

A TRANSFORMAÇÃO DIGITAL COMO RESPOSTA ÀS AFLUÊNCIAS INDEVIDAS: APLICAÇÃO DE DIGITAL TWINS E MACHINE LEARNING NO SISTEMA DE SANEAMENTO DO VALE DO AVE

António PEREIRA¹

RESUMO

As aflúências indevidas representam uma das principais ameaças à eficiência, sustentabilidade e resiliência dos sistemas de saneamento. Resultantes de infiltrações ou ligações indevidas de águas pluviais e/ou industriais nas redes de drenagem de águas residuais. Estas aflúências aumentam significativamente os volumes transportados e tratados, com impacto direto em custos energéticos, capacidade hidráulica e risco ambiental.

A operação de SSAR apresenta múltiplos desafios, incluindo garantir um transporte eficaz e seguro desde as habitações até à restituição dos efluentes tratados no meio ambiente. Durante a ocorrência de eventos extremos obriga ainda à mobilização de meios técnicos e humanos que permitam uma resposta eficaz, sempre que surjam obstruções ou se ultrapasse a capacidade de drenagem instalada.

No âmbito do sistema de saneamento do Vale do Ave, foi desenvolvido um projeto pioneiro que recorre à integração de dados de monitorização, Digital Twins (DT) e Machine Learning (ML) como ferramentas de digitalização e otimização operacional. O objetivo foi criar um sistema inteligente de apoio à decisão, capaz de identificar, prever e mitigar os efeitos das aflúências indevidas, promovendo uma gestão mais eficiente e orientada por dados. Foram instalados sensores de nível e velocidade em pontos estratégicos da rede, permitindo recolher dados em tempo real. Através destes dados, foi possível construir modelos de ML supervisionados (com destaque para o algoritmo XGBoost), calibrados para prever o comportamento hidráulico da rede em diferentes condições meteorológicas.

Os modelos permitiram distinguir padrões de caudal associados a escoamento doméstico e pluvial, detetar anomalias e quantificar volumes anómalos. Simultaneamente, o DT foi desenvolvido como uma réplica digital dinâmica da rede, integrada com dados históricos, sensores e previsões, capaz de simular cenários operacionais e antecipar sobrecargas. A implementação deste sistema trouxe resultados concretos: identificação de zonas críticas com elevado contributo de águas parasitárias, apoio à definição de prioridades de intervenção e otimização da operação de estações de tratamento, promovendo a eficiência energética. Este caso demonstra como a transformação digital, quando alavancada por ciência de dados e modelação avançada, pode ser um catalisador para sistemas de saneamento mais resilientes, sustentáveis e preparados para os desafios climáticos e operacionais. A experiência no Vale do Ave evidencia também o potencial de replicação desta abordagem noutros contextos urbanos, reforçando o papel da inovação na melhoria contínua da gestão da água.

Palavras-chave: Aflúências indevidas, *digital twins*, *machine learning*, realidade aumentada, sustentabilidade

1 - Doutor em Engenharia Civil, Especialista em Hidroinformática e Machine Learning, Responsável de Soluções de Performance, Aquapor e Professor Auxiliar (conv.) do Departamento de Engenharia Civil da Universidade do Minho, antoniopereira@aquaporservicos.pt

1. INTRODUÇÃO

Os sistemas de saneamento urbano desempenham um papel central na garantia da saúde pública, da qualidade ambiental e da sustentabilidade das cidades. A sua função principal é recolher, transportar e tratar as águas residuais domésticas, industriais e, em alguns casos, pluviais, assegurando que o efluente final cumpre os padrões legais de descarga e não compromete os ecossistemas recetores. No entanto, estes sistemas enfrentam desafios estruturais e operacionais persistentes, que comprometem a sua eficiência e resiliência. Entre estes, as afluências indevidas — isto é, a entrada de caudais adicionais de origem pluvial, industriais ou por infiltração em redes destinadas apenas ao transporte de águas residuais — destacam-se como uma das principais fragilidades.

As afluências indevidas resultam de múltiplas fontes: infiltrações difusas em coletores envelhecidos ou danificados, ligações clandestinas de águas pluviais e/ou industriais à rede de saneamento, ou mesmo transbordos de sarjetas durante eventos extremos de precipitação. O impacto é significativo: caudais muito superiores ao projetado para a rede e estações de tratamento, aumento do consumo energético associado ao bombeamento e tratamento de volumes adicionais de água, e maior risco de descargas não tratadas para o meio recetor. Em Portugal, este problema é particularmente relevante devido à idade avançada de grande parte da infraestrutura, à pressão urbanística desordenada e às alterações climáticas, que intensificam a frequência e a intensidade de eventos extremos.

Com efeito, a problemática das AI aos sistemas de saneamento de águas residuais (SSAR) tem sido frequentemente negligenciada. As entidades gestoras destes sistemas têm-se focado essencialmente na expansão e, conseqüentemente, no aumento de cobertura do serviço. No entanto, o envelhecimento das infraestruturas e sua natural deterioração têm conduzido a um aumento significativo dos volumes infiltrados, refletindo-se negativamente no desempenho dos sistemas. Em segundo o RASAARP (2023), registaram-se cerca de 270 milhões de m³ em afluências indevidas, um problema, que nos dias de hoje já deveria ser, manifestamente, inferior.

No entanto, nos últimos anos, algumas entidades gestoras têm vindo a investir em diagnósticos e planos de mitigação de afluências indevidas, através de campanhas de monitorização, modelação hidráulica e reabilitação de troços críticos. Contudo, estas abordagens, embora úteis, revelam-se limitadas pela sua natureza reativa e pontual. Tipicamente, os problemas só são identificados após a ocorrência de falhas, descargas ou custos excessivos, e a capacidade de previsão permanece reduzida. A gestão moderna de sistemas complexos exige ferramentas que permitam passar de uma lógica reativa para uma lógica proativa e preditiva.

Contudo, o combate às AI é uma das prioridades patentes na nova DARU e que está alinhada com o PEENSAARP2030 que visa o incremento na monitorização de forma a obter um conhecimento mais efetivo das condições de escoamento, nomeadamente em cenários ou condições de incerteza o que revela ser um fator potenciador de melhoria.

No período pandémico de 2020, que despoletou uma Hecatombe Sanitária ficou demonstrado que a tecnologia revelou ser prática intrínseca ao quotidiano e com potencial para desmaterializar operações num tempo muito curto. A revolução tecnológica já estava a acontecer, mas de uma forma silenciosa.

No âmbito dos avanços tecnológicos mais relevantes no domínio do digital, foi, indubitavelmente, a chegada da IA. Não sendo, de todo, um domínio novo, é, no entanto, incipiente no que a Recursos Hídricos diz respeito, principalmente em sistemas de águas residuais.

É neste contexto que a transformação digital do setor da água surge como uma oportunidade concreta e necessária. A evolução das tecnologias de Internet of Things (IoT), Machine Learning (ML), Big Data e Realidade Aumentada (RA) abre novas possibilidades para gerir os sistemas de saneamento de forma

mais inteligente, resiliente e eficiente. Entre estas tecnologias, destacam-se os Gémeos Digitais (Digital Twins, DT, Figura 1), que constituem representações digitais dinâmicas das infraestruturas físicas, continuamente alimentadas por dados em tempo real. Um DT não é apenas uma maquete digital estática: é um modelo vivo, que replica o comportamento do sistema, permite simular cenários, prever falhas e testar estratégias de operação sem risco.

Quando combinados com algoritmos de ML, os DT tornam-se ainda mais poderosos: em vez de apenas refletirem o estado atual, conseguem antecipar comportamentos futuros, identificando padrões e anomalias em séries temporais complexas. Por sua vez, a integração de Realidade Aumentada acrescenta uma camada essencial de operacionalização, levando a inteligência do DT diretamente ao terreno. A RA permite que operadores visualizem no local, sobrepostos à infraestrutura física, dados em tempo real, previsões de caudal, mapas subterrâneos da rede e instruções de intervenção. Esta combinação reduz drasticamente o tempo de diagnóstico, aumenta a segurança em operações de campo e melhora a confiança nas decisões.

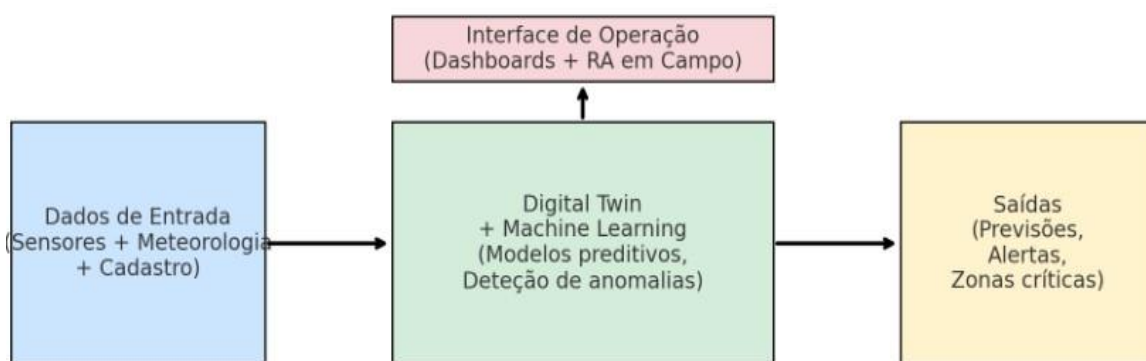


Figura 1 - Esquema concetual do DT.

A investigação e aplicação de tecnologias digitais no setor da água têm crescido na última década. (Pinho e Vieira, 2014; Joseph-Duran et al., 2014 e Lowe et al., 2016). Existe um objetivo premente em desenvolver e aplicar metodologias inovadoras de gestão operacional de recursos hídricos, especialmente utilizando ferramentas de controlo em tempo real (RTC) e de digitalização dos sistemas (DT) (Fuchs e Beeneken, 2005; Puig, 2009; Pinho et al., 2011; Lowe et al., 2016; Van Daal-Rombouts et al., 2016; Shishegar et al., 2018; Pereira et al., 2018; Pereira et al., 2019). Organizações internacionais como a International Water Association (IWA) e a Water Europe têm promovido ativamente a digitalização como um dos pilares da transição do setor, através da visão de “Digital Water”. Esta visão assenta em três eixos fundamentais: (i) integração de dados provenientes de sensores e plataformas IoT; (ii) utilização de modelos dinâmicos e de Machine Learning para previsão e apoio à decisão; e (iii) interfaces avançadas para interação com utilizadores e operadores, incluindo RA.

Em termos internacionais, vários exemplos ilustram o potencial dos Gémeos Digitais. Em Singapura, a entidade PUB implementou o Smart Water Grid, um sistema que integra monitorização em larga escala com modelos digitais para gerir o abastecimento e a drenagem urbana. Este sistema permite prever consumos, detetar fugas e antecipar riscos de inundação. Nos Países Baixos, projetos-piloto em Roterdão e Copenhaga utilizam DTs para modelar a resposta das redes pluviais a eventos extremos, permitindo ajustar em tempo real o funcionamento de estações de bombagem e reservatórios. Em Barcelona, a Agbar/SUEZ tem investido em DTs para integrar dados de qualidade da água, sensores em rede e previsões meteorológicas, com vista à otimização da drenagem e do abastecimento.

No contexto do saneamento, os exemplos ainda são mais limitados. Embora haja projetos relevantes em Espanha e Alemanha, a maior parte das aplicações documentadas incide sobre abastecimento de água ou controlo de cheias. Esta lacuna justifica a importância de estudos como o presente, que exploram a aplicação de DTs especificamente à mitigação de afluências indevidas.

Relativamente ao Machine Learning, a literatura reporta múltiplas aplicações: previsão de caudais em redes de drenagem, deteção de anomalias em séries temporais de qualidade da água, estimativa de contribuições pluviais com base em dados de condutividade e precipitação, e previsão de volumes afluentes em ETAR. Algoritmos como XGBoost, Random Forests e Support Vector Machines apresentam bons resultados, com capacidade de lidar com séries temporais irregulares e variáveis altamente correlacionadas. Trabalhos recentes demonstram que, quando integrados em DTs, estes algoritmos permitem melhorar significativamente a capacidade de previsão, reduzindo erros e aumentando a confiança dos operadores.

A Realidade Aumentada tem sido mais explorada noutros setores, como eletricidade, gás e indústria automóvel. Nestes contextos, a RA é usada para apoiar operações de manutenção, fornecer instruções visuais em tempo real e aumentar a segurança dos técnicos. No setor da água, aplicações iniciais em abastecimento já mostram ganhos concretos, nomeadamente na rapidez de localização de válvulas e na visualização de condutas subterrâneas. Contudo, no saneamento urbano, a RA encontra-se ainda numa fase embrionária, sendo esta investigação uma das primeiras a propor a sua aplicação de forma sistemática.

Em Portugal, os avanços mais relevantes na gestão operacional de sistemas de drenagem incluem o Plano Geral de Drenagem de Lisboa (PGDL), que já combina monitorização intensiva, modelação e previsão meteorológica (esta parte menos desenvolvida), e o SIDVA, onde a TRATAVE, em colaboração com a Universidade do Minho, desenvolveu e aplicou um sistema de suporte à decisão (SSD) com módulo de DT. Este módulo já está a ser preparado para ser testado em diferentes casos de estudo, evidenciando ganhos claros na previsão de caudais e mitigação de afluências indevidas. O módulo de RA, por sua vez, encontra-se em fase de protótipo, sendo alvo de ensaios exploratórios para avaliar a sua aplicabilidade e usabilidade.

Conclui-se que existe uma tendência clara de integração tecnológica, mas também revela uma lacuna significativa: há poucos exemplos concretos que combinem de forma operacional os três elementos — DT, ML e RA — num ciclo completo de aquisição de dados, previsão, apoio à decisão e execução assistida no terreno. O presente artigo procura contribuir para colmatar esta lacuna, apresentando um caso de estudo estruturado no contexto português.

O presente artigo pretende apresentar e discutir esta experiência, estruturando-se da seguinte forma: a Secção 2 detalha a metodologia adotada para a construção e integração do módulo DT+ML+RA; a Secção 3 apresenta os resultados obtidos e discussão em termos de previsão de caudais, identificação de afluências indevidas ganhos operacionais; a Secção 4 as conclusões.

2. METODOLOGIA DE INVESTIGAÇÃO

2.1 Caracterização da Área em Estudo

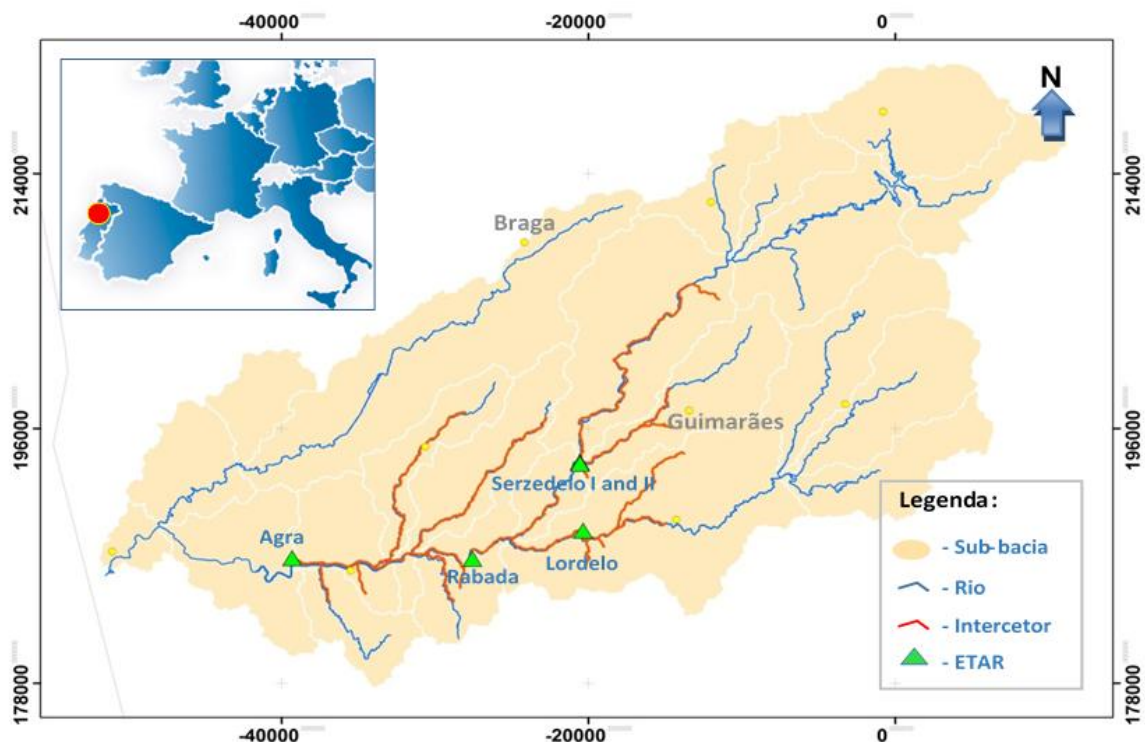


Figura 2 - Localização da bacia do rio Ave. Fonte: Autor, 2025.

O âmbito geográfico (Figura 2) deste trabalho é a bacia hidrográfica do rio Ave (figura 1), que inclui uma densa rede de linhas de água delimitadas a Norte pela bacia do rio Cávado, a Este pela bacia do rio Douro e a Sul pelas bacias dos rios Leça e Douro. A bacia possui uma área de 1469 km², sendo que 247 km² e 380 km² correspondem às áreas das bacias de seus principais afluentes, os rios Este e Vizela.

Em relação ao clima, a região é caracterizada por verões muito quentes e levemente húmidos e invernos frios e chuvosos, apresentando precipitações médias anuais superiores a 1500 mm. A intensa precipitação que ocorre nos períodos chuvosos, juntamente com as características orográficas, a constituição do solo das bacias e sua permeabilidade, tornam esta zona propícia a inundações e cheias.

O Rio Ave nasce na serra da Cabreira, concelho de Vieira do Minho, a 1200 m de altitude, percorrendo 94 km no sentido nascente-poente até à foz no Oceano Atlântico junto à cidade de Vila do Conde. Esta massa de água engloba, total ou parcialmente, as áreas de jurisdição de 16 concelhos. Os concelhos de Guimarães e Vila Nova de Famalicão ficam totalmente integrados na bacia hidrográfica do rio Ave, enquanto Braga, Barcelos, Vizela, Celorico de Basto, Fafe, Póvoa do Lanhoso e Vieira do Minho, todos do distrito de Braga e ainda parte dos concelhos de Felgueiras, Lousada, Maia, Paços de Ferreira, Póvoa do Varzim, Santo Tirso e Vila do Conde do distrito do Porto.

Trata-se de uma região caracterizada por boas capacidades de recarga natural provenientes dos cursos de água ou por infiltração da água da chuva. No caso dos aquíferos formados por estratos distintos a recarga acontece a partir da drenagem de camadas adjacentes. A paisagem ribeirinha do Rio Ave é ainda extremamente valorizada pelo grande número de pontes romanas e medievais que o atravessa,

pela existência de elementos com interesse arquitetónico e paisagístico, bem como pelos inúmeros exemplares de azenhas e açudes.

A bacia hidrográfica é monitorizada por uma rede meteorológica que serve para acompanhar o ramo aéreo do ciclo hidrológico (precipitação, temperatura, humidade do ar, direção e velocidade do vento, insolação, radiação solar, evaporação e evapotranspiração), sendo constituída por 59 estações, das quais cinco são climatológicas e as restantes são udométricas (neste trabalho utilizam-se 35 destas estações). Existe ainda uma rede hidrométrica utilizada para monitorizar o nível hidrométrico, a partir do qual são recolhidos dados de escoamento (caudal, nível, velocidade), sendo possível acompanhar de forma mais eficaz o comportamento dos rios. São monitorizados os níveis freáticos, para posterior classificação do estado quantitativo e da condutividade e concentrações de poluentes para determinação do estado químico.

Embora a rede de monitorização seja extensa muitas das estações encontram-se desativadas e outras em fase de manutenção, incluindo apenas dados históricos. Por outro lado, o desenvolvimento de um sistema de suporte à decisão necessita de dados disponíveis em tempo real com frequências de registo iguais ou inferiores à horária.

Assim, instalaram-se quatro estações meteorológicas em locais distintos da bacia, três equipamentos para monitorização de caudal em diferentes trechos dos rios Ave e Vizela e diversos equipamentos de medição da altura de escoamento e velocidades nos intercetores de drenagem de águas residuais (figura 3). Estes sensores permitem acompanhar a evolução dos caudais recolhidos, incluindo os que resultam de ligações indevidas de águas pluviais à rede de drenagem de águas residuais. Foram ainda configurados procedimentos que permitem processar os dados medidos pela rede de monitorização instalada, sendo a informação disponibilizada em tempo real.

2.2 Componentes principais do sistema

Os dados utilizados neste trabalho foram processados no âmbito do desenvolvimento de um sistema de apoio à decisão (SAD) que deu, entretanto, origem a um digital twin (DT) para a gestão operacional do sistema de drenagem e tratamento localizado no Vale do Ave. Foi utilizada uma plataforma desenvolvida para o efeito (AQUAFLOW - módulo Waste) que para além dos dados de base (componente de informação) está a ser utilizada para implementar uma componente de modelação que permita gerar previsões com base em estados atualizados dos sistemas simulados recorrendo-se a técnicas de assimilação de dados. Foram inseridos os dados históricos disponíveis no SNIRH e os dados de quatro estações meteorológicas instalados no decurso do projeto. Foram ainda utilizados os dados de Radar, fornecidos pelo servidor THREDDS (Thematic Realtime Environmental Distributed Data Service) da Meteogalicia.

Os dados são armazenados numa base de dados criadas em PostgreSQL que alimentam o sistema. O sistema apresenta quatro componentes principais: (i) componente de informação que inclui todos os dados de monitorização, características das infraestruturas e os dados relevantes de gestão operacional; (ii) construção do Gémeo Digital (DT); (iii) desenvolvimento de modelos de Machine learning (ML) e iv) Aplicação de Realidade Aumentada (RA).

Estas componentes foram concebidas de forma integrada, assegurando a interoperabilidade entre dados, modelos e interfaces, de modo a criar um ciclo contínuo de observação–previsão–decisão–ação.

Monitorização e Aquisição de Dados

A monitorização (Figura 3) constitui a base de qualquer DT. No caso em análise, foram instalados diferentes tipos de sensores ao longo da rede de drenagem, nomeadamente:

- Sensores de nível: baseados em pressão hidrostática e radar, com resolução milimétrica, utilizados para registar variações rápidas em pontos críticos.
- Medidores de caudal: de tecnologia ultrassónica e eletromagnética, aplicados em coletores principais, permitindo estimar o caudal instantâneo e acumulado.
- Medidores de velocidade: sensores de velocidade baseados em tecnologia Doppler ultrassónica, adequados à medição em condutas parcialmente cheias ou canais abertos.
- Sondas de condutividade elétrica: fundamentais para distinguir entre caudais domésticos (alta condutividade) e pluviais (baixa condutividade).
- Pluviómetros locais: distribuídos pela bacia de drenagem, complementados por dados de radar meteorológico.

A escolha e distribuição dos sensores teve em conta:

- A necessidade de identificar zonas críticas de aflúências indevidas;
- A possibilidade de calibração cruzada entre diferentes tipos de dados (ex.: nível + velocidade + caudal);
- A redundância em locais estratégicos para assegurar continuidade de medição em caso de falha.

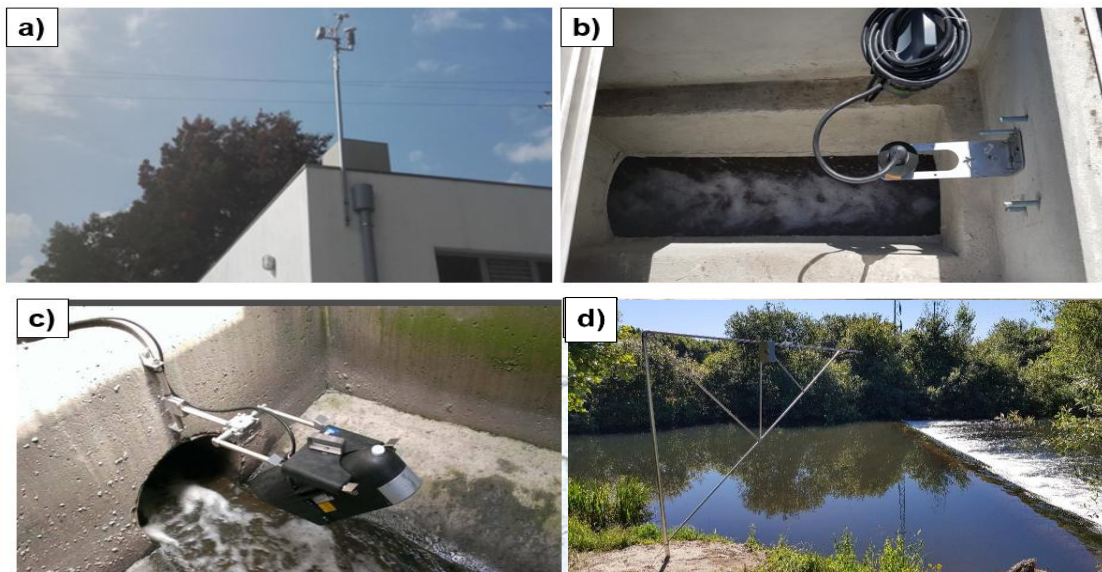


Figura 3 - Equipamento de monitorização: a) estação meteorológica PCE, b) sensor ultrassónico Sofrel LT US com acoplamento de sensor de velocidade de escoamento (VRAD), c) equipamento *doppler Laserflow* e d) estrutura com instalação de sensor Sofrel LT US em estação fluvial. Fonte: Autor, 2025.

Transmissão de Dados

Os sensores foram ligados a módulos de telemetria com comunicação redundante:

- LoRaWAN (baixo consumo, adequado para longas distâncias e zonas rurais), 4G/5G (maior largura de banda, usado em zonas urbanas). A taxa de amostragem variou entre 2 minutos

(eventos críticos) e 5 minutos (monitorização de base). Todos os dados foram enviados para uma plataforma central de armazenamento em nuvem, com capacidade de validação automática e integração em tempo real.

Custos de Instrumentação

Os custos médios registados foram:

- Conjuntos de sensores de nível + velocidade + condutividade: 3.500 € / conjunto;
- Sensores de nível simples: 700 € / unidade;
- Custos de comunicação e manutenção anual: cerca de 2000€.

Estes valores permitiram elaborar um plano de instrumentação sustentável e escalável.

Construção do Gémeo Digital

O Gémeo Digital foi concebido como uma plataforma interoperável, com as seguintes camadas (Figura 4):

- Camada física: representação do sistema de drenagem (coletores, estações elevatórias, ETAR), construída a partir do cadastro GIS existente.
- Camada de dados: integração de medições IoT, séries históricas, dados meteorológicos e características da rede.
- Camada de simulação: modelos hidráulicos baseados em EPA-SWMM, calibrados com as medições obtidas, e preparados para simulação em tempo quase real.
- Camada de decisão: módulos analíticos e dashboards (figura 5), que permitem avaliar cenários e propor estratégias de operação.

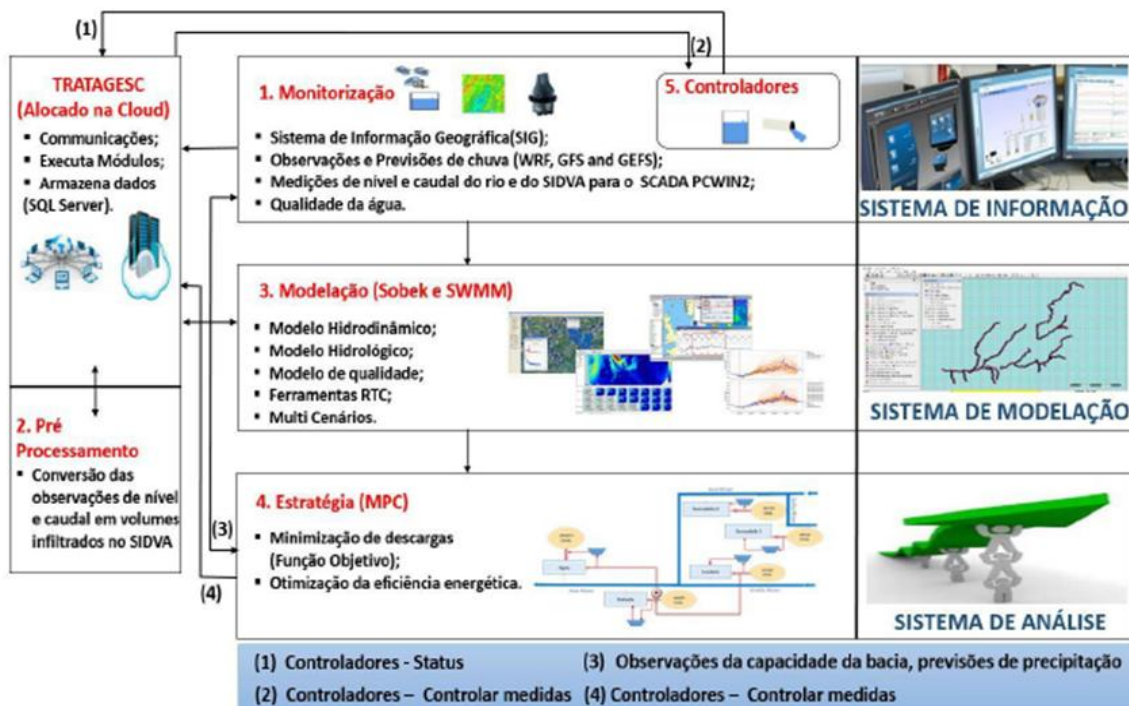


Figura 4 - Componentes do DT.



Figura 5 - Dashboard desenvolvido em DT com KPIs de caudal, nível, velocidade e previsões meteorológicas.

Calibração e Validação

Foram realizadas campanhas de calibração com séries temporais dos últimos 10 anos, em condições de tempo seco e húmido. O modelo calibrado foi validado em eventos extremos, com boa concordância entre valores observados e simulados.

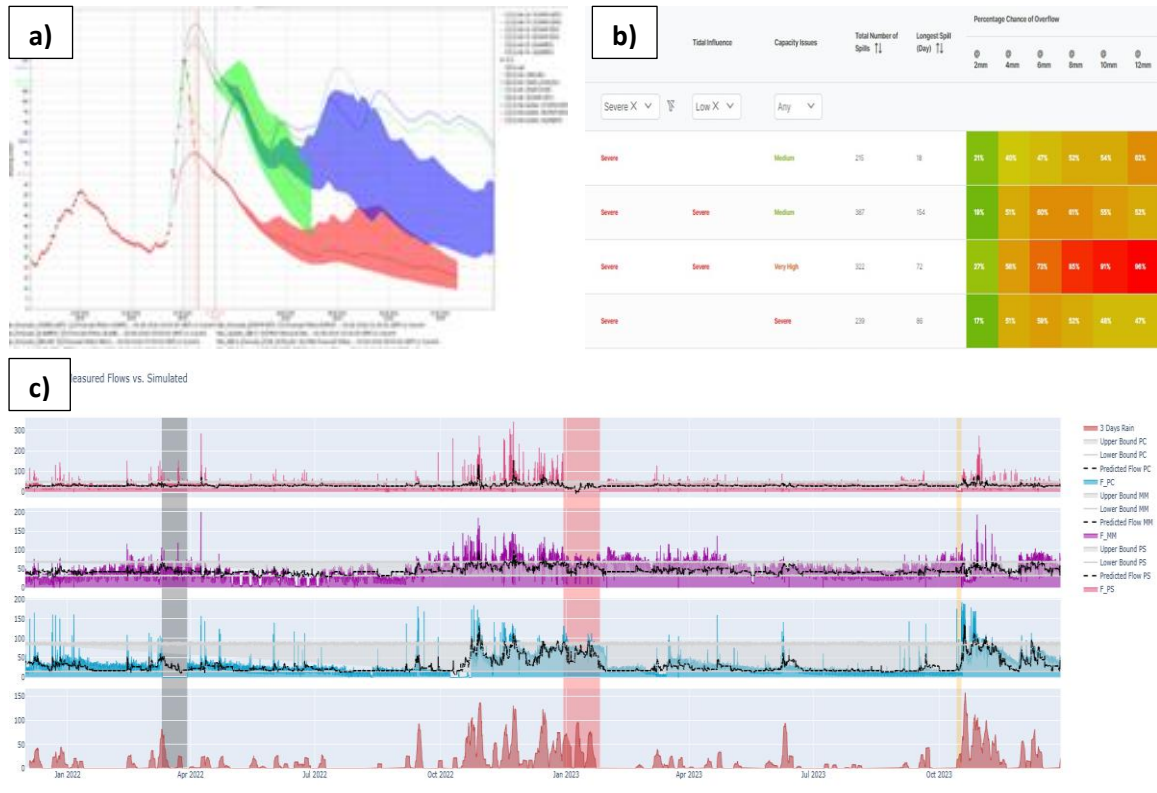


Figura 6 - Apresentação de resultados.

A Figura 6 a) apresenta a previsão de caudais para um ponto específico (neste caso, o IG do Ave), considerando diferentes modelos de previsão meteorológica. As previsões foram realizadas para horizontes temporais de 48 h, 96 h e 120 h, configurando um conjunto de múltiplos cenários ('what-if scenarios'). A Figura 6 b) mostra a probabilidade de ocorrência de uma determinada anomalia em diferentes microbacias, para eventos meteorológicos caracterizados por intensidades de precipitação de 2 mm/h, 4 mm/h, 6 mm/h, 8 mm/h, 10 mm/h e 12 mm/h. Por fim, a Figura 6 c) apresenta as séries temporais em três bacias (Ave, Nespereira e Vizela), com a respetiva previsão obtida através do modelo XGBoost (linha a preto) e os limites de significância representados a cinzento.

Algoritmos Utilizados

- XGBoost: escolhido pela sua elevada capacidade preditiva e robustez em séries temporais complexas.
- Random Forests: utilizados como modelo comparativo, pela sua interoperabilidade.
- Redes Neurais Artificiais (ANNs): testadas em fase exploratória, mas menos eficazes devido à dimensão limitada da base de dados.

Foram analisados outros modelos, contudo os suprarreferidos foram os que obtiveram melhores resultados.

Variáveis Explicativas

As variáveis utilizadas incluíram:

- Caudal e nível a montante;
- Condutividade elétrica (para diferenciar contributos pluviais/domésticos);
- Intensidade e acumulação de precipitação;
- Hora do dia e dia da semana;
- Características geométricas dos coletores.
- Avaliação de Desempenho
- Os modelos foram avaliados com:
 - RMSE (Root Mean Squared Error);
 - MAE (Mean Absolute Error);
 - R^2 (Determination Coefficient).

Resultados típicos indicaram $R^2 = 0,85-0,90$, demonstrando boa capacidade de previsão.

Aplicação de Realidade Aumentada

A Realidade Aumentada foi concebida como interface final entre o DT e o operador no terreno.

Desenvolvimento Tecnológico

Plataforma utilizada: Unity 3D com AR Foundation (compatível com ARKit/ARCore).

Dispositivos: tablets e óculos de RA (ex.: Microsoft HoloLens em fase exploratória).

Funcionalidades principais:

Visualização da rede subterrânea sobreposta ao ambiente real;
Acesso a medições em tempo real (caudal, velocidade, nível, condutividade);

Alertas visuais codificados por cores (verde, amarelo, vermelho);
 Simulação de cenários de intervenção.

Estado Atual

O módulo de RA encontra-se em protótipo, com apoio científico da Universidade do Minho. Testes preliminares foram realizados em pontos críticos da rede do Vale do Ave, com feedback positivo das equipas operacionais, sobretudo em termos de rapidez de diagnóstico e segurança em espaço confinado.

2.3 Processamento e tratamento de dados

Um dos maiores produtos deste estudo é, indubitavelmente, a disponibilidade de uma vasta quantidade de informação. Nos últimos 10 anos criamos um histórico extremamente representativo e fidedigno do sistema, com uma elevada resolução temporal e espacial. Tal feito, tem sido perentório na qualidade dos resultados dos modelos desenvolvidos, nomeadamente nos seguintes. O trabalho de investigação e desenvolvimento está afeto à incorporação de novas bibliotecas para manipulação de dados (Figura 7). Utilizamos as seguintes, que são programáveis no Python: O NumPy; Matplotlib; Pandas; Seaborn e Scikit Learn. O Scikit Learn é uma das várias bibliotecas de ML (que é um ramo da IA) utilizada no Python, e que permite a implementação de diversos modelos, nomeadamente para estabelecer previsões.

<<<<Neste trabalho apresentar-se-ão resultados de modelos regressores.

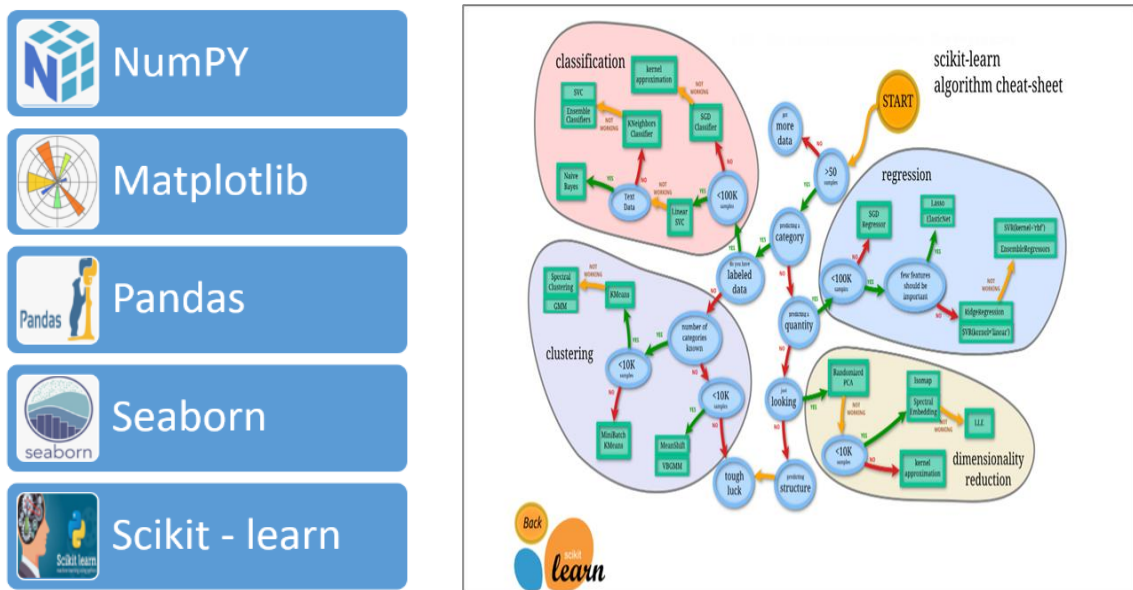


Figura 7 - Bibliotecas utilizadas.

Após configuração dos procedimentos é possível a geração de informação que se transforma numa interpretação mais acessível para os operadores do sistema, possibilitando uma visualização gráfica praticamente sem limites (Figura 8).

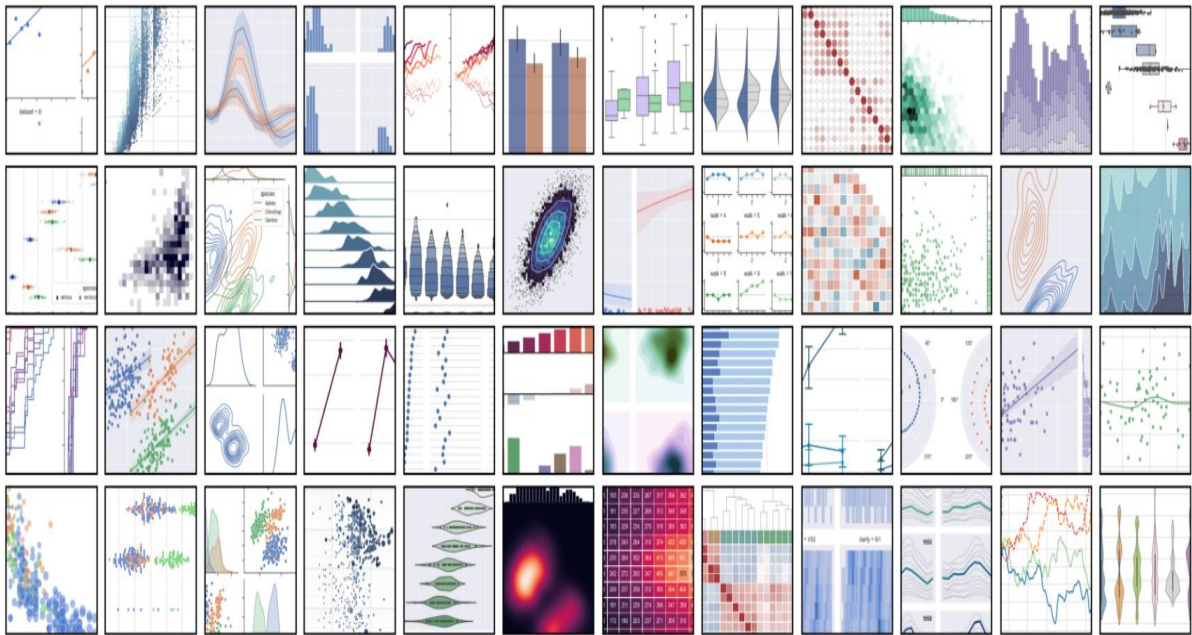


Figura 8 - Possibilidades gráficas.

Os modelos de machine learning utilizados/ testados (Figura 9) para treino foram os seguintes: Dummy Regressor (DR), Multiple Regressor (MR), Support Vector Machines Regressor (SVR), K-Nearest Neighbour Regressor (KNR), Decision Tree Regressor (DTR), Random Forest Regressor (RFR), LightGBM Regressor (LGR) e XGBoost. Todos os modelos pertencem à biblioteca scikit-learn, com exceção do LightGBM, pertencente à biblioteca LGBM, e do XGBoost.

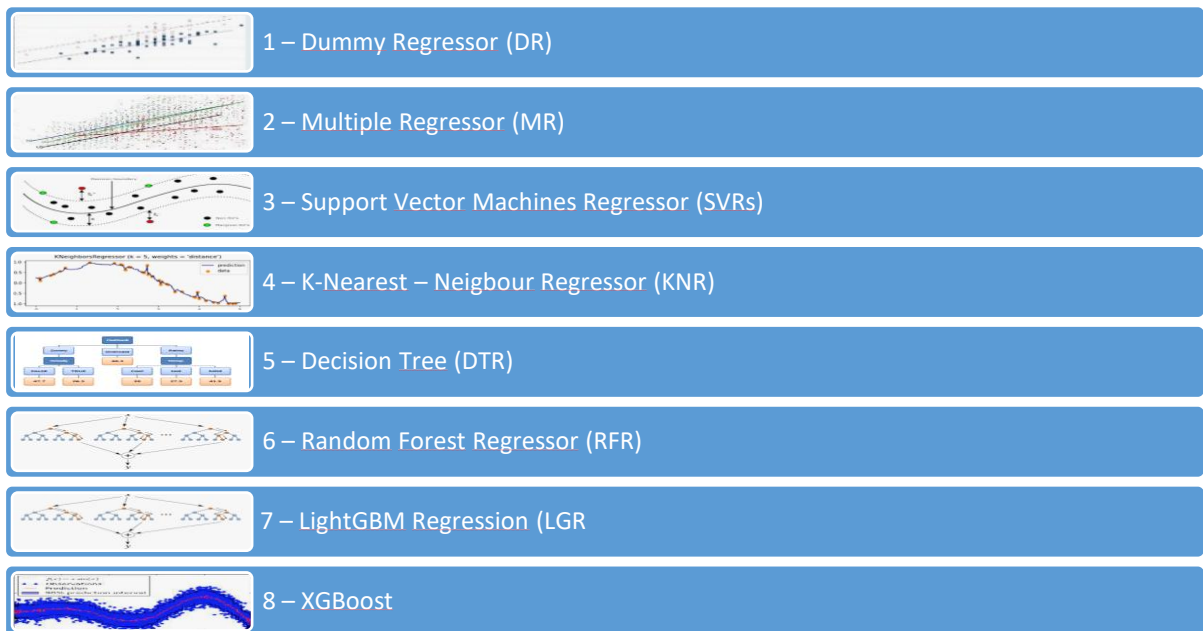


Figura 9 - Modelos de ML testados.

3. RESULTADOS E DISCUSSÃO

3.1. Afluências indevidas ao sistema de drenagem

Um dos problemas que afeta a operação do sistema de drenagem de águas residuais domésticas e industriais da bacia do rio Ave é a afluência indevida de caudais de águas pluviais que entram no sistema ou através de infiltrações nas infraestruturas ou através de ligações indevidas de redes de águas pluviais à rede de esgotos e também ilícitos por parte de clientes industriais de grande envergadura.

Tendo por base os dados de precipitação e caudais medidos com resolução horária no decénio de 2015 a 2024, num dos intercetores, é possível estimar padrões de caudais diários para dias secos e dias com ocorrência de precipitação, conforme apresentado na figura 10.

Padrões horários por dia da semana foram também obtidos (figura 11) e ainda correlações de caudais entre diferentes intercetores, avaliando o impacto da precipitação nesses mesmos caudais através de análise gráfica em 2.5D (figura 12).

Como se pode constatar os valores dos caudais transportados neste intercetor têm um incremento significativo em períodos em que ocorre precipitação. Praticamente duplicam o seu valor durante os dias de chuva. Este aumento é bastante problemático, uma vez que todas as ETAR foram concebidas para receber efluentes domésticos e industriais com volumes pouco significativos de águas pluviais infiltradas e como tal, não existindo capacidade de tratamento do volume total afluente, parte deste volume é descarregado para meio recetor sem tratamento. Por outro lado, a aumento das afluências implica que a rede entre em pressão originando derrames não controlados através das tampas das caixas de visita.

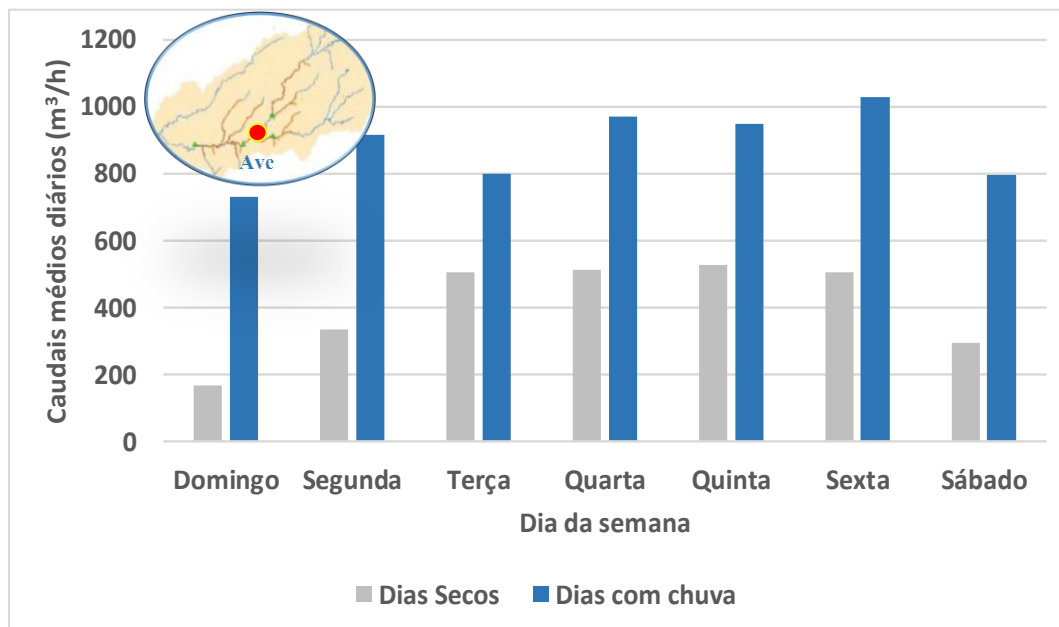


Figura 10 - Caudais médios diários no intercetor do Ave para dias secos e dias com precipitação nos diferentes dias da semana. Fonte: Autor, 2024.

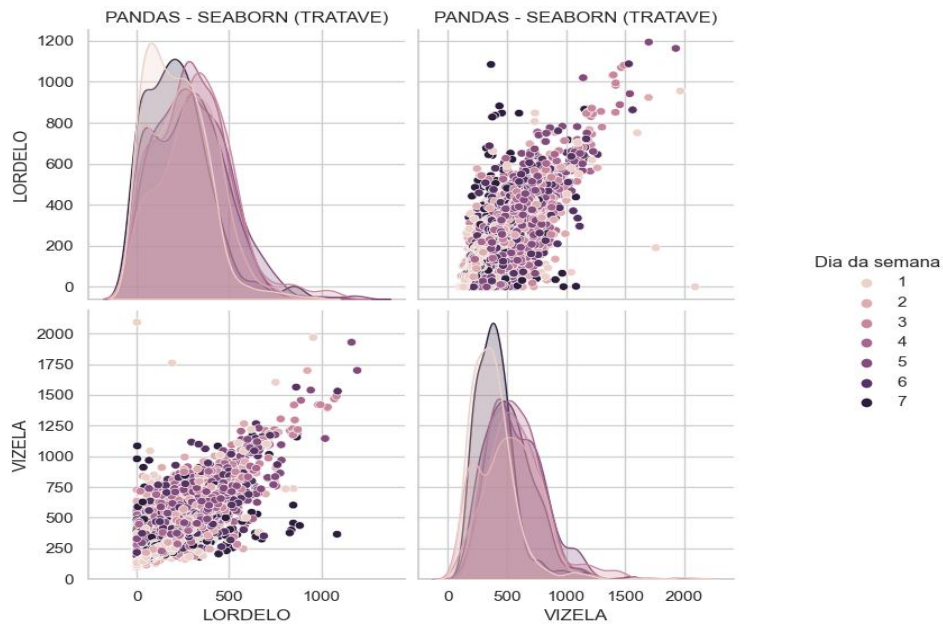


Figura 11 - Exemplos de geração de padrões de caudais horários e por dia da semana (Domingo 1 a Sábado 7) para tempo seco. Fonte: Autor, 2024.

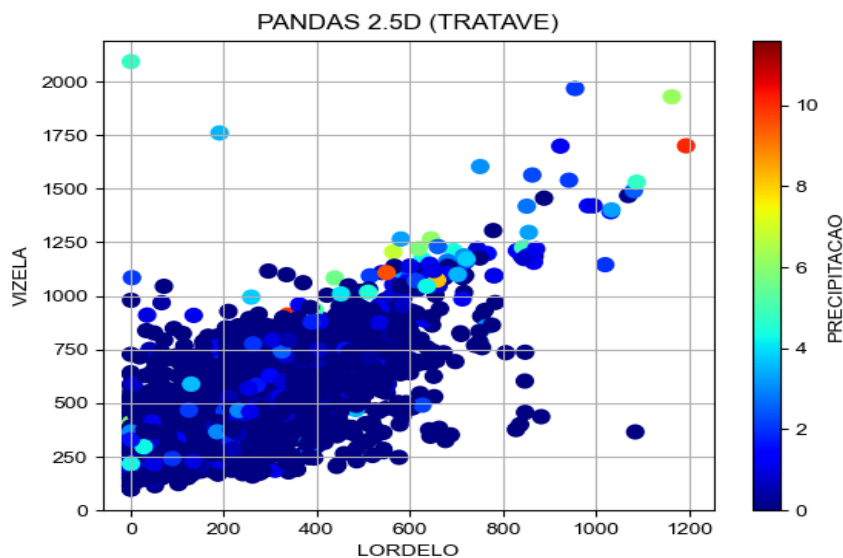


Figura 12 - Geração de gráfico 2.5D relacionando 2 sistemas intercetores e a respetiva precipitação, Fonte: Autor, 2024.

Com a disponibilização do sistema de apoio à decisão a operação do sistema tornou-se mais eficiente. De uma operação reativa, tendo por base a comunicação à posteriori de derrames ao longo da rede, passou-se a uma operação preventiva, em que com base na previsão das precipitações se adotam medidas para minorar e controlar as descargas que ocorrem na rede em períodos de chuva intensa. Os sistemas SCADA (caudais) e FEWS (previsões de precipitação) são utilizados para emissão de alertas. Estes alertas foram configurados através de ferramentas *Python*, utilizando 8 modelos de regressão de *machine learning* (*Dummy Regressor (DR)*, *Linear Regression (LR)*, *Support Vectorial Machines Regressor (SVR)*, *KNeighborsRegressor (KNNR)*, *Decision Tree Regressor (DTR)*, *Random Forest Regressor (RFR)*, *XGboost* e o *LightGBM Regression (LGR)*) ou seja, estes modelos, após o respetivo treino/teste, permitiram a produção de estimativas futuras. As métricas estatísticas para avaliação do desempenho

dos modelos foram: coeficiente de determinação (**R²**), Média do erro absoluto (**MEA**) e erro quadrático médio (**EQM**). No quadro 1 apresentam-se os resultados das métricas estatísticas dos 7 modelos.

Quadro 1- Resultados das métricas estatísticas

Modelo	R ²	MEA	EQM
DR	0.5585	6.5616	240.8178
LR	0.7216	8.4513	213.3312
SVR	0.6171	6.5527	125.0679
KNR	0.7109	6.7052	86.5423
DTR	0.8868	4.5213	22.3214
RFR	0.9187	3.2213	27.3512
XGboost	0.9493	1.721	19.1712
LGR	0.8566	5.1658	77.2314

Foi considerado um histórico de 10 anos, conforme já referido, disponibilizado pela vasta rede de monitorização na tentativa de compreender o impacto futuro. Todos os modelos estimaram assertivamente “*overflows*” nos dias representados na figura 13. Não obstante, o RFR e o XGboost foram os que obtiveram melhor desempenho.

A título exemplificativo apresentam-se na figura 13 registos de precipitações e caudais no interceptor do Ave para o mês de outubro de 2023. Durante este mês foram emitidos seis alertas (dias 13, 14, 16, 19, 29 e 30). Os alertas foram emitidos com uma antecedência de 48 horas e confirmados com a previsão a 24 horas, tendo por base valores de intensidade de precipitação superiores a 5 mm/h. Todos os alertas foram verdadeiros conforme se comprova pelas medições de precipitação através do radar.

A partir destes valores de intensidade de precipitação o sistema entra em pressão e o caudal ultrapassa a capacidade máxima de transporte estimada em 500 m³/h, a partir do modelo hidrodinâmico da rede.

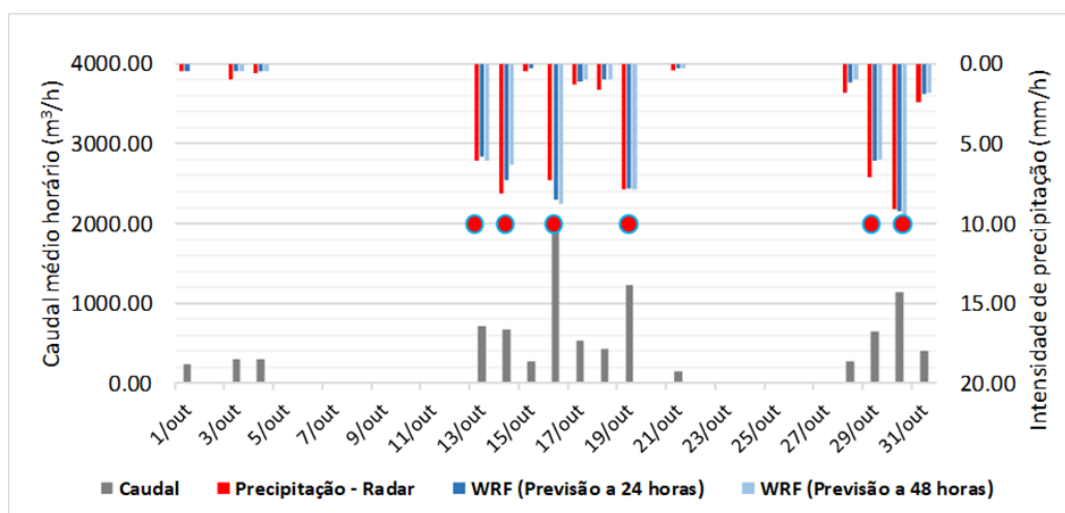


Figura 13 - Registos utilizados para emissão de alertas (círculos vermelhos) incluindo as previsões WRF de precipitação, de precipitações do radar meteorológico e de caudal durante o mês de outubro de 2023 no interceptor do Ave. Fonte: Autor, 2023.

3.2. Localização de pontos sensíveis a derrames através de modelo hidrodinâmico

A correção das AI na rede de drenagem e a estimativa de volumes pluviais infiltrados fornecem as condições necessárias para simular adequadamente as descargas e os níveis de água ao longo da rede de interceptores usando o modelo hidrodinâmico. Este modelo foi aplicado para analisar as câmaras de visita mais vulneráveis a derrames durante eventos de chuva intensa e decidir os locais mais adequados para instalar descargas controladas, evitando descargas em áreas urbanas.

O desempenho do modelo foi quase perfeito (coeficiente de determinação de 0,95) na simulação dos níveis de água, numa simulação de um mês com dados com resolução horária.

Simulações envolvendo outros períodos correspondentes a eventos de precipitação mais intensa foram realizadas para identificar os interceptores/ câmaras de visita mais vulneráveis (figura 14 a)) e ainda locais alternativos para descargas controladas (figura 14 b)) também foram testados no modelo. Essas alternativas mostraram uma melhora significativa nas condições de descarga. Portanto, tanques temporários de armazenamento (figura 14 b)) podem ser uma solução viável, assim como descargas controladas em meio hídrico, sendo para esta necessária, monitorização qualitativa como incremento à monitorização quantitativa. Para a calibração do modelo da rede, foram analisados dados de monitorização de forma a estimar os coeficientes de rugosidade das tubagens. Esses dados incluíram medições simultâneas de velocidades de escoamento e níveis de água. Foram calibrados valores de coeficientes de rugosidade para 15 locais distintos da rede e ainda relação entre níveis, velocidades e caudais (Quadro 2).

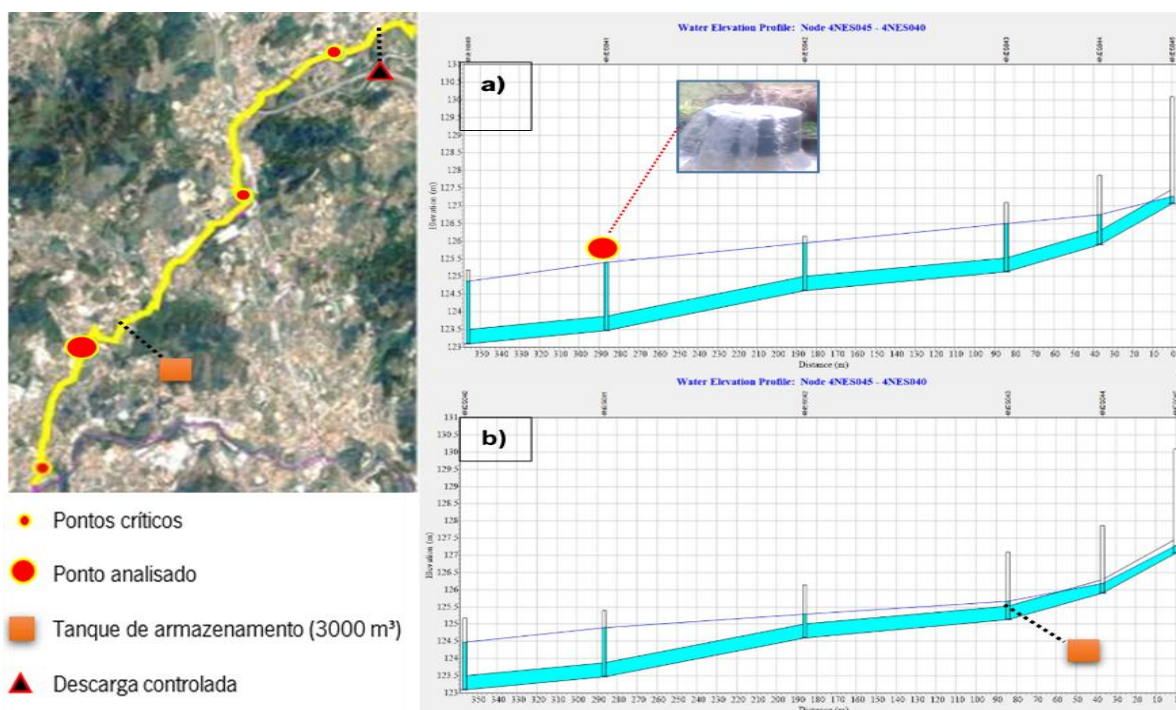


Figura 14 - Simulação hidrodinâmica - Nespereira: a) simulação em cenário de 9mm ocorridos em uma hora – ‘status quo’, identificando um ponto de derrame e b) dimensionamento de um tanque de armazenamento, evitando derrames.

São recorrentes eventos de precipitação intensos (6-10mm) em curto espaço temporal (1h) nos dias de hoje. Neste sentido e conforme figura precedente, em pontos de derrame recorrente, é possível mitigar este efeito com criação de um interceptor inteligente, no sentido de conhecer os pontos críticos

e criar pontos de descarga controlada e/ ou tanques de armazenamento. Quando estes fenómenos surgem o caudal perante um evento de precipitação duplica o seu valor, ou seja, metade do caudal drenado é água de origem pluvial, portanto, é perfeitamente viável, avaliando alguns parâmetros de qualidade que se façam descargas controladas onde o custo ambiental seria eventualmente bastante reduzido (avaliando a capacidade de autodepuração das linhas de água em função dos parâmetros de descarga). Um segundo cenário (figura 14 b)) é o dimensionamento de tanques de armazenamento. No caso do Intercetor de Nespereira para cenários de 9mm/h durante um período de 6h um tanque com capacidade de 3000 m³ seria uma solução ao problema e de fácil implementação no local sugerido na figura. Não obstante, o dimensionamento destes órgãos, devido aos grandes volumes infiltrados (sem possibilidade de execução em local apropriado) e ainda ao custo associado de execução nem sempre é viável. Com efeito, as descargas controladas, poderão ser uma boa solução e já se encontram a ser implementadas em diversos pontos do SIDVA, com a devida monitorização (tempo real), conhecimento e autorização da entidade legisladora, portanto a APA (agência portuguesa do ambiente).

Quadro 2 - Modelo hidrodinâmico da rede – Coeficiente de rugosidade e relação entre nível de água e caudal estimado.

Estação de monitorização (Id)	Coeficiente de rugosidade (Manning) (m ^{1/3} s)	Nível (x) - Caudal (Q) relação entre variáveis $Q = ax^2 + bx$		R ² (-)
		x (m) - Q (m ³ /h)		
		a	b	
1	0.0123	7422.50	-303.08	0.8585
2	0.0132	7492.60	-375.86	0.9916
3	0.0113	23040.00	150.12	0.9971
4	0.0115	-958.22	2129.70	0.9109
5	0.0112	5659.60	710.19	0.9868
6	0.0134	9239.80	596.94	0.9987
7	0.0128	4289.60	394.01	0.9966
8	0.0130	20167.00	-737.24	0.9705
9	0.0132	7058.10	698.10	0.9246
10	0.0130	15806.00	-356.05	0.9902
11	0.0120	16400.00	156.42	0.9737
12	0.0136	3111.30	-545.39	0.9839
13	0.0125	4133.90	525.56	0.9799
14	0.0130	4448.20	-66.12	0.9027
15	0.0128	5891.40	631.52	0.9186

3.3. Detecção de ligações ilícitas

A metodologia adotada para estimar AI revelou os padrões típicos de descargas horárias ao longo da rede de drenagem. Esses padrões também permitem identificar desvios significativos causados por falhas nas infraestruturas de drenagem ou ligações ilícitas. As principais fontes tóxicas na bacia do rio Ave são as indústrias têxteis que exigem grandes volumes de água para os processos industriais de tingimento. Antes da instalação de equipamentos de monitorização ao longo da rede, só era possível avaliar o volume total de águas residuais nas ETAR, conforme já se verificou anteriormente. Comparações sistemáticas dos caudais integrais e a soma das contribuições individuais de clientes

(indústrias e municípios) permitiram detetar que os volumes tratados eram maiores do que os faturados. No entanto, a localização e identificação de ligações indevidas era um processo extremamente complexo. Para aumentar o desempenho do sistema em termos de deteção de ligações ilegais, foi realizada uma normalização dos perfis de descargas industriais ao longo dos dias da semana, juntamente com balanços hídricos redundantes relativos ao incremento de sensores de monitorização instalados em diferentes locais da rede de drenagem. Essa abordagem permitiu fazer uma vigilância mais assertiva.

A Figura 15 representa a normalização dos padrões típicos de caudal horário para os diferentes dias da semana. É perceptível na primeira anomalia (figura 15 a)) a variação do caudal entre descarga ilegal (traço cheio) e descarga permitida (traço pontilhado), onde um volume horário não autorizado de aproximadamente 100 m³ foi detetado. Após a análise dos padrões horários típicos, foi realizada uma inspeção e detetado um desvio ilegal do efluente numa câmara de visita a jusante, não passando pelo medidor de caudal. A figura 15 b) mostra outra anomalia que ocorreu durante os domingos e segundas-feiras pela manhã, após inspeção verificou-se que a empresa em questão esvaziava o tanque de armazenamento (5.000m³) durante o fim-de-semana, desviando esses volumes para uma câmara de visita a jusante debelando assim a medição de caudal. Esta metodologia baseada na medição on-line de descargas permite a identificação de várias fraudes. Após a correção das três ligações ilegais (figura 13), houve um aumento no volume de efluentes faturados de 45,81%, o que representa 5,81% dos volumes não faturados de toda a rede. Durante o período de instalação do SAD (último quadriénio), procedeu-se à correção de cerca de vinte anomalias, permitindo uma diminuição de caudais não faturados na ordem dos 16%, resultando num aumento extremamente representativo, no que concerne ao desempenho do sistema. Toda esta análise encontra-se implementada no DT o que permite a geração de alertas em tempo real, suportando a equipa técnica na decisão.

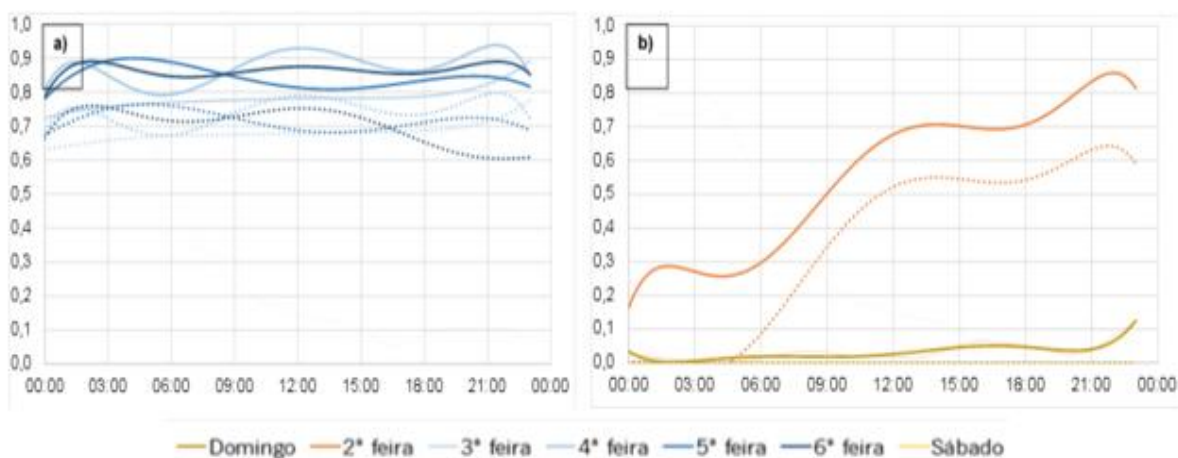


Figura 15 - Deteção de ligações ilícitas através de padrões horários normalizados para os diferentes dias da semana.

Impacto Energético

Um dos principais benefícios do DT foi a otimização do consumo energético. Ao prever picos de caudal e ajustar o acionamento das estações elevatórias, foi possível evitar bombagens em simultâneo ou em horários de maior custo energético. Em média, registou-se uma redução de ≈15% no consumo energético em comparação com a operação tradicional.

Ganhos Operacionais com Realidade Aumentada

A aplicação de RA, embora ainda em protótipo, é possível estimar ganhos imediatos:

- Redução média de 30% no tempo de intervenção em situações de anomalia.
- Maior segurança em operações em espaços confinados, uma vez que os operadores tinham acesso imediato a informação crítica sem necessidade de consultar relatórios ou comunicar com a central.
- Melhoria da formação de equipas, através da simulação de cenários em ambiente de RA.

A Figura 15 apresenta como os sistemas podem tornar-se reais através da RA, podendo o operador aceder ao histórico de um determinado ponto da rede e, perceber em tempo real, se podemos estar perante uma anomalia ou não.

Sem RA: operações mais demoradas, dependentes de comunicação e consulta de plantas.

Com RA: operações mais rápidas, com dados e instruções diretamente sobrepostos ao campo de visão.

Estes resultados foram reforçados por inquéritos às equipas, que referiram maior confiança e clareza na execução das tarefas.

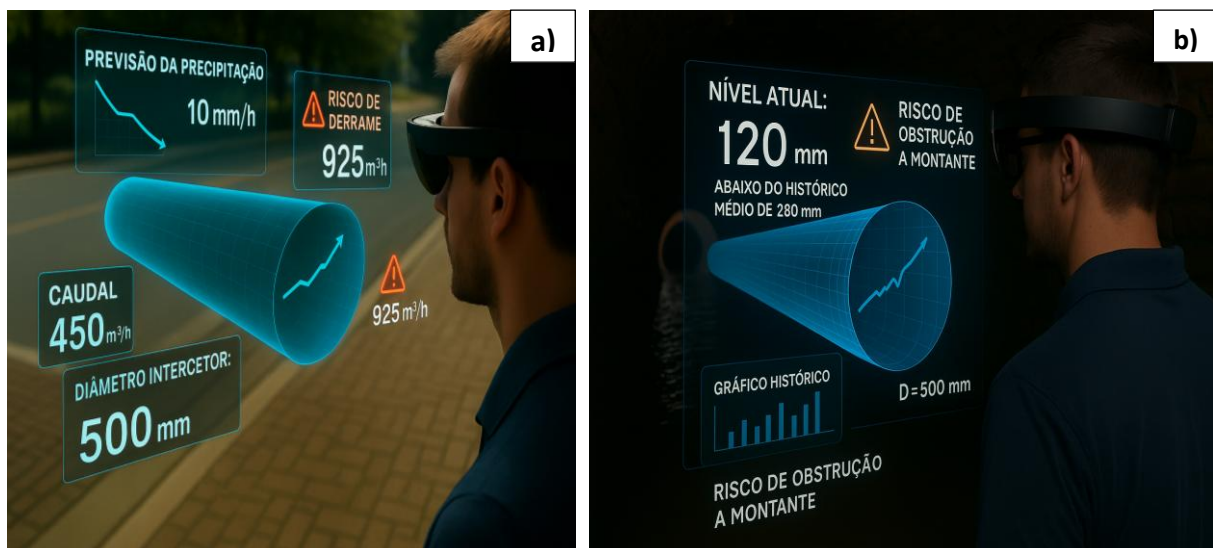


Figura 16 - a) Eventual anomalia despoletada por evento de precipitação intenso e b) eventual anomalia relacionada com obstrução.

O que se verifica em 16 a) é que o operador tem acesso em tempo real aos dados de um determinado ponto da bacia onde o sistema de previsão lhe indica a possibilidade de precipitação nas próximas 6h, sendo que o evento poderá ser de 10mm em apenas uma hora. Indica ainda que tendo em conta os 10 anos de histórico é provável que naquele o ponto o caudal evolua de 450 m³/h para 925 m³/h, o que despoletará inúmeros problemas. Com isto, aconselha opções de manutenção e intervenção na rede para mitigar o problema. Em 16 b) o sistema alerta para uma eventual obstrução a montante. O nível no ponto atual é de 120 mm, sendo que o histórico para aquele ponto, na hora e dia de semana deveria estar nos 280mm. O sistema apresenta 87% de probabilidade da ocorrência de uma obstrução a montante.

4. CONCLUSÕES

O trabalho apresentado, pioneiro a nível nacional e ainda raro em termos internacionais, permitiu aumentar consideravelmente o desempenho económico e ambiental da entidade responsável pela operação de um conjunto de infraestruturas de tratamento e drenagem complexos.

Pela primeira vez, desenvolveram-se e aplicaram-se ferramentas hidroinformáticas para a gestão operacional de um conjunto de infraestruturas de drenagem na bacia do rio Ave. A integração de diferentes tecnologias e ferramentas hidroinformáticas foi conseguida com sucesso como se pode comprovar pelos resultados apresentados anteriormente.

Foram diagnosticadas as causas responsáveis pelas afluências excessivas durante eventos de precipitação, resultando de forma clara que afluem à rede de recolha de águas residuais domésticas e industriais importantes volumes de águas pluviais provenientes das redes secundárias ligadas à rede principal e que foram quantificados neste trabalho. A análise simultânea de medições de precipitação e hidrogramas medidos na rede principal permite antecipar como a sua origem provável, sistemas de recolha e drenagem de águas pluviais indevidamente ligados à rede de águas residuais domésticas e industriais.

A utilização de modelos de *machine learning*, *digital twins* e até mesmo *realidade aumentada* revelaram ser uma boa opção para perspetivar eventos futuros, sendo os modelos RFR e XGboost os que obtiveram melhor desempenho.

O Aquaflow, plataforma utilizada como DT permitiu aceder, integrar e processar um conjunto de informação relevante que constitui uma mais-valia para os operadores do sistema, salientando-se o acesso a resultados de modelos de previsão meteorológica. Destes, as previsões de precipitação permitem, por exemplo, antecipar a programação de intervenções das equipas de manutenção sempre que as condições meteorológicas assim o exigem.

Salienta-se o interesse do sistema implementado do ponto de vista operacional, ao permitir antecipar com antecedência de dois dias eventos extremos com um nível de incerteza razoável.

A ferramenta de apoio à decisão desenvolvida para a bacia do Rio Ave constitui uma plataforma tecnológica robusta e eficiente para apoio à gestão de recursos hídricos. Os resultados obtidos mostram que esta nova ferramenta será extremamente eficaz e importante para atingir os objetivos da gestão operacional à escala da bacia hidrográfica. Nos próximos anos, a utilização deste tipo de ferramentas avançadas de controlo operacional será um requisito essencial para todas as entidades responsáveis pela gestão de recursos hídricos, sendo que cada vez mais, gerir este tipo de recursos se torna mais desafiante e complexo.

Em síntese, o trabalho aqui apresentado confirma que a transformação digital do saneamento não é apenas uma tendência, mas uma necessidade estratégica. A integração DT+ML+RA representa um caminho concreto para redes mais eficientes, resilientes e sustentáveis, alinhado com os objetivos nacionais e internacionais de gestão da água.

AGRADECIMENTOS

O autor agradece à TRATAVE S.A., empresa onde o projeto vem a ser desenvolvido e à empresa Aquapor, pela disponibilidade de recursos, sejam eles humanos, tecnológicos ou financeiros e ainda à Universidade do Minho pela parceria, onde as sinergias foi complementadas com conhecimento acrescido e diferenciador.

BIBLIOGRAFIA

JOSEPH-DURAN, B.; CARLOS, O.-M.; CEMBRANO, G. – "Hybrid modeling and receding horizon control of sewer networks". *Water Resources Research*, 50, 2014, pp. 8497–8514.

LOWE, R.; VEZZARO, L.; STEEN, P.; GRUM, M. – "Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems". *Environmental Modelling & Software*, 80, 2016, pp. 143–158.

PEREIRA, A.; PINHO, J.; FARIA, R.; VIEIRA, J.; COSTA, C. – "Improving operational management of wastewater systems: A case study". *Water Science and Technology*, 8(1), 2019, pp. 173–183.

PINHO, J. L.; VIEIRA, J. M.; PINHO, R. P.; ARAÚJO, J. P. – "Plataforma Web de suporte ao sistema de modelação da rede hidrográfica da bacia do rio Ave (Web based technological platform to support river basin modelling system)". *Revista Recursos Hídricos (Portuguese Water Resources Journal)*, 32(1), 2011, pp. 13–28.

PINHO, J. L.; VIEIRA, J. M. – "A web-based hydroinformatic platform for water quality modelling in a river basin". *11th International Conference on Hydroinformatics*, 2014, pp. 1–8.

PEREIRA, A.; PINHO, J.; FARIA, R.; VIEIRA, J. – "A DSS for operational management of wastewater systems". In *HIC 2018 – 13th International Conference on Hydroinformatics*, Palermo (Itália), 2018, pp. 1616–1623.

NICOLAU, R. – *Modelação e Mapeamento da Distribuição Espacial da Precipitação – Uma Aplicação a Portugal Continental*. Lisboa (Portugal), Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2002.