702
	0	59	3599	215999	12959999	777599999	3592
	1	1	1	1	1	1	1
	0	4	24	124	624	3124	6
$L_{n+2} = L_{n+2} - \left(\frac{a_{(n+2),2}}{a_{2,2}}\right) \times L_2$	0	0	285	7410	157035	3185160	1,5
	0	0	1020	41820	1495320	5,2E+07	158
	0	0	2205	123480	6242355	3,1E+08	1628,5
	0	0	3245	214170	1,3E+07	7,8E+08	3503,5
	1	1	1	1	1	1	1
	0	4	24	124	624	3124	6
$L_{n+3} = L_{n+3} - \left(\frac{a_{(n+3),3}}{a_{3,3}}\right) \times L_3$	0	0	285	7410	157035	3185160	1,5
	0	0	0	15300	933300	41095800	152,6316
	0	0	0	66150	5027400	287818650	1616,895
	0	0	0	129800	11162800	741287800	3486,421
	1	1	1	1	1	1	1
	0	4	24	124	624	3124	6
$L_{n+4} = L_{n+4} - \left(\frac{a_{(n+4),4}}{a_{4,4}}\right) \times L_4$	0	0	0	15300	933300	3185160	1,5
	0	0	0	0	992250	4,1E+07	152,63
	0	0	0	0	0	1,1E+08	956,99
	0	0	0	0	3245000	3,9E+08	2191,5
	1	1	1	1	1	1	1
	0	4	24	124	624	3124	6
$L_{n+5} = L_{n+5} - \left(\frac{a_{(n+5),5}}{a_{5,5}}\right) \times L_5$	0	0	285	7410	157035	3185160	1,5
	0	0	0	15300	933300	41095800	152,6316
	0	0	0	0	992250	110139750	956,9876
	0	0	0	0	0	32450000	-938,133

Fonte: Autores (2020).

Após realizar a transformação triangular da matriz os valores dos coeficientes $a_1, a_2, a_3, a_4, a_5, a_6$ foram calculados e tabelados conforme a Tabela 4.

Tabela 4 – Valores dos Coeficientes

COEFICIENTE	VALOR
a_1	6,9636
a_2	-8,17
a_3	2,3696
a_4	-0,167
a_5	0,0042
a_6	-0,00003

Fonte: Autores (2020).

Com os valores dos coeficientes a_1 , a_2 , a_3 , a_4 , a_5 , a_6 definidos foi montada a Equação (2.2) que relaciona o número de pessoas infectadas (y) com a variação de dias transcorridos (x).

$$y = 6,9636 - 8,17x + 2,3696x^2 - 0,167x^3 + 0,0042x^4 - 0,00003x^5 \quad (2.2)$$

O modelo matemático obtido, representado pela Equação (2.2), possibilitará analisar e realizar estimativas de previsão de casos da COVID-19 no estado de Alagoas.

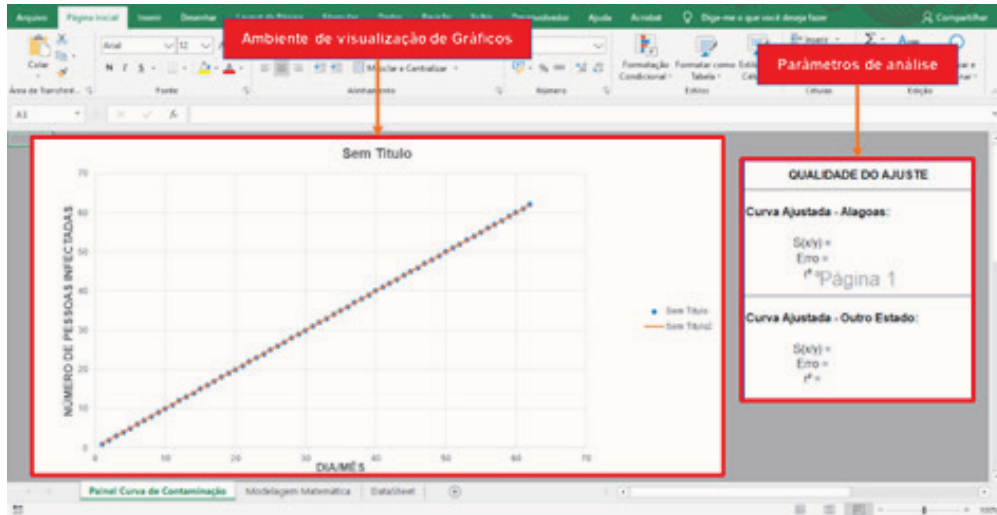
2.3 DESENVOLVIMENTO DA INTERFACE DO PROGRAMA NO EXCEL

Com o objetivo de desenvolver uma interface de programa facilitada ao usuário que lhe permita realizar a análise e avaliação do modelo matemático, utilizou-se dos recursos de programação do *Visual Basics for Applications* (VBA) disponíveis no Excel. O VBA é uma linguagem de programação orientada a objetos, com foco no desenvolvimento de aplicativos por meio dos programas da Microsoft. Nas Figuras 5 e 6, é mostrado a interface do programa desenvolvido com recursos do VBA.

Figura 5 – Painel Interativo



Fonte: Autores (2020).

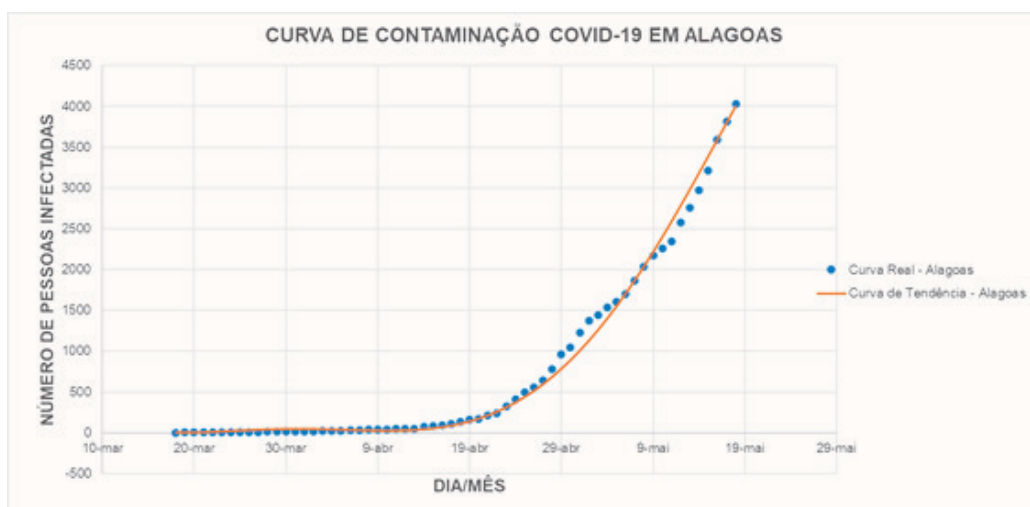
Figura 6 – Ambientes de Visualização

Fonte: Autores (2020).

O programa pode ser utilizado como uma ferramenta pelo usuário, que lhe permitirá realizar análises sobre a situação da COVID-19 em qualquer estado do Brasil de uma forma simples, interativa e objetiva.

3 RESULTADOS E DISCUSSÕES

O modelo matemático gerado possibilitou a criação de uma curva de tendência, mostrada na Figura 7, que descreve a evolução no número de casos confirmados da COVID-19 em Alagoas durante o período analisado.

Figura 7 – Curva de contaminação da COVID-19 em Alagoas

Fonte: Autores (2020).

A partir da curva de tendência é possível realizar análises e fazer estimativas para o crescimento no número de casos confirmados no estado. A Tabela 5 apresenta os valores dos coeficientes $s_{x/y}$, Erro Padrão e r^2 calculados, comparando os valores reais com os descritos pela curva de tendência. Esses coeficientes são importantes parâmetros para avaliar a qualidade do modelo matemático gerado, quanto mais próximo o valor de r^2 estiver de 1 melhor será a sua confiabilidade.

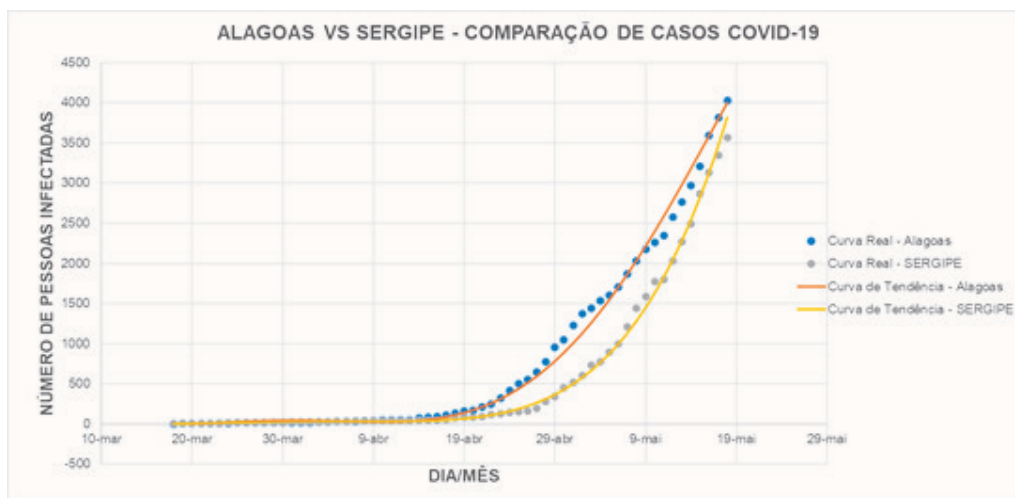
Tabela 5 – Parâmetros do Modelo Matemático

PARÂMETROS DA CURVA (ALAGOAS)	
$S_{x/y}$	671,7
Erro Padrão	72,782
r^2	0,9885

Fonte: Autores (2020).

É de fundamental importância analisar a situação de Alagoas perante os outros estados do Brasil. A partir de dados disponíveis sobre a situação epidemiológica nos outros estados, foi possível a geração de seus respectivos modelos matemáticos para fins de comparações. Nas Figura 8, foi realizada uma comparação entre os modelos matemáticos de Alagoas e Sergipe e por meio dessa comparação é possível analisar as respectivas evoluções que esses estados tiveram no número de casos confirmados da COVID-19.

Figura 8 – Comparação entre os estados de Alagoas e Sergipe



Fonte: Autores (2020).

Tabela 6 – Parâmetros dos Modelos Matemáticos

COMPARAÇÃO ENTRE OS MODELOS MATEMÁTICOS			
ALAGOAS		SERGIPE	
$S_{x/y}$	671,7	$S_{x/y}$	475,9
<i>Erro Padrão</i>	72,782	<i>Erro Padrão</i>	41,262
r^2	0,9885	r^2	0,9926

Fonte: Autores (2020).

A análise comparativa da situação epidemiológica de Alagoas e Sergipe revela que os dois estados possuem tendências de crescimento parecidas em relação ao número de pessoas infectadas pela COVID-19.

4 CONCLUSÕES

Os resultados obtidos com o modelo matemático desenvolvido neste trabalho comprovam a eficiência do método numérico utilizado na sua obtenção. A partir dos dados coletados pelas secretarias estaduais de saúde e com a aplicação do método numérico foram desenvolvidos excelentes modelos matemáticos que possibilitaram realizar a análise das situações epidemiológica dos estados do Brasil.

Entre os fatores principais que contribuíram para a obtenção da excelente qualidade dos modelos matemáticos gerados estão a quantidade de dados utilizados e a escolha do grau do polinômio adotado na Etapa 2 do processo, quanto maior for o grau adotado maior será a precisão do modelo matemático. O método numérico demonstrado neste trabalho com o auxílio do Excel poderá ser utilizado como uma ferramenta de análise de dados com o objetivo de auxiliar os governos estaduais do Brasil na elaboração de políticas públicas de enfrentamento a cenários de crise como o atual momento da pandemia da COVID-19.

REFERÊNCIAS

CHAPRA, S. C. **Métodos numéricos aplicados com matlab para engenheiros e cientistas**. 3. ed. [S.l.]: Bookman.

FORTUNA, F. A importância da análise de dados no combate à COVID-19. **Saúde Business**, 2020. Disponível em: <https://saudebusiness.com/ti-e-inovacao/a-importancia-da-analise-de-dados-no-combate-a-covid-19/>. Acesso em: 30 jun. 2020.

MATEMÁTICA UNIVERSITÁRIA. **Interpolação Polinomial**, 2020. Disponível em: <https://youtu.be/-xGbO2BhL9I>. Acesso em: 18 abr. 2020.

MINISTÉRIO DA SAÚDE. **Coronavírus Brasil**, 2020. Disponível em: <https://covid.saude.gov.br/>. Acesso em: 15 abr. 2020.

ROCK CONTENT. **A teoria do cisne negro aplicada ao coronavírus**. 2020. Disponível em: <https://inteligencia.rockcontent.com/teoria-cisne-negro-coronavirus/>. Acesso em: 21 abr. 2020.

SILVA, J. P. D. **Dashboard análise da COVID19 em Alagoas, arquivo do excel para download**. 2020. Disponível em: <https://drive.google.com/drive/folders/1wDCI3DxF8CSpcp7GvHBS2i4BCO0Y3Tg?usp=sharing>. Acesso em: 18 abr. 2020.

TALEB, N. N. **A lógica do cisne negro**. 9. ed. [S.l.]: Best Business, 2007.

Data do recebimento: 25 de novembro de 2020

Data da avaliação: 10 de dezembro de 2020

Data de aceite: 14 de dezembro de 2020

1 Acadêmico de Engenharia Mecatrônica do Centro Universitário Tiradentes – UNIT/AL.

E-mail: alan.soares@souunit.com.br

2 Acadêmico de Engenharia Mecatrônica do Centro Universitário Tiradentes – UNIT/AL.

E-mail: anderson.augusto@souunit.com.br

3 Acadêmico de Engenharia Mecatrônica do Centro Universitário Tiradentes – UNIT/AL.

E-mail: juliano.pedro@souunit.com.br

4 Acadêmica de Engenharia Mecatrônica do Centro Universitário Tiradentes – UNIT/AL.

E-mail: maria.lreis@souunit.com.br

5 Acadêmico de Engenharia Mecatrônica do Centro Universitário Tiradentes – UNIT/AL.

E-mail: sidicley.borba@souunit.com.br