

PLATAFORMA DECIdE – TRANSFORME DADOS EM CONHECIMENTO

Bruno FERREIRA¹; Nelson CARRIÇO¹; Ana MENDES¹; Miguel BAIÓ¹; Raquel BARREIRA¹; Cédric GRUEAU¹; André ANTUNES¹; João GOMES¹; Miguel GONÇALVES¹; Dídía COVAS²; João FERREIRA²; Sílvia FERNANDES²; João SANTOS³; Isabel BRITO³; Cristina BOGA³

RESUMO

As Entidades Gestoras (EG) de sistemas de abastecimento de água são responsáveis por assegurar água em quantidade e qualidade para as populações que servem, conforme o 6º Objetivo de Desenvolvimento Sustentável da ONU [1] sem, no entanto, comprometerem a sua integridade económico-financeira. Estas EG são responsáveis pela gestão e operação destes sistemas, recorrendo regularmente a um numeroso conjunto de sistemas de informação (SI) para recolher, armazenar e gerir o vasto conjunto de dados gerados por esta atividade.

No âmbito do projeto de investigação DECIdE - Plataforma de Apoio à Decisão em Infraestruturas Urbanas de Água – foi desenvolvida uma plataforma informática que integra dados provenientes de diferentes SI da EG, incluindo ainda algumas ferramentas para o apoio à tomada de decisão. A plataforma foi desenvolvida em colaboração com cinco EGs de pequena e média dimensão, com modelos de gestão (i.e., gestão direta e gestão delegada), maturidade digital e processos de gestão da informação diferentes.

A integração de informação proveniente de diversos SI numa base de dados unificada é um processo complexo, em parte devido à disparidade de processos de gestão de informação e maturidade digital presentes nas EGs parceiras. Neste sentido, foi desenvolvido um modelo de dados abrangente, assim como um mecanismo de análise, mapeamento e indexação entre o conteúdo a importar e o modelo de dados unificado, permitindo assim que diferentes entidades utilizem a plataforma.

A plataforma, de forma a apoiar o processo de apoio à tomada de decisão, incorpora um conjunto de indicadores de desempenho [2], selecionados pelas EGs parceiras, de forma a refletir as necessidades por estas sentidas. Adicionalmente, foram incorporados módulos de cálculo de balanços hídrico e energético. Estes balanços permitem a avaliação da eficiência hídrica [3] e energética [4] dos sistemas. A visualização gráfica deste conjunto de indicadores de desempenho pode tornar-se de alguma forma difícil, especialmente quando se pretende realizar análises ao nível setorial (i.e., subsistema ou Zona de Medição e Controlo). Neste sentido, e com o objetivo de facilitar o processo de apoio à tomada de decisão, foi desenvolvido um *dashboard* baseado em filtros customizados pelo utilizador (i.e., por área de análise, conjunto de dados e indicadores, e período temporal), incluindo ainda diferentes tipos de representação gráfica (e.g., gráfico de barras, linhas, circular).

O presente artigo aborda o desenvolvimento das diversas funcionalidades, módulos e ferramentas da plataforma DECIdE.

Palavras chave: Sistemas de informação, Integração de dados, Apoio à tomada de decisão

¹ Instituto Politécnico de Setúbal, Campus do IPS – Estefanilha, 2910-761 Setúbal, wisdomproject.rd@gmail.com

² Instituto Superior Técnico, Av. Rovisco Pais, 1, 1049-001 Lisboa, wisdomproject.rd@gmail.com

³ Instituto Politécnico de Beja, Campus do IPB - Rua Pedro Soares, 7800-295 Beja, wisdomproject.rd@gmail.com

1. INTRODUÇÃO

Para garantir um serviço de abastecimento de água, em quantidade e qualidade, as Entidades Gestoras (EG) necessitam de recolher inúmeros dados (e.g., caudal, pressão, cloro) de diferentes fontes que, posteriormente, são utilizadas em diversas áreas de atividade, nomeadamente, cadastro, contabilidade, operação e manutenção, engenharia (e.g., diagnóstico, previsão e modelação), entre outros (Carriço *et al.* 2018). Estes dados são, posteriormente, armazenados e acedidos através de diversos sistemas de informação (SI), sendo os mais comuns os seguintes: Sistemas de Informação Geográfica (SIG), *Enterprise Resource Planning* (ERP), sistemas de gestão de clientes, sistemas de apoio à contabilidade e ainda sistemas de telemedicação e de telegestão, mais comumente, conhecido por SCADA.

A implementação de um processo de gestão patrimonial de infraestruturas (GPI), de acordo com a atual legislação em vigor, requer que as EG disponham de informação sobre a situação atual e projetada das infraestruturas, a sua caracterização e a avaliação do seu estado funcional e de conservação. Algumas das ferramentas para esta caracterização e avaliação assenta em diferentes técnicas, como por exemplo, cálculo de indicadores de desempenho, simulação hidráulica do sistema e previsão de falhas com recurso a técnicas de Inteligência Artificial (IA), carecem de interoperabilidade e integração entre os diversos SI. No entanto, devido à complexidade associada aos SI existentes e modelos de dados por vezes desadequados, a interoperabilidade e a integração não são facilmente alcançáveis (Vieira *et al.* 2018).

Em Portugal, a grande maioria dos sistemas de abastecimento de água em baixa são geridos diretamente pelos municípios, que têm uma ampla missão que não se restringe aos serviços de água como acontece nas restantes EG existentes. Como tal, a gestão é estruturada de acordo com modelos de convenção de administrações públicas e em muitos casos a eficiência da gestão da informação fica frequentemente aquém do pretendido. Para além disso, a maioria dos SI não partilham facilmente os seus dados, criando-se silos de informação que deixam diversos departamentos sem uma visão holística da operação e do desempenho do sistema.

Considerando a falta de meios de comunicação entre os diferentes SI da EG, bem como as reduzidas políticas e práticas para transferência de informação entre departamentos, o simples acesso a dados pode revelar-se uma atividade bastante morosa, tipicamente envolvendo diversos intervenientes, ficheiros de diversos formatos e processos intermédios (Grueau *et al.* 2019). Diversas soluções têm sido desenvolvidas de forma a resolver a problemática da integração e análise de dados em EG. Este desenvolvimento pode partir da própria EG que, dispondo de meios financeiros e/ou tecnológicos, desenvolve uma solução à medida das suas necessidades, ou através da aquisição e utilização de *softwares* comerciais.

A recolha e tratamento dos dados, bem como a manutenção e gestão dos SI constituem uma parte significativa dos recursos humanos e tecnológicos das EG, sendo este um dos grandes desafios com que se deparam atualmente (Vieira *et al.* 2018). Não existindo uma forma otimizada para a gestão de dados numa EG, e tendo em conta a impossibilidade de se implementar um modelo de gestão de dados de raiz, é de extrema importância o correto planeamento e uso dos SI existentes e a adquirir, tendo em conta a sua interligação e o seu proveito para alcançar os objetivos da EG.

No âmbito do projeto de investigação Plataforma de Apoio à Decisão em Infraestruturas Urbanas de Água – DECIdE – foi desenvolvida uma plataforma informática que pretende integrar dados provenientes de diferentes SI da EG, incluindo ainda algumas ferramentas para o apoio à tomada de decisão. O projeto foi coordenado pelo Instituto Politécnico de Setúbal e desenvolvido em colaboração com o Instituto Superior Técnico e o Instituto Politécnico de Beja. A plataforma foi desenvolvida em

colaboração com cinco EG de pequena e média dimensão, com diferentes modelos de gestão (i.e., gestão direta e gestão delegada), maturidade digital e processos de gestão da informação.

O projeto teve início em janeiro de 2018 e desenvolveu-se ao longo de 18 meses em três atividades principais, nomeadamente, A1 – Caracterização dos sistemas de informação; A2 – Desenvolvimento da plataforma; e A3 – Desenvolvimento de relatórios. A primeira atividade envolveu a caracterização e levantamento dos SI existentes entre as entidades parceiras do projeto. A segunda atividade consistiu no desenvolvimento da plataforma que integra os dados dos diferentes SI identificados na primeira atividade. Nesta atividade foram desenvolvidos os diversos módulos da plataforma. Por fim, a terceira atividade consistiu na preparação de diretrizes de práticas recomendadas para a gestão eficiente da informação.

O presente artigo aborda o desenvolvimento das diversas funcionalidades, módulos e ferramentas da plataforma DECIdE.

2. CARACTERIZAÇÃO DOS SISTEMAS DE INFORMAÇÃO EXISTENTES

A primeira tarefa do projeto teve como objetivo recolher e compilar o máximo de informação possível acerca dos sistemas de informação, de forma a proceder à sua caracterização. As cinco EG participantes do projeto dispõem de SIG, o que é natural dado que a gestão dos sistemas de abastecimento de água é marcada pela elevada complexidade das infraestruturas geridas, em grande parte devido à sua acentuada dispersão geográfica. Desta forma, o SIG é uma das ferramentas de apoio à gestão mais relevantes.

Nas cinco EG participantes, foram identificadas três marcas diferentes de SIG, assim como processos de recolha e atualização de informação bastante distintos. A falta de recursos humanos especializados para trabalhar com este tipo de tecnologia é também uma realidade para algumas das EG.

Não é, portanto, de estranhar que nos últimos anos tenham sido desenvolvidas diversas soluções por parte das EG e empresas de software, de forma a utilizar o SIG como meio de ligação entre os seus diversos sistemas (Reis *et al*, 1998). Silva *et al* (2017) e Sousa *et al* (2017) apresentam dois exemplos de EG que resolveram o seu problema de integração e análise de dados através de uma solução de interligação entre o SIG e o sistema de gestão comercial, sistema de gestão de perdas e sistema de gestão de equipas e operações em campo.

O sistema de gestão de clientes, onde se inclui a faturação, foi indicado como sendo importante de integrar dado que é nesta aplicação que estão registados os consumos ao longo da rede de distribuição de água. Algumas EG possuem um código que permite a associação do consumo registado a um determinado ramal de ligação, que por sua vez está georreferenciado no SIG. No entanto, esta realidade não é comum a todas as EG. À semelhança do SIG, também aqui, há diferentes marcas de software e processos de recolha e atualização da informação. Um aspeto que importa referir, contudo, é que há *softwares* cuja base de dados não está alojada localmente, mas, sim, num servidor remoto alojado no fornecedor de *software*. Esta situação implica que cada vez se pretenda aceder aos dados é necessário solicitá-los ao fornecedor do *software* de faturação. A questão que se levanta, neste caso, e que se deixa aqui para reflexão é a de quem é o real proprietário dos dados, é a EG ou o fornecedor de *software*?

Para a GPI é importante a integração das chamadas ordens de trabalho/ordens de serviço (OT) e, neste campo, há várias realidades entre as EG. Há as que dispõem de um software próprio para o registo das OT e as que dispõem do tradicional registo em papel que, posteriormente, são registadas numa folha de cálculo.

Os sistemas de telemedição e telegestão produzem um volume elevado de dados que, no entanto, necessita de algum tratamento (i.e., tratamento de *outliers*, dos nulos, preenchimento de falhas) de forma a ser utilizado em técnicas mais robustas de Inteligência Artificial. Por se encontrar em desenvolvimento o módulo de tratamento de séries temporais, não foram integrados os sistemas de Telemedição e telegestão, tendo, no entanto, sido feita a sua caracterização.

Nesta fase, a integração do sistema ERP não foi equacionada.

3. SOLUÇÕES EXISTENTES PARA INTEGRAÇÃO DE DADOS

A estruturação e desenvolvimento interno de uma plataforma tecnológica é outro tipo de solução que permite resolver o problema da integração. No entanto, a utilização deste tipo de solução é, usualmente, restrita à EG que a desenvolveu.

Vieira *et al.* 2018 apresenta uma solução que, através da integração dos dados de diversas fontes (i.e., sensores, SIG) com os resultados da simulação hidráulica, permite a melhoria da gestão eficiente de todo o Ciclo Urbano da Água. Adicionalmente, permite a criação de uma base de trabalho transversal a toda a organização, tornando-se assim num importante instrumento de tomada de decisão para a EG.

Uma vez integradas as diversas fontes de informação, é possível proceder à análise destes dados, com recurso a diversos métodos (e.g., cálculo de indicadores de desempenho, auditorias hídricas e energéticas, entre outros), produzindo assim informação útil e valiosa ao processo de apoio à tomada de decisão.

Diversas soluções comerciais de *Business Intelligence* permitem a ligação a diversas fontes de informação, possibilitando, entre outras funções, o cálculo de indicadores de desempenho, construção de *dashboards* dinâmicos ou produção de relatórios. Adicionalmente, permitem a manipulação dos dados e o desenvolvimento de código, permitindo assim a realização de análises customizáveis.

O cálculo do balanço hídrico (Alegre *et al.* 2000) e energético (Mamade *et al.* 2017) pode ser realizado através de aplicações baseadas em folhas de cálculo desenvolvidas para o efeito. No entanto, estas folhas exigem uma preparação exigente dos dados de entrada, dificultando assim sua utilização regular para análise do sistema.

Um conjunto limitado de soluções comerciais permite a avaliação do desempenho do sistema de abastecimento, através do cálculo de indicadores de desempenho, auditorias hídricas e energéticas. No entanto, não permitem a integração das fontes de informação. Desta forma, o tempo de preparação da informação poderá injustificar a utilização da solução.

Existe assim a necessidade de uma solução que permita às EG integrar as diversas fontes de informação, permitindo a produção de informação útil ao processo de apoio à tomada de decisão.

4. PLATAFORMA DECIdE

4.1. Modelo conceptual e desenho

No âmbito do projeto de investigação DECIdE foi desenvolvida uma plataforma informática com intuito de integrar dados provenientes de diferentes SI das EG parceiras. Adicionalmente, disponibilizaram-se algumas ferramentas para o apoio à tomada de decisão, nomeadamente, indicadores de desempenho,

cálculo de balanços hídrico e energético. As cinco EG parceiras são de pequena e média dimensão, com modelos de gestão (i.e., gestão direta e gestão delegada), maturidade digital e processos de gestão da informação bastante diferentes.

A integração de informação proveniente de diversos SI numa base de dados unificada é um processo complexo, em parte devido à disparidade de processos de gestão de informação e maturidade digital presentes nas EG parceiras. Neste sentido, foi desenvolvido um modelo de dados unificado, assim como um mecanismo de análise, mapeamento e indexação entre o conteúdo a importar e o modelo de dados unificado, permitindo assim que diferentes entidades utilizem a plataforma.

A plataforma, de forma a apoiar o processo de apoio à tomada de decisão, incorpora um conjunto de indicadores de desempenho (ERSAR 2019), selecionados pelas EGs parceiras, de forma a refletir as necessidades por estas sentidas. Adicionalmente, foram incorporados módulos de cálculo de balanços hídrico e energético. Estes balanços permitem a avaliação da eficiência hídrica (Alegre *et al.* 2000) e energética (Mamade *et al.* 2017) dos sistemas.

A visualização gráfica deste conjunto de indicadores de desempenho pode tornar-se de alguma forma difícil, especialmente quando se pretende realizar análises ao nível setorial (i.e., subsistema ou zona de medição e controlo). Neste sentido, e com o objetivo de facilitar o processo de apoio à tomada de decisão, foi desenvolvido um *dashboard* baseado em filtros customizados pelo utilizador (i.e., por área de análise, conjunto de dados e indicadores, e período temporal), incluindo ainda diferentes tipos de representação gráfica (e.g., gráfico de barras, linhas, circular).

A plataforma DECIdE, disponível em <https://decide.ips.pt/>, apresenta uma estrutura modular, constituída por diversos módulos que comunicam com a base de dados unificada. Na Figura 1, apresenta-se uma imagem ilustrativa da Plataforma DECIdE.

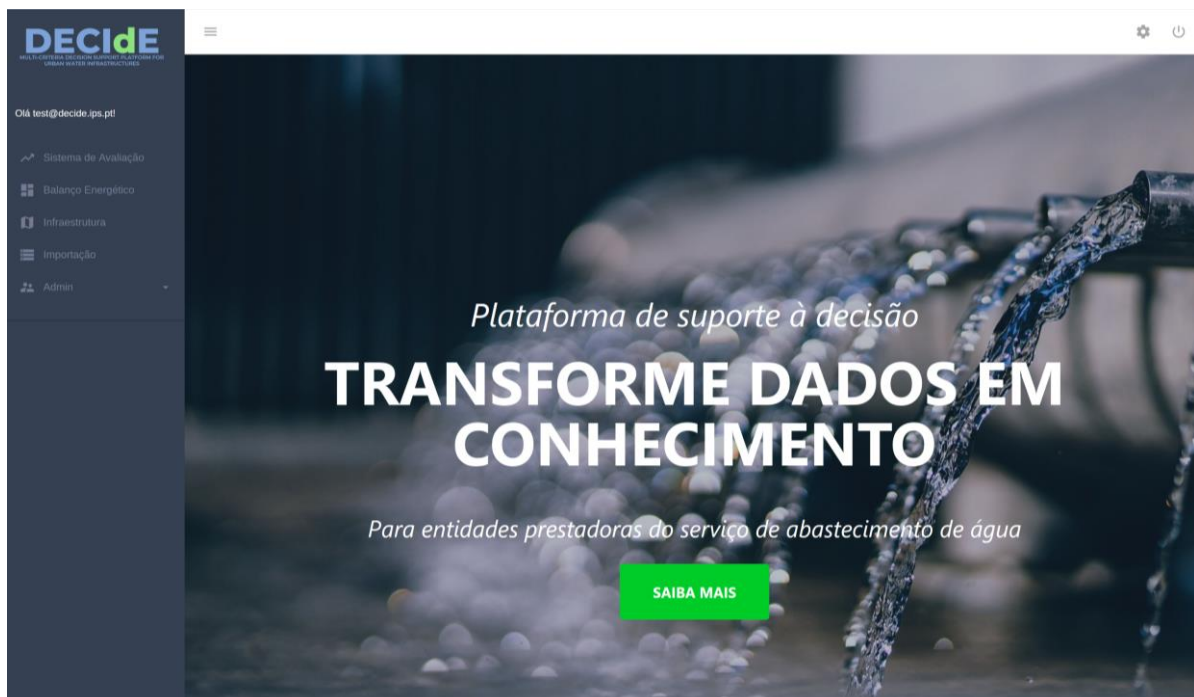


Figura 1 – Imagem ilustrativa da plataforma DECIdE

De seguida apresenta-se sumariamente cada um dos módulos desenvolvidos.

4.2. Módulos desenvolvidos

4.2.1. Importação e atualização de dados

O módulo de importação e atualização de dados permite a importação e atualização das formas e atributos referentes às diversas infraestruturas e áreas de análise, bem como a importação da faturação, medições e ordens de trabalho associadas ao sistema.

A disparidade de modelos de gestão da informação entre EG apresenta-se como o maior desafio ao desenvolvimento de uma base de dados unificada. Assim, para cada tipo de elemento a ser importado (e.g., condutas ou reservatórios) foi definido um modelo de dados. Este modelo é constituído por um conjunto de atributos mínimo que deve constar em cada elemento para responder aos requisitos de cálculo dos diversos módulos.

A partir deste modelo de dados, e através da interface da plataforma, é possível realizar o mapeamento entre o modelo de dados original e o modelo unificado, integrando assim os dados. Adicionalmente, o modelo de dados está preparado para manter um registo histórico dos diversos elementos infraestruturais, permitindo assim reproduzir o estado da rede num determinado período do tempo.

4.2.2. Visualização de dados

O módulo de visualização de dados permite agregar um conjunto de dados pré-estabelecidos, considerando diferentes áreas de análise e um determinado período temporal.

O resultado deste cálculo pode ser visualizado de forma customizada ao nível do tipo de gráfico, filtragem e ordenação. Desta forma é possível analisar o sistema de forma setorial ou integrada, permitindo ainda analisar a evolução temporal dos indicadores de desempenho, contribuindo com informação útil e importante para o processo de apoio à tomada de decisão.

4.2.3. Visualização da representação da infraestrutura

O módulo de visualização da representação da infraestrutura apresenta graficamente os elementos georreferenciados, importados através do módulo de importação. Cada elemento é assim apresentado num mapa a duas dimensões, disponibilizado na plataforma. São ainda disponibilizadas um conjunto de funcionalidades que facilitam a manipulação do mapa, como ferramentas de aproximação e movimentação.

4.2.4. Cálculo do balanço hídrico

O módulo de cálculo do balanço hídrico, através da definição do período temporal e das áreas de análise, prepara os dados de entrada para validação por parte do utilizador, possibilitando a alteração de qualquer valor. Uma vez validados os dados de entrada, o módulo permite o cálculo do balanço hídrico de acordo com Alegre *et al.* 2000.

4.2.5. Cálculo do balanço energético

À semelhança do cálculo do balanço hídrico, o cálculo do balanço energético exige uma preparação rigorosa dos dados de entrada. Este módulo, através da definição do período temporal e das áreas de análise, prepara os dados de entrada para validação por parte do utilizador, possibilitando a alteração

de qualquer valor. Uma vez validados os dados de entrada, o módulo permite o cálculo do balanço energético de acordo com Mamade *et al.* 2017.

4.3. Módulos em desenvolvimento

4.3.1. Ligação às fontes de informação através de serviços web

Pretende-se desenvolver um mecanismo de ligação entre a base de dados unificada e as diversas fontes de informação da EG. Desta forma, e através de atualizações periódicas e automáticas (e.g., uma vez por dia) é possível manter o modelo de dados atualizado, permitindo a avaliação do desempenho com a periodicidade desejada.

4.3.2. Incorporação do módulo de produção do relatório anual

A Entidade Reguladora dos Serviços de Água e Resíduos (ERSAR) solicita, anualmente, às EG um alargado conjunto de dados sobre as infraestruturas que gerem, servindo posteriormente para a elaboração do “Relatório Anual dos Serviços de Águas e Resíduos em Portugal (RASARP)”, onde se caracterizam os serviços de águas e resíduos em Portugal continental.

A recolha e tratamento do conjunto de dados solicitados exige das EG um elevado esforço ao nível de recursos humanos, podendo esta atividade prolongar-se por diversos meses.

Neste sentido, está previsto o desenvolvimento de um módulo para produção do relatório com o conjunto de dados solicitado pela ERSAR.

4.3.3. Módulo de tratamento de séries temporais

A diminuição do custo de aquisição dos sensores de medição, da tecnologia de comunicação e do armazenamento permitirá a obtenção de um alargado volume de dados que, através de um conjunto de inovações tecnológicas, poderá gerar insights bastante significativos. No entanto, uma parte dos dados recolhidos de forma automática necessitam, posteriormente, de tratamento de *outliers*, dos nulos e preenchimento de falhas.

Sem este tratamento prévio não é possível aplicar técnicas robustas de Inteligência Artificial (IA), tais como, reconhecimento automático de padrões de consumo, modelos preditivos de consumos de água e localização espacial de roturas. Desta forma, encontra-se em desenvolvimento o módulo de tratamento de séries temporais.

4.3.4. Módulo de deteção e categorização de eventos anómalos

O comportamento da rede pode ser analisado através dos sensores de medição instalados e, perante a inexistência de eventos anómalos (e.g., roturas ou abertura de marcos de incendio), o comportamento segue padrões específicos (i.e., diário, semanal).

A ocorrência de um evento anómalo gera um incremento de caudal e variações de pressão que, com recurso a técnicas estatísticas, pode ser detetado como *outlier*.

Uma vez identificado o evento anómalo, e através de técnicas de *machine learning* e inteligência artificial, é possível proceder à sua categorização.

4.3.5. Módulo de localização espacial de roturas

Continua a ser um desafio para as EG localizar um evento de rotura na rede após a sua deteção.

O frequente uso de tubagens em Polietileno dificulta o uso das técnicas realizadas com recurso a equipamentos acústicos. Embora nos últimos anos se tenha assistido a uma evolução ao nível dos equipamentos não acústicos, o uso deste tipo de técnicas é dispendioso e requer um esforço intensivo da equipa para localizar a rotura, até em redes de pequena dimensão.

5. CONSIDERAÇÕES FINAIS

A integração das diversas fontes de informação da EG é essencial à produção de informação útil ao processo de apoio à tomada de decisão. O presente artigo abordou o desenvolvimento de uma plataforma informática com o intuito de integrar os dados das diversas fontes, assim como o desenvolvimento das suas diversas funcionalidades, módulos e ferramentas.

A solução permite uma fácil integração dos dados por parte da EG, possibilitando a avaliação do desempenho global e setorial do sistema através de um conjunto de indicadores de desempenho, cálculo do balanço hídrico e energético.

Esta plataforma constitui-se como o primeiro passo no desenvolvimento de uma ferramenta de referência para o apoio à tomada de decisão em EGs de serviços urbanos de água. Os trabalhos futuros incidem sobre área da ciência de dados, com especial enfoque para o tratamento de séries temporais e a deteção, categorização e localização de eventos anómalos.

AGRADECIMENTOS

Os autores agradecem ao Programa Operacional Regional de Lisboa e ao Programa Operacional Regional do Alentejo, na componente FEDER, bem como à Fundação para a Ciência e Tecnologia pelo suporte financeiro ao projeto nº 024135.

BIBLIOGRAFIA

Alegre, H., Hirner, W., Baptista, J.M., Parena, R. "Performance indicators for water supply services", Manual of Best Practice Series, IWA Publishing, London, 2000.

Carriço, N., Ferreira, B., Monteiro, L., Mendes, A., Gureau, C., Covas, D., Santos, J., Figueira, J., Baio, M., Barreira, R., Silvestre, I., Isidro, R., Mestrinho, J., Faím, J., Traitolas, S., Chaveiro, P. "Plataforma de apoio á decisão em infraestruturas urbanas de água", no 18.º ENASB/18.º SILUBESA, Cidade: Porto, 2018.

ERSAR. "Guia de avaliação da qualidade dos serviços de águas e resíduos prestados aos utilizadores - 3.ª geração do sistema de avaliação". ERSAR, Lisboa, 2019.

Grueau, C., Antunes, A., Ferreira, B., Gonçalves, M., Gomes, M., Carriço, N. "Towards an integrated platform for decision support in water utility management", no 12th IADIS International Conference on Information Systems, Cidade: Utrecht, Holanda, 2019.

Mamade, A., Loureiro, D., Alegre, H., Covas, D. "A comprehensive and well tested energy balance for water supply systems", no Urban Water Journal, 2017. DOI: 10.1080/1573062X.2017.1279189

Reis, P., Gonçalves, J., Rodrigues, P. "O SIG como interligação de sistemas de informação", no 4º Congresso da água, Cidade: Lisboa, 1998.

Silva M. H., Vilela J., Morais M. "Interligação dos sistemas de informação nos SMAS de Sintra. Contributo na prestação de serviço público de qualidade", no ENEG 2017, Cidade: Évora, 2017.

Sousa C., Marques A., Fernandes J. "O contributo da informação em SIG para uma gestão integrada e eficiente dos sistemas", no ENEG 2017, Cidade: Évora, 2017.

Vieira, P., Costa, F., Cunha, R., Marques, R., Fernandes, M. "Plataforma tecnológica do ciclo urbano de água. Da conceção à concretização", no 18.º ENASB/18.º SILUBESA, Cidade: Porto, 2018.