University of London
Imperial College of Science, Technology and Medicine
Department of Civil and Environmental Engineering

Data-driven approach to modelling bearing behaviors of OWT foundations

Ningxin Yang

Submitted on July 27, 2023

Abstract

Proper estimation of offshore piles is vital to the life and permanence of the foundation in question. However, due to the uncertainties of soil parameters in the field, the offshore piles are greatly affected. The source of soil uncertainties may come from various reasons, such as lack of uniformity between in-situ test and laboratory experiment; spatial variability of soil profile and rationality of the constitutive model, etc. Traditional statistical analysis which is based on Monte Carlo, is time-consuming and laborious. Though Bayesian theorem provides ways to understand and update the uncertainties, the amount of the inference analysis is still computationally heavy, thus bringing big challenges for the pile design. Additionally, to mimic the geotechnical structures and bearing behaviors as accurate as possible, digital twin is seeping into all kinds of engineering problems, enabling to evolve over time to persistently represent a unique physical asset and achieving data-driven decision making process. However, state-of-the-art digital twins are still relying on considerable expertise and deployment resources, leading to an only one-off implementation and remaining limitation on providing adaptive digital models on unique offshore piles.

In this thesis, we hope to introduce Markov Chain process involving in Bayesian inverse analysis to speed up the calculation and provide reasonable posterior results for the soil parameters. Besides, as a mathematical and rigorous foundation, probabilistic graphical model is introduced to support the transition from custom defined model towards accessible digital twins at scale. Based on such flexible asset-specific models, the entire loading life-cycle can be incorporated into a digital twin forming a unified and accessible foundation for a wide range of offshore piles. Combined with monitored data, the proposed dynamic updated digital twin provides rapid analysis results for reliable soil parameters and enables intelligent decision making on the pile bearing behaviors.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Modern pile installation and proper estimation is becoming increasingly complex and vital to the reliability and permanence of the foundation in question. However, in the construction process, uncertainties and insufficient information about the soil parameters lead to inaccurate predictions of pile-soil response and bearing capacities. The source of soil uncertainties may come from various reasons, such as lack of uniformity between in-situ test and laboratory experiment; spatial variability of soil profile and rationality of the constitutive model, etc. One typical soil profile can be illustrated in Figure 1.1, which shows:

- Fluctuating curve indicates the spatial variability
- Non-uniformity exists between in-situ test and laboratory experiment

Traditional design process usually does not allow leaner designs at the start of the project, which can bring difficulties to the understanding of the bearing behaviors of offshore piles. Fortunately, in recent years, with the rapid development of computational tools and techniques, the design process can obtain insights based on the accurate physical-based models to infer the underlying soil parameters, and thus provide reasonable predictions for pile behaviors. However, current statistical analysis is based on Monte Carlo, which is time-consuming and laborious. Though Bayesian theorem provides a possible tool to understand and update the uncertainties for the priors (Tarantola, 2005), the amounts of the inference analysis are still computationally heavy, thus bring big challenges for understanding the uncertainties and provide timely predictions for the pile design. Thus, it is necessary to reduce the analysis time and give instant response for pile design.

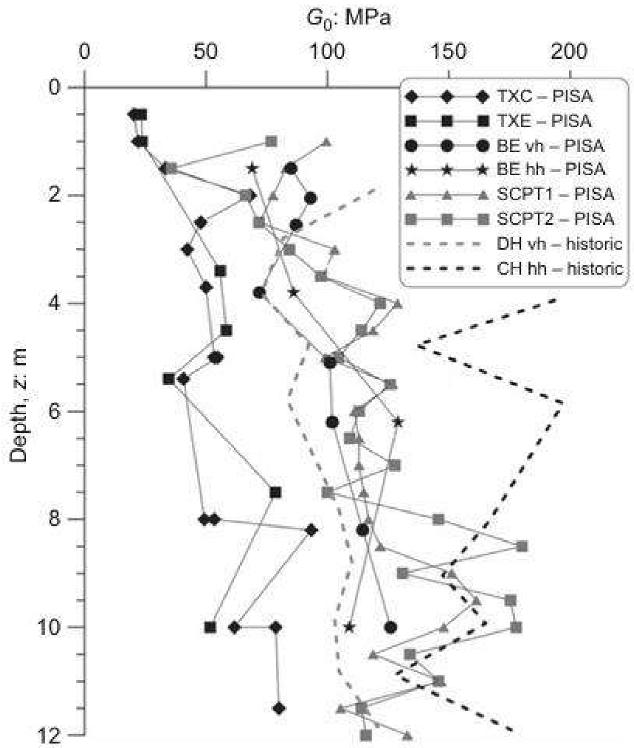

Figure 1.1: Stiffness characteristics at Cowden (Zdravković et al., 2020)

At the same time, with the increasing complexity and high-fidelity in geotechnical engineering problems, digital twin (DT) has gained popularity in handling abundant data and predict the structure's response in a more organized and accurate manner (Wang et al., 2021). DT makes full use of data such as physical models, sensor updates and operating history, and integrates simulation processes to real-time reproduce the dynamics of a physical system in the virtual space. More importantly, the DT model cannot only describe the current state of the physical entity, but also predict the future state. However, state-of-the-art digital twins are still relying on considerable expertise and deployment resources (Kapteyn et al., 2021), leading to an only one-off implementation and remaining limitation on providing adaptive digital models on unique offshore piles. Thus, scalable and unified models should be developed and incorporated into digital twin to enable a intelligent decision making.

Based on above, in practice, geotechnical engineers will benefit from having a site-specific digital model with a rapid analysis and intelligent decision making process.

1.2 Urgent need for a unified and scalable digital twins for piles

The demand for efficiency, reliability, and safety continues to grow in offshore wind foundation constructions. Computational models are an invaluable tool for understanding complex pile behaviors; they can be used to simulate new pile designs, operating conditions, or control strategies to explore their bearing behavior. This reduces the need for costly experiments or field tests. However, insights gained from a computational model are contingent on the model being an accurate reflection of the underlying soil parameters. Moreover, real-world pile foundations are constantly changing and evolving throughout their lifecycle. Using a single static computational model that ignores these differences fundamentally limits the specificity, and thus the accuracy, of the model and any insights gained through its use.

While the value proposition of digital twin has become widely appreciated in geotechnical engineering, the pile design process remains in a custom production phase. Current digital twin for offshore piles are still bespoken, relying on highly specialized implementations and thus requiring considerable resources and expertise to deploy and maintain. Therefore, it is necessary to move toward digital twins at scale by developing a rigorous and unified mathematical foundation. Meanwhile, based on the computational approaches and probabilistic graphical model, scholars in aerospace engineering (Kapteyn et al., 2021) proposed a declarative and general digital twin model by leveraging techniques from computational science and engineering, autonomous systems, and machine learning. In this way, this mathematical foundation enables a promising application at scale in the offshore pile bearing response.

1.3 Objectives and outline

The primary motivation of this thesis is to develop feasible and scalable digital twin model for offshore piles, based on the state-of-the-art techniques from aerospace engineering and computer science in order to enable predictive digital twin at scale. In particular, the specific goals of this thesis are:

- Develop a methodology for back calculating soil parameters enabling to speed up the analysis and give in-time response for understanding the underlying uncertainties of soil parameters.
- Develop a unifying mathematical foundation for predictive digital twins for offshore piles in the form of a probabilistic graphical model.

Chapter 2

State of the Art

2.1 Scale of fluctuation

Due to the complex geological and environmental processes involved, the soil characteristics in situ are rarely homogeneous. The soil characteristics can be highly variable and spatially correlated in the vertical and horizontal directions. As shown in Figure 2.1, a soil property $g(z)$ can be decomposed into a deterministic trend component $t(z)$ and a stationary random function $w(z)$ as follows:

$$g(z) = t(z) + w(z) \tag{2.1}$$

This concept of the Scale of fluctuation (SoF) was firstly proposed by Vanmarcke (1977), which provides an indicator of the estimated distance over a soil property. It is a convenient measure for describing the spatial variability of a soil property in a random field. A small SoF indicates that the soil oscillates quickly around its mean trend, whereas a large SoF indicates that the property is nearly spatially homogeneous.

For the values of the SoF, scales of fluctuation reported in the literature generally indicate that the horizontal SoF is larger than the vertical SoF. And the range for the vertical SoF is relatively narrow, from 0.06 to 2.6 m, for cone penetration test (CPT) results. However, the range for the horizontal SoF for CPT is fairly broad, from 0.14 to 80 m.

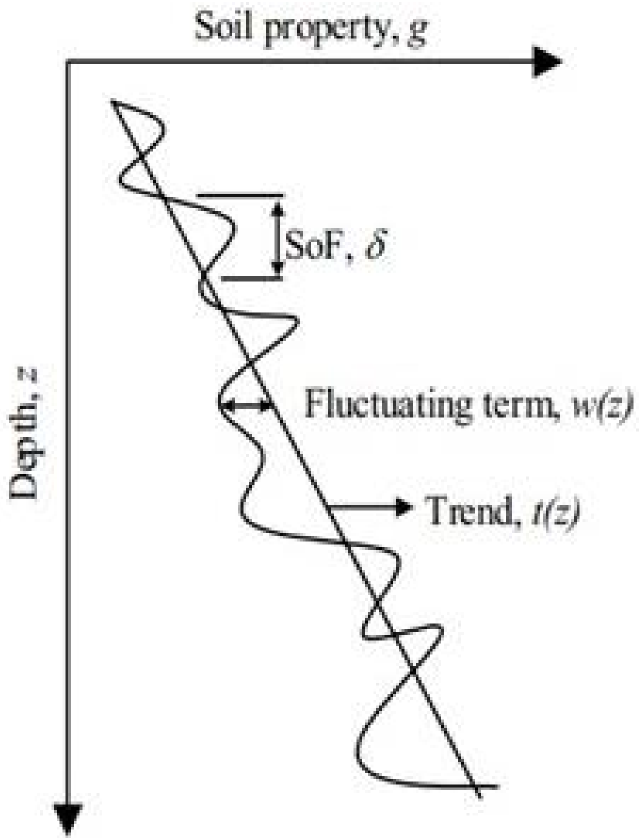

Figure 2.1: Illustration of the soil inherent variability (follow Nie et al. (2015))

For assessment, in the present study, the autocorrelation fitting method (ACFM) appears to be one of the most widely used methods for estimating SoF. The main idea of ACFM is to fit theoretical models to the sample autocorrelation function $\hat{\rho}\left(\tau\right)$based on an ordinary least squares approach.

Let $\bar{w}$ and $\hat{\sigma}^2$ denote the sample mean and the sample variance of $w\left(z\right)$. where $n\left(\tau\right)$denotes the number of pairs that are separated by the distance τ.The sample autocorrelation function can be obtained from the following equation:

$$\hat{\rho}\left(\tau\right) = \frac{\sum\limits_{i=1}^{n(\tau)} \left[w\left(z_i\right) - \bar{w}\right]\left[w\left(z_i + \tau\right) - \bar{w}\right]}{\left[n\left(\tau\right) - 1\right]\hat{\sigma}^2} \tag{2.2}$$

Chapter 3

Active learning for Bayesian inference

3.1 Reduced-order surrogate model

Computer modelling is used in nearly every field of science and engineering. Often, these computer codes model complex phenomena, have many input parameters, and are expensive to evaluate. In order to explore the behavior of the model under uncertainty (e.g., uncertainty propagation, parameter calibration from data or sensitivity analysis), many model runs are required. However, if the model is costly, only a few model evaluations can be afforded, which often do not suffice for thorough uncertainty quantification. In engineering and applied sciences, a popular work-around in this situation is to construct a reduced-order surrogate model. A reduced-order surrogate model is a cheap-to-evaluate proxy of the original model, which typically can be constructed from a relatively small number of model evaluations and approximates the input-output relation of the original model well. Since the surrogate model is cheap to evaluate, uncertainty quantification can be performed at a low cost by using the surrogate model instead of the original model. Therefore, surrogate modelling aims at constructing a metamodel that provides an accurate approximation to the original model while requiring as few model evaluations as possible for its construction.

3.2 Bayesian inference framework

3.2.1 Bayes theorem for soil parameter estimate

Bayesian theorem provides a possible tool to understand and update the uncertainties for the priors as shown in Equation (3.1).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{3.1}$$

$P(A|B)$: posterior: Distribution of soil parameters;$P(B|A)$:likelihood: observed data and FE simulation based on priors; $P(A)$: prior: Soil parameters from lab/field.

3.2.2 Maximise a posterior estimation (MAP)

3.2.3 Sampling methods

3.2.4 Maximal likelihood estimation (MLE)

3.3 Sequential enrichment for surrogate model

Instead of sampling the whole experimental design at once, it has been proposed to use sequential enrichment. Starting with a small experimental design, additional points are chosen based on the last computed sparse solution. In the context of machine learning, sequential sampling is also known as active learning. In all cases, numerical examples show that the sequential strategy generally leads to solutions with a smaller validation error compared to non-sequential strategies

3.4 Sequential Bayesian inference

Chapter 4

A unified framework for digtital twins for piles

This chapter develops a mathematical and computational foundation for digital twins of piles.

4.1 Markov Chain

As shown in Figure 4.1, Markov Chain is involved in Bayesian inference calculation and can therefore reduce the analysis time and provide reasonable posterior results.

$$X_0 \xrightarrow{P(X_1|X_0)} X_1 \xrightarrow{P(X_2|X_1)} X_2 \longrightarrow \cdots \longrightarrow X_n$$

Time t_0 Time t_1 Time t_2 $\cdots$ Time t_n

Figure 4.1: Markov Chain process

Markov Chain has two main features:

- State transition of Markov chains depends only on the current state, thus simplifying the modeling process
- Given the current state, a state transition probability distribution can be used to infer the likelihood of the next state

Thus, the properties of Markov Chain will bring two advantages:

- Adapt to more complex models (with increasing soil parameters)
- Iterative updating of parameter estimates because of Markov Chain's property

(enables updating the priors through Bayesian inference in stage)

4.2 Probabilistic graphical model of the pile-digital twin system

As shown in Figure 4.2, based on Markov Chain, with introducing Rewards and Actions, it can form the basis of Partially observed Markov decision process.

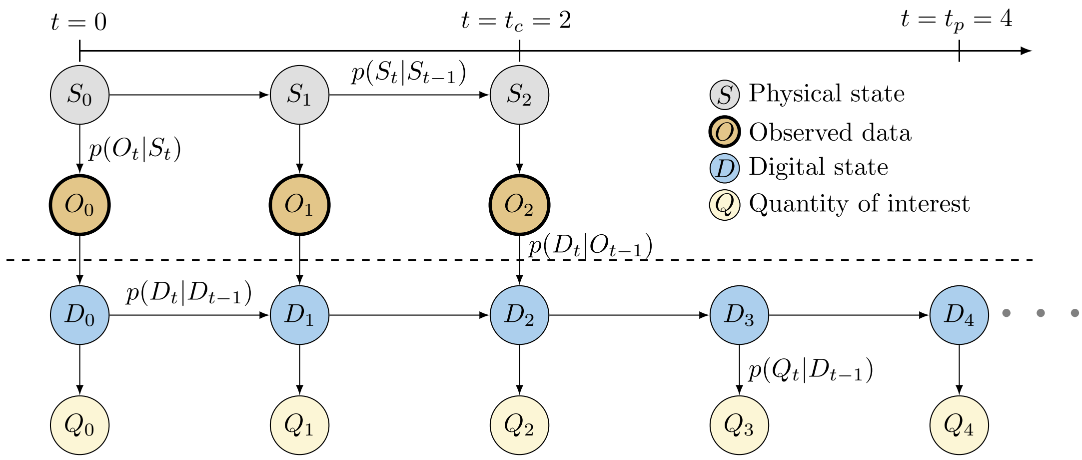

Figure 4.2: Untrained strength profile at Cowden Kapteyn et al. (2021)

Generally, Digital Twin can be divided into two main parts, including (1) calibration and assimilation (2) Prediction, as shown in Equation (4.1) and Equation (4.2).

$$
\begin{aligned}
&p(D_0, ..., D_{t_c}, Q_0, ..., Q_{t_c}, R_0, ..., R_{t_c}|o_0, ..., o_{t_c}, u_0, ..., u_{t_c}) \\
&= \prod_{t=0}^{t_c}[\phi_t^{update}\phi_t^{QoI}\phi_t^{evaluation}]
\end{aligned}
\tag{4.1}
$$

$$
\begin{aligned}
&p(D_0, ..., D_{t_p}, Q_0, ..., Q_{t_p}, R_0, ..., R_{t_p}, U_{t_c+1}, ..., U_{t_p}|o_0, ..., o_{t_c}, u_0, ..., u_{t_c}) \\
&\propto \prod_{t=0}^{t_p}[\phi_t^{dynamics}\phi_t^{QoI}\phi_t^{evaluation}]\prod_{t=0}^{t_c}\phi_t^{assimilation}\prod_{t=t_c+1}^{t_p}\phi_t^{control}
\end{aligned}
\tag{4.2}
$$

4.3 Planning and prediction via digital twin

Chapter 5

Focus of the Work

5.1 Objectives

Introduce Markov Chain process involving in Bayesian inverse analysis to reduce the inverse analysis time and provide reasonable posterior results for the soil parameters. Based on Markov Chain process, we also introduce the Reward function and Action function to form the basis of digital twin. Combined with monitored data, the proposed probabilistic graph model-Partially Observed Markov Decision Process can be used to mimic and predict the pile bearing behaviors.

Chapter 6

Work Plan

6.1 Stage 1

Calibrate the models in Figure 6.1 with partially observed Markov decision process.

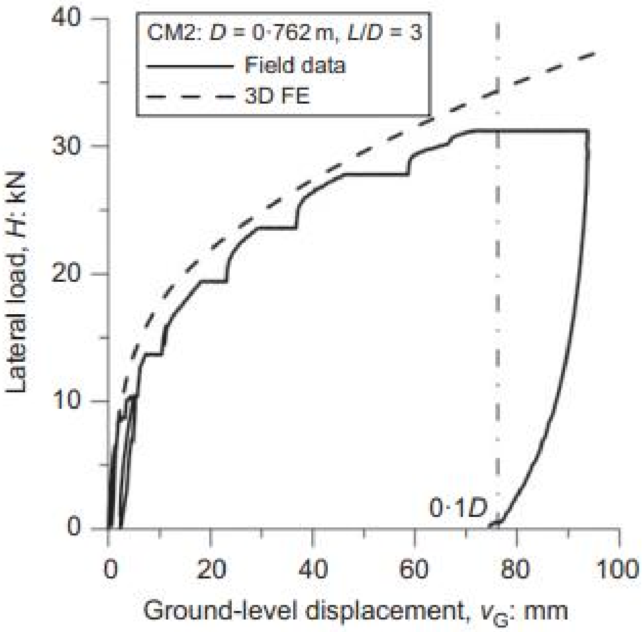

Figure 6.1: CM2 pile load displacement in Zdravković et al. (2020)

- Software: ICFEP-Likelihoods and observed data.
- Constitutive model: clay in Zdravković et al. (2020) and sand in Taborda et al. (2020).
- Consider the soil profile variance-Create the random field (scale of fluctuation in

ICFEP).

Objective: Ensure the soil parameters in digital model can reveal unique characteristics of piles.

6.2 Stage 2

In operational Phase, Based on Partially observed Markov decision process method, continue the assimilation process: extend the digital twin capability to capture the piles response during loading.

6.3 Stage 3

Extension to Prediction

6.4 Time plan

Table 6.1: PhD timeline

month	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48
Literature review	✓	✓	✓														
Numerical modelling (Data collection)		✓	✓	✓	✓	✓	✓										
Statistics Methods learning		✓	✓	✓	✓	✓	✓	✓	✓								
Statistics analysis calibration			✓	✓	✓												
Statistics analysis assimilation						✓	✓	✓	✓	✓	✓						
Statistics analysis prediction												✓	✓	✓			
Thesis writing															✓	✓	✓
Journal/Conference								✓				✓					✓

References

Kapteyn, M. G., Pretorius, J. V., & Willcox, K. E. (2021). A probabilistic graphical model foundation for enabling predictive digital twins at scale. *Nature Computational Science*, *1*(5), 337–347.

Nie, X., Zhang, T., Huang, H., Liu, Z., & Lacasse, S. (2015). Scale of fluctuation for geotechnical probabilistic analysis scale of fluctuation for geotechnical probabilistic analysis. In *Proc. of the 5th international symposium on geotechnical safety and risk* (Vol. 1, pp. 834–40).

Taborda, D. M., Zdravković, L., Potts, D. M., Burd, H. J., Byrne, B. W., Gavin, K. G., … others (2020). Finite-element modelling of laterally loaded piles in a dense marine sand at dunkirk. *Géotechnique*, *70*(11), 1014–1029.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.

Vanmarcke, E. H. (1977). Probabilistic modeling of soil profiles. *Journal of the geotechnical engineering division*, *103*(11), 1227–1246.

Wang, M., Wang, C., Hnydiuk-Stefan, A., Feng, S., Atilla, I., & Li, Z. (2021). Recent progress on reliability analysis of offshore wind turbine support structures considering digital twin solutions. *Ocean Engineering*, *232*, 109168.

Zdravković, L., Jardine, R. J., Taborda, D. M., Abadias, D., Burd, H. J., Byrne, B. W., … others (2020). Ground characterisation for pisa pile testing and analysis. *Géotechnique*, *70*(11), 945–960.