

RESUMO - CIÊNCIAS EXATAS E DA TERRA - MATEMÁTICA

**ANÁLISE DE AGRUPAMENTOS E COMPONENTES PRINCIPAIS
APLICADAS ÀS SÉRIES TEMPORAIS DE ENERGIA NATURAL AFLUENTE
NO BRASIL**

Thales Da Silva Barroso Pessoa (thasbp@gmail.com)

Felipe Leite Coelho Da Silva (felipeleite@ufrj.br)

A matriz energética brasileira, fortemente dependente de fontes renováveis, tem na geração hidrelétrica seu principal pilar, sendo suscetível às vulnerabilidades climáticas (1). Em 2024, as hidrelétricas representaram cerca de 60% da capacidade de geração do país (2), o que torna o estudo da Energia Natural Afluente (ENA) um fator determinante para a segurança energética e o planejamento de longo prazo (3). A ENA, que quantifica o volume de água disponível para geração nas usinas e é diretamente influenciada pelos regimes de chuva, tem seus dados históricos disponibilizados pelo Operador Nacional do Sistema Elétrico (ONS). Neste contexto, o presente estudo propõe-se a investigar os padrões de comportamento da ENA nas bacias hidrográficas brasileiras por meio de uma abordagem de análise multivariada. Para a análise, foram empregadas as técnicas de Análise de Componentes Principais (PCA) e de agrupamento (clustering), implementadas com o suporte das linguagens de programação R (4) e Python (5). A PCA foi utilizada como etapa de pré-processamento para reduzir a dimensionalidade dos dados, condensando a variabilidade original em um conjunto menor de componentes não correlacionados. Essa abordagem otimizou a subsequente aplicação dos algoritmos de agrupamento, realçando os padrões estruturais latentes nos

dados. A etapa de agrupamento utilizou os métodos hierárquico aglomerativo e K-médias, aplicados sobre as componentes principais com base na distância euclidiana. A escolha do número ótimo de agrupamentos foi validada pelo método do cotovelo (elbow method), que apontou $k=3$ como a configuração ideal, resultando na segmentação das bacias em três grupos com perfis hidrológicos distintos. Os resultados revelaram a formação de agrupamentos consistentes, evidenciando, por exemplo, um grupo distinto composto pelas bacias b9, b4 e b13, que apresentaram maior dissimilaridade em relação às demais. A análise do dendrograma e dos grupos formados corrobora a existência de macrorregiões com comportamento hidrológico similar, cujo entendimento é valioso para otimizar a gestão compartilhada de reservatórios e a alocação de energia. Conclui-se que a metodologia combinada de PCA e clustering se mostrou robusta e eficaz para desvendar padrões de semelhança entre as bacias hidrográficas, fornecendo uma ferramenta analítica poderosa para o planejamento do setor elétrico. As conclusões deste trabalho oferecem insights para o desenvolvimento de estratégias de gestão de recursos hídricos mais eficientes e sustentáveis, aumentando a resiliência do sistema elétrico nacional diante das incertezas hidrológicas e climáticas.

(1)LUCENA, A. F. P. et al. The vulnerability of renewable energy to climate change in Brazil. *Energy Policy*, Amsterdam, v. 37, n. 3, p. 879-889, 2009.

(2)EMPRESA DE PESQUISA ENERGÉTICA (EPE). Expansão da geração. [S. l.], 2024. Disponível em: <https://www.epe.gov.br/pt/areas-de-atuacao/energia-eletrica/expansao-da-geracao/fontes>. Acesso em: 4 maio 2025.

(3)SILVA, F. L. C. et al. A bottom-up methodology for long term electricity consumption forecasting of an industrial sector – application to pulp and paper sector in Brazil. *Energy*, [S. l.], v. 144, p. 1107-1118, 2018. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0360544217321217>. Acesso em: [inserir data].

(4)R CORE TEAM. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2022. Disponível em: <http://www.R-project.org>. Acesso em: [inserir data].

(5)PYTHON SOFTWARE FOUNDATION. Python Language Site: Documentation. [S. l.], 2022. Disponível em: <https://www.python.org/doc/>. Acesso em: 10 nov. 2022.

Palavras-chave: k-médias; energia natural afluyente; análise de componentes principais.