



# Modelagem de Séries Temporais de temperatura com aplicação de Deep Learning - LSTM: Um estudo de caso em Rivera, Uruguai

**Leticia Barros Dias Soares<sup>\*1</sup>**; Geilson de Almeida Soares<sup>1</sup>; Pamela Gabriella Barboza Pereira<sup>2</sup>; João Inácio Moreira Bezerra<sup>3</sup>; Jairo Valões de Alencar Ramalho<sup>1</sup>; Marlon Mauricio Hernandez Cely<sup>4</sup>; Honório Joaquim Fernando<sup>5</sup>; Nibia Raquel Navarro Pastor<sup>2</sup>; Marcos Rafael Burlón Olivera<sup>2</sup>; Wilmar Armando Pineda Castiblanco<sup>2</sup>; Cindy Jireth Ortiz Gamba<sup>2</sup>

**Resumo:** A energia geotérmica superficial é uma fonte promissora de energia renovável, oferecendo soluções sustentáveis para o aquecimento e resfriamento de edificações. Um exemplo disso são os Trocadores de Calor Solo-Ar (TCSA), que utilizam ventiladores de baixa potência para circular ar por dutos subterrâneos, reduzindo o consumo de energia de sistemas convencionais de ar-condicionado. A eficiência dos TCSA depende de dados sobre as temperaturas do ar e do solo, obtidos por meio de modelagem matemática e técnicas de inteligência artificial. Esses modelos, inspirados no funcionamento do cérebro humano, são capazes de aprender padrões complexos em grandes volumes de dados. A previsão de séries temporais é essencial para a análise de dados, pois envolve sequências coletadas ao longo do tempo, permitindo identificar padrões e prever eventos futuros. Com o avanço da aprendizagem profunda, modelos como Redes Neurais Recorrentes

---

<sup>\*</sup> [leticiabarros1996@yahoo.com.br](mailto:leticiabarros1996@yahoo.com.br)

<sup>1</sup> Universidade Federal de Pelotas (UFPel)/Instituto de Física e Matemática (IFM).

<sup>2</sup> Universidad Tecnológica (UTEC) del Uruguay/Instituto Tecnológico Regional del Norte (ITRN).

<sup>3</sup> Universidade Federal de Pelotas (UFPel)/Centro de Desenvolvimento Tecnológico (CDTEC).

<sup>4</sup> Universidade Federal de Pelotas (UFPel)/Centro de Engenharia (CENGE).

<sup>5</sup> Universidade Federal Fluminense (UFF)/ Instituto de Ciências Exatas (ICEx).

(RNN) e Long Short-Term Memory (LSTM) têm superado métodos tradicionais na modelagem de dependências temporais em dados sequenciais. Este trabalho busca prever o comportamento das temperaturas do ar e do solo em Rivera (Uruguai), com dados coletados entre julho de 2023 e abril de 2024 por um sistema de monitoramento remoto, em parceria com a Universidad Tecnológica del Uruguay (UTEC). Em particular, são avaliadas as técnicas de LSTM e Regressão Linear para prever e recuperar dados de temperatura ausentes devido a falhas nos sensores ou manutenção. Os resultados indicam que ambas as técnicas recuperam os dados de forma satisfatória, sendo a regressão linear mais precisa nos casos testados.

**Palavras-chave:** Inteligência Artificial; LSTM; Regressão Linear; Modelagem de Temperatura; Energias Renováveis.

## 1 INTRODUÇÃO

Segundo os dados da Organização Meteorológica Mundial (OMM)<sup>6</sup>, os anos de 2020 e subsequentes estão entre os mais quentes já registrados, refletindo um padrão alarmante de aquecimento global. Este aumento nas temperaturas médias globais está associado a uma série de consequências ambientais, incluindo o derretimento das calotas polares, o aumento do nível do mar e a intensificação de eventos climáticos extremos, como furacões e secas severas. Essas características não impactam apenas o meio ambiente, mas também têm repercussões diretas na economia global e no bem-estar humano.

Segundo RABINO et al. (2022), diante desse cenário, a transição para fontes de energia limpa torna-se importante. A busca por energias renováveis, como a energia geotérmica, hidrelétrica e biomassa, está em crescente desenvolvimento. Essas fontes oferecem uma alternativa viável às energias fósseis, que são grandes responsáveis pelas emissões de gases de efeito estufa. A utilização de energias renováveis não apenas ajuda a mitigar os impactos das mudanças climáticas, mas também movimenta o setor econômico.

A energia geotérmica superficial é uma fonte de energia renovável promissora, oferecendo soluções sustentáveis para o aquecimento e resfriamento de edificações. Por exemplo, o sistema de Trocador de Calor Solo-Ar (TCSA) aproveita essa energia utilizando ventiladores de baixa potência que circulam ar por dutos subterrâneos, contribuindo para a redução do consumo de energia de sistemas convencionais de ar-condicionado (Liu et al., 2023). A eficiência desse sistema depende de dados das temperaturas do ar e do solo, que podem ser obtidos através de modelagem matemática e técnicas de inteligência artificial.

A previsão de séries temporais é uma área importante na análise de dados, com aplicações que vão desde a economia até a meteorologia. Séries temporais consistem em sequências de dados coletados ao longo do tempo, permitindo que analistas identifiquem padrões e tendências para prever eventos futuros. Nos últimos anos, redes neurais têm

---

<sup>6</sup> Ver site: <https://portal.inmet.gov.br/noticias/2023-%C3%A9-o-mais-quente-em-174-anos-confirma-relat%C3%B3rio-da-omm> Acesso em: 03/10/2024.

emergido como uma potente ferramenta para melhorar soluções específicas nesse campo (Rizvi, 2024).

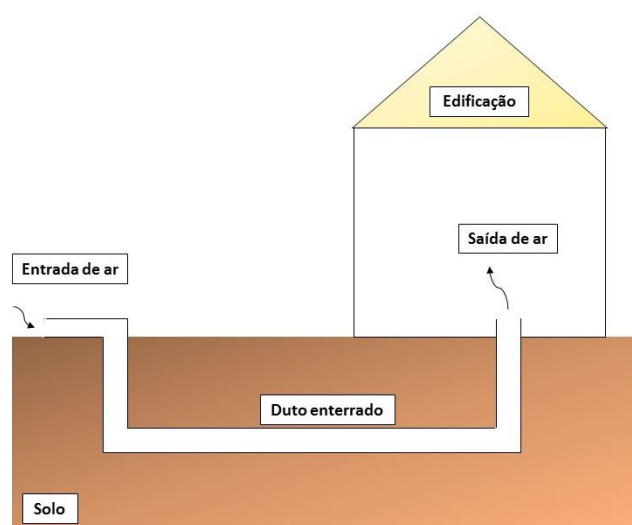
Redes neurais são modelos computacionais inspirados no funcionamento do cérebro humano, capazes de aprender padrões complexos em grandes volumes de dados. Zheng (2019) e Graves (2012) salientam que com o advento da aprendizagem profunda, modelos como Redes Neurais Recorrentes (RNN) e Long Short-Term Memory (LSTM) se destacaram na tarefa de modelar dependências temporais em dados sequenciais, oferecendo resultados superiores em comparação com métodos tradicionais.

Este trabalho tem como objetivo prever o comportamento da temperatura do ar e do solo na cidade de Rivera (Uruguai). Tais temperaturas foram coletadas através de um sistema de monitoramento remoto, no período de julho de 2023 até abril de 2024, em parceria com a Universidad Tecnológica del Uruguay (UTEC). Este trabalho avalia as técnicas de LSTM e Regressão Linear para prever e recuperar dados de temperatura ausentes devido a falhas e/ou manutenção dos sensores no período citado.

## 2 REFERENCIAL TEÓRICO

### 2.1 Sistema de Trocador de Calor Solo-Ar

Conforme Rodrigues et al. (2018), os Trocadores de Calor Solo-Ar (TCSA) consistem em um ou mais dutos enterrados, nos quais o ar é impulsionado por ventiladores. Nesse sistema, o solo atua como fonte ou sumidouro de calor para o ar que circula pelo TCSA. Um esquema dos TCSA é mostrado na Figura 1. O ar que emerge do dispositivo pode ser utilizado para melhorar as condições térmicas de edificações, especialmente durante o verão e o inverno. Além disso, essa tecnologia contribui para a eficiência energética dos edifícios, reduzindo a necessidade de sistemas de aquecimento de alto custo.



**Figura 1:** Esquema de funcionamento dos TCSA. Fonte: Domingues (2023)

A transição para energias limpas traz diversas vantagens. Em primeiro lugar, a utilização dessas fontes pode resultar na redução das importações de energia, o que fortalece a autossuficiência energética de um país. Além disso, contribui para o abatimento do esgotamento dos recursos não renováveis, promovendo uma gestão mais sustentável dos recursos naturais (Zhou, 2024).

A partir dessa questão, os trocadores de calor solo-ar (TCSA) se apresentam como uma alternativa favorável para melhorar o conforto térmico por diversas razões: seus princípios de funcionamento são relativamente simples; utilizam energia geotérmica renovável; operam com ventiladores de baixa potência, resultando em um consumo reduzido de energia elétrica; o investimento inicial na instalação pode ser recuperado por meio da economia gerada na conta de energia; podem ser integrados a outros sistemas de climatização, ajudando ainda mais a reduzir o consumo de energia ao fornecer um pré-aquecimento ou resfriamento do ar (Brum, 2016).

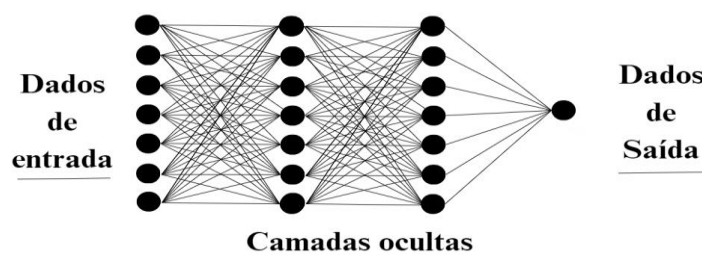
## 2.2 Séries Temporais

As séries temporais são caracterizadas por mudanças intrínsecas que modificam as propriedades do processo de geração de dados, alterando então sua distribuição de probabilidade subjacente ao longo do tempo. Uma série temporal não estacionária é definida em termos de sua média ou variância (ou ambas) variando ao longo do tempo e as mudanças podem assumir várias formas, um fenômeno conhecido como “deriva de conceito”. A “deriva de conceito” pode deteriorar a precisão da previsão do modelo ao longo do tempo, o que requer estratégias de adaptação permanentes.

Em trabalhos anteriores, Bastos (2016) empregou modelos que fazem uso de técnicas de modelagem preditiva, como a regressão linear, para a previsão de séries temporais de temperatura, apresentando resultados satisfatórios. Este presente estudo envolve a utilização de uma série temporal  $T(t_1), T(t_2), \dots, T(t_n)$ , onde as observações são realizadas em instantes de tempo específicos  $t_1, t_2, \dots, t_n$ . A série é composta por valores medidos em intervalos de uma hora, que representam a evolução das temperaturas medidas na UTEC de Rivera.

## 2.3 Redes Neurais Artificiais

Uma rede neural artificial (RNA), ou um aproximador universal para mapeamento não linear, classicamente compreende três camadas, incluindo camadas de entrada, ocultas e de saída (Géron, 2019), conforme mostrado na Figura 2. Apesar das variações nas estruturas dos modelos de Redes Neurais Artificiais (ANN), todos possuem dois componentes fundamentais: um processo de aprendizado e um de otimização, utilizados para classificar as entradas nas camadas ocultas e gerar as saídas, respectivamente.



**Figura 2:** Arquitetura de uma RNA

O primeiro componente, conhecido como camada de entrada oculta, é responsável por identificar a não linearidade nos dados de entrada e classificá-los em camadas ocultas por meio de um algoritmo de aprendizado, seja supervisionado ou não supervisionado. O segundo componente, denominado camada de saída oculta, implementa um processo de otimização, buscando a melhor correspondência entre as entradas processadas e os alvos pretendidos. Essas entradas classificadas são então mapeadas para as saídas por uma projeção linear, ajustando os parâmetros do modelo através de um algoritmo de gradiente descendente, de forma a minimizar o erro quadrático médio entre a previsão da ANN e os valores reais. As ANNs conseguem lidar com grandes quantidades de dados devido à separação dos processos de aprendizado e otimização. Além disso, pode-se utilizar um mecanismo adaptativo que atualiza recursivamente os parâmetros à medida que novos dados observacionais se tornam disponíveis (Géron, 2019).

## 3 METODOLOGIA

### 3.1 Coleta de dados

Este estudo teve como finalidade prever a temperatura com a aplicação de modelos de séries temporais com regressão linear e redes neurais recorrentes. Os dados de temperatura foram coletados por sensores instalados em diferentes profundidades: à superfície do solo e nas camadas de 0,4; 0,8; 1,2; 1,6 e 2 metros. Devido à necessidade de manutenções periódicas do sistema de monitoramento, houve alguns períodos em que os dados não puderam ser registrados. A base de dados gerada foi organizada com base na hora cumulativa, resultando em um total de 6.056 registros contínuos ao longo de dez meses de coleta.

Antes da análise, os dados passaram por um processo de pré-tratamento que incluiu a remoção de valores nulos e inconsistentes, de modo a garantir a qualidade e a integridade das informações para os modelos preditivos. Esse processo é importante para assegurar que os modelos desenvolvidos sejam capazes de fornecer previsões confiáveis e precisas, evitando interferências causadas por falhas no monitoramento. Assim, garantiu-se uma base de dados consistente e robusta para a aplicação dos métodos de previsão, aumentando a confiabilidade dos resultados alcançados ao final da análise.

## 3.2 Escolha do Modelo

Para a construção dos modelos preditivos, optou-se por empregar o método de Regressão Linear com aplicação de médias móveis, foi adotada para explorar uma abordagem mais simples e direta para a previsão, aproveitando a linearidade observada em partes do conjunto de dados, onde foram acumuladas 1816 horas.

Além desse método tradicional, utilizou-se também uma rede neural recorrente do tipo LSTM (Long Short-Term Memory), reconhecida por sua habilidade em capturar padrões complexos em séries temporais, especialmente aqueles que envolvem dependências de longo prazo. A escolha do LSTM deve-se à sua arquitetura interna, que permite o armazenamento seletivo de informações importantes ao longo de períodos de tempo, garantindo uma análise robusta de fenômenos que apresentam variabilidade temporal significativa. Combinando essas abordagens, o estudo buscou alcançar previsões mais precisas e abrangentes.

## 3.3 Programação

As implementações foram feitas na linguagem de programação *Python*. No processo de desenvolvimento, foram empregadas diversas ferramentas e bibliotecas específicas para análise de dados. Entre elas, destacam-se o *Numpy* e o *Pandas*, que proporcionam estruturas de dados para armazenamento e manipulação de grandes volumes de informações.

Além disso, a biblioteca *Keras* foi utilizada para o desenvolvimento dos modelos de redes neurais, oferecendo uma interface para a construção de redes complexas. O *Scikit-Learn*<sup>7</sup>, por sua vez, foi aplicado para a implementação de algoritmos de aprendizado de máquina, possibilitando a criação de modelos preditivos com técnicas de regressão e otimização. Essas ferramentas, em conjunto, formaram a base para o desenvolvimento e a execução dos modelos propostos, permitindo o tratamento e na análise dos dados.

## 3.4 Desenvolvimento da Rede

Para realizar as previsões de temperatura, foram utilizados dois modelos preditivos: Regressão Linear e uma Rede Neural Recorrente com técnica do LSTM. O modelo Regressão Linear, foi empregada a média móvel de 24 horas como variável preditora, o que permitiu capturar flutuações locais de curto prazo nos dados de temperatura. Para garantir suavização e preenchimento de eventuais lacunas nos dados, foi utilizada a técnica de interpolação polinomial de ordem 12.

---

<sup>7</sup> Ver, por exemplo, o site: <https://scikit-learn.org/stable> Acessado em 03/10/2024

Os dados foram divididos em dois conjuntos: um de treino e outro de teste. Utilizando as últimas 168 horas (equivalente há 7 dias) de observações para validação do desempenho dos modelos preditivos, enquanto restante foi reservado para o treinamento. O conjunto de treino continha 70% dos dados, totalizando 4239 horas. As últimas 1816 horas foram previstas. Durante o treinamento, a média móvel foi calculada ao longo de uma janela de 24 horas, proporcionando uma visão mais equilibrada das variações diárias.

No modelo de Rede Neural Recorrente LSTM, os dados de temperatura foram inicialmente normalizados, com o intuito de escalonar os valores para o intervalo entre 0 e 1, o que contribui para melhorar a performance do modelo. Foram criadas sequências de 24 horas, correspondentes a um dia, para treinar a rede neural. Essas sequências serviram como entradas ( $x$ ), enquanto o valor da temperatura da hora seguinte foi utilizado como saída ( $y$ ). A arquitetura do modelo LSTM incluiu uma camada de entrada com 100 unidades LSTM, configurada para retornar sequências, seguida de uma camada de *Dropout* (“desativação de neurônios”) de 20% para evitar o chamado *Overfitting* (“sobreajuste”). A camada final (densa) tem uma unidade e é responsável por gerar a previsão final.

O modelo LSTM foi compilado utilizando o otimizador Adam (Géron, 2019) e a função de perda Erro Médio Quadrático (MSE). A divisão dos dados para validação seguiu a proporção de 70% para treino e 30% para teste. O treinamento do modelo foi realizado em 50 épocas, com um *batch size* (“tamanho do lote”) de 32, possibilitando ajustes iterativos nos pesos da rede neural para reduzir o erro. Após o processo de treinamento, as previsões foram geradas no conjunto de teste e posteriormente desnormalizadas para retornar à escala original dos dados. O desempenho dos modelos foi avaliado utilizando métricas de erro como o Erro Médio Absoluto (MAE) e Raiz do Erro Médio Quadrático (RMSE).

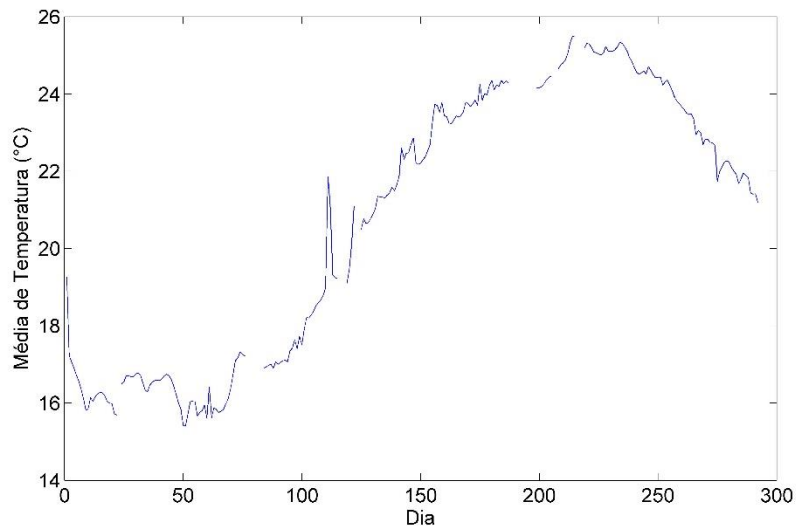
## 4 RESULTADOS

A título de ilustração, a Figura 3 mostra as médias de temperatura diárias no solo da UTEC, em uma profundidade de 1,2 m. Essa foi a profundidade tomada para análise de dados nesse artigo. Como se pode perceber, há sete períodos onde as temperaturas estão faltando. Isso ocorreu devido a falhas pontuais nos sensores, onde ocorreram, por exemplo, discrepâncias perceptíveis nos dados, além de momentos em que ocorreram desligamentos para manutenção. Os resultados ficaram dentro do esperado para o local, com a temperatura do solo variando entre 15 e 26°C.

Como observado, na análise dos dados, as temperaturas foram processadas em períodos de hora em hora, usando as técnicas de Regressão Linear e LSTM. Com base nos dados apresentados na Tabela 1, observou-se que os dois modelos obtiveram Erros Médios Absolutos inferiores a 0,2°C. Considerando que a temperatura média registrada foi de aproximadamente 20°C, esses valores indicam que os erros relativos foram inferiores a 1%.

A Figura 4 ilustra os gráficos de previsão da Regressão Linear e do LSTM para as temperaturas nas últimas 1816 horas de medição (correspondendo aos 30% de previsão).

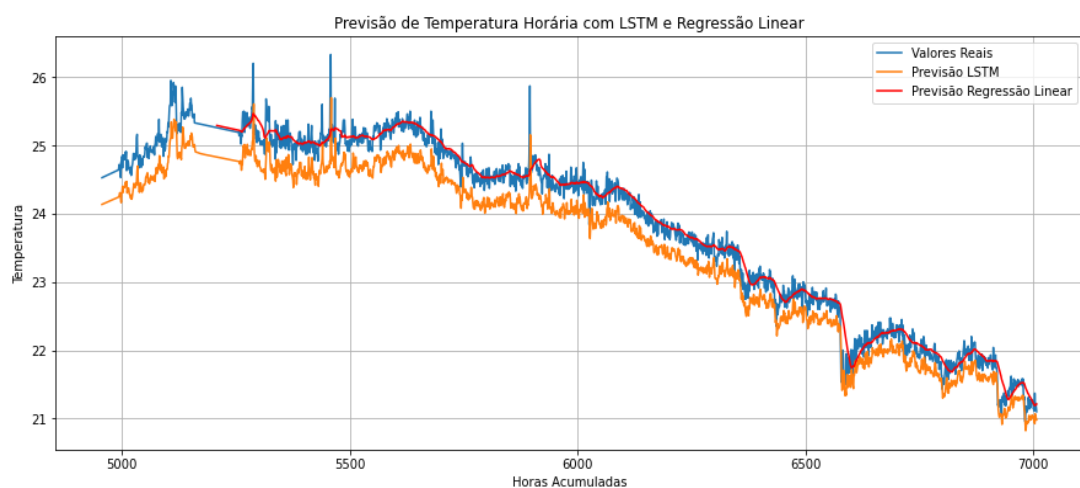
Pode-se perceber que ambas as técnicas fornecem resultados qualitativamente satisfatórios, seguindo os padrões de variação das temperaturas. Nessa análise, a Regressão Linear apresentou um desempenho ligeiramente superior, com sua curva seguindo mais próxima a dos dados reais.



**Figura 3:** Gráfico das médias diárias de temperatura do solo na UTEC a 1,2m de profundidade

**Tabela 1:** Avaliação do desempenho dos modelos

Modelo	Erro Absoluto Médio (°C)	Raiz do Erro Quadrático Médio (°C)
Regressão Linear	0,108	0,157
RNA(LSTM)	0,352	0,377



**Figura 4:** Gráfico previsão LSTM



## 5 CONCLUSÕES

A análise das temperaturas do ar e do solo na cidade de Rivera, utilizando as técnicas de LSTM e Regressão Linear, demonstrou-se efetiva para a previsão e recuperação de dados faltantes. A eficiência de ambos os modelos, evidenciada pelos Erros Médios Absolutos inferiores a  $0,2^{\circ}\text{C}$ , reflete a capacidade dessas abordagens em lidar com a variabilidade dos dados em um contexto real, onde falhas e manutenção de sensores são ocorrências eventuais.

Os dados coletados durante o período de estudo indicaram que a temperatura média diária do solo variou entre  $15$  e  $26^{\circ}\text{C}$ , alinhando-se com as expectativas para a região. As representações gráficas das previsões mostraram que, embora ambas as técnicas tenham fornecido resultados qualitativos satisfatórios, a Regressão Linear apresentou um desempenho ligeiramente superior em relação ao LSTM.

Além disso, a utilização de técnicas de inteligência artificial, como a modelagem de séries temporais, se revela um recurso valioso para a otimização de sistemas de monitoramento e para a implementação de soluções sustentáveis em edificações, como os Trocadores de Calor Solo-Ar. Os resultados obtidos reforçam a importância de investir em tecnologias que integrem dados meteorológicos e técnicas de previsão, promovendo uma gestão mais eficiente dos recursos energéticos e contribuindo para a mitigação dos efeitos das mudanças climáticas.

## AGRADECIMENTOS

Letícia B. D. Soares agradece a UFPel por sua bolsa de iniciação científica, Geilson A. Soares agradece ao CNPq por sua bolsa de iniciação científica. João I. M. Bezerra e Marlon M. H. Cely agradecem a Petrobrás (Projeto: Análise Hidrodinâmico de Escoamentos Multifásicos Usando Técnicas de Inteligência Artificial). Todos os autores agradecem o apoio da UTEC de Rivera pelo apoio a construção do experimento.

## REFERÊNCIAS

- Brum, R. S. (2016). Teoria construtal e desempenho térmico de trocadores de calor solo-ar (Tese de doutorado). Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.
- Caineng, Z., Xiong, B., Xue, H., Zheng, D., Ge, Z., Wang, Y., Jiang, L., Pan, S., & Wu, S. (2021). The role of new energy in carbon neutral. *Petroleum Exploration and Development*, 48(2), 480–491. [https://doi.org/10.1016/S1876-3804\(21\)60039-3](https://doi.org/10.1016/S1876-3804(21)60039-3).
- Géron, A. (2019). *Mãos à obra: Aprendizado de máquina com Scikit-Learn & TensorFlow: Conceitos, ferramentas e técnicas para a construção de sistemas inteligentes* (R. Contatori, Trad.). Alta Books. (Trabalho original publicado em 2017).

- Graves, A. (2012). Long short-term memory. Em F. Seibt (Ed.), Rotulagem de sequência supervisionada com redes neurais recorrentes (Estudos em Inteligência Computacional, v. 385, pp. 37–57). Springer. [https://doi.org/10.1007/978-3-642-24797-2\\_4](https://doi.org/10.1007/978-3-642-24797-2_4).
- Instituto Nacional de Meteorologia - Organização Meteorológica Mundial. (2024). 2023 é o mais quente em 174 anos, confirma relatório da OMM. Disponível em <https://portal.in-met.gov.br/noticias/2023-%C3%A9-o-mais-quente-em-174-anos-confirma-relat%C3%B3rio-da-omm>. Acesso em 3 de outubro de 2024.
- Liu, Z., Xie, M., Zhou, Y., He, Y., Zhang, L., Zhang, G., & Chen, D. (2023). 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*, 4(2), 148–168.
- Rabino, M. F., Popp, J., Máté, D., & Kovács, S. (2022). Segurança energética e transição energética para atingir a neutralidade de carbono. *Energies*, 15(8126), 1–16. <https://doi.org/10.3390/en15218126>.
- Rizvi, M. F., Sahu, S., & Rana, S. (2024). ARIMA model time series forecasting. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 12(V), 3782–3791. Disponível em <https://www.ijraset.com>. Acesso em 23 de setembro de 2024.
- Rodrigues, M. K., Coswig, F. S., Camargo, K. R., Isoldi, L. A., Brum, R. S., Ramalho, J. V. A., Vaz, J., Rocha, L. A. O., & dos Santos, E. D. (2018). Thermal performance simulations of Earth-air heat exchangers for different soils of a coastal city using in-situ data. *Sustainable Energy Technologies and Assessments*, 30, 224–229. <https://doi.org/10.1016/j.seta.2018.10.003>.
- Zheng, S., & Zhou, Y. (2019). Numerical study on the thermal and optical performances of an aerogel glazing system with the multivariable optimization using an advanced machine learning algorithm. *Advanced Theory and Simulations*, 2(9), Article 1900092. <https://doi.org/10.1002/adts.201900092>
- Zhou, Y., Zheng, S., & Hensen, J. L. M. (2024). Machine learning-based digital district heating/cooling with renewable integrations and advanced low-carbon transition. *Renewable and Sustainable Energy Reviews*, 199, Article 114466. <https://doi.org/10.1016/j.rser.2024.114466>.