

REGIONALIZAÇÃO DE VAZÕES DE ESTIAGEM NO ESTADO DO RIO GRANDE DO SUL EMPREGANDO TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL

MARCELLE MARTINS VARGAS¹; LEONARDO DE LIMA CORRÊA²; MATHEUS FONSECA DURAES³; MARILTON SANCHOTENE DE AGUIAR⁴; TAMARA LEITZKE CALDEIRA⁵; SAMUEL BESKOW⁶

¹Discente UFPel/Engenharia Hídrica– marcellevarg@gmail.com

²Discente UFRGS/PPG Computação– skllc@hotmail.com

³Docente UFPR/ Engenharia Agrícola– duraes@ufpr.br

⁴Docente UFPel/Ciência da Computação–.marilton@inf.ufpel.edu.br

⁵Discente UFPel/PPG Recursos Hídricos–tamaraleitzkecaldeira@gmail.com

⁶Docente UFPel/Engenharia Hídrica– samuelbeskow@gmail.com

1. INTRODUÇÃO

A tomada de decisões em bacias hidrográficas não monitoradas, ou que dispõem de uma série histórica curta, é considerada um dos problemas na análise de frequência de estiagens. Neste contexto, é comum empregar métodos de regionalização visando enquadrar bacias hidrográficas em uma região homogênea com relação à variável de interesse.

De acordo com RAO; SRINIVAS (2006a) regiões geograficamente contíguas nos limites geográficos, políticos, administrativos ou fisiográficos têm sido usadas em hidrologia. Todavia, estes pesquisadores afirmam que esta prática não é tecnicamente aceitável, visto que estes procedimentos não garantem homogeneidade hidrológica. As informações hidrológicas contidas em uma região homogênea são utilizadas para estimar os quantis da variável de interesse na bacia hidrográfica não monitorada ou para melhorar as estimativas em bacias que tenham série histórica curta (RAO; SRINIVAS, 2006a).

A clusterização é um processo pelo qual um conjunto de vetores é dividido em clusters ou grupos de tal forma que os vetores dentro de um cluster são similares e os vetores de diferentes clusters são dissimilares (RAO; SRINIVAS, 2006b). Para a análise de cluster em hidrologia é necessário definir atributos para permitir diferenciar bacias hidrográficas visando agrupá-las. Comumente são empregados como atributos diferentes características fisiográficas e localização geográfica, além de indicadores relacionados às séries de vazões, sendo bastante variável com o foco do estudo.

O objetivo deste estudo foi avaliar diferentes técnicas de Inteligência Artificial (IA) para clusterização, incluindo os algoritmos K-means, Partioned Around Medoids (PAM), K-Harmonic Means (KHM), Genetic K-Means (GKA) e Fuzzy C-Means (FCM), para determinar regiões hidrologicamente homogêneas no Rio Grande do Sul visando à regionalização hidrológica da vazão que é igualada ou superada em 90% do tempo (Q_{90}), baseando-se em atributos de localização das bacias hidrográficas e atributos associados à sazonalidade de vazões de estiagem.

2. METODOLOGIA

As análises expostas neste trabalho foram realizadas em bacias hidrográficas de diferentes tamanhos, sendo as séries históricas de vazão utilizadas, obtidas através do portal HidroWeb da Agência Nacional de Águas (ANA). Em primeiro momento, considerou-se os 108 postos empregados por BESKOW et al. (2014) no estudo de sazonalidade de vazões de estiagem no Rio

Grande do Sul. Estas séries foram submetidas ao teste de estacionariedade de Mann-Kendall, seguindo recomendações de MALAMUD et. al (2011). Na sequência, as vazões de estiagem foram calculadas a partir de vazões diárias, empregando-se o quantil de vazão Q_{90} , conforme sugerem NAGHETTINI; PINTO (2007). Para avaliar a sazonalidade de vazões de estiagem a fim de diferenciá-las entre bacias hidrográficas, aplicou-se o método de Seasonality Histogram (SH) (LAAHA; BLÖSCHL, 2006), que consiste em um histograma de sazonalidade em escala mensal. Foram separados os dados de vazão média diária em que se observou valores inferiores à Q_{90} e suas respectivas datas de ocorrência. Desta forma, foi derivado o histograma de sazonalidade de vazões de estiagem em cada seção de controle, contendo a frequência de cada mês do ano (12 atributos), como recomendado em LAAHA; BLÖSCHL (2006) e VEZZA et. al (2010). Além destes 12 atributos, foram consideradas a latitude e a longitude do centroide de cada bacia para a análise.

Com o propósito de agrupar bacias hidrográficas de comportamento hidrológico semelhante em relação à Q_{90} , com base nos resultados do método SH, avaliou-se as seguintes técnicas de IA (CORRÊA, 2014): PAM, KHM, GKA e FCM. As métricas de distâncias Euclidiana, Manhattan e Mahalanobis foram empregadas para avaliar o grau de similaridade de clusters, enquanto K-Means ++, Aleatório (Random) e Origem foram os métodos de inicialização adotados.

Para cada cluster gerado com diferentes estações fluviométricas, foi ajustado um modelo matemático para relacionar a Q_{90} com a área de drenagem das bacias hidrográficas. O índice de confiança “c”, bem como seu enquadramento, descritos em CAMARGO; SENTELHAS (1997), foram utilizados para avaliar o desempenho dos modelos ajustados. O valor de “c” foi calculado utilizando o coeficiente de correlação “r” e o coeficiente de exatidão (d), através das Equações 1 e 2.

$$d = 1 - \left[\frac{\sum(P_i - O_i)^2}{\sum(|P_i - O_i| + |O_i - O|)^2} \right] \quad (1)$$

em que P_i são valores estimados de Q_{90} , O_i são valores observados de Q_{90} e O é o valor médio observado de Q_{90} .

$$c = r \cdot d \quad (2)$$

Para verificar a homogeneidade desses grupos, foi utilizado o teste H, desenvolvido por HOSKING; WALLIS (1993) baseado nos momentos-L. A importação, manuseio e elaboração de séries hidrológicas, assim como o teste de homogeneidade H, foram realizados no aplicativo computacional “System of Hydrological Data Acquisition and Analysis” (SYHDA) (BESKOW et al., 2013), o qual está sendo desenvolvido no Laboratório de Hidrologia e Modelagem Hidrológica, vinculado ao curso de Engenharia Hídrica e ao Programa de Pós-Graduação em Recursos Hídricos, da Universidade Federal de Pelotas (UFPEl).

3. RESULTADOS E DISCUSSÃO

Através do teste de Mann-Kendall, aplicado nas 108 séries históricas de vazão analisadas, a um nível de significância de 5% para as vazões médias anuais, observou-se que desse total, 78 séries foram consideradas estacionárias e empregadas neste trabalho.

Através dos valores médios do índice “c”, considerando 6 regiões e as respectivas equações de Q_{90} versus área, para cada um dos cenários analisados, os quais combinam diferentes técnicas de IA, de distâncias métricas e de inicialização, observou-se que a técnica FCM proporcionou ajuste “Ótimo” para as

diferentes combinações, de acordo com a classificação de CAMARGO; SENTELHAS (1997), seguido pela técnica GKA. Dentre todos os cenários analisados, o que melhor apresentou resultados para as regiões formadas foi a técnica de IA FCM com inicialização K-means ++ e distância métrica Euclidiana.

Na Tabela 1 são apresentados os detalhes acerca dos resultados obtidos para cada região formada pela melhor combinação citada. Analisando os resultados do índice de confiança “c”, observa-se que estes foram enquadrados como “Ótimo” ($c > 0,85$) para as 6 regiões, de acordo com a classificação de CAMARGO; SENTELHAS (1997).

Tabela 1 – Informações detalhadas para cada região estabelecida pela combinação da técnica de IA “FCM”, método de inicialização “K-means++” e distância métrica “Euclidiana”

Região	Função $Q_{90}(A)^*$	Intervalo de área das bacias hidrográficas (km ²)	Número de postos	Índice de confiança “c”	H**
1	$Q_{90} = 0,0042A^{0,8628}$	57,48 - 30702,24	17	0,97	-0,04
2	$Q_{90} = 0,001A^{1,1769}$	435,58 - 27299,63	7	0,96	-0,14
3	$Q_{90} = 0,00007A^{1,3043}$	596,28 - 5941,40	9	0,98	24,16
4	$Q_{90} = 0,0049A^{0,9983}$	314,15 - 17126,19	21	0,98	-0,27
5	$Q_{90} = 0,0006A^{1,1322}$	62,09 - 15889,49	9	0,99	-0,06
6	$Q_{90} = 0,0007A^{1,1789}$	102,36 - 12399,9	15	0,92	-0,04

* Q_{90} em $m^3.s^{-1}$ e A em km^2 ; ** Resultados do teste H de HOSKING; WALLIS (1993)

Pode-se constatar na Tabela 1 que, de acordo com o teste H (HOSKING; WALLIS, 1993), 5 regiões podem ser consideradas homogêneas ($H < 1$). Ou seja, apenas a região 3 foi considerada heterogênea ($H \geq 2$), sendo esta alterada com a exclusão de 3 bacias para melhor enquadramento da metodologia. Desta forma, o novo modelo matemático para a região 3 ($Q_{90} = 0,0001A^{1,2496}$) resultou no índice “c” igual a 0,97 e esta região pôde ser considerada homogênea, pois o valor de H foi de 0,46. A distribuição espacial das regiões homogêneas, formadas pela melhor combinação, pode ser visualizada na Figura 1.

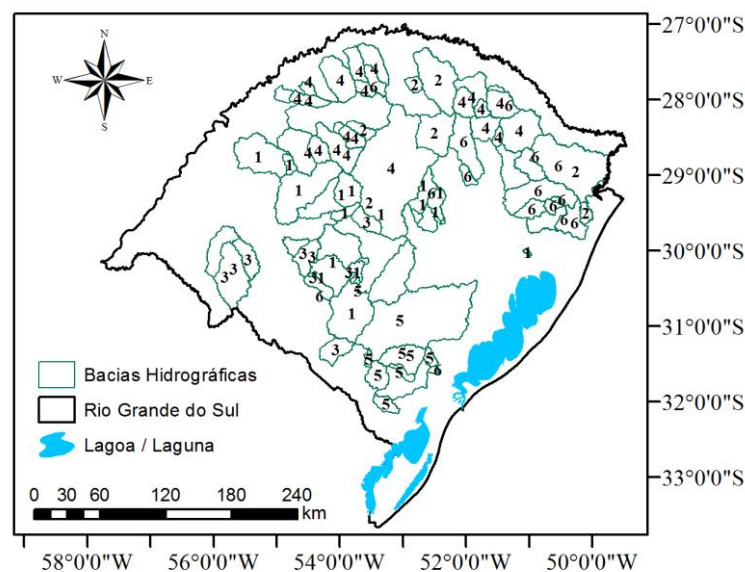


Figura 1 – Bacias hidrográficas integrantes das 6 regiões determinadas pela combinação da técnica de IA “FCM”, método de inicialização “K means ++” e a distância métrica “Euclidiana”

4. CONCLUSÕES

Através dos resultados apresentados neste trabalho foi possível concluir que: i) técnicas de IA têm potencial para definir regiões homogêneas para regionalização da Q_{90} no estado do Rio Grande do Sul, em especial os métodos FCM, baseado em lógica fuzzy; ii) os atributos relacionados à sazonalidade de vazões de estiagem, através do método SH, agregaram informações importantes e representativas às bacias hidrográficas, facilitando a clusterização; iii) os modelos matemáticos ajustados tiveram ótimo desempenho, mesmo com apenas uma variável explicativa (área) bastante fácil de ser obtida; desta forma, tais modelos podem ser utilizados no estado para estimativa de Q_{90} em locais sem monitoramento.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- BESKOW, S.; CORRÊA, L.; MAHL, M.; SIMÕES, M. C.; CALDEIRA, T. L.; NUNES, G. S.; LUCAS, E. H.; FARIA, L. C.; MELLO, C. R. desenvolvimento de um sistema computacional de aquisição e análise de dados hidrológicos. In: **Anais do XX Simpósio Brasileiro de Recursos Hídricos**. Bento Gonçalves, 2013.
- BESKOW, S.; MELLO, C. R. de; FARIA, L. C.; SIMOES, M. C.; CALDEIRA, T. L.; NUNES, G. S. Índices de sazonalidade para regionalização hidrológica de vazões de estiagem no Rio Grande do Sul. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.18, n.7, p. 740 – 746, 2014.
- CAMARGO, A. P; SENTELHAS, P. C. Avaliação do desempenho de diferentes métodos de estimativas da evapotranspiração potencial no Estado de São Paulo, Brasil. **Revista Brasileira de Agrometeorologia**, v.5, n.1, p. 89 – 97, 1997.
- CORRÊA, L. L. **Implementação e análise de técnicas de inteligência artificial aplicados à clusterização em recursos hídricos**. 2014, 96 p. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação), Universidade Federal de Pelotas.
- HOSKING, J. R. M.; WALLIS, J. R. Some statistics useful in regional frequency analysis. **Water Resources Research**, v. 29, n.2, p. 271 – 281, 1993.
- LAHA, G.; BLÖSCHL, G. Seasonality indices for regionalizing low flows. **Hydrological Processes**, v. 20, p. 3851 – 3878, 2006.
- MALAMUD, B.D.; TURCOTTE, D.L.; GRIMMOND, C. S. B. Temperature trends at the Mauna Loa observatory, Hawaii. **Clima of the Past**, v.7, p. 975 – 983, 2011.
- NAGHETTINI, M.; PINTO, E. J. de A. **Hidrologia estatística**. Belo Horizonte: CPRM, 2007.
- RAO, A. R.; SRINIVAS, V. V. Regionalization of watersheds by fuzzy cluster analysis. **Journal of Hydrology**, v. 318, p. 57 – 79, 2006a.
- RAO, A. R.; SRINIVAS, V. V. Regionalization of watersheds by hybrid-cluster analysis. **Journal of Hydrology**, v.318, p. 37 – 56, 2006b.
- VEZZA, P.; COMOGLIO, C.; ROSSO, M.; VIGLIONE, A. Low flows regionalization in North-Western Italy. **Water Resources Management**, v.24, p.4049 – 4074, 2010.