

MODELAGEM DA PERSISTÊNCIA DE PERÍODOS SECOS POR CADEIAS DE MARKOV PARA SUPORTE À GESTÃO DA DISPONIBILIDADE HÍDRICA

Gabriel Taquichiri Vargas^{1, 2}; Nury Bertini Bernal²; Dário Hachisu Hossoda^{1, 2}; Pedro Ludovico Bozzini²; Carla Voltarelli Franco da Silva²; Alexandre Nunes Roberto²; Luciana Campos de Oliveira²; Sidnei Ono³; Joaquin Ignacio Bonnacarrère Garcia¹

RESUMO - Este trabalho apresenta um modelo hidrometeorológico baseado em cadeias de Markov de primeira ordem para caracterização da persistência de períodos secos e suporte à gestão de riscos em sistemas de captação de água. O método utiliza séries históricas de precipitação diária para classificar estados secos e úmidos a partir de um limiar operacional de 2 mm/dia, estimando probabilidades de transição entre estados consecutivos. A persistência da seca foi representada pelo parâmetro P_{SS} , utilizado para calcular a duração média esperada das sequências secas e a probabilidade de ocorrência de eventos prolongados. A partir dessas probabilidades, foi desenvolvido um índice de criticidade capaz de converter a raridade estatística dos eventos em níveis graduais de alerta operacional, compatível com protocolos de resposta emergencial urbana. Os resultados obtidos nas estações piloto demonstraram boa aderência entre o modelo teórico e as frequências observadas, com erros médios inferiores a 5%. O método mostrou potencial para aplicação em sistemas de monitoramento hidrometeorológico e suporte à gestão da disponibilidade hídrica em tempo quase real, contribuindo para o fortalecimento da resiliência hídrica urbana frente às mudanças climáticas.

Palavras-chave: modelagem de secas, disponibilidade hídrica, cadeias de Markov.

ABSTRACT - The increasing occurrence of prolonged droughts in urban areas represents a direct threat to urban water security, requiring monitoring tools capable of anticipating and grading operational risks in water intake systems. This study presents a hydrometeorological model based on first-order Markov chains to characterize dry-period persistence and support risk management in water intake systems. The method uses historical daily precipitation series to classify dry and wet states based on an operational threshold of 2 mm/day, and estimate transition probabilities between consecutive states. Drought persistence is represented by the parameter P_{SS} , which is used to estimate the expected duration of dry sequences and the probability of prolonged events. Based on these probabilities, a criticality index was developed to convert the statistical rarity of events into graded levels of operational alert, compatible with urban emergency response protocols. Results obtained for pilot stations showed good agreement between theoretical and observed distributions, with mean errors below 5%. The proposed approach demonstrated potential for real-time hydrometeorological monitoring and for supporting near -real time water availability management, contributing to strengthening urban water resilience under climate change conditions.

Keywords: drought modelling, water availability management, Markov chains.

¹Departamento de Engenharia Hidráulica e Ambiental da Escola Politécnica da Universidade de São Paulo. gabriel.taquichiri.vargas@usp.br; dario.hossoda@usp.br; joaquinbonne@usp.br.

²Laboratório de Sistemas de Suporte a Decisão (LabSid), Fundação Centro Tecnológico de Hidráulica. nury@labsid.eng.br; pedro@labsid.eng.br; carla@labsid.eng.br; alexandre@labsid.eng.br; luciana@labsid.eng.br.

³Hidroinfo. sidnei@hidroinfo.com.br.

INTRODUÇÃO

A ocorrência de períodos com baixa precipitação prolongados representa um dos principais desafios para a gestão operacional de sistemas de captação de água e consequentemente para a segurança hídrica, especialmente em regiões sujeitas à elevada variabilidade hidrometeorológica. Diferente de eventos de cheia, que respondem rapidamente a episódios intensos de precipitação, as secas hidrológicas se desenvolvem de forma gradual e muitas vezes pouco perceptível, propagando-se pelo sistema hidrológico até afetar escoamento, níveis de rios, reservatórios e aquíferos (Van Loon, 2015). Em contextos de alta variabilidade hidrometeorológica, como a Região Metropolitana de São Paulo, a incapacidade de identificar e graduar esse risco em tempo hábil pode resultar em crises operacionais de grande impacto social e econômico, segundo Tercini e Mélo Júnior (2024).

Tundisi (2007) afirma que a adaptação das cidades às mudanças climáticas requer sistemas de monitoramento que transformem dados pluviométricos em inteligência operacional. A persistência temporal das secas influencia diretamente a disponibilidade hídrica, o planejamento operacional e a necessidade de adoção de medidas preventivas em situações críticas (Fernández *et al.*, 2023). Tradicionalmente, análises hidrológicas utilizam indicadores baseados em volumes acumulados ou médias climatológicas. Entretanto, tais abordagens nem sempre representam adequadamente a continuidade temporal dos eventos secos, aspecto fundamental para a avaliação do risco hidrológico operacional. Métodos clássicos da hidrologia apresentam limitações quando há baixa disponibilidade de dados, dificultando a obtenção de conclusões estatisticamente robustas e a adequada representação da variabilidade hidrológica, especialmente na estimativa da frequência de secas (Akyuz *et al.*, 2012). Nesse contexto, modelos probabilísticos baseados em Cadeias de Markov têm sido amplamente empregados na representação de sequências seco-úmido, devido à sua simplicidade, robustez estatística e interpretabilidade (Ochola; Kerkides, 2003; Cao *et al.*, 2023; Avilés *et al.*, 2016; Sonnadara; Jayewardene, 2015).

Este trabalho apresenta a aplicação de um modelo hidrometeorológico baseado em cadeias de Markov de primeira ordem para a caracterização da persistência de períodos secos em captações urbanas, com foco na geração de alertas operacionais graduais. O método utiliza séries históricas de precipitação diária para estimar probabilidades de transição entre estados secos e úmidos, convertendo parâmetros probabilísticos em um índice de criticidade compatível com protocolos de resposta emergencial. A abordagem é diretamente aplicável à gestão de riscos hídricos em tempo quase real, em alinhamento com as estratégias de resiliência e adaptação climática para áreas urbanas.

MATERIAIS E MÉTODOS

Base de dados

Foram utilizados dados históricos de precipitação diária provenientes da rede telemétrica do Sistema de Alerta a Inundações de São Paulo com a Companhia de Saneamento de São Paulo

(SAISP - Sabesp), abrangendo o período entre maio de 2021 e fevereiro de 2026 . Este período é particularmente relevante do ponto de vista da gestão de riscos, pois a partir do final de 2025 a região de estudo foi caracterizada com a intensidade S2 referente à “seca grave” de acordo com o Monitor de Secas da Agência Nacional de Águas e Saneamento Básico (2026). As séries históricas foram organizadas em uma base integrada contendo informações pluviométricas e metadados operacionais das captações analisadas.

As estações com maior consistência temporal foram utilizadas na calibração e validação do método, com destaque para as captações a fio d’água da Sabesp nos rios Tatuí, rio Sarapuí, rio Mandacaru e captação Sub-Álvea, conforme Figura 1.

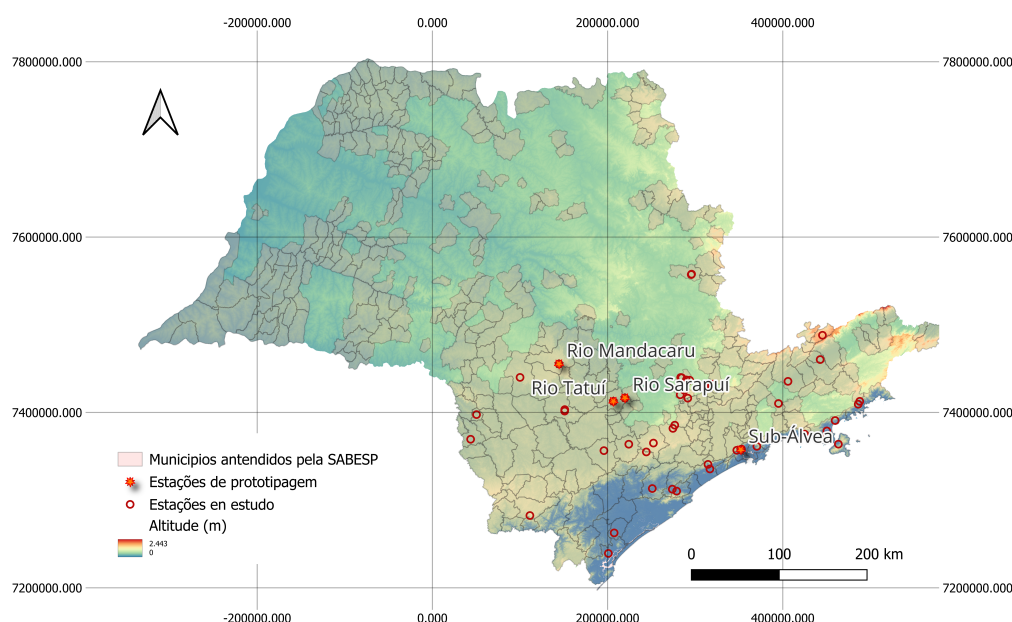


Figura 1: Região de Estudo

A Figura 1 apresenta a localização da região de estudo inserida no estado de São Paulo, no qual se identificam os 376 municípios atendidos pela Sabesp. Dentre esses, destacam-se 51 captações a fio d’água que dispõem recentemente de dados de pluviometria diária provenientes da rede SAISP. Entre elas, as quatro captações selecionadas para o modelo possuem séries temporais com extensão suficientemente longa (4 anos), permitindo um ajuste adequado da metodologia proposta. As demais 47 captações possuem o mesmo potencial para serem analisadas por meio da mesma metodologia proposta, sendo que os resultados podem ser progressivamente recalibrados à medida que houver ampliação dos registros disponíveis, assim como a incorporação de melhorias no ajuste.

Ambiente Computacional

O modelo foi desenvolvido em linguagem Python (versão 3.13), em ambiente interativo Jupyter Notebook. A manipulação e a análise das séries históricas de precipitação foram realizadas com as bibliotecas *pandas* (McKinney, 2010) e *NumPy* (Harris *et al.*, 2020). A geração das visualizações estáticas, incluindo as distribuições acumuladas complementares (CCDF) e os gráficos de validação, foi realizada com *Matplotlib* (Hunter, 2007), enquanto os painéis operacionais interativos de monitoramento em tempo quase real foram produzidos com a biblioteca *Plotly* (Plotly Technologies Inc., 2015). Todas as ferramentas são de código aberto, o que assegura a reprodutibilidade integral dos resultados e a portabilidade do método a outros contextos operacionais.

Classificação de estados secos e úmidos

A série de precipitação diária ($P_{\text{diária}}$) foi convertida em uma sequência binária de estados hidrometeorológicos. Considerou-se como dia seco toda precipitação diária inferior a 2 mm, enquanto valores iguais ou superiores foram classificados como dias úmidos.

$$P_{\text{diária}} < 2 \text{ mm} \quad (1)$$

O limiar adotado foi definido a partir da análise conjunta entre precipitação e resposta hidrológica observada nas estações piloto, sendo compatível com condições de baixa contribuição hidrológica nas bacias analisadas (Ross *et al.*, 2021).

Cadeia de Markov de primeira ordem

A persistência dos períodos secos foi modelada por meio de uma cadeia de Markov de primeira ordem composta por dois estados: seco e úmido. A matriz de transição (Equação 2) foi estimada a partir das frequências observadas de transição entre dias consecutivos.

$$P = \begin{pmatrix} P_{SS} & P_{SC} \\ P_{CS} & P_{CC} \end{pmatrix} \quad (2)$$

onde P_{SS} é probabilidade de um dia seco ser seguido por outro dia seco, P_{SC} é a probabilidade de interrupção da sequência seca, P_{CS} é probabilidade de transição de úmido para seco e P_{CC} é probabilidade de continuidade de dias úmidos. E por definição, a soma das probabilidades de eventos complementares (Equação 3 e Equação 4) é igual a 1 (ou 100%):

$$P_{SS} + P_{SC} = 1 \quad (3)$$

$$P_{CS} + P_{CC} = 1 \quad (4)$$

A duração média esperada das sequências secas ($E[L_{\text{seco}}]$) foi estimada pela Equação 5

$$E[L_{\text{seco}}] = \frac{1}{1 - P_{SS}} \quad (5)$$

A probabilidade de ocorrência de sequências secas com duração igual ou superior a k dias foi ($P(X \geq k)$) calculada assumindo distribuição geométrica, dada pela Equação 6:

$$P(X \geq k) = P_{SS}^{k-1} \quad (6)$$

Índice de criticidade

A conversão das probabilidades de persistência em níveis operacionais de alerta constitui o núcleo da contribuição deste trabalho para a gestão de riscos urbanos. Foi definido um índice de criticidade operacional (Equação 7) capaz de traduzir a raridade estatística das sequências secas em níveis graduais de alerta aplicáveis a protocolos de resposta operacional e emergencial.

$$IC = 1 - P(X \geq k) \quad (7)$$

Substituindo a Equação 6 na Equação 7, obtém-se a Equação 8:

$$IC = 1 - P_{SS}^{k-1} \quad (8)$$

Os valores do índice foram classificados em cinco níveis operacionais: Normal, Observação, Alerta, Crítico e Extremo, conforme os critérios da matriz de risco propostos por (Bittencourt *et al.*, 2021). Essa escala gradual permite que gestores de sistema de abastecimento urbano adotem medidas preventivas proporcionais ao nível de risco identificado, desde o monitoramento reforçado até ativação de protocolos emergenciais de restrição hídrica (Figura 2).

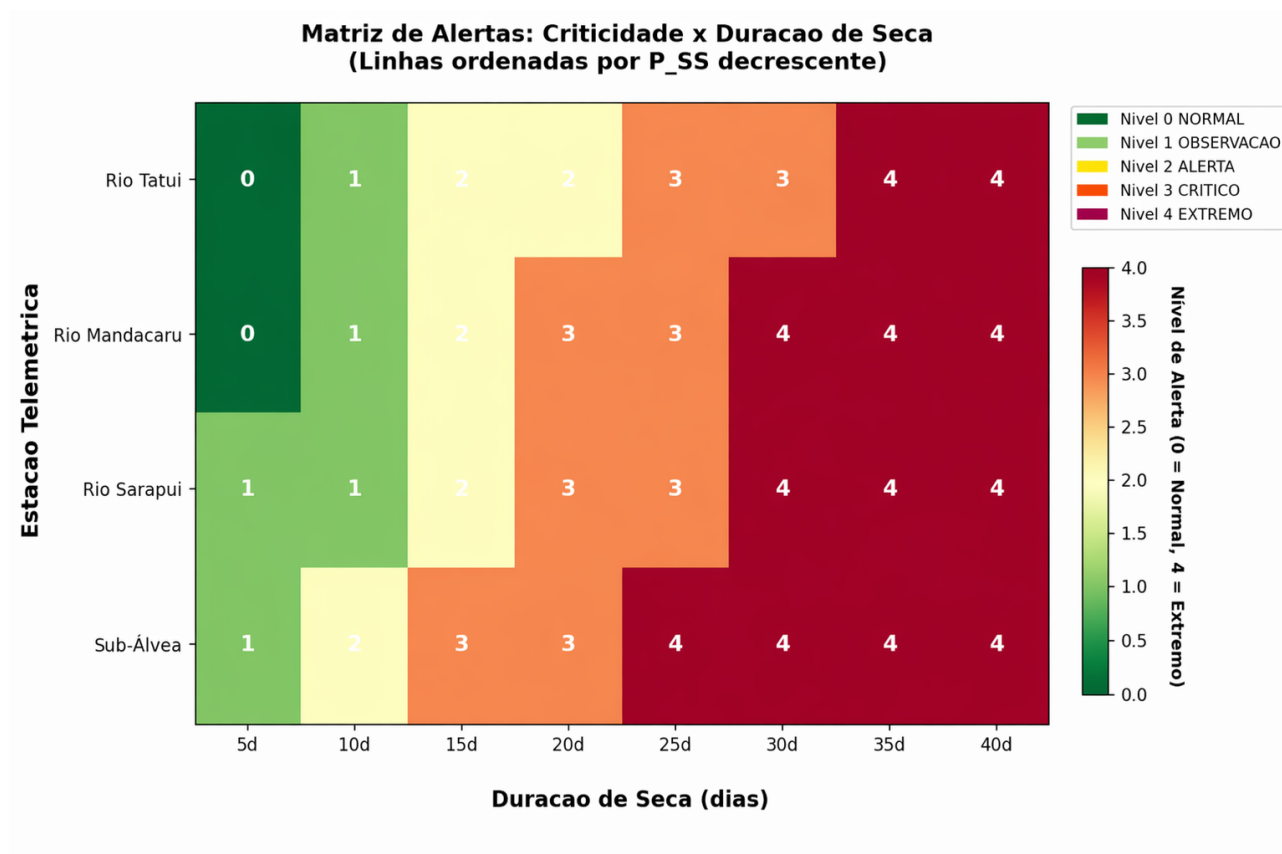


Figura 2: Cenários de criticidade por estação

RESULTADOS E DISCUSSÃO

Parâmetros de persistência

A aplicação do modelo evidenciou diferenças significativas na persistência de períodos secos entre as captações analisadas. As estações Rio Tatuí e Rio Mandacarú apresentaram elevados valores de P_{SS} , indicando maior tendência à continuidade de eventos secos consecutivos. A Tabela 1 apresenta os parâmetros estimados para as estações piloto.

Tabela 1: Parâmetros de transição estimados para as estações piloto.

Estação	P_{SS}	P_{SC}
Rio Tatuí	0.8731	0.1269
Rio Mandacarú	0.8510	0.1490
Rio Sarapuí	0.8391	0.1609
Sub-Álvea	0.7931	0.2069

Os resultados demonstram comportamento hidrológico distinto entre as captações, refletindo diferenças estruturais associadas às características físicas e hidrológicas das bacias contribuintes. Do ponto de vista da gestão de riscos urbanos, essa heterogeneidade reforça a necessidade de calibração estação-específica do índice de criticidade, em vez de adotar limiares genéricos para toda a rede.

Validação do modelo

A aderência da cadeia de Markov foi avaliada por meio da comparação entre as probabilidades teóricas estimadas e as frequências observadas das sequências secas. Os erros médios permaneceram inferiores a 5% nas estações piloto, indicando boa representação estatística da persistência observada.

A estação Rio Tatuí apresentou um erro médio de aproximadamente 1,09%, evidenciando elevada concordância entre o modelo e os dados históricos. Resultados semelhantes foram observados para o Rio Sarapuí, o Rio Mandacaru e a Captação Sub-Álvea.

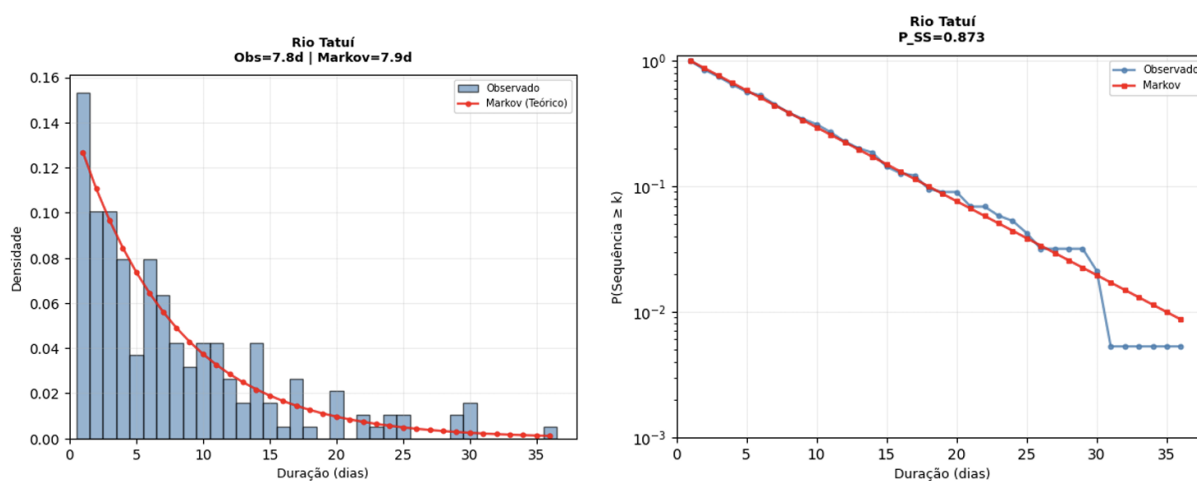


Figura 3: Distribuições acumuladas complementares do Rio Tatuí

A análise das distribuições acumuladas complementares (CCDF) (Figura 3) mostra que o modelo reproduz adequadamente o comportamento das sequências secas, com melhor concordância para durações mais curtas, particularmente até aproximadamente 15 dias.

Para durações mais longas, observa-se um aumento gradual da divergência entre as probabilidades teóricas e observadas, com tendência à subestimação das frequências associadas a eventos mais raros. Ainda assim, o modelo preserva a tendência geral da cauda da distribuição.

Criticidade operacional

A conversão da persistência probabilística em níveis operacionais de criticidade permitiu estabelecer limiares específicos por estação. Dessa forma, uma mesma duração de seca pode representar níveis distintos de risco operacional em diferentes captações.

Captações com maior persistência seca apresentaram limiares mais elevados para transição entre níveis de alerta, refletindo maior recorrência histórica de sequências prolongadas. Em contraste, estações com menor persistência atingem níveis críticos em intervalos mais curtos de estiagem.

O método demonstrou potencial para aplicação em sistemas de monitoramento contínuo, possibilitando a atualização automática do nível de criticidade com base na evolução diária da sequência seca observada. Essa característica o torna compatível com sistemas de alerta precoce

para secas urbanas, uma vez que permite incorporar prognósticos de diferentes modelos meteorológicos de curto prazo, viabilizando análises de até quinze dias sobre a situação operativa na bacia de captação. A Figura 4 apresenta um exemplo de monitoramento operacional em tempo quase real utilizando previsões hidrometeorológicas integradas.

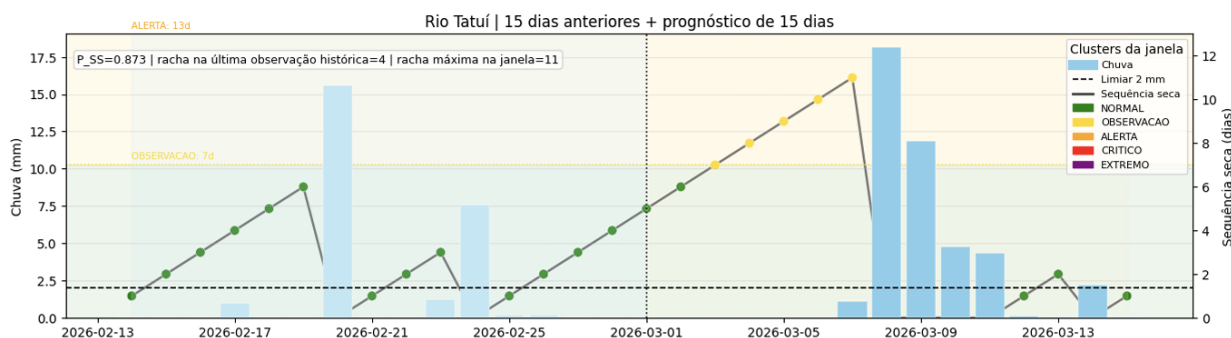


Figura 4: Monitoramento de evento de seca com protótipo hidrometeorológico utilizando dados SAISP/Sabesp e previsões FCTH/SAISP/GFS28 na captação do Rio Tatuí.

CONCLUSÕES

O modelo desenvolvido demonstrou capacidade satisfatória para representar a persistência de períodos secos por meio de cadeias de Markov de primeira ordem, com aplicação direta ao monitoramento de riscos hídricos em sistemas urbanos de captação. Os resultados indicaram boa concordância entre as probabilidades teóricas estimadas e as frequências observadas nas séries históricas analisadas. Em especial, o parâmetro P_{SS} mostrou-se eficiente como indicador da persistência da seca, permitindo estimar a duração média esperada das sequências secas e a probabilidade de ocorrência de eventos prolongados.

O índice de criticidade proposto representa uma contribuição metodológica relevante para a gestão de riscos hídricos urbanos: ao converter probabilidades de persistência em cinco níveis graduais de alerta — Normal, Observação, Alerta, Crítico e Extremo —, o método oferece uma interface direta entre a análise hidrometeorológica e os protocolos operacionais de resposta emergencial. Essa capacidade de traduzir dados estatísticos em informação acionável é fundamental para o fortalecimento da resiliência urbana frente a eventos extremos de seca, em especial em um contexto de intensificação dos efeitos das mudanças climáticas sobre o ciclo hidrológico das cidades.

Os resultados obtidos indicam que o método apresenta potencial para integração em sistemas operacionais de monitoramento hidrometeorológico e alerta de secas em tempo quase real, contribuindo para a antecipação de crises hídricas e o suporte à tomada de decisão em sistemas de abastecimento urbano.

Embora os resultados demonstrem boa aderência estatística, o método depende diretamente do limiar adotado para a classificação dos estados secos e úmidos. Adicionalmente, a utilização de uma cadeia de Markov de primeira ordem simplifica parcialmente a dinâmica hidrometeorológica

real, que pode apresentar dependências temporais mais complexas. Nesse sentido, aperfeiçoamentos metodológicos progressivos poderão ampliar a representatividade do modelo à medida que as séries históricas das captações se tornarem mais extensas e consistentes.

Entre as evoluções mais promissoras, destacam-se: o acoplamento com variáveis de vazão e nível d'água, que permitiria refletir a resposta hidrológica efetiva da bacia e não apenas a persistência pluviométrica; a adoção de cadeias de Markov de segunda ou terceira ordem, capazes de incorporar dependências temporais mais complexas e representar com maior fidelidade o regime hidrológico local; e o ajuste do defasamento temporal entre precipitação e resposta hidrológica, especialmente relevante em captações com períodos de retorno mais longos. Essas evoluções não requerem reestruturação metodológica, uma vez que os parâmetros da cadeia podem ser recalibrados de forma contínua conforme novos registros operacionais são incorporados à base de dados de cada captação, conforme discutido por Van Loon (2015).

AGRADECIMENTOS

Os autores agradecem o apoio do Programa de Pós-Graduação em Engenharia Civil da Universidade de São Paulo (PPGEC-USP), à Fundação Centro Tecnológico de Hidráulica (FCTH) e à Companhia de Saneamento Básico do Estado de São Paulo (Sabesp). O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO. *Monitor de Secas do Brasil*.

Agência Nacional de Águas e Saneamento Básico. 2026. Disponível em:

<https://monitordesecas.ana.gov.br>. Acesso em: 8 maio 2026.

AKYUZ, D. E.; BAYAZIT, M.; ONOZ, B. Markov chain models for hydrological drought characteristics. en. *J. Hydrometeorol.*, American Meteorological Society, v. 13, n. 1, p. 298–309, fev. 2012.

AVILÉS, A. *et al.* Probabilistic Forecasting of Drought Events Using Markov Chain- and Bayesian Network-Based Models: A Case Study of an Andean Regulated River Basin. *Water*, v. 8, n. 2, 2016. ISSN 2073-4441. DOI: 10.3390/w8020037.

BITTENCOURT, R. P. *et al.* Gerenciamento de riscos operacionais aplicado aos recursos hídricos. *Revista Ibero-Americana de Ciências Ambientais*, v. 12, n. 8, p. 498–511, ago. 2021. DOI: 10.6008/CBPC2179-6858.2021.008.0041. Disponível em: <https://sustenere.inf.br/index.php/rica/article/view/6080>.

CAO, J. *et al.* Multiple Markov Chains for Categorical Drought Prediction on the U.S. Drought Monitor at Weekly Scale. *Journal of Applied Meteorology and Climatology*, American Meteorological Society, Boston MA, USA, v. 62, n. 10, p. 1415–1435, 2023. DOI: 10.1175/JAMC-D-23-0061.1.

FERNÁNDEZ, F. J. *et al.* The economics impacts of long-run droughts: Challenges, gaps, and way forward. *Journal of Environmental Management*, v. 344, p. 118726, 2023. ISSN 0301-4797. DOI: 10.1016/j.jenvman.2023.118726.

HARRIS, C. R. *et al.* Array programming with NumPy. *Nature*, v. 585, p. 357–362, 2020. DOI: 10.1038/s41586-020-2649-2.

HUNTER, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, v. 9, n. 3, p. 90–95, 2007. DOI: 10.1109/MCSE.2007.55.

MCKINNEY, W. Data Structures for Statistical Computing in Python. In: PROCEEDINGS of the 9th Python in Science Conference. [S. l.: s. n.], 2010. p. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

OCHOLA, W. O.; KERKIDES, P. A Markov chain simulation model for predicting critical wet and dry spells in Kenya: analysing rainfall events in the Kano Plains. *Irrigation and Drainage*, v. 52, n. 4, p. 327–342, 2003. DOI: 10.1002/ird.94.

PLOTLY TECHNOLOGIES INC. *Collaborative Data Science*. [S. l.: s. n.], 2015. <https://plotly.com>. Montreal, QC.

ROSS, C. A. *et al.* Evaluating the Ubiquity of Thresholds in Rainfall-Runoff Response Across Contrasting Environments. *Water Resources Research*, v. 57, n. 1, 2021. DOI: 10.1029/2020WR027498.

SONNADARA, D. U. J.; JAYEWARDENE, D. R. A Markov chain probability model to describe wet and dry patterns of weather at Colombo. *Theoretical and Applied Climatology*, v. 119, n. 1-2, p. 333–340, 2015. DOI: 10.1007/s00704-014-1117-z.

TERCINI, J. R. B.; MÉLLO JÚNIOR, A. V. Impact of hydroclimatic changes on the operation of water resources systems: a case study of the Cantareira Water Production System. *RBRH*, Associação Brasileira de Recursos Hídricos, v. 29, e19, 2024. ISSN 2318-0331. DOI: 10.1590/2318-0331.292420230132.

TUNDISI, J. Water resources in the future: problems and solutions. *Estudos Avançados*, v. 22, p. 7–16, dez. 2007.

VAN LOON, A. F. Hydrological drought explained. *WIREs Water*, v. 2, n. 4, p. 359–392, 2015. DOI: 10.1002/wat2.1085.