

DESEMPENHOS DE UMA REDE NEURAL RECORRENTE E DE UM MODELO NUMÉRICO NA PREVISÃO DE NÍVEIS DE ÁGUA NO ESTUÁRIO DO RIO DOURO, PORTUGAL

Willian Melo*, José Pinho, Isabel Iglesias, Ana Bio, Paulo Avilez-Valente, José Vieira, Luísa Bastos, Fernando Veloso-Gomes

Palavras-chave: Redes neurais recorrentes, modelação numérica, aprendizagem profunda.

INTRODUÇÃO

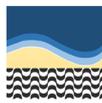
A monitorização e previsão de variáveis hidrodinâmicas é essencial para otimizar a gestão dos recursos hídricos e para implementar sistemas de previsão e alerta. Deste modo, contribui-se para aumentar a resiliência de locais em risco de inundação e para o planeamento territorial pela antecipação de potenciais problemas futuros. Normalmente, utilizam-se metodologias baseadas em modelos numéricos, que já demonstraram a sua eficiência na previsão de impactos de eventos extremos, como as cheias (Iglesias et al. 2019). Contudo, a utilização de modelos numéricos de alta resolução, que são adequados para analisar cenários futuros em escalas locais, exige elevados recursos computacionais, como capacidade de processamento e armazenamento de dados e resultados, o que limita a aplicação destes métodos em plataformas de previsão e alerta, sobretudo para previsão em tempo real. Assim, este trabalho propõe o estudo de modelos baseados em dados, também conhecidos como emuladores de modelos, como as redes neurais recorrentes (RNN - Recurrent Neural Network), com o objetivo de comparar o desempenho de um modelo numérico com dois modelos de RNNs para a previsão do nível de água em Cais dos Banhos, no estuário do rio Douro, Portugal.

METODOLOGIA

Foram utilizados registos de níveis da água na estação de Cais dos Banhos e de caudais efluentes na barragem de Crestuma-Lever (base de dados do Sistema Nacional de Informação de Recursos Hídricos da Agência Portuguesa do Ambiente), necessários para configurar os modelos criados. Foram ainda utilizados dados da batimetria do estuário do Douro, provenientes do Instituto Hidrográfico.

O modelo numérico do estuário do rio Douro foi configurado no *software* Delft3D (D3D) e apresenta duas fronteiras abertas, uma fluvial localizada em Crestuma-Lever, onde se impõem os caudais efluentes da barragem e a outra no Oceano Atlântico, onde se consideram os níveis de maré estimados com base nos respetivos constituintes harmónicos. Ressalta-se que esta definição da fronteira oceânica não considera os níveis reais que podem ser afetados pelas marés meteorológicas, que, em situações de baixas pressões e/ou agitação energética, pode alcançar níveis pelo menos 0.5 metro superior. Para calibrar o modelo, utilizou-se a ferramenta OpenDA (Garcia et al. 2015), que permite ajustar os valores dos parâmetros automaticamente, através da minimização das diferenças entre os dados observados e os resultados do modelo. Foram calibrados constituintes harmónicos da maré e o coeficiente de rugosidade de Manning.

Para os emuladores, optou-se por utilizar a arquitetura de RNN com memória de longo prazo, conhecida por LSTM, atendendo à sua adaptabilidade para tratar séries de dados temporais (Aggarwal 2018). Utilizou-se, de forma a facilitar a configuração das RNNs, o framework *Deep Learning Toolbox*, disponível no MATLAB. As RNNs foram configuradas com cinco camadas, sendo uma camada de entrada de dados, duas camadas totalmente conectadas, uma camada LSTM e uma camada de regressão.



A primeira RNN (WL) prevê o nível da água no Cais dos Banhos a partir da elevação de água da hora anterior e a segunda RNN (WL-Q) considera, além do nível anterior, o caudal de CrestumaLever do mesmo período. Para o modelo WL, utilizaram-se cerca de 200 mil observações de nível no Cais dos Banhos, com frequência horária, para o período de 2002 a 2015, sendo que 75% destas foram utilizadas para treinar a rede e 25% para a validar. Para o modelo WL-Q, utilizaram-se cerca de 70 mil observações, mantendo as proporções de dados de treino/validação. Ressaltase que o número de dados utilizados no segundo modelo foi menor, pois havia menos dados disponíveis de caudal, pelo que a análise se restringiu apenas aos dados de Cais dos Banhos com datas coincidentes.

Por fim, os modelos foram testados na simulação de uma cheia ocorrida a 11 de janeiro de 2016, cujo caudal de pico foi de 7300 m³/s, o equivalente a uma cheia com período de retorno de cerca de 9 anos, em conformidade com a respetiva distribuição de Gumbel dos caudais de cheia.

RESULTADOS

A figura 1 apresenta os resultados de níveis da água em Cais dos Banhos obtidos pelos modelos durante a cheia de janeiro de 2016. Observa-se que o modelo que mais se aproximou dos valores observados foi o WL. O D3D foi o que apresentou pior resultado nas previsões, porém foi o segundo melhor a prever o nível máximo alcançado. Apesar disso, verifica-se um desvio significativo dos resultados do D3D após a ocorrência do nível máximo, o que não ocorreu com os resultados obtidos pelos emuladores.

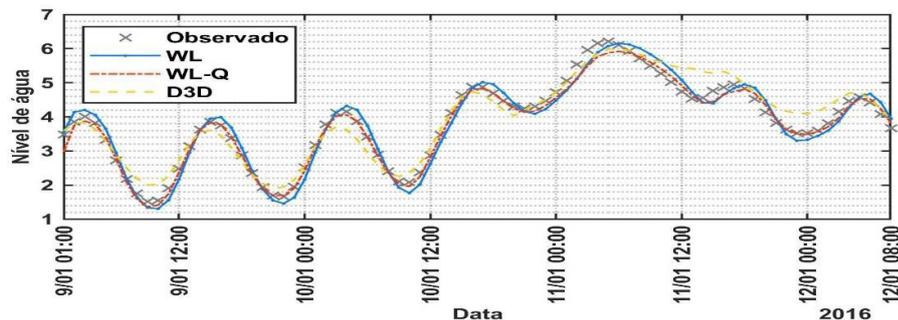


Figura 1. Comparação dos resultados de níveis da água dos modelos com os níveis observados.

A figura 2 apresenta a diferença entre os níveis previstos pelos modelos e o nível observado, bem como três métricas que permitem comparar os modelos: a raiz quadrada da média dos erros quadrados (RMSE), o viés e o coeficiente de correlação de Pearson (p). No gráfico é possível notar que o modelo D3D sobrestimou a maioria dos valores observados, principalmente após o nível máximo. Esta afirmação é confirmada pelo viés do modelo de 0.099 m. Os dois emuladores, por sua vez, subestimaram os níveis observados, sendo que o viés do modelo WL foi de -0.017 m e o do WL-Q foi de -0.064 m. Além disso, a comparação dos coeficientes de Pearson confirma que o modelo que melhor simulou o período de cheia analisado foi o WL.

O melhor desempenho na previsão pelo modelo WL deve-se muito provavelmente à maior quantidade de dados disponíveis para treinar o algoritmo. O desempenho de uma rede neural é melhor quanto maior for o conjunto de dados disponíveis na etapa de treino. Deste modo, apesar do modelo WL considerar apenas o nível da água na hora anterior, a quantidade de dados disponíveis foi suficiente para compensar a falta de outras informações que influenciam o nível em Cais dos Banhos, como o caudal efluente em Crestuma, considerado pelo modelo WL-Q. O desempenho do modelo WL-Q foi superior ao do D3D, o que mostra que a aplicação de métodos de emulação pode alcançar bons resultados na previsão de variáveis hidrodinâmicas.

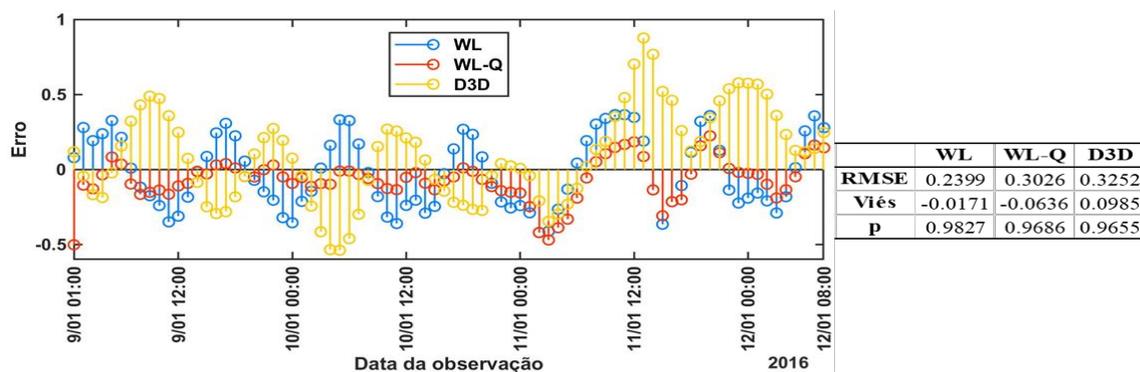
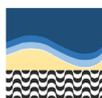


Figura 2. Erro dos modelos nas previsões de nível em Cais dos Banhos.

CONCLUSÕES

Este trabalho comparou o desempenho de um modelo numérico com o desempenho de duas RNNs com memória de longo prazo na previsão de níveis de água no Cais dos Banhos, no estuário do rio Douro. Enquanto os modelos numéricos consideram o processo físico subjacente ao fenómeno em estudo, dependendo de informações como a batimetria, coeficiente de atrito, entre outros parâmetros, os emuladores consideram apenas a relação matemática entre os dados de entrada e saída, considerando apenas de forma implícita as leis físicas que regem o fenómeno em estudo.

Assim, afirma-se que as RNNs apresentaram desempenho superior ao modelo numérico aplicado na previsão de nível no Cais dos Banhos, sendo alternativas interessantes para estudar variáveis hidrodinâmicas em áreas onde não existam informações suficientes que permitam configurar um modelo numérico. Contudo, o desempenho dos emuladores é afetado diretamente pela quantidade de dados disponíveis para treinar o algoritmo, o que realça a importância da monitorização de variáveis hidrodinâmicas para a configuração de modelos preditivos.

AGRADECIMENTOS

Esta contribuição foi financiada pelo projeto EsCo-Ensembles (PTDC/ECI-EGC/30877/2017), cofinanciado pelo Norte2020, por Portugal 2020 e pela União Europeia através do FEDER e pela FCT através de fundos nacionais e da bolsa no âmbito da bolsa de doutoramento com referência SFRH/BD/151383/2021 de Willian Melo. Foi ainda apoiada por verbas nacionais através da FCT – Fundação para a Ciência e Tecnologia, no âmbito de UIDB/04423/2020 e UIDP/04423/2020.

REFERÊNCIAS

- Aggarwal, Charu C. 2018. *Neural Networks and Deep Learning. Artificial Intelligence*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-94463-0>.
- Garcia, Mariangel, Isabel Ramirez, Martin Verlaan, and Jose Castillo. 2015. 'Application of a Three-Dimensional Hydrodynamic Model for San Quintin Bay, B.C., Mexico. Validation and Calibration Using OpenDA'. *Journal of Computational and Applied Mathematics* 273 (January): 428–37. <https://doi.org/10.1016/J.CAM.2014.05.003>.
- Iglesias, Isabel, S. Venâncio, J. L. Pinho, P. Avilez-Valente, and J. M. P. Vieira. 2019. 'Two Models Solutions for the Douro Estuary: Flood Risk Assessment and Breakwater Effects'. *Estuaries and Coasts* 42 (2): 348–64. <https://doi.org/10.1007/s12237-018-0477-5>.