

Deteção de Tendência nas Séries de Afluência das Estações Fluviométricas Ponte do Piquiri (PR) e UHE Santo Antônio do Jari (AP)

Resumo

A identificação de comportamento não estacionário em séries hidrológicas tem vital importância dentro da Hidrologia Estocástica, muito devido às restrições impostas para a utilização dos modelos do tipo ARIMA. Este estudo busca detectar tendências nas séries de vazão média anual das estações Ponte do Piquiri (PR) e UHE Santo Antônio do Jari (AP), que se localizam em áreas de diferentes usos do solo e comportamento hidrológico no país. Para tal, são aplicados os testes de Mann-Kendall, Pettitt e F de Snedecor para variância, com o requisito de normalidade dos dados sendo aferido pelo teste de Lilliefors, todos a um nível de significância de 5%. Como resultado, obteve-se que a hipótese nula foi aceita para ambas as estações no teste de Mann-Kendall e Pettitt. A única rejeição da hipótese de estacionariedade se deu no teste F de Snedecor para UHE Santo Antônio do Jari.

1 Introdução

Desde a incorporação do monitoramento hidrológico como serviço público federal entre 1900 e 1920 (Agência Nacional de Águas, 2021) [1], a população urbana teve um grande aumento. Impulsionadas pelo êxodo rural, as cidades cresceram, mudando o uso e a cobertura de muitas regiões do país. Tal fato provocou uma alteração no escoamento superficial das bacias, potencialmente modificando a sua resposta frente a diferentes estímulos.

Quando se olha somente para a afluência, contudo, é usual que tal alteração seja refletida como um comportamento não estacionário. Este estudo procura detectar possíveis tendências na série de vazão média anual de duas estações fluviométricas brasileiras, Ponte do Piquiri (PR) e UHE Santo Antônio do Jari (AP). Três testes, de naturezas díspares, são aplicados com esse fim: Mann-Kendall, Pettitt e F de Snedecor para variância.

O trabalho está subdividido em 5 seções distintas, sendo esta introdução a primeira delas. A segunda seção apresenta uma breve revisão bibliográfica sobre o tema. A seção 3 identifica as estações fluviométricas utilizadas e descreve os testes de hipótese empregados na análise. Por sua vez, a seção 4 apresenta os resultados obtidos, enquanto

a seção 5 discute o seu significado e eventuais fatores de influência, além de sugerir aperfeiçoamentos ao trabalho executado.

2 Revisão bibliográfica

A seguir, apresenta-se a abordagem de dois artigos acerca da detecção de tendência em séries hidrológicas.

2.1 (Hodgkins et al) Effects of climate, regulation, and urbanization on historical flood trends in the United States

Motivados pelo impacto econômico e social que as cheias possuem, e levando em consideração o processo de dimensionamento de estruturas hidráulicas no país, os autores [3] se propõem a estudar as tendências nas séries de vazão máxima anual a nível nacional.

O estudo apresenta um refinamento ao conhecimento comum de que a ação antrópica modifica o comportamento da bacia hidrográfica, pois nem todas as alterações se manifestam de forma igual na série de afluências. Algumas alterações humanas

aumentam as vazões de pico, como a urbanização e sua consequente impermeabilização da superfície. Contudo, outras estruturas também transformam a bacia, mas podem abater os valores observados, como os reservatórios.

A desconsideração de tais particularidades pode levar a conclusões errôneas. Uma tendência de aumento das cheias, por exemplo, pode não se estender para o futuro caso as bacias tenham findado o seu desenvolvimento antrópico. Da mesma forma, a influência de outros fatores pode ser mascarada quando mistos com mudanças no uso do solo da bacia.

Para aferir o impacto das alterações por meio de tendências, o estudo considerou 2681 séries de 50 anos, 1465 de 75 anos e 257 de 100 anos de duração, distribuídas ao longo dos Estados Unidos da América. Séries de vazão máxima anual e média diária foram utilizadas, após passarem por um critério de completude definido pelo autor. A última foi empregue na aplicação do método *peaks-over-threshold* (POT), para eventos que ocorrem duas vezes por ano e uma vez a cada 5 anos.

Os dados de urbanização das bacias foram retirados do banco de dados *geospatial attributes of gages for evaluating streamflow* (GAGES II), categorizando as bacias entre aquelas com capacidade de regularização, urbanizadas e inalteradas. Os casos onde as características não se enquadraram completamente em nenhum critério estabelecido foram estudadas em um grupo a parte.

A magnitude das vazões máximas anuais foi estudada com o método não paramétrico da curva-tura de Sen, enquanto a existência e significância de tendências foi aferida pelo teste de Mann-Kendall. Sendo o método sensível à existência de persistência de curto/longo termo e da independência dos dados, todas as séries tiveram 3 aplicações do teste de Mann-Kendall (original, e com modificações para séries com persistência).

Como resultado, foi concluído que as mudanças na vazão de pico dependem fortemente do tipo de bacia e região geográfica em que está inserida. De forma geral, bacias não modificadas tiveram poucas tendências significativas, tanto para aumento quanto para diminuição das vazões. Por outro lado, as bacias com regularização tiveram diminuição de cheias, e as urbanizadas aumento, assim como conjecturado inicialmente.

A localização dos postos também foi significativa

nessa classificação: as bacias no quadrante nordeste dos EUA apresentaram aumentos significativos, enquanto o sudoeste exibiu muitas bacias com abatimentos. Contudo, tal fator foi considerado menos significativo que o uso e cobertura da bacia, visto que esta classificação é dependente da primeira. A intensidade das precipitações também foi elencada como fator preponderante na aparição de tendências nas bacias, pois por mais que a capacidade de armazenamento de algumas bacias tenha aumentado, as chuvas também apresentaram acréscimos significativos.

Tais resultados foram verificados para as séries anuais, mas podem não se replicar em outros casos. Particularmente, os autores citam que a capacidade de regularização não acumula significância para eventos muito extremos, que excedem o armazenamento da bacia. De tal forma, ressaltam que as considerações para o cálculo da vazão de projeto não devem ser mecânicas, pois precisam levar em consideração uma série de fatores intervenientes ao processo, e que os resultados obtidos permitem afirmar com clareza que as medições passadas não foram representativas para o presente, em muitos dos casos estudados.

2.2 (Sharma et al) Modified signal-to-noise ratio method for early detection of climate change

O estudo das tendências em séries no contexto das mudanças climáticas possui o papel de identificar alterações significativas no comportamento das variáveis frente a variabilidade natural que apresentam. Já é de conhecimento na literatura que o desempenho de muitas metodologias não paramétricas é condicionado ao comprimento da série e a localização do ponto de quebra.

A relação sinal-ruído (SNR) detecta tendências através da comparação da variação de causa antrópica da série com um ruído que representa a oscilação natural da variável. Assim, a proposta do trabalho [12] é modificar o método SNR de forma a otimizar a aplicação do mesmo, e corrigir outros aspectos que não se verificam nas aplicações. O desempenho do SNR modificado, chamado pelos autores de *noise-based change point* (NBCP), é então comparado ao próprio SNR, e ao teste de Pettitt, que também é vastamente utilizado na área.

A formulação do SNR se baseia na comparação

entre a razão entre o sinal (série) e o seu ruído, acusando como ponto de quebra o primeiro instante onde tal razão exceda um determinado nível de significância estipulado. O NBCP constrói sua metodologia de forma a evitar dois pontos problemáticos do SNR: a hipótese de que o ruído apresenta distribuição normal, e o fato de o SNR acumular peso para os anos iniciais da série, sem tendência, para a detecção da mesma. Tal fato afeta a sensibilidade do parâmetro, impactando a capacidade de identificação das mudanças.

Para a avaliação da performance do NBCP foram geradas e/ou modificadas 500 séries sintéticas de comprimento 2000 para 6 diferentes cenários: o caso geral (1), sem tendência (2), baixa tendência (3), alta tendência (4), diferentes localizações do ponto de quebra dentro da série (5), e diferentes magnitudes de coeficiente de variação (6). O NBCP se provou mais eficiente e preciso para todos os cenários simulados, a um nível de significância de 5% e 10%. Também foi realizado um estudo de caso real com a série de precipitação mensal de 13 localizações na Índia, com o NBCP mais uma vez apresentando um resultado superior aos demais métodos.

Assim como conjecturado inicialmente, os resultados permitiram aos autores concluir que a abordagem utilizada pelo SNR não é otimizada, sendo as tendências identificadas de forma tardia pelo modelo. Tal característica é partilhada com o NBCP, principalmente em casos onde a variabilidade natural da série é elevada, contudo a versão modificada apresentou detecção antecipada quando comparado ao original.

3 Materiais e métodos

Os ensaios desse artigo foram realizados com a versão 3.8 da linguagem Python. A manipulação de dados e plotagem de gráficos se deu com as bibliotecas Pandas [14], NumPy [2] e Matplotlib [4]. Outrossim, as bibliotecas statsmodels [11], pyMannKendall [6], pyhomogeneity [5] e SciPy [15] foram usadas para a análise propriamente dita.

3.1 Dados utilizados

O estudo foi realizado com dados de vazão afluente anual média, cujas estações fluviométricas

são identificadas na Tabela 1. Nenhuma das séries apresentou falhas de medição, sendo UHE Santo Antônio do Jari totalmente consistida e Ponte do Piquiri parcialmente consistida (até dez/2014). A Figura 1 apresenta o gráfico das séries em análise.

A estação UHE Santo Antônio do Jari encontra-se dentro da Amazônia Legal, com uso e ocupação do solo sendo majoritariamente florestal. Por outro lado, a estação Ponte do Piquiri é localizada no terceiro planalto paranaense, sendo parte da Bacia do Paraná, cuja área de drenagem é muito superior do que a que contribui diretamente para o posto em questão. A bacia como um todo é dominada pela agropecuária, sendo possível aferir, particularmente para a área contribuinte da estação, o uso mais intenso como agricultura. (Projeto MapBio-mas, 2021) [9].

Enquanto a série de Ponte do Piquiri foi obtida convencionalmente por meio da Hidroweb [1], cabe ressaltar que a série de UHE Santo Antônio do Jari faz parte do conjunto de séries naturais de vazão do Operador Nacional do Sistema Elétrico (ONS) [13], apresentadas desde o ano de 1931, e que desconsideram a existência de barramento no local.

3.2 Testes de Hipótese

A seguir, apresenta-se uma breve descrição das premissas requeridas por cada teste de hipótese. Uma exposição mais completa sobre o tema pode ser vista em Naghettini (2007) [8], Pettitt (1979) [10] e Lilliefors (1967) [7].

3.2.1 Teste de Mann-Kendall

É um teste não paramétrico, no qual a série não deve apresentar correlação significativa, com o risco de alteração do p -valor em caso contrário. Como resultado tem-se o p -valor e a estatística do teste. Foi aplicado por meio da função `original_test` em `pymannkendall v1.4.2`.

3.2.2 Teste de Pettitt para duas amostras

O teste não paramétrico de Pettitt busca um ponto de mudança de comportamento dentro da série analisada, subdividindo a mesma em duas amostras distintas e averiguando a homogeneidade entre elas. Como resultado tem-se o p -valor, o provável

Tabela 1 – Estações fluviométricas estudadas

Estação	UHE Santo Antônio do Jari	Ponte do Piquiri
Bacia	1 - Rio Amazonas	6 - Rio Paraná
Rio	Rio Jari	Rio Piquiri
Unidade da Federação	AP	PR
Município	Laranjal do Jari	Ubiratã
Fonte dos dados	ONS	Hidroweb
Operador	ONS	IAT-PR
Código	290	64795000
Latitude	-0,65	-24,56
Longitude	-52,53	-53,13
Área de drenagem (km ²)	52643,00	11200,00
Início da série	Jan/1931	Jan/1971
Fim da série	Dez/2018	Dez/2019
Nível de consistência	Consistido	Consistido até 2014
Quantidade de medições	89	48
Média (m ³ /s)	1092	314
Mediana (m ³ /s)	1076	304
Amplitude (m ³ /s)	1566	489
Variância (m ⁶ /s ²)	84295	10057
Desvio padrão (m ³ /s)	290	100
Coeficiente de variação	27%	32%
Coeficiente de assimetria	0,42	0,67

Fonte: Adaptado de ANA (2021) e ONS (2021).

ponto de quebra e a estatística do teste. Foi executado por meio da função `pettitt.test` em `pyhomogeneity` v1.1.

models v0.14.0.

3.2.3 Teste de aderência de Lilliefors

É uma modificação do teste de Kolmogorov-Smirnov para o caso onde a média e a variância populacional são desconhecidas.

Foi proposto por Hubert Lilliefors em 1967 [7], de forma que dada uma amostra da variável aleatória X de tamanho N , calcula-se:

$$D = \max_{x \in X} |F^*(x) - S_N(x)|,$$

sendo $S_N(X)$ a distribuição empírica de probabilidades e $F^*(X)$ a função de densidade acumulada da distribuição para média e desvio padrão iguais aos da amostra estudada. A distância máxima D é então comparada aos valores críticos propostos por Lilliefors, estes obtidos por meio de simulações de Monte Carlo. Tem como característica ser mais conservativo que o teste clássico de Kolmogorov-Smirnov.

Sua utilização nesse trabalho se deu por meio da função `stats.diagnostic.lilliefors` em `stats-`

3.2.4 Teste F de Snedecor para variâncias

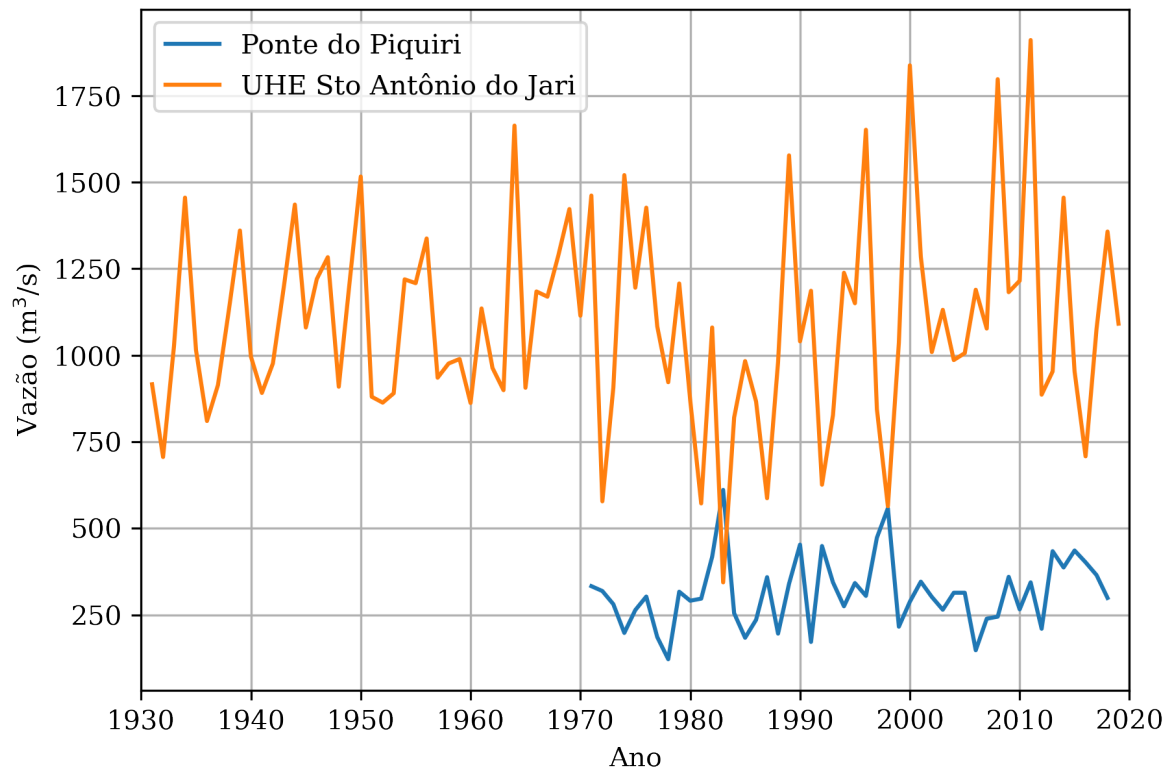
Ao contrário dos dois primeiros testes, requer normalidade dos dados para ser aplicado. Compara a variância de duas subamostras retiradas da série original. Sendo um teste bicaudal, possui um limite inferior e superior para aceitação da hipótese nula, sendo as estatísticas retiradas da distribuição F de Fischer e Snedecor.

4 Resultados

A função de autocorrelação (FAC) de ambas as séries estudadas é apresentada na Figura 2. O lag máximo apresentado é de 20 anos, sendo a FAC determinada pela função `api.graphics.tsa.plot.acf` da biblioteca `statsmodels` v0.14.0. A implementação utiliza a transformada rápida de Fourier para o cálculo das correlações, sendo o intervalo de confiança computado por meio da fórmula de Bartlett.

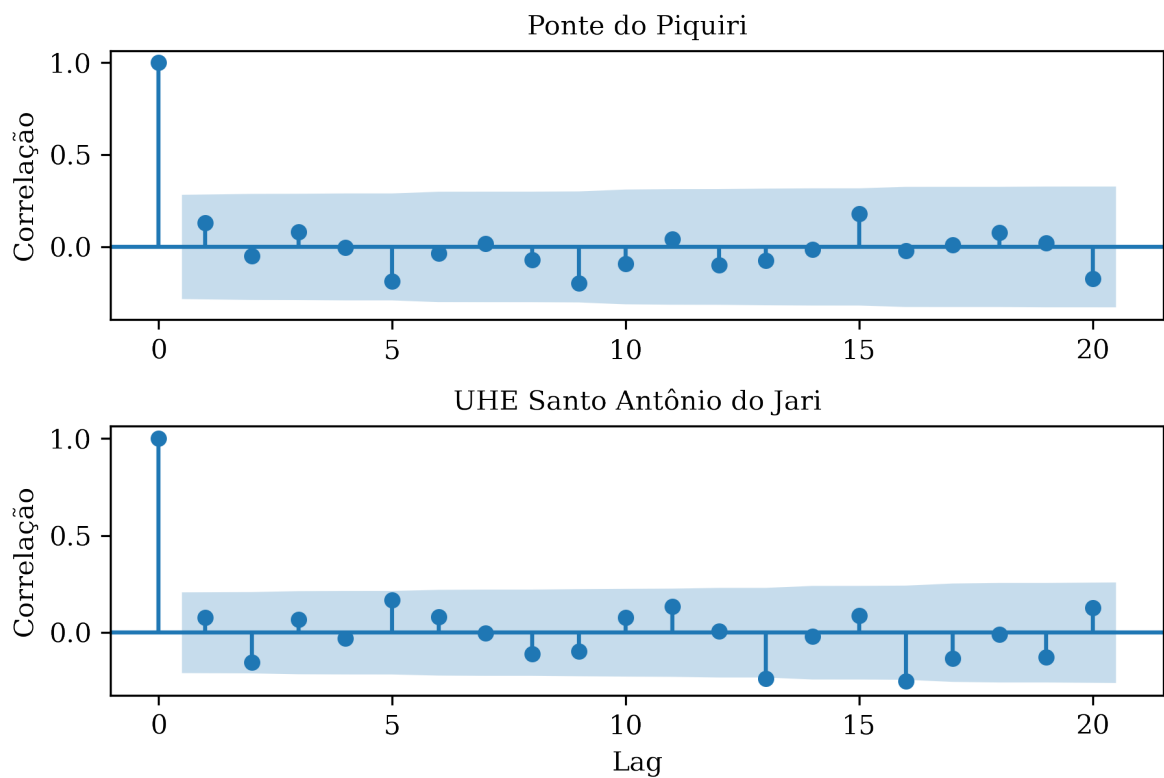
Como é possível observar no correlograma obtido,

Figura 1 – Vazão afluyente anual média



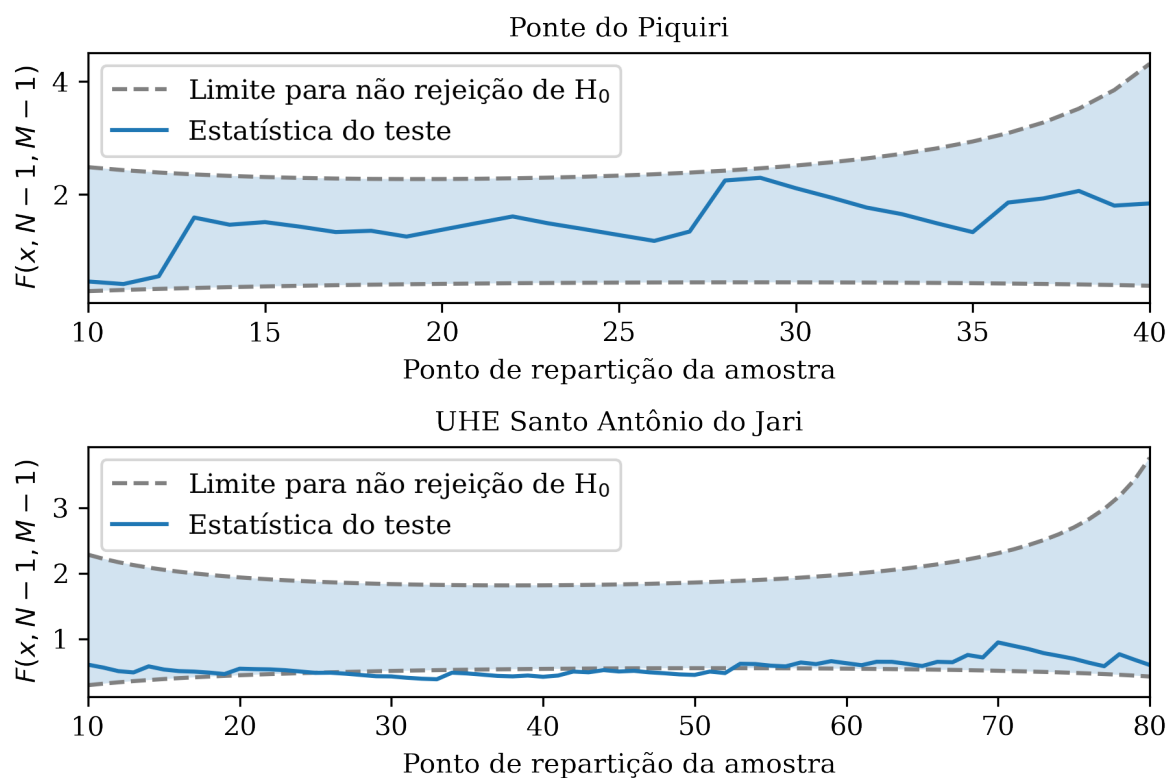
Fonte: A Autora (2021).

Figura 2 – Correlograma das séries de afluência anual média



Fonte: A Autora (2021).

Figura 3 – Valores de F de Snedecor



Fonte: A Autora (2021).

Tabela 2 – Estatísticas por teste de hipótese aplicado

Estação	Mann-Kendall		Pettitt			Lilliefors	
	p-valor	Score	p-valor	Pt. de quebra	U	p-valor	KS
Ponte do Piquiri	0,152	162	0,307	18	172	0,103	230
UHE Santo Antônio do Jari	0,632	136	0,667	69	328	0,094	66

Fonte: A Autora (2021).

ambas as séries não apresentam correlações significativas, elegendo-se para o teste de Mann-Kendall.

A Tabela 2 apresenta os resultados do teste de Mann-Kendall, Pettitt e Lilliefors.

Constata-se que ambas as estações foram aceitas pelos testes aplicados, ou seja, não apresentaram tendências significativas no estudo. Também pode-se notar que os p -valores obtidos foram significativamente superiores ao nível de confiança estipulado de $\alpha = 5\%$.

Por sua vez, o teste de Lilliefors indicou aceitação da hipótese de normalidade dos dados. Notavelmente, observa-se que o p -valor de UHE Santo Antônio do Jari foi o que esteve mais próximo do limiar estipulado, mas ainda assim significativamente distante do valor limite.

Desta forma, aplicou-se o teste paramétrico de Snedecor para variâncias, cujo resultado é mostrado na Figura 3. Como a metodologia de aplicação do teste prevê a divisão da amostra em duas subamostras, este trabalho optou por realizar o teste para todas as subamostras sequenciadas completas possíveis dentro do espaço amostral. Desta forma, o gráfico apresenta o resultado do teste, limitado inferior e superiormente pelos valores retirados da distribuição, para todos os pontos de divisão possíveis. Os valores de F para subamostras de tamanho inferior a 10 medições foram omitidos, para maior clareza na representação.

É possível notar que a série de Ponte do Piquiri se mantém dentro da zona de aceitação do teste, enquanto, para valores centrais, isto é, para subamostras de tamanho balanceado, o teste de F de Snedecor acusa a existência de tendência para variâncias. Observa-se que não existe uma relação clara com tais valores de divisão com o provável ponto de quebra obtido com a aplicação do teste de Pettitt.

5 Discussões e conclusão

Os testes aplicados permitiram concluir que as séries não apresentam tendências significativas quanto aos valores médios de vazão. No entanto, com relação à variância das séries, o teste de Snedecor apresenta argumentos contra a hipótese nula. Tal fato pode ser empiricamente confirmado com uma inspeção visual da Figura 1, onde nota-se que

os eventos extremos (cheia e estiagem), a partir da década de 70, tornam-se mais intensos, indicando uma maior variabilidade após esse período, assim como detectado pelo teste.

Contudo, tal hipótese carece de estudos adicionais para uma conclusão mais efetiva dos resultados. Uma sugestão é a aplicação do mesmo teste para variâncias, mas na série logarítmica, visto que trabalhos anteriores sugerem que tal posto é melhor representado por uma distribuição log-normal, embora tenha sido aceito como normalmente distribuído. A própria aceitação de tal hipótese, por sua vez, pode ser melhor investigada, dado que é conhecido que o teste de Kolmogorov-Smirnov seja muito permissivo em seus resultados.

Referências

- [1] Agência Nacional de Águas (ANA). *Hidroweb*. URL: <https://www.snirh.gov.br/hidroweb/apresentacao>.
- [2] Charles R. Harris et al. “Array programming with NumPy”. Em: *Nature* 585.7825 (set. de 2020), pp. 357–362. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2). URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- [3] G. A. Hodgkins et al. “Effects of climate, regulation, and urbanization on historical flood trends in the United States”. Em: *Journal of Hydrology* 573 (2019), pp. 697–709. ISSN: 0022-1694. DOI: <https://doi.org/10.1016/j.jhydrol.2019.03.102>. URL: <https://www.sciencedirect.com/science/article/pii/S0022169419303312>.
- [4] J. D. Hunter. “Matplotlib: A 2D graphics environment”. Em: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [5] Md. Manjurul Hussain. *pyHomogeneity: A python package for homogeneity test of time series data*. Versão v1.1. Mai. de 2020. DOI: [10.5281/zenodo.3785287](https://doi.org/10.5281/zenodo.3785287). URL: <https://github.com/mmhs013/pyHomogeneity>.
- [6] Md. Manjurul Hussain e Ishtiaq Mahmud. “pyMannKendall: a python package for non parametric Mann Kendall family of trend tests.” Em: *Journal of Open Source Software* 4.39 (25 de jul. de 2019), p. 1556. ISSN: 2475-9066. DOI: [10.21105/joss.01556](https://doi.org/10.21105/joss.01556).

- URL: <http://dx.doi.org/10.21105/joss.01556>.
- [7] Hubert W Lilliefors. “On the Kolmogorov-Smirnov test for normality with mean and variance unknown”. Em: *Journal of the American statistical Association* 62.318 (1967), pp. 399–402.
 - [8] Naghettini M. e Pinto E. J. de A. *Hidrologia Estatística*. CPRM, 2007.
 - [9] Projeto MapBiomas. *Coleção 6.0 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil*. URL: <https://plataforma.brasil.mapbiomas.org/>.
 - [10] Anthony N Pettitt. “A non-parametric approach to the change-point problem”. Em: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28.2 (1979), pp. 126–135.
 - [11] Skipper Seabold e Josef Perktold. “statsmodels: econometric and statistical modeling with python”. Em: *9th Python in Science Conference*. 2010.
 - [12] C. Sharma e C. S. P. Ojha. “Modified Signal-to-Noise Ratio Method for Early Detection of Climate Change”. Em: *Journal of Hydrologic Engineering* 25.8 (2020), p. 04020032. DOI: [10.1061/\(ASCE\)HE.1943-5584.0001943](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001943). eprint: <https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29HE.1943-5584.0001943>. URL: <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0001943>.
 - [13] Operador Nacional do Sistema Elétrico (ONS). *Séries de vazão naturalizadas - 1931 a 2018*. URL: <https://dados.ons.org.br/>.
 - [14] The pandas development team. *pandas-dev/pandas: Pandas*. Versão 1.3.5. Dez. de 2021. DOI: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134). URL: <https://doi.org/10.5281/zenodo.3509134>.
 - [15] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. Em: *Nature Methods* 17 (2020), pp. 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).