

IMPACTOS POTENCIAIS DA INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL (XAI) NA GESTÃO DE RECURSOS HÍDRICOS: UMA REVISÃO SISTEMÁTICA DA LITERATURA

Elioenai Wilcesky Tosini Neves ¹; José Anderson do Nascimento Batista ²

Abstract: Effective and sustainable water resource management is fundamental for human and environmental development. In this context, Explainable Artificial Intelligence (XAI) presents itself as a thriving approach to making AI models more transparent and understandable, particularly in critical systems such as water resource management. This study proposes a systematic literature review to analyze the state-of-the-art application of XAI in water resource management, identifying benefits, challenges, and knowledge gaps. The objectives include understanding the application of XAI in water management, synthesizing current knowledge, and discussing benefits and limitations. Thus, based on the PRISMA methodology for systematic reviews, thematic filters were developed, followed by careful reading and categorization by techniques, algorithms, and application areas, complemented by a bibliometric analysis of author keywords using the VOSviewer tool. A total of 314 studies published between 2021 and April 2025 were analyzed, with a predominance of the SHAP technique (in 247 studies) and the XGBoost and Random Forest algorithms. The most frequent themes were water quality, aquifer management, and extreme events. The main reported benefits include increased transparency, reduced subjectivity, greater institutional acceptance, and operational optimization. Among the challenges, the most prominent are methodological concentration, the lack of standardized evaluation metrics, and the need for greater interdisciplinary integration. This study reinforces the potential of XAI as a technical and ethical support for data-driven water governance, contributing to more informed, participatory, and resilient decision-making.

Resumo: A gestão eficaz e sustentável dos recursos hídricos é fundamental para o desenvolvimento humano e ambiental. Nesse contexto, a Inteligência Artificial Explicável (XAI) desponta como uma abordagem promissora para tornar os modelos de IA mais transparentes e compreensíveis, especialmente em sistemas críticos como a gestão de recursos hídricos. Este estudo propõe uma revisão sistemática da literatura para analisar o estado-da-arte do emprego da XAI na gestão de recursos hídricos, identificando benefícios, desafios e lacunas de conhecimento. Os objetivos incluem entender a aplicação da XAI na gestão hídrica, sintetizar o conhecimento atual e discutir benefícios e limitações. Assim, com base no método PRISMA para revisões sistemáticas, foram elaborados filtros temáticos, leitura criteriosa e categorização por técnicas, algoritmos e áreas de aplicação, complementadas por uma análise bibliométrica das palavras-chave autorais com o auxílio da ferramenta VOSviewer. Foram analisados 314 trabalhos publicados entre 2021 e abril de 2025, com predominância da técnica SHAP (em 247 estudos) e dos algoritmos XGBoost e Random Forest. As temáticas mais frequentes foram qualidade da água, gestão de aquíferos e eventos extremos. Os principais benefícios incluem aumento da transparência, redução da subjetividade e otimização operacional. Entre os desafios, destacam-se a concentração metodológica e a necessidade de maior integração interdisciplinar. O estudo reforça o potencial da XAI como suporte técnico e ético para a governança hídrica baseada em dados, contribuindo para decisões mais fundamentadas, participativas e resilientes.

Palavras-chave: Inteligência Artificial Explicável, XAI, Gestão de Recursos Hídricos.

¹) Afiliação: R. Saturnino de Brito, 224 - Cidade Universitária, Campinas - SP, 13083-889. (19) 3521-2308. e224097@dac.unicamp.br

²) Afiliação: R. Saturnino de Brito, 224 - Cidade Universitária, Campinas - SP, 13083-889. (19) 3521-2916. nbatista@unicamp.br

1. INTRODUÇÃO

A gestão sustentável dos recursos hídricos enfrenta desafios intensificados pelas mudanças climáticas, urbanização acelerada e crescente demanda por água [DELANKA-PEDIGE *et al.*, (2021)], sendo o acesso à água segura reconhecido como direito humano e reforçado pelo ODS nº 6 [UN-WATER, (2020)]. Nesse cenário, técnicas de Inteligência Artificial (IA) têm se destacado pela capacidade de processar grandes volumes de dados hidrológicos, gerar previsões e apoiar decisões operacionais em tempo real [DANISH; MOHD, (2022)]; contudo, a natureza opaca de muitos modelos compromete sua aplicabilidade em setores sensíveis, como o da gestão pública e ambiental [INNERARITY, (2021)]. A Inteligência Artificial Explicável (Explainable Artificial Intelligence – XAI) busca superar essa limitação ao fornecer previsões interpretáveis, auditáveis e mais aceitáveis a técnicos, gestores e à sociedade [ANGELOV *et al.*, (2021); RIBEIRO *et al.*, (2016)]. Apesar do interesse crescente, ainda são escassas as sistematizações sobre sua aplicação na gestão hídrica, especialmente quanto às técnicas explicativas mais utilizadas, aos algoritmos predominantes e aos principais benefícios e desafios [CHEN *et al.*, (2023)], o que motiva esta revisão sistemática voltada a consolidar o conhecimento existente e contribuir para práticas mais transparentes, resilientes e baseadas em evidência científica.

2. OBJETIVOS DA PESQUISA

Este estudo tem como objetivo geral analisar sistematicamente o estado da arte sobre a utilização da Inteligência Artificial Explicável (XAI) na gestão dos recursos hídricos, identificando os métodos explicativos mais utilizados, os algoritmos predominantes e os principais benefícios e limitações reportados na literatura acadêmica recente. Especificamente, busca-se: investigar e esclarecer os principais conceitos relacionados à Inteligência Artificial e à Inteligência Artificial Explicável (XAI); identificar as técnicas explicativas mais aplicadas em estudos voltados à gestão hídrica; mapear os algoritmos de aprendizado de máquina mais utilizados, destacando os critérios que orientam sua escolha; sintetizar os benefícios observados com a aplicação da XAI; discutir as limitações e desafios metodológicos enfrentados nas aplicações identificadas; examinar as tendências temporais das publicações sobre o tema, evidenciando o crescimento do interesse acadêmico; e, por fim, apontar lacunas existentes e sugerir direções para futuras pesquisas, com vistas a fortalecer a governança hídrica digital, transparente e baseada em dados robustos.

3. METODOLOGIA

Esta pesquisa adotou o método de Revisão Sistemática da Literatura, conhecido por sua capacidade de sintetizar de maneira estruturada, transparente e replicável o conhecimento científico existente em um determinado campo [TRANFIELD, (2003)].

A pesquisa foi conduzida utilizando a base de dados Scopus, devido à sua abrangência e relevância acadêmica. Os dados utilizados nesta análise estão disponíveis publicamente no Repositório de Dados de Pesquisa da Unicamp (<https://doi.org/10.25824/redu/NMBSHT>). Foram incluídos artigos científicos, livros e trabalhos apresentados em conferências, publicados exclusivamente em inglês, com foco em estudos que relacionassem claramente as aplicações de Inteligência Artificial Explicável (XAI) à gestão hídrica.

A estratégia de busca utilizou operadores booleanos (AND, OR, NOT), combinando termos específicos relacionados à XAI, tais como “explainable artificial intelligence”, “XAI”, “SHAP”, entre outros, com termos diretamente relacionados à gestão hídrica, como “water quality”, “flood”, “groundwater”, “hydrology”, “water resources”, entre outros.

Após a busca inicial, os artigos foram organizados em planilhas para verificação e exclusão de duplicatas. Em seguida, todos os títulos e resumos foram analisados manualmente para garantir

aderência ao tema proposto. Por fim, os artigos selecionados passaram por leitura completa, assegurando rigor na seleção e inclusão apenas daqueles que claramente contribuíram para os objetivos desta revisão.

O processo de extração e organização dos dados incluiu categorização detalhada por temas, tipos de técnicas explicativas, algoritmos utilizados e benefícios identificados. Este procedimento permitiu análises qualitativas e quantitativas aprofundadas, oferecendo uma visão abrangente e sólida sobre o estado atual das aplicações da XAI na gestão hídrica.

4. FUNDAMENTAÇÃO TEÓRICA

4.1. Inteligência Artificial (IA) e Aprendizado de Máquina (ML)

A Inteligência Artificial (IA) consiste em um campo interdisciplinar da ciência da computação dedicado à criação de sistemas capazes de realizar tarefas que tradicionalmente requerem inteligência humana, tais como reconhecimento de padrões, raciocínio lógico, aprendizado e tomada de decisão [SILVA *et al.*, (2021)]. Desde seu surgimento formal na década de 1950, a IA evoluiu significativamente graças ao avanço tecnológico, sobretudo em poder computacional e na abundância de dados disponíveis para treinamento de modelos inteligentes [DANISH; MOHD, (2022)].

Entre as abordagens da IA, destaca-se o Aprendizado de Máquina (Machine Learning - ML), uma subárea que permite aos sistemas computacionais identificar padrões, aprender a partir de grandes volumes de dados e realizar previsões ou decisões sem serem explicitamente programados para tal. Essa capacidade torna o ML especialmente valioso em áreas complexas, como a gestão hídrica, onde os fenômenos naturais são altamente não lineares, imprevisíveis e dinâmicos [HAMIDREZA *et al.*, (2022); GHOBADI; KANG, (2023)].

As técnicas de ML são tradicionalmente classificadas em três categorias principais: aprendizado supervisionado, não supervisionado e semi-supervisionado. O aprendizado supervisionado utiliza dados rotulados para treinar modelos a realizar previsões ou classificações precisas [RYAN HAGGERTY, (2023)]. Já o aprendizado não supervisionado busca identificar padrões ocultos ou relações intrínsecas em dados não rotulados [GHOBADI; KANG, (2023)]. Por fim, o aprendizado semi-supervisionado explora conjuntos que combinam dados rotulados e não rotulados para aperfeiçoar o desempenho preditivo, especialmente em contextos onde a rotulagem completa é inviável ou dispendiosa [YOON, (2023)].

Recentemente, técnicas mais sofisticadas, como Deep Learning (aprendizado profundo), surgiram com força, devido à sua capacidade superior em tratar grandes volumes de dados não estruturados, como imagens ou séries temporais complexas. Exemplos proeminentes são as Redes Neurais Recorrentes (RNN) e Convolucionais (CNN), que revolucionaram previsões em diversas áreas, incluindo a hidrologia, por capturarem dependências espaciais e temporais profundas [LECUN, (2015); CHIDEPUDI, (2023)].

4.2. Inteligência Artificial Explicável (XAI)

Apesar das vantagens técnicas do ML, especialmente do Deep Learning, muitos modelos são considerados "caixas-pretas", isto é, apresentam resultados precisos, porém pouco transparentes quanto às razões por trás das decisões tomadas [DOSHI-VELEZ; KIM, (2017)]. Essa falta de transparência torna-se problemática em áreas críticas como saúde, segurança e meio ambiente, incluindo gestão de recursos hídricos, onde decisões equivocadas podem acarretar consequências graves [GUNNING; AHA, (2019)].

Como resposta, emergiu a Inteligência Artificial Explicável (XAI), uma área dedicada a tornar compreensíveis as decisões dos algoritmos complexos por meio de técnicas que fornecem explicações claras, auditáveis e acessíveis para especialistas e usuários comuns [GILPIN *et al.*, (2018); ANGELOV *et al.*, (2021)]. A XAI busca conciliar o alto desempenho preditivo dos modelos complexos com a exigência ética e regulatória por transparência, contribuindo para maior aceitação institucional e confiança pública [GOODMAN; FLAXMAN, (2017)].

Esses objetivos da XAI estão alinhados aos princípios estabelecidos pelo Instituto Nacional de Padrões e Tecnologia dos Estados Unidos (NIST), que propõe quatro princípios fundamentais para sistemas explicáveis: explicação, relevância, precisão e reconhecimento das limitações do modelo [PHILLIPS *et al.*, (2020)]. Nesse contexto, a explicação refere-se à capacidade do modelo fornecer justificativas claras sobre suas decisões. A relevância demanda que as explicações sejam significativas e compreensíveis para o usuário final. A precisão estabelece que as explicações devem refletir fielmente os mecanismos internos do modelo. Por fim, reconhecer os limites significa que o sistema deve indicar quando suas previsões são incertas ou excedem suas competências [EHSAN *et al.*, (2023); DECK *et al.*, (2023)].

Ao seguir esses princípios, a XAI fortalece três pilares centrais da IA ética destacados por ARRIETA *et al.* (2020): a robustez, ao destacar perturbações adversas potenciais que poderiam alterar a previsão; a justiça, ao auxiliar na garantia de imparcialidade na tomada de decisões, permitindo identificar e corrigir possíveis vieses nos dados de treinamento; e a causalidade, ao permitir a rastreabilidade das variáveis que impactam os resultados.

Para operacionalizar esses conceitos na prática, os métodos Explicativos da XAI são categorizados conforme o momento e abrangência das explicações geradas. Os métodos ante-hoc são modelos naturalmente interpretáveis, como árvores de decisão ou regressões, projetados para serem transparentes desde o início [VILONE; LONGO, (2021)]. Por outro lado, os métodos pós-hoc usam abordagens complementares para explicar decisões feitas por modelos originalmente opacos, como redes neurais e máquinas de vetor de suporte [ISLAM, (2022)].

Já os métodos pós-hoc, destacam-se técnicas agnósticas ao modelo, como SHAP (SHapley Additive exPlanations) e LIME (Local Interpretable Model-agnostic Explanations). O SHAP utiliza a teoria dos jogos para atribuir importâncias às variáveis e explicar decisões de forma robusta, tanto local quanto globalmente [LUNDBERG; LEE, (2017)]. Já o LIME fornece explicações locais intuitivas, geralmente aplicadas em decisões individuais para modelos complexos [RIBEIRO; SINGH; GUESTRIN, (2016)].

5. RESULTADOS

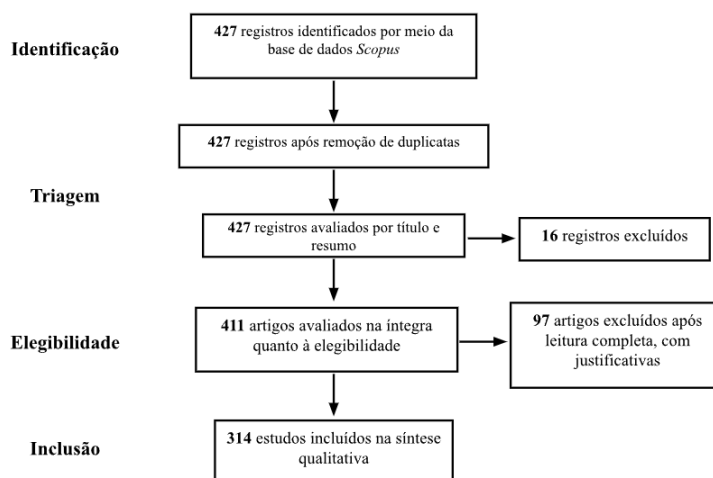
A revisão sistemática identificou e analisou 314 artigos publicados entre 2021 e abril de 2025 na base de dados Scopus (Figura 1), os quais abordam a aplicação da Inteligência Artificial Explicável (XAI) na gestão de recursos hídricos. Os resultados revelam padrões importantes em relação às áreas temáticas de concentração, técnicas explicativas mais utilizadas, algoritmos aplicados e benefícios relatados.

5.1. Análise Temporal:

O crescimento expressivo no número de publicações também merece destaque. Houve um salto considerável no volume de estudos a partir de 2023, sinalizando o amadurecimento da XAI como área emergente e estratégica para a gestão hídrica. Em 2021, apenas 6 artigos atenderam aos critérios da revisão; em 2022, esse número subiu para 25; em 2023, chegou a 43; e, em 2024, saltou para 175 publicações. Até abril de 2025, já haviam sido identificados 65 artigos, mantendo a tendência de expansão contínua. Tal evolução reflete o reconhecimento crescente da importância da

explicabilidade algorítmica para aplicações críticas, sobretudo em um contexto global marcado por exigências normativas, impactos climáticos intensificados e transformação digital dos sistemas de monitoramento e gestão da água [AZEEM *et al.*, (2024)].

Figura 1 – Diagrama de fluxo PRISMA



5.2. Distribuição temática

Em termos de distribuição temática, observou-se uma predominância significativa de estudos voltados à qualidade da água, que totalizaram 118 artigos e se subdividiram em diferentes subtemas: 51 artigos abordaram nutrientes e eutrofização; 45 trataram de processos físico-químicos; 22 enfocaram o tratamento de contaminantes e poluentes, e 19 investigaram a qualidade da água para uso humano e saúde. Esse conjunto representa o maior volume da amostra analisada, refletindo a prioridade conferida à sustentabilidade ambiental e à segurança hídrica, especialmente em contextos urbanos, agrícolas e industriais. Além disso, estudos voltados à gestão de águas subterrâneas somaram 40 publicações, dividindo-se entre qualidade e monitoramento (21 artigos) e processos de recarga e manejo de aquíferos (19 artigos), destacando a crescente atenção dada aos sistemas aquíferos como reservas estratégicas em cenários de escassez. Também foi expressiva a presença de 26 estudos sobre tratamento de águas residuais e esgoto, voltados à otimização operacional de estações de tratamento e ao uso eficiente de recursos em contextos urbanos. A gestão do escoamento superficial e da vazão em bacias hidrográficas apareceu em 32 artigos, enquanto a análise de eventos hidrológicos extremos e riscos associados, como inundações e secas, foi abordada em 24 estudos, evidenciando a preocupação global com os impactos das mudanças climáticas. Outros macrotemas relevantes incluíram o ciclo hidrológico, sedimentos e serviços ecossistêmicos (22 artigos), gestão urbana da água e infraestrutura (18), e uso agrícola, com foco em irrigação e eficiência hídrica (15). Essa concentração temática indica que a aplicação da XAI tem se consolidado prioritariamente em domínios diretamente afetados por fatores climáticos, pressões antrópicas e demandas crescentes por eficiência, resiliência e transparência na governança hídrica.

5.3. Análise das Técnicas XAI utilizadas:

Quanto aos métodos explicativos utilizados nos estudos revisados, observou-se um predomínio claro da técnica SHAP (SHapley Additive Explanations), presente em 247 dos 314 artigos analisados. Essa hegemonia se justifica por sua compatibilidade com modelos de árvore, como XGBoost, Random Forest e LightGBM, bem como por sua capacidade de fornecer explicações locais e globais de forma robusta e visualmente acessível [LUNDBERG; LEE, (2017)]. O SHAP foi empregado em contextos diversos, incluindo previsão de poluentes [DOMINKOVIĆ *et*

al., (2024)], tratamento de efluentes [HOU *et al.*, (2025)], avaliação da qualidade da água subterrânea [GUO *et al.*, (2023)] e modelagem de eventos extremos [HEUER *et al.*, (2024)].

Outras técnicas apareceram em frequência significativamente menor. O Partial Dependence Plot (PDP), com 27 ocorrências, tem sido usado principalmente em análises globais sobre influência de variáveis climáticas [ZHANG *et al.*, (2025)], enquanto o LIME, com 20 estudos, foi mais comum em aplicações com redes neurais e dados espaciais [GHOLAMI *et al.*, (2024)]. Métodos como ICE (Individual Conditional Expectation), ALE (Accumulated Local Effects) e IG (Integrated Gradients) foram empregados de forma complementar em modelos complexos, como LSTM e CNN, permitindo análises mais refinadas em contextos específicos [SHADKANI *et al.*, (2024)].

Apesar da predominância do SHAP, observam-se iniciativas pontuais de diversificação metodológica com técnicas como Deep SHAP, EG (Expected Gradients), Occlusion, Saliency Maps, GAMI-net e EBM (Explainable Boosting Machine). Embora menos frequentes, essas abordagens vêm sendo adaptadas a cenários mais complexos, como análises visuais por imagens satelitais e exigências de interpretabilidade com vistas a fornecer evidência científica para ambientes regulatórios [NEOPHYTIDES *et al.*, (2024)].

5.4. Análise dos Algoritmo de Tarefa Primária utilizados:

No que diz respeito aos algoritmos de aprendizado utilizados como tarefa primária, observou-se forte predominância de modelos baseados em árvores de decisão, notadamente o XGBoost, empregado em 121 estudos, e o Random Forest, presente em 102. A escolha desses algoritmos relaciona-se à sua robustez preditiva, resistência à multicolinearidade e facilidade de integração com técnicas explicativas como SHAP e PDP [CHEN; GUESTRIN, (2016)]. Outros modelos de boosting, como LightGBM e CatBoost, usados em 35 e 24 estudos respectivamente, também figuram com destaque, especialmente em contextos que exigem maior eficiência computacional. Além disso, modelos de redes neurais profundas, como LSTM (Long Short-Term Memory) e CNN (Convolutional Neural Networks), foram largamente utilizados em tarefas de previsão temporal e espacial, sendo identificados em 49 e 31 estudos, respectivamente [HEUER *et al.*, (2024); CHEN *et al.*, (2024)]. A combinação desses algoritmos com técnicas de explicação visual, como Grad-CAM, LayerCAM e SHAP, evidencia uma tendência à hibridização entre modelos preditivos sofisticados e abordagens interpretativas.

5.5. Análise dos Benefícios Associados:

A análise dos estudos revisados demonstra que os benefícios da XAI na gestão dos recursos hídricos vão além do ganho técnico, gerando impactos positivos nas dimensões operacional, institucional e social. Entre os aspectos mais citados estão a maior clareza sobre variáveis-chave, a ampliação da confiança nos modelos, a transparência das decisões e o suporte a políticas públicas mais fundamentadas [HEUER *et al.*, (2024)].

Técnicas como SHAP e PDP permitiram identificar os fatores mais relevantes para eventos extremos, como secas e inundações, contribuindo para interpretações mais acessíveis e estratégias de mitigação mais eficazes [(FENG *et al.*, (2024)]. A possibilidade de justificar previsões com base em variáveis compreensíveis tem favorecido a aceitação prática de modelos por técnicos e gestores, sobretudo em temas como gestão de aquíferos e previsão de qualidade da água subterrânea [JUNG *et al.*, (2024)].

Além disso, a explicabilidade tem facilitado a comunicação com stakeholders não técnicos, aumentando o engajamento em decisões sobre alocação hídrica, priorização de intervenções e monitoramento participativo [MAUSSNER *et al.*, (2025)]. Em contextos operacionais, como saneamento e abastecimento, a XAI tem permitido otimizar processos, prever falhas e reduzir o uso de insumos, promovendo maior eficiência [CHENG *et al.*, (2025)].

Por fim, ao revelar de forma auditável a lógica interna dos modelos, a XAI tem se mostrado essencial para decisões regulatórias e formulação de políticas públicas baseadas em evidências, reforçando os princípios de governança adaptativa e justiça ambiental [CHENG *et al.*, (2024)]. Assim, os benefícios relatados confirmam que a XAI contribui não apenas para a interpretação técnica de modelos, mas para uma gestão hídrica mais transparente, participativa e orientada por dados [GULSHAD *et al.*, (2024)].

Figura 2 – Mapa de Coocorrência de Palavras-chave Autorais no VOSviewer

Para aprofundar a compreensão das relações temáticas identificadas na revisão sistemática, foi realizada uma análise bibliométrica utilizando o software VOSviewer, por meio da técnica de coocorrência das palavras-chave autorais dos artigos selecionados. Para a análise de coocorrência, foram selecionadas as palavras-chave com, no mínimo, três ocorrências na base de dados, resultando em um total de 77 termos analisados a partir de um conjunto inicial de 1.000. O mapa gerado (Figura 2) evidencia uma rede fortemente estruturada em torno de métodos explicativos e algoritmos específicos, destacando-se um núcleo altamente centralizado em termos como "machine learning", "SHAP", "XGBoost" e "explainable artificial intelligence". Essa centralidade ilustra não apenas a importância técnica desses elementos, mas também sua ampla aplicabilidade em diversos contextos ambientais, tais como a gestão da qualidade da água, a previsão de eventos extremos, a gestão de aquíferos e o tratamento de águas residuais, reforçando os achados da revisão sistemática.

pesquisa, sugerindo oportunidades para estudos futuros focados na exploração dessas fronteiras metodológicas e temáticas emergentes.

6. LIMITAÇÕES E DESAFIOS

Embora os avanços na aplicação da Inteligência Artificial Explicável (XAI) na gestão de recursos hídricos sejam notáveis, a literatura revisada evidencia um conjunto relevante de limitações e desafios que ainda precisam ser enfrentados para que a incorporação desses modelos ocorra de forma plena, ética e eficaz nos sistemas decisórios ambientais.

Uma das limitações mais evidentes está na concentração excessiva em torno de uma única técnica explicativa: o SHAP. Apesar de sua robustez, adaptabilidade a diversos algoritmos e capacidade de produzir explicações tanto locais quanto globais, o predomínio quase absoluto dessa abordagem — identificada em 247 dos 314 estudos — levanta preocupações quanto à diversidade metodológica e à possível dependência de um único paradigma interpretativo [HEUER *et al.*, (2024)]. Em muitos casos, outras técnicas, como LIME, ICE, PDP e métodos visuais como Grad-CAM, poderiam enriquecer a análise, oferecendo perspectivas complementares sobre o funcionamento dos modelos. A limitação reside, portanto, não na técnica em si, mas na ausência de pluralidade metodológica que favoreça abordagens híbridas e adaptadas ao contexto da aplicação.

7. CONSIDERAÇÕES FINAIS

A presente revisão sistemática confirma o avanço da Inteligência Artificial Explicável (XAI) como ferramenta estratégica na gestão dos recursos hídricos, especialmente diante da intensificação de eventos extremos e da demanda por decisões públicas mais eficientes. A análise de 314 estudos identificou a predominância do método SHAP, o uso recorrente de algoritmos como XGBoost, Random Forest e LSTM, e a aplicação concentrada em temas como qualidade da água, eventos extremos e águas subterrâneas [CHEN *et al.*, (2023)]. Verificou-se que a explicabilidade algorítmica contribui para aumentar a confiança nos modelos, promover maior auditabilidade e facilitar a integração entre ciência, gestão e sociedade [MAUSSNER *et al.*, (2025)]. No entanto, desafios persistem, como a concentração metodológica, a escassez de dados em determinadas regiões, a ausência de métricas padronizadas e as barreiras interdisciplinares, indicando a necessidade de consolidar um ecossistema de XAI mais diverso, interoperável e orientado por valores, capaz de sustentar uma governança hídrica mais ética, transparente e resiliente.

REFERÊNCIAS

1. ANGELOV, P. P.; SOARES, E. A.; JIANG, R.; ARNOLD, N. I.; ATKINSON, P. M. (2021). “Explainable artificial intelligence: an analytical review”. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, v. 11, n. 5, e1424.
2. ARRIETA, A.B.; DÍAZ-RODRÍGUEZ, N.; DEL SER, J.; BENNETOT, A.; TABIK, S.; BARBADO, A.; GARCIA, S.; GIL-LOPEZ, S.; MOLINA, D.; BENJAMINS, R.; CHATILA, R.; HERRERA, F. (2020). “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI”. Information Fusion, 58, pp. 82–115.
3. AZEEM, M.; DEV, S. (2024). “A performance and interpretability assessment of machine learning models for rainfall prediction in the Republic of Ireland”. Decision Analytics Journal, v. 12.
4. CHEN, T.; GUESTRIN, C. (2016). “XGBoost: A scalable tree boosting system”. In: Anais do 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco. New York: ACM, pp. 785–794.

5. CHENG, X.; KE, X.; WEI, T.; CHEN, A.; PANG, Z.; QIN, Z.; CHEN, Y.; TIAN, Y.; WANG, Q.; WU, H.; QIU, G.; WEI, C. (2025). "Optimized strategies for achieving nitrosative accumulation based on a comprehensive understanding of nitrogen form transformation". *Chemical Engineering Journal*, v. 503.
6. CHIDEPUDI, S. K. R.; MASSEI, N.; JARDANI, A.; HENRIOT, A.; ALLIER, D.; BAULON, L. A. (2023). "A wavelet-assisted deep learning approach for simulating groundwater levels affected by low-frequency variability". *Science of The Total Environment*, v. 865.
7. DANISH, M. (2022). "Artificial intelligence and machine learning in water resources engineering". In: ZAKWAN, M. et al. (Orgs.). *Current Directions in Water Scarcity Research*. Elsevier, v. 7, pp. 3–14.
8. DECK, L.; SCHOEFFER, J.; DE-ARTEAGA, M.; KÜHL, N. (2023). "A critical survey on fairness benefits of explainable AI". *arXiv preprint*, arXiv:2310.13007.
9. DELANKA-PEDIGE, H. M.; MUNASINGHE-ARACHCHIGE, S. P.; ABEYSIRIWARDANA-ARACHCHIGE, I. S.; NIRMALAKHANDAN, N. (2021). "Wastewater infrastructure for sustainable cities: Assessment based on UN sustainable development goals (SDGs)". *International Journal of Sustainable Development and World Ecology*.
10. DOSHI-VELEZ, F.; KIM, B. (2017). "Towards a rigorous science of interpretable machine learning". *arXiv preprint*, arXiv:1702.08608.
11. EHSAN, U.; SAHA, K.; DE CHOUDHURY, M.; RIEDL, M. O. (2023). "Charting the sociotechnical gap in explainable AI: A framework to address the gap in XAI". *arXiv preprint*, arXiv:2302.00799.
12. GHOBADI, F.; KANG, D. (2023). "Application of machine learning in water resources management: A systematic literature review". *Water*, v. 15, n. 4.
13. GHOLAMI, H.; JALALI, M.; REZAEI, M.; MOHAMADIFAR, A.; SONG, Y.; LI, Y.; WANG, Y.; NIU, B.; OMIDVAR, E.; KASKAOUTIS, D. (2024). "An explainable integrated machine learning model for mapping soil erosion by wind and water in a catchment with three desiccated lakes". *Aeolian Research*, v. 67–69.
14. GILPIN, L. H.; BAU, D.; YUAN, B. Z.; BAJWA, A.; SPECTER, M.; KAGAL, L. (2018). "Explaining explanations: An overview of interpretability of machine learning". In: *Anais do 5th IEEE International Conference on Data Science and Advanced Analytics*, Turim. Piscataway: IEEE, pp. 80–89.
15. GOODMAN, B.; FLAXMAN, S. (2017). "European Union regulations on algorithmic decision-making and a 'right to explanation'". *AI Magazine*, v. 38, n. 3, pp. 50–57.
16. GULSHAD, K.; YASEEN, A.; SZYDŁOWSKI, M. (2024). "From data to decision: Interpretable machine learning for predicting flood susceptibility in Gdańsk, Poland". *Remote Sensing*, v. 16.
17. GUO, X.; XIONG, H.; LI, H.; GUI, X.; HU, X.; LI, Y.; CUI, H.; QIU, Y.; ZHANG, F.; MA, C. (2023). "Designing dynamic groundwater management strategies through a composite groundwater vulnerability model: Integrating human-related parameters into the DRASTIC model using LightGBM regression and SHAP analysis". *Environmental Research*, v. 236.
18. HAGGERTY, R.; SUN, J.; YU, H.; LI, Y. (2023). "Application of machine learning in groundwater quality modeling: A comprehensive review". *Water Research*, v. 233.
19. HAMIDREZA, M.; SADEGHI, M.; MALLAKPOUR, I.; JAHROMI, M. N.; POURGHASEMI, H. R. (2022). "Application of machine learning algorithms in hydrology". In: *Computers in Earth and Environmental Sciences*. Elsevier.

20. HEUER, H.; SCHWABE, M.; GENTINE, P.; GIORGETTA, M. A.; EYRING, V. (2024). "Interpretable multiscale machine learning-based parameterizations of convection for ICON". *Journal of Advances in Modeling Earth Systems*, v. 16.
21. HOU, P.; LIU, S.; HU, D.; ZHANG, J.; LIANG, J.; LIU, H.; ZHANG, J.; ZHANG, G. (2025). "Predicting biomass conversion and COD removal in wastewater treatment by phototrophic bacteria with interpretable machine learning". *Journal of Environmental Management*, v. 375.
22. INNERARITY, D. (2021). Making the black box society transparent. *AI & Society*, v. 36, p. 975–981, 2021.
23. NEOPHYTIDES, S. P.; MAVROVOUNIOTIS, M.; ELIADES, M.; BACHOFER, F.; HADJIMITSIS, D. G. (2024). "Tree-based explainable clustering for drought severity predictions in United States". In: *Anais do SPIE – The International Society for Optical Engineering*, 2024.
24. PHILLIPS, P. J.; HAHN, C. A.; FONTANA, P. C.; YATES, A. N.; GREENE, K.; BRONIATOWSKI, D. A.; PRZYBOCKI, M. A. (2020). "Four principles of explainable artificial intelligence". NIST Interagency/Internal Report 8312, Gaithersburg: National Institute of Standards and Technology.
25. RIBEIRO, M. T.; SINGH, S.; GUESTIN, C. (2016). "Why Should I Trust You?: Explaining the Predictions of Any Classifier". In: *Anais do 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco. New York: ACM, pp. 1135–1144.
26. SELVARAJU, R. R.; COGSWELL, M.; DAS, A.; VEDANTAM, R.; PARIKH, D.; BATRA, D. (2020). "Grad-CAM: Visual explanations from deep networks via gradient-based localization". *International Journal of Computer Vision*, v. 128, pp. 336–359.
27. SILVA, F. W.; SILVA, S. F.; RABÊLO, S. O. (2021). "Tendências no uso da inteligência artificial e sua influência na requalificação da força de trabalho no setor público". *Cadernos de Prospecção*, v. 14, n. 3, p. 824.
28. TRANFIELD, D.; DENYER, D.; SMART, P. (2003). "Towards a methodology for developing evidence-informed management knowledge by means of systematic review". *British Journal of Management*, v. 14, n. 3, pp. 207–222.
29. UN-WATER. (2020). UN World Water Development Report 2020: Water and Climate Change. Disponível em: <https://www.unwater.org/news/un-world-water-development-report-2020-%E2%80%98water-and-climate-change%E2%80%99>.
30. VILONE, G.; LONGO, L. (2021). "Classification of explainable artificial intelligence methods through their output formats". *Machine Learning and Knowledge Extraction*, v. 3, pp. 615–661.
31. WANG, S.; PENG, H.; HU, Q.; JIANG, M. (2022). "Analysis of runoff generation driving factors based on hydrological model and interpretable machine learning method". *Journal of Hydrology: Regional Studies*, v. 42.
32. YOON, S.; AHN, K.-H. (2023). "Semi-supervised learning approach to improve the predictability of data-driven rainfall-runoff model in hydrological data-sparse regions". *Hydrology and Earth System Sciences Discussions*.
33. ZHANG, H.; WANG, G.; LI, S.; CABRAL, P. (2025). "Understanding evapotranspiration driving mechanisms in China with explainable machine learning algorithms". *International Journal of Climatology*.