

PREVISÃO DE DEMANDA DE ÁGUA EM REGIÃO SEMIÁRIDA UTILIZANDO REDES NEURAIS COM PEQUENA BASE DE DADOS

MIRELLE TAINÁ VIEIRA LIMA¹; MARIA LUCINEIDE GOMES DA SILVA²; INGRID DE OLIVEIRA CAVALCANTE LIMA³; JULIANA PERTILLE DA SILVA⁴; FABRÍCIO DA SILVA TERRA⁵

¹Universidade Federal de Pelotas (UFPel) – mirellet.vieira@gmail.com

²Universidade Federal da Paraíba (UFPB) – lucineidegmd@gmail.com

³Universidade Federal do Rio Grande (FURG) – ingrid.limaoc@hotmail.com

⁴Universidade Federal de Pelotas (UFPel) – juliana.pertill@gmail.com

⁵Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) – fabricio.terra@ufvjm.edu.br

1. INTRODUÇÃO

O aumento da demanda humana por recursos hídricos tem impulsionado estudos de previsão de curto e longo prazo do consumo de água. Em nível local, esses estudos têm se mostrado primordiais para otimização dos processos de operação, planejamento, gestão, sustentabilidade e modernização dos sistemas públicos de abastecimento e distribuição de água (BATA et al., 2020, ZUBAIDI et al., 2019).

Existem diversas abordagens para a modelagem do consumo de água. Dentre essas, estão os algoritmos de inteligência artificial, cuja eficiência na previsão e modelagem de séries temporais tem sido demonstrada em diversos estudos científicos em diversos campos de aplicações (BATA et al., 2020, COMESAÑA et al., 2020; GUO et al, 2023; PEREIRA et al. 2019, ZUBAIDI et al., 2019). As Redes Neurais Artificiais (RNAs) têm sido amplamente utilizadas para a previsão da demanda por água, por apresentar a vantagem de modelar dados não lineares (HORNIK et al., 1989).

Neste contexto, o município de Juazeiro do Norte destaca-se por apresentar a maior taxa de crescimento urbano, demográfico, comercial e industrial do interior do estado do Ceará, com forte apelo ao turismo religioso e ecológico (PEREIRA, CARDOSO, 2020). Todo esse crescimento tem impulsionado o consumo de água nesta localidade de clima semiárido, que depende quase que exclusivamente da água subterrânea para seu abastecimento. De modo que, realizar previsões da demanda por água torna-se essencial para a adequada gestão deste recurso. Assim, o presente trabalho busca realizar previsões de curto prazo do consumo de água no município de Juazeiro do Norte, utilizando uma base de dados pequena, testando os modelos de Redes Neurais Artificiais (RNAs): *multi-layered perceptron* (MLP) e *long short-term memory* (LSTM), bem como, a influência de bases de dados univariados e multivariados.

2. METODOLOGIA

Os dados mensais de consumo de água entre 2004 e 2022 foram eletronicamente solicitados à Companhia de Gestão dos Recursos Hídricos (COGERH). Considerando que a previsão da demanda urbana de água pode envolver diferentes variáveis e periodicidades, neste estudo foram admitidas as possíveis variáveis exógenas, levantadas em escala mensal no mesmo período da série de consumo hídrico, com o objetivo de adicionar mais informações aos modelos: número de outorgas e ligações de esgoto ativas (COGERH, 2023), população estimada (IPECE, 2018), área urbanizada km² (MAPBIOMAS, 2022),

pluviosidade média (FUNCEME, 2023) e o número de romarias¹ e quantidade de dias de romaria definidas pela lei estadual nº 16.927 de julho de 2019.

A identificação dos outliers foi realizada pelo método dos quartis, em que se identificou a necessidade de correção de *outliers* na série histórica de pluviosidade. Para este estudo, os dados foram padronizados entre os valores de 0 e 1 pelo método *z-score*, a fim de equiparar variáveis de unidades e escalas muito diferentes (GUO; XIE; LI, 2023). A identificação da tendência nas séries temporais deu-se pelo teste Mann Kendall (pacote R "Kendall"), onde realizaram-se diferenciações nas séries temporais com p-valor inferior a 0,05.

Neste trabalho, foram utilizadas as seguintes arquiteturas de redes neurais: rede *Feedforward Multi-layered Perceptron* (MLP) e rede recorrente *Extreme Learning Machine* (ELM), ou máquina de aprendizagem extrema. Os modelos foram treinados de forma univariada (apenas com a variável consumo) e multivariada (complementando a base de dados com as variáveis exógenas). Tendo em vista a pequena quantidade de dados e o objetivo de previsões a curto prazo, as séries foram divididas em 95% para treino e 5% para validação. Todo o procedimento foi executado no ambiente RStudio (RSTUDIO TEAM, 2020), utilizando principalmente o pacote R keras (KALINOWSKI et al., 2023).

Visando comparar a eficiência dos modelos em prever a demanda por água, foram utilizadas as seguintes métricas estatísticas: coeficiente de determinação (R^2), média do erro absoluto (Mean Absolute Error - MAE) e raiz do erro médio quadrado (Root Mean Square Error - RMSE). O método com as melhores métricas foi utilizado para prever o consumo de água no período de 12 meses.

3. RESULTADOS E DISCUSSÃO

A série temporal de consumo de água apresentou valores mínimo e máximo de 680.695 e 1.088.127 m³, respectivamente, com média mensal de 832.500,6 m³ e desvio padrão igual a 86.038,7 m³. De modo geral, as variáveis consumo de água, ligações ativas, outorgas ativas, área urbana e população total apresentaram tendência positiva (conforme teste Mann Kendall) e foram corrigidas no processo de diferenciação. Em relação aos modelos implementados, os melhores resultados foram alcançados pelo modelo LSTM (Tabela 1).

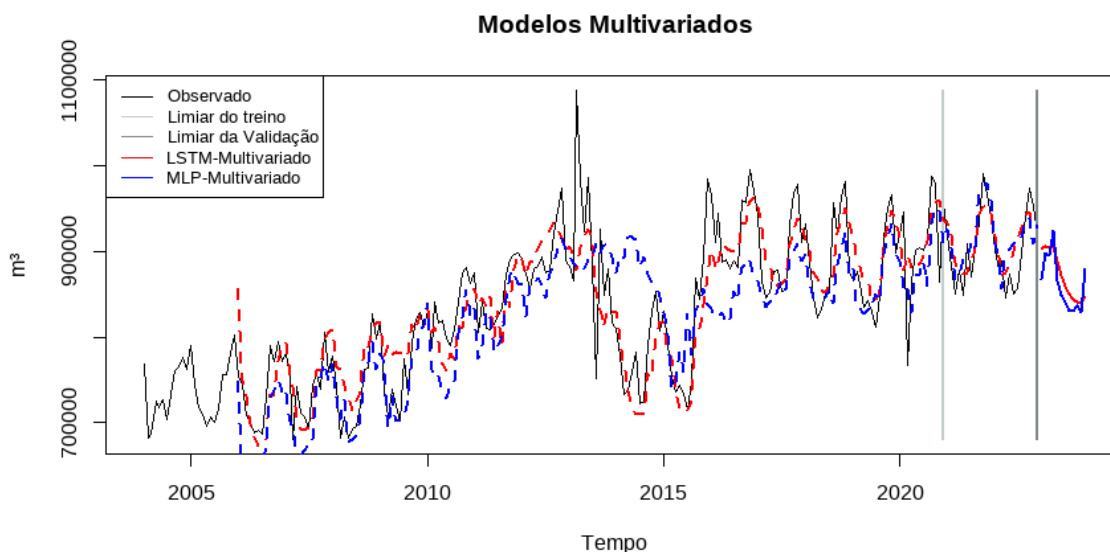
Tabela 1. Métricas de desempenho dos modelos.

Modelo	RMSE		R²		MAE	
	Treino	Validação	Treino	Validação	Treino	Validação
<i>MLP</i>						
Univariado	73245	44128	0.25	0.06	59408	36173
Multivariado	57449	35261	0.59	0.29	42424	29626
<i>LSTM</i>						
Univariado	67191	24931	0.39	0.69	52603	20705
Multivariado	40699	28303	0.77	0.62	30254	23571

Verificou-se que os conjuntos de dados multivariados apresentaram melhor desempenho que os univariados, por se ajustarem melhor ao comportamento da série original, aumentando a correlação e minimizando os erros (Figura 1).

¹ Compreendendo romarias como um evento religioso de devoção, marcada pela peregrinação, viagens individuais ou coletivas a lugares considerados sagrados, estando ausente nos anos de pandemia de 2020 e 2021.

Figura 1. Previsões de consumo (m^3) com a rede LSTM.



Nota-se que o consumo de água caracterizou-se pelo aumento da tendência positiva entre o início de 2010 até março de 2013, quando iniciou uma redução até meados de 2015. Vale ressaltar que, o período entre 2012 e 2017 marcou uma das grandes secas da história recente do estado do Ceará, onde a demanda foi reprimida inclusive devido à redução de liberação de outorgas para o uso da água. Além disso, a série apresenta comportamento sazonal, em que os picos de consumo, geralmente, ocorrem nos meses mais secos e quentes (setembro a novembro). Complementa-se que este período coincide com 4 das principais romarias realizadas anualmente na cidade (Romaria de Nossa Senhora das Dores; de São Francisco; de Finados; e Ordenação do Padre Cícero). O destaque para tais eventos decorre do consequente aumento na demanda de recursos hídricos, visto a expressividade da explosão de consumidores, decorrente do número de participantes, a citar, na romaria de finados, estima-se um fluxo de cerca de 600 mil romeiros (VASCONCELOS, 2015).

4. CONCLUSÕES

O presente trabalho apresentou a comparação entre modelos de redes neurais dos tipos *feedforward* (MLP) e recorrente (LSTM), para a previsão do consumo de água em localidade semiárida, treinada com um conjunto de dados univariado de consumo de água) e um conjunto multivariado (onde foram adicionadas variáveis exógenas prováveis de explicar o comportamento do consumo de água). No geral, o modelo LSTM apresentou melhor desempenho que o MLP. Nota-se que a inserção das variáveis exógenas contribuiu para o desempenho da modelagem, especialmente na fase de testes. Dessa forma, este trabalho evidencia a importância de se agregar variáveis exógenas a pequenas bases de dados, especialmente, verificando variáveis ligadas à cultura e realidade local, como foi incluído aqui o aspecto religioso. Ressalta-se que, apesar de ser um assunto de extrema preocupação, não foram identificados estudos científicos prevendo a demanda por água na presente área de estudo.



5. REFERÊNCIAS BIBLIOGRÁFICAS

- BATA, M. H.; CARRIVEAU, R.; TING, D. S.-K. Short-Term Water Demand Forecasting Using Nonlinear Autoregressive Artificial Neural Networks. **Journal of Water Resources Planning and Management**, v.146, n.3, p.1-9, 2020. doi:10.1061/(asce)wr.1943-5452.0001165
- COGERH. Companhia de Gestão dos Recursos Hídricos. Outorgas concedidas e vigentes. Disponível em: http://outorgasvigentes.cogerh.com.br/paginaSemValidacao/outorgaVigente/outorg_as_fh.xhtml. Acesso em: agosto de 2023.
- COMESAÑA, M. M.; FEBRERO-GARRIDO, L.; TRONCOSO-PASTORIZA, F.; MARTÍNEZ-TORRES, J. Prediction of building's thermal performance using LSTM and MLP neural networks. **Applied Sciences**, v.10, n.21, p.7439; doi:10.3390/app10217439
- FUNCENE. Fundação Cearense de Meteorologia e Recursos Hídricos. Postos pluviométricos. [S.I.]. Disponível em: http://www.funceme.br/produtos/script/chuvas/Download_de_series_historicas/DownloadChuvasPublico.php. Acesso em: Agosto de 2023.
- GUO, X.; XIE, W.; LI, X. Spatial-Temporal Correlation Neural Network for Long Short-Term Demand Forecasting During COVID-19. **IEEE Access**, 2023.
- IBGE. Instituto Brasileiro de Geografia e Estatística. Censo demográfico 2022: População e domicílios, 2023.
- IPECE. Instituto de Pesquisa e Estratégia Econômica do Ceará. Perfil Municipal. IPECE: Fortaleza, 2018.
- KALINOWSKI, T.; FALBEL, D.; ALLAIRE, J. J.; CHOLLET, F.; TANG, Y.; VAN DER BIJL, W.; ... KEYDANA, S. R Interface to 'Keras'. 2023.
- MAPBIOMAS. Série Anual de Mapas de Cobertura e Uso do Solo do Brasil. Coleção 7, 2022.
- PEREA, R.G.; CAMACHO POYATO, E.; MONTESINOS, P.; DÍAZ, J.A.R. Optimisation of water demand forecasting by artificial intelligence with short data sets. **Biosystems Engineering**. v. 177, p. 59-66, 2019. doi:10.1016/j.biosystemseng.2018.03.011
- PEREIRA, C. E.; CARDOSO, P. H. G. Contribuição das Instituições de Ensino Superior no processo de crescimento e desenvolvimento da Região Metropolitana do Cariri Ceará Brasil. **Research, Society and Development**, v. 9, n. 3, p. e142932531-e142932531, 2020.
- RStudio Team. RStudio: Integrated Development for R. RStudio, PBC, Boston, 2020.
- VASCONCELOS, M. P.; BRAGA, C.; GOUVEIA, G. C. e SOUZA, W. V. Romarias no município de Juazeiro do Norte, Ceará: perfil da demanda por atendimento de saúde e sazonalidade de internações. **Epidemiol. Serv. Saúde**, Brasília , v. 24, n. 1, p. 39-48, mar. 2015 .
- ZUBAIDI, S. L.; AL-BUGHARBEE, H.; MUHSEN, Y. R.; HASHIM, K.; ALKHADDAR, R. M.; HMEESH, W. H. The prediction of municipal water demand in Iraq: a case study of Baghdad governorate. In: **2019 12th International Conference on Developments in eSystems Engineering (DeSE)**. IEEE, 2019. p. 274-277.