PROCEEDINGS OF THE
INTERNATIONAL CONFERENCE ON ENERGY GEOTECHNICS
17 - 20 JUNE 2025 PARIS, FRANCE

A data-driven approach to predicting the long-term thermal performance of thermo-active piles

Javier Sanchez Fernandez[1,2,*], Alison Provost[1], Agustin Ruiz Lopez[3,1] and David M G Taborda[1]

[1]*Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK*
[2]*Science and Solutions for a Changing Planet DTP, Imperial College London, SW7 2AZ, UK*
[3]*Seequent – The Bentley Subsurface Company, Delft, The Netherlands*
[*]*Corresponding author: js8918@ic.ac.uk*

Summary

The analysis and optimisation of large ground source energy systems involving the use of thermo-active piles requires the ability to predict the thermal performance of these geothermal structures using limited computational resources, as many different configurations need to be tested. Current design methods are either based on empirical expressions, the accuracy of which is necessarily limited, or on thermo-hydraulic modelling which is computationally expensive and hence of difficult integration with optimisation procedures. In this paper, a surrogate model of a single thermo-active pile is established by running multiple thermo-hydraulic finite element analyses using different combinations of thermal ground properties, pipe arrangement, fluid temperature, pile length and pile diameter determined using a Latin hypercube sampling approach. The database of results is then used to train an artificial neural network (ANN), which is shown to produce accurate predictions of the thermal performance of a thermo-active pile given its characteristics and those of the surrounding ground. Given the low computational cost of surrogate models, this approach enables the design optimisation of large systems with greater confidence than previously possible using empirical relationships and a fraction of the resources required by thermo-hydraulic finite element models.

1 Introduction

Thermo-active structures provide low-carbon heating and cooling to buildings, offering significant potential to advance the transition towards net-zero carbon emissions [1]. By embedding heat exchange pipes within structural elements such as piles, these become capable of exchanging heat with the surrounding ground without incurring significant additional costs [2]. To facilitate wider adoption, fast and accurate methods for estimating thermal performance can help engineers improve their designs and assessments.

This study presents a surrogate modelling approach using an ANN to predict the power output per unit length of a single thermo-active pile. The model is trained on a dataset generated through parametric finite element (FE) simulations, incorporating varying pile geometries, soil stratigraphies, and material properties. By replacing costly numerical simulations, the surrogate model enables rapid assessment and design of pile output performance under a wide range of conditions.

2 Methodology

2.1 Finite element model

A parametrised 3D model of a thermo-active pile equipped with heat exchanger pipes was developed using COMSOL Multiphysics. This model was used to generate the dataset required for training the ANN. The geometry consisted of a single pile with varying radius and length, embedded in a soil domain measuring $80 \times 80 \times (L + 20)$ m (i.e. the depth of the model always exceeded the length of the pile L by 20 m to limit boundary effects). To expand the scope of the dataset, a layered stratigraphy with three materials is adopted, with thickness values L_1, L_2 and L_3, each with its own thermal conductivity, k_1, k_2 and k_3. The absorber pipes, arranged in either one, two, three or four loops, were modelled with linear elements capable of simulating heat through diffusion and advection. The flow inside the pipe of area A_{pipe} was simulated with a constant injection temperature (T_{inlet}) and velocity (v_{fluid}), both defined parametrically for each of the samples. The meshing process was carried out using COMSOL built-in mesher, producing approximately 100,000 elements per sample. The outlet temperature was computed from the initial conditions, continuing up to 360 days. A final, steady state solution was also calculated. The power output per unit length, P (W/m), was calculated using Eq. (1).

$$P = \frac{\rho_w c_w v_{fluid} A_{pipe}(T_{initial} - T_{inlet})}{L} \tag{1}$$

2.2 Numerical data

A total of 800 parametric combinations were generated for the FE model, with 200 samples assigned to each pipe loop configuration. The combinations were generated using the Latin Hypercube Sampling approach [3], which improves sampling efficiency by allowing a more uniform coverage of the input space compared to conventional random sampling techniques. The range of geometric and material parameters is listed in Tab. 1. The parametric range was chosen to be deliberately wide to capture a broad set of scenarios, thereby improving model generalisation. The outputs from the FE model were then used as database for the ANN.

Table 1: Numerical model parameters

Parameter	Range	Description	Parameter	Range	Description
L	15-45 m	Pile length	v_{fluid}	0.1-1 m/s	Fluid velocity
D	0.5-2.0 m	Pile diameter	k_{pile}	0.5-5.0 W/mK	Pile thermal conductivity
L_1/L	5%-45%	Layer 1 thickness ratio	k_1, k_2, k_3	0.5-5.0 W/mK	Thermal conductivities for layers 1, 2 and 3
L_2/L	5%-45%	Layer 2 thickness ratio	T_{inlet}	0-8 degC	Inlet temperature
n_{loops}	1-4	Number of pipe loops	$T_{initial}$	10-20 degC	Initial soil temperature

2.3 Artificial Neural Network

An ANN regressor was trained to predict the power output per unit length, P, at different times over the first year of operation. The input features included pile length and diameter, the proportions and thermal conductivities of the surrounding soil layers, concrete conductivity, and the number of pipe loops. The model predicts the power output at 16 discrete time steps, along with the final steady-state value, corresponding to the long-term thermal performance of the thermo-active pile. The model was built using the deep learning package Keras [4], and was trained using a 80-10-10 train-validation-test split. The architecture consisted of three fully connected layers with ReLU activation functions, containing 1200, 600, and 120 neurons, respectively. A dropout rate of 0.4 was included after each layer to reduce model overfitting. A starting learning rate of 0.01 was used, with a reduced learning rate technique applied for better, more robust learning. Early stopping was implemented, with a patience of 20 epochs monitoring the validation loss. The model was trained using a mean squared error (MSE) loss function.

3 Results

The results from the ANN surrogate model show a strong predictive capability in estimating the power per unit length prediction for a single pile under transient conditions. Tab. 2. presents the standard metrics used for regressor output evaluation, including the Mean Absolute Error (MAE) and the coefficient of determination (R^2), for all the model outputs. Overall, the model effectively captures variations in power output per unit length with high accuracy, with the average error for the steady state solution being limited to just 2.4 W/m.

Table 2: Model accuracy metrics for the test dataset

Time [day]	MAE [W/m]	R^2	Time [day]	MAE [W/m]	R^2
0.01	24.8	0.968	10.0	3.9	0.985
0.05	12	0.986	20.0	3.4	0.984
0.1	10.9	0.982	30.0	3.2	0.983
0.25	10.8	0.973	60.0	2.9	0.983
0.5	9.7	0.973	90.0	2.7	0.983
0.75	8.5	0.976	180.0	2.5	0.983
1.0	7.6	0.98	360.0	2.4	0.983
2.0	5.7	0.985	Steady State	2.4	0.983
5.0	4.4	0.987			

Fig. 1. shows a scatter plot comparing the real versus the predicted power per unit length for the test samples for three representative times: 1, 30 and 360 days. The results indicate high predictive accuracy, with a balanced distribution of underpredictions and overpredictions around the 1:1 reference line.

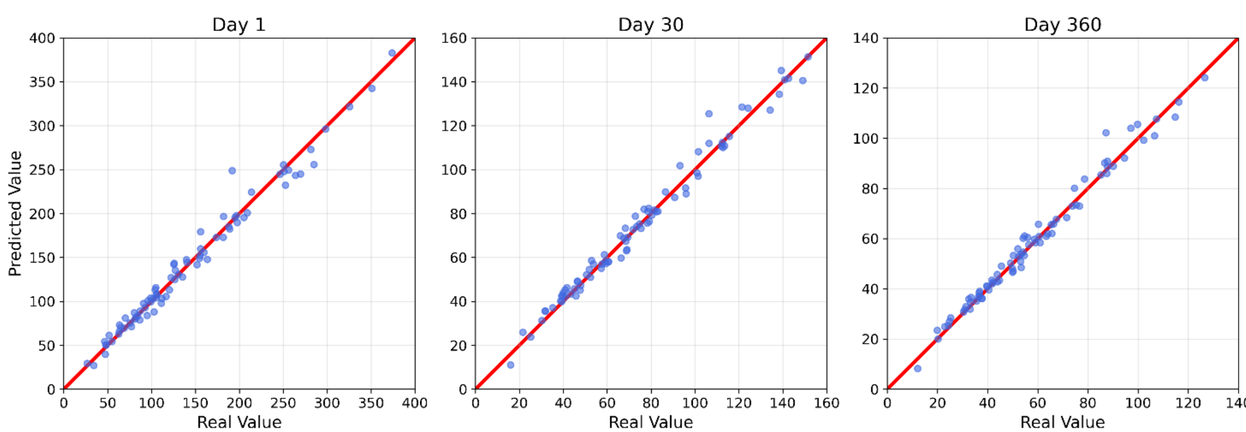

Figure 1: Real vs predicted model outputs for three representative times

Fig. 2. illustrates the predicted and actual power per unit length over time for three samples corresponding to the 25th, 50th, and 75th percentiles in MAE across all time steps. These demonstrate that, even for the less accurate 75th percentile case, the model captures the overall trend with consistent accuracy for all times, highlighting the robustness of the surrogate across the wide range of evaluated thermo-active piles.

A SHAP (SHapley Additive exPlanations) analysis was also carried out to improve the interpretability of the ANN by quantifying the relative importance of each of the input features. The results are consistent with the expectations: the temperature difference between the inlet and the initial ground temperature, number of loops and thermal conductivity of concrete are the most important factors in determining the power output. Fluid velocity plays a significant role during the initial transient period, but its importance diminishes for the steady-state, agreeing with available literature [5,6]. Since the power output is expressed per unit length, the influence of pile length appears reduced compared to a formulation based on total nominal power.

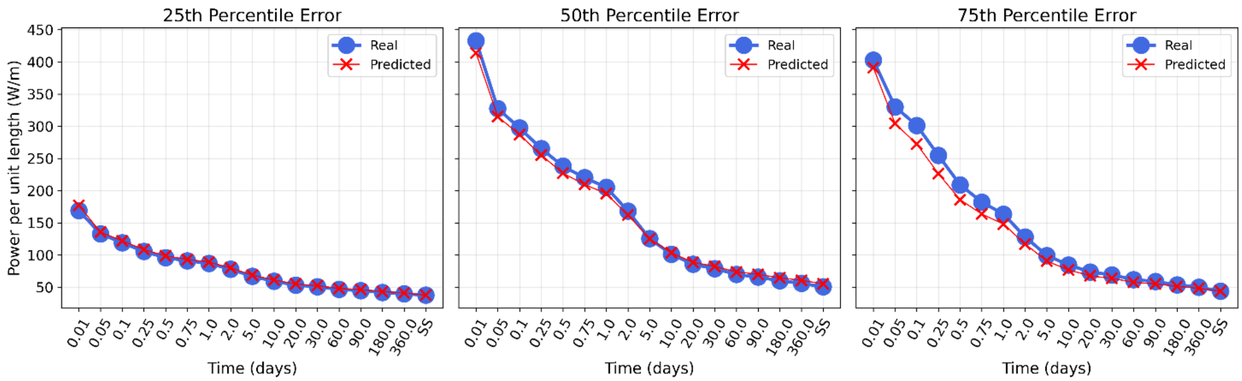

Figure 2: Predicted and actual power per unit length over time for samples at the 25th, 50th, and 75th percentiles of mean absolute error (MAE).

4 Conclusions

This study illustrates the development of an ANN-based surrogate model for predicting the power output per unit length of a single thermo-active pile over time. A wide range of ground conditions, pile geometries, material properties and pipe arrangements were included in order to establish a general model. The obtained surrogate model, trained on a dataset generated via FE simulations, accurately captures both transient and steady-state thermal behaviours. Performance metrics show the model's high predictive capability across a wide range of input conditions. A SHAP analysis was also carried out to provide further insight into feature importance of the model.

Acknowledgements

This research is part of the SaFEGround (Sustainable, Flexible and Efficient Ground-source heating and cooling systems) project, which is funded by the Engineering and Physical Sciences Research Council (EPSRC) (grant number: EP/V042149/1), and part of the SIDeTools (Sustainable Infrastructure Design Tools) project, funded by the Engineering and Physical Sciences Research Council (EPSRC) (grant number: EP/X52556X/1). This work was also supported by the Natural Environment Research Council [NE/S007415/1].

References

[1] R. Y. W. Liu and D. M. G. Taborda, 'The effects of thermal interference on the thermal performance of thermo-active pile groups', *Renew Energy*, vol. 225, no. March, p. 120357, 2024.

[2] H. Brandl, 'Thermo-active ground-source structures for heating and cooling', *Procedia Eng*, vol. 57, pp. 9–18, 2013.

[3] A. I. J. Forrester, A. Sóbester, and A. J. Keane, *Engineering Design via Surrogate Modelling*. Wiley, 2008.

[4] F. Chollet, *Deep Learning with Python*. New York, United States: Manning Publications, 2021.

[5] F. Loveridge and F. Cecinato, 'Thermal performance of thermoactive continuous flight auger piles', *Environmental Geotechnics*, vol. 3, no. 4, pp. 265–279, 2016.

[6] F. Cecinato and D. Salciarini, 'Energy performance assessment of thermo-active micro-piles via numerical modeling and statistical analysis', *Geomechanics for Energy and the Environment*, vol. 29, p. 100268, 2022.