Conjecturas

DOI: 10.53660/CONJ-527-814

Estimativa de parâmetros usando a técnica Monte Carlo via Cadeia de Markov aplicados a modelos epidemiológicos para Covid-19 no estado do Pará

Parameter estimation using Markov Chain Monte Carlo technique applied to epidemiological models for Covid-19 in the state of Pará

Thiago Moreira Pinto¹, Salomão Braga Santos¹, Camila Santana Dias¹, Carlos Henrique Rodrigues de Moura¹, Bruno Marques Viegas¹,Emanuel Macêdo Negrão¹, Diego Cardoso Estumano^{1*}

RESUMO

Neste artigo foi selecionada a técnica bayesiana de Monte Carlo via Cadeia de Markov (MCMC) para estimativa dos parâmetros das equações diferenciais dos modelos compartimentais SQUIDER e SEIR, buscando compreender a propagação da Covid-19 no estado do Pará. Foi elaborado um algoritmo em Matlab reproduzindo a técnica de MCMC que utiliza processos estocásticos e simula um passeio aleatório, onde temos os possíveis valores dos parâmetros amostrados aleatoriamente. Os parâmetros estimados neste trabalho para os modelos SQUIDER e SEIR foram comparados aos dados reais e aplicadas as métricas de Akaike Information Criterion (AIC) e Bayesian Information Criteria (BIC) para definir o modelo que melhor representa o fenômeno de propagação da Covid-19 no estado do Pará. Como resultados foram obtidos histogramas que indicam uma convergência de parâmetros no modelo SQUIDER, o que não aconteceu com o modelo SEIR. Pela aplicação do AIC e BIC foi demonstrado que o modelo SQUIDER é o modelo que melhor representeou a propagação da Covid-19 no estado do Pará e possui um potencial de ser utilizado como modelo preditivo.

Palavras-chave: COVID-19; Monte Carlo via Cadeia de Markov; Modelos Epidemiológicos.

ABSTRACT

In this study, the Bayesian Monte Carlo technique via Markov Chain (MCMC) was selected to estimate the parameters of the differential equations of the SQUIDER and SEIR compartmental models, seeking to undestand the propagation of Covid-19 in the state of Pará. An algorithm was developed in Matlab, reproducing the MCMC technique that uses stochastic processes and simulates a random selection of values of each parameter. The parameters estimated in this paper for the SQUIDER and SEIR compartmental models were compared to real data using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) metrics. Both of these were applied for a better definition of the model that best represents the Covid-19 propagation phenomenon in the state of Pará. As a result, histograms were obtained that indicate a convergence of parameters in the SQUIDER model, which did not happen in the SEIR model. By applying the AIC and BIC, it was demonstrated that the SQUIDER model was the best model to represent the phenomena, i.e. the propagation of Covid-19 in the state of Pará), and has the potential to be used as a predictive model.

Keywords: COVID-19; Markov Chain Monte Carlo; Epidemiological Models.

¹ Universidade Federal do Pará (UFPA), Belém, Pará, Brasil

^{*}E-mail: dcestumano@ufpa.br

INTRODUÇÃO

No ano de 2019 na província de Wuhan na China foi descoberto uma nova doença causada por um vírus da família dos coronavírus, o SARS-CoV-2. Este vírus que é do tipo corona e recebeu o nome de Covid-19 que significa *Corona Virus Disease 19* (CARCIONE et al, 2019) e em virtude de sua rápida disseminação geográfica alcançou o nível de pandemia.

A Covid-19 apresenta um quadro clínico que varia de infecções assintomáticas a quadros respiratórios graves. De acordo com a Organização Mundial de Saúde (OMS), a maioria dos pacientes com Covid-19 (cerca de 80%) podem ser assintomáticos e cerca de 20% dos casos podem requerer atendimento hospitalar por apresentarem dificuldade respiratória e desses casos aproximadamente 5% podem necessitar de suporte para o tratamento de insuficiência respiratória (suporte ventilatório) (BARRETO et al, 2020). Aproximadamente 1,5% dos infectados vão a óbito.

O Brasil foi um dos países mais afetados pela pandemia de Covid-19, onde até julho de 2021 foram mais de 19 milhões de casos confirmados que causaram mais de 500 mil óbitos. Na região norte do Brasil, especificamente o estado do Pará, no mesmo período foram registrados mais de 550 mil casos de contaminação por Covid-19 e 15 mil mortes (BRASIL, 2021a). Contudo este resultado pode ter sido ainda pior devido a uma dificuldade na mensuração das populações expostas e infectadas.

A mensuração da extensão de uma pandemia pode ser expressa através de um modelo matemático, que pode prever a dinâmica da contaminação, o comportamento da doença e simular ações de prevenção como isolamento e vacinação em massa. (KEELING e ROHANI, 2018). Neste sentido, este trabalho utilizou dois modelos o SEIR e o SQUIDER para analisar os dados e a abrangência da disseminação de Covid-19 no estado do Pará.

Os modelos foram solucionados sob a perspectiva da estatística bayesiana, que possibilita o uso de todo o conhecimento previamente disponível associado aos dados existentes, bem como considerar as incertezas a eles associadas e dessa forma buscar maior aproximação com o cenário real (LEE, 1989). O método de Monte Carlo via Cadeias de Markov (MCMC) foi utilizado para a estimação dos parâmetros de interesse. As métricas bayesianas Akaike (AIC) e o critério de informação bayesiano (BIC), além

do R^2 foram adotados para demonstrar o modelo que apresentou melhor ajuste em descrever a curva de infectados.

MODELO MATEMÁTICO

A complexidade da doença e seus meios de contaminação definem qual modelo matemático seria o mais adequado para representá-la. Muitos estudos para tentar modelar a epidemia Covid-19 usaram modelos compartimentais do tipo SIR (susceptível, infectado, recuperado) ou uma variação desse modelo chamada SEIR que introduz um grupo E (exposto e não infectado) (RUSTAN e ANDAYANI, 2020). Este modelo compartimental pode ser representado conforme a Figura 1.

Figura 1 – Diagrama de blocos do modelo SEIR com os parâmetros a serem estimados



 β = taxa de transmissão ξ = taxa de recontaminação σ = taxa de infecção γ = taxa de recuperação υ = taxa de morte

A relação entre cada grupo da população e cada parâmetro é descrito pelo conjunto de equações diferenciais a seguir:

$$\frac{dS}{dt} = \mu N - \frac{\beta SI}{N} + \xi R - \upsilon S$$
 (Eq. 1)

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E - \nu E$$
 (Eq. 2.)

$$\frac{dI}{dt} = \sigma E - \gamma I - \upsilon I \tag{Eq. 3}$$

$$\frac{\mathrm{dR}}{\mathrm{dt}} = \gamma \mathbf{I} - \xi \mathbf{R} - \upsilon \mathbf{R} \tag{Eq. 4}$$

Devido à complexidade e mortalidade da Covid-19, modelos mais complexos foram desenvolvidos como o SQUIDER que leva em conta também uma população Q (quarentena), U (infectada e não-detectada) e D (óbitos) (KHAN; VAN BUSSEL; HUSSAIN, 2020). Este modelo compartimental é representado na Figura 2.



Figura 2 – Diagrama de blocos do modelo SQUIDER com os parâmetros a serem estimados

Semelhante ao modelo SEIR, o modelo SQUIDER também tem um conjunto de equações diferenciais que mostram a relação entre os grupos da população e os parâmetros que foram calculados neste artigo:

$$\frac{dS}{dt} = -\beta SU\alpha - \lambda S + \rho E + \rho R$$
 Eq. (5)

$$\frac{\mathrm{dQ}}{\mathrm{dt}} = \lambda \mathrm{U} + \lambda \mathrm{S}$$
 Eq. (6)

$$\frac{\mathrm{dI}}{\mathrm{dt}} = \delta \mathbf{U} - \mathbf{I} \boldsymbol{\gamma} + \mathbf{I} \boldsymbol{\alpha}$$
 Eq. (8)

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \mathrm{I}\gamma$$
 Eq. (9)

$$\frac{dE}{dt} = U\varepsilon - \rho E$$
 Eq. (10)

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \mathrm{I}\alpha - \rho \mathrm{R}$$
 Eq. (11)

INFERÊNCIA BAYESIANA

A inferência bayesiana permite obter a análise de determinado sistema em termos de uma distribuição de probabilidade. Esta abordagem possibilita associar todo o conhecimento prévio sobre o fenômeno abordado aos dados experimentais existentes e as incertezas a eles associadas possibilitando maior aproximação do cenário real. O teorema de Bayes mostrado na Equação (12) resume matematicamente este conceito (KAIPIO; SOMERSALO, 2004; NAVEIRA-COTTA et al., 2010; ORLANDE et al., 2012; KHAN et al., 2021)

$$\pi(\mathbf{P}|\mathbf{Y}) = \frac{\pi(\mathbf{P})\pi(\mathbf{Y}|\mathbf{P})}{\pi(\mathbf{Y})}$$
(12)

Sendo $\pi(\mathbf{P})$ a distribuição de probabilidade a priori e nela estão incluídas todas as informações previamente disponíveis a respeito dos parâmetros. $\pi(\mathbf{Y}|\mathbf{P})$ é a verossimilhança, onde os dados experimentais e suas incertezas estão inseridos. $\pi(\mathbf{P}|\mathbf{Y})$ é a distribuição de probabilidade posteriori, que é a distribuição de interesse e $\pi(\mathbf{Y})$ é a distribuição de probabilidade marginal das medidas e tem a função de uma constante normalizadora (BECK; ARNOLD, 1977).

A fim de conhecer informações da posteriori é necessário que diversas amostras desta distribuição sejam obtidas de modo a possibilitar a sua caracterização. Para este fim, o método de amostragem de Monte Carlo via Cadeias de Markov (MCMC) tem sido amplamente adotado na literatura (ESTUMANO et al., 2014; PASQUALETTE et al., 2017; OLIVEIRA et al., 2020; NUNES et al., 2021; MOURA et al., 2021).

3.1. MONTE CARLO VIA CADEIAS DE MARKOV (MCMC)

Monte Carlo via Cadeias de Markov (MCMC) é um método estocástico que repetidamente gera amostras aleatórias que caracterizam a distribuição de probabilidade de um parâmetro de interesse (HAMRA; MACLEHOSE; RICHARDSON, 2013). É um método de amostragem robusto devido a associação das cadeias de Markov e do método de Monte Carlo e frequentemente vem sendo utilizado para estimação de parâmetros em modelos epidemiológicos (WU; LEUNG; LEUNG, 2020) (HAMRA; MACLEHOSE; RICHARDSON, 2013).

A ideia de cadeia no MCMC se desenvolve a partir da estrutura de que cada amostra randômica gerada depende apenas daquela imediatamente anterior e, nesta abordagem sequencial, percorre todo o espaço paramétrico até que uma distribuição estacionária ou de equilíbrio seja alcançada (LUENGO et al., 2020). A implementação do MCMC pode ser executada de maneira simples e eficaz por meio do algoritmo de Metropolis-Hastings (MH), utilizado neste trabalho, a partir das seguintes etapas (GAMERMAN e LOPES, 2006):

- 1. Inicializa-se o contador de iterações i = 1 e estabeleça um valor inicial arbitrário $\mathbf{P}^{(0)}$.
- 2. Gera-se um valor candidato \mathbf{P}^* a partir de uma distribuição auxiliar $q(\mathbf{P}^*|\mathbf{P})$.
- 3. Gera-se um número aleatório u a partir de uma distribuição uniforme, $u \sim U(0,1)$.
- 4. Avalia-se a probabilidade de aceitação $\alpha(\mathbf{P}, \mathbf{P}^*)$ dada por:

$$\alpha(\mathbf{P}, \mathbf{P}^*) = \min\left\{1, \frac{\pi(\mathbf{P}^*|\mathbf{Y})q(\mathbf{P}^*|\mathbf{P})}{\pi(\mathbf{P}|\mathbf{Y})q(\mathbf{P}|\mathbf{P}^*)}\right\}$$
(13)

Se $u \le \alpha(\mathbf{P}, \mathbf{P}^*)$ o movimento é aceito e faz-se $\mathbf{P}^{(i+1)} = \mathbf{P}^*$. Caso contrário, cadeia não se move e faz-se $\mathbf{P}^{(i+1)} = \mathbf{P}^{(i)}$.

 Incrementa o contado i para i + 1. Retorna-se ao passo 2 até que a convergência seja alcançada ou o número de estados definidos para a cadeia seja atingido.

O algoritmo de MH foi implementado neste trabalho utilizando o software Matlab em que as Equações 1–11 foram solucionadas de maneira interativa com base no trabalho de (KHAN; VAN BUSSEL; HUSSAIN, 2020) e os parâmetros, identificados ao lado das Figuras 1 e 2, estimados com MCMC.

Uma vez os parâmetros das equações calculados, é necessária a validação dos valores encontrados, comparando os resultados do modelo com as curvas geradas pelos dados reais disponíveis na literatura (BRASIL, 2021a).

Neste sentido, para efeito comparativo dos dados da literatura com os dados gerados pelo modelo, foi escolhido a variação do grupo de pessoas infectadas em base acumulativa. O método utilizado para comparação foram dois critérios descritos por (EMILIANO et al., 2020), sendo eles o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (BIC).

3.2. CRITÉRIO DE INFORMAÇÃO DE AKAIKE

A seleção de modelos usando o critério de informação de Akaike (Akaike Information Criteria - AIC) (Akaike, 1974), é feita com base em uma equação que considera:

- "p" o número de parâmetros da equação
- "L" valor máximo da função de verossimilhança, que permite inferir sobre o valor dos parâmetros a partir de um conjunto de observações.

O AIC leva em conta que um fenômeno nunca é perfeitamente representado e algumas informações serão perdidas quando de sua modelagem estatística. A relação entre essas duas variáveis (equação 14), mostra que o modelo com um menor valor de AIC é o escolhido como melhor representando o fenômeno em estudo:

$$AIC = 2p - 2\log(L)$$
(Eq. 14)

O critério de informação de Akaike bonifica a qualidade de ajuste (altos valores para a função de verossimilhança) e, por outro lado, penaliza a quantidade de parâmetros do modelo.

3.3. CRITÉRIO BAYESIANO DE SCHWARZ

O Critério Bayesiano de Schwarz (*Bayesian Information Criteria* - BIC) (Schwarz et al., 1978) tem a premissa da existência de um "modelo verdadeiro" que descreva uma relação perfeita entre as variáveis dependentes e independentes do modelo.

Esse critério de mensuração BIC estabelece uma relação entre o número de parâmetros de um modelo (n), o número de variáveis (p) e a função de verossimilhança do modelo selecionado (f):

$$BIC = -2 \log(f) + p \log(n)$$
(Eq. 15)

METODOLOGIA

Foi construído um algoritmo no Matlab usando a versão estudante 2016 deste software. O algoritmo construído foi divido em três partes principais:

- Calcular o mínimo das funções: utilizando a função "*fminsearch*" no Matlab, para calcular o valor da variável independente onde ocorre um mínimo da função com múltiplas variáveis de cada modelo. As equações utilizadas foram as descritas no trabalho de (KHAN; VAN BUSSEL; HUSSAIN, 2020).
- 2. Estimativa dos parâmetros das equações diferenciais dos modelos SEIR e SQUIDER (equações 1 a 11), pelo método de MCMC descrito na seção anterior. Como saída foram gerados histogramas de cada um dos parâmetros para ver seu comportamento, e calculado o R² da curva teórica comparando-as com a curva real de infectados acumulado. Ao fazer uma amostragem proporcional à

probabilidade dos valores, o método tenta chegar numa distribuição de probabilidade que se aproxime dos dados para conseguir ajustar os parâmetros do modelo e convergir para a distribuição estacionária de interesse. Neste trabalho foram utilizados os algoritmos de *Metropolis-Hastings* (TOWERS, 2020) 1000 estados. Como saída obteve-se um histograma com o valor mais provável das variáveis de interesse, e foi utilizada aquela com maior probabilidade para a etapa posterior.

3. Verificação do modelo mais adequado: foi modelado em Matlab os critérios de Akaike e o Bayesiano de Schwartz, para identificar quantitativamente qual o modelo compartimental mais adequado para representar a propagação da Covid-19 no estado do Pará, de acordo com o trabalho de (EMILIANO et al, 2010). Estes critérios foram modelados em Matlab onde para cada um deles obteve-se um valor numérico baseado no modelo específico em estudo e o número de variáveis. Este valor numérico é mais um indicativo de qual modelo matemático mais se adequa aos dados reais do fenômeno em estudo.

A estruturação lógica da metodologia utilizada na programação do código de Matlab é descrita na Figura 3 a seguir:



Figura 3 – Fluxograma da metodologia adotada.

RESULTADOS

Esta sessão traz a comparação entre as curvas real e estimada de infectados por Covid-19 da população no estado do Pará utilizando os modelos SEIR e SQUIDER e o método de MCMC para estimação dos parâmetros. A Figuras 4 mostra a evolução cumulativa do número de infectados reais representada por uma linha preta e número de infectados simulados representado por uma linha tracejada vermelha. Foi possível observar que o modelo SEIR não acompanhou as variações de eventos como queda ou aumento no número de indivíduos infectados, entretanto foi capaz de acompanhar a tendência macro da evolução que foi de cunho ascendente e o alcançou R^2 de 0,9955.

Figura 4 – Infectados Cumulativo do Modelo SEIR no estado do Pará no período de 15 de março de 2020 até 20 de dezembro de 2020.



SEIR: Dados do grupo de "Infectados Cumulativo"

O valor médio estimado para os parâmetros do modelo SEIR obtidos com o método de MCMC estão mostrados na Tabela 1 e a partir dela, foi possível notar que todos os valores obtidos estão dentro do intervalo de credibilidade de 95%.

SEIR	Média	Intervalo de Credibilidade
β	8,3352	8,1336 - 8,8391
γ	3,7861	3,7160 - 3,9613
σ	12,9018	11,5926 - 13,4254
γ	0,0396	0,0360 - 0,0410
ν	0,0241	0,0231 - 0,0245
γc	1,2355	1,2250 - 1,2617

Tabela 1 – Valores para cada parâmetro dos modelos SEIR.

A Figura 5 mostra as curvas real e estimada com o modelo SQUIDER. Em que foi possível simular inclusive as ondas de infectados que ocorreram durante a pandemia, observando que a curva simulada sobrepôs a curva real obtendo R² igual a 1,00. Este comportamento pode ser justificado em virtude da maior complexidade deste modelo e do número de variáveis que ele apresenta, refletindo na representatividade da oscilação

no número de infectados. O valor médio para os parâmetros estimados do modelo SQUIDER com o método MCMC estão dispostos na Tabela 2.

Figura 5 – Infectados Cumulativo do Modelo SQUIDER no estado do Pará no período de 15 de março de 2020 até 20 de dezembro de 2020.

SQUIDER: Dados do grupo de "Infectados Cumulativo"



Tabela 2 – Valores para cada parâmetro do modelo SQUIDER.

SQUIDER	Média	Intervalo de Credibilidade
В	9,6558 x 10 ⁻⁶	2,8056 x 10 ⁻⁶ – 1,610 x 10 ⁻⁵
A	0,0274	0,0162 - 0,5789
Λ	7,0740 x 10 ⁻⁸ – 9,6569 x 10 ⁻⁵	2,21 x 10 ⁻⁸ – 1,2845 x 10 ⁻⁷
Р	0,2992	0,0771 - 1,0895
Е	2,4054 x 10 ⁻⁶	1,4403 x 10 ⁻⁶ – 3,6681 x 10 ⁻⁶
δ	2,9110 x 10 ⁻⁵	1,2468 x 10 ⁻⁵ – 7,7189 x 10 ⁻⁵
γ	7,9417 x 10 ⁻⁵	3,0587 x 10 ⁻⁵ – 1,821 x 10 ⁻⁴

Ainda de acordo com os dados de saída do código de Matlab para os modelos SEIR e SQUIDER, foram plotados os histogramas da distribuição de probabilidade posterior dos parâmetros utilizados. A Figura 6 mostra como exemplo, os histogramas encontrados para o parâmetro β (taxa de transmissão), para os modelos SEIR e SQUIDER.







No modelo SQUIDER, a variação dos parâmetros conforme os histogramas se aproximam mais de uma distribuição normal, que reflete um comportamento de um fenômeno natural. Em probabilidade e estatística, a distribuição normal é uma das distribuições de probabilidade mais utilizadas para modelar fenômenos naturais. Isso se deve ao fato de que um grande número de fenômenos naturais apresenta sua distribuição de probabilidade tão proximamente normal, que a ela pode ser com sucesso referida, e, portanto, com adequado acerto por ela representada como se normal fosse (BITTENCOURT & VIALI, 2006).

Já o histograma obtido pelo modelo SEIR, mostra uma distribuição em apenas duas faixas para todos os parâmetros, indicando um fenômeno não-natural refletindo na não capacidade desse modelo em refletir adequadamente o comportamento da doença em estudo.

A escolha de um modelo que consiga simular dados reais da forma mais fidedigna possível é de extrema importância para que previsões mais assertivas consigam ser realizadas sobre o problema estudado (EMILIANO, 2009). Por isso, o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (BIC), foram utilizados.

Dessa forma, a seleção de modelos realizada neste trabalho foi baseada na teoria da informação, onde mostra a perda de informação ao fazer aproximações de dados reais com um modelo, tanto o método AIC quanto o BIC mostram essa perda de informação, onde o melhor modelo será aquele que consegue aproximar os dados reais dos simulados com uma menor perda de informação (EMILIANO et al., 2010).

Os critérios de mensuração de qualidade dos modelos SEIR e SQUIDER estão mostrados na Tabela 3 (AIC e BIC) apontaram o modelo SQUIDER como o mais indicado para estimação da curva de infectados por Covid-19 no estado do Pará.

Tabela 3 – Valores de taxa de aceitação, e dos critérios de mensuração de qualidade				
para os modelos SEIR e SQUIDER.				

	Modelo SEIR	Modelo SQUIDER
Taxa de aceitação do MCMC	1%	23%
Critério de informação de Akaike (AIC)	-4,0777	-12,2215
Critério Bayesiano de Schwarz (BIC)	15.7122	10,8667

Analisando o critério AIC para mensuração da qualidade dos modelos matemáticos, obteve-se o valor de -4,0777 para o modelo SEIR e de -12,2215 para o modelo SQUIDER. A premissa é de quanto menor o valor, mais adequado é o modelo em comparação com os dados reais do estudo. Sendo assim, mais uma vez o modelo SQUIDER representou melhor a curva de contaminação. O menor valor encontrado para o modelo SQUIDER é devido à relação matemática do resultado numérico desse modelo e a função verossimilhança, expressos na Equação 14. Quando maior a semelhança entre os valores de saída do modelo e os dados reais, mais negativo é este índice e mais representativo do fenômeno em estudo.

Em relação ao critério bayesiano de Schwarz, quanto menor o resultado encontrado, mais adequado é o modelo. Apesar de este critério levar em conta que quanto mais parâmetros a serem estimados, pior o modelo; existe também um peso grande em função da equação representativa do modelo em si (Equação 15). Os valores encontrados usando o critério BIC foram 15,7192 para o modelo SEIR e 10,8667 para o modelo SQUIDER reforçam a constatação de que o modelo SQUIDER se mostrou mais adequado que o modelo SEIR para representar a propagação da Covid-19 no estado do Pará.

A Tabela 3 mostra a taxa de aceitação da estimativa de parâmetros obtida como resultado da aplicação do Método de Monte Carlo via Cadeias de Markov calculada adotando um intervalo de credibilidade de 95%. Os resultados obtidos pela taxa de aceitação para o modelo SQUIDER foi de 23%, significativamente maior que a taxa de aceitação para o modelo SEIR que foi de 1%. O significado matemático da taxa de aceitação é que pelo passeio aleatório do número de variáveis aceitas no modelo SQUIDER foi 23 vezes maior que no modelo SEIR quando comparados aos dados reais. Novamente, mais um indicativo que o modelo matemático SQUIDER é superior ao modelo SEIR, na representação do fenômeno de propagação da Covid-19.

SQUIDER	Valor encontrado nesse estudo	Média valores encontrados por (KHAN; VAN BUSSEL; HUSSAIN, 2020) nos EUA
В	9,6558 x 10 ⁻⁶	0,7510
А	0,0274	0,9937
q(t)	7,0740 x 10 ⁻⁸ – 9,6569 x 10 ⁻⁵	0,1423
Р	0,2992	0,0213
E	2,4054 x 10 ⁻⁶	0,1546
Δ	2,9110 x 10 ⁻⁵	0,5134
Г	7,9417 x 10 ⁻⁵	0,0263

Tabela 4 – Comparativo dos valores dos parâmetros encontrados nesse estudo com a literatura.

Outra análise importante foi que alguns parâmetros encontrados para o modelo SQUIDER neste estudo para o estado do Pará são significativamente diferentes dos encontrados na literatura. De acordo com (KHAN; VAN BUSSEL; HUSSAIN, 2020), o modelo SQUIDER foi utilizado para modelar a propagação da Covid-19 em alguns estados norte-americanos e foram encontrados com uma variação dos parâmetros entre 0,02 e 1,00. A Tabela 4 mostra que cinco dos sete parâmetros calculados neste artigo são da ordem de 10^{-5} , (parâmetros β , λ , ε , $\delta \in \gamma$).

CONCLUSÃO

Em resumo, com estre trabalho conclui-se que o modelo SQUIDER é mais representativo do fenômeno em estudo do que o modelo SEIR. Ao analisar o R², as curvas plotadas mostram que o modelo SQUIDER (R-quadrado de 1,000) se moldam com mais exatidão aos dados reais do que as curvas do modelo SEIR (R-quadrado 0,9955).

Os histogramas obtidos na saída de estimativa de parâmetros por MCMC para modelo SQUIDER se aproximam de uma curva normal, indicando a ocorrência de um fenômeno natural. Já os histogramas do modelo SEIR apresentaram variando em somente duas faixas.

Por fim, os critérios de mensuração de qualidade pelos métodos AIC e BIC foram menores para o modelo SQUIDER do que para o modelo SEIR, indicando que este primeiro modelo melhor reflete o fenômeno de propagação da Covid-19 no estado do Pará.

REFERÊNCIAS

AKAIKE, Hirotugu. A new look at the statistical model identification. **IEEE transactions on automatic control**, v. 19, n. 6, p. 716-723, 1974.

BARRETO, M et al. O que é urgente e necessário para subsidiar as políticas de enfrentamento da pandemia de COVID-19 no Brasil? **Revista Brasileira de Epidemiologia**, V .23, 2020. Disponível em: https://doi.org/10.1590/1980-549720200032>. Acesso em 15 de outubro de 2021

BECK, James Vere; ARNOLD, Kenneth J. **Parameter estimation in engineering and science**. James Beck, 1977.

BITTENCOURT, H. R.; VIALI, L. Contribuições para o ensino da distribuição normal ou curva de Gauss em cursos de graduação. In: **III Seminário Internacional de Pesquisa em Educação Matemática**, 2006, Águas de Lindóia. Anais do III Seminário Internacional de Pesquisa em Educação Matemática. Brasília : SBEM, 2006.

BRASIL, MINISTÉRIO DA SAÚDE. **Boletim epidemiológico especial - Doença pelo Coronavírus COVID-19** nº 71, 2021a. Secretaria de Vigilância em Saúde. Disponível em https://www.gov.br/saude/pt-br/media/pdf/2021/julho/16/boletim_epidemiologico _covid_71.pdf.. Acesso em 20 de julho de 2021.

BRASIL, MINISTÉRIO DA SAÚDE. **Painel Coronavírus . Secretaria de Vigilância em Saúde.** Disponível em https://covid.saude.gov.br/. Acesso em 10 de outubro de 2021.

CARCIONE, J. M. et al. A Simulation of a COVID-19 Epidemic Based on a Deterministic SEIR Model. **Frontiers in Public Health**, v. 8, n. May, 2020.

EMILIANO, P. C. Fundamentos e aplicações dos critérios de Informação: Akaike e **Bayesiano**. Ufla, p. 105, 2009.

EMILIANO, P. C. et al. Critérios de informação de Akaike versus Bayesiano: análise comparativa. (Poster session). 19º Simpósio Nacional de Probabilidade e Estatística, p. 6, 2010.

ESTUMANO, Diego; HAMILTON, F.C.; COLACO, Marcelo ; LEIROZ, Albino ; ORLANDE, Helcio ; CARVALHO, Rogério; DULIKRAVICH, George. (2014). Bayesian estimate of mass fraction of burned fuel in internal combustion engines using pressure measurements. **Engineering Optimization IV - Proceedings of the 4th International Conference on Engineering Optimization, ENGOPT** 2014. 997-1004. 10.1201/b17488-181.

GAMERMAN, D.; LOPES, H. F. Markov chain Monte Carlo: stochastic simulation for **Bayesian inference**. CRC Press, 2006.

HAMRA, G; MACLEHOSE, R; RICHARDSON, D. Markov Chain Monte Carlo: an introduction for epidemiologists. **International Journal of Epidemiology**, v.42, p.627-634, 2013.

KEELING, M; ROHANI, P. "What are mathematical models?". Modeling Infectious Diseases in Humans and Animals, 2nd ed., chapter 1.4, New Jersey, USA, Princeton University Press, 2008.

KAIPIO, Jari; SOMERSALO, Erkki. **Statistical and computational inverse problems**. Springer Science & Business Media, 2006.

KHAN, Z. S.; VAN BUSSEL, F.; HUSSAIN, F. A predictive model for COVID-19 spread – with application to eight US states and how to end the pandemic. **Epidemiology and Infection**, 2020.

KHAN, Aminul Islam et al. Bayesian Method for Parameter Estimation in Transient Heat Transfer Problem. **International Journal of Heat and Mass Transfer**, v. 166, p. 120746, 2021.

LEE, Peter M. Bayesian statistics. London:: Oxford University Press, 1989.

LUENGO, D., MARTINO, L., BUGALLO, M., ELVIRA, V., SÄRKKÄ, S. A survey of Monte Carlo methods for parameter estimation. **EURASIP Journal on Advances in Signal Processing**, v. 2020, p. 1-62, 2020.

MOURA, Carlos HR et al. Parameter Estimation in Population Balance through Bayesian Technique Markov Chain Monte Carlo. **Journal of Applied and Computational Mechanics**, v. 7, n. 2, p. 890-901, 2021.

NAVEIRA-COTTA, Carolina P.; COTTA, Renato M.; ORLANDE, Helcio RB. Inverse analysis of forced convection in micro-channels with slip flow via integral transforms and Bayesian inference. **International journal of thermal sciences**, v. 49, n. 6, p. 879-888, 2010.

NUNES, Keila Guerra Pacheco et al. Evaluation of zinc adsorption through batch and continuous scale applying Bayesian technique for estimate parameters and select model. **Journal of Environmental Science and Health, Part A**, p. 1-15, 2021.

OLIVEIRA, Renata Farias et al. Cr (VI) adsorption in batch and continuous scale: A mathematical and experimental approach for operational parameters prediction. **Environmental Technology** & **Innovation**, v. 20, p. 101092, 2020.

ORLANDE, Helcio RB. Inverse problems in heat transfer: new trends on solution methodologies and applications. **Journal of Heat Transfer**, v. 134, n. 3, 2012.

PASQUALETTE, Marcelo A. et al. Bayesian estimate of pre-mixed and diffusive rate of heat release phases in marine diesel engines. Journal of the Brazilian Society of Mechanical Sciences and Engineering, v. 39, n. 5, p. 1835-1844, 2017.

RUSTAN, R.; HANDAYANI, L. the Outbreak'S Modeling of Coronavirus (Covid-19) Using the Modified Seir Model in Indonesia. **Spektra: Jurnal Fisika dan Aplikasinya**, v. 5, n. 1, p. 61–68, 2020.

SCHWARZ, Gideon. Estimating the dimension of a model. **The annals of statistics**, p. 461-464, 1978.

TOWERS, S. "**Markov Chain Monte Carlo parameter optimization method**", 2014. Disponível em: <sherrytowers.com/2014/07/15/ markov-chain-monte-carlo-parameter-optimization-method/>. Acessado em: 4 de agosto de 2020.

WANG, K. et al. Current trends and future prediction of novel coronavirus disease (COVID-19) epidemic in China: A dynamical modeling analysis. Mathematical Biosciences and Engineering, v. 17, n. 4, p. 3052–3061, 2020.

WHO, World Health Organization. Glossary of humanitarian terms. Disponível em: <<u>https://www.who.int/hac/about/definitions/en/#:~:text=Checchi></u>. Acesso em 30 de julho de 2020.

WU, T; LEUNG K; LEUNG, G. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. **Lancet**, v.395, p.689-697, 2020.

Recebido em: 20/12/2021 Aprovado em: 21/01/2022 Publicado em: 25/01/2022