*XXII Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica*
*XI Simpósio Brasileiro de Engenheiros Geotécnicos Jovens*
*24 a 28 de agosto de 2026 – Brasília/DF*

**DOI: 10.47094/COBRAMSEG2026.2/25**

# Previsão da Variação de Temperatura do Solo em Sistema GSHP Vertical usando Redes Neurais Artificiais (RNA)

Letícia Menezes Santos Sá
Professora Substituta, Universidade Federal de Sergipe, Aracaju, Brasil, leticiamenezes0512@gmail.com

Gabrielle Maria Romeiro Lombardi
Professora, PECEGE – ESALQ/USP, Piracicaba, Brasil, glombardi.pecege@gmail.com

Alberto Hernandez Neto
Professor, Universidade de São Paulo, São Paulo, Brasil, ahneto@usp.br

Cristina de Hollanda Cavalcanti Tsuha
Professora, Escola de Engenharia de São Carlos da Universidade de São Paulo, São Carlos, Brasil, chctsuha@sc.usp.br

RESUMO: Diante da crescente demanda por fontes de energia renováveis e do aumento do consumo energético para climatização de edificações no Brasil, esta pesquisa investiga o uso da geotermia superficial por meio de estacas trocadoras de calor, com o objetivo de prever o desempenho térmico de sistemas "Ground Source Heat Pump (GSHP)". Para isso, este estudo propôs e avaliou um modelo de rede neural artificial capaz de estimar parâmetros térmicos desses sistemas acoplados a fundações com estacas trocadoras de calor. Utilizando resultados obtidos a partir de simulações realizadas com o software PILESIM 2.1, o modelo previu três variáveis principais: coeficiente de performance (COP), temperatura da superfície do solo e temperatura média ao longo do perfil de solo. Os resultados evidenciaram a eficácia do modelo, com redução consistente do erro absoluto médio e da função de perda ao longo do treinamento. A análise dos coeficientes de determinação ($R^2$) e do erro relativo médio confirmou a capacidade preditiva da rede, com maior precisão na estimativa do coeficiente de desempenho e maior variabilidade nas previsões de temperaturas do solo. Com base nestes resultados, a aplicação de redes neurais artificiais apresentou-se como uma alternativa promissora para complementar a análise e a modelagem térmica de longo prazo de sistemas GSHP acoplados a fundações profundas no Brasil. Pesquisas futuras poderão incluir variáveis adicionais, otimizar a arquitetura da rede neural e explorar técnicas avançadas de pré-processamento de dados.

PALAVRAS-CHAVE: Estacas Trocadoras de Calor, Modelagem Preditiva, Geotermia Superficial, Simulação Computacional.

ABSTRACT: Considering the growing demand for renewable energy sources and the increasing energy consumption for climatization in Brazil, this research investigates the use of shallow geothermal energy through energy piles, aiming to predict the thermal performance of Ground Source Heat Pump (GSHP) systems. To this end, the study proposed and evaluated an artificial neural network model capable of estimating thermal parameters of such systems coupled with energy pile foundations. Using results obtained from simulations performed with the PILESIM 2.1 software, the model predicted three main variables: coefficient of performance (COP), ground surface temperature, and average ground temperature along the soil profile. The results demonstrated the effectiveness of the model, showing a consistent reduction in mean absolute error and loss function throughout the training process. The analysis of the coefficients of determination ($R^2$) and mean relative error confirmed the predictive capability of the network, with higher accuracy in estimating the coefficient of performance and greater variability in ground temperature predictions. Based on these findings, the application of artificial neural networks was shown to be a promising alternative to complement the long-term thermal analysis and modeling of GSHP systems coupled with deep foundations in Brazil. Future studies may incorporate additional variables, optimize the neural network architecture, and explore advanced data preprocessing techniques.

KEYWORDS: Energy Piles, Predictive Modeling, Shallow Geothermal Energy, Computational Simulation.

## 1 INTRODUÇÃO

A busca por fontes de energia limpas e renováveis tem crescido globalmente, impulsionada pelos problemas ambientais enfrentados nos últimos anos. Nesse contexto, a energia geotérmica superficial emerge como uma alternativa sustentável para reduzir o consumo proveniente de outras fontes poluidoras.

O sistema Ground Source Heat Pump (GSHP) realiza trocas térmicas entre o solo e a superfície por meio da circulação de um fluido em tubos instalados no subsolo e conectados a uma bomba de calor (Amatya et al., 2012). A integração desses tubos com fundações profundas torna o sistema mais economicamente atrativo, originando as estacas trocadoras de calor (Brandl, 2006). Durante a operação do sistema, as trocas térmicas podem alterar a temperatura do solo ao redor das estacas ao longo do tempo, especialmente quando há predominância de resfriamento, como ocorre no Brasil (Olgun et al., 2015). Nessas condições, pode ocorrer acúmulo de calor no subsolo e redução da eficiência do sistema.

No Brasil, onde a demanda por resfriamento dos ambientes (rejeição de calor no solo) é predominante ao longo do ano, pode haver uma perda de eficiência do sistema GSHP após alguns anos de operação, devido ao acúmulo de calor no solo ao redor das estacas geotérmicas. Porém, obter medições precisas e detalhadas das variações de temperatura do solo a longo prazo pode ser desafiador, devido a falhas que vão desde problemas nos sensores instalados em campo até interrupções na aquisição de dados. Assim, o uso de Redes Neurais Artificiais (RNA) para prever o comportamento da temperatura do solo durante a operação de sistemas GSHP, bem como o desempenho desses sistemas ao longo do tempo, surge como uma alternativa promissora para apoiar sua aplicação eficaz. Nesse sentido, alguns estudos têm utilizado RNAs para prever a variação da temperatura do solo a longo prazo associada à operação de sistemas GSHP (Xu et al., 2020; Zhou et al., 2021; Liu et al., 2023).

Diante desse contexto, o principal objetivo deste trabalho foi desenvolver e avaliar um modelo de RNA para prever parâmetros térmicos relacionados a um sistema GSHP acoplado a fundações por estacas trocadoras de calor, utilizando dados provenientes do banco de dados gerados através de simulação com o "software" PILESIM 2.1. As variáveis previstas foram o coeficiente de desempenho [COP], a temperatura da superfície do solo [TmGrndTop] e a temperatura do média do perfil de solo [TmGround]. Para isso, utilizou-se uma rede neural "feedforward" e foi considerada uma edificação hipotética na cidade de São Paulo.

## 2 MATERIAIS E MÉTODOS

A área de estudo localiza-se na cidade de São Paulo (23°33'15.8"S; 46°43'51.2"W), com temperatura média anual de aproximadamente 19 °C. O estudo baseou-se em simulações numéricas realizadas com o software PILESIM 2.1 por Sá et al. (2022), sendo comparados os resultados dessas simulações com aqueles obtidos por um modelo de Rede Neural Artificial.

O estudo original avaliou a variação da temperatura do solo ao longo de 20 anos de operação de estacas trocadoras de calor utilizadas como fundação de um hotel de negócios hipotético. O solo local é predominantemente arenoso, com lentes de argila orgânica silto-arenosa e nível do lençol freático entre 2 e 4 m de profundidade. Nesta pesquisa foram consideradas apenas estacas do tipo hélice contínua. A fundação analisada é composta por 68 estacas com diâmetro de 0,70 m e comprimento de 20 m.

A base de dados foi obtida a partir das saídas das simulações no PILESIM 2.1, considerando características do solo, parâmetros das estacas, dados da bomba de calor e demandas térmicas da edificação (Sá et al., 2022). As simulações consideraram ciclos de seis meses de resfriamento alternados com seis meses de aquecimento ao longo de 20 anos, resultando em um banco de dados com 12 variáveis e 240 observações.

O estudo foi dividido em duas etapas: análise exploratória dos dados e modelagem preditiva com RNA, ambas realizadas em Python (Rossum, 1995). Na análise exploratória foram realizadas limpeza do banco de dados, estatísticas descritivas e análise de correlação entre variáveis.

Para a modelagem preditiva foram selecionadas como variáveis de entrada tempo_acum_h/mes, QHeat, QCold, QElecTot, num_mes e num_ano. As variáveis numéricas foram normalizadas com MinMaxScaler e as categóricas transformadas por one-hot encoding. As variáveis de saída foram COP, TmGrndTop e TmGround.

Os dados foram divididos em treinamento (70%), validação (15%) e teste (15%) utilizando train_test_split (Brownlee, 2020). Foi implementada uma rede neural "feedforward" com três camadas ocultas

(16, 12 e 8 neurônios) e funções de ativação tanh, tanh e sigmoide, respectivamente, e camada de saída com três neurônios e ativação linear (Cybenko, 1989).

O treinamento foi realizado utilizando função de perda MSE, otimizador Adam e taxa de aprendizado de 0,001. O desempenho do modelo foi avaliado por meio das métricas MAE, RMSE, erro relativo médio (MRE) e coeficiente de determinação ($R^2$).

## 3 RESULTADOS E DISCUSSÃO

### 3.1 Análise Exploratória

Alguns gráficos foram elaborados a fim de permitir uma visualização inicial do comportamento das variáveis ao longo do tempo. A Figura 1 apresenta a variação da demanda de energia para diferentes finalidades — aquecimento (QHeat), resfriamento (QCold) e energia elétrica total (QElectTot) - ao longo do tempo. Observa-se que a demanda total de energia para aquecimento (QHeat) apresenta oscilações menores durante a operação do sistema GSHP em comparação com a demanda total de energia para resfriamento (QCold), que apresenta variações mais expressivas. Essas variações aumentam consideravelmente nos seis meses em que há resfriamento dos ambientes da edificação e diminuem abruptamente nos períodos de utilização do sistema para aquecimento da água. Além disso, a energia elétrica consumida pelo sistema também apresenta um padrão cíclico, com picos de demanda nos períodos de resfriamento.

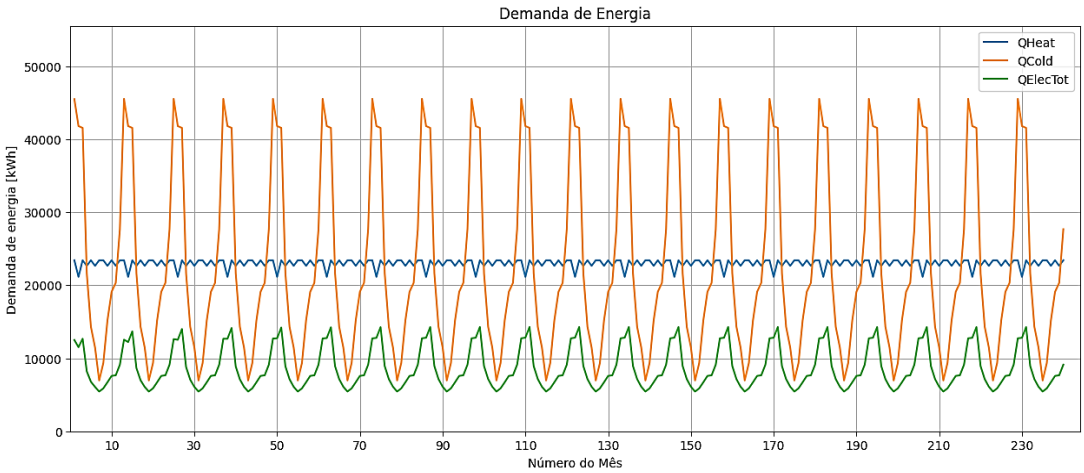



Figura 1. Demanda de energia do sistema ao longo do tempo: QHeat, QCold e QElectTot.

A Figura 2 mostra a variação do COP ao longo do tempo. O gráfico revela que o COP acompanha um padrão semelhante ao da demanda de energia, atingindo valores mais elevados durante os períodos de resfriamento. Ainda assim, o intervalo de variação do COP é pequeno, com valores entre 5,38 (mínimo) e 5,50 (máximo). Esses resultados indicam que, ao longo de todo o período analisado, o COP manteve-se acima da recomendação de Brandl (2006) para sistemas GSHP, que sugere valores superiores a 4. Destaca-se também que, nos períodos de aquecimento, os valores de COP apresentaram uma redução gradual até atingir certa estabilidade.

Por fim, a Figura 3 exibe a variação da temperatura média do volume de solo que é atribuída às estacas trocadoras de calor (TmGround) bem como do solo nos 30 cm iniciais da superfície, onde estão os tubos de conexão horizontal (TmGrndTop). Percebe-se que a temperatura varia de acordo com os ciclos de aquecimento e resfriamento, conforme esperado. No caso do TmGrndTop, provavelmente devido à interferência do ambiente, os valores de temperatura oscilaram em intervalos menores. Verifica-se que há uma tendência de crescimento da temperatura do solo ao longo do tempo, possivelmente devido ao acúmulo de calor ocasionado pela operação do sistema, até atingir uma certa estabilidade, aproximadamente após o mês 90 (~ 7,5 anos de operação).

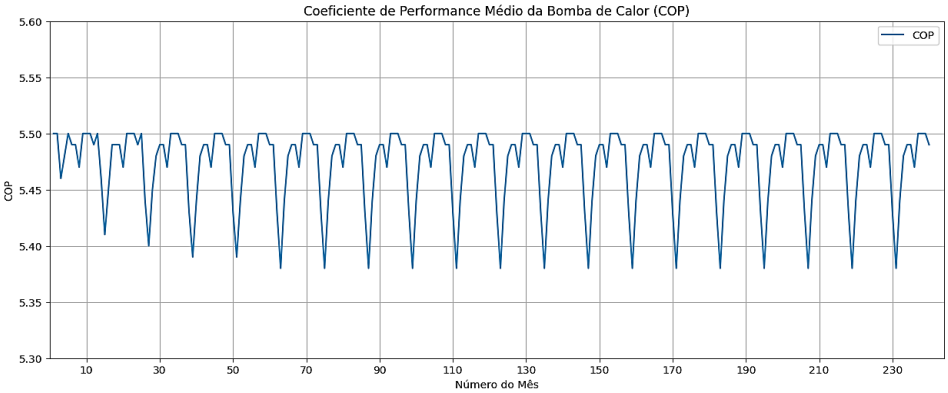



Figura 2. Variação do COP ao longo do tempo.

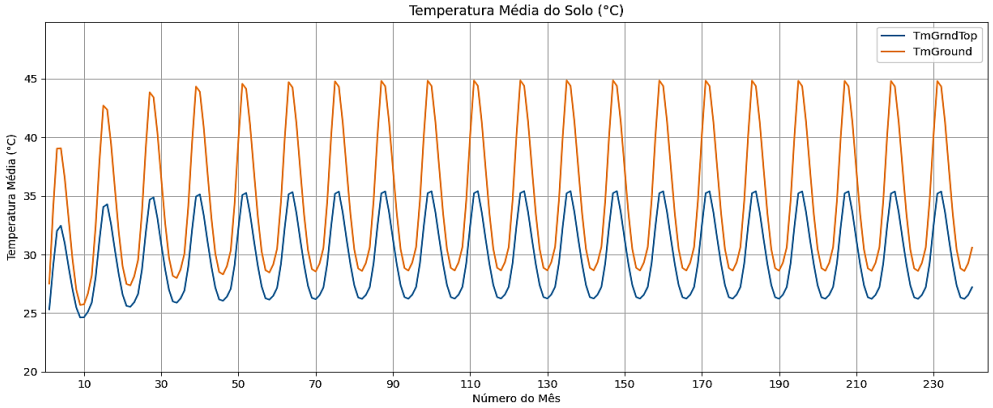



Figura 3. Temperatura média do solo ao longo do tempo de operação do sistema GSHP.

Após a visualização inicial dos dados, foi realizada uma análise descritiva das variáveis quantitativas de interesse, cujos resultados estão apresentados na Tabela 1. Foram excluídas dessa análise a variável "mês", por se tratar de uma categoria nominal, bem como as variáveis "número do mês" e "ano", que não foram consideradas relevantes para a obtenção de medidas descritivas.

Tabela 1. Estatística descritiva das variáveis estudadas

| | QHeat | QCold | QElecTot | COP | TmGrndTop | TmGround |
|---|---|---|---|---|---|---|
| Contagem | 240 | 240 | 240 | 240 | 240 | 240 |
| Média | 22977.50 | 22906.91 | 8665.13 | 5.47 | 29.80 | 34.94 |
| Desvio Padrão | 654.15 | 12821.31 | 2836.81 | 0.03 | 3.33 | 5.85 |
| Mínimo | 21150 | 6960 | 5440 | 5.38 | 24.62 | 25.69 |
| 1º Quartil | 22660 | 13600 | 6557 | 5.47 | 26.57 | 29.31 |
| 2º Quartil | 23420 | 19715 | 7622 | 5.49 | 29.17 | 33.50 |
| 3º Quartil | 23420 | 31140 | 9716.25 | 5.50 | 32.43 | 40.14 |
| Máximo | 23420 | 45530 | 14280 | 5.50 | 35.40 | 44.85 |

Em seguida, foi analisada a correlação entre as variáveis estudadas, com exceção do mês (por se tratar de uma variável categórica) e do número do ano, considerados não relevantes para essa etapa (Figura 4). A correlação foi avaliada por meio do coeficiente de correlação de Pearson. Observa-se que as variáveis TmGround e TmGrndTop apresentam uma correlação positiva forte, indicando que a temperatura do solo em profundidade acompanha a variação da temperatura na superfície ao longo do tempo.

O COP apresenta uma correlação negativa forte com TmGround e TmGrndTop, o que é esperado, considerando que a temperatura do solo influencia a eficiência e o desempenho da bomba de calor. Isto é, temperaturas mais altas do solo tendem a reduzir a eficiência do sistema, refletida na diminuição do COP. Além disso, o COP também exibe uma correlação negativa moderada com QElectTOt,o que é esperado, visto

que maior consumo de energia elétrica está associado a menor eficiência do sistema. Há também correlação negativa com QCold, embora de menor magnitude.

Por fim, as variáveis QCold e QElectTot apresentaram correlações fracas com a temperatura do solo, tanto na superfície quanto no volume de solo das estacas, sugerindo influência limitada desses fatores sobre o consumo de energia para resfriamento. As demais variáveis analisadas apresentaram correlações desprezíveis ou nulas entre si. Vale destacar que a análise se baseia exclusivamente em correlações lineares, o que limita a identificação de possíveis relações não lineares entre as variáveis.

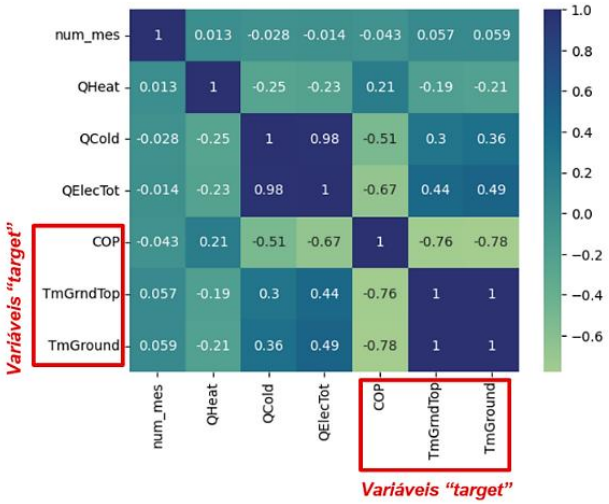



Figura 4. Cálculo das correlações entre as variáveis estudadas.

### 3.2 Arquitetura do Modelo

A Tabela 2 apresenta um resumo da arquitetura da RNA implementada neste estudo. O modelo é composto por uma camada de entrada, três camadas ocultas densas e uma camada de saída. Cada camada foi configurada com uma quantidade específica de neurônios e funções de ativação distintas, conforme descrito a seguir. A camada de entrada possui um número de neurônios correspondente ao total de variáveis preditoras utilizadas no modelo, a saber, 06 (seis). A primeira e a segunda camadas ocultas empregam a função de ativação tanh, com 16 e 12 neurônios, respectivamente. Já a terceira camada oculta utiliza a função de ativação sigmoid, com oito neurônios. Por fim, a camada de saída é composta por três neurônios com função de ativação linear, adequada para a predição de variáveis contínuas.

Tabela 2. Tabela resumo da arquitetura do modelo.

| Camada | Tipo | Número de Neurônios | Função de Ativação |
|---|---|---|---|
| Entrada | Densa | 06 | - |
| Oculta 1 | Densa | 16 | tanh |
| Oculta 2 | Densa | 12 | tanh |
| Oculta 3 | Densa | 8 | sigmoid |
| Saída | Densa | 3 | linear |

### 3.3 Evolução da Função de Perda

Analisando a evolução da função de perda (Figura 5a) - expressa pelo erro quadrático médio (“mean squared error” – MSE) - ao longo das épocas de treinamento e validação da Rede Neural Artificial, observa-

se uma redução progressiva da perda durante o processo de treinamento, refletindo a capacidade do modelo em ajustar seus parâmetros e minimizar o erro preditivo. As curvas correspondentes aos conjuntos de treinamento e validação apresentam trajetórias semelhantes de decaimento, o que é um indicativo de que o modelo está aprendendo os padrões dos dados de forma eficiente, sem sinais evidentes de "overfitting". Essa coerência entre as perdas nos conjuntos de treinamento e validação reforça a estabilidade e a generalização do modelo, aspectos fundamentais para garantir sua aplicabilidade a novos conjuntos de dados.

### 3.4 Evolução do Erro Absoluto Médio (MAE)

A Figura 5b apresenta a evolução do erro absoluto médio ("mean absolute error" – MAE) ao longo das épocas de treinamento e validação. Essa métrica corresponde à média das diferenças absolutas entre os valores previstos pelo modelo e os valores observados, sendo uma medida direta da precisão das previsões.

Observa-se que os valores de MAE diminuem progressivamente com o avanço das épocas, tanto para o conjunto de treinamento quanto para o de validação. Esse comportamento indica que a rede neural está aprendendo de forma consistente, reduzindo os erros ao longo do tempo e aprimorando sua capacidade de generalização. A trajetória semelhante das curvas também indica possivelmente a ausência de sobreajuste ("overfitting") e sugere um bom equilíbrio entre aprendizado e desempenho preditivo.

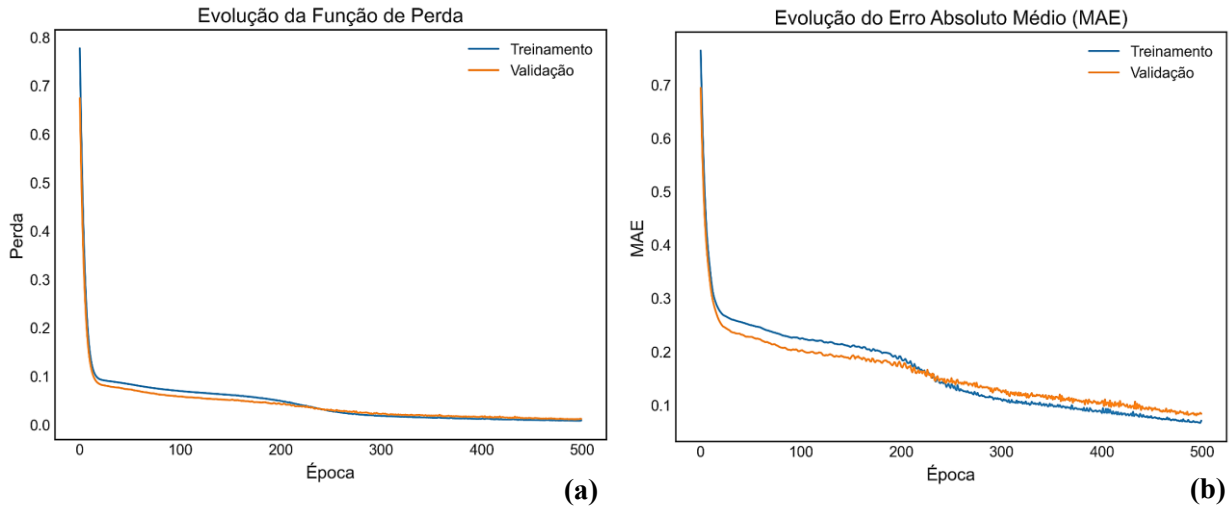



Figura 5. (a) Evolução da Função de Perda e (b) Evolução do erro absoluto médio (MAE).

### 3.5 Coeficiente de Determinação ($R^2$)

A Figura 6a apresenta os valores do coeficiente de determinação ($R^2$) obtidos para cada uma das variáveis dependentes do modelo: COP, TmGrndTop e TmGround. O $R^2$ representa a proporção da variância total das variáveis-alvo que é explicada pelas predições do modelo, sendo uma métrica amplamente utilizada para avaliar a qualidade do ajuste em modelos de regressão.

Os resultados obtidos indicam que os valores de $R^2$ se aproximam de um para todas as variáveis preditas, o que sugere que o modelo apresenta elevado poder explicativo. Em outras palavras, a maior parte da variabilidade observada nos dados é capturada pelo modelo, evidenciando sua eficácia na previsão das variáveis de interesse. Esses resultados reforçam o bom desempenho do modelo de rede neural desenvolvido.

### 3.6 Erro Relativo Médio (MRE)

O erro relativo médio em termos percentuais para cada uma das variáveis "target" pode ser visualizado na Figura 6b. Como o MRE quantifica o erro relativo entre as previsões e os valores reais, deseja-se obter valores menores de MRE, que indicam um erro proporcional menor. Neste caso, o MRE ficou entre 5% e 10%, indicativo de uma boa precisão. Além disso, pode-se notar que o modelo está se comportando de forma semelhante para as variáveis TmGrndTop e TmGround, que tratam da temperatura do solo, e que a previsão é mais simples para a variável COP, com um menor valor de MRE.

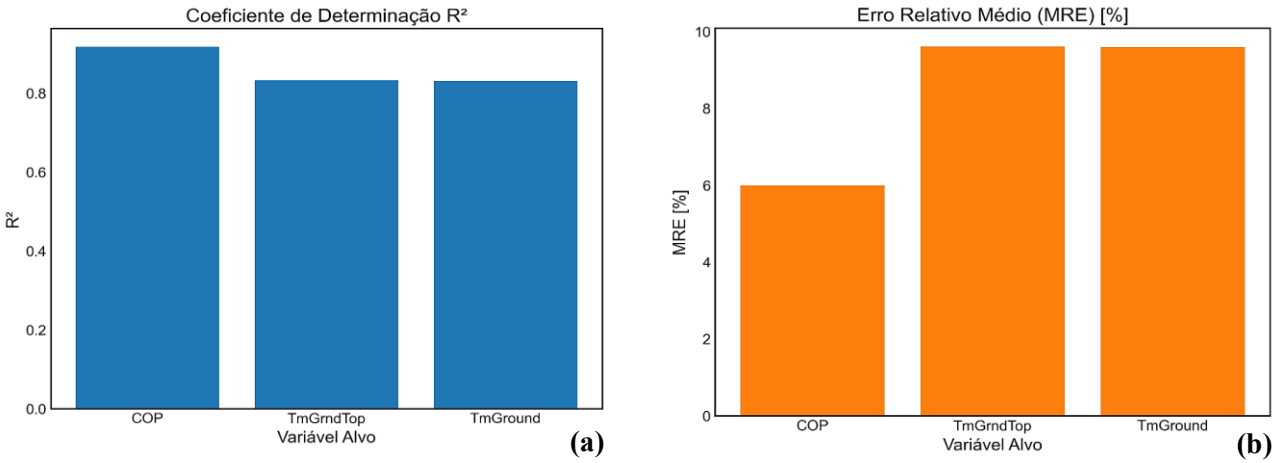



Figura 6. Métricas de avaliação: (a) Coeficiente de determinação ($R^2$); (b) Erro Relativo Médio (MRE)

### 3.7 Gráficos de Comparativos entre Valores Reais vs. Previstos

Finalmente, foram elaborados gráficos de dispersão mostrando a relação entre os valores reais e os valores previstos para cada variável "target" (Figura 7). A linha diagonal indica uma previsão ideal (perfeita) em que os valores reais sejam iguais aos previstos. Idealmente, todos os pontos deveriam coincidir com a linha diagonal. Numa situação real, como é o caso, quanto mais próximos os pontos estiverem dessa linha, mais preciso será o modelo. Analisando o gráfico referente ao COP, observa-se uma correlação linear positiva forte, com os pontos distribuídos próximos à linha de tendência. Já no caso do TmGrndTop e do TmGround, há maior dispersão dos pontos. Nesse caso, o modelo de previsão para COP apresentou melhor desempenho do que os modelos para as temperaturas do solo, que mostram maior variabilidade. Com base nas métricas analisadas, de modo geral, é possível concluir que o modelo está performando bem.

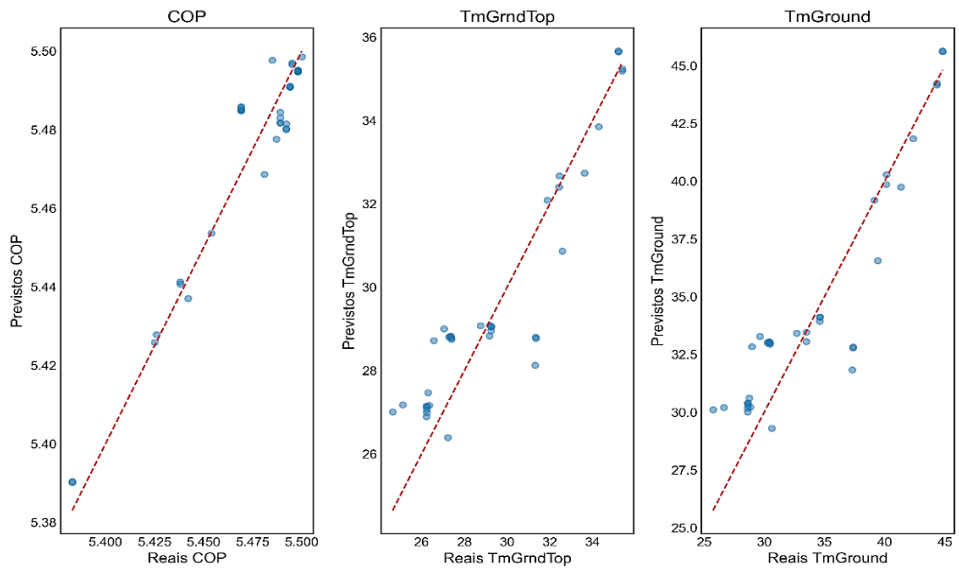



Figura 7. Comparação entre os valores reais e previstos para as variáveis 'target' (COP, TmGrndTop e TmGround).

## 4 CONSIDERAÇÕES FINAIS

Este trabalho desenvolveu e avaliou um modelo de Rede Neural Artificial para prever parâmetros térmicos em sistemas GSHP acoplados a fundações por estacas trocadoras de calor, utilizando um banco de dados gerado por simulações no "software" PILESIM 2.1.

A análise dos resultados demonstra o bom desempenho geral do modelo. Observou-se uma redução consistente da perda e do MAE ao longo das épocas de treinamento evidenciando o refinamento gradual dos parâmetros do modelo. Já os elevados valores de $R^2$ e os baixos índices de MRE confirmam a eficácia preditiva para cada variável "target". Os gráficos de dispersão revelam que o modelo apresentou maior precisão na previsão do COP em comparação com as temperaturas do solo, onde se observou maior variabilidade entre valores previstos e reais.

Para trabalhos futuros, recomenda-se ampliar o conjunto de variáveis preditoras, aprimorar a arquitetura da rede neural e adotar técnicas mais avançadas de pré-processamento e seleção de atributos, com o objetivo de aumentar a acurácia do modelo e sua aplicabilidade em contextos reais de engenharia e eficiência energética. Além disso, algumas limitações do estudo atual podem ser abordadas em pesquisas subsequentes, tais como: realizar a divisão do conjunto de dados em treinamento, validação e teste, respeitando a dependência temporal das séries ao invés de utilizar a divisão aleatória dos dados; aplicar "*early stopping*" para determinar automaticamente o número ideal de épocas; e validar o modelo no conjunto de teste.

## REFERÊNCIAS BIBLIOGRÁFICAS

Amatya, B. L., Soga, K., Bourne-Webb, P. J., Amis, T., Laloui, L. (2012) Thermo-mechanical behaviour of energy piles. *Géotechnique*, 62 (6), p. 503–519.

Brandl, H. (2006) Energy foundations and other thermo-active ground structures. *Géotechnique*, 56 (2), p. 81–122.

Brownlee, J. (2018) *Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python. Machine learning mystery*. Disponível em: https://www.inf.u-szeged.hu/~korosig/teach/books/Jason%20Brownlee%20-%20Deep%20Learning%20for%20Time%20Series%20Forecasting%20-%20Predict%20the%20Future%20with%20MLPs,%20CNNs%20and%20LSTMs%20in%20Python%20(2018).pdf. Acesso em: 21 de agosto de 2025.

Cybenko, G. (1989) Approximations by superpositions of sigmoidal functions. *Mathematics of Control, Signals, and Systems*, 2 (4), p. 303-314.

Liu, Y.; Mei, X.; Zhang, G.; Cao, Z. 2023. Long-term performance prediction of ground source heat pump system based on co-simulation and artificial neural network. *Journal of Building Engineering*, 79, p. 107949.

Olgun, C. G., Ozudogru, T. Y., Abdelaziz, S. L., Senol, A. (2015) Long-term performance of heat exchanger piles. *Acta Geotechnica*¸10, p.553–569.

Rossum, G. V.; Jr, F. L. D. (1995) Python tutorial. [S.l.: s.n.]: Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.

Sá, L. M. S., Hernandez Neto, A., Tsuha, C.H.C., Pessin, J., Freitas, M. C., Morais, T.S.O. (2022) Thermal design of energy piles for a hotel building in subtropical climate: a case study in São Paulo, Brazil. *Soils and Rocks*, 45(1), p. 1–14.

Xu, X.; Liu, J., Wang, Y.; Xu, J.; Bao, J. (2020) Performance evaluation of ground source heat pump using linear and nonlinear regressions and artifical neural networks. *Applied Thermal Engineering* 180, p. 115914.

Zhou, S., Li, J., Zhang, Y., Liu, X., Zhang, W. (2021) Prediction of the ground temperature variations caused by the operation of GSHP system with ANN. *Geothermics* 95, p. 102140.