



Discriminante não linear baseado em mistura de distribuições: uma aplicação no tempo de retorno de observações de fluídos máximos heterogêneos na bacia do rio Paranapanema-Brasil

Non linear discriminant based on mixture of distributions: an application in return time of heterogeneous maximum flows observed in the Paranapanema river basin – Brazil

Cira E. G. Otiniano¹, Evelyn C. Cruvinel² and Carlos H. Lima³

Received, Aug. 08, 2020

Accepted, Dec. 09, 2020



How to cite this article:

Otiniano CEG, Cruvinel EC, Lima CH. *Discriminante não linear baseado em mistura de distribuições: uma aplicação no tempo de retorno de observações de fluídos máximos heterogêneos na bacia do rio Paranapanema-Brasil*. *Selecciones Matemáticas*. 2020;7(2):242–249. <http://dx.doi.org/10.17268/sel.mat.2020.02.06>

Resumo

Neste trabalho, vazões máximas observadas na bacia do rio Paranapanema são modelados por uma mistura de duas distribuições GEV cujas componentes correspondem às subpopulações de verão e inverno. A função discriminante não linear permitiu confirmar a hipótese do modelo de mistura. O modelo proposto é então usado para determinar as probabilidades de excedências e períodos de retorno associados a vazões extremas, que são de fundamental importância para projetos hidráulicos. Os resultados mostram diferenças significativas quando um modelo de mistura é usado e, portanto, uma abordagem mais coerente com os dados observados.

Palavras-chave. MGEV, Análises de discriminante, Cheia de bacias hidrográficas, Período de retorno.

Abstract

In this work, maximum flows observed in the Paranapanema river basin are modeled by a mixture of two GEV distributions whose components correspond to the summer and winter subpopulations. The nonlinear discriminant function allowed to confirm the assumption of mixture model. The proposed model is then used to determine probabilities of exceedance and return periods associated with extreme flows, which are of fundamental importance for hydraulic projects. The results show significant differences when a mixture model is used and therefore an approach more coherent with the observed data.

Keywords . MGEV, Discriminant Analysis, Flows, Return Period.

1. Introdução. Eventos de cheia causam diversos impactos na sociedade, por exemplo, transbordamentos de rios e deslizamentos. Segundo a Defesa Civil do Estado de São Paulo, do 01 de dezembro de 2019 a 3 de março de 2020, as chuvas torrenciais deixaram 35 mortos, 50 feridos, 578 desabrigados e 5.150 desalojados na Grande São Paulo e 14 municípios do estado de São Paulo, [7]. A determinação do risco de cheia associado a uma dada localidade requer a análise e modelagem estatística da frequência de ocorrência desses eventos. Tradicionalmente, tal análise requer assumir que os dados observados são independentes e identicamente distribuídos. Dessa forma, assume-se homogeneidade e que cada evento pertence a uma

*Departamento de Estatística, UnB, 70910-900 Brasília, DF, Brazil (cira@unb.br).

†Departamento de Estatística, UnB, 70910-900 Brasília, DF, Brazil (evelyncruvinel@hotmail.com).

‡Departamento de Engenharia Civil and Ambiental, UnB, 70910-900 Brasília, DF, Brazil (chr12107@gmail.com).

mesma população. Entretanto, cheias podem ser causadas por chuvas originadas por diferentes mecanismos (convecção, frentes, ciclones, etc). Nesse caso, as características estocásticas da chuva podem variar de acordo com o mecanismo causador e a suposição de homogeneidade pode não ser mais válida. Como forma de lidar com a violação dessa suposição, pode-se assumir que os eventos de cheia são gerados por distribuições heterogêneas compostas por uma mistura de duas ou mais populações. [3] descrevem algumas técnicas para analisar frequência de vazões de populações heterogêneas, dentre elas se destaca a mistura de distribuições com diferentes distribuições homogêneas.

[14] usou uma mistura de duas distribuições normais para se adequar a uma série anual de inundações log-transformadas no estado de Illinois. [12] e [8] usaram uma mistura de duas distribuições Gumbel para se ajustarem dados de inundações regionais na Itália.

O modelo de misturas finitas cujas componentes são da família de distribuição de Valor Extremo Generalizada (GEV) abrangem diversos fenômenos. [3] ajustaram dados de vazões anuais de 10 bacias do rio Gila usando uma mistura de duas componentes GEV, estimando seus parâmetros pelo método dos L momentos. [6] utilizou misturas das componentes GEV para avaliar dados sobre a frequência de vazões de uma região do Noroeste do México, estimando os parâmetros por máxima verossimilhança. [11] aplicaram mistura de distribuições GEV para ajustar dados de energia e os estimadores de máximo verossimilhança dos parâmetros da mistura de duas componentes GEV via algoritmo EM.

Quando se trata de modelos de mistura de k componentes, dois principais problemas emergem. O primeiro refere-se ao método de estimação dos parâmetros. Nesse caso, dados não-classificados são utilizados para esse propósito. O segundo problema aparece no estudo do discriminante de dados não-classificados, onde é necessário classificar a população homogênea que um certo valor amostral procede.

Uma abordagem de análise de discriminante, no contexto de mistura finita de distribuições, está baseada no discriminante não-linear de Bayes. Essa função discriminante não linear tem mostrado bons resultados incluindo para pequenas amostras de misturas de diversas distribuições conhecidas. [4] utilizou em mistura de duas distribuições inversa gaussiana, [1] em uma mistura de duas distribuições Weibull, [2] em uma mistura de duas distribuições Gumbel, e [13] em uma mistura de duas distribuições inversa Weibull.

Neste trabalho é aplicada a técnica de discriminante não-linear de Bayes, baseada no trabalho de [5], para classificar dados de uma população heterogênea correspondente a dados de vazão máxima anual das usinas da bacia do rio Paranapanema, Brasil.

O restante deste artigo tem a seguinte organização. Na Seção 2, é apresentada a técnica de discriminante não-linear para a mistura de duas distribuições GEV. Na Seção 3, são apresentados os resultados dos testes de independência e ajuste dos dados, e em seguida os resultados das estimativas de máxima verossimilhança dos parâmetros desconhecidos do modelo de mistura GEV e a estimativa da função discriminante.

2. Discriminante não linear. A distribuição limite de estatísticas de máximos normalizadas é bem conhecida como a distribuição de valor extremo generalizado (GEV), ele e sua função de densidade de probabilidade, introduzida por Jenkinson [9]), são definidas por

$$(2.1) \quad G(x; \gamma, \sigma, \mu) = \begin{cases} \exp \left\{ - \left[1 + \frac{\gamma}{\sigma} (x - \mu) \right]^{-\frac{1}{\gamma}} \right\}, & \gamma \neq 0 \\ \exp \left[- \exp \left(-\frac{x-\mu}{\sigma} \right) \right], & \gamma = 0 \end{cases},$$

e

$$(2.2) \quad f(x; \gamma, \sigma, \mu) = \begin{cases} \frac{1}{\sigma} \left[1 + \frac{\gamma}{\sigma} (x - \mu) \right] \exp \left\{ - \left[1 + \frac{\gamma}{\sigma} (x - \mu) \right]^{-\frac{1}{\gamma}} \right\}, & \gamma \neq 0 \\ \frac{1}{\sigma} \exp \left(-\frac{x-\mu}{\sigma} \right) \exp \left[- \exp \left(-\frac{x-\mu}{\sigma} \right) \right], & \gamma = 0 \end{cases},$$

respectivamente, sendo $\gamma \in \mathbb{R}$ o parâmetro de forma (índice de cauda), $\sigma > 0$ o parâmetro de escala e $\mu \in \mathbb{R}$ o parâmetro de locação. A distribuição GEV está definida para x tal que $1 + \frac{\gamma}{\sigma}(x - \mu) \geq 0$.

Por outro lado, uma variável aleatória X que resulta de uma mistura de k distribuições (componentes) diz-se que sua distribuição é heterogênea ou é uma mistura de componentes homogêneas. Isto é, a função de distribuição acumulada de X é dada por:

$$(2.3) \quad F(x; \theta) = \sum_{i=1}^k p_i \cdot F_i(x; \theta_i),$$

em que $p_i > 0$ é a i -ésima proporção da mistura e $\sum_{i=1}^k p_i = 1$, F_i é a i -ésima distribuição da população homogênea chamada de componente da mistura e $i = 1, \dots, k$, McLachlan e Peel [10].

Ao considerar X uma população mistura de duas populações homogêneas GEV (MGEV), de (2.3) e (2.2), sua função de densidade de probabilidade (fdp) é dada por

$$(2.4) \quad \begin{aligned} h(x; \theta) &= p_1 f_1(x; \gamma_1, \sigma, \mu_1) + p_2 f_2(x; \gamma_2, \sigma, \mu_2) \\ &= \frac{1}{\sigma} [p_1 g_1(x_j; \gamma_1, \sigma, \mu_1) + p_2 g_2(x_j; \gamma_2, \sigma, \mu_2)], \quad p_1 + p_2 = 1, \end{aligned}$$

sendo

$$g_i(x_j; \gamma_i, \sigma, \mu_i) = \begin{cases} \left[1 + \frac{\gamma_i}{\sigma} (x_j - \mu_i) \right] \exp \left\{ - \left[1 + \frac{\gamma_i}{\sigma} (x_j - \mu_i) \right]^{-\frac{1}{\gamma_i}} \right\}, & \gamma_i \neq 0 \\ \exp \left(-\frac{x_j - \mu_i}{\sigma} \right) \exp \left[-\exp \left(-\frac{x_j - \mu_i}{\sigma} \right) \right], & \gamma_i = 0 \end{cases},$$

e $\theta = (p_1, \gamma_1, \gamma_2, \sigma, \mu_1, \mu_2)$ o vetor dos parâmetros de (2.4). A função de distribuição acumulada correspondente a MGEV é

$$H(x; \theta) = p_1 G_1(x; \gamma_1, \sigma, \mu_1) + p_2 G_2(x; \gamma_2, \sigma, \mu_2),$$

em que G_i são distribuições GEV definidas em (2.1).

Como os valores de γ_i , $i = 1, 2$ podem ser positivos, negativos ou nulos, há seis maneiras de combinar esses parâmetros dentro do modelo de MGEV. As curvas na Figura (2.1) são ilustrações de alguns desses casos. Dados heterogêneos de hidrologia e finanças frequentemente mostram densidades desses três tipos.

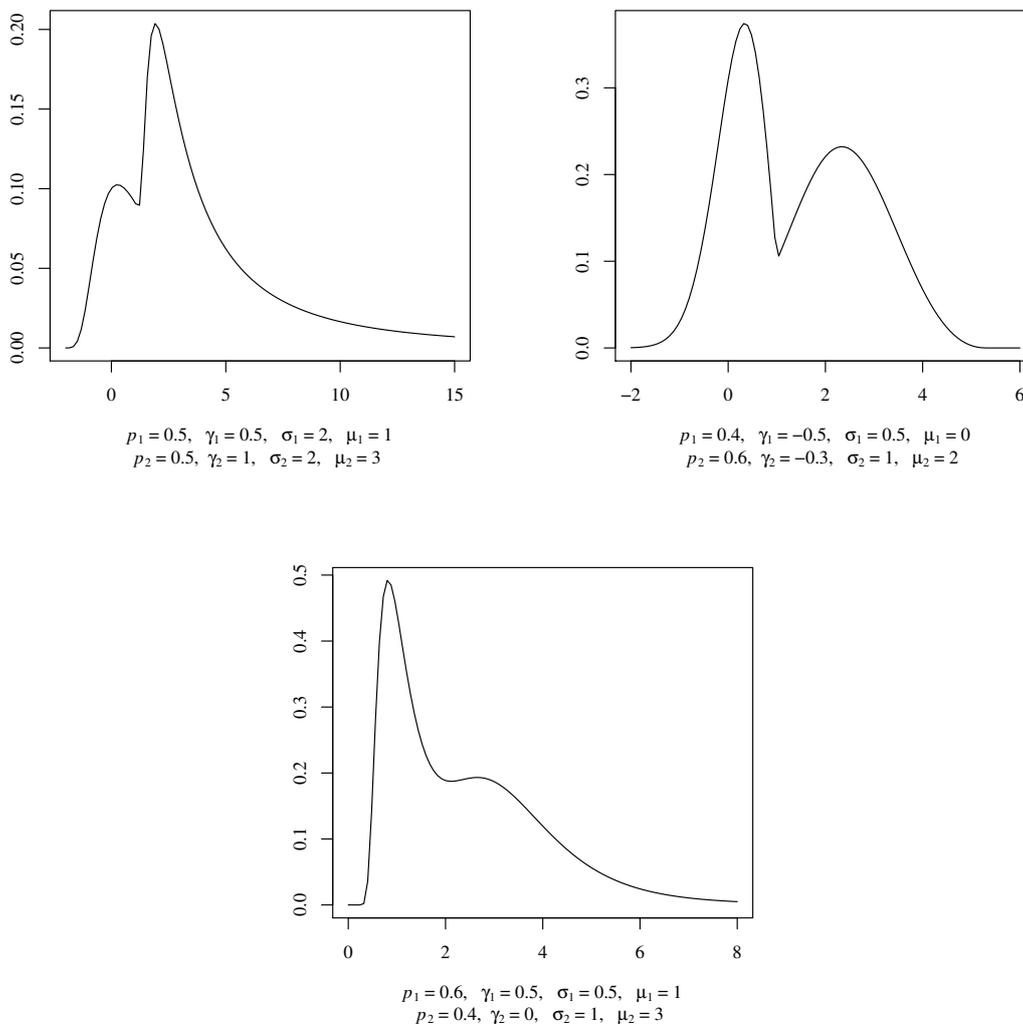


Figura 2.1: Gráfico da densidade (2.4) para quatro conjuntos de parâmetros.

Quando dados reais apresentam densidade empírica como da Figura 2.1, temos indícios que a população em estudo é heterogênea então a densidade correspondente deve ser do tipo (2.4). Para comprovar essa

suposição, um estudo de classificação dos dados é realizado para identificar o agrupamento dos dados em duas subpopulações. Neste trabalho utiliza-se o classificador de Bayes como ferramenta na classificação de uma observação x que pode vir da primeira subpopulação (primeira componente) f_1 ou segunda subpopulação f_2 (segunda componente). O procedimento ocorre em duas etapas:

Na primeira etapa, considera-se o modelo (2.4) com vetor de parâmetros $\theta = (p_1, \gamma_1, \gamma_2, \sigma, \mu_1, \mu_2)$. Para valores amostrais x_1, \dots, x_n de X com densidade (2.4), o método de máxima verossimilhança é utilizado para obter as estimativas $\hat{\Theta} = (\hat{p}_1, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\sigma}, \hat{\mu}_1, \hat{\mu}_2)$ de Θ ao otimizar numericamente a função de verossimilhança

$$(2.5) \quad L(X, \theta) = \prod_{j=1}^n h(x_j; \theta).$$

Para a maximização, nós utilizamos a função Optim do software R versão 3.3.1. Procedimento que já foi testado sua eficácia, via simulações, no trabalho de Cruvinel (2017).

Na segunda etapa, no análise de discriminante utilizam-se as estimativas $\hat{\Theta} = (\hat{p}_1, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\sigma}, \hat{\mu}_1, \hat{\mu}_2)$ de Θ , obtidas na primeira etapa. Para isso, suponha que Π_1 e Π_2 sejam as subpopulações homogêneas correspondentes as densidades $f_1(x)$ e $f_2(x)$, respectivamente, do modelo (2.4). O classificador de Bayes é definido pela função

$$(2.6) \quad NLD(x) = a - bz,$$

em que a e b são funções do vetor de parâmetros θ e z uma função de x e de θ . A função NLD definida em (2.6) é chamada de discriminante não linear de Bayes e é obtida ao considerar que uma observação x de X (não classificada) vem de Π_1 se

$$(2.7) \quad \begin{aligned} P(x \in \Pi_1) &= \frac{p_1 f_1(y)}{\sum_{i=1}^k p_i f_i} \\ &= \{1 + \exp[a - bz]\}^{-1} \\ &= \{1 + \exp[NLD(x)]\}^{-1}. \end{aligned}$$

Para $\gamma_1 > 0$ e $\gamma_2 < 0$

$$(2.8) \quad a = \ln\left(\frac{p_2}{p_1}\right) + \left(\frac{1}{\gamma_2} - \frac{1}{\gamma_1}\right) \ln(\sigma),$$

$$(2.9) \quad b = 1,$$

e

$$(2.10) \quad \begin{aligned} z &= \left(\frac{1}{\gamma_2} + 1\right) \ln[\sigma + \gamma_2(x_j - \mu_2)] - \left(\frac{1}{\gamma_1} + 1\right) \ln[\sigma + \gamma_1(x_j - \mu_1)] + \\ &+ \left[1 + \frac{\gamma_2}{\sigma}(x_j - \mu_2)\right]^{-\frac{1}{\gamma_2}} - \left[1 + \frac{\gamma_1}{\sigma}(x_j - \mu_1)\right]^{-\frac{1}{\gamma_1}} \}. \end{aligned}$$

Após calcular $\hat{\Theta} = (\hat{p}_1, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\sigma}, \hat{\mu}_1, \hat{\mu}_2)$ obtem-se a função $\widehat{NLD}(x) = \hat{a} - \hat{b}z$ que permite classificar um valor amostral x de X em Π_1 ou Π_2 da seguinte maneira:

x é classificado em Π_1 se $\widehat{NLD}(x) < 0$ pois $P(x \in \Pi_1) < 1/2$ e x é classificado em Π_2 se $\widehat{NLD}(x) \geq 0$.

Repetindo esse mesmo procedimento para todas as observações, são classificadas as observações da sub-população Π_1 e Π_2 . Se as proporções de dados para cada grupo estão próximas dos valores \hat{p}_1 e \hat{p}_2 , pode-se concluir que a população é de fato heterogênea e está composta por duas subpopulações. Com isto, concluí-se que o tempo de retorno dos dados deve ser calculado a partir de do modelo (2.4).

3. Aplicação. Nesta seção, usando a metodologia descrita na Seção 3, calcula-se o tempo de retorno de vazões máximas da bacia hidrográfica do rio Paranapanema, localizada no sul do Brasil.

Os dados de vazão diária são medidos em metros cúbicos por segundo (m^3/s) e cobrem o período de 1937 a 2008. As vazões são caracterizadas por duas estações distintas: verão (novembro a abril) e inverno (maio a outubro). Cheias podem ocorrer em ambos os períodos e são caracterizadas por diferentes

Antes de aplicar a metodologia descrita na seção anterior, realizou-se o teste Ljung - Box para verificar a hipótese nula de independência serial nos máximos anuais. Os resultados são mostrados na Tabela 3.2, onde o p -valor indica falha em rejeitar a hipótese nula de independência para um nível de significância de 5%.

| Estação | Estatística Ljung-Box | g.l. | p -valor |
|---------|-----------------------|------|------------|
| Verão | 1.1515 | 1 | 0.2832 |
| Inverno | 2.2292 | 1 | 0.1354 |
| Anual | 1.8564 | 1 | 0.173 |

Tabela 3.2: Ljung-Box Test - Peak River Flow Series.

Com a hipótese de independência serial, as estimativas dos parâmetros das distribuições GEV e MGEV calculados pelo método da máxima verossimilhança são apresentados na Tabela 3.3.

A adequação do ajuste dos modelos é avaliada a partir do teste de Kolmogorov-Smirnov (K-S), da avaliação do Critério de informação Akaike (AIC) e do critério de informação bayesiano (BIC). Com base nesses resultados, dados na Tabela 3.4, pode-se concluir que o modelo MGEV apresenta um ajuste melhor aos dados quando comparado com o modelo GEV simples.

| Model | | | | | | |
|-------|-------------------------|------------------------|------------------------|------------------------|----------------------|----------------------|
| GEV | $\hat{\gamma} = 0.0181$ | $\sigma = 148.2102$ | $\mu = 295.0978$ | | | |
| MGEV | $\hat{p}_1=0.36$ | $\hat{\gamma}_1= 0.12$ | $\hat{\gamma}_2=-0.69$ | $\hat{\sigma}= 103.83$ | $\hat{\mu}_1=508.65$ | $\hat{\mu}_2=248.17$ |

Tabela 3.3: Estimativas dos parâmetros

| Modelo | l | AIC | BIC | p -valor (teste K-S) |
|-------------|---------|----------|----------|-------------------------|
| GEV | -488 | 981.324 | 988.2362 | 2.2^{-16} |
| Mixture GEV | -488.82 | 989.6587 | 1003.483 | 0.9661 |

Tabela 3.4: $\log(L(\theta)) = l$ e resultado dos testes

Na Figura 3.2 mostra-se também que o modelo MGEV se ajusta melhor ao histograma dos dados do que a distribuição GEV. O que confirma os resultados da Tabela 3.4.

O qq-plot e a função de distribuição acumulada, quando comparadas a versão empírica com a teórica, são ferramentas úteis para medir a adequação do modelo. Na Figura 3.3, gráfico a esquerda, é mostrado o qq-plot das vazões máximas anuais(empírico) versus qq-plot teórico dos modelos GEV (preto) e MGEV (cinza). No gráfico a esquerda mostra-se a fda empírica comparada com a fda teórica dos modelos GEV e MGEV. Nos dois gráficos o modelo MGEV (pontos cinza) fica mais próxima da função empírica (preto).

Uma vez que o modelo MGEV mostra se ajustar bem aos dados, procede-se a calcular a função de discriminantes não linear, $NLD(x_j)$ dos dados x_j . Ao utilizar a equação (2.6), obtém-se:

$$\begin{aligned}
 NLD_m(x_j) = & -45,93 - 0,44 \ln [103,83 - 0,69(x_j - 248,17)] \\
 & - 9,57 \ln [103,83 + 0,12(x_j - 508,65)] \\
 & + [1 - 0,007(x_j - 248,17)]^{\frac{1}{0,69}} \\
 & - [1 + 0,001(x_j - 508,65)]^{-\frac{1}{0,12}}.
 \end{aligned}
 \tag{3.1}$$

Assim, do total de 74 observações, 26 foram classificadas na primeira população (inverno) e 48 na segunda população (verão). Como o número de observações de cada classe é conhecida, a proporção total de erro de classificação foi calculada. Assim, tem-se que a classificação das observações das vazões máximas anuais

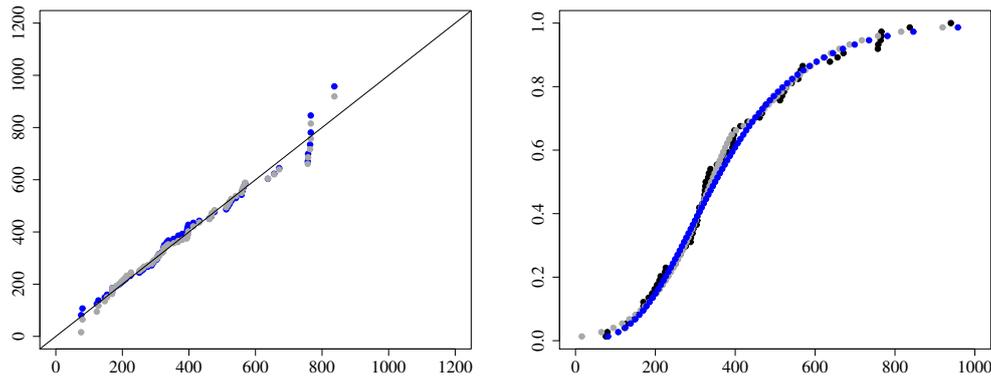


Figura 3.3: QQ-plot (esquerda) e fda (direita). Pontos azuis (GEV), pontos cinzas (MGEV) e pontos pretos (empírico).

cujas subpopulações são inverno (Π_1) e verão (Π_2) foram bem classificadas com aproximadamente 98,8 % de erro, pois a proporção de erro foi de 0,01245. Com isto concluímos que o modelo (2.4) adotado é apropriado para ajustar os dados de vazão máxima.

3.1. Estimativa do Período de Retorno. Em estudos hidrológicos, torna-se fundamental calcular os valores de vazões associados a diferentes probabilidades de excedência. Tal procedimento, também conhecido em hidrologia como estimativa do período de retorno, que se obtém ao calcular o intervalo médio de ocorrência para um dado evento extremo de vazão. Assim, $(1 - F(x; \theta))$ é a probabilidade de excedência de uma vazão x e o período de retorno associado a x é determinado pela função

$$(3.2) \quad T(x) = \frac{1}{1 - F(x; \theta)}, \quad F \text{ é fda da população.}$$

Ao considerar as vazões máximas anuais do rio Paranapanema foram calculadas as funções de distribuição $G(x; \hat{\theta})$ (distribuição GEV) e $H(x; \hat{\theta})$ (distribuição MGEV), em seguida utilizou-se (3.2) para estimar $T(x)$. Os resultados obtidos para os modelos GEV e MGEV são mostrados na Figura 3.4. Nessa figura observa-se que as vazões máximas são subestimadas pela distribuição GEV, principalmente para o período de retorno maior do que o tamanho da série observada.

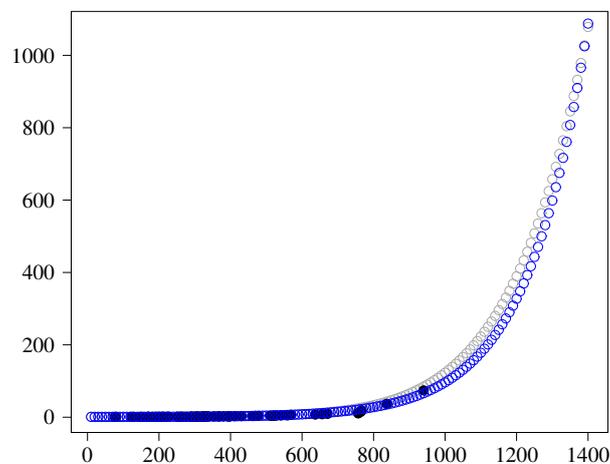


Figura 3.4: $T(x)$: empírico (pontos azul escuro), com GEV (pontos azuis), e com MGEV (pontos cinza).

4. Conclusões. Utilizou-se a função de discriminante não-linear de Bayes, definida por Cruvinel[5], para classificar as vazões máximas observadas na Bacia do Rio Paranapanema- Brasil. Com um erro de 0,012 o classificador de Bayes conseguiu classificar corretamente os dados. Assim, pode-se concluir que uma boa estrutura probabilística para esses dados é um modelo de mistura de distribuições GEV. Com isto, mostra-se que a probabilidades de excedência e respectivos períodos de retorno associados a vazões extremas são melhor ajustados considerando o modelo de mistura de distribuições GEV. Os resultados mostram que para vazões entre 900 e 1400 m^3/s , aproximadamente, o tempo de retorno é subestimado se na modelagem das vazões máximas for utilizado o modelo GEV simples em vez do modelo MGEV. Para a modelagem foi considerada hipótese de independência serial dos dados, porém de acordo com o teste Ljung-Box os dados possuem dependência serial. Para um trabalho futuro fica aberto a modelagem desses dados e a estimação do tempo de retorno correspondente.

5. Reconhecimentos. Somos gratos aos referees que com seus comentários contribuíram na melhoria da versão final deste artigo. Esta pesquisa teve o suporte parcial do DPP/UnB, FINATEC e FAPDF.

ORCID and License

Cira E. G. Otíniano <https://orcid.org/0000-0002-5619-0478>

Evelyn C. Cruvinel <https://orcid.org/0000-0002-0145-5678>

Carlos H. Lima <http://orcid.org/0000-0003-2947-7956>

This work is licensed under the [Creative Commons Attribution-NoComercial-ShareAlike 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Referências

- [1] Ahmad KE, Abd-Elrahman AM. Updating a nonlinear discriminant function estimated from a mixture of two Weibull distributions. *Math. Comput. Modelling*. 1994; 19(11):41–51.
- [2] Ahmad KE, Jaheen ZF, Modhesh AA. Estimation of a discriminant function based on small sample size from a mixture of two Gumbel distributions. *Comm. Statist. Simulation Comput*. 2010; 39(4):13–725.
- [3] Alila Y, Mtiraoui A. Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques. *Hydrological Process*. 2002; 16:1065–1084.
- [4] Amoh RK. *Estimation of a discriminant function from a mixture of two inverse Gaussian distributions when the sample size is small*. *J. Statist. Comp. Simul*. 1985; 20:275–286.
- [5] Cruvinel EC. Discriminante para mistura de distribuições GEV.[Dissertação de Mestrado]. Brasília: Universidade de Brasília. 2017. Disponível em: <http://repositorio.unb.br/handle/10482/24006>.
- [6] Escalante-Sandoval C. A Mixed distribution with EV1 and GEV components for analyzing heterogeneous samples. *Ingeniería, investigación y tecnología*. 2007; 8(3):1405-7743.
- [7] Girardo G. Volume de chuvas em fevereiro em São Paulo já é maior desde 1999[Intenet]. Sao Paulo: O Estado de S. Paulo; 2020[Consultado em 09 dezembro de 2020]. Recuperado de <https://sao-paulo.estadao.com.br/noticias/geral,volume-de-chuvas-em-fevereiro-em-sp-ja-e-o-maior-desde-1999,70003194244>.
- [8] Fiorentino M, Arora K, Singh VP. The two-component extreme value distribution for flood frequency analysis: Derivation of a new estimation method. *Stochastic Hydrology and Hydraulics*. 1987; 1(3):199–208.
- [9] Jenkinson AF. *The frequency distribution of annual maximum (or minimum) values of meteorological elements*. *Quarterly J. of the Royal Met. Society*. 1995; 8:145–158.
- [10] McLachlan GJ, Peel D. *Finite Mixture Models*. Australia: John Wiley, 2004.
- [11] Otíniano CEG, Teixeira ECM. Estimação dos parâmetros da mistura de duas componentes GEV via Algoritmo EM. *Tendencias em Matemática Aplicada e Computacional*. 2014; 15(1):59–71.
- [12] Rossi F, Fiorentino M, Versace P. Two-component extreme value distribution for flood frequency analysis. *Water Resour. Res*. 1994; 20(7):847–856.
- [13] Sultan KS, Al-Moisheer AS. *Estimation of a discriminant function from a mixture of two inverse weibull distributions*. *J. of Statistical Computation and Simulation*. 2013; 83(3):405–416.
- [14] Singh KP. *Hydrologic distributions resulting from mixed populations and their computer simulations*. *I.A.S.H*. 1968; 8:671–681.