



08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

MODELAGEM DE DADOS VIA DISTRIBUIÇÕES DE VALORES EXTREMOS E COMPUTAÇÃO BAYESIANA APROXIMADA - ABC

Andrew Nery da Silva Custódio¹ - a.nery1997@gmail.com

José Mir Justino da Costa¹ - zemir@ufam.edu.br

Wellington Betencurte da Silva² - wellingtonufes@gmail.com

Julio César Sampaio Dutra² - juliosdutra@yahoo.com.br

Alexandre Toman³ - alexandre.toman@cefet-rj.br

¹Universidade Federal do Amazonas - Manaus, AM, Brazil

²Universidade Federal do Espírito Santo - Alegre, ES, Brazil

³Centro Federal de Educação Tecnológica Celso Suckow da Fonseca - CEFET/RJ, Brazil

Resumo. *Enchentes são fenômenos extremos e por isso a modelagem por distribuições de valores extremos é uma forma razoável de abordar tais dados para suas análises de comportamento. O presente trabalho tem como finalidade ajustar dados de enchentes no Rio Negro, na cidade de Manaus - AM, durante o período de 2000 a 2015, com as distribuições de valores extremos Weibull, Gumbel e Fréchet, estimando os parâmetros utilizando computação bayesiana aproximada (ABC). Especificamente, é aplicado um método melhorado do algoritmo ABC original, chamado ABC-SMC, no qual é usado para a estimação de parâmetros das distribuições propostas. Além disso, o ABC-SMC também foi utilizado para seleção de modelos.*

Keywords: *Distribuições de valores extremos, Algoritmo ABC, Estatística Bayesiana, Seleção de modelos*

1. INTRODUÇÃO

Como a teoria de valores extremos tem sido estudada e aplicada em vários campos da ciência que envolvem máximos e mínimos, resolvemos fazer uma aplicação desse estudo em dados de enchente do Rio Negro, na cidade de Manaus, ajustando as distribuições de valores extremos aos dados de enchente anual da série histórica de 2000-2015. As distribuições Weibull, Gumbel e Fréchet são os modelos abordados e através de um dos métodos da Computação Bayesiana Aproximada - ABC, o ABC-SMC, estimamos os parâmetros de cada distribuição.

Em se tratando de três possíveis modelos para ajustar tais dados, o algoritmo ABC-SMC foi utilizado com a finalidade de verificar qual modelo se mostra mais adequados para representar a física do problema e como consequência disso estimar seus referidos parâmetros.

Os resultados obtidos usando as distribuições de valores extremos e Computação Bayesiana Aproximada, mostram que trata-se de um ferramenta interessante, e que apesar do algoritmo

ABC-SMC ter mostrado apenas evidências em favor de uma dada distribuição o mesmo pode ser usado para tais fins.

2. METODOLOGIA

Será apresentado um breve estudo sobre cada distribuição a fim de conhecer seus principais aspectos e, posteriormente, discutido o algoritmo ABC e o ABC-SMC.

2.1 Distribuições de valores extremos

De acordo com Kotz & Nadarajah (2000), temos as seguintes distribuições de valores extremos

Distribuição Weibull. A distribuição Weibull é uma distribuição de probabilidade contínua. Se X é uma variável aleatória que assume distribuição Weibull com parâmetros α e β , então sua função densidade de probabilidade é dada por

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0, \\ 0 & x < 0 \end{cases} \quad (1)$$

Na Eq. (1), $\alpha \in (0, +\infty)$ é o parâmetro de forma e $\beta \in (0, +\infty)$ é o parâmetro de escala.

Distribuição Gumbel A distribuição Gumbel é uma distribuição de probabilidade contínua. Se X é uma variável aleatória que assume distribuição Gumbel com parâmetros μ e σ , então sua função densidade de probabilidade é dada por

$$f(x; \mu, \sigma) = \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\} \quad x \in \mathbb{R} \quad (2)$$

Na Eq. (2), $\mu \in \mathbb{R}$ é o parâmetro de locação e $\sigma \in (0, +\infty)$ é o parâmetro de escala.

Distribuição Fréchet A distribuição Fréchet é uma distribuição de probabilidade contínua. Se X é uma variável aleatória que assume distribuição Gumbel com parâmetros α , β e s , então sua função densidade de probabilidade é dada por

$$f(x; \alpha, \beta, s) = \frac{s}{\beta} \left(\frac{x-\alpha}{\beta}\right)^{-1-s} \exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{-\alpha}\right) \quad x > \alpha \quad (3)$$

Na Eq. (3), $\alpha \in \mathbb{R}$ é o parâmetro de locação do mínimo, $\beta \in (0, +\infty)$ é o parâmetro de escala e $s \in (0, +\infty)$ é o parâmetro de forma.

Algoritmo 1: Algoritmo ABC-SMC (Toni & Stumpf (2010))

- 1 Determine $\epsilon_1, \dots, \epsilon_T$.
- 2 Faça o indicador da população $t = 0$.
- 3 Se $t = 0$, gere θ^{**} de $\pi(\theta)$ de forma independente.
 Senão, selecione aleatoriamente um θ_* da população anterior, $\{\theta_{t-1}^{(i)}\}$, com pesos w_{t-1} e perturbe a partícula para obter $\theta^{**} \sim K_t(\theta|\theta^*)$, onde K_t é um kernel de perturbação.
 Se $\pi(\theta^{**}) = 0$, reinicie o passo 3.
 Simule uma amostra candidata $x^* \sim f(x|\theta^{**})$.
 Se $d(x_0, x^*) \geq \epsilon_t$, reinicie o passo 3.
- 4 Faça $\theta_t^{(i)} = \theta^{**}$ e calcule o peso para a partícula $\theta_t^{(i)}$,

$$w_t^{(i)} = \begin{cases} 1, & \text{se } t = 0, \\ \frac{\pi(\theta_t^{(i)})}{\sum_{j=1}^n w_{t-1}^{(j)} K_t(\theta_{t-1}^{(j)}, \theta_t^{(i)})}, & \text{se } t > 0. \end{cases}$$

- Se $i < N$, faça $i = i + 1$, volte para o passo 3.
- 5 Normalize os pesos.
 Se $t < T$, faça $t = t + 1$, volte para o passo 2.
-

Algoritmo 2: Algoritmo ABC-SMC para seleção de modelo (Toni & Stumpf (2010))

- 1 Determine $\epsilon_1, \dots, \epsilon_T$. Faça o indicador da população $t = 0$.
- 2 Faça o indicador da partícula $i = 1$.
- 3 Gere m^* de $\pi(m)$.
 Se $t = 0$, gere θ^{**} de $\pi(\theta(m^*))$.
 Senão, selecione aleatoriamente um θ_* da população anterior, $\{\theta(m^*)_{t-1}\}$, com pesos $w(m^*)_{t-1}$. Perturbe a partícula para obter $\theta^{**} \sim K_t(\theta|\theta^*)$, onde K_t é um kernel de perturbação.
 Se $\pi(\theta^{**}) = 0$, reinicie o passo 3.
 Simule uma amostra candidata $x^* \sim f(x|\theta^{**}, m^*)$.
 Se $d(x_0, x^*) \geq \epsilon_t$, reinicie o passo 3.
- 4 Faça $m_t^{(i)} = m^*$ e adicione θ^{**} para a população de partículas $\{\theta(m^*)_t\}$ e calcule seu peso como,

$$w_t^{(i)} = \begin{cases} 1, & \text{se } t = 0, \\ \frac{\pi(\theta_t^{(i)})}{\sum_{j=1}^n w_{t-1}^{(j)} K_t(\theta_{t-1}^{(j)}, \theta_t^{(i)})}, & \text{se } t > 0. \end{cases}$$

- Se $i < N$, faça $i = i + 1$, volte para o passo 3.
- 5 Para cada m , normalize os pesos.
 Se $t < T$, faça $t = t + 1$, volte para o passo 2.
-

2.2 Métodos de computação bayesiana aproximada (ABC)

Métodos de Computação Bayesiana Aproximada - ABC visam aproximar a distribuição a posteriori sem fazer uso da verossimilhança. Neste trabalho, apresentamos a versão ABC-SMC proposta por Toni & Stumpf (2010).

No ABC-SMC, um número de valores de parâmetros amostrados (chamados de partículas), $\{\theta^{(1)}, \dots, \theta^{(N)}\}$, vindos de uma priori $\pi(\theta)$, são propagados em uma sequência de distribuições intermediárias, $\pi(\theta|d(x_0, x^*) \leq \epsilon_i)$, $i = 1, \dots, T - 1$, até que isso represente uma amostra da *distribuição alvo*, $\pi(\theta|d(x_0, x^*) \leq \epsilon_T)$. As tolerâncias são dadas tais que $\epsilon_1 > \epsilon_2 > \dots > \epsilon_T \geq 0$, ou seja, as distribuições gradualmente vão chegando mais próximas da posteriori aproximada. Posto isto, o Algoritmo 1 é proposto por Sisson et al. (2008) e Moral et al. (2006).

Partículas amostradas da população anterior são denotadas por um único asterisco, e após a perturbação se tornam partículas com dois asteriscos. É possível escolher um passeio aleatório (Gaussiano ou Uniforme) para K_t .

Nosso objetivo também é selecionar um modelo que é melhor ajustado aos dados. Utilizamos o Algoritmo 2 adaptado para seleção de modelos proposto por Toni & Stumpf (2010) e Costa et al. (2017).

Seja $m \in \{1, \dots, M\}$, onde M é o número de modelos, e denotando os parâmetros do modelo m como $\theta(m) = (\theta(m)^{(1)}, \theta(m)^{(2)}, \dots, \theta(m)^{(k_m)})$, onde k_m é o número de parâmetros no modelo m . Proposto por Toni & Stumpf (2010).

Digamos que possuímos dois modelos m_1 e m_2 e queremos saber qual modelo explica melhor os dados. Então, vamos utilizar o fator de Bayes para esse propósito e Toni & Stumpf (2010) definiu como

$$B_{12} = \frac{P(m_1|x)/P(m_2|x)}{P(m_1)/P(m_2)} \quad (4)$$

onde, $P(m_i)$ é a priori e $P(m_i|x)$ é a posteriori marginal do modelo m_i , $i = 1, 2$. Com priori uniforme, teremos

$$B_{12} = P(m_1|x)/P(m_2|x) \quad (5)$$

A seguinte tabela nos mostra como interpretar o fator de Bayes

Tabela 1- Interpretação do fator de Bayes B_{12}

valor do fator de Bayes B_{12}	evidência a favor do modelo m_1
<1	negativa
1 - 3	pobre
3 - 20	substancial
20 - 150	positiva
>150	decisiva

Portanto, podemos verificar qual modelo melhor se ajusta aos dados, já que o fator de Bayes é um resumo da evidência dos dados em favor de um modelo em relação ao outro.

3. ESTUDO DAS ENCHENTES DO RIO NEGRO (2000 - 2015)

O objetivo principal desse trabalho é modelar dados de enchente (máxima) do Rio Negro, na cidade de Manaus. Portanto, utilizamos ABC-SMC para ajustar os dados com as distribuições Weibull, Gumbel e Fréchet estimando o vetor de parâmetros θ em cada distribuição. A tabela 2 mostra a máxima das enchentes ocorridas de 2000 a 2015.

Tabela 2- Tabela de níveis máximos de enchente no Rio Negro com seus respectivos anos.

Ano	2000	2001	2002	2003	2004	2005	2006	2007
Nível Máximo	28,62	28,21	28,91	28,27	27,13	28,10	28,84	28,18
Ano	2008	2009	2010	2011	2012	2013	2014	2015
Nível Máximo	28,62	29,77	27,96	28,62	29,97	29,33	29,50	29,66

O algoritmo (ABC-SMC) foi implementado no software R, um software gratuito de código livre mantido pela *R Foundation for Statistical Computing*.

3.1 Inferência através do ABC-SMC

Na seção de metodologia, mostramos o passo-a-passo da estimação da distribuição à posteriori através do algoritmo ABC-SMC. Porém, melhorias e extensões se fizeram necessárias, tais como um esquema de tolerâncias adaptativas, preservando a estrutura fundamental do algoritmo.

As principais modificações entre o algoritmo ABC-SMC proposto por Toni & Stumpf (2010) e a versão otimizada para o nosso caso são dados a seguir.

- Na versão original precisamos determinar um conjunto de tolerâncias $\epsilon_1, \dots, \epsilon_T$. Na versão otimizada fizemos essa tolerância de forma adaptativa de forma que na população i , a tolerância será dada pela média das distâncias $d(x_0, x^*)$ das partículas aceitas na população anterior.
- Se fizermos a substituição anterior, precisaremos de uma condição de parada já que não teremos uma quantidade específica de populações. Essa condição de parada é definida como a diferença entre a média das distâncias aceitas da população atual e a média das distâncias aceitas da população anterior menor que certo ϵ .

Aplicação na distribuição Weibull Primeiramente, escolhemos como priori para os parâmetros α e β uma distribuição uniforme, onde $\alpha, \beta \sim U(0, 50)$, e o kernel de perturbação K_t , também, dado por uma uniforme, onde para α , $K_t \sim U(-m1, m1)$, e para β , $K_t \sim U(-m2, m2)$,

$$m1 = \frac{1}{2} \left(\max \left\{ \alpha_t^{(1)}, \dots, \alpha_t^{(N)} \right\} - \min \left\{ \alpha_t^{(1)}, \dots, \alpha_t^{(N)} \right\} \right)$$

$$m2 = \frac{1}{2} \left(\max \left\{ \beta_t^{(1)}, \dots, \beta_t^{(N)} \right\} - \min \left\{ \beta_t^{(1)}, \dots, \beta_t^{(N)} \right\} \right)$$

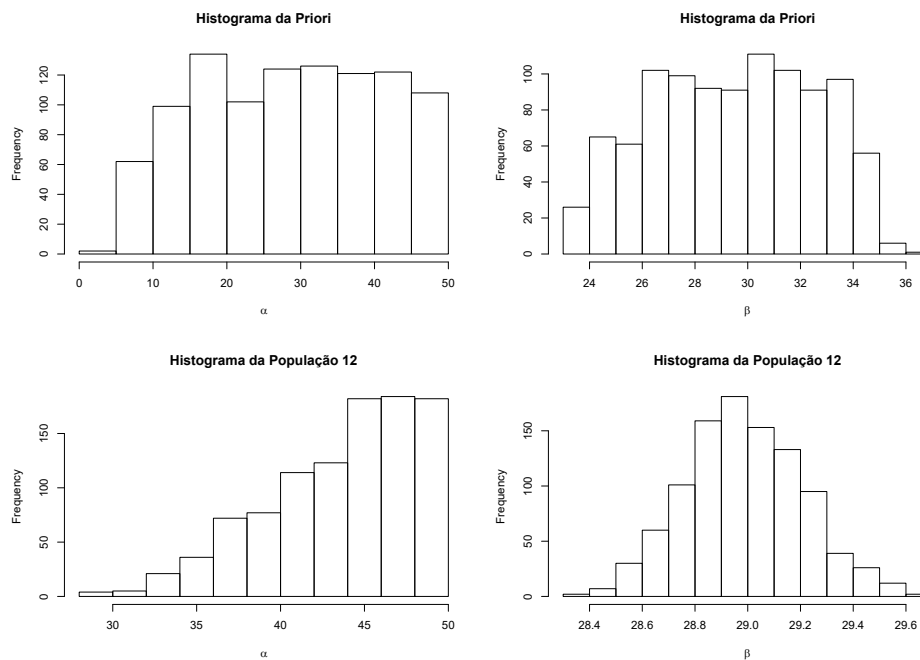


Figura 1- Histogramas das partículas selecionadas pela população priori e a população 12.

O número de partículas para cada população foi definido em $N = 1000$ e a tolerância para população priori como $e_0 = 500$.

Na Fig. 1, temos os histogramas das partículas selecionadas tanto da população priori quanto da população 12, no caso, a última população com partículas aceitas. Observa-se que as estimativas na última população são melhores quando comparadas com a população *a priori*. A tolerância usada na população 12, como definido anteriormente foi 5, 51. Algo notório é que as partículas selecionadas em relação a α , na população 12, não são tão concentradas quanto as de β , o que exige uma investigação futura. As estimativas, baseadas na média das partículas selecionadas na população 12 são $\hat{\alpha} = 43,61$ e $\hat{\beta} = 28,98$.

Aplicação na distribuição Gumbel Escolhemos como priori para μ e σ uma uniforme, onde $\mu, \sigma \sim U(0, 30)$, e o kernel de perturbação K_t é, também, dado por uma uniforme, onde para μ , $K_t \sim U(-m1, m1)$, e para σ , $K_t \sim U(-m2, m2)$,

$$m1 = \frac{1}{2} \left(\max \left\{ \mu_t^{(1)}, \dots, \mu_t^{(N)} \right\} - \min \left\{ \mu_t^{(1)}, \dots, \mu_t^{(N)} \right\} \right)$$

$$m2 = \frac{1}{2} \left(\max \left\{ \sigma_t^{(1)}, \dots, \sigma_t^{(N)} \right\} - \min \left\{ \sigma_t^{(1)}, \dots, \sigma_t^{(N)} \right\} \right)$$

Na Fig. 2, o histograma da população 13 mostra um conjunto de partículas mais concentradas para ambos parâmetros. A tolerância para população 13 é 5, 33. Desta forma, as estimativas, baseadas na média das partículas selecionadas na população 13 são $\hat{\mu} = 28,5$ e $\hat{\sigma} = 0,41$.

Aplicação na distribuição Fréchet Escolhemos como priori para α , β e s uma uniforme, onde $\alpha, \beta \sim U(0, 50)$, $s \sim U(-10, 10)$, e o kernel de perturbação K_t é, também, dado por uma uniforme, onde para α , $K_t \sim U(-m1, m1)$, para β , $K_t \sim U(-m2, m2)$ e para s , $K_t \sim$

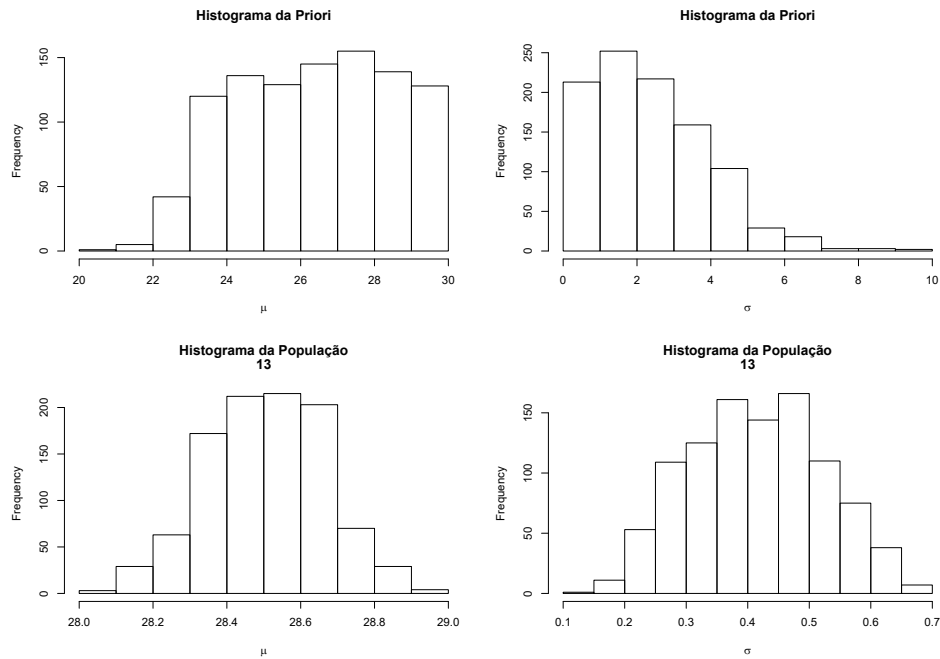


Figura 2- Histogramas das partículas selecionadas pela população priori e a população 13.

$U(-m3, m3),$

$$m1 = \frac{1}{2} \left(\max \left\{ \alpha_t^{(1)}, \dots, \alpha_t^{(N)} \right\} - \min \left\{ \alpha_t^{(1)}, \dots, \alpha_t^{(N)} \right\} \right)$$

$$m2 = \frac{1}{2} \left(\max \left\{ \beta_t^{(1)}, \dots, \beta_t^{(N)} \right\} - \min \left\{ \beta_t^{(1)}, \dots, \beta_t^{(N)} \right\} \right)$$

$$m3 = \frac{1}{2} \left(\max \left\{ s_t^{(1)}, \dots, s_t^{(N)} \right\} - \min \left\{ s_t^{(1)}, \dots, s_t^{(N)} \right\} \right)$$

Assim como no caso da distribuição Weibull, a Fig. 3 mostra que as partículas para população 12 na distribuição Fréchet não são tão concentradas, porém, aqui, é em relação a todos os parâmetros. A tolerância da população 12 foi 5.57. Portanto, as estimativas, baseadas na média das partículas selecionadas na população 12 são $\hat{\alpha} = 4,55$, $\hat{\beta} = 23,92$ e $\hat{s} = 42,51$.

3.2 Seleção de modelos através do ABC-SMC

Aplicamos o algoritmo ABC-SMC para seleção de modelos em relação às três distribuições propostas, Fréchet, Gumbel e Weibull, representadas, respectivamente, através do vetor $m \in \{1, 2, 3\}$, e seus respectivos vetores de parâmetros $\theta(m)$: $\theta(1) = (\alpha, \beta, s)$; $\theta(2) = (\mu, \sigma)$ e $\theta(3) = (\alpha, \beta)$.

Na população à priori escolhemos 10000 partículas distribuídas de forma uniforme entre os três modelos. Estamos usando duas versões do algoritmo, uma que aceita nas populações intermediárias 2000 partículas, e outra que aceita 3000 partículas, vamos chamar, respectivamente, algoritmo A e algoritmo B.

A tolerância inicial para ambos casos foi $\varepsilon^{(0)} = 500$ e após 21 populações intermediárias a tolerância final foi de $\varepsilon_A^{(21)} = 2.6332$ e $\varepsilon_B^{(21)} = 2.6144$. O fator de Bayes foi utilizado para

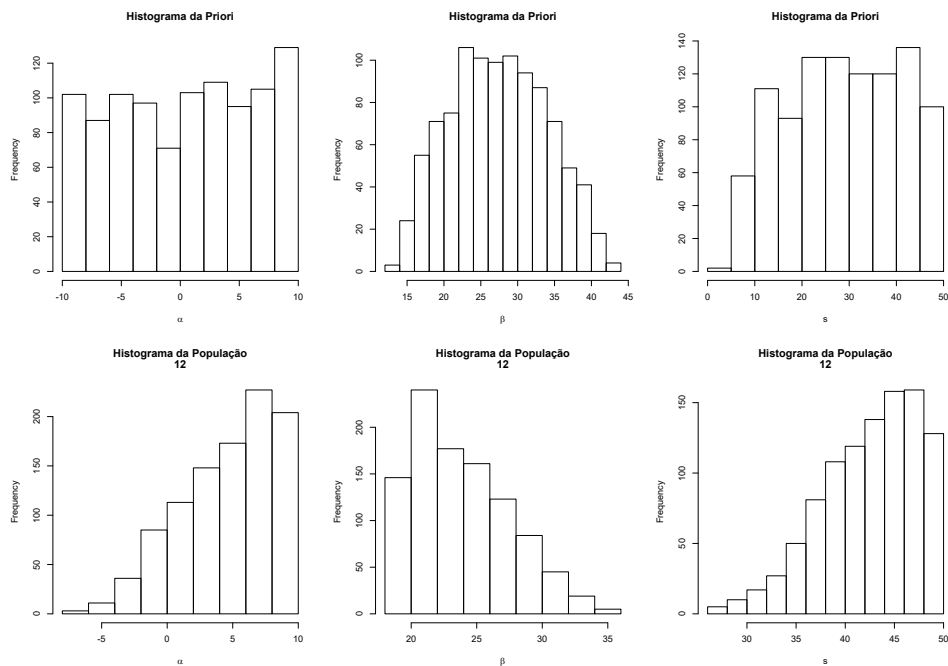


Figura 3- Histogramas das partículas selecionadas pela população priori e a população 12.

avaliarmos as evidências entre os modelos concorrentes. No caso do algoritmo A, das 2000 partículas aceitas na população final, como pode ser observado na Fig. 4, o modelo m_1 (Frechét) foi selecionado 148 vezes, o modelo m_2 (Gumbel) foi selecionado 816 vezes e o modelo m_3 (Weibull) foi selecionado 1036 vezes. Portanto, podemos concluir a partir dos fatores de Bayes,

$$B_{12} = \frac{148}{816} = 0,19; \quad B_{21} = \frac{816}{148} = 5,51; \quad B_{13} = \frac{148}{1036} = 0,14;$$

$$B_{31} = \frac{1036}{148} = 7; \quad B_{32} = \frac{1036}{816} = 1,28; \quad B_{23} = \frac{816}{1036} = 0,78$$

que existe uma evidência negativa do modelo m_1 em relação ao outros e evidência pobre do modelo m_3 em relação ao modelo m_2 , apenas quando comparamos os modelos m_2 e m_3 com o modelo m_1 é que obtemos uma evidência substancial.

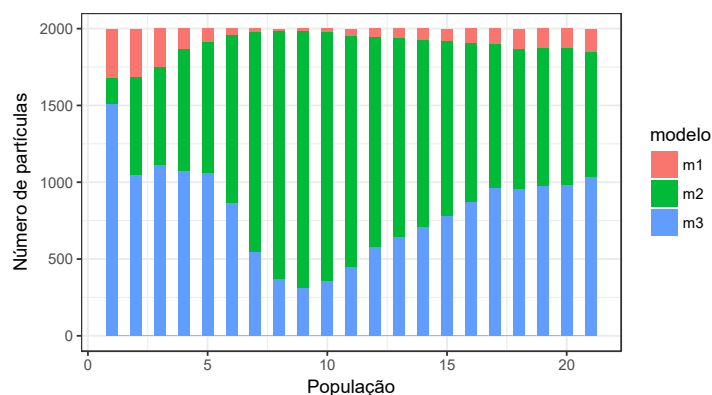


Figura 4- Histograma da distribuição das partículas para os modelos em relação a cada população.(algoritmo A (2000 partículas))

No caso do algoritmo B, das 3000 partículas aceitas na população final, como pode ser observado na Fig. 5, o modelo m_1 (Frechét) foi selecionado 136 vezes, o modelo m_2 (Gumbel) foi selecionado 1322 vezes e o modelo m_3 (Weibull) foi selecionado 1542 vezes. Portanto, podemos concluir a partir dos fatores de Bayes,

$$B_{12} = \frac{136}{1322} = 0,10; \quad B_{21} = \frac{1322}{136} = 9,72; \quad B_{13} = \frac{136}{1542} = 0,089;$$

$$B_{31} = \frac{1542}{136} = 11,33; \quad B_{32} = \frac{1542}{1322} = 1,16; \quad B_{23} = \frac{1322}{1542} = 0,85$$

que existe uma evidência negativa do modelo m_1 em relação ao outros e evidência pobre do modelo m_3 em relação ao modelo m_2 , apenas quando comparamos os modelos m_2 e m_3 com o modelo m_1 é que obtemos uma evidência substancial.

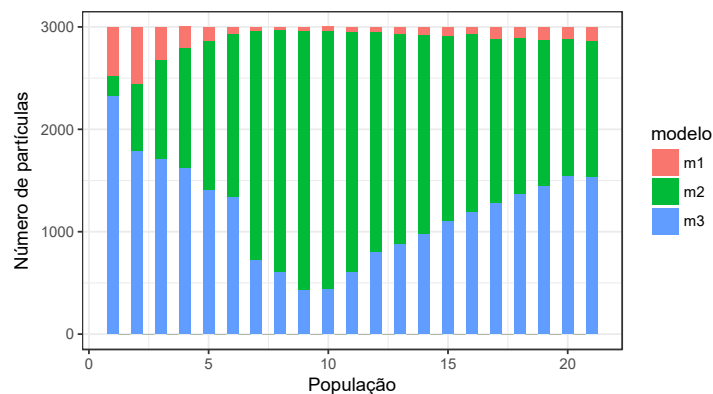


Figura 5- Histograma da distribuição das partículas para os modelos em relação a cada população.(algoritmo B (3000 partículas))

Em relação à última população no algoritmo A, temos as seguintes estimativas de cada modelo m

$$\hat{\theta}(1) = (\hat{\alpha} = -4,83; \hat{\beta} = 33,21; \hat{s} = 43,98)$$

$$\hat{\theta}(2) = (\hat{\mu} = 28,43; \hat{\sigma} = 0,5833)$$

$$\hat{\theta}(3) = (\hat{\alpha} = 44,31; \hat{\beta} = 29,025)$$

Em relação à última população no algoritmo B, temos as seguintes estimativas de cada modelo m

$$\hat{\theta}(1) = (\hat{\alpha} = -4,9461; \hat{\beta} = 33,6467; \hat{s} = 42,7487)$$

$$\hat{\theta}(2) = (\hat{\mu} = 28,4958; \hat{\sigma} = 0,6223)$$

$$\hat{\theta}(3) = (\hat{\alpha} = 44,5510; \hat{\beta} = 29,09)$$

Podemos concluir que os dois algoritmos foram similares em seus resultados, ou seja, em ambos os casos, não podemos concluir haja um dos três modelos que se sobressaia fortemente em relação aos demais para explicar este conjunto de dados. Dois fatores podem ter colaborado para isto: a amostra observada é pequena (15 pontos) e o número de partículas definido para ser aceito em cada população intermediária pode ter sido insuficiente.

4. CONCLUSÕES

Neste trabalho, iniciamos a aplicação da teoria de valores extremos em dados de enchente no Rio Negro, na cidade de Manaus, ajustando as distribuições de valores extremos aos dados de enchente anual da série histórica de 2000-2015. Aplicamos a metodologia ABC, mais especificamente o algoritmo ABC-SMC, para estimação de parâmetros. No caso, usamos os dados para três diferentes modelos probabilísticos, a saber, Weibull, Gumbel e Fréchet. Quando utilizado o algoritmo ABC-SMC para selecionar simultaneamente modelo e parâmetros da distribuição que melhor ajuste os dados em estudo, verificou-se apenas evidência pobre em favor da distribuição Weibull quando comparada com a distribuição de Fréchet e nada pode-se dizer quando comparada com a distribuição Gumbel já que a quantidade de partículas aceitas na última população por ambos os modelos (Weibull e Gumbel) está muito próxima. Como proposta de trabalho futuro, faz-se necessário uma investigação quanto ao número de partículas a serem aceitas em cada população, bem como um conjunto de dados de dimensão maior.

Acknowledgements

Agradecimentos à Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) pelo financiamento da pesquisa, através do Programa Institucional de Bolsas de Iniciação Científica da Universidade Federal do Amazonas.

REFERÊNCIAS

- Costa, Jose M. J., Orlande, Helcio R. B. O., & Silva, Wellington B. (2017), *Model selection and parameter estimation in tumor growth models using approximate Bayesian computation-ABC*, Computational and applied Mathematics, (Accepted).
- Kotz, S., & Nadarajah, S. (2000), *Extreme value distributions: theory and applications*. World Scientific.
- Del Moral, P., Doucet, A., & Jasra, A. (2006), *Sequential monte carlo samplers*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68(3), 411-436.
- Sisson, S., Fan, Y., & Tanaka, M. M. (2008), *A note on backward kernel choice for sequential Monte Carlo without likelihoods*, Technical report, University of New South Wales.
- Toni, T., & Stumpf, M. P. (2010), *Simulation-based model selection for dynamical systems in systems and population biology*, Bioinformatics, 26(1), 104-110.

APPENDIX A

DATA MODELING VIA EXTREME VALUE DISTRIBUTIONS AND APPROXIMATE BAYESIAN COMPUTATION - ABC

Resumo. *Floods are extreme phenomena, therefore the modeling by extreme value distributions is a reasonable way to approach such data for its behavior analysis. This study goal is to fit flood data from Rio Negro, in Manaus - AM, during the period from 2000 to 2015, using the Weibull, Gumbel, and Fréchet distributions, estimating the parameters utilizing approximate bayesian computation (ABC). Specifically, it is applied a improved method from the original ABC algorithm, called ABC-SMC, which is used for parameter estimation from the proposed distributions. In addition, the ABC-SMC was also used for the selection of models.*

Keywords: *Extreme Value Distributions, ABC Algorithm, Bayesian Statistics, Selection Models*