Practical Approach for Data-efficient Metamodeling and Real-time Modeling of Monopiles using Physics-Informed Multi-fidelity Data Fusion

Stephen K. Suryasentana[1], Brian B. Sheil[2], Bruno Stuyts[3,4]

Affiliations

[1] Lecturer, Department of Civil and Environmental Engineering, University of Strathclyde, 75 Montrose St, Glasgow G1 1XJ, UK.

[2] Laing O'Rourke Associate Professor in Construction Engineering, Department of Engineering, University of Cambridge, Trumpington St, Cambridge CB2 1PZ, UK.

[3] Postdoctoral researcher, OWI-Lab, Vrije Universiteit Brussel, Pleinlaan 2, Elsene,1050 Brussels, Belgium.

[4] Visiting Professor, Geotechnical Laboratory, Ghent University, Technologiepark 68 B-9052 Gent, Belgium.

Corresponding author information

Stephen Suryasentana

stephen.suryasentana@strath.ac.uk

Main text word count: 2498

Figures: 11

Tables: 4

Oct 14, 2023

This is a peer-reviewed, accepted author manuscript of the following article: Suryasentana, S. K., Sheil, B. B., & Stuyts, B. (2024). Practical approach for data-efficient metamodeling and real-time modeling of monopiles using physics-informed multifidelity data fusion. Journal of Geotechnical and Geoenvironmental Engineering, 150(8), Article 06024005. https://doi.org/10.1061/jggefk.gteng-12395

Abstract

This paper proposes a practical approach for data-efficient metamodeling and real-time modeling of laterally loaded monopiles using physics-informed multi-fidelity data fusion. The proposed approach fuses information from 1D beam-column model analysis, 3D finite element analysis and field measurements (in order of increasing fidelity) for enhanced accuracy. It uses an interpretable ‘scale factor’-based data fusion architecture within a deep learning framework and incorporates physics-based constraints for robust predictions with limited data. The proposed approach is demonstrated for modeling monopile lateral load-displacement behavior using data from a real-world case study. Results show that the approach provides significantly more accurate predictions compared to a single-fidelity metamodel and a widely used multi-fidelity data fusion model. The model's interpretability and data efficiency make it suitable for practical applications.

Keywords

Foundations, Soil-structure interaction, Piles

List of notation

H	lateral load applied to pile
u_H	ground-level lateral displacement of pile
D	pile outer diameter
L	pile embedded length
T_w	pile wall thickness
y_L	low-fidelity data
y_H	high-fidelity data
y_{HH}	data that is higher fidelity than y_H
α	scale factor between low- and high-fidelity data
NN_H	neural network approximating high-fidelity data
NN_L	neural network approximating low-fidelity data
s_{u}	undrained shear strength
$s_{\mathrm{u}}^{\mathrm{av}}$	average undrained shear strength over the pile length

Introduction

Accurate pile design models are needed for optimizing foundation design and achieving cost-efficiency and reliability. The state-of-the-art in offshore pile design involves: (i) the use of simplified design models based on a one-dimensional (1D) beam-column monopile structure and non-linear soil springs (e.g., API 2011; DNV 2014; Suryasentana and Lehane 2014; Byrne et al. 2020a), which are fast and interpretable but may have limited accuracy for complex seabeds, and (ii) advanced three-dimensional (3D) finite element analysis (FEA), which can provide more realistic estimations (Jostad et al. 2020; Zdravković et al. 2020) but is computationally expensive, complex to set up and requires a suite of advanced laboratory tests for the calibration of the constitutive model. Often, a combination of both methods, conducted independently, is used to balance efficiency and accuracy.

Metamodeling has been proposed as a technique to reduce FEA computational costs for offshore pile design (Mentani et al. 2023). This approach develops fast surrogate models that can approximate the FEA results. Various techniques can be used for metamodeling, such as polynomial chaos expansion (Mentani et al. 2023), radial basis functions (Wang and Owens 2022), response surface methods (Khuri and Mukhopadhyay 2010), artificial neural networks (Shen et al. 2022) and kriging (Kang et al. 2015, Soubra et al. 2019). However, existing metamodels have limitations. First, they require up to hundreds of FEA simulations (Mentani et al. 2023) for training, which limits their applicability when FEA data is scarce and expensive to generate (especially if FEA modeling is outsourced). Moreover, the interpretability of these models may be limited. In the current paper, interpretability is defined as the extent to which a human can understand and trace the decision-making process of a model (Murdoch et al., 2019). For example, a linear regression model has high interpretability because the decision-making process is straightforward and can be easily traced. High interpretability allows users to have more confidence in the model predictions, which is vital for offshore foundation design certification.

Metamodeling is relevant for the design phase, where the objective is to use available site investigation data to optimize the pile foundation design. Real-time modeling is relevant for the operational phase, where the objective is to use available field measurement data to update the foundation performance predictions and ensure that the design is still safe. This is critical for foundation design security and will influence decision-making on foundation inspection and maintenance. However, existing model updating techniques (e.g., Conde et al. 2021; Sheil et al. 2022; Buckley et al. 2023) to assimilate available field measurement data to update the design predictions tend to be computationally demanding. Recent advances in machine learning provide significant opportunities to enhance the accuracy and efficiency of both offshore and onshore geotechnical engineering (Stuyts and Suryasentana 2023, Sheil et al. 2020a, b).

To address the above challenges, this paper introduces a practical approach for data-efficient metamodeling and computationally efficient real-time modeling of laterally loaded monopiles using physics-based multi-fidelity data fusion. Multi-fidelity data fusion is defined as the integration of data from a multitude of sources, each varying in accuracy and data volume (Peherstorfer et al. 2018). This improves the quality and robustness of the synthesized information by leveraging the strengths of each source. Physics-based multi-fidelity data fusion integrate physics-based laws to enhance the robustness of the fusion process. The proposed approach is based on a physics-informed multi-fidelity neural network (PIMFNN) that fuses three data sources: 1D beam-column model predictions, 3D FEA predictions and field measurements. PIMFNN incorporates physics-based constraints to ensure physically reasonable predictions. Moreover, PIMFNN provides uncertainty quantification using the Monte Carlo (MC) dropout technique (Gal and Ghahramani, 2016). This paper evaluates PIMFNN using data from a real-world case study (Byrne et al. 2020a, b). The performance of PIMFNN is compared against existing approaches such as single-fidelity metamodeling and a widely used multi-fidelity data fusion model (Kennedy and O'Hagan 2000).

Case study

This paper demonstrates PIMFNN as a proof-of-concept using data from the PISA case study (Byrne et al. 2020a, b), which involved lateral load field tests on monopiles in Cowden till. Four of the tested piles were analyzed with advanced 3D FEA, where the soil is modeled using an enhanced modified Cam Clay model. Detailed FEA implementation can be found in Zdravkovic et al. (2020b). Table 1 summarizes the pile dimensions and availability of 3D FEA results. Detailed ground conditions, including the undrained shear strength profile, can be found in Zdravkovic et al. (2020a).

The primary task is to predict the ground level lateral load-displacement behavior of the piles, up to a ground-level lateral displacement u_H of $0.1D$, where D is the outer pile diameter. The predictive output is the applied lateral load H and the inputs are D, L/D and u_H/D, where L is the pile embedded length. In this study, the pile wall thickness T_w was not included as an input as initial analyses suggested that T_w has negligible effect on the predictive output. This may be due to the limited range of T_w/D examined in this study. However, in larger-scale studies with a broader range of values, T_w could be a significant factor and should be duly noted.

This case study provides three data sources with different fidelities and sparsity. The lowest fidelity data are the 1D beam-column model predictions of the $H - u_H$ behaviour for all the piles in Table 1, based on the API soil reaction model (API 2011). The API soil reaction model is adopted as the purpose of the case study is to illustrate how the predictive task can still benefit from inaccurate low-fidelity predictions. The next higher fidelity data are the 3D FEA predictions for the four piles marked in Table 1, which were digitized from Zdravkovic et al. (2020a). The highest fidelity data are the field measurements of the $H - u_H$ behaviour for all the piles in Table 1, which were digitized from Byrne et al. (2020b). Detailed information on how the field measurements were obtained can be found in Byrne et al. (2020b).

Evaluation

The paper evaluates PIMFNN for two tasks: metamodeling and real-time modeling. Metamodeling is relevant for the design phase of monopiles, where the API model predictions are available for all design cases, but 3D FEA predictions are available only for a few cases. PIMFNN uses this small set of 3D FEA predictions to improve the API model predictions and estimate the '3D FEA-equivalent' predictions for all design cases. Real-time modeling uses available field measurements to correct any potential inaccuracies with the predictions made during the design phase. PIMFNN uses information from all three data sources (API model, FEA, field measurements) for this purpose.

Data fusion methodology

The motivation for multi-fidelity data fusion (MFDF) is that high-fidelity data (e.g. FEA predictions) are scarce due to cost and time constraints, while low-fidelity data (e.g. 1D beam-column model predictions) are abundant. MFDF uses information from all data sources to provide more accurate predictions than using any single data source alone. PIMFNN learns the relation between data sources of different fidelities using a scaling relationship:

$$y_H(\boldsymbol{x}) = \alpha(\boldsymbol{x})\, y_L(\boldsymbol{x}) \tag{1}$$

where y_L and y_H are the low- and high-fidelity output data corresponding to some inputs $\boldsymbol{x}$. For the current study, the output is H and the inputs are D, L/D and u_H/D.

$\alpha(\boldsymbol{x})$ in Eq. 1 is a scale factor that may vary with $\boldsymbol{x}$. This allows Eq. 1 to model non-linear scaling relationships between y_H and y_L. PIMFNN can fuse any number of data sources by applying Eq. 1 autoregressively e.g., $y_H = \alpha_L\, y_L$, $y_{HH} = \alpha_H\, y_H$, where y_{HH} is a data source that is higher fidelity than y_H.

This 'scale factor'-based fusion architecture is chosen for the following advantages: (i) it offers high interpretability as the concept of scale factors are aligned with common geotechnical engineering practice e.g., the shaft friction and end bearing capacity of piles in cohesive soil are calculated by applying a scale factor to the undrained shear strength of the soil (API 2011); and

(ii) it allows the incorporation of a physics-based hard constraint to guarantee the physical reasonableness of the model predictions, as will be explained later.

PIMFNN is based on a neural network framework, which consist of multiple layers. Each layer transforms the layer input $\boldsymbol{x}$ to the layer output $\boldsymbol{y} = f(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b})$, where $\boldsymbol{W}$ and $\boldsymbol{b}$ are the weights and biases of the layer, and f is a non-linear activation function. $\boldsymbol{W}$ and $\boldsymbol{b}$ in each layer are adjusted during training to minimize a loss function. Fig. 1a shows the core architecture of PIMFNN, which has two neural networks NN_L and NN_H. NN_L approximates $y_L(\boldsymbol{x})$, while NN_H approximates $\alpha(\boldsymbol{x})$ in Eq. 1. Both networks have two fully connected hidden layers with equal number of neurons and hyperbolic tangent activation. MC dropout, which randomly zeros some neurons during training, is applied to the last hidden layer of NN_H to quantify the model uncertainty and minimize overfitting. The dropout rate is 0.1, following Zhang et al. (2021, 2023). The networks are trained sequentially (NN_L, followed by NN_H) to save computation, as NN_L does not need retraining if only the high-fidelity data changes.

Physics-based or empirical constraints (Cuomo et al. 2022) can be added to neural networks to make predictions more robust and physically reasonable. These constraints can be soft or hard. Soft constraints are penalized in the loss function. They are easy to implement but may not satisfy the constraints everywhere. Hard constraints are enforced by modifying the network architecture itself; they satisfy the constraints everywhere, but implementation can be intractable for complex constraints. PIMFNN incorporates the hard constraint that its output must match the sign of the low-fidelity data, which prevents physically unreasonable predictions such as negative H when u_H is positive; this can happen if there are very limited high-fidelity training data. The hard constraint is enforced by adopting the data fusion architecture per Eq. 1 and applying an exponential activation at the last layer of NN_H to make the scale factor always positive-valued.

Metamodeling

The PIMFNN implementation used for the metamodeling task is called ‘Dual-Fidelity Neural Network’ (DFNN), which is represented by Fig. 1a. Here, the low- and high-fidelity data are the

API and 3D FEA model predictions, respectively. Since FEA predictions are available for only four piles in Table 1, the paper uses the 'leave-one-out' method to test DFNN's performance. It uses the FEA prediction for one pile as test data and the rest as training data. It does this for each of the four piles once. Table 2 shows the FEA training and test data for each iteration. DFNN uses the FEA training data in Table 2 as high-fidelity data and the API model predictions for all the piles in Table 1 as low-fidelity data.

The loss functions to train NN_L and NN_H of DFNN are standard data loss functions, as follows:

$$L_{DATA,API} = \frac{1}{N_{y_{API}}} \sum_{i=1}^{N_{y_{API}}} ((\tilde{y}_{API} - y_{API})^2) \quad (2)$$

$$L_{DATA,FEA} = \frac{1}{N_{y_{FEA}}} \sum_{i=1}^{N_{y_{FEA}}} ((\tilde{\alpha}_{FEA} - y_{FEA}/\tilde{y}_{API})^2) \quad (3)$$

$N_{y_{API}}$ and $N_{y_{FEA}}$ are the number of API and FEA training data (denoted as y_{API} and y_{FEA}), respectively. $\tilde{y}_{API}$ and $\tilde{\alpha}_{FEA}$ are the predictions of NN_L and NN_H, respectively. Eqs. 2 and 3 serve two distinct purposes. Eq. 2 trains NN_L to approximate the API data, which allows interpolation of the API data. Eq. 3 trains NN_H to learn the scaling relationship between the API and FEA data, which allows the leveraging of information from the API data to make improved predictions of the FEA data.

Real-time modeling

The PIMFNN implementation used for the real-time modeling task is called 'Tri-Fidelity Neural Network' (TFNN), which is represented by Fig. 1b. TFNN combines information from all three data sources (API, FEA, field measurements) by applying the PIMFNN framework autoregressively.

First, a DFNN is trained using all available FEA training data as high-fidelity data and the API model predictions for all the piles in Table 1 as low-fidelity data. Then, TFNN uses the trained DFNN as NN_L and uses an initial portion (e.g., $0 \leq u_H \leq 0.04D$) of the field measurements for all the piles in Table 1 as high-fidelity training data for NN_H; the remaining portion of the field measurements are used as test data to evaluate the forecast performance of TFNN.

Eqs. 2 and 3 are used to train the DFNN component of TFNN. The loss function L_H to train NN_H of TFNN adds a soft constraint L_P to the data loss $L_{DATA,FIELD}$, as follows:

$$L_H = w_0\, L_{DATA,FIELD} + w_1\, L_P \tag{4}$$

where

$$L_{DATA,FIELD} = \frac{1}{N_{y_{FIELD}}} \sum_{i=1}^{N_{y_{FIELD}}} ((\tilde{\alpha}_{FIELD} - y_{FIELD}/\tilde{y}_{FEA})^2) \tag{5}$$

$$L_P = \frac{1}{N_{y_{FORECAST}}} \sum_{i=1}^{N_{y_{FORECAST}}} \left(\left(\frac{\partial \tilde{\alpha}_{FIELD}}{\partial (u_H/D)} - 0 \right)^2 \right) \tag{6}$$

$N_{y_{FIELD}}$ is the number of field measurements training data, $\tilde{y}_{FEA} = \tilde{\alpha}_{FEA}\tilde{y}_{API}$ is the prediction of the DFNN component of TFNN and $\tilde{\alpha}_{FIELD}$ is the scale factor learned by NN_H of TFNN. Eq. 6 mitigates excessive scale factor fluctuations in the forecast region by driving the partial derivative of $\tilde{\alpha}_{FIELD}$ with respect to u_H/D (denoted as $\frac{\partial \tilde{\alpha}_{FIELD}}{\partial (u_H/D)}$) to be close to zero in the region. This involves calculating $\frac{\partial \tilde{\alpha}_{FIELD}}{\partial (u_H/D)}$ at evaluation points uniformly spaced in the forecast region, and guiding the network to learn a relationship for $\tilde{\alpha}_{FIELD}$ where $\frac{\partial \tilde{\alpha}_{FIELD}}{\partial (u_H/D)}$ are close to zero at those points. $N_{y_{FORECAST}}$ refers to the number of evaluation points. For example, if the forecast region is $u_H/D > 0.04$, $\frac{\partial \tilde{\alpha}_{FIELD}}{\partial (u_H/D)}$ values are calculated, using automatic differentiation, at $u_H/D =$ 0.045, 0.055, 0.065, 0.075, 0.085, 0.095, for each set of D and L/D in the training dataset. w_0 and w_1 in Eq. 4 are self-adaptive weights to improve convergence rate (Liu and Wang, 2019). After some experimentation on a subset of the training data, w_0 and w_1 are set as:

$$w_0 = \frac{L_{DATA,FIELD}}{L_{DATA,FIELD} + 0.001 L_P} \tag{7}$$

$$w_1 = \frac{0.001 L_P}{L_{DATA,FIELD} + 0.001 L_P} \tag{8}$$

Applying the 0.001 factor to L_p results in w_0 being significantly larger than w_1 initially. As the training progresses and $L_{DATA,FIELD}$ decreases, w_1 increases. This approach prioritizes achieving a good fit with the training data first, before applying the soft constraint.

Comparison with existing approaches

Kriging, a widely used metamodeling technique (Kang et al. 2015; Soubra et al. 2019), typically relies on a single data source for model training, referred to as single-fidelity kriging (SFK) in this paper. However, SFK's effectiveness in extrapolation is limited due to its dependence on spatial correlation assumptions that may not hold outside the sampled area (Oliver and Webster 2014). To address these challenges, multi-fidelity kriging (MFK) (Kennedy and O'Hagan 2000) can be used. MFK is a popular MFDF technique, which has been used in many scientific applications (e.g., Liu et al. 2022; Han et al. 2022). It operates on the principle of linear correlation between data sources. Fig. 2 shows a strong linear correlation (correlation coefficient of 0.997) between FEA and API model predictions, suggesting MFK's applicability.

For the metamodeling task, the paper compares DFNN with SFK and MFK. SFK uses only the FEA training data from Table 2, while MFK uses the same API and FEA training data as DFNN. SFK and MFK are used as benchmark models for standard metamodeling and MFDF, respectively, due to their extensive use in these areas. For the real-time modeling task, the paper compares TFNN with MFK (which uses the same API, FEA and field training data as TFNN), and SFK (which uses only the field training data).

Table 3 summarizes the neural network configurations for DFNN and TFNN. All neural networks are implemented using Pytorch (Paszke et al. 2019), while MFK and SFK are implemented in SMT (Saves et al. 2023).

Results

All load-displacement results are presented in the figures in normalized form $(\frac{H}{s_{\mathrm{u}}^{\mathrm{av}} D^2}, u_H/D)$, where $s_{\mathrm{u}}^{\mathrm{av}}$ is the average undrained shear strength along the pile length. The paper first evaluates all the models' abilities to approximate the training data for metamodeling. Fig. 3 compares the predictions using DFNN with those determined by MFK and SFK for CM2, CM9,

CM3 and CL2, when they are included in the training dataset. All models are shown to match the FEA data well. Fig. 3 also shows the model uncertainty of the DFNN predictions through the shaded bounds that represent the 5th to 95th percentiles (P5-P95) of 1000 random samples of the DFNN predictions with dropout (the same dropout rate of 0.1 that was used for training is also used for the sampling of the predictions). Fig. 4 shows the scale factors that DFNN has learnt, which are not very sensitive to u_H.

Fig. 5 shows the predictions of DFNN, MFK and SFK for each test case in Table 2. For all test cases, DFNN demonstrates higher accuracy than MFK and SFK, as corroborated by the root-mean-square (RMS) errors detailed in Table 4. The ability of DFNN to produce reasonable predictions given limited high-fidelity training data is notable, especially since most test cases are extrapolation tasks, as outlined in Table 2. MFK is more accurate than SFK, which highlights the benefits of fusing information with another data source, even if it is low in accuracy. However, MFK can produce negative-valued predictions (see Fig. 5a, c), which are not physically reasonable. This highlights the benefits of incorporating physics-based hard constraints in PIMFNN.

For the real-time modeling evaluation, Figs. 6 and 7 compare the predictions of TFNN with MFK, SFK and DFNN for the piles where FEA data is available and not available, respectively. TFNN's forecasts improve upon DFNN's predictions for all cases, as detailed in Table 4. Notably, TFNN provides reasonably accurate forecasts, even without FEA data (see Fig. 7). In contrast, MFK only offers reasonably accurate forecasts when FEA data is available, as can be seen by comparing Figs. 6 and 7. SFK is generally inaccurate for all cases.

As more field data becomes available, TFNN's forecasts will improve, as shown in Fig. 8 and Table 4 for the example pile case of CM3. Fig. 9 shows the scale factors learned by TFNN as more field data becomes available, which suggests that they generally do not vary much after $u_H/D > 0.06$.

Discussion

PIMFNN is generally more accurate than single-fidelity metamodeling and a widely used MFDF model. The physics-based constraints are critical for robust predictions with limited data. The reduced data requirements of PIMFNN enable practical applications with feasible data volumes. Interpretability via scale factors provides transparency and promotes engineering confidence. Note that the physics-based constraints are not limited to neural networks and may also be implemented in other machine learning models.

A major benefit of PIMFNN for real-time modeling is its non-intrusive and efficient nature. Model updating techniques (e.g., Conde et al. 2021) require intrusive access to update numerical models like FEA to match field measurements, which may not be feasible for users who outsource the modeling. PIMFNN does not require access to the numerical model. It is also efficient to fine-tune with new field data, compared to the time-consuming parameter optimization inverse analysis for FEA-based model updating. Thus, PIMFNN enables assimilation of new field measurements for real-time modeling, with low computational cost. However, a drawback of PIMFNN is that it does not support the identification of input parameters, a feature that is available in model updating and can be important for certain applications. Additionally, the neural network architecture behind PIMFNN requires more hyperparameter tuning (e.g., number of layers, type of activation function etc.) than other models such as kriging.

Some limitations remain to be addressed in future work. First, PIMFNN outperforms existing models with very limited data, but some prediction errors persist and the minimum FEA data requirements to attain a target accuracy remains unclear. Second, the performance of PIMFNN is uncertain for non-linearly correlated data sources, which could potentially apply for cyclic loading (e.g., Byrne et al. 2020c). Third, further investigation is needed to determine if PIMFNN can be applied to other foundation types like suction caissons, using similar low-fidelity 1D beam-column model predictions (Suryasentana et al. 2017, 2018, 2022, 2023a, 2023b; Suryasentana and Mayne 2022), and for other design tasks such as predicting failure envelopes

(e.g., Suryasentana et al. 2020, 2021) for ultimate capacity analysis under combined loading. Finally, this paper examines one case study with relatively uniform ground conditions as a proof-of-concept; more validation across a wider range of pile dimensions, ground conditions, and loading regimes would test the model generalization and enable further refinement of the model. To facilitate this, the PIMFNN code is available open-source at https://github.com/autogeolab/PIMFNN/.

Conclusion

This paper describes the use of physics-informed multi-fidelity data fusion techniques for investigating the response of monopiles subjected to monotonic lateral loading. The proposed approach combines information from three sources with varying levels of accuracy - a 1D beam-column model, FEA data, and field measurements (the latter two sourced from existing literature) - to enhance prediction accuracy, particularly in scenarios with limited high-fidelity data availability. The approach demonstrates good predictive capabilities in both interpolation and extrapolation scenarios, and holds potential for more effective interpretation of data from multiple sources for practical applications.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

The second author is supported by the Royal Academy of Engineering under the Research Fellowships scheme.

References

API (2011). Recommended practice 2GEO geotechnical and foundation design considerations. Washington (DC, USA): American Petroleum Institute.

Buckley, R., Chen, Y. M., Sheil, B., Suryasentana, S., Xu, D., Doherty, J., and Randolph, M. (2023). Bayesian optimization for CPT-based prediction of impact pile drivability. *Journal of Geotechnical and Geoenvironmental Engineering*, 149(11), 04023100.

Byrne, B.W., Houlsby, G.T., Burd, H.J., Gavin, K.G., Igoe, D.J., Jardine, R.J., Martin, C.M., McAdam, R.A., Potts, D.M., Taborda, D.M. and Zdravković, L., (2020a). PISA design model for monopiles for offshore wind turbines: application to a stiff glacial clay till. *Géotechnique*, 70(11), pp.1030-1047.

Byrne, B.W., McAdam, R.A., Burd, H.J., Beuckelaers, W.J., Gavin, K.G., Houlsby, G.T., Igoe, D.J., Jardine, R.J., Martin, C.M., Muir Wood, A. and Potts, D.M. (2020b). Monotonic laterally loaded pile testing in a stiff glacial clay till at Cowden. *Géotechnique*, *70*(11), pp.970-985.

Byrne, B.W., Aghakouchak, A., Buckley, R.M., Burd, H.J., Gengenbach, J., Houlsby, G.T., McAdam, R.A., Martin, C.M., Schranz, F., Sheil, B.B. and Suryasentana, S.K., (2020c). PICASO: Cyclic lateral loading of offshore wind turbine monopiles. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics* (ISFOG 2021).

Conde López, E. R., Toledo Municio, M. Á., and Salete Casino, E. (2021). Optimization of numerical models through instrumentation data integration: Digital twin models for dams. *Computational and Mathematical Methods*, 3(6), e1205.

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. (2022). Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3), 88.

DNV (2014). OS-J101 - Design of Offshore Wind Turbine Structures. Oslo: Det Norske Veritas.

Gal, Y., and Ghahramani, Z. (2016, June). Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In *The 33rd International Conference on Machine Learning*. New York, NY, USA2016. p. 1050-1059.

Han, T., Ahmed, K. S., Gosain, A. K., Tepole, A. B., and Lee, T. (2022). Multi-Fidelity Gaussian Process Surrogate Modeling of Pediatric Tissue Expansion. *Journal of Biomechanical Engineering*, 144(12), 121005.

Jostad, H. P., Dahl, B. M., Page, A., Sivasithamparam, N., and Sturm, H. (2020). Evaluation of soil models for improved design of offshore wind turbine foundations in dense sand. *Géotechnique*, 70(8), 682-699.

Kang, F., Han, S., Salgado, R., and Li, J. (2015). System probabilistic stability analysis of soil slopes using Gaussian process regression with Latin hypercube sampling. *Computers and Geotechnics*, 63, 13-25.

Kennedy, M.C. and O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1), pp.1-13.

Khuri, A. I., and Mukhopadhyay, S. (2010). Response surface methodology. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(2), 128-149.

Liu, D. and Wang, Y. (2019). Multi-fidelity physics-constrained neural network and its application in materials modeling. *J. Mech. Des*. 141 (12), 121403.

Liu, X., Zhao, W., and Wan, D. (2022). Multi-fidelity Co-Kriging surrogate model for ship hull form optimization. *Ocean Engineering*, 243, 110239.

Mentani, A., Govoni, L., Bourrier, F., and Zabatta, R. (2023). Metamodelling of the load-displacement response of offshore piles in sand. *Computers and Geotechnics*, 159, 105490.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, 116(44), 22071-22080.

Oliver, M. A., and Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, *113*, 56-69.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. and Desmaison, A. (2019). Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, *32*.

Peherstorfer, B., Willcox, K., and Gunzburger, M. (2018). Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3), 550-591.

Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J., Lefebvre, T., Hwang, J.T., Morlier, J. and Martins, J.R. (2023). SMT 2.0: A Surrogate Modeling Toolbox with a focus on Hierarchical and Mixed Variables Gaussian Processes. *arXiv preprint* arXiv:2305.13998.

Sheil, B. B., Suryasentana, S. K., Templeman, J. O., Phillips, B. M., Cheng, W. C., and Zhang, L. (2022). Prediction of pipe-jacking forces using a Bayesian updating approach. *Journal of Geotechnical and Geoenvironmental Engineering*, 148(1), 04021173.

Sheil, B. B., Suryasentana, S. K., Mooney, M. A., and Zhu, H. (2020a). Machine learning to inform tunnelling operations: Recent advances and future trends. *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, 173(4), 74-95.

Sheil, B. B., Suryasentana, S. K., Mooney, M. A., Zhu, H., McCabe, B. A., & O'Dwyer, K. G. (2020b). Discussion: Machine learning to inform tunnelling operations: recent advances and future trends. *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, 173(1), 180-181.

Shen, Q., Vahdatikhaki, F., Voordijk, H., van der Gucht, J., and van der Meer, L. (2022). Metamodel-based generative design of wind turbine foundations. *Automation in construction*, *138*, 104233.

Soubra, A. H., Al-Bittar, T., Thajeel, J., and Ahmed, A. (2019). Probabilistic analysis of strip footings resting on spatially varying soils using kriging metamodeling and importance sampling. *Computers and Geotechnics*, 114, 103107.

Stuyts, B., and Suryasentana, S. (2023). Applications of data science in offshore geotechnical engineering: state of practice and future perspectives. In *Proceedings of SUT OSIG 9th International Conference*, London, UK.

Suryasentana, S., and Lehane, B. M. (2014). Verification of numerically derived CPT based py curves for piles in sand. In *Proceedings of 3rd International Symposium on Cone Penetration Testing* (pp. 3-29). Las Vegas, Nevada USA.

Suryasentana, S. K., Byrne, B. W., Burd, H. J., and Shonberg, A. (2017). Simplified model for the stiffness of suction caisson foundations under 6 dof loading. In *Proceedings of SUT OSIG 8th International Conference*, London, UK.

Suryasentana, S. K., Byrne, B. W., Burd, H. J., and Shonberg, A. (2018). An elastoplastic 1D Winkler model for suction caisson foundations under combined loading. In *Numerical Methods*

in Geotechnical Engineering IX, Volume 2: Proceedings of the 9th European Conference on Numerical Methods in Geotechnical Engineering (NUMGE 2018), June 25-27, 2018, Porto, Portugal (p. 973). CRC Press.

Suryasentana, S. K., Burd, H. J., Byrne, B. W., and Shonberg, A. (2020). A systematic framework for formulating convex failure envelopes in multiple loading dimensions. *Géotechnique*, *70*(4), 343-353.

Suryasentana, S. K., Burd, H. J., Byrne, B. W., and Shonberg, A. (2021). Automated procedure to derive convex failure envelope formulations for circular surface foundations under six degrees of freedom loading. *Computers and Geotechnics*, 137, 104174.

Suryasentana, S. K., Burd, H. J., Byrne, B. W., and Shonberg, A. (2022). A Winkler model for suction caisson foundations in homogeneous and non-homogeneous linear elastic soil. *Géotechnique*, 72(5), 407-423.

Suryasentana, S. K., Burd, H. J., Byrne, B. W., and Shonberg, A. (2023a). Small-strain, non-linear elastic Winkler models for uniaxial loading of suction caisson foundations. *Géotechnique Letters*, 13(4), 170-181.

Suryasentana, S. K., Burd, H. J., Byrne, B. W., and Shonberg, A. (2023b). Modulus weighting method for stiffness estimations of suction caissons in layered soils. *Géotechnique Letters*, 13(2), 1-8.

Suryasentana, S. K., and Mayne, P. W. (2022). Simplified method for the lateral, rotational, and torsional static stiffness of circular footings on a nonhomogeneous elastic half-space based on a work-equivalent framework. *Journal of Geotechnical and Geoenvironmental Engineering*, 148(2), 04021182.

Wang, Q., and Owens, P. (2022). Reliability-based design optimisation of geotechnical systems using a decoupled approach based on adaptive metamodels. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *16*(3), 470-488.

Zdravković, L., Jardine, R.J., Taborda, D.M., Abadias, D., Burd, H.J., Byrne, B.W., Gavin, K.G., Houlsby, G.T., Igoe, D.J., Liu, T. and Martin, C.M., (2020a). Ground characterisation for PISA pile testing and analysis. *Géotechnique*, 70(11), pp.945-960.

Zdravković, L., Taborda, D.M., Potts, D.M., Abadias, D., Burd, H.J., Byrne, B.W., Gavin, K.G., Houlsby, G.T., Jardine, R.J., Martin, C.M. and McAdam, R.A., (2020b). Finite-element

modelling of laterally loaded piles in a stiff glacial clay till at Cowden. *Géotechnique*, 70(11), pp.999-1013.

Zhang, P., Jin, Y. F., and Yin, Z. Y. (2021). Machine learning–based uncertainty modelling of mechanical properties of soft clays relating to time-dependent behavior and its application. *International Journal for Numerical and Analytical Methods in Geomechanics*, 45(11), 1588-1602.

Zhang, P., Yin, Z. Y., and Sheil, B. (2023). Interpretable data-driven constitutive modelling of soils with sparse data. *Computers and Geotechnics*, 160, 105511.

Table 1. Summary of unique pile dimensions that were field tested in the Cowden till case study, where h is the height above ground for the lateral load application.

Pile name	D (m)	L/D	T_w/D	h (m)	3D FEA available?
CM2	0.762	3	0.013	10	Yes
CM9	0.762	5.25	0.014	10	Yes
CM3	0.762	10	0.033	10	Yes
CL2	2	5.25	0.012	10	Yes
CS2	0.273	5.25	0.026	5	No
CS3	0.273	8	0.026	5	No
CS4	0.273	10	0.026	5	No

Table 2. Summary of training and test sets of 3D FEA predictions from the Cowden till case study, where the test type is considered an extrapolation if the test pile diameter or L/D falls out of the respective range for the training piles.

Test case	Training piles	Test pile	D range	L/D range	Test type
T1	CL2, CM9, CM3	CM2	0.765 – 2	5.25 – 10	Extrapolation
T2	CL2, CM2, CM3	CM9	0.765 – 2	3 – 10	Interpolation
T3	CL2, CM2, CM9	CM3	0.765 – 2	3 – 5.25	Extrapolation
T4	CM2, CM9, CM3	CL2	0.765	3 – 10	Extrapolation

Table 3. Neural network configuration for the PIMFNN models (DFNN and TFNN) as per Fig. 1.

	NN_L of DFNN	NN_H of DFNN	NN_H of TFNN
No. of hidden layers	2	2	2
No. of neurons in each hidden layer	32	128	256
Activation function	Tanh	Tanh	Tanh
Optimiser	Adam	Adam	Adam

Table 4. Root-mean-square (RMS) errors between the test data and the predictions of DFNN, TFNN, SFK and MFK for Figs. 5 to 8.

Figure	DFNN	TFNN	SFK	MFK
5a	2.42	-	36.96	29.21
5b	2.06	-	25.87	19.88
5c	26.16	-	143.82	74.62
5d	114.48	-	1409.55	353.75
6a	3.65	0.84	298.55	2.74
6b	9.95	1.96	238.55	9.01
6c	22.03	21.94	121.35	21.96
6d	161.21	35.48	879.37	39.63
7a	2.25	0.14	295.89	1.62
7b	1.23	0.56	282.91	3.14
7c	4.14	1.28	307.31	5.29
8a	22.03	10.00	61.82	13.36
8b	22.03	4.45	17.66	9.28

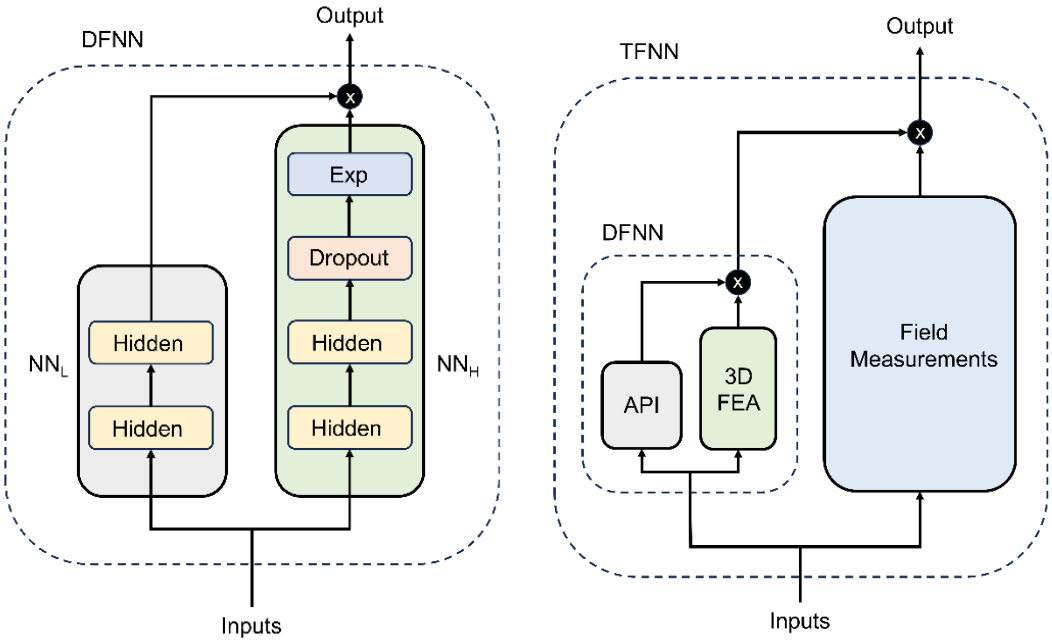


Figure 1. Schematic diagram of the PIMFNN architecture for: (a) DFNN model that integrates two data sources; (b) TFNN model that integrates three data sources. TFNN is made up of two PIMFNN units. First, a PIMFNN unit is trained on the API model and 3D FEA predictions to produce DFNN. Then, another PIMFNN unit employs DFNN as its trained NN_L and uses the field measurements to train its NN_H, resulting in TFNN.

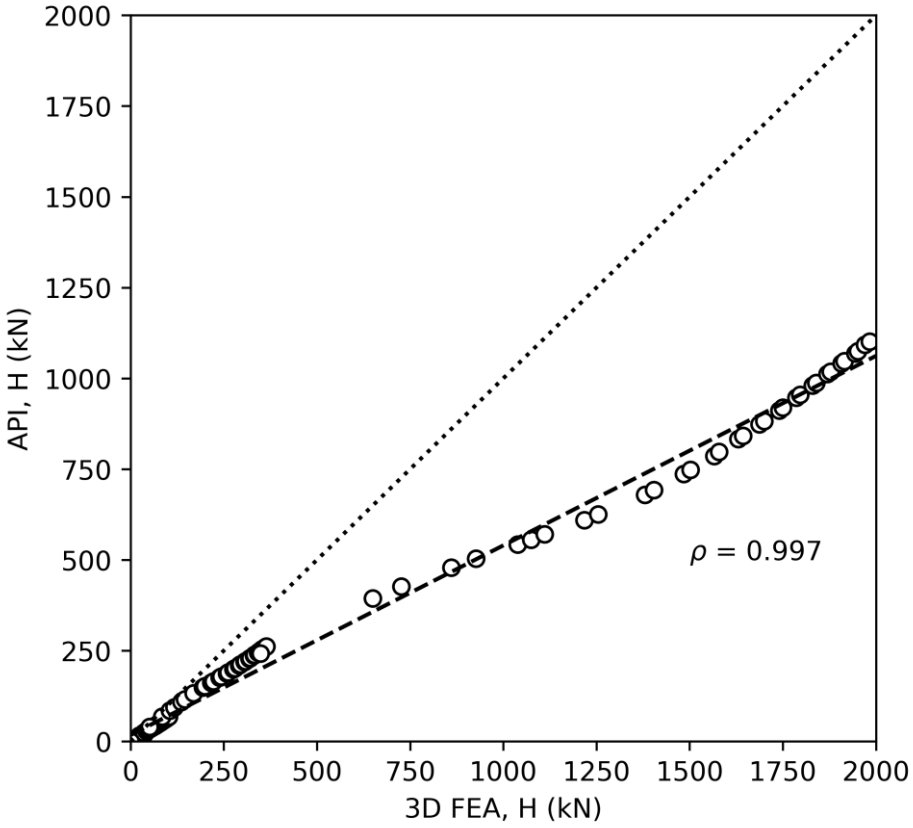


Figure 2. Comparison of the 3D FEA predictions with the corresponding API model predictions (see white markers), which shows a strong linear correlation (correlation coefficient of 0.996) between the two datasets. The dotted line is a 1:1 line for reference and the dashed line is the best linear fit between the datasets.

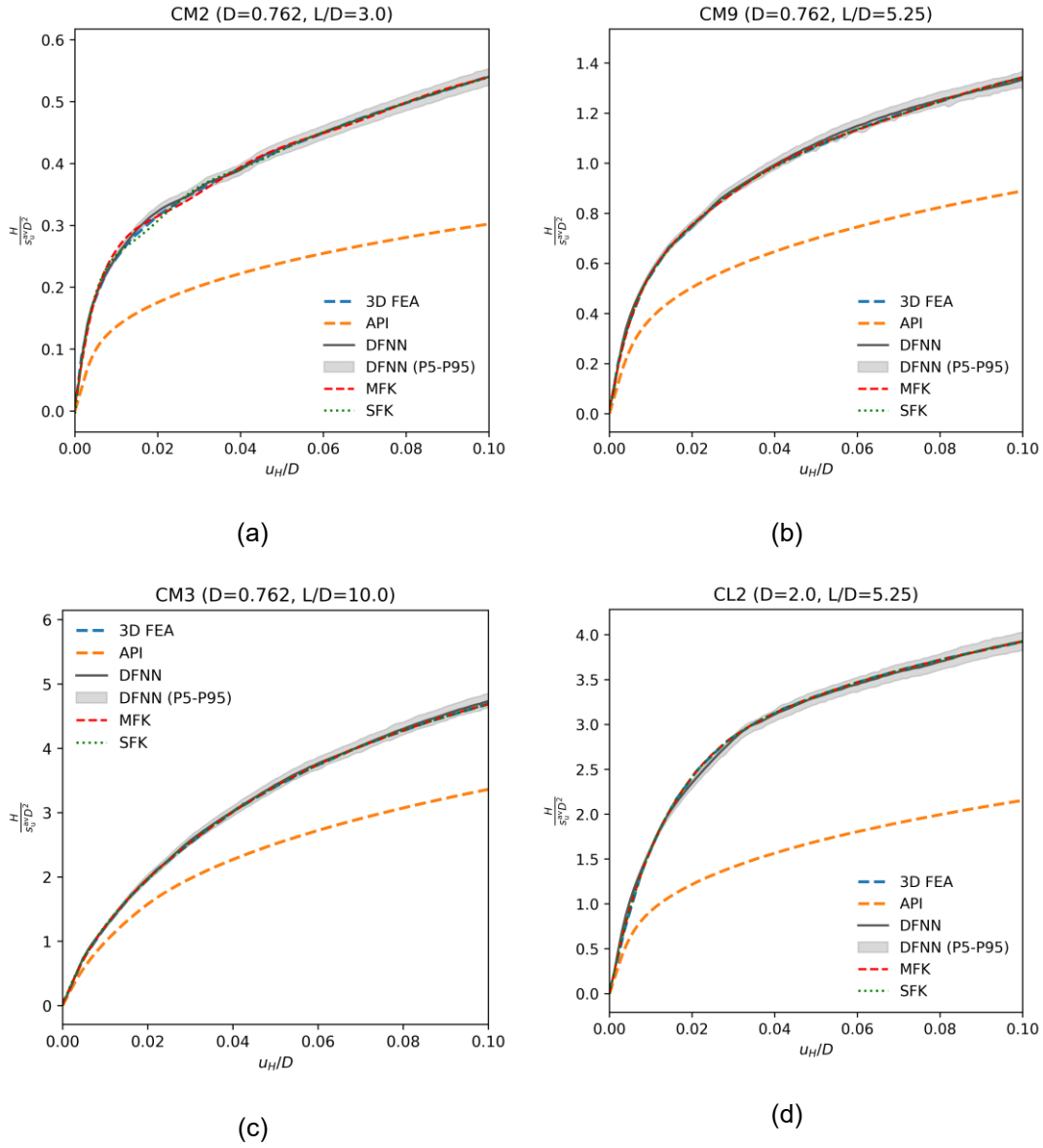


Figure 3. Comparison of normalized 3D FEA predictions with those determined using the DFNN, SFK and MFK models for (a) CM2 (b) CM9 (c) CM3 (d) CL2, when they are included in the training dataset. The API predictions are also included for comparison.

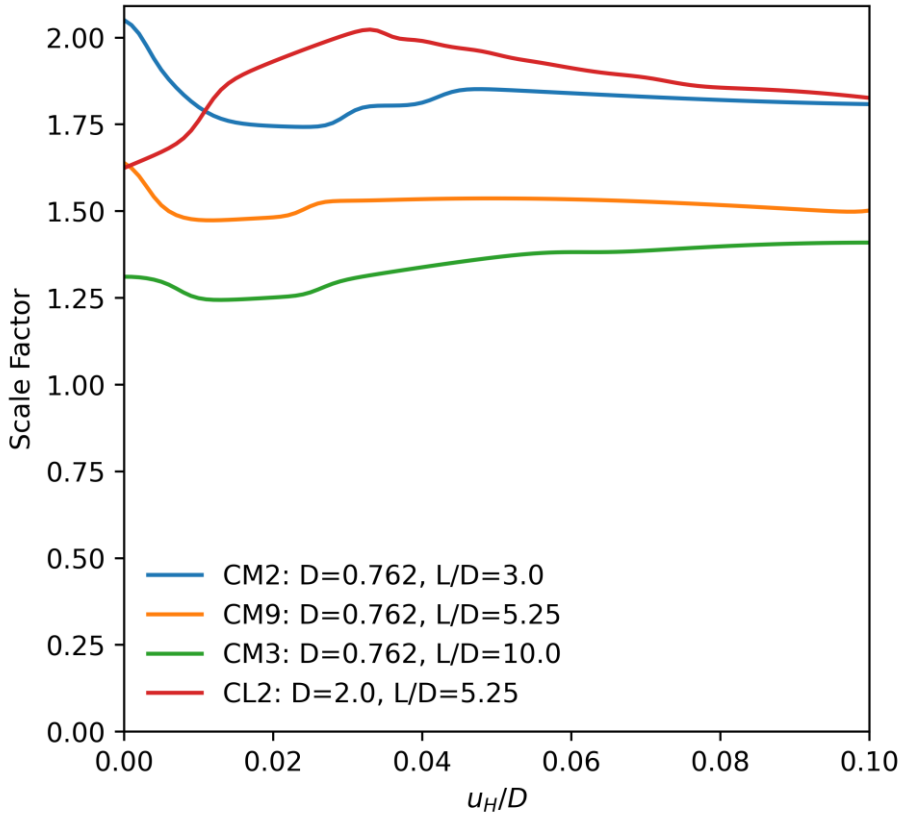


Figure 4. Comparison of the scale factors learned by DFNN

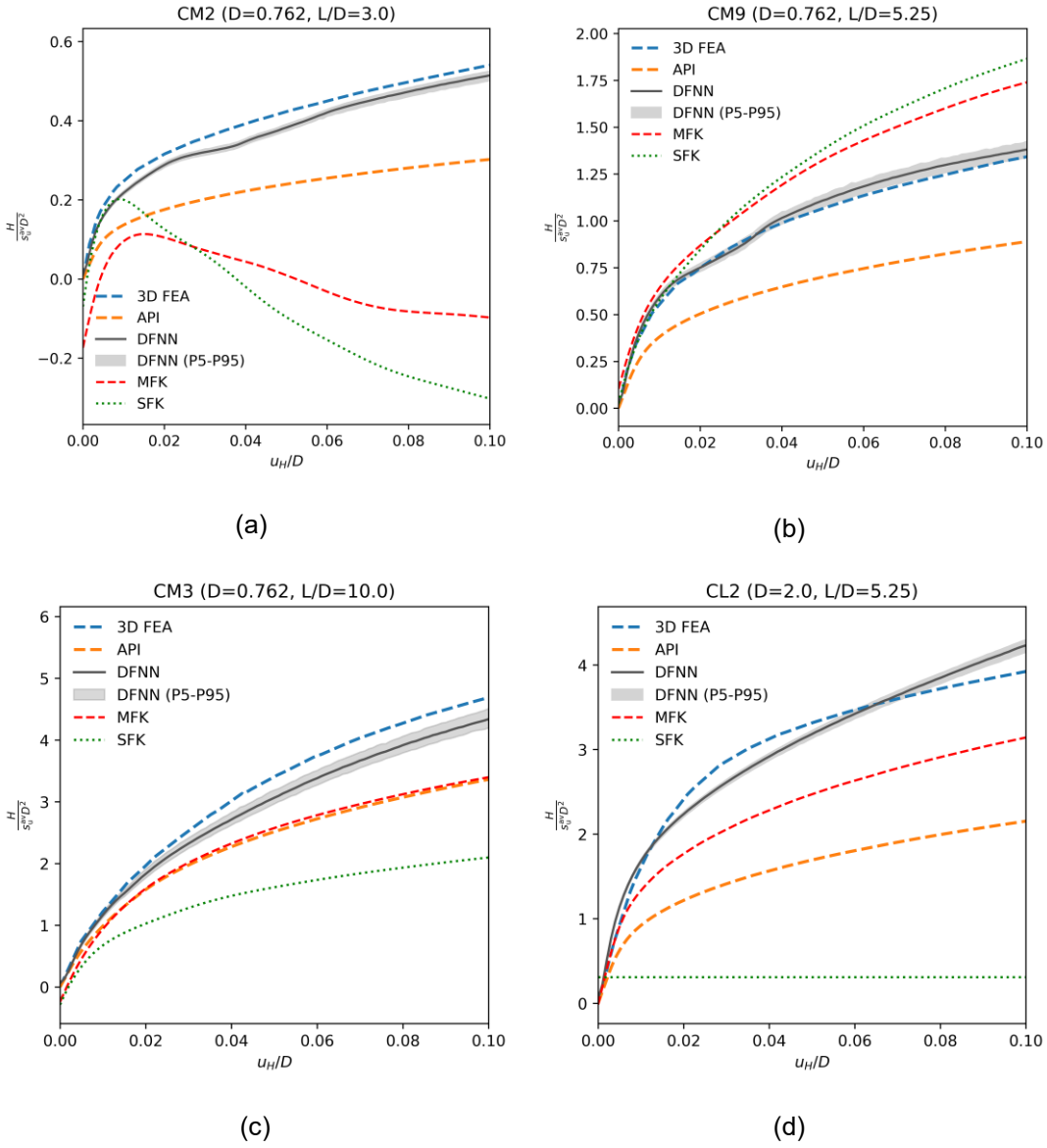


Figure 5. Comparison of normalized 3D FEA predictions with those determined using the DFNN, SFK and MFK models for test cases (a) T1 (b) T2 (c) T3 (d) T4 in Table 2. The API predictions are also included for comparison.

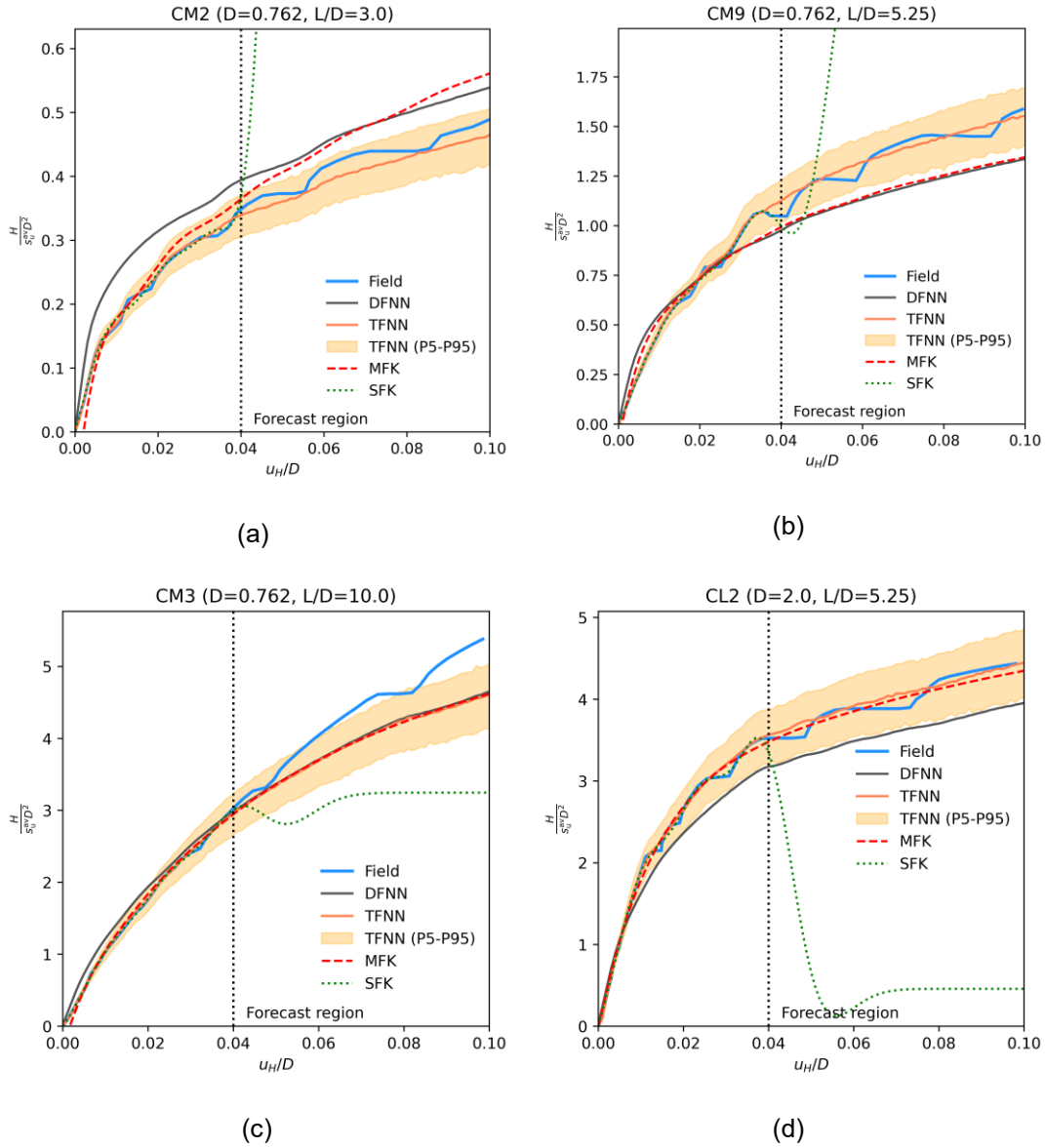


Figure 6. Comparison of normalized field test measurements with the predictions of the TFNN, DFNN, SFK and MFK for the cases where 3D FEA training data is available: (a) CM2 (b) CM9 (c) CM3 (d) CL2. Only field test data up to u_H/D = 0.04 are used to train TFNN, SFK and MFK.

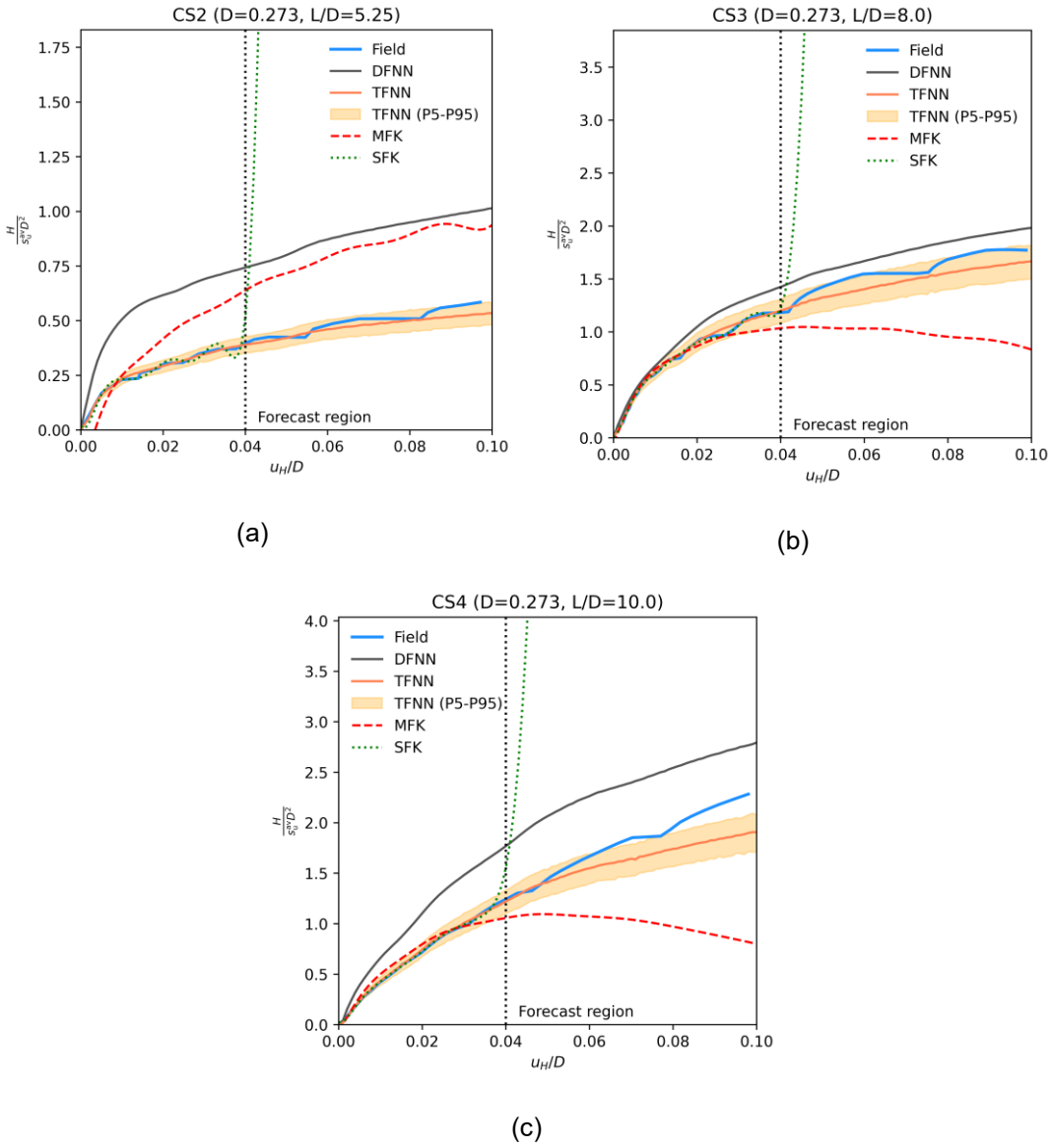
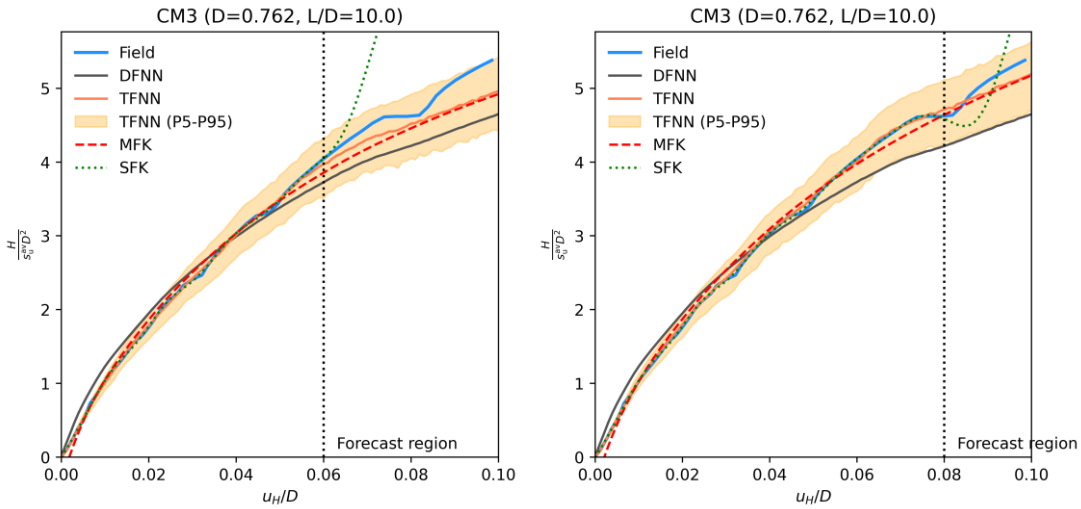


Figure 7. Comparison of normalized field test measurements with the predictions of TFNN, DFNN, SFK and MFK for the cases where 3D FEA training data is not available: (a) CS2 (b) CS3 (c) CS4. Only field test data up to u_H/D = 0.04 are used to train TFNN, SFK and MFK.



(a) (b)

Figure 8. Comparison of normalized field test measurements with the predictions of TFNN, DFNN, SFK and MFK for pile case CM3 as more field test data becomes available for training: (a) up to u_H/D = 0.06; (b) up to u_H/D = 0.08

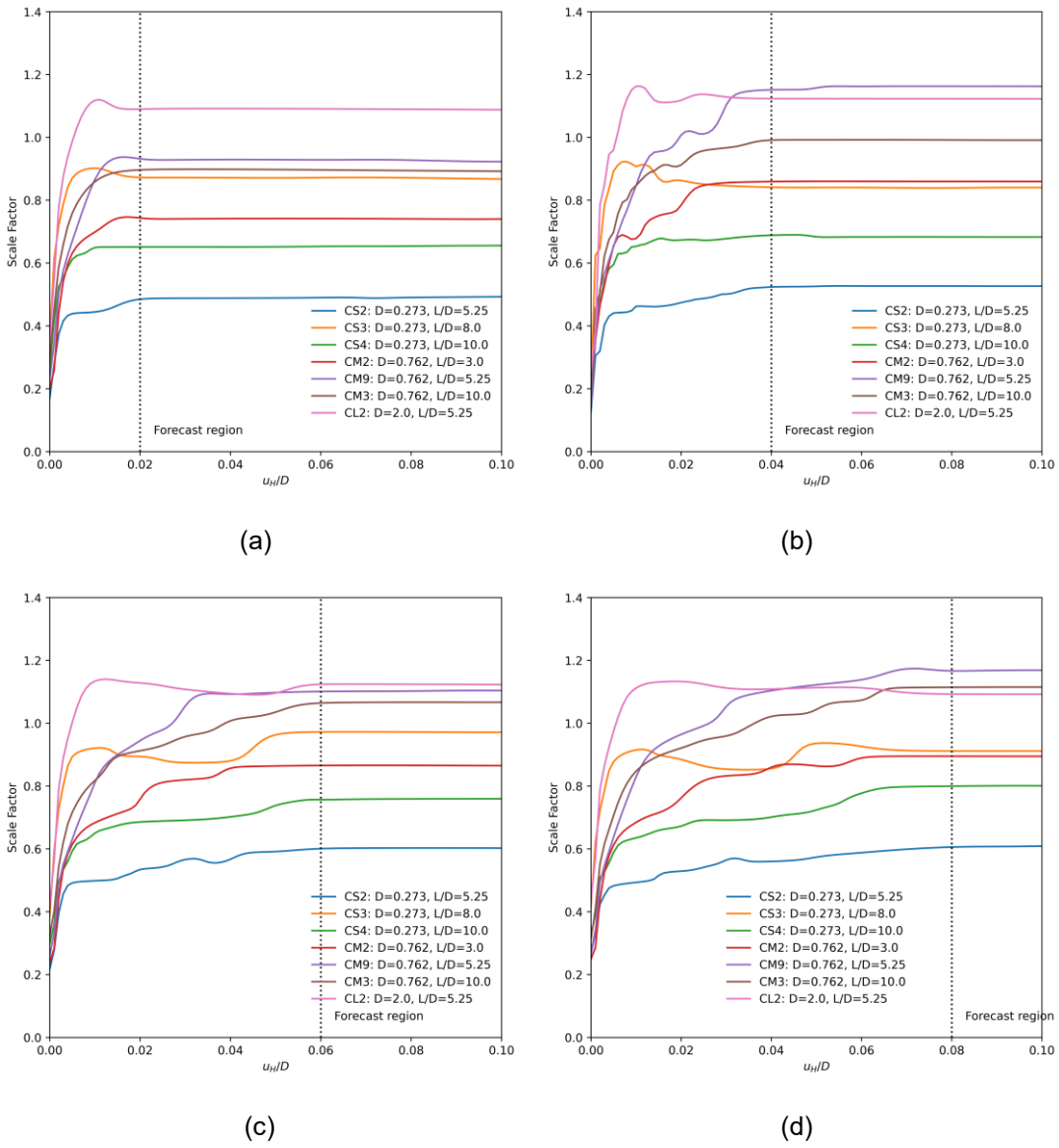


Figure 9. Comparison of the scale factors learned by TFNN as more field test data becomes available. The field data available for training is marked by the dotted vertical line in the above figures, up to u_H/D of (a) 0.02; (b) 0.04; (c) 0.06; (d) 0.08.