

Artigo Técnico

Tamanho ideal de séries históricas de vazões mínimas anuais e suas aplicações em estudos hidrológicos: estudo de caso nas regiões hidrográficas Paraná e São Francisco

Ideal size of minimum annual flow historical series and its applications in hydrological studies: case study in the Paraná and São Francisco Hydrographic Regions

Thamyres Cardoso Chaves Oliveira^{1*} , Flávio Aparecido Gonçalves¹ 

RESUMO

O uso de séries históricas de registros hidrológicos é uma das ferramentas base para a gestão dos recursos hídricos, contudo diversas regiões não possuem registros históricos longos. Dessa forma, o objetivo geral deste trabalho foi definir o tamanho de série amostral ideal para séries históricas de vazões mínimas. Foram utilizadas 19 estações fluviométricas da Região Hidrográfica Paraná (RHP) e outras 35 da Região Hidrográfica São Francisco (RHFSF), com o critério de possuir no mínimo 50 anos de registros de dados diários. Geraram-se séries de mínimas anuais para cada uma das estações, nas quais se aplicaram os testes não paramétricos de aleatoriedade, estacionariedade, independência e homogeneidade, com a finalidade de verificar a possibilidade de aplicar distribuições de probabilidade aos dados. Em seguida, o teste de aderência de Kolmogorov-Smirnov foi aplicado com o propósito de constatar a melhor distribuição ajustada para cada estação. As séries de Q_7 e de vazão de permanência foram constituídas para as séries completas e específicas (parciais), possibilitando obter os valores médios, além dos valores da Q_{70} e da $Q_{90\%}$ de cada série. Os resultados obtidos por meio das séries completas e específicas foram confrontados estatisticamente por meio do PBIAS e do teste t de Student. O estudo mostrou que, no geral, os resultados obtidos por meio das séries com tamanho médio de sete a 15 anos se apresentaram satisfatórios, ou estatisticamente iguais, aos resultados da série completa, podendo indicar que nem sempre possuir poucos dados de registro diário de uma série hidrológica inviabilize os estudos necessários.

Palavras-chave: ciclo hidrológico; gestão de recursos hídricos; bacia hidrográfica.

ABSTRACT

The use of historical series of hydrological records is one of the basic tools for the management of water resources. However, several regions do not have long historical records. Thus, the general objective of this study was to define the ideal sample series size for historical series of minimum flows. Nineteen fluviometric stations from the Paraná Hydrographic Region (RHP) and another 35 from the São Francisco Hydrographic Region (RHFSF) were used, with the criterion of having at least 50 years of daily data records. Annual minimum series were generated for each of the stations, to which the non-parametric tests of randomness, stationarity, independence and homogeneity were applied, in order to verify the possibility of applying probability distributions to the data. Then, the Kolmogorov-Smirnov adherence test was applied in order to verify the best fitted distribution for each station. The Q_7 and the permanence flow series were constituted for the complete and specific (partial) series, making it possible to obtain the values of means, Q_{70} and $Q_{90\%}$ of each series. The results obtained through the complete and specific series were statistically compared using the PBIAS and Student's t test. The study showed that, in general, the results obtained through series of seven to fifteen years were satisfactory, or statistically equal, to the results of the complete series, which may indicate that not always having few daily record data of hydrological series makes the necessary studies unfeasible.

Keywords: hydrological cycle; water resource management; hydrographic basin.

INTRODUÇÃO

As séries históricas de dados hidrológicos têm sido utilizadas em diversos estudos, com diferentes aplicações e interesses. Na detecção de tendências nas séries de vazão e precipitação, na construção de obras hidráulicas como barragens,

reservatórios e sistemas de drenagem urbana (ULIANA *et al.*, 2015), na elaboração de modelos com base em séries temporais (DETZEL, 2015), no planejamento ambiental, na tomada de decisões de uma gestão hídrica adequada (ROCHA; SANTOS, 2018), na análise de disponibilidade hídrica (BARBAROTTO

¹Universidade Federal de Alfenas - Poços de Caldas (MG), Brasil.

*Autora correspondente: thamy.oli@hotmail.com

Conflitos de interesse: não há conflitos de interesse.

Financiamento: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Recebido: 20/08/2021 - Aceito: 10/02/2022 - Reg. ABES: 20210232

JUNIOR, 2014), entre tantos outros estudos e projetos, é imprescindível ter em mãos um histórico hidrológico da região, um banco de dados com qualidade e quantidade adequadas conhecido como séries históricas.

As séries históricas de vazões mínimas são muito úteis na aplicação de licenças ambientais, na outorga e no uso da água, na avaliação da disponibilidade hídrica natural, bem como no planejamento de bacias hidrográficas e nos limites de despejo de cargas poluidoras tolerados pelo corpo hídrico (LUIZ, 2013; SMAKHTIN, 2001 *apud* OLIVEIRA; PRUSKI; NUNES, 2013). As vazões mínimas de referência são amplamente usadas para esse fim, sendo muito comuns a vazão mínima de referência com sete dias de duração e período de retorno de 10 anos ($Q_{7,10}$) e as vazões de permanência em 90 ($Q_{90\%}$) e em 95% ($Q_{95\%}$).

O estudo de caso realizado no Rio Paraguaçu (BA) por Genz, Lessa e Cirano (2008) mostra a importância do conhecimento da vazão mínima para o equilíbrio de salinidade e, conseqüentemente, da vida num corpo hídrico. Eles avaliaram o impacto sofrido por um estuário desde a construção da Barragem de Pedra do Cavalo, a qual por 12 anos impôs ao estuário redução de vazão que ocasionou no aumento da salinidade, e como a realidade mudou (para melhor) quando passaram a considerar a vazão mínima estabelecida pela licença ambiental de operação.

A Organização Meteorológica Mundial (OMM) recomenda a utilização de um período de 30 anos de dados no caso de monitoramentos climáticos e especificou o período de 1º de janeiro de 1961 a 31 de dezembro de 1990 como de referência, sob a denominação de normal climatológica de referência. Essa padronização possibilita a comparação de dados e observações ao redor de todo o mundo (WMO, 2017). Em 2019, a OMM publicou a Regulamentação Técnica nº 49, atualizando o conceito já outrora existente de normal climatológica padrão, o qual agora é composto dos 30 anos completos consecutivos mais recentes, e o 30º ano deve terminar com o dígito 0 (WMO, 2017; 2019). Dessa maneira, a normal climatológica padrão atual (na data de escrita deste artigo) é de 1991-2020, e a próxima será de 2001-2030.

Segundo Gomes e Blanco (2021), as análises realizadas com os dados hidrológicos podem ser limitadas por alguns fatores, incluindo a qualidade e o tamanho da série histórica disponível. A dificuldade, porém, encontrada na prática é que nem toda localidade possui séries históricas longas, ou pelo menos com os 30 anos de dados recomendados pela OMM, e por isso diversos modelos de previsão têm sido criados para contornar essas dificuldades (DARBANDSARI; COULIBALY, 2020), tais como correlação espacial (BETTERLE; SCHIRMER; BOTTER, 2019), regressão linear múltipla, modelo autorregressivo de médias móveis, rede neural artificial (SOUSA; SOUSA, 2010) e os modelos hidrológicos que consideram características físicas da região (MEDINA; MUÑOZ, 2020).

Aqui no Brasil, há registros pluviométricos na estação Morro Velho, Minas Gerais, desde 1855, e fluviométricos, de níveis, na estação Rio Negro, Manaus, desde 1902 (ANA, 2009), contudo a realidade da maior parte do país são estações antigas que hoje não se encontram mais em funcionamento, ou estações com várias falhas de registros ou grandes interrupções ao longo dos anos, ou estações novas, com poucos anos de dados registrados. Assim, muitas vezes não é possível a utilização de séries com grandes períodos em localidades específicas.

Embora existam estações com muitos anos de registro de dados hidrológicos no Brasil (algumas com 100 anos), a grande maioria ainda apresenta várias falhas nesses registros, interrupções de coleta de dados por longo período de tempo e até mesmo de forma definitiva, não sendo possível, muitas vezes, o uso de séries com grandes períodos em localidades específicas, de forma a atender às recomendações da OMM.

Um estudo semelhante ao proposto neste trabalho foi feito com dados da Austrália. Nele, foram analisados os seguintes parâmetros: precipitação total, número de dias com precipitação superior ou igual a 1 mm, temperatura média máxima, temperatura média mínima, pressão média ao nível do mar, duração total solar e pressão de vapor média (WMO, 2007; 2017). Os autores chegaram ao resultado de que, no geral, séries menores, com 10-12 anos de dados, fornecem respostas similares à série maior de 30 anos (com exceção dos parâmetros de extremos e outras propriedades estatísticas de ordem superior, tais quais os quantis). Ressaltaram que, em estudos cujo interesse está nos valores extremos, quanto mais dados, melhor, porém, nessas situações, observaram que entre 10 e 15 anos de dados, no geral, já se obtém uma informação dos extremos prováveis de ocorrer nos próximos anos naquela região.

Esse resultado, uma vez também observado em outros estudos e localidades, permite que normais menores sejam calculadas, de forma segura, para um número muito maior de estações, já que, conforme dito, nem toda localidade possui registros suficientes para a utilização da normal climatológica de referência ou da normal climatológica padrão.

Outro estudo, realizado com séries de vazões máximas e mínimas do Rio Allipén, no Chile (MEDINA; MUÑOZ, 2020), buscou analisar um método para estimativas de tais vazões em bacias hidrográficas cuja base de dados era precária, ou seja, com poucos anos de dados disponíveis. Como resultado, recomendaram um tamanho mínimo de 16 anos de dados para que as estimativas realizadas estivessem em conformidade com uma incerteza controlável.

Huang e Bardossy (2020) também estudaram o impacto da qualidade e da quantidade dos dados disponíveis das séries históricas na calibração de modelos hidrológicos, da região do Médio Atlântico, Estados Unidos. Observaram que ao usar séries de 10 anos de dados se podem alcançar parâmetros estimados confiáveis, com *performance* de 90% de validação.

São poucos os estudos existentes relacionados a tamanhos de série histórica de dados hidrológicos. Assim, nota-se a importância de estudos tais como neste trabalho, que tem como objetivo definir o tamanho de série amostral ideal para séries históricas de dados fluviométricos de mínimas, por meio dos seguintes objetivos específicos:

- obter séries históricas de longo período de dados de vazão;
- constituir séries históricas amostrais de vazão;
- correlacionar os parâmetros estatísticos das séries de longo período com os dados das séries amostrais;
- indicar o tamanho de série amostral ideal considerando as séries de vazões mínimas.

MATERIAL E MÉTODOS

Área de estudo

Ressalta-se que a metodologia aqui proposta é aplicável e reproduzível em qualquer bacia ou região hidrográfica desejável, não sendo exclusiva para as duas regiões hidrográficas aqui escolhidas. Nas Figuras 1 e 2, encontram-se dois fluxogramas sintetizando a metodologia aplicada.

O presente estudo foi realizado em duas regiões hidrográficas: a Paraná (RHP) e a São Francisco (RHFSF). Juntas, elas abrangem 17% do território brasileiro em 10 estados, além do Distrito Federal: São Paulo, Paraná, Mato Grosso do Sul, Minas Gerais, Goiás e Santa Catarina (RHP, conforme Figura 3A) e Bahia, Minas Gerais, Pernambuco, Alagoas, Sergipe e Goiás (RHFSF, conforme Figura 3B).

Seleção das estações fluviométricas

Com base no Portal HidroWeb, disponível no *site* da Agência Nacional de Águas e Saneamento Básico (ANA, 2020b), foram selecionadas 54 estações de séries históricas fluviométricas (19 da RHP e 35 da RHSF), com o critério de possuir, pelo menos, 50 anos de dados completos, ou seja, excluindo-se os anos com falhas de preenchimento.

Na Tabela 1 estão apresentadas as estações selecionadas com as informações dos códigos, bem como seu nome, a sub-bacia hidrográfica a que pertencem, rio em que a estação está inserida, localização geográfica, município e estado, altitude e área de drenagem (A_d). Todos os códigos iniciados pelo dígito 6 dizem respeito à RHP, enquanto os códigos que começam com o dígito 4 se referem à RHSF.

Testes não paramétricos

Segundo Mello e Silva (2013), a série histórica estudada precisa ter os testes não paramétricos aceitos para que se possam aplicar as distribuições de probabilidade. Por isso, aplicou-se a metodologia descrita na bibliografia supracitada para o cálculo dos testes não paramétricos de aleatoriedade, estacionariedade, independência e homogeneidade, com suas respectivas estatísticas T calculadas tal como se segue.

Aleatoriedade

A estatística T para o teste de aleatoriedade é calculado pela Equação 1:

$$T = \frac{N_i - E(N_i)}{\sqrt{\text{var}(N_i)}} \quad (1)$$

Em que:

N_i = o número de inflexões principais observadas na série histórica;
 $E(N_i)$ = a esperança;

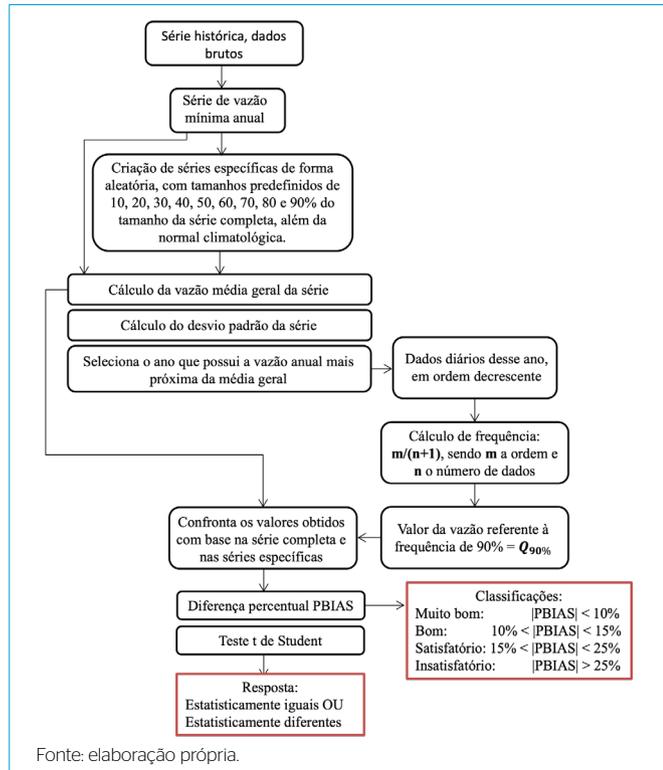


Figura 1 - Q90% e média das vazões mínimas anuais.

$\text{var}(N_i)$ = a variância de N_i .

Esses elementos são calculados pelas Equações 2 e 3, respectivamente:

$$E(N_i) = \frac{2(N - 2)}{3} \quad (2)$$

$$\text{var}(N_i) = \frac{16N - 29}{90} \quad (3)$$

Em que:

N = o número de dados da série.

Estacionariedade

A estatística T para o teste de estacionariedade é calculado pela Equação 4:

$$T = \frac{cs}{\sqrt{\text{var}(cs)}} \quad (4)$$

O coeficiente cs é determinado pela Equação 5, e a variância de cs , pela Equação 6:

$$cs = 1 - \frac{6 \sum_{i=1}^N (F_i - f_i)^2}{N^3 - N} \quad (5)$$

$$\text{var}(cs) = \frac{1}{N - 1} \quad (6)$$

O valor de f_i é obtido pela posição temporal do dado na série histórica, e F_i é a posição desse mesmo dado com a série organizada em ordem crescente.

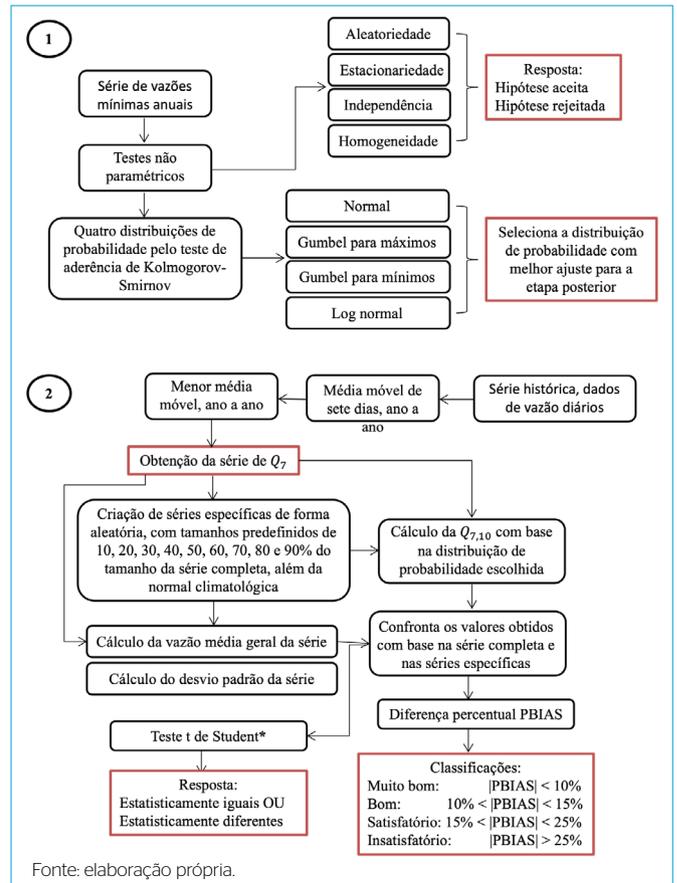
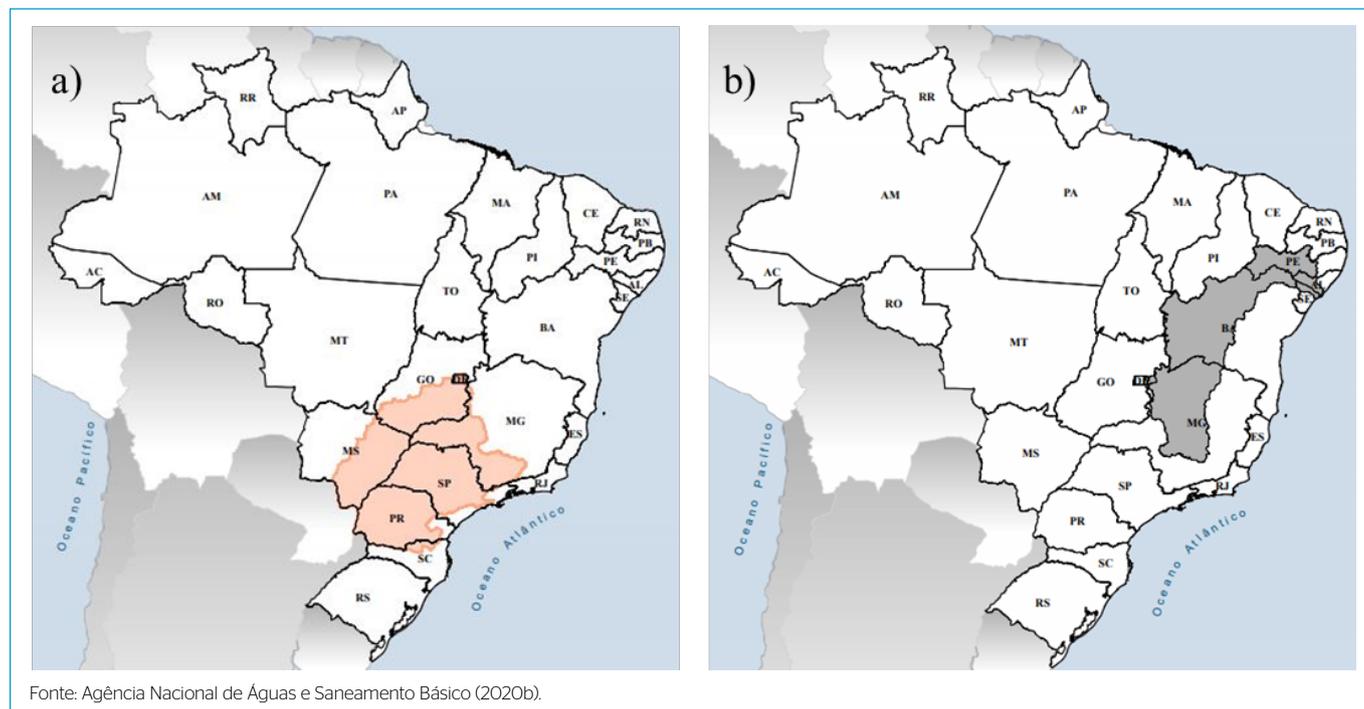


Figura 2 - Série de Q_7 e vazão $Q_{7,10}$.



Fonte: Agência Nacional de Águas e Saneamento Básico (2020b).

Figura 3 - (A) Região Hidrográfica Paraná; (B) Região Hidrográfica São Francisco.

Independência

A estatística T para o teste de independência é calculado pela Equação 7:

$$T = \frac{r - E(r)}{\sqrt{\text{var}(r)}} \tag{7}$$

r é calculado pela Equação 8:

$$r = \left(\sum_{i=1}^{N-1} Y_i \times Y'_{i+1} \right) + Y'_1 \times Y'_N \tag{8}$$

Em que:

Y'_i = a subtração entre o valor de Y_i e a média de todos os valores da série \bar{Y} , como na Equação 9:

$$Y'_i = Y_i - \bar{Y} \tag{9}$$

A esperança $E(r)$ e a variância $\text{var}(r)$ de r são calculados pelas Equações 10 e 11:

$$E(r) = -\left(\frac{s_2}{N-1} \right) \tag{10}$$

$$\text{var}(r) = \left(\frac{s_2^2 - s_4}{N-1} \right) + \left(\frac{s_2^2 - 2s_4}{(N-1)(N-2)} \right) - \left(\frac{s_2^2}{(N-1)^2} \right) \tag{11}$$

sendo s_2 e s_4 momentos amostrais de segunda e de quarta ordem, obtidos pelas Equações 12 e 13:

$$s_2 = \sum_{i=1}^N (Y'_i)^2, \tag{12}$$

$$s_4 = \sum_{i=1}^N (Y'_i)^4 \tag{13}$$

Homogeneidade

A estatística T para o teste de homogeneidade é calculado pela Equação 14:

$$T = \frac{V - E(V)}{\sqrt{\text{var}(V)}} \tag{14}$$

O valor de V é determinado calculando-se os valores de V_1 e V_2 das subamostras 1 e 2, respectivamente, conforme Equações 15 e 16:

$$V_1 = N_1 N_2 + \frac{N_1(N_1 + 1)}{2} - H_1 \tag{15}$$

$$V_2 = N_1 N_2 - V_1 \tag{16}$$

Em que:

N_1 e N_2 = o tamanho das subamostras 1 e 2 (números inteiros, sendo o N_2 maior que N_1 em caso de divisão de número ímpar);

H_1 = a soma das ordens de classificação da série histórica da subamostra 1 (ou seja, a soma de F_i do teste de estacionariedade correspondente à primeira subamostra). Adota-se como V na Equação 14 o menor valor entre V_1 e V_2 .

A esperança e a variância de V são obtidas pelas Equações 17 e 18, respectivamente.

$$E(V) = \frac{N_1 N_2}{2}, \tag{17}$$

$$\text{var}(V) = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12} \tag{18}$$

Para todos os testes não paramétricos aqui citados, a hipótese inicial (isto é, que a série possui dados aleatórios, estacionários, independentes e homogêneos) é rejeitada caso $|T| > z_{(1-\frac{\alpha}{2})}$, em que α é o nível de significância (de 5%

Tabela 1 - Estações fluviométricas selecionadas e informações pertinentes.

Código	Nome da estação	Sub-bacia	Rio	Município	Latitude	Longitude	Altitude (m)	A _d (km²)
68100000	Corrientes (408-MOSP)	Rios Paraná-Corrientes	Rio Paranaíba	Argentina	-2745	-58.85	42	1950000
64843000	Guaira Porto	Rios Paraná-Paranapanema	Rio Paraná	Guaira (PR)	-24.0694	-54.2489	218	804.000
61052000	Andrelândia	Rio Grande	Rio Turvo Pequeno	Andrelândia (MG)	-21.7381	-44.3056	951	274
61250000	Fazenda da Guarda	Rio Grande	Rio Sapucaí	Campos do Jordão (SP)	-22.6878	-45.4797	1501	109
64465000	Tibagi	Rios Paraná-Paranapanema	Rio Tibagi	Tibagi (PR)	-24.5094	-50.4097	750	8.840
60850000	Fazenda Buriti do Prata	Rio Paranaíba	Rio da Prata	Prata (MG)	-19.3597	-49.1803	517	2.460
62715000	Artemis	Rios Paraná-Tietê	Rio Piracicaba	Piracicaba (SP)	-22.6792	-47.7753	445	10.900
61122000	Vila Rio das Mortes	Rio Grande	Rio das Mortes Pequeno	São João Del Rei (MG)	-21.1956	-44.3286	870	272
61100000	Ibertioga	Rio Grande	Rio Elvas	Ibertioga (MG)	-21.4444	-43.9633	980	186
60150000	Estrela do Sul	Rio Paranaíba	Rio Bagagem	Estrela do Sul (MG)	-18.7389	-47.6897	720	868
65365000	Porto Vitória	Rio Paraná-Iguaçu	Rio Espingarda	Porto Vitória (PR)	-26.1775	-51.22	790	165
60635000	Inhumas	Rio Paranaíba	Rio Meia Ponte	Inhumas (GO)	-16.3464	-49.4942	900	532
60250000	Fazenda São Mateus	Rio Paranaíba	Rio Quebra Anzol	Ibiá (MG)	-19.5175	-46.57	852	1.300
60130000	Fazenda Cachoeira	Rio Paranaíba	Rio Perdizes	Monte Carmelo (MG)	-18.7808	-47.4089	616	131
60265000	Ibiá	Rio Paranaíba	Rio Misericórdia	Ibiá (MG)	-19.475	-46.5419	880	1.370
60615000	Fazenda Cachoeira	Rio Paranaíba	Rio Pousou Alegre	Tupaciguara (MG)	-18.6983	-48.7806	793	195
60100000	Charqueada do Patrocínio	Rio Paranaíba	Rio Dourados	Patrocínio (MG)	-18.9011	-46.9658	906	78
60220000	Desemboque	Rio Paranaíba	Rio Araguari	Sacramento (MG)	-20.0139	-47.0172	960	1.070
60145000	Irai de Minas	Rio Paranaíba	Rio Bagagem	Irai de Minas (MG)	-18.9772	-47.4575	946	93
40025000	Vargem Bonita	Rios São Francisco-Paraopeba	Rio São Francisco	Vargem Bonita (MG)	-20.3278	-46.3658	744	301
40040000	Fazenda Ajudas	Rios São Francisco-Paraopeba	Rio Ajudas	Bambuí (MG)	-20.0953	-46.0622	681	244
40050000	Iguatama	Rios São Francisco-Paraopeba	Rio São Francisco	Iguatama (MG)	-20.17	-45.7158	624	5.560
40100000	Porto das Andorinhas	Rios São Francisco-Paraopeba	Rio São Francisco	Abaeté (MG)	-19.2814	-45.2753	532	14.000
40150000	Carmo do Cajuru	Rios São Francisco-Paraopeba	Rio Pará	Carmo do Cajuru (MG)	-20.1811	-44.7939	719	2.500
40330000	Velho da Taipa	Rios São Francisco-Paraopeba	Rio Pará	Conceição do Pará (MG)	-19.6939	-44.9308	575	7.350
40400000	Estação Álvaro da Silveira	Rios São Francisco-Paraopeba	Rio Lambari	Bom Despacho (MG)	-19.7536	-45.1167	648,5	1.820
40680000	Entre Rios de Minas	Rios São Francisco-Paraopeba	Rio Brumado	Entre Rios de Minas (MG)	-20.6611	-44.0722	840	487
40800001	Ponte Nova do Paraopeba	Rios São Francisco-Paraopeba	Rio Paraopeba	Juatuba (MG)	-19.9489	-44.305	695	5.690
40930000	Barra do Funchal	Rios São Francisco-Paraopeba	Rio Indaiá	Serra da Saudade (MG)	-19.3956	-45.8864	679	897
41250000	Vespasiano	Rios São Francisco-das Velhas	Ribeirão da Mata	Vespasiano (MG)	-19.6875	-43.9206	674	709
41300000	Taquaraçu	Rios São Francisco-das Velhas	Rio Taquaraçu	Taquaraçu de Minas (MG)	-19.6536	-43.6875	757	618
41340000	Ponte Raul Soares	Rios São Francisco-das Velhas	Rio das Velhas	Lagoa Santa (MG)	-19.5597	-43.9111	637	4.860
41380000	Ponte Preta	Rios São Francisco-das Velhas	Rio Jaboticatubas	Jaboticatubas (MG)	-19.4619	-43.9036	752	564
41818000	Santo Hipólito (ANA/Cemig)	Rios São Francisco-das Velhas	Rio das Velhas	Santo Hipólito (MG)	-18.3061	-44.2258	499	16.600
41990000	Várzea da Palma	Rios São Francisco-das Velhas	Rio das Velhas	Várzea da Palma (MG)	-17.5961	-44.7133	488	26.500
42210000	Cachoeira da Manteiga	Rios São Francisco-Paracatu	Rio São Francisco	Buritizero (MG)	-16.6572	-45.0808	461	107.000
42395000	Santa Rosa	Rios São Francisco-Paracatu	Rio Paracatu	Paracatu (MG)	-17.255	-46.4728	490	12.800
42600000	Porto dos Poções	Rios São Francisco-Paracatu	Rio Preto	Dom Bosco (MG)	-16.8397	-46.3572	512	9.400
42690001	Porto da Extrema	Rios São Francisco-Paracatu	Rio Paracatu	Brasilândia de Minas (MG)	-17.0303	-46.0136	484	30.100
43200000	São Romão	Rios São Francisco-Uruçuaia	Rio São Francisco	São Romão (MG)	-16.3725	-45.07	457,3	154.000
44200000	São Francisco	Rios São Francisco-Verde Grande	Rio São Francisco	São Francisco (MG)	-15.9494	-44.8678	448	184.000
44500000	Manga	Rios São Francisco-Verde Grande	Rio São Francisco	Manga (MG)	-14.7572	-43.9322	433,8	202.000
45131000	São Gonçalo	Rios São Francisco-Carinhanha	Rio Carinhanha	Montalvânia (MG)	-14.3136	-44.4594	449	6.020
45298000	Carinhanha	Rios São Francisco-Carinhanha	Rio São Francisco	Carinhanha (BA)	-14.3044	-43.7633	427	254.000
46150000	Ibotirama	Rios São Francisco-Grande	Rio São Francisco	Ibotirama (BA)	-12.1825	-43.2231	408	323.000
46360000	Morpará	Rios São Francisco-Grande	Rio São Francisco	Morpará (BA)	-11.5583	-43.2825	399	345.000
46550000	Barreiras	Rios São Francisco-Grande	Rio Grande	Barreiras (BA)	-12.1525	-45.0094	437	24.400
46650000	Taguá	Rios São Francisco-Grande	Rio Grande	Cotegipe (BA)	-11.7208	-44.5022	459	36.000
46675000	Fazenda Macambira	Rios São Francisco-Grande	Rio Grande	Cotegipe (BA)	-11.6106	-44.1567	420	39.600
46790000	Formosa do Rio Preto	Rios São Francisco-Grande	Rio Preto	Formosa do Rio Preto (BA)	-11.0514	-45.1969	479	14.300
46830000	Ibipetuba	Rios São Francisco-Grande	Rio Preto	Santa Rita de Cássia (BA)	-11.0058	-44.5242	432	17.900
46902000	Boqueirão	Rios São Francisco-Grande	Rio Grande	Mansidão (BA)	-11.345	-43.8261	401	65.900
48020000	Juazeiro	Rios São Francisco-Pajeú	Rio São Francisco	Juazeiro (BA)	-9.4064	-40.5036	357,74	516.000
48290000	Santa Maria da Boa Vista	Rios São Francisco-Pajeú	Rio São Francisco	Santa Maria da Boa Vista (PE)	-8.8097	-39.8239	343,67	535.000

ANA: Agência Nacional de Águas e Saneamento Básico; Cemig: Companhia Energética de Minas Gerais.

Fonte: elaboração própria.

neste estudo), e z , o valor tabelado encontrado na literatura para distribuição normal. Portanto, aqui, a hipótese foi rejeitada sempre que o módulo de T era maior que 1,96.

Tais testes foram realizados para as séries de vazões mínimas anuais de cada uma das 54 estações fluviométricas selecionadas, obtendo-se como resposta a aceitação ou rejeição da hipótese de cada teste.

Teste de aderência e distribuição de probabilidade

Em seguida, aplicou-se o teste de aderência de Kolmogorov-Smirnov para quatro distribuições de probabilidade, a fim de verificar se a amostra analisada (os dados das séries de vazões mínimas de cada uma das estações) pode ser considerada como proveniente de uma população com determinada distribuição de probabilidade. Tal teste calcula a diferença entre as frequências observadas e as frequências teóricas, sendo estas calculadas por meio de distribuição de probabilidade. Seleciona-se o módulo da maior diferença entre as frequências, e este é comparado com o valor obtido por meio da tabela de Kolmogorov-Smirnov, disponível em literatura, a qual considera o número de dados e o nível de significância desejado. Se o módulo da maior diferença for menor que o valor tabelado, tem-se que a distribuição de probabilidade analisada é adequada à série de dados (MELLO; SILVA, 2013).

As distribuições de probabilidade analisadas foram: normal; Gumbel para máximos; Gumbel para mínimos; e *log normal*. Todos os cálculos foram realizados para o caso de não excedência.

Distribuição normal

Com os dados ordenados de forma crescente, calcularam-se média e desvio padrão dos dados, e aplicou-se a função de distribuição normal do Excel [=DIST.NORM.N(x; média; desv_padrão; 1)].

Gumbel para máximos

Com os dados ordenados de forma crescente, calcularam-se média \bar{X} , desvio padrão σ , além de dois parâmetros α e μ pelo método dos momentos (Equações 19 e 20):

$$\alpha = \frac{1,2826}{\sigma}, \quad (19)$$

$$\mu = \bar{x} - 0,45\sigma. \quad (20)$$

A frequência teórica F com Gumbel para máximos é dada pela Equação 21:

$$F = \exp\{-\exp[-\alpha(x - \mu)]\}. \quad (21)$$

Gumbel para mínimos

Com os dados ordenados de forma crescente, calcularam-se os parâmetros de média \bar{X} , desvio padrão σ , e α (Equação 19) e μ pelo método dos momentos (Equação 22):

$$\mu = \bar{x} + 0,45\sigma. \quad (22)$$

A frequência teórica F com Gumbel para máximos é dada pela Equação 23:

$$F = 1 - \exp\{-\exp[\alpha(x - \mu)]\}. \quad (23)$$

Log normal

Aplicaram-se a transformação logarítmica aos dados e, em seguida, a função do Excel: [=DIST.LOGNORMAL.N(x; média; desv_padrão; 1)]. Nessa função, o valor de x é sem a transformação logarítmica, enquanto a média e o desvio padrão são calculados com os dados transformados.

Dessa forma, foi possível verificar se há ajuste das distribuições aos dados das séries de vazões mínimas e, em caso afirmativo, qual delas melhor se ajusta para cada uma das estações.

Série Q_7 (completa e específica) e respectiva vazão $Q_{7,10}$

A série de Q_7 foi constituída ano a ano com base nos dados diários disponíveis na série histórica bruta. Cada ano recebeu um único valor de vazão, referente à menor média de sete dias consecutivos do ano em questão.

Assim, a série de Q_7 completa refere-se à série feita com todos os anos sem falhas disponíveis na série histórica. A série de Q_7 específica diz respeito às séries geradas com base na completa, porém com tamanhos variados – 10, 20, 30, 40, 50, 60, 70, 80 e 90% do tamanho original –, além da normal climatológica (contendo 30 anos consecutivos, segundo as especificações da OMM).

Os anos que compuseram as séries específicas foram escolhidos com o auxílio de uma ferramenta que fornecia números aleatórios. A cada porcentagem, novos números foram gerados, garantindo assim a aleatoriedade dos dados para fins de análise. Em relação à normal climatológica, seguiu-se o subsequente critério (de acordo com as recomendações da OMM): deu-se prioridade para a normal climatológica de referência, de 1961 a 1990; para as séries que não possuíam esses anos sem falhas, selecionou-se a normal climatológica padrão mais recente, de 1981 a 2010 (conforme Instituto Nacional de Meteorologia — INMET, 2020); e no caso de não ter também esses anos completos, selecionaram-se os 30 anos consecutivos mais recentes, com o primeiro ano terminando com o dígito 1. No caso de uma das estações que não pôde atender a nenhum dos critérios supracitados, foram selecionados os 30 anos consecutivos mais recentes.

A vazão $Q_{7,10}$ foi calculada tanto para a série completa de Q_7 quanto para as séries específicas geradas com base nela, para cada uma das estações. O cálculo foi realizado seguindo a metodologia da distribuição de probabilidade mais bem ajustada para cada estação.

Vazão mínima anual (séries completa e específica) e respectiva vazão de permanência $Q_{90\%}$

Para a vazão de permanência, utilizou-se a série de vazões mínimas anuais, e encontrou-se o ano que possuía a vazão mais próxima da média da série de cada estação, ou seja, o ano que se mostrou mais próximo do comportamento médio da respectiva série completa. Em seguida, usaram-se os dados diários desse ano específico para a obtenção da vazão de permanência desejada, que neste trabalho foi a $Q_{90\%}$.

Esse mesmo procedimento foi realizado tanto para a série completa quanto para as séries específicas. Vale ressaltar que as séries específicas foram feitas seguindo os mesmos critérios da Q_7 , ou seja, aleatoriamente.

Diferenças percentuais e teste t de Student

Por fim, para verificar se as informações obtidas por meio das séries específicas eram condizentes ou semelhantes às informações da série completa, calcularam-se as diferenças percentuais, PBIAS, entre a vazão média da série completa

Q_{comp} ($m^3.s^{-1}$) e as vazões das séries específicas Q_{esp} ($m^3.s^{-1}$), conforme Equação 24. Aplicou-se essa equação, também, nos valores de $Q_{7,10}$ e de $Q_{90\%}$, comparando-se assim a diferença entre tais vazões obtidas por meio da série completa e das séries específicas. O PBIAS é considerado muito bom quando seu módulo é inferior a 10%; bom, entre 10 e 15%; satisfatório, entre 15 e 25%; e insatisfatório, quando superior a 25% (VALLE JUNIOR; RODRIGUES; OLIVEIRA, 2019).

$$PBIAS = \frac{Q_{comp} - Q_{esp}}{Q_{comp}} \times 100. \quad (24)$$

Também se analisaram os resultados estatisticamente pelo teste *t* de Student, com nível de significância de 5%.

Vale ressaltar que a análise de PBIAS foi realizada para todas as análises (série de Q_7 , vazões $Q_{7,10}$, série de vazões mínimas anuais – ou série de vazão de permanência – e vazões $Q_{90\%}$). O teste *t* de Student, por sua vez, foi aplicado apenas nas séries de Q_7 e de vazões mínimas anuais, já que foi feito utilizando-se média e desvio padrão, e não comparando-se valores pontuais.

RESULTADOS E DISCUSSÃO

Por causa de falhas nos registros de dados fluviométricos, houve variações do período utilizado para o cálculo das normais climatológicas. Na Tabela 2 estão apresentados a relação das estações e o respectivo período de 30 anos usado, seguindo recomendações da OMM, com exceção da estação 14, que não tem 30 anos de dados sem falhas que atenda a tais critérios. As estações 26, 29 e 32 não possuem nenhuma sequência de 30 anos consecutivos, e por isso, para elas, não foi feita a análise das normais climatológicas. Os números entre parênteses e em negrito são para facilitar a identificação e citação das estações nas tabelas e textos posteriores. Os números de 1 a 19 são referentes à RHP e de 20 a 54 à RHSE.

Testes não paramétricos e distribuição de probabilidade

Na Tabela 3, estão apresentados os testes não paramétricos aceitos em cada uma das 54 estações, além da distribuição de probabilidade escolhida com base no melhor ajuste em cada situação. Para a RHP, pode-se observar que apenas duas estações (estações 14 e 15) tiveram todos os testes não paramétricos aceitos, enquanto nove delas tiveram apenas um ou nenhum teste aceito. Já a RHSE não possui nenhuma estação que aceitou todos os testes e 21 estações que aceitaram nenhum ou um teste apenas. Nota-se, assim, o baixo número de estações com resultados positivos para os testes não paramétricos de aleatoriedade, estacionariedade, independência e homogeneidade. Resultado semelhante foi obtido por Rubatino *et al.* (2017) em um estudo de cotas e vazões na Bacia do Rio Doce, concluindo que isso pode apontar para constantes variações dos dados, seja em razão da forma como os dados são coletados, seja pela periodicidade com que são feitos, além de se tratar de dados de cota e vazão, os quais sofrem variações de cheias e secas ao longo dos anos, em intensidades diferentes.

Vale lembrar que, segundo Mello e Silva (2013), quando um teste não paramétrico não é aceito, a aplicação de distribuições de probabilidade não é recomendada. Para exemplificar, a hipótese de aleatoriedade assume que as variações ocorridas em determinada série histórica são todas decorrentes de causas naturais, e não por influência antrópica (como ocorre, por exemplo, na construção de barramentos). Se tal hipótese é rejeitada, provavelmente as variações observadas deixaram de ser aleatórias, ou seja, naturais, e a utilização da distribuição de probabilidade na previsão de alguma variável pode não ser

realista ou tão confiável, já que o comportamento hidrológico tem sido afetado pelo homem e pode variar além do previsto historicamente.

Em relação às distribuições de probabilidade, apenas três estações obtiveram melhor ajuste com Gumbel para mínimos, o que contrariou a expectativa inicial, já que se esperava que valores mínimos (vazões mínimas) se ajustassem melhor a esse tipo de distribuição de probabilidade, conforme observado no estudo de Melo (2016), realizado na Bacia Hidrográfica do Rio das Velhas, em que a distribuição Gumbel apresentou melhores ajustes na maioria dos testes realizados com vazões mínimas de referência. As distribuições de melhor ajuste foram normal e log normal, com 22 e 15 estações, respectivamente. Resultado semelhante foi apresentado por Uliana *et al.* (2011) e por Granemann e Mine (2015). Ao se obter o melhor ajuste para cada estação, tem-se a forma mais adequada para realizar estimativas da frequência ou probabilidade de ocorrência de determinados valores.

Embora os testes não paramétricos não tenham sido, em sua grande maioria, aceitos para as séries de vazões mínimas aqui analisadas, realizaram-se mesmo assim os cálculos das vazões de $Q_{7,10}$, a fim de se comparar os resultados obtidos por meio das séries longas e das séries menores. Esse era o objetivo principal do trabalho. Tais resultados podem não ser tão confiáveis para estudos que seriam aplicados na prática, como na construção de obras hidráulicas, por exemplo. Contudo, como aqui se objetivou comparar tamanhos de séries, supõe-se que a mesma incerteza embutida na série longa se encontra também na série menor, não atrapalhando a comparação aqui pretendida.

Série Q_7 (completa e específica) e respectiva vazão $Q_{7,10}$

A Figura 4 apresenta as respostas obtidas para a análise de diferença percentual para os vários tamanhos de série analisados, considerando-se o parâmetro média das séries de Q_7 , tanto para a RHP quanto para a RHSE. Nota-se a predominância dos pontos localizados abaixo da linha do PBIAS de 10% (classificação muito bom), representando mais de 87% dos casos analisados.

Foi constatado que séries com tamanho médio de 15 anos representam mais de 70% da classificação muito bom, ou seja, séries em torno de 15 anos de dados fornecem resultados semelhantes aos das séries longas. Além disso, é interessante notar que as normais climatológicas apresentam comportamento similar ao das séries de 10%, como pode ser visto também na Figura 4.

Observa-se, ainda, a correlação entre o aumento do tamanho da série e a predominância de ocorrência de diferença percentual inferior a 6%.

Em relação ao teste *t*, são estes os resultados de diferença estatística entre as séries específicas e a série completa: 10 ocorrências de diferença estatística nas normais climatológicas; duas ocorrências na série específica de 10%; uma em 20%; e uma em 50%. Dessa forma, no caso das regiões hidrográficas estudadas, trabalhar com apenas 10% dos anos disponíveis (o que significa utilizar sete anos de dados) para o cálculo da Q_7 não se mostrou, no geral, estatisticamente diferente de trabalhar com a série completa.

Na Figura 5 está a diferença percentual para as vazões $Q_{7,10}$ para os vários tamanhos de séries parciais geradas. Nela, retiraram-se cinco pontos localizados acima de 80% de diferença percentual para melhorar a visualização do comportamento gráfico dos dados. Esses cinco pontos possuíam séries parciais de seis, 30, 30, 54 e 57 anos. Semelhantemente ao observado na série de Q_7 , há a predominância da classificação muito bom também às vazões $Q_{7,10}$, com mais de 71% de todos os casos analisados assim classificados. Novamente, a

Tabela 2 - Período considerado como normal climatológica (NC) para cada estação deste estudo.

Nomenclatura	Estação	NC	Nomenclatura	Estação	NC
(1)	68100000	1961-1990	(28)	40800001	1961-1990
(2)	64843000	1961-1990	(29)	40930000	Não possui
(3)	61052000	1961-1990	(30)	41250000	1961-1990
(4)	61250000	1951-1980	(31)	41300000	1961-1990
(5)	64465000	1981-2010	(32)	41340000	Não possui
(6)	60850000	1961-1990	(33)	41380000	1961-1990
(7)	62715000	1961-1990	(34)	41818000	1961-1990
(8)	61122000	1961-1990	(35)	41990000	1961-1990
(9)	61100000	1961-1990	(36)	42210000	1961-1990
(10)	60150000	1961-1990	(37)	42395000	1961-1990
(11)	65365000	1961-1990	(38)	42600000	1961-1990
(12)	60635000	1981-2010	(39)	42690001	1961-1990
(13)	60250000	1961-1990	(40)	43200000	1961-1990
(14)	60130000	1985-2014	(41)	44200000	1961-1990
(15)	60265000	1981-2010	(42)	44500000	1961-1990
(16)	60615000	1981-2010	(43)	45131000	1961-1990
(17)	60100000	1961-1990	(44)	45298000	1961-1990
(18)	60220000	1961-1990	(45)	46150000	1961-1990
(19)	60145000	1981-2010	(46)	46360000	1981-2010
(20)	40025000	1961-1990	(47)	46550000	1961-1990
(21)	40040000	1981-2010	(48)	46650000	1961-1990
(22)	40050000	1981-2010	(49)	46675000	1961-1990
(23)	40100000	1961-1990	(50)	46790000	1961-1990
(24)	40150000	1951-1980	(51)	46830000	1961-1990
(25)	40330000	1961-1990	(52)	46902000	1961-1990
(26)	40400000	Não possui	(53)	48020000	1961-1990
(27)	40680000	1981-2010	(54)	48290000	1961-1990

Fonte: elaboração própria.

normal climatológica apresentou comportamento similar ao das séries próximas de 10 anos de dados.

Embora as séries de normal climatológica estejam bem presentes nas classificações de insatisfatório e satisfatório (ou seja, com diferença percentual superior a 15%), não se notou nenhum padrão que pudesse indicar o melhor período de 30 anos a ser usado.

Não se viu correlação entre as diferenças percentuais nos valores calculados de $Q_{7,10}$ e a rejeição dos testes não paramétricos.

Vazão mínima anual (séries completa e específica) e respectiva vazão de permanência $Q_{90\%}$

Em relação ao parâmetro média das séries de vazões mínimas anuais, observou-se que mais de 87% dos casos estudados apresentaram resultados classificados como muito bom e menos de 1% como insatisfatório. Com séries de tamanho médio a partir de sete anos já se tem frequência de ocorrência classificada como muito boa próxima de 63%. Aumentando o tamanho médio da série para 15 anos, a frequência de ocorrência como muito boa sobe para quase 78% dos casos. A Figura 6 mostra visualmente o comportamento dos resultados obtidos.

Mais uma vez, as séries das normais climatológicas apresentam resultados gerais muito próximos aos das séries específicas de 10 e 20%, com tamanhos médios entre sete e 15 anos de dados.

Em relação ao teste t de Student, os resultados das diferenças estatísticas para a série de vazões mínimas são: 10 ocorrências na normal climatológica; três em 10%; e uma em 20%. Nota-se a predominância de ocorrência de diferença estatística nas séries das normais climatológicas no que tange às demais séries específicas.

A vazão $Q_{90\%}$ foi a que apontou a maior variação nos resultados, tendo ocorrência de todas as classificações em todas as séries específicas estudadas, como visto na Figura 7. Para melhor visualização do comportamento dos dados do gráfico da Figura 7, excluíram-se um ponto com 33 anos de dados, cuja diferença percentual era de 99%, e outros dois pontos com 30 e 36 anos de dados, cujo PBIAS era de 100% (muito maior que a diferença percentual dos demais pontos, fazendo com que o gráfico se achatasse).

As diferenças percentuais entre as séries completa e específicas ocorrem nas várias porcentagens analisadas, mesmo nas maiores. Foram 42 casos em que a diferença percentual foi superior a 25% – classificação insatisfatória –, correspondendo a menos de 8% dos casos analisados; e 61 casos com diferença

Tabela 3 - Testes não paramétricos aceitos e distribuição de probabilidade com melhor ajuste para cada estação. Série de vazões mínimas anuais.

Nomenclatura	Estação	Testes não paramétricos aceitos	Distribuição de probabilidade
(1)	68100000	-	Log normal
(2)	64843000	-	Log normal
(3)	61052000	E; H	Log normal
(4)	61250000	A	Log normal
(5)	64465000	A	Normal
(6)	60850000	-	Gumbel para máximos
(7)	62715000	H	Log normal
(8)	61122000	E; I; H	Normal
(9)	61100000	A; E	Gumbel para máximos
(10)	60150000	H	Gumbel para máximos
(11)	65365000	H	Log normal
(12)	60635000	A; E; H	Normal
(13)	60250000	A; E	Normal
(14)	60130000	A; E; I; H	Normal
(15)	60265000	A; E; I; H	Normal
(16)	60615000	A	Normal
(17)	60100000	A; E; H	Gumbel para máximos
(18)	60220000	A; E; H	Log normal
(19)	60145000	A; E; H	Gumbel para máximos
(20)	40025000	A; H	Gumbel para máximos
(21)	40040000	-	Log normal
(22)	40050000	A	Log normal
(23)	40100000	A; E; H	Log normal
(24)	40150000	E	Gumbel para mínimos
(25)	40330000	E; H	Normal
(26)	40400000	H	Gumbel para máximos
(27)	40680000	H	Gumbel para máximos
(28)	40800001	-	Normal
(29)	40930000	A	Normal
(30)	41250000	H	Gumbel para máximos
(31)	41300000	A; H	Gumbel para máximos
(32)	41340000	E; H	Gumbel para máximos
(33)	41380000	H	Normal
(34)	41818000	H	Gumbel para máximos
(35)	41990000	A	Gumbel para máximos
(36)	42210000	E; H	Gumbel para mínimos
(37)	42395000	A; E	Gumbel para máximos
(38)	42600000	E; H	Log normal
(39)	42690001	-	Normal
(40)	43200000	E; H	Gumbel para mínimos
(41)	44200000	-	Normal
(42)	44500000	-	Normal
(43)	45131000	-	Normal
(44)	45298000	-	Normal
(45)	46150000	E; H	Normal
(46)	46360000	-	Normal
(47)	46550000	-	Normal
(48)	46650000	H	Normal
(49)	46675000	-	Normal
(50)	46790000	E; H	Log normal
(51)	46830000	E	Normal
(52)	46902000	H	Log normal
(53)	48020000	-	Log normal
(54)	48290000	E; H	Log normal

A: aleatoriedade; E: estacionariedade; I: independência; H: homogeneidade.
 Fonte: elaboração própria.

entre 15 e 25% – classificação satisfatório –, equivalente a pouco mais de 11%. Ainda assim, a classificação muito bom foi maioria, com cerca de 70% dos casos.

Para a série de 10%, mais de 57% dos casos obtiveram a classificação muito bom, com tamanho médio da série de 7,42 anos. Se aumentar o tamanho médio para 15 anos, essa porcentagem sobe para mais de 68%. Mais uma vez, as normais climatológicas tiveram comportamento muito similar ao das séries de 10%, ou seja, séries de 30 anos de dados com resultados similares aos das séries de sete anos de dados.

Tal diversidade no resultado talvez seja explicada pela forma como a $Q_{90\%}$ foi calculada, já que ela foi obtida de um único ano representativo dos demais anos da estação, contudo a maior parte dos resultados ainda assim foi classificada como boa e muito boa.

Não foi possível verificar relação entre as estações com maior diferença percentual e a rejeição dos testes não paramétricos.

Apesar de se tratar de metodologias diferentes, resultados finais similares aos do presente estudo foram obtidos por Huang e Bardossy (2020) e Medina e Muñoz (2020), nas áreas de estudo do Médio Atlântico dos Estados Unidos e do Rio Allipén, Chile, respectivamente. Ambas as pesquisas utilizaram como base do trabalho o modelo hidrológico Hydrologiska Byråns Vattenbalansavdelning (HBV) para análise de resultados em localidades com dados hidrológicos limitados, mas com nuances na metodologia, além de objetivos diferentes: o primeiro, para realizar estimativas de eventos extremos de máximas e mínimas

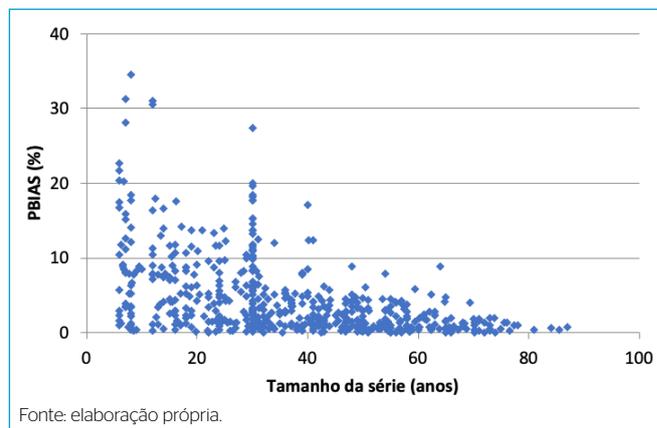


Figura 4 - Diferenças percentuais para os variados tamanhos de séries parciais analisados: parâmetro média, série de Q_7 .

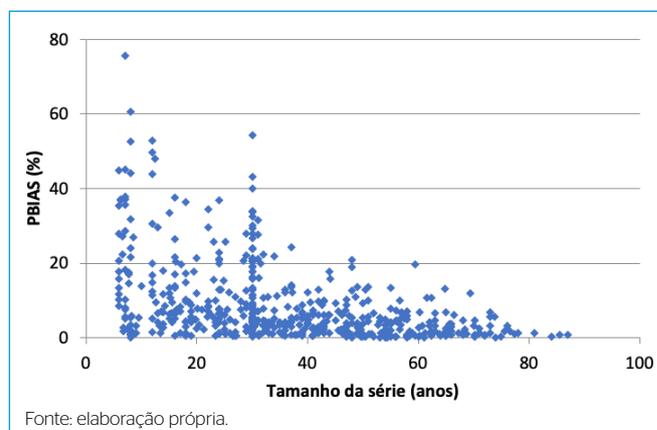


Figura 5 - Diferenças percentuais para os vários tamanhos de séries parciais analisados: vazão Q_{90} .

anuais; e o segundo, para verificar o impacto da qualidade e da quantidade dos dados na calibração do modelo, além da regionalização dos parâmetros calibrados em duas décadas distintas. Os dois utilizaram dados de precipitação média diária, temperatura e evapotranspiração mensal como entrada no modelo HBV. Medina e Muñoz (2020) concluíram que, para se obter resultados com incertezas controladas, é recomendado o tamanho mínimo de 16 anos de dados no uso de previsões por intermédio do modelo HBV. Huang e Bardossy (2020), por sua vez, verificaram que, em geral, oito anos de dados são suficientes para se estimar os parâmetros do modelo de maneira confiável, embora tenham observado que o aumento de dados na calibração leve à melhor *performance* na regionalização dos parâmetros. Nesse caso, utilizar séries de 10 anos de dados na calibração se mostra bastante estável para esse fim de uso para regionalização.

CONCLUSÕES

Este trabalho teve como objetivo verificar o tamanho ideal de séries históricas fluviométricas nos estudos hidrológicos, por meio da análise da diferença percentual e do teste *t* de Student entre os valores médios obtidos com as séries de Q_7 e as vazões mínimas anuais completas (ou seja, com todos os anos disponíveis na série histórica) e específicas (com tamanhos preestabelecidos de 10, 20, 30, 40, 50, 60, 70, 80 e 90% do tamanho completo além da normal climatológica). Além das médias, fez-se essa mesma análise com os valores de $Q_{7,10}$ e $Q_{90\%}$ calculados. O estudo foi realizado com o total de 54

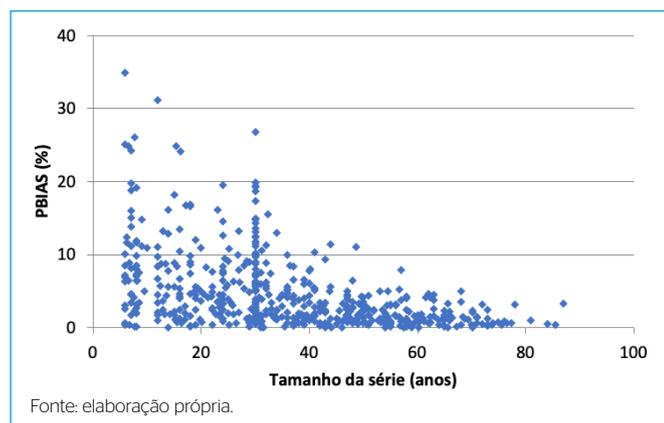


Figura 6 - Diferenças percentuais para os vários tamanhos de séries parciais analisados: parâmetro média, série de vazões mínimas anuais.

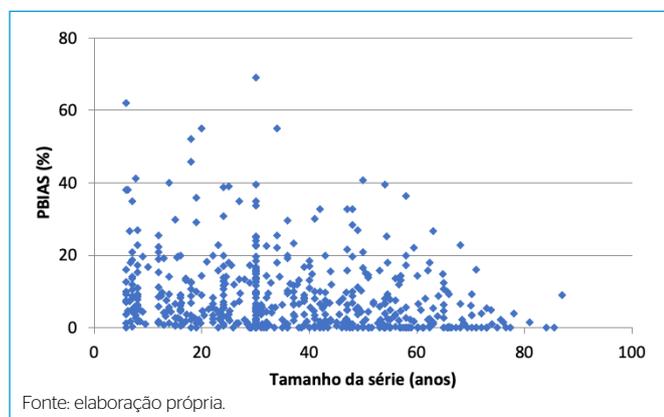


Figura 7 - Diferenças percentuais para os vários tamanhos de séries parciais analisados: vazão $Q_{90\%}$.

estações fluviométricas pertencentes a duas regiões hidrográficas: RHP (19 estações) e RHSF (35 estações).

Aplicaram-se os testes não paramétricos de aleatoriedade, estacionariedade, independência e homogeneidade às estações, verificando que 30 das 54 estações aceitaram apenas um ou nenhum teste, enquanto apenas duas estações aceitaram todos os testes não paramétricos. Isso mostra um baixo número de estações com resultados positivos aos testes, seja em razão da influência antrópica na região hidrográfica, seja pela forma como os dados são coletados, seja por causa da periodicidade com que são feitos, seja até mesmo pelas variações de cheias e secas ao longo dos anos, com diferentes intensidades.

As distribuições de probabilidade mais bem ajustadas às estações selecionadas para as séries de vazões mínimas anuais foram normal e log normal, com 22 e 15 estações, respectivamente. Contrariando o esperado, apenas três estações se ajustaram melhor com Gumbel para mínimos.

Em relação aos valores médios da série de Q_7 , houve predominância da classificação muito bom pela diferença percentual, ocorrendo em mais de 87% dos casos analisados. Tal classificação foi obtida satisfatoriamente de séries com tamanho médio próximo de oito anos de dados. O teste *t* de Student acusou diferença estatística em apenas 14 casos.

Os valores de $Q_{7,10}$ foram classificados como insatisfatórios em 50 casos, correspondendo a pouco mais de 9% dos casos analisados. Desses 50 casos, 18 foram em séries com tamanho médio de sete anos, e 12, nas normais climatológicas.

A análise das médias da série de vazões mínimas anuais mostrou que apenas cinco casos foram insatisfatórios, 22 satisfatórios, 40 casos bons e 470 muito bons. A classificação muito bom correspondeu a mais de 87% dos casos, ocorrendo com frequência a partir de um tamanho médio de sete anos de dados. O teste *t* de Student mostrou diferença estatística em apenas 14 casos.

As vazões $Q_{90\%}$ foram classificadas como insatisfatórias e satisfatórias em 42 e 61 situações, respectivamente, ao longo de todas as séries específicas analisadas, representando juntas pouco mais de 19% de todos os casos. Em 59 casos foram classificadas como bom e 375 como muito bom, representando em conjunto quase 81% do total e ocorrendo em séries de tamanho médio de oito anos.

Assim, para as estações fluviométricas aqui analisadas, utilizar séries históricas com tamanhos a partir de oito anos de dados possibilitou obter resultados semelhantes, ou estatisticamente iguais, aos obtidos por séries longas (no presente estudo, entre 59 e 97 anos de dados). Dessa maneira, com base nos resultados obtidos por meio das 54 estações fluviométricas pertencentes às RHP e RHSF, recomenda-se a utilização de séries históricas com no mínimo oito anos de dados para as análises de vazões mínimas anuais, Q_7 , $Q_{7,10}$ e $Q_{90\%}$.

Outra observação relevante é em relação às normais climatológicas, as quais mostraram resultados muito semelhantes aos das séries específicas de 10%.

Ressalta-se a dificuldade de se encontrar séries históricas de longo período para a realização do estudo. Mais de 1.700 estações fluviométricas foram baixadas para selecionar as 54 que atendiam ao pré-requisito do número de anos de dados. Além disso, o assunto ainda é escasso, com um número reduzido de trabalhos similares para confrontar resultados.

Uma variação do presente trabalho, já realizada pelos autores deste artigo e também em preparação para publicação, foi a metodologia aplicada para as mesmas 54 estações fluviométricas considerando vazões de máximas e médias anuais.

Muito ainda se pode fazer para compreender melhor a temática. Algumas sugestões de variações a serem consideradas em futuros trabalhos:

- Aplicar a mesma metodologia em outras estações fluviométricas pertencentes às RHP e RHSF não consideradas aqui;
- Aplicar a metodologia para as demais regiões hidrográficas do Brasil;
- Analisar períodos específicos em vez de anos aleatórios;
- Considerar os tamanhos das séries parciais fixos, independentemente do tamanho da série original;
- Focar nas séries parciais de curto período, desconsiderando as análises para as séries de longo período;
- Aprofundar um estudo com relação às normais climatológicas, visto que as normais padrão, com 30 anos de dados, apresentaram resultados similares aos das séries parciais de tamanhos próximos a sete anos;

- Realizar um estudo semelhante com dados pluviométricos, considerando séries de dados diários, mensais e anuais, por exemplo.

CONTRIBUIÇÕES DOS AUTORES

Oliveira, T. C. C.: Conceituação, Curadoria de Dados, Análise Formal, Obtenção de Financiamento, Investigação, Metodologia, Validação, Visualização, Escrita — Primeira Redação, Escrita — Revisão e Edição. Gonçalves, F. A.: Conceituação, Análise Formal, Investigação, Metodologia, Administração do Projeto, Supervisão, Validação, Visualização, Escrita — Primeira Redação, Escrita — Revisão e Edição.

REFERÊNCIAS

- AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). *As doze regiões hidrográficas brasileiras*. Brasil: ANA. Disponível em: <https://www.ana.gov.br/as-12-regioes-hidrograficas-brasileiras/links-12-regioes>. Acesso em: 4 abr. 2020a.
- AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). *Inventário estações pluviométricas*. Brasília: ANA, 2009. 196 p.
- AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). *Séries históricas de estações*. Brasil: ANA. Disponível em: <http://www.snirh.gov.br/hidroweb/apresentacao>. Acesso em: 6 abr. 2020b.
- BARBAROTTO JUNIOR, J. L. *Análise da disponibilidade hídrica da bacia do Rio Jundiá por meio de simulações hidrológicas de cenários prováveis*. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Estadual de Campinas, Campinas, 2014. Disponível em: http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/258741/1/BarbarottoJunior_JorgeLuiz_M.pdf. Acesso em: 13 jul. 2021.
- BETTERLE, A.; SCHIRMER, M.; BOTTER, G. Flow dynamics at the continental scale: Streamflow correlation and hydrological similarity. *Hydrological Processes*, v. 33, n. 4, p. 627-646, 2019. <https://doi.org/10.1002/hyp.13350>
- DARBANDSARI, P.; COULIBALY, P. Inter-comparison of lumped hydrological models in data-scarce watersheds using different precipitation forcing data sets: Case study of Northern Ontario, Canada. *Journal of Hydrology: Regional Studies*, v. 31, 100730, 2020. <https://doi.org/10.1016/j.ejrh.2020.100730>
- DETZEL, D. H. M. *Modelagem de séries hidrológicas: uma abordagem de múltiplas escalas temporais*. Tese (Doutorado em Engenharia de Recursos Hídricos e Ambiental) – Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Ambiental, Universidade Federal do Paraná, Curitiba, 2015. Disponível em: <https://acervodigital.ufpr.br/bitstream/handle/1884/41305/R%20-%20T%20-%20DANIEL%20HENRIQUE%20MARCOS%20DETZEL.pdf?sequence=2&isAllowed=y>. Acesso em: 13 jul. 2021.
- GENZ, F.; LESSA, G. C.; CIRANO, M. Vazão mínima para estuários: um estudo de caso no Rio Paraguaçu/BA. *Revista Brasileira de Recursos Hídricos*, v. 13, n. 3, p. 73-82, 2008. Disponível em: <http://eventos.abrh.org.br/xiisrhn/anais/papers/PAP017690.pdf>. Acesso em: 13 jul. 2021.
- GOMES, E. P.; BLANCO, C. J. C. Daily rainfall estimates considering seasonality from a MODWT-ANN hybrid model. *Journal of Hydrology and Hydromechanics*, v. 69, n. 1, p. 13-28, 2021. <https://doi.org/10.2478/johh-2020-0043>
- GRANEMANN, A. R. B.; MINE, M. R. M. Distribuições de probabilidade para vazões mínimas: uma revisão bibliográfica. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 21., 2015, Brasília. *Anais [...]*. Brasília, 2015.
- HUANG, Y.; BARDOSSY, A. Impacts of data quantity and quality on model calibration: implications for model parameterization in data-scarce catchments. *MDPI Water*, Suíça, v. 12, n. 9, 2352, 2020. <https://doi.org/10.3390/w12092352>
- INSTITUTO NACIONAL DE METEOROLOGIA (INMET). *Normais climatológicas do Brasil*. Brasil: INMET. Disponível em: <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>. Acesso em: 26 nov. 2020.
- LUIZ, S. F. *Regionalização hidrológica de vazões mínimas por meio do método dos mínimos quadrados generalizados aplicada à bacia do Alto São Francisco*. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) – Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte, 2013. Disponível em: <http://www.smarh.eng.ufmg.br/defesas/1053M.PDF>. Acesso em: 14 jul. 2021.
- MEDINA, Y.; MUÑOZ, E. Estimation of annual maximum and minimum flow trends in a data-scarce basin. Case study of the Allipén river watershed, Chile. *MDPI Water*, Suíça, v. 12, n. 1, 162, jan. 2020. <https://doi.org/10.3390/w12010162>
- MELLO, C. R.; SILVA, A. M. *Hidrologia: princípios e aplicações em sistemas agrícolas*. Lavras: Editora UFLA, 2013. 455 p.
- MELO, L. S. *Distribuições e tendências de vazões na bacia hidrográfica do Rio das Velhas*. Dissertação (Mestrado em Ciências Agrárias) – Programa de Pós-Graduação em Ciências Agrárias, Universidade Federal de São João del-Rei, Sete Lagoas, 2016. Disponível em: https://www.ufsj.edu.br/portal2-repositorio/File/ppgca/Dissertacao%20Larissa%20Melo_29_07_2016.pdf. Acesso em: 27 nov. 2020.
- OLIVEIRA, J. R.; PRUSKI, F. F.; NUNES, A. A. Otimização do aproveitamento da disponibilidade de águas superficiais na Bacia do Ribeirão entre Ribeiros. *Revista Brasileira de Recursos Hídricos*, v. 18, n. 4, p. 157-172, 2013. <http://doi.org/10.21168/rbrhv18n4.p157-172>

ROCHA, P. C.; SANTOS, A. A dos. Hydrological analysis in water basins. *Revista Mercator*, v. 17, e17025, 2018. <https://doi.org/10.4215/rm2018.e17025>

RUBATINO, F. S.; NASCIMENTO, P. C.; ROENICK, R. G. A.; BATISTA, S. H. *Análise de homogeneidade, aleatoriedade, independência e estacionariedade das séries de dados de cota e vazão de estações fluviométricas da bacia do Rio Doce*. Relatório Final (Bacharelado em Ciência e Tecnologia) – Universidade Federal de São João Del Rei, Ouro Branco, 2017. Disponível em: <https://ufsj.edu.br/portal2-repositorio/File/labhidromec/Analise%20de%20homogeneidade,%20aleatoriedade,%20independencia%20e%20estacionariedade%20das%20series%20de%20dados%20de%20cota%20e%20vazao%20de%20estacoes%20fluviometricas%20da%20bacia%20do%20Rio%20Doce.pdf>. Acesso em: 26 mar. 2021.

SOUSA, W. dos S.; SOUSA, F. de A. S. de. Rede neural artificial aplicada à previsão de vazão da Bacia Hidrográfica do Rio Piancó. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 14, n. 2, p. 173-180, 2010. <https://doi.org/10.1590/S1415-43662010000200008>

ULIANA, E. M.; MARTINS, C. A. S.; SILVA, J. G. F.; REIS, E. F. Análise estatística para determinação da Q_{70} na Bacia do Rio São Mateus-ES. *In: ENCONTRO LATINO-AMERICANO DE INICIAÇÃO CIENTÍFICA*, 15; ENCONTRO LATINO AMERICANO DE PÓS-GRADUAÇÃO, 11., 2011. *Anais* [...]. Universidade do Vale do Paraíba, 2011.

ULIANA, E. M.; SILVA, D. D.; ULIANA, E. M.; RODRIGUES, B. S.; CORRÊDO, L. P. Análise de tendência em séries históricas de vazão e precipitação: uso de teste estatístico não paramétrico. *Revista Ambiente & Água*, v. 10, n. 1, p. 82-88, 2015. <https://doi.org/10.4136/ambi-agua.1427>

VALLE JUNIOR, L. C. G.; RODRIGUES, D. B. B.; OLIVEIRA, P. T. S. Initial abstraction ratio and Curve Number estimation using rainfall and runoff data from a tropical watershed. *Revista Brasileira de Recursos Hídricos*, v. 24, e5, 2019. <https://doi.org/10.1590/2318-0331.241920170199>

WORLD METEOROLOGICAL ORGANIZATION (WMO). *Guidelines on the calculation of climate normals*: nº 1.203. Genebra: Secretariat of the World Meteorological Organization, 2017. 29 p.

WORLD METEOROLOGICAL ORGANIZATION (WMO). *Technical Regulations*: nº 49. Genebra: Secretariat of the World Meteorological Organization, 2019. 71 p.

WORLD METEOROLOGICAL ORGANIZATION (WMO). *The role of climatological normals in a changing climate*: nº 1.377. Genebra: Secretariat of the World Meteorological Organization, 2007. 130 p.

