

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COECI – COORDENAÇÃO DO CURSO DE ENGENHARIA CIVIL
CURSO DE ENGENHARIA CIVIL

BRUNO TOHORU SHIOTANI

**ESTUDO DAS VAZÕES MÁXIMAS DA BACIA DO RIO IGUAÇU UTILIZANDO A
DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS E MOMENTOS-LH**

TRABALHO DE CONCLUSÃO DE CURSO

TOLEDO
2019

BRUNO TOHORU SHIOTANI

**ESTUDO DAS VAZÕES MÁXIMAS DA BACIA DO RIO IGUAÇU UTILIZANDO A
DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS E MOMENTOS-LH**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel, no curso de Engenharia Civil, da Universidade Tecnológica Federal do Paraná – Campus Toledo.

Orientador: Prof. Dr. Wagner Alessandro Pansera

TOLEDO

2019



TERMO DE APROVAÇÃO

Título do Trabalho de Conclusão de Curso de Nº 187

ESTUDO DAS VAZÕES MÁXIMAS DA BACIA DO RIO IGUAÇU UTILIZANDO A DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS E MOMENTOS-LH

por

Bruno Tohoru Shiotani

Este Trabalho de Conclusão de Curso foi apresentado às 13:50 h do dia **11 de Junho de 2019** como requisito parcial para a obtenção do título **Bacharel em Engenharia Civil**. Após deliberação da Banca Examinadora, composta pelos professores abaixo assinados, o trabalho foi considerado **APROVADO**.

Prof Dr. Benedito Martins Gomes
(UNIOESTE - CV)

Profª Mª Ana Claudia Bergmann
(UTFPR – TD)

Prof Dr. Wagner Alessandro Pansera
(UTFPR – TD)
Orientador

Visto da Coordenação
Prof. Dr Fúlvio Natércio Feiber
Coordenador da COECI

AGRADECIMENTOS

Primeiramente, ao meu professor orientador, Prof. Dr. Wagner Alessandro Pansera, que se dedicou à realização desse estudo, aconselhando e corrigindo várias vezes para que mantivesse um trabalho de qualidade.

Aos meus pais, Elizabeth Miyuki Mitui Shiotani e Ricardo Massahiro Shiotani, que nunca deixaram de apoiar minhas decisões e sempre me acalmaram em momentos difíceis.

Aos meus amigos, que sempre me ajudaram e me acompanharam, sempre me mantendo feliz por tê-los ao meu lado.

Aos meus companheiros de classe, por terem trilhado esse caminho junto comigo.

E por último, porém não menos importante, ao meu irmão mais velho, André Kazuhiro Shiotani, que serviu como fonte de inspiração durante todo o curso.

RESUMO

SHIOTANI, Bruno Tohoru. **Estudos das vazões máximas da bacia do rio Iguaçu utilizando a distribuição generalizada de valores extremos e momentos-Lh.** 40 f. Trabalho de conclusão de curso (Graduação) – Bacharelado em Engenharia Civil, Universidade Tecnológica Federal do Paraná. Toledo, 2019.

O rio Iguaçu é considerado o maior curso de água no estado do Paraná, além disso, ele contém vários barramentos ao longo de sua extensão. Portanto, estudos sobre as vazões máximas no rio Iguaçu são muito importantes. Por isso, este trabalho fez um estudo das vazões máximas anuais utilizando a distribuição Generalizada de Valores Extremos (GEV), sendo seus parâmetros e teste de qualidade de ajuste realizados utilizando Momentos-LH. Foram estudadas séries históricas de vazões máximas de 14 estações fluviométricas. A distribuição GEV ajustou-se em todas as séries estudadas ao nível de 5% de significância. Também, foram estimados os quantis para a distribuição Gumbel para efeito de comparação com a distribuição GEV. Além disso, foi avaliada a capacidade preditiva das distribuições usando bootstrapping. Observou-se que para os dados obtidos pela bacia hidrográfica do rio Iguaçu, até 100 anos de período de retorno, é possível modelar utilizando as distribuições GEV e Gumbel, e para período de retornos maiores deve ser dada preferência a distribuição GEV devido sua melhor capacidade preditiva.

Palavras-chaves: Distribuição generalizada de valores extremos. Momentos-LH. Vazões de projeto. Teste de Wang. Capacidade preditiva.

ABSTRACT

SHIOTANI, Bruno Tohoru. **Studies of the maximum flows of the Iguaçu river basin using the generalized distribution of extreme values and moments-lh.** 40 p. Graduation Work - Bachelor of Science in Civil Engineering, Federal Technological University of Paraná. Toledo, 2019.

The Iguaçu River is considered the largest watercourse in the state of Paraná, in addition, it contains several buses along its length. Therefore, studies on the maximum flows in the Iguaçu river are very important. Therefore, this study made a study of the maximum annual flows using the Generalized Distribution of Extreme Values (GEV), and its parameters and quality of fit test were performed using Moments-LH. Fourteen fluvimetric stations were studied. From these, it was estimated the parameters of the GEV distribution and the Wang test, which showed that all the stations were adjusted with 5% significance. Next, the quantiles for the Gumbel distribution were estimated for comparison with the GEV distribution. In addition, the predictive capacity of the distributions was evaluated using bootstrapping. Finally, the study showed that up to 100 years of return period, it is possible to model the data obtained in the region of the Iguaçu river basin using the GEV and Gumbel distributions, and for a longer return period the GEV distribution should be given preference due to its better capacity predictive.

Keywords: Generalized Extreme Value distribution. LH-moments. Project flow. Wang's test. Predictive capacity.

LISTA DE FIGURAS

Figura 1 - Exemplo de hidrograma para vazões máximas anuais.....	12
Figura 2 - Mapa de enchentes no mundo.....	13
Figura 3 - Inferência estatística.	15
Figura 4 - Distribuição das três formas de valores extremos representados pela GEV.	18
Figura 5 - Bacia do rio Iguaçu.	21
Figura 6 - Roteiro do estudo.....	22
Figura 7 - Postos fluviométricos e usinas hidrelétricas na bacia do Rio Iguaçu.	23
Figura 8 – Análise gráfica da estação 2.	28
Figura 9 – Número de séries para cada nível dos momentos-LH.	30
Figura 10 - Comparação entre as distribuições GEV e Gumbel na estação 2.	33
Figura 11 - Comparação entre as distribuições GEV e Gumbel na estação 3.	33
Figura 12 - Comparação entre as distribuições GEV e Gumbel na estação 11.	34
Figura 13 - Comparação entre as distribuições GEV e Gumbel na estação 5.	34
Figura 14 - Comparação entre as distribuições GEV e Gumbel na estação 12.	35
Figura 15 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 2.	36
Figura 16 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 3.	37
Figura 17 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 5.	37
Figura 18 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 11.	38
Figura 19 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 12.	38

LISTA DE TABELAS

Tabela 1 – Informações dos postos fluviométricos.....	23
Tabela 2 – Resultados da análise estatística.	28
Tabela 3 - Ajustes da distribuição GEV às séries de vazão máxima anual.....	29
Tabela 4 - Tipo de distribuição de valores extremos.	31
Tabela 5 - Resultados da determinação dos quantis da distribuição GEV e Gumbel.	32

SUMÁRIO

1 INTRODUÇÃO	10
1.1 JUSTIFICATIVA	11
1.2 OBJETIVOS	11
1.2.1 Geral.....	11
1.2.2 Específicos	11
2 REVISÃO BIBLIOGRÁFICA	12
2.1 AS VAZÕES MÁXIMAS ANUAIS	12
2.2 ANÁLISE DE FREQUÊNCIA.....	14
2.3 DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS (GEV)	16
2.4 ESTIMATIVA DE PARÂMETROS	18
2.5 SELEÇÃO DO MELHOR MODELO DE DISTRIBUIÇÃO	19
3 MATERIAIS E MÉTODOS	21
3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO.....	21
3.2 MÉTODOS	21
3.2 POSTOS FLUVIOMÉTRICOS.....	22
3.3 ANÁLISE DE FREQUÊNCIA.....	24
3.3.1 Distribuição Generalizada de Valores Extremos com Momentos-LH	24
3.3.2 Distribuição Gumbel pelo teorema de Ven Te Chow.....	25
3.4 CAPACIDADE PREDITIVA	26
4. RESULTADOS E DISCUSSÕES	27
4.1 ESTATÍSTICA DESCRITIVA.....	27
4.2 ESTIMATIVA DOS PARÂMETROS DA DISTRIBUIÇÃO GEV	29
4.3 DETERMINAÇÃO DOS QUANTIS	31
4.4 CAPACIDADE PREDITIVA DAS DISTRIBUIÇÕES GEV E GUMBEL	36
CONCLUSÃO	39
REFERÊNCIAS	40

1 INTRODUÇÃO

O rio Iguaçu é considerado o maior rio totalmente paranaense, tendo 1320 quilômetros de extensão e possui vários rios adjacentes contribuindo com sua vazão (ÁGUAS PARANÁ, 2019). Nele, foram instalados vários barramentos espalhados pela sua extensão e em seus contribuintes, dos quais se identificam 10 pequenas centrais hidrelétricas (PCH's) (COPEL, 2019).

Para a operação dessas obras hidráulicas é importante realizar a análise de frequência de vazões máximas. Ou seja, necessita-se utilizar métodos estatísticos para prever valores que são utilizados para dimensionamento, como a vazão de operação.

Sendo assim, esse trabalho realizou um estudo utilizando séries históricas de vazões máximas, em postos fluviométricos localizados na bacia do rio Iguaçu. A modelagem estocástica para as séries de vazões máximas foi realizada utilizando a distribuição Generalizada de Valores Extremos (GEV) utilizando os estimadores de momentos-LH.

É muito comum o uso da distribuição Gumbel para se dimensionar obras hidráulicas, mediante a aplicação do Teorema de Ven Te Chow, devido a sua praticidade e facilidade computacional. Porém, Watanabe (2013) indica que existem distribuições que podem apresentar desempenho melhor que a Gumbel.

Watanabe (2013) também mostra em seu estudo, que a distribuição Gumbel é capaz de trazer bons resultados para o dimensionamento de estruturas de baixo período de retorno, como pontes e microdrenagens (100 e 25 anos de período de retorno, respectivamente). Entretanto, para dimensionar estruturas de período de retorno alto, a distribuição GEV com momentos-LH pode gerar resultados mais precisos, como apontado por do Queiroz e Chaudhry (2006).

Dessa forma, foi realizado um comparativo entre as distribuições GEV e Gumbel no estudo das vazões máximas na bacia do rio Iguaçu. Além disso, foi avaliada a capacidade preditiva das distribuições Gumbel e GEV usando bootstrapping.

Para isto, este trabalho foi dividido em cinco etapas: coleta de dados, análise descritiva, seleção do modelo, estimativa de quantis e capacidade preditiva. Onde cada etapa pode indicar resultados que ajudarão na compreensão do estudo.

1.1 JUSTIFICATIVA

Na bacia do Rio Iguaçu há dez reservatórios de usinas hidrelétricas, sendo elas: usina hidrelétrica Governador José Richa; Arturo Andreoli; Salto Osório; Salto Santiago; Cavernoso II; Governador Ney Aminthas de Barros Braga; Derivação do Rio Jordão; Fundão; Santa Clara e Governador Bento Munhoz da Rocha Netto.

No dimensionamento de reservatórios é feito uma previsão de vazão máxima para operação, para isso é considerado uma vazão máxima para um determinado período de retorno associado, pois apresenta um risco menor em casos de eventos extremos.

A importância em identificar as vazões máximas, estudadas neste trabalho, é em função da quantidade de usinas hidrelétrica (UH) que se encontram ao longo da bacia do rio Iguaçu, tal importância é fundamentada na hipótese de haver um evento extremo causando o rompimento de uma dessas UH, ocasionando um efeito em cadeia ao longo de toda a extensão da bacia.

Tal desastre poderá resultar em perdas materiais e no pior dos casos perdas de vidas, arruinando a fauna e flora de rios adjacentes e podendo afetar de forma desastrosa a vida daqueles que foram afetados pelo rompimento de uma barragem.

Assim, o foco deste trabalho é analisar e modelar as vazões do rio Iguaçu utilizando a distribuição GEV com momentos-LH e comparando com o teorema de Ven Te Chow. Oferecendo assim, um estudo capaz de minimizar riscos para o dimensionamento de uma obra hidráulica localizada na região de estudo.

1.2 OBJETIVOS

1.2.1 Geral

Modelar as vazões máximas na bacia hidrográfica do rio Iguaçu utilizando a distribuição GEV e momentos-LH.

1.2.2 Específicos

- Avaliar a capacidade preditiva da distribuição GEV; e
- Comparar as distribuições Gumbel e GEV.

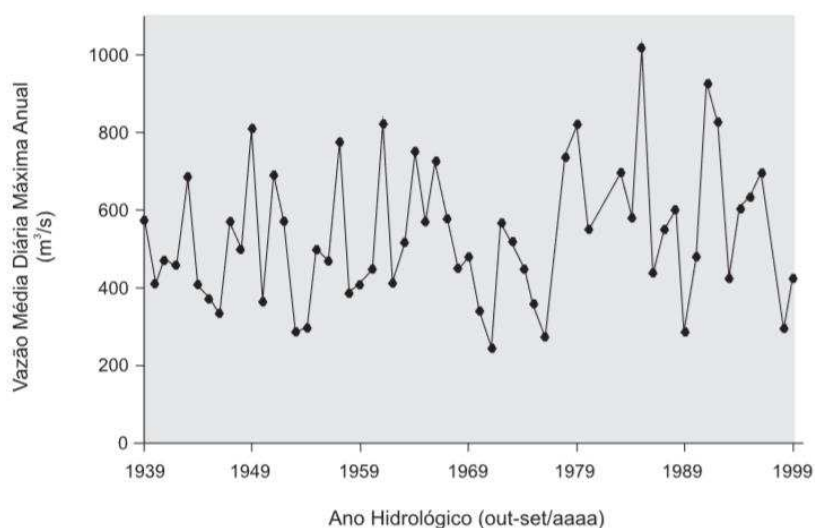
2 REVISÃO BIBLIOGRÁFICA

Neste capítulo apresenta-se um resumo da bibliografia estudada que se considerou mais relevante para o desenvolvimento do trabalho. Assim, serão abordados os temas: as vazões máximas anuais, análise de frequência, generalizada de valores extremos (GEV), momentos LH e escolha da melhor distribuição.

2.1 AS VAZÕES MÁXIMAS ANUAIS

Collischonn e Dornelles (2015) conceituam vazão, ou caudal, como o volume de água que passa por uma determinada seção de um rio ao longo de uma unidade de tempo, causada pela interação entre a precipitação e a bacia, e depende das características da bacia que influenciam a infiltração, o armazenamento e a evapotranspiração. Sendo necessário também, realizar medições para eventos extremos, como cheias, permitindo o dimensionamento de estruturas ou medidas preventivas. Os autores indicam que os métodos de estimativa de vazões máximas a partir da chuva dependem do tamanho da bacia e que a vazão de um rio é uma variável que se modifica de forma contínua no tempo, podendo ser representado por um hidrograma como o da Figura 01.

Figura 1 - Exemplo de hidrograma para vazões máximas anuais.

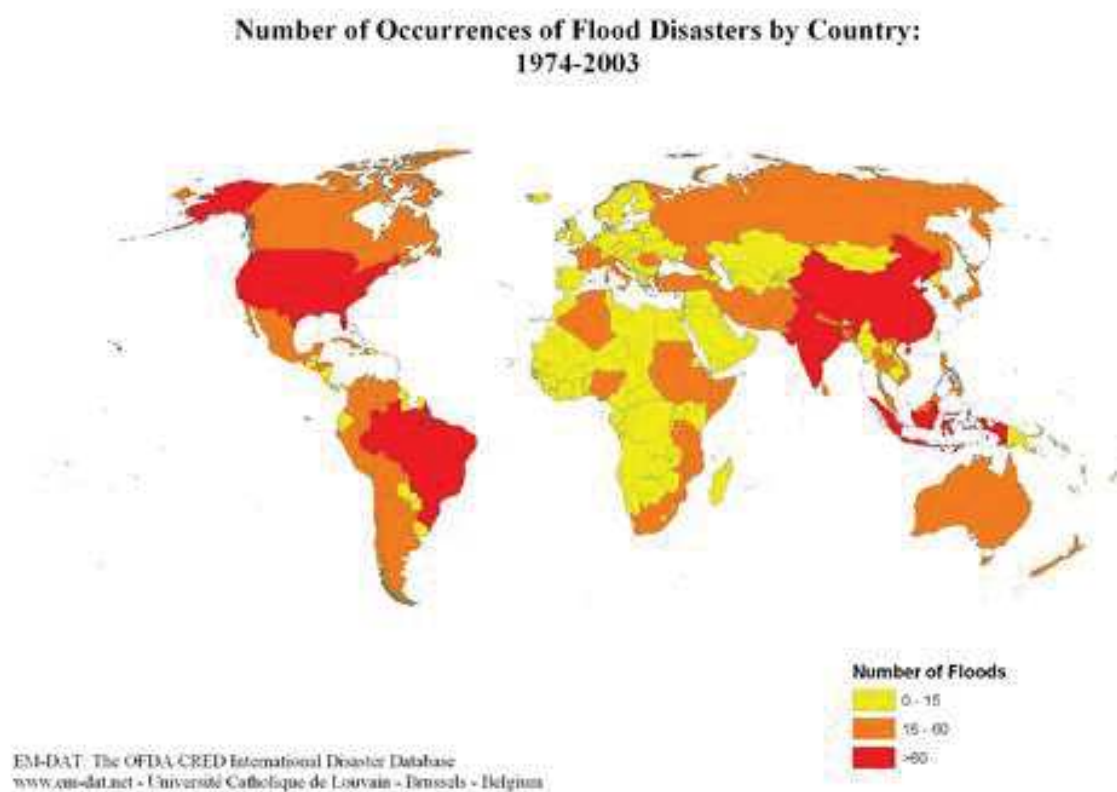


Fonte – Adaptado Naguettini e Pinto (2007).

Quando se dispõe de séries históricas de vazão de cursos d'água, tem-se a necessidade de trabalhá-las adequadamente por meio de análises probabilísticas e / ou estocásticas (Haan, 2002; Douglas & Vogel, 2006).

De acordo com a base de dados disponibilizado no site Emergency Disasters Data Base (EM-DAT) (2019), desastres por cheias são bem frequentes em determinadas regiões.

Figura 2 - Mapa de enchentes no mundo.



Fonte - EM-DAT (2019).

O mapa apresentado na Figura 02, retirado do site Emergency Disasters Data Base (EM-DAT) descreve casos de enchente de 1974 até 2003, evidenciando que há uma frequência de cheias. Podemos destacar também o Brasil que representou um dos países com mais ocorrências de cheias, tendo mais que 60 cheias dentro deste período de análise. Isso demonstra a importância na análise de dados de forma mais minuciosa para dimensionamento de barragens e outras estruturas relacionadas aos eventos hidrológicos.

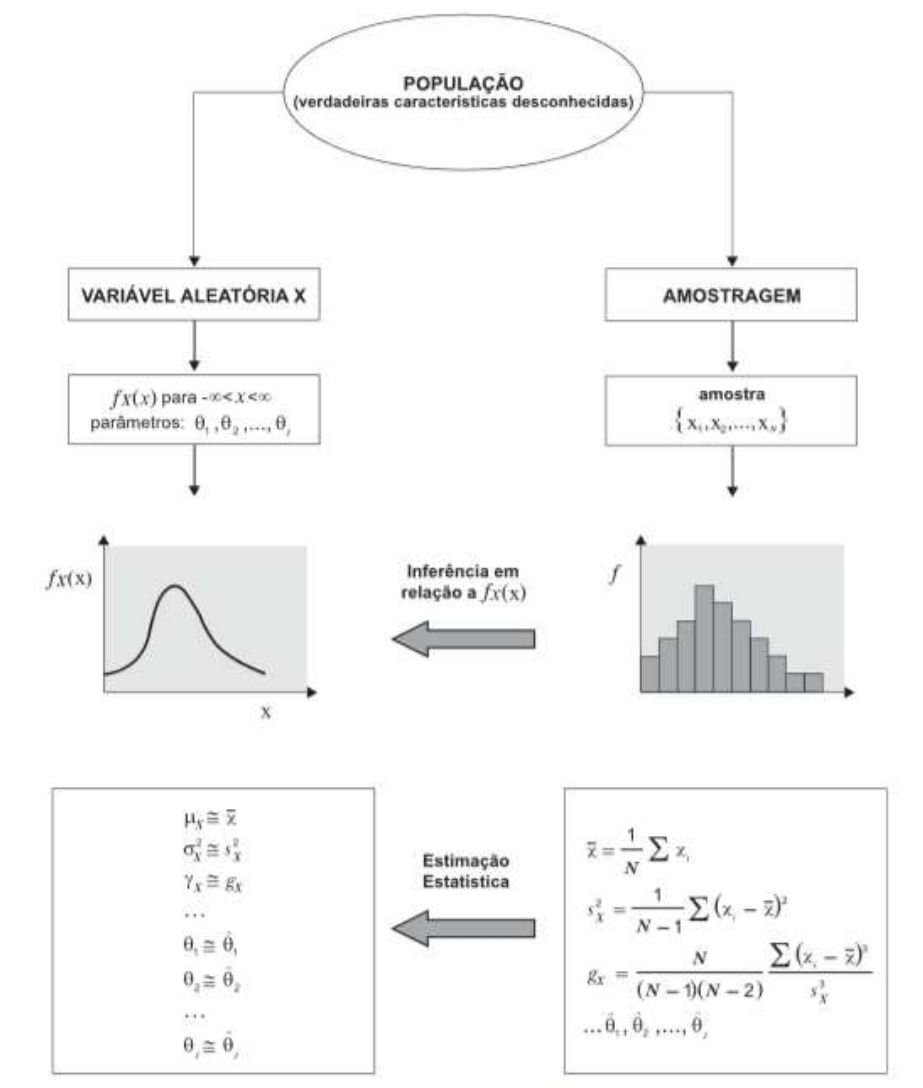
2.2 ANÁLISE DE FREQUÊNCIA

Como o comportamento exato das vazões de rios não podem ser previstos, as leis da probabilidade são empregadas de modo a se estabelecer as prováveis variações de vazões. A análise estatística para a determinação de vazões de eventos extremos emprega os vários dados amostrais observados num dado período, utilizando-os como meio para se efetuar a probabilidade de ocorrência em um período diferente, assim determinando possíveis previsões para futuros eventos (MARTINS et al., 2011).

De acordo com Naghettini e Pinto (2007) assim que se tem posse de apenas amostras finitas de variável aleatória, comum em hidrologia, deve-se obter conclusões sobre o modelo distributivo da população que contém a amostra e as estimativas dos valores numéricos dos parâmetros que descrevem o modelo distributivo.

A Figura 03 representa a estimação de parâmetros, onde o termo 'estimação' é usado livremente para significar o ato de produzir estimativas de parâmetros populacionais, a partir de uma amostra (Naghettini e Pinto, 2007). Esses autores dizem que eventos extremos, tais como tempestades, grandes cheias e secas impactam os sistemas hidrológicos e que a magnitude de um evento extremo é inversamente relacionada à sua frequência de ocorrência. O objetivo da análise de frequência das variáveis hidrológicas é relacionar a magnitude dos eventos com sua frequência da ocorrência por meio do uso de uma distribuição de probabilidade.

Figura 3 - Inferência estatística.



Fonte- Adaptado Naghettini e Pinto (2007).

Liou et al. (2008) explicam que existem três etapas quando for realizar a análise de frequência. Primeiramente obter uma amostra aleatória; segundo, achar a distribuição de melhor ajuste para a amostra a partir de teste de qualidade de ajuste, segundo os autores, esta etapa é subdividida em outras duas etapas: escolher o método estimativa de parâmetros e selecionar a melhor distribuição com o melhor ajuste; e terceiro, determinar o quantil de não-excedência utilizando a melhor distribuição.

2.3 DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS (GEV)

A distribuição de probabilidade generalizada de valores extremos (GEV) é utilizada para modelar uma grande variedade de extremos naturais, como cheias, chuvas, velocidade do vento, temperaturas e outros (Martins e Stedinger, 2000; Queiroz, 2002).

Jenkinson (1955) introduziu a GEV, descrevendo que esta distribuição combina os três possíveis tipos de distribuição de valores extremos em uma única forma, sendo utilizada principalmente em análise de frequência de cheias, onde é bem aceita por descrever, de forma fiel, o fluxo máximo de cheias anuais.

A distribuição GEV engloba três formas assintóticas de distribuição de valores extremos conhecidos como valor extremo tipo I (VEI), valor extremo tipo II (VEII) e valor extremos tipo III (VEIII) (Naghetini e Pinto, 2007).

Queiroz (2002) adaptou as equações (1) e (2) apresentando a função da probabilidade acumulada da distribuição GEV, onde a primeira é dada a partir de valores do parâmetro de forma igual diferente de zero (ou seja, $k \neq 0$);

$$F_y(y) = \exp \left[- \left(1 - k \frac{y - \beta}{\alpha} \right)^{\frac{1}{k}} \right] \quad (1)$$

Caso o parâmetro de forma assintótica seja igual a zero (ou seja, $k = 0$), tem-se a segunda equação;

$$F_y(y) = \exp \left[- \exp \left(- \frac{y - \beta}{\alpha} \right) \right] \quad (2)$$

Em que:

$F_y(y)$ = função probabilidade acumulada da distribuição GEV;

α = parâmetro de escala;

β = parâmetro de posição;

k = parâmetro de forma.

O parâmetro de forma (k) indica a forma assintótica da distribuição e é caracterizado por:

a) $k = 0$, distribuição Tipo I ou Gumbel;

- b) $k < 0$, distribuição Tipo II, definida para $y > \frac{\beta + \alpha}{k}$;
- c) $k > 0$, distribuição Tipo III, definida para $y < \frac{\beta + \alpha}{k}$.

Stedinger et al. (1993 apud Queiroz, 2002) dizem que na prática, quando o parâmetro de forma (k) se encontra entre os intervalos de $-0,03$ e $0,03$ a forma da distribuição é a de Gumbel, classificando-se na forma VEI.

Queiroz (2002) adaptou as equações (3) e (4) para a determinação do quantil da distribuição generalizada de valores extremos respectivamente para $k \neq 0$ e $k = 0$.

$$x(F) = \beta + \frac{\alpha}{k} \{1 - [-\ln(F)]^k\} \quad \text{para } k \neq 0, 0 < F < 1 \quad (3)$$

$$x(F) = \beta - \alpha \ln[-\ln(F)] \quad \text{para } k = 0, 0 < F < 1 \quad (4)$$

Em que:

$x(F)$ = quantil;

α = parâmetro de escala;

β = parâmetro de posição;

k = parâmetro de forma.

Considerando a equação 2, define-se uma variável reduzida (z), chamada variável reduzida de Gumbel, como $z = \frac{x-\beta}{\alpha}$ substituindo z em (2), obtém-se:

$$F_y(y) = \exp[-\exp(-z)]$$

Logo

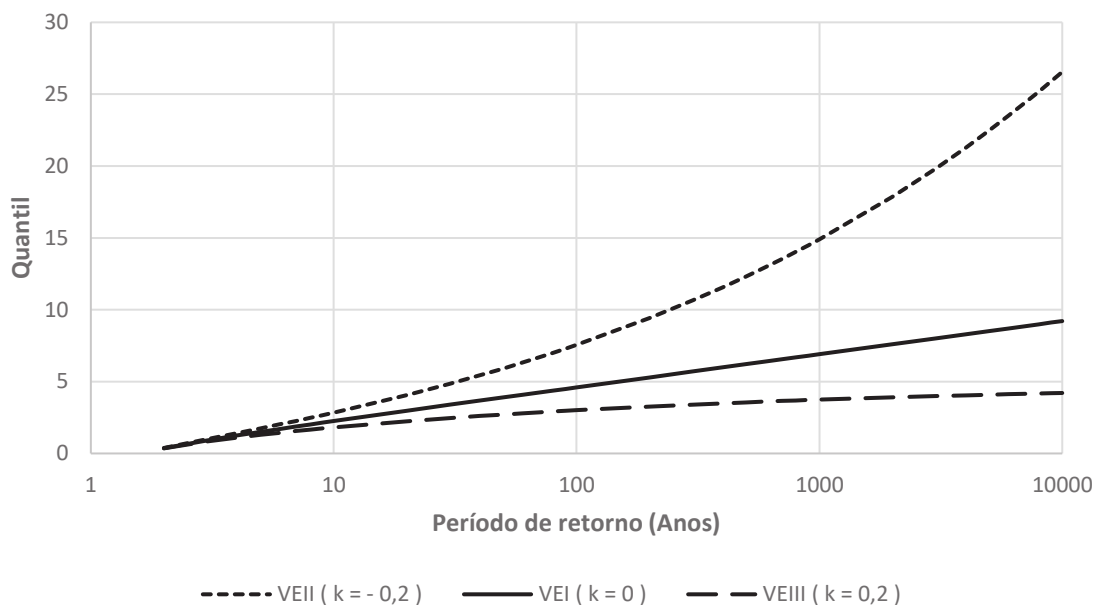
$$z = \ln\{-\ln[F_y(y)]\} \quad (5)$$

Onde, de acordo com Queiroz (2002), neste caso, a variável reduzida de Gumbel (z), também se relaciona com o período de retorno (T), $T = 1/F_y(y)$. Além disso, a equação (5) pode ser utilizada para definir z com respeito às distribuições VEI, VEII e VEIII.

Tais parâmetros são bem representados na Figura 04 a seguir, onde o gráfico demonstra as três formas que a GEV pode obter de acordo com o parâmetro de forma (k), para $k = -0,2$ (VEII), $k = 0$ (VEI ou Gumbel) e $k = 0,2$ (VEIII), parâmetro de escala

$\alpha = 1$ e parâmetro de posição $\beta = 0$, tal gráfico define o comportamento das três formas de distribuições de valores extremos.

Figura 4 - Distribuição das três formas de valores extremos representados pela GEV.



Fonte – Autor (2019).

Com o aumento dos valores do período de retorno, os valores de quantil do tipo VEI aumentam linearmente, para o tipo VEII os valores de x crescem de forma mais acelerada formando uma curva côncava, ambos aumentando de forma ilimitada na parte superior, e para o tipo VEIII, os valores de x crescem seguindo uma curva convexa para um limite finito (Queiroz, 2002).

2.4 ESTIMATIVA DE PARÂMETROS

Os parâmetros das distribuições são estimados de forma pontual ou por intervalos, onde o primeiro método é o mais utilizado na hidrologia. Aqueles que mais se destacam na aquisição pontual dos parâmetros são: momentos, máxima verossimilhança, momentos-L, máxima entropia, mínimos quadrados, generalizado dos momentos e momentos mistos (Naghetini e Pinto, 2007, p. 206). Esses autores dizem que o método de momentos-L (MML) produzem parâmetros de qualidade comparáveis ao método da máxima verossimilhança (MVS), com um fator positivo em exigir menor esforço computacional para a solução de sistemas de equações menos complexas. Ainda esses mesmos autores conceituam esses métodos:

- a) O método dos momentos se resume em igualar os momentos amostrais aos populacionais. O fruto desse tratamento de dados resultará nas estimativas dos parâmetros da distribuição de probabilidades em questão.
- b) O método da máxima verossimilhança consiste basicamente em maximizar uma função dos parâmetros da distribuição, conhecida como função de verossimilhança. O equacionamento para a condição de máximo resulta em um sistema de igual número de equações e incógnitas, cujas soluções produzem os estimadores de máxima verossimilhança.
- c) O método dos momentos-L (MML), para a estimação de parâmetros de distribuições de probabilidades é semelhante ao método dos momentos convencionais. O método MML de estimação de parâmetros consiste em igualar os momentos-L populacionais aos momentos-L amostrais. O resultado dessa operação produzirá as estimativas dos parâmetros da distribuição de probabilidades em questão.

Katz et al. (2002) e Queiroz & Chaudhry (2006) evidenciam que variáveis hidrológicas com distribuição assintótica, como vazões máximas e mínimas, devem ser adequadamente tratadas do ponto de vista probabilístico, com destaque nas metodologias de maior exatidão, como os dos momentos LH, o qual constituem em momentos ponderados por probabilidades, sendo uma generalização de momentos de combinações lineares de estatísticas de ordem superior.

Wang (1997) inseriu os momentos-LH como generalização dos momentos-L, com a finalidade de melhorar a estimativa de quantis, quando são usados níveis mais elevados dos momentos-L. Pansera (2013) descreveu que se desenvolveram cinco níveis para os momentos-LH (L0, L1, L2, L3 e L4), sendo que L0 corresponde aos momentos-L.

2.5 SELEÇÃO DO MELHOR MODELO DE DISTRIBUIÇÃO

Quando se ajusta a GEV à uma série de dados amostrais, além de estimar os parâmetros através de métodos como o de momentos-LH, é fundamental a execução de testes da qualidade desse ajuste. Estes são denominados, também, como testes de aderência e são essenciais para definir se uma distribuição é apropriada, ou não, para uma dada amostra. Dentre os testes de aderência, Naghettini e Pinto (2007)

frisam: qui-quadrado, de Kolmogorov-Smirnov, de Anderson-Darling e de Filliben. Já Queiroz e Chaudhry (2006) apresentam o teste de qualidade de Wang (1998) como opção para o ajuste da GEV.

Naguettini e Pinto (2007) descrevem que ao se utilizar testes de aderência tem-se que analisar duas linhas de raciocínio: a primeira é avaliar o teste de aderência, ou seja, analisar se os ajustes feitos pelo teste foram bons ou não; e a segunda é avaliar quais dos modelos que já tem aderência se demonstram mais adequado. Esses autores frisam que no segundo caso os testes de aderência só indicam se existe ajuste ou não, ou seja, não é possível fazer comparações entre testes de aderência e que sendo assim, não adéquam à seleção do melhor modelo distributivo mais adequado para uma certa amostra de dados.

Naguettini e Pinto (2007) indicam que o teste de aderência de Kolmogorov-Smirnov (KS) é um teste não paramétrico, cuja estatística de teste tem como base a diferença máxima entre as funções de probabilidades acumuladas, empírica e teórica, de variáveis aleatórias contínuas. O teste não é aplicável a variáveis aleatórias discretas. Estes mesmos autores indicam que o poder dos testes de aderência do Qui-Quadrado e de Kolmogorov-Smirnov, de discriminar entre hipóteses falsas e verdadeiras, é bastante diminuído nas caudas inferior e superior, tanto em função do reduzido número de observações amostrais, quanto em decorrência dos maiores erros de estimação, nessas partições do espaço amostral. Alternativamente, o teste de aderência de Anderson-Darling é um teste não-paramétrico que procura ponderar mais fortemente as caudas das distribuições, nas quais, as maiores (ou as menores) observações da amostra podem alterar sobremaneira a qualidade do ajuste.

Wang (1998) produziu um teste de qualidade de ajuste da GEV com base nos dados de cada estação, seus passos são: estimar os parâmetros da GEV para $\eta = 0, 1, 2, 3$ e 4 ; calcular os taxas de momentos-LH; calcular o teste de Wang da distribuição GEV; e escolher o conjunto de parâmetros em função do menor valor do teste de Wang.

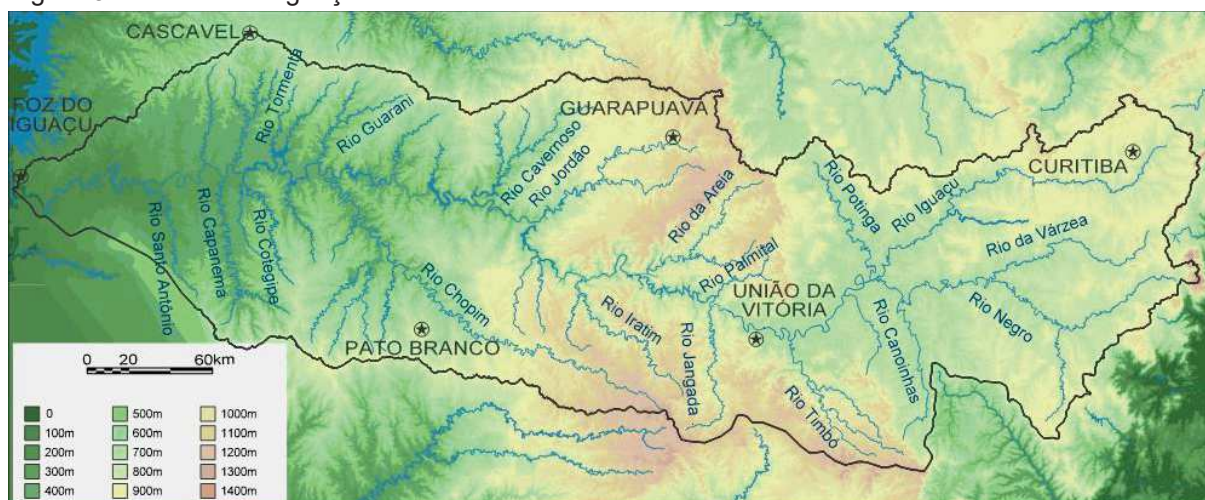
Neste trabalho optou-se por utilizar o teste de Wang, por resultar em ajustes de qualidade como demonstrado por Queiroz e Chaudhry (2006).

3 MATERIAIS E MÉTODOS

3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A bacia do Rio Iguaçu, apresentada na Figura 05, encontra-se tanto no território do Brasil quanto na Argentina cobrindo uma área de superfície aproximada de 70.800 km². O rio Iguaçu é considerado o maior rio totalmente paranaense formado por vários contribuintes, os principais são: Iraí, Atuba, Passaúna, Barigui, Verde, Passa Dois, da Várzea, Chopin, Palmital, Cavernoso, Adelaide, Gonçalves Dias, Castro Alves, Ampére e Silva Jardim.

Figura 5 - Bacia do rio Iguaçu.



Fonte- Adaptado COPEL (2018).

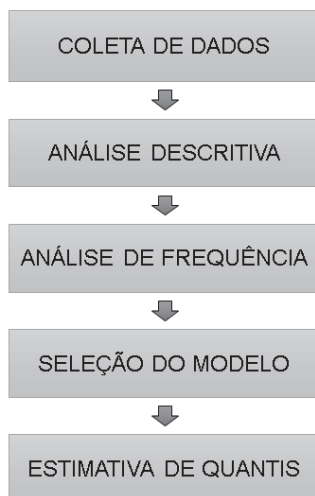
A região possui temperatura média anual inferior a 20° C e chuvas bem distribuídas ao ano, com regiões na costa oeste que podem chegar a 40° C.

As formações geológicas apresentam diversidade com predomínio de Serra Geral no terceiro planalto, Itararé no segundo e Sedimentos Recentes no primeiro.

3.2 MÉTODOS

Para o andamento deste estudo e finalmente a obtenção dos resultados e objetivos desejados, será seguido o roteiro na Figura 06. Este estudo será subdividido em cinco partes principais: coleta de dados, análise descritiva, análise de frequência, seleção do modelo e estimativa de quantis.

Figura 6 - Roteiro do estudo.



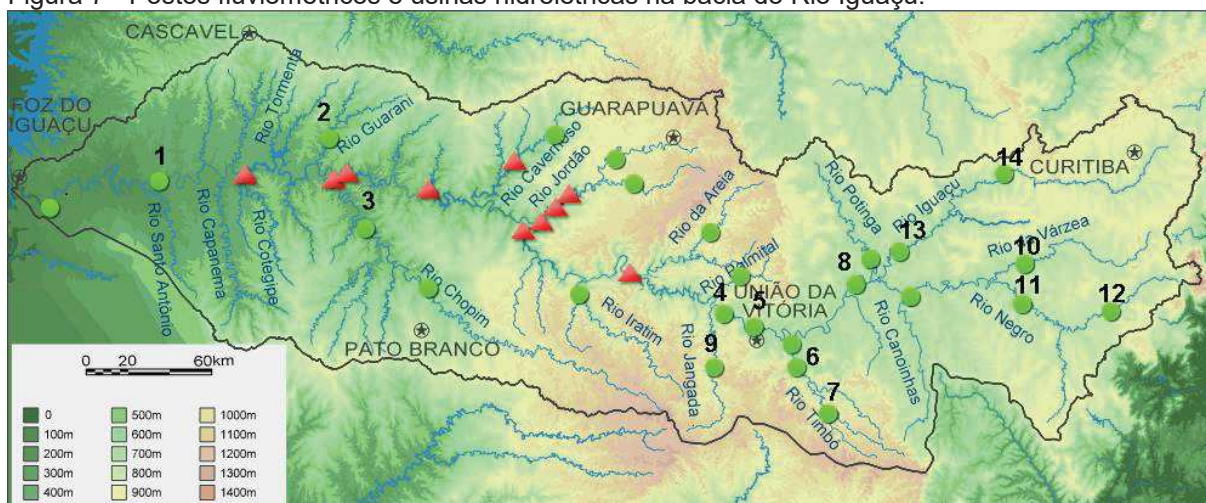
Fonte – Autor (2018).

3.2 POSTOS FLUVIOMÉTRICOS

Os dados foram coletados a partir do sistema do site Hidroweb, onde há a disposição de dados de vazões máximas, já coletados a partir de postos fluviométricos na região da bacia do Rio Iguaçu.

A Figura 07 apresenta a localização dos postos fluviométricos, que são representados com círculos verdes, e a localização das usinas hidrelétricas, que são representadas por triângulos vermelhos.

Figura 7 - Postos fluviométricos e usinas hidrelétricas na bacia do Rio Iguaçu.



Fonte- Adaptado COPEL (2019).

Destes postos fluviométricos foram obtidas informações relevantes para este trabalho, como: número da estação, código ANA, longitude, altitude, início da leitura de dados, final da leitura de dados e número de dados. Estas informações são demonstradas na Tabela 01, porém nota-se que algumas estações não constam nas tabelas. Isso se dá pelo fato de haver estações cujo número de dados amostrais se encontravam menor que 20, podendo comprometer na análise dos dados em casos como do momento-LH, onde são excluídos dados para se fazer a análise.

Tabela 1 – Informações dos postos fluviométricos.

ID	Cód.	Lat.	Long.	Alt.(m)	I	F	N
1	65987000	S 25° 34' 0.12"	W 53° 55' 59.16"	180.0	1978	2007	30
2	65970000	S 25° 23' 38.04"	W 53° 6' 14.04"	430.0	1978	2005	28
3	65960000	S 25° 46' 9.12"	W 52° 55' 36.12"	390.0	1957	2014	58
4	65365000	S 26° 10' 37.92"	W 51° 31' 10.92"	790.0	1946	2014	69
5	65310000	S 26° 13' 41.16"	W 51° 4' 49.08"	670.0	1931	2014	84
6	65295000	S 26° 23' 2.04"	W 50° 52' 41.88"	750.0	1975	2005	31
7	65260000	S 26° 35' 15.00"	W 50° 44' 42.00"	895.0	1986	2005	20
8	65220000	S 26° 1' 9.12"	W 50° 35' 33.00"	770.0	1964	2014	51
9	65370000	S 26° 23' 12.84"	W 51° 16' 18.84"	800.0	1946	2005	59
10	65155000	S 25° 56' 44.16"	W 49° 47' 29.04"	750.0	1930	2005	76
11	65100000	S 26° 6' 34.92"	W 49° 48' 3.96"	766.0	1950	2014	65
12	65090000	S 26° 9' 16.92"	W 49° 22' 50.16"	790.0	1968	2005	38
13	65060000	S 25° 52' 32.88"	W 50° 23' 22.92"	745.0	1931	2014	84
14	65035000	S 25° 32' 53.16"	W 49° 53' 18.96"	780.0	1936	2005	70

ID - Identificação da estação; Cód. – Código da Agência Nacional de Águas (ANA); Lat.- Latitude; Long. – Longitude; Alt. – Altitude; I – Ano inicial das medições; F – Ano final das medições; N – Número de amostras.

Fonte: Autor (2018).

3.3 ANÁLISE DE FREQUÊNCIA

A análise de frequência engloba toda a parte de tratamento de dados, nela foram realizados processos computacionais que identificaram resultados para este estudo. Por isso optou-se por utilizar duas distribuições para efeito de comparação.

3.3.1 Distribuição Generalizada de Valores Extremos com Momentos-LH

Será utilizada a distribuição GEV para o ajuste de dados, conforme o valor estimado do parâmetro de forma (k), que engloba as três formas assintóticas de distribuição de valores extremos, podendo ser três tipos: tipo I (VEI), valor extremo do tipo II (VEII) e valor extremo do tipo III (VEIII), onde para o caso do tipo I resulta na Gumbel.

Ao modelar a distribuição Generalizada de Valores Extremos aos dados amostrais, busca-se estimar os parâmetros de forma, posição e escala. Nesta pesquisa o processo de obtenção dos parâmetros da GEV será realizado através de uma rotina adaptada para utilização no Software R Core Team (2017). Essa rotina foi implementada computacionalmente, no software em Matlab, e apresentada por Queiroz e Chaudhry (2006), podendo ser utilizada para proceder ao ajuste de amostras de qualquer tamanho.

Então, através da rotina, alguns passos podem ser seguidos e aplicados para os dados de cada estação: estimar os parâmetros da GEV para $\eta = 0, 1, 2, 3$ e 4 ; calcular as taxas de momento-LH; calcular o teste de Wang (1998) da distribuição GEV; e escolher o conjunto de parâmetros em função do menor valor do teste Wang (1998).

Na sequência, com os parâmetros de escala, posição e forma numericamente determinados para cada uma das estações fluviométricas de estudo, substituem-se os mesmos nas Equações (3) ou (4), de acordo com o parâmetro k . Então, geram-se os quantis da distribuição GEV, considerando os tempos de retornos estipulados entre 100 a 10000 anos, para a melhor visualização da variação de erro causado por períodos de retornos altos.

3.3.2 Distribuição Gumbel pelo teorema de Ven Te Chow

Para a distribuição Gumbel necessita-se obter informações a respeito da série de dados amostrais de cada posto fluviométrico, como a média, desvio padrão e coeficiente de variação. Tais dados já serão obtidos a partir da análise descritiva feito um processo anterior.

Em seguida serão utilizadas as equações 6, 7 e 8 sugeridas pelo teorema de Ven Te Chow para distribuição Gumbel:

$$Y_t = -\ln \left[-\ln \left(1 - \frac{1}{T} \right) \right] \quad (6)$$

$$K_t = 0,7797 * Y_t - 0,45 \quad (7)$$

$$X_t = X_m + K_t * S \quad (8)$$

Onde:

Y_t = Variável reduzida de Gumbel;

K_t = Fator de frequência;

X_t = Quantil;

T = Período de retorno (anos);

X_m = Média das amostras;

S = Desvio padrão das amostras.

Para o período de retorno aplicado na equação 6, será determinado entre 100 a 10000 anos para que seja parâmetro de comparação entre as distribuições.

3.4 CAPACIDADE PREDITIVA

Para avaliar a capacidade preditiva dos modelos estocásticos, será utilizado o método bootstrapping que produz múltiplas amostras sintéticas contendo o mesmo tamanho amostral da série estudada (Efron e Tibshirani, 1994). A característica mais atrativa dessa metodologia é sua capacidade de descrever o comportamento da distribuição apenas com os valores amostrais, mesmo quando a distribuição verdadeira é desconhecida. Além disso, pode ser utilizada para avaliar a incerteza amostral (Vogel, 1995).

O bootstrapping será realizado com 1000 amostras com a metade do tamanho amostral original. Na sequência, ajustam-se os modelos estocásticos para as amostras de bootstrap. Por fim, a extrapolação é realizada no caudal direito da distribuição, correspondendo aos quatro maiores valores da série original, ou seja, dos valores efetivamente observados.

Nguyen et al. (2017) descreve que para visualização dos resultados será utilizado uma forma modificada do boxplot (Helsel e Hirsch, 2002). A linha do meio é a média amostral, o tamanho do boxplot é de duas vezes o desvio padrão, o limite superior é representado pelo maior valor e o limite inferior pelo menor valor. O tamanho do boxplot, bem como a extensão dos limites, indicam grande incerteza na estimativa dos valores extremos. Caso sejam observados valores fora do boxplot, então a distribuição ajustada por bootstrap superestima ou subestima os quantis e, portanto, não é recomendável.

4. RESULTADOS E DISCUSSÕES

Para este trabalho foram coletados os dados das vazões máximas anuais dos 14 postos fluviométricos da bacia do rio Iguaçu, tais dados apresentaram grandes variações de amostras de acordo com o posto analisado, variando de 20 dados amostrais (estação 7) até 84 dados amostrais (estações 5 e 13), como apresentado na Tabela 1.

4.1 ESTATÍSTICA DESCRITIVA

A Tabela 2 apresenta os resultados obtidos a partir da análise descritiva. A maior e menor média é dada pelas estações 1 e 4, respectivamente. Já o desvio padrão, a estação 1 com 8195,5 m³/s e a estação 12 com 27,8 m³/s; para o coeficiente de variação as estações 13 e 6 apresentam o maior e o menor valor, respectivamente, tendo uma diferença de 56,6% entre eles; a grande maioria das estações mostraram assimetria positiva, porém apenas a estação 2 apresentou ser negativo, sendo um valor de -0,7; assim como a assimetria, a curtose apresentou vários dados positivos, apenas as estações 1 e 2 apresentaram valores negativos de curtose.

Além disso, a Tabela 2 também mostrou o número de outliers por estação que foram identificados a partir da construção de boxplots. Notou-se que uma grande parcela das estações obteve outliers, porém a estação 13 se destaca por ter 9 outliers, quase 10% dos seus dados amostrais, isso se dá pelo fato do coeficiente de variação da estação 13 ser o maior entre as estações (86,1%).

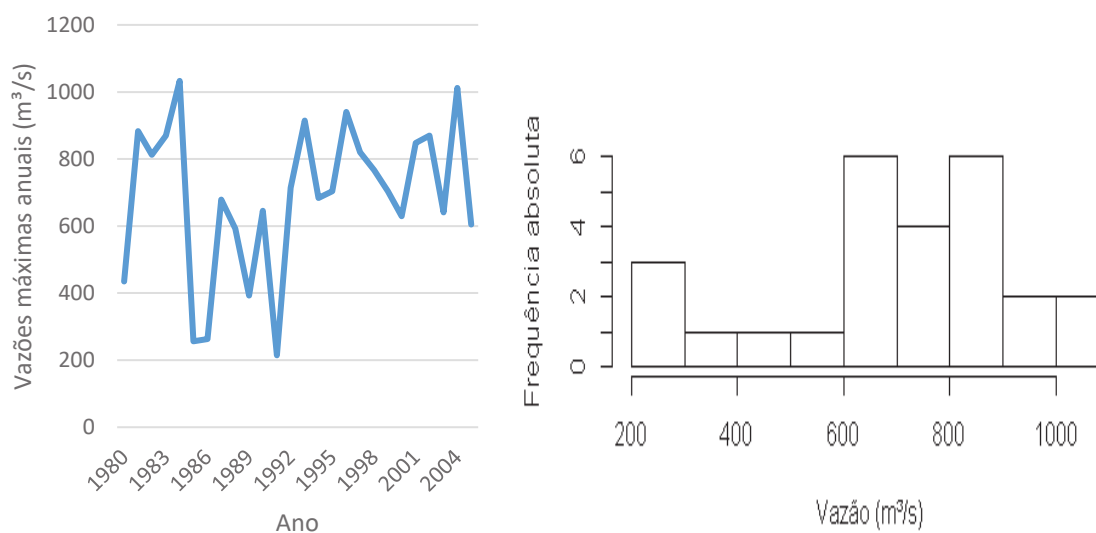
Tabela 2 – Resultados da análise estatística.

ID	N	M (m ³ /s)	DP (m ³ /s)	CV (%)	AS	Curt	Mín (m ³ /s)	Q1 (m ³ /s)	Q2 (m ³ /s)	Q3 (m ³ /s)	Máx (m ³ /s)	NO
1	30	13845,6	8195,5	59,2	0,3	-0,8	1944,0	7007,0	14067,0	19419,0	31660,0	0
2	26	689,5	225,9	32,8	-0,7	-0,1	214,4	610,4	704,0	864,4	1033,0	0
3	58	1601,8	811,7	50,7	1,8	5,8	402,2	1051,4	1428,2	1954,8	5220,2	1
4	69	54,2	28,5	52,6	1,8	5,5	8,9	34,9	49,5	67,7	168,7	2
5	84	1710,9	767,7	44,9	1,5	3,8	568,7	1174,6	1594,6	2088,5	5072,6	2
6	31	380,1	112,1	29,5	0,4	1,8	147,2	304,5	388,2	436,6	710,2	1
7	20	219,3	89,3	40,7	0,7	-0,1	89,6	158,4	196,6	263,8	412,6	0
8	51	1425,2	703,0	49,3	1,6	3,3	448,6	938,5	1331,9	1742,8	3819,4	3
9	59	248,3	124,3	50,1	3,1	15,2	101,9	165,0	230,6	281,9	935,3	3
10	76	135,2	78,8	58,3	2,8	12,2	45,5	85,1	117,3	163,7	766,2	2
11	65	332,7	151,6	45,6	1,6	4,0	87,1	238,8	319,5	387,1	564,0	3
12	39	84,8	27,8	32,8	0,2	0,3	26,8	69,9	86,5	96,4	151,8	4
13	84	461,5	397,5	86,1	2,6	8,1	120,2	250,0	314,6	514,2	2403,7	9
14	70	326,0	188,6	57,8	2,1	6,2	105,8	211,4	283,4	396,6	1167,0	3

ID – Identificação da estação; N – Tamanho da série amostral; M – Média; DP – Desvio padrão; CV – Coeficiente de variação; AS – Assimetria; Curt – Curtose; Mín – Mínimo; Q1 – Primeiro quadrante; Q2 – Segundo quadrante; Q3 – Terceiro quadrante; Máx – Máximo; NO – Número de Outliers.

Fonte: Autor (2019).

Será destacada a estação 2 (tabela 2), por apresentar valores negativos para assimetria e curtose. Portanto, foi construído o histograma e plotado o gráfico da série, conforme a Figura 8, para a visualização das suas características.

Figura 8 – Análise gráfica da estação 2.

Fonte: Autor (2019).

4.2 ESTIMATIVA DOS PARÂMETROS DA DISTRIBUIÇÃO GEV

Neste trabalho foi utilizada a rotina sugerida por Queiroz e Chaudhry (2006), pois ela estima os parâmetros da distribuição GEV por meio do método de momentos-LH, indicando quais dos níveis (L0, L1, L2, L3 e L4) foi selecionado e determina o valor do teste de qualidade de ajuste de Wang (1998). O resultado dessa rotina apresenta o teste de Wang para cada um dos níveis de momentos-LH para cada estação, porém, para a análise desses dados foi escolhido o menor valor de teste de Wang, sendo assim, obtido o melhor momento-LH para esse ajuste. O resultado desse processo resultou na Tabela 3. Para efeito de comparação foi incluído o Apêndice A com todos os 5 níveis de momentos-LH com seus respectivos ajustes de teste de Wang.

Tabela 3 - Ajustes da distribuição GEV às séries de vazão máxima anual.

Estação	Zw	LH	Parâmetros da distribuição GEV			Taxas de momentos-LH		
			k	α	β	cv	cc	ca
1	0,217	2	0,337	9806,947	10028,080	0,156	0,087	0,112
2	0,012	3	0,430	214,074	648,450	0,061	0,061	0,096
3	0,297	1	-0,112	535,157	1231,119	0,192	0,194	0,306
4	1,396	0	-0,029	20,515	41,741	0,270	0,210	0,187
5	0,189	0	-0,037	568,947	1361,086	0,239	0,154	0,192
6	1,340	4	-0,296	36,406	377,037	0,083	0,322	0,431
7	0,186	0	0,016	74,943	177,220	0,233	0,131	0,158
8	0,278	0	-0,098	481,288	1095,875	0,258	0,191	0,233
9	1,241	0	-0,134	72,879	195,245	0,234	0,242	0,257
10	0,411	2	-0,256	38,541	101,174	0,191	0,253	0,403
11	0,294	4	-0,246	72,602	283,654	0,140	0,245	0,408
12	0,350	2	-0,013	18,583	76,682	0,101	0,168	0,279
13	1,562	0	-0,450	134,381	277,324	0,383	0,310	0,494
14	0,329	0	-0,184	110,993	237,525	0,289	0,223	0,291

Zw (0,05) = 1,96 – Teste de Wang; LH – Níveis dos momentos-LH; k – Parâmetro de forma; α – Parâmetro de escala; β – Parâmetro de posição; cv – Coeficiente de variação; cc – Coeficiente de curtose; ca – coeficiente de assimetria.

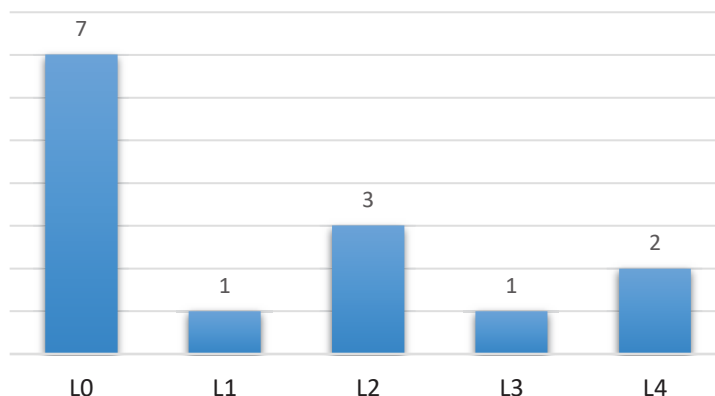
Fonte: Autor (2019).

Analisando os valores de Zw obtidos pelo teste de Wang, observou-se que todos os dados tiveram seus valores abaixo de 1,96, isso significa que todas as estações podem ser modeladas pela GEV, segundo teste de Wang a 5% de significância.

A Figura 9 apresentou a quantidade de estações por nível de momentos-LH sugerido por Wang (1997). Observa-se que das 14 séries estudadas, 7 delas são do

nível L0, correspondendo a metade das séries, sendo assim, esses momentos-LH se igualam aos momentos-L. Enquanto que apenas uma série (7,14%) se enquadra no nível L1, 3 séries (21,43%) se enquadra no nível L2, uma serie (7,14%) se enquadra no nível L3 e 2 series (14,28%) se enquadram no nível L4.

Figura 9 – Número de séries para cada nível dos momentos-LH.



Fonte: Autor (2019).

Em seguida, analisou-se os parâmetros de forma (k) de cada uma das estações. Observou-se que o menor e o maior valor para o parâmetro de forma é de $-0,450$ (estação 17) e de $0,430$ (estação 3), respectivamente. De acordo com Hosking, Wallis e Wood (1985), normalmente o parâmetro permanece entre os intervalos de $-0,5 < k < 0,5$, sendo assim, todas estações estão dentro do intervalo sugerido.

Complementando, Stedinger et al. (1993) aponta que entre os intervalos $-0,03 < k < 0,03$, a forma geral da distribuição GEV se assemelha a distribuição Gumbel, sendo do tipo VEI. Desse modo, foram classificadas as estações conforme o tipo de forma, podendo ser do tipo VEI, VEII e VEIII. A Tabela 4 mostra o número de estações para cada tipo. Sendo assim, 21,43% é representado pelo tipo VEI, 64,28% é representado pelo tipo VEII e 14,28% é representado pelo tipo VEIII.

Tabela 4 - Tipo de distribuição de valores extremos.

Tipo	Parâmetro k	Número de séries	Representatividade (%)
VEI	$-0,03 < k < 0,03$	3	21,44
VEII	$k < -0,03$	9	64,28
VEIII	$k > 0,03$	2	14,28

Fonte: Autor (2019).

4.3 DETERMINAÇÃO DOS QUANTIS

Algumas das estações foram escolhidas para as etapas seguintes em função dos níveis de momentos-LH e pelo melhor ajuste do teste de Wang. As estações são: estação 2, por demonstrar o melhor ajuste pelo teste de Wang e nível L3 de momentos-LH; estação 3, por ter nível L1 de momentos-LH; estação 5 por ter um bom ajuste, boa quantidade de dados amostrais e nível L0; estação 11, por ser o melhor ajuste entre os de nível L4; e estação 12, por destacar o parâmetro de forma ($k=-0,013$) mais próximo de zero entre todos e nível L2.

Dando sequência, realizou-se os estudos do cálculo dos quantis da distribuição de valores extremos (GEV) para os períodos de retorno de 2, 10, 100, 1000 e 10000 anos, em sequência, calculou-se os quantis da distribuição Gumbel para os mesmos períodos de retorno, e então compará-los. Para isso utilizam-se as Equações (3) e (4) para a distribuição GEV junto com os parâmetros da GEV da tabela 4, e as Equações (6), (7) e (8) para a distribuição Gumbel.

Obtidos os resultados dos quantis para cada distribuição, organizaram-se os dados na Tabela 5 para a melhor comparação. Comparando os resultados das distribuições, notou-se que as estações 5 e 12 não apresentaram uma grande diferença, assim como na estação 11, que no período de retorno de 10000 anos, apresentou quantil pela GEV duas vezes maior que por Gumbel.

Tabela 5 - Resultados da determinação dos quantis da distribuição GEV e Gumbel.

ID	Período de Retorno (Anos)									
	2		10		100		1000		10000	
	GEV	Gum.	GEV	Gum.	GEV	Gum.	GEV	Gum.	GEV	Gum.
2	721,0	652,4	957,1	984,2	1077,3	1398,1	1120,5	1804,5	1136,6	2210,1
3	1431,4	1468,5	2602,0	2660,8	4457,6	4148,0	6826,2	5608,2	9892,4	7065,8
5	1571,0	1584,9	2696,0	2712,4	4213,1	4118,9	5836,0	5499,7	7599,6	6878,2
11	311,5	307,8	501,8	530,5	903,2	808,3	1601,2	1081,1	2828,9	1353,4
12	83,5	80,2	119,1	121,1	164,9	172,1	211,2	222,2	259,0	272,1

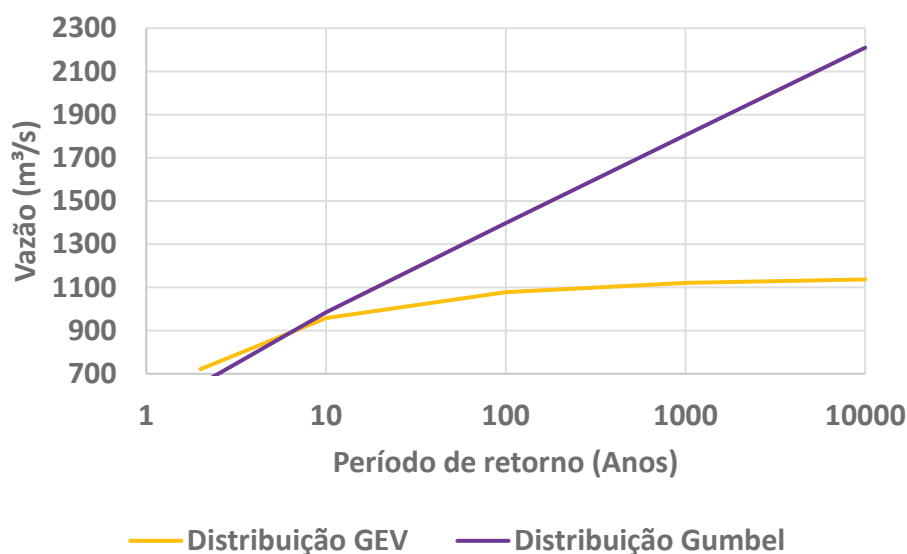
Fonte: Autor (2019).

Teoricamente a distribuição generalizada de valores extremos (GEV) deve apresentar as estimativas de valores corretos para altos períodos de retorno, sendo assim, espera-se que os resultados obtidos tenham uma confiabilidade maior conforme o teste de Wang apresente valores menores.

Portanto, criaram-se gráficos que ajudaram a comparar os dados das distribuições. Para melhorar a análise dos gráficos, optou-se por separar os gráficos que demonstraram comportamentos parecidos.

Na Figura 10 percebe-se que a estação 3 demonstra uma grande diferença entre as duas distribuições, sendo que, também é importante ressaltar que seu valor do teste de Wang foi de 0,012, o menor valor obtido entre as estações, indicando que a confiabilidade do ajuste pode ser a melhor neste trabalho. Também pode-se notar que em caso de dimensionamento de uma barragem, utilizando-se um período de retorno de 10000 anos, se utilizar a distribuição Gumbel há a possibilidade de estar superdimensionando a obra hidráulica.

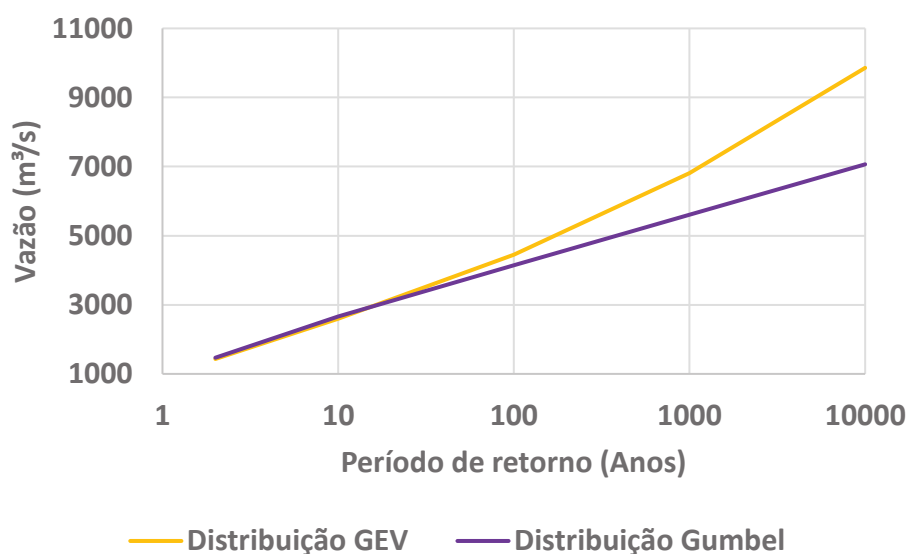
Figura 10 - Comparação entre as distribuições GEV e Gumbel na estação 2.



Fonte: Autor (2019).

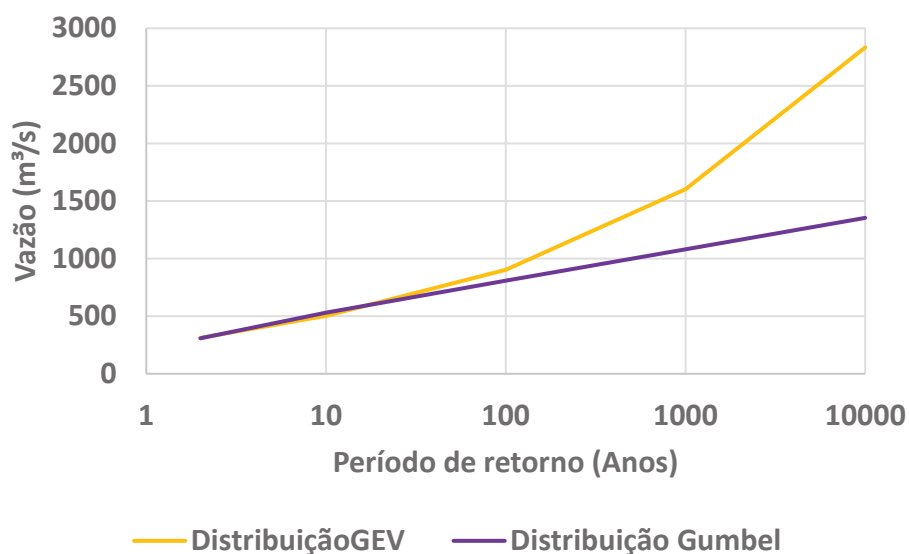
Nas Figuras 11 e 12 percebe-se que as estações 4 e 11, respectivamente, demonstram um comportamento parecido. Nota-se que, aproximadamente até ao período de retorno de 100 anos, as duas distribuições apresentam um comportamento similar, porém, conforme o período de retorno aumenta a distribuição GEV ascende, enquanto a Gumbel continua com a sua linearidade. Podendo assim, inferir que caso haja um dimensionamento de uma obra hidráulica, e for utilizado a distribuição Gumbel, há a probabilidade de subdimensionamento.

Figura 11 - Comparação entre as distribuições GEV e Gumbel na estação 3.



Fonte: Autor (2019).

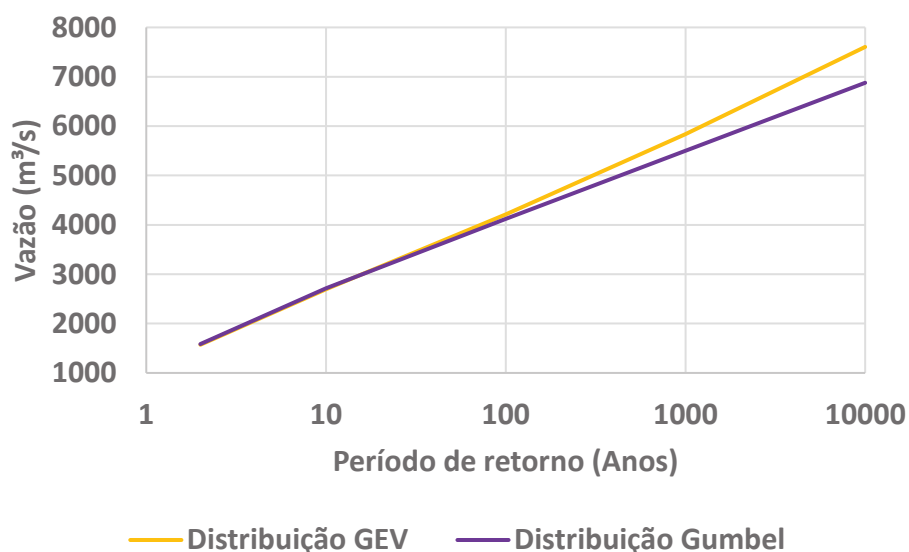
Figura 12 - Comparação entre as distribuições GEV e Gumbel na estação 11.



Fonte: Autor (2019).

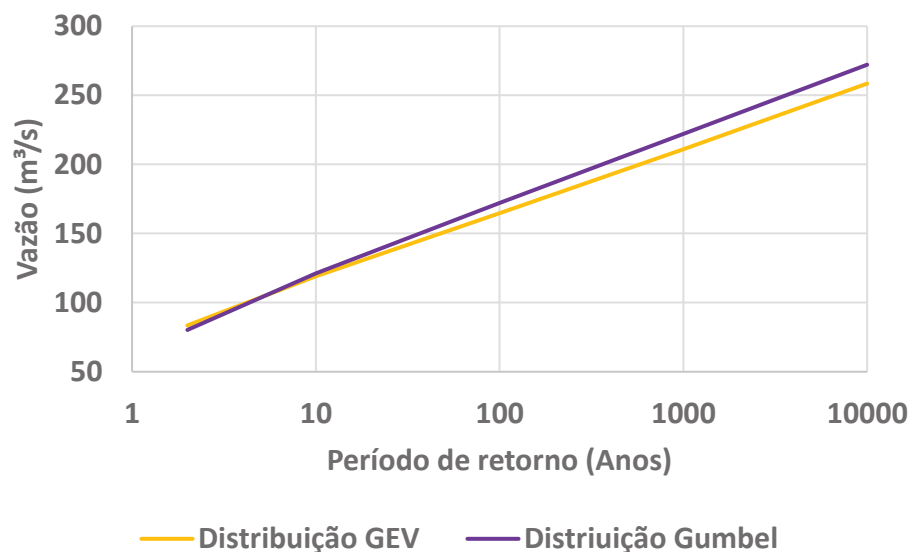
Nas Figuras 13 e 14 percebe-se que as estações 5 e 12, respectivamente, demonstram um comportamento semelhante entre eles. Tal comportamento é resultado de um parâmetro de forma próximo de zero. Sendo assim, é possível notar que conforme o período de retorno aumenta, as duas distribuições se distanciam bem pouco, resultado de o parâmetro de forma ter valores próximos ou entre o intervalo de $-0,03 < k < 0,03$ demonstrado por Queiroz e Chaudhry (2006). Portanto, para o dimensionamento de uma obra hidráulica, a diferença entre as distribuições não interfere consideravelmente, podendo escolher tanto a GEV, quanto a Gumbel.

Figura 13 - Comparação entre as distribuições GEV e Gumbel na estação 5.



Fonte: Autor (2019).

Figura 14 - Comparação entre as distribuições GEV e Gumbel na estação 12.



Fonte: Autor (2019).

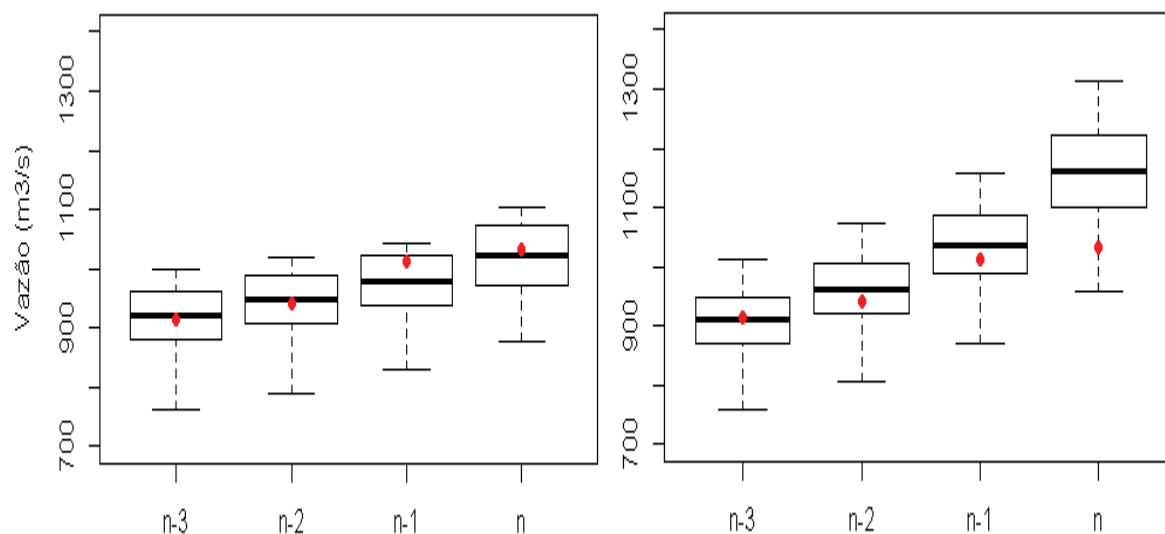
As estimativas de vazões de projeto são valores presumidos a partir de teorias probabilísticas, ou seja, os valores de vazões que vão realmente acontecer podem não ser os valores calculados a partir dessas teorias. Portanto, não se pode afirmar qual distribuição melhor representa a realidade, mas teoricamente a distribuição GEV apresenta as estimativas mais próximas dos valores reais.

4.4 CAPACIDADE PREDITIVA DAS DISTRIBUIÇÕES GEV E GUMBEL

Assim como descrito na metodologia, a capacidade preditiva foi realizada utilizando o bootstrapping, realizando 1000 amostras com metade da série amostral, para cada uma das estações. Para a visualização dos resultados, utilizou-se o método do boxplot modificado, sugerido por Nguyen et al. (2017), onde valores fora do boxplot indicam que a distribuição ajustada pelo bootstrap superestima ou subestima os quantis, não sendo recomendável caso aconteça.

Para a estação 2, a Figura 15 demonstrou uma boa capacidade preditiva para a distribuição GEV, tendo seus valores dentro do boxplot e perto da média amostral. Já a distribuição Gumbel pelo teorema de Ven Te Chow, apresentou baixa capacidade preditiva, pois seus dados se distanciaram da média amostral e saíram da caixa do boxplot.

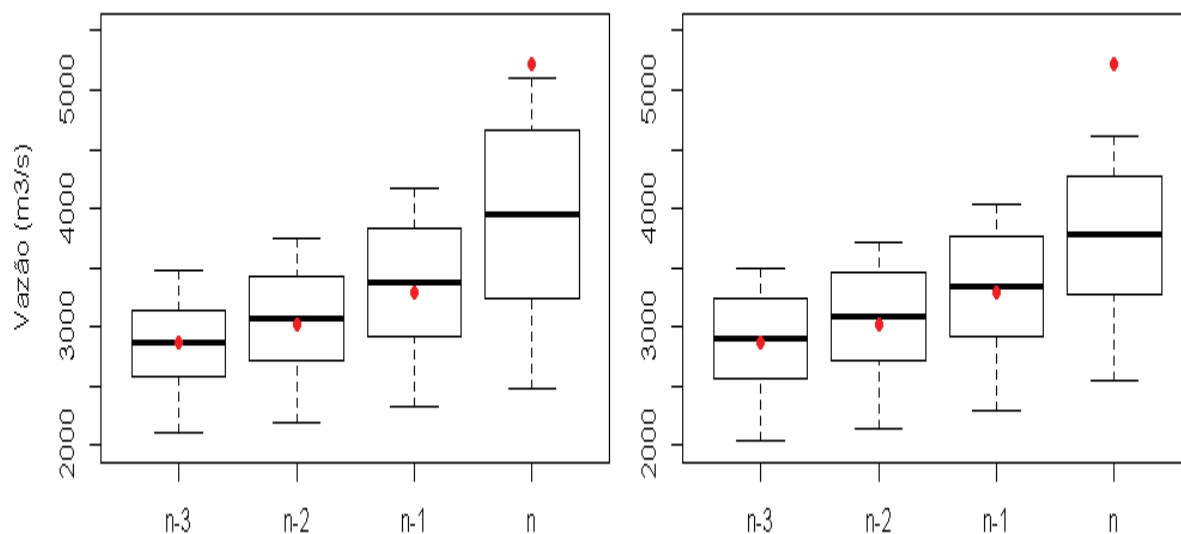
Figura 15 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 2.



Fonte: Autor (2019).

Para a estação 3, a Figura 16 mostrou que ambas distribuições tiveram baixa capacidade preditiva em seus bootstrap de série amostral total ($n= 58$), pois demonstraram valores fora dos limites do boxplot.

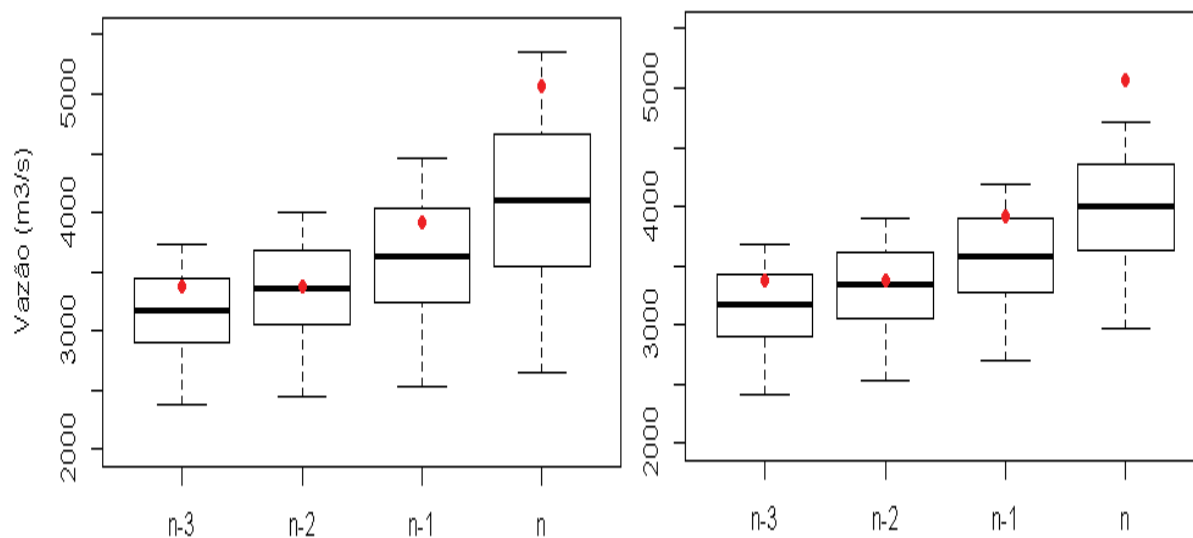
Figura 16 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 3.



Fonte: Autor (2019).

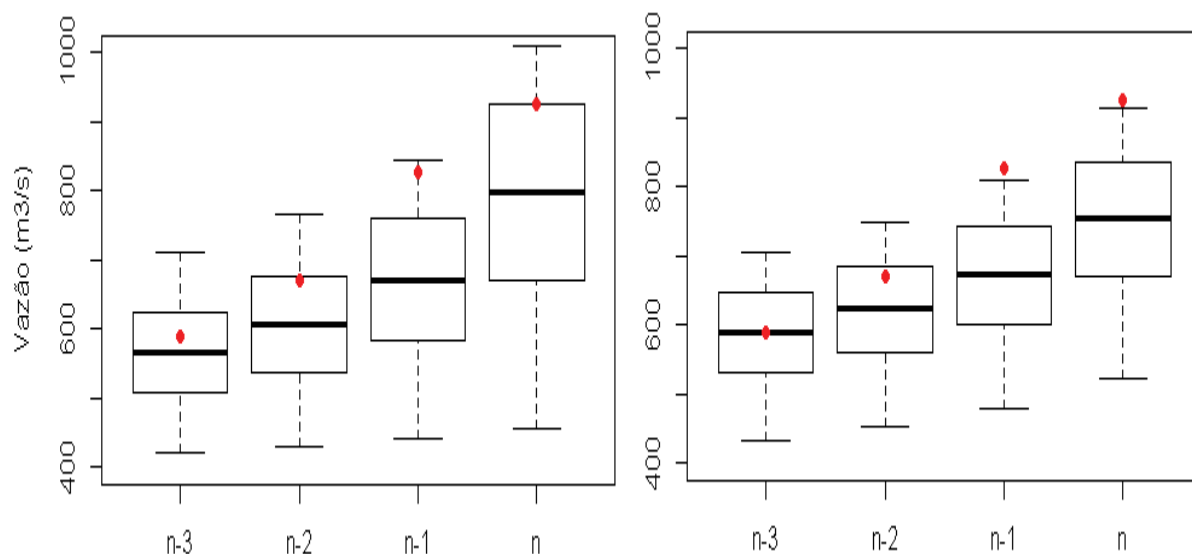
Para as estações 5 e 11, as Figuras 17 e 18 apresentaram uma baixa capacidade preditiva para a distribuição GEV, por ter valores fora da caixa do boxplot. Porém, a capacidade preditiva para a distribuição Gumbel foi pior, pois seu ponto de referência ficou fora dos limites, tendo uma distância maior da média amostral comparando com a distribuição GEV.

Figura 17 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 5.



Fonte: Autor (2019).

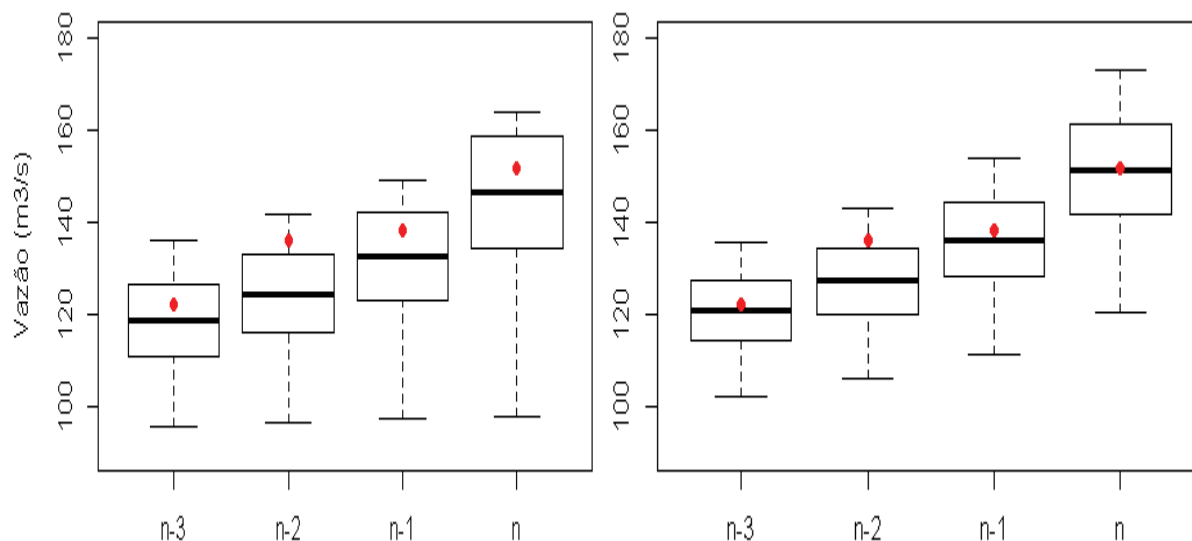
Figura 18 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 11.



Fonte: Autor (2019).

Para a estação 12, a Figura 19 apresentou uma boa capacidade preditiva para ambas distribuições, podendo perceber que quase todos os dados ficaram dentro dos boxplots.

Figura 19 - Bootstrapping da GEV e da Gumbel, respectivamente, para a estação 12.



Fonte: Autor (2019).

Devido aos resultados da capacidade preditiva, pode-se dizer que, para as séries estudadas na bacia do rio Iguaçu, a distribuição GEV apresentou de forma geral uma melhor capacidade preditiva em relação a distribuição Gumbel.

CONCLUSÃO

Neste trabalho foi utilizada a distribuição Generalizada de Valores Extremos (GEV), com seus parâmetros estimados por momentos-LH, para a análise das vazões máximas de séries históricas da bacia do rio Iguaçu. Dentre as 14 séries analisadas, 4 pertenceram a distribuição do tipo VEI; 8 ao tipo VEII; e, 2 ao tipo VEIII. Sendo que, pelo teste de Wang (1998), todas as 14 séries tiveram seus ajustes aceitos com nível de significância de 5%. Indicando que, foi possível modelar as vazões máximas na bacia hidrográfica do rio Iguaçu utilizando a distribuição GEV com momentos-LH.

Neste trabalho, também foram estimadas as vazões de projeto utilizando a distribuição Gumbel. Podendo assim, serem comparados as duas distribuições, que demonstrou pouca diferença para pequenos períodos de retorno (até aproximadamente 100 anos), e, conforme o período de retorno foi aumentando as distribuições do tipo VEII foram se distanciando tendo seu caudal cada vez mais pesado, para o tipo VEIII, seu caudal ficou cada vez mais leve e para o tipo VEI, seus dados não diferenciaram muito com a Gumbel.

Também foram analisadas as capacidades preditivas das duas distribuições utilizando o método de bootstrapping, sendo possível notar que, a distribuição GEV possuiu uma capacidade preditiva melhor que a Gumbel, pois seus dados mantiveram-se mais próximos da média amostral do que na Gumbel. Porém, apesar de apresentar dados melhores que a Gumbel, a distribuição GEV também demonstrou valores fora da caixa do boxplot.

Sendo assim, concluiu-se que, para este trabalho, a distribuição Generalizada de Valores Extremos (GEV) com momentos-LH conseguiu modelar os dados das séries amostrais na bacia do rio Iguaçu. Além disso, apresentou uma pequena vantagem em relação a distribuição Gumbel pelo teorema de Ven Te Chow, vantagem essa sustentada pela sua capacidade preditiva.

REFERÊNCIAS

AGUASPARANA – **INSTITUTO DAS ÁGUAS DO PARANÁ**. Disponível em: <<http://www.aguasparana.pr.gov.br/arquivos/File/BACIAS/iguacu.pdf>>. Acesso em: 27 set. 2018.

COLLISCHONN, W.; DORNELLES, F. **Hidrologia para engenharia e ciências ambientais**. 2. ed. Revisada e ampliada. Porto Alegre: Associação Brasileira de Recursos Hídricos (ABRH), 2015.

COPEL – **COMPANHIA PARANAENSE DE ELETRICIDADE**. Disponível em: <<https://www.copel.com/mhbweb/paginas/bacia-iguacu.jsf>>. Acesso em: 28 de set. 2018.

DOUGLAS, E. M.; VOGEL, R. M. Probabilistic behavior of floods of record in the United States. **Journal of Hydrologic Engineering**, Reston, v.11, n.5, 2006.

EFRON, B.; TIBSHIRANI, R. J. **An introduction to the bootstrap Chapman & Hall/ CRC Monographs on Statistics & Applied Probability**. CRC Press, New York, USA, 1994.

EM-DAT – **EMERGENCY DISASTERS DATA BASE**. Disponível em: <<https://www.emdat.be/sites/default/files/Maps/Disaster/map5.jpg>>. Acesso em: 28 set. 2018.

HAAN, C.T. **Statistical methods in hydrology**. 2.ed. Ames: The Iowa State University, 2002.

HELSEL, D. R., HIRSCH, R. M. **Statistical Methods in Water Resources Techniques of Water Resources Investigations**, Book 4, chapter A3. U.S Geological Survey, US, 2002.

HOSKING, J. R.; WALLIS, J.r.; WOOD, E.f.. Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. **Technometrics**, Alexandria, v. 27, n. 3, p.251-261, ago. 1985.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) of meteorological elements, **Quarterly Journal of the Royal Meteorological Society**, London, v.81, 1955.

KATZ, R.W.; PARLANGE, M.B.; NAVEAU, P. **Statistics of extremes in hydrology**. Advances in Water Resources, Amsterdam, v.25, 2002.

LIU, J-J; WU, Y-C; CHENG, K-S. Establishing acceptance regions for L-moments based goodness-of-fit tests b stochastic simulation. **Journal of Hydrology**, v.355, n.1, 2008.

MARTINS, C. A. SILVA.; BORGES, T. S.; SILVEIRA. F. C; CECÍLIO, R. A. Caracterização da precipitação da bacia hidrográfica do rio Benevente. In: XII Encontro Latino Americano de Iniciação Científica e VIII Encontro Latino Americano de Pós-Graduação – Universidade do Vale do Paraíba. 2011. **Anais...** Paraíba: 2011.

MARTINS, E. S.; STEDINGER, J. R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. **Water Resources**, Research, Washington, v.36, n.3, 2000.

NAGHETTINI, M; PINTO, É. J. A. **Hidrologia Estatística**. Belo Horizonte: CPRM, 2007.

NGUYEN, T. H.; OUTAYEK, S. E.; LIM, S. H.; NGUYEN, V. T. V. A systematic approach to selecting the best probability models for annual maximum rainfalls – A case study using data in Ontario (Canada). **Journal of Hydrology**, v. 553, 2017.

PANSERA, W. A. **Distribuição generalizada de chuvas máximas no estado do Paraná**. UNIOESTE, 2013.

QUEIROZ, M. M. F. de. **Análise de cheias anuais segundo distribuição generalizada**: EESC/USP, 2002.

QUEIROZ, M. M. F. de; CHAUDHRY, F.H. Análise de eventos hidrológicos extremos usando a distribuição GEV e momentos LH. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.10, n.2, 2006.

STEDINGER, J. R; Vogel, R. M.; FOUDOULA-GEORGIOUS, E. (1993). Frequency analysis of extreme events. In: **Handbook of Hydrology**, McGraw-Hill, New York.

VOGEL, R. M. Recent advances and themes in hydrology. **Rev. Geophys.** 33 (S2), 1995.

WANG, Q. J. LH moments for statistical analysis of extreme events, **Water Resources Research**, Washington, v.33, n.12, 1997.

WANG, Q. J. Approximate goodness-of-fit test of fitted generalized extreme value distribution using LH moments, **Water Resources Research**, Washington, v.34, n.12, 1998.

WATANABE, F. M. **Análise do método de Gumbel para o cálculo de vazões de dimensionamento de vertedouros**. USP, 2013.

APÊNDICE

Apêndice A – Tabela de parâmetros e teste de Wang da GEV.

									(continua)
ID	Zw	LH	K	A	U	CV	CC	CA	
1	0,217	2	0,337	9806,947	10028,080	0,156	0,087	0,112	
	0,470	4	0,288	8986,789	10480,530	0,113	0,112	0,172	
	0,487	3	0,321	9552,565	10147,370	0,129	0,105	0,143	
	0,511	1	0,294	9243,586	10206,990	0,208	0,055	0,090	
	1,656	0	0,131	7643,374	10322,210	0,344	0,022	0,088	
2	0,012	3	0,430	214,074	648,450	0,061	0,061	0,096	
	0,055	2	0,435	215,692	647,765	0,073	0,065	0,068	
	0,257	4	0,429	213,711	648,640	0,053	0,079	0,114	
	0,441	0	0,584	251,036	643,030	0,185	0,138	-0,153	
	0,757	1	0,517	238,284	641,866	0,099	0,114	-0,016	
3	0,297	1	-0,113	535,157	1231,119	0,192	0,194	0,306	
	0,362	0	-0,094	552,569	1226,290	0,263	0,193	0,230	
	0,474	2	-0,124	522,628	1237,850	0,167	0,205	0,335	
	0,622	3	-0,140	503,947	1251,621	0,155	0,216	0,353	
	0,731	4	-0,158	481,076	1272,407	0,147	0,227	0,368	
4	1,396	0	-0,029	20,515	41,741	0,270	0,210	0,187	
	1,540	4	-0,276	12,012	46,904	0,149	0,288	0,422	
	1,820	3	-0,232	13,582	45,293	0,154	0,280	0,398	
	2,072	2	-0,174	15,706	43,604	0,164	0,266	0,360	
	2,120	1	-0,099	18,352	42,194	0,189	0,245	0,298	
5	0,189	0	-0,037	568,947	1361,086	0,239	0,154	0,192	
	0,764	1	-0,027	576,387	1360,738	0,172	0,176	0,258	
	1,237	2	-0,053	549,223	1373,803	0,147	0,197	0,299	
	1,375	4	-0,119	471,133	1434,307	0,128	0,222	0,350	
	1,382	3	-0,086	509,895	1400,838	0,135	0,212	0,328	
6	1,340	4	-0,296	36,406	377,037	0,083	0,322	0,431	
	1,921	0	0,288	108,608	342,201	0,161	0,220	-0,002	
	1,992	3	-0,205	46,822	366,407	0,084	0,318	0,385	
	2,446	1	0,091	85,804	344,060	0,104	0,254	0,194	
	2,466	2	-0,075	63,512	354,058	0,089	0,296	0,310	
7	0,186	0	0,016	74,943	177,220	0,233	0,131	0,158	
	0,472	1	0,038	77,251	176,973	0,164	0,102	0,222	
	0,650	4	0,200	105,094	160,919	0,111	0,047	0,210	
	0,696	2	0,082	83,248	174,462	0,138	0,075	0,232	
	0,701	3	0,141	93,209	167,769	0,123	0,059	0,223	
8	0,278	0	-0,098	481,288	1095,875	0,258	0,192	0,233	
	0,584	4	-0,208	375,296	1168,956	0,147	0,242	0,391	
	0,787	1	-0,113	468,671	1099,562	0,190	0,213	0,306	
	0,861	3	-0,181	403,078	1142,389	0,154	0,241	0,373	
	0,970	2	-0,146	437,553	1116,144	0,166	0,232	0,346	
9	1,241	0	-0,134	72,879	195,245	0,234	0,242	0,257	
	1,892	4	-0,410	36,551	221,973	0,154	0,357	0,485	
	2,310	3	-0,353	43,740	213,292	0,155	0,348	0,457	
	2,373	1	-0,195	65,243	197,263	0,179	0,295	0,353	
	2,581	2	-0,279	53,684	204,233	0,161	0,329	0,416	
10	0,411	2	-0,256	38,541	101,174	0,191	0,253	0,403	
	0,452	3	-0,267	37,512	102,054	0,181	0,258	0,415	
	0,517	1	-0,241	39,961	100,291	0,212	0,251	0,380	
	0,573	0	-0,218	41,815	99,675	0,274	0,246	0,316	
	0,639	4	-0,277	36,466	103,152	0,174	0,265	0,422	

ID – Identificação da estação fluviométrica; **Zw** – Teste de Wang; **K** – Parâmetro de forma; **A** – Parâmetro de escala; **B** – Parâmetro de posição; **CV** – Coeficiente de variação; **CC** – Coeficiente de curtose; **CA** – coeficiente de assimetria.

									(conclusão)
ID	Zw	LH	K	A	U	CV	CC	CA	
11	0,294	4	-0,246	72,602	283,654	0,140	0,245	0,408	
	0,843	3	-0,224	76,998	279,245	0,145	0,252	0,394	
	1,192	0	-0,061	107,325	263,842	0,237	0,214	0,208	
	1,430	2	-0,182	85,417	272,544	0,154	0,253	0,365	
	1,698	1	-0,121	97,201	266,134	0,174	0,241	0,310	
12	0,350	2	-0,013	18,583	76,682	0,101	0,168	0,279	
	0,591	3	-0,030	17,945	77,100	0,094	0,136	0,302	
	1,036	4	-0,005	19,028	76,245	0,089	0,110	0,299	
	1,693	1	0,085	22,118	75,143	0,119	0,202	0,197	
	1,912	0	0,251	27,086	74,690	0,184	0,207	0,019	
13	1,562	0	-0,450	134,381	277,324	0,383	0,310	0,494	
	2,125	4	-0,239	238,079	192,903	0,244	0,190	0,405	
	2,355	1	-0,402	141,784	270,273	0,310	0,258	0,478	
	2,388	3	-0,288	207,583	224,515	0,260	0,203	0,425	
	2,524	2	-0,344	177,774	251,626	0,279	0,224	0,450	
14	0,329	0	-0,184	110,993	237,525	0,289	0,223	0,291	
	0,329	4	-0,256	93,704	249,486	0,173	0,250	0,412	
	0,499	3	-0,244	96,844	246,279	0,180	0,250	0,403	
	0,698	2	-0,225	101,634	242,319	0,192	0,248	0,387	
	0,836	1	-0,198	108,043	238,545	0,217	0,242	0,354	

ID – Identificação da estação fluviométrica; **Zw** – Teste de Wang; **K** – Parâmetro de forma; **A** – Parâmetro de escala; **B** – Parâmetro de posição; **CV** – Coeficiente de variação; **CC** – Coeficiente de curtose; **CA** – coeficiente de assimetria.

Fonte: Autor (2019).