

SÍNTESE DE ESTUDO COMPARATIVO DE ESTRATÉGIAS DE SÉRIES TEMPORAIS NA PREVISÃO DE DADOS DE TEMPERATURA

LETICIA BARROS DIAS SOARES¹; JOÃO INÁCIO MOREIRA BEZERRA²;
NATASSIA RAFAELLE MEDEIROS SIQUEIRA³
SOFIA PAGLIARINI⁴; MARLON MAURICIO HERNANDEZ CELY⁵; JAIRO VALÕES
DE ALENCAR RAMALHO⁶

¹*Universidade Federal de Pelotas – LeticiaBarros1996@yahoo.com.br*

²*Universidade Federal de Pelotas – jimbezerra@inf.ufpel.edu.br*

³*Universidade Federal de Pelotas – nrmsiqueira@inf.ufpel.edu.br*

⁴*Universidade Federal de Pelotas – sofiapagliarini@gmail.com*

⁵*Universidade Federal de Pelotas – marlon.cely@ufpel.edu.br*

⁶*Universidade Federal de Pelotas – jairo.ramalho@ufpel.edu.br*

1. INTRODUÇÃO

A energia geotérmica é uma forma de energia renovável que é gerada a partir do calor armazenado no interior da Terra. Esse calor é uma consequência dos processos geológicos naturais. O principal uso da energia geotérmica é a geração de eletricidade. Por exemplo, o sistema de Trocador de Calor Solo-Ar (TCSA), aproveita essa energia ao utilizar ventiladores de baixa potência que circulam por dutos subterrâneos, para o aquecimento e resfriamento de residências e edificações, e contribui para a redução do consumo de energia dos sistemas convencionais de ar-condicionado (Liu et al., 2023). A eficiência desse sistema depende de dados sobre as temperaturas do ar e do solo, que podem ser obtidos através de modelagem matemática ou aplicação de algoritmos computacionais.

A previsão de séries temporais constitui um campo de relevância central na análise de dados, pois trata de informações organizadas sequencialmente em função do tempo. Nesse contexto, valores históricos servem como base para inferir tendências e comportamentos futuros, explorando a dependência temporal que caracteriza esse tipo de dado. Quando o objetivo é estimar apenas o valor imediatamente subsequente, trata-se de um problema de previsão de etapa única (*one-step*); já quando múltiplos pontos futuros são projetados simultaneamente, o problema é classificado como previsão de múltiplas etapas (*multi-step*) Navarro (2023). As redes neurais artificiais podem ser compreendidas como modelos computacionais inspirados em mecanismos biológicos de processamento de informação Graves (2024). Arquiteturas específicas como as Redes Neurais Recorrentes (RNN) e, em especial, a variante Long Short-Term Memory (LSTM), consolidaram-se como alternativas robustas para a modelagem de dependências temporais em sequências de dados, apresentando desempenho consistentemente.

Este trabalho tem como objetivo prever e comparar os dados de temperatura utilizando modelos de séries temporais com aplicação de redes neurais recorrentes. Os dados utilizados foram obtidos por meio de um sistema de monitoramento remoto, no intervalo compreendido entre julho de 2023 e abril de 2024, em parceria com a Universidad Tecnológica del Uruguay (UTECH), localizada na cidade de Rivera, Uruguai.

2. METODOLOGIA

2.1 Coleta de dados

Os dados de temperatura foram coletados por sensores instalados próximos à superfície do solo e na profundidade de 2 metros. Como o sistema de monitoramento requer manutenção periódica, pode haver períodos em que os dados não estejam disponíveis. Os dados foram organizados por hora cumulativa, resultando em uma base de dados composta por 6056 registros, abrangendo um período contínuo de dez meses. Além disso, os dados passaram por pré-tratamento para remover valores inconsistentes e identificação de períodos de falha no monitoramento.

2.2 Escolha do modelo

Para a construção dos modelos preditivos, optou-se por empregar o método de LSTM (Long Short-Term Memory), devido à sua capacidade de capturar padrões de longo e curto prazo em séries temporais complexas. A abordagem multi-step com adição de features temporais (seno/cosseno da hora) e estatísticas (médias móveis) foi adotada para explorar relações não-lineares e sazonais presentes nos dados.

2.3 Pré-processamento

O processo de preparação dos dados incluiu etapas fundamentais para otimizar o desempenho dos modelos: Classificação de Furos: Os dados faltantes foram categorizados conforme sua extensão temporal: microfuros (1-48 horas), pequenos (49-120 horas), médios (121-216 horas) e grandes furos (>216 horas). Interpolação Inicial: Para furos de menor dimensão (micro e pequenos), aplicou-se interpolação linear bidirecional, preservando os padrões locais imediatos. Features Temporais: Implementou-se transformação cíclica das horas através de funções seno e cosseno para representar adequadamente a natureza circular do tempo, evitando descontinuidades entre 23h e 0h. Features Estatísticas: Calculou-se médias móveis de 6 e 24 horas, além da diferença de primeira ordem, para capturar tendências de curto e médio prazo.

A arquitetura neural desenvolvida comprehende: Camada de Input: Recebe sequências temporais de 24-72 horas com 6 features. Camadas LSTM: Duas camadas recorrentes (128 e 64 unidades) com dropout de 20% para regularização. Camada de Output: Saída multi-step (3-6 horizontes) para previsão simultânea de múltiplos pontos. Otimização: Algoritmo Adam com early stopping baseado em validação monitorando perda MSE

Foram implementadas e comparadas duas abordagens comparativas e diferentes para preenchimento de dados faltantes, como mostra a tabela 1.

Tabela 1: Configurações dos Modelos Comparados

Aspectos	Abordagem 1	Abordagem 2
Tipo de previsão	Univariada (temperatura)	Multivariada (temperatura + features temporais)
Features	Temperatura	Temperatura, sen/cos hora/dia, médias móveis, diferenças
Janela Temporal	70 horas	70 horas

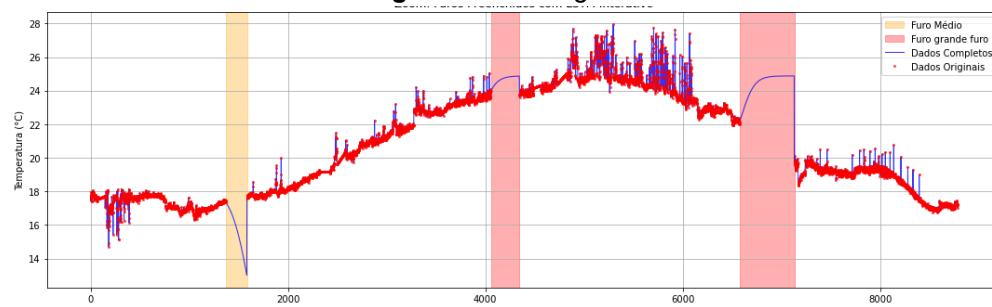
Arquitetura (LSTM)	LSTM simples, 50+50 unidades, dropout 0.2	Bidirectional LSTM, camadas profundas, dropout 0.3/0.2
Divisão dos dados	80% treino / 20% teste	80% treino / 20% teste
Métrica de avaliação	MAE, RMSE	MAE, RMSE

Fonte: Presente Estudo

3. RESULTADOS E DISCUSSÃO

A título de ilustração, na figura 1, a análise dos resultados evidencia diferenças substanciais entre os modelos implementados. A Abordagem 1, de caráter univariado, apresentou limitações na representação dos dados de temperatura, conforme ilustrado na Figura 1. Observa-se que o modelo não conseguiu captar de forma satisfatória as flutuações inerentes à série temporal. Esse comportamento sugere que a ausência de variáveis auxiliares compromete a capacidade preditiva da rede, resultando em previsões com elevada discrepância em relação aos valores observados.

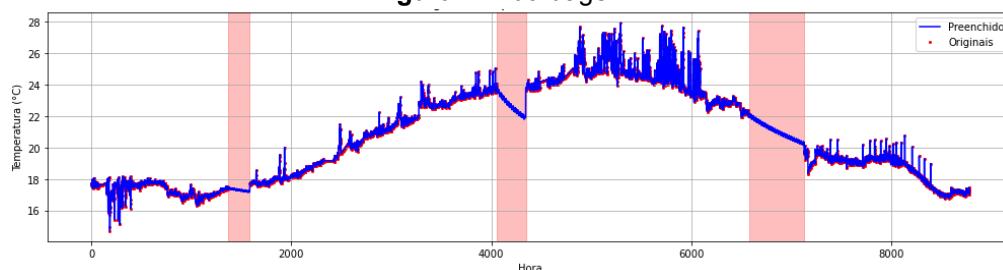
Figura 1: Abordagem 1



Fonte: Presente estudo

Em contraste, a Abordagem 2, de natureza multivariada, revelou desempenho superior (Figura 2). Ao implementar features temporais, como transformações cíclicas das horas e médias móveis, o modelo foi capaz de identificar padrões mais complexos presentes na série, acompanhando com maior fidelidade as variações reais da temperatura ao longo do tempo.

Figura 2: Abordagem 2



Fonte: Presente estudo

Os resultados quantitativos corroboram essa superioridade, conforme apresentado na Tabela 2, houve uma redução de 47% no erro médio absoluto (MAE) e de 38% na raiz do erro quadrático médio (RMSE), indicadores que

atestam ganhos expressivos de acurácia. Do ponto de vista teórico, os resultados alinham-se a investigações recentes que destacam a superioridade de arquiteturas profundas no tratamento de séries complexas (Navarro et al., 2023). A capacidade das redes LSTM de capturar dependências de longo e curto prazo torna-se ainda mais relevante quando associada a variáveis sazonais e estatísticas, como demonstrado neste estudo. Com base nos resultados, temos então na tabela 2, a comparativa detalhada entre as duas abordagens.

Tabela 2: Análise Comparativa

Métricas	Abordagem 1	Abordagem 2	Melhoria (%)
MAE	0.2953°C	0.1562°C	47
RMSE	0.3257°C	0.2017°C	38

Fonte: Presente Estudo

4. CONCLUSÕES

Este estudo apresenta sistematização crítica de dados e na forma como representam os componentes temporais, que fomentam a importância de sistematizar os dados para uma qualidade da previsão em séries temporais. A implementação de features temporais cíclicas e estatísticas móveis supera abordagens mais básicas, atingindo índices de precisão maiores. Apesar dos resultados serem satisfatórios ao presente estudo, reforçam que a pesquisa está contribuindo para o desenvolvimento de modelos mais eficientes na detecção de mudanças nos padrões de temperatura. Futuramente, os modelos serão testados em diferentes algoritmos e base de dados, para validar a robustez das previsões, proporcionando um melhor entendimento das variações de temperatura e ampliando suas aplicações práticas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

CHOLLET, F. **Deep Learning with Python**. 2. ed. Manning Publications, 2022.

GRAVES, A. Long Short-Term Memory. In: **SUPERVISED SEQUENCE LABELLING WITH RECURRENT NEURAL NETWORKS**. Studies in Computational Intelligence, v. 385. Berlin, Heidelberg: Springer, 2012. Available at: https://doi.org/10.1007/978-3-642-24797-2_4. Accessed on: 23 Jul. 2024.

JIMÉNEZ-NAVARRO, M.J.; MARTÍNEZ-BALLESTEROS, M.; MARTÍNEZ-ÁLVAREZ, F.; ASENCIO-CORTÉS, G. PHILNet: A novel efficient approach for time series forecasting using deep learning. *Information Sciences*, Volume 632, p.815-832, 2023.

LIU, Z.; XIE, M.; ZHOU, Y.; HE, Y.; ZHANG, L.; ZHANG, G.; CHEN, D. A state-of-the-art review on shallow geothermal ventilation systems with thermal performance enhancement system classifications, advanced technologies and applications. *Energy and Built Environment*, v.4, n.2, p.148–168, 2023.