

XIII CONGRESSO DA GEOGRAFIA PORTUGUESA

O COMPROMISSO DA GEOGRAFIA PARA TERRITÓRIOS EM MUDANÇA

Coimbra, 18 a 20 de novembro de 2021

LIVRO DE ATAS

Maio de 2022



FACULDADE DE LETRAS
UNIVERSIDADE D
COIMBRA

CEGOT
Centro de Estudos de Geografia
e Ordenamento do Território



MODELO PREDITIVO DE FUGAS NO SISTEMA DE DISTRIBUIÇÃO MUNICIPAL DE ÁGUA DE VILA NOVA DE GAIA

Oliveira, André ^{1*}; Gomes, Alberto ²; Baptista, Ricardo ³

1* Faculdade de Letras da Universidade do Porto e Águas de Gaia; refandreoliveira@gmail.com

2 Departamento de Geografia, Faculdade de Letras da Universidade do Porto; Centro de Estudos em Geografia e Ordenamento do Território (CEGOT); atgomes@letras.up.pt

3 Departamento de Geografia, Faculdade de Letras da Universidade do Porto; ricardojose@fe.up.pt

Resumo: A gestão eficiente dos recursos hídricos é um dos principais desafios que a sociedade atual enfrenta. É essencial dotar as entidades gestoras com ferramentas de apoio à tomada de decisão, de forma a reduzirem a perda de água no momento do seu transporte para abastecimento das populações. A partir das características geográficas da rede de distribuição de água no município de Vila Nova de Gaia e dos registos de fugas de água ocorridas entre os anos de 2017 e 2019, este estudo desenvolve um modelo da suscetibilidade à ocorrência de fugas baseado em algoritmos *machine learning* e ferramentas SIG. O algoritmo utilizado foi o *random forest regressor*, e permitiu delimitar áreas mais suscetíveis à ocorrência de fugas na área de estudo.

Palavras-chave: Distribuição de água; fugas de água; random forest regressor; modelo preditivo

Abstract: An efficient management of water resources is one of the main society's challenges. Therefore, it is essential that entities arm themselves with tools that support decision-making in order to reduce water losses on the transportation moment to supply populations. This study proposes to analyse Vila Nova de Gaia's water distribution network, on a geographic point of view, and the occurred leaks in the years of 2017 and 2019. From the collected data and analysis, it was developed a model with machine learning algorithms and GIS tools to predict future leaks. The used algorithm was random forest regressor and it allowed to bound the most susceptible areas for the occurrence of water leaks in this municipality.

Keywords: Water distribution, water leaks, random forest regressor, predictive model

1. Introdução

A água é um recurso vital da Terra, no entanto, é um dos recursos mais ineficazmente gerido. Atualmente, o mundo enfrenta uma crise na gestão dos recursos hídricos que se prevê intensificar-se (Rogers, 2014). Portugal não é exceção a esta tendência. A Entidade Reguladora de Águas e Resíduos (ERSAR), estima que em 2018 ocorreu uma perda de 172 milhões de m³ de água potável, o que se traduz, sensivelmente, numa perda económica de 85.5M€. Considerando que cada residente em Portugal consome em média 192 litros por dia, significa que esta perda seria suficiente para abastecer ¼ da população. Cerca de 79 municípios portugueses tiveram perdas de águas superiores a 50%, ou seja, perderam mais água do que aquela que serviram à população (ERSAR, 2019). As perdas de água podem ser categorizadas em função do momento da sua ocorrência (Liemberger & Farley, 2004). Estas podem ser perdas no momento de tratamento de água, perdas através de uso não autorizado, perdas devido a erros de medição e perdas devido a fugas. As perdas devido a fugas podem ainda ser divididas em subcategorias, como fugas e extravasamentos nos reservatórios, fugas nos ramais de ligação e fugas nas condutas de distribuição. Este estudo foca-se no último caso.

O investimento na manutenção da infraestrutura de abastecimento é uma constante, mas paralelamente, tenta-se desenvolver novos métodos capazes de reduzir, direta ou indiretamente, fugas na rede. Este estudo foca-se no município de Vila Nova de Gaia, pertencente à Área Metropolitana do Porto, que segundo os censos 2011 contava com 303430 habitantes (4º mais populoso do país), sendo, no entanto, apenas o 11º relativamente à densidade populacional (1799 hab/km²). Assim, a dispersão do povoamento traduz-se numa rede de distribuição de água bastante extensa e ramificada. Tendo isto em conta, e o ciclo de vida de toda a infraestrutura de abastecimento de água é natural e previsível a ocorrência de fugas de água.

Os métodos frequentemente utilizados na identificação de fugas de água da rede são acústicos. No entanto, nem sempre é fácil e economicamente viável recorrer a eles, por isso, é importante identificar as áreas mais suscetíveis à ocorrência de fugas para reduzir a área de procura do incidente, principalmente, quando se lida com uma área extensa, sendo mais eficiente a manutenção e a procura de eventuais fugas. Neste estudo, através de vários parâmetros geográficos, das características da própria rede, e de um registo temporal de ocorrências, pretende-se compreender os padrões de ocorrência de fugas e prever os locais onde possam ocorrer, mediante a obtenção de áreas mais suscetíveis à ocorrência de fugas de água na rede de abastecimento de água de Vila Nova de Gaia.

2. Métodos

O registo temporal de ocorrências que serviu para o desenvolvimento do modelo preditivo, são as fugas ocorridas nos anos de 2017, 2018 e 2019. Os dados reportam-se às ordens de serviço, que se encontravam em tabela, o que implicou a sua geocodificação, e a associação das fugas à conduta de distribuição. Com este passo foi possível identificar as condutas onde ocorreram fugas no período de estudo e aquelas em que não ocorreram. O passo seguinte foi perceber quais as variáveis geográficas que podem afetar negativamente a rede de distribuição de água. Jenkins et al. (2014) desenvolveram um estudo sobre os parâmetros de vulnerabilidade utilizados nas principais publicações científicas em que elaboram modelos para prever fugas de água com dados geográficos, entre outros. Foi com base nestas publicações que se identificaram os parâmetros a usar e se procedeu à sua recolha, em função da sua existência e acesso. A partir destes dados, foi desenvolvido uma base de dados geográfica, que permitiu ter uma análise descritiva de cada fuga ocorrida, com base nos vários parâmetros de suscetibilidade.

O algoritmo utilizado para desenvolver o modelo preditivo foi o *random forest regressor* (RFR). Este tipo de metodologia permite gerar informação que não é conhecida a partir de dados conhecidos. O RFR é um método *ensemble* e segue a tática de “dividir e conquistar”, sendo uma estratégia comum em modelos *machine learning* (Polikar, 2006). Ou seja, neste caso, através dos vários parâmetros de suscetibilidade, o modelo irá separar cada conduta através das suas características geográficas e no final atribuir um valor de suscetibilidade à ocorrência de fugas. Uma das vantagens na utilização de métodos *ensemble* é a possibilidade de utilizar várias “árvores” de decisão. Isto faz particularmente sentido quando os dados não são todos iguais. Voltando a tomar como exemplo o caso estudo, a geografia da área urbana de Vila Nova de Gaia é completamente diferente das áreas rurais, não fazendo sentido utilizar um único classificador para todo o município.

O desenvolvimento do modelo, pressupõe o tratamento inicial dos dados de entrada. Este tratamento foi realizado em SIG, com recurso ao software *ArcGIS Pro*. A modelação foi realizada em linguagem *python*, e com recurso à biblioteca *scikit-learn*. Para este processo, foi necessário delimitar áreas de treino e utilizar os dados nessas áreas para treinar e avaliar o modelo. Este tipo de metodologia pressupõe a repetição do processo continuamente, ajustando os vários parâmetros até que os resultados sejam satisfatórios. A partir do momento que a capacidade preditiva do modelo corresponde ao valor desejado, os resultados foram mapeados novamente em ambiente SIG, identificando e delimitando as áreas mais suscetíveis à ocorrência de fugas (Figura 1).

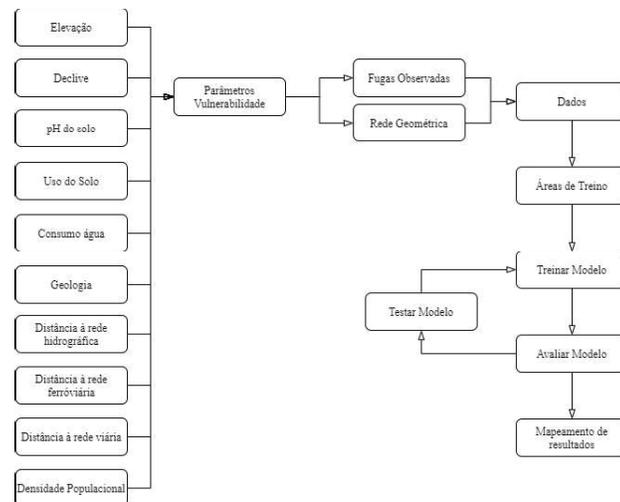


Figura 1. Parâmetros de suscetibilidade utilizados e processo de modelação.

3. Resultados e discussão

A partir da geocodificação das fugas registadas é desde logo possível identificar focos de ocorrência (Figura 2a). Assim, verifica-se uma grande concentração na área norte do município, nas freguesias mais urbanas. Identificam-se ainda outras concentrações, ainda que menos intensas, no sul do município. Esta análise preliminar é importante, uma vez que fornece um ponto de partida de comparação com os resultados produzidos pelo modelo. Durante o desenvolvimento de um modelo RFR, é possível calcular a significância de cada variável no modelo desenvolvido. A Figura 2b, representa isso mesmo no modelo final. No município, o algoritmo determinou que a distância à ferrovia é o parâmetro geográfico com maior peso para ocorrência de fugas. Aparentemente, as vibrações no solo provocadas por este tipo de transporte poderá provocar desgaste na infraestrutura e eventualmente desencadear fugas de água (Xiaozhen, 2019). Algo curioso, foi verificar a importância do pH no solo, que apesar do intervalo de variação na área de estudo ser bastante reduzido, o modelo considerou essa variação relevante. Este parâmetro tem a sua importância, uma vez que quando se trata de solos ácidos, estes aceleram o processo de corrosão na infraestrutura, tornando-se ainda mais impactantes quando a rede é constituída por materiais ferrosos. O valor mediano do pH do solo em Vila Nova de Gaia é de 4,8 (segundo os dados disponibilizados pelo ESDAC). A análise permitiu ainda identificar quais os parâmetros com pouca significância, como é o caso da litologia ou o uso do solo. Este tipo de análise é importante para identificar os parâmetros que eventualmente causam “ruído” no modelo, podendo diminuir a sua eficácia e robustez. A informação sobre a extensão da conduta, é o que assume maior significância no modelo.

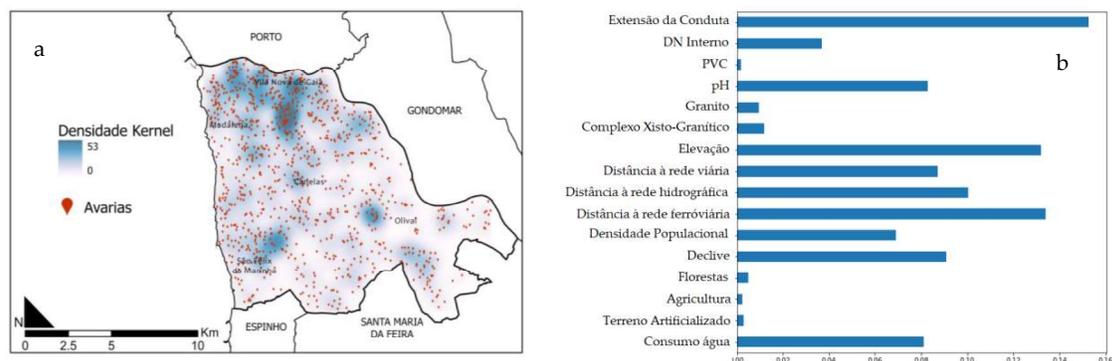


Figura 2. Análise preliminar das fugas de água na rede de distribuição de VN de Gaia: a) distribuição das fugas observadas; b) significância atribuída pelo modelo a cada parâmetro de vulnerabilidade.

Os resultados obtidos (Figura 3) atestam a utilidade das ferramentas SIG, permitindo a delimitação de áreas de interesse. Começando pela figura 3b, que apresenta os resultados diretos atribuídos pelo modelo, tendo em conta o universo de condutas, este tipo de cartografia torna-se de difícil leitura. Por isso, desenvolveu-se uma grelha onde estão associadas e contabilizadas as condutas com elevada suscetibilidade prevista (≥ 0.7). O tipo de cartografia da figura 3a permite à entidade gestora identificar as áreas que, segundo o modelo, exigem uma maior atenção. Deste modo, é possível identificar três focos principais de suscetibilidade no município, sendo eles: a área urbana a norte do município (freguesias de Mafamude, Santa Marinha e Oliveira do Douro), uma área mais a sul e litoral (freguesia de S. Félix da Marinha) e, segundo o modelo, a área mais crítica, abrange o setor este do município (União das Freguesias de Sandim, Oliva, Lever e Crestuma). Os dois principais focos acabam por ser bastante semelhantes à análise da distribuição de avarias observadas, no entanto, isto não acontece no terceiro foco, na área este do município. Este tipo de análise produzida pelo modelo acrescenta valor aos dados já existentes, identificado assim um setor que, possivelmente, é altamente suscetível a fugas e não tinha sido previamente identificado.

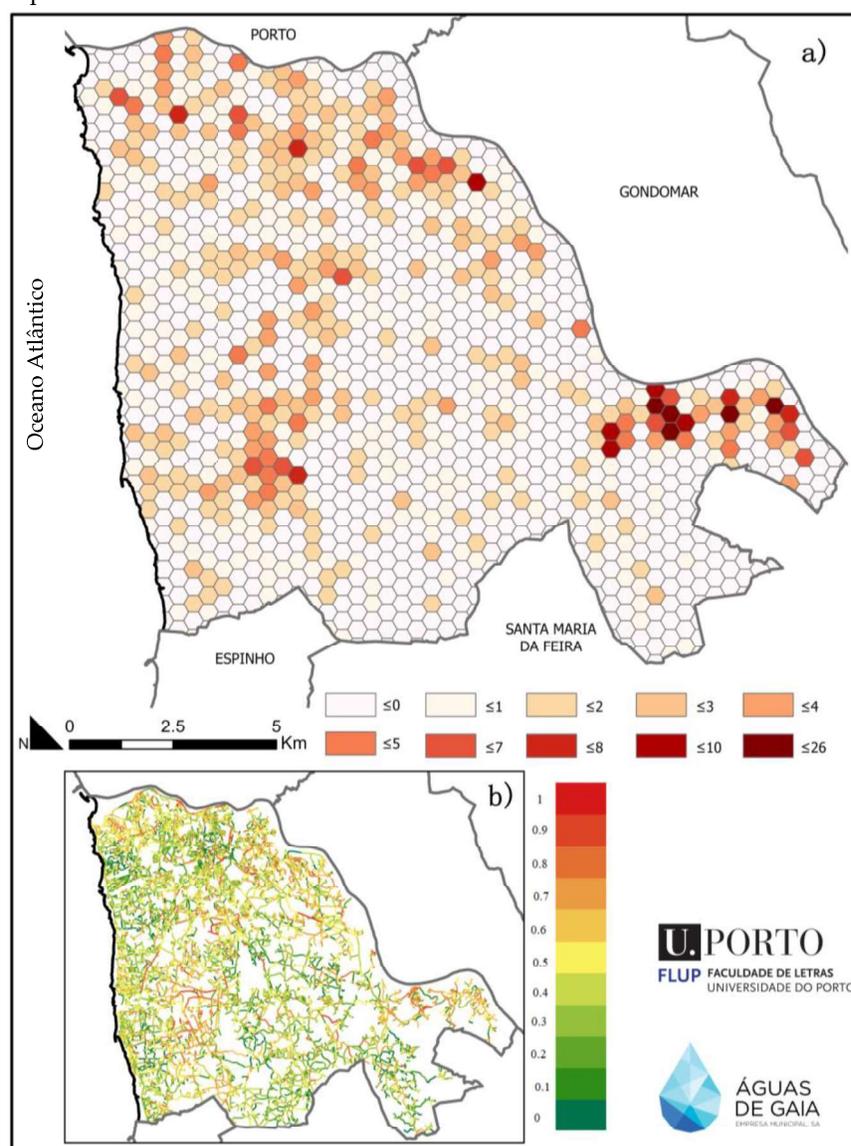


Figura 3. Suscetibilidade de ocorrência de fugas: a) agregação de condutas com elevado valor de suscetibilidade prevista; b) suscetibilidade de cada conduta prevista pela aplicação do modelo.

4. Conclusões

O algoritmo *random forest regressor* permitiu antever a suscetibilidade da infraestrutura de abastecimento de água, o que acaba por ser bastante útil de um ponto de vista de manutenção preventiva de infraestrutura. Com o modelo produzido foi possível identificar e delimitar três áreas principais com elevada suscetibilidade, estando elas localizadas na área norte do município, nas freguesias de Crestuma e Lever e a área entre a freguesia de Serzedo e o litoral-sul do município.

O desenvolvimento desta investigação permitiu entender algumas lacunas de informação importante, como a ausência de dados sobre idade da infraestrutura e a pressão em cada tubagem. No entanto, este tipo de metodologia, mesmo com ausência de alguns dados, permite a criação de modelos dinâmicos, prontos a serem melhorados a qualquer momento com o acréscimo de novos dados, com aplicação a outras áreas de estudo. Avaliando a suscetibilidade, com recurso à idade, e tendo em conta o tempo estimado de vida de cada infraestrutura, e como esta se comporta face aos parâmetros de suscetibilidade, seria possível prever com alguma certeza a ocorrência de fugas. Com este tipo de melhorias, o modelo ampliaria o poder de decisão uma vez que possibilita uma perspetiva temporal e facilita o agendamento de manutenções de rotina para evitar fugas e conseqüentemente, reduzir as perdas de água. O desenvolvimento de ferramentas de apoio à tomada de decisão, são cruciais para o bom funcionamento das infraestruturas, bem como a redução de custos e perdas de água. A análise efetuada e o modelo desenvolvido ajuda a perceber a importância dos parâmetros de suscetibilidade, mas também a identificar e delimitar áreas prioritárias de potencial ocorrência de fugas e de investimento na manutenção.

Este estudo fornece uma visão geográfica a uma área que é tradicionalmente gerida pela Engenharia, nomeadamente, a componente Civil. A Geografia e os SIG podem ser uma ferramenta importante para assistir as entidades gestoras de abastecimento de água, e a sociedade em geral, a enfrentar a crise hídrica, evitar escassez de água e contribuir para comunidades mais sustentáveis. Este estudo acaba por servir de protótipo daquilo que a Geografia consegue oferecer. Integrando este tipo de ferramentas com dados em tempo-real (de telemetria, por exemplo), conseguirá produzir resultados ainda melhores e permitir às entidades gestoras agir mais rapidamente e preventivamente. A gestão preventiva é mais eficiente que a gestão reativa.

Bibliografia

- Entidade Reguladora de Serviços de Águas e Resíduos - ERSAR. (2019). *Relatório Anual dos Serviços de Águas e Resíduos em Portugal (2020) Volume 1- Caracterização do setor de águas e resíduos*. Entidade Reguladora de Serviços de Águas e Resíduos.
- Jenkins, L., Gokhale, S., & McDonald, M. (Setembro de 2014). Comparison of Pipeline Failure Prediction Models for Water Distribution Networks with Uncertain and Limited Data. *Journal of Pipeline Systems and Engineering and Practice*.
- Liemberger, R., & Farley, M. (2004). Developing a Non-Revenue Water Reduction Strategy, Part 1: Investigating and Assessing Water Losses.
- Polikar, R. (2006). Ensemble Based Systems in Decision Making. *IEEE Circuits and Systems Magazine*, 21-45.
- Rogers, D. (2014). Leaking water networks: an economic and environmental disaster. *12th International Conference on Computing and Control for Water Industry, CCWI2013* (pp. 1421-1429). Perugia, Itália: Brunone, B.