

Regiões Hidrologicamente Homogêneas no Brasil: Uma abordagem baseada em índice de sazonalidade para regionalização de vazões baixas

Alexandre Sokoloski de Azevedo Delduque de Macedo¹ ; Michael Männich²

Palavras-Chave – Vazões baixas, regionalização, sazonalidade.

INTRODUÇÃO

Um dos principais desafios para a gestão de recursos hídricos é a estimativa de vazões em bacias não monitoradas. A regionalização de vazões baixas— como a Q95 (vazões baseadas na curva de permanência, que ocorrem em pelo menos 95% do tempo no corpo hídrico, considerando a hipótese de estacionaridade), torna-se estratégica para fins de planejamento da alocação de água e concessão de outorgas. Este estudo propõe uma abordagem baseada em índice de sazonalidade, especificamente o histograma de sazonalidade (SH) (Laaha, 2002), visando: (i) identificar padrões sazonais de vazões baixas em todo o Brasil; (ii) agrupar bacias com comportamento similar por meio de clustering (k-means); (iii) aplicar o algoritmo de regressão Lasso para desenvolver equações específicas para cada região sazonal hidrologicamente homogênea (RHH); (iv) possibilitar o uso dessas equações em sub-bacias não monitoradas, contribuindo para a regionalização das vazões mínimas com base em variáveis fisiográficas e climáticas.

METODOLOGIA

Para a realização dos cálculos e elaboração das análises, foram utilizados os dados de vazão provenientes da CAMELS-BR, que inicialmente possuem 4.025 estações de monitoramento no território nacional (Chagas *et al.*, 2020). Dois critérios de filtragem foram adotados: i. um período mínimo de 10 meses de dados anuais; ii. estações com mais de 10 anos de dados diários. Após essa seleção, 2.758 estações de monitoramento foram obtidas.

O indicador de sazonalidade utilizado neste estudo é o Histograma de Sazonalidade (SH) (Laaha, 2002). Esse método busca contabilizar os dias do ano em que ocorrem vazões iguais ou inferiores a Q95, calculando a frequência de ocorrência dessas vazões em cada mês do ano, resultando em 12 parâmetros mensais (Jan, Fev, ..., Dez) médios, conforme descrito por Laaha & Blöschl (2006).

Para identificar as RHH, em termos de sazonalidade, foi aplicado o método k-means (MacQueen, 1967). O método trata-se de um algoritmo não supervisionado que classifica observações multivariadas, a partir da distância euclidiana. A escolha e validação do número de clusters (k) foram feitas pelos métodos *Elbow* e *Silhouette*.

Para investigar os fatores que influenciam as vazões específicas regionais, foi aplicado um modelo de regressão (Lasso) com o objetivo de estimar os valores de q95 (L/s/km²). A análise foi realizada tanto para um modelo global quanto individualmente para cada RHH identificada a partir do agrupamento, utilizando 12 variáveis preditoras, que incluem uma combinação de características fisiográficas, de uso do solo, geológicas e climáticas, todas disponibilizadas na base CAMELS-BR.

1) Programa de Pós-graduação em Engenharia Ambiental, Universidade Federal do Paraná, e-mail: alexandresokoloski@ufpr.br
2) Departamento de Engenharia Ambiental, Universidade Federal do Paraná, Brasil, e-mail: mannich@ufpr.br

Para evitar valores de vazões negativas, a variável resposta foi transformada para uma escala log, conforme a equação 1, em que x é o vetor das variáveis preditoras, e α são os coeficientes de regressão estimados por Lasso.

$$\ln(q_{95} + 1) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n, \quad (1)$$

RESULTADOS

A aplicação da regressão Lasso às quatro RHH resultou nos valores apresentados na Tabela 1, com modelos ajustados individualmente para cada região. Conforme esperado, observou-se uma melhora nos indicadores de desempenho para a maioria dos parâmetros, especialmente em determinadas regiões, com características mais homogêneas e menos estações presentes no cluster.

Tabela 1 – Métricas de performance global e para cada cluster.

Model	Estações	R ²	NSE	KGE	Bias (%)	RMSE	NRMSE
Global	897	0.57	0.57	0.7	-9.93	2.56	0.08
Cluster 1	50	0.96	0.96	0.9	-9.12	1.69	0.05
Cluster 2	238	0.54	0.54	0.57	-9.42	1.99	0.13
Cluster 3	148	0.79	0.79	0.73	-12.59	2.06	0.08
Cluster 4	461	0.51	0.51	0.59	-8.8	2.43	0.09

CONCLUSÕES

Essa abordagem mostrou-se promissora para a estimativa de vazões específicas (q_{95}) em bacias não monitoradas, com base em variáveis fisiográficas e climáticas disponíveis. Os resultados destacam a relevância da sazonalidade como critério de regionalização, e reforçam a importância de métodos de seleção de variáveis, como o Lasso, para lidar com grandes conjuntos de dados.

Como perspectivas futuras, o foco será na investigação de outros modelos de aprendizado de máquina – como Random Forest, K-Nearest Neighbours (KNN), Support Vector Machines (SVM) e redes neurais artificiais – para comparação e aprimoramento da performance preditiva. Esses métodos poderão ser aplicados com base nos mesmos indicadores de sazonalidade, ou outros, buscando uma classificação mais refinada e precisa de Regiões Hidrologicamente Homogêneas no Brasil.

REFERÊNCIAS

- CHAGAS, V. B.; CHAFFE, P. L.; ADDOR, N.; FAN, F. M.; FLEISCHMANN, A. S.; PAIVA, R. C.; SIQUEIRA, V. A. (2020). “CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil”. *Earth System Science Data*, 12(3), pp. 2075–2096.
- LAAHA, G. (2002). Modelling summer and winter droughts as a basis for estimating river low flows. *International Association of Hydrological Sciences, Publication*, (274), 289-296.
- LAAHA, G.; BLÖSCHL, G. (2006). “Seasonality indices for regionalizing low flows”. *Hydrological Processes: An International Journal*, Wiley Online Library, 20(18), pp. 3851–3878.
- MACQUEEN, J. (1967, January). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics* (Vol. 5, pp. 281-298). University of California press.
- ## AGRADECIMENTOS
- O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.