


VARIABILIDADE DA ESTIMATIVA DE EVENTOS EXTREMOS DE PRECIPITAÇÃO NO MUNICÍPIO DE MACAÉ – RJ *VARIABILITY OF THE ESTIMATE OF EXTREME PRECIPITATION EVENTS IN THE CITY OF MACAÉ - RJ*

Isabor Viana Sant'Ana Mendes¹ Aline Míriam Marques² Danilo de Assis Amaral³ Flávia Ribeiro Villela Fernandes⁴ Marcos Antônio Cruz Moreira⁵ Bruno Luis da Silva Maia⁶ 

Resumo: O estudo de eventos extremos de precipitação torna-se essencial e relevante devido aos impactos e implicações que provocam na sociedade. No município de Macaé esses eventos adversos são constantes, sendo o que mais registrou ocorrência de alagamentos caracterizados como desastres no estado do Rio de Janeiro, no período de 1991 a 2012, no Atlas Brasileiro de Desastres Naturais. Assim, o objetivo deste trabalho foi analisar a variabilidade dos valores obtidos através dos métodos *Generalized Extreme Value (GEV)* e *Generalized Pareto Distribution (GPD)* para determinação do nível de retorno de eventos extremos de precipitação no município de Macaé. O modelo estatístico *Analysis of Variance (ANOVA)* foi adotado e o diagnóstico do modelo validado. Adicionalmente, a análise *post hoc* foi realizada através do *Tukey Honest Significant Difference*. As análises estatísticas foram implementadas no ambiente de *software R*. De fato, constata-se que existem diferenças significativas entre as estimativas dos níveis de precipitação segundo métodos e estação do ano. Os resultados podem auxiliar análises futuras e corroboram com a escolha do

¹ Mestranda em Engenharia Ambiental, Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFFluminense. isaborvsm@gmail.com.

² Engenheira Civil. Mestranda em Engenharia Ambiental, Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFFluminense. alinemarqueseng.civil@gmail.com.

³ Mestrando em Engenharia Ambiental, Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFFluminense. daniloamaral@hotmail.com.

⁴ Mestre em Engenharia Elétrica. Professora assistente da Universidade Federal do Rio de Janeiro (UFRJ, Campus Macaé). flaviavillela@macaeufrj.br.

⁵ Doutor em Engenharia Elétrica. Professor Titular do Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFFluminense. macruz@iff.edu.br

⁶ Engenheiro de Produção. Mestrando em Engenharia Ambiental, Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFFluminense. brunoluismaia@gmail.com.

melhor método para entender a variação e o comportamento de eventos com grandes volumes pluviométricos em curta duração de tempo ou enxurradas. Sendo portanto, um estudo relevante para auxiliar no planejamento e na implementação de ações de políticas públicas a serem aplicadas em problemas oriundos de tais eventos.

Palavras-chave: ANOVA. Eventos Extremos. Distribuição GEV. Distribuição GPD.

Abstract: The study of extreme precipitation events becomes essential and relevant due to the impacts and implications that these events have on society. In the city of Macaé, these adverse events are constant, the most frequent occurrence of floods characterized as disasters in the state of Rio de Janeiro, from 1991 to 2012, in the Brazilian Atlas of Natural Disasters. Thus, the objective of this work was to analyze the variability of the values obtained through the Generalized Extreme Value (GEV) and Generalized Pareto Distribution (GPD) methods to determine the level of return of extreme precipitation events in Macaé city. The Analysis of Variance (ANOVA) statistical model was adopted and the model's diagnosis validated. Additionally, post hoc analysis was performed using the Tukey Honest Significant Difference. Statistical analyzes were implemented in the software environment R. In fact, it appears that there are significant differences between the estimates of rainfall levels according to methods and season. The results can assist future analyzes and corroborate the choice of the best method to understand the variation and behavior of events with large rainfall volumes in a short time or downpour. Therefore, it is relevant study to assist in the planning and implementation of public policy actions to be applied to problems arising from such events.

Keywords: ANOVA. Extreme Events. GEV Distribution. GPD Distribution.

1 INTRODUÇÃO

Conhecer a variação da precipitação e o seu comportamento, tanto no espaço quanto no tempo, é essencial para o planejamento e implementação de ações de políticas públicas para evitar os problemas ocasionados por eventos com grandes volumes pluviométricos em curta duração de tempo ou enxurradas. Segundo Souza e Azevedo (2012), a discussão e a preocupação frequente sobre as mudanças climáticas está associada diretamente aos impactos sobre o meio ambiente e as diversas atividades humanas no planeta, e, principalmente aos efeitos, cada vez mais frequentes, dos extremos climáticos.

Para determinação destes eventos utiliza-se a metodologia da Teoria dos Valores Extremos (TVE), que tem como objetivo estudar eventos raros (KOTZ; NADARAJAH, 2000) que são compostas por duas outras teorias: i) Distribuição Generalizada de Valor Extremo (GEV, sigla em inglês), onde a amostra é dividida em subperíodos (blocos) que podem ser mensais, sazonais, etc, e é extraído o valor máximo ou mínimo cada bloco, para formar o conjunto de eventos extremos (SANTOS; LUCIO; SILVA, 2015); e ii) Distribuição Generalizada de Pareto (GPD, sigla em inglês), que objetiva a aproximação da distribuição assintótica dos excessos de uma variável aleatória acima de um valor limite pré-determinado (PICKANDS, 1975).

Os eventos extremos de precipitação podem causar danos tanto em áreas urbanas quanto rurais, além de ocasionar vários problemas para a sociedade. De acordo com Bautista; Zocchi; Angelocci (2004), estudar estes eventos é de grande relevância, uma vez que suas estimativas probabilísticas são essenciais para planejar e desenvolver as atividades sujeitas às adversidades. Possibilitando assim, a aplicação de medidas preventivas, minimizando os danos através do planejamento baseado nos estudos prévios.

No município de Macaé a recorrência desses eventos adversos são constantes e conforme descrito no Atlas Brasileiro de Desastres Naturais, foi o município que mais registrou ocorrência de alagamentos caracterizados como desastres, 8 de um total de 56 registros no estado do Rio de Janeiro, no período de 1991 a 2012 (CEPED UFSC, 2013).

Segundo Castro (1998), o desastre é resultado de eventos adversos, naturais ou provocados pelo homem em ambientes vulneráveis, possibilitando causar danos humanos, ambientais, materiais, além de prejuízos econômicos e sociais. Buriol *et al.* (2006) comentaram que as chuvas intensas ocasionam alagamentos, escorregamento das áreas susceptíveis, intensificam a erosão do solo e o transporte de sedimentos que assoreiam os leitos dos reservatórios de água e dos rios. Assim como períodos longos sem chuva acarretam outros impactos negativos, como deficiência hídrica, abaixamento do nível das barragens e redução da vazão nos sistemas fluviais.

Deste modo, torna-se de fundamental importância analisar o comportamento das chuvas intensas, contribuindo na tomada de decisões do planejamento urbano e no planejamento de outras atividades como o turismo e a construção civil, que são influenciadas pela magnitude e frequência dos eventos extremos.

A partir da análise dos dados de comportamento e tendências das chuvas intensas e o período de retorno que identificou as características, o comportamento, a tendência das chuvas intensas e os principais impactos causados, o objetivo deste trabalho foi identificar as diferenças significativas entre os resultados dos métodos GEV (*Generalized Extreme Value*) e GPD (*Generalized Pareto Distribution*) utilizados para determinar o tempo de retorno de eventos extremos de precipitação no município de Macaé.

2 DESENVOLVIMENTO

2.1 Materiais e Método

Os dados utilizados neste trabalho foram retirados dos resultados finais obtidos do trabalho de Mendes (2019) como mostrado na Tabela 1. Trata-se de dados de nível de retorno obtidos através da Teoria dos Valores Extremos.

Tabela 1: Nível de retorno (mm/dia) por tempo segundo método e estação do ano.

Estação	Método	Tempo de Retorno (anos)				
		2	5	10	50	100
DJF	GPD – P95	138.32	156.57	169.58	197.34	208.31
	GEV	45.66	68.39	85.83	132.33	155.94
MAM	GPD – P95	183.39	235.53	281.24	412.8	482.61
	GEV	64	104.89	134.84	209.92	245.9
JJA	GPD – P95	87.75	104.37	117.35	148.92	163.16
	GEV	26.22	44.7	53.53	101.42	123.86
SON	GPD – P95	135.72	157.28	173.14	208.49	223.1
	GEV	65.63	97.35	117.31	158.5	174.83

DJF: Verão, MAM: Outono, JJA: Inverno e SON: Primavera

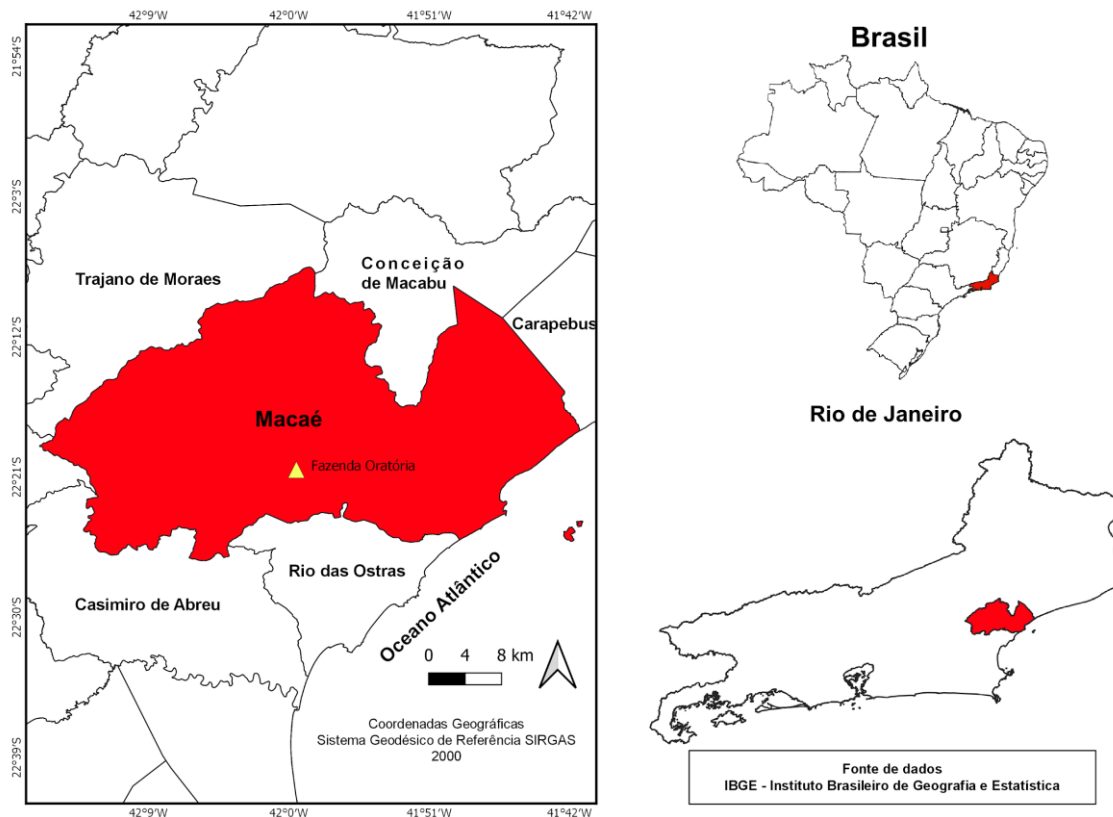
GEV: Generalized Extreme Value e GPD: Generalized Pareto Distribution

Fonte: Adaptado de Mendes (2020)

Para o método GPD, em Mendes (2019), foram feitas análises utilizando como limites os percentis 90, 95 e 99. E para este presente trabalho foi utilizado o resultados obtidos para o percentil 95, pois foi o que apresentou melhores resultados no trabalho de referência.

Os dados de nível de retorno foram alcançados através da análise de dados de precipitação em nível diário obtidos de uma estação pluviométrica localizada na Fazenda Oratória, no município de Macaé, a uma latitude S 22°15'33,12" e longitude W 41°59'3,12", na estação código 2241004, da rede hidrometeorológica da Agência Nacional de Águas – ANA, como pode ser observado na Figura 1, que destaca o município de Macaé e a localização da estação pluviométrica Fazenda Oratória. O intervalo dos dados de precipitação são de 50 anos, que compreende o período de análise de 1968 a 2018 e 1,48% de falhas.

Figura 1: Destaque do município de Macaé e da localização da estação pluviométrica fazenda Oratória.



Fonte: Elaborado pelos autores

Adotou-se o modelo estatístico *Analysis of Variance* (ANOVA), definindo por amostragem nível de retorno como a variável aleatória contínua (variável dependente), num número fixo de níveis do fator método e estação do ano. A componente de Erro ou Residual, incorporada em todo e qualquer modelo estatístico, representa o efeito de todas as variáveis não incluídas no design experimental e reflete toda a variação cuja fonte é desconhecida (não explicada pelo modelo). A abordagem estatística de Análise de Variância testou a significância de cada um dos efeitos, verificando quais os componentes que definem o modelo mais ajustado aos dados, à um nível de significância pré-estabelecido (α).

O modelo estatístico ANOVA pressupõe que os erros (residuais) são variáveis aleatórias independentes (não covariam com as variáveis estudadas), cuja distribuição é assumida como normal, centrada em zero, e com variância constante em todas as condições estudadas. Estes pressupostos têm consequências diretas

no grau de adequação (validade e precisão) com que o modelo ANOVA representa os dados coletados, pelo que se tornou necessário garantir a sua verificação.

Adicionalmente, foi realizada a análise *post hoc* através do Teste de Tukey de comparações múltiplas, também conhecido como *Tukey Honest Significant Difference (HSD)*, procedimento estatístico necessário para entendimento de padrões que não foram determinados a priori na ANOVA.

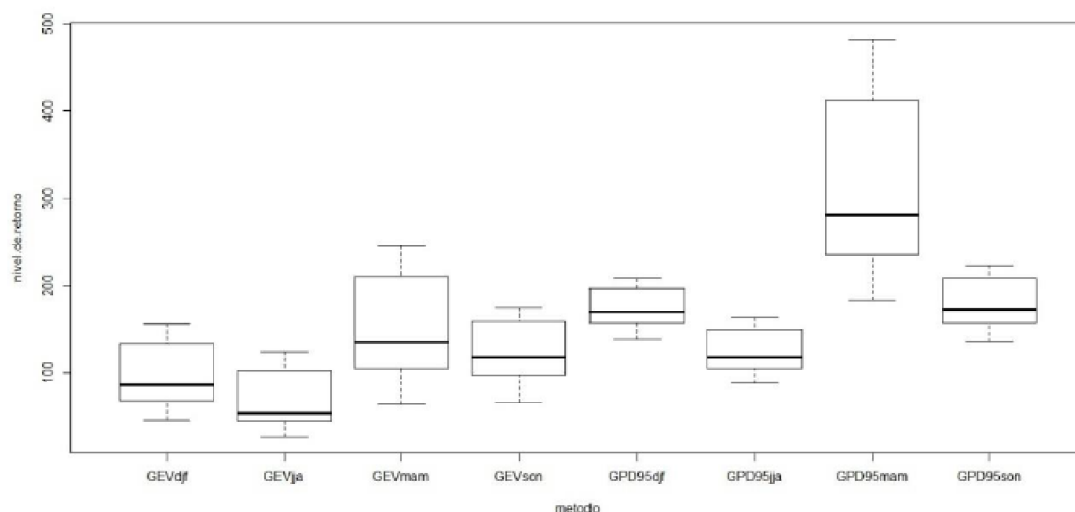
As análises estatísticas foram implementadas no ambiente *R statistical software*, que se torna atrativo por ser interativo, gratuito, *open source* e permitir aos usuários acesso às novas metodologias de ciência de dados desenvolvidas.

Para o leitor interessado na discussão sobre ANOVA, análise *post hoc* e linguagem R recomendamos a leitura dos trabalhos de: Scheffé (1959); Altman e Krzywinski (2015); Tukey (1949) e Becker, Chambers e Wilkset (1988).

2.2 Resultados e Discussão

O Boxplot na Figura 2 apresenta a análise comparativa dos métodos de estimação de eventos extremos de precipitação no município de Macaé – RJ. O método GPD 95 para a estação do outono teve a maior média do nível de retorno, igual a 319,14 mm, e o método GEV para o inverno a menor média de nível de retorno, no valor de 69,95 mm. A análise por estação revela que os valores são maiores no método GPD95. Podemos observar que a estação do outono apresenta uma variabilidade dos dados muito alta nos dois métodos, principalmente no método GPD 95, se comparado com as outras estações. Nota-se que existe uma diferença significativa entre os dois métodos para todas as estações.

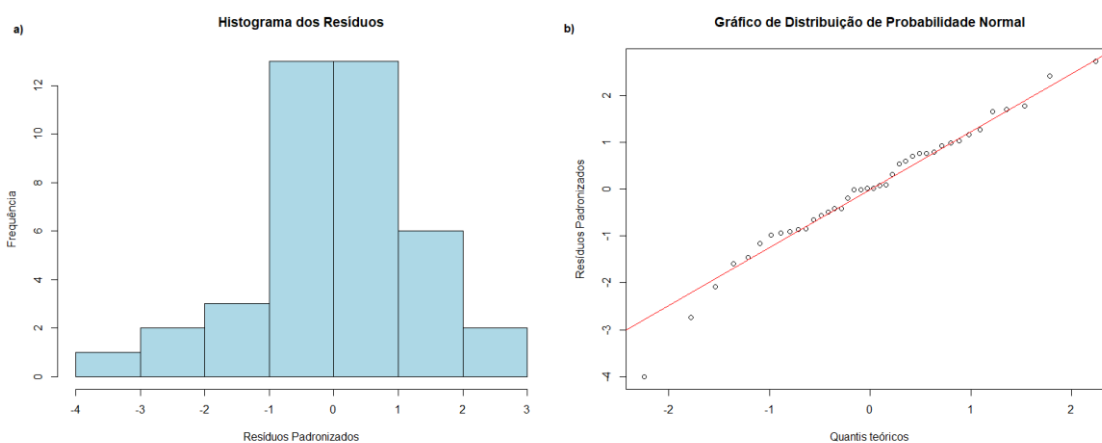
Figura 2: Análise de Variabilidade dos métodos por estação.



Fonte: Elaborado pelos autores (2020).

O Histograma dos Resíduos e o Gráfico de Probabilidade Normal (Figura 3), demonstram o ajuste dos resíduos à Distribuição de Probabilidade Normal. A hipótese de normalidade foi confirmada através do Teste de Shapiro-Wilk ao nível de significância de 5% ($p\text{-value}=0,7168$).

Figura 3: Avaliação da distribuição dos dados - a) Histograma e b) Normal QQ Plot.



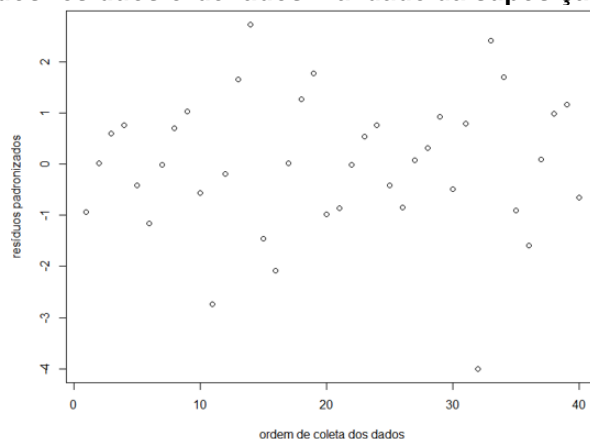
Fonte: Elaborado pelos autores (2020).

O Teste de Bartlett de homogeneidade de variâncias forneceu estatísticas para afirmar um padrão de variabilidade semelhante dos resíduos ($p\text{-value} =$

0,06035), validando a suposição de homocedasticidade ao um nível de significância de 5%.

O Gráfico dos resíduos ordenados (Figura 4), para detectar autocorrelação residual, apresenta distribuição aleatória dos resíduos em torno de zero, havendo ausência de padrão sistemático.

Figura 4: Gráfico dos resíduos ordenados: validade da suposição de independência.



Fonte: Elaborado pelos autores (2020).

A Análise de Variância, resumida na Tabela 2, evidencia que existe diferença estatisticamente significativa entre os níveis de precipitação nos dois tipos de métodos e por estação (p-valor = 0,00005264) considerando um nível de significância de 5%.

Tabela 2. ANOVA comparando as médias dos níveis de precipitação por método e estação.

Fatores	df	SQ	MQ	F	p-valor
Método/Estação	7	198903	28414,7	6,8581	5,264e-05
Resíduos	32	132584	4143,3		
Total	38	331.487			

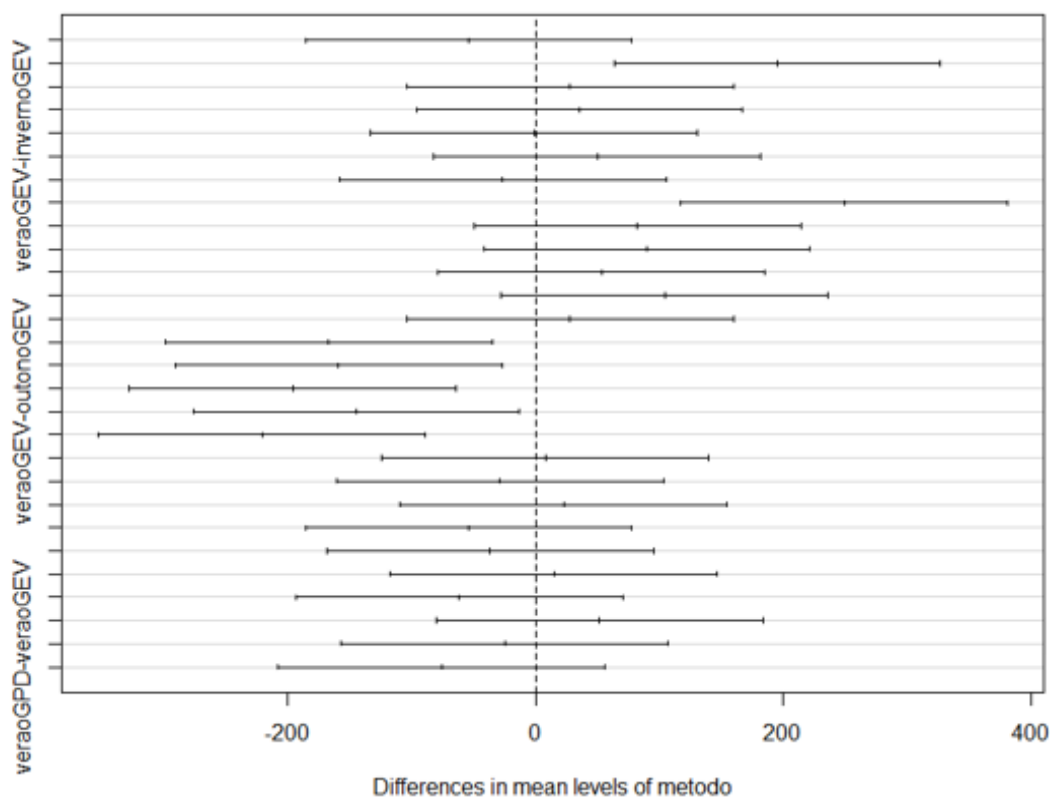
df: graus de liberdade, SQ: soma quadrática; MQ: média quadrática, F: F calculado

Fonte: Elaborado pelos autores (2020).

A Análise *post hoc* realizada através do *Tukey Honest Significant Difference (HSD)* está apresentada na Figura 5. As comparações múltiplas entre os métodos

GEV e GPD, por estação, demonstraram onde se encontram as diferenças entre os níveis médios de precipitação. Podemos observar as diferenças estatisticamente significativas entre os níveis médios de precipitação quando o Intervalo de Confiança 95% não contém o 0.

Figura 6. Análise post hoc de comparações múltiplas entre os métodos/estação



Fonte: Elaborado pelos autores (2020).

É possível observar na Tabela 3, que destaca as comparações que apresentaram diferenças significativas, que a estação outono pelo método GEV esteve presente em todas as comparações, mostrando que este método apresentou valores significativamente diferentes em comparação com as outras estações e métodos.

Tabela 3. Comparações do Teste Tukey que apresentaram diferenças significativas.

	Diferença	LimInf	LimSup	p-valor
outonoGEV- invernoGEV	194.804	62.93228	326.67572	0.0008717
outonoGEV- invernoGPD	249.168	117.29628	381.03972	0.0000195
outonoGPD- outonoGEV	-167.204	-299.07572	-35.33228	0.0056159
primaveraGEV- outonoGEV	-159.568	-291.43972	-27.69628	0.0092266
primaveraGPD- outonoGEV	-196.390	-328.26172	-64.51828	0.0007814
veraoGEV- outonoGEV	-145.090	-276.96172	-13.21828	0.0228874
veraoGPD- outonoGEV	-221.484	-353.35572	-89.61228	0.0001360

Fonte: Elaborado pelos autores (2020).

Tal resultado corrobora os resultados obtidos no trabalho de Mendes (2019), que mostrou em seus testes (*boxplot* e *QQplot*) que o outono foi a estação que apresentou maior variabilidade em seus valores analisados. Mendes (2019) também descobriu que a estação do outono é a que apresenta os eventos mais extremos para o município de Macaé, o que mais uma vez mostra o porquê do outono se destacar nos dois trabalhos.

Os resultados dos testes de ANOVA e Tukey se mostraram muito relevantes, uma vez que permitem identificar a existência de diferenças significativas entre os métodos GEV e GPD, o que já tinha sido observado por Mendes (2019), porém não estatisticamente. Mendes (2019) concluiu em seu trabalho que o melhor método para representar os níveis e tempo de retorno seria o método GEV, uma vez que este se apresentou mais próximos dos valores reais observados e o GPD se apresentava com valores discrepantes em comparação com o que já fora observado.

3 CONSIDERAÇÕES FINAIS

As análises estatísticas utilizadas no presente trabalho foram pertinentes para a compreensão da variabilidade entre os resultados de níveis de precipitação e tempo de retorno calculados através dos métodos utilizados na Teoria dos Valores Extremos, GEV (*Generalized Extreme Value*) e GPD (*Generalized Pareto Distribution*). Na qual foi possível constatar que existe diferença significativa entre os níveis de precipitações calculados pelos métodos e por estação, considerando um nível de significância de 5%, recorrendo à análise de variância ANOVA.

Também se observou que todas essas diferenças significativas envolvem a estação do outono. Esse resultado corrobora os resultados obtidos nos testes realizados no trabalho de Mendes (2019), onde se avaliou que a estação de outono possuía uma maior variabilidade dos dados, tanto no método de GEV e quanto no GPD. Recomenda-se estudos similares com outros métodos que envolvam a Teoria dos Valores Extremos.

Com efeito, os resultados do estudo podem auxiliar em análises futuras similares servindo de base e corroborando com a escolha do melhor método para entender a variação e o comportamento de eventos com grandes volumes pluviométricos em curta duração de tempo. Sendo portanto, relevante para o planejamento e implementação de ações de políticas públicas para evitar problemas oriundos de tais eventos.

REFERÊNCIAS

Altman & Krzywinski (2015). **Points of Significance:** Simple linear regression. *Nature Methods* 12, 999–1000.

BAUTISTA, E. A. L.; ZOCCHI, S. S.; ANGELOCCI, L. R. **A distribuição generalizada de valores extremos aplicada ao ajuste dos dados de velocidade máxima do evento em Piracicaba**, São Paulo, Brasil. *Rev. Mat. Estat.*, São Paulo, v.22, n.1, p.95-111, 2004.

BECKER, R.A., CHAMBERS, J.M., & WILKS, A.R. (1988). **The new S language:** a programming environment for data analysis and graphics.

Revista Mundi Engenharia, Tecnologia e Gestão. Paranaguá, PR, v.6, n.2, p. 334-01, 334-14, 2021.

DOI: 10.21575/25254782rmetg2021vol6n21381

BURIOL, G.A.; ESTEFANEL, V.; SWAROWSKY, A.; D'AVILA, R.F.; HELDWEIN, A.B. **Homogeneidade e estatísticas descritivas dos totais mensais e anuais de chuva de Santa Maria, estado do Rio Grande do Sul**. Revista Brasileira de Recursos Hídricos, v.11, n.4, p.89-97, 2006.

CASTRO, A. L. C. Secretaria Nacional De Defesa Civil. Glossário De Defesa Civil Estudos De Riscos E Medicina De Desastres, 2ª Edição (1998).

CEPED UFSC – CENTRO UNIVERSITÁRIO DE ESTUDOS E PESQUISAS SOBRE DESASTRES DA UNIVERSIDADE FEDERAL DE SANTA CATARINA. Atlas brasileiro de desastres naturais 1991 a 2012: volume Rio de Janeiro, 2ª ed, Florianópolis, 2013.

FARIA, T. P.; TOUGEIRO, J. V. **Conflitos socioambientais motivados por ocupação de manguezais e restingas para fins habitacionais no espaço urbano de Macaé, RJ**. R. Inter. Interdisc. INTERthesis, Florianópolis, v.7, n.1, p. 242-271, jan./jul. 2010.

KOTZ, S; NADARAJAH, S. Extreme Value Distributions: Theory and Applications. London; Imperial College Press, 2000

LIMA, ROVERBAL MONTEIRO BEZERRA DE. **Técnicas de experimentação estatística** [Em linha]: planos factoriais e aplicações. Lisboa: [s.n.], 2010. 85 p.

MENDES, I. V. S. **Tendência e Período de Retorno das Chuvas Intensas em Macaé-Rj: Detecção e Impactos**. Dissertação (Especialização em Clima, Água e Energia: uma Abordagem Técnica e Integrada) – Laboratório de Meteorologia do Centro de Ciência e Tecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro. Rio de Janeiro, 2019.

PICKANDS, J. Statistical inference using extreme order statistics. **Annals of Statistics**, v. 3, n. 1, p. 119–131, 1975.

SANTOS, E. B.; LUCIO, P. S.; SILVA, C. M. S. Analysis of return periods for maximum daily precipitation in the Brazilian Amazon. **Journal of Hydrometeorology**, v. 16, p. 973–984, 2015.

SCHEFFÉ, H. (1959). Analysis of variance. New York: Wiley.

SOUZA, W. M. de; AZEVEDO, P. V. de. Índices de Detecção de Mudanças Climáticas Derivadas da Precipitação Pluviométrica e das Temperaturas em Recife - PE. **Revista Brasileira de Geografia Física**, v. 5, n. 1, p. 143-159, 2012.

TUKEY, J. W. Comparing individual means in the analysis of variance. Biometrics vol. 5,2 (1949): 99-114.

Enviado em: 24 jul. 2020

Aceito em: 19 mai. 2021

Editores responsáveis: Bianca Neves Machado / Mateus das Neves Gomes