Generalised methods for the development of 1D models of monopiles for monotonic lateral loading



Ioannis Kamas

St. Catherine's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2023

Abstract

Generalised methods for the development of 1D models of monopiles for monotonic lateral loading

Ioannis Kamas

St Catherine's College, University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*

Michaelmas 2023

The offshore wind energy market has grown remarkably over the last decade. Monopiles are the most widely used foundation system employed for offshore wind turbines. Current design guidelines related to the ultimate limit state (ULS) and fatigue limit state (FLS) design of laterally loaded monopiles recommend the use of 1D models to make rapid calculations for the preliminary design of laterally loaded monopiles. Then a more complex time-consuming 3D finite element analyses can be undertaken to finalise and validate the ULS and FLS design of the monopile.

During the PISA project, a new rule-based design method was introduced together with a 1D finite element (FE) framework where the pile is modelled as a beam, for generating 1D models appropriate for rapid calculations corresponding to the ULS and FLS design of laterally loaded monopiles. In recent years there has been a shift towards using artificial intelligence to ease and assist the design procedures for geotechnical projects. This thesis explores the capabilities of a new data-driven design approach developed in Oxford University during the PICASO project. This approach utilises a new parametrisation of the soil reaction curves that employs a piecewise parameter spline function. A machine learning model – based on Gaussian Process Regression – is used to predict each soil reaction curve parameter. Subsequently, the machine learning model predictions can be employed in the 1D FE framework which was developed during the PISA project along with spline functions to regenerate the predicted soil reaction curves and compute the monopile's monotonic lateral behaviour. The data-driven 1D design models form an alternative to the PISA design models improving the accuracy, generality and applicability of the PISA design models.

This thesis outlines a methodology for developing high fidelity data-driven 1D design models using soil reaction curves data extracted by calibration 3D FE analyses. The data-driven 1D design models can compute the monopile's backbone curve and induced bending moments for different type of soil profiles including: (1) clay soils with arbitrary strength and stiffness profiles; (2) sand soils with arbitrary relative density and stiffness profiles; (3) layered soils with arbitrary number of layers and spatial variation of the layer thickness and soil strength and stiffness properties; (4) chalk profiles where the de-structured (due to installation effects) annulus thickness around the pile varies with the pile wall thickness.

Finally, the data-driven 1D design models are assessed in terms of assisting the practical design of monopile foundations for a realistic scenario within an offshore wind farm site. Integrating more data-driven 1D design models in the design procedure of monopiles can improve, speed and ease the ULS and FLS design.

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Harvey Burd and Prof. Byron Byrne, with whom I have been extremely privileged to work with throughout the course of my DPhil, which would not have been possible without their continuous guidance, valuable advice and support. In particular, I would like to thank Prof. Harvey Burd for constantly supporting me to overcome any difficulties I had come across and for encouraging me to develop my research and writing skills through his invaluable input and knowledge. I would also like to sincerely thank Prof. Byron Byrne as his motivation and advice has been extremely helpful and has significantly contributed to this thesis development.

This research project would not have been possible without the financial support of Ørsted, who have given me the opportunity to study and live in Oxford during the last four years. I would like to sincerely thank Vasiliki Tsaparli and Vasileios Avgerinos from Ørsted for taking such a keen interest in my work, for their assistance through providing me with valuable data, as well as for all the time they protected for our online meetings.

Also, I will never thank enough my best friend and colleague, Dimitris, who has always been more than willing to support me around any difficulties faced during this project and who has always encouraged me at every stage of this DPhil. My deep gratitude goes as well to my friends and colleagues Miad & Andrew, whose advice and feedback has been extremely valuable. Also, I am grateful for being part of the Civil Engineering Group and the CDT, as this has given me the opportunity to meet with so many inspiring people.

Finally, I would like to thank my parents for their encouragement and support throughout my studies as well as my girlfriend Olympia for her patience and for always encouraging me at every stage of my DPhil studies.

Contents

1 Introduction 1
1.1 Background 1
1.2 Monopile design requirements 4
1.2.1 Former design procedures 8
1.2.2 The PISA design model 9
1.2.3 PICASO project 12
1.2.4 Finite element modelling 14
1.3 Motivation 14
1.4 Terminology 16
1.5 Chapter summary 18
2 Literature Review 20
2.1 Behaviour of the monopile under monotonic response 20
2.2 Analysis procedures 22
2.2.1 1D Level 22
2.2.2 3D Level 31
2.3 Artificial intelligence applications in geotechnical engineering 33
2.4 Machine learning models 37
2.4.1 Gaussian Process Regression 37
2.4.2 Sparse Gaussian process regression 41
2.4.3 Artificial neural network 43
3 Finite Element Analyses of Laterally Loaded Monopiles 46
3.1 Introduction 46
3.2 Pile geometry calibration space 47
3.3 Generation of a 3D finite element model 48
3.4 Soil constitutive model selection 50
3.4.1 Numerical analysis of monopile embedded in the PISA Cowden till 51
3.4.1.1 Calibration procedure for the NGI-ADP model 53
3.4.1.2 Calibration procedure for the PIMS model 57
3.4.1.3 Validation of NGI-ADP and PIMS models with PISA FEA 58
3.4.2 Numerical analysis of monopile embedded in the PISA Dunkirk sand 61
3.4.2.1 Calibration procedure for the HS-Small model 64
3.4.2.2 Calibration procedure for the Norsand model 68
3.4.2.3 Validation of HS-Small and Norsand with PISA FEA 71

3.4.3 PISA offshore soil profiles constitutive model parameters 74
3.4.3.1 Calibration procedure for PISA Bothkennar offshore clay 74
3.4.3.2 Calibration procedure for PISA Dunkirk sand of D_R=45, 60, 90% 77
3.4.4 Analysis of the PISA field test piles 80
3.5 Summary 83
4 Data-Driven Method for Laterally Loaded Piles 84
4.1 Introduction 84
4.2 Soil reaction curves extraction 85
4.2.1 PLAXIS 3D interface element formulation 85
4.2.2 Soil reaction curves extraction procedure 87
4.2.3 Boundary equilibrium check on 3D model convergence 90
4.3 Data processing of soil reaction curves 90
4.4 Normalisation framework of soil reaction curves 91
4.5 Eight 'knot point' soil reaction parametrisation 92
4.6 Machine learning model selection 94
4.6.1 Feature selection for clay/sand soil profiles 96
4.6.2 Gaussian Process Regression training procedure 98
4.6.2.1 Assessment of GPR model 100
4.6.3 Feedforward Neural Network regression training procedure 109
4.6.3.1 Assessment of FNN regression model 110
4.6.4 Conclusion 117
4.7 Implementation of the machine learning predictions in the PISA 1D finite element framework 117
4.8 Application of the data-driven method through calibration analyses 120
4.9 Application of the data-driven method through design cases 124
4.10 Summary 127
5 Development of Generalised Data-Driven Models for Clay and Sand Sites 129
5.1 Introduction 129
5.2 Data-driven Method for Clay with Arbitrary Strength and Stiffness 130
5.2.1 Generation of representative offshore glacial clay till sites 130
5.2.2 Soil distributions and pile geometry calibration space 134
5.2.3 3D finite element analyses calibration procedure 137
5.2.4 Site-specific glacial clay till machine learning models 139
5.2.4.1 Training of the site-specific clay till machine learning models 139
5.2.4.2 Validation of the site-specific clay till machine learning models 140
5.2.5 Generation of data-driven general Cowden clay model (GCCM) 144
5.2.6 Applications of the GCCM 146

5.2.6.1 Design offshore monopiles embedded in trained clay till profiles..... 147
5.2.6.2 Offshore monopiles embedded in unseen clay till profiles 150
5.2.6.3 Offshore monopiles geometries falling out of the trained *L/D* geometry calibration space embedded in a trained site 152
5.2.6.4 Smaller diameter monopile geometries falling out of the geometrical calibration space 153
5.2.7 Assessment metrics..... 156
5.3 Data-Driven Method for Sand with Arbitrary Relative Density..... 159
5.3.1 Soil distribution and pile geometry calibration space 160
5.3.2 Generation of the GDSM V2..... 161
5.3.3 Applications of the GDSM V2 163
5.3.3.1 Design piles embedded in sand profiles with trained relative density . 163
5.3.3.2 Design piles embedded in sand profiles with unseen relative density . 165
5.4 Summary 166
6 Data-Driven Method for Layered Soils 168
6.1 Introduction..... 168
6.2 Specification of the layered soil study 169
6.3 Independent layer method..... 170
6.3.1 Parametric studies..... 172
6.3.2 Key observations 179
6.4 Layering effects on the soil reaction curves 180
6.5 Rule-based method for layered soils..... 185
6.5.1 Case studies 187
6.5.2 Key observations 188
6.6 Data-driven general layered soil models..... 188
6.6.1 General two-layered soil model..... 189
6.6.2 Extending the model to three layers 193
6.6.3 Key observations 198
6.7 Site-specific layered soil data-driven 1D design models..... 198
6.7.1 Site-specific layered soil at a single location..... 199
6.7.2 Site-specific layered soil 1D models accounting for spatial variations..... 203
6.7.3 Key observations 208
6.8 Summary 208
7 Layered Soil Modelling Approach for Practical Design 210
7.1 Introduction..... 210
7.2 Representative offshore site across the offshore wind farm 211
7.3 Pile geometry calibration space 214

7.4 3D Finite element modelling procedure 214
7.5 Calibration space of the data-driven layered soil models 215
7.6 GPR models training procedure 217
7.6.1 Training procedure 217
7.6.2 GPR model predictions 218
7.7 Application of the data-driven layered soil models to design cases 220
7.8 Summary 225
8 Data-Driven Method for Laterally Loaded Piles in Chalk 227
8.1 Introduction 227
8.2 3D finite element analyses 228
8.3 Soil reaction curves for field test piles 236
8.4 Influence of the de-structured chalk zones thickness 238
8.5 Calibration of a data-driven chalk model 240
8.5.1 Calibration finite element analyses 241
8.5.2 GPR model training procedure 242
8.5.2.1 Training Inputs 242
8.5.2.2 Training procedure 243
8.5.2.3 GPR model predictions 244
8.6 Application of the chalk model 247
8.7 Refining the chalk model for different thicknesses for the de-structured chalk zones.. 249
8.8 Summary 251
9 Conclusions 252
9.1 Introduction 252
9.2 Key Contributions 252
9.3 Integrating data-driven method in design procedure 256
9.4 Future Work 259
9.5 Summary 260
References 261
Appendix A 270
Appendix B 278

Nomenclature

Variables – Greek alphabet

β	Parameter controlling strength anisotropy (PIMS)
B	Parameter calibrated through the PISA seismic cone penetration tests for the Dunkirk sand test site
γ	Soil unit weight Shear strain
Γ	Void ratio corresponding to a mean pressure equal to 1 kPa (Norsand) Gamma function employed in the Matérn kernel
$\gamma_{0.7}$	Threshold shear strain for 70% reduction of G_o (HS-Small)
γ_f^C	Failure shear strain in compression (NGI-ADP)
γ_f^E	Failure shear strain in extension (NGI-ADP)
γ_f^{DSS}	Failure shear strain in direct shear stress (NGI-ADP)
δ	Pile-soil interface friction angle
ε	Error between the machine learning model predictions and the actual data
ε_α	Axial strain
ε_i	Independent and identically distributed Gaussian noise random variables
ε_q	Deviatoric strain
ε_{vol}	Volumetric strain
η	Local coordinates in PLAXIS elements
η_{sd}	Small displacement quality of fit
η_{ult}	Ultimate response quality of fit
θ	Hyperparameter values
κ	Slope of swelling line (MCC)
λ	Slope of virgin consolidation line (MCC)
μ	Mean of the posterior distribution
ν	Poisson's ratio Non-negative parameter controlling smoothness of the function employed in the Matérn kernel

ξ	Local coordinates in PLAXIS elements
ρ_d	Dry density
ρ_{sd}	Ultimate response ratio metric
ρ_{ult}	Small displacement ratio metric
Σ	Variance of the posterior distribution
σ	General stress-type variable Noise variance
σ'_c	Preconsolidation stress
σ_f	Signal variance
σ_n	Variance of the white noise kernel
σ'_v	Vertical effective stress
σ'_h	Horizontal effective stress
τ_0	Initial shear stress mobilisation (NGI-ADP)
ϕ'	Effective friction angle
ϕ_{cs}	Critical state friction angle
χ_{tc}	Material parameter governing the slope of the minimum dilatancy (Norsand)
ψ	Angle of dilation State parameter Angle of rotation of the monopile
ψ_B	Rotation at the base of the pile
ψ_0	State parameter at the start of the loading path

Variables – Latin alphabet

a_r	Strength reduction factor for chalk *p-y* curves
a_{str}	Positive parameter determining the transition depth of a strong layer
a_w	Positive parameter determining the transition depth of a weak layer
α	Regularisation term that controls the amount of regularisation in neural networks
b	Bias
c'	Effective cohesion
D	Pile's diameter
D^p_{min}	Maximum dilation (Norsand)
D_R	Relative density
d_{ice}	Thickness of ice sheet layer
E	Young's modulus Cost function
E_d	Generalised deviatoric strain

E_{50}^{ref}	Reference secant stiffness in triaxial compression (HS-Small)
E_{oed}^{ref}	Reference tangent stiffness in primary oedometric loading (HS-Small)
E_{ur}^{ref}	Reference triaxial unloading/reloading stiffness (HS-Small)
e	Void ratio
e_{cs}	Critical state void ratio at current mean effective stress
e_{max}	Maximum void ratio
e_{min}	Minimum void ratio
f_0	Natural frequency
F_x, F_y, F_z	Nodal forces
G_o	Small strain shear modulus
G_{ur}	Unload-reload small shear stiffness (NGI-ADP)
H	Applied horizontal force Hardening modulus (Norsand)
H_B	Base horizontal force (soil reaction component)
H_0 , H_ψ	Hardening parameters (Norsand)
h	Pile's eccentricity load
I	Identity matrix
J	Jacobian matrix
$k(x, x')$	Covariance function in Gaussian Process Regression (also referred as kernel)
K	Covariance matrix
K_f	Calibration parameter (NGI-ADP)
K_N	Elastic normal stiffness
K_o	Earth pressure coefficient
K_S	Elastic shear stiffness
K_v	Modified Bessel function
K_{xx}	Covariance matrix computed for the training inputs
L	Pile embedment length
l_i	Length scales of GPR covariance function (kernel)
M	Applied moment Number of inducing points or variables in sparse Gaussian Process regression models
M_B	Base moment (soil reaction component)
M_{tc}	Friction ratio at critical state in triaxial conditions (Norsand)
m	Distributed moment reaction (soil reaction component) Stress-dependency of stiffness (HS-Small) Mean function in Gaussian Process Regression.

N	Material parameter controlling the maximum stress ratio as a function of the minimum dilatancy (Norsand) Parameter indicating normal or Gaussian distribution Number of datapoints
n	Constant depending on the type of layered configuration
n^G	Stress-dependency of stiffness (Norsand)
OCR	Overconsolidation ratio
p	Distributed lateral reaction (soil reaction component)
p'	Mean effective stress
p'_{ref}	Reference pressure
p_{ult}	Ultimate pressure of the pile embedded in chalk
p_y	Initial yield load of the pile embedded in chalk
$p_\xi(Z)$	Prior information about the location of inducing inputs in sparse Gaussian Process Regression models
$p_\psi(\theta)$	Prior information of kernel hyperparameters in sparse Gaussian Process Regression models
q	Deviatoric stress
q_c	Cone penetration test tip resistance
q_u	Compressive strength
$q(u)$	Variational posterior in sparse Gaussian Process Regression
R	Rotation matrix
R_f	Failure ratio (HS-Small)
$r(x, x')$	Covariance matrix
s	Standard deviation
S	Softening parameter to model undrained conditions (Norsand)
s_u	Undrained shear strength
s_{uinc}	Increase of active shear strength with depth
t_1	Thickness of remoulded annulus
t_2	Thickness of fractured annulus
t_w	Pile's wall thickness
u	Pore water pressure profile
$u_{i=1,N}$	Inducing points for sparse Gaussian Process Regression models
u_x	Applied prescribed displacement
V	Applied vertical force
v_G	Lateral displacement at ground level
v	Lateral displacement
v_B	Lateral displacement of pile's base

w_i	Weights of the Gauss points Weight of micromodels in PIMS model Weight of each neuron in artificial neural network
x_r	Depth below rock surface
$\boldsymbol{x}, \boldsymbol{x}'$	Training or test inputs employed in the machine learning models
x, y, z	Coordinate system
y	Lateral pile displacement (*p*-*y* curves)
y_r	Lateral deflection in remoulded material (*p*-*y* curves)
y_s	Lateral deflection in structured material (*p*-*y* curves)
z	Depth below ground
$z_{i=1,N}$	Inducing inputs for sparse Gaussian Process Regression models
z_b	Depth of the boundary between two adjacent layers
z_L	Thickness of a soil layer
z_T	Transition zone depth of layering effects

Pile name abbreviations

C1, C2, C4, C5, C6, C8, C9, C10, C11	Offshore calibration piles adopted for PISA Cowden till model in Byrne et al. (2020b) and PISA general Dunkirk sand model in Burd et al. (2020b)
D1, D2, D2t	Design piles adopted in Byrne et al. (2020b) and Burd et al. (2020b)
CL1	PISA Cowden large diameter piles (Byrne et al., 2020a)
CM9	PISA Cowden medium diameter piles (Byrne et al., 2020a)
DL2	PISA Dunkirk large diameter piles (McAdam et al., 2020)
DM4	PISA Dunkirk medium diameter piles (McAdam et al., 2020)
P01	PICASO field test pile (Byrne 2021c)
A	ALPACA+ field test pile – dimensions provided by Prof. Byrne
B	ALPACA field test pile – dimensions provided by Prof. Byrne
C	ALPACA field test pile – dimensions provided by Prof. Byrne
D	Wind Support field test pile (Ciavaglia et al., 2017b)

Other common abbreviation

1D	1-dimensional
3D	3-dimensional
ALPACA	Axial-Lateral Pile Analysis for Chalk
API	American Petroleum Institute
ARD	Automatic Relevance Determination
CO_2	Carbon Dioxide
CapEx	Capital Expenditure
CoV	Coefficient of variance
CPT	Cone penetration test
DNVGL	Det Norske Veritas Germanischer Lloyd
ELBO	Evidence lower bound
FEA	Finite element analyses
FITC	Fully Independent Training Conditional
FNN	Feedforward Neural Network
GCCM	General Cowden clay model
GDSM	General Dunkirk sand model
GL	General layered
GPR	Gaussian Process Regression
GWEC	Global Wind Energy Council
HARM	Hyperplastic Accelerated Ratcheting Model
ILM	Independent layer method
MCC	Modified Cam Clay
ML	Machine learning
MSE	Mean square error
NC	Normally consolidated soil profiles
OC	Overconsolidated soil profiles
OCR	Overconsolidation ratio
OWF	Offshore wind farm
OWT	Offshore wind turbine
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PICASO	Pile Cyclic Analysis: Oxford and Ørsted
PISA	Pile Soil Analyses
PWP	Pore water pressure

RBLM	Rule-based layered method
RQD	Rock quality designation
SLS	Serviceability Limit State
TXC	Triaxial Compression
TXE	Triaxial extension
UCS	Unconfined compression strength
ULS	Ultimate Limit State
VFE	Variational Free Energy

Common subscripts, superscripts and diacritics

$\bar{X}$	Normalised value
X_0	Initial value
$\boldsymbol{X}$	Vector or matrix
X_n	Updated value referring to number of steps
$\boldsymbol{X}^{\mathrm{T}}$	Transpose vector
X^{Total}	Total unit when used in soil mechanics
X'	Effective unit when used in soil mechanics Training or test input when used in machine learning models
$\widehat{\boldsymbol{X}}$	Normal vector
X_i	The ith training input employed in machine learning models training procedure
X_{min}	Maximum value
X_{max}	Minimum value
$\boldsymbol{X}_{xx}$	Covariance matrix computed for the training inputs
X_*	Test input in machine learning models
$\boldsymbol{X}_{**}$	Covariance matrix computed for the test inputs
X_G	Value of displacement at ground level
$[X]$	Training input (feature) used in a machine learning model

Chapter 1

1 Introduction

1.1 Background

A worldwide effort for a more sustainable future by reducing the effects of climate change has motivated many countries to invest in various forms of renewable energy. The historic Paris Agreement in 2015 has set the scene for reducing the global average temperature rise to a limit of 2°C or preferably to 1.5°C compared to pre-industrial levels (United Nations, 2015). It is widely recognised that the substantial increase of annual carbon dioxide (CO_2) emissions (Figure 1.1) over the last 60 years is the primary driver of global climate change and temperature rise. To avoid the worst impacts of climate change, many countries including the United Kingdom have set the goal to reach 'net zero' emissions by the second half of this century.

Offshore wind energy constitutes a renewable unlimited clean form of energy that has been growing significantly in the last two decades. According to GWEC (2022) in 2021, a total capacity of 21.1 GW was added to the global grid from offshore wind farms (OWF). Figure 1.2 demonstrates the increase of the global capacity of large-scale offshore wind farms. Worldwide the aim is to install 380 GW of new offshore wind capacity by 2030 to achieve net zero targets (GWEC 2022). At the present time, the leading market for offshore wind energy is in China and Europe and up to 2030 it is expected that this market will

maintain a steady growth with additional new wind farm installations occurring in many other countries. The cost of offshore wind energy has been historically higher than onshore wind. However, offshore wind energy has several advantages over onshore wind. Offshore wind speeds are higher and steadier than the inland winds; the wind turbines are not limited by planning factors relating to the needs and preferences of local communities together with environmental concerns. This provides the possibility to construct larger turbines. The key to expanding the development of offshore wind energy is bringing down development, operating and maintenance costs.

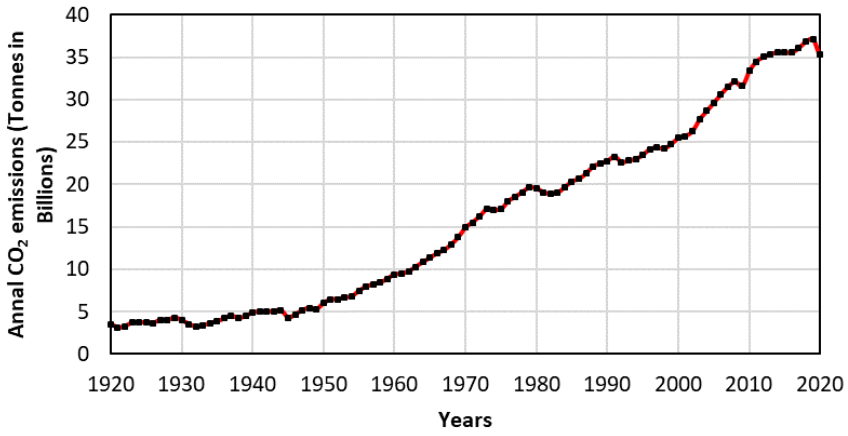


Figure 1.1: Annual carbon dioxide emissions worldwide from 1920 to 2021 (Ritchie et al., 2020).

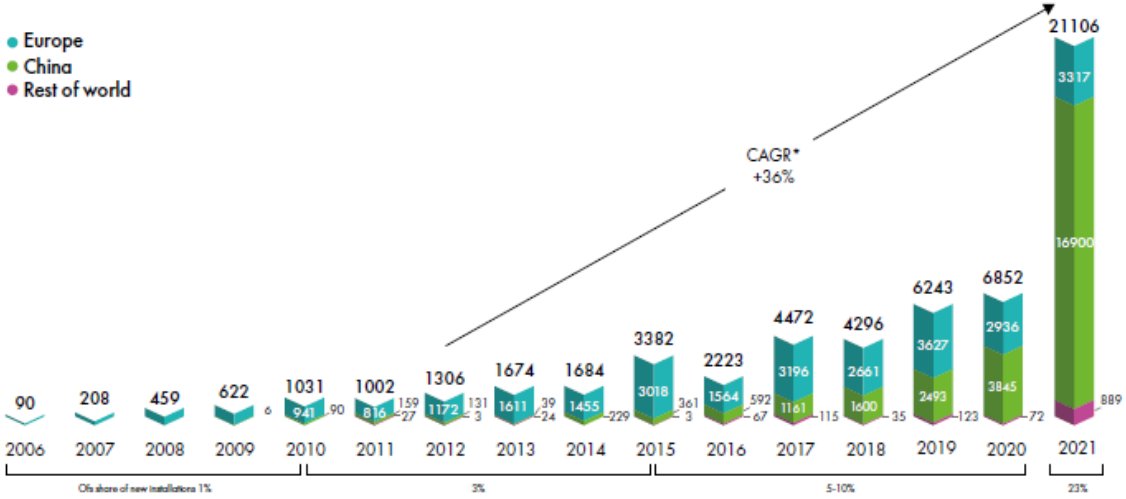


Figure 1.2: Total amount of worldwide grid connected offshore wind energy measured in Megawatts. (GWEC 2022).

The foundation system of an offshore wind turbine (OWT) is estimated to comprise approximately 20% of the total CapEx cost (Byrne et al., 2019). Several different foundation systems have been implemented worldwide depending whether the OWT is fixed-bottom or floating. Figure 1.3 illustrates possible foundation systems that can be used for an OWT. Typical fixed-bottom foundation examples are monopiles, gravity base, jackets and suction caissons, each having their own advantages and disadvantages based on the cost, drivability, structural and geotechnical bearing capacity (Wu et al., 2019). Currently, the leading foundation system in Europe is the monopile. According to Manceau et al. (2019) up to 87% of the OWT employ a monopile as their foundation system due to its simplicity in designing and manufacturing, its capability to be installed in various soil profiles as well as established knowledge of using this technology existing in the industry.

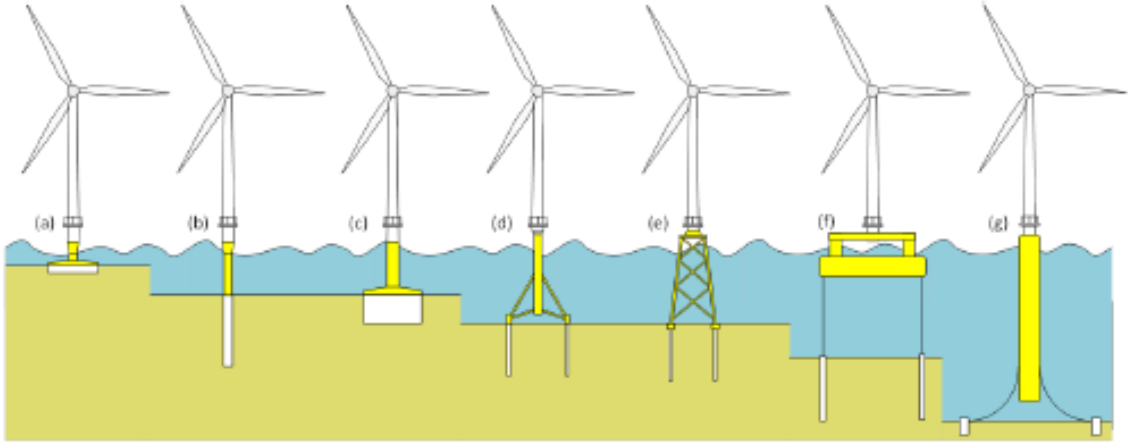


Figure 1.3: Possible foundation systems for OWT. (a) gravity-based foundation, (b) monopile, (c) suction caisson, (d) tripod, (e) jacket, (f) tension leg platform, (g) ballast stabilised spar buoy. (private communication with Prof. Byrne).

Monopiles employed for OWTs are hollow steel tubes (Figure 1.4). Typically, they that are driven by a hammer into the soil. They are connected to the tower of the OWT through a transition piece made of steel. Monopiles can be installed in different soil types such as sand, clays or gravels. Monopiles used for OWT are relatively short and have large diameters in contrast to the piles that have been used for many years in the oil and gas industry that are

slender and long. The OWT monopiles employ L/D ratios ranging between $2 \leq L/D \leq 6$. According to Byrne (2021a), recent monopile foundations typically have 8m to 10m of diameter, 30m to 50m embedment length; they weigh up to 1000 tonnes and have a wall thickness that varies with depth that can reach up to 150mm. Currently, monopiles are typically employed in water depths of up to 50m. The requirement to increase the amount of wind energy harvested is leading to a tendency of increasing the dimensions of turbines and of moving the structure into deeper water, increasing the magnitude of the loadings that are applied on the foundation system. As the OWTs move into deeper water extra-large monopiles are expected to be designed reaching to diameters up to 12-14m and water depths up to 60m.



Figure 1.4: A 7.5m diameter monopile foundation (see Kellahave et al. 2015).

1.2 Monopile design requirements

Lateral loads due to environmental loading conditions dominate the design of the monopile foundations for OWTs (Houlsby et al. 2016) in contrast to other conventional offshore foundations like jackets where axial loading dominates. OWT monopile foundations are subjected to significant loads and overturning moments from the action of wind on the turbine and tower and from the action of current and waves on the OWT substructure. These

loads result in a combined moment M and horizontal loading H on the monopile foundation at the ground level (Figure 1.5). The actual combined loading that is applied on the monopile is cyclic and pseudo-random since the environmental load components vary with time and direction. The cyclic loading shows variations in:

- The magnitude of loading, where millions of load amplitude cycles or hundreds of extreme storm events are subjected to the foundation over the lifetime of the turbine (Leblanc et al., 2010).
- The directionality of loading due to the spatial variations of the wind and wave loading that are not always in-plane (Rudolph et al., 2014).
- The loading frequency which can vary significantly in order of magnitude between 0.001–10 Hz (Arany et al., 2017).
- The resulting eccentricity ratio M/H as the relative contribution of the load components varies with time.

According to DNV GL (2018) the foundation design considerations for an OWT need to satisfy the following criteria:

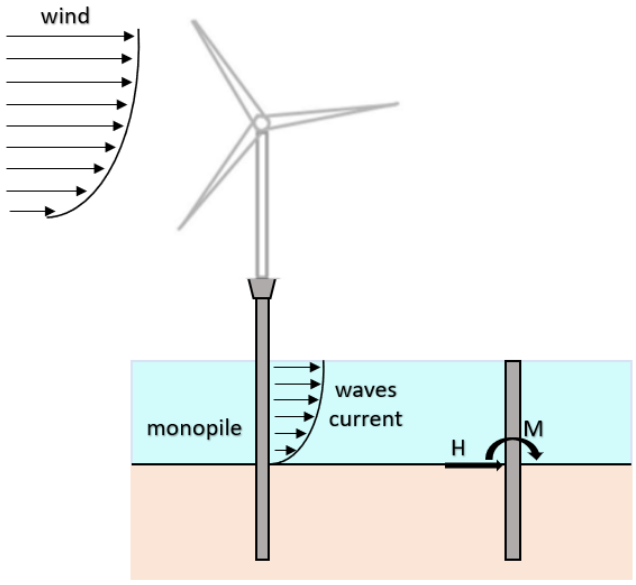


Figure 1.5: Loads applied on the offshore wind turbine substructure and distributed to monopile showing in-plane loading.

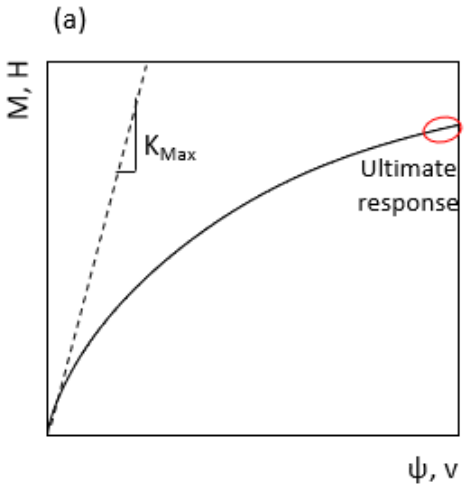
Ultimate Limit State (ULS): The maximum overturning moment M, lateral load H and the vertical load V, on the monopile foundation are estimated and they are compared with the calculated capacity of the selected foundation to ensure sufficient pile lateral resistance and to avoid both soil and pile failure. Therefore, ULS design considerations are focused on the monotonic response of the monopile and provide the minimum dimensions (embedment length, diameter) of the monopile. The ultimate capacity of the foundation is frequently estimated at a value of predetermined maximum allowed displacement v or rotation θ (refer to Figure 1.6(a)). DNV GL (2018) does not specify explicitly these limits, but typical values used in the design are: (i) a maximum displacement at ground level of $v_{G,ULS} = 0.1D$ (where D is the diameter of the pile) or (ii) a maximum rotation of the pile at ground level of $\theta_{G,ULS} = 2°$ (Byrne et al., 2020b; Burd et al., 2020a).

Serviceability Limit State (SLS): The effect of the long-term cyclic loading on the foundation is assessed by estimating the accumulated rotation on the foundation (ratcheting) with the number of loading cycles. The permanent deformation should not exceed a design limit set by the manufacturers of the turbine. Peralta et al. (2017) set an independent limit for the permanent rotation criterion at 0.25°. The permanent rotations govern together with the ULS design the determination of the embedded length and the diameter of the monopile. Cyclic loading influences also the soil strength and stiffness and the evolution of the hysteresis loop per loading cycle (refer to Figure 1.6(b)). This phenomenon should also be evaluated during the design procedure.

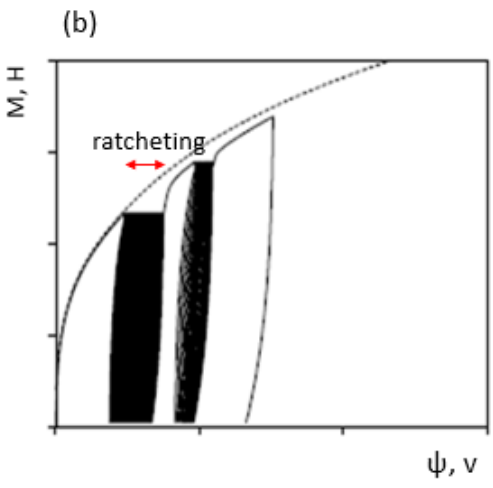
Fatigue Limit State (FLS): The fatigue life of the foundation depends on the prediction of the natural frequency f_0 of the whole turbine support system which is related to small amplitude vibrations. The required monopile's wall thickness depends on the target value of f_0 calculated for the foundation as well as on the pile's induced ULS bending moments (to avoid the yielding of the pile's steel material). Kellahave et al. (2015) presented measured

values of f_0 for various designed OWT. The value of f_0 depends on the initial foundation stiffness at very small displacements (e.g. $v_G = D/10000$ presented in Byrne et al. (2020b) and Burd et al. (2020a)) (Figure 1.6(a)). OWT structures are designed such as the f_0 falls between the f_{1P} (rotor) and f_{3P} (blade passing) excitation frequency (Figure 1.6(c)) to minimise the dynamic amplification and reduce the fatigue damage. The dynamic response of the structure is also affected by the aerodynamic, hydrodynamic, structural and soil damping; damping decays the amplitude of the induced vibrations. Typical design values of the structural critical damping according to Kellahave et al. (2015) range from 0.9 to 1.2%.

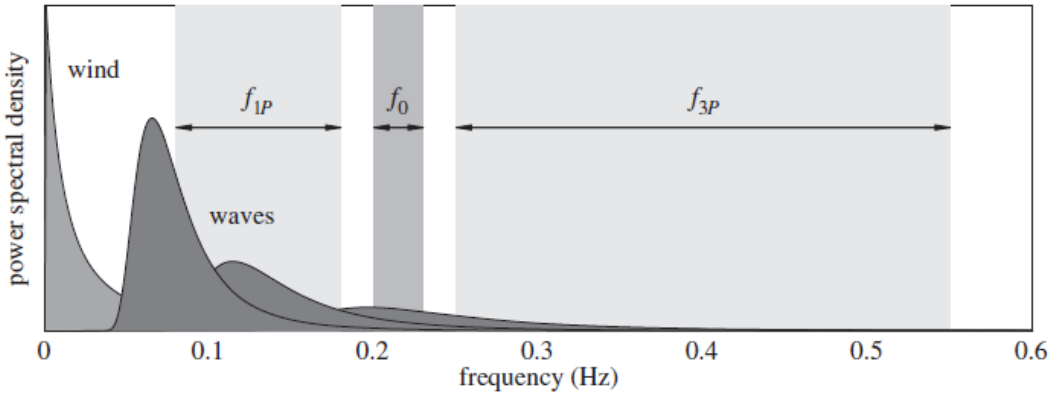
Additional design considerations: According to DNV GL (2018), a design check is needed to confirm that the foundation can be installed at the site successfully. Another important design check is the need to avoid the possibility of any damage or buckling of the foundation structure during installation (e.g. Randolph, 2018). API (2011) provides a guideline to determine the minimum allowed dimensions of the pile wall thickness compared to the pile diameter to avoid buckling. The effect of scouring is an additional design consideration and in many wind turbines preventative scour protection is installed; scour and scour protection influence the structure's natural frequency and lateral response (Mayall, 2019).



(a) Illustration of pile's backbone curve showing pile's initial stiffness and ultimate capacity.



(b) Illustration of pile's cyclic response showing the terminal deformation – ratcheting (Byrne et al., 2021b).



(c) Typical excitation ranges of an OWT (Kellahave et al., 2015).

Figure 1.6: Illustration of different foundation design considerations for an OWT.

1.2.1 Former design procedures

Early OWT monopile foundations employed a simplified 1D analysis design approach known as the *p-y* method based on the Winkler assumption (API, 2011; DNV GL, 2018). In the *p-y* method illustrated in Figure 1.7, the pile is modelled as an embedded beam and the lateral soil response is represented by *p-y* curves, which are non-linear relationships between the distributed lateral load acting on the pile and the local lateral pile displacement. The Winkler assumption considers that the lateral reaction p of each spring depends solely on the local spring lateral displacement y, not allowing coupling between the spring elements. This method was initially developed for slender flexible piles that were employed in the oil and gas industry in the 1950-1960s. The method is based on measured data from field tests on slender flexible piles. Matlock (1970) and Reese et al. (1975) developed *p-y* curves for soft and stiff clay respectively. Cox et al. (1974) and O'Neill and Murchison (1983) developed *p-y* curves for sand profiles. The *p-y* method provides unreliable design for the shorter rigid monopiles that are used for OWT foundations which employ a relatively low pile slenderness ratio, L/D (Doherty and Gavin, 2012; Byrne et al., 2017; Duroska et al., 2021).

Cyclic loading is accounted for in the offshore design codes guidelines (API, 2011; DNV GL, 2018) by simply using coefficients to reduce the lateral reaction capacity p of the static

p-y curves to evaluate the maximum pile displacement or rotation developed due to cyclic loading. This approach does not account for the continuing degradation observed under cyclic loading. It assumes that steady state conditions are reached after a limited number of load cycles; it is straightforward to implement as a modification to the static *p-y* curves. However, it does not provide an effective means of calculating the pile cyclic response since it does not account for the soil damping, loading rate, change in soil stiffness or the accumulated pile rotation (ratcheting) with cycle number.

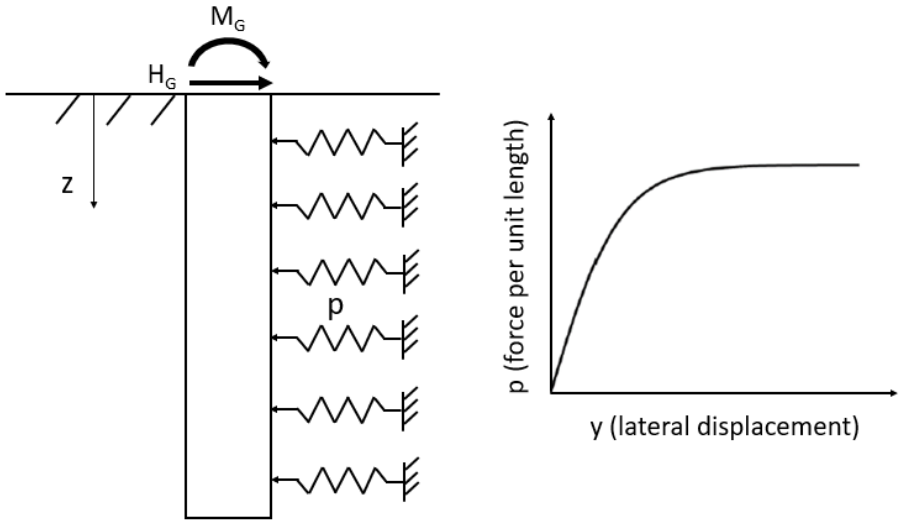


Figure 1.7: Framework for the *p-y* model.

1.2.2 The PISA design model

Oxford University participated in the PISA (Pile Soil Analysis) project (2013-2018). The PISA project was a research joint industry project involving detailed site investigation and laboratory testing (Zdravkovic et al., 2020a), medium scale field testing (Byrne et al., 2020a; McAdam et al., 2020) and numerical analyses (Zdravkovic et al., 2020b; Taborda et al., 2020) aiming to improve the design of low slenderness ratio, L/D, laterally loaded monopiles that are employed as a foundation option for OWTs. The PISA project established the 'PISA design model' for calculating the monotonic behaviour of laterally loaded monopiles (Byrne et al., 2020b; Burd et al., 2020a; Burd et al., 2020b). The PISA model

incorporates four soil reaction components; this concept predicts accurately the response of low slenderness ratio (L/D) relatively rigid piles within a pre-defined calibration space. The PISA soil reaction components (Figure 1.8(a)) are: (a) distributed lateral loads; (b) vertical shear tractions resulting into a distributed moment; (c) a horizontal force at the pile base; and (d) a moment at the pile base.

The PISA design model is implemented in the 1D finite element (FE) framework illustrated in Figure 1.8(b). The pile is simulated by a line mesh of beam elements employing Timoshenko theory. The soil is modeled using a separate set of finite elements, employing displacement interpolation functions identical to those utilised for the pile elements. Each embedded pile element is paired with a corresponding soil element, connected to its two nodes. These procedures ensure consistency in displacement and rotation between the pile and the soil along each embedded pile element's length. Each soil reaction component in the model correlates with local lateral displacement or rotation, following a Winkler approach, through a calibrated parametric function known as a 'soil reaction curve'. These curves serve a similar purpose to conventional *p-y* curves in a *p-y* analysis but are extended in the PISA model to incorporate additional soil reaction components (Figure 1.8(b)). The PISA design model is capable of conducting rapid calculations to determine the pile load-displacement response and the distribution of bending moment and shear force acting in the pile. In the 1D model a monotonic lateral load H, corresponding to the environmental load, is applied to the pile at a height h above the ground level. This applies a lateral load ($H_G = H$) and a moment load ($M_G = hH$) acting at ground level. Vertical loads applied to the foundation are neglected.

The soil reaction components used in the PISA design model are calibrated by 3D finite element analyses (FEA) where advanced constitutive models are employed to capture accurately the soil behaviour. The soil reaction data are normalised using the dimensionless

groups in Byrne et al. (2020b) and Burd et al. (2020a). The normalisation process allows the soil reaction curves to be computed within the 1D FE model for local values of the soil strength and stiffness properties. A four-parameter conic function is used to represent each soil reaction curve relating the normalised forms of pile displacement/rotation to the local traction/moment (Figure 1.8(c)). Simple functions are fitted to represent the depth variation of the 4-conic function parameters along the pile for each of the soil reaction components. The PISA design models incorporate a further optimisation stage; this was performed to improve the accuracy of the 1D model's Winkler framework, by attempting to minimize an error metric which incorporates the accuracy at small and ultimate displacements equally weighted for each pile (Byrne et al., 2020b).

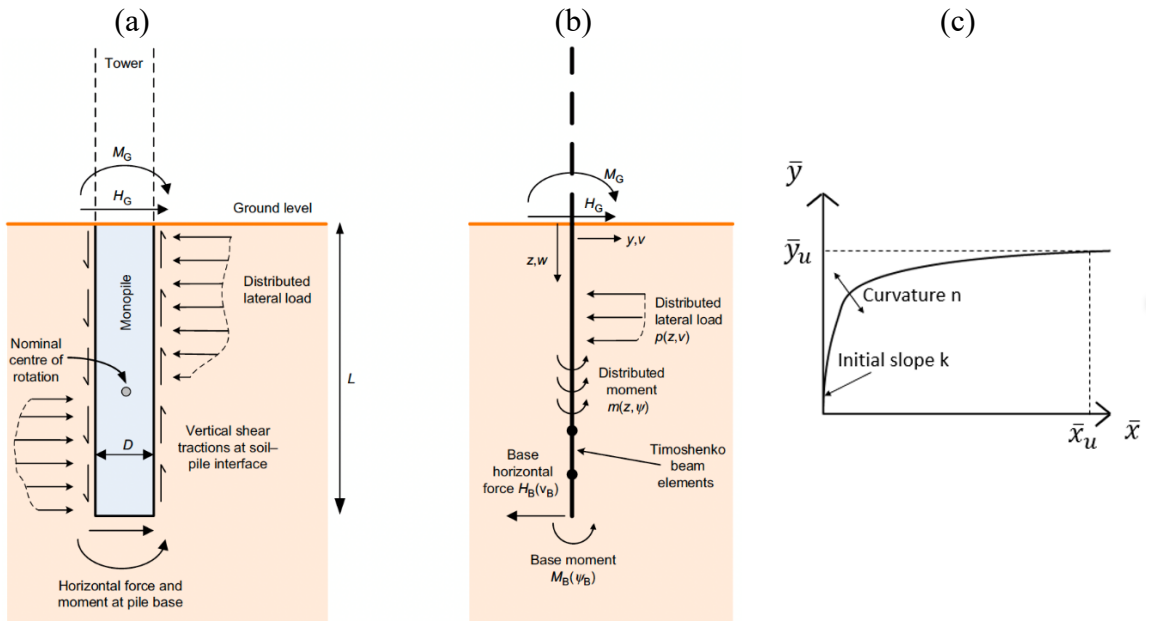


Figure 1.8: PISA design model taken from Byrne et al. (2020b). (a) Soil reaction components acting on the pile. (b) 1D finite element model framework for the embedded monopile, (c) 4-parameter conic function used to represent the PISA soil reaction curves.

One of the outputs of the PISA project was the general Dunkirk sand model (GDSM) presented in Burd et al. (2020a) which is a 1D model developed for sands with arbitrary values of relative density in the range of $45\% \leq D_R \leq 90\%$. Furthermore, during the PISA project three additional 1D models were developed suitable for clay soil profiles:

- the PISA Cowden till model (Byrne et al., 2020b) developed for a stiff overconsolidated low-plasticity glacial clay till site.
- the PISA Bothkennar clay model (AWG 2018) developed for a normally consolidated soft deposit as a result of post-glacial sedimentation.
- the PISA London clay model (AWG 2018) developed for a stiff overconsolidated high-plasticity clay.

PLAXIS used the PISA design model to develop PLAXIS Monopile designer Tool (MoDeTo) (Panagoulias et al., 2018). PLAXIS MoDeTo is a software tool which focuses on the design of OWT monopiles under monotonic lateral loading. The software incorporates both PLAXIS 3D software with the PISA design model allowing the designer to calibrate PISA rule-based design 1D models for new offshore soil profiles.

1.2.3 PICASO project

PICASO (Pile Cyclic Analysis: Oxford and Ørsted) is an ongoing project being conducted by Oxford University and Ørsted. The PICASO research is intended to build on the PISA project methods and extend them to cyclic loading conditions. An overview of the PICASO project as presented in Byrne et al. (2021b) is illustrated in Figure 1.9. The HARM (Hyperplastic Accelerated Ratcheting Model) (Houlsby et al., 2017; Abadie et al., 2019) – to be developed in WP2 and WP3 as illustrated in Figure 1.9 – is a model being developed at Oxford University that is capable of capturing cyclic loading at soil element, 1D model and complete pile level. Detailed laboratory and field tests (Richards et al., 2020; Buckley et al., 2021; Martin et al., 2023; Qiu et al., 2023) were conducted during WP4 in Figure 1.9 to investigate the ratcheting phenomenon, stiffness evolution and rate effects related to the cyclic loading of monopiles. The results were used to calibrate and validate HARM (Beuckelaers et al., 2021) to establish a consistent framework for the calibration and use of HARM for cyclic loading design considerations (Abadie et al., 2023). Separately, during the

PICASO project a new effective stress constitutive model HySand within the framework of hyperplasticity has been developed (Simonin et al., 2022); HySand can be used for cyclic calculations on sand soil profiles.

In addition, to develop new cyclic models, PICASO is also concerned with expanding the scope and accuracy of the PISA design model through the WP1 task presented in Figure 1.9. A new data-driven method is introduced in Suryasentana et al. (2022); Kamas et al. (2023) which employs the basic 1D finite element framework of the PISA design model but with Gaussian process regression (GPR) machine learning (ML) models (Williams and Rasmussen, 2006) to formulate the soil reaction curve data. A new parametrisation of the soil reaction curves employing a piecewise parameter spline function was introduced in Suryasentana et al. (2021) which represents more accurately the soil reaction curves and is capable of capturing softening effects within the soil reaction curves.

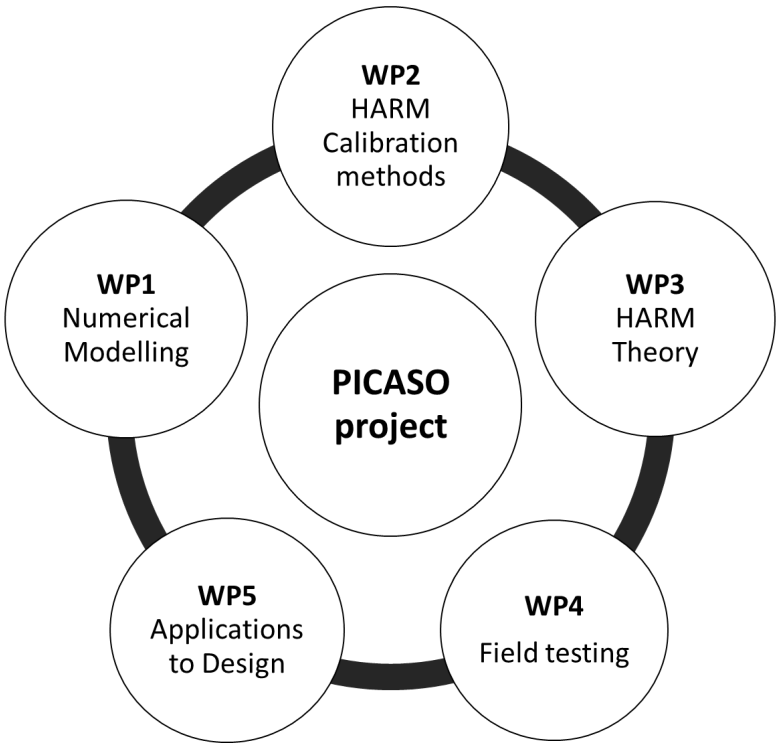


Figure 1.9: Overview of the PICASO project structure (presented in Byrne et al., (2021b)).

1.2.4 Finite element modelling

A bespoke 3D finite element analysis (FEA), where the soil is modelled as continuum, is acknowledged as the most accurate way to predict realistically the response of a laterally loaded monopile (DNV GL 2018). Although 3D FEA offers high accuracy, it has the disadvantages of being highly time consuming and more complex and demanding than a 1D model. The accuracy of FEA predictions depends on the selection of an appropriate constitutive model to capture the soil behaviour. Using an advanced constitutive model may require lengthy computational times and typically require a large number of calibrated parameters potentially requiring non-routine laboratory testing. In contrast, selecting a simple constitutive model might result in crude predictions of the pile behaviour. Usually, designers prefer to compromise by using constitutive models which only capture the most salient features of the soil behaviour and have been shown to predict a realistic foundation response. A crucial consideration influencing the accuracy of the validity of the predictions of a 3D FEA is the calibration procedure of the constitutive model parameters.

1.3 Motivation

The PISA project established a novel approach for developing calibrated 1D models for monotonic laterally loaded monopiles, referred to as the PISA design model. The PISA design model can provide rapid calculations of the foundation's small displacement, ultimate response and bending moment induced in the pile. The application of the PISA framework can lead up to a 30% of reduction in the foundation's steel material, resulting in significant cost savings compared to the traditional *p-y* method used in early OWFs (Manceau et al., 2019). However, the PISA design model has a number of shortcomings, as outlined in Byrne et al. (2020b), Burd et al. (2020a) and Burd et al. (2020b) which reduce the fidelity of the calibrated 1D models such as: (i) the conic function cannot be accurately fitted to the soil reaction curves data, (ii) the use of simple functions to represent the complex variation of

the conic function parameters is inadequate, (iii) the Winkler approach employed in the PISA 1D FE framework neglects various pile-soil interaction mechanisms developed in a 3D FE model, (iv) the calibrated 1D models do not perform well for layered soils that involve neighbouring soil layers with significantly different soil strength and stiffness characteristics, (v) the PISA method is not applicable for certain soil profiles – for example to generate 1D models for chalk soils or a general clay 1D model for clay soils with different undrained shear strength (s_u) profiles.

The work described in this thesis is linked to the numerical modelling work package of the PICASO project (WP1) (Figure 1.9). It aims to further develop a new approach to generate data-driven 1D design models, building on the methodology proposed by Suryasentana et al. (2021). This data-driven method employs Gaussian process regression (GPR) machine learning models to formulate the soil reaction curve data that are required to drive the PISA 1D finite element (FE) framework. The data-driven method is based on the hypothesis that if there is good fit between the soil reaction curves predicted by a 1D model (using the Winkler approach) and the numerical soil reaction curves extracted from a 3D FEA, there is implicitly a good match between the 1D model and the 3D model predictions of pile monotonic lateral behaviour. This new approach comes as an alternative to the PISA design model and aims to:

- **Simplify the generation of calibrated 1D models**: The machine learning model is solely data-driven and scales well with the data, eliminating the need for manual adjustment of 'rules' (e.g. optimisation stages in PISA design model).
- **Capture complex depth variations with high fidelity**: The advanced GPR machine learning model accurately captures the complex depth variation of soil reaction curve parameters along the pile.

- **Enhance the generality of the PISA design model**: This method improves the applicability of the PISA design model to different soil types (e.g., general clay 1D models, layered soil 1D models, chalk 1D models).
- **Accurately represent soil reaction curves**: The piecewise spline function provides a new parameterisation that represents any soil reaction curve shape or non-linearity, including potential softening effects.
- **Estimate prediction uncertainty**: The probabilistic nature of the GPR model offers an initial estimation of the uncertainty in the predictions of the soil reaction curves parameters. Typically, when the pile geometry or design soil profile significantly deviates from previous calibration data, the GPR model will show a high standard deviation, indicating that a 3D FEA is required.

Overall, this thesis contributes to a broader effort by Oxford University to enhance the outcomes of the PISA project. It establishes a framework for generating highly accurate data-driven 1D design models for various soil configurations found in the North Sea. These 1D models can provide rapid calculations of the OWT monopile's backbone curve and the pile's induced bending moments, assisting both ULS and FLS design procedures.

1.4 Terminology

This section explains the terminology used for the calibrated 1D models presented in this thesis. Figure 1.10 outlines the key components, soil profiles examined and the calibrated 1D models generated using the data-driven method. The data-driven 1D design models shown in Figure 1.10 were developed during this DPhil project and are discussed in the following Chapters of this thesis. Figure 1.11 lists the key components, soil profiles studied and the calibrated 1D models developed using the PISA design model. These aspects were developed during the PISA project and are employed in this thesis to validate the newly developed data-driven method.

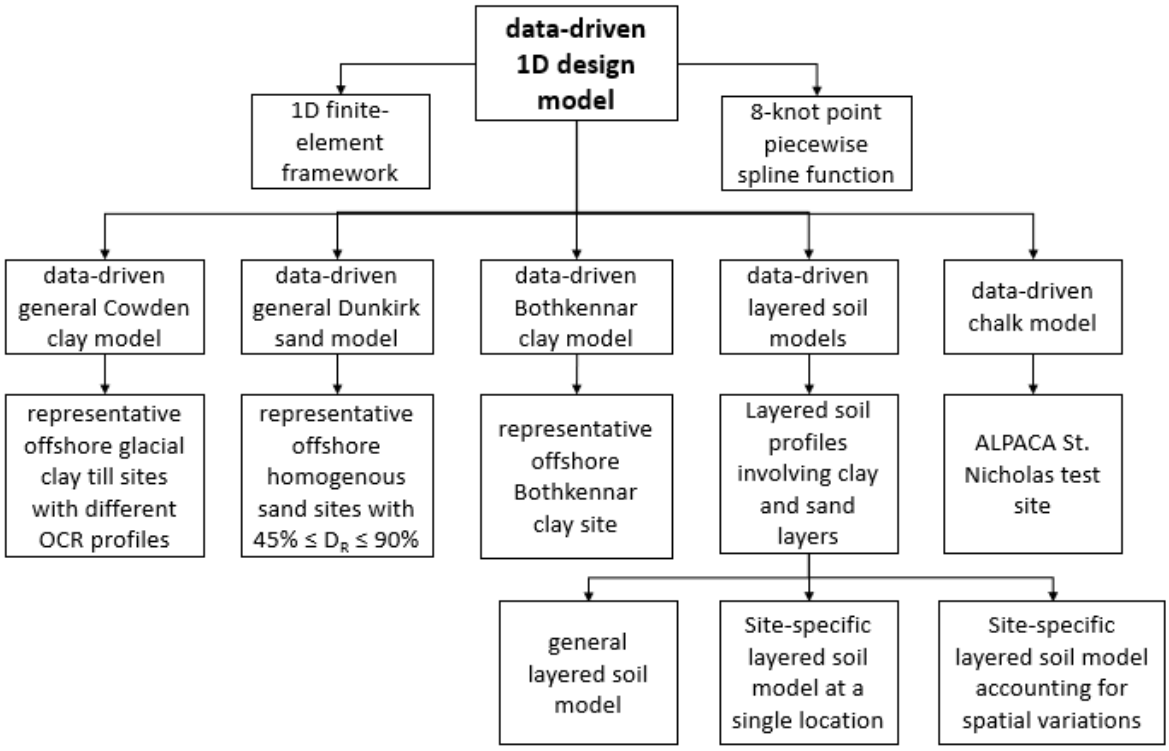


Figure 1.10: Terminology employed for the data-driven design 1D model. The model is implemented in the MATLAB code 'OxPile' which was developed during the PISA project. The OxPile software is updated to incorporate this data-driven model in the current thesis.

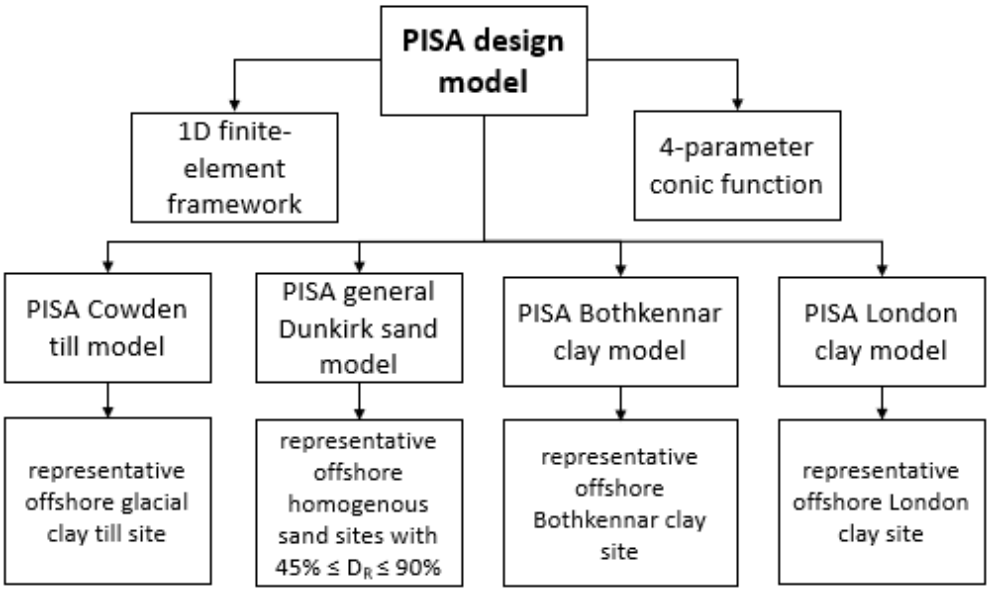


Figure 1.11: Terminology employed for the PISA design model. The model was implemented in the MATLAB code 'OxPile' during the PISA project.

1.5 Chapter summary

The structure of the thesis is outlined below:

Chapter 1 (*Introduction*) describes the design requirements for monopile foundations. It provides a brief overview of current design practices, including the outcomes of the PISA project, the aims of the current PICASO project, and the capabilities of using 3D Finite Element Analysis (FEA) for laterally loaded piles. This chapter sets out the problem definition for this thesis.

Chapter 2 (*Literature Review*) reviews the research landscape relating to the different tools that are used in this thesis such as: the monotonic behaviour of laterally loaded monopiles; examples of 1D and 3D analysis procedures; surrogate modelling theory; geotechnical studies that utilised machine learning models; the theory behind two machine learning regression models used in this thesis.

Chapter 3 *(Finite Element Analyses of Laterally Loaded Monopiles)* summarises the numerical analyses of monopile foundations conducted in this research project for piles embedded in clay or sand soil profiles.

Chapter 4 (*Data-Driven Method for Laterally Loaded Piles*) presents the framework of the data-driven method developed during the PICASO project as an alternative to the PISA design model. Design examples for homogenous sand and clay soils are provided.

Chapter 5 (*Development of Generalised Data-Driven Models for Clay and Sand Sites*) outlines a methodology for creating general 1D models applicable to both clay and sand sites. This chapter demonstrates the approach using two case studies: (i) the development of a new data-driven general clay model capable of predicting the monotonic lateral response of piles embedded in representative offshore glacial clay till sites with varying undrained shear strength and small shear stiffness profiles, and (ii) the generation of a data-driven

general sand model for predicting the monotonic lateral response of piles embedded in sand profiles with different values of relative density.

Chapter 6 (*Data-Driven Method for Layered Soils*) presents different procedures for applying the data-driven method to layered soils. The advantages and disadvantages of each procedure are shown through realistic design examples.

Chapter 7 (*Layered Soil Modelling Approach for Practical Design*) presents a design example for generating data-driven layered soil models to predict the monotonic backbone curve of monopiles over a range of OWT locations across an OWF. Realistic soil conditions were employed for the different wind turbine locations, in collaboration with Ørsted, the industrial partner which funds this research project).

Chapter 8 (*Data-Driven Method for Laterally Loaded Piles in Chalk*) presents a novel framework to generate data-driven calibrated 1D models for laterally loaded piles embedded in chalk. A chalk model is developed and validated using the ALPACA, ALPACA + and Wind Support field test piles.

Chapter 9 (*Conclusions*) summarises the outcome of this research.

Chapter 2

2 Literature Review

2.1 Behaviour of the monopile under monotonic response

A rigorous estimation of the monotonic lateral behaviour of the monopile used as a foundation for an OWT is required during the design procedure. The calculations focus on determining: (i) the initial stiffness of the monopile at small displacements, which is used during the FLS design calculations; (ii) the ultimate capacity of the monopile at a large predefined displacement, which is required during the ULS design calculations.

Small displacement response

The small strain shear stiffness, G_o, is the soil parameter governing the pile's initial stiffness. As the values of the G_o profile increase, so does the pile's initial stiffness. Moreover, the initial stiffness is influenced by the slenderness ratio L/D of the pile. Typically, stiffness increases with a higher slenderness ratio. However, according to AWG (2016), there is a threshold L/D ratio beyond which the initial stiffness remains constant. At small displacement range, the deeper parts of the pile are not mobilised for longer piles (Figure 2.1(c)) and therefore the initial stiffness depends solely on the flexure of the pile at shallow depths. For rigid monopiles as those employed for the OWTs, the initial stiffness depends both on the flexure and rotation of the entire pile. Several research studies such as AWG

(2016), Achmus et al. (2019) focused on the precise calculation of the pile's initial stiffness. Generally, the pile's initial stiffness can be accurately calculated through a fast 3D FEA where the soil is modelled using an elastic or elastoplastic model which can accurately reproduce the depth variation of the G_o profile across the pile's embedment length. The uncertainty embedded in the measured values of the G_o profile through the laboratory and field test experiments should be also accounted for in the determination of the pile's initial stiffness.

Ultimate response

Figure 2.1(c) illustrates the deflection mode of both rigid and flexible laterally loaded piles at ultimate state condition. For the long slender piles that are employed in the oil and gas industry with $L/D > 20$, the pile ultimate capacity is developed by the formation of a plastic hinge in the pile (Randolph and Gouvernec, 2011). The ultimate soil resistance can, in this case be estimated by integrating the lateral resistance from the soil above the plastic hinge. The monopiles employed for the OWTs have much lower pile slenderness ratios ($L/D \leq 6$) and deform as a rigid body. The bottom of the pile tends to 'kick out' in the opposite direction of the loading (Murff and Hamilton, 1993) and the pile rotates about a point which is located at an approximate pile depth of $z = 2/3L$ (where L is the embedment length of the pile). The ultimate resistance is determined by the total soil resistance and can be estimated by integrating the lateral resistance from the soil along the pile embedment length.

In the API (2011) model (*p-y* method) proposed for high L/D flexible piles, the ultimate resistance at shallow depths follows a wedge-type failure mechanism. This mechanism shifts to a 'flow around' failure in the horizontal plane at deeper depths where the soil is assumed to flow horizontally around the pile, as illustrated in Figure 2.1(a). Rigid monopiles with low L/D values, exhibit a deep wedge failure mode in the upper section of the pile, followed by

rotational soil flow in the vertical plane below this section (see Figure 2.1(b)), instead of in the horizontal plane assumed in the API (2011) model. For cohesive soils such as clays a gap can form behind the pile, reducing the soil resistance. In the PISA field tests, gap formation was observed in piles embedded in both the stiff, overconsolidated Cowden till test site (Byrne et al., 2020a) and the dense Dunkirk sand test site (McAdam et al., 2020). However, while this phenomenon was attributed to the unsaturated conditions of the top 5.4m of the sand in the latter case, it is not typical for offshore monopiles embedded in fully saturated sand conditions. Gap formation significantly influences the lateral behaviour of a monopile (AWG 2016) and needs to be accurately accounted for during the design process.

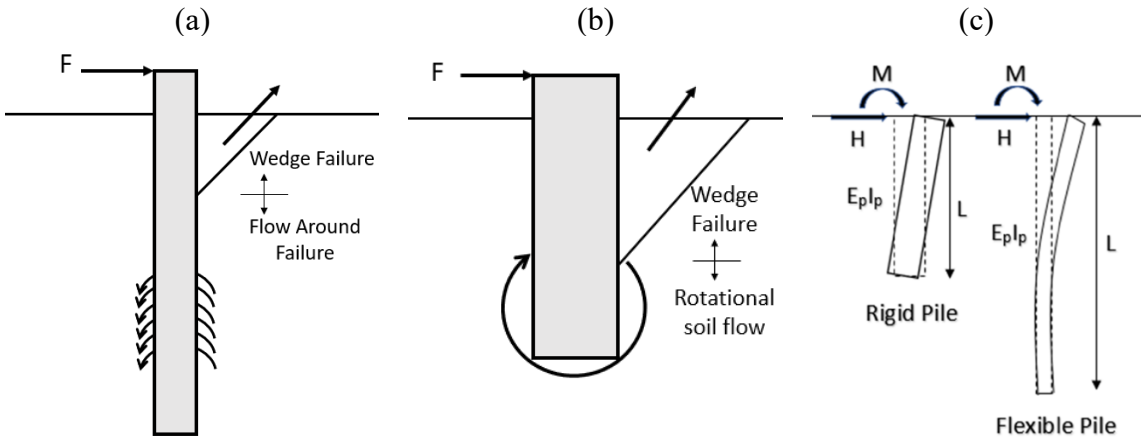


Figure 2.1: Lateral loaded pile failure mechanism (Randolph and Gouvernec, 2011) for: (a) slender piles with high L/D, (b) rigid piles with low L/D (typically employed for OWTs). (c) Deflection mode of rigid and slender laterally loaded piles.

2.2 Analysis procedures

2.2.1 1D Level

p-y curves for homogenous soil profiles

The *p-y* method is a classical approach initially developed for oil and gas structures as described in Section 1.2.1. A review of developments of the *p-y* method for application to the design of OWT monopiles is provided below. Zhang and Andersen (2017) proposed a *p-*

y framework that is generally applicable for long slender monopiles where the flow around failure mechanism dominates; the method assumes plane strain conditions around the pile (which are more applicable to slender piles) and estimates the *p-y* curves by scaling the stress-strain curves calculated from direct simple shear tests. Suryasentana and Lehane (2014) proposed a formulation of *p-y* curves suitable for large diameter piles embedded in sand using the cone penetration test (CPT) tip resistance q_c as the measure of soil strength to determine the relationship between the distributed lateral load p acting on the pile and the local lateral pile displacement y; they implemented 3D FEA using the Hardening Soil model on laterally loaded monopiles to determine the pile's behaviour and derived the q_c profile by simulating the cavity expansion problem. Thieken et al. (2015) performed a series of 3D FEA for laterally loaded piles embedded in sandy profiles using the HS-Small model to derive a *p-y* formulation which utilizes a 'stretching' factor to modify the ultimate bedding resistance depending on the pile's deflection line and the distance of the *p-y* curve's depth to the pile tip. Generally, as it was shown in the PISA project and presented in Section 1.2.2, the low L/D ratio piles that are employed for the OWT, require the use of four different soil reaction components. Therefore, the majority of the *p-y* methods presented above are more suitable for slender monopiles and will most likely underpredict the OWT pile's capacity.

After the PISA project, there have been some new 1D analysis models which incorporate the additional soil reaction components apart from the distributed lateral soil reaction. Siegl et al. (2020) proposed a 1D analysis model where four soil reaction components are extracted from FEA (consistent with the PISA framework). The extracted soil reaction curves are normalised and a quadratic/hyperbolic framework (originally developed for non-linear earthquake analyses) is adopted to reproduce the soil reaction curves. Wang et al. (2020) and Wang et al. (2023) proposed a 1D analysis model for computing the lateral behaviour of monopiles embedded in soft clay and sand profiles, with varied pile L/D slenderness ratios.

In their approach, the lateral soil resistance above the rotation point is described by a *p-y* model; the total resistance below the rotation point is calculated with a moment spring (*M-θ*) located at the rotation point including the shear and moment reactions at pile base.

p-y curves for layered soils

OWT foundations are typically embedded in layered soils. The soil layering will have a significant influence on the *p-y* curves. For example, Yang and Jeremic (2005) conducted an FEA study on two- and three-layered profiles showing that soil layering can have a significant influence on the monopile lateral behaviour and as a result on the *p-y* curves. McGann et al. (2011) used the results of 3D FEA on slender piles embedded in layered soils to evaluate the ultimate resistance and initial stiffness of the conventional *p-y* curves. They concluded that due to the interactions between neighbouring layers the ultimate resistance calculated by the *p-y* curves does not agree well with the FEA results, whereas a better agreement was found for the initial stiffness. Reese et al. (1981) conducted small-scale laboratory tests and a series of field tests employing slender piles with high L/D slenderness ratio embedded in layered soils and concluded that there is a relatively good match between the measured deflections in the tests and the deflections calculated using the *p-y* curves developed for homogeneous soil at small loads.

Georgiadis (1983) developed a simplified method to modify the standard *p-y* calculation to account for the interactions between neighbouring layers. This method is well-suited for slender piles, as it does not account for the effect of layering on additional soil reaction components, such as distributed moment reaction, base horizontal force, and base moment, which are generated on low L/D slenderness ratio piles commonly used for offshore wind turbines. The approach is applicable to layered soils containing both clay and sand layers. Georgiadis (1983) approach is based on defining the 'equivalent depth' of every soil layer

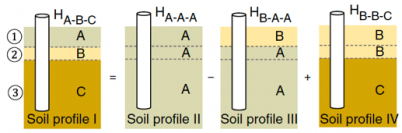
beneath the surface layer. This is important because *p-y* curves are constructed based on soil depth. To find the equivalent depth (z_2) of the layer beneath the surface, the integrals of ultimate soil resistance over depth for both layers are equated, with z_1 denoting the depth of the surface layer. Subsequently, the values of the pile's ultimate lateral resistance, p_{ult}, are derived from the provided soil properties.

$$F_1 = \int_0^{z_1} p_{ult1} dz \quad (2.1) \qquad F_2 = \int_0^{z_2} p_{ult2} dz \quad (2.2)$$

The *p-y* curves for the second layer are calculated, starting from z_2 (the actual depth, is at z_1) and proceeding downward until another layer is encountered in terms of actual depth. This method can be extended to the subsequent layers that the pile penetrates.

Hu et al. (2022) introduced a series of design equations to determine the lateral capacity and ground level load-rotation response (for $\theta_G \leq 2°$) of monopiles embedded in uniform, two-layer, and three-layer sand profiles (Figure 2.2) under fully saturated conditions and monotonic loading. These equations are based on a large number of finite element analysis (FEA) using the two-surface-plasticity constitutive model developed by Loukidis and Salgado (2009) to represent sand behaviour. One of the biggest advantages of this method is that engineers don't require access to a finite element code with a sophisticated constitutive model incorporated into it to achieve comparable results. The design equations can be employed over a wide range of parameters, including pile diameters $1 \leq D(m) \leq 10$, pile slenderness ratios $2 \leq L/D \leq 20$, pile wall thicknesses ($D/100 \leq t_w \leq D/50$), load eccentricities ($15 \leq h(m) \leq 30$) and various sand types such as Ottawa sand and Toyoura sand with sand relative densities (D_R) spanning from 40% to 95%. It's important to note that this method is applicable only for layered soil profiles with sand layers and is recommended for design cases falling within the calibration space. Hu et al. (2022) proposed approach is summarised in the Table 2.1 adopted from Hu et al. (2022).

Table 2.1: Methodology adopted in Hu et al. (2022) to calculate the H-θ_G response of piles embedded in sandy layered soil profiles.

Calculation stages	Equations	Comments
Uniform Sand		
(1): Calculate L_{crit} for $\theta_G = 0.5°$ and $\theta_G = 1°$	$L_{crit}/{I_p}^{0.25}{}_{\|0.5°} = -12.5\frac{D_R}{100} - 3\ln\left(\frac{{I_p}^{0.25}}{L_R}\right) + 37.9$ $L_{crit}/{I_p}^{0.25}{}_{\|1°} = -12\frac{D_R}{100} - 3.3\ln\left(\frac{{I_p}^{0.25}}{L_R}\right) + 40.3$ (2.3)	L_{crit}: critical length $L_R = 1m$ θ_G: pile rotation at ground level $I_p = \pi(D^4 - {D_i}^4)/64$ D_i: inner diameter
(2) Calculate H_{crit} for $\theta_G = 0.5°$ and $\theta_G = 1°$	$\frac{H_{crit}}{p_A {L_R}^2} = a(D_R, h)\left(\frac{I_p}{{L_R}^4}\right)^{0.72} + b$ $a(D_R, h) = \left(a_1\frac{h}{L_R} + a_2\right)\frac{D_R}{100\%} + a_3\frac{h}{L_R} + a_4$ (2.4)	H_{crit}: Pile lateral capacity when $L \geq L_{crit}$ a_1, a_1, a_1, a_1, b are calculated from Table 2.2 for $\theta_G = 0.5°$ and $\theta_G = 1°$
(3) Calculate H/H_{crit} for $\theta_G = 0.5°$ and $\theta_G = 1$	$\frac{H}{H_{crit}} = \begin{cases}\left[\frac{1}{2}\sin\left(\pi\frac{L}{L_{crit}} - \frac{\pi}{2}\right) + 0.5\right]^{\beta} & if\ L < L_{crit} \\ 1 & if\ L \geq L_{crit}\end{cases}$ (2.5) $\beta(D_R, I_p)_{\|0.5°} = 0.084\left(L_{crit}/{I_p}^{0.25}\right) - 1.2$ $\beta(D_R, I_p)_{\|1°} = 0.08\left(L_{crit}/{I_p}^{0.25}\right) - 1.25$	
(4) Calculate H for $\theta_G = 0.5°$ and $\theta_G = 1$	$H_{\|0.5°} = \frac{H}{H_{crit}}\|_{0.5°} \times H_{crit}\|_{0.5°}$ $H_{\|1°} = \frac{H}{H_{crit}}\|_{1°} \times H_{crit}\|_{1°}$ (2.6)	H_{crit} is calculated from Eq. 2.4 H/H_{crit} is calculated from Eq. 2.5
Two-layer sand profile		
Calculate capacity H_{A-B} for a two-layered sand profile where layer A is located over layer B (see Figure 2.2(a))	$\frac{H_{A-B}}{H_B} = \frac{H_A}{H_B} + \left(1 - \frac{H_A}{H_B}\right)\exp\left[a\left(\frac{d_1}{D}\right)^{1.5}\right]$ $a = \begin{cases}-0.23 & if\ L/D \geq 5 \\ 1/(-1.07\frac{L}{D} + 1) & if\ L/D < 5\end{cases}$ (2.7)	d_1: thickness of top layer (see Figure 2.2(a)) H_A: lateral capacity of pile embedded in homogenous sand A H_B: lateral capacity of pile embedded in homogenous sand B
Three-layer sand profile		
Calculate capacity H_{A-B-C} for a three-layered sand profile (see Figure 2.2(c))	$H_{A-B-C} = H_{A-A-A} - H_{B-A-A} + H_{B-B-C}$ (2.8) 	H_{A-A-A}: Pile lateral capacity for homogenous sand A H_{B-A-A} and H_{B-B-C}: Pile lateral capacities for two-layer sand profile (calculated from Eq. 2.7)
Load-Rotation curve		
Calculate load-rotation (H-θ_G) response at ground level	$H = \theta/(k + \eta\theta)$ $\begin{cases}\eta = 2/H\|_{1°} - 1/H\|_{0.5°} \\ k = 1/H\|_{0.5°} - 1/H\|_{1°}\end{cases}$ (2.9)	$H_{0.5°}$ and $H_{1°}$ are calculated by Eqs. 2.3 to 2.8 Pile lateral capacity for any rotation $\theta < 2°$ is calculated using Eq. 2.9

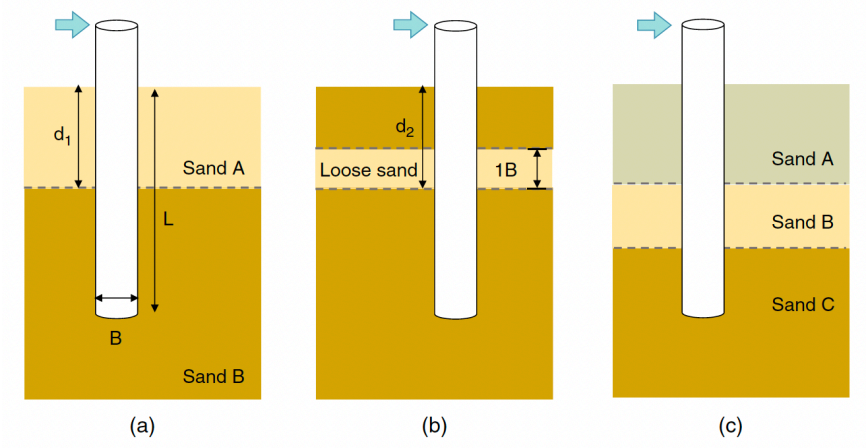


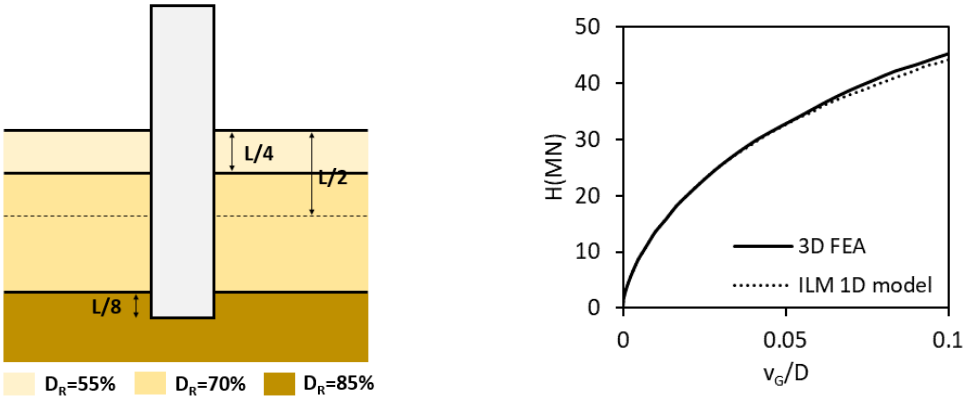
Figure 2.2: The layered soil profiles considered in Hu et al. (2022) including: (a) a two-layer sand profile; (b) a three-layer sand profile comprising a loose sand layer embedded within two dense sand layers; (c) a more generalised three-layer sand profile.

Table 2.2: Values of coefficients to calculate H_{crit} (taken by Hu et al. (2022)).

Case	a_1	a_2	a_3	a_4	b
Ottawa sand $\theta_G = 0.5°$	-0.95	49	-0.34	25.8	-0.01
Ottawa sand $\theta_G = 1°$	-2.05	107	-0.35	33.5	0.07
Toyoura sand $\theta_G = 0.5°$	-0.85	32	-0.35	33.5	-0.15
Toyoura sand $\theta_G = 1°$	-1.30	65	-0.42	43.0	-0.02

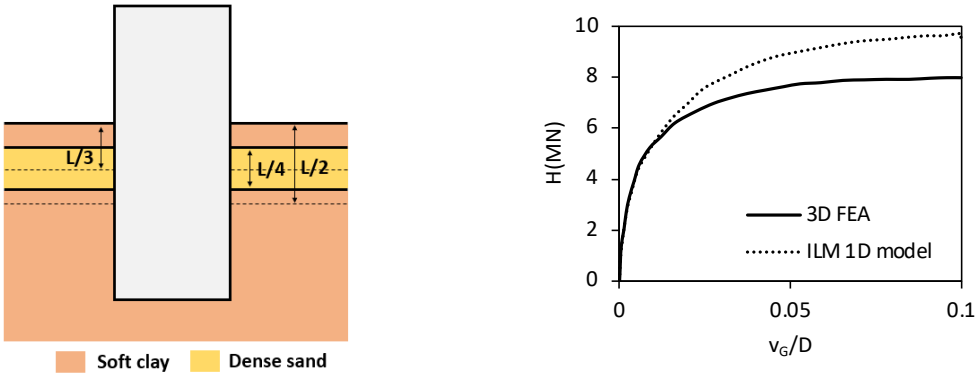
Burd et al. (2020b) extended the applicability of the PISA design model to layered soils by using the hypothesis that the soil reaction curves extracted from FEA conducted with homogeneous soil profiles can be used for calculations in layered soil profiles. Consequently, 1D models calibrated for homogeneous soil profiles in prior studies (Byrne et al., 2020b; Burd et al., 2020a; AWG, 2018) were employed by Burd et al. (2020b) for calculations involving layered soil profiles. In the current thesis the approach adopted by Burd et al. (2020b) is referred to as the 'independent layer method' (ILM). Through extensive testing with various layered soil configurations provided by PISA industrial partners, Burd et al. (2020b) demonstrated the effectiveness of the ILM in soil conditions approaching realistic offshore wind turbine (OWT) locations. The ILM is straightforward to implement and accurately predicted the pile's ultimate response for layered configurations where interactions between neighbouring layers were not dominant such as: (i) layered soils

involving multiple sand soil layers of different relative density, (ii) multi-layered soils involving many thin layers (less than 5m), (iii) two-layered soils where the base layer is located near the tip of the pile, (iv) layered soils with dense clay layers adjacent to loose sand layers. Additionally, it was found that the ILM provides reliable results for the pile's small displacement response regardless of the different soil layering.



(a) (b)

Figure 2.3: ILM applied for a three-layered soil with sand layers of different relative density where the interactions between neighbouring layers are not dominant (Burd et al., 2020b). (a) Layered soil configuration, (b) Comparison between the ultimate responses from the 1D and 3D models.



(a) (b)

Figure 2.4: ILM applied for neighbouring soil profiles with different strength characteristics (Burd et al., 2020b). (a) Layered soil configuration, (b) Comparison between the ultimate responses from the 1D and 3D models.

Figure 2.3 illustrates the high accuracy of the ILM when applied to a pile with a diameter $D = 8.75m$, slenderness ratio $L/D = 4$ and load eccentricity ratio $h/D = 10$, embedded in a three-layered soil consisting of sand layers with different relative density (refer to Case

E1B in Burd et al. (2020b)). However, Burd et al. (2020b) demonstrated that the ILM did not perform well in cases where the neighbouring soil profiles had severely different strength characteristics (e.g. soft clay with dense sand). An example from Burd et al. (2020b), where the ILM does not perform well is presented in Figure 2.4; a pile with a diameter $D = 10m$, slenderness ratio $L/D = 2$ and eccentricity load ratio $h/D = 5$ is embedded in a layered soil comprising soft Bothkennar clay and dense Dunkirk sand layers (refer to Case F3 in Burd et al. (2020b)). For instances where the ILM significantly overpredicts the pile's ultimate response, Burd et al. (2020b) suggest the possibility of developing rules to degrade the strength of the stronger material. However, specific advice on this matter is not provided in the paper.

p-y curves for piles embedded in chalk

Low to medium density chalk material is found in many offshore wind farm sites. Currently there is no established design method for analysing the behaviour of lateral loaded piles driven in chalk. Chalk is a challenging material for engineers; field tests of laterally loaded piles undertaken in low to medium density chalk (Buckley et al., 2020b; Ciavaglia et al., 2017b) showed that the installation procedure severely de-structures the chalk around the pile generating a de-structured remoulded chalk annulus around the pile and inducing fractures around that area to the intact chalk material. There are very few proposed *p-y* curve formulations when coming to chalk profiles.

Muir Wood et al. (2015) proposed a *p-y* curve formulation which was utilised in the design of the monopiles deployed in the Westermost Rough offshore wind farm. The upper layer of the chalk profile, classified as extremely weak chalk due to full weathering, prompted the adoption of a *p-y* curve formulation typically used for clay soils. The layer below was classified as a low to medium density chalk. For that layer it was assumed that

the pile driving created a remoulded chalk annulus around the pile. Hence, the overall deflection y of the pile under an applied load was assumed to be represented by two *p-y* springs in series ($y = y_r + y_s$) generated by: (i) the remoulded chalk y_r and (ii) the structured (fractured or intact) chalk y_s. The deformation of the remoulded chalk layer was calculated using the one-dimensional consolidation settlement theory assuming that the compression of the layer was analogous to the compression of a finite clay layer beneath a spread footing.

The ultimate lateral resistance p_{ult} of the low to medium density chalk layer was calculated according to the *p-y* curves method proposed by Reese (1997) for weak rocks which takes into consideration the quality and strength of the rock material through the measured rock quality designation (RQD) and unconfined compression strength (UCS). A tri-linear model was employed for the *p-y* curves to capture both the pre and post-yield behaviour of the chalk prior to the ultimate state p_{ult}. The initial stiffness of the *p-y* curve was set equal to the chalk's Young's modulus values based on the results of drained triaxial testing. The initial yield load p_y was defined by pressumeter test results and available data of plate load tests published by Matthews and Clayton (2004). The post yield stiffness of the *p-y* curve was calculated based on data from drained triaxial tests and published plate load test data by Matthews and Clayton (2004). The *p-y* model proposed by Muir Wood et al. (2015) is the first 1D analysis model that incorporates to some extent (only for the pile's deflection calculation) the effect of the de-structured chalk zone. During the ALPACA and ALPACA + projects (Buckley et al., 2020b) it was found that the de-structured chalk zone has a significant influence on the pile's lateral response. Therefore, this effect needs to be accounted for in more detail in the developed *p-y* curves formulation.

2.2.2 3D Level

In recent decades numerous papers have been published that describe FEA applications for predicting the pile monotonic lateral behaviour. In the PISA project, bespoke soil constitutive models were employed. The constitutive model parameters were calibrated by advanced laboratory testing and validated by medium scale field tests. The PISA FEA provided highly accurate results of the pile's monotonic lateral response. However, they were also highly computational demanding. For the clay soils, an enhanced Modified Cam Clay model (Zdravkovic et al., 2020b) was adopted, comprising a non-linear stress-strain pre-yield behavior for the elastic shear modulus (Taborda et al., 2016), a non-linear Hvorslev surface (Tsiampousi et al., 2013) and a generalised shape for the yield and plastic potential surfaces in the deviatoric plane (Van Eekelen, 1980) to account for strength anisotropy. For the sand profiles (Taborda et al., 2020) a state parameter-based boundary surface plasticity model (Taborda et al., 2014) was used capable of reproducing the effects of stress level and void ratio on the sand behavior using a single set of model parameters.

Grammatikopoulou et al. (2021) successfully simulated the PISA field test results for the Dunkirk test site using a Mohr-Coulomb softening model coupled with the ICL non-linear elastic model (Taborda et al., 2016), achieving relatively accurate results. Castillo-Fuentes et al. (2023) employed critical state Norsand constitutive model to compute with good accuracy the PISA field test results for the Dunkirk test site. Hu et al. (2022) validated the use of the two-surface-plasticity constitutive model developed by Loukidis and Salgado (2009) with the PISA Dunkirk sand field test results and then employed the constitutive model for offshore calculations. Whyte et al. (2020) used a multi-surface Parallel Iwan approach, the 'PIMS' model, to reproduce the PISA Cowden test site field test piles response; this model also facilitates cyclic loading calculations. Page et al. (2018) proposed a macro-element model approach where the response of the pile and the surrounding soil is

reduced to a force-displacement relation at the seabed. This model follows the multi-surface plasticity framework and is calibrated by FEA results; it can be used for highly computationally time demanding cyclic loading calculations. This macro-model calibrated by an FEA employing the NGI-ADP constitutive model, reproduces the PISA Cowden field test pile monotonic response with good accuracy. Liu and Kaynia (2022a) used the SANISAND-MSu model (Liu et al., 2020c) to compute the laterally loaded monopile's response embedded in a dense sand profile; both the drained and undrained monotonic response as well as the undrained cyclic response of the pile were evaluated.

Several additional studies employed more straightforward constitutive models such as Mohr Coulomb (Abdel-Rahman and Achmus, 2005), NGI-ADP (Zhang et al., 2016; Minga and Burd, 2019a), Hardening Soil (Murphy et al., 2019) and HS-Small model (Igoe and Jalilvand, 2020; Minga and Burd, 2019b) to calculate the monotonic lateral behaviour of monopiles embedded in homogenous clay or sand soil profiles. These studies have demonstrated that such simpler models can provide robust analyses of pile behaviour without requiring an extensive number of laboratory tests or including many model parameters. However, these simpler models may have limitations in reproducing certain aspects of soil response. For example, Mohr-Coulomb, being a very simple linear elastic perfectly plastic model, includes only a limited number of features that soil exhibits in reality. NGI-ADP cannot accurately capture the non-linear shear stiffness degradation observed in clay soils, while the HS-Small model fails to capture the post-peak softening observed in dense sand. Additionally, simpler models may lack generality, especially when large-scale validation data is unavailable. For instance, HS-Small requires different parameter sets for varying void ratios. Finally, it is important to note that simpler models are not suitable for cyclic calculations due to their inherent limitations.

2.3 Artificial intelligence applications in geotechnical engineering

In recent years there has been a notable shift towards utilising surrogate models and machine learning (ML) methods to ease and assist the design process for geotechnical studies. The data-driven 1D design models developed in this thesis employ surrogate models to compute the soil reaction curves parameters for different pile geometries and soil profile properties. A surrogate model functions as a rapid approximation model designed to replace a high-order and computationally intensive model, such as a 3D finite element simulation. Its objective is to learn a mapping rule, denoted as $y = f(\boldsymbol{x})$, which includes the physics that convert the vector $\boldsymbol{x}$ (inputs) into a scalar output y (Forrester et al., 2008). In machine learning terminology, this process aligns with supervised learning. The efficacy of a surrogate model depends on several factors, including selecting a ML model capable of accurately capturing the complexity of the observed data and exhibiting robust generalisation properties for unseen data. Furthermore, it necessitates identifying the most critical input features or observations that have a significant influence on the function f, thereby facilitating the training of the ML model.

The ML models can be described as mathematical models that formulate the target values as a function f of the training inputs (the inputs are commonly referred as features). The prediction made with the ML model may differ from the actual value. This error (ε) between the predictions of the ML model and the actual data is minimised during the training procedure by optimising the model coefficients; this is typically implemented through minimising a loss function. The ML model has also a set of hyperparameters which are parameters that control the behaviour of the model. The hyperparameters can be tuned to improve the model performance. There are two groups of supervised ML models: (i) the regression problems, where the value of a continuous output is predicted based on a number of training inputs, (ii) the classification problems, where the ML model is trained to predict

categorical targets. In recent years, there has been a notable increase in the digitalisation of various geotechnical data obtained from laboratory tests, field monitoring, and computationally expensive numerical analyses. Consequently, there has been a surge in studies exploring the application of machine learning methods to facilitate the fitting of complex mathematical models to the available geotechnical data.

The studies by Sheil et al. (2020), Jong et al. (2021), Zhang et al. (2021a), Ebid (2021), Baghbani et al. (2022), Stuyts and Suryasentana (2023), and Phoon et al. (2023) collectively provide a comprehensive review of machine learning applications in geotechnical studies, emphasising the significance of dataset quantity and quality, feature selection, and machine learning model selection. Together, they underscore the potential of machine learning methods to address various geotechnical challenges effectively. In this thesis, supervised regression machine learning methods are employed for developing surrogate models. A summary of key insights from studies (see Table 2.3) employing these methods on geotechnical data is provided below:

- ML regression models, including artificial neural networks (ANNs), support vector machines (SVMs), multivariate adaptive regression splines (MARS), Gaussian Process Regression (GPR), random forest regression (RFR), and eXtreme Gradient Boosting (XGB), have been successfully applied in geotechnical applications.
- ML methods offer an alternative to computationally expensive numerical analyses, time-consuming laboratory tests, and expensive field tests, particularly for handling complex datasets efficiently.
- The studies underscore the importance of having sufficient and high-quality datasets for training machine learning models, as the model's validity is directly linked to the data quality.

- Proper selection of training inputs (features) is crucial for the accuracy and effectiveness of machine learning models. Features should be relevant to the problem and may require pre-processing or scaling before training. For example, ANN and SVM models require feature scaling.
- Data is typically divided into training and testing sets, with various evaluation metrics such as: coefficient of determination (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE)) used to assess the ML model's accuracy.
- Candidate ML models should be evaluated thoroughly before final selection, considering factors such as dataset complexity, size, and number of features. For example, certain machine learning models such as SVRs work well with small datasets and less well for large datasets whereas others such as ANN work better with having larger datasets.
- Hyperparameter optimisation techniques such as Grid Search, Random Search Bayesian optimisation, Particle Swarm Optimization (PSO), Symbiotic Organism Search (SOS) are commonly employed to find the best model configurations. Grid Search is the least efficient method in a high dimensional space.
- Trained ML models may have limitations in providing predictions beyond specified input variable ranges, necessitating their use within the calibration space. However, some ML models such as GPR can be employed successfully for extrapolation (Suryasentana et al., 2023).
- Most of the studies employing ANN architectures often utilise feedforward neural networks with Rectified Linear Unit (*ReLU*) activation function and *adam* solver for weight optimisation.

Table 2.3: Summary of previous studies on employing supervised regression machine learning methods on geotechnical applications.

Authors	Geotechnical application	Proposed ML method	Number of training inputs	Hyperparameters optimisation	Number of samples in dataset
Suryasentana et al. (2022)	Lateral loaded offshore monopiles	GPR (Matern ν=5/2)	3	Maximising log marginal likelihood	330
Ferrero et al. (2023)	Wall movements of a deep braced excavation in clay	ANN (*ReLu*)	4	Grid Search method	300
Yin et al. (2023)	Load-deflection response of caisson foundations in clay	ANN (*ReLu*)	5	Grid Search method	192600
Momeni et al. (2020)	Pile bearing capacity	GPR (Matern ν=5/2)	5	Maximising log marginal likelihood	296
Tejada and Antolin (2021)	Mechanical behaviour of granular materials	ANN (*ReLu*)	10	1 hidden layer with 32 neurons	92738
Nguyen et al. (2023a)	Bearing capacity of conical foundation in clay	MARS	4	Grid Search method	1296
Diaz and Spagnoli et al. (2024)	Calculation of compression index c_c	RFR	4	PSO method	1008
Nguyen et al. (2024)	Load-displacement behaviour of piles	ANN	24	SOS method	1650
Pirrone and Taborda (2023)	Load-displacement curve of laterally loaded monopile	SVM (*RBF* kernel)	3	Grid Search method	30
Kardani et al. (2020)	Bearing capacity of piles	XGB	5	PSO method	59
Liu et al. (2021)	Horizontal displacements of soil nail walls	ANN	8	Grid Search method	376
Samui (2008)	Settlement prediction of shallow foundations	SVM (*RBF* kernel)	5	Grid Search method	174
Zhang et al. (2021b)	Lateral deformation of a diaphragm wall in clays	XGB	7	Not mentioned	1778
Nguyen et al. (2023b)	Uplift resistance of circular anchors	MARS	3	Grid Search method	720
Bui et al. (2019)	Slope stability	ANN	4	Not mentioned	630

2.4 Machine learning models

2.4.1 Gaussian Process Regression

Gaussian process regression (GPR) is a machine learning model that assumes a Gaussian probability distribution over a range of possible random functions which is constantly updated by the observed data according to the Bayes rule. A detailed description of the GPR machine learning models is provided by Rasmussen and Williams (2006). GPR models provide a built-in variance measure which quantifies the prediction uncertainty (i.e. from the model itself and the noise). This uncertainty becomes large in regions away from the training dataset and shrinks close to the observations. GPR is a non-parametric model meaning that its expressiveness (ability to reproduce complex functions) increases with the amount of data. Non-parametric models have potentially unlimited parameters and the model's complexity grows with the data. However, non-parametric models have a portability issue since after the training of the model they require access to the training dataset to make future predictions.

A GPR model is fully characterized by its mean function $m(x)$ and its covariance function (also referred as kernel) $k(\boldsymbol{x}, \boldsymbol{x}')$:

$$f(x) \sim GP(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \tag{2.10}$$

$$m(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})], \;\; k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x}') - m(\boldsymbol{x}'))] \tag{2.11}$$

where $(\boldsymbol{x}, \boldsymbol{x}')$ are inputs in either the training or the test sets (where $\boldsymbol{x}$ refers to training and $\boldsymbol{x}'$ to test data). In Equations (2.10), (2.11) the operator $\mathbb{E}$ denotes the expectation of $f(x)$. For any finite set of inputs $\{\boldsymbol{x}(1), \dots, \boldsymbol{x}(N)\}$, the probability distribution of the outputs $p(f(\boldsymbol{x}(1)), \dots, f(x(N)))$ is multivariate Gaussian, with a mean m and a covariance matrix K (where $m_i = m(\boldsymbol{x}_{(i)})$ and $K_{ij} = k(\boldsymbol{x}_{(i)}, \boldsymbol{x}_{(j)})$).

Typically, a zero mean function is employed where $m(\boldsymbol{x}) = 0$; GPR models are expressive enough to model the true mean well using just the kernel function. Hence, it is unnecessary to employ complex mean functions. The kernel function is one of the most important parts to define the whole GPR; it specifies the similarity between the inputs based on the assumption that if the $(\boldsymbol{x}, \boldsymbol{x}')$ inputs are similar then the $y(\boldsymbol{x})$, $y(\boldsymbol{x}')$ predicted outputs will be strongly correlated. These two GPR aspects (mean, kernel functions) are the prior properties of a Bayesian model, and they need to be defined before the use of the GPR. By providing this prior information, the GPR is restricted to possible candidate function f in respect to the final prediction.

Two commonly employed kernel functions are the Squared Exponential kernel and the Matérn kernel. The Squared Exponential kernel is defined by:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp\left(-\frac{r(\boldsymbol{x}, \boldsymbol{x}')^2}{2l^2}\right) \tag{2.12}$$

where $r(\boldsymbol{x}, \boldsymbol{x}') = \|x - x'\|$; σ_f is the signal variance determining the average distance of the data-generating function from its mean function; l is the characteristic length-scale indicating how close should two inputs to be correlated.

The Matérn kernel is defined by:

$$k_{Matern}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \times \frac{2^{1-\nu}}{\Gamma(\nu)} \times \left(\frac{\sqrt{2\nu} r(\boldsymbol{x}, \boldsymbol{x}')}{l}\right)^{\nu} \times K_{\nu} \times \left(\frac{\sqrt{2\nu} r(\boldsymbol{x}, \boldsymbol{x}')}{l}\right) \tag{2.13}$$

where $\Gamma(\nu)$ is the gamma function; K_{ν} is the modified Bessel function of the second kind; ν is a positive parameter of the covariance that controls the smoothness of the function. Overall, the σ_f, l are the hyperparameters of the kernel. The hyperparameters term in the machine learning framework is used to describe the parameters whose values control the learning process and determine the values of the ML model parameters.

In a GPR model the objective function is assumed to be a sample from a Gaussian distribution such as:

$$y_i = f(\boldsymbol{x}_i) + \varepsilon_i \tag{2.14}$$

where $f(x_i)$ is a random function and ε_i are independent and identically distributed Gaussian noise random variables with mean zero and variance ($p(\varepsilon_i) = N(\varepsilon_i|0, \sigma^2)$.

For a given dataset $(\boldsymbol{X}, \boldsymbol{y})$ the probability density of the observations given the parameters is calculated by $p(\boldsymbol{y}|\boldsymbol{X}, f) = N(\boldsymbol{y}|f, \sigma^2\boldsymbol{I})$. In the Bayesian theory a prior is specified over the parameters, expressing the general belief about the prior parameters before the observations are taken into consideration. The prior is assumed over f using a Gaussian distribution with mean zero and covariance matrix $\boldsymbol{K}$ such that $p(f|\boldsymbol{X}) = N(f|0, \boldsymbol{K})$ providing that for any finite set of inputs $\{\boldsymbol{x}(1), \ldots, \boldsymbol{x}(N)\}$:

$$\begin{bmatrix} f(\boldsymbol{x}_1) \\ \ldots \\ f(\boldsymbol{x}_N) \end{bmatrix} \sim N\left(0, \begin{bmatrix} k(\boldsymbol{x}_{(1)}, \boldsymbol{x}_{(1)}) & \ldots & k(\boldsymbol{x}_{(1)}, \boldsymbol{x}_{(N)}) \\ \ldots & \ldots & \ldots \\ k(\boldsymbol{x}_{(N)}, \boldsymbol{x}_{(1)}) & \ldots & k(\boldsymbol{x}_{(N)}, \boldsymbol{x}_{(N)}) \end{bmatrix}\right) \tag{2.15}$$

Next the joint distribution of the training outputs $\boldsymbol{y}$ and the output y_* corresponding to a test input x_* is defined as:

$$\begin{bmatrix} \boldsymbol{y} \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \boldsymbol{K}_{xx} + \sigma^2 I & \boldsymbol{k}_* \\ \boldsymbol{k}_*^T & k_{**} \end{bmatrix}\right) \tag{2.16}$$

where $\boldsymbol{K}_{xx}$ is the covariance matrix computed for the training inputs $\boldsymbol{X}$; $\boldsymbol{k}_*$ is the covariance vector computed between the test input $\boldsymbol{x}_*$ and the training dataset $\boldsymbol{X}$ and k_{**} is the covariance between $\boldsymbol{x}_*$ and itself. For a given test input $\boldsymbol{x}_*$, the output y_* is predicted by the predictive distribution computed by:

$$p(y_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}) = N(y_*|\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \tag{2.17}$$

The posterior distribution is fully described by its mean $\boldsymbol{\mu}_*$ and variance $\boldsymbol{\Sigma}_*$. These are evaluated from:

$$\boldsymbol{\mu}_* = \boldsymbol{k}_*^T[\boldsymbol{K}_{xx} + \sigma^2\boldsymbol{I}]^{-1}\boldsymbol{y} \tag{2.18}$$

$$\boldsymbol{\Sigma}_* = \boldsymbol{k}_{**} - \boldsymbol{k}_*^T[\boldsymbol{K}_{xx} + \sigma^2\boldsymbol{I}]^{-1}\boldsymbol{k}_* \tag{2.19}$$

The mean prediction is a linear combination of observations $\boldsymbol{y}$. If w (where w are the weights) is set as $w = \boldsymbol{k}_*^T[\boldsymbol{K}_{xx} + \sigma^2\boldsymbol{I}]^{-1}$ then the posterior mean is equal to $\boldsymbol{\mu}_* = \boldsymbol{w}^T\boldsymbol{y}$; this indicates that for a given point the GPR predicts an output by computing a weighted average of the neighbouring outputs. The outputs of Equation (2.17) depend on the values of hyperparameters of the kernel. The best values of these hyperparameters θ are usually estimated by maximising the marginal likelihood (or typically its log). The marginal likelihood is the integral of the likelihood multiplied with the prior defined by:

$$p(\boldsymbol{y}|\boldsymbol{X}) = p(\boldsymbol{y}|f, \boldsymbol{X})p(f|\boldsymbol{X})df \tag{2.20}$$

The log marginal likelihood is evaluated by:

$$log(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}) = -\frac{1}{2}\boldsymbol{y}^T[\boldsymbol{K}_{xx} + \sigma^2\boldsymbol{I}]^{-1}\boldsymbol{y} - \frac{1}{2}log|\boldsymbol{K}_{xx} + \sigma^2\boldsymbol{I}| - \frac{N}{2}log2\pi \tag{2.21}$$

The first term assesses the degree to which the kernel-optimised hyperparameters align with the dataset. The second term serves as a regularisation factor, addressing noise within the dataset. Lastly, the third term functions as a normalisation factor, ensuring that the resulting distribution or objective function is appropriately scaled.

2.4.2 Sparse Gaussian process regression

Gaussian process regression models are impractical for large datasets due to the computational and memory cost of inverting and storing the covariance matrix $\boldsymbol{K}_{xx}$. The training of GPR models requires $O(N^3)$ time and $O(N^2)$ memory space where N is the number of datapoints. Hence, sparse approximations of the GPR models have been widely used to address this issue. There are well-defined sparse GPR models such as: (i) the Nyström approximation (Williams and Seeger 2000) which attempts to decompose the $\boldsymbol{K}_{xx}$ covariance matrix, (ii) the Fully Independent Training Conditional (FITC) (Quinonero-Candela and Rasmussen, 2005; Snelson and Ghahramani, 2006) which approximates the prior distribution, (iii) the Variational Free Energy (VFE) (Titsias 2009; Hensman et al., 2013; Bauer et al., 2016; Liu et al., 2020b; Burt et al., 2020 and Rossi et al., 2021) which approximates the posterior distribution. The sparse GPR models introduce a small number of supporting variables also referred as inducing variables, $\boldsymbol{f}_m$, evaluated at inputs $\boldsymbol{X}_M$. These inducing variables are assumed to describe well the function $\boldsymbol{f}$ and are employed as approximation to $\boldsymbol{f}$ and $\boldsymbol{X}$ to define an approximate posterior $q(\boldsymbol{f}_*)$ where:

$$q(\boldsymbol{f}_*) = \int p(\boldsymbol{f}_*|\boldsymbol{f}_m)\phi(\boldsymbol{f}_m)\, d\boldsymbol{f}_m \tag{2.22}$$

where $\phi(\boldsymbol{f}_m)$ approximates the intractable $p(\boldsymbol{f}_m|\boldsymbol{y})$.

$$\phi(\boldsymbol{f}_m) = N(\boldsymbol{f}_m|\boldsymbol{\mu}_M, \boldsymbol{A}_M) \tag{2.23}$$

The objective of the sparse GPR model is to determine the optimal values for the mean, $\boldsymbol{\mu}_M$, and covariance matrix, $\boldsymbol{A}_M$. Additionally, the effectiveness of the $\phi(\boldsymbol{f}_m)$ function depends on the location of the inducing inputs $\boldsymbol{X}_M$, hence, optimising their positions is also important. The mean and covariance matrix of the Gaussian approximate posterior $q(\boldsymbol{f}_*)$ are defined in terms of $\boldsymbol{\mu}_M$, $\boldsymbol{A}_M$ and $\boldsymbol{X}_M$.

$$q(\boldsymbol{f}_*) = \left(\boldsymbol{f}_* \middle| \boldsymbol{\mu}_*^q, \boldsymbol{\Sigma}_*^q\right) \tag{2.24}$$

$$\boldsymbol{\mu}_*^q = \boldsymbol{K}_{*m}\boldsymbol{K}_{mm}^{-1}\boldsymbol{\mu}_m \tag{2.25}$$

$$\boldsymbol{\Sigma}_*^q = \boldsymbol{K}_{**} - \boldsymbol{K}_{*m}\boldsymbol{K}_{mm}^{-1}\boldsymbol{K}_{m*} + \boldsymbol{K}_{*m}\boldsymbol{K}_{mm}^{-1}\boldsymbol{A}_M\boldsymbol{K}_{mm}^{-1}\boldsymbol{K}_{m*} \tag{2.26}$$

Here, $\boldsymbol{K}_{mm}$ represents the covariance matrix computed for the inducing inputs, denoted by $\boldsymbol{m}$, $\boldsymbol{k}_{*m}$ is the covariance vector computed between the test input x_* and the inducing inputs $\boldsymbol{m}$. Additionally, $\boldsymbol{K}_{*m}$ represents the transpose of $\boldsymbol{K}_{m*}$. For a single test input, this approximate posterior can be computed in $O(NM^2)$ time complexity, where N is the number of training data points and M is the number of inducing inputs. This computational approach significantly reduces the training time compared to a typical Gaussian Process Regression (GPR), particularly after optimal values for $\boldsymbol{\mu}_M$, $\boldsymbol{A}_M$ and $\boldsymbol{X}_M$ have been determined.

In this thesis, it was observed that the Variational Free Energy (VFE) sparse GPR model yielded the highest accuracy predictions for the soil reaction curves dataset. The VFE approach utilises variational methods to optimise $\boldsymbol{\mu}_M$, $\boldsymbol{A}_M$ and $\boldsymbol{X}_M$ by minimising the Kullback-Leibler (KL) divergence between the approximate posterior $q(\boldsymbol{f}_*)$ and the exact posterior $p(f|\boldsymbol{y})$ over training latent variables f. This minimisation of KL divergence is equivalent to maximising a lower bound $L(\boldsymbol{\mu}_M, \boldsymbol{A}_M, \boldsymbol{X}_M)$ on the true log marginal likelihood $logp(\boldsymbol{y})$. Analytically solving for $\boldsymbol{\mu}_M$ and $\boldsymbol{A}_M$ enables the optimisation of this lower bound. The resulting lower bound after optimisation becomes a function of X_M:

$$L(\boldsymbol{X}_M) = logN\left(\boldsymbol{y}\middle|0, {\sigma_y}^2\boldsymbol{I} + \boldsymbol{Q}_{nn}\right) - \frac{1}{2{\sigma_y}^2}Tr(\boldsymbol{K}_{nn} - \boldsymbol{Q}_{nn}) \tag{2.27}$$

In Equation (2.27), $\boldsymbol{Q}_{nn} = \boldsymbol{K}_{nm}\boldsymbol{K}_{mm}^{-1}\boldsymbol{K}_{mn}$ where $\boldsymbol{K}_{nn}$ is the covariance matrix computed for the training inputs and $\boldsymbol{K}_{nm}$=${\boldsymbol{K}_{mn}}^T$. This equation can be optimised in $O(NM^2)$ time complexity and is used to optimise the inducing inputs jointly with the kernel hyperparameters. The first term in Equation (2.27) serves as an approximation to the log-

likelihood term, while the second term acts as a regularisation term resulting from the variational approach, which helps prevent overfitting. The optimisation process aims to find the optimal values of $\boldsymbol{X}_M$ to optimise $\boldsymbol{\mu}_M$ and $\boldsymbol{A}_M$ (which are parameters of $\phi(\boldsymbol{f}_m)$.

$$\boldsymbol{\mu}_M = \frac{1}{{\sigma_y}^2}\boldsymbol{K}_{mm}\boldsymbol{\Sigma}\boldsymbol{K}_{mn}\boldsymbol{y} \tag{2.28}$$

$$\boldsymbol{A}_M = \boldsymbol{K}_{mm}\boldsymbol{\Sigma}\boldsymbol{K}_{mm} \tag{2.29}$$

where $\boldsymbol{\Sigma} = (\boldsymbol{K}_{mm} + {\sigma_y}^2\boldsymbol{K}_{mn}\boldsymbol{K}_{nm})^{-1}$. $\boldsymbol{\mu}_M$ and $\boldsymbol{A}_M$ are then substituted in Equations (2.25) and (2.26) to compute the approximate posterior $q(\boldsymbol{f}_*)$ at new inputs $\boldsymbol{x}_*$.

2.4.3 Artificial neural network

A neural network is a widely used regression technique in artificial intelligence renowned for its expressive capabilities, demonstrated by its ability to approximate nearly any function (Cybenko, 1989). It belongs to the domain of machine learning, specifically deep learning, and functions through interconnected nodes or neurons arranged in layered structures resembling the human brain. While individual neurons are relatively simple, the collective strength of numerous neurons in a network enables the execution of complex tasks. A neural network model comprises two key elements: (i) the network architecture, which dictates the number of hidden layers, the quantity of neurons, and the connectivity patterns among them, and (ii) the network parameters, encompassing weights and biases (Goodfellow et al., 2016). Figure 2.5 illustrates an example of a simple neural network featuring one hidden layer. By incorporating multiple neurons and hidden layers, a neural network can evolve into a more complicate model capable of handling complex functions.

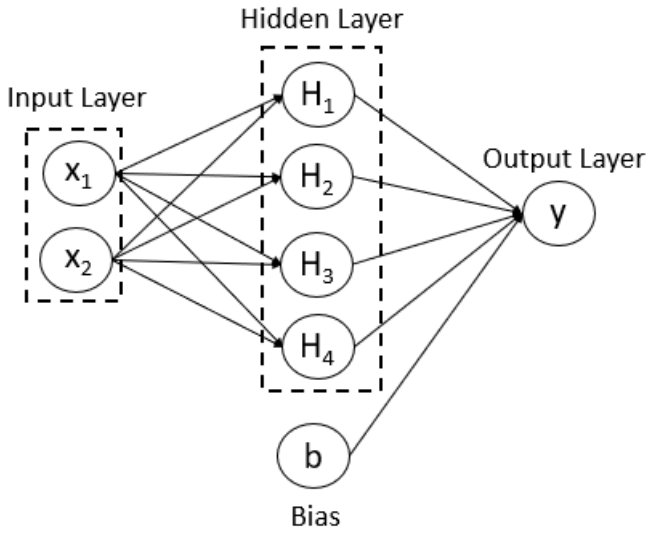


Figure 2.5: Simple neural network with two input neurons, a single hidden layer with four neurons and a single output.

A feedforward neural network (FNN), also known as a multi-layer perceptron, comprises interconnected neurons, in which information only flows forward (Heaton, 2008). An FNN establishes a mapping $\boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{\theta})$ and learns the optimal values of $\boldsymbol{\theta}$ (weights and biases) to achieve the best function approximation. Initially, an input (feature) vector $[x]$ is introduced into an input layer, where each neuron represents an input feature. Subsequently, the data moves through hidden layers before reaching the output layer. Each neuron forwards its output to every neuron in the subsequent layer, with each connection characterized by a weight, $\boldsymbol{w}$. Additionally, every neuron has a bias $\boldsymbol{b}$, which applies uniformly to all inputs linking to the neuron. These weights and biases represent a typically non-linear transformation applied to the values of the neurons being forwarded. Without employing non-linear activation functions, the neural network will simply perform linear regression. Commonly utilised activation functions (Heaton, 2008) are: (i) the identity function $g(x) = x$, (ii) hyperbolic tanh function $g(x) = tanhx$, (iii) the logistic or sigmoid function $g(x) = (1 + exp(-x)) - 1$ and (iv) the rectified linear unit (ReLu) function $g(x) = max(x, 0)$.

During the training of a neural network (Bishop, 2006), a search is carried out to optimise the weights w of the neural network, by minimising the sum of squares cost function $E(\boldsymbol{w})$ with respect to w for the training set of N training examples (Equation (2.30)), where $\{\boldsymbol{x}_n\}$ are the input vectors for $n = 1, \ldots, N$ and $\{\boldsymbol{t}_n\}$ are the target vectors.

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N} \|y(\boldsymbol{x}_n, w) - \boldsymbol{t}_n\|^2 \tag{2.30}$$

The error function $E(\boldsymbol{w})$ is not straightforward to optimise. $E(\boldsymbol{w})$ typically has a high non-linear dependence on the weights and bias parameters, generating a vanishing gradient problem of many points in the weight space. Therefore, the optimisation w is typically done using backpropagation and gradient descent (or its variants stochastic gradient descent or mini-batch gradient descent) (Bishop, 2006). Overfitting may occur reducing the generalisation properties of the neural network (usually when the magnitude of the weights become relatively large). A regularisation term (bias) can then be added to the error function to prevent overfitting (Equation (2.31)); where a is the regularisation term that controls the amount of regularisation (the relative importance of the data-dependent error $E(\boldsymbol{w})$ and the regularisation term $\tilde{E}(\boldsymbol{w})$. The simplest regulariser form is given by the sum-of-squares.

$$\tilde{E}(\boldsymbol{w}) = E(\boldsymbol{w}) + \frac{a}{2}\boldsymbol{w}^T\boldsymbol{w} \tag{2.31}$$

Chapter 3

3 Finite Element Analyses of Laterally Loaded Monopiles

3.1 Introduction

This chapter outlines the framework of the numerical analysis conducted as part of this theses. This work involved the selection and calibration of a constitutive model that is capable to reproduce the lateral monotonic behaviour of a monopile embedded in homogenous clay or sand sites. The soil configurations (homogenous or layered) examined in this thesis comprise the homogenous representative offshore clay and sand sites developed during the PISA project. The PISA representative offshore glacial clay till site representing a stiff overconsolidated low-plasticity glacial clay till soil profile (Zdravkovic et al., 2020a) and the PISA representative offshore Bothkennar clay site representing a soft clay deposit as a result of post-glacial sedimentation (AWG, 2018) are adopted as the baseline clay ground profiles. The four representative offshore homogeneous sand sites described in Burd et al. (2020a) and based on the Dunkirk test site soil profile (Zdravkovic et al., 2020a) with relative densities D_R of 45%, 60%, 75% and 90% were selected as the baseline sand ground profiles.

The aim of the work presented is to establish that the 3D FEA implemented in this thesis deliver a trustworthy database of soil reaction curves. To validate this, the computed pile lateral response over both small and large displacement range using the PLAXIS 3D software is compared with the PISA 3D FE results presented by Byrne et al. (2020b) and Burd et al. (2020a). The accuracy of the data-driven 1D design model relies on the fidelity of the calibration 3D FEA. Given that multiple 3D FEA are necessary to properly calibrate an accurate data-driven 1D design model, it is wise to optimise computational time during numerical analyses. Therefore, it is advisable to select a constitutive model that balances simplicity and accuracy. In this thesis, the selected constitutive models aim to ensure that each analysis is completed within a reasonable timeframe (approximately 30 hours) while upholding the reliability of the results.

3.2 Pile geometry calibration space

The pile geometries adopted for the validation of the finite element results are listed in Table 3.1; this employs the calibration pile geometries adopted in the development of the PISA design model (Byrne et al., 2020b; Burd et al., 2020a). These calibration piles represent typical offshore pile configurations, with slenderness ratios (L/D) ranging from 2 to 6 and pile diameters (D) varying between 5 and 10 meters. For wave-dominated loading, the normalized load eccentricity (h/D) is set to 5, while for wind-dominated loading, it is set to 15. Piles C3 and C7 showed similar lateral response to piles C1 and C6, respectively, differing only in pile wall thickness. Therefore, they are not included in this thesis.

Table 3.1: Pile geometry calibration space. D is the pile diameter; L is the embedded length; t_w is the pile wall thickness; h is the load eccentricity.

Pile	$\boldsymbol{D(m)}$	$\boldsymbol{L(m)}$	$\boldsymbol{t_w(mm)}$	$\boldsymbol{h(m)}$	$\boldsymbol{L/D}$	$\boldsymbol{t_w\ /D}$	$\boldsymbol{h/D}$
C1	10	20	91	50	2	0.0091	5
C2	10	20	91	150	2	0.0091	15
C3	10	20	125	150	2	0.0125	15
C4	10	60	91	50	6	0.0091	5
C5	10	60	91	150	6	0.0091	15
C6	5	10	45	25	2	0.009	5
C7	5	10	83	25	2	0.0166	5
C8	5	30	45	25	6	0.009	5
C9	5	30	45	75	6	0.009	15
C10	7.5	15	68	37.5	2	0.00907	5
C11	7.5	45	68	37.5	6	0.00907	5

3.3 Generation of a 3D finite element model

This section outlines the procedures for generating 3D finite element models for the laterally loaded monopiles examined in this thesis. The PLAXIS 3D software (Brinkgreve et al. 2018) was used to conduct the 3D FEA. The 3D models incorporate a plane of symmetry, simulating only half of the problem. The bottom boundary is set at a distance of at least $1.5L$ to $2L$ where L is the embedded length of the pile; the other mesh dimensions are $12D$, in the X direction and $4D$ in the Y direction, where D is the monopile diameter (Figure 3.1). These dimensions are chosen to avoid potential boundary effects. A fixed constraint is applied at the bottom boundary of the model. The side boundaries and the plane of symmetry are fixed in the normal direction only. The total number of elements is determined after a convergence study to avoid any potential effect of the element size in the final produced results. The element distribution is optimised to provide local refinement in the zones around the pile and close the pile tip to accurately represent the soil reaction curves. The lateral loads that

are applied to the monopile are simulated by applying a prescribed displacement u_x in the X direction at the required load eccentricity height h.

The soil is represented using 10-noded tetrahedral elements, as described by Brinkgreve et al. (2018). Each soil element is equipped with three degrees of freedom per node (u_x, u_y, u_z), allowing for displacement in three dimensions. Additionally, a 4-point Gaussian integration scheme is employed for accurate numerical integration within each soil element. Since the analyses represent an offshore situation, the soil is fully saturated, and the effective bulk unit weight is adopted for the ground profile. Therefore, the water table is placed at the bottom of the models. The pile was modelled using 6-noded triangular plate elements (Brinkgreve et al., 2018). The embedded part of the pile is generated using a Python script to divide the pile into 9 segments of 20° each; this provides a suitably refined and structured mesh around the pile (Panagoulias et al., 2018). The steel pile material is assumed to be linear elastic with a Young's modulus $E = 200\text{GPa}$ a Poisson's ratio $\nu = 0.30$ and a unit weight of $\gamma_p = 78.5 - \gamma'_s (kN/m^3)$ (where γ'_s is the submerged unit weight of the soil). The pile was wished in place; hence any installation effects are disregarded.

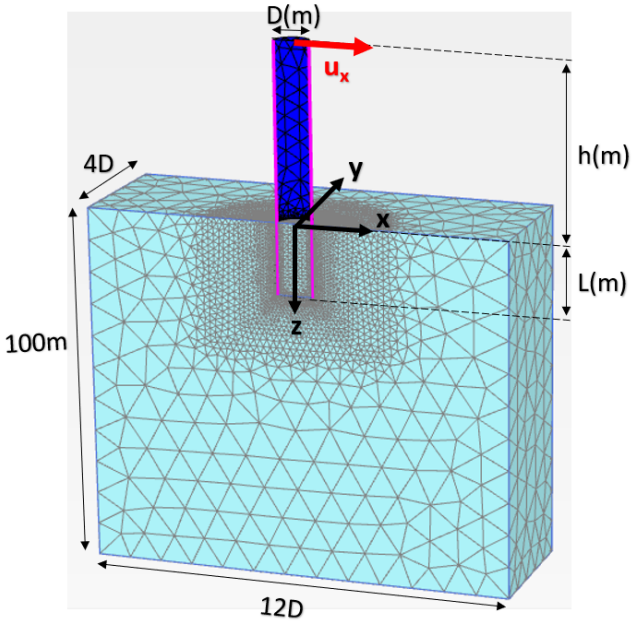


Figure 3.1: Example finite element mesh of pile C1 (Table 3.1) comprising 95200 elements.

Interface elements are used around the pile to model the pile-soil interaction and accommodate potential pile-soil slipping and gapping. Additionally, horizontal interface elements are employed at the bottom of the pile to retrieve the horizontal base force, H_B, and base moment, M_B, soil reaction curves. The interface elements are 12-node triangular zero-thickness elements. A value of $10^5 kN/m^3$ was adopted for the elastic normal (K_N) and shear stiffness (K_s) of the interface elements (Zdravkovic et al., 2020b; Taborda et al., 2020). Further details regarding the constitutive properties of the interface elements are provided in Section 3.4, covering both clay and sand profiles.

Four calculation stages are performed in each FEA. The initial stage calculates the K_o soil stresses. The second stage simulates the wished in place monopile (installation effects are disregarded). The third stage applies a small prescribed lateral displacement with a small size of load step to accurately capture the foundation initial stiffness. The final stage is used to apply a prescribed lateral displacement u_x that is sufficient to cause a ground level pile displacement of about $v_G = 0.2D$. The $0.2D$ target displacement is adopted to ensure that a clear representation of the soil reaction curves is obtained.

3.4 Soil constitutive model selection

This section explores the constitutive model selection and the calibration procedures for each clay and sand ground profile used in this thesis. Selecting an appropriate constitutive model that accurately captures the key aspects of soil behaviour is essential for computing soil reaction curves that accurately reproduce pile behaviour. The PISA representative offshore glacial clay till site (Byrne et al., 2020) and the PISA representative offshore homogeneous sand site with a relative density D_R of 75% (Burd et al., 2020a) were selected as the baseline test soil profiles for choosing an appropriate constitutive model to represent the behaviour

of the clay and sand soils examined in this thesis. Clay is modelled as undrained and sand as fully drained. The constitutive models that have been considered are listed below:

- NGI-ADP (Grimstad et al., 2012). The PIMS (Parallel Iwan Multi-Surface) model (Whyte et al., 2020) for clay.
- HS-Small (Brinkgreve et al., 2018). Norsand (Jefferies and Been, 2015) for sand.

Procedures to calibrate the models are described in the following sections. The performance of each model is compared with the PISA field test results (Byrne et al. 2020a; McAdam et al., 2020) and the PISA 3D FEA results (Byrne et al. 2020b, Burd et al. 2020a) concerned with the calibration piles presented in Table 3.1.

3.4.1 Numerical analysis of monopile embedded in the PISA Cowden till

The PISA Cowden test site is a stiff overconsolidated low-plasticity glacial clay till (Zdravkovic et al., 2020a) that is similar to clay till soils found in the North Sea. To address the influence of hydrostatic pressure on the gap formation behind the pile, which reduces the pile capacity (AWG, 2016), the submerged unit weight was included as input for the clay soil profiles in each PLAXIS 3D FEA. The submerged unit weight for the PISA Cowden till is $\gamma' = 11.38\ kN/m^3$ (Zdravkovic et al., 2020a). Two different constitutive models were tested to evaluate their performance in modeling the representative offshore glacial clay till site as presented in Byrne et al. (2020b): the NGI-ADP model (Grimstad et al., 2012) and the PIMS (Parallel Iwan Multi-Surface) model (Whyte et al., 2020). NGI-ADP is a built-in model in PLAXIS 3D. PIMS is a user defined model that was developed in Oxford University. The PLAXIS 3D Fortran DLL subroutine of PIMS was developed by Scott Whyte and provided to the author by Prof. Burd. Both constitutive models were demonstrated by Minga and Burd (2019a) and Whyte et al. (2020) to accurately represent the monotonic behavior of the PISA Cowden field test piles. In this thesis, 3D FEA were conducted to simulate the PISA field test piles embedded in the Cowden test site. The results

presented in Section 3.4.4, indicate that the NGI-ADP and PIMS models, with their final chosen sets of calibrated parameters, can accurately reproduce the field test piles response.

NGI-ADP (Grimstad et al., 2012) is a total stress model formulated for simulating clay soils for undrained conditions; it is capable of capturing the soil strength anisotropy by employing different undrained shear strength response for the different stress paths. The model receives as an input three different values of undrained shear strength values for the triaxial compression, triaxial extension, direct shear stress, which are used to construct a modified Tresca yield surface. In the PLAXIS implementation the sharp corners of the Tresca yield surface are rounded to avoid any numerical instabilities and to simplify the numerical implementation (Brinkgreve et al., 2018). The model is capable of capturing an approximate fit to the elastoplastic soil response measured from triaxial or direct shear stresses by receiving the small strain shear stiffness value G_o as an input and by including three additional input parameters γ_f^C, γ_f^E, γ_f^{DSS} (shear strains at failure), which define a bespoke hardening law.

PIMS is suitable for simulating both monotonic and cyclic loading (Whyte et al., 2020). The model employs a 'Parallel Iwan' approach (Iwan 1966) using multiple elastic-perfectly-plastic micro models employing the Von Mises yield surface. By selecting an appropriate number of micro models, the model is able to provide an accurate fit to the soil non-linear elastoplastic response measured in triaxial tests. The implementation of PIMS in Whyte et al. (2020) incorporates 12 micro models. PIMS can capture the non-linear shear stiffness degradation from small to large strains by incorporating a weight w_i for each micro model; the weight defines the contribution of each micro-model to the stiffness backbone curve. PIMS is able to capture soil strength anisotropy by incorporating a model parameter β which is referred as the 'triaxial extension factor' and controls the shape of the yield surface in the deviatoric plane.

3.4.1.1 Calibration procedure for the NGI-ADP model

The shear strain at failure parameters ($\gamma_f^C, \gamma_f^E, \gamma_f^{DSS}$) are estimated by performing simulations of triaxial compression (TXC) tests using the built-in soil testing facility in PLAXIS 3D (Figure 3.2). The parameter γ_f^C is determined by (Minga and Burd, 2019a):

$$\gamma_c^f = K_f \times 100 / \frac{G_{ur}}{s_u^A} \tag{3.1}$$

where K_f is a calibration parameter, s_u^A is the plane strain undrained shear strength ($s_u^A = s_u^{TXC}/0.99$ where s_u^{TXC} is the triaxial compression shear strength) and G_{ur} is the unloading shear modulus (where $G_{ur} = G_o$). A value of $K_f = 110$ was estimated to provide a good fit to the PISA TXC tests and the ICFEP simulated lab test responses (Zdravkovic et al., 2020b).

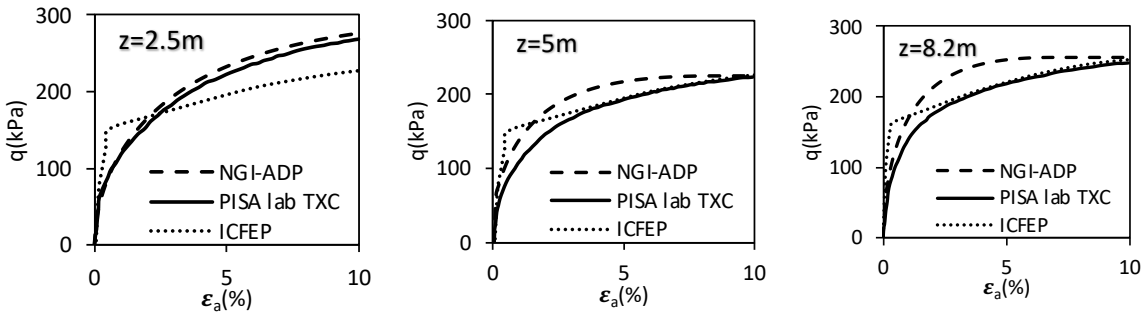


Figure 3.2: Simulation of undrained TXC tests in PLAXIS 3D using NGI-ADP with Imperial College PISA TXC tests and ICFEP simulated (Zdravkovic et al., 2020b) $q - e_a$ response for Cowden till soil samples retrieved at 2.5m, 5m, 8.2m.

The effective bulk unit weight was used as input value for the determination of the soil's initial stresses. As a result, the effective stress K_0 profile was used as input. The Poisson's ratio adopted is equal to $v = 0.495$. For the triaxial compression condition, the undrained shear strength ratio was calculated using a fixed ratio of $s_u^{TXC}/s_u^A = 0.99$ (Brinkgreve et al., 2018). For the triaxial extension conditions, the undrained shear strength s_u^{TXE} was calculated by defining the NGI-ADP model's input ratio as $s_u^{\mathrm{TXE}}/s_u^A = 0.833$ (Minga and Burd, 2019). This s_u^{TXE}/s_u^A value was derived based on the values of the average ultimate

stress ratios q/p determined for triaxial compression and extension conditions for Cowden till as presented in Zdravkovic et al. (2020a). The undrained shear strength for direct simple shear s_u^{DSS} was determined by the NGI-ADP model's input ratio $s_u^{DSS}/s_u^A = 0.911$. This was calculated by averaging the active and passive shear strengths as proposed in Brinkgreve et al. (2018). This approximation is confirmed by Liu et al. (2020a) where it was found that the torsional shear strength of the Cowden till samples fall in the middle of the passive and the active shear strength. The τ_0/s_u^A ratio is determined from Equation (3.2) (Panagoulias et al., 2018); a value of zero is adopted if the calculated value of τ_0/s_u^A is negative (in cases where $K_0 > 1$).

$$\tau_0/s_u^A = 0.5 \times (1 - K_0) \times \sigma'_{v0}/s_u^A \tag{3.2}$$

The selected parameters for NGI-ADP used in the 3D FEA for the PISA representative offshore glacial clay till site introduced in Byrne et al. (2020b) are presented in Table 3.2.

Table 3.2: NGI-ADP input parameters used for the PISA representative offshore glacial clay till site.

$z(m)$	K_o	s_u^A	s_{uinc}^A	G_{ur}/s_u^A	τ_0/s_u^A	γ_f^C	γ_f^E	γ_f^{DSS}
0-1	1.5	53.9	53.8	104.2	0	67.2	134.3	100.7
1-2	1.5	107.7	57.0	183.9	0	48.9	97.9	73.4
2-3	1.5	164.7	-9.7	261.1	0	42.1	84.2	63.2
3-4	1.5	154.9	-37.5	429.0	0	25.6	51.3	38.5
4-5	1.485	117.4	-10.3	663.7	0	16.6	33.1	24.9
5-6	1.441	107.1	2.9	819.8	0	13.4	26.8	20.1
6-8	1.353	110.0	3.1	952.8	0	11.5	23.1	17.3
8-10	1.235	116.2	3.1	1088.3	0	10.1	20.2	15.2
10-11	1.147	122.5	4.0	1158.8	0	9.5	19.0	14.2
11-12	1.088	126.4	2.8	1191.4	0	9.2	18.5	13.8
12-14	1.020	129.2	2.8	1245.9	0	8.8	17.7	13.2
14-16	0.962	134.9	3.8	1319.0	0.023	8.3	16.7	12.5
16-20	0.910	142.4	3.0	1422.3	0.062	7.7	15.5	11.6
20-25	0.848	154.6	3.2	1553.9	0.120	7.1	14.2	10.6
25-35	0.783	170.5	3.2	1720.5	0.199	6.4	12.8	9.6
35-65	0.699	202.6	3.1	2000	0.343	5.5	11.0	8.2
65-100	0.645	296.9	3.1	2000	0.473	4.9	9.8	7.4

The G_o, s_u, K_o profiles simulated by the NGI-ADP model are compared with the PISA representative offshore glacial clay till site distributions (Byrne et al., 2020b), in Figure 3.3. The NGI-ADP model in PLAXIS 3D adopts a linear representation of s_u, G_o profiles for each sublayer, whereas the K_o profile remains constant with depth within each sublayer. The model succeeds in reproducing accurately the required s_u profile. In the NGI-ADP model the G_o profile is calculated based on the G_{ur}/s_u^A ratio leading to a piecewise representation of the G_o profile resulting into a slightly different distribution from the G_o profile corresponding to the representative offshore glacial clay till site. Moreover, in the PLAXIS implementation, the G_{ur}/s_u^A ratio in the NGI-ADP model is limited a maximum of 2000; this results in a slight underprediction of the G_o profile at the bottom sub-layer.

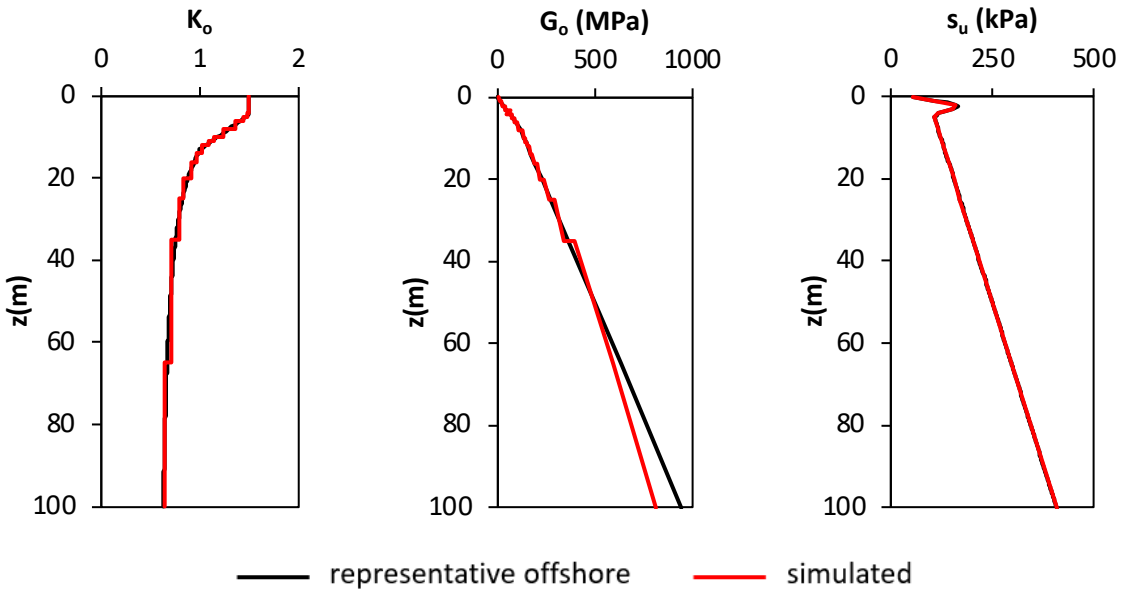


Figure 3.3: Simulated K_o, G_o, s_u profiles by the NGI-ADP model in comparison to the PISA representative offshore glacial clay till site (Byrne et al., 2020b).

The interface elements located around the pile were modelled as a separate material using the Mohr-Coulomb model with a 'Drained Calculation Type' and a tension cut-off at a tensile strength of $\sigma_t = 0kPa$ to accommodate potential gap opening at the back of the pile. Employing a different model for the interface elements was essential because the NGI-ADP

model does not have a tension cut-off criterion. The undrained shear strength of the interface elements was reduced to 0.65 times the undrained shear strength of the surrounding soil elements ($s_u^{int} = 0.65 s_u^{TXC}$) (Panagoulias et al., 2018). These values were implemented in the Mohr-Coulomb model using the cohesion model inputs (c, c_{inc}). The stiffness profile of the interface elements simulated using the Young modulus model inputs (E, E_{inc}) with $E = 2(1+\nu)G_o$ and the K_o profile adopted, were carefully selected to match the G_o and K_o profiles simulated by the NGI-ADP model. The Poisson ratio was set at $\nu = 0.495$. The horizontal interface elements located at the bottom of the pile were modelled using the same material properties as their adjacent soil elements. The Mohr-Coulomb parameters employed for the interface elements surrounding the pile are presented in Table 3.3.

Table 3.3: Mohr-Coulomb input parameters used for the interface elements.

$z(m)$	K_o	$s_u(kPa)$	$s_{uinc}(kPa)$	$E(kPa)$	$E_{inc}(kPa)$
0-1	1.5	35.0	34.2	16909	16524.7
1-2	1.5	69.264	34.1	60066	29582.3
2-3	1.5	103.4	-4.3	127406	-5287.2
3-4	1.5	99.1	-23.6	19828.5	-47229.3
4-5	1.485	75.5	-5.6	231292	-17108.4
5-6	1.441	69.9	1.7	262824	6427.6
6-8	1.353	71.6	1.99	313461	8720.7
8-10	1.235	75.6	2.1	377725.5	10506.8
10-11	1.147	79.8	2.5	424560	13210.2
11-12	1.088	82.3	1.8	450456	9856.7
12-14	1.020	84.1	1.8	481439.5	10532.4
14-16	0.962	87.8	2.5	531632	15198.5
16-20	0.910	92.8	1.9	606501	12513.3
20-25	0.848	100.4	2.04	718615	14621.6
25-35	0.783	110.65	2.1	872075	14888.9
35-60	0.699	129.54	2.04	1191844	19346.3

3.4.1.2 Calibration procedure for the PIMS model

The calibration of parameters for PIMS model is achieved by using data from isotropically consolidated undrained triaxial compression and extension tests. For the current analyses, the calibrated parameters proposed in Whyte et al. (2020) for the PISA Cowden test site were adopted. The normalized values of strain $\bar{\varepsilon}_q$ and deviatoric stress $\bar{q}$ are determined from:

$$\bar{\varepsilon}_q = I_r^* \times \varepsilon_q \quad (3.3) \qquad I_r^* = \frac{3G_o}{2s_{uc}} \quad (3.4) \qquad \bar{q} = \frac{q}{q_{uc}} \quad (3.5)$$

where $s_{uc} = q_{uc}/2$ in the undrained shear strength for triaxial compression. For the current project these parameters (Table 3.4) were defined by the undrained triaxial compression data for Bolders Bank soil profile (a stiff glacial clay till similar to the PISA Cowden test site clay till) presented in Whyte et al. (2020). Following Whyte et al. (2020) a triaxial extension factor $\beta = 0.7$ was adopted to account for the soil strength anisotropy.

Table 3.4. PIMS calibrated parameters for Bolders bank soil profile (see Whyte et al., 2020).

Micro-model	**1**	**2**	**3**	**4**	**5**	**6**
$\bar{\varepsilon}_q$	0.0075	0.0225	0.075	0.225	4.0	10.0
$\bar{q}$	0.02	0.05	0.10	0.16	0.26	0.45
w	0.06333	0.18486	0.27928	0.4333	0.01923	0.00547
Micro-model	**7**	**8**	**9**	**10**	**11**	**12**
$\bar{\varepsilon}_q$	16.88	33.75	56.25	90.0	135.0	180.0
$\bar{q}$	0.61	0.69	0.77	0.86	0.95	1.00
w	0.00624	0.00208	0.00267	0.002	0.00089	0.00066

Eighteen sub-layers were generated in PLAXIS 3D to simulate the PISA representative offshore glacial clay till site. The Poisson's ratio adopted is $v = 0.495$ for undrained conditions. The K_o ,G_o, s_u profiles employed as inputs to the PIMS model are plotted with the PISA representative offshore glacial clay till site distributions (Byrne et al., 2020b), in Figure 3.4. PIMS model provides a piecewise representation of the s_u and G_o distributions adopting the value at the middle of each sub-layer. The K_o profile is $K_o = 1$ in the provided PIMS implementation that is employed in this thesis.

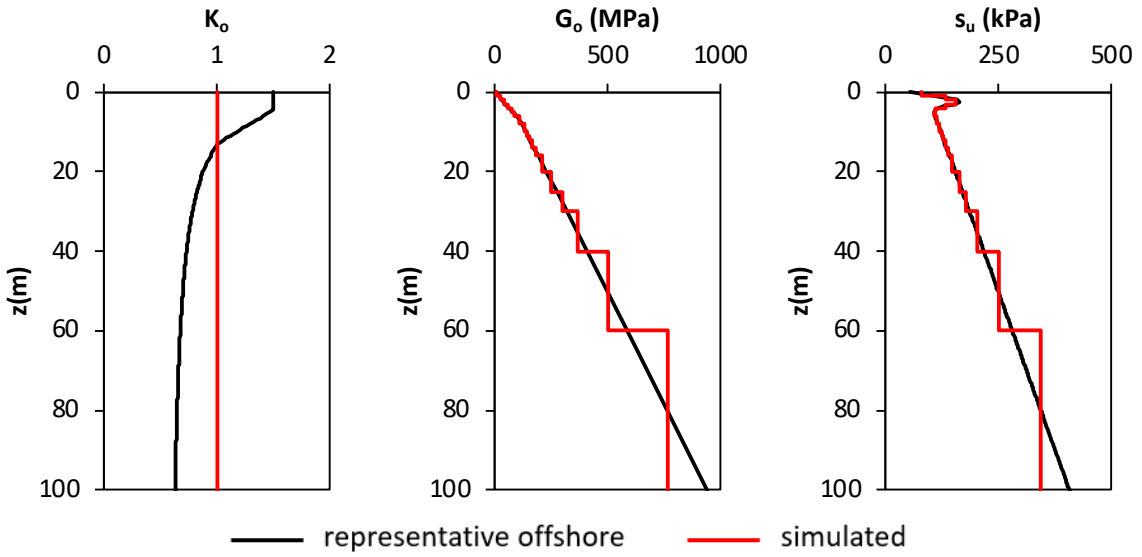


Figure 3.4: Simulated K_o, G_o, s_u profiles by the PIMS model in comparison to the PISA representative offshore glacial clay till site (Byrne et al., 2020b).

The interface element properties employed in these analyses are similar to those described analytically in Section 3.4.1.1. The interface elements surrounding the pile were modelled using the Mohr-Coulomb model with a 'Drained Calculation Type' and a tension cut-off at a tensile strength of $\sigma_t = 0kPa$. The undrained shear strength of the interface elements was set as $s_u^{int} = 0.65 s_u^{TXC}$. The G_o and K_o profiles adopted, were carefully selected to match those simulated by the PIMS model. The Poisson ratio was set at $\nu = 0.495$. The horizontal interface elements located at the bottom of the pile were modelled using the same material properties as their adjacent soil elements.

3.4.1.3 Validation of NGI-ADP and PIMS models with PISA FEA

To assess the performance of the NGI ADP and PIMS models, analyses were conducted to compare with the finite element analyses – based on an extended modified Cam Clay effective stress model – described in Zdravkovic et al. (2020b). These analyses are concerned with the representative offshore glacial clay till site specified in Byrne et al. (2020b) related to the calibration piles originally adopted to calibrate the PISA design model presented in Table 3.1. The load-displacement response using the PIMS model has been computed solely

for the piles C6, C8 and C9. A comparison between two sets of calculations; 'ultimate response' (ground level pile displacement v_G up to $v_G = D/10$) and 'small displacement response' (ground level pile displacements up to $v_G = D/10000$) is shown in Figure 3.5 and Figure 3.6 respectively. Both the NGI ADP and the PIMS results compare well with the ICFEP data in Byrne et al. (2020b). However, the NGI-ADP model is computationally less demanding than the PIMS model, hence NGI-ADP has been chosen due to its inherent simplicity to represent the behaviour of the different clay soil profiles examined in this thesis.

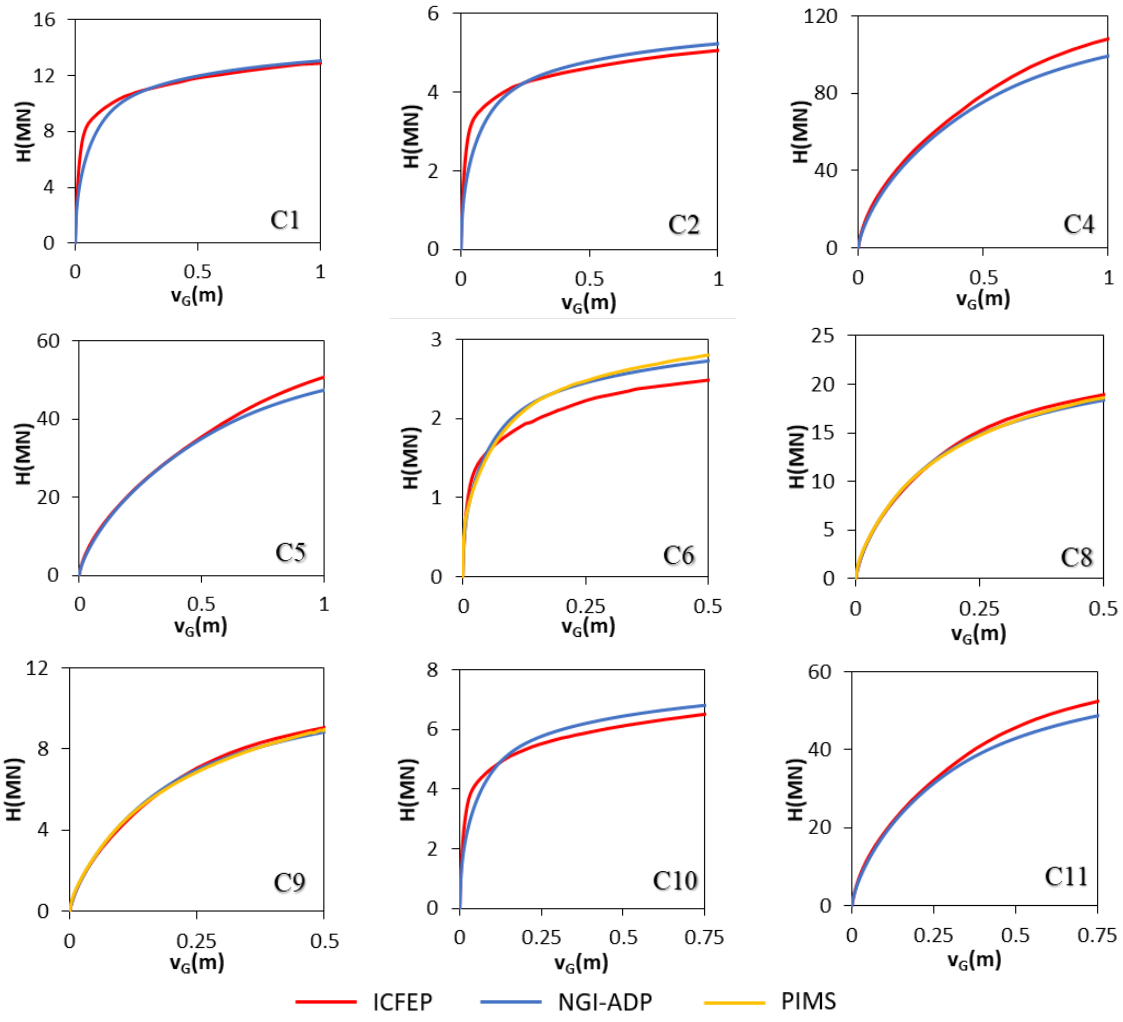


Figure 3.5: Load-displacement responses at ground level for the ultimate response of the nine calibration piles (Table 3.1) embedded in the PISA representative offshore glacial clay till site computed by the 3D FE model in Byrne et al. (2020b) and the current NGI-ADP and PIMS PLAXIS 3D results. PIMS has been employed solely for the piles C6, C8 and C9.

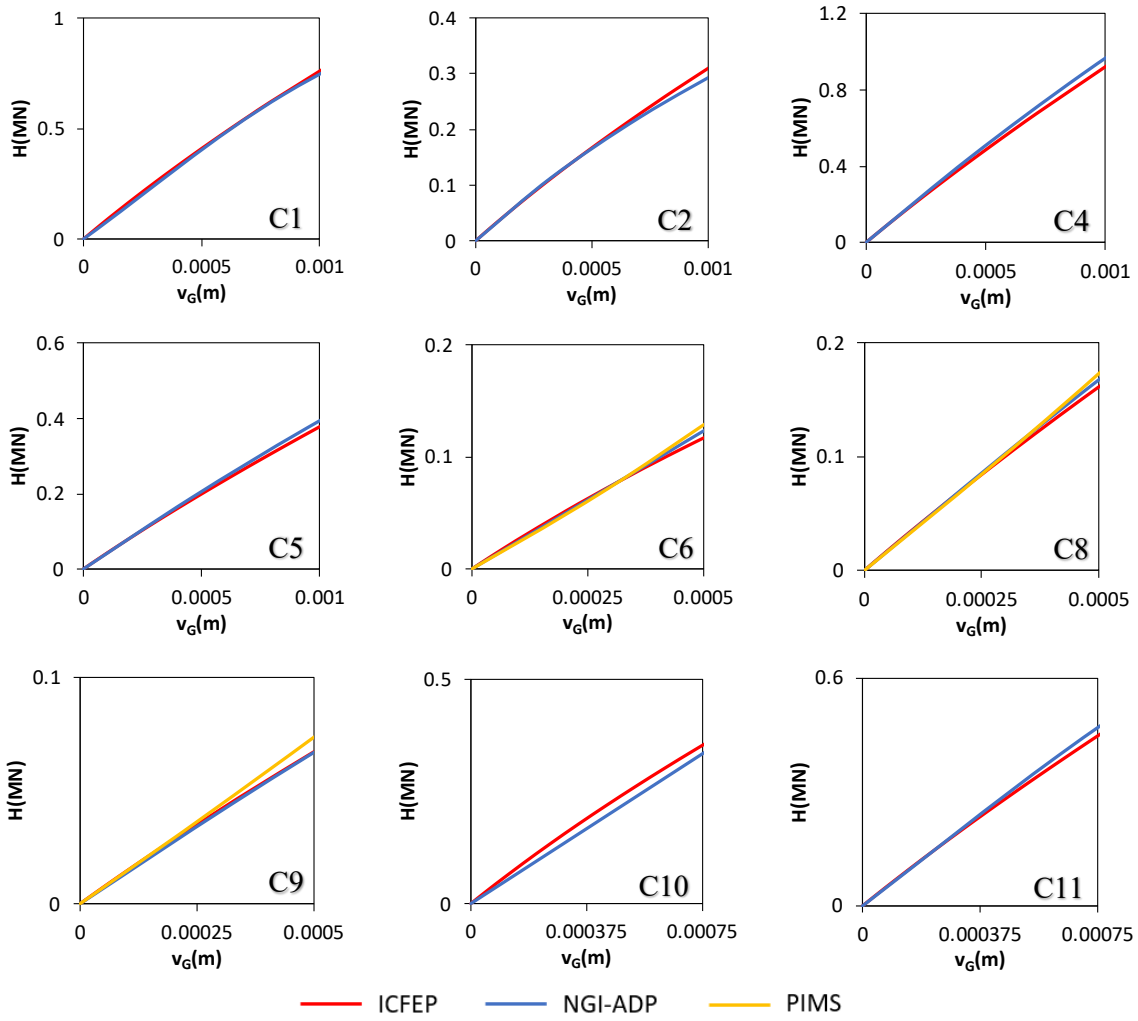


Figure 3.6: Load-displacement responses at ground level for the small displacement response of the nine calibration piles (Table 3.1) embedded in the PISA representative offshore glacial clay till site computed by the 3D FE model in Byrne et al. (2020b) and the current NGI-ADP and PIMS PLAXIS 3D results. PIMS has been employed solely for the piles C6, C8 and C9.

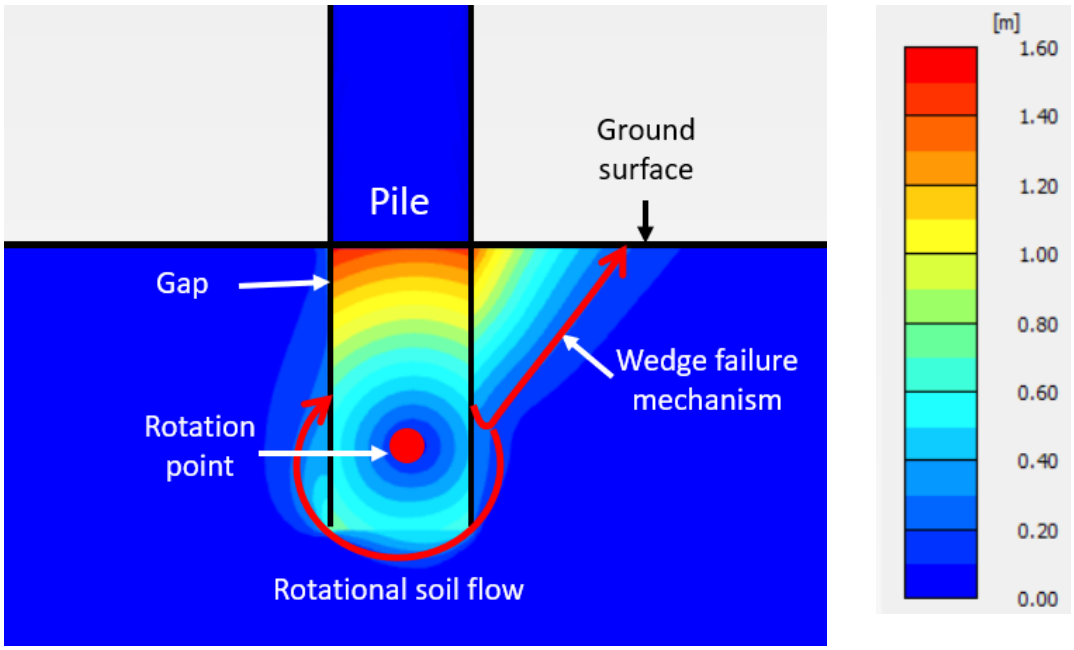


Figure 3.7: Total displacement contours extracted by PLAXIS 3D FEA for the representative offshore glacial clay till site for an applied load H_{ult}^{3D} at a ground level displacement of v_G=$D/10$.

Figure 3.7 shows the displacement contours extracted by PLAXIS 3D for pile C1 at a ground level displacement close to $0.1D$. A gap is formed behind the pile due to the tension cut-off criterion employed for the interface elements properties. This agrees with PISA field test findings (Byrne et al., 2020a) and affects the pile's lateral behaviour (reduces the pile's capacity). Furthermore, it can be observed that there are two distinct failure mechanisms formed around the pile (wedge failure, rotation soil flow mechanisms) as mentioned in Randolph and Gouvernec (2011). Also, the rotation point is located at a depth z of approximately $z = 0.7L$ below the ground level.

3.4.2 Numerical analysis of monopile embedded in the PISA Dunkirk sand

The PISA representative offshore homogenous sand site with relative density of $D_R = 75\%$ is the baseline sand soil profile adopted in the PISA project. The soil characteristics adopted for this offshore site are based on the PISA Dunkirk test site presented in Zdravkovic et al. (2020a). The submerged unit weight of the Dunkirk sand is $\gamma' = 10.09\,kN/m^3$; the

coefficient of earth pressure at rest is $K_0 = 0.4$; the Poisson ratio is $v = 0.17$ (Zdravkovic et al. 2020). The small strain shear modulus G_0 profile adopted for the idealised offshore dense sand site in Burd et al. (2020) was used. The profile is determined from the local values of mean effective stress p' and the initial void ratio e_0 by using the formula proposed by Hardin (1978).

$$G_0 = \frac{B \times {p'}_{ref}}{0.3 + 0.7e^2} \times \sqrt{\frac{p'}{{p'}_{ref}}} \tag{3.6}$$

In this equation, ${p'}_{ref}$ is a reference pressure ${p'}_{ref} = 101.3kP$, $e = 0.629$ is the void ratio for Dunkirk sand layer with $D_R = 75\%$ (Burd et al. 2020); p' is the local value of mean effective stress and B is a parameter that was calibrated through the seismic cone penetration tests conducted at the PISA Dunkirk test site, where $B = 875$ (Taborda et al. 2020).

Two different constitutive models: HS-Small (Brinkgreve et al., 2018) and Norsand (Jefferies and Been, 2016) have been tested to evaluate their performance. HS-Small is a built-in model in PLAXIS 3D software. Norsand is available in PLAXIS 3D as a user defined model. Both constitutive models were shown by Minga and Burd (2019b) and Castillo-Fuentes et al. (2023) to effectively capture the monotonic behaviour of the Dunkirk sand field test piles described in McAdam et al. (2020).

HS-small model (Benz, 2007) is an extension to the Hardening Soil model (Schanz et al., 1999) for reproducing the stiffness variation with strain and the hysteretic response of soils. It is an elastoplastic, multi-yield surface effective stress soil model incorporating separate families of yield surfaces to account for volumetric and shear plastic strains and using the Mohr-Coulomb criterion to define failure. The model uses as input parameters cohesion c', friction angle ϕ' and dilation angle ψ. Three stiffness parameters are defined by E_{50}^{ref}, E_{oed}^{ref}, E_{ur}^{ref}; controlling plastic straining from deviatoric loading, primary compression and elastic

unloading-reloading. The model employs a pressure dependent elasticity through a power-law expression which uses an exponent m to correlate the elasticity with the mean effective stress. The degradation of shear stiffness with increasing strains is defined by using G_o (small strain shear stiffness) and threshold shear strain $\gamma_{0.7}$ which is the value of shear strain at which the shear stiffness is reduced to 70% of the initial shear stiffness.

Norsand (Jefferies and Been, 2016) is an isotropic elastoplastic model that is formulated within the critical state soil mechanics (CSSM) framework. The CSSM is incorporated in the model by using the critical friction ratio in triaxial conditions M_{tc} and the state parameter ψ introduced by Been and Jefferies (1985) calculated by $\psi = e - e_{cs}$ where e is the void ratio and e_{cs} is the critical state void ratio at current mean effective stress. In the PLAXIS implementation two different options to represent the variation of critical state void ratio with pressure are present: (i) the conventional semi-log form (Schofield and Wroth, 1968), (ii) the power-law law relationship proposed by Li and Wang (1998). For the current analyses the semi-log representation of the critical state line was adopted. The model employs an input ψ_0 which defines the state parameter at the start of the loading path. This parameter controls the magnitude of dilatancy of the soils and it is incorporated in Norsand to discriminate a 'loose' from a 'dense' soil. A soil with a positive value of ψ_0 is classified as 'loose' and is expected to contract whereas, a soil with a negative value of ψ_0 is classified as 'dense' and is expected to dilate. Norsand is characterised by a pressure dependent elasticity; the small strain shear modulus G_o is defined through a power-law expression which uses an exponent to n^G to correlate the initial stiffness with the mean effective stress. The parameters describing the plasticity behaviour of the model are H, the hardening modulus, χ_{tc}, which reflects the variation of D^p_{min} (maximum dilation) with the state parameter ψ and N, which describes the strength and dilatancy relationship. The model

includes also a softening parameter S to model undrained loading conditions; however, for the current analyses this term is $S = 0$, since all the 3D FEA relate to drained loading.

3.4.2.1 Calibration procedure for the HS-Small model

A friction angle $\phi' = 43.5°$ was employed; this friction angle was found to accurately reproduce the PISA field test piles response (refer to Section 3.4.4). The cohesion was taken as $c' = 0.1kPa$ (Panagoulias et al., 2019). The formula proposed by Bolton (1986) was used to calculate the dilation angle ψ:

$$\psi = \frac{\phi' - \phi'_{cs}}{0.8} \tag{3.7}$$

where ϕ'_{cs} is the critical state friction angle for triaxial compression taken as $\phi'_{cs} = 32°$ (Zdravkovic et al., 2020a). The dilation cut-off option has been deactivated in the current set of analyses. The simulated small strain shear stiffness, G_0, varies with mean effective stress (Brinkgreve et al., 2018):

$$G_{0ref} = \frac{G_0}{\left(\frac{c'_{ref} \times cos\,\phi' - \sigma'_3 \times sin\,\phi'}{c'_{ref} \times cos\,\phi' - p'_{ref} \times sin\,\phi'}\right)^m} \tag{3.8}$$

where $m = 0.5$ is the coefficient that controls the stress-dependency of stiffness with depth; $p'_{ref} = 100kPa$ is the reference pressure; G_0 is the small strain shear stiffness at the mid-depth of each soil layer; σ'_3 is the local horizontal effective stress. The simulated G_0 profile by the HS-Small is presented in Figure 3.16(b) in Section 3.4.3.2. The HS-Small achieves an excellent fit to the G_0 ground profile adopted in the PISA project.

The secant stiffness in standard drained triaxial test, $E_{50}^{ref} = 45000kPa$ is determined by (Brinkgreve et al., 2010):

$$E_{50}^{ref} = 60000 \times D_R/100 \tag{3.9}$$

where D_R is the relative density of the sand ($D_R = 75\%$). The adopted E_{50}^{ref} value agrees well with the E_{50}^{ref} calculated by the PISA drained triaxial laboratory test at a confining pressure of $p' = 100kPa$, as shown in Figure 3.8. The secant stiffness at a deviatoric stress $q = 0.5q_{ult}$ determined by the PISA lab tests presented in Taborda et al. (2020) is:

$$E_{50}^{ref} = \frac{160}{0.0034} = 47000kPa \tag{3.10}$$

The reference tangent stiffness in primary oedometric loading E_{oed}^{ref} is assumed to be equal to E_{50}^{ref} (Brinkgreve et al., 2010), giving $E_{oed}^{ref} = 45000kPa$.

The reference unloading-reloading stiffness E_{ur}^{ref} is calculated by the formula below (Brinkgreve et al., 2010):

$$E_{ur}^{ref} = 3E_{50}^{ref} = 135000\text{kPa} \tag{3.11}$$

The $\gamma_{0.7}$ for $D_R = 75\%$ is calculated by (Brinkgreve et al., 2010):

$$\gamma_{0.7} = (2 - D_R/100) \times 10^{-4} = 0.000125 \tag{3.12}$$

The failure ratio R_f for $D_R = 75\%$ is determined by (Brinkgreve et al., 2010):

$$R_f = 1 - D_R/800 = 0.90625 \tag{3.13}$$

The normally consolidated coefficient of earth pressure K_0^{NC} is set to 0.4 for the normally consolidated sand found in the Dunkirk test site (Zdravkovic et al., 2020a). The HS-Small parameters employed to simulate the response of the PISA representative offshore homogenous sand site with a relative density D_R of 75% are listed in Table 3.5.

The value of the friction angle $\phi' = 43.5°$ was determined by fitting the simulated pile response by HS-Small to the measured field test results in McAdam et al. (2020). Calibrating

the value of ϕ' using the PISA laboratory tests presented in Taborda et al. (2020) results in a significant underprediction of the pile's lateral capacity. Taborda et al. (2020) also demonstrated an underprediction of the pile's lateral capacity when calibrating an advanced bounding surface plasticity model in the framework of critical state soil mechanics (Taborda et al., 2014) on the PISA laboratory tests. This discrepancy was attributed to the difference between the measured profile of G_o from the field tests and the lab experiments. Figure 3.8 presents the drained laboratory tests presented in Taborda et al. (2020) (see Table 3.7) simulated using the built-in soil testing facility in PLAXIS 3D by HS-Small model, in terms of deviatoric stress-axial strain $(q - \varepsilon_a)$, axial strain–volumetric strain $(\varepsilon_a - \varepsilon_{vol})$. It is evident that the HS-Small model overpredicts the $(q - \varepsilon_a)$ response due to: (i) the high peak friction angle adopted, ${\phi_p}' = 43.5°$, which is higher compared to the peak friction angle determined by the PISA laboratory tests, ${\phi_p}' = 42°$, and (ii) the inability of the model to capture the post-peak softening response. Additionally, including a proper dilation cut-off criterion may improve the overprediction of the dilational response over large axial strains (ε_a).

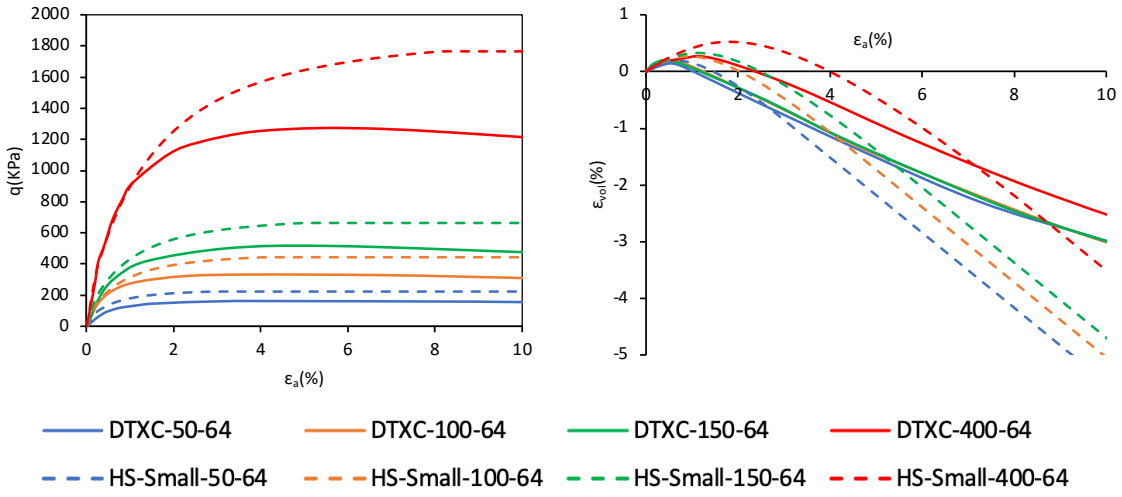


Figure 3.8: Simulation of drained TXC tests (Table 3.6) using HS-Small with Imperial College PISA TXC tests for Dunkirk sand samples presented in Taborda et al. (2020). Comparison of $q - e_a$, $e_a - e_{vol}$ responses for pre-shear mean effective stress $p' = 50, 100, 150, 400\ kPa$.

Table 3.5: HS-Small parameters employed for the representative offshore homogenous sand site with D_R=75%.

Parameters	**Symbols**	**Values**	**Units**
Reference secant stiffness in triaxial compression	E_{50}^{ref}	45000	kPa
Reference tangent stiffness in primary oedometric loading	E_{oed}^{ref}	45000	kPa
Reference triaxial unloading/reloading stiffness	E_{ur}^{ref}	135000	kPa
Stress-dependency of stiffness	m	0.5	$(-)$
Effective cohesion	c'	0.1	kPa
Effective friction angle	ϕ'	43.5	°
Dilatancy angle	ψ	14.375	°
Threshold shear strain for 70% reduction of G_o	$\gamma_{0.7}$	0.00012	$(-)$
Reference small-strain shear modulus	G_{0ref}	187000	kPa
Reference stress	p'_{ref}	100	kPa
Unloading/reloading Poisson's ratio	ν_{ur}	0.17	$(-)$
K_o stress ratio in normally consolidated condition	K_0^{NC}	0.4	$(-)$
Failure ratio	R_f	0.90625	$(-)$
Interface strength reduction factor	R_{inter}	0.5841	$(-)$

The reduction in interface shear strength during slip occurrence is considered by utilising the interface strength reduction factor $R_{inter} = \tan\delta \,/\, \tan\phi'$. A value of $R_{inter} = 0.5841$ is calculated such as the interface friction angle is set at $\delta = 29°$ (Panagoulias et al., 2018). This value differs marginally from the interface friction angle $\delta = \phi'_{cs} = 32°$ adopted in Taborda et al. (2020). However, according to Zdravkovic et al. (2021), this slight discrepancy in the selected value of δ will have a negligible effect on the pile's behaviour. The horizontal interface elements located at the bottom of the pile were modelled using a separate material adopting the HSS model with the same parameters as those adopted for the soil profile (see Table 3.5).

3.4.2.2 Calibration procedure for the Norsand model

The Γ, λ_e parameters define the semi-log critical state line. Adopting the critical state line defined by Aghakouchak (2015) for the Dunkirk sand, provided $\Gamma = 0.92$ and $\lambda_e = 0.032$. The friction ratio at critical state is $M_{tc} = 1.28$ (Zdravkovic et al., 2020a). The parameters N and χ_{tc} are soil properties measured by triaxial compression tests. Shuttle and Jefferies (2010) after investigating drained triaxial compression tests on different sands indicated that N varies between 0.3 and 0.4 and χ_{tc} varies between 3 and 4.5. For the Dunkirk sand N and χ_{tc} are taken as $N = 0.3$ and $\chi_{tc} = 4$. The overconsolidation ratio parameter, is taken as $R = 1$ for the normally consolidated Dunkirk sand (Zdravkovic et al., 2020a).

The simulated small strain shear stiffness G_0 varies with mean effective stress by using (Jefferies and Been, 2016):

$$G_0 = G_{0ref} \times (\frac{p'}{p'_{ref}})^{n_G} \tag{3.14}$$

where $n_G = 0.5$ controls the stress-dependency of stiffness with depth; $p'_{ref} = 100kPa$ is the reference pressure; G_0 is taken at the mid-depth of each soil layer; p' is the local mean effective stress. The simulated G_0 profile by Norsand is presented in Figure 3.16(b) in Section 3.4.3.2; a precise fit is achieved between the simulated G_0 profile and the G_0 ground profile adopted in the PISA project.

The hardening parameter H is determined by (Jefferies and Been, 2016):

$$H = H_0 - H_\Psi \times \psi_0 \tag{3.15}$$

where ψ_0 is the state parameter and H_0, H_Ψ are the hardening parameters that are calibrated by TXC tests; $H_0 = 400$ was found to provide a good fit to the PISA triaxial laboratory tests (Figure 3.9); $H_\Psi = 0$ was adopted to ease the calibration procedure. A negative value of ψ_0

needs to be adopted to model the behaviour of the 'dense' representative offshore homogenous sand site with $D_R = 75\%$. The state parameter can vary between 0 to -0.3 (Jefferies and Been, 2016) for a dense sand profile and depends on the mean effective pressure p' based on the Equation (3.16). For the current set of analyses, for simplicity, a single sand layer was employed. Therefore, it was decided to calculate the value of ψ_0 for the $D_R = 75\%$ Dunkirk sand profile with $e = 0.629$ based on the p' resulting at the depth z located at the middle of each pile.

The Norsand parameters employed to simulate the response of the PISA representative offshore homogenous sand site with a relative density D_R of 75% are listed in Table 3.6.

Table 3.6: Norsand parameters employed for the PISA representative offshore homogenous sand site with D_R=75%.

Parameters	**Symbols**	**Values**	**Units**
Reference shear modulus	G_{ref}	153631.5	kPa
Reference mean pressure	p_{ref}	101.3	kPa
Exponent of the power-law elasticity	n_G	0.5	(-)
Poisson's ratio	v	0.17	(-)
Void ratio corresponding to a $p' = 1$kPa	Γ	0.92	(-)
Slope of the critical state (e-$ln(p)$)	λ_e	0.032	(-)
Friction ratio at critical state in triaxial conditions	M_{tc}	1.28	(-)
Material parameter controlling the maximum stress ratio with the minimum dilatancy	N	0.3	(-)
Material parameter governing the slope of the minimum dilatancy with the state parameter	χ_{tc}	4	(-)
Hardening parameter	H_0	400	(-)
Hardening parameter	H_ψ	0	(-)
Over-consolidation ratio	R	1	(-)
Softening flag	S	0	(-)
Initial value of the state parameter: Piles C1, C2	ψ_0	-0.16	(-)
Initial value of the state parameter: Piles C4, C5	ψ_0	-0.125	(-)

The interface elements surrounding the pile were modeled as Mohr-Coulomb frictional material, utilizing the same stiffness profile as the adjacent soil elements. The cohesion was set to $c' = 1kPa$ and the friction angle to $\varphi' = 29°$. Horizontal interface elements situated at the base of the pile were modeled separately using the Norsand model, with parameters identical to those adopted for the soil profile (refer to Table 3.6).

To demonstrate the performance of the model for the calibrated parameters in Table 3.6, four PISA drained isotropically consolidated laboratory tests in triaxial compression with different initial mean effective stress and void ratio (Taborda et al., 2020) are simulated using the built-in soil testing facility in PLAXIS 3D. The tests are summarised in Table 3.7 and comparisons shown in Figure 3.9 in terms of axial strain–deviatoric stress ($q - e_a$), axial strain–volumetric strain ($e_a - e_{vol}$). To confirm the Norsand calibration, for the simulation of laboratory tests the state parameter ψ_0 is calculated from the measured void ratio and fitted critical state line for the pre-shear mean effective stress p'_o.

$$\psi_0 = e - \ (\Gamma - \lambda_e \times lnp'_o) \tag{3.16}$$

Table 3.7: Initial conditions of the triaxial tests presented in Taborda et al. (2020) selected for evaluating Norsand performance.

Test code	**Pre-shear void ratio e_0**	**Pre-shear p'_o (kPa)**	**State parameter ψ_0**
DTXC-50-64	0.64	50	-0.1548
DTXC-100-64	0.639	100	-0.1336
DTXC-150-64	0.637	150	-0.1227
DTXC-400-64	0.633	400	-0.0953

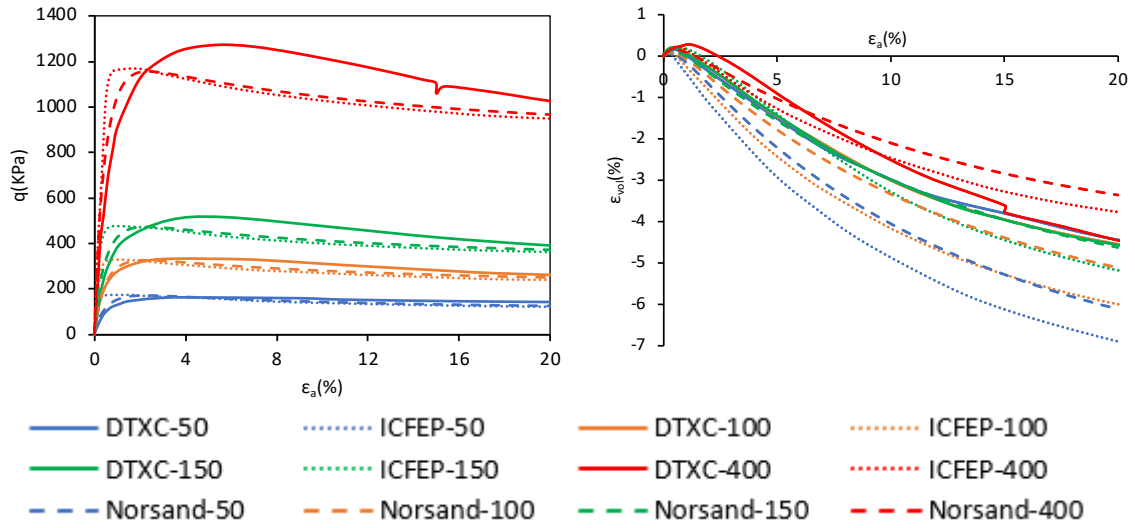


Figure 3.9. Simulation of drained TXC tests (Table 3.7) using Norsand with Imperial College PISA TXC tests and ICFEP simulated for Dunkirk sand samples (Taborda et al., 2020). Comparison of $q - e_a$, $e_a - e_{vol}$ responses for pre-shear mean effective stress $p' = 50, 100, 150, 400\ kPa$.

3.4.2.3 Validation of HS-Small and Norsand with PISA FEA

To assess the performance of the HS-Small and Norsand models, analyses were conducted to compare with the finite element analyses – based on the state parameter-based bounding surface plasticity model - described in Taborda et al. (2020). These analyses are concerned with the representative offshore dense sand site with $D_R = 75\%$ specified in Burd et al. (2020a). The load-displacement response using the Norsand model has been computed solely for the piles C1, C2, C4 and C5. A comparison between two sets of calculations; 'ultimate response' (ground level pile displacement v_G up to $v_G = D/10$) and 'small displacement response' (ground level pile displacements up to $v_G = D/10000$) is shown in Figure 3.10 and Figure 3.11. The HS-Small and the Norsand results both compare relatively well with the data in Burd et al. (2020a). Norsand outperforms the HS-Small, however it is not very computationally stable when reaching near the target ground level displacement of $0.2D$ – which is required to extract the soil reaction curves. Also, Norsand is much more computationally demanding than HS-Small. Hence, HS-Small has been chosen to represent the behaviour of the different sand soil profiles examined in this thesis.

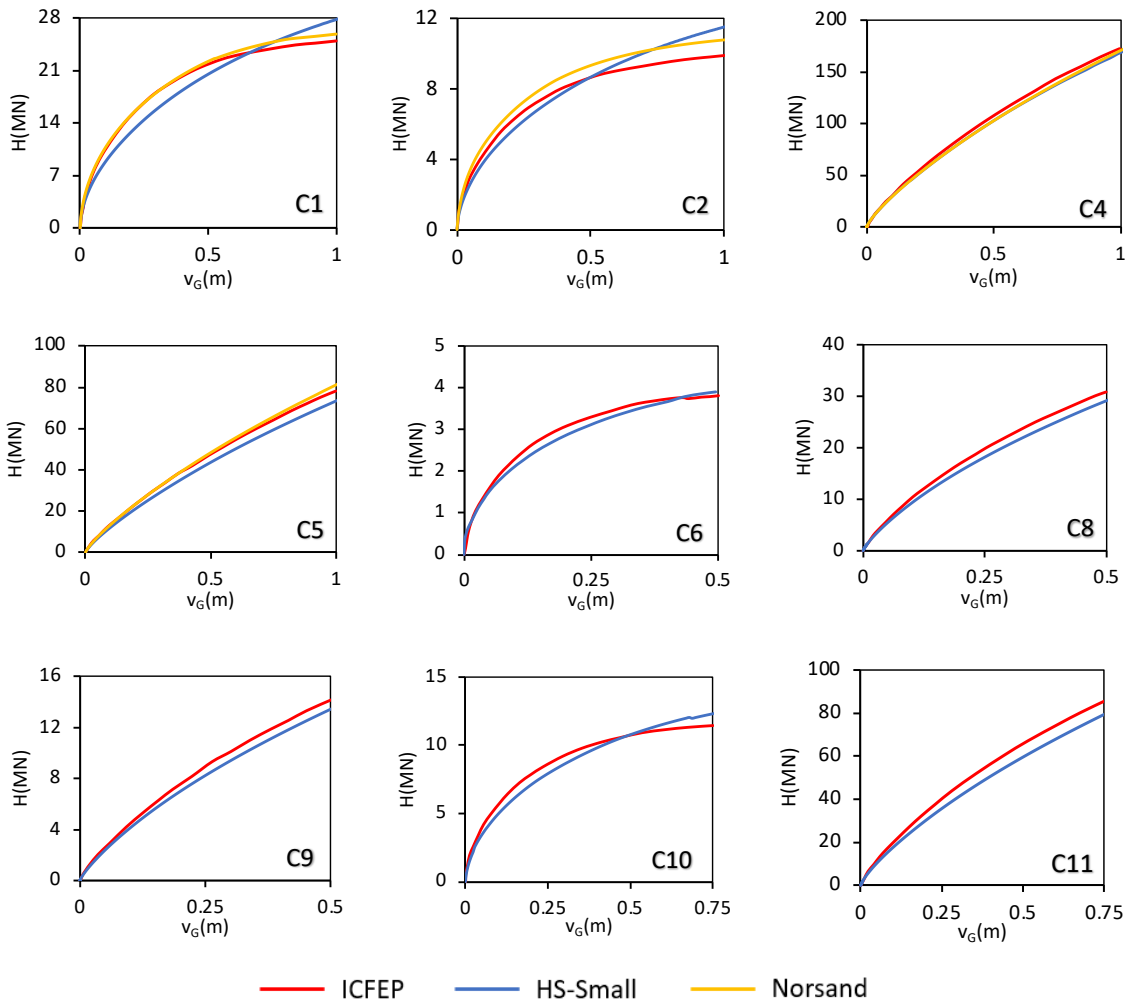


Figure 3.10: Load-displacement responses at ground level for the ultimate response of the nine calibration piles (Table 3.1) embedded in the PISA representative offshore homogenous sand site with $D_R = 75\%$ computed by the 3D FE model in Burd et al., (2020b) and the HS-Small and Norsand PLAXIS 3D results. Norsand has been employed solely for the piles C1, C2, C4 and C5.

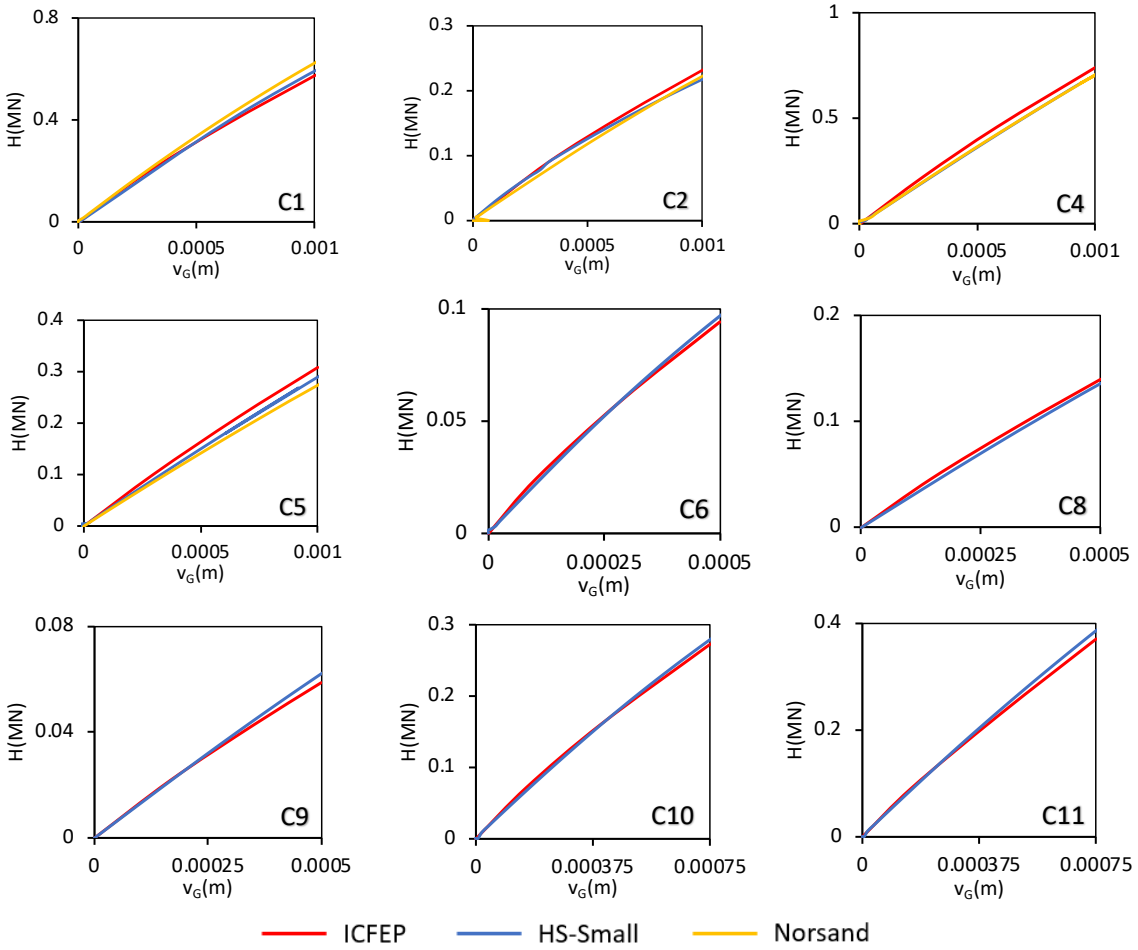


Figure 3.11: Load-displacement responses at ground level for the small displacement response of the nine calibration piles (Table 3.1) embedded in the PISA representative offshore homogenous sand site with $D_R = 75\%$ computed by the 3D model in Burd et al., (2020b) and the HS-Small and Norsand PLAXIS 3D results. Norsand has been employed solely for the piles C1, C2, C4 and C5.

Figure 3.12 shows the displacement contours extracted by PLAXIS 3D for pile C1 embedded in PISA representative offshore homogenous sand site with $D_R = 75\%$ at a ground level displacement close to $0.1D$. The HS-Small was employed to simulate the sand soil. The gap is not evident behind the pile. This contrasts with the PISA field tests (McAdam et al., 2020) where a gap was formed. This phenomenon occurred in the field tests because the top 5.4m of the sand were unsaturated (McAdam et al., 2020). As for the clay soil profile presented in Section 3.4.1.3, there are two distinct failure mechanisms formed around the pile (wedge failure, rotational soil flow mechanisms). The rotation point is located at a depth z below the ground level of approximately 2/3 of the pile's length.

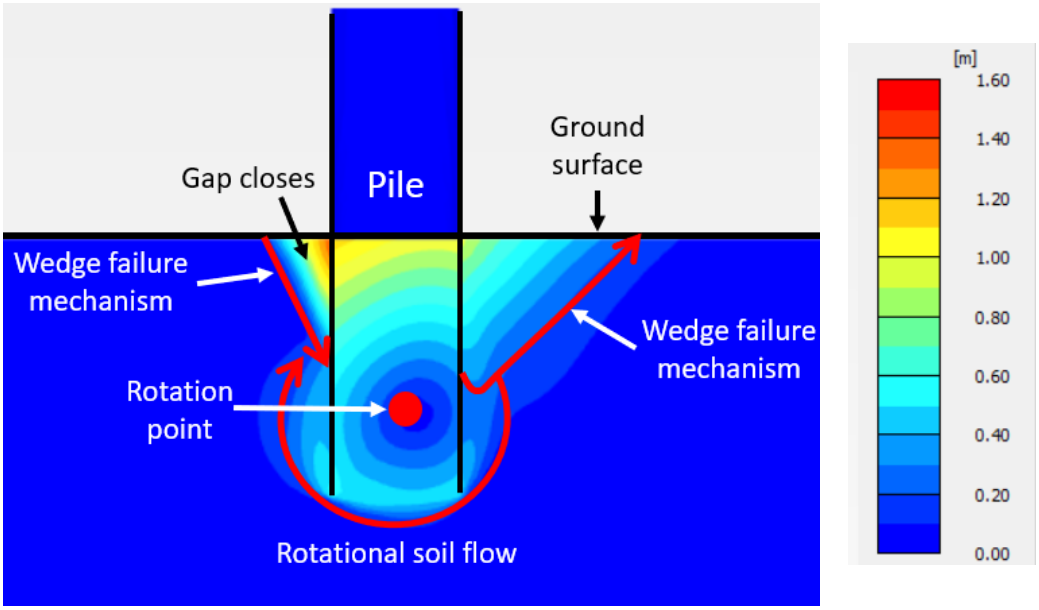


Figure 3.12: Total displacement contours extracted by PLAXIS 3D FEA for the PISA representative offshore homogenous sand site with D_R=75% for an applied load H_{ult}^{3D} at a ground level displacement of v_G=*D/10*.

3.4.3 PISA offshore soil profiles constitutive model parameters

This section outlines the selected NGI-ADP or HS-Small model parameters utilised for the the PISA representative offshore Bothkennar clay site (AWG, 2018) and the PISA representative offshore homogenous sand sites with relative densities of $D_R = 45, 60, 90\%$. The PISA representative offshore glacial clay till and homogenous sand with $D_R = 75\%$ were presented in the previous section and therefore, are omitted. These soil profiles are employed in Chapters 5, 6 and 7 of this thesis for analyses concerned with layered soil configurations. To demonstrate the performance of the selected model parameters, four laterally loaded offshore piles (C1, C2, C4, C5) summarised in Table 3.1 are simulated in PLAXIS 3D and the results are compared with the ICFEP data presented in AWG (2018).

3.4.3.1 Calibration procedure for PISA Bothkennar offshore clay

The NGI-ADP model is used to simulate the behaviour of the normally consolidated PISA representative offshore Bothkennar clay soil profile. The submerged unit weight is $\gamma' = 6.19\ kN/m^3$ (AWG, 2018). The Bothkennar clay is described by a constant K_o profile ($K_o =$

0.65) and a linear s_u profile distribution, hence a single sub-layer in PLAXIS 3D is sufficient to reproduce the K_o and s_u distributions. The NGI-ADP soil parameters were calibrated following the procedure described in Section 3.4.1.1. The s_u^{TXE}/s_u^A ratio employed as in input parameter in the NGI-ADP model was determined by:

$$s_u^{TXE}/s_u^A = \frac{s_u^{TXE}}{s_u^{TXC}/0.99} = \frac{\left(\frac{q}{p}\right)^{TXE}}{\left(\frac{q}{p}\right)^{TXC}/0.99} = \frac{1}{1.38/0.99} = 0.7174 \qquad (3.17)$$

where (q/p) are the values of the average ultimate stress ratios as indicated in AWG (2018) for triaxial extension (TXE) and compression (TXC) conditions. The s_u^{DSS}/s_u^A ratio was calculated by averaging the s_u^{TXE}/s_u^A and $s_u^{TXC}/s_u^A = 0.99$ ratios (Brinkgreve et al., 2018). The γ_f^C parameter was determined by performing undrained triaxial compression/extension tests using the built-in soil testing facility in PLAXIS 3D (Figure 3.13) and seeking the best fit to the triaxial test results simulated by ICFEP using the enhanced MCC model (AWG, 2018). Selecting a value of $K_f = 22.5$ resulted in a good match. The NGI-ADP input parameters adopted in the 3D FEA for the simulation of the PISA representative offshore Bothkennar clay are presented in Table 3.8. The simulated soil profile G_o, s_u, K_o distributions by the NGI-ADP model are presented in Figure 3.14 together with the PISA representative offshore Bothkennar clay soil profile (AWG, 2018). The Mohr-Coulomb properties adopted for the interface elements located around the pile are presented in Table 3.9.

Table 3.8: NGI-ADP parameters used for Bothkennar clay simulation.

K_0	s_u^A	s_{uinc}^A	G_{ur}/s_u^A	τ_0/s_u^A	γ_f^C	γ_f^E	γ_f^{DSS}	s_u^{TXE}/s_u^A	s_u^{DSS}/s_u^A
0.65	5.758	2.677	849.9	0.388	2.35	4.71	3.53	0.7174	0.8537

Table 3.9: Mohr-Coulomb parameters adopted for interface elements.

K_o	$s_u(kPa)$	$s_{uinc}(kPa)$	$E(kPa)$	$E_{inc}(kPa)$
0.65	3.7	1.72	14631	6802

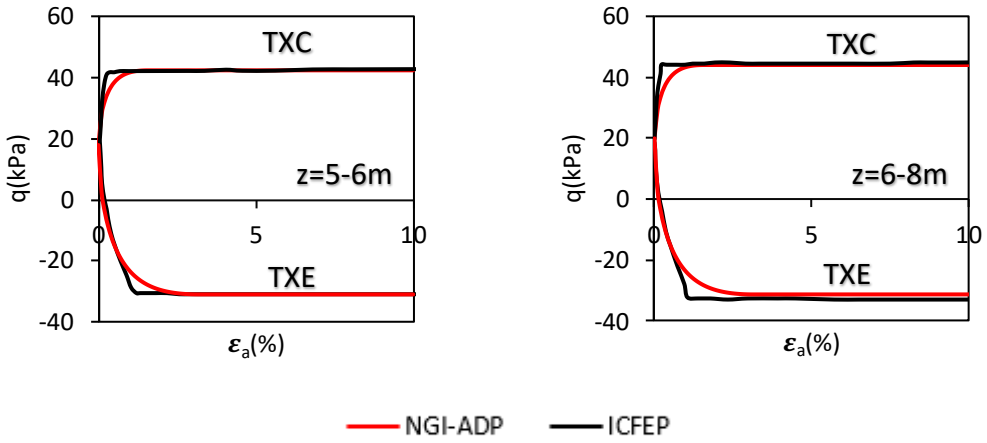


Figure 3.13: Simulation of undrained TXC and TXE tests using NGI-ADP model with MCC model simulated results in ICFEP (AWG, 2018) for Bothkennar clay soil samples retrieved at 5-8m.

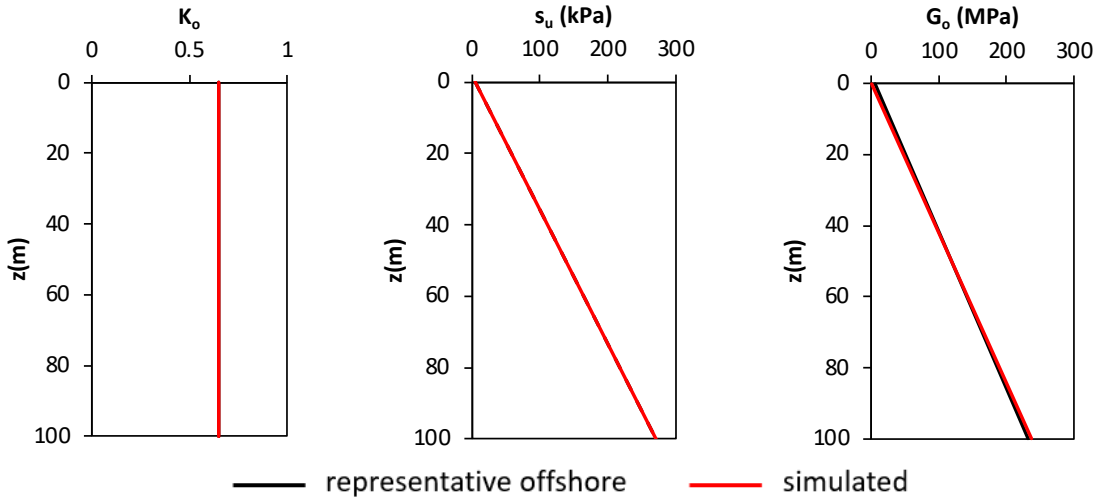


Figure 3.14: Simulated K_o, s_u, G_o profiles by the NGI-ADP model in comparison to the PISA representative offshore Bothkennar clay site (AWG 2018).

Figure 3.15 illustrates the ultimate response (ground level pile displacement v_G up to $v_G = D/10$) of the C1, C2, C4 and C5 piles (Table 3.1) predicted by the calibrated NGI-ADP parameters. The current NGI-ADP results compare well with the data in AWG (2018) produced by the ICFEP software.

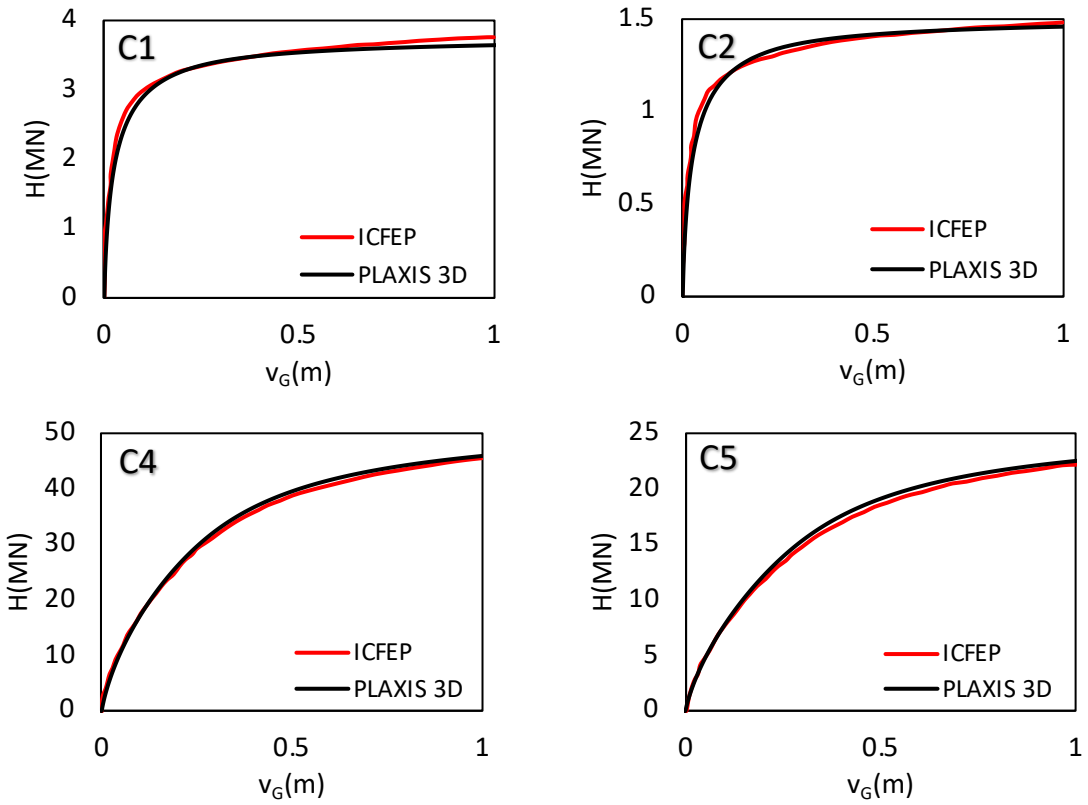


Figure 3.15. Load-displacement responses at ground level for the ultimate response of the four calibration piles (Table 3.1) embedded in the PISA representative offshore Bothkennar clay site computed by the 3D FE model in AWG (2018) denoted as ICFEP and the current NGI-ADP PLAXIS 3D results.

3.4.3.2 Calibration procedure for PISA Dunkirk sand of D_R=45, 60, 90%

The HS-Small model is used to simulate the representative offshore homogenous sand sites of different relative densities which are based on the properties of the Dunkirk test site. The submerged unit weight of the Dunkirk sand is $\gamma' = 10.09\, kN/m^3$; the coefficient of earth pressure at rest is $K_0 = 0.4$; the Poisson ratio is $v = 0.17$ (Zdravkovic et al. 2020a). Equation (3.6) is employed to calculate the G_0 profile of the three sand sites; Table 3.10 summarises the void ratio e_0 values for the different relative densities D_R ground conditions.

Table 3.10 Initial D_R and e_0 for the representative offshore sand sites (Burd et al., 2020a).

D_R	45	60	75	90
e_0	0.741	0.685	0.629	0.573

The calibration procedure presented in Section 3.4.2.1 was adopted to generate the model parameters for the PISA representative offshore homogenous sand sites of $D_R = 45, 60, 90\%$. Separate parameters are required for each relative density. A key concern when calibrating the HS-Small model is to select the appropriate value of ϕ'. The PISA TXC tests presented in Taborda et al. (2020), were conducted for samples with void ratios corresponding to $D_R = 75\%$ so there are no available lab tests to determine a relationship between ϕ' and D_R. It was decided to calibrate the model and define the most consistent value of ϕ' for each sand profile using the PISA ICFEP results reported in AWG (2018), where the critical state constitutive model that was adopted employs the state parameter framework and inherently accounts for the different values of D_R without the need to adjust the model parameters. The HS-Small parameters employed to simulate the response of the PISA representative offshore homogenous sand sites are listed in Table 3.11. Figure 3.16(a) illustrates the dependency of the effective friction angle ϕ' with the change of relative density D_R for the Dunkirk sand profiles.

Table 3.11. HS-Small parameters used for the PISA representative offshore homogenous sand sites.

Parameter	**D_R=45%**	**D_R=60%**	**D_R=90%**	**Unit**
E_{50}^{ref}	27000	36000	54000	kPa
E_{oed}^{ref}	27000	36000	54000	kPa
E_{ur}^{ref}	81000	108000	162000	kPa
m	0.5	0.5	0.5	$(-)$
ϕ'	38	41	45.5	°
ψ	7.5	11.25	16.25	°
$\gamma_{0.7}$	0.000155	0.00014	0.00011	$(-)$
G_{0ref}	157660	171676	203627	kPa
R_f	0.94375	0.925	0.8875	$(-)$
R_{inter}	0.7095	0.6377	0.5447	$(-)$

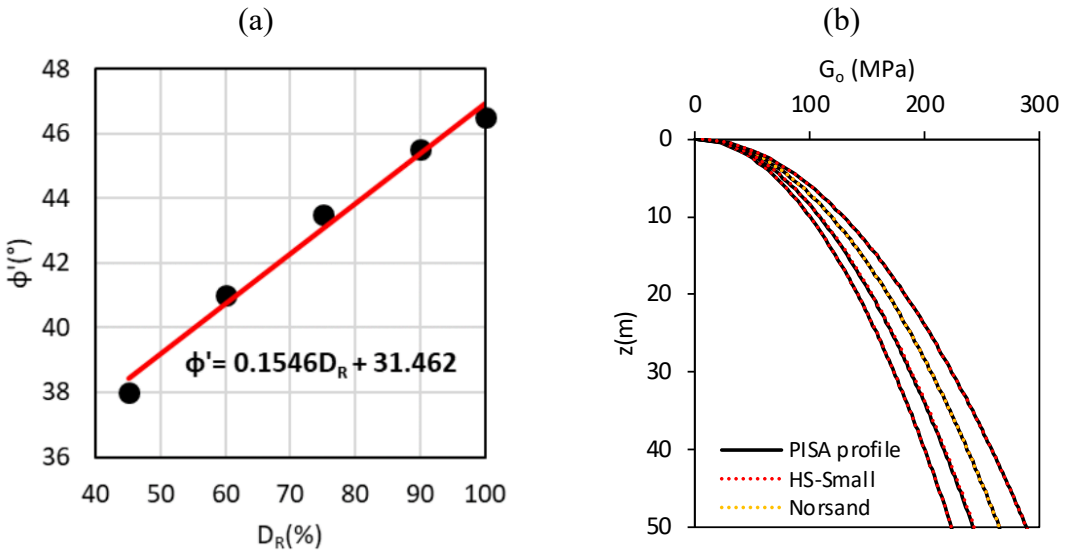


Figure 3.16: (a) Dependency of ϕ' with D_R for Dunkirk sand. The red line indicates a linear fit to the D_R-ϕ' data. (b) Simulated G_o profiles by the HS-Small and Norsand for the representative offshore Dunkirk sand sites of different relative density (referred as PISA profile in the legend).

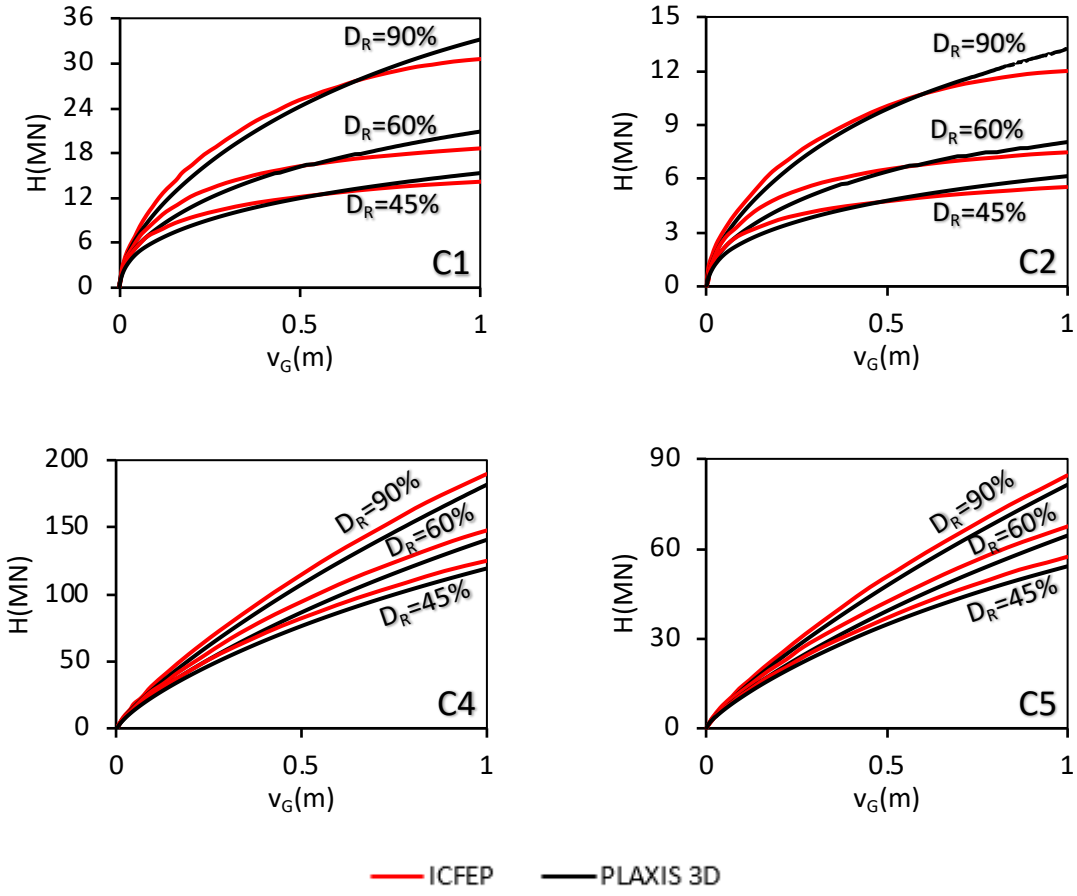


Figure 3.17. Load-displacement responses at ground level for the ultimate response of the C1, C2, C4 and C5 calibration piles (Table 3.1) embedded in the representative offshore homogenous sand sites of $D_R = 45, 60, 90\%$ computed by the 3D FE model in AWG (2018) denoted as ICFEP and the current PLAXIS 3D results.

Figure 3.17 shows the ultimate response of the C1, C2, C4 and C5 piles (Table 3.1) predicted by the calibrated HS-Small parameters presented in Table 3.11 for the PISA representative offshore homogenous sand sites of $D_R = 45, 60, 90\%$. The HS-Small results compare relatively well with the ICFEP data presented in AWG (2018) confirming the selected ϕ' values for each sand soil profile.

3.4.4 Analysis of the PISA field test piles

The selected constitutive model parameters employed for the offshore pile analyses were validated with the PISA field test piles measured response. This section provides a brief overview of these validation FEA undertaken for the PISA field test piles embedded in the Cowden test site (Byrne et al., 2020a) and the Dunkirk test site (McAdam et al., 2020). The calibration procedures for determining each constitutive model parameters are similar to the procedures presented previously in the current chapter. The water table in both the Cowden till and Dunkirk sand test sites is not located at ground level as in offshore sites. Therefore, the resultant pore water pressure (PWP) profile u of each site reported in Zdravkovic et al. (2020a) needs to be simulated and the total bulk unit weight is employed for the determination of the initial soil stresses in each layer.

Table 3.12 Field test piles geometry as reported for pile CM9 and CL1 in Byrne et al. (2020a) and for piles DM4 and DL1 in McAdam et. (2020).

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D	h/D
CM9	0.762	4	11	10	5.25	13.12
CL1	2	10.35	25	10	5.175	5
DM4	0.762	4	14	10	5.25	13.12
DL1	2	10.6	38	10	5.3	5

Cowden test site

The water table is located 1m below the ground surface (Zdravkovic et al., 2020a). However, due to the seasonal water table fluctuation the PWP profile are under-drained from 5 to 10m

depth below the ground surface. The saturated bulk unit weight is $\gamma = 21.19\ \mathrm{kN/m^3}$ (Zdravkovic et al., 2020a). The profiles of G_o, s_u, K_o and u established in the PISA project from the Cowden test site characterisation (Zdravkovic et al., 2020a) were employed for the analyses. The NGI-ADP and PIMS models are both total stress models. Hence, K_o^{total} relevant to the total stress condition was employed for the analyses calculated by $K_o^{total} = \sigma_h/\sigma_v$. Twelve sublayers were employed in PLAXIS for both constitutive models to simulate the depth variation of the G_o, s_u, K_o profiles.

Figure 3.18 illustrates the load-displacement ground level response for piles CM9, CL1 (Table 3.12) embedded in the PISA Cowden test site computed by the NGI-ADP and PIMS models. Both NGI-ADP and PIMS results indicate a good match between the analyses and the measured field test pile response in Byrne et al. (2020a). In both PLAXIS FEA a gap was formed behind the pile agreeing with the observations reported by Byrne et al. (2020a).

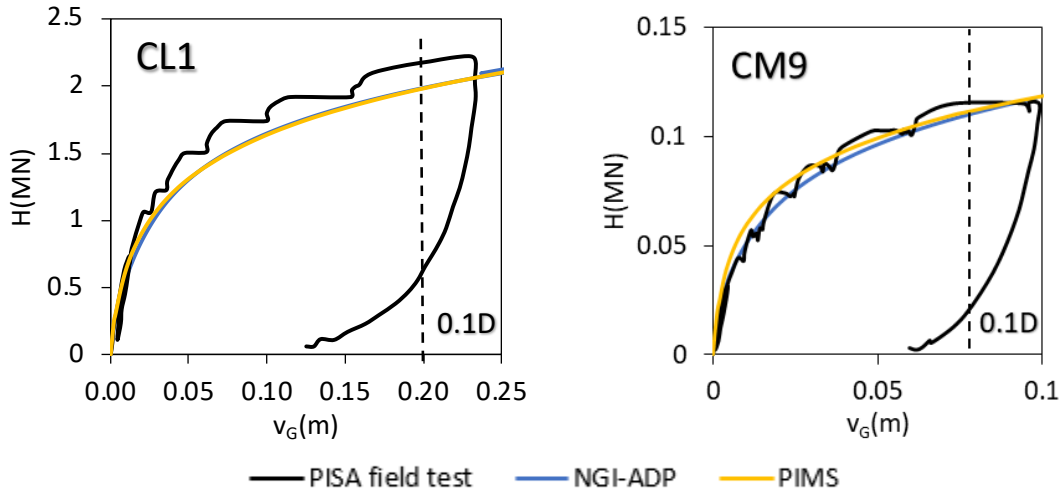


Figure 3.18: Pile load-displacement ground level response for the CL1 and CM9 piles (Table 3.12).

Dunkirk test site

The PISA Dunkirk test site comprises a top layer of hydraulic fill of $D_R = 100\%$ to a depth of 3m and a homogenous sand layer below with $D_R = 75\%$ (Zdravkovic et al., 2020a). The water table is located at a depth of 5.4m generating unsaturated conditions on the sand layers

above. Above the water table the unit bulk weight is $\gamma = 17.1\, kN/m^3$ and below the water table it is $\gamma = 19.9\, kN/m^3$ as reported in Zdravkovic et al. (2020a). K_0 is set to 0.4 for all the sand layers. To account for the increased strength generated by the unsaturated conditions an artificial cohesion has been implemented in the top layers consistent with Minga and Burd (2019b). A cohesion $c' = 10kPa$ has been employed for the HS-Small calculations for the sand layers located above a depth of 5.4m. The numerical implementation of Norsand in PLAXIS does not include a cohesion parameter hence the Norsand calculations are deemed as conservative compared to the HS-Small.

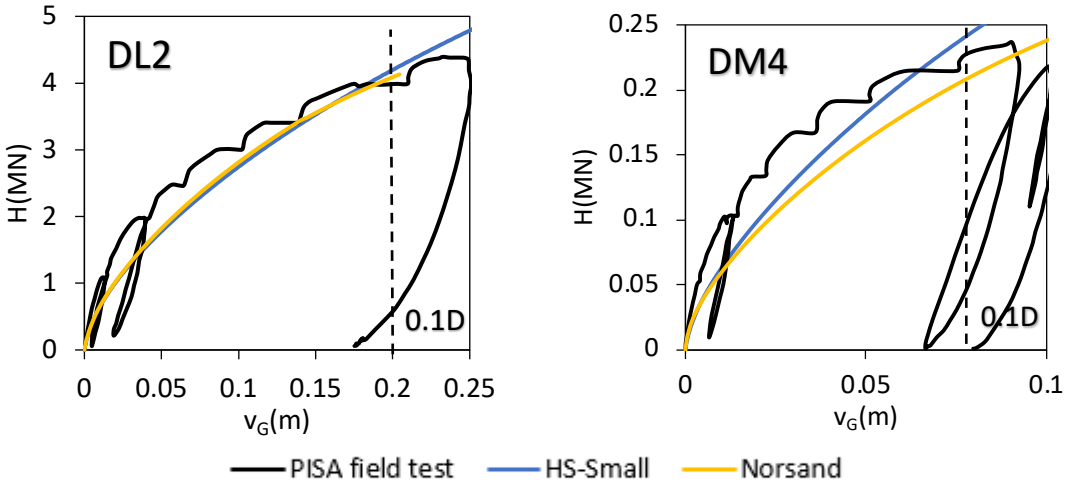


Figure 3.19: Pile load-displacement ground level response for the DL1 and DM4 piles (Table 3.12).

Figure 3.19 illustrates ground level load-displacement response of piles DL1, DM4 (Table 3.12) computed by the HS-Small and Norsand models compared to the field test response in McAdam et al. (2020). The 3D FEA results match in relatively good accuracy the field test pile response. In both FEA a gap was formed behind the pile up to the depths where unsaturated sand is found. This agrees with the findings reported in McAdam et al. (2020). In the HS-Small analysis, a friction angle of $\phi' = 46.5°$ and $\phi' = 43.5°$ is adopted for the $D_R = 100\%$ and $D_R = 75\%$ sand layers respectively. Additionally, in the Norsand analysis the initial state parameter value was set equal to $\psi_o = -0.282$ for the sand layer with $D_R = 100\%$ and $\psi_o = -0.163$ for the sand layers with $D_R = 75\%$.

3.5 Summary

This chapter presents the framework of the numerical analysis performed in this thesis for monotonic laterally loaded piles embedded in clay or sand sites. All analyses presented were performed with the PLAXIS 3D software. The selection and calibration of a constitutive model capable of delivering trustworthy results was undertaken through comparing the PLAXIS 3D results with the PISA 3D FEA results presented in Byrne et al. (2020b) and Burd et al. (2020a).

For clay soils, NGI-ADP was selected over PIMS since it is computationally less demanding and the analyses indicated that both models provide broadly similar results. NGI-ADP and PIMS are total stress models and can be used only for undrained loading. They are not critical state dependent, hence they do not provide any means of the pore water pressure evolution during undrained loading and cannot simulate accurately the stress path of the overconsolidated clays. However, they have the advantage that they use s_u as a direct input and are faster to compute than an effective stress model. Additionally, NGI-ADP and PIMS can capture well the different undrained shear strengths for different stress paths. NGI-ADP cannot capture the non-linear shear stiffness degradation from small to large strains that is well captured by PIMS.

For sand soils, HS-Small was selected over Norsand; HS-Small is computationally more stable and less demanding. The current implementation of Norsand in PLAXIS 3D produces convergency issues at large displacements when reaching near the target ground level displacement of $0.2D$ – which is required to extract the soil reaction curves. However, since HS-Small is an isotropic hardening model and is not critical state dependent as Norsand in terms of the pile's large displacement response, Norsand slightly outperforms HS-Small due to: (i) Norsand can accurately capture the post-peak softening found in the dense sand profiles, (ii) Norsand provides a more realistic dilational response.

Chapter 4

4 Data-Driven Method for Laterally Loaded Piles

4.1 Introduction

This chapter presents a new method to generate data-driven 1D design models (Suryasentana et al. 2020; Suryasentana et al. 2022) that was developed in Oxford University during the PICASO project for predicting the monotonic behaviour of laterally loaded monopiles. The data-driven 1D design models work as an alternative design approach to the PISA rule-based design models developed by Byrne et al. (2020b) and Burd et al. (2020a). The data-driven 1D design models employ a ML model to formulate the soil reaction curve data that are required to drive the 1D FE framework (Figure 1.8(b)) developed in the PISA project to predict the monopile's monotonic behaviour. The 1D FE software OxPile (implemented in MATLAB) provided by Prof. Burd was employed to utilise the 1D FE framework. The ML models employed in the data-driven 1D design models can be categorised as supervised ML models, where the ML model attempts to find the relationship that links the inputs to the outputs. The pile geometry and soil profile characteristics are employed as the ML model training inputs and the soil reaction curves data comprise the model outputs. The data-driven 1D design models enable the 1D FE framework developed in the PISA project (Byrne et al., 2020; Burd et al., 2020) to adapt relatively easily to new soil profiles and pile geometries

and to incorporate a high-fidelity representation of the soil reaction curves as computed from the 3D FEA. The key stages in the data-driven method are summarised in Figure 4.1.

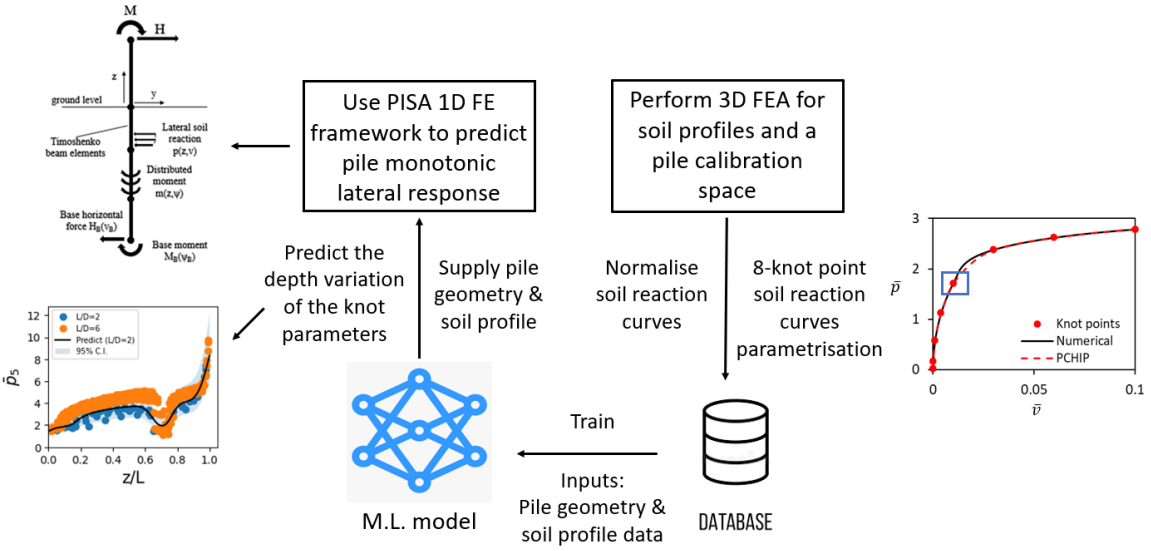


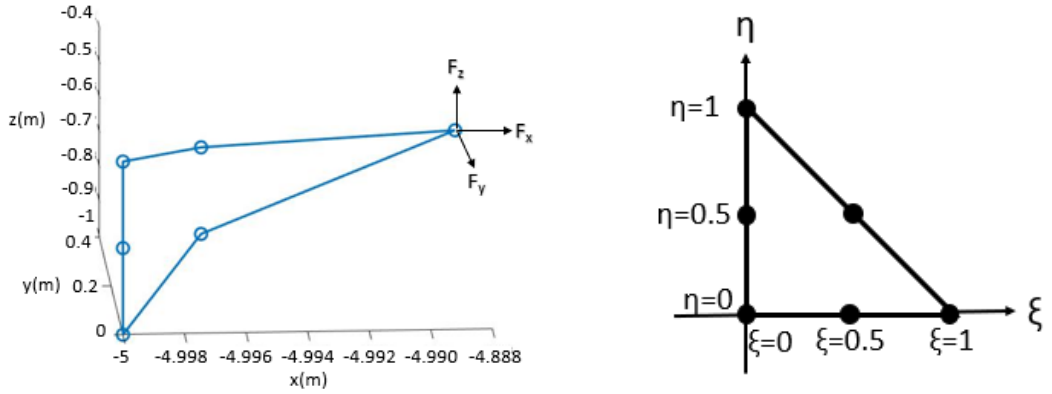
Figure 4.1: Key stages in the generation of a data-driven 1D design model to predict the monotonic behaviour of a laterally loaded monopile. Each stage is analysed in the following sections.

4.2 Soil reaction curves extraction

This section describes the procedure followed after each successful calibration FEA in PLAXIS 3D to determine sets of numerical soil reaction curves for the four soil reaction components acting on the pile, illustrated in Figure 1.8(a). A bespoke code has been developed that uses the computed tractions at the interface element Gauss points and the displacements at the plate element nodes from each calculated step in PLAXIS to compute the soil reaction curves.

4.2.1 PLAXIS 3D interface element formulation

The pile/soil interface is represented by 12-noded triangular elements; the 12 nodes are arranged in pairs of 6-noded triangular elements (Figure 4.2(a)). The distance between two nodes of a node pair is zero. The shape functions and their derivatives with respect to the local coordinates (ξ, η) are presented in Table 4.1. The numerical integration uses 6 Gauss points; the position and weights of the Gauss points are presented in Table 4.2.



(a) (b)

Figure 4.2: (a) Example 6-noded pair PLAXIS interface element in global coordinates. (b) Local numbering and nodes positions in local coordinates (ξ, η) of 6-noded pair triangular element (Brinkgreve et al., 2018).

Table 4.1: Shape functions and shape function derivatives for the 6-noded pair elements (Brinkgreve et al., 2018).

Shape functions	**Shape functions derivatives**	
$N_1 = (1-\xi-\eta)(1-2\xi-2\eta)$	$\frac{\partial N_1}{\partial \xi} = -3+4\xi+4\eta$	$\frac{\partial N_1}{\partial \eta} = -3+4\xi+4\eta$
$N_2 = \xi(2\xi-1)$	$\frac{\partial N_2}{\partial \xi} = -1+4\xi$	$\frac{\partial N_2}{\partial \eta} = 0$
$N_3 = \eta(2\eta-1)$	$\frac{\partial N_3}{\partial \xi} = 0$	$\frac{\partial N_3}{\partial \eta} = -1+4\eta$
$N_4 = 4\xi(1-\xi-\eta)$	$\frac{\partial N_4}{\partial \xi} = 4-8\xi-4\eta$	$\frac{\partial N_4}{\partial \eta} = -4\xi$
$N_5 = 4\xi\eta$	$\frac{\partial N_5}{\partial \xi} = 4\eta$	$\frac{\partial N_5}{\partial \eta} = 4\xi$
$N_6 = 4\xi(1-\xi-\eta)$	$\frac{\partial N_6}{\partial \xi} = -4\eta$	$\frac{\partial N_6}{\partial \eta} = 4-4\xi-8\eta$

Table 4.2: 6-point Gaussian integration for triangular interface elements (Brinkgreve et al., 2018).

Point	$\boldsymbol{\xi_i}$	$\boldsymbol{\eta_i}$	$\boldsymbol{w_i}$
1	0.091576	0.816848	0.109952
2	0.091576	0.091576	0.109952
3	0.816848	0.091576	0.109952
4	0.108103	0.445948	0.223382
5	0.445948	0.108103	0.223382
6	0.445948	0.445948	0.223382

4.2.2 Soil reaction curves extraction procedure

Data were obtained from each monopile slice and from the base surface on the computed tractions at the interface element Gauss points and the displacements at the plate element nodes. This information was used to determine sets of numerical soil reaction curves for the four soil reaction components (Figure 1.8(a)) from each of the calibration FEA. The thickness of each slice is related to the total length of the pile; for offshore monopile geometries the thickness is approximately 1m; for field test monopiles the selected thickness is required to generate at least 15 soil reaction curves for both the distributed lateral load (p) and moment (m) reactions along the monopile. The displacements (v, v_B) are calculated by averaging the values of lateral displacement for each slice/base surface. The rotations (ψ, ψ_B) are determined by employing a least-square linear fit to the computed vertical displacements of each slice/base surface. The distributed lateral load reaction p (units of force / length) for each slice is calculated by summing all the resultant horizontal F_x nodal forces and by dividing the total sum by the thickness of the slice. The distributed moment reaction m (units of force × length / length) for each slice is computed by summing all the resultant vertical F_z nodal forces times the lateral distance from the axis of symmetry; the final product is divided by the thickness of the slice. The base horizontal force reaction H_B (units of force) is calculated by summing all the resultant horizontal F_x nodal forces of the base interface. The base moment reaction M_B (units of forces × length) is computed by summing all the resultant vertical F_z nodal forces of the base interface times the lateral distance from the axis of symmetry.

The interface element resultant nodal forces (F_x, F_y, F_z) are computed from the extracted tractions (σ_n, τ_1, τ_2) at the interface element Gauss points. These nodal forces are calculated using the following formula from Segurado and LLorca (2004):

$$\boldsymbol{f}_N{}^{\mathrm{el}} = \int_0^1 \int_0^1 \boldsymbol{N}^{\mathrm{T}} \boldsymbol{R}^{\mathrm{T}} \boldsymbol{t}_{\mathrm{loc}} J d\xi d\eta = \sum_j w_j \boldsymbol{N}^{\mathrm{T}} \boldsymbol{R}^{\mathrm{T}} \boldsymbol{t}_{\mathrm{loc}} J \tag{4.1}$$

where $\boldsymbol{N}$ is a 3×18 matrix used for the shape function matrix; w_j are the Gauss point weights presented in Brinkgreve et al. (2018) (see Table 4.2); $\boldsymbol{R}$ is the rotational matrix; $\boldsymbol{t}_{\mathrm{loc}}$ is a 3×1 vector representing the computed tractions (one normal and two tangential) and J is the Jacobian of the transformation between local and global coordinate systems calculated using the formula from Nikishkov (2010). The calculated nodal force values are then multiplied by 2 due to symmetry, as only half of the model is simulated in the 3D FEA.

To integrate the computed tractions over the element area, the outward normal vector at any point of the element needs to be determined. The tangent vectors $\boldsymbol{e}_\xi$ and $\boldsymbol{e}_\eta$ in the local coordinate system ξ and η, respectively, have components in the global coordinate system:

$$\boldsymbol{e}_\xi = \left\{\frac{\partial x}{\partial \xi} \;\; \frac{\partial y}{\partial \xi} \;\; \frac{\partial z}{\partial \xi}\right\} \qquad \boldsymbol{e}_\eta = \left\{\frac{\partial x}{\partial \eta} \;\; \frac{\partial y}{\partial \eta} \;\; \frac{\partial z}{\partial \eta}\right\} \tag{4.2}$$

The derivatives of global coordinates with respect to local coordinates can be calculated using shape function derivatives presented in Table 4.1:

$$\frac{\partial x}{\partial \xi} = \sum \frac{\partial N_m}{\partial \xi} x_m \qquad \frac{\partial y}{\partial \xi} = \sum \frac{\partial N_m}{\partial \xi} y_m \qquad \frac{\partial z}{\partial \xi} = \sum \frac{\partial N_m}{\partial \xi} z_m \tag{4.3}$$

$$\frac{\partial x}{\partial \eta} = \sum \frac{\partial N_m}{\partial \eta} x_m \qquad \frac{\partial y}{\partial \eta} = \sum \frac{\partial N_m}{\partial \eta} y_m \qquad \frac{\partial z}{\partial \eta} = \sum \frac{\partial N_m}{\partial \eta} z_m \tag{4.4}$$

The normal vector $\boldsymbol{n}$ to the surface presented in Fig. 4.3 is obtained by the vector product of the tangential vectors $\boldsymbol{e}_\xi$ and $\boldsymbol{e}_\eta$:

$$\boldsymbol{n} = \boldsymbol{e}_\xi \times \boldsymbol{e}_\eta \tag{4.5}$$

The components of the normal vector can be computed using:

$$n_x = \frac{\partial y}{\partial \xi}\frac{\partial z}{\partial \eta} - \frac{\partial z}{\partial \xi}\frac{\partial y}{\partial \eta} \qquad n_y = \frac{\partial z}{\partial \xi}\frac{\partial x}{\partial \eta} - \frac{\partial x}{\partial \xi}\frac{\partial z}{\partial \eta} \qquad n_z = \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial y}{\partial \xi}\frac{\partial x}{\partial \eta} \tag{4.6}$$

The Jacobian, J, of transformation between global Cartesian coordinates and local coordinates is determined as:

$$J = \sqrt{{n_x}^2 + {n_y}^2 + {n_z}^2} \tag{4.7}$$

Once the reference surface is established, the local coordinate system in the current configuration is defined by three perpendicular vectors, each with a unit modulus, represented by:

$$\hat{\boldsymbol{n}} = \frac{1}{\|\boldsymbol{e}_\xi \times \boldsymbol{e}_\eta\|}(\boldsymbol{e}_\xi \times \boldsymbol{e}_\eta);\ \hat{\boldsymbol{e}}_\xi = \frac{1}{\|\boldsymbol{e}_\xi\|}\boldsymbol{e}_\xi;\ \hat{\boldsymbol{e}}_\eta = \hat{\boldsymbol{n}} \times \hat{\boldsymbol{e}}_\xi \tag{4.8}$$

where $\hat{\boldsymbol{n}}$ is the unit vector of the normal direction and $\hat{\boldsymbol{e}}_\xi$, $\hat{\boldsymbol{e}}_\eta$ are the unit vectors of the first and second tangential directions to the surface presented in Figure 4.3 (Segurado and LLorca, 2004). The rotation matrix $\boldsymbol{R}$ is determined as follows:

$$\boldsymbol{R} = \begin{Bmatrix} \hat{\boldsymbol{n}}^{\mathrm{T}} \\ \hat{\boldsymbol{e}}_\xi \\ \hat{\boldsymbol{e}}_\eta \end{Bmatrix} \tag{4.9}$$

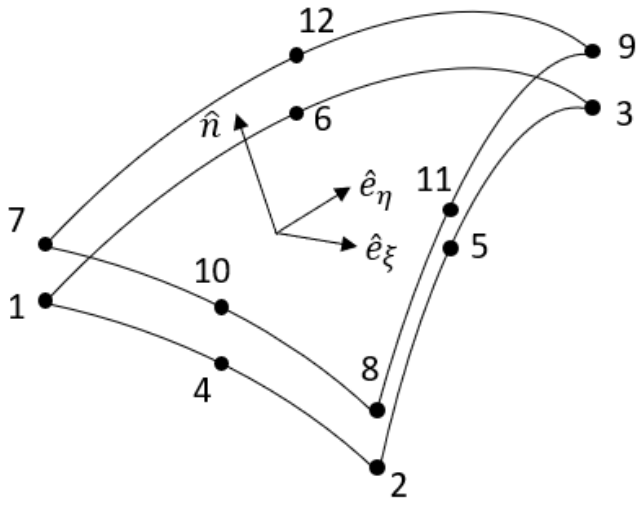


Figure 4.3: Zero thickness interface elements with the normal $\hat{\boldsymbol{n}}$ and the first $\hat{\boldsymbol{e}}_\xi$ and second $\hat{\boldsymbol{e}}_\eta$ tangential directions (adopted by Segurado and LLorca, 2004).

4.2.3 Boundary equilibrium check on 3D model convergence

After every completed FEA, the extracted soil reaction curves are checked for every successful step of the FEA to verify that the computed nodal forces in the pile are in equilibrium. The total applied horizontal load F_x to the pile needs to be equal (to within an acceptable tolerance) to the total sum of the horizontal distributed lateral reaction p and base shear force H_B. The total moment load applied at ground level calculated as $M = F_x h$ (where h is the load eccentricity) needs to be equal (to within an acceptable tolerance) to the resultant bending moment of the pile at ground level calculated by the computed soil reaction curves. The overall convergence is affected by the presence of a gap in the interface elements (for clay layers) behind the pile; for every FEA the average error for the computed forces and moments for all the successful implemented steps should be below 5%. Otherwise, particularly tight tolerances are required which would severely increase the computational demand of the analyses.

4.3 Data processing of soil reaction curves

The soil reaction curves employed in the PISA 1D FE framework are required to: (i) retain positive values of forces/moments and displacements/rotations, (ii) be arranged to retain a monotonic response, so that a given displacement corresponds to a single load (AWG 2018). Therefore, the absolute values of the extracted soil reactions curves are employed. However, at some depths near the pivot point of the pile, the displacements/rotations are not organised monotonically; this behavior is exacerbated for longer piles that have a more flexible deflection mode. The absolute soil reaction curve data that include non-monotonic displacements/rotation values, are restructured to obtain an increasing monotonic response maintaining the initial stiffness gradient up to the following force value from which the maximum non-monotonic displacement is encountered (Figure 4.4). Finally, the soil

reaction curves extracted from the monopile slice near the pivot point of the pile are disregarded.

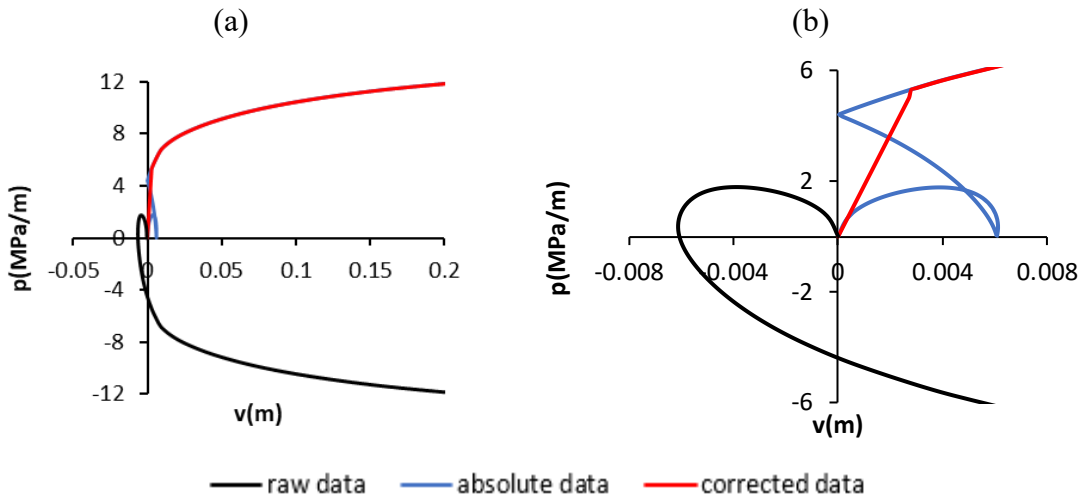


Figure 4.4: Corrected distributed lateral load *p* soil reaction curve slightly above the pile pivot point: (a) large displacement response, (b) small displacement response.

4.4 Normalisation framework of soil reaction curves

The data-driven 1D design models adopt the normalisation framework presented in Table 4.3. This system differs from the PISA approach (Byrne et al. 2020b; Burd et al. 2020a) which seeks to map all the soil reaction curves extracted from different pile geometries to a single function. In the data-driven normalisation, the soil strength and stiffness parameters are excluded from the kinematic non-dimensional forms and are instead incorporated in the ML model via the training features. Suryasentana et al. (2021) proposed a normalisation framework for clay and sand soils where: (i) the normalised lateral displacement and pile rotation only incorporate kinematic variables to bound their normalised values between 0 and 0.1 which is favourable for a machine learning implementation, (ii) the normalised distributed moment is uncoupled from the distributed load, allowing the possibility of non-zero distributed moment with zero distributed load and hence easing the machine learning implementation. The 1D models that are calibrated directly for layered soils require a single normalisation approach consistent for all soil types to capture the difference in the

normalised force/moment values between the different soil layers. Therefore, the 'sand' normalisation was designated for this task (s_u is equal to zero for sands) as it is applicable to all range of soil types (clay, sand, chalk).

Table 4.3**:** Non-dimensional forms adopted in the data-driven approach where σ'_{vi} represents the local effective stress, s_u denotes the local undrained shear strength, G_o signifies the local small strain shear stiffness and D is the pile's diameter.

Soil Reactions	**Data-driven normalisation**		**PISA normalisation**	
Soil Types	**'clay'**	**'sand'**	**'clay'**	**'sand'**
Distributed lateral load, $\bar{p}$	$\frac{p}{s_u D}$	$\frac{p}{\sigma'_{\mathrm{vi}} D}$	$\frac{p}{s_u D}$	$\frac{p}{\sigma'_{\mathrm{vi}} D}$
Lateral displacement, $\bar{v}$	$\frac{v}{D}$	$\frac{v}{D}$	$\frac{v}{D} \times \frac{G}{s_u}$	$\frac{v}{D} \times \frac{G}{\sigma'_{\mathrm{vi}}}$
Distributed moment, $\bar{m}$	$\frac{m}{s_u D^2}$	$\frac{m}{\sigma'_{\mathrm{vi}} D^2}$	$\frac{m}{s_u D^2}$	$\frac{m}{pD}$
Pile rotation, $\bar{\psi}$	ψ	ψ	$\psi \times \frac{G}{S_u}$	$\psi \times \frac{G}{\sigma'_{\mathrm{vi}}}$
Base horizontal force, $\bar{H}_B$	$\frac{H_B}{s_u D^2}$	$\frac{H_B}{\sigma'_{\mathrm{vi}} D^2}$	$\frac{H_B}{s_u D^2}$	$\frac{H_B}{\sigma'_{\mathrm{vi}} D^2}$
Base moment, $\bar{M}_B$	$\frac{M_B}{s_u D^3}$	$\frac{M_B}{\sigma'_{\mathrm{vi}} D^3}$	$\frac{M_B}{s_u D^3}$	$\frac{M_B}{\sigma'_{\mathrm{vi}} D^3}$

4.5 Eight 'knot point' soil reaction parametrisation

The normalised soil reaction curves are used to define an eight-knot point piecewise parameter spline function of the form $\bar{y} = g(\bar{x})$ introduced in Suryasentana et al. (2021) and presented in Table 4.4; where $\bar{x}$ is the normalised displacement or rotation value and $\bar{y}$ is the normalised force or moment value of each soil reaction curve.

Table 4.4: Eight-knot point piecewise soil reaction curves formulation.

0	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$	$\bar{y}_7$	$\bar{y}_8$
$\bar{x} = 0$	$\bar{x} =$ 0.00004	$\bar{x} =$ 0.0001	$\bar{x} =$ 0.001	$\bar{x} =$ 0.004	$\bar{x} =$ 0.01	$\bar{x} =$ 0.03	$\bar{x} =$ 0.06	$\bar{x} = 0.1$

This function requires the specification of eight-knot point force-moment parameters which represent the values of $\bar{y}$ at eight predefined normalised displacement values $\bar{x}$. These predefined displacement values were selected arbitrarily after a trial and error procedure as they were found to capture well the non-linearity of the PISA numerical soil reactions extracted by the ICFEP software for different soil types (clay, sand) (Suryasentana et al., 2021). The same conclusion was also drawn by the PLAXIS 3D FEA data that were undertaken in this thesis for clay, sand and chalk soil profiles. The first predefined displacement value has been modified from a value of $\bar{x} = 0.00001$ in Suryasentana et al. (2021) to $\bar{x} = 0.00004$. This change was applied to push the first predefined displacement at the upper part of the pile to acquire a higher value than the displacement calculated by the first step in PLAXIS 3D. The values of the knot point forces/moments at each predefined displacement were obtained by linear interpolating between the normalised values of each numerical soil reaction curve. Some of the numerical soil reactions near the pile pivot point are not mobilised up to $\bar{x} = 0.1$. In these cases, extrapolation is applied by setting the force/moment value equal to the last calculated numerical knot point value. This procedure is considered conservative when the soil reaction curves exhibit hardening behavior (e.g. Figure 4.5(c)), but it is deemed unconservative when the soil reaction curves show softening behavior, as illustrated in Figure 4.5(b). The Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation framework is used then to define the points between the initial pre-selected eight-knot point soil reaction curves data (Suryasentana et al., 2021). Figure 4.5 illustrates the PCHIP representation of three different parametrised eight-knot point soil reaction curves. The current parametrisation can reproduce with great accuracy the numerical soil reaction curves; the PCHIP interpolation is able to capture softening effects; this feature is not reproduced by the PISA 4-parameter conic function (Byrne et al., 2020b; Burd et al.; 2020a).

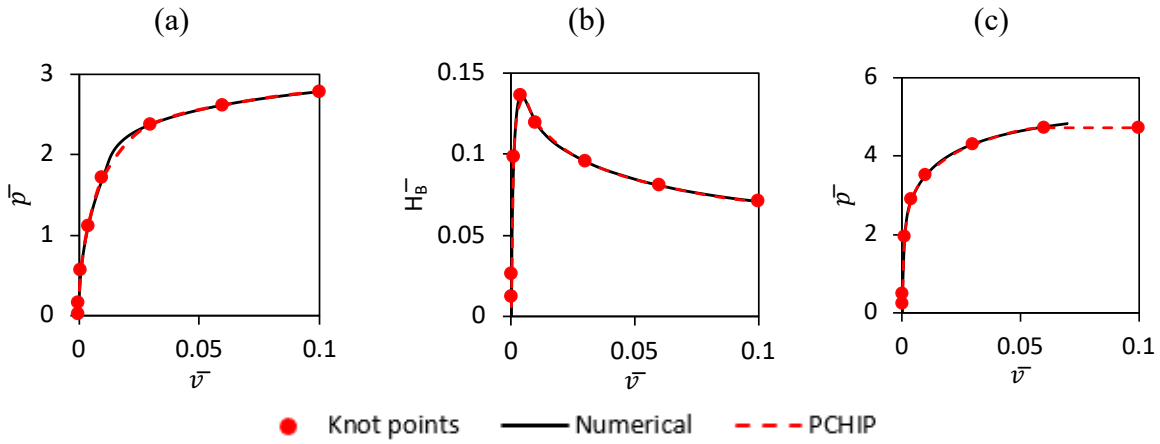


Figure 4.5: Examples of parametrised eight-knot point soil reaction curves with PCHIP interpolation. (a) typical form of soil reaction curve, (b) soil reaction curve with softening effects, (c) soil reaction curve near the rotation point (not mobilised beyond $\bar{x} = 0.1$) showing the extrapolation procedure that was applied.

4.6 Machine learning model selection

Each soil reaction curve in the data-driven method is formulated as $\bar{y} = g(\bar{x})$ where $\bar{x}$ is a displacement/rotation variable and $\bar{y}$ is a load/moment variable. The values $\bar{y}_i$ at each of the knot points are a function of the engineering features f_j that are selected to represent the problem. The soil reaction curves are therefore implicit functions of the relevant displacement/rotation variable and the features as $\bar{y} = g(\bar{x}|f_1, f_{2,\dots}, f_n)$ where n is the number of features being employed. To generate a prediction model for unseen cases a suitable machine learning (ML) regression model is required that can accurately capture the complexity of the data and have good generalisation properties.

This section explores the suitability of the Gaussian Process Regression (GPR) and the feedforward neural network (FNN) regression ML models. The performance of both models was assessed in terms of generating a site-specific data-driven 1D design model for: (i) the PISA representative offshore glacial clay till site (Byrne et al., 2020b), (ii) the PISA representative offshore homogenous sand site with a relative density D_R of 75% (Burd et al., 2020a). The soil reaction curves data used to train the ML models were extracted from the PLAXIS 3D calibration FEA presented in Sections 3.4.1 and 3.4.2 for the piles listed in

Table 3.1. Both ML models were implemented using the open-source package Scikit-learn Python library (Pedregosa et al., 2011). An independent ML model is trained for each of the eight-knot point parameters for each soil reaction component ($\bar{p}$, $\bar{m}$, $\bar{H}_B$, $\bar{M}_B$). The data-driven normalisation (Table 4.3) was employed for the clay and sand soils accordingly. The dataset of independent knot points was split randomly assigning 80% of it as the training set, leaving the rest 20% as the test set. The training set was used to calibrate the ML model parameters while the test set was used to evaluate the generalisation properties of the models on unseen data. To evaluate the performance of the trained ML models the mean square error (MSE) and the coefficient of determination (R^2) metrics are employed. The MSE is characterised by:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.10}$$

where y_i is the actual value of the ith sample in the dataset and $\hat{y}_i$ is the corresponding predicted value from the ML model. Small values of MSE means that the ML model is performing well.

The R^2 measures how well the predictions describe the real values and is defined by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.11}$$

where $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is the sum of squares of the predicted errors; $\sum_{i=1}^{n}(y_i - \bar{y})^2$ is the sum of the squared difference between the predictions and the actual values mean $\bar{y}$ which is calculated by $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. R^2 values usually range from 0 to 1; higher R^2 values indicate a better match between the predictions and the observations.

4.6.1 Feature selection for clay/sand soil profiles

A systematic procedure is employed to select the appropriate training inputs (engineering features f_j) to describe the problem. In this thesis the training inputs are presented in square brackets (e.g. $[f_j]$). The key geometric variables defining the problem are (L, D, t_w, h, z) (where L is the pile's length; D is the pile's diameter; t_w is the pile's wall thickness; h is the load eccentricity and z in the depth below the ground level). These variables are grouped into a set of potential dimensionless quantities $[t_w/D]$, $[h/D]$, $[z/D]$, $[z/L]$, $[L/D]$ referred to as synthetic features in machine learning terminology. To identify the features that have the most influence on the computed soil reaction curves and to exclude any redundant features, feature selection was conducted between $[t_w/D]$, $[h/D]$, $[z/D]$, $[z/L]$, $[L/D]$ using multi-task LASSO regression (Obozinski et al., 2006) for the distributed lateral load and moment $\bar{p}$, $\bar{m}$ and base $\bar{H}_B$, $\bar{M}_B$ knot point parameters. The multi-task LASSO regression is typically used to identify a common set of features that are relevant to the simultaneous predictions of multiple tasks (Obozinski et al., 2006). A similar task was employed in Suryasentana et al. (2020) to identify the key geometric features corresponding to the calibration parameters that have been previously developed for the PISA design model.

To prevent the multi-task LASSO regression from being distorted by the varying magnitudes of each feature, all features are standardised using the following formula:

$$x_{st} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4.12}$$

where x represents any value of the features and x_{max} and x_{min} represent the maximum and minimum value of x. The $L1$ regularisation (Tibshirani, 1996) in multi-task LASSO regression encourages sparsity in the weights associated with each feature across multiple prediction tasks. The weights associated with each standardised feature are obtained using

10-fold cross validation. The closer these weights are to zero, the less relevant the features are for the prediction tasks.

For brevity, Table 4.5 presents the average value of the weights (the modulus of each weight was used) determined by multi-task LASSO regression for $\bar{p}$, $\bar{m}, \bar{H}_B$ and $\bar{M}_B$ knot point parameters corresponding to: (i) the PISA representative offshore glacial clay till site (Byrne et al., 2020b), (ii) the PISA representative offshore homogenous sand site with a relative density D_R of 75% (Burd et al., 2020a). The most informative set of features (having high values of weights) are $[z/D]$, $[z/L]$, $[L/D]$ for the distributed soil reactions and $[L/D]$ for the base soil reactions. Both $[z/D]$ and $[z/L]$ describe the depth variation of the knot point parameters; hence only one is required for the training procedure of the ML models. The $[z/L]$ input has been selected over the $[z/D]$ to describe the depth variation of the knot point soil reaction curve parameters as it has a standardised range of 0 to 1 which is preferable for a ML model. Therefore, $[z/L]$ and $[L/D]$ are employed as training inputs for $\bar{p}$, $\bar{m}$ and $[L/D]$ is employed as a training input for $\bar{H}_B$ and $\bar{M}_B$. The weights determined by multi-task LASSO regression for $[t_w/D]$ and $[h/D]$ were close to zero for most of the knot point parameters; hence they are disregarded from the training procedure of the ML models.

Table 4.5: Average values of weights determined by multi-task LASSO regression for the knot point parameters associated with the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75%. The average weights values presented in Table 4.5 were rounded to the second digit.

Soil Profile	**Knot-points**	z/L	z/D	L/D	h/D	t_w/D
PISA glacial clay till	$\bar{p}$	0.28	0.05	0.09	≈0	≈0
PISA glacial clay till	$\bar{m}$	0.11	0.30	0.18	≈0	≈0
PISA glacial clay till	$\bar{H}_B$	≈0	≈0	0.22	0.03	≈0
PISA glacial clay till	$\bar{M}_B$	≈0	≈0	0.35	0.05	≈0
Sand D_R=75%	$\bar{p}$	0.58	0.29	0.18	0.01	≈0
Sand D_R=75%	$\bar{m}$	0.62	0.30	0.19	0.01	≈0
Sand D_R=75%	$\bar{H}_B$	≈0	≈0	0.66	≈0	≈0
Sand D_R=75%	$\bar{M}_B$	≈0	≈0	0.7	≈0	≈0

Additional features need to be specified to represent the soil conditions. A distinction needs to be made between site-specific models (i.e. a calibration that is specific to a particular site with a single set of soil strength and stiffness profile) or generalised models (in which the soil strength and stiffness profiles span a range of potential sites). To develop site-specific models, additional features $[G_o/s_u]$ for clay and $[G_o/\sigma'_v]$ for sand are adopted (Suryasentana et al., 2021); these features provide a means of incorporating soil stiffness into the model and discriminate locally the soil reaction curves with the same values of z/L and L/D. For generalised models trained for a range of different soil profiles, additional features are required to represent the 'relative strength' of the soil (Suryasentana et al., 2021). For the generalised clay models $[\sigma'_v/s_u]$ is employed to discriminate between clays with different OCR profiles. For the generalised sand models $[D_R]$ is employed to discriminate between sands with different relative density.

4.6.2 Gaussian Process Regression training procedure

The GPR model properties are presented in detail in Section 2.4.1. A zero-mean function $m(x) = 0$ is adopted (where x is a training input); the kernel function is adequate to capture the complexity of the data. The covariance $r(\boldsymbol{x}, \boldsymbol{x}')$ in the kernel is specified by:

$$r(\boldsymbol{x}, \boldsymbol{x}') = \sqrt{\sum_{i=1}^{d} \frac{(x_i - x'_i)^2}{l_i^2}} \tag{4.13}$$

where l_i is the characteristic length scale and x_i, x'_i are the training and test inputs for the ith input dimension. If l_i is large all the data points in the ith dimension are correlated with each other (resulting into a flatter function), whereas if l_i is small only the data points that are close to each other are correlated (resulting into a function that exhibits rapid variations over small distances). Suryasentana et al. (2020) showed that the Matérn kernel ($\nu = 5/2$) marginally outperformed the Squared Exponential (SE) kernel in terms of the R^2 metrics for

the highly non-linear distribution of the PISA 4-parameters conic function data. After using both kernels in the current work the same conclusion was drawn. Therefore, it was decided to employ the Matérn kernel with $\nu = 5/2$ for this study. Figure 4.6 shows the depth variation predictions corresponding to the $\bar{p}_8$ knot point parameter for piles embedded in the PISA representative offshore glacial clay till site; both kernels capture the high non-linearity of the data. The R^2 train score and test score are 0.92 and 0.91 when using the Matérn kernel and 0.88 and 0.87 when employing the SE kernel. The Automatic Relevance Determination (ARD) is adopted during the hyperparameters optimisation, meaning that different l_i for each input dimension can be employed. The σ_f (signal variance) and l_i hyperparameters are optimised by maximising the log marginal likelihood. The white noise kernel $k_{WN}(\boldsymbol{x}, \boldsymbol{x}')$ is added to the Matérn kernel (Equation (2.13) in Section 2.4.1) to account for possible noise in the dataset:

$$k_{WN}(\boldsymbol{x}, \boldsymbol{x}') = \begin{cases} \sigma_n^2 & : if\ \boldsymbol{x} = \boldsymbol{x}' \\ 0 & : if\ \boldsymbol{x} = 0 \end{cases} \tag{4.14}$$

where σ_n^2 is the variance of the noise.

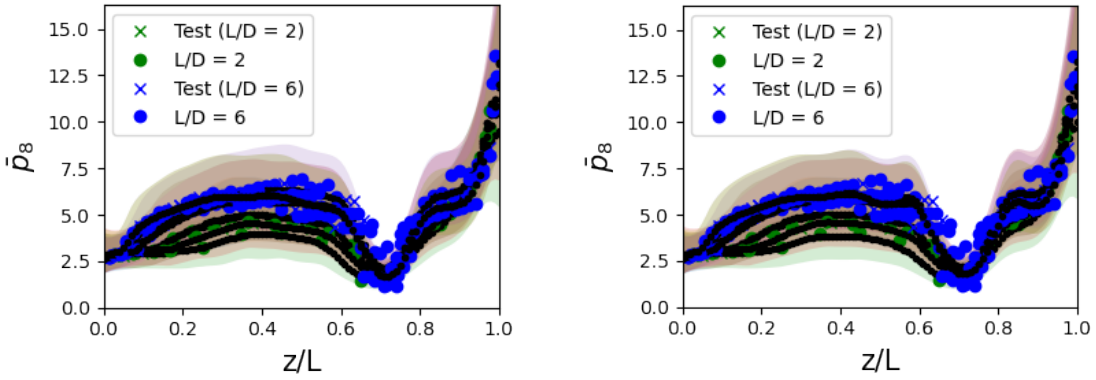


Figure 4.6: Example depth variation of the mean GPR predictions employing (a) the Matérn kernel, (b) the Squared exponential kernel, for the $\bar{p}_8$ distributed lateral load knot point parameter for piles embedded in the PISA representative offshore glacial clay till site. The mean GPR predictions are indicated by the black point markers. The 95% confidence interval is shown (shaded bounds). The green and blue point markers are the trained datapoints and the green and blue cross markers are the test datapoints extracted from the calibration FEA.

4.6.2.1 Assessment of GPR model

Table 4.6, Table 4.7 illustrate the optimised hyperparameters obtained for the knot point soil reaction parameters after the GPR models training for the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75%. Some of the knot point parameters for the $[G_o/s_u]$ and $[G_o/\sigma'_v]$ inputs have a lengthscale of the defined maximum limit of 1000. This indicates that the inputs are irrelevant to the final prediction. Table 4.8 to Table 4.10 show the MSE and R^2 scores for the trained GPR models. The similarity in the achieved training and test metric values indicate that the GPR models have good generalisation properties – the ML models can be used for unseen inputs inside the calibration space. The high train and test R^2 scores indicate that the GPR models can accurately capture the complexity of the data. However, some of the GPR models trained for the base soil reaction parameters provide negative R^2 values. These R^2 values are not reliable due to the small dataset (9 points) employed for each model.

The predictions – with the pile geometry, soil profile properties as inputs – provide: (i) the depth variation of the knot point parameters of the distributed lateral load $\bar{p}$ and distributed moment $\bar{m}$; (ii) the point predictions of the knot point parameters for the base horizontal force $\bar{H}_B$ and base moment $\bar{M}_B$. Figure 4.7 to Figure 4.10 present the GPR models' predictions for the calibration piles listed in Table 3.1 (with $L/D = 2, 6$) embedded in the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site of $D_R = 75\%$. The GPR predictions are compared with the training and test data. The 95% confidence intervals (shown as shaded bounds) provided by the GPR models are also presented. In general, the GPR model can accurately capture the highly non-linear nature of both the clay and sand data. The 95% confidence interval is higher for the sand than the clay data due to the noisier knot point soil reaction data of the sand soil. The uncertainty associated with the additional noise is reflected by the larger confidence interval.

The 95% confidence interval becomes relatively high for the GPR models trained for the sand soil at depths where $z/L < 0.05$ due to the lack of soil reaction curves (the first soil reaction curve is always located at $z = 0.5m$). In that region the $[G_o/\sigma'_v]$ input retains extremely high values as σ'_v approaches to zero, hence the GPR model is practically extrapolating far away from its calibration space.

The GPR models can predict accurately the base soil reaction knot point data. The datapoints indicate that piles with different load eccentricity h (e.g. piles C1, C2 in Table 3.1) have a slightly different normalised $\bar{H}_B, \bar{M}_B$ knot point value. An additional training input $[h/D]$ could be introduced for the base soil reactions to account for this difference. However, for the current GPR models as it was also indicated by the multi-task LASSO regression in Section 4.6.1, the influence of the different load eccentricity in the final predictions is almost negligible, hence this effect was disregarded.

Table 4.6: Optimised Matérn kernel parameters for the GPR models as applied to the knot point soil reaction parameters associated with the PISA representative offshore glacial clay till site (σ_f is the signal variance of the Matérn kernel describing how much span the function has vertically, σ_n is the variance of the white noise kernel, l is the lengthscale for every input dimension).

Knot point	$\bar{p}$	$\bar{m}$	$\bar{H}_B$	$\bar{M}_B$
1	σ_f=1.16, σ_n^2=0.00337, $l_{z/L}$=0.262, $l_{L/D}$=16.4, l_{G_o/s_u}=1000	σ_f=2.04, σ_n^2=0.037, $l_{z/L}$=0.226, $l_{L/D}$=10.5, l_{G_o/s_u}=1000	σ_f=2.23, σ_n^2=0.0279, $l_{L/D}$=16.4, l_{G_o/s_u}=1000	σ_f=2.52, σ_n^2=0.0445, $l_{L/D}$=100, l_{G_o/s_u}=1000
2	σ_f=0.928, σ_n^2=0.00319, $l_{z/L}$=0.218, $l_{L/D}$=10.2, l_{G_o/s_u}=567	σ_f=1.75, σ_n^2=0.033, $l_{z/L}$=0.225, $l_{L/D}$=8.82, l_{G_o/s_u}=1000	σ_f=1.59, σ_n^2=0.0064, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=2.74, σ_n^2=0.837, $l_{L/D}$=100, l_{G_o/s_u}=1000
3	σ_f=0.757, σ_n^2=0.0116, $l_{z/L}$=0.2, $l_{L/D}$=11.8, l_{G_o/s_u}=1000	σ_f=1.24, σ_n^2=0.0188, $l_{z/L}$=0.496, $l_{L/D}$=10.3, l_{G_o/s_u}=1000	σ_f=0.803, σ_n^2=0.000603, $l_{L/D}$=64.6, l_{G_o/s_u}=1000	σ_f=1.38, σ_n^2=0.0145, $l_{L/D}$=72.9, l_{G_o/s_u}=1000
4	σ_f=0.842, σ_n^2=0.0156, $l_{z/L}$=0.2, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.753, σ_n^2=0.0552, $l_{z/L}$=0.295, $l_{L/D}$=11.7, l_{G_o/s_u}=1000	σ_f=0.596, σ_n^2=1.06e^{-5}, $l_{L/D}$=25.5, l_{G_o/s_u}=1000	σ_f=0.992, σ_n^2=0.0059, $l_{L/D}$=100, l_{G_o/s_u}=1000
5	σ_f=1.05, σ_n^2=0.0224, $l_{z/L}$=0.2, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.698, σ_n^2=0.00195, $l_{z/L}$=0.454, $l_{L/D}$=11.8, l_{G_o/s_u}=1000	σ_f=0.502, σ_n^2=0.000208, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.842, σ_n^2=0.00383, $l_{L/D}$=100, l_{G_o/s_u}=1000
6	σ_f=1.09, σ_n^2=0.0227, $l_{z/L}$=0.2, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.835, σ_n^2=0.00244, $l_{z/L}$=0.369, $l_{L/D}$=12.1, l_{G_o/s_u}=890	σ_f=0.475, σ_n^2=0.0004, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.724, σ_n^2=0.00209, $l_{L/D}$=100, l_{G_o/s_u}=1000
7	σ_f=1.10, σ_n^2=0.0232, $l_{z/L}$=0.2, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.832, σ_n^2=0.00254, $l_{z/L}$=0.379, $l_{L/D}$=11.7, l_{G_o/s_u}=892	σ_f=0.454, σ_n^2=0.000916, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.708, σ_n^2=0.00298, $l_{L/D}$=100, l_{G_o/s_u}=1000
8	σ_f=1.12, σ_n^2=0.0233, $l_{z/L}$=0.2 $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.835, σ_n^2=0.00254, $l_{z/L}$=0.38 $l_{L/D}$=11.8, l_{G_o/s_u}=894	σ_f=0.453, σ_n^2=0.000882, $l_{L/D}$=100, l_{G_o/s_u}=1000	σ_f=0.699, σ_n^2=0.00305, $l_{L/D}$=37.4, l_{G_o/s_u}=1000

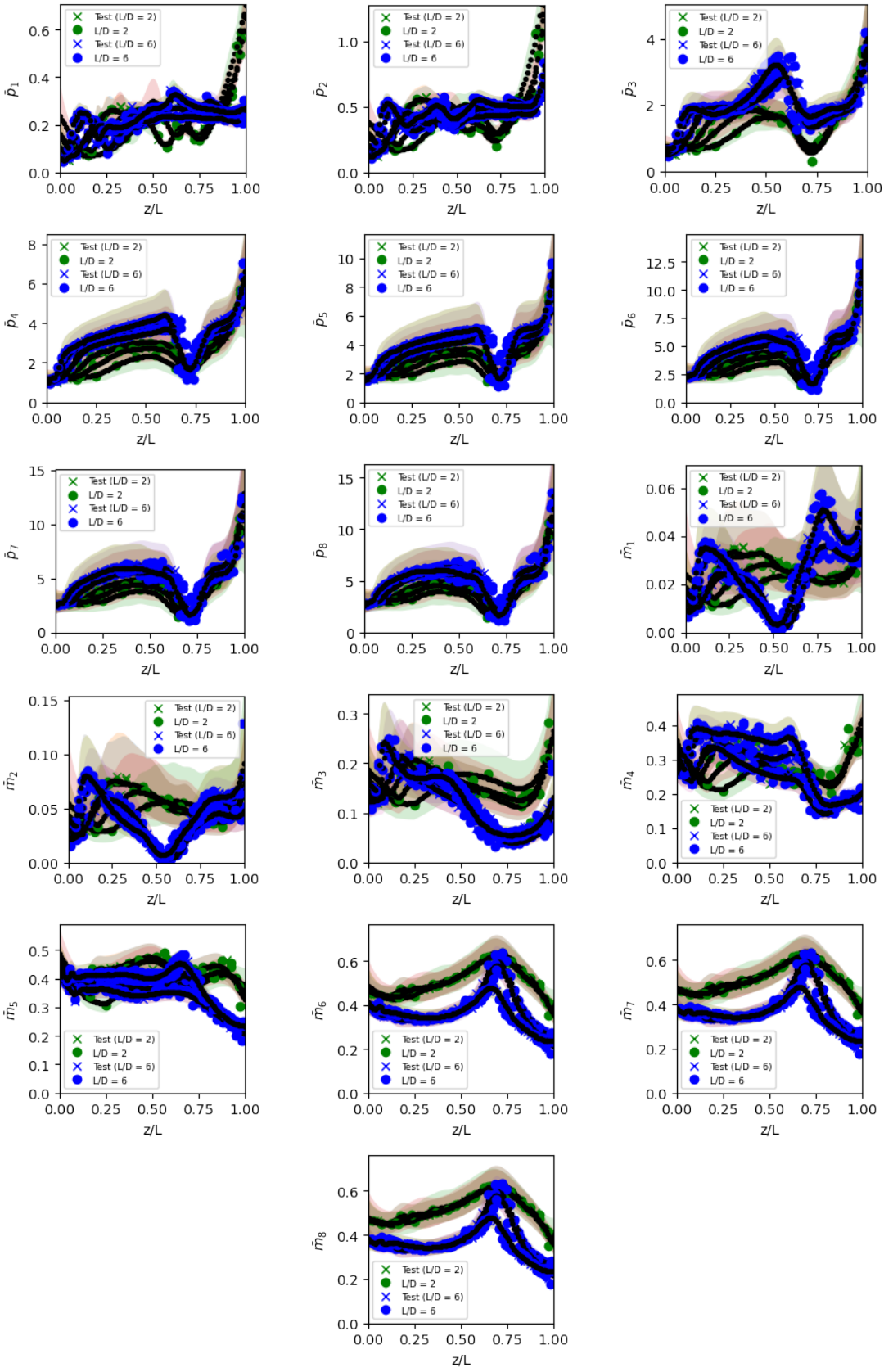


Figure 4.7: Example GPR outputs for the distributed lateral $\bar{p}$ and moment $\bar{m}$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore glacial clay till site. The depth variation of the mean GPR prediction for the knot points parameters is indicated by the black markers. The 95% confidence interval is shown (shaded bounds). The green and blue markers are the datapoints extracted from the calibration FEA.

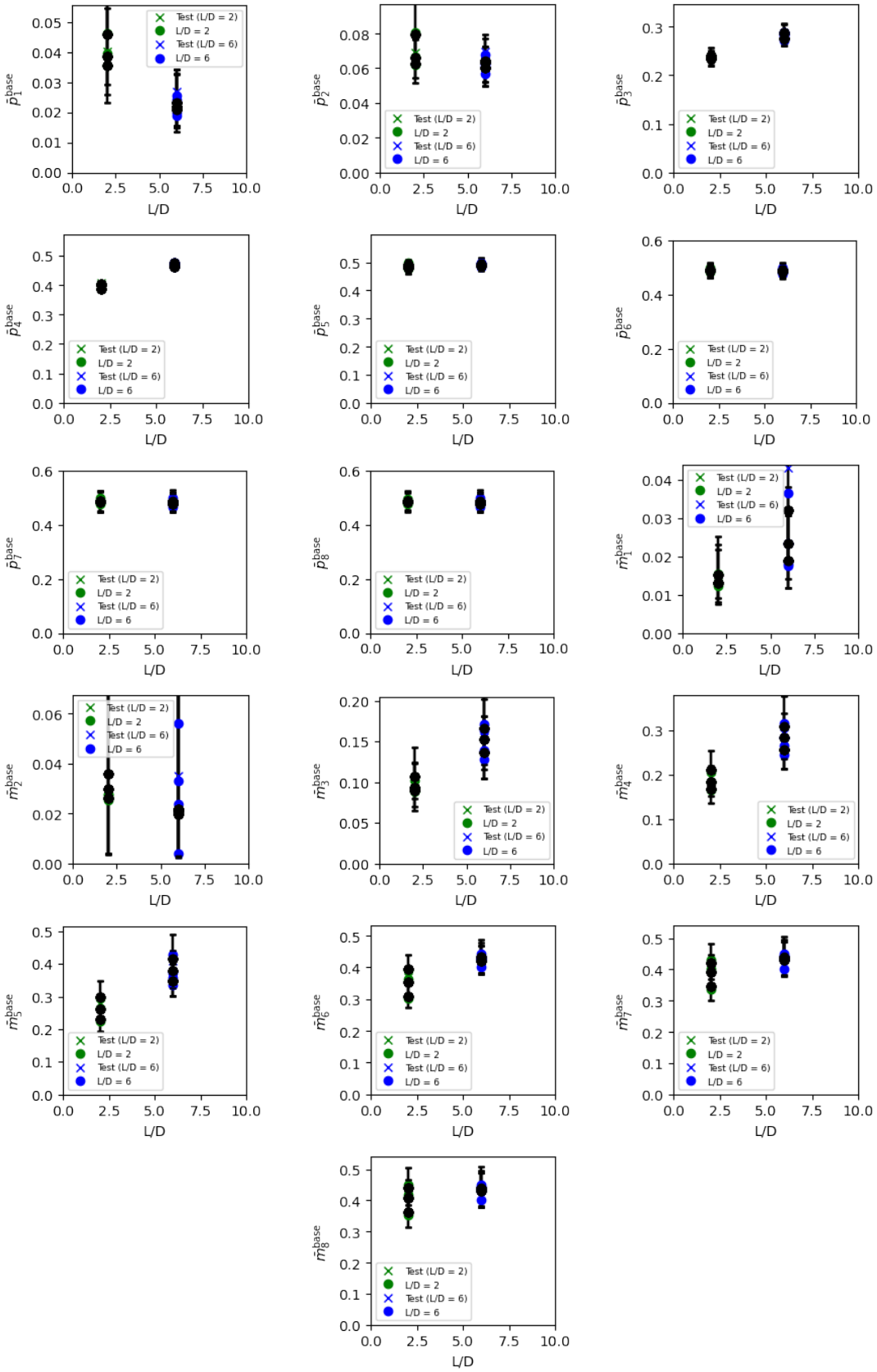


Figure 4.8: Example GPR outputs for the base horizontal force $\overline{H}_B$ and base moment $\overline{M}_B$ knot point parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore glacial clay till site. The point prediction of the mean GPR prediction for the knot points parameters is indicated by the black markers. The 95% confidence interval is shown by the black error bars. The green and blue markers are the datapoints extracted from the calibration FEA.

Table 4.7: Optimised kernel parameters for the GPR models as applied to the knot point soil reaction parameters associated with the PISA representative offshore homogenous sand site with a relative density D_R of 75% (σ_f is the signal variance of the Matérn kernel, σ_n is the variance of the white noise kernel, l is the lengthscales for every input dimension).

Knot point	$\bar{p}$	$\bar{m}$	$\bar{H}_B$	$\bar{M}_B$
1	σ_f=1.06, σ_n^2=0.0601, $l_{z/L}$=0.312, $l_{L/D}$=6.46, $l_{G_o/\sigma_v'}$=1000	σ_f=1.99, σ_n^2=0.0684, $l_{z/L}$=0.865, $l_{L/D}$=24.4, $l_{G_o/\sigma_v'}$=1000	σ_f=2.66, σ_n^2=0.0112, $l_{L/D}$=22.1, $l_{G_o/\sigma_v'}$=1000	σ_f=3.42, σ_n^2=0.014, $l_{L/D}$=100, $l_{G_o/\sigma_v'}$=1000
2	σ_f=0.86, σ_n^2=0.177, $l_{z/L}$=0.2, $l_{L/D}$=10.2, $l_{G_o/\sigma_v'}$=1000	σ_f=1.67, σ_n^2=0.0661, $l_{z/L}$=0.848, $l_{L/D}$=19, $l_{G_o/\sigma_v'}$=1000	σ_f=2.02, σ_n^2=0.0051, $l_{L/D}$=20.6, $l_{G_o/\sigma_v'}$=1000	σ_f=3.53, σ_n^2=0.041, $l_{L/D}$=100, $l_{G_o/\sigma_v'}$=1000
3	σ_f=0.84, σ_n^2=0.0318, $l_{z/L}$=0.232, $l_{L/D}$=6.42, $l_{G_o/\sigma_v'}$=1000	σ_f=0.99, σ_n^2=0.0484, $l_{z/L}$=0.857, $l_{L/D}$=8.05, $l_{G_o/\sigma_v'}$=1000	σ_f=0.93, σ_n^2=4.83e-5, $l_{L/D}$=2.5, $l_{G_o/\sigma_v'}$=887	σ_f=2.17, σ_n^2=2.94e-3, $l_{L/D}$=19.2, $l_{G_o/\sigma_v'}$=1000
4	σ_f=1.08, σ_n^2=0.047, $l_{z/L}$=0.231, $l_{L/D}$=13.3, $l_{G_o/\sigma_v'}$=1000	σ_f=0.82, σ_n^2=0.0186, $l_{z/L}$=0.857, $l_{L/D}$=8.05, $l_{G_o/\sigma_v'}$=1000	σ_f=0.60, σ_n^2=0.0155, $l_{L/D}$=2.39, $l_{G_o/\sigma_v'}$=1000	σ_f=1.69, σ_n^2=2.12e-4, $l_{L/D}$=7.08, $l_{G_o/\sigma_v'}$=1000
5	σ_f=1.28, σ_n^2=0.0593, $l_{z/L}$=0.236, $l_{L/D}$=19.7, $l_{G_o/\sigma_v'}$=1000	σ_f=0.87, σ_n^2=0.0522, $l_{z/L}$=0.2, $l_{L/D}$=5.88, $l_{G_o/\sigma_v'}$=1000	σ_f=0.60, σ_n^2=0.0281, $l_{L/D}$=1.80, $l_{G_o/\sigma_v'}$=1000	σ_f=1.86, σ_n^2=0.021, $l_{L/D}$=3.29, $l_{G_o/\sigma_v'}$=1000
6	σ_f=1.5, σ_n^2=0.0811, $l_{z/L}$=0.236, $l_{L/D}$=16.2, $l_{G_o/\sigma_v'}$=1000	σ_f=0.93, σ_n^2=0.0586, $l_{z/L}$=0.201, $l_{L/D}$=6.25, $l_{G_o/\sigma_v'}$=1000	σ_f=0.62, σ_n^2=0.0453, $l_{L/D}$=0.5, $l_{G_o/\sigma_v'}$=1000	σ_f=1.84, σ_n^2=0.023, $l_{L/D}$=2.17, $l_{G_o/\sigma_v'}$=1000
7	σ_f=1.62, σ_n^2=0.0867, $l_{z/L}$=0.254, $l_{L/D}$=14.5, $l_{G_o/\sigma_v'}$=1000	σ_f=0.99, σ_n^2=0.0697, $l_{z/L}$=0.205, $l_{L/D}$=5.78, $l_{G_o/\sigma_v'}$=1000	σ_f=0.63, σ_n^2=0.0464, $l_{L/D}$=0.5, $l_{G_o/\sigma_v'}$=1000	σ_f=1.84, σ_n^2=0.023, $l_{L/D}$=1.72, $l_{G_o/\sigma_v'}$=1000
8	σ_f=1.61, σ_n^2=0.0872, $l_{z/L}$=0.252, $l_{L/D}$=14.2, $l_{G_o/\sigma_v'}$=1000	σ_f=1.03, σ_n^2=0.0704, $l_{z/L}$=0.203, $l_{L/D}$=5.7, $l_{G_o/\sigma_v'}$=1000	σ_f=0.63, σ_n^2=0.0464, $l_{L/D}$=0.5, $l_{G_o/\sigma_v'}$=1000	σ_f=1.84, σ_n^2=0.023, $l_{L/D}$=1.34, $l_{G_o/\sigma_v'}$=1000

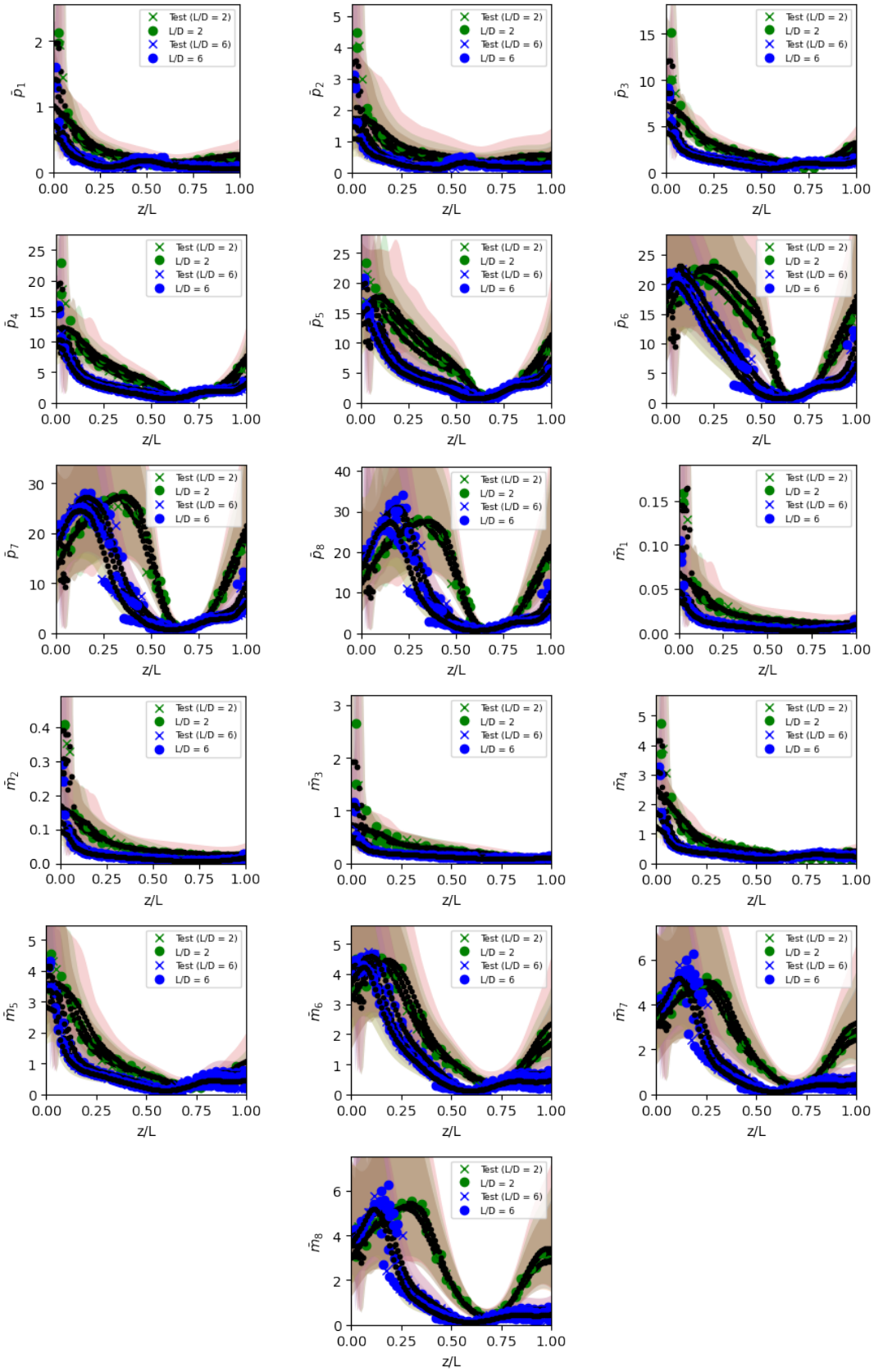


Figure 4.9: Example GPR outputs for the distributed lateral $\bar{p}$ and moment $\bar{m}$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore homogenous sand site with D_R of 75%. The depth variation of the mean GPR prediction for the knot point parameters is indicated by the black markers. The 95% confidence interval is shown (shaded bounds). The green and blue markers are the datapoints extracted from the calibration FEA.

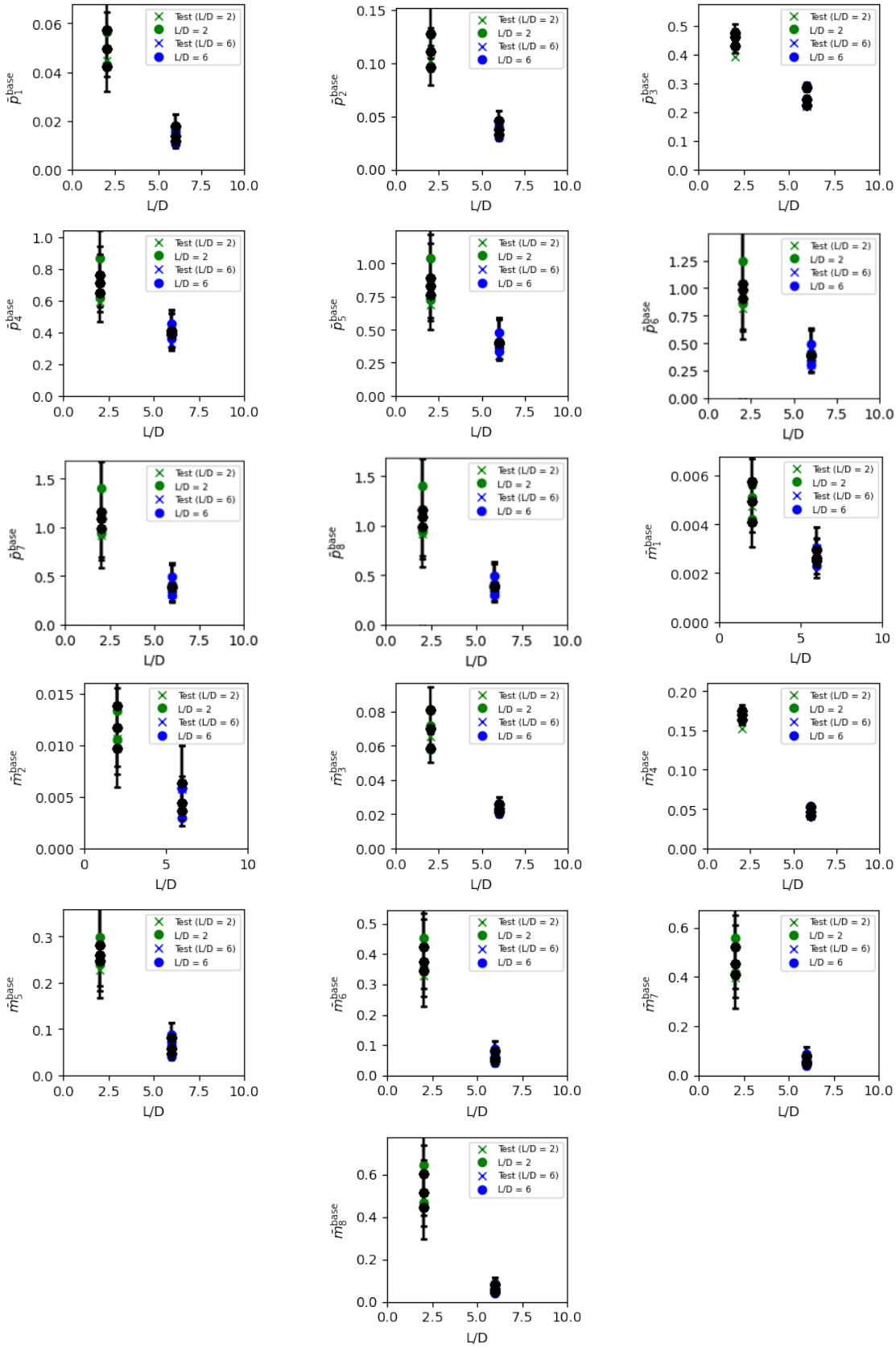


Figure 4.10: Example GPR outputs for the base horizontal force $\overline{H}_B$ and base moment $\overline{M}_B$ knot point soil reaction parameters for the calibration piles embedded in the PISA representative offshore homogenous sand site with D_R of 75%. The point prediction of the mean GPR prediction for the knot points parameters is indicated by the black markers. The 95% confidence interval is shown by the black error bars. The green and blue markers are the datapoints extracted from the calibration FEA.

Table 4.8 Regression performance metrics for the Matérn GPR models as applied to the knot point soil reaction parameters associated with the PISA representative offshore glacial clay till site.

Distributed knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Base knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.00	0.01	0.99	0.96	$\overline{H}_{B_1}$	0.01	0.04	0.89	0.07
$\bar{p}_2$	0.00	0.01	0.98	0.95	$\overline{H}_{B_2}$	0.00	0.01	0.78	-35.59
$\bar{p}_3$	0.01	0.01	0.95	0.95	$\overline{H}_{B_3}$	0.00	0.00	0.98	0.99
$\bar{p}_4$	0.01	0.01	0.97	0.92	$\overline{H}_{B_4}$	0.00	0.00	1.00	0.97
$\bar{p}_5$	0.01	0.01	0.95	0.89	$\overline{H}_{B_5}$	0.00	0.00	0.76	-13.18
$\bar{p}_6$	0.01	0.02	0.93	0.90	$\overline{H}_{B_6}$	0.00	0.00	0.90	0.09
$\bar{p}_7$	0.01	0.02	0.92	0.90	$\overline{H}_{B_7}$	0.00	0.00	0.94	-2.15
$\bar{p}_8$	0.01	0.02	0.91	0.89	$\overline{H}_{B_8}$	0.00	0.00	0.95	-2.08
$\bar{m}_1$	0.03	0.03	0.95	0.91	$\overline{M}_{B_1}$	0.02	0.04	0.83	0.84
$\bar{m}_2$	0.03	0.03	0.94	0.93	$\overline{M}_{B_2}$	0.05	0.06	0.92	-6.36
$\bar{m}_3$	0.02	0.02	0.94	0.90	$\overline{M}_{B_3}$	0.01	0.01	0.89	0.86
$\bar{m}_4$	0.00	0.00	0.95	0.94	$\overline{M}_{B_4}$	0.00	0.01	0.94	0.88
$\bar{m}_5$	0.00	0.00	0.95	0.95	$\overline{M}_{B_5}$	0.00	0.00	0.96	0.91
$\bar{m}_6$	0.00	0.00	0.97	0.96	$\overline{M}_{B_6}$	0.00	0.00	0.94	-0.18
$\bar{m}_7$	0.00	0.00	0.96	0.96	$\overline{M}_{B_7}$	0.00	0.00	0.84	-22.57
$\bar{m}_8$	0.00	0.00	0.96	0.96	$\overline{M}_{B_8}$	0.00	0.00	0.80	-40.61

Table 4.9: Regression performance metrics for the Matérn GPR models as applied to the knot point soil reaction parameters corresponding to the distributed lateral load $\bar{p}$ the base horizontal force $\overline{H}_B$ associated with the PISA representative offshore homogenous sand site with D_R of 75%.

Distributed knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Base knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.05	0.24	0.90	0.66	$\overline{H}_{B_1}$	0.00	0.00	0.99	1.00
$\bar{p}_2$	0.16	0.23	0.76	0.62	$\overline{H}_{B_2}$	0.00	0.00	0.99	0.99
$\bar{p}_3$	0.03	0.03	0.94	0.92	$\overline{H}_{B_3}$	0.00	0.02	1.00	0.76
$\bar{p}_4$	0.04	0.05	0.94	0.92	$\overline{H}_{B_4}$	0.01	0.04	0.91	0.48
$\bar{p}_5$	0.05	0.05	0.94	0.94	$\overline{H}_{B_5}$	0.02	0.06	0.89	0.61
$\bar{p}_6$	0.07	0.07	0.95	0.95	$\overline{H}_{B_6}$	0.03	0.06	0.89	0.77
$\bar{p}_7$	0.07	0.09	0.95	0.95	$\overline{H}_{B_7}$	0.03	0.05	0.91	0.82
$\bar{p}_8$	0.07	0.09	0.95	0.95	$\overline{H}_{B_8}$	0.03	0.05	0.91	0.82

Table 4.10: Regression performance metrics for the Matérn GPR models as applied to the knot point soil reaction parameters corresponding to the distributed moment $\bar{m}$ and base moment $\bar{M}_B$ associated with the PISA representative offshore homogenous sand site with D_R of 75%.

Distributed knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Base knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{m}_1$	0.06	0.08	0.92	0.89	$\bar{M}_{B_1}$	0.01	0.03	0.94	0.41
$\bar{m}_2$	0.06	0.10	0.92	0.87	$\bar{M}_{B_2}$	0.02	0.12	0.92	-0.26
$\bar{m}_3$	0.04	0.03	0.89	0.91	$\bar{M}_{B_3}$	0.00	0.01	1.00	0.97
$\bar{m}_4$	0.02	0.03	0.96	0.93	$\bar{M}_{B_4}$	0.00	0.01	1.00	0.99
$\bar{m}_5$	0.04	0.08	0.95	0.89	$\bar{M}_{B_5}$	0.01	0.04	0.98	0.92
$\bar{m}_6$	0.05	0.08	0.96	0.94	$\bar{M}_{B_6}$	0.01	0.05	0.99	0.94
$\bar{m}_7$	0.06	0.10	0.96	0.94	$\bar{M}_{B_7}$	0.01	0.05	0.99	0.95
$\bar{m}_8$	0.06	0.10	0.96	0.94	$\bar{M}_{B_8}$	0.01	0.06	0.99	0.95

4.6.3 Feedforward Neural Network regression training procedure

A shallow neural network with one hidden layer is employed; exploratory studies using more hidden layers increase unnecessary the complexity and do not boost the performance of the artificial neural network model due to the relatively small training database employed. Suryasentana et al. (2021) showed that the $tanh$ activation function significantly outperformed the $ReLu$ activation function for the highly non-linear distribution of the PISA 4-parameters conic function data; therefore, the $tanh$ activation function was employed. In the current work a careful exercise of hyperparameter tuning was carried out to maximise the performance of the FNN. The '*lbfgs*' solver was adopted for the weight optimisation; this solver is appropriate for small datasets and converges faster. The strength of the regularization term (a) was set to $a = 10^{-3}$ to account for the possibility of noisy data and avoid overfitting. The optimal number of neurons in the hidden layer for the distributed soil reaction $\bar{p}$, $\bar{m}$ knot point parameters was obtained after implementing 5-fold cross validation over the possible choices of 10 to 120 neurons. Table 4.11 presents the optimum number of neurons obtained for the $\bar{p}$, $\bar{m}$ knot point parameters. These neurons were used for the

training of the FNN regression models corresponding to the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75%. For the $\bar{H}_B, \bar{M}_B$ knot point parameters where the number of available training data is markedly reduced, 10 neurons were employed. Before the training, the inputs were standardised by scaling their values between 0 and 1 to improve the generalisation properties of the final trained neural network ML models.

Table 4.11: Optimum number of neurons determined by the 5-fold cross validation for the FNN regression models as applied to the $\bar{p}$, $\bar{m}$ knot point parameters associated with the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75%.

Knot point	**Glacial clay till**		**Sand (D_R=75%)**	
	$\bar{p}$	$\bar{m}$	$\bar{p}$	$\bar{m}$
1	10	20	35	20
2	20	20	10	50
3	25	20	65	20
4	25	20	80	15
5	50	40	10	15
6	40	10	15	40
7	55	20	10	40
8	55	10	15	10

4.6.3.1 Assessment of FNN regression model

As there are many parameters for each FNN regression model, the optimised parameters obtained after training are not listed here. The difficulty in interpreting the optimised weights due to the large number of parameters that they incorporate is one of the main disadvantages of neural networks. Table 4.12 and Table 4.13 show the MSE and R^2 scores achieved by the training of the FNN regression models. For the clay data the achieved train MSE and R^2 scores show that the predictions describe the real values well. Also, the similarity in the achieved training and test metric values indicate that the FNN regression models have good generalisation properties and the predictions for unseen inputs inside the calibration space

should be reliable. For the noisier sand data, the achieved train MSE and R^2 scores are slightly worse than the clay data, but the predictions are still seen to describe the real values quite well. However, for the $\bar{p}_7$, $\bar{p}_8$ knot point parameters where the data are very noisy, the MSE test scores are much higher than the train scores (this suggests the possibility of overfitting). The predictions for these two FNN regression models for unseen inputs are not highly reliable, which can be attributed to the small number of hidden neurons utilised in the shallow neural networks for the $\bar{p}_7$, $\bar{p}_8$ knot point parameters. Similar, to the GPR models, the base knot point parameters for some of the FNN regression models provide negative R^2 values; this can be explained due to the small dataset (9 points) for each ML model. The R^2 values for this base knot points are not reliable.

Table 4.12: Regression performance metrics for the FNN regression models as applied to the knot-point parameters associated with the PISA representative offshore glacial clay till site.

Distributed knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Base knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.00	0.00	0.89	0.84	$\bar{H}_{B_1}$	0.00	0.00	0.87	0.67
$\bar{p}_2$	0.00	0.00	0.94	0.87	$\bar{H}_{B_2}$	0.00	0.00	0.28	-27.77
$\bar{p}_3$	0.02	0.08	0.93	0.88	$\bar{H}_{B_3}$	0.00	0.00	0.97	0.97
$\bar{p}_4$	0.01	0.01	0.94	0.94	$\bar{H}_{B_4}$	0.00	0.00	0.94	0.98
$\bar{p}_5$	0.02	0.03	0.96	0.94	$\bar{H}_{B_5}$	0.00	0.00	0.15	-39.62
$\bar{p}_6$	0.06	0.10	0.96	0.98	$\bar{H}_{B_6}$	0.00	0.00	0.02	0.07
$\bar{p}_7$	0.07	0.15	0.97	0.94	$\bar{H}_{B_7}$	0.00	0.00	0.09	-0.06
$\bar{p}_8$	0.07	0.12	0.97	0.94	$\bar{H}_{B_8}$	0.00	0.00	0.06	0.48
$\bar{m}_1$	0.07	0.12	0.98	0.93	$\bar{M}_{B_1}$	0.00	0.00	0.55	0.19
$\bar{m}_2$	0.00	0.00	0.82	0.86	$\bar{M}_{B_2}$	0.00	0.00	0.01	-0.63
$\bar{m}_3$	0.00	0.00	0.85	0.82	$\bar{M}_{B_3}$	0.00	0.00	0.85	0.96
$\bar{m}_4$	0.00	0.00	0.89	0.89	$\bar{M}_{B_4}$	0.00	0.00	0.92	0.96
$\bar{m}_5$	0.00	0.00	0.85	0.88	$\bar{M}_{B_5}$	0.00	0.00	0.93	0.97
$\bar{m}_6$	0.00	0.00	0.91	0.96	$\bar{M}_{B_6}$	0.00	0.00	0.83	-0.39
$\bar{m}_7$	0.00	0.00	0.95	0.95	$\bar{M}_{B_7}$	0.00	0.00	0.69	-35.40
$\bar{m}_8$	0.00	0.00	0.92	0.96	$\bar{M}_{B_8}$	0.00	0.00	0.53	-72.80

Table 4.13: Regression performance metrics for the FNN regression models as applied to the knot-point parameters associated with the PISA representative offshore homogenous sand site with D_R of 75%.

Distributed knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Base Knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.00	0.00	0.98	0.98	$\overline{H}_{B_1}$	0.00	0.00	0.99	0.99
$\bar{p}_2$	0.00	0.01	0.99	0.98	$\overline{H}_{B_2}$	0.00	0.00	0.99	1.00
$\bar{p}_3$	0.07	0.05	0.98	0.98	$\overline{H}_{B_3}$	0.00	0.00	0.95	0.77
$\bar{p}_4$	0.10	0.07	0.99	0.99	$\overline{H}_{B_4}$	0.00	0.03	0.95	-0.56
$\bar{p}_5$	0.18	0.19	0.99	0.99	$\overline{H}_{B_5}$	0.00	0.05	0.95	-0.34
$\bar{p}_6$	0.58	0.76	0.99	0.99	$\overline{H}_{B_6}$	0.00	0.08	0.96	-0.17
$\bar{p}_7$	1.43	3.66	0.98	0.96	$\overline{H}_{B_7}$	0.00	0.10	0.98	-0.13
$\bar{p}_8$	1.70	5.59	0.98	0.99	$\overline{H}_{B_8}$	0.00	0.11	0.99	-0.14
$\bar{m}_1$	0.00	0.00	0.81	0.74	$\overline{M}_{B_1}$	0.00	0.00	0.96	0.79
$\bar{m}_2$	0.00	0.00	0.98	0.96	$\overline{M}_{B_2}$	0.00	0.00	0.96	0.79
$\bar{m}_3$	0.00	0.00	0.93	0.95	$\overline{M}_{B_3}$	0.00	0.00	0.98	0.96
$\bar{m}_4$	0.00	0.00	0.99	0.99	$\overline{M}_{B_4}$	0.00	0.00	1.00	0.99
$\bar{m}_5$	0.01	0.01	0.99	0.99	$\overline{M}_{B_5}$	0.00	0.00	0.97	0.89
$\bar{m}_6$	0.01	0.01	0.99	0.99	$\overline{M}_{B_6}$	0.00	0.00	0.97	0.83
$\bar{m}_7$	0.07	0.37	0.98	0.89	$\overline{M}_{B_7}$	0.00	0.01	0.98	0.81
$\bar{m}_8$	0.07	0.50	0.98	0.87	$\overline{M}_{B_8}$	0.00	0.01	0.98	0.77

The predictions – with the pile geometry, soil profile properties as inputs – provide: (i) the depth variation of the knot point parameters of the distributed lateral load $\bar{p}$ and distributed moment $\bar{m}$; (ii) the predictions of the knot point parameters for the base horizontal force $\overline{H}_B$ and base moment $\overline{M}_B$. Figure 4.11 to Figure 4.14 compare the predictions by the FNN regression models with the knot point parameters employed as the ML models' training and test datapoints. The results correspond to the calibration piles (Table 3.1) embedded in the Cowden till and Dunkirk sand ($D_R = 75\%$) soil profiles respectively. In general, as indicated from the train MSE and R^2 scores, the FNN regression models can accurately capture the highly non-linear nature of both the trained clay and sand distributed soil reactions knot point data. Furthermore, similarly to the GPR models, the FNN regression models can accurately predict the base knot point data.

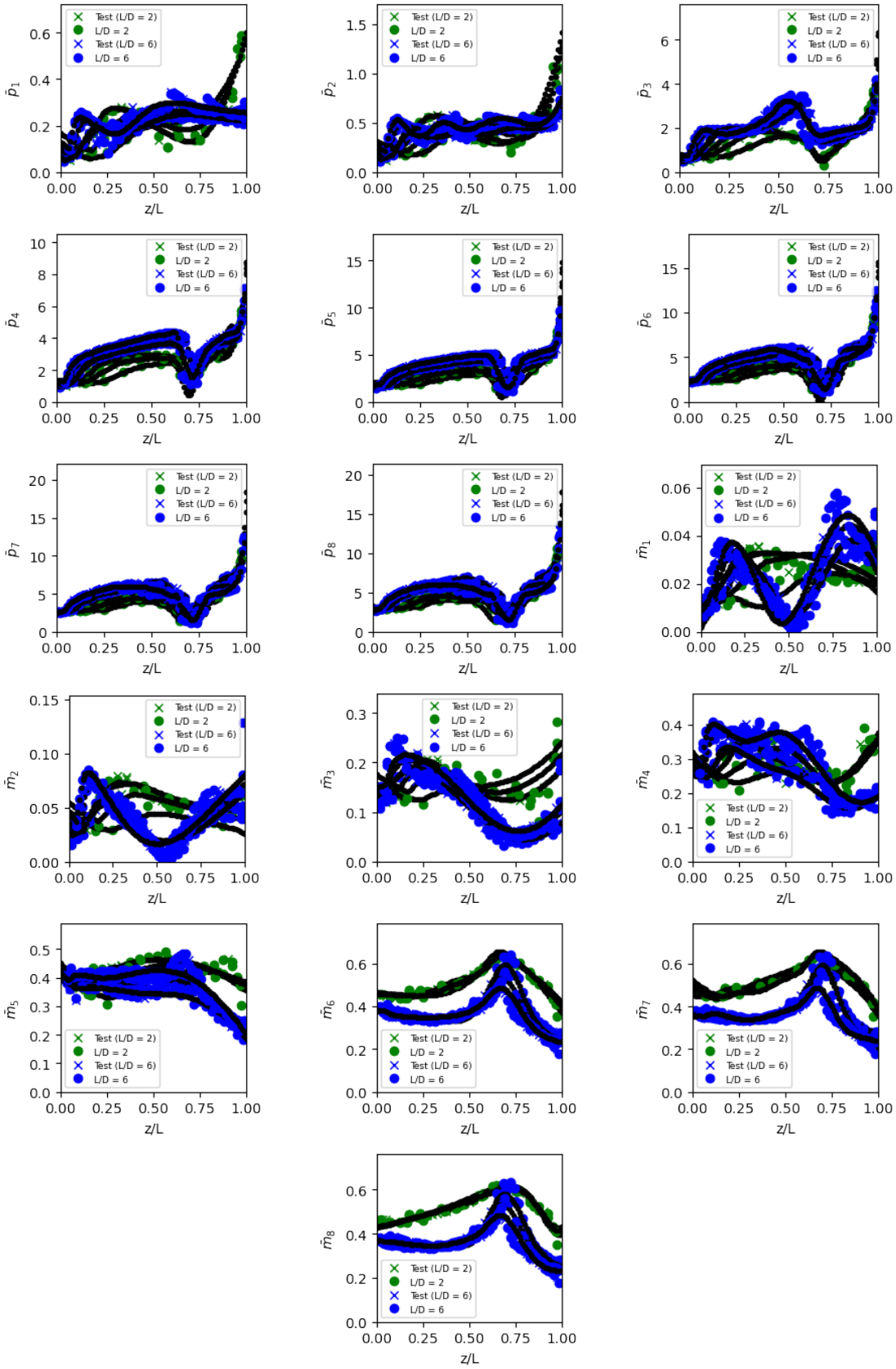


Figure 4.11: Example FNN regression ML outputs for the distributed lateral load $\bar{p}$ and moment $\bar{m}$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore glacial clay till site. The depth variation prediction of the FNN regression ML models for the knot points parameters is indicated by the black markers. The green and blue markers are the datapoints extracted from the calibration FEA.

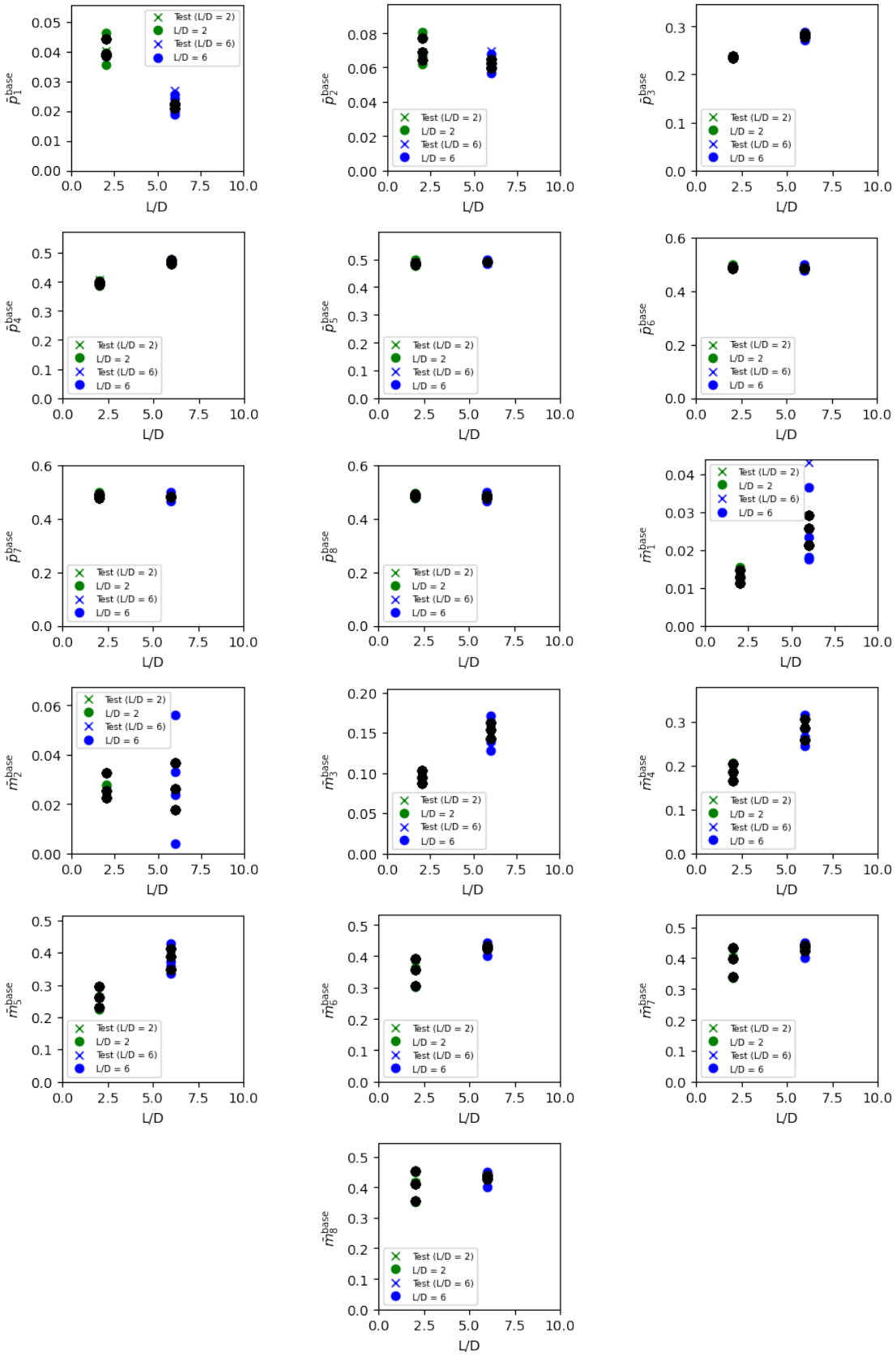


Figure 4.12: Example FNN regression ML outputs for the base horizontal force $\bar{H}_B$ and base moment $\bar{M}_B$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore glacial clay till site. The point prediction of the FNN regression ML models for the knot points parameters is indicated by the black markers. The green and blue markers are the datapoints extracted from the calibration FEA.

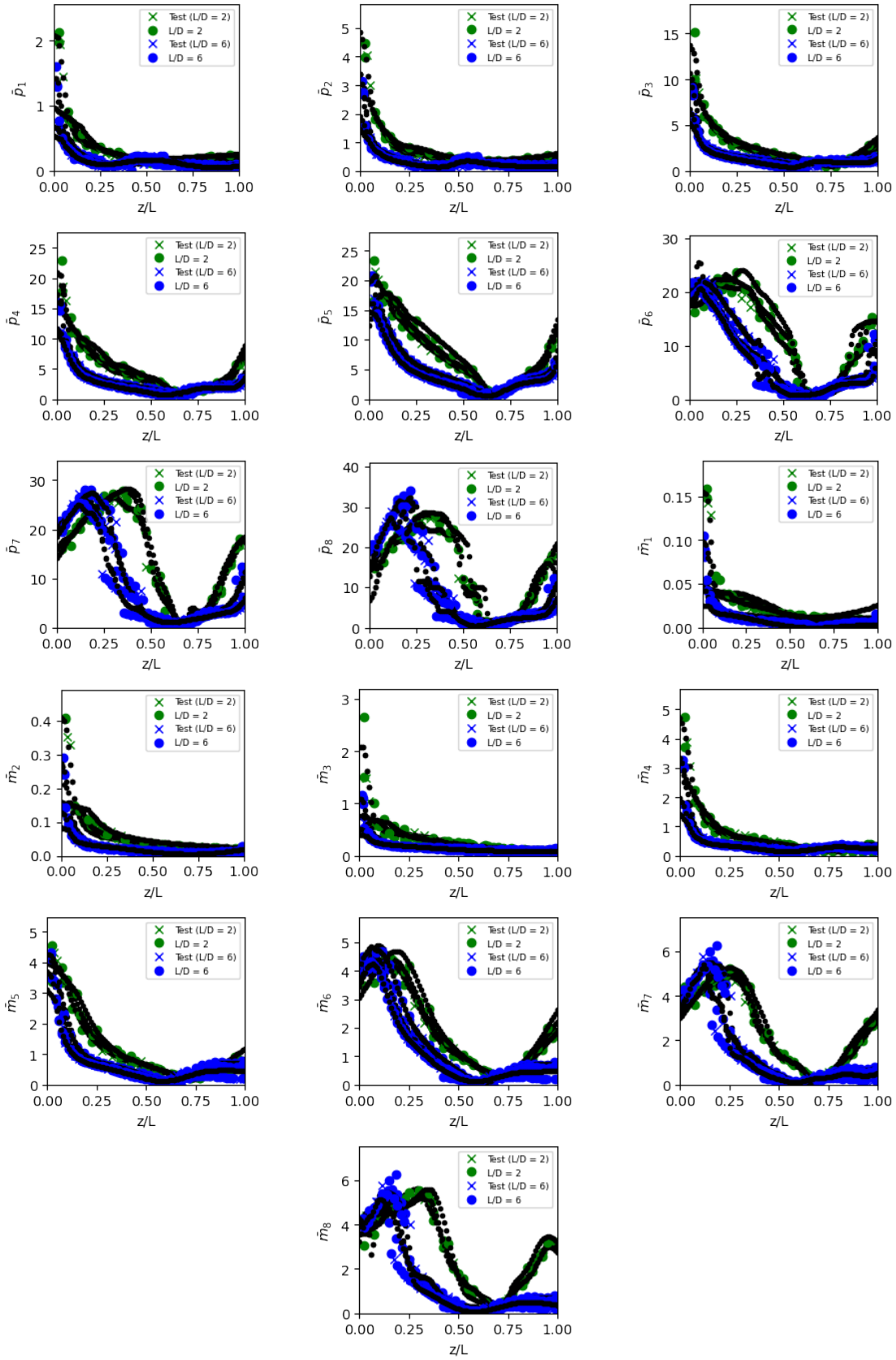


Figure 4.13: Example FNN regression ML outputs for the distributed lateral $\bar{p}$ and moment $\bar{m}$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore homogenous sand site with D_R of 75%. The depth variation prediction of the FNN regression ML models for the knot points parameters is indicated by the black markers. The green and blue markers are the datapoints extracted from the calibration FEA.

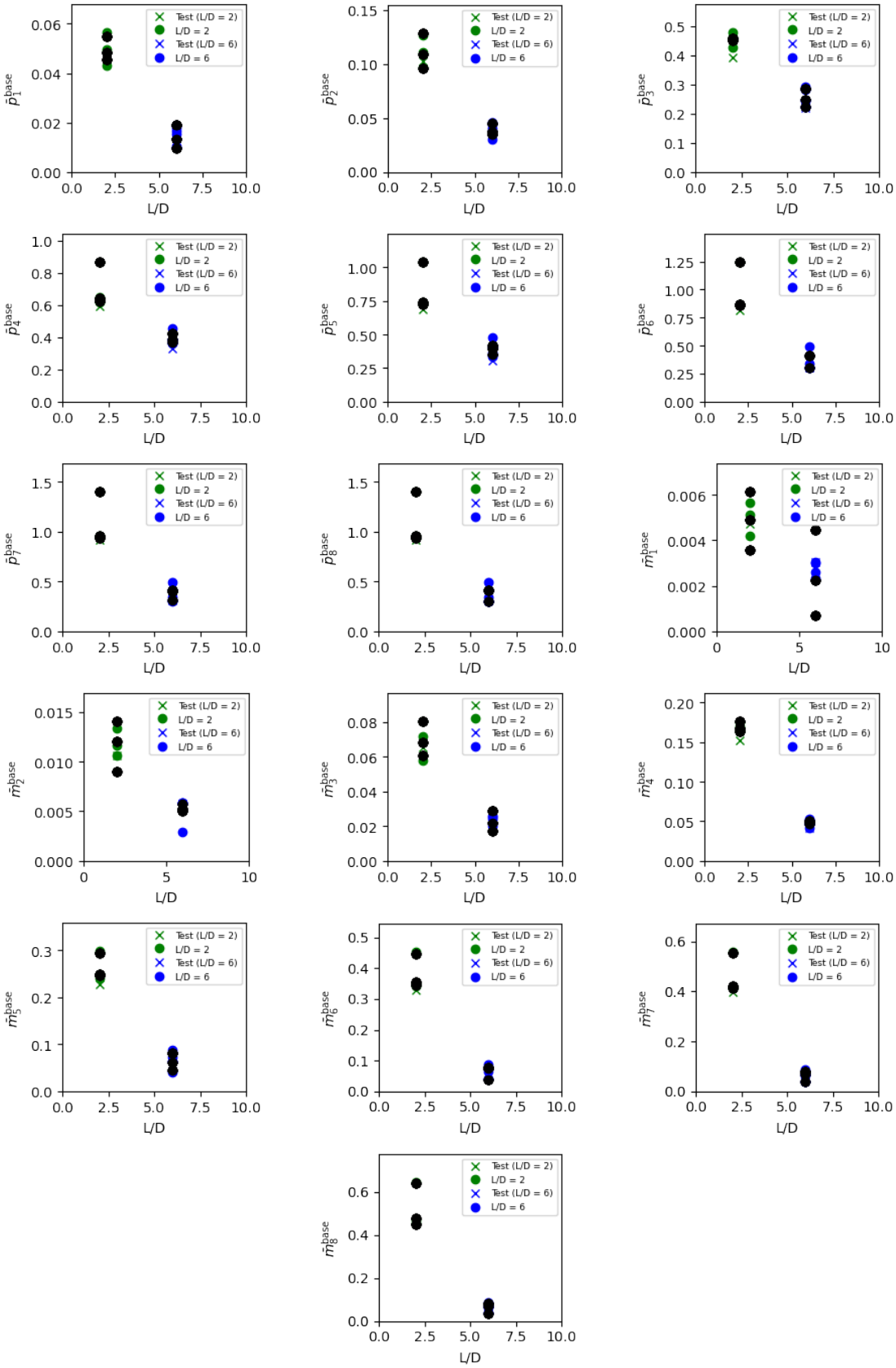


Figure 4.14: Example FNN regression ML outputs for the base horizontal force $\bar{H}_B$ and base moment $\bar{M}_B$ knot point soil reaction parameters for the calibration piles listed in Table 3.1 embedded in the PISA representative offshore homogenous sand site with D_R of 75%. The point prediction of the FNN regression ML models for the knot points parameters is indicated by the black markers. The green and blue markers are the datapoints extracted from the calibration FEA.

4.6.4 Conclusion

The GPR model was found to be the best performing of the two candidate ML models in terms of the MSE and R^2 train and test scores. Additionally, its built-in uncertainty measure, provided through its confidence interval, offers an estimation of the uncertainty associated with the final predictions of the knot point parameters. The GPR model has also the advantage that it can be interpreted in a straightforward manner via its optimised hyperparameters; the predictions can be viewed as a weighted average of its 'nearby' data points, where the definition of 'nearby' depends on the lengthscales. In contrast, the FNN regression model is harder to interpret, as it is not obvious to see which inputs contribute most to the output predictions. Both ML models are evaluated in Section 4.9 in terms of their ability to predict the knot point parameterised soil reaction curves for unseen pile geometries

4.7 Implementation of the machine learning predictions in the PISA 1D finite element framework

1D Models adopting the eight-knot point piecewise soil reaction curves parametrisation have been generated in OxPile software. These 1D models use the PCHIP interpolation scheme in MATLAB to calculate the forces values in the entire displacement domain generating a suitable approach to accurately reproduce the soil reaction curves as presented in Section 4.5. The 1D models are connected with the data-driven machine learning models' predictions and can provide the pile monotonic lateral behaviour. PCHIP interpolation (Fritsch and Carlson, 1980) contains a cubic polynomial function $P(x)$. On the $x_k \leq x \leq x_{k+1}$ interval expressed in terms of local variables $s = x - x_k$ and $h = h_k$, $P(x)$ is determined by:

$$P(x) = \frac{3hs^2 - 2s^3}{h^3} y_{k+1} + \frac{h^3 - 3hs^2 + 2s^3}{h^3} y_k + \frac{s^2(s-h)}{h^2} d_{k+1} + \frac{s(s-h)^2}{h^2} d_k \tag{4.15}$$

Here $d(k)$ is the slope of the function at kth interval $d(k) = P'(x_k)$ and h_k denotes the length of the kth subinterval where $h_k = x_{k+1} - x_k$. This cubic polynomial satisfies four interpolation conditions two on function values and two on the possibly unknown derivative values:

$$\begin{cases} P(x_k) = y_k \\ P(x_{k+1}) = y_{k+1} \\ P'(x_k) = d_k \\ P'(x_{k+1}) = d_{k+1} \end{cases} \tag{4.16}$$

PCHIP interpolation uses a special shape preserving procedure to define the gradient $d(k)$ (Moler, 2004) which is incorporated in the 1D models for the calculation of the constitutive matrix of the stress update procedure in OxPile software. In the case where between two neighboring points the linear gradients δ_{k-1} and δ_k have opposite signs then $d(k) = 0$. If δ_k and δ_{k-1} have the same sign and the two intervals have equal length then $d(k)$ is calculated from the harmonic mean of the two discrete slopes by:

$$\frac{1}{d_k} = \frac{1}{2} \times \left(\frac{1}{\delta_{k-1}} + \frac{1}{\delta_k} \right) \tag{4.17}$$

Finally, if δ_k and δ_{k-1} have the same sign and intervals with different lengths, $d(k)$ is a weighted harmonic mean, with weights determined by the lengths of the two intervals.

$$\frac{w1+w2}{d_k} = \frac{w1}{\delta_{k-1}} + \frac{w2}{\delta_k}, \qquad w_1 = 2h_k + h_{k-1}, \qquad w_2 = h_k + 2h_{k-1} \tag{4.18}$$

The slopes at the breakpoints of the interval d_1 and d_n are calculated by a slightly different procedure. For the first point initially, the slope d_1 is determined by:

$$d_1 = \frac{\big(2h(1) + h(2)\big) \times \delta_1 - h(1) \times \delta_2}{h(1) + h(2)} \tag{4.19}$$

In case that the sign of the calculated d_1 is the opposite of δ_1 then $d_1 = 0$. If the sign of d_1 is not the opposite from d_1 and the condition $|d_1| > |3\delta_1|$ is true then d_1 is calculated by

$d_1 = 3\delta_1$. The same procedure is followed for the determination of d_n where the indexes 1 and 2 are substituted by $n-1$ and $n-2$ respectively.

Figure 4.15 presents the load displacement response for C1 pile (Table 3.1) embedded in the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75% generated by using the 1D models adopting the eight-knot point parametrised soil reaction curves in OxPile software. The PLAXIS 3D results are also presented for comparison. The 1D models match accurately the PLAXIS 3D results. The small differences between the responses of the 3D and 1D models are attributed to a combination of factors including minor differences in the interpreted soil reaction from the eight-knot point soil reaction curves and any inaccuracies in the assumed physics of the Timoshenko beam elements of the 1D model.

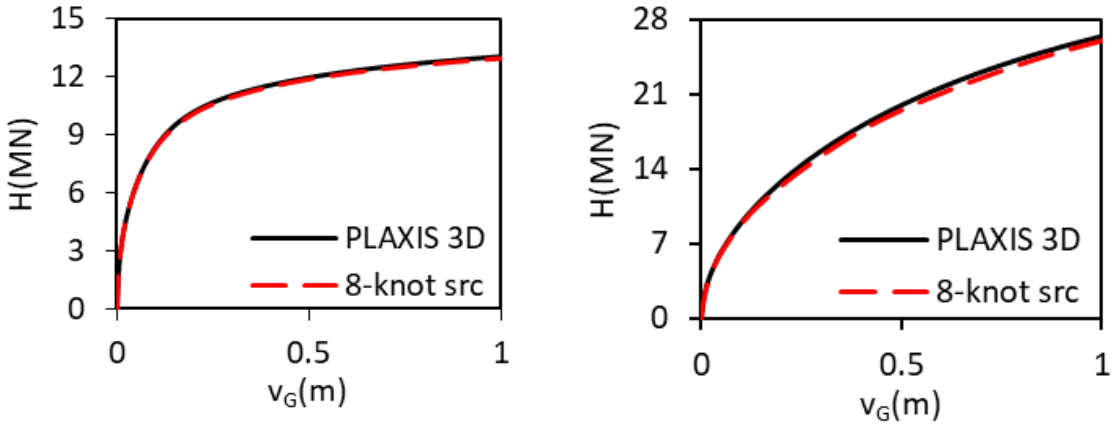


Figure 4.15: Comparison between computed responses determined from the 1D models adopting the eight-knot point piecewise spline function with the PCHIP interpolation scheme (indicated as '8-knot src' in the legend) and equivalent PLAXIS 3D FEA for pile C1. (a) the PISA representative offshore glacial clay till site, (b) the PISA representative offshore homogenous sand site with D_R of 75%.

4.8 Application of the data-driven method through calibration analyses

An initial exercise has been conducted in which predictions from each data-driven 1D design model are compared with the original training data (see

Figure 4.16). This comparison signified that 1D models employing the GPR or the FNN regression ML models, as indicated from the resulted high R^2 train scores, provide a good match to the PLAXIS 3D calibration analyses for both the clay and the sand soil profiles.

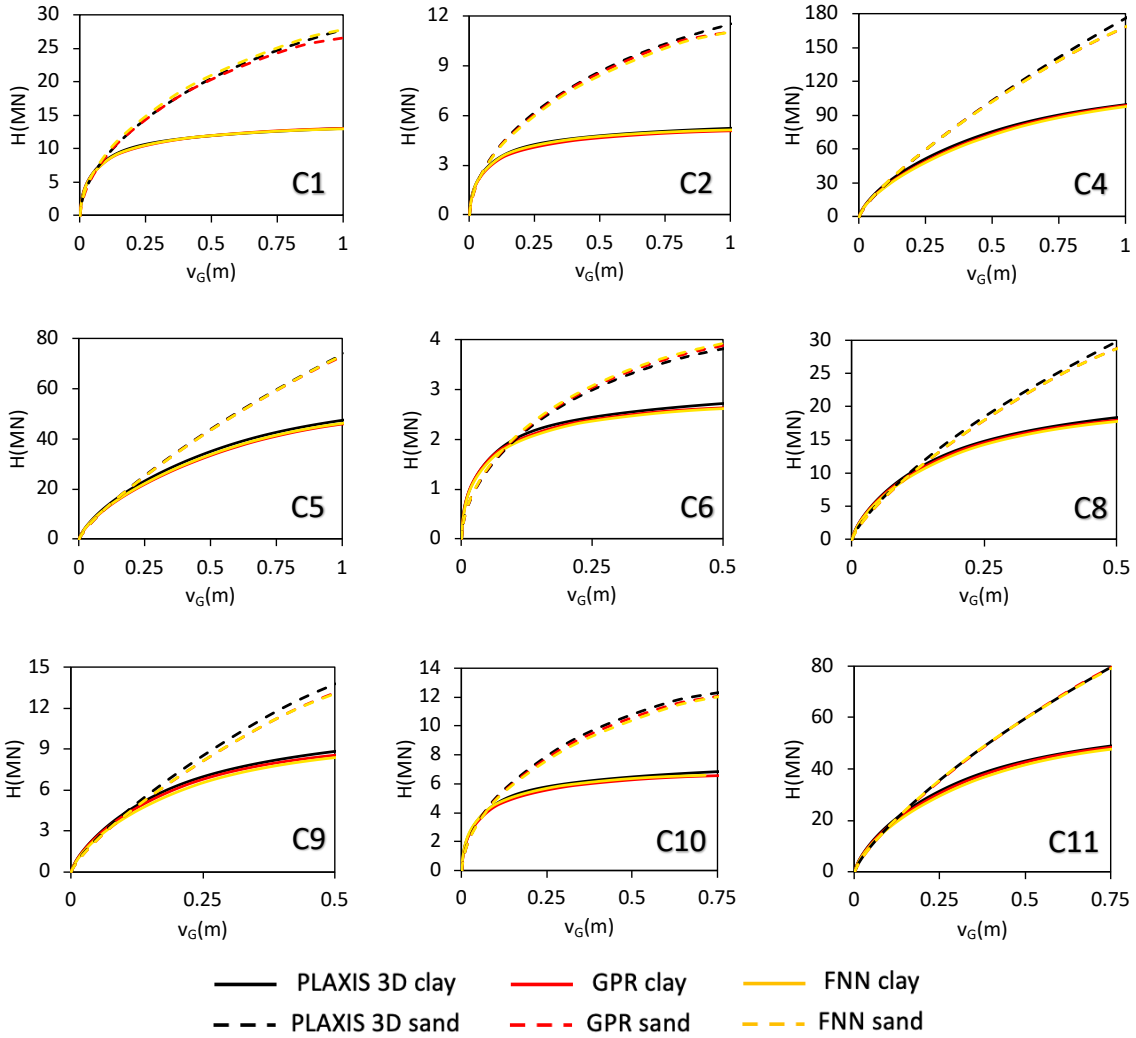


Figure 4.16: Load-displacement responses at ground level of the nine calibration piles (Table 3.1) embedded in the PISA representative offshore glacial clay till site (denoted by the straight lines) and the PISA representative offshore homogenous sand site of $D_R = 75\%$ (denoted by the dashed lines) computed by the 3D FE PLAXIS model and the data-driven 1D design model employing the GPR or the FNN regression ML models.

The high fidelity of the data-driven 1D design models is attributed to their ability to accurately predict numerical soil reaction curves. Figure 4.17 presents the distributed lateral load $\bar{p}$ and distributed moment $\bar{m}$ numerical soil reaction curves extracted from the PLAXIS 3D FEA at an arbitrary local level below the ground level $z/L = 0.225$ for the calibration piles in Table 3.1. Additionally, Figure 4.18 illustrates the base horizontal force $\bar{H}_B$ and base moment $\bar{M}_B$ numerical soil reaction curves extracted from the PLAXIS 3D FEA for the calibration piles in Table 3.1.

In Figure 4.17 and Figure 4.18, the numerical soil reaction curves are compared with the predicted soil reaction curves using the data-driven 1D design models which employ: (i) GPR models (which outperformed the FNN regression models in terms of the R^2 score), (ii) the eight-knot point piecewise spline function with the PCHIP interpolation scheme. It can be observed that the data-driven 1D design models accurately compute the numerical soil reaction curves at a local level for both clay and sand soil profiles. This feature is a significant advantage of the data-driven method because a 1D model employing numerical soil reaction curves consistently matches the predictions of the 3D model. Consequently, there is no need for additional optimisation stages (e.g., 2nd optimisation stage in PISA design model) to enhance the performance of the data-driven 1D design models.

It should be noted that the data-driven model (DDM) computes the same soil reaction curves for piles with the same length, L, and diameter, D, and different eccentricity load, h, (e.g. piles C1, C2) since the h/D eccentricity ratio was not included in the training inputs. In these cases, the GPR model interpolates between the knot point data of piles with different eccentricity ratios. Furthermore, as observed in Figure 4.17 and Figure 4.18 and discussed in Section 4.5, the soil reaction curves predicted by the data-driven 1D design models that are not mobilised up to $\bar{x} = 0.1$ are extrapolated by setting the force/moment value equal to the last calculated knot point value.

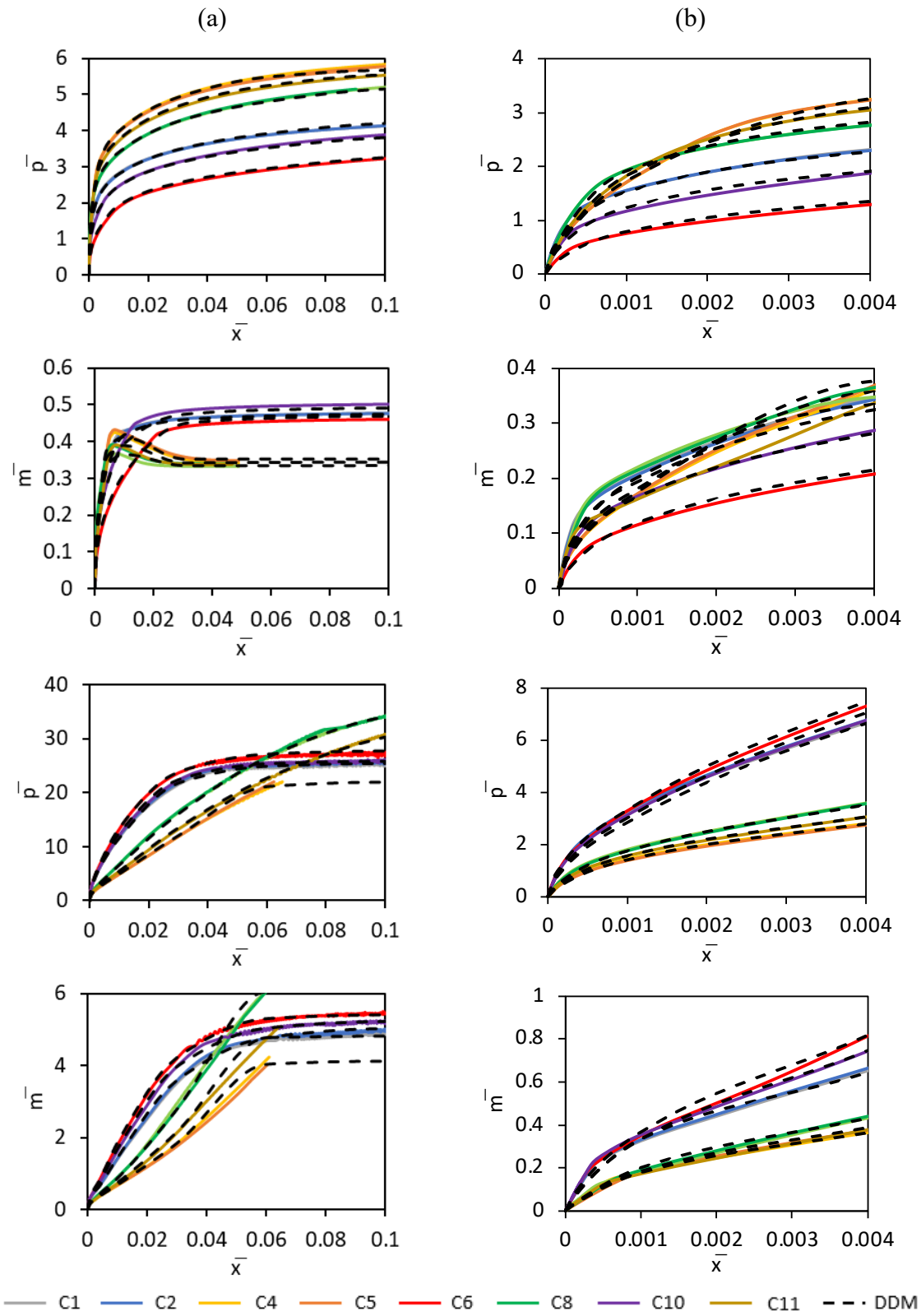


Figure 4.17: Normalised $\bar{p}$, $\bar{m}$ soil reaction curves at an arbitrary local level below the ground level $z/L = 0.225$ for piles C1 to C11 (Table 3.1) embedded in the PISA representative offshore glacial clay till (1st an 2nd row) and the PISA representative offshore homogenous sand site with D_R of 75% (3rd and 4th row) calculated by the PLAXIS 3D FEA and the data-driven 1D design models (referred as DDM): (a) overall response, (b) small displacement response.

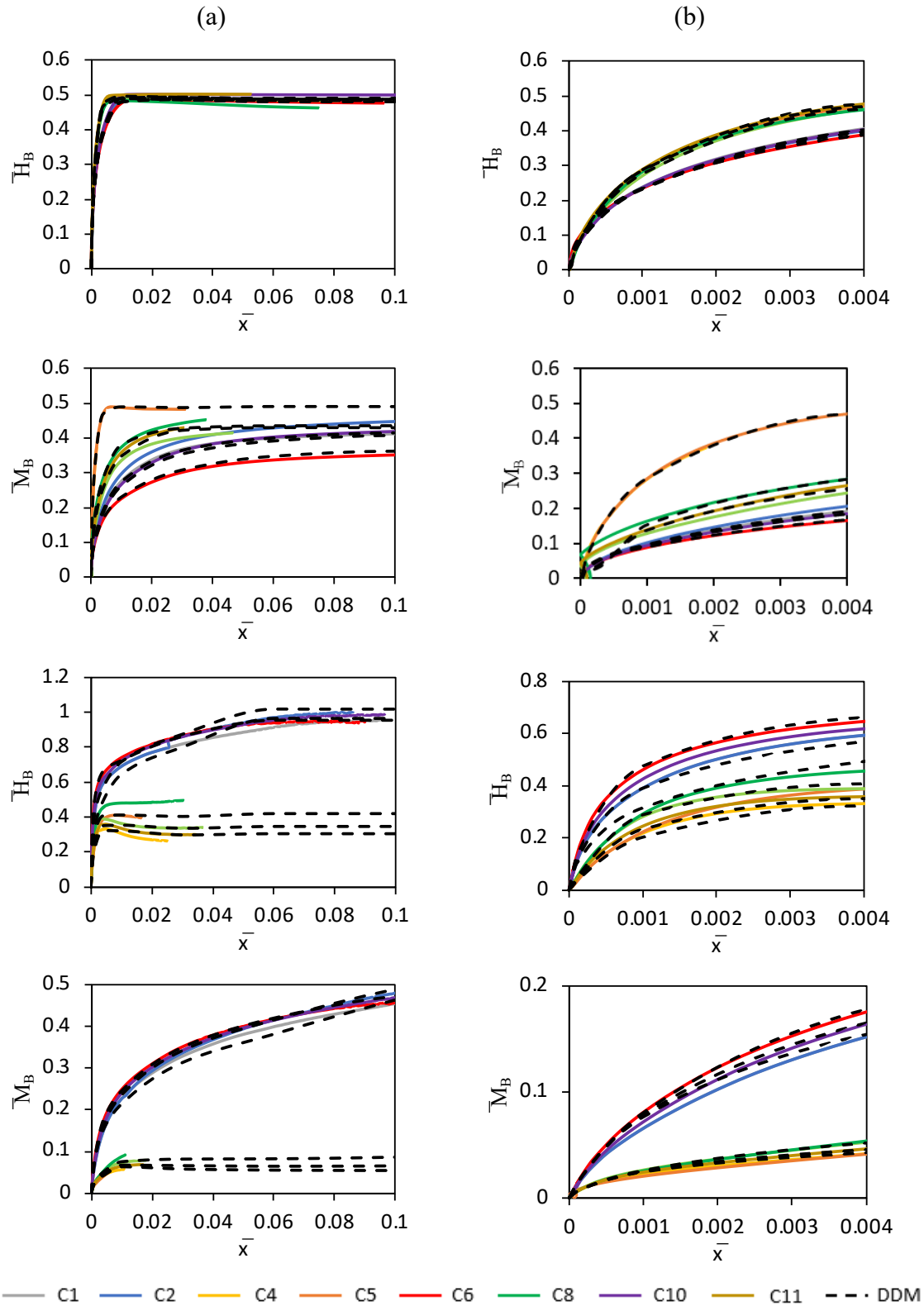


Figure 4.18: Normalised $\bar{H}_B$, $\bar{M}_B$ soil reaction curves for piles C1 to C11 (Table 3.1) embedded in the PISA representative offshore glacial clay till (1st an 2nd row) and the PISA representative offshore homogenous sand site with D_R of 75% (3rd and 4th row) calculated by the PLAXIS 3D FEA and the data-driven 1D design models (referred as DDM): (a) overall response, (b) small displacement response.

4.9 Application of the data-driven method through design cases

To demonstrate the predictive capability of the data-driven 1D design models, two design examples have been examined for the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with $D_R = 75\%$. The design examples are concerned with the design piles D1 and D2 (Table 4.14) employed in the PISA project (Byrne et al., 2020b; Burd et al., 2020a). These two pile geometries employ unseen $[L/D]$ inputs falling inside the trained calibration space. These same pile configurations were separately analysed using PLAXIS 3D FEA for validation purposes. The PISA accuracy metric η and ratio metric ρ (Byrne et al., 2020b; Burd et al., 2020a) were employed to quantify the precision of the two selected ML models in respect to the computed responses determined from the data-driven 1D design models. The two PISA metrics are defined as:

$$\eta = 1 - A_{diff}/A_{ref}\,, \quad \rho = H_{1D}/H_{3D} \tag{4.20}$$

where A_{ref} is the area below the H-v_G curve computed using the 3D model (for a specified value of v_G); A_{diff} is the corresponding area indicating the difference between the H-v_G curve computed using the 3D model with the H-v_G curve computed using the 1D model; H_{1D} and H_{3D} are the lateral loads computed with the 1D and 3D models respectively at a specific value of v_G.

Table 4.14: Design pile geometries; D is the pile diameter; L is the embedded length; t_w is the pile wall thickness; h is the load eccentricity.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D
D1	7.5	22.5	68	37.5	3
D2	8.75	35	91	87.5	4

Figure 4.19(a), Figure *4.19*(b) show the load-displacement response (H-v_G) over the small and large displacements range for the design piles D1, D2 embedded in the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with a relative density D_R of 75%. The results were computed using data-driven 1D design models which employ the GPR or the FNN regression ML models; corresponding data from PLAXIS 3D FEA are also shown. The results demonstrate a close fit between the 1D and 3D FE data across the full range of applied loading for both ML models used with the clay data. However, in terms of the PISA accuracy metrics, the GPR model slightly outperforms the FNN regression model, particularly for the pile's small displacement response (refer to Table 4.15). In contrast, for the sand data, which exhibit higher levels of noise, the GPR model notably outperforms the FNN regression model in terms of the PISA accuracy metrics. Furthermore, the load-displacement curve calculated for pile D2 at both the clay and sand sites displays irregularities, with non-monotonic values of tangent stiffness. This suggests that the shallow neural networks employed are insufficient to capture the depth variation of the knot point parameters for unseen pile geometries with high fidelity. This limitation may come from the relatively small database used for the site-specific data-driven 1D design models. Based on the study findings, it is concluded that the non-parametric GPR model will outperform a shallow neural network when trained with a database of approximately 250 to 300 points and describing four training inputs.

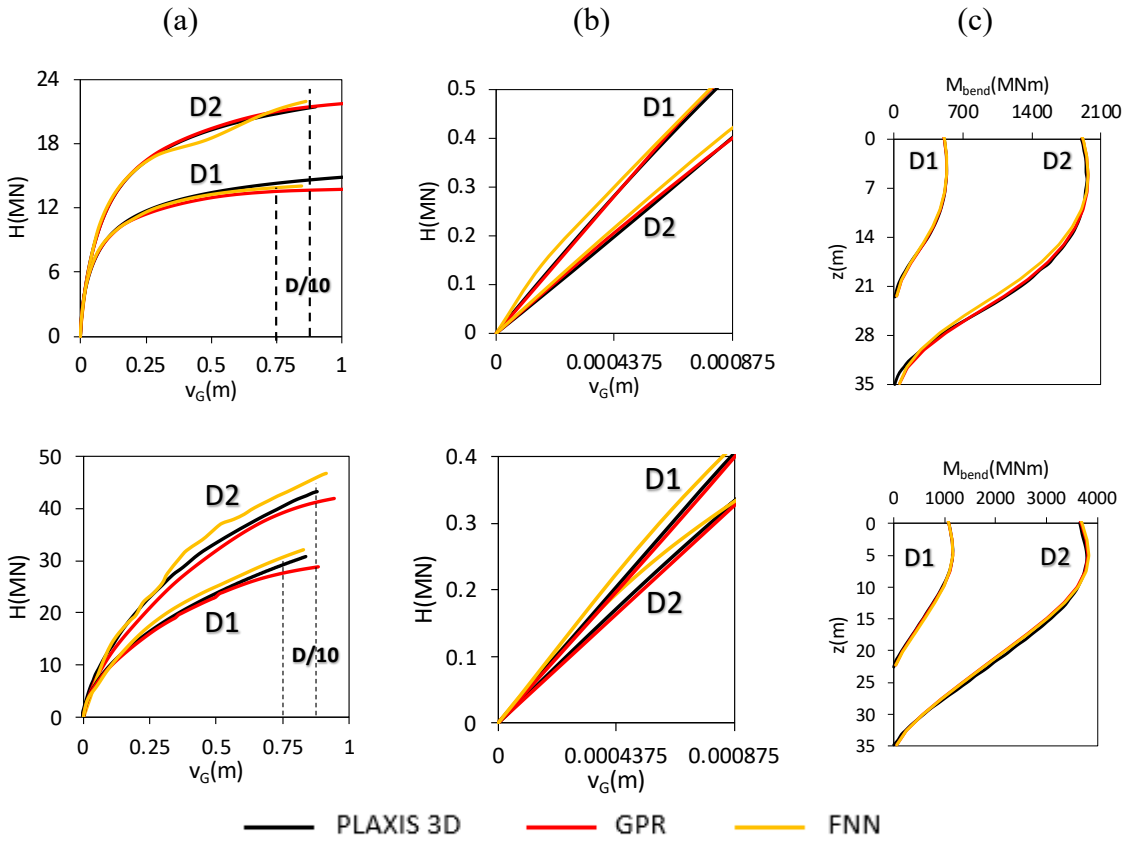


Figure 4.19: Comparison between computed responses determined from the data-driven 1D design models and PLAXIS results for piles D1, D2 emdedded: (i) in the PISA representative offshore glacial clay till site (1st row), (ii) in the PISA representative offshore homogenous sand site with D_R of 75% (2nd row); (a) ultimate response, (b) small displacement response, (c) bending moment profile with pile depth for an applied load H_{ult}^{3D} at ground-level displacement of $v_G = D/10$. The black line represents the 3D FE results, the red line the data-driven 1D design model employing the GPR and the yellow line the data-driven 1D design model employing the FNN regression.

Figure 4.19(c) illustrate the pile's induced bending moments computed by the 1D and 3D models; these bending moments were determined for an applied load H_{ult}^{3D} determined by the 3D FEA at a ground level displacement of $v_G = D/10$. The results indicate a good match between the bending moments determined from the 1D and 3D models. As indicated in Burd et al. (2020b), the magnitudes of the bending moments induced in the pile are strongly dependent on the applied moment at the pile head at ground level; since the pile head moment is determined solely by the applied loading the PISA 1D FE framework (OxPile software) will provide a good match to the computed bending moments from a 3D FEA for the same applied pile head moment. Hence, the data-driven 1D design models will consistently

compute accurately the pile's bending moment distribution and the bending moment data for other cases are not further discussed in this thesis.

Table 4.15: Values of accuracy metric η and ratio metric ρ for the small displacement response (η_{sd}, ρ_{sd}) and ultimate response (η_{ult}, ρ_{ult}) employing the Gaussian process regression and feedforward neural network regression machine learning models.

Soil Profiles	**Pile**	**Gaussian Process Regression**				**Feedforward Neural Network Regression**			
		η_{sd}	η_{ult}	ρ_{sd}	ρ_{ult}	η_{sd}	η_{ult}	ρ_{sd}	ρ_{ult}
PISA glacial clay till	D1	0.984	0.969	1.018	0.950	0.931	0.986	1.024	0.972
PISA glacial clay till	D2	0.955	0.994	1.026	1.004	0.910	0.977	1.082	1.031
Sand D_R=75%	D1	0.969	0.970	1.009	0.945	0.936	0.947	1.070	1.048
Sand D_R=75%	D2	0.946	0.952	0.948	0.956	0.931	0.929	1.042	1.060
Average	(-)	0.964	0.971	1.001	0.967	0.927	0.959	1.055	1.032

4.10 Summary

This chapter provides a detailed description of the key stages implemented to develop a data-driven 1D design model. The method is evaluated through two design examples concerned with generating a site-specific data-driven 1D design model for the PISA representative offshore glacial clay till site and the PISA representative offshore homogenous sand site with D_R of 75%. The soil reaction curves data used to train the ML models were extracted from the PLAXIS 3D calibration FEA presented in Sections 3.4.1 and 3.4.2.

The eight-knot point piecewise spline function with the PCHIP interpolation scheme can represent in high fidelity the numerical soil reaction curves. The relatively high number of parameters introduced using the new parametrisation method require the use of a ML regression model to generate a practical surrogate model. The results indicate that for the small training knot point database used for a site-specific data-driven 1D design model, the

GPR model using the Matérn kernel ($\nu = 5/2$) slightly outperforms a FNN regression model employing the $tanh$ activation function and a single hidden layer of neurons in terms of: (i) the R^2 scores, (ii) accuracy in unseen data and (iii) interpretability. Furthermore, the GPR model's built-in uncertainty measure, through its confidence interval, offers an effective means of estimating the uncertainty associated with the final predictions of the knot point parameters. Consequently, if the confidence interval is wide, indicating high uncertainty, the predicted soil reaction curves used by the 1D model may be deemed unreliable. In such cases, a 3D FEA is required to confirm the design.

Therefore, the GPR model was chosen as the regression method in the developed data-driven 1D design models in the following chapters. The non-parametric nature of the GPR model helps the model to easily adjust and develop with additional data and training inputs. It should be noted that the training computational time of a GPR model increases significantly with the amount of data (due to its non-parametric nature). The GPR model is not advised to be used for training databases consisted of tens of thousands of data. In such cases, the training time can be made faster by using sparse GPR models. Finally, the non-parametric GPR models – once the parameters have been trained – require the training data to be kept around to make future predictions. Therefore, it is advised to store the database (e.g. knot points, training inputs) in SQL format. A SQL database can be easily generated using for example the SQLite (DB4S) open source tool (Hipp, 2020).

Chapter 5

5 Development of Generalised Data-Driven Models for Clay and Sand Sites

5.1 Introduction

This chapter introduces a new method to generate generalised data-driven 1D design models suitable for: (i) clay sites with arbitrary strength and stiffness properties and, (ii) for sand sites with arbitrary relative density. These generalised data-driven 1D design models employ GPR machine learning models to predict the depth variation of eight-knot point soil reaction curves parameters along the pile. Subsequently, these predictions from the GPR models are integrated into the OxPile software, combined with the PCHIP interpolation 1D models outlined in Section 4.7, to compute the monotonic loading response of laterally loaded monopiles. The key steps involved in generating and validating these generalised data-driven 1D design models are summarised in Figure 5.1. The soil reaction curves utilised in this chapter to construct the generalised data-driven 1D design models were obtained from 3D finite element analysis (FEA) conducted using PLAXIS 3D software, following the methodology detailed in Chapter 3.

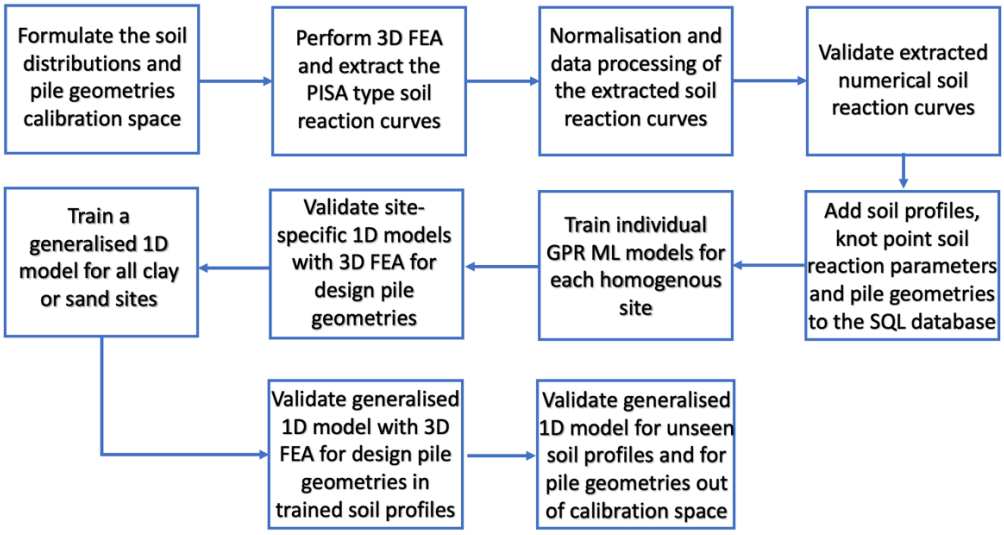


Figure 5.1: Stages implemented for generating a generalised data-driven 1D design model.

5.2 Data-driven Method for Clay with Arbitrary Strength and Stiffness

The new methodology is assessed in terms of developing a general clay model for a range of clay sites with different undrained shear strength and small shear stiffness profiles. This is achieved by developing a general Cowden clay model (GCCM), where 'Cowden' denotes that the soil properties incorporated in the model are based on the PISA Cowden test site utilised in the PISA project (Zdravkovic et al., 2020a, b). The new GCCM model is a development of the PISA Cowden till model (Byrne et al., 2020b) and demonstrates the potential of using data-driven 1D design models to develop general clay models.

5.2.1 Generation of representative offshore glacial clay till sites

The PISA representative offshore glacial clay till site, initially derived from the PISA Cowden clay site (Zdravkovic et al., 2020a), underwent slight adjustments to better reflect offshore conditions. These modifications, as detailed in Byrne et al. (2020b), primarily involved adapting the OCR, K_o and G_o profiles to account for the hydrostatic pore water pressure distribution found in offshore conditions. Additionally, the s_u profile of the representative offshore site was modified to maintain a constant gradient below the soil crust layer of 5m thickness (AWG, 2016).

The glacial clay tills occurring in the North Sea are typically overconsolidated as a result of past glaciation events. In this project, it is assumed consistent with previous approaches (e.g. Boulton and Dobbie, 1993) that this overconsolidation is solely attributed to the surcharge loading caused by the weight of the ice. To define the stress history of the soil, one-dimensional (1D) compression theory is employed. This approach is simplified and does not account for possible stress history effects relating to geological processes such as erosion and deposition or the influence of shear stresses applied to the soil by glacier movements. Several previous studies such as Gareau et al. (2004) and Sauer et al. (1993) suggested that 1D compression theory may not adequately capture all aspects of this phenomenon.

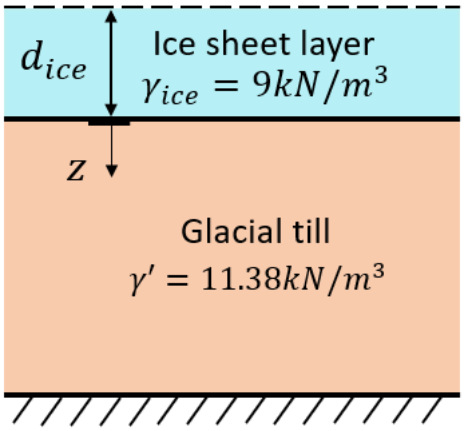


Figure 5.2: OCR determination of the proposed representative glacial clay till profiles where z describes the value of the current soil layer depth and d_{ice} is the depth of the removed ice sheet layer which caused the overconsolidation effect.

Employing 1D compression models the overconsolidation ratio, OCR, at a particular soil depth is determined as the ratio between the preconsolidation stress σ'_c and the current vertical effective stress σ'_v at that depth. The preconsolidation stress is estimated by applying an overburden pressure – corresponding to the weight of an assumed thickness of ice sheet – to the current soil surface (Figure 5.2). A similar approach was employed in Le et al. (2014) to estimate the OCR profile of the Bolders bank glacial tills at the Hornsea offshore wind farm where the OCR profile was estimated assuming a preconsolidation stress equivalent to

the weight of 25 to 75m of ice sheet thickness. The OCR profile at a depth z below the seabed is therefore defined as follows:

$$OCR = \frac{\sigma'_c}{\sigma'_v} = \frac{\gamma_{ice} \times d_{ice} + \gamma' \times z}{\gamma' \times z} \tag{5.1}$$

where $\gamma' = 11.38\ \mathrm{kN/m^3}$ (Zdravkovic et al., 2020a) is the effective unit weight of Cowden till; $\gamma_{ice} = 9\ \mathrm{kN/m^3}$ is the unit weight of an ice sheet with assumed thickness d_{ice}. The soil located below the ice sheet is assumed to be submerged (non-frozen pore water) during the assumed idealised geological process.

The coefficient of lateral earth pressure K_o is estimated by Mayne and Kulhawy (1982):

$$K_o^{OC} = K_o^{NC} \times OCR^{sin\phi'_{cs}} \tag{5.2}$$

where K_o^{NC} is defined by Jaky's formula $K_o^{NC} = 1 - sin\,\phi'_{cs}$; based on laboratory data presented in Zdravkovic et al. (2020a), the critical state friction angle is assumed to be $\phi'_{cs} = 27°$. Equation (5.2) predicts unrealistically high values of K_o for high-OCR marine clays at depths of 10m or less (Brosse et al., 2017). In the current work, therefore, the K_o profile is limited to a maximum value of 1.5-1.75.

The profile of undrained shear strength s_u under triaxial compression conditions is determined using Equation (5.3) (Potts and Zdravkovic, 1999), which is derived from the Modified Cam Clay (MCC) framework utilised in the PISA project (Zdravkovic et al., 2020a). The MCC model provides a mapping between the OCR and the s_u profile.

$$s_u = OCR \times \sigma'_{vo} \times g(\theta) \times cos(\theta) \times \frac{1+2K_o^{NC}}{6} \times (1 + B^2) \times \left[\frac{2\times(1+2K_o^{OC})}{(1+2K_o^{NC})\times OCR\times(1+B^2)}\right]^{\frac{\kappa}{\lambda}} \tag{5.3}$$

where $g(\theta)$ is a function used to define a Mohr-Coulomb hexagon for the yield surface in the deviatoric plane and $B(\theta)$ calculated by Equation (5.4) is a parameter depending on the values of $g(\theta)$ and K_o^{NC}. For triaxial compression conditions the Lode's angle is $\theta = -30°$. The intact compressibility coefficient is $\lambda = 0.115$; and the swelling coefficient is $\kappa = 0.021$ (Zdravkovic et al., 2020a).

$$B(\theta) = \frac{\sqrt{3}(1 - K_o^{NC})}{g(\theta) \times (1 + 2K_o^{NC})} \tag{5.4}$$

The small strain shear stiffness G_0 profile was determined by applying the PISA formulation for the Cowden clay site, where G_0 was correlated with the mean effective stress using the relationship $G_0 = 1100p'$ (Zdravkovic et al., 2020a). This relationship was established based on measurements of seismic cone penetration tests and bender element tests. However, it's worth noting that according to Liu et al. (2020a), the Cowden till sampled from the PISA site demonstrates stiffness anisotropy, with values of $G_{hh}/G_{vh} = 1.25$ observed up to depths of 10m. Their experiments suggest that this anisotropy is more pronounced at shallow depths and diminishes significantly with increasing depth. Nevertheless, for the newly generated representative offshore glacial clay till sites considered in this chapter, the phenomenon of stiffness anisotropy in Cowden till has been disregarded due to its occurrence at shallow depths compared to the typical length of offshore monopiles, which usually range from 20 to 60m.

Figure 5.3 illustrates the OCR, K_o, s_u and G_o distributions of a newly generated representative offshore glacial clay till site; this homogenous soil profile represents an overconsolidated clay till corresponding to a previous ice sheet of thickness $d_{ice} = 50m$. It can be observed that the generated soil profile distributions correlate well with the PISA representative offshore glacial clay till site in Byrne et al. (2020b). The main difference

between the two profiles is encountered in the top 5m, where in the PISA project a crust formation is simulated due to the discovery of a crust layer during the Cowden test site ground characterisation (Zdravkovic et al., 2020a). However, crust formation is a phenomenon usually encountered onshore due to the fluctuation of water table or to the repeated freezing-thawing cycles (Lutenegger, 1995). Additionally, it is not common to be formed up to an extent of 5m at intermediate to deep water depth at offshore sites and therefore it has been decided not to include a crust layer to the newly generated soil profiles.

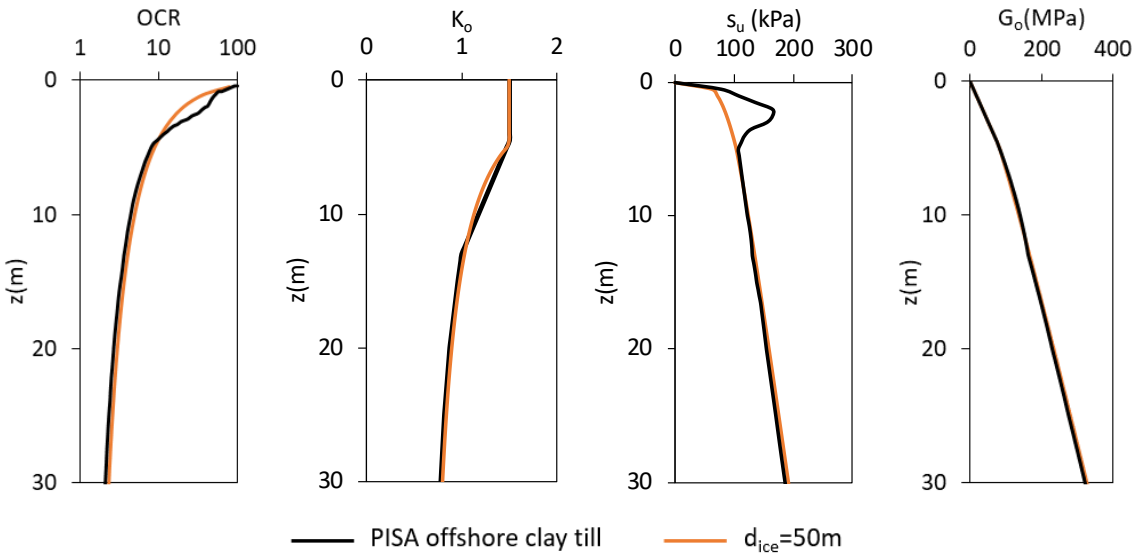


Figure 5.3: OCR, K_o, s_u and G_o distributions of a newly generated representative offshore glacial clay till site corresponding to a previous ice sheet of thickness $d_{ice} = 50m$ with the PISA representative offshore glacial clay till site (Byrne et al., 2020b).

5.2.2 Soil distributions and pile geometry calibration space

Four distinct homogenous representative offshore glacial clay till sites based on the Cowden test site employed in the PISA project (Zdravkovic et al., 2020a) with different OCR profiles have been used to formulate the GCCM, introducing:

1. The PISA representative offshore glacial clay till site (Byrne et al., 2020b) representing an overconsolidated glacial clay till with crust formation.

2. A heavily overconsolidated (OC) representative offshore glacial clay till site corresponding to a previous ice sheet of thickness $d_{ice} = 100m$.
3. A slightly overconsolidated representative offshore glacial clay till site corresponding to a previous ice sheet of thickness $d_{ice} = 25m$.
4. A normally consolidated (NC) representative offshore glacial clay till site with $OCR = 1$.

The three new representative offshore glacial clay till sites have been generated using the procedure described in Section 5.2.1. Due to the PLAXIS implementation of the NGI-ADP constitutive model implemented for the simulation of the behavior of the glacial clay till soil profiles, the G_o profile for the normally consolidated soil profile has been capped at $G_o = 2000s_u$. This is because the G_{ur}/s_u^A ratio in the NGI-ADP model is limited to 2000, meaning that it is impossible to simulate higher values of G_o for the NC offshore glacial clay till site. The K_o profile is limited near to the soil surface to a value of $K_o = 1.75$ for the heavily overconsolidated glacial clay till and it is limited to a value of $K_o = 1.5$ in the case of the slightly overconsolidated offshore glacial clay till (Brosse et al., 2017). The higher limit value of K_o profile adopted for the more heavily overconsolidated glacial clay till sites that maintain higher OCR values, will demonstrate higher values of G_o. The K_o profile for the NC soil profile was determined by using the Jaky's formula presented in Section 5.2.1. For all the representative offshore glacial clay till sites fully saturated soil conditions were established. The submerged unit weight was set equal to $\gamma' = 11.38\ kN/m^3$ based on the PISA Cowden test site (Zdravkovic et al., 2020a).

Figure 5.4 presents the OCR, K_o, s_u and G_o distributions of the four offshore representative glacial clay till sites, which form the GCCM. These four profiles span a range of typical clay till profiles at offshore wind farms (Le et al., 2014). It is expected that the glacial till profiles will show significant variability across a wind farm site (Clarke, 2018).

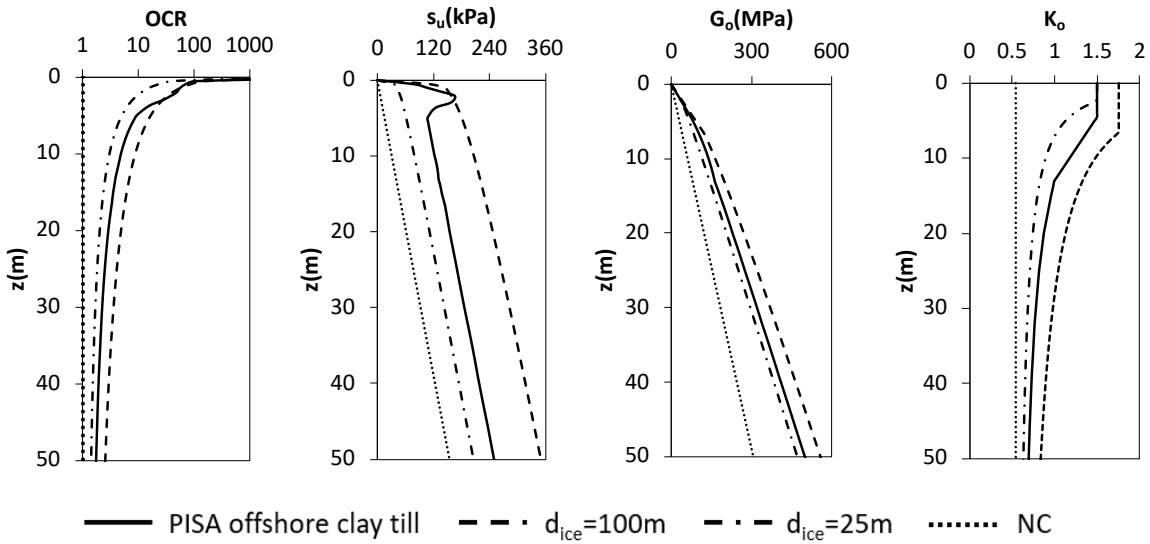


Figure 5.4: OCR, s_u, G_o and K_o, profiles adopted for the four offshore representative offshore glacial clay till sites used to calibrate the GCCM. This set of profiles includes the representative offshore glacial clay till site employed in the PISA project (Byrne et al., 2020b) referred as PISA offshore clay till.

The pile geometry calibration space adopted for each representative glacial clay till site is presented in Table 5.1, employing nine of the eleven calibration piles employed in Byrne et al. (2020b) which were used calibrate the PISA Cowden till model. Two of the monopiles presented in Byrne et al. (2020b) were not considered, since they only differed in the monopile wall thickness, and it was found that the monopile thickness does not significantly influence the soil reaction curves.

Table 5.1: PISA offshore pile geometries (Byrne et al., 2020b). D is the pile diameter; L is the embedded length; t_w is the pile wall thickness; h is the load eccentricity.

Pile	$D(m)$	$L(m)$	$t_w(m)$	$h(m)$	L/D	h/D
C1	10	20	0.091	50	2	5
C2	10	20	0.091	150	2	15
C4	10	60	0.091	50	6	5
C5	10	60	0.091	150	6	15
C6	5	10	0.045	25	2	5
C8	5	30	0.045	25	6	5
C9	5	30	0.045	75	6	15
C10	7.5	15	0.068	37.5	2	5
C11	7.5	45	0.068	37.5	6	5

5.2.3 3D finite element analyses calibration procedure

The representative offshore glacial clay till sites were simulated using the NGI-ADP constitutive model. A detailed constitutive model calibration procedure corresponding to the PISA representative offshore glacial clay till site (Byrne et al., 2020b) is presented in Chapter 3, where it was shown that the NGI-ADP model is capable of accurately capturing the pile monotonic lateral response for both the PISA Cowden field test piles and the offshore pile geometries that were specified in Byrne et al. (2020b). Therefore, the same calibration procedure was followed to determine the NGI-ADP parameters for the three new representative offshore glacial clay till sites. The selection of the total number of sublayers generated in PLAXIS 3D was carefully implemented for each soil profile to capture a close fit of the simulated G_o, s_u, K_o profiles. Thirteen sub-layers were selected to simulate the two new overconsolidated representative offshore glacial clay till sites corresponding to $d_{ice} = 100m$ and $d_{ice} = 25m$. A single sublayer is sufficient for the normally consolidated representative offshore glacial clay till since both the G_o and s_u profiles increase linearly with depth and the K_o profile is constant.

The NGI-ADP soil parameters used in PLAXIS 3D for the simulation of the three clay profiles are presented in Table 5.2 to Table *5.4*. Figure 5.5 illustrates the simulated G_o, s_u, K_o profiles in comparison to the representative offshore glacial clay till sites. The s_u, K_o profiles simulated by the NGI-ADP model match accurately the profiles for all new representative offshore glacial clay tills. However, since the G_o profile simulated by the NGI-ADP model depends on G_{ur}/s_u^A (where $s_u^A = s_u/0.99$), which is limited to a maximum of 2000, the simulated G_o profile underpredicts the representative G_o profile at depths where $G_o > 2000s_u$ (e.g. depth below than 30m at the slightly OC clay till with $d_{ice} = 25m$). The G_o profile dictates the pile's initial stiffness response which is determined at the small displacements range where the pile deflection is mostly concentrated at the upper part of the

pile. Therefore, the small underprediction in the simulated G_o profile for the slightly OC clay till can be considered as insignificant for the computed pile behaviour.

Table 5.2: NGI-ADP input parameters for the heavily OC representative offshore glacial clay till site for $d_{ice} = 100m$.

$z(m)$	K_o	s_u^A	s_{uinc}^A	G_{ur}/s_u^A	τ_0/s_u^A	γ_f^C	γ_f^E	γ_f^{DSS}
0-1	1.75	0.1	146.1	128.3	0	54.5	109.1	81.8
1-2	1.75	146.2	13.3	184.3	0	43.4	86.8	65.1
2-4	1.75	159.5	8.4	335.5	0	32.8	65.6	49.2
4-6	1.75	176.3	6.2	480.1	0	22.9	45.8	34.4
6-8	1.75	188.7	5.25	602.4	0	18.3	36.5	27.4
8-10	1.538	199.2	4.75	732.6	0	15.0	30.0	22.5
10-12	1.481	208.7	4.4	828.7	0	13.3	26.5	19.9
12-16	1.296	217.5	4.1	928.5	0	11.8	23.7	17.8
16-20	1.179	233.9	3.83	1043.1	0	10.5	21.1	15.8
20-30	1.048	249.2	3.59	1206.1	0	9.1	18.2	13.7
30-45	0.917	285.1	3.38	1426.3	0.057	7.7	15.4	11.6
45-60	0.831	335.8	3.27	1617.9	0.140	6.8	13.6	10.2
60-100	0.748	384.8	3.18	1855.7	0.256	5.9	11.9	8.9

Table 5.3. NGI-ADP input parameters for the slightly OC representative glacial clay till site for $d_{ice} = 25m$.

$z(m)$	K_o	s_u^A	s_{uinc}^A	G_{ur}/s_u^A	τ_0/s_u^A	γ_f^C	γ_f^E	γ_f^{DSS}
0-1	1.5	0.1	41.7	398.0	0	27.6	55.3	41.5
1-2	1.5	41.8	7.5	549.1	0	20.0	40.1	30.0
2-4	1.377	49.3	4.6	844.9	0	13.0	26.0	19.5
4-6	1.134	58.6	3.8	1090.8	0	10.1	20.2	15.1
6-8	1.008	66.1	3.5	1262.9	0	8.7	17.4	13.1
8-10	0.929	73.2	3.4	1399.4	0.048	7.9	15.7	11.8
10-12	0.874	80.1	3.3	1511.7	0.095	7.3	14.6	10.9
12-16	0.817	86.8	3.3	1647.9	0.156	6.7	13.4	10.0
16-20	0.767	99.9	3.2	1788.9	0.225	6.1	12.3	9.2
20-30	0.713	112.8	3.2	1965.5	0.317	5.6	11.2	8.4
30-45	0.663	144.6	3.1	1986.8	0.428	5.5	11.1	8.3
45-60	0.632	191.8	3.1	1985.4	0.510	5.5	11.1	8.3
60-100	0.604	238.8	3.1	1983.9	0.598	5.5	11.1	8.3

Table 5.4. NGI-ADP input parameters for the NC representative offshore glacial clay till site.

$z(m)$	K_o	s_u^A	s_{uinc}^A	G_{ur}/s_u^A	τ_0/s_u^A	γ_f^C	γ_f^E	γ_f^{DSS}
0-100	0.546	0.1	3.08	2000	0.8378	5.5	11.0	8.25

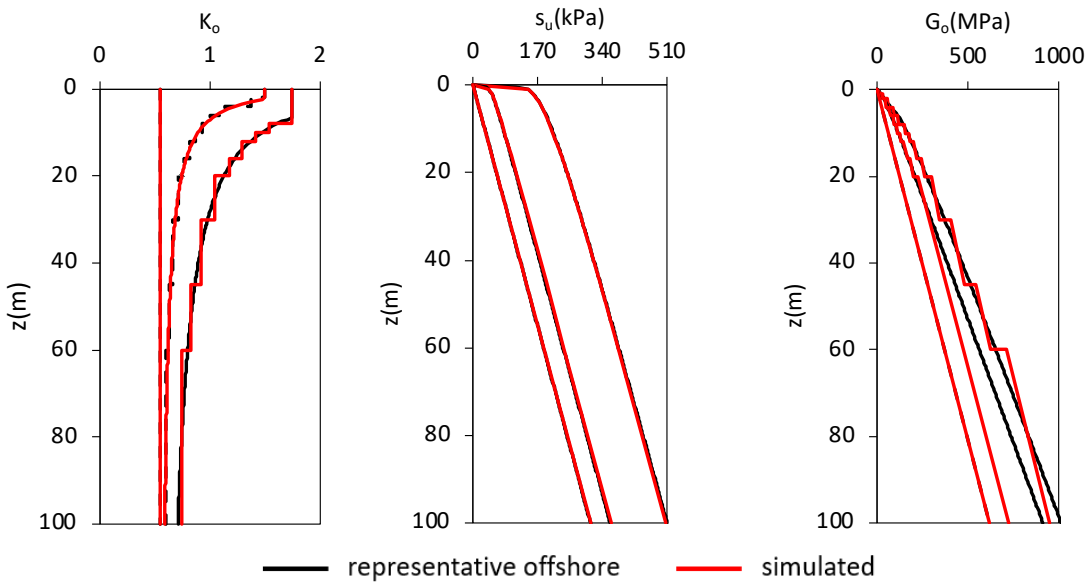


Figure 5.5: Simulated K_o, s_u, G_o profiles by the NGI-ADP model in comparison to the newly generated representative offshore glacial clay till sites presented in Section 5.2.2.

5.2.4 Site-specific glacial clay till machine learning models

Site-specific machine learning models have been generated by using the data-driven method presented in Chapter 4 for each representative offshore glacial clay till site that is presented in Section 5.2.2. The training inputs for a clay site-specific machine learning model are $[z/L]$, $[L/D]$, $[G_o/s_u]$. The G_o and s_u values of the representative offshore glacial clay till sites were employed for the training procedure.

5.2.4.1 Training of the site-specific clay till machine learning models

The four representative offshore glacial clay till sites soil profile distributions, the pile geometries specified in Section 5.2.2 and the normalised knot point soil reaction parameters were added to a newly generated SQL database (Suryasentana et al., 2019) by using the DB Browser for SQLite (DB4S) open source tool (Hipp, 2020). Four site-specific clay machine learning models were trained using the procedure in Section 4.6.2. The Gaussian process regression (GPR) machine learning model employing the Matérn kernel ($v = 5/2$) plus the white noise kernel were adopted. An 80:20 training-test set split ratio was adopted. The MSE and R^2 train and test scores were used to assess each GPR trained model.

5.2.4.2 Validation of the site-specific clay till machine learning models

The GPR ML models produce predictions of the depth variation for each of the knot point parameters of the PISA type four soil reaction components. By using the 1D FE framework together with the eight-knot point piecewise spline function with PCHIP interpolation scheme in the OxPile software, presented in Section 4.7, the site-specific machine learning models' predictions can compute the pile monotonic lateral behaviour. For each one of the nine monopile geometries used for the training of the site-specific (ssp) machine learning models, the generated pile response has been validated with PLAXIS 3D FEA calculations.

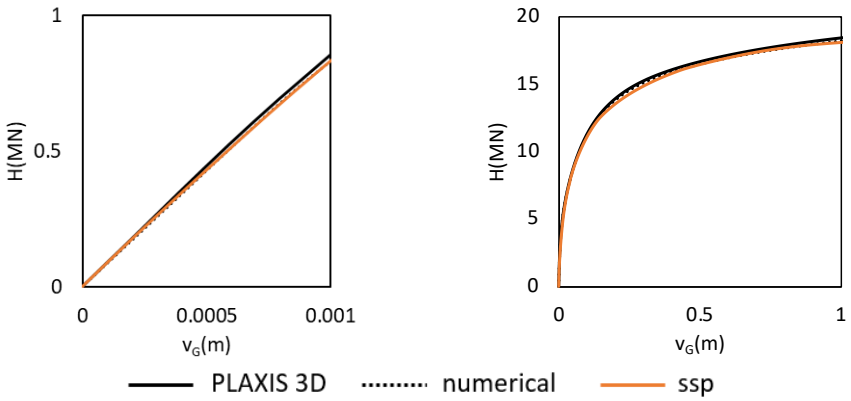


Figure 5.6: Load ground-level displacement response for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacement ranges for the C1 pile embedded in a highly OC representative offshore glacial clay till site corresponding to $d_{ice} = 100m$.

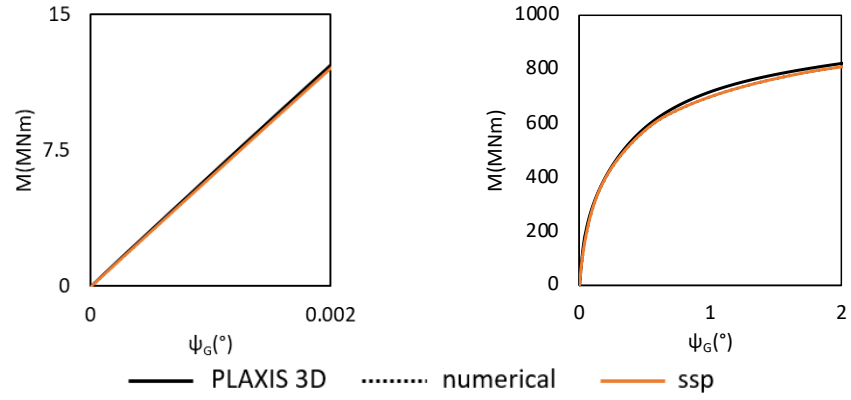


Figure 5.7: Moment ground-level rotation response for the small ($\psi_G = 0.002°$) and ultimate ($\psi_G = 2°$) rotation ranges for the C1 monopile embedded in a highly OC representative offshore glacial clay till site corresponding to $d_{ice} = 100m$.

Figure 5.6 and Figure 5.7 provide comparisons of the pile load-displacement and moment-rotation response at ground level computed from the data-driven 1D design model and 3D FEA regarding the C1 pile embedded in the heavily OC representative glacial clay till with $d_{ice} = 100m$. The pile response calculated by the numerical soil reaction curves in OxPile software are also added in the graphs. The site-specific (ssp) data-driven 1D design model computes precisely both the pile load-displacement and moment-rotation response at ground level. A good match was detected for the total set of monopile geometries. Due to the high-fidelity representation of the soil reaction curves with the new piecewise soil reaction curve representation and the capability of the GPR machine learning model to closely represent the non-linear depth variation of the soil reaction curve knot point parameters, the data-driven site-specific 1D design models will consistently predict with good accuracy both the pile load-displacement and moment-rotation response.

To quantify the accuracy of the site-specific data-driven 1D design models' predictions with the PLAXIS 3D results, the assessment metrics adopted in the PISA project (Equation (4.20)) have been employed. Table 5.5 and Table 5.6 present the η accuracy metric and ρ ratio metric (Equation (4.20)) of each site-specific clay till 1D model for all the tested pile geometries for small displacement ($v_G = D/10000$) and ultimate ($v_G = D/10$) response comparing the load-displacement response predicted from the data-driven 1D design models to the PLAXIS 3D calculations. The small displacement metric for all calibration piles is in the range $0.923 \leq \eta_{sd} \leq 0.991$; the ultimate displacement metric is in the range $0.947 \leq \eta_{ult} \leq 0.995$; these results indicate a good match with the calibration data. Furthermore, the ρ metric data indicate that there is no significant overall bias. At the small displacement range the coefficient of variance (CoV) ranges between $1.479 \leq CoV \leq 4.133\%$ and the mean μ ranges between $0.967 \leq \mu \leq 0.999$; At the ultimate load CoV ranges between $0.945 \leq CoV \leq 1.559\%$ and the mean ranges between $0.972 \leq \mu \leq 0.995$.

Table 5.5: Values of η accuracy metric for each site-specific representative offshore glacial clay till data-driven 1D design model for the small displacement and ultimate response.

Pile	PISA Cowden clay		d_{ice}=100m		d_{ice}=25m		N.C. Cowden clay	
	η_{sd}	η_{ult}	η_{sd}	η_{ult}	η_{sd}	η_{ult}	η_{sd}	η_{ult}
C1	0.946	0.991	0.980	0.980	0.939	0.984	0.981	0.991
C2	0.957	0.970	0.978	0.962	0.952	0.986	0.950	0.979
C4	0.965	0.972	0.986	0.954	0.989	0.985	0.935	0.986
C5	0.953	0.965	0.964	0.947	0.923	0.975	0.931	0.973
C6	0.987	0.977	0.990	0.962	0.992	0.954	0.962	0.995
C8	0.987	0.978	0.957	0.963	0.982	0.984	0.953	0.994
C9	0.973	0.965	0.960	0.957	0.975	0.969	0.956	0.979
C10	0.956	0.967	0.961	0.986	0.967	0.987	0.962	0.990
C11	0.952	0.967	0.939	0.961	0.972	0.991	0.956	0.995
Average	0.964	0.972	0.968	0.964	0.966	0.979	0.954	0.987

Table 5.6: Values of ρ ratio metric for each site-specific representative offshore glacial clay till data-driven 1D design model for the small displacement and ultimate response.

Pile	PISA Cowden clay		d_{ice}=100m		d_{ice}=25m		N.C. Cowden clay	
	ρ_{sd}	ρ_{ult}	ρ_{sd}	ρ_{ult}	ρ_{sd}	ρ_{ult}	ρ_{sd}	ρ_{ult}
C1	1.004	0.995	0.980	0.978	1.060	0.986	0.963	1.003
C2	0.968	0.970	0.986	0.968	1.042	0.988	0.958	0.985
C4	0.970	0.976	0.994	0.966	1.008	0.990	0.944	0.992
C5	0.960	0.986	0.966	0.962	0.928	0.984	0.956	0.984
C6	1.005	0.967	0.993	0.970	0.992	0.946	0.966	1.005
C8	0.993	0.975	0.980	0.970	0.981	0.984	0.973	0.993
C9	0.977	0.963	0.971	0.964	0.975	0.968	0.966	0.977
C10	1.015	0.963	0.965	0.992	1.033	0.985	1.014	1.02
C11	0.968	0.963	0.950	0.976	0.970	0.998	0.964	0.997
Average	0.984	0.973	0.976	0.972	0.999	0.981	0.967	0.995
CoV (%)	2.037	1.155	1.479	0.945	4.133	1.559	2.000	1.306

Finally, a comparison between the performance of the 1D models developed by the data-driven method and the PISA design model in terms of predicting the pile load-displacement response is presented for the PISA representative offshore glacial clay till site (Byrne et al., 2020b). Figure 5.8 and Figure 5.9 present the values of the η accuracy metric and ρ ratio metric (Equation (4.20)) of the nine calibration piles for the load-displacement response determined by the site-specific data-driven 1D design model compared to the accuracy metrics presented in Byrne et al. (2020b) calculated by the PISA Cowden till model.

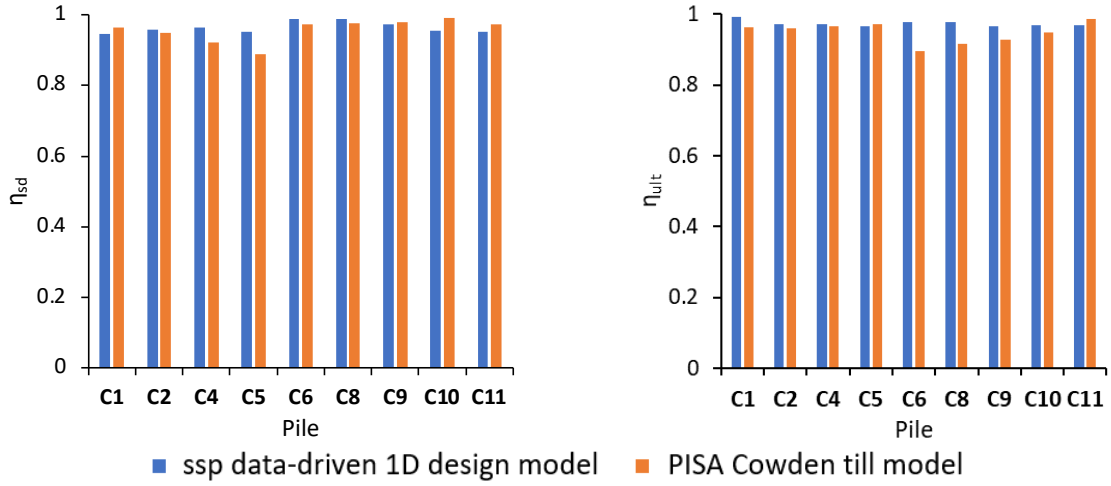


Figure 5.8: Accuracy metrics of the site-specific (ssp) data-driven 1D design model and the PISA Cowden till model (Byrne et al., 2020b) for the nine calibration piles embedded in PISA representative offshore glacial clay till site for small displacement and ultimate response.

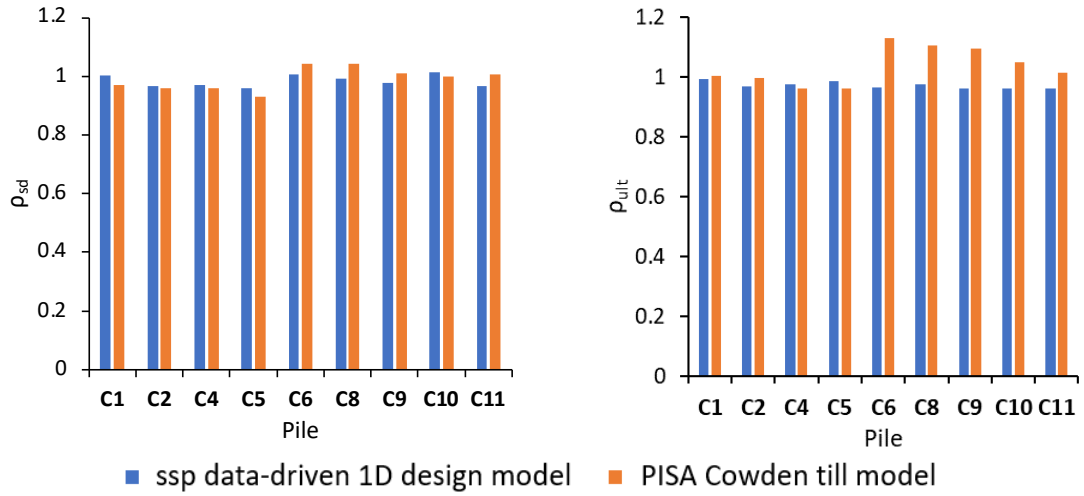


Figure 5.9: Ratio metrics of the site-specific (ssp) data-driven 1D design model and the PISA Cowden till model (Byrne et al., 2020b) for the nine calibration piles embedded in PISA representative offshore glacial clay till site for small displacement and ultimate response.

Both 1D model approaches provide high accuracy metrics, with the data-driven 1D design model slightly outperforming the PISA design model. The data-driven 1D design model achieves average accuracy metrics of $\eta_{sd} = 0.964$ and $\eta_{ult} = 0.972$ for the nine calibration piles, while the PISA rule-based design method attains average values of $\eta_{sd} = 0.957$ and $\eta_{ult} = 0.948$. This is based: (i) on the ability of the GPR machine learning model to rigorously capture the depth variation of the soil reaction parameters, (ii) on the better representation of the soil reaction curves which is achieved by the eight-knot point soil reaction curve parametrisation. Finally, it is important to mention the data-driven method will consistently reproduce the monopile's moment rotation response generated by the 3D FEA in similar accuracy to the pile's load-displacement response. This is not straightforward with the PISA Cowden till model where a second stage optimisation is applied using solely the pile's load-displacement response (Byrne et al., 2020b) to improve the model's accuracy.

5.2.5 Generation of data-driven general Cowden clay model (GCCM)

The soil reaction curves used to generate the four site-specific offshore glacial clay till data-driven 1D design models presented in the previous section were used to develop a general Cowden clay model (GCCM). The GCCM incorporates the soil reaction curves extracted from a total of 36 FEA of laterally loaded monopiles embedded in different OCR distributions of representative offshore glacial clay till sites. Two different versions of the GCCM were generated: (i) GCCM V1 where the soil reaction curves were normalised using the 'clay' data-driven normalisation framework in Table 4.3, (ii) GCCM V2 where the soil reaction curves were normalised using the 'sand' data-driven normalisation framework in Table 4.3; for this case s_u is incorporated solely in the machine learning model. GCCM V2 was mainly employed to check whether the sand normalisation can be applied successfully to clay soils and to confirm its use within the clay layers in layered soils (presented in Chapters 6, 7 of this thesis). The general clay data-driven 1D design models incorporate an

additional training input $[\sigma'_v/s_u]$. This additional feature is closely associated with the overconsolidation ratio of the soil. The GPR ML model with the same kernel function used for the site-specific data-driven 1D design models was employed to train the GCCM. An 80:20 training-test set split ratio was adopted. All the trained GPR models were assessed using the MSE and R^2 train and test scores. Figure 5.10 presents the depth variation of G_o/s_u and σ'_v/s_u which comprise the soil strength inputs of the machine learning model for the four representative sites used to calibrate the GCCM.

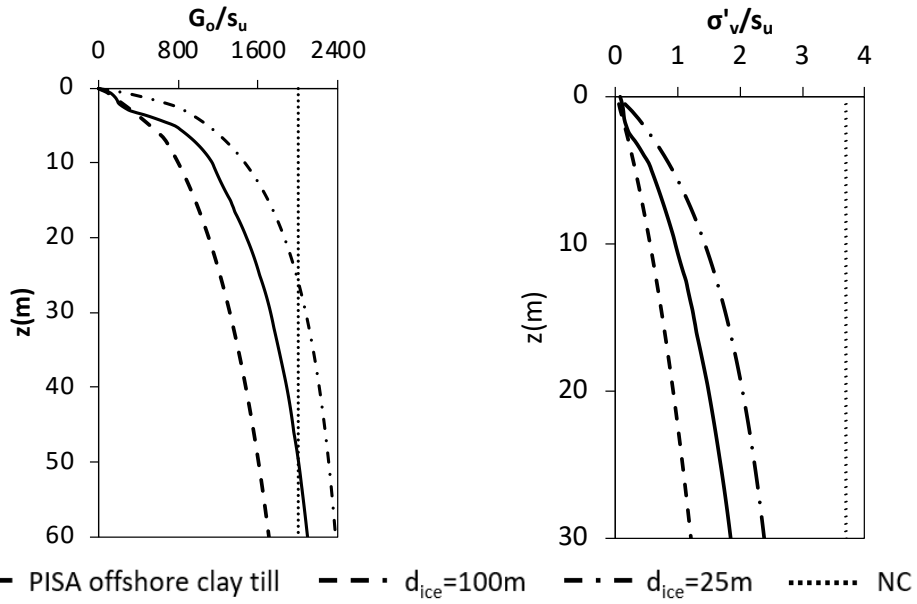


Figure 5.10: Variation with depth of G_o/s_u and σ'_v/s_u for the representative offshore glacial clay till sites which are used to calibrate the GCCM.

Table 5.7 illustrates the optimised hyperparameters obtained for the knot point soil reaction curve parameters of the distributed soil reactions $\bar{p}$, $\bar{m}$ after the GPR models training corresponding to the GCCM V1. The low value of the optimised lengthscale $l_{\sigma'_v/s_u}$ indicates that the GPR model identifies the $[\sigma'_v/s_u]$ input as important to discriminate the soil reaction curves of clay soils with different stress history. The same conclusion came out of the training procedure of GCCM V2, where a low value of $l_{\sigma'_v/s_u}$ was also achieved.

Table 5.7: Optimised kernel parameters for the GPR models as applied to the knot point parameters associated with the GCCM V1. (σ_f is the signal variance of the Matérn kernel, σ_n is the variance of the noise kernel, l is the lengthscale for every input dimension).

Knot point	$\bar{p}$	$\bar{m}$
1	σ_f=0.29, σ_n^2=5.64e-4, $l_{z/L}$=0.23, $l_{L/D}$=4.94, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=2.42	σ_f=0.0473, σ_n^2=2.64e-5, $l_{z/L}$=0.323, $l_{L/D}$=15.8, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=3.92
2	σ_f=0.493, σ_n^2=1.88e-3, $l_{z/L}$=0.203, $l_{L/D}$=7.21, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=3.46	σ_f=0.123, σ_n^2=5.86e-5, $l_{z/L}$=0.309, $l_{L/D}$=21.0, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=4.43
3	σ_f=1.56, σ_n^2=0.0399, $l_{z/L}$=0.16, $l_{L/D}$=15.7, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=4.91	σ_f=0.188, σ_n^2=2.43e-4, $l_{z/L}$=0.240, $l_{L/D}$=8.53, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=7.64
4	σ_f=1.95, σ_n^2=0.0714, $l_{z/L}$=0.116, $l_{L/D}$=17.4, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=3.89	σ_f=0.188, σ_n^2=5.19e-4, $l_{z/L}$=0.219, $l_{L/D}$=9.73, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=6.14
5	σ_f=2.65, σ_n^2=0.125, $l_{z/L}$=0.111, $l_{L/D}$=15.5, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=3.48	σ_f=0.239, σ_n^2=6.80e-4, $l_{z/L}$=0.257, $l_{L/D}$=15.7, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=11.0
6	σ_f=3.55, σ_n^2=0.200, $l_{z/L}$=0.123, $l_{L/D}$=20.1, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=4.10	σ_f=0.331, σ_n^2=9.15e-4, $l_{z/L}$=0.244, $l_{L/D}$=16.3, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=1.45
7	σ_f=3.96, σ_n^2=0.245, $l_{z/L}$=0.133, $l_{L/D}$=19.5, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=4.06	σ_f=0.337, σ_n^2=9.01e-4, $l_{z/L}$=0.246, $l_{L/D}$=17.1, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=1.42
8	σ_f=4.11, σ_n^2=0.261, $l_{z/L}$=0.137, $l_{L/D}$=18.6, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=4.06	σ_f=0.342, σ_n^2=9.08e-4, $l_{z/L}$=0.245, $l_{L/D}$=17.5, l_{G_o/s_u}=1000, $l_{\sigma'_v/s_u}$=1.40

5.2.6 Applications of the GCCM

An initial exercise has been conducted in which predictions from the GCCM V1 and GCCM V2 data-driven 1D design models are compared with the original training data. This comparison indicated that both models provide a close fit to the calibration analyses. This section demonstrates the capabilities of both 1D models to accurately compute the pile load-displacement response with respect to:

1. Design offshore monopiles embedded in a soil profile that has been used for the training of the ML models falling within the L/D geometrical calibration space.
2. Design offshore monopiles geometries embedded in unseen representative offshore glacial clay till sites which have not been used for the training of the ML models.
3. Offshore monopiles geometries which fall outside of the trained L/D geometrical calibration space of the GCCM embedded in a trained soil profile.

4. Smaller monopile geometries such as the PISA and PICASO Cowden field test monopiles that fall out of the geometrical calibration space of the GCCM.

For all cases the computed load-displacement pile response from the predicted soil reaction curves has been validated with the calculations of 3D FEA. The finite element mesh for all the validation 3D analyses consisted approximately of 100000 to 120000 elements. For all the application cases the results of the PISA design model employing the PISA Cowden till model (Byrne et al., 2020b) were added to the figures to check the efficiency of the pre-existing PISA Cowden till model to compute the monopile's response for piles embedded in offshore glacial clay tills with different OCR profiles; the confidence in using the PISA Cowden till model along with 3D FE calculations computed by the NGI-ADP model comes from the relatively good capability of the NGI-ADP model to capture the PISA FEA pile's monotonic response (presented in Chapter 3).

The pile moment-rotation response at ground level is usually used for design purposes of the monopiles foundations of the OWTs. It was found for all examined design cases that the moment ground-level rotation response shows a similar level of accuracy to the computed pile load ground-level displacement response. Therefore, for brevity the moment-rotation graphs are not presented in this thesis although the computed accuracy metric η and ratio metric ρ for the piles' moment-rotation response are displayed in the Section 5.2.7.

5.2.6.1 Design offshore monopiles embedded in trained clay till profiles

As a starting check the data-driven 1D design models are tested with respect to offshore design monopile geometries falling within the L/D geometrical calibration space embedded in a trained offshore glacial clay till site. A design monopile with $L/D = 3$ was adopted. The monopile's geometrical characteristics were introduced in the PISA project (Byrne et al., 2020b; Burd et al., 2020) and are presented in Table 5.8. For all the examined design cases,

the pile load-displacement response computed by using the site-specific (ssp) 1D model produced for the particular trained soil profile will be also presented for further validation of the GCCM V1 and GCCM V2 general clay 1D models.

Table 5.8: D1 Design offshore monopile geometry.

Pile	$D(m)$	$L(m)$	$t_w(m)$	$h(m)$	L/D	h/D	t_w/D
D1	7.5	37.5	0.068	37.5	3	5	0.00907

Figure 5.11 shows the load ground-level displacement response for pile D1 for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements ranges computed by the GCCM V1 and GCCM V2. As can be observed, both versions of the GCCM match the 3D FEA predictions. Furthermore, it can be seen that the results of the GCCM V1 and GCCM V2 match the behaviour of the individual site-specific data-driven 1D design models.

The PISA Cowden till model (Byrne et al., 2020b) is capable of reproducing the pile load-displacement response for the D1 pile geometry embedded in an overconsolidated representative offshore glacial clay till site. It lacks some accuracy in terms of calculating the pile's small displacement response for overconsolidated glacial clay till sites that have a different G_o distribution compared to the PISA representative offshore glacial clay till site. Notably, the model significantly underestimates the pile's ultimate response for the NC representative offshore glacial clay till site. This discrepancy is closely linked to the undrained shear strength of the soil, suggesting that a PISA design model should not be applied to clay soil profiles with substantially different undrained shear strength profiles compared to the original calibration site. Additionally, the PISA Cowden till model tends to overestimate the pile's small displacement response for the NC representative offshore glacial clay till site, potentially due to the limited G_o profile in the soil model, capped at $G_o = 2000 s_u$.

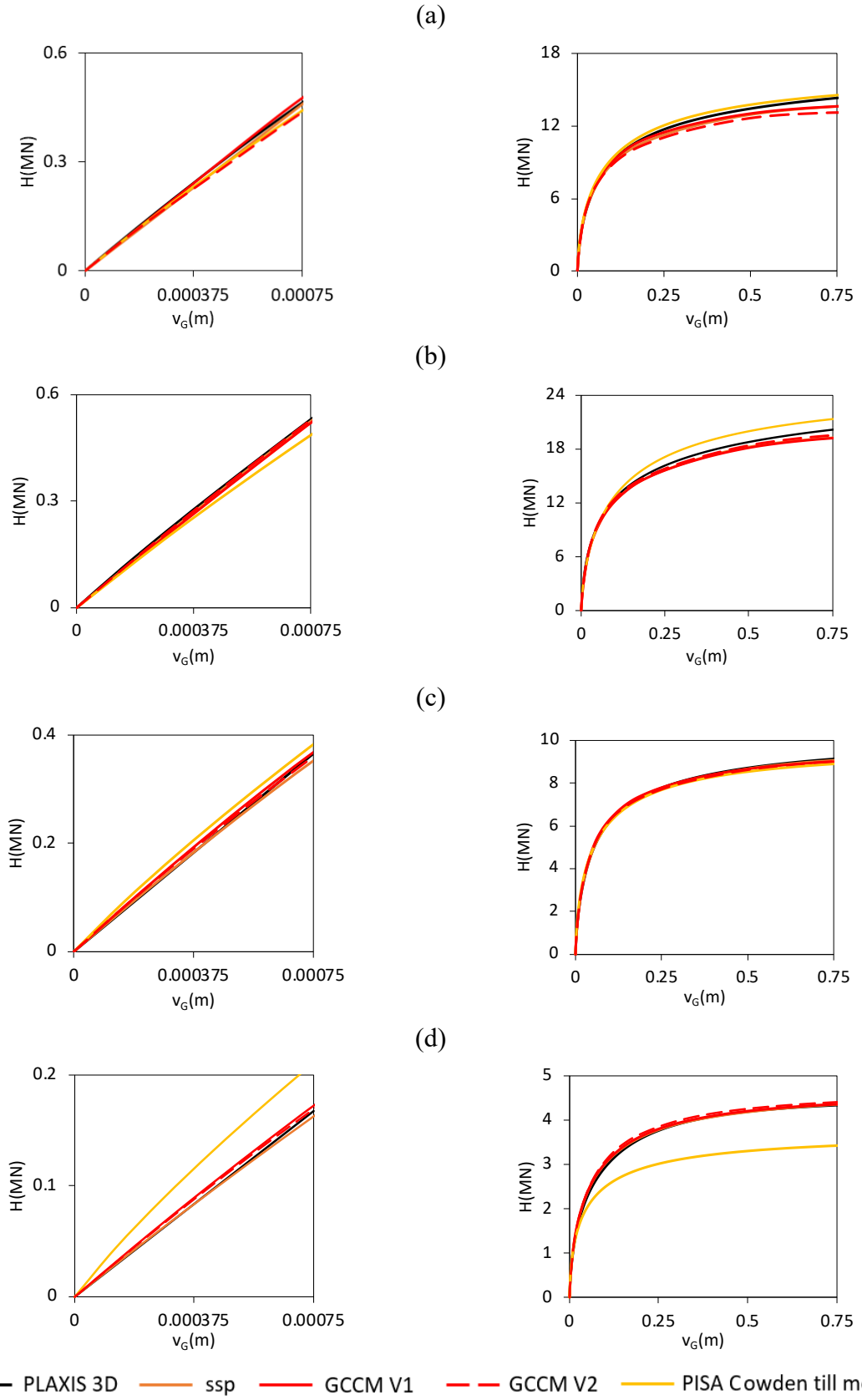


Figure 5.11: Load ground-level displacement response for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements ranges for the pile D1 embedded in: (a) the PISA representative offshore glacial clay till site, (b) the heavily OC representative offshore glacial clay till site corresponding to d_{ice}=100m, (c) the slightly OC representative offshore glacial clay till site corresponding to d_{ice}=25m, (d) the NC representative offshore glacial clay till site.

5.2.6.2 Offshore monopiles embedded in unseen clay till profiles

Two new representative offshore glacial clay till sites are introduced using the procedure analysed in Section 5.2.1; an overconsolidated representative offshore glacial clay till site corresponding to a previous ice sheet of thickness $d_{ice} = 75m$ and a heavily overconsolidated representative offshore glacial clay till site corresponding to $d_{ice} = 125m$. Figure 5.12 shows the variation with depth of the G_o/s_u and σ'_v/s_u ratios of the two unseen representative offshore glacial clay till sites compared to the corresponding ratios of the sites which form the GCCM. These two ratios comprise the soil strength and stiffness training inputs of the GCCM. For both design soil profiles, K_o was limited to a maximum value of 1.75 in the top 10 metres. As can be observed, both unseen soil profiles demonstrate ratios which approximately follow the trend of the trained soil strength inputs of the overconsolidated Cowden clay profiles. However, the glacial clay till profile with an OCR ratio corresponding to $d_{ice} = 125m$ defines an unusual case where the soil strength and stiffness falls outside of the calibration space of the GCCM.

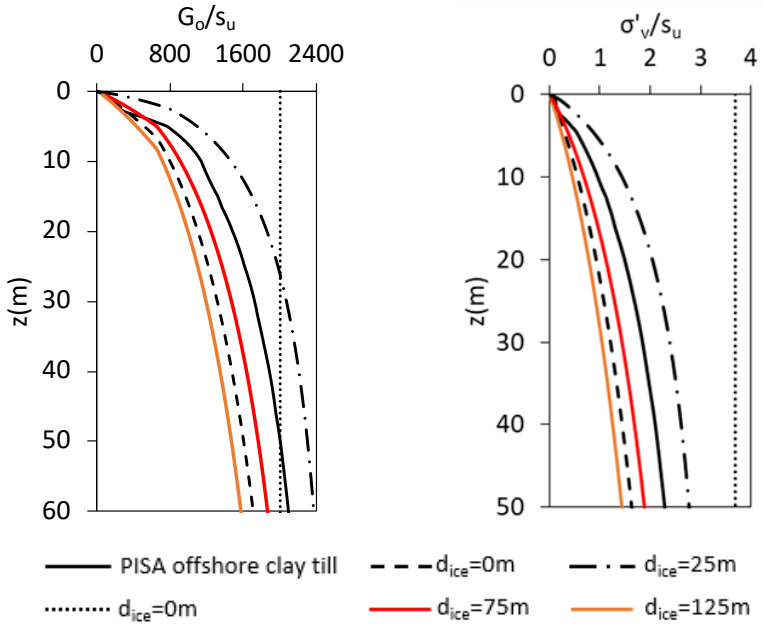


Figure 5.12: Variation with depth of G_o/s_u and σ'_v/s_u of the two unseen representative offshore glacial clay till sites with OCR ratios corresponding to $d_{ice} = 75m$ and $d_{ice} = 125m$ with the representative offshore glacial clay till sites consisting the GCCM.

Figure 5.13 presents the load-displacement response at ground level for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements range computed by the GCCM V1 and GCCM V2 for the D1 pile (Table 5.8) embedded in the unseen representative offshore glacial clay till sites. Both versions of the GCCM data-driven 1D design model provide a good fit to the PLAXIS 3D calculations. The PISA Cowden till model slightly overpredicts the ultimate pile response for both overconsolidated offshore glacial clay till sites; however, it still provides acceptable accuracy metrics in the range of $92.7\% \leq \eta_{ult} \leq 93.9\%$. Furthermore, it can be noticed that the PISA Cowden till model slightly underpredicts the foundation small displacement response for both unseen overconsolidated offshore glacial clay till sites.

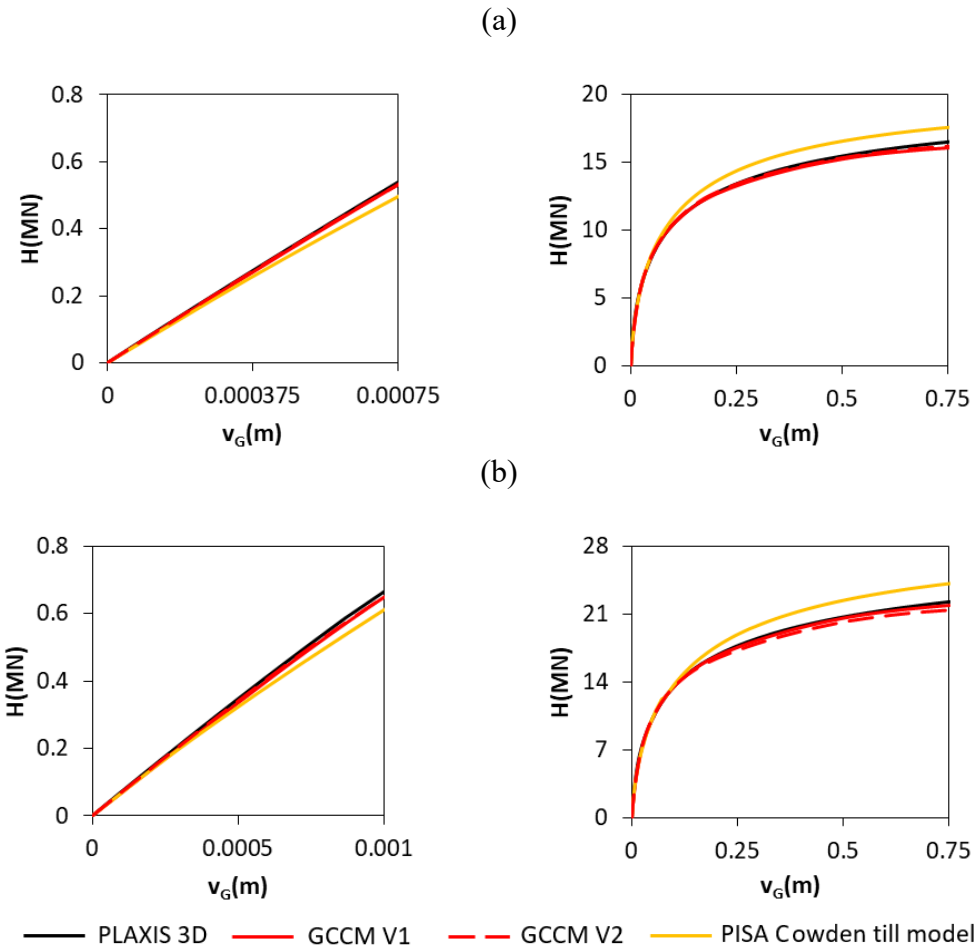


Figure 5.13 Load ground-level displacement response for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements range for the D1 pile embedded in: (a) the OC representative offshore glacial clay till site corresponding to $d_{ice} = 75m$, (b) the OC representative offshore glacial clay till site corresponding to $d_{ice} = 125m$.

5.2.6.3 Offshore monopiles geometries falling out of the trained *L/D* geometry calibration space embedded in a trained site

In this section, the GCCM is evaluated for offshore monopiles geometries that fall out of the trained L/D calibration space and are embedded in a trained site. For this design scenario the GPR model has to extrapolate in the $[L/D]$ input dimension. Generally, GPR models are able to provide reliable predictions by extrapolating up to one unit of lengthscale away from the given data. For example, Suryasentana and Sheil (2023) showed the extrapolating capabilities of the GPR model in terms of forecasting the accumulated axial strains of a clay sample under undrained cyclic triaxial loading.

A slender monopile (S1), with $L/D = 8$ has been selected for this task and its geometry is presented in Table 5.9. Two sets of analyses were conducted for the S1 monopile embedded in the heavily OC representative offshore glacial clay till site corresponding to $d_{ice} = 100m$ and in the NC representative offshore glacial clay till site. Monopiles with this type of geometry are not very commonly designed in the new offshore wind farms where more rigid monopiles are favoured. However, monopiles with geometries with L/D values of up to 10 or more were adopted for early wind farms (Byrne, 2021a).

Table 5.9: Slender offshore design monopile geometry.

Pile	$D(m)$	$L(m)$	$t_w(m)$	$h(m)$	L/D	h/D	t_w/D
S1	5	40	0.045	75	8	15	0.009

Figure 5.14 presents the load-displacement response at ground level for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements range calculated by the GCCM V1 and GCCM V2 data-driven 1D design models for the S1 monopile embedded in the representative offshore glacial clay till site corresponding to $d_{ice} = 100m$ and the NC representative offshore glacial clay till site. The GCCM V1 and GCCM V2 results are

compared to the PLAXIS 3D calculations. Both GCCM versions can calculate with good accuracy the pile's load-displacement response. Therefore, the advanced GPR model allows the designer to make predictions for pile geometries with L/D falling marginally outside of the trained geometry calibration space; this feature is not available with the PISA design models which are not advised to be used outside the pile geometry calibration space. However, it is advised not to use the data-driven 1D design models for pile geometries falling outside the calibration space without a validation 3D FEA.

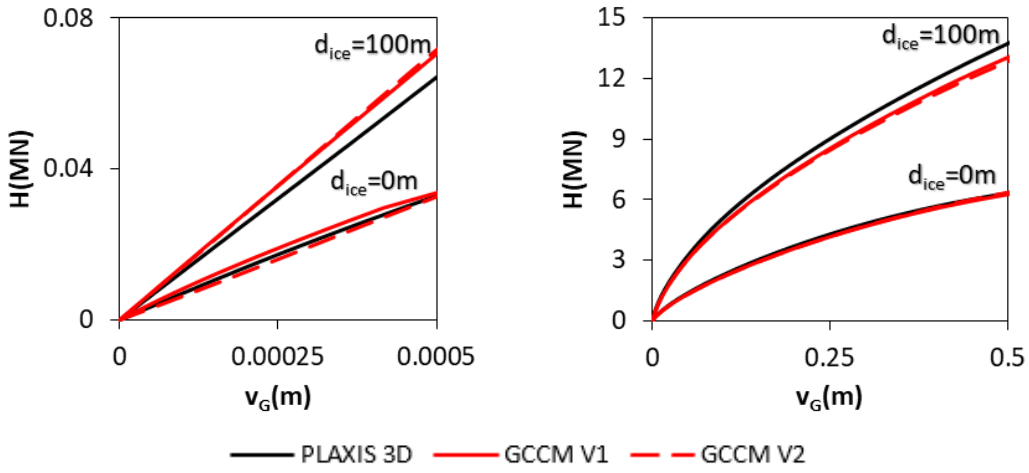


Figure 5.14: Load ground-level displacement response for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacements range for the S1 pile embedded in a heavily OC representative offshore glacial clay till site with $d_{ice} = 100m$ and a NC representative offshore glacial clay till site.

5.2.6.4 Smaller diameter monopile geometries falling out of the geometrical calibration space

In the following section, GCCM V1 and GCCM V2 data-driven 1D design models are evaluated for their ability to compute the monotonic lateral behaviour of two of the PISA Cowden field test piles (Byrne et al., 2020a) and one of the PICASO Cowden field test piles (Byrne, 2021c). The monopiles have L/D values falling inside the calibration space but they have diameters significantly smaller than the offshore monopiles that were employed to train the data-driven 1D design models. Furthermore, the soil profile distributions of both the PISA and PICASO Cowden test sites are slightly different as a consequence of: (i) the water

table level in the PISA Cowden test site is not located at ground level ($z = 0m$), hence it has slightly different σ'_v, G_o, K_o and s_u distributions than the representative offshore glacial clay till site, (ii) the PICASO test site has s_u, G_o and K_o characteristics that are different from the PISA site, even though the sites are in close proximity. The examined field test pile geometries are presented in Table 5.10.

Table 5.10: PISA and PICASO field test monopile geometries according to Byrne et al. (2020a) and Byrne (2021c).

Project	**Pile**	$D(m)$	$L(m)$	$t_w(m)$	$h(m)$	L/D	h/D
PISA	CL1	2.0	10.3	0.025	10.10	5.15	5.05
PISA	CM9	0.762	3.98	0.011	9.98	5.22	13.10
PICASO	P01	1.22	3.66	0.016	4.86	3	3.98

Figure 5.15 shows the variation with depth of G_o/s_u and σ'_v/s_u for the PISA and PICASO Cowden test sites compared to the corresponding ratios of the representative offshore glacial clay till sites which form the GCCM. The PISA Cowden test site soil strength ratios are close to the representative offshore glacial clay till site distributions and within the trained calibration space of the GPR models. The minor discrepancies are mainly due to the non-hydrostatic pore water distribution at the PISA Cowden test site (Zdravkovic et al. 2020a). Furthermore, both G_o/s_u and σ'_v/s_u values of the PICASO test site fall within the trained calibration space.

Figure 5.16 illustrates the load-displacement response at ground level computed with the two versions of GCCM for CL1 and CM9 PISA field test monopiles. The GCCM results are compared to the PLAXIS 3D calculations, the measured PISA Cowden field test results (Byrne et al., 2020a) and the results calculated by the PISA Cowden till model. It can be observed that the GCCM V1 data-driven 1D design model accurately reproduces the load-displacement response for both field test piles; the calculations for the longer pile are slightly

more precise. The GCCM V2 data-driven 1D design model predictions are inconsistent; the 1D model calculates with relatively decent accuracy the PLAXIS 3D results for the CL1 pile, however it predicts badly for the CM9 pile. The high uncertainty of the GCCM V2 calculations regarding the field test pile geometries was evident in the large band of the 95% confidence interval produced by the GPR model predictions of the knot point parameters. Finally, as also mentioned in Beuckelaers (2017), the PISA Cowden till model calculations are not quite accurate for piles falling outside the pile geometry calibration space.

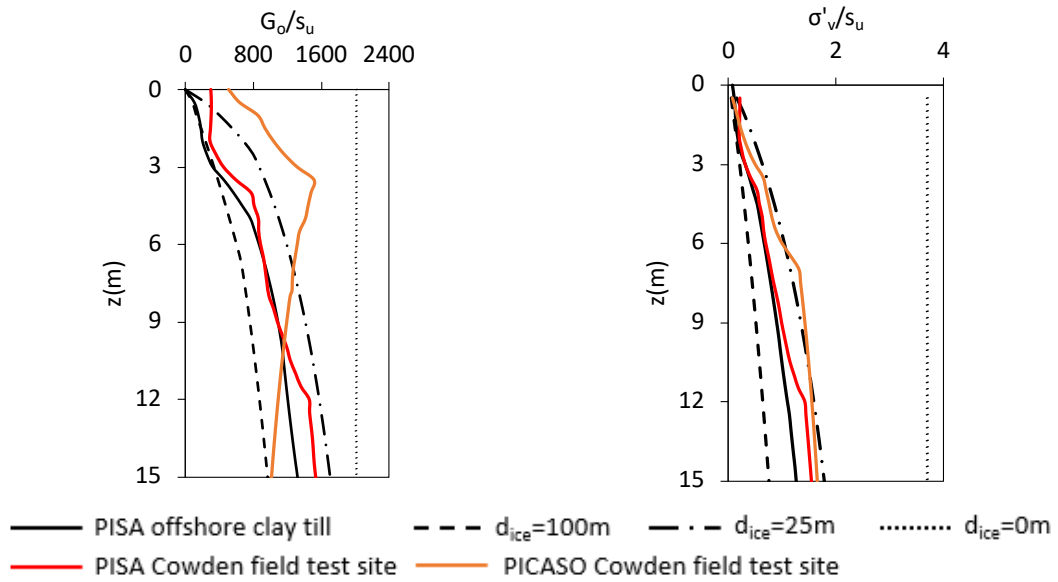


Figure 5.15: Variation with depth of G_o/s_u and σ'_v/s_u of the PISA Cowden site (Byrne et al., 2020a) and PICASO Cowden site (Byrne 2021c) with the representative offshore glacial clay till sites consisting the GCCM.

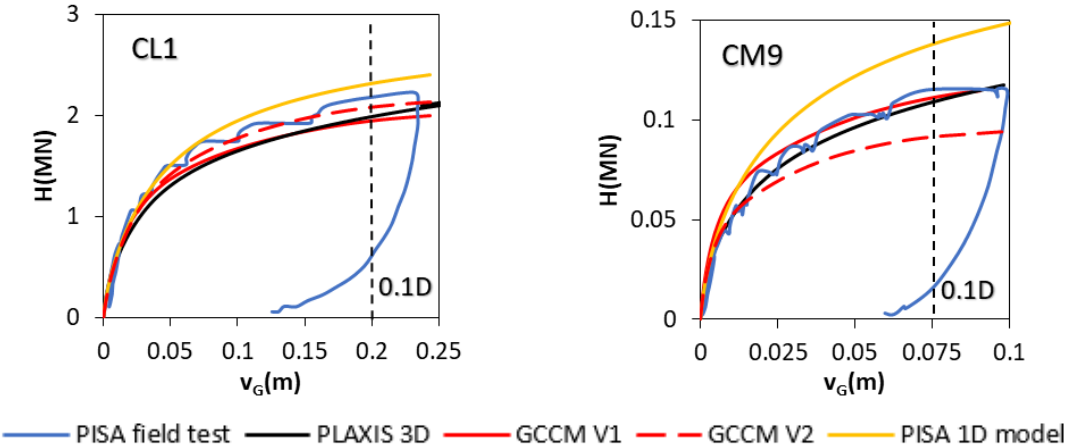


Figure 5.16: Load ground-level displacement response for the ultimate displacements range ($v_G = D/10$) for the CL1 and CM9 PISA field test monopiles.

Figure 5.17 presents the normalised load-displacement response of P01 PICASO field test pile at ground level computed from the GCCM V1 1D model, which has demonstrated greater accuracy for the PISA field test piles. The results from GCCM V1 are compared against the predictions from PLAXIS 3D and the measured P01 field test results (provided by Dr. Crispin). Similar to the findings observed for the PISA Cowden field test piles, it is evident that the GCCM V1 1D model accurately computes the load-displacement response.

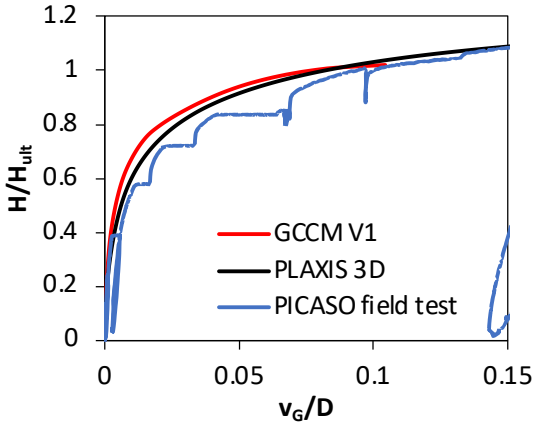


Figure 5.17: Pile normalised load-displacement response at the ground level for the ultimate displacements range ($v_G = D/10$) for the P01 PICASO field test monopile.

5.2.7 Assessment metrics

To quantify the performance of the data-driven general Cowden Clay models (GCCM V1, GCCM V2) for the design scenarios presented in Section 5.2.6, the η and ρ metrics adopted in Byrne et al. (2020b) (Equation (4.20)) were employed These metrics were also applied to assess the accuracy of the PISA Cowden till model. It's important to note that all 3D FEA conducted for the design cases employed the NGI-ADP model to simulate soil behaviour. The PISA Cowden till model was developed using numerical soil reaction curves extracted from 3D FEA, utilising the Modified Cam Clay model as detailed in Byrne et al. (2020b). However, as demonstrated in Chapter 3, the adopted NGI-ADP model parameters can effectively reproduce the pile response observed in the PISA 3D FEA conducted by Byrne et al. (2020b), validating this comparison.

Table 5.11 and Table *5.12* present the accuracy metric, η, and ratio metric, ρ, (calculated using Equation (4.20)) for the monopile's load-displacement response across small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacement ranges. These metrics were computed by the GCCM V1, GCCM V2 data-driven 1D design models, and the PISA Cowden till model for all design cases outlined in Section 5.2.6. It is important to note that the for the smaller diameter PISA (CM9, CL1) and PICASO (P01) field test piles, the values of η and ρ were calculated solely for the pile's ultimate response. As can be observed for the larger diameter offshore pile geometries the values of accuracy metric for both GCCM V1 and GCCM V2 data-driven 1D design models are in range $90.5\% \leq \eta_{sd} \leq 98.5\%$ for small displacements and in range of $92.5\% \leq \eta_{ult} \leq 99.5\%$ for the ultimate response. These data indicate a close match between the GCCM V1 and GCCM V2 calculations and the corresponding 3D FEA predictions. However, for the smaller diameter onshore pile geometries, only the GCCM V1 data-driven 1D design model provides $\eta_{ult} > 90\%$ indicating a good accuracy with FEA results.

Table 5.11: Values of accuracy metric η for the small displacement and ultimate response.

Glacial clay till sites	**Pile**	**GCCM V1 with 3D predictions**		**GCCM V2 with 3D predictions**		**PISA Cowden till model with 3D FEA**	
		η_{sd}	η_{ult}	η_{sd}	η_{ult}	η_{sd}	η_{ult}
OC – d_{ice}=100m	D1	0.973	0.964	0.996	0.990	0.916	0.946
OC – d_{ice}=25m	D1	0.964	0.985	0.989	0.985	0.902	0.981
NC – d_{ice}=0m	D1	0.950	0.988	0.966	0.975	0.662	0.755
PISA offshore till	D1	0.984	0.969	0.932	0.937	0.949	0.978
OC – d_{ice}=125m	D1	0.982	0.988	0.992	0.971	0.919	0.927
OC – d_{ice}=75m	D1	0.964	0.976	0.979	0.991	0.923	0.939
OC – d_{ice}=100m	S1	0.905	0.942	0.950	0.941	-	-
NC – d_{ice}=0m	S1	0.954	0.976	0.931	0.984	-	-
PISA Cowden test site	CL1	-	0.976	-	0.929	-	0.833
PISA Cowden test site	CM9	-	0.925	-	0.877	-	0.754
PICASO Cowden test site	P01	-	0.966	-	-	-	-
Average		0.960	0.969	0.969	0.958	0.879	0.889

Table 5.12: Values of ratio metric ρ for the small displacement and ultimate response.

Glacial clay till sites	**Pile**	**GCCM V1 with 3D predictions**		**GCCM V2 with 3D predictions**		**PISA Cowden till model with 3D predictions**	
		ρ_{sd}	ρ_{ult}	ρ_{sd}	ρ_{ult}	ρ_{sd}	ρ_{ult}
OC – d_{ice}=100m	D1	0.979	0.953	0.994	0.966	0.916	1.056
OC – d_{ice}=25m	D1	1.006	0.985	0.987	1.013	1.047	0.974
NC – d_{ice}=0m	D1	1.026	1.007	1.030	1.019	1.265	0.790
PISA offshore till	D1	1.019	0.947	0.940	0.914	0.952	1.028
OC – d_{ice}=125m	D1	0.983	0.980	0.986	0.958	0.92	1.081
OC – d_{ice}=75m	D1	0.972	0.990	0.980	1.002	0.918	1.065
OC – d_{ice}=100m	S1	0.911	0.947	0.966	0.946	-	-
NC – d_{ice}=0m	S1	0.976	0.990	0.918	1.01	-	-
PISA Cowden test site	CL1	-	0.981	-	1.048	-	1.165
PISA Cowden test site	CM9	-	1.019	-	0.836	-	1.264
PICASO Cowden test site	P01	-	0.987	-	-	-	-
Average		0.984	0.981	0.975	0.971	1.003	1.053
CoV(%)		3.65	2.27	3.29	6.07	12.54	12.29

Table 5.13 presents η and ρ for the moment-rotation response determined by the GCCM V1 data-driven 1D design model (which performed better for the field test piles). These metrics are provided for all the pile design cases discussed in Section 5.2.6. Similar to Table 5.11 and Table *5.12*, the values of η and ρ corresponding to the PISA and PICASO field test piles were calculated solely for the pile's ultimate response. As can be observed the values of accuracy metric are in range of $92.3\% \leq \eta_{sd} \leq 98.9\%$ for small rotation response and in range of $92.8\% \leq \eta_{ult} \leq 99.5\%$ for ultimate response, indicating a close match between the GCCM predictions and the corresponding 3D FEA monopile's moment-rotation response. The data-driven 1D design models can predict the pile moment rotation response at ground level at a similar level of accuracy to the computed pile load-displacement response for both the pile small displacement and ultimate response.

Table 5.13: Values of accuracy metric η and ratio metric ρ for small rotation and ultimate response.

Glacial clay till sites	**Pile**	**GCCM V1 with 3D predictions**			
		η_{sd}	η_{ult}	ρ_{sd}	ρ_{ult}
PISA offshore till	D1	0.983	0.966	1.011	0.947
OC – d_{ice}=100m	D1	0.960	0.976	0.965	0.971
OC – d_{ice}=25m	D1	0.979	0.989	0.978	0.981
NC – d_{ice}=0m	D1	0.957	0.984	0.971	1.008
OC – d_{ice}=125m	D1	0.970	0.976	0.975	0.976
OC – d_{ice}=75m	D1	0.972	0.978	0.968	0.979
OC – d_{ice}=100m	S1	0.960	0.994	0.962	0.995
NC – d_{ice}=0m	S1	0.989	0.995	1.012	1.008
PISA Cowden test site	CL1	-	0.974	-	0.979
PISA Cowden test site	CM9	-	0.928	-	1.011
PICASO Cowden test site	P01	-	0.964	-	0.982
Average		0.971	0.975	0.980	0.985
CoV(%)		-	-	1.904	1.770

5.3 Data-Driven Method for Sand with Arbitrary Relative Density

The new methodology is assessed for various sand sites with different relative density profiles through the development of a general Dunkirk sand mode (GDSM). The term 'Dunkirk' refers to the soil characteristics derived from the Dunkirk test site utilised in the PISA project (Zdravkovic et al., 2020a; Taborda et al., 2020). The data-driven GDSM developed in this study is based on the methodology outlined by Suryasentana et al. (2021), who used soil reaction curves extracted from the PISA FEA. These soil reaction curves correspond to representative offshore homogenous sand sites with relative density values ranging from $45\% \leq D_R \leq 90\%$ presented in Burd et al. (2020a), AWG (2018) enabling the development of a data-driven GDSM design model. The data-driven GDSM model is an alternative to the GDSM PISA design model described in Burd et al. (2020a). Suryasentana

et al. (2021) showed that the data-driven GDSM model slightly outperforms the GDSM developed by the PISA design model for unseen design cases.

In this chapter, a new data-driven general Dunkirk sand model named 'GDSM V2' is introduced. GDSM V2 is constructed by utilising soil reaction curves extracted from PLAXIS 3D FEA employing the HS-Small constitutive model. The calibration procedure for the HS-Small parameters is presented in Chapter 3. The effective friction angle ϕ' employed as an input parameter for HS-Small was determined for each sand profile used in this chapter by linearly interpolating between the values of (D_R, ϕ') presented in Figure 3.16(a) in Section 3.4.3.2.

5.3.1 Soil distribution and pile geometry calibration space

The GDSM V2 adopts a similar pile geometry calibration space to the GCCM data-driven 1D design model presented in Section 5.2.2. The pile geometries employed for each sand profile are presented in Table 5.1, employing nine of the eleven calibration pile geometries introduced in Burd et al. (2020a) which were used to form the calibration space of the GDSM developed by the PISA design model.

The four PISA representative offshore homogenous sand sites with a relative density of $D_R = 45, 60, 75, 90\%$ presented in Burd et al. (2020a) were employed for the development of the GDSM V2. The soil properties are based on the PISA Dunkirk test site; the submerged unit weight is $\gamma' = 10.09\, kN/m^3$ and $K_o = 0.4$. The G_0 profile for each site was calculated using Equation (3.6); the void ratio e for each sand profile is determined by:

$$e = e_{max} - D_R \times (e_{max} - e_{min}) \tag{5.5}$$

where $e_{max} = 0.91$ and $e_{min} = 0.54$ (Zdravkovic et al. 2020a). The G_0 profiles of the four PISA representative offshore homogenous sand sites along with their relative densities are presented in Figure 5.18.

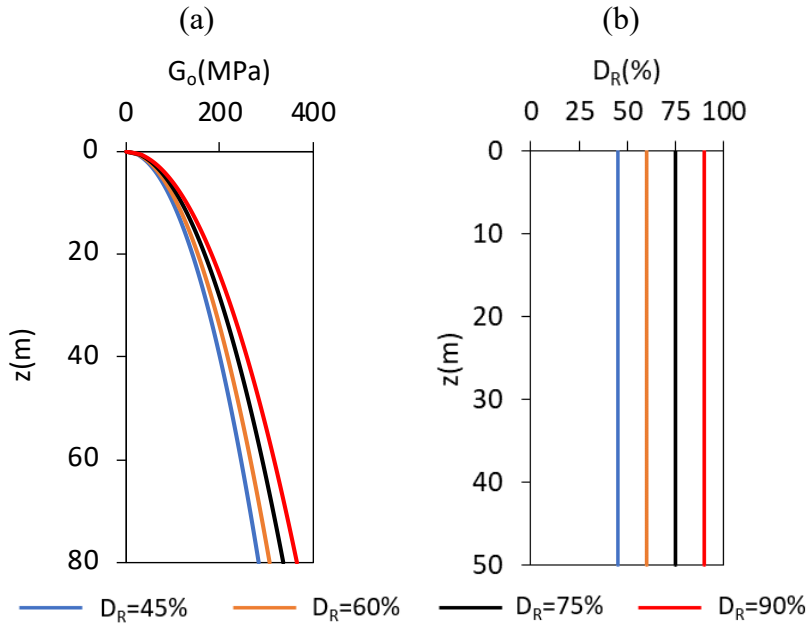


Figure 5.18: Profiles of (a) G_o, (b) D_R, employed for PISA representative offshore homogenous sand sites used to calibrate the GDSM V2 data-driven 1D design model.

5.3.2 Generation of the GDSM V2

Soil reaction curves were extracted from each calibration FEA adopting the procedure described in Section 4.2. Subsequently, the numerical soil reaction curves were processed and normalised following the procedures outlined in Sections 4.3, 4.4. Finally, the extracted soil reaction curves have been validated with the PLAXIS 3D predictions for each monopile geometry. All the pile geometries, soil profiles and knot point soil reaction parameters data from the 36 FEA were added to a SQL database by using the DB Browser for SQLite (DB4S) open source tool (Hipp, 2020).

The GDSM V2 data-driven 1D design model was trained using as training inputs data for $[L/D]$, $[z/L]$, $[G_o/\sigma'_v]$ and $[D_R]$. The Gaussian process regression (GPR) machine learning model was employed. The GPR model used the Automatic Relevance Determination (ARD) Matérn ($\nu = 5/2$) kernel function plus the white noise kernel. An 80:20 training-test set split ratio was adopted. The MSE and R^2 train and test scores were used to evaluate the performance of each GPR trained model.

Table 5.14: Optimised kernel parameters for the GPR models as applied to the knot point parameters associated with the GDSM V2 GPR models. (σ_f is the signal variance of the Matérn kernel, σ_n is the variance of the noise kernel, l is the lengthscale for every input dimension).

Knot point	$\bar{p}$	$\bar{m}$
1	σ_f=0.497, σ_n^2=3.35-3, $l_{z/L}$=0.586, $l_{L/D}$=3.14, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=2.78	σ_f=0.0458, σ_n^2=7.66-6, $l_{z/L}$=1.73, $l_{L/D}$=4.89, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=6.07
2	σ_f=1.02, σ_n^2=1.18e-2, $l_{z/L}$=0.203, $l_{L/D}$=2.44, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=2.14	σ_f=0.136, σ_n^2=8.49-5, $l_{z/L}$=2.95, $l_{L/D}$=4.2, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=4.56
3	σ_f=2.67, σ_n^2=0.0642, $l_{z/L}$=0.394, $l_{L/D}$=4.89, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=5.78	σ_f=0.309, σ_n^2=3.23-3, $l_{z/L}$=2.9, $l_{L/D}$=6.97, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=0.36
4	σ_f=4.45, σ_n^2=0.0731, $l_{z/L}$=0.29, $l_{L/D}$=6.78, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=1.88	σ_f=0.921, σ_n^2=4.21e-3, $l_{z/L}$=0.498, $l_{L/D}$=5.57, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=1.69
5	σ_f=5.52, σ_n^2=0.106, $l_{z/L}$=0.295, $l_{L/D}$=6.04, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=1.3	σ_f=1.38, σ_n^2=5.34e-3, $l_{z/L}$=0.511, $l_{L/D}$=3.15, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=0.954
6	σ_f=7.53, σ_n^2=0.425, $l_{z/L}$=0.288, $l_{L/D}$=4.48, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=0.704	σ_f=2.75, σ_n^2=6.37e-3, $l_{z/L}$=0.244, $l_{L/D}$=16.3, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=2.42
7	σ_f=9.67, σ_n^2=1.38, $l_{z/L}$=0.288, $l_{L/D}$=4.35, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=0.537	σ_f=3.28, σ_n^2=0.0748, $l_{z/L}$=0.327, $l_{L/D}$=3.12, $l_{G_o/\sigma_v'}$=867, l_{D_R}=0.817
8	σ_f=12.1, σ_n^2=1.83, $l_{z/L}$=0.313, $l_{L/D}$=4.31, $l_{G_o/\sigma_v'}$=1000, l_{D_R}=0.576	σ_f=0.32, σ_n^2=0.0731, $l_{z/L}$=0.32, $l_{L/D}$=3.28, $l_{G_o/\sigma_v'}$=772, l_{D_R}=0.707

Table 5.14 lists the optimised hyperparameters for the knot point soil reaction parameters of the distributed soil reactions $\bar{p}$, $\bar{m}$ after the GPR models training. The optimised lengthscale l_{DR} retains a low value indicating that the GPR model identifies the [D_R] input as important to discriminate the soil reaction parameters for sands with different relative density.

5.3.3 Applications of the GDSM V2

An initial exercise has been conducted in which predictions from the GDSM V2 are compared with the original training data. This comparison indicates that the GDSM V2 provides a good match to the calibration analyses. This section evaluates the performance of the GDSM V2 in respect to predicting the pile's monotonic lateral behaviour for:

1. design offshore monopiles embedded in trained sand profiles.
2. design offshore monopiles embedded in sand profiles with unseen relative density.

The computed load-displacement pile response at ground level from the GDSM V2 data-driven 1D design model has been validated with the predictions of PLAXIS 3D FEA. The pile moment-rotation response at ground level was also examined and the results indicated that it shows a similar level of accuracy to the computed pile load-displacement response for both the pile's small ($\psi_G = 0.002°$) and large ($\psi_G = 2°$) rotation response. Therefore, moment-rotation results are not presented in this section.

5.3.3.1 Design piles embedded in sand profiles with trained relative density

Initially the GDSM V2 is evaluated with respect to unseen offshore design monopile geometries falling within the L/D calibration space embedded in a trained sand profile. The D1 design monopile geometry with $L/D = 3$ defined in Table 5.8 was employed. Figure 5.19 presents the load-displacement response at ground level for pile D1 for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacement range developed by the GDSM V2 data-driven 1D design model. As can be observed, the 1D model matches the 3D predictions for all the design cases.

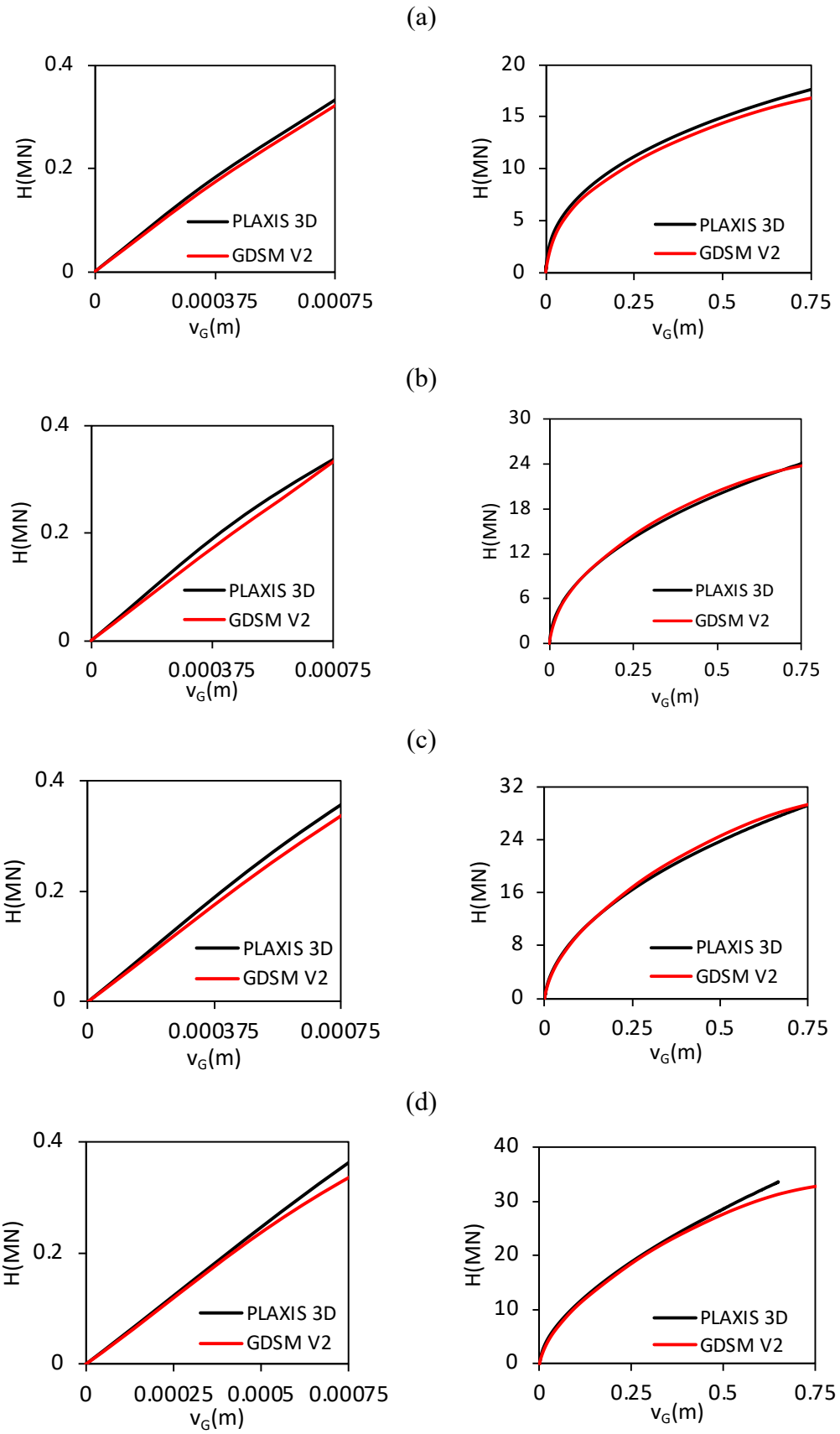


Figure 5.19: Load ground-level displacement response for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacement ranges for the D1 pile embedded in the PISA representative offshore homogenous sand site with a relative density of: (a) $D_R = 45\%$, (b) $D_R = 60\%$, (c) $D_R = 75\%$, (d) $D_R = 90\%$.

5.3.3.2 Design piles embedded in sand profiles with unseen relative density

Three new representative offshore homogenous sand sites were generated for a relative density of $D_R = 50, 80$ and 100%. Figure 5.20 illustrates the depth variation of G_o/σ'_v and D_R of the three unseen sand profiles compared to the corresponding ratios of the sand profiles that form the GDSM V2. The sand profile with $D_R = 100\%$ falls outside the calibration space of the 1D model.

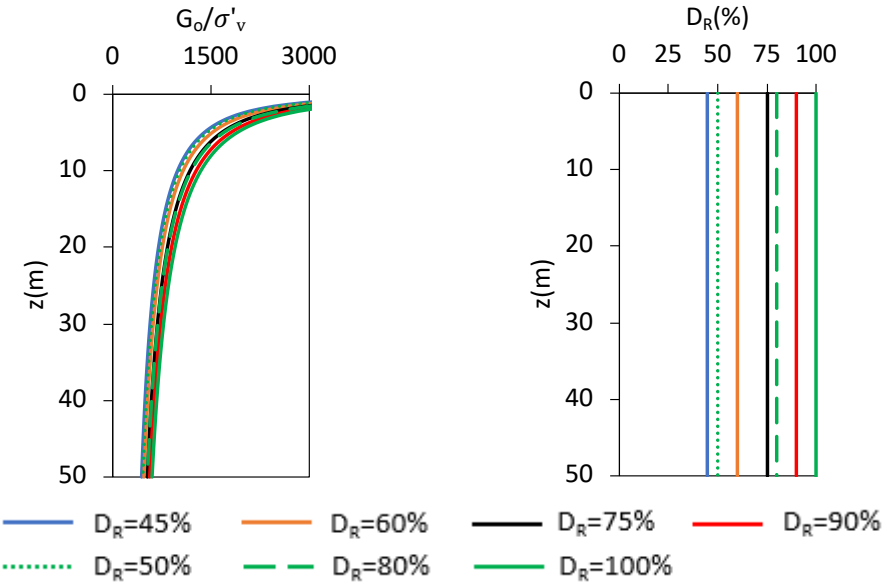


Figure 5.20: Variation with depth of D_R and G_o/σ'_v of the three unseen representative offshore homogenous sand sites with the homogenous sand sites consisting the GDSM V2.

Figure 5.21 and Figure 5.22 present the pile's load-displacement response at ground level for the small ($v_G = D/10000$) and ultimate ($v_G = D/10$) displacement range computed by the GDSM V2 data-driven 1D design model for the D1 (Table 5.8) monopile embedded in the three unseen representative offshore homogenous sand sites. The results from the 1D model are compared to validated PLAXIS 3D calculations. The 1D model fits accurately the 3D calculations for all the cases. The effective friction angle ϕ' employed as an input parameter for HS-Small of the sand profile with $D_R = 100\%$ was assumed to be $\phi' = 46.5°$. This value is similar to that employed for the hydraulic fill present in the PISA Dunkirk sand site field test results in Section 3.4.4, which was also characterised by $D_R = 100\%$.

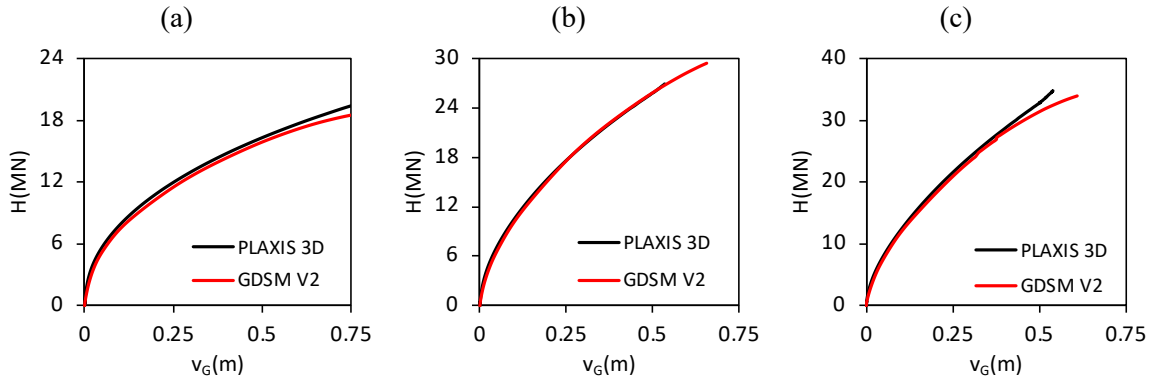


Figure 5.21: Load ground-level displacement response for the ultimate displacement range ($v_G = D/10$) for D1 pile embedded in the three unseen representative offshore Dunkirk sand site with: (a) $D_R = 50\%$, (b) $D_R = 80\%$, (c) $D_R = 100\%$.

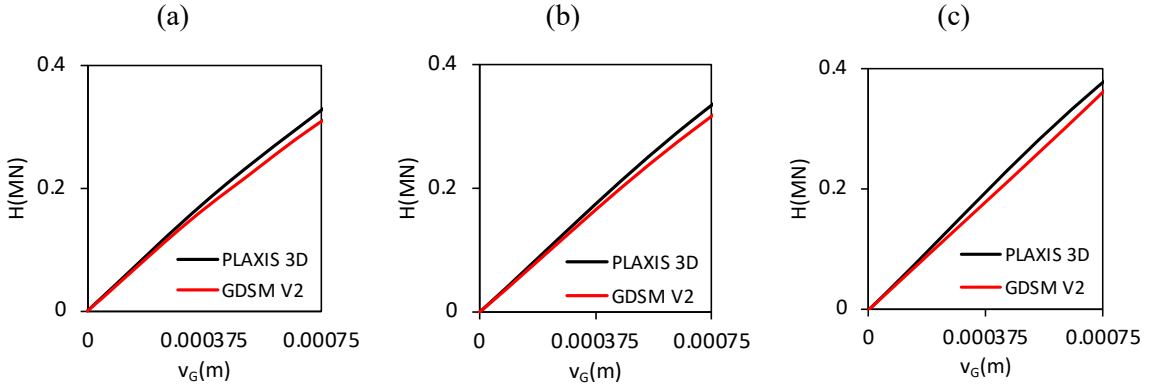


Figure 5.22: Load ground-level displacement response for the small displacement range ($v_G = D/10000$) for D1 pile embedded in in the PISA representative offshore Dunkirk sand site with: (a) $D_R = 50\%$, (b) $D_R = 80\%$, (c) $D_R = 100\%$.

5.4 Summary

This chapter explores the effectiveness of a data-driven method suitable for generating generalised clay or sand models. Key observations are:

- The data-driven method is suitable for generating general clay 1D models for offshore clay sites with different s_u and G_o distributions. This facility is not available for clay soils by using the PISA design model (Byrne et al., 2020b).
- Similarly, the data-driven method is effective for generating general sand 1D models for offshore sand sites with different D_R distributions. The results align well with the findings in Suryasentana et al. (2021).

- The data-driven 1D design models generally outperform the corresponding PISA design models. The eight-knot parametrisation of the soil reaction curves provides higher fidelity than the 4-parameter conic function, and the GPR models can rigorously capture the depth variation of the soil reaction parameters.
- The data-driven 1D design models accurately predict both the pile's load-displacement and moment-rotation response at ground level across small and large displacement and rotation ranges.
- The GPR models can be reliably used for extrapolation to some extent beyond the available data. The data-driven 1D design models can provide reliable predictions for pile geometries and unseen soil profiles falling outside the calibration space. This feature is not available with the PISA design models.
- Both the clay and sand normalisation frameworks in Table 4.3 can be utilised for developing a general clay model. However, it was found that the general clay model employing the sand normalisation does not always provide reliable results for piles with diameters falling significantly away the calibration space. Therefore, it is advised to favour the clay normalisation when generating general clay models.
- Both GCCM V1 and GCCM V2 generally outperform the PISA Cowden till model (Byrne et al., 2020b). Overall, while the PISA Cowden till model predicts the response of piles embedded in overconsolidated (OC) representative glacial clay till sites accurately at the ultimate response ($v_G = D/10$), it is less accurate at predicting the pile's small displacement response for OC glacial clay till sites with different G_o profiles compared to the PISA representative offshore glacial clay till site (i.e. the calibration soil profile). However, the model is not suitable for use with normally consolidated (NC) representative glacial clay till sites or for pile geometries falling outside the 1D model's calibration space.

Chapter 6

6 Data-Driven Method for Layered Soils

6.1 Introduction

Offshore wind farm sites typically consist of interbedded layers of soil with different geological origins. Most of the past (e.g. *p-y* method) and more advanced (e.g. PISA) 1D design models were developed for analysing the monotonic lateral behaviour of monopiles embedded in homogenous soil. Typically, the same 1D models are used for piles embedded in layered soils. For example, Burd et al. (2020b) extended the PISA design model to layered soils. They suggest that the 1D models that were developed using the PISA design model for homogenous soil profiles can be used directly for calculations involving to layered soils. This method is referred as the 'independent layer method' (ILM) in this thesis. However, it is widely accepted that there is interaction between the layers in layered soils which can have a significant effect on pile lateral response (Georgiadis, 1985; Yang and Jeremic, 2005; Burd et al., 2020b; Hu et al., 2022).

This chapter considers new approaches to develop accurate data-driven 1D design models for layered soil configurations. PLAXIS 3D is used to conduct the calibration calculations. The finite element meshes for the 3D FEA comprise approximately 80000 to 100000 elements. The HS-Small constitutive model is employed for sands and the NGI-ADP model is adopted for clays. In both cases, constitutive model parameters were determined following

the calibration procedures described in Chapter 3. The soil reaction curves extracted from the calibration analyses are normalised using the approach presented in Section 4.4 in which clay and sand profiles are treated in a consistent way employing the 'sand' data-driven normalisation presented in Table 4.3.

6.2 Specification of the layered soil study

This section describes the layered soil properties that were adopted for the study. For all the layered soils the initial pore pressures were assumed to be hydrostatic corresponding to an offshore environment. The layered soil properties were based on the reference soil properties of the representative offshore clay and sand profiles developed during the PISA project (Byrne et al., 2020b; Burd et al. 2020a; Burd et al. 2020b; Kamas et al. 2023).

Profiles of K_o values, small-strain shear modulus G_o and undrained shear strength s_u for the reference homogeneous soil models are illustrated in Figure 6.1. The sand layers properties are based on the representative offshore homogenous sand sites with submerged unit weight $\gamma' = 10.09\,\mathrm{kN/m^3}$ presented in Burd et al., (2020a). The clay soil properties are based on the PISA representative offshore glacial clay till site (Byrne et al., 2020) with $\gamma' = 11.38\,kN/m^3$. The stress history of each clay layer is determined based on the procedure presented in Section 5.2.1 where it was assumed that the overconsolidation effect was caused solely by the prior action of the weight of an ice sheet layer with depth d_{ice}.

The calibration layered soil distributions were assigned following Burd et al. (2020b); σ'_{vi} is determined from the submerged unit weight γ' from the reference soil profiles; σ'_{hi} and p' are calculated using the K_o of the reference soil profiles; G_o and s_u are considered to be functions of p' and are estimated at the local p' of each layered soil profile based on the G_o and s_u distributions of the reference soil profiles. The relationship between G_o and p' is determined by the reference homogenous soil models.

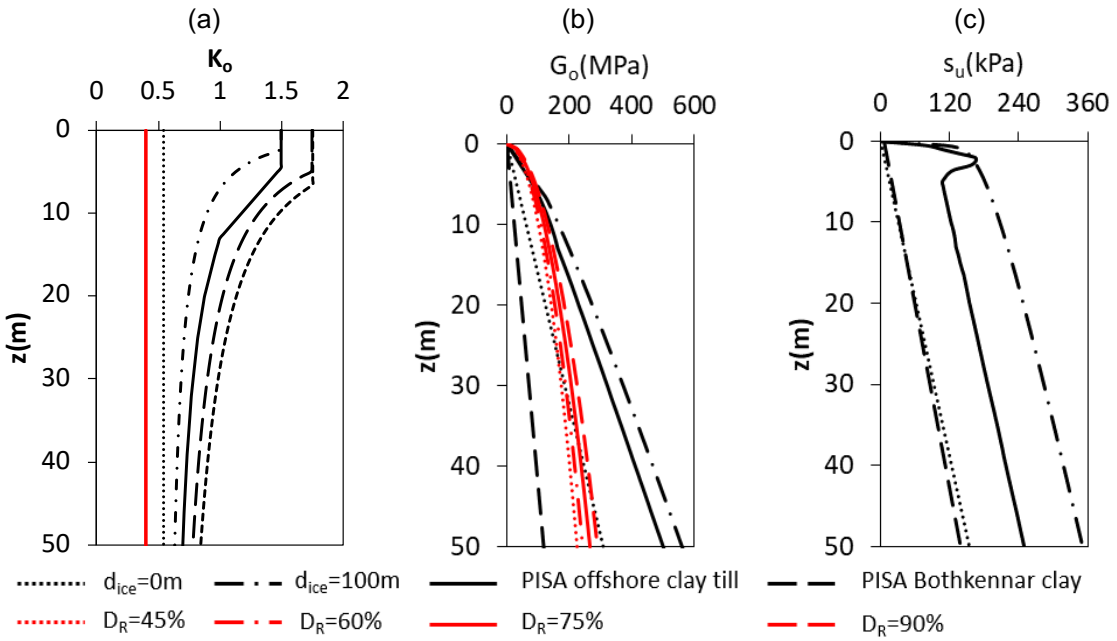


Figure 6.1: Profiles of (a) K_o values, (b) G_0, and (c) s_u, employed in the reference homogeneous soil models (where clay profiles are indicated in black and sand profiles in red).

6.3 Independent layer method

The independent layer method (ILM) presented by Burd et al., (2020b) was developed during the PISA project to extend the PISA design method to layered soils. In the ILM the soil reaction curves extracted from 1D models calibrated by FEA for homogenous soil profiles are used for predictions concerned with layered soils. This procedure, coupled with the PISA design model, was shown to provide reliable results for many layered configurations (Burd et al., 2020b). However, ILM lacked accuracy when neighbouring soil layers had significantly different strength characteristics, such as soft clay adjacent to dense sand.

The data-driven 1D design models as described in Chapter 4 can predict with high-fidelity the soil reaction curves for homogenous soil profiles. In the following section, the ILM is tested for different layered configurations introduced in the PISA project using soil reaction curves predicted by data-driven 1D design models developed for homogenous soil profiles. The analyses focus on pile C1 (Table 3.1) with $L/D = 2$, a geometry previously shown by Burd et al. (2020b) to yield the worst quality results. The layered configurations considered

are illustrated in Figure 6.2, representing idealised soil profiles of realistic layered configurations that were introduced in Burd et al. (2020b). These layered soil profiles models are generated using the procedure outlined in Section 6.2, employing the reference homogeneous soil models illustrated in Figure 6.1. The term 'matrix' presented in Figure 6.2 is used to describe the principal soil type and the term 'layer' characterises any embedded or base layers that are present. Case A describes a two-layered soil with a boundary located at the pile mid-depth. Case B employs a relatively thin embedded layer within a homogeneous matrix. Case C characterises a two-layered soil profile where the second layer is placed near the pile base. Case D comprises the combination of Cases B and C. Case E represents a multi-layer system consisting of many different soil types. Finally, it should be noted that the analyses for each case (e.g. Case A1, A2) use the same nomenclature as Burd et al. (2020b) but the analyses are not the same, because: (i) representative offshore glacial clay till sites with different values of d_{ice} are used instead of the PISA representative offshore Bothkennar clay and glacial clay till sites, (ii) a pile with $L/D = 2$ is employed instead of pile D2 (refer to Table 4.14) with $L/D = 4$ that was used in Burd et al. (2020b).

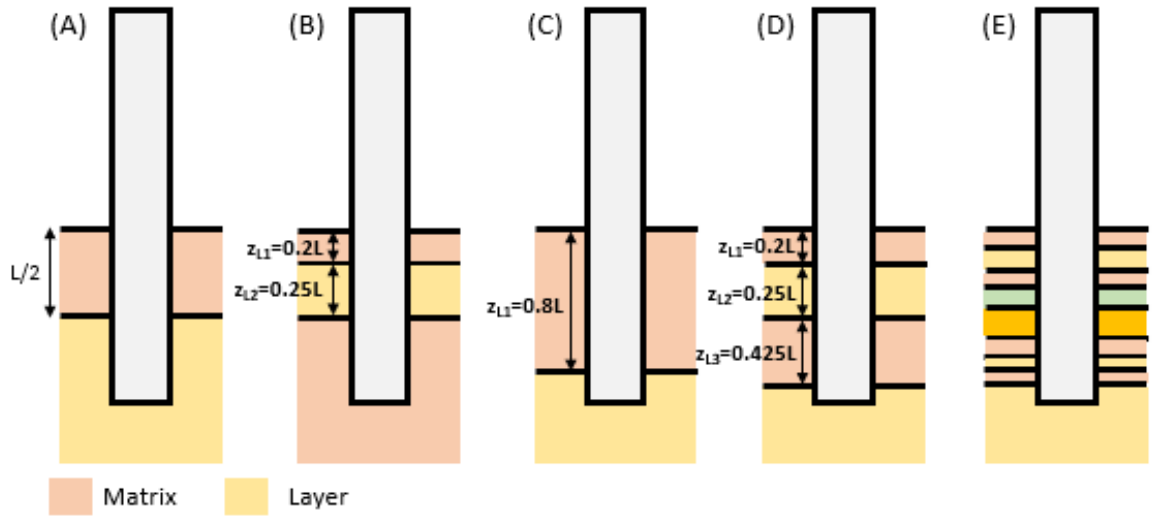


Figure 6.2: Layered soil configurations adapted form Burd et al. (2020b): (a) case A, two layers; (b) case B, single embedded layer; (c) case C, base layer; (d) case D, conceptual multi-layer systems (Case B + Case C); (e) case E, multi-layer systems.

6.3.1 Parametric studies

Case A: two-layer system

Five analyses employing the layered configuration A (Figure 6.2) were conducted as specified in Table 6.1. Case A1 and A2 involve similar type (clay or sand) of soil layers where a dense layer underlies a soft layer. In case A1 a stiff clay is underlying a layer of soft clay; in case A2 a dense sand underlies a layer of loose sand. Cases A3, A4, A5 introduce neighbouring soil layers of different type. Case A3 specifies a special case where the neighbouring layers differ significantly in their strength and stiffness characteristics.

Table 6.1: Case A analyses- two-layer systems.

Case	Upper layer	Lower layer
A1	NC glacial clay till	OC glacial clay till with d_{ice}=100m
A2	Dunkirk sand (D_R=45%)	Dunkirk sand (D_R=90%)
A3	Bothkennar clay	Dunkirk sand (D_R=90%)
A4	OC glacial clay till with d_{ice}=100m	Dunkirk sand (D_R=90%)
A5	Dunkirk sand (D_R=45%)	OC glacial clay till with d_{ice}=100m

Figure 6.3 and Figure 6.4 compare the load-displacement response at ground level predicted by the ILM 1D model and the PLAXIS 3D calculations for the pile's ultimate and small displacement response. The ILM computes accurately the pile's small displacement response for all examined cases. Furthermore, the ILM calculates in relatively decent accuracy the pile's ultimate response for the layered soil Cases A1, A2 and A5 where the neighbouring layers do not have essentially different strength characteristics (e.g. similar soil type or stiff clay neighbouring with loose sand). However, for layered soil Cases A3 and A4 where the pile neighbouring soil layers have considerably different strength and stiffness characteristics (e.g. soft or stiff clay with dense sand) the ILM severely underpredicted the pile capacity.

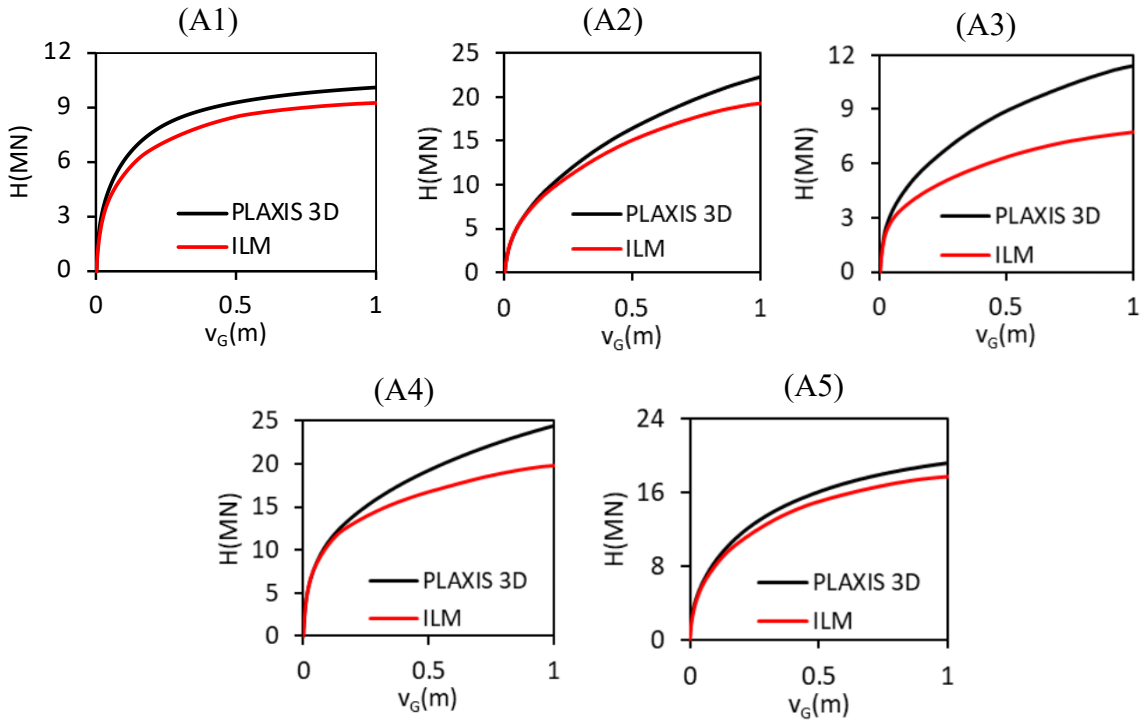


Figure 6.3: Comparison between the load-displacement responses at ground level for the ultimate response from the 1D model using the ILM and the 3D FEA for Case A analyses.

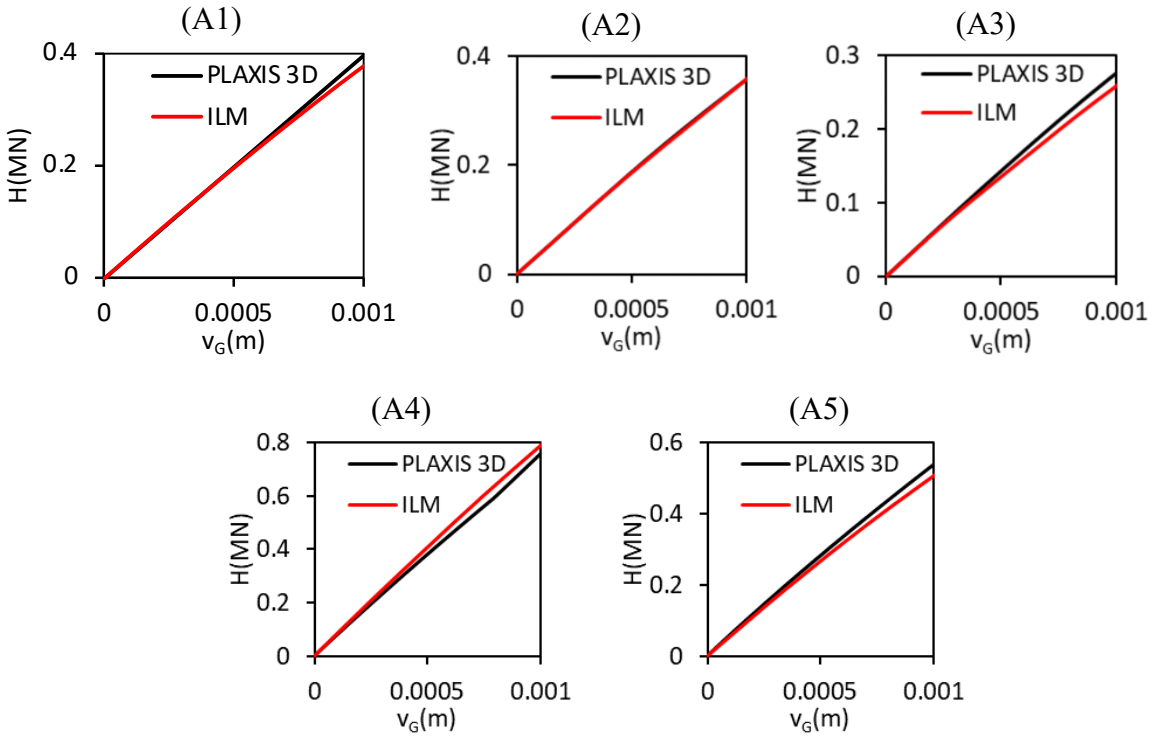


Figure 6.4: Comparison between the load-displacement responses at ground level for the small displacement response from the 1D model using the ILM and the 3D FEA for Case A analyses.

Case B: single embedded layer

Two different layered configurations were considered; introducing a clay layer embedded in a sand matrix and a sand layer embedded in a clay matrix (Figure 6.2(b)). A preliminary study showed that the pile's lateral behaviour is mostly affected when the layer is located in the upper half of the pile; if the layer is located near the rotation point of the pile it does not contribute significantly to the pile lateral response. Two different thicknesses were considered for the embedded layer, $t_{layer} = 5m$ and $t_{layer} = 2m$. The analyses examined for layered Configuration B are presented in Table 6.2.

Table 6.2: Case B analyses - single embedded layer system (z_{L1}, z_{L2} are the first- and second-layer depths presented in Figure 6.2).

Case	Matrix	Embedded layer	z_{L1}/L	z_{L2}/L
B1	OC glacial clay till with d_{ice}=100m	Dunkirk sand (D_R=90%)	0.2	0.25
B2	Bothkennar clay	Dunkirk sand (D_R=90%)	0.2	0.25
B3	Dunkirk sand (D_R=45%)	OC glacial till clay with d_{ice}=100m	0.2	0.25
B4	Dunkirk sand (D_R=90%)	Bothkennar clay	0.2	0.25
B5	Bothkennar clay	Dunkirk sand (D_R=90%)	0.2	0.1
B6	Dunkirk sand (D_R=90%)	OC glacial clay till with d_{ice}=100m	0.2	0.1

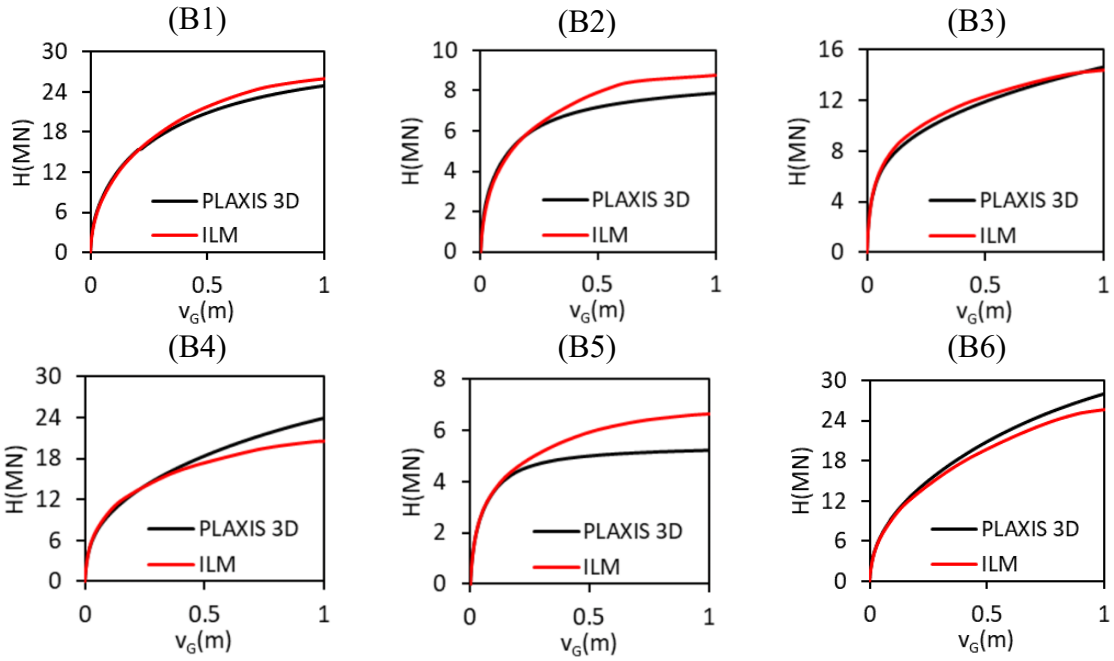


Figure 6.5: Comparison between the load-displacement responses at ground level for the ultimate response from the 1D model using the ILM and the 3D FEA for Case B analyses.

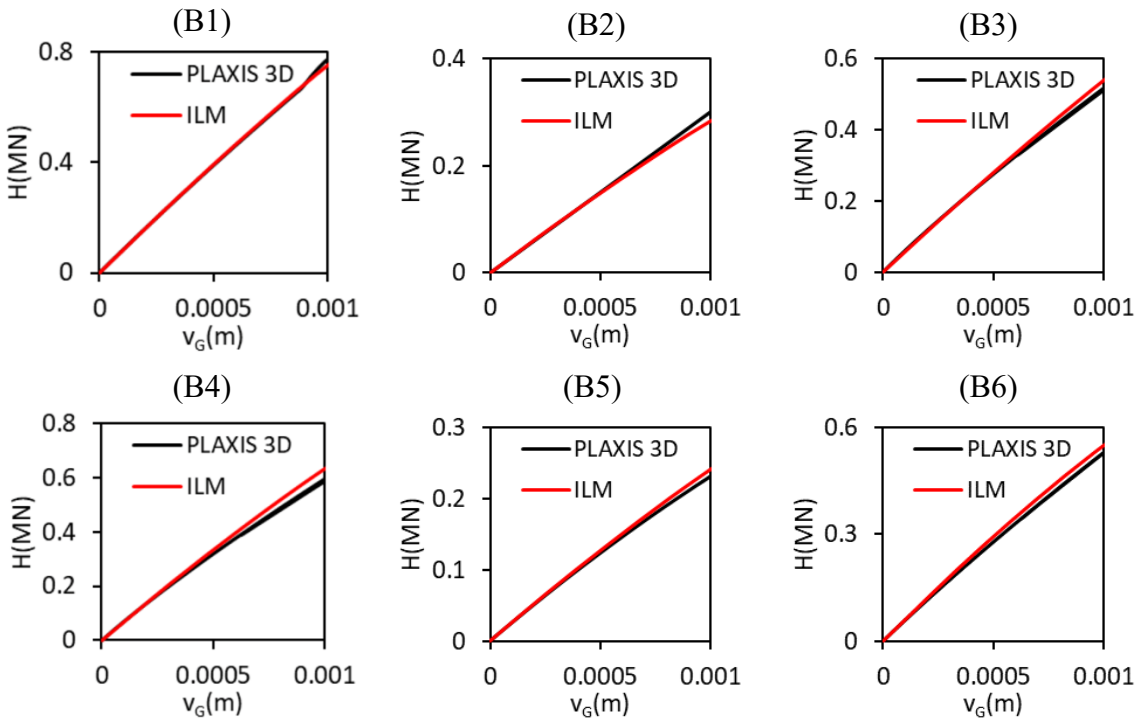


Figure 6.6: Comparison between the load-displacement responses at ground level for the small displacement response from the 1D model using the ILM and the 3D FEA for Case B analyses.

Figure 6.5 and Figure 6.6 show the comparison between the load-displacement responses at ground level predicted by the ILM 1D model and the 3D FEA. The ILM computes accurately the pile's small displacement response. Furthermore, the ILM calculates with a good accuracy the pile's lateral behaviour for layered soil cases B1, B3 and B6 where a very stiff clay is adjacent to a loose or dense sand. The ILM does not compute accurately the pile's ultimate response for cases B2, B4 and B5 where the neighbouring soil layers have significantly different strength characteristics (soft clay with dense sand). Overall, the results indicate that the accuracy of the ILM drops as the thin embedded layer thickness decreases.

Case C: base layer

The layered configuration C represents a two-layer system where the base layer is located near the tip of the pile. The top layer has depth of $z_{L1} = 0.8L$ (Figure 6.2(c)). Burd et al. (2020b) showed that if a deep layer is located below the tip of the pile it does not significantly

affect the pile lateral behaviour; therefore, this case has not been examined in the current case study. The layered soils that have been examined are specified in Table 6.3.

Table 6.3: Case C analyses - base layer.

Case	Upper layer	Lower layer
C1	NC glacial clay till	OC glacial clay till with d_{ice}=100m
C2	Dunkirk sand (D_R=45%)	Dunkirk sand (D_R=90%)
C3	Bothkennar clay	Dunkirk sand (D_R=90%)

Figure 6.7 and Figure 6.8 illustrate the comparison between the load-displacement responses at ground level computed by the ILM 1D model with the PLAXIS 3D analyses for case C. The ILM 1D model computes with good accuracy the pile behaviour (both pile's small displacement and ultimate response) for all the examined cases.

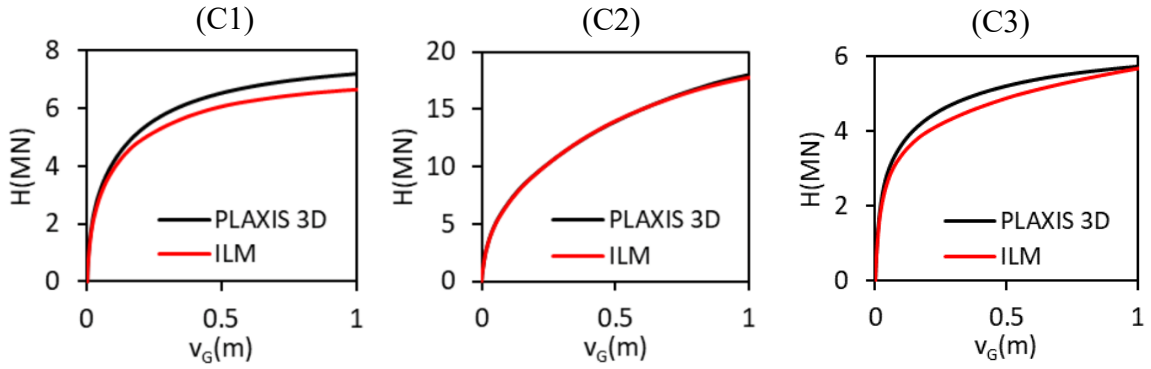


Figure 6.7: Comparison between the load-displacement responses at ground level for the ultimate response from the 1D model using the ILM and the 3D FEA for Case C analyses.

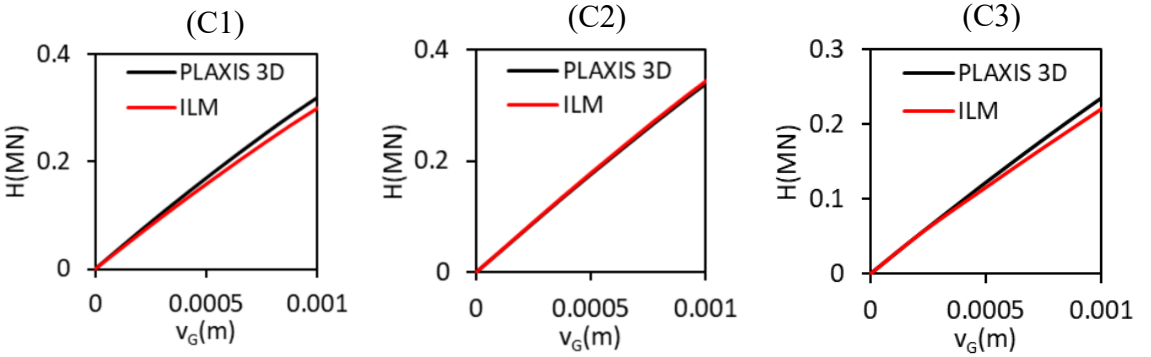


Figure 6.8: Comparison between the load-displacement responses at ground level for the small displacement response from the 1D model using the ILM and the 3D FEA for Case C analyses.

Case D: conceptual multi-layer systems

Case D explores the combined effect of Case B with Case C comprising a relatively thin embedded layer with a base layer of the same soil type embedded in a soil matrix of different soil type. Three layered soils were considered presented in Table 6.4; $z_{L1} = 0.2L$, $z_{L2} = 0.25L$ and $z_{L3} = 0.425$ are the first-, second- and third-layer depths presented in Figure 6.2(d).

Table 6.4: Case D analyses - conceptual multi-layer systems.

Case	Soil Matrix	Embedded + base layers
D1	OC glacial clay till with d_{ice}=100m	Dunkirk sand (D_R=90%)
D2	Dunkirk sand (D_R=75%)	PISA offshore glacial clay till
D3	Bothkennar clay	Dunkirk sand (D_R=90%)

Figure 6.9 and Figure 6.10 present the comparison between the load-displacement responses at ground level calculated by the ILM 1D model with the PLAXIS 3D results for layered configuration D. The ILM 1D models compute accurately the pile's small displacement response. However, they provide inaccurate calculations in terms of the ultimate response. Hence, the results indicated that for layered configurations similar to Case D, the ILM should not be employed. New 1D models need to be introduced that are capable of incorporating the effect of the interaction between the neighbouring layers on the soil reaction curves.

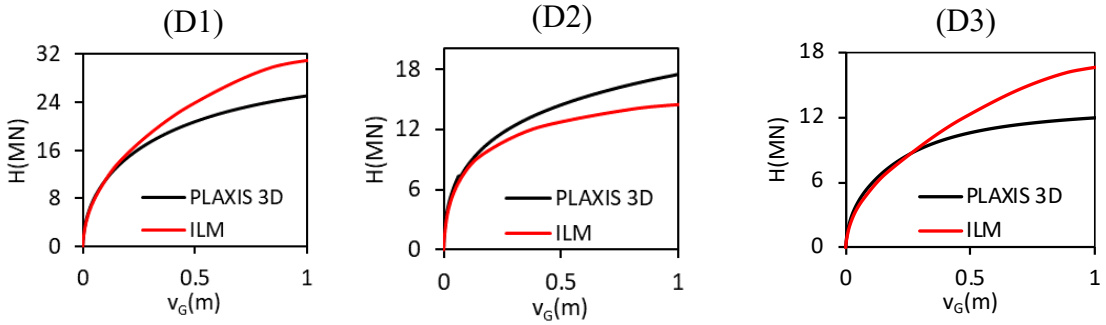


Figure 6.9: Comparison between the load-displacement responses at ground level for the ultimate response from the 1D model using the ILM and the 3D FEA for Case D analyses.

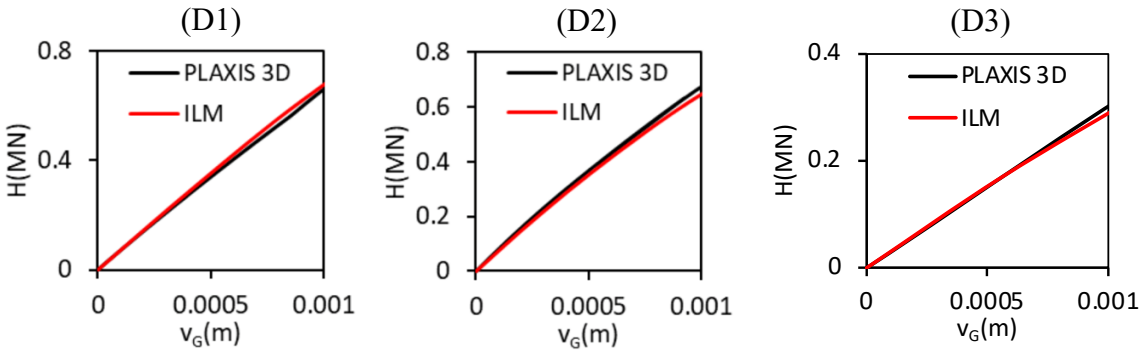


Figure 6.10: Comparison between the load-displacement responses at ground level for the small displacement response from the 1D model using the ILM and the 3D FEA for Case D analyses.

Case E – multi-layered systems

The configuration E examined the lateral behaviour of a monopile embedded in a layered configuration comprising three or more layers with different strength/stiffness characteristics. The configurations to be considered are illustrated in Figure 6.11. Case E1 is a three-layered sand soil profile with varying relative densities; Case E2 shows a three-layered soil profile where a stiff clay is located between two sand layers with different relative density; Case E3 is a multi-layered system introduced in the PISA project (Burd et al., 2020b) introducing many different clay and sand layers.

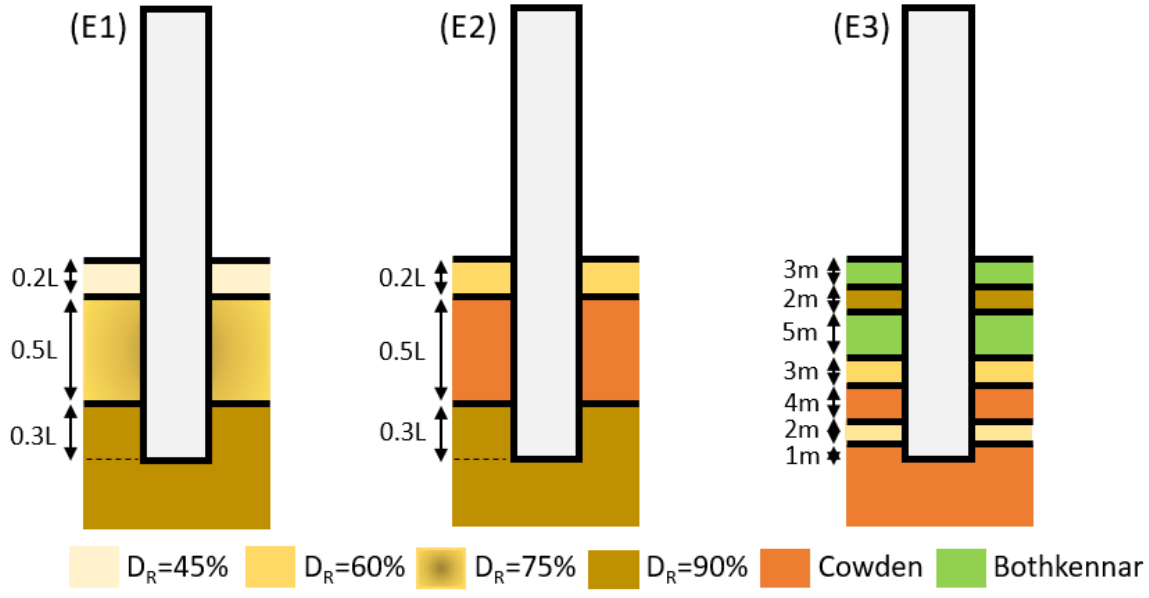


Figure 6.11:Multi-layered configurations case studies examined to validate ILM.

Figure 6.12 and Figure 6.13 illustrate the comparison between the pile response computed by the ILM 1D model and the PLAXIS 3D analyses for configuration E. The ILM 1D model computes quite accurately the pile's lateral behaviour for Cases E1, E2 and E3. The results indicated that thin soil layers (less than 5m), have a small or negligible effect on the developed soil reaction curves of their neighbouring layers. Also, ILM method is reliable for sandy layered soil profiles with neighbouring layers of different relative density.

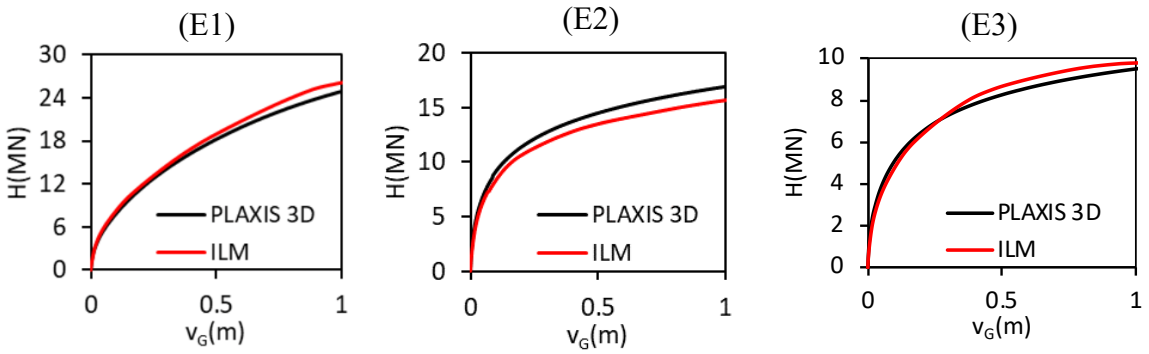


Figure 6.12: Comparison between the load-displacement responses at ground level for the ultimate response from the 1D model using the ILM and the 3D FEA for Case E analyses.

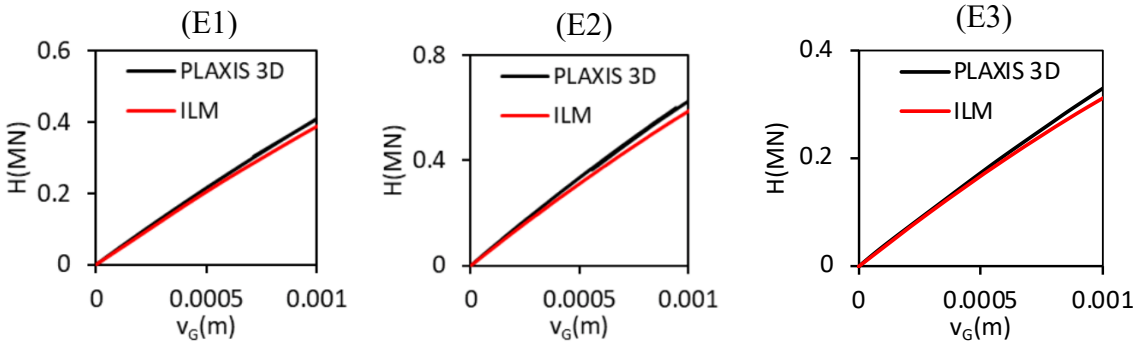


Figure 6.13: Comparison between the load-displacement responses at ground level for the small displacement response from the 1D model using the ILM and the 3D FEA for Case E analyses.

6.3.2 Key observations

This section explores the performance of the ILM proposed in Burd et al. (2020b) for layered soils coupled with predictions computed by data-driven 1D design models developed for homogenous soil profiles. Key observations are:

- The ILM calculates with good accuracy the pile's small displacement response agreeing with the results presented in Burd et al (2020b).

- The ILM calculates with relatively good accuracy the pile's ultimate response for layered soil profiles with layers of the same soil type (e.g. Cases A1, A2, C1, C2, E1) regardless the number of layers that are present.
- The ILM predictions are not consistently reliable for layered soils with neighbouring soil layers of different type (e.g. Case A4, A5, B3, B6, D1, D2, E2).
- The ILM is inaccurate for layered soils with neighbouring soil layers of considerably different strength and stiffness characteristics (e.g. Case A3, B2, B5, D3).
- The ILM provides good predictions for multi-layered soils with multiple thin layers (e.g. Case E3).

Overall, the results indicate that there is a need to incorporate the interactions between neighbouring layers in the developed 1D models to ensure that the predictions will be always consistent and reliable.

6.4 Layering effects on the soil reaction curves

This section explores the effect of layering on the extracted soil reaction curves; the aim is to understand why the ILM is inaccurate for certain layered configurations and to highlight the importance of incorporating interactions between neighbouring layers in data-driven 1D design models. Therefore, soil reaction curves extracted by FEA implemented for piles embedded in the reference homogenous soil profiles as presented in Section 6.2, are compared with soil reactions curves extracted directly by FEA for piles embedded in layered soils. For brevity, three layered soil configurations will be discussed (Cases A3, B2, B5 as presented in Table 6.1 and Table 6.2). In these cases: (i) the ILM fails to accurately predict the pile's ultimate response and (ii) the layered configurations are relatively simple, consisting of two soil layers, making it easier to observe the layering effects on the soil reaction curves.

For these study cases, the distributed lateral load p soil reaction curves (which is the most significant soil reaction component) are plotted. However, layering effects occur for all soil reaction components. The main interest of this study is concerned with soil reactions located at a depth near the boundary between two adjacent layers (where the interactions between neighbouring layers are more significant) to investigate the effect of interactions between neighbouring layers on the soil reaction curves.

Two-layered system - Case A3

Figure 6.14(a) and Figure 6.14(b) compare soil reaction curves extracted from reference homogenous analyses (Bothkennar clay and Dunkirk sand with $D_R = 90\%$) with those from a layered analysis for the Case A3. Figure 6.14(a) indicates that, for a certain transition depth above the layer boundary, the soil reaction curves extracted by the layered soil profile are greatly affected by the neighbouring dense sand profile and develop much higher values than the homogenous Bothkennar clay soil reactions; the plotted curves illustrate soil reactions at the depths of $z/L = 0.275, 0.375, 0.475$. It can be seen that the distributed lateral reaction curves are affected up to a maximum distance of approximately $0.5D$ from the layer boundary. As expected, the layering effect is larger at a depth near the layer boundary and decays as it moves away from the boundary. Figure 6.14(b) demonstrates that just below the layer boundary, for a certain depth above the pile's rotation point, the p soil reaction curves extracted from the layered soil FEA develop slightly higher values than those from the homogenous reference sand FEA. This phenomenon can be explained by the pile's deflection shape (Figure 6.14(c)). In the layered FEA, the pile's rotation point is located at a lower depth due to the presence of soft clay as the top layer, influencing all the soil reaction curves below the layer boundary. Overall, it is obvious that the calculated p soil reactions from the homogenous FEA differ significantly from those computed by a layered FEA. This

discrepancy underscores the inability of the ILM to accurately predict the pile's lateral behaviour, leading to a severe underestimation of the pile capacity.

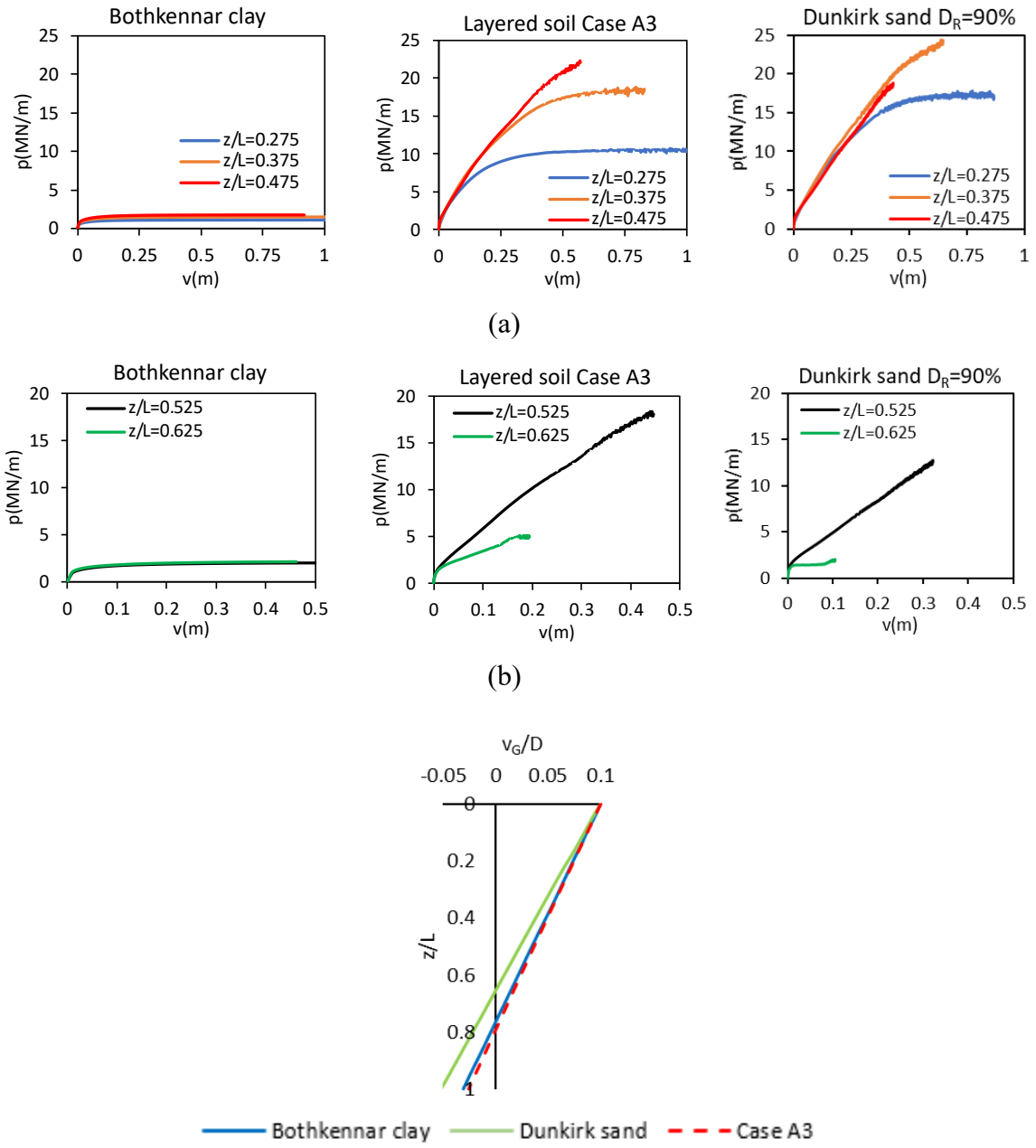
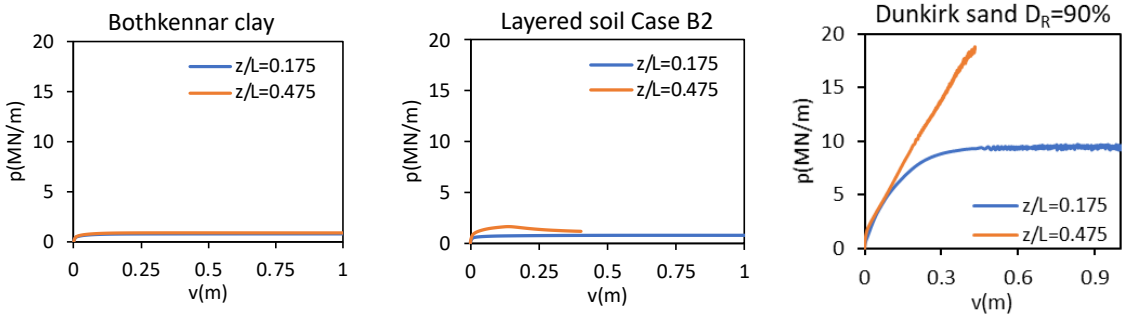


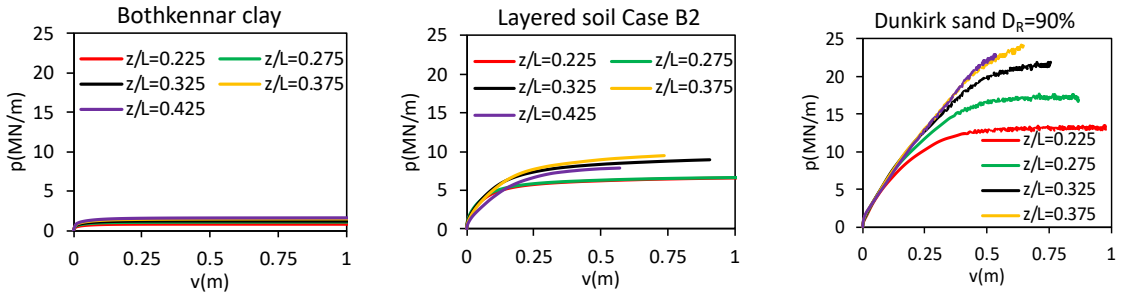
Figure 6.14: Effect of layering on soil reaction curves for a two-layered soil profile of a Bothkennar clay layer overlying a dense sand layer of D_R=90% (Case A3 presented in Table 6.1). (a) Soil reactions above the layer boundary at normalised depths of z/L=0.275, 0.375, 0.475. The first column indicates the Bothkennar clay soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. (b) Soil reactions below the layer boundary at normalised depths of z/L=0.525, 0.625. The third column indicates the Dunkirk sand soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. (c) Pile deflection curves for the layered analyses and the two reference homogenous analyses. For all soil reaction curves the magnitude of p increases as the value of z/L increases.

Single embedded layer – Cases B2 & B5

Figure 6.15 and Figure 6.16 present the impact of introducing a relatively thin embedded layer of dense sand within a soft clay matrix on soil reaction curves extracted from analyses on the reference homogenous soil profiles. These figures correspond to Cases B2 and B5 presented in Table 6.2. From Figure 6.15(b) and Figure 6.16(b), it can be observed that the soft clay matrix induces a significant reduction in the magnitude of the p soil reaction curves within the thin embedded dense sand layer compared to the reference homogeneous soil reactions. This reduction is mainly caused by the substantial difference in soil strength between the two neighbouring soil layers. This phenomenon cannot be captured by the ILM resulting in the significant overprediction of the pile capacity for Case B2 and B5. Interestingly, the results also indicated that as the thickness of the embedded layer decreases, the layering effects in the embedded layer will increase and the magnitude of the soil reaction curves will drop further. Figure 6.15(a) and Figure 6.16(a) indicate that the thin embedded layer of dense sand has a negligible effect on the soil reaction curves extracted from the adjacent soft clay layer; the soil reaction curves are broadly similar to those extracted from the reference homogenous analysis. This observation implies that thin embedded layers only have a small effect on the soil reactions in neighbouring soil layers. Therefore, as demonstrated in Burd et al. (2020b), the ILM can accurately calculate the pile's ultimate response for multi-layered configurations comprising multiple thin layers.

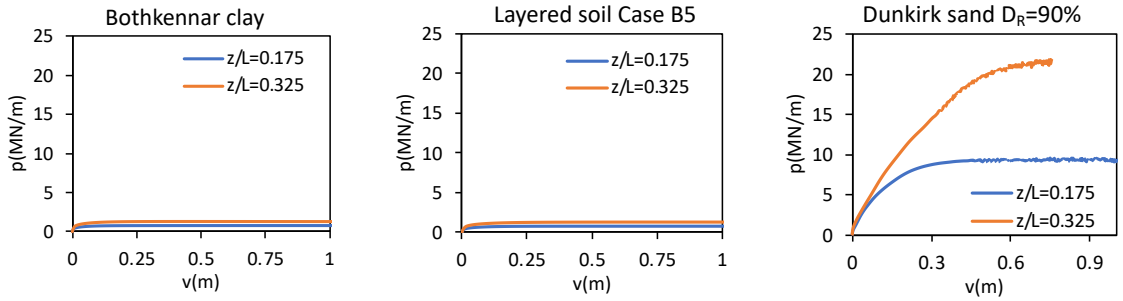


(a)

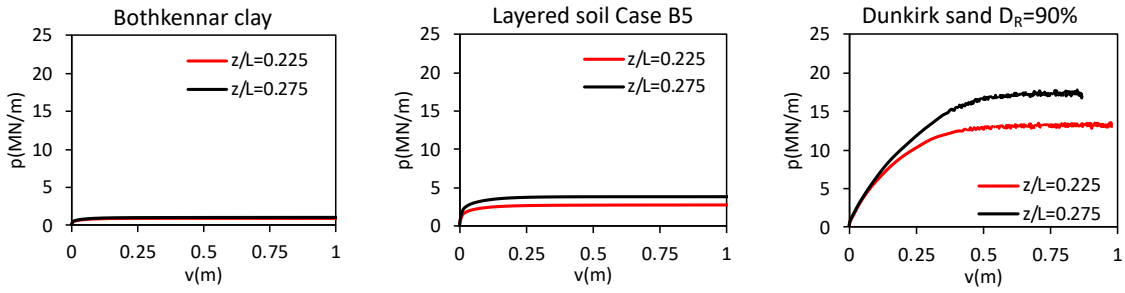


(b)

Figure 6.15: Effect of layering on soil reaction curves for a relatively thin embedded layer of dense sand of D_R=90% within a Bothkennar clay matrix (Case B2 presented in Table 6.2). (a) Soil reactions located just above (z/L=0.175) and below (z/L=0.475) the embedded sand layer. The first column indicates the Bothkennar clay soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. (b) Soil reaction curves located within the embedded sand layer at normalised depths of z/L=0.225, 0.275, 0.325, 0.375, 0.425. The third column indicates the Dunkirk sand soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. For all soil reaction curves the magnitude of p increases as the value of z/L increases.



(a)



(b)

Figure 6.16: Effect of layering on soil reaction curves for a very thin embedded layer of dense sand of D_R=90% within a Bothkennar clay matrix (Case B5 presented in Table 6.2). (a) Soil reactions located just above (z/L=0.175) and below (z/L=0.325) the embedded sand layer. The first column indicates the Bothkennar clay soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. (b) Soil reaction curves located within the embedded sand layer at normalised depths of z/L=0.225, 0.275. The third column indicates the Dunkirk sand soil reaction curves that would have been used by the ILM instead of the layered soil profile soil reaction curves in the second column. For all soil reaction curves the magnitude of p increases as the value of z/L increases.

6.5 Rule-based method for layered soils

This section explores a simple rule-based method to adjust the soil reaction curves computed by the data-driven 1D design models developed for the reference homogenous soil profiles to improve their accuracy in layered soil predictions. As shown in Section 6.4, even for simple configurations, the interactions between neighbouring layers are highly complex and depend on multiple factors (e.g. type of layered configuration, number of layers, thickness of layers, pile geometry, soil strength). It is challenging to quantify these different factors in a rule-based method. For this scope, a rule-based method is examined for a two-layered soil profile.

The proposed approach employs a simple set of rules to scale the layers' strength and stiffness properties over an assumed transition depth. The transition depth is defined as the depth above and below the boundary between two adjacent layers where the layering effects are evident. The scaling rules are applied over the s_u and G_o profiles for clays and D_R and G_o profiles for sands by:

$$\begin{aligned} s_u &= [1/2 \times ({s_{u_w}}^n + {s_{u_{str}}}^n)]^{1/n}; \\ G_o &= [1/2 \times ({G_{o_w}}^n + {G_{o_{str}}}^n)]^{1/n}; \\ D_R &= [1/2 \times ({D_{R_w}}^n + {D_{R_{str}}}^n)]^{1/n} \end{aligned} \tag{6.1}$$

where n is a constant depending on the type of layered configuration; s_{u_w}, G_{o_w}, D_{R_w} are the soil properties of the weak layer; $s_{u_{str}}$, $G_{o_{str}}$, $D_{R_{str}}$ are the soil properties of the strong layer. It was found that for a two-layered soil profile $n = 2$ provided generally accurate results of the updated pile lateral behaviour.

Therefore, the general clay or sand data-driven 1D design models can be used to compute the soil reactions of the new equivalent scaled soil profiles. However, Equation (6.1) only applies to adjacent layers of the same type. There is no straightforward method to relate s_u

with D_R; hence this approach has limited application for layered soils that incorporate either clay or sand layers.

A transition depth z_T related to the weak and the strong layer can be determined above and below the layer boundary by inspecting the soil reaction curves computed by 3D FEA for the layered soil and comparing them to the soil reaction curves extracted by FEA implemented on the reference homogenous soil profiles.

Since the layering effects are affected by L/D, to simplify the problem the transition depth is related to the pile length as $z_T = aL$; where a is a parameter determined by the type of layered soil and the depth of the layer boundary z_b. Figure 6.17 illustrates the depth variation of parameter a along the pile length for a two-layered profile. As it can be observed a depends on the position of the layer boundary across the pile's length ($0 < \frac{z_b}{L} < 1$). The rules were designed for the more realistic scenario where a strong layer underlies a weak layer. The appropriate value of a was estimated by observing up to which depth the interactions between adjacent layers influenced the soil reaction curves.

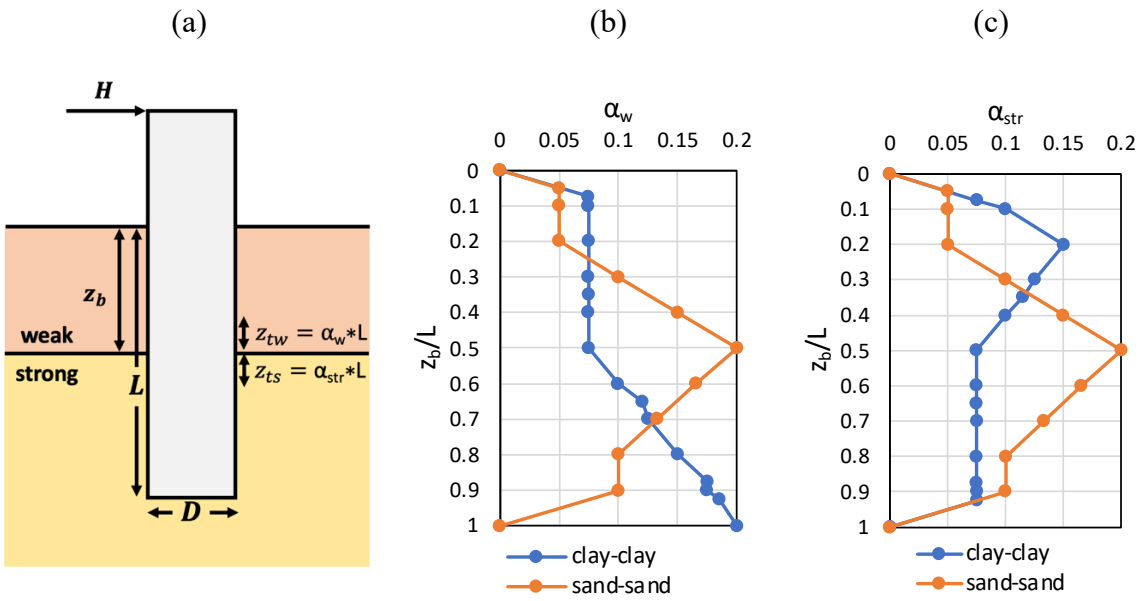


Figure 6.17: (a) Two-layered soil profile where a weak clay or sand layer is located on top of a strong clay or sand layer. (b) Values of a_w determining the transition depth on the top weak layer depending on the position of the layer boundary z_b. (c) Values of a_{str} determining the transition depth on the bottom strong layer depending on the position of the layer boundary z_b.

6.5.1 Case studies

Case A1- two-layered clay profile

Figure 6.18(a) and Figure 6.18(b) present the equivalent s_u and G_o profiles for Case A1 (Table 6.1) calculated by Equation (6.1). Based on Figure 6.17(b) and Figure 6.17(c), it is determined that $a_w = a_{str} = 0.075$. The equivalent soil properties were used to drive the GCCM data-driven 1D design model presented in Chapter 5. Figure 6.18(c) illustrates the comparison between the load-displacement response at ground level predicted by the rule-based layered method (RBLM), the ILM 1D models and the PLAXIS 3D FEA results. It can be observed that the RBLM outperforms the ILM predicting the pile's lateral behaviour.

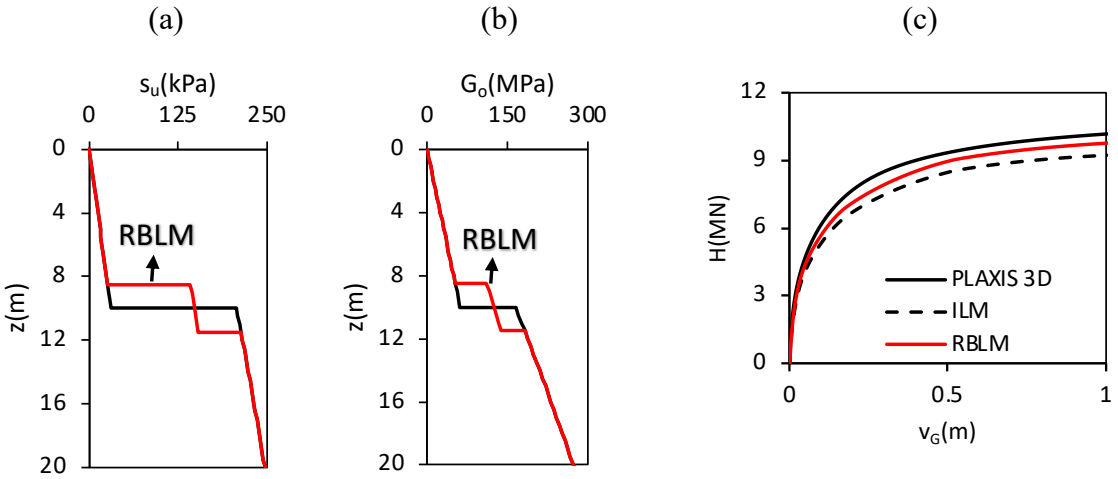


Figure 6.18: Application of the RBLM on the Case A1 layered clay (Table 6.1). Equivalent profile of: (a) s_u, (b) G_o. (c) Comparison between the load-displacement responses at ground level for the ultimate response using the RBLM, the ILM and the 3D FEA.

Case A2- two-layered sand profile

Figure 6.19(a) and Figure 6.19(b) present the equivalent D_R and G_o profiles for Case A2 (Table 6.1) calculated by Equation (6.1). Based on Figure 6.17(b) and Figure 6.17(c), it is determined that $a_w = a_{str} = 0.2$. These soil profiles are used to drive the GDSM V2 data-driven 1D design model presented in Chapter 5. Figure 6.19(c) shows the comparison between the load-displacement response at ground level predicted by the RBLM, the ILM

1D models and the PLAXIS 3D results. The RBLM improves the accuracy of the ILM results and provides a good match to the PLAXIS 3D results.

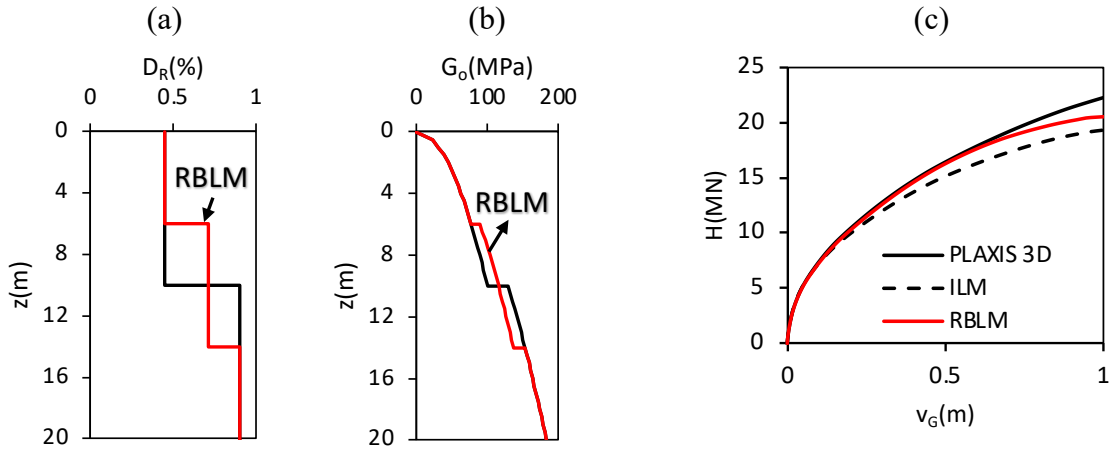


Figure 6.19: Implementation of the RBLM on the Case A2 layered sand (Table 6.1). Equivalent profile of: (a) D_R, (b) G_o. (c) Comparison between the load-displacement responses at ground level for the ultimate response using the RBLM, the ILM and the 3D FEA.

6.5.2 Key observations

A rule-based layered method (RBLM) is proposed for layered soils with layers of the same soil type. The RBLM has been tested on two-layered soil configurations, and the results indicate that the RBLM enhances the accuracy of the ILM 1D model. However, the RBLM is not scalable because it requires a large number of FEA to extend the method to configurations with more layers. Furthermore, the RBLM has limited application since it cannot be applied to configurations that include both clay and sand layers.

6.6 Data-driven general layered soil models

This section explores the feasibility of developing general layered soil models by using as a database the soil reaction curves that have been extracted from FEA implemented for piles embedded in layered soils. The aim is to develop a framework to generate a general layered (GL) model using the data-driven method that is capable of calculating the pile's monotonic lateral behavior over a range of different layered soils.

The soil reaction curves extracted from the 3D FEA are normalised using the 'sand' data-driven normalisation framework in Section 4.4. The GPR model is adopted to train the ML models that drive the general layered model; GPR was shown in Chapter 4 to form high fidelity data-driven 1D design models for homogenous soils. The depth variation of the eight-knot point soil reaction curves (GPR predictions) are connected to the eight-knot point piecewise spline function in the OxPile software to predict the monotonic loading response of laterally loaded monopiles embedded in layered soils.

The number of training inputs is now increased to include both the clay and sand soil strength properties and the soil layering effects. GPR models employ the following training inputs; $[L/D]$ is a proxy for the pile slenderness; $[z/L]$ characterises the depth variation of the knot; $[G_o/\sigma'_v]$ discriminates locally the soil reaction curves with the same values of $[z/L]$ and $[L/D]$; $[s_u/\sigma'_v]$ is employed to discriminate clay profiles with different strength properties; $[D_R]$ is used to discriminate sand profiles with different relative density. It should be noted that for sands $[s_u/\sigma'_v]$ and for clays $[D_R]$ are always set to zero assisting the GPR model to discriminate the strength between a clay and a sand soil profile. The layering effects are incorporated in the model by employing $[z_{Li}/L]$ (where z_{Li} is the thickness of each individual layer); the number of the $[z_{Li}/L]$ inputs adopted depends on the number of layer boundaries z_b within the pile's embedment length.

6.6.1 General two-layered soil model

A two-layered soil profile employs a single layering input $[z_{L1}/L]$, as there is only a single layer boundary found within the pile's embedment length. This section presents the generation of a general two-layered 1D model for two specific soil layers (Bothkennar clay over Dunkirk sand with $D_R = 90\%$) for all depths of the layer boundary. This soil configuration was shown in Section 6.3 to be particularly challenging for the ILM due to the

different soil strength and stiffness characteristics of the two neighboring layers. The proposed method is also applicable for two-layered configurations involving only clay layers or sand layers with different strength and stiffness properties.

The pile geometries adopted for the general two-layered model are presented in Table 6.5. The selected pile geometry calibration space includes 12 piles with three different L/D values to accurately capture the dependency between the pile geometry and the depth at which interactions occur between neighbouring layers. The FEA performed to develop this model are detailed in Appendix A (refer to Table Table A.*3*).

Table 6.5: Pile geometry calibration space adopted for the general layered 1D models. L is the embedded length; D is the pile diameter; h is the load eccentricity; t_w is the pile wall thickness.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D	Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$
LC1	10	20	91	50	2	**LC7**	5	10	45	25
LC2	10	20	91	150	2	**LC8**	5	10	45	75
LC3	10	40	91	50	4	**LC9**	5	20	45	25
LC4	10	40	91	150	4	**LC10**	5	20	45	75
LC5	10	60	91	50	6	**LC11**	5	30	45	25
LC6	10	60	91	150	6	**LC12**	5	30	45	75

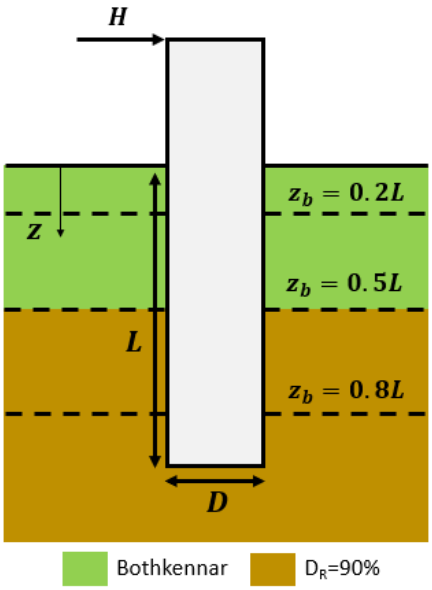


Figure 6.20: Possible depths of layer boundary z_b adopted for the calibration analyses for the general two-layered model with Bothkennar clay located over Dunkirk sand with $D_R = 90\%$.

A total of 36 FEA were employed to extract the calibration soil reaction curves to train the two-layered general model. The GPR model employing the Matérn kernel ($v = 5/2$) and the white noise kernel was adopted for the training process. An 80:20 training-test set split ratio was adopted. The MSE and R^2 metrics were employed to assess the trained GPR models. The location of the layer boundary z_b was found to govern the depth up to which there are interactions between neighbouring layers. Three different locations were employed for the layer boundary (Figure 6.20) where; the layer boundary is near the head of the pile at a depth of $z_b = 0.2L$; the layer boundary is located at the mid-depth of the pile where $z_b = 0.5L$; the layer boundary is located near the pile tip at a depth of $z_b = 0.8L$. In addition, the soil reaction curves extracted from the FEA for the homogenous Bothkennar clay profile were included in the training of the general two-layered soil model representing the case where $z_b = L$. The general two-layered soil model can be used to predict the pile's lateral behaviour for $0 < \frac{z_b}{L} \leq 1$. It should be noted that for a two layered soil profile the layer boundary depth coincides with the thickness of the top layer ($z_b = z_{L1}$).

Table 6.6 presents the optimised ($l_{z_{L1}/L}$) lengthscale kernel hyperparameter obtained for the distributed lateral load and moment knot point parameters $\bar{p}$, $\bar{m}$ after the GPR model training. The low values of the ($l_{z_{L1}/L}$) shows how sensitive the GPR predictions are to the value of z_{L1}/L, meaning that moving the layer boundary will have a significant effect on the $\bar{p}$, $\bar{m}$ knot point parameters.

Table 6.6: Optimised ($l_{z_{L1}/L}$) lengthscale kernel hyperparameter for the GPR models as applied to the knot points associated with the GPR model.

Knot point	$\bar{p}$	$\bar{m}$	**Knot point**	$\bar{p}$	$\bar{m}$
1	$l_{z_{L1}/L}$=0.957	$l_{z_{L1}/L}$=1.62	**5**	$l_{z_{L1}/L}$=0.667	$l_{z_{L1}/L}$=1.72
2	$l_{z_{L1}/L}$=0.896	$l_{z_{L1}/L}$=1.33	**6**	$l_{z_{L1}/L}$=0.522	$l_{z_{L1}/L}$=1.19
3	$l_{z_{L1}/L}$=0.662	$l_{z_{L1}/L}$=1.05	**7**	$l_{z_{L1}/L}$=0.870	$l_{z_{L1}/L}$=2.78
4	$l_{z_{L1}/L}$=0.756	$l_{z_{L1}/L}$=2.37	**8**	$l_{z_{L1}/L}$=0.913	$l_{z_{L1}/L}$=2.55

Figure 6.21 illustrates the depth variation predictions by the GPR models of the distributed lateral load $\bar{p}$ knot point parameters for pile LC1 $(L/D = 2)$ embedded in the two-layered soil for $z_{L1} = 0.5L$. The $\bar{p}$ knot point parameters for the same problem extracted from the 3D FEA used for the training of the ML models are also presented. The 95% confidence interval (shown as shaded bounds) is also shown. The GPR model can accurately capture the highly non-linear nature of the data. At the layer boundary $z_b = 0.5L$ there is a 'jump' between the distributed soil reactions indicating the transition from the soft clay to the dense sand. This 'jump' is introduced in the GPR model using the $[z_{L1}/L]$ training input; omitting this training input would result in consecutive 'jumps' at all the depths that the training points indicate that there is a layer boundary.

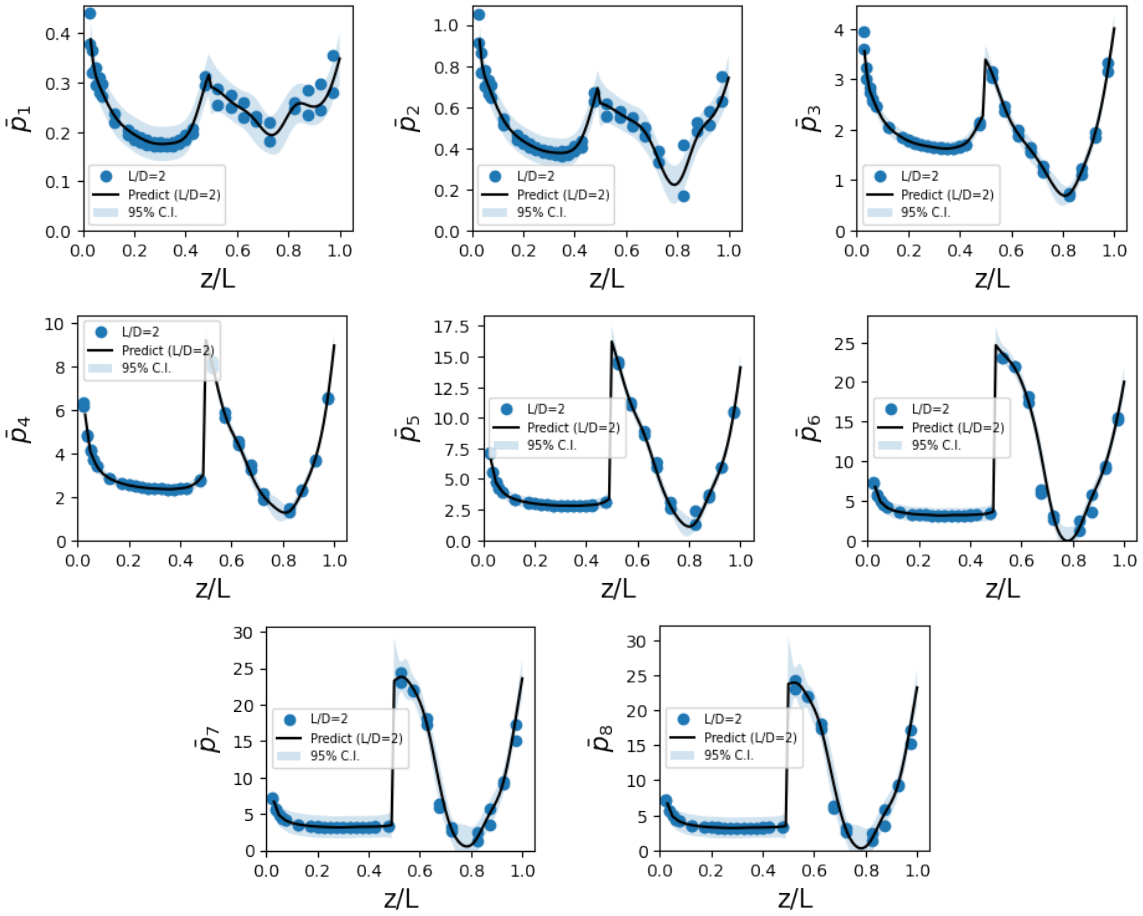


Figure 6.21: Example GPR ML outputs for the distributed lateral load $\bar{p}$ soil reactions knot points for the LC1, LC2 piles $(L/D = 2)$ embedded in the two-layered soil where $z_{L1} = 0.5L$. The depth variation of the mean GPR prediction for the knot points is indicated by the black line. The 95% confidence interval is shown (shaded bounds). The blue markers are the trained datapoints for the calibration analysis LC1, LC2 with $z_{L1} = 0.5L$.

Two design cases were selected to validate the 1D model, where both the pile geometry and the layer boundary depth are unseen for the trained GPR model. The pile for both design cases has a length of $L = 24m$ and a diameter of $D = 8m$; the load eccentricity is $h = 40m$ and the pile wall thickness is $t_w = 91mm$. In the first analysis, the layer boundary is located at $z_b = 0.3L$ ($z_{L1} = 0.3L$); in the second analysis, the layer boundary is located $z_b = 0.6L$ ($z_{L1} = 0.6L$). Figure 6.22 compares the load-displacement response at ground level computed by the 1D model and the PLAXIS 3D results for the pile's ultimate and small displacement response. The 1D model matches well the 3D results.

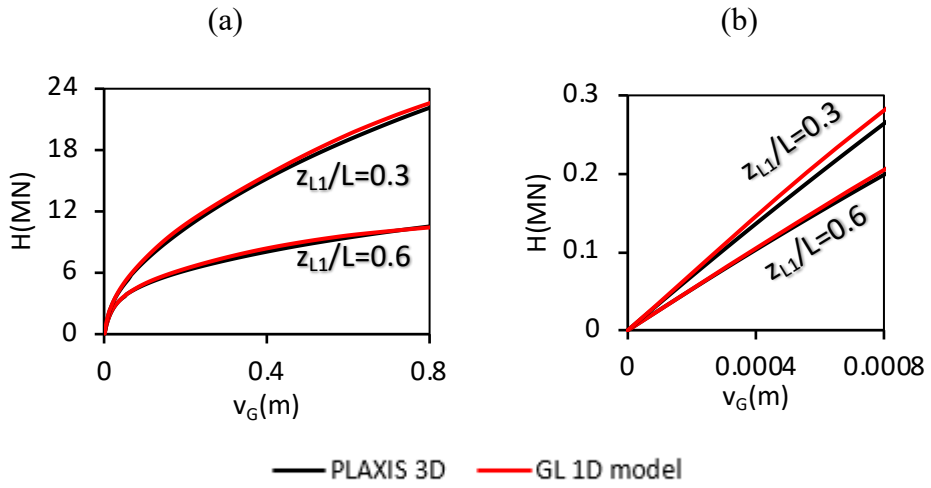


Figure 6.22: Comparison between the load-displacement responses at ground level for: (a) the ultimate response, (b) small displacement response for the general two-layered soil model and the 3D FEA.

6.6.2 Extending the model to three layers

This section explores the possibility of extending the general two-layered soil model presented in the previous section to account for more complicated layering configurations. The current study is concerned with the layered configuration B presented in Figure 6.2(b) which comprises an embedded layer in a soil matrix. The soil matrix is a Bothkennar clay layer and the embedded layer is a Dunkirk sand layer with $D_R = 90\%$. This layered configuration presented in Figure 6.23, employs two layering inputs $[z_{L1}/L]$, $[z_{L2}/L]$.

The pile geometries in Table 6.5 were employed for the pile calibration space. This new layered configuration defines a rather complex problem as both the position and thickness of the embedded layer will have a significant effect on the pile's lateral behaviour. Seventy-two additional FEA were conducted to extend the general two-layered soil model presented in Section 6.6.1 to account for the new layered configuration. Six different locations were employed for the position and thickness of the embedded layer (Table 6.7). These six different layered soil cases introduce to the general layered soil model the possibility of having a very thin, relatively thin or thick embedded layer located at the upper or the lower part of the pile. The FEA performed to develop this model are detailed in Appendix A (refer to Tables A.3 and A.4).

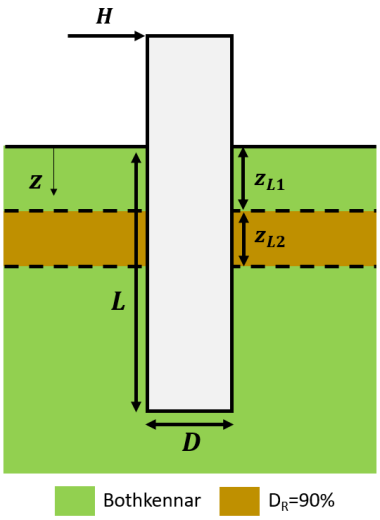


Figure 6.23:Embedded sand layer within a clay matrix employing two-layer boundaries.

Table 6.7: Layers boundaries depth adopted for the calibration FEA.

Embedded Case	z_{L1}/L	z_{L2}/L
I	0.2	0.25
II	0.2	0.1
III	0.6	0.25
IV	0.6	0.1
V	0.2	0.6
VI	0.3	0.4

The updated training database (i.e. incorporating the data from the three-layer case) contains data on about 4000 soil curves for $\bar{p}, \bar{m}$ (this corresponds to 32000 knot points). This large amount of data, employed with the non-parametric form of the GPR model, requires substantial computational training time (more than 12 hours for the total of 18 GPR models) and memory demand. A sparse GPR model (presented in more detail in Section 2.4.2) was therefore employed for the training of the $\bar{p}, \bar{m}$ knot point parameters to reduce the computational training time and demand. Sparse GPR models are GPR approximations that do not form the full covariance matrix but instead they use a number of inducing points ('pseudo' points) which are strategically located throughout the training domain to approximate the covariance matrix and reduce both the computational time and the memory requirements. However, the sparse GPR models have the disadvantage that they reduce the expressiveness of the GPR models. Two sparse GPR approaches were trialled: (i) the Fully Independent Training Conditional (FITC) (Snelson Ghahramani 2006), (ii) the Variational Free Energy (VFE) (Titsias 2009). For the training database employed, the results along with the MSE and R^2 train and test scores showed that VFE significantly outperformed the FITC sparse GPR model agreeing with Van de Wilk (2018) and Rossi et al., (2021), hence the VFE results are presented. The sparse GPR models were employed using the *GPflow* GPR Python libraries (Matthews et al., 2017).

The total number of inducing points has a significant effect on the accuracy of the predictions and the training time. For the current work it was decided to use 450 inducing variables which correspond to approximately 10% of the datapoints. Another key factor for VFE sparse GPR models, is the initial selection of the inducing points which has a significant influence on the final inducing points that are determined by the optimisation of the GPR model's hyperparameters. Two methods were trialled to determine the 450 initial inducing variables that were used for the VFE sparse GPR models: (i) randomly selecting 450

datapoints from the dataset, (ii) using the K-means algorithm. The MSE and R^2 train and test scores showed that using the K-means algorithm is the best approach to determine the initial inducing variables for the specific database. K-means is a fast-unsupervised machine learning algorithm; the algorithm separates the datapoints into clusters and updates the cluster centroids until the position of the centroids is stable over a successive number of iterations (Hartigan and Wong, 1979). The K-means cluster centroids were employed as the initial inducing variables.

Consistent with the GPR models presented in this thesis, the σ_f^2 and l_i hyperparameters as well as the inducing variables of the sparse VFE GPR models were optimised by maximising the log marginal likelihood. The Matérn kernel ($v = 5/2$) is employed for the VFE GPR model. The white noise kernel was not added to the Matérn kernel, since it was found that for sparse GPR models it had a negligible influence on the final results. The final training time of the 16 VFE sparse GPR models trained for the distributed knot point soil reactions reduced to approximately 30 minutes.

Figure 6.24 presents the distributed lateral load $\bar{p}$ depth variation predictions by the VFE sparse GPR models for pile LC1 ($L/D = 2$) embedded in a layered profile where the thickness of the two top layers are $z_{L1} = 0.2L$, $z_{L2} = 0.25L$. The $\bar{p}$ knot point parameters are also presented (black crosses) along with the 95% confidence interval (shaded bounds). The VFE sparse GPR models employing 450 inducing variables and using the Matérn kernel can accurately reproduce the non-linear nature of the data. The results are indicative of the capabilities of the VFE sparse GPR; the model does not reduce the expressiveness of the GPR models and reduces substantially the training computational time. Thus, the results are highly dependent on the initial determination and number of inducing variables.

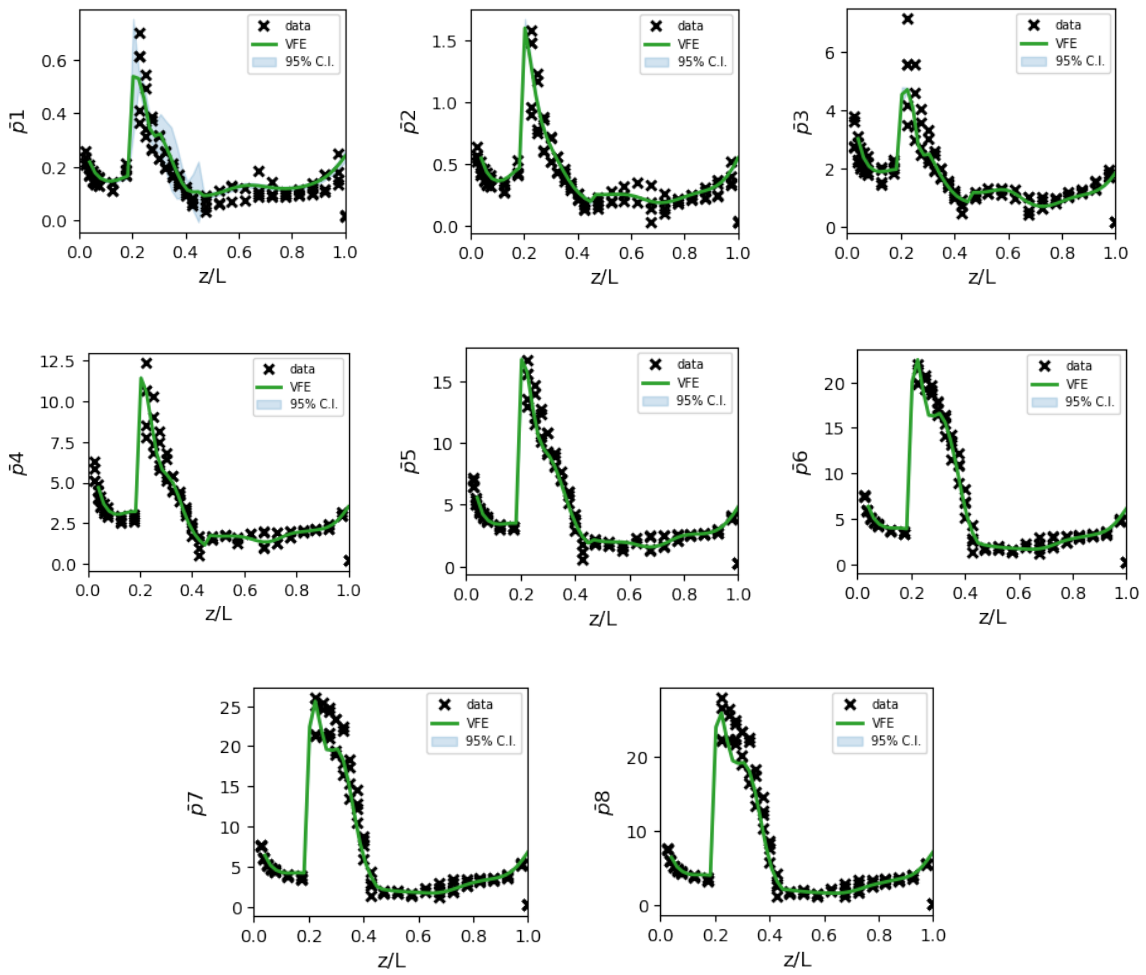


Figure 6.24: Example VFE sparse GPR ML outputs for the distributed lateral load $\bar{p}$ soil reactions knot points for the LC1 pile *($L/D = 2$)* embedded in a layered soil where $z_{L1} = 0.2L$, $z_{L2} = 0.25L$. The depth variation of the mean sparse GPR prediction for the knot points is indicated by the green line. The 95% confidence interval is shown (shaded bounds). The black crosses markers are the trained datapoints for the calibration analysis LC1, LC2, LC9, LC10 with $z_{L1} = 0.2L$, $z_{L2} = 0.25L$.

Two design cases were selected to check the accuracy of the VFE sparse GPR models for unseen layer boundaries. For both cases the LC1 pile was employed with $L/D = 2$. In the first case only the second- and third-layer thickness is unseen ($z_{L1} = 0.2L$, $z_{L2} = 0.4L$); in the second case all three-layer thicknesses are unseen ($z_{L1} = 0.35L$, $z_{L2} = 0.15L$). Figure 6.25 compares the load-displacement response computed by the general layered (GL) soil model and the PLAXIS 3D results for the pile's small displacement and ultimate response. The 1D model reproduces the 3D FE results with good accuracy.

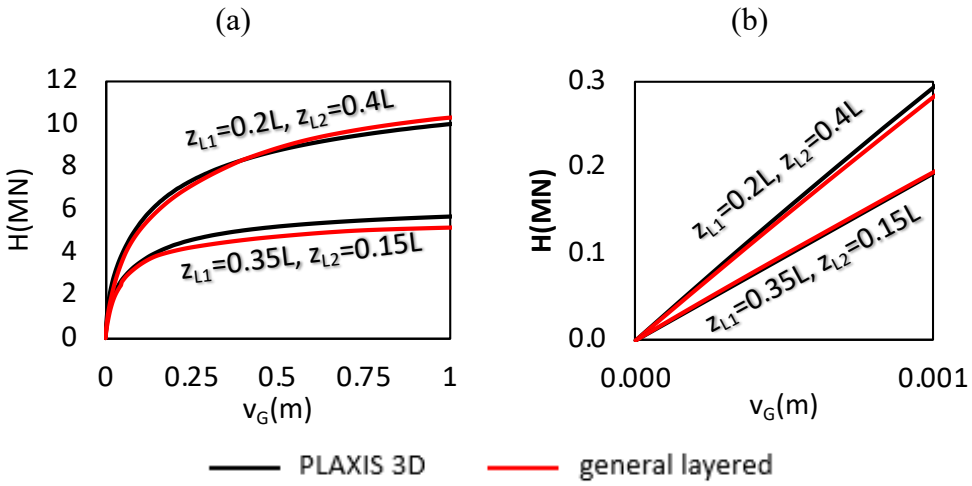


Figure 6.25: Comparison between the load-displacement responses at ground level for: (a) the ultimate response, (b) small displacement response from the general layered soil model that employs the VFE sparse GPR models and the 3D FEA.

6.6.3 Key observations

The results show that it is feasible to develop a general layered soil model for two- or three-layered soil configurations. Such a model computes high fidelity soil reaction curves accurately capturing the layering effects and the soil reaction physics since the model is calibrated directly from layered analyses. However, typically the layered configurations that are found at offshore locations involve more than two or three soil layers. Also, within an offshore wind farm there is spatial variability in the soil strength and stiffness characteristics. It would be practically impossible to develop a general layered soil model for all possible layered configurations, as this would require a considerably high number of expensive calibration 3D FEA. Hence, this design approach is not practical for conventional design.

6.7 Site-specific layered soil data-driven 1D design models

This section explores the feasibility of generating site-specific layered soil models using the data-driven method. These 1D models are intended for use exclusively with layered soil profiles falling within the calibration space. This approach differs from the general layered models developed in Section 6.6, where the aim was to create a 1D model capable of making

predictions regardless of layer thicknesses. The site-specific layered soil models are calibrated by soil reaction curves that are extracted from FE analyses of piles embedded in layered soil profiles. The soil reactions curves are normalised using the 'sand' data-driven normalisation framework presented in Section 4.4. The GPR machine learning model is adopted to train the machine learning models that drive the data-driven 1D design model; GPR predictions are coupled with the OxPile software to predict the monotonic loading response of laterally loaded monopiles.

The case study presented in this section is concerned with layered Case D3 in Table 6.4 which employs a multi-layered soil profile with the combination of Bothkennar clay and Dunkirk sand with $D_R = 90\%$; this combination was found to be problematic for the ILM design approach. The procedure of generating site-specific layered soil models was tested also for configurations E2 and E3 (Figure 6.11) (for brevity the results are not presented) yielding consistent results. Therefore, the method is readily extended for different layered configurations which involve clay or sand layers of different strength and stiffness characteristics. Two variations of the site-specific layered design approach are presented examining the capabilities of the method: (i) at a single location, (ii) accounting for spatial variations.

6.7.1 Site-specific layered soil at a single location

Site-specific layered soil models are designed for a single location. A pile geometry calibration space represented by 8 to 12 piles is defined similarly to the PISA project (Byrne et al., 2020b; Burd et al., 2020a), involving piles of different geometry characteristics (L/D, h/D, D/t). Finite element analyses are then undertaken to determine the soil reaction curves. The ML model is trained using these soil reaction curves and is used to provide predictions for piles that fall within this geometrical calibration space.

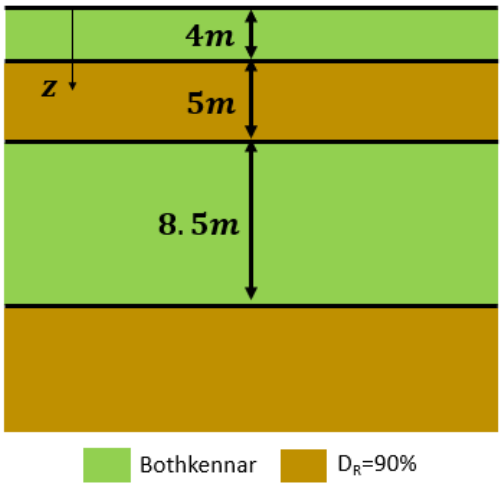


Figure 6.26: Adopted soil profile to test the site-specific layered soil method at a single location. The combination of Bothkennar clay and Dunkirk sand with $D_R = 90\%$ is considered as this was found to be a problematic combination for the ILM.

Figure 6.26 illustrates the soil configuration that was employed for this study. The pile geometry calibration space adopted for the site-specific layered soil model at a single location is presented in Table 6.8 (LC1-LC10). The selected pile geometry calibration space includes piles with three different values of L/D to accurately capture the dependency between the pile geometry and the extent of layering effects. The offshore piles employ L/D ranging between 1.5 and 3. Two eccentricity loads ($h/D = 5, 15$) were considered. The FEA performed to develop this model are detailed in Appendix A (refer to Table A.5).

Table 6.8: Pile geometry calibration space adopted for the site-specific layered 1D model. L is the embedded length; D is the pile diameter; h is the load eccentricity; t_w is the pile wall thickness.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D
LC1	12	18	125	60	1.5
LC2	12	18	125	180	1.5
LC3	12	36	125	60	3
LC4	12	36	125	180	3
LC5	8	12	91	40	1.5
LC6	8	24	91	40	3
LC7	8	24	91	120	3
LC8	10	15	91	50	1.5
LC9	10	15	91	150	1.5
LC10	10	30	91	50	3

Different training inputs are employed for the distributed lateral load and moment soil reactions than those used for the base soil reactions. The $\bar{p}$, $\bar{m}$ employ as training inputs the $[L/D]$, $[z/L]$, $[G_o/\sigma'_v]$, $[s_u/\sigma'_v]$, $[D_R]$, $[z_{L1}/L]$ ratio (where z_{L1} is the thickness of the top layer) and $[nlay]$ (which is an integer number indicating the number of layers that the pile has penetrated). For each of the base soil reactions, $\bar{H}_B$, $\bar{M}_B$, the GPR model employs a small training database of 10 points. Hence, it was decided to reduce the number of training inputs since employing 6 inputs for a 10 points database would reduce the expressiveness of the GPR which would effectively predict by averaging the values of all the datapoints. Hence, the base soil reactions employ four training inputs: $[L/D]$, $[h/D]$, $[G_o/\sigma'_v]$, $[D_R]$. The GPR models employ the Matérn ($v = 5/2$) and the white noise kernel for the training process of the $\bar{p}, \bar{m}, \bar{H}_B, \bar{M}_B$. An 80:20 training-test set split ratio was used. All the trained GPR models were assessed by employing the MSE and R^2 metrics.

The base soil reaction curves of the site-specific layered soil models employ the $[h/D]$ training input. This input was found to be redundant for homogenous soil profiles in Chapter 4 and was disregarded from the training procedure of the GPR models trained for homogenous soil profiles. However, it was found that for the layered soil (Figure 6.26) employed, the $[h/D]$ is a redundant proxy only for the distributed lateral load and moment soil reactions and has a significant influence on the base soil reactions. Figure 6.27 illustrates the base and moment forces extracted for piles LC1 ($h/D = 5$) and LC2 ($h/D = 15$). The $h/D = 15$ generates approximately double values of $\bar{H}_B$, $\bar{M}_B$ compared to the $h/D = 5$. Hence, the $[h/D]$ proxy is required for the GPR model training or else the GPR model would simply interpolate between the two base soil reactions producing false predictions. For the current set of analyses most of the calibration piles have their tip embedded in the dense sand layer. Therefore, the base soil reactions contribute a significant part of the pile's lateral capacity and the site-specific layered model needs to predict $\bar{H}_B$, $\bar{M}_B$ accurately.

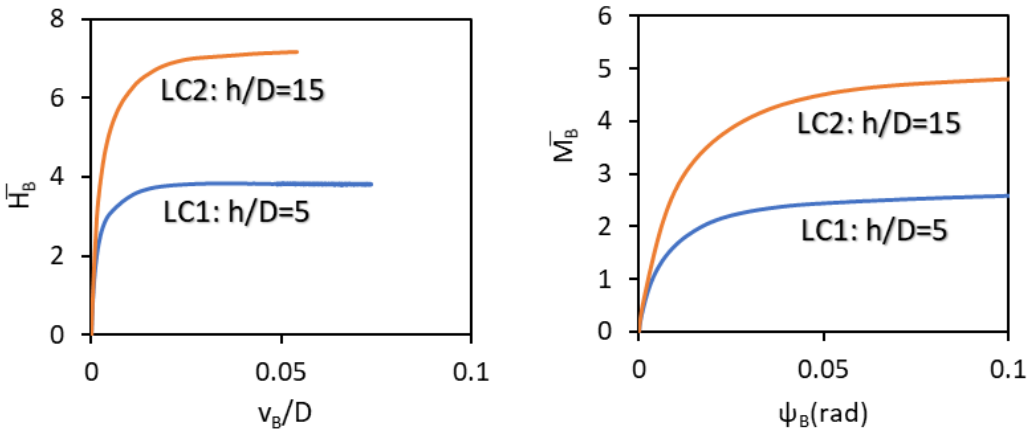


Figure 6.27: Base shear soil reaction (on the left) and base moment soil reaction (on the right) for piles LC1 and LC2.

Two design cases were selected to validate the site-specific layered soil model. The first case comprises case D3 in Table 6.4 where the pile LD1 has a length of $L = 20m$, a diameter of $D = 10m$, the lateral load is applied at $h = 50m$ and $t_w = 91mm$. The second case employs a pile LD2 with a length of $L = 22.5m$, a diameter of $D = 9m$, $h = 90m$ and $t_w = 91mm$. Figure 6.28 compares the load-displacement response at ground level predicted by the site-specific layered soil model and the PLAXIS 3D results for the pile's small displacement and ultimate response. The 1D model accurately calculates the pile's lateral behaviour.

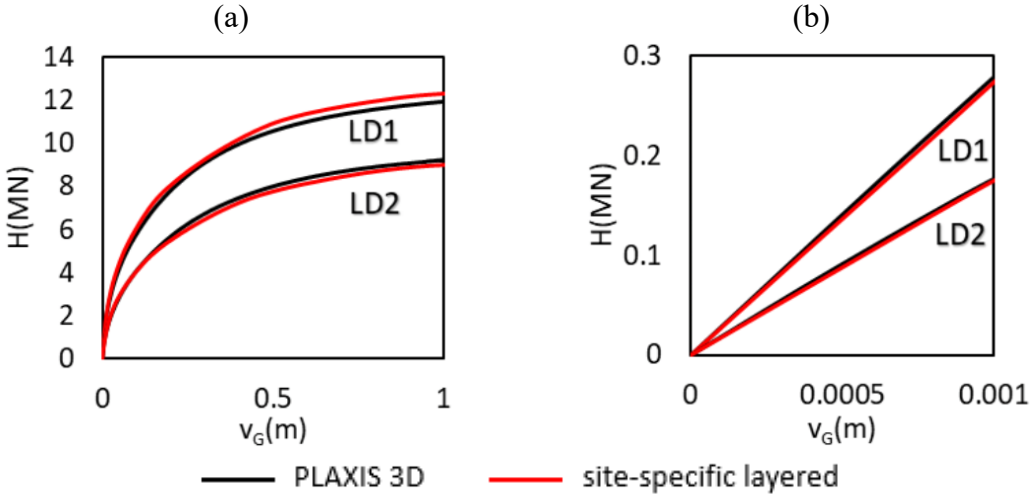


Figure 6.28: Comparison between the load-displacement responses at ground level for: (a) the ultimate response, (b) small displacement response from the site-specific layered soil model for a single location and the 3D FEA.

6.7.2 Site-specific layered soil 1D models accounting for spatial variations

This method is an extended version of the site-specific layered soil model at a single location. This approach accounts for possible spatial variations of the soil layers' strength and stiffness properties and the layers' thickness across a wind farm site. To determine the soil reaction curves, a series of finite element calibration analyses is undertaken for piles embedded in two or three different calibration layered soil profiles. The calibration layered soil profiles are selected by employing: (i) the minimum, (ii) the average and (iii) the maximum possible values of depth for each individual layer, that are expected across the wind farm. The required number of calibration layered soil profiles is mainly determined by the range of variability in each soil layer thickness. For example, if the layer thickness shows high variability across the wind farm more calibration layered soil variations are required to develop a 1D model on the safe side, since big changes in the layer thickness are expected to change significantly the pile's lateral behaviour. The pile geometrical calibration space remains the same for all different calibration layered soil profiles.

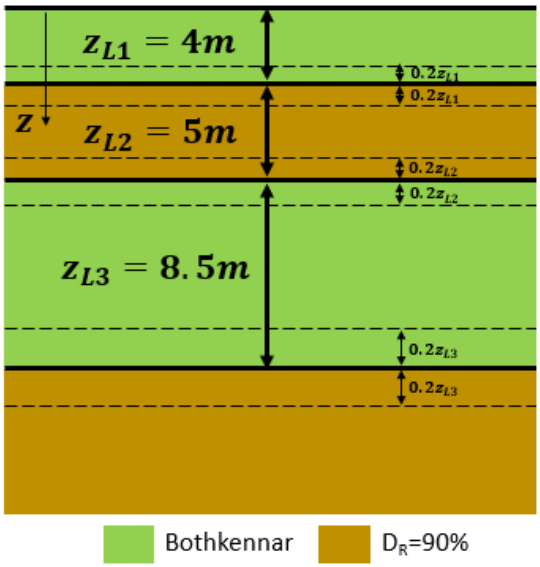


Figure 6.29: Adopted soil profile to test the site-specific layered soil model accounting for layer thickness spatial variation. The combination of Bothkennar clay and 90% relative density Dunkirk sand is considered as this was found to be a problematic combination for the ILM. The calibration space incorporates variations of the layer thicknesses (z_{L1}, z_{L2}, z_{L3}) by $\pm 20\%$ to account for possible variations in layer thickness across an offshore site.

Table 6.9: Layer thickness variations adopted for the calibration layered soil analyses.

Calibration layered soils	Layer Thickness
Min (z_{Li})	z_{L1}=3.2m, z_{L2}=4m, z_{L3}=6.8m
Max (z_{Li})	z_{L1}=4.8m, z_{L2}=6m, z_{L3}=10.2m

Similar to the site-specific layered soil model for a single location presented in Section 6.7.1, the analyses focus on Case D3 in Table 6.4. The site-specific layered soil model accounts for spatial variations of the layers thickness adopting the hypothesis that each one of the layers depth can vary up to 20% above or below the layered soil D3 configuration (Figure 6.29). This hypothesis is a simplified case where there is little variation in the layer thickness across the wind farm site. Therefore, two calibration layered soil profiles are sufficient to form the layered soil calibration space. The adopted layer thickness for the two calibration layered soils are presented in Table 6.9 employing the minimum and maximum possible layer thickness. The pile geometry calibration space adopted for the site-specific layered soil model is presented in Table 6.8. In total 20 FEA (10 piles × 2 layered soil profiles) were conducted to determine the calibration soil reaction curves. The FEA performed to develop this model are detailed in Appendix A (refer to Table A.6).

The GPR model employing the Matérn kernel ($\nu = 5/2$) plus the white noise kernel were used for the training process. An 80:20 training-test set split ratio was adopted and the GPR models were evaluated using the MSE and R^2 train and test scores. Similar to the procedure presented in Section 6.7.1 different training inputs are employed for the distributed lateral and moment load soil reactions compared to the base soil reactions. The distributed $\bar{p}$, $\bar{m}$ soil reactions employ as training inputs the $[L/D]$, $[z/L]$, $[G_o/\sigma'_v]$, $[s_u/\sigma'_v]$, $[D_R]$, $[z_{Li}/L]$ (where z_{Li} is the thickness of each layer and the total number of these inputs depend on the layered configuration). The $[z_{Li}/L]$ proxy is used to account for the spatial variation of the

layer thickness across the wind farm site. The base soil reactions $\bar{H}_B$, $\bar{M}_B$ employ five training inputs: $[L/D]$, $[h/D]$, $[G_o/\sigma'_v]$, $[s_u/\sigma'_v]$, $[D_R]$ ratios.

Table 6.10: Layer thickness variations adopted for the design layered soil analyses.

Design layered soils	Pile geometry	Layer thickness
DLS1	LD1, LD2	z_{L1}=4m, z_{L2}=5m, z_{L3}=8.5m
DLS2	LD1, LD2	z_{L1}=4.5m, z_{L2}=4.5m, z_{L3}=9m
DLS3	LD1, LD2	z_{L1}=3.8m, z_{L2}=5.3m, z_{L3}=7.3m

Six design cases were selected to validate the site-specific layered soil model accounting for spatial variations of the layer thickness. The unseen inputs for the GPR models are $[L/D]$, $[z/L]$, $[z_{L1}/L]$, $[z_{L2}/L]$, $[z_{L3}/L]$. The design piles LD1, LD2 geometries were employed presented in Table 6.8. The thickness of each soil layer for the design cases is presented in Table 6.10 where $3.2m \leq z_{L1} \leq 4.8m,\ 4m \leq z_{L2} \leq 6m,\ 6.8m \leq z_{L3} \leq 10.2m$. Figure 6.30 illustrates the predictions of the GPR models for pile LD1 embedded within the layered soil DLS1. The knot point soil reaction parameters extracted from a validation 3D FEA are also added to the graphs for validation purposes; these datapoints were not included in the training procedure of the site-specific model 1D model. The GPR models can predict with high accuracy the depth variation of the unseen knot point parameters. The GPR models that incorporate the $[z_{Li}/L]$ training inputs can estimate precisely the depth that a 'jump' occurs in the soil reaction parameters magnitude which is found at the layer boundaries indicating the transition from a clay to sand layer and vice versa; the data also showed that the biggest 'jump' is found at the first layer boundary z_{L1} which is located in the upper half of the pile. Hence, the datapoints are noisier at normalised depths of $0.16 \leq z/L \leq 0.24$ which formed the training dataset. This causes a higher 95% confidence interval uncertainty around that area. Figure 6.31 to Figure 6.33 compare the load-displacement response at ground level predicted by the 1D model and the PLAXIS 3D results for the pile's small displacement and ultimate response. The 1D model calculates accurately the pile's lateral behaviour.

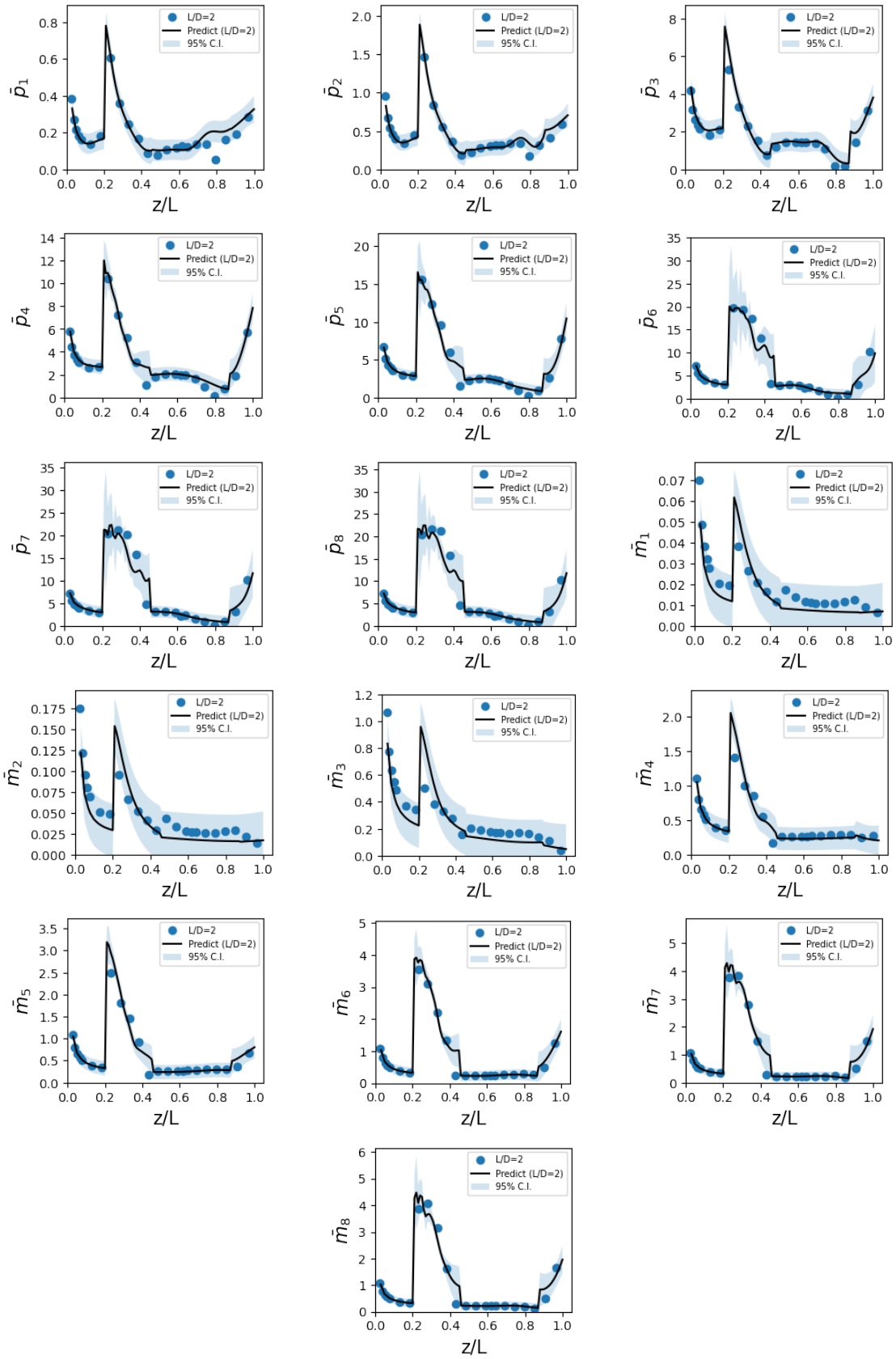


Figure 6.30: Example GPR ML outputs for the soil reactions knot points for the DL1 pile ($L/D = 2$, $h/D = 5$) embedded in DLS1 layered soil. The depth variation of the mean GPR prediction for the knot points is indicated by the black line. The 95% confidence interval is shown (shaded bounds). The blue markers are datapoints that have not been included in the training procedure with unseen (L/D, z_{L1}/L, z_{L2}/L, z_{L3}/L).

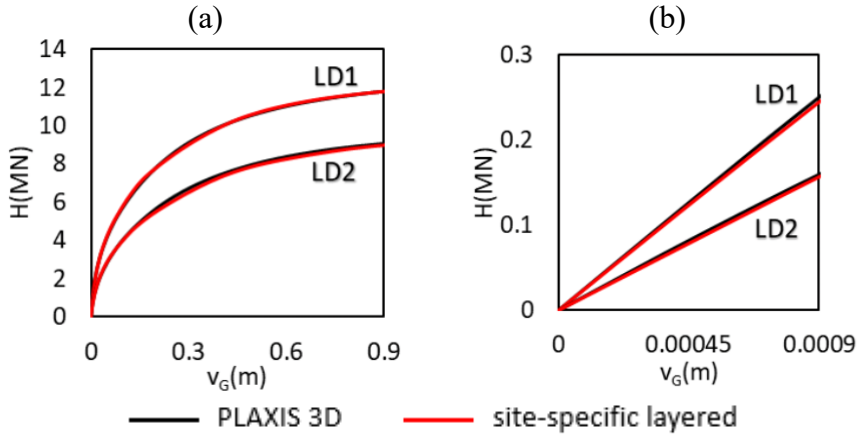


Figure 6.31: Piles embedded in design layered soil DLS1. Comparison between the load-displacement responses at ground level for: (a) ultimate response, (b) small displacement response from site-specific layered soil model accounting for spatial variations and the 3D FEA.

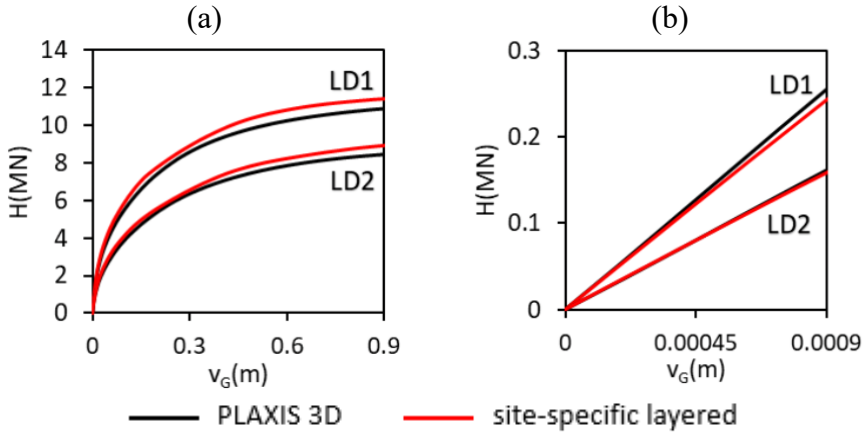


Figure 6.32: Piles embedded in design layered soil DLS2. Comparison between the load-displacement responses at ground level for: (a) ultimate response, (b) small displacement response from the site-specific layered soil model accounting for spatial variations and the 3D FEA.

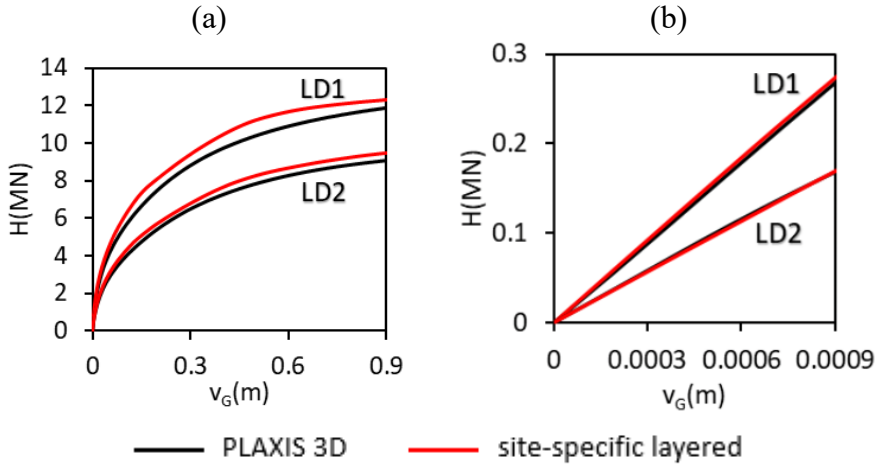


Figure 6.33: Piles embedded in design layered soil DLS3. Comparison between the load-displacement responses at ground level for: (a) ultimate response, (b) small displacement response from the site-specific layered soil model accounting for spatial variations and the 3D FEA.

6.7.3 Key observations

The developed site-specific layered soil models can account for possible spatial variations of the layer thickness across the wind farm site. The method was tested for a configuration involving the combination of Bothkennar clay and Dunkirk sand with $D_R = 90\%$ which was found to be a highly problematic combination for the ILM. It is readily extended for all possible different layered configurations irrespective the number of soil layers and has the advantage that relatively few calibration analyses are required to form the 1D model.

6.8 Summary

This chapter explores different methods for implementing the data-driven method for piles embedded in layered soils. Key observations are:

- **Independent Layer method**: The method is relatively straightforward to use since it employs 1D models calibrated for homogenous soils. It provides relatively accurate results for many layered soil configurations such as: (i) configurations involving layers of the same soil type, (ii) multi-layered soil profiles that employ many thin layers, (iii) soil configurations with loose sands ($D_R \leq 50\%$) and highly overconsolidated stiff clays. However, ILM produces inaccurate results for many layered soil profiles involving both clay and sand layers (e.g. where a soft clay layer is adjacent to a dense sand layer).
- **Rule-based method for layered soils:** This method is an extension to the ILM. Simplified rules are developed in which the soil strength and stiffness in each layer is modified – according to certain rules – to incorporate the influence of interaction between neighbouring layers. The rules are determined by inspecting the soil reaction curves computed by PLAXIS 3D for layered soils and comparing them to the soil reaction curves extracted by the reference homogeneous soils. The rules increase the

accuracy of the ILM. However, this method is not straightforward to implement for layered soils that employ both clay and sand soil layers. Hence, it is difficult to establish rules that are valid for all possible layered soil configurations.

- **General layered soil models**: This method is used to develop 1D models from calibration analyses performed directly for layered soils; hence, the interactions between adjacent layers are introduced directly into the 1D model. The method provides accurate results for the trained layered soil configurations irrespectively to the depth of the layer boundaries over the pile calibration space. However, the method is not scalable. A large number of calibration analyses is needed to create a general layered soil model that will cover all possible layered soil profiles which limits the application of the approach for practical design applications.
- **Site-specific layered soil models**: This method is used to develop 1D models from calibration analyses performed for layered soils; hence, the interactions between adjacent layers are introduced directly into the 1D model. The 1D models that are developed are not general; site-specific calibration is needed for each layered soil configuration. This method can be used for all soil configurations irrespectively of the layers soil type (clay, sand) and the number of soil layers. This method can account for spatial variations of the layer thickness across a complete wind farm site. An advantage of this method is that it employs relatively few calibration analyses to form the 1D models.

Overall, to establish a design approach to generate data-driven 1D design models for layered soil profiles which are both practical and calculate accurately the pile's lateral behaviour, the site-specific layered soil models seem the most promising option. This method is readily extended: (i) to account for spatial variations of the soil strength and stiffness properties and (ii) for all different layered soil configurations.

Chapter 7

7 Layered Soil Modelling Approach for Practical Design

7.1 Introduction

In reality OWF sites typically consist of layered soil configurations of layers with different geological origins (e.g. clay, sand). Possible site-wide variations of the layer thickness and soil properties occur. This chapter presents a study on developing data-driven layered soil models. The site-specific layered soil modelling approach accounting for spatial variations described in Section 6.7.2 is employed to calibrate the data-driven layered soil models. The 1D model developed in this chapter accounts both for spatial variation of the layer thickness and soil properties across the site. Initially, GPR models are trained using the soil reaction knot points extracted from a series of calibration FEA; the sand normalisation framework presented in Section 4.4 is employed which is appropriate for layered soils. The GPR model's predictions are connected to the OxPile software to predict the monotonic response of the laterally loaded monopiles. The aim of this chapter is to present the capabilities of this new design approach for the practical design of monopile foundations within an OWF site.

Technical advice on the choice of the soil stratigraphy dataset, the calibration piles geometry and the layer thickness variations across the wind farm site were provided by the industry partner (Ørsted). The soil properties incorporated in the model, which were used to

characterise the clay and sand layers of the site, for anonymity purposes were modified based on the representative offshore sites developed during the PISA project.

7.2 Representative offshore site across the offshore wind farm

The soil stratigraphy dataset indicated that the OWF site mainly consisted of two soil configurations; realistic site-wide variations of the layer thickness and soil properties are accounted for. Figure 7.1 demonstrates the two main layered soil configurations found across the site. *Configuration I* and *Configuration II* are found at approximately 75% of the site. The soil properties employed for each layer, are based on the PISA representative offshore sites. It should be noted that for the anomalous cases that do not employ the soil configurations presented in Figure 7.1 (e.g. 25% of the site) different 1D models should be developed consistent with the approaches presented in Chapter 6. For example, if the layered soil configuration employs sand layers of different relative density, the ILM method can be used in combination with the GDSM presented in Chapter 5.

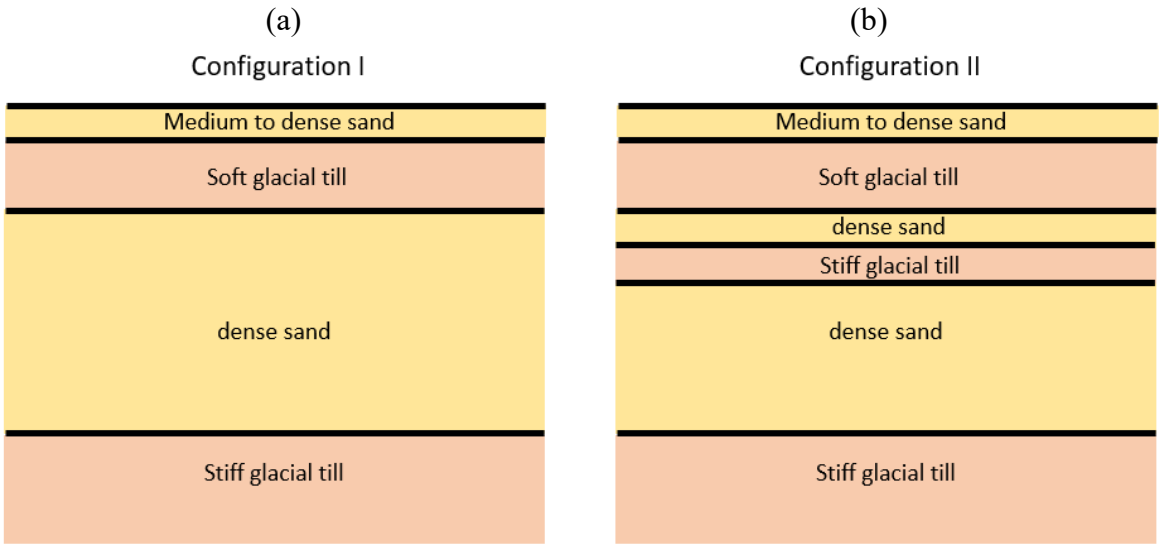


Figure 7.1: Layered soil configurations within the offshore wind farm site: (a) four-layer case; (b) six-layer case. The soil strength and stiffness and the layer thickness vary across the site.

Configuration I is the baseline layered soil configuration (it is found at approximately 50% of the site); it employs an embedded soft glacial till layer within a medium to dense

sand matrix overlying a stiff overconsolidated glacial clay till layer. *Configuration II* has the same geological origin as *Configuration I*. Due to past geological processes a thin stiff glacial till layer was embedded within the bottom dense sand layer. The soil stratigraphy dataset indicates that there is a spatial variation of the clay and sand soil properties as well as the layer thickness across the site.

Configuration I

Configuration I incorporates four independent soil layers where:

- The top layer comprises a medium to dense sand layer based on the PISA representative offshore homogenous sand site (Burd et al., 2020a) with a relative density D_R in the range 45% to 75%. The layer thickness z_L varies across the OWF site from 0.5m to 4.8m.

- The second soil layer comprises a soft NC to slightly OC representative offshore glacial clay till ($d_{ice} = 25m$); the soil profile properties (σ'_v, s_u, K_o, G_o) of the glacial till are determined by the procedure shown in Section 5.2.1. The layer thickness z_L varies across the OWF site from 3.8m to 12.6m.

- The third layer comprises a dense sand layer based on the PISA representative offshore homogenous sand site (Burd et al., 2020a) with a relative density D_R in the range 80% to 100%. The layer thickness z_L varies across the OWF site from 19m to 30.6m.

- The base layer is the PISA representative offshore glacial clay till soil (Byrne et al., 2020); no soil variability has been adopted since this layer is relatively deep. For brevity in Figure 7.2 and Table 7.7 and Table 7.9 the PISA representative offshore glacial clay till layer is referred to as 'PISA offshore clay till'.

Configuration II

Configuration II is based on *Configuration I*; an additional thin stiff glacial till layer is embedded within the dense sand matrix of the third layer where:

- The thin embedded layer is a highly OC representative offshore glacial clay till; the variation of the soil profile properties (σ'_v, s_u, K_o, G_o) correspond to $75m \leq d_{ice} \leq 100m$ determined using the procedure described in Section 5.2.1. The layer thickness z_L varies across the OWF site from 2.5m to 5m and the thin embedded layer is located between the depths of 6m to 15.2m below the ground level across the site.

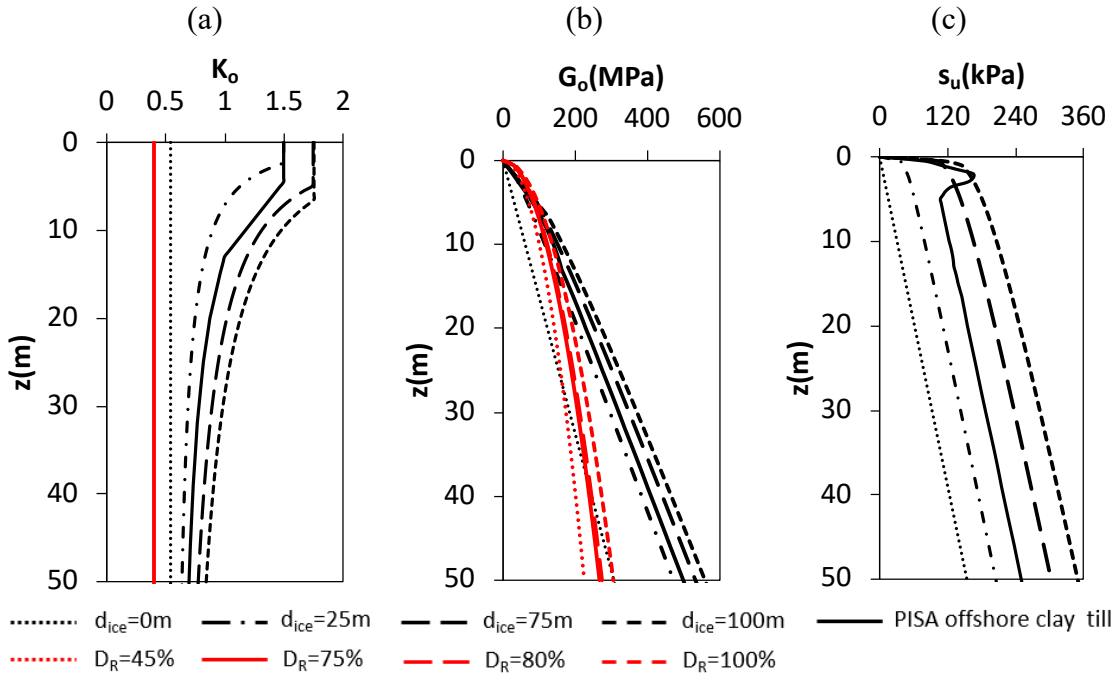


Figure 7.2: Profiles of: (a) K_o, (b) G_o, and (c) s_u, employed in the reference homogeneous soil models (where clay profiles indicated in black and sand profiles are indicated in red).

Profiles of K_0, G_0 and s_u for the reference homogeneous soil models are illustrated in Figure 7.2. The sand layers properties are based on the PISA representative offshore homogenous sand site presented in Burd et al., (2020a). The clay soil properties are based on the PISA representative offshore glacial clay till site (Byrne et al., 2020b); the stress history of each clay layer is determined using the procedure presented in Section 5.2.1.

7.3 Pile geometry calibration space

The offshore pile geometries adopted for developing the current model are presented in Table 7.1. The calibration piles geometry was selected in communication with the industry partner (Ørsted). The piles adopt geometries that are representative of typical offshore pile configurations with a single load eccentricity. The pile wall thickness is 125mm; it is selected to avoid unrealistically high bending stresses developing in the pile wall and it does not influence the soil reaction curves.

Table 7.1: Pile geometry calibration space.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D
CP1	10	32	125	40	3.2
CP2	10	20	125	40	2
CP3	8.5	27.2	125	40	3.2
CP4	8.5	17	125	40	2

7.4 3D Finite element modelling procedure

3D FEA were conducted using PLAXIS 3D to extract the numerical soil reaction curves that are used to generate the data-driven layered soil models. Clay layers are modelled as undrained; sand layers are drained. The NGI-ADP constitutive was employed to model the glacial till layers. The sand layers are modelled with the HS-Small model. The use of both constitutive models to develop calibrations for the representative offshore glacial clay till and representative offshore homogenous sand layers are described in Chapter 3. The effective friction angle ϕ' employed as an input parameter for HS-Small was determined for each sand layer by using the linear equation which correlates D_R with ϕ' presented in Figure 3.16(a) in Section 3.4.3.2.

7.5 Calibration space of the data-driven layered soil models

The selection of a well-defined calibration space is vital for the accuracy of the data-driven layered soil model. The calibration space needs to incorporate the spatial variation of the soil properties and thickness of each soil layer for the selected pile calibration space. The calibration layered soil profile distributions were developed following the procedures in Burd et al. (2020b); σ'_{vi} is determined from γ' from the reference soil profiles; σ'_{hi} and p' (mean effective stress) are calculated from the K_o of the reference soil profiles; G_o and s_u are considered to be functions of p' and are estimated on the basis of the local p' of each layered soil profile based on the G_o and s_u distributions of the reference soil profiles.

Configuration I

Configuration I incorporates three-layer boundaries within the piles' embedment length; the thickness of each layer varies across the site. Two different approaches for selecting the total number of calibration analyses were trialled (Table 7.2).

- **Model V1**: Two variations of the possible thickness of each layer were considered (minimum or maximum) as determined from the soil stratigraphy dataset.
- **Model V2**: Three values of the possible thickness of each layer were considered minimum, average or maximum) as determined from the soil stratigraphy dataset. The additional FEA were added to address the high variability of the layer thickness that was found in the upper two layers across the OWF site.

Table 7.2: Layer thickness values adopted for the calibration analyses.

Layer Thickness	Model V1	Model V2
Min (z_{Li})	z_{L1}=0.5m, z_{L2}=4.5m, z_{L3}=15m	z_{L1}=0.5m, z_{L2}=4.5m, z_{L3}=15m
Ave (z_{Li})	-	z_{L1}=2.1m, z_{L2}=8.6m, z_{L3}=22.8m
Max (z_{Li})	z_{L1}=3.7m, z_{L2}=12.7m, z_{L3}=30.6m	z_{L1}=3.7m, z_{L2}=12.7m, z_{L3}=30.6m

To incorporate the spatial variation of soil strength and stiffness in the 1D model, two different sets of calibration analyses were employed. The strength and stiffness profiles of each layer can be obtained from either the minimum ($D_R = 45\%, d_{ice} = 0m, D_R = 80\%$) or maximum ($D_R = 75\%, d_{ice} = 25m, D_R = 100\%$) values. The total number of calibration FEA required to train the site-specific layered soil models (Model V1, Model V2) is presented in Table 7.3 and detailed in Appendix A (refer to Tables A.7 and A.8).

Table 7.3: Total number of calibration FEA conducted for the soil *Configuration I*.

Calibration FEA	**Model V1**	**Model V2**
Pile geometry	4	4
Layer thickness	2	3
Soil properties variability	2	2
Total FEA	4×2×2=16	4×3×2=24

Configuration II

Configuration II includes five-layer boundaries within the piles' embedment length. The soil in this configuration has similar geological origins as *Configuration I*, therefore soil reaction curves extracted from the *Configuration I* calibration FEA can be used for the training of an extended site-specific layered soil model. Eight additional calibration analyses were conducted to account for the additional thin emdedded stiff glacial till layer. Two soil layers incorporating the minimum ($d_{ice} = 75m$) and maximum ($d_{ice} = 100m$) cases were considered. The additional eight calibration analyses are conducted for a constant layer depth; the layer thickness variation is introduced in the extended 1D model from the previous 24 FEA. In the eight additional calibration FEA, the new embedded OC glacial clay till layer is located between 6.2m and 15.2m. These depths are indicated from the soil stratigraphy dataset and form the minimum and maximum values of the upper- and lower-layer

boundaries of this additional stiff glacial till layer. The FEA performed to develop this model are detailed in Appendix A (refer to Tables A.8, A.9).

7.6 GPR models training procedure

As outlined in Section 6.7.2, the number of inputs that the GPR model employs depends on the particular layered soil configuration being considered. *Configuration I* and *II* employ for the training of the distributed soil reaction parameters the $[L/D]$, $[z/L]$, $[G_o/\sigma'_v]$, $[s_u/\sigma'_v]$, $[D_R]$ training inputs. The GPR models trained for the *Configuration I* employs three additional inputs $[z_{L1}/L]$, $[z_{L2}/L]$, $[z_{L3}/L]$ whereas the GPR model trained for the *Configuration II* uses five additional inputs $[z_{L1}/L]$, $[z_{L2}/L]$, $[z_{L3}/L]$, $[z_{L4}/L]$, $[z_{L5}/L]$. The base soil reaction parameters employ five training inputs for both layered soil configurations: the $[L/D]$, $[h/D]$, $[G_o/\sigma'_v]$, $[s_u/\sigma'_v]$, $[D_R]$ values.

7.6.1 Training procedure

Consistent with the approach followed in the previous chapters, a GPR model is trained for each of the eight knot points for each soil reaction component ($\bar{p}$, $\overline{m}$, $\overline{H}_B$, $\overline{M}_B$). The Matérn ($\nu = 5/2$) covariance function with ARD is employed. The white noise kernel is added to the Matérn kernel to account for possible noise in the dataset. An 80:20 training-test set split ratio was adopted. For each trained GPR model the MSE and R^2 trained and test scores were used to check whether the GPR models can accurately represent the knot point parameters. Table 7.4 and Table 7.5 show the MSE and R^2 scores for the trained GPR models corresponding to $\bar{p}$, $\overline{m}$ knot point parameters associated with the *Configuration I* and *Configuration II* respectively. The similarity in the achieved training and test metric values indicate that the GPR models have good generalisation properties – the ML models can be used for unseen inputs inside the calibration space. The high train and test R^2 scores indicate that the GPR models can accurately capture the complexity of the data. The base knot point

parameters $\overline{H}_b$, $\overline{M}_b$ employ a small dataset for each GPR ML model; the MSE and R^2 values for $\overline{H}_b$, $\overline{M}_b$ are not reliable and are not presented in this thesis.

Table 7.4: Regression performance metrics for the GPR models as applied to the $\bar{p}$, $\overline{m}$ knot point soil reaction parameters associated with the layered soil *Configuration I*.

Knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.00	0.02	0.98	0.85	$\overline{m}_1$	0.00	0.00	1.00	0.99
$\bar{p}_2$	0.01	0.09	0.99	0.86	$\overline{m}_2$	0.00	0.00	1.00	0.91
$\bar{p}_3$	0.00	0.53	1.00	0.89	$\overline{m}_3$	0.00	0.02	1.00	0.90
$\bar{p}_4$	0.04	0.94	0.99	0.91	$\overline{m}_4$	0.00	0.04	1.00	0.91
$\bar{p}_5$	0.02	0.94	1.00	0.95	$\overline{m}_5$	0.00	0.07	1.00	0.90
$\bar{p}_6$	0.27	1.8	0.99	0.96	$\overline{m}_6$	0.00	0.07	1.00	0.94
$\bar{p}_7$	0.65	2.65	0.99	0.96	$\overline{m}_7$	0.00	0.06	1.00	0.97
$\bar{p}_8$	0.65	2.82	0.99	0.96	$\overline{m}_8$	0.20	0.33	0.90	0.89

Table 7.5: Regression performance metrics for the GPR models as applied to the $\bar{p}$, $\overline{m}$ knot point soil reaction parameters associated with the layered soil *Configuration II*.

Knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test	Knot point parameters	MSE Train	MSE Test	R^2 Train	R^2 Test
$\bar{p}_1$	0.00	0.00	0.98	0.98	$\overline{m}_1$	0.00	0.00	1.00	0.99
$\bar{p}_2$	0.00	0.01	0.99	0.99	$\overline{m}_2$	0.00	0.00	1.00	1.00
$\bar{p}_3$	0.00	0.03	1.00	0.99	$\overline{m}_3$	0.00	0.00	1.00	0.99
$\bar{p}_4$	0.03	0.18	1.00	0.98	$\overline{m}_4$	0.00	0.00	1.00	0.99
$\bar{p}_5$	0.02	0.27	1.00	0.98	$\overline{m}_5$	0.00	0.00	1.00	0.99
$\bar{p}_6$	0.23	1.12	0.99	0.97	$\overline{m}_6$	0.00	0.01	1.00	0.99
$\bar{p}_7$	0.56	1.77	0.99	0.96	$\overline{m}_7$	0.00	0.05	1.00	0.97
$\bar{p}_8$	0.57	2.18	0.99	0.96	$\overline{m}_8$	0.01	0.11	1.00	0.94

7.6.2 GPR model predictions

The predictions – with the pile geometry, soil profile properties, layered configuration as inputs – provide: (i) the depth variation of the knot point parameters of the distributed lateral load $\bar{p}$ and normalised distributed moment $\overline{m}$; (ii) predictions of the knot point parameters for the base horizontal force $\overline{H}_B$ and base moment $\overline{M}_B$.

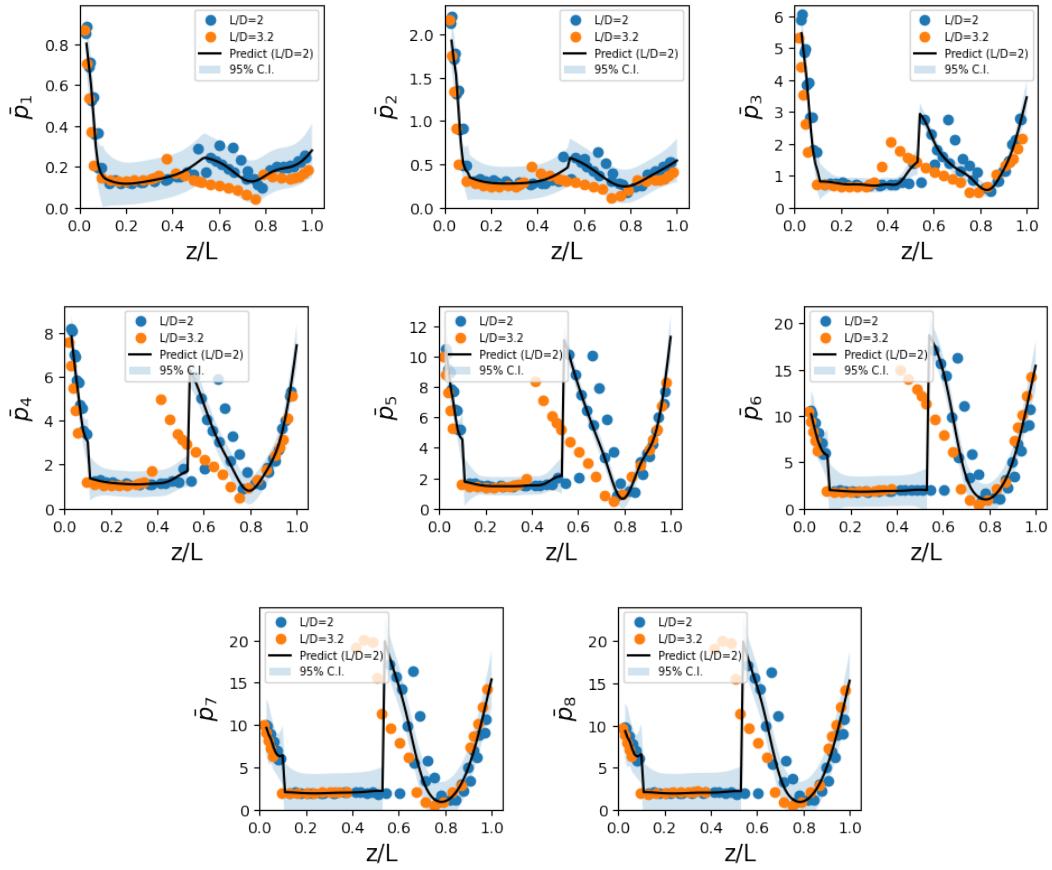


Figure 7.3: Example GPR ML outputs for the distributed load $\bar{p}$ knot points for the C1 pile ($L/D = 2$) embedded in a layered profile which employs the minimum possible strength and stiffness values and employ the average values of layer thickness z_{Li}. The depth variation of the mean GPR prediction for the knot points is indicated by the black line. The 95% confidence interval is shown (shaded bounds). The blue markers are the knot points extracted from the calibration analysis.

Figure 7.3 presents depth variation predictions for the $\bar{p}$ soil reaction parameters determined from GPR models for a trained calibration analysis. These predictions describe the CP1 pile (Table 7.1) embedded in a layered soil where its soil properties employ the minimum strength and stiffness values and the layer thickness employs the average values of z_{Li}. The GPR models can accurately predict the complex depth variation of the distributed lateral reaction data along the monopile capturing the soil reaction curves 'jumps' between the neighbouring clay and sand layers. The GPR model also provides the 95% confidence interval, serving as a measure of uncertainty for the predictions of knot point parameters. In this example, the uncertainty is relatively small, as the results are predicted for a trained layered calibration analysis.

7.7 Application of the data-driven layered soil models to design cases

To demonstrate the predictive capability of the new site-specific layered soil models accounting for spatial variations, various design examples have been examined. These examples employ the *Configurations I, II* and are concerned with monopile geometries within or outside the trained L/D calibration space embedded in layered soils with unseen values of layer thickness z_{Li}/L and unseen soil properties G_o/σ'_v, s_u/σ'_v, D_R. The design pile geometries employed for the design examples are listed in Table 7.6.

Table 7.6: Design pile geometries. L is the embedded length; D is the pile diameter; h is the load eccentricity; t_w is the pile wall thickness.

Pile	$\boldsymbol{L(m)}$	$\boldsymbol{D(m)}$	$\boldsymbol{h(m)}$	$\boldsymbol{t_w(mm)}$	$\boldsymbol{L/D}$
DP1	25	10	40	125	2.5
DP2	28.5	9.5	40	125	3
DP3	30	10	40	125	3
DP4	36	9	40	125	4

Configuration I

Six design cases were examined for *Configuration I*. The variations of layer thickness and the characterisation of the soil properties of each layer are presented in Table 7.7 together with the design pile geometry codes. For the design case D different values of D_R (different strength and stiffness) within a sand layer have been considered; this scenario is commonly found at offshore sites. For this design case it is assumed that a sand layer with $D_R = 45\%$ is found from 0 to 1m depth and a layer with $D_R = 60\%$ comprises the rest of the top layer. A sand layer with $D_R = 85\%$ is located from 14 to 18m depth and below that D_R is equal to $D_R = 95\%$.

Table 7.7: Selected pile geometry, variations of layer thickness and characterisation of the soil profile properties of each layer for each design case examined for *Configuration I*.

Design Cases	Pile	z_{Li}(m)	Layer 1	Layer 2	Layer 3	Layer 4
A	DP2	z_{L1}=3 z_{L2}=7 z_{L3}=28	Dunkirk sand D_R=55%	offshore till d_{ice}=10m	Dunkirk sand D_R=95%	PISA offshore clay till
B	DP1	z_{L1}=1.8 z_{L2}=9.7 z_{L3}=17.5	Dunkirk sand D_R=50%	offshore till d_{ice}=15m	Dunkirk sand D_R=90%	PISA offshore clay till
C	DP1	z_{L1}=2.5 z_{L2}=11.5 z_{L3}=22	Dunkirk sand D_R=45, 60%	offshore till d_{ice}=20m	Dunkirk sand D_R=85, 95%	PISA offshore clay till
D	DP1	z_{L1}=3.5 z_{L2}=4.8 z_{L3}=31.5	Dunkirk sand D_R=70%	offshore till d_{ice}=10m	Dunkirk sand D_R=95%	PISA offshore clay till
E	DP2	z_{L1}=0.8 z_{L2}=8.2 z_{L3}=25	Dunkirk sand D_R=60%	offshore till d_{ice}=20m	Dunkirk sand D_R=95%	PISA offshore clay till
F	DP3	z_{L1}=1 z_{L2}=5 z_{L3}=21	Dunkirk sand D_R=50%	offshore till d_{ice}=15m	Dunkirk sand D_R=90%	PISA offshore clay till

The computed H-v_G response for the design cases using the data-driven layered soil models (V1, V2) and, separately, with the corresponding validation PLAXIS 3D FE calculations are shown in Figure 7.4. The 1D model predictions computed by using the ILM employing the data-driven general Dunkirk sand model GDSM V2 presented in Chapter 5 and data-driven general Cowden clay model GCCM presented in Chapter 5 are also provided. The V2 model significantly outperforms the V1 model for most of the design cases; this can be explained by the high variability of the thickness of the upper three layers across the wind farm site. Overall, V2 model provides a close match to the PLAXIS 3D data set for all the examined design cases for both the pile's ultimate and small displacement response. The close fit that is obtained between the 1D and 3D FE results supports the use of the calibrated data-driven layered soil model across the wind farm site to conduct design calculations. Consistent with the results in Section 6.3, the ILM does not provide reliable

predictions of the pile's ultimate response for piles with small L/D embedded in layered soil profiles which incorporate neighbouring soil layers with different strength characteristics.

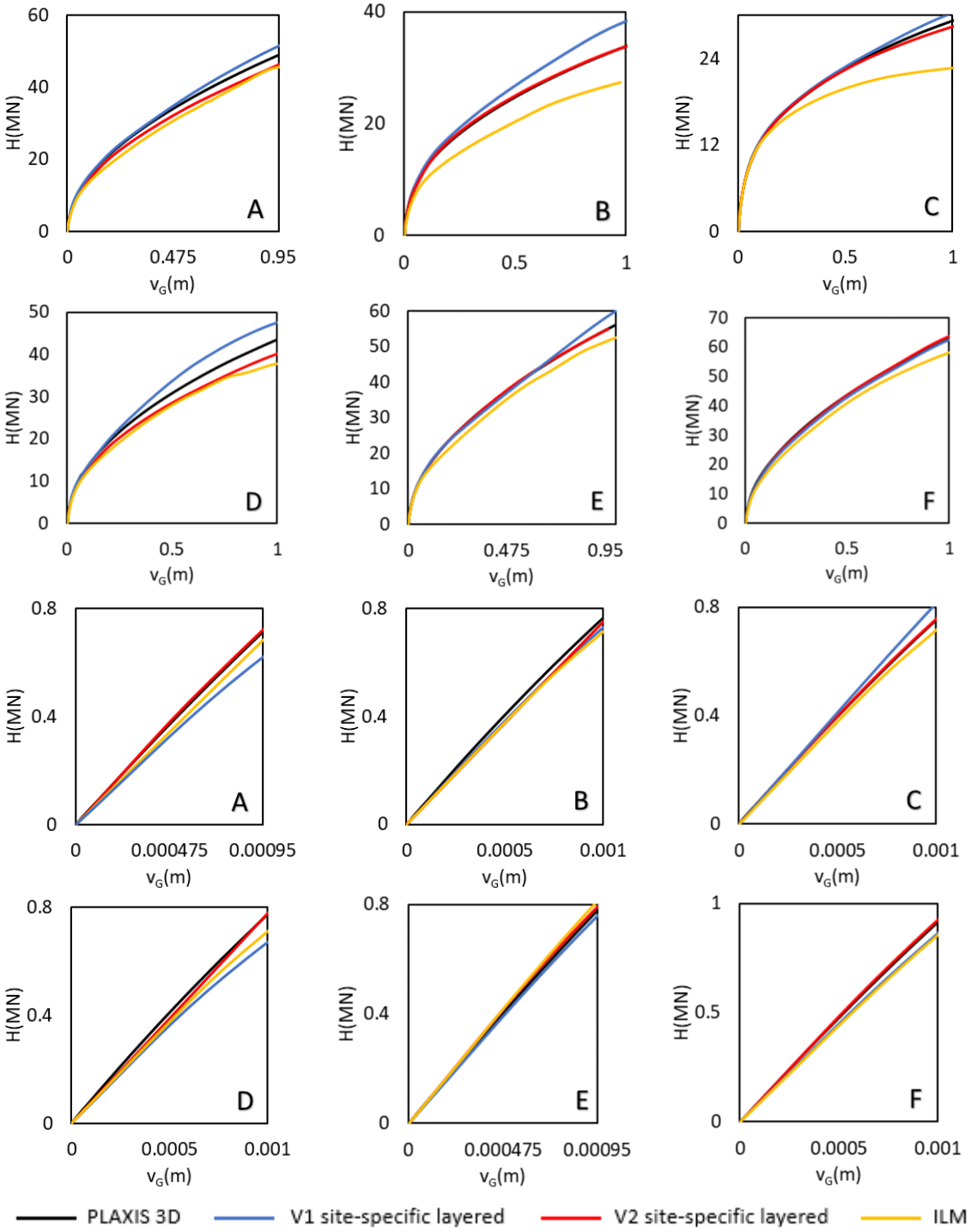


Figure 7.4 Comparison between computed responses determined from the data-driven layered soil model and equivalent PLAXIS 3D finite element analyses for the design cases in Table 7.7 for: (a) the ultimate response (rows 1, 2), (b) the small displacement response (rows 3, 4). The solid line represents the 3D finite element results and the dashed/dotted lines the data-driven layered soil (V1, V2) models and ILM results.

Configuration II

Three design cases were considered for *Configuration II*. The values of layer thickness together with the design pile reference are presented in Table 7.8. The characterisation of the soil properties of each layer are specified in Table 7.9.

Table 7.8: Values of layer thickness for each design case for *Configuration II*.

Design Cases	Pile	z_{L1}(m)	z_{L2}(m)	z_{L3}(m)	z_{L4}(m)	z_{L5}(m)
G	DP1	0.7	6.8	1.2	3	28.3
H	DP1	1	5	3	4	22
I	DP2	1.2	5.8	2	2	25

Table 7.9: Characterisation of soil properties for each design case for *Configuration II*.

Design Cases	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6
G	Dunkirk sand D_R=50%	offshore till d_{ice}=20m	Dunkirk sand D_R=90%	offshore till d_{ice}=85m	Dunkirk sand D_R=90%	PISA offshore clay till
H	Dunkirk sand D_R=55%	offshore till d_{ice}=10m	Dunkirk sand D_R=90%	offshore till d_{ice}=90m	Dunkirk sand D_R=90%	PISA offshore clay till
I	Dunkirk sand D_R=60%	offshore till d_{ice}=15m	Dunkirk sand D_R=85%	offshore till d_{ice}=90m	Dunkirk sand D_R=95%	PISA offshore clay till

The data-driven layered soil model results presented for *Configuration II* are concerned with the V2 model (which outperformed the V1 model for *Configuration I*). The computed H-v_G response for the three design cases using the 1D model and, separately, with the corresponding validation PLAXIS 3D calculations, are shown in Figure 7.5. The 1D model achieves a close match to the PLAXIS 3D data. Once a data-driven layered soil model has been calibrated for a specific layered soil configuration, it can be readily enhanced to account for additional thin embedded layers with a small number of supplementary calibration FEA.

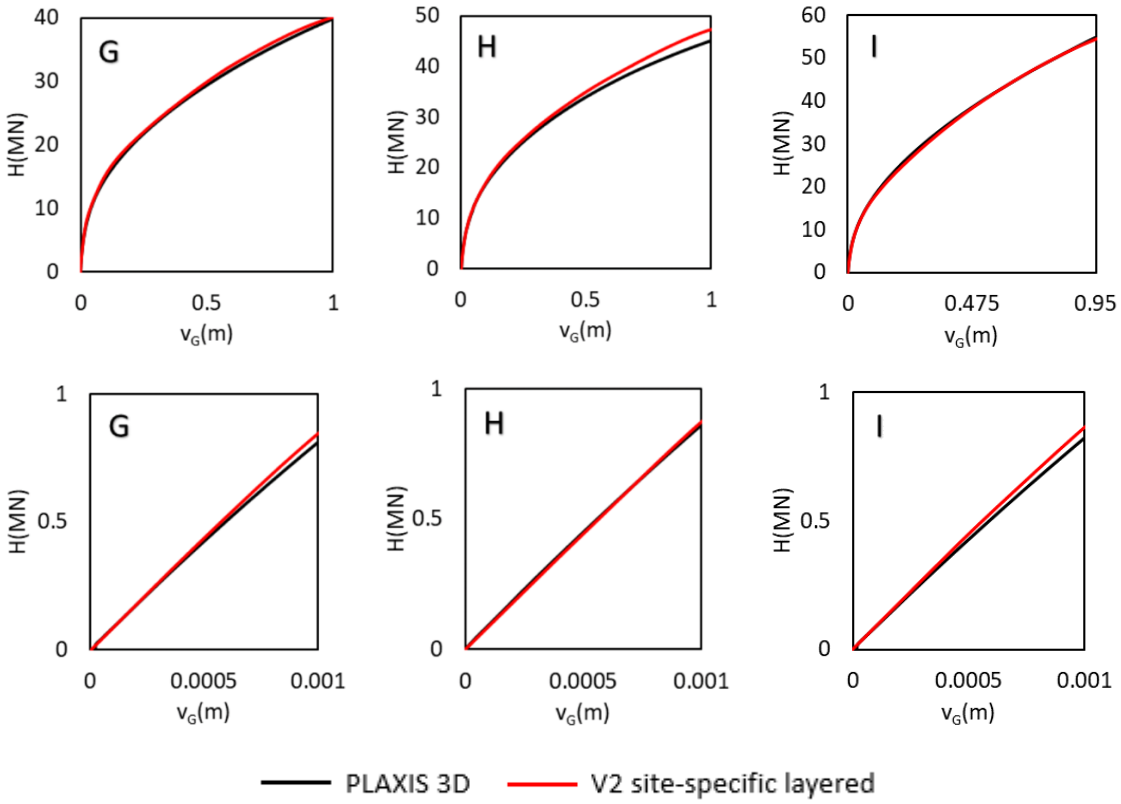


Figure 7.5: Comparison between computed responses determined from the data-driven layered soil model and equivalent PLAXIS 3D finite element analyses for the design cases G, H, I for: (a) the ultimate response (row 1), (b) the small displacement response (row 2). The solid line represents the 3D FE results and the dashed line the data-driven layered soil V2 model results.

L/D pile geometries outside the trained calibration space

During the design procedure of the OWF foundations, it is advantageous to develop a 1D model that can provide reliable predictions for pile geometries falling outside the L/D calibration space. To assess the developed data-driven layered soil model for this scenario, the V2 model (which is more accurate than V1) was tested for the D4 pile embedded in the A, B and C design layered soil profiles (Table 7.7). The design pile employs a L/D value slightly higher than the calibration piles' geometry. Figure 7.6 illustrates the layered soil 1D model predictions compared with the PLAXIS 3D computed results for the pile's H-v_G response. The results indicate that the data-driven layered soil models provide reliable predictions for piles with L/D values that do not depart significantly from the trained L/D

calibration space. However, the 1D models should not be used for piles with L/D values departing significantly from the trained calibration space.

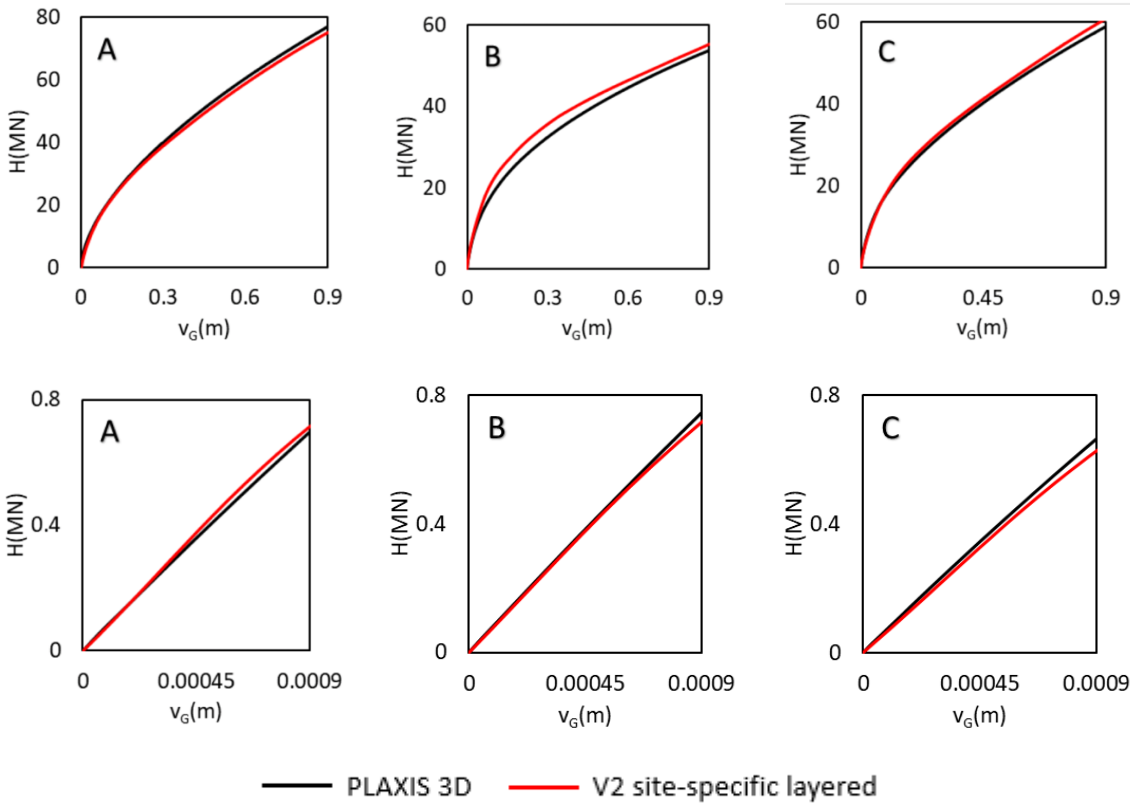


Figure 7.6: Comparison between computed responses determined from the data-driven layered soil model and equivalent PLAXIS 3D finite element analyses for the design cases A, B, C for: (a) the ultimate response (rows 1), (b) the small displacement response (rows 2). The solid line represents the 3D finite element results and the dashed line the V2 data-driven layered soil model results.

7.8 Summary

Two site-specific layered soil models are generated accounting for spatial variations of the thickness and soil properties of each layer and soil properties across the site. These 1D models describe a design scenario that is likely to be encountered in an offshore wind farm site. The 1D models are calibrated using as a database the soil reaction curves that were extracted by analyses implemented for piles embedded in layered soils.

A baseline site-specific layered soil model was calibrated using 24 FEA to account for possible spatial variations of the thickness and soil properties of each layer over a realistic

pile geometry calibration space (accounting for four different piles and a single load eccentricity). The site-specific layered soil model was tested for several realistic design scenarios involving unseen pile geometries embedded in layered soils with unseen values of layer thickness z_{Li}/L and unseen soil properties G_o/σ'_v, s_u/σ'_v, D_R. The site-specific layered soil model captures in good accuracy the pile's monotonic lateral behaviour for both the small displacement and ultimate response. A site-specific layered soil model can be easily extended with a small number of calibration analyses to account for layered soil configurations which include an additional embedded thin layer. Furthermore, a site-specific layered soil model was shown to provide reliable predictions for pile geometries falling not significantly outside the L/D calibration space.

The study suggests that for many practical layered soil configurations involving clay with sand soil layers the new data-driven method can be used with high confidence as a surrogate model. The developed 1D models include inherently the interactions between adjacent soil layers. The method is readily extended for layered configurations consisting of fewer or more layers than the soil configurations examined in this study. Also, the method does not have any restrictions in the context of the soil strength and stiffness employed for each soil layer (e.g. a soft clay layer can be adjacent to a dense sand layer).

Chapter 8

8 Data-Driven Method for Laterally Loaded Piles in Chalk

8.1 Introduction

Low to medium density chalk material is widespread across northern Europe, under the English Channel as well as the North and Baltic Sea (Mortimore, 2012). Many offshore wind farms are or will be developed at chalk sites. In chalk, engineers frequently encounter difficulties in designing the pile foundations due to the fact that the installation procedure de-structures the chalk around the pile. There are no established design methods for analysing the lateral behaviour of driven piles in chalk.

This chapter presents a new approach to develop data-driven 1D design models to determine the monotonic lateral behaviour of monopiles driven in chalk. A chalk model has been calibrated for the Kent St. Nicholas site (Vinck et al., 2022) for a parameter space that is relevant to the ALPACA, ALPACA+ (Buckley et al., 2020b) and Wind Support (Ciavaglia et al., 2017b) field tests. The proposed chalk 1D model utilises a GPR model that provides the depth variation predictions of the knot point parameters of a pile embedded in a homogeneous chalk site. The GPR predictions are employed with the PISA 1D FE framework OxPile software. The chalk models developed using the data-driven method can

incorporate the possibility of different de-structured zone thicknesses caused by the pile installation.

8.2 3D finite element analyses

3D finite element 'validation' models – employing PLAXIS 3D – have been developed for the three ALPACA test piles and the Wind Support test pile. The purpose of these validation models is to develop an analysis procedure that can subsequently be employed – with some confidence – to calibrate a data-driven 1D design model.

Pile driving activities cause the chalk near to the pile to become de-structured (Muir Wood et al., 2015; Buckley et al., 2018; Jardine et al., 2018). As a consequence, a remoulded chalk annulus (putty zone) is developed around the pile. Additionally, the installation process causes a volume of the chalk to fracture (fractured chalk zone) beyond the putty zone. The thickness of each zone is considered to be dependent on the pile wall thickness t_w (Muir Wood et al., 2015; Buckley et al., 2018).

The 3D FEA implemented in this chapter adopt the three distinct chalk zones illustrated in Figure 8.1; different strength and stiffness properties are assigned to each zone. A similar approach was adopted by Pedone et al. (2023) for the ALPHA project. On the basis of data in Buckley et al. (2018), the thickness of the remoulded chalk zone is considered to be in the region $t_1/t_w = 0.6 - 1.6$ and the thickness of the fractured chalk zone is in the region $t_2/t_w = 1.65 - 5.9$ where t_w is the pile wall thickness. For the current analyses, the internal soil plug consists entirely of remoulded chalk ('remoulded plug' model); this conservative assumption is considered realistic for the small diameter ALPACA and Wind Support test piles (Ciavaglia et al., 2017a; Pedone et al., 2023). The field test pile configurations adopted for the validation FEA are specified in Table 8.1.

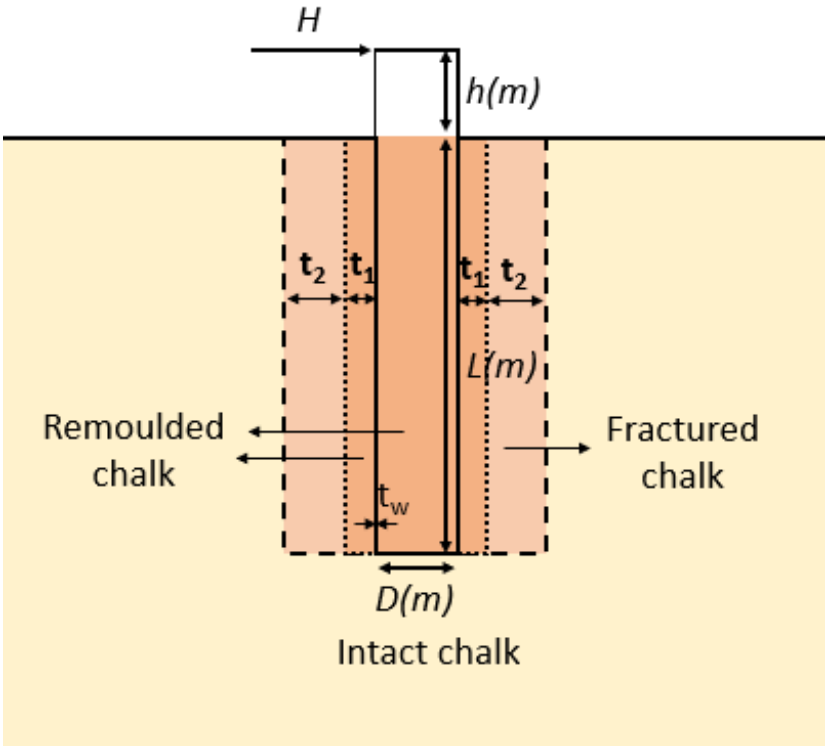


Figure 8.1: Assumed configuration of the remoulded, fractured and intact chalk regions for the remoulded plug model. t_1 is the remoulded chalk (putty) thickness; t_2 is the fractured chalk thickness; L is the embedded length; D is the pile diameter; h is the load eccentricity; t_w is the pile wall thickness.

Table 8.1: Specification of field test piles. L is the embedded length; D is the pile diameter; h is the load eccentricity; t_w is the pile wall thickness.

Field test	Pile Ref	$L(m)$	$D(m)$	$h(m)$	$t_w(m)$	L/D	t_w/D
ALPACA+	A	7.3	1.22	1	0.0246	6	0.02016
ALPACA	B	3.05	0.508	1	0.0206	6	0.04055
ALPACA	C	10.16	0.508	1	0.0206	20	0.04055
Wind Support	D	4	0.762	1	0.0445	5.25	0.0584

Finite element calculations were conducted with PLAXIS 3D employing the chalk properties specified in Table 8.2. In all cases the water table was placed at a depth of 5.51m (Buckley et al., 2020b). The tests were simulated as fully drained due to the relatively high permeability of the chalk (Bundy 2013; Buckley et al., 2020a; Pedone et al., 2023). For comparison, Table 8.2 also includes a specification of the chalk properties adopted in the published FEA of the ALPACA (pile B) field test pile (Pedone et al., 2023).

Remoulded (putty) chalk zone

The bulk unit weight of the remoulded chalk was taken as $\gamma = 19.02\ kN/m^3$ (Buckley et al., 2020a). The HS-Small model was employed to simulate the remoulded chalk behaviour. Drained conditions were assumed following Bundy (2013) who suggested that the remoulded chalk behaves like a non-cohesive material such as sand although consisting of silt to clay size particles. Bialowas et al. (2018) demonstrated that drained triaxial putty samples show a strain hardening and compressive volumetric response; this response is well represented by the HS-Small model. Alvarez-Borges et al. (2020) showed that the failure state of the remoulded chalk profiles is represented by the critical state friction angle ϕ'_{CS} of the intact chalk profile. Strength and dilation data adopted in the current analyses ($c' =$ 1 kPa, $\phi' = 31.5^{\circ}$, $\psi = 0^{\circ}$) were based on the post-yield de-structured chalk behaviour presented in Pedone et al. (2023). The small strain shear modulus is considered to increase with time after pile installation as the putty consolidates (Doughty et al., 2018; Liu et al., 2022b). The small strain shear modulus also increases with depth. For example, Bialowas and Diambra (2018) found that the G_o of reconstituted chalk samples increased with an increase in the applied isotropic pressure. In the current analyses, the simulated G_o varies with mean effective stress using the following formula (Brinkgreve et al., 2018):

$$G_{oref} = \frac{G_o}{(\frac{c'_{ref} \times cos\,\phi' - \sigma'_3 \times sin\,\phi'}{c'_{ref} \times cos\,\phi' - p'_{ref} \times sin\,\phi'})^m} \tag{8.1}$$

where $G_{oref} = 750\text{MPa}$, $m = 0.15$, $\text{p}'_{ref} = 100kPa$ and σ'_3 is the local horizontal effective stress (see Figure 8.3(a)). The simulated G_o data are broadly consistent with those reported in Liu et al. (2022b) who found that the average maximum Young's modulus for remoulded chalk was $E_{umax} = 1320\text{MPa}$ (resulting into $G_o = 440\text{MPa}$) on the basis of undrained monotonic laboratory tests. The stiffness degradation curve simulated by the HS-

Small model – employing $\gamma_{0.7} = 0.00003$ as adopted in the 3D FEA – is illustrated in Figure 8.2(b). K_o for the remoulded chalk is calculated by the Jaky's formula where $K_o = 1 - sin\varphi = 0.478$. Few drained triaxial laboratory tests exist in the literature exploring the shearing response of remoulded chalk samples; Bialowas et al. (2018) showed that the strength and shear response of remoulded chalk samples was unaffected by sample ageing for medium to large strain levels. Bialowas et al. (2018) also noticed that the secant stiffness in standard drained triaxial test E_{50}^{ref} can be affected by the sample's ageing and OCR. For the current study, it was decided to assume conservatively that E_{50}^{ref} =20MPa a value observed for samples with no significant ageing. The reference tangent stiffness in primary oedometric loading E_{oed}^{ref} was assumed to be equal to E_{50}^{ref}. The reference unloading-reloading stiffness E_{ur}^{ref} is assumed to be $E_{ur}^{ref} = 5E_{50}^{ref}$. The failure ratio is $R_f = 0.9$.

Fractured chalk zone

The bulk unit weight of the fractured chalk was taken as $\gamma = 19.02\ kN/m^3$ (Buckley et al., 2020a). The fractured chalk exhibits a ductile response when loaded (Pedone et al., 2023). The HS-Small model was selected to simulate this behaviour. The small strain shear modulus, G_o, of the fractured chalk was considered to be a constant with depth, determined as 50% of the small strain shear modulus for the intact chalk at ground level (see Figure 8.3(b)) (Muir Wood et al., 2015). The selected strength parameters ($c' = 159$ kPa, $\phi' = 29^{\circ}$) were taken directly from the post-rupture strength data in Pedone et al. (2023). The value of $\gamma_{0.7}$ (the shear strain at 70% reduction of G_o) was set to $\gamma_{0.7} = 0.0009$. The stiffness degradation curve simulated with the HS-Small model for this value of $\gamma_{0.7}$ is shown in Figure 8.2(a) compared to the stiffness degradation curve simulated in Pedone et al. (2023). The K_o profile of the fractured chalk zone is calculated by the Jacky's formula where $K_o = 1 - sin\varphi = 0.5152$. Due to the lack of drained triaxial tests data in the literature on fractured

chalk samples E_{50}^{ref}, E_{oed}^{ref} and E_{ur}^{ref} of the fractured chalk profile are assumed to be equal to E_{50}^{ref}, E_{oed}^{ref} and E_{ur}^{ref} of the remoulded chalk profile. The failure ratio is $R_f = 0.9$.

Intact chalk zone

The intact chalk zone was simulated using the elastic-perfectly plastic Mohr Coulomb model. Lower bound values of the small shear stiffness profile G_0, as provided in Buckley et al., (2020a) for the St. Nicholas site, were employed as illustrated in Figure 8.3(a). The adopted strength parameters ($c' = 190$ kPa, $\phi' = 29^{\circ}$) were determined by obtaining a conservative fit to the brittle intact chalk behaviour presented in Vinck et al. (2022) (see Figure 8.3(b)). The bulk unit weight of the chalk was taken as $\gamma = 19.02\ kN/m^3$ (Buckley et al., 2020a). To determine the initial horizontal stresses, a constant value of $K_o = 0.6$ (Vinck et al., 2022) was employed. The Poisson ratio for intact chalk zone ranges from 0.2 to 0.3 (Le et al., 2014; Carotenuto et al., 2018); $v = 0.2$ was adopted for the current analyses.

Interface elements

Interface elements are used around and below the pile to model the pile-soil interaction and to allow for the possibility of soil-pile slipping and gapping. The HS-Small model was employed to simulate the properties of the interface elements. The selected parameters ($c' = 1$ kPa, $\phi' = 34^{\circ}$, $\psi = 0^{\circ}$) were taken directly from data in Pedone et al., (2023). The rest of the HS-Small parameters of the interface properties are the same as the remoulded chalk zone HS-Small model parameters. The adopted interface normal and shear stiffness are $k_n = 2400$ MPa and $k_s = 300$ MPa respectively (Pedone et al., 2023).

Table 8.2: Modelling procedures and parameters employed in the current PLAXIS model. Also shown, for comparison purposes, are the modelling procedures and parameters employed in the Pedone et al. (2023) analyses of the ALPACA field test for pile B.

Chalk model aspects	**Current PLAXIS model**	**ALPHA modelling (Pedone et al., 2023)**
Zone thicknesses	$t_1 = 1.5t_w$, $t_2 = 5t_w$	$t_1 = 1.25t_w$, $t_2 = 4.85t_w$
Remoulded chalk (putty) model	Drained HS-Small (non-linear elasticity)	Drained Mohr Coulomb with non-linear elasticity
Putty strength and stiffness	$c' = 1$ kPa, $\phi' = 31.5^{\circ}$, $\psi = 0^{\circ}, G_{\text{oref}} = 750$ MPa (stress-dependent stiffness)	$c' = 1$ kPa, $\phi' = 31^{\circ}$, $G_o = 1000$MPa
Fractured chalk model	Drained HS-Small (non-linear elasticity)	Drained Mohr Coulomb with non-linear elasticity
Fractured chalk strength and stiffness	$c' = 159$ kPa, $\phi' = 29^{\circ}$, $G_{\text{oref}} = 380$ MPa	$c' = 159$ kPa, $\phi' = 29^{\circ}$, $G_{\text{o}} = 50$ MPa
Intact chalk model	Drained Mohr Coulomb	Drained Mohr Coulomb with non-linear elasticity
Intact chalk strength and stiffness	$c' = 159$ kPa, $\phi' = 29^{\circ}$, $G_{\text{oref}} = 760$ MPa	$c' = 159$ kPa, $\phi' = 29^{\circ}$, $G_{\text{o}} = 500$ MPa

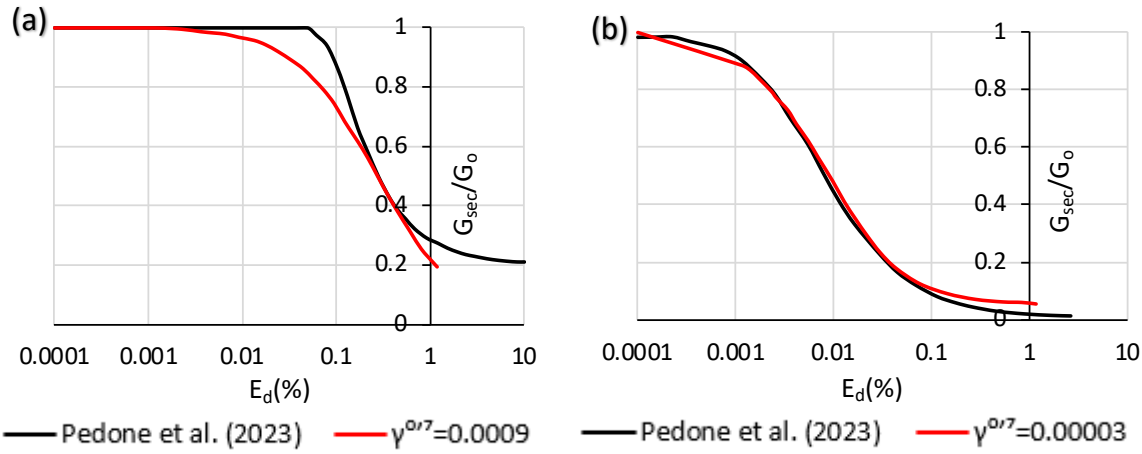


Figure 8.2: Stiffness degradation curves simulated by the HS-Small model (shown in red) compared to the simulated stiffness degradation curve (shown in black presented in Pedone et al. (2023) for: (a) the fractured chalk zone, (b) remoulded chalk zone (putty). The secant stiffness modulus is plotted with the deviatoric strain. The generalised deviatoric strain E_d is related with the shear strain γ as $E_d = 2/\sqrt{3}\,\gamma$. It is important to note that while the similarity between the stiffness reduction curves is apparent, this does not imply that similar stiffness values are being used in this study compared to

Pedone et al. (2023). This discrepancy arises because the small shear stiffness values G_o employed differ between the current PLAXIS 3D models and the ALPHA modelling (refer to Table 8.2).

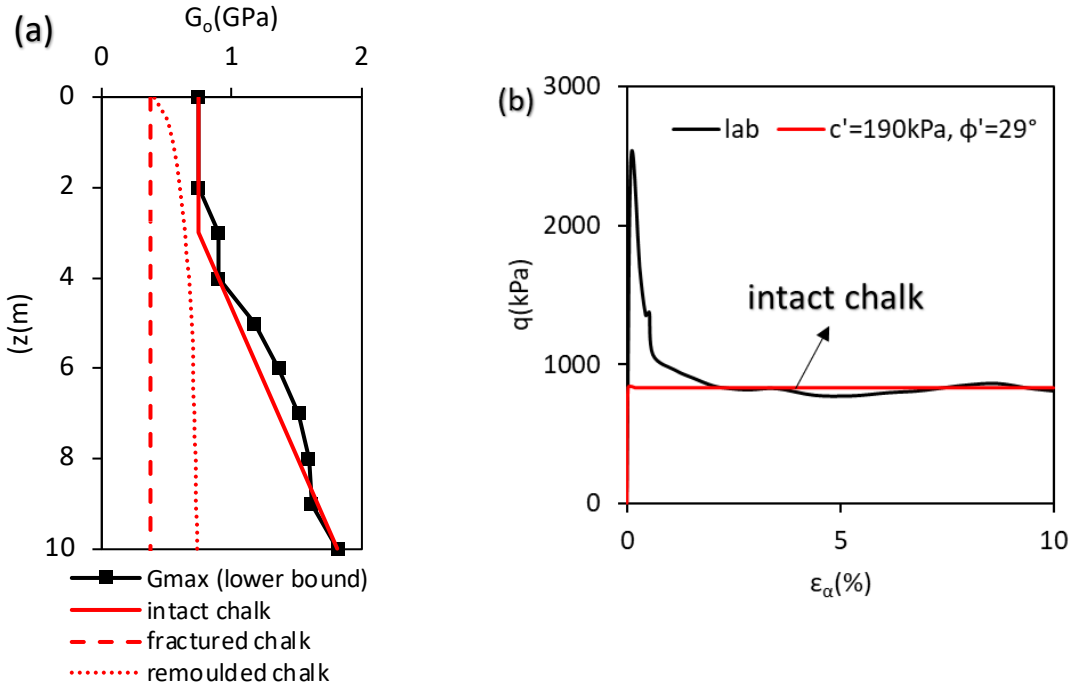


Figure 8.3: (a) G_o profiles for the three chalk zones as adopted in the PLAXIS analyses, shown as red lines. Also shown are the estimated lower bound of G_{max} values for intact chalk measured by seismic CPT as presented in Buckley et al. (2020a). (b) Simulated behaviour of the intact chalk (solid red line) computed with the Mohr Coulomb model employed in the PLAXIS 3D analyses. Also shown (black line) are results from drained triaxial tests reported by Vinck et al. (2022).

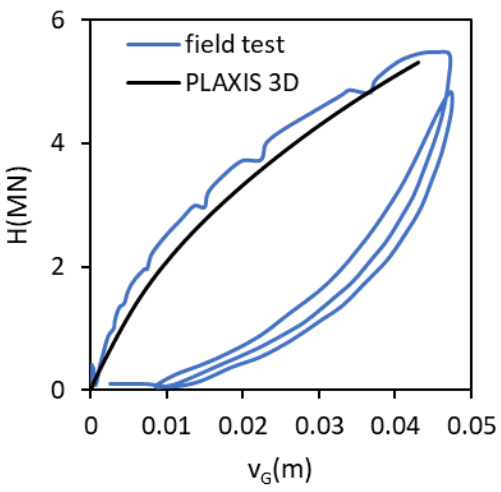
The chalk zone thickness normalised with pile wall thickness values t_1/t_w and t_2/t_w employed in the FEA are listed in Table 8.3; these values fall in the range proposed by Buckley et al. (2018). The two ALPACA and the ALPACA+ field test piles (piles A, B, C), have the same values of t_1/t_w and t_2/t_w since the same pile installation procedure and pile instrumentation was used. However, the Wind Support test (pile D) employed a pile with: (i) a thicker pile wall than adopted for the ALPACA field test piles and (ii) a different pile instrumentation configuration. Ciavaglia et al. (2017a, b) observed that – for the Wind Support pile test – the normalised remoulded chalk zone thickness t_1/t_w ranged from 0.5 to 0.9; a value of $t_1/t_w = 0.7$ was selected for the current analysis. The normalised fractured chalk zone thickness t_2/t_w for the Wind Support pile is not fully defined in Ciavaglia et al.

(2017a, b); for this study t_2/t_w was estimated by performing a range of trial and error FEA to achieve a good fit with the field test response in Figure 8.4(d); a value of $t_2/t_w = 2.5$ was considered to provide a good fit.

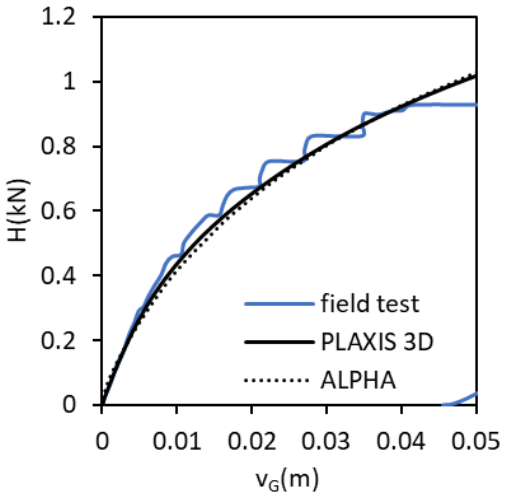
Table 8.3: Thickness of the de-structured chalk zones employed in the FEA. t_1 is the remoulded chalk (putty) thickness; t_2 is the fractured chalk thickness; t_w is the pile wall thickness.

Field test	**Pile Ref**	t_1/t_w	t_2/t_w
ALPACA+	A	1.5	5
ALPACA	B	1.5	5
ALPACA	C	1.5	5
Wind Support	D	0.7	2.5

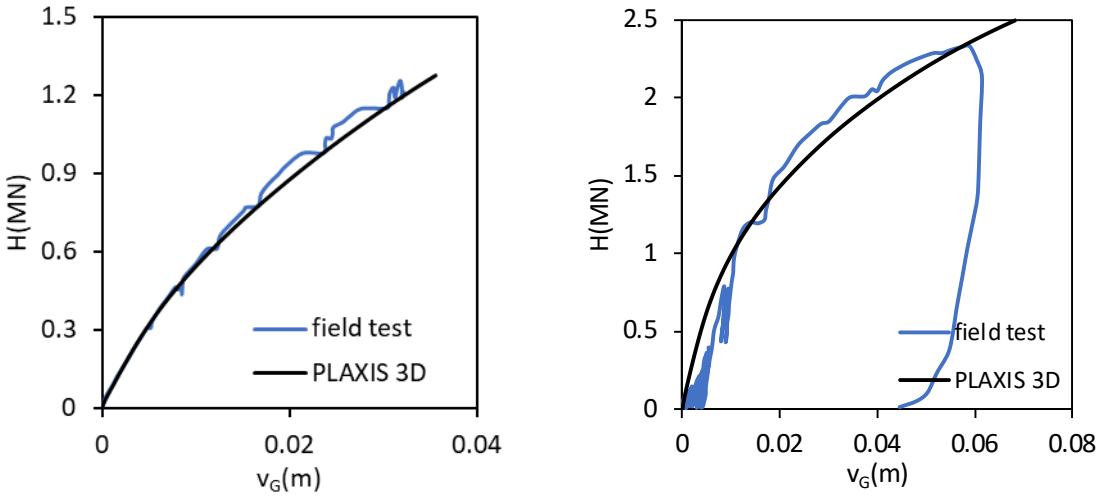
Comparisons between the PLAXIS 3D results and the field test data are shown in Figure 8.4. The close comparison between the data sets suggests that the PLAXIS 3D model is capable of delivering realistic predictions of the monotonic lateral load performance of piles embedded in the St. Nicholas chalk profile. Figure 8.4(b) presents for further comparison the pile load-displacement response computed in the ALPHA project (Pedone et al., 2023) adopting a drained Mohr Coulomb model with non-linear elasticity to simulate the behaviour of the intact chalk zone; the current PLAXIS 3D results compare well with the Pedone et al., (2023) data.



(a)



(b)



(c) (d)

Figure 8.4: Comparison between the PLAXIS results (shown in black) and field test data (in blue): (a) ALPACA+ ($L = 7.3m$); (b) ALPACA ($L = 3.05m$); (c) ALPACA ($L = 10.16m$); (d) Wind Support ($L = 4m$) (Ciavaglia et al., 2017b). The black dotted line indicates the results presented in Pedone et al. (2023) during the ALPHA project.

8.3 Soil reaction curves for field test piles

Figure 8.5 shows an example of the distributed and base numerical soil reaction curves extracted at selected depths for the ALPACA pile (pile B). The numerical soil reactions were extracted from the PLAXIS 3D analysis using the procedure described in Section 4.2.

For the Pile B ($L/D = 6$), the numerical soil reactions coupled with OxPile software provide a close fit with the 3D FEA results (Figure 8.6). To investigate the contributions of the individual soil reaction components (p, m, H_B, M_B) separate analyses were conducted in which the distributed moment and the base horizontal force and moment are removed from the OxPile analysis. The results presented in Figure 8.6 indicate that the distributed lateral load p contributes more than 97% of the total lateral capacity and distribution of bending moment at a ground level lateral displacement of 40mm. This behaviour can be explained due to: (i) the relatively high pile L/D slenderness ratio employed for the field test piles, (ii) the soil plug is fully remoulded affecting the base soil reactions as observed in Figure 8.9.

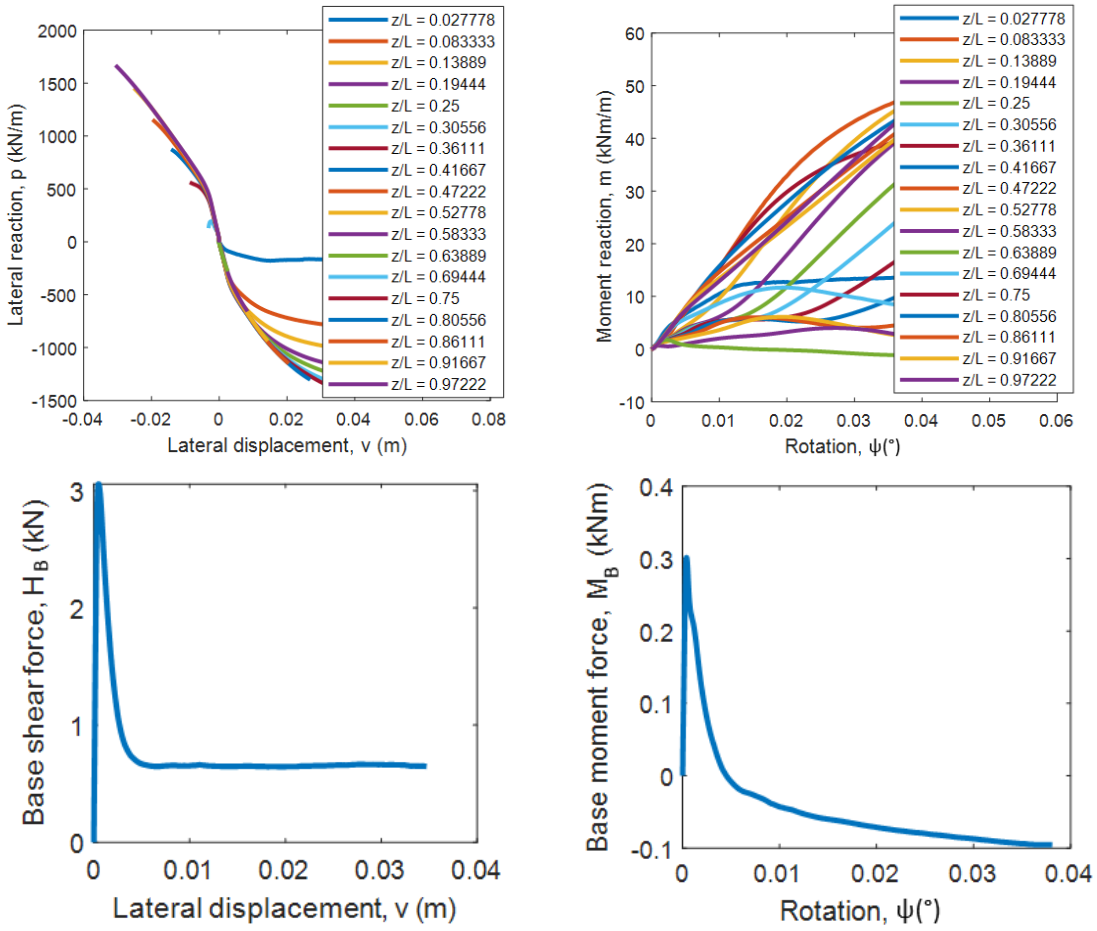


Figure 8.5: Numerical soil reaction generated by the PLAXIS 3D FEA for pile B.

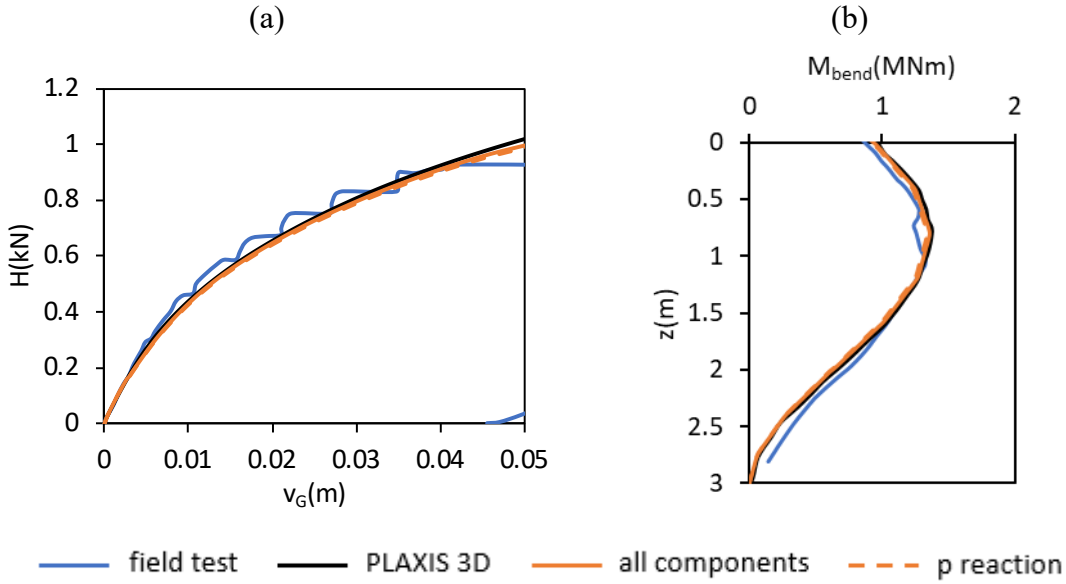


Figure 8.6: Comparison of pile response determined from the 3D FEA and the corresponding results employing the numerical soil reaction curves obtained using OxPile software. (a) ground level load-displacement response, (b) bending moment profiles profile with pile depth at 40mm of lateral displacement at ground level. The contributions of p, m, H_B and M_B are indicated by the orange line denoted by 'all components' in the legend; the contributions of p are indicated by the orange dashed line.

8.4 Influence of the de-structured chalk zones thickness

Field test information provided by pressumeter tests undertaken during the ALPACA project showed that the thickness of the de-structured chalk zones had a considerable impact on the lateral behaviour of the monopiles. To further investigate this issue 3D FEA were implemented for piles B and D (Table 8.1) exploring: (i) the influence of the remoulded chalk zone thickness on the lateral behaviour of pile B, (ii) the influence of the fractured chalk zone thickness on the lateral behaviour of pile D. For both piles the thickness of the other de-structured chalk zone was held constant to examine the thickness effect for a single de-structured zone. The normalised de-structured chalk zone thickness values t_1/t_w and t_2/t_w employed in the FEA are presented in Table 8.4. The examined values of t_1/t_w, t_2/t_w fall within the range proposed by Buckley et al. (2018).

Table 8.4: Thickness of the de-structured chalk zones employed in the FEA. t_1 is the remoulded chalk (putty) thickness; t_2 is the fractured chalk thickness; t_w is the pile wall thickness.

Field test	**Pile Ref**	t_1/t_w	t_2/t_w
ALPACA	B	0.9, 1,2, 1.5	5
Wind Support	D	0.7	2.5, 3.5, 5.8

The PLAXIS results in Figure 8.7(b) illustrate the influence of the selected fractured chalk zone thickness t_2 on the computed response of pile D; it is evident that the lateral behaviour of a monopile embedded in chalk is strongly dependent on the assumed value of t_2. Figure 8.7(a) shows the influence of the selected remoulded chalk zone thickness t_1 on the computed response of pile B; the lateral behaviour of a monopile embedded in chalk is also affected, thus in a smaller magnitude, by the value of t_1. This suggests that it may be appropriate to incorporate t_1, t_2 as training features in the chalk model so that values of t_1 , t_2 normalised by the pile wall thickness t_w are treated as input parameters to the model (rather than being 'hard-wired' within the model).

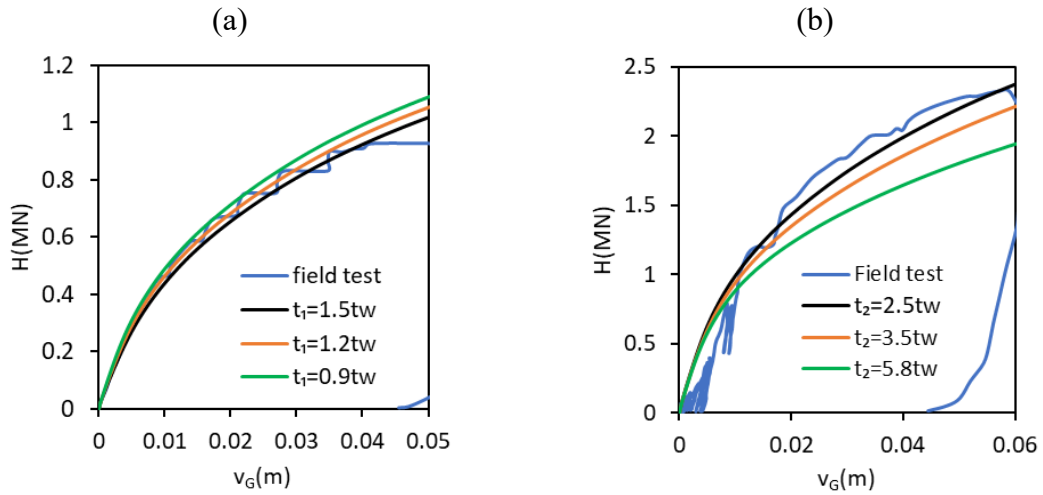


Figure 8.7: Influence of de-structured chalk zones thickness on pile lateral behaviour. (a) Effect of thickness of remoulded chalk zone t_1 on the pile B (Table 8.1) behaviour, (b) Effect of thickness of fractured chalk zone t_2 on pile D (Table 8.1) behaviour.

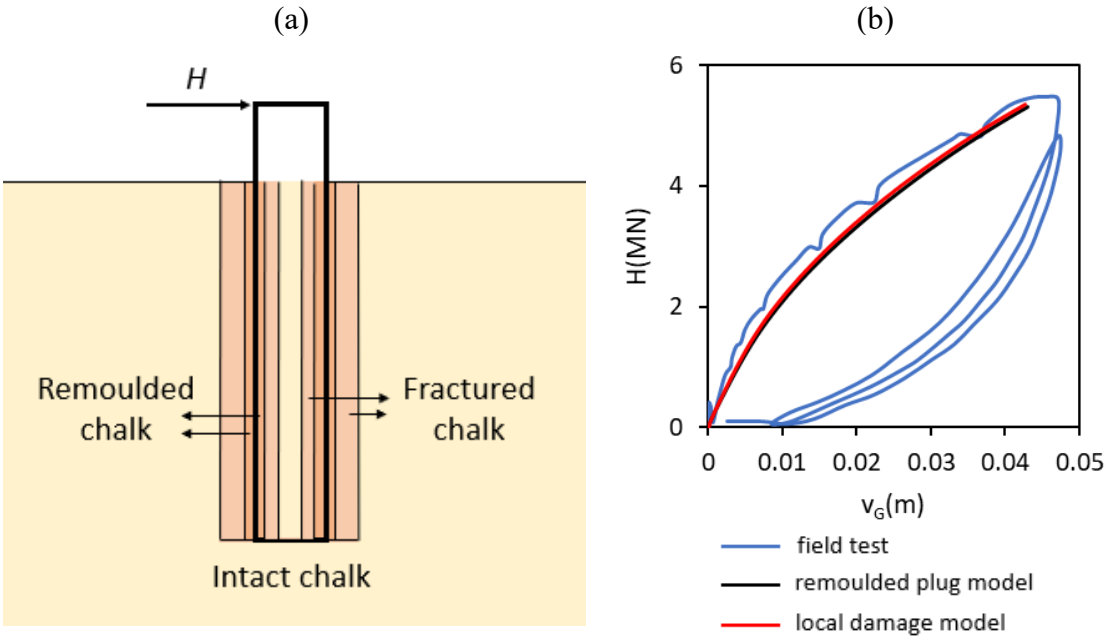


Figure 8.8: (a) Assumed configuration of the intact, fractured and remoulded chalk regions in the local damage model where the thickness of the de-structured zones outside the pile is equal to that inside the pile. t_1 is the remoulded chalk thickness; t_2 is the fractured chalk thickness; (b) Influence of de-structured chalk zones thickness of the plug-in soil on the pile A (Table 8.1) lateral behaviour.

Additional field test information from the ALPACA+ field tests, where the normalised ratio t_w/D (D is the pile diameter) is lower than the other field test piles and closer to typical values for offshore monopiles, showed that the soil plug is not fully converted to remoulded chalk. To examine the influence of chalk de-structuring in the plug on the overall behaviour of the pile, an additional FEA was implemented for the ALPACA+ field test pile (Pile A)

adopting a 'local damage' model where the thickness of the remoulded and the fractured zones inside the pile are equal to the thicknesses outside of the pile (Figure 8.8(a)). The two modelling approaches for the soil plug are compared in Figure 8.8(b). The local damage model has a negligible influence on the pile's lateral behaviour. The differences between the local damage and the remoulded plug models are concerned with the computed base soil reactions (Figure 8.9); as can be observed both the horizontal force and base moments soil reactions are significantly increased accounting for a local damage model.

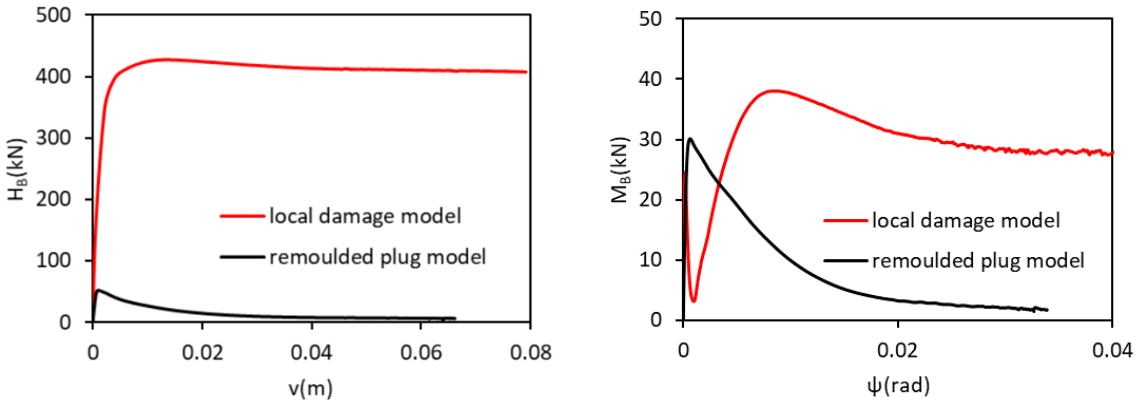


Figure 8.9: Comparison between the two plug models on the resulted base soil reaction curves.

8.5 Calibration of a data-driven chalk model

This section describes the development of a data-driven 1D design model for the four field test piles presented in Table 8.1. The calibration analyses were performed using the PLAXIS 3D software. The 3D models were generated using a Python code which constructs the model following the stages outlined in Section 3.2; the pile was divided into 20 slices of equal length to form a sufficient database of soil reaction curves. The soil reaction curves were produced by employing the procedure described in Section 4.2. The numerical soil reaction curve data were normalised using the 'sand' data-driven non-dimensional forms in Table 4.3. All the extracted soil reaction curves were validated by comparison with the PLAXIS 3D predictions for every monopile geometry by using the numerical soil reaction curves in OxPile.

8.5.1 Calibration finite element analyses

A set of 20 calibration analyses were conducted using the procedures and parameters presented in Section 8.2. The calibration piles are specified in Table 8.5. Calibration piles CST1 to CST20 were devised so that groups of piles occupy a parameter space in proximity to the field test piles specified in Table 8.1. Calibration piles CST1-CST6 occupy a portion of parameter space close to piles A and B; calibration piles CST7-CST10 occupy a portion of parameter space close to pile C; calibration piles CST11-CST20 occupy a portion of calibration space close to pile D. The calibration space employs all the training inputs $[L/D]$, $[t_W/D]$, $[t_1/D]$, $[t_2/D]$ that are considered to govern the soil reaction curves for the field test piles. With these calibration piles it is possible to treat the four field test piles as 'unseen cases'; the performance of the chalk model can therefore be assessed in terms of the extent to which it can successfully reproduce the field test data.

Table 8.5: Calibration piles employed for the chalk model.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D	t_w/D	t_1/t_w	t_2/t_w
CST1	1.4	9.1	28	1	6.5	0.02	1.5	5
CST2	1.4	6.3	28	1	4.5	0.02	1.5	5
CST3	0.9	5.85	24	1	6.5	0.0267	1.5	5
CST4	0.9	4.05	24	1	4.5	0.0267	1.5	5
CST5	0.4	2.6	20	1	6.5	0.05	1.5	5
CST6	0.4	1.8	20	1	4.5	0.05	1.5	5
CST7	0.9	19.8	24	1	22	0.0267	1.5	5
CST8	0.9	14.4	24	1	16	0.0267	1.5	5
CST9	0.4	8.8	20	1	22	0.05	1.5	5
CST10	0.4	6.4	20	1	16	0.05	1.5	5
CST11	0.9	5.85	54	1	6.5	0.06	1.5	5
CST12	0.9	4.05	54	1	4.5	0.06	1.5	5
CST13	0.4	2.6	24	1	6.5	0.06	1.5	5
CST14	0.4	1.8	24	1	4.5	0.06	1.5	5
CST15	0.9	5.85	45	1	6.5	0.05	0.6	5
CST16	0.9	4.05	45	1	4.5	0.05	0.6	5
CST17	0.9	5.85	45	1	6.5	0.05	0.6	2
CST18	0.9	4.05	45	1	4.5	0.05	0.6	2
CST19	0.4	2.6	24	1	6.5	0.06	0.6	2
CST20	0.4	1.8	24	1	4.5	0.06	0.6	2

8.5.2 GPR model training procedure

The numerical soil reaction curves are normalised using the 'sand' data-driven normalisation approach presented in Table 4.3. The knot point parameters are considered to be functions of training inputs.

8.5.2.1 Training Inputs

The GPR models trained for the distributed soil reactions $\bar{p}, \bar{m}$ employ the following inputs: $[z/L]$ where z is distance along the embedded pile; $[L/D]$ is a proxy for the pile slenderness; $[G_o/1000\sigma'_v]$ discriminates locally the soil reaction curves with the same values of $[z/L]$ and $[L/D]$; $[t_W/D]$ is a proxy for the pile wall thickness ratio; $[t_1/t_W]$ represents the thickness of the remoulded chalk zone; $[t_2/t_W]$ describes the thickness of the remoulded chalk zone. Finally, an additional input has been identified to train general chalk models accounting for chalk profiles with different strength and stiffness properties. The UCS (unconfined compression strength), q_u or the dry density, ρ_d, of the intact chalk zone (Bowden et al. 2002) could be employed to develop an appropriate training input. However, this training input is inactive for the current data-driven 1D design model since it is developed for a site-specific homogenous chalk profile.

The GPR models trained for the base soil reactions $\bar{H}_B, \bar{M}_B$ utilise a smaller database and employ three training inputs: (i) $[L/D]$, (ii) $[G_o/1000\sigma'_v]$, (iii) $[t_W/D]$. As mentioned above the q_u or ρ_d values can be used as an additional fourth input to account for chalk profiles of different strength and stiffness properties.

8.5.2.2 Training procedure

The normalised distributed lateral reaction $\bar{p}$ and moment $\bar{m}$ knot points were standardised before the training of the GPR models by using the formula:

$$\bar{y} = \frac{y - u}{s} \tag{8.2}$$

where $\bar{y}$ is the standardised value of each knot point output; u is the mean value and s is the standard deviation calculated from the data for each knot point. By using the standardised knot points, the magnitude of the signal variance found by the $\bar{y}$ output datapoints is significantly reduced, easing the training procedure and the kernel's optimisation of the GPR models. The high covariance between the $\bar{p}$, $\bar{m}$ knot points is found due to:

1. The selected 'sand' normalisation framework of the soil reaction curves which includes the σ'_v instead of s_u. This normalisation procedure is consistent with the normalisation adopted for layered soils. However, it cannot incorporate the relatively high undrained shear strength s_u of the chalk profiles (much higher than the s_u of soil profiles) resulting in high magnitude of normalised reaction values and relatively high covariance of the knot points.
2. The knot points of the piles that employ L/D ratios ranging from 16 to 22 which form a plastic hinge (refer to Section 2.1) at approximately a depth of $z = 0.3L$; the $\bar{p}$, $\bar{m}$ data below this depth have low values compared to the less slender piles with L/D values ranging from 4.5 to 6.5.

The final predictions of the $\bar{p}$, $\bar{m}$ soil reaction knot point data are scaled back using Equation (8.2) to the original representation before using them in OxPile software to determine the actual distributed soil reactions. Near the pile rotation points, where the magnitude of the soil reactions is small, the GPR models might predict negative values of

knot point data; these predictions have small magnitudes; They are physically unrealistic and cannot be employed in OxPile. Therefore, they are removed from the data. A GPR model is trained for each of the knot points for each soil reaction component ($\bar{p}$, $\bar{m}$, $\bar{H}_B$, $\bar{M}_B$). The Matérn ($v = 5/2$) covariance function with Automatic Relevance Determination (ARD) is employed where the σ_f^2 and l_i hyperparameters are optimised by maximising the log marginal likelihood; different σ_f^2, l_i values typically emerge for each input dimension. The white noise kernel $k_{WN}(x, x')$ is added to the Matérn kernel to account for possible noise in the dataset. An 80:20 training-test set split ratio was adopted. All the trained GPR models were assessed using the MSE and R^2 train and test scores.

8.5.2.3 GPR model predictions

Table 8.6 lists the optimised hyperparameters obtained for the $\bar{p}$, $\bar{m}$ knot point parameters after the GPR model training. The training of the soil reaction curve knot points employs a small database and the trained GPR models provide point predictions (Section 4.6.2.1) easing the GPR model implementation; therefore, their results are not presented in this thesis. The relatively low values of the optimised normalised pile wall thickness [t_W/D] lengthscale hyperparameter indicate that the GPR predictions are sensitive to the pile wall thickness which determines the extent of the de-structured chalk zones.

Figure 8.10 compares the depth variation predictions for pile CST4 ($L/D = 4.5$) by the GPR models with the knot points for the piles with $L/D = 4.5, 6.5$ extracted from the 3D FEA and used for the training of the ML models. The 95% confidence interval (shown as shaded bounds) is also presented; the interval is relatively small since the GPR models predict for data that have been included in the training of the GPR models. From all the graphs in Figure 8.10, it can be observed that generally the GPR model can accurately capture the non-linear nature of the data. Furthermore, piles with the same value of L, D and

h but with different t_w do not have the same soil reactions at the same normalised depth z/L. This observation confirms the importance of the pile wall thickness on the de-structured chalk zones and consequently on the pile's lateral behaviour. In the data in Figure 8.10 a smaller value of t_w correlates with a greater magnitude of soil reactions. Besides, the graphs in Figure 8.10 show that it would be practically impossible to define a unique normalisation framework appropriate for a chalk model to develop a PISA design model for chalk.

Table 8.6: Optimised kernel parameters for the GPR models for the knot point parameters associated with the chalk GPR model. (σ_f is the signal variance of the Matérn kernel, σ_n is the variance of the noise kernel, l is the lengthscale for every input dimension).

Knot point	$\bar{p}$	$\bar{m}$
1	σ_f=6.4, σ_n^2=0.00553, $l_{z/L}$=3.11, $l_{L/D}$=19.2, $l_{t_w/D}$=0.0747, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=2.92, l_{t_2/t_w}=8.39	σ_f=3.46, σ_n^2=0.0111, $l_{z/L}$=1.22, $l_{L/D}$=5.79, $l_{t_w/D}$=0.0473, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=2.15, l_{t_2/t_w}=5.91
2	σ_f=5.41, σ_n^2=0.00687, $l_{z/L}$=0.991, $l_{L/D}$=8.64, $l_{t_w/D}$=0.31, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=3.92, l_{t_2/t_w}=9.31	σ_f=3.51, σ_n^2=0.00662, $l_{z/L}$=1.37, $l_{L/D}$=5.82, $l_{t_w/D}$=0.0617, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=2.56, l_{t_2/t_w}=6.78
3	σ_f=31.6, σ_n^2=0.00083, $l_{z/L}$=6.51, $l_{L/D}$=28.7, $l_{t_w/D}$=0.232, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=19.7, l_{t_2/t_w}=37.4	σ_f=3.81, σ_n^2=0.000902, $l_{z/L}$=0.359, $l_{L/D}$=34.7, $l_{t_w/D}$=0.123, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=2.71, l_{t_2/t_w}=6.17
4	σ_f=13.3, σ_n^2=0.00167, $l_{z/L}$=1.61, $l_{L/D}$=18.1, $l_{t_w/D}$=0.245, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=17.7, l_{t_2/t_w}=27.2	σ_f=10.5, σ_n^2=0.000215, $l_{z/L}$=1.38, $l_{L/D}$=14.9, $l_{t_w/D}$=0.304, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=12.3, l_{t_2/t_w}=19.4
5	σ_f=8.82, σ_n^2=0.00158, $l_{z/L}$=0.542, $l_{L/D}$=90.7, $l_{t_w/D}$=0.143, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=19.3, l_{t_2/t_w}=32.3	σ_f=31.6, σ_n^2=8.68e-5, $l_{z/L}$=2.15, $l_{L/D}$=41.5, $l_{t_w/D}$=0.681, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=24.8, l_{t_2/t_w}=41.9
6	σ_f=8.07, σ_n^2=0.00359, $l_{z/L}$=0.731, $l_{L/D}$=100, $l_{t_w/D}$=0.130, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=20.5, l_{t_2/t_w}=53.3	σ_f=12.1, σ_n^2=0.000182, $l_{z/L}$=1.21, $l_{L/D}$=28.1, $l_{t_w/D}$=0.393, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=24.5, l_{t_2/t_w}=27.5
7	σ_f=10.5, σ_n^2=0.00429, $l_{z/L}$=0.984, $l_{L/D}$=100, $l_{t_w/D}$=0.123, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=27.0, l_{t_2/t_w}=84.2	σ_f=8.77, σ_n^2=0.000543, $l_{z/L}$=1.1, $l_{L/D}$=22.6, $l_{t_w/D}$=0.18, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=11.1, l_{t_2/t_w}=20.9
8	σ_f=11.4, σ_n^2=0.00395, $l_{z/L}$=1.05, $l_{L/D}$=100, $l_{t_w/D}$=0.119, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=26.7, l_{t_2/t_w}=91.8	σ_f=8.77, σ_n^2=0.000543, $l_{z/L}$=1.1, $l_{L/D}$=22.6, $l_{t_w/D}$=0.18, $l_{G_o/\sigma_v'}$=100, l_{t_1/t_w}=11.1, l_{t_2/t_w}=20.9

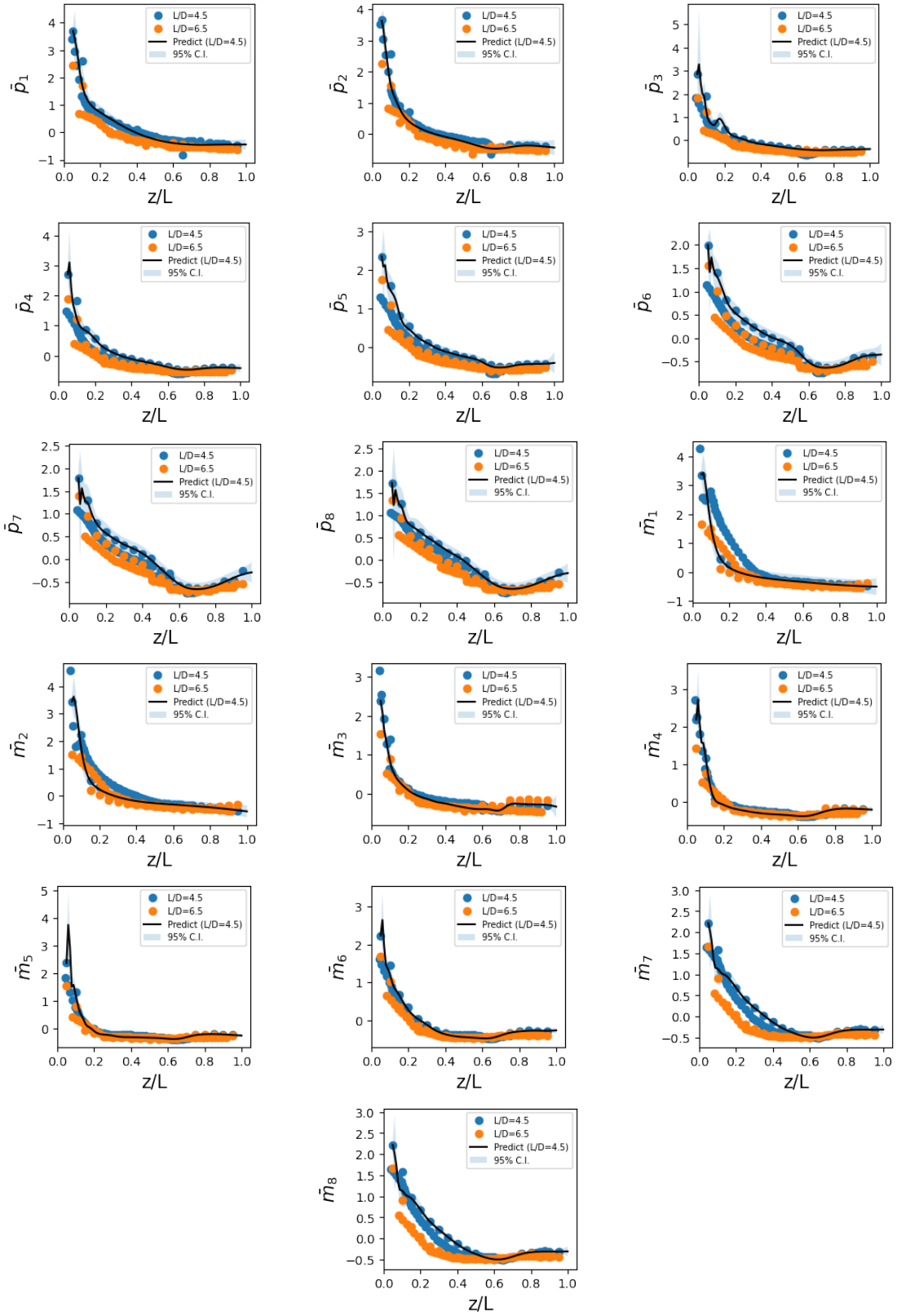


Figure 8.10: Example GPR outputs for the soil reaction knot points for the CST4 pile ($L/D = 4.5$, $t_w/D = 0.0267$) embedded in St. Nicholas chalk site. The depth variation of the mean GPR prediction for the knot points is indicated by the black line. The 95% confidence interval is shown (shaded bounds). The blue and the orange markers are the knot points for the calibration analysis C3, C4, C11, C12 with similar values of (D, t_1/t_w , t_2/t_w). All the distributed lateral load $\bar{p}$ and moment $\bar{m}$ knot points and GPR predictions were standardised by using Equation (8.2).

8.6 Application of the chalk model

To conduct an analysis using the chalk model for unseen cases (e.g. the ALPACA configurations), GPR is used to predict the depth variation of the soil reaction knot point data along the embedded pile. Soil reactions curves determined in this way are employed in the PISA 1D FE framework to compute the monopile's response. An initial exercise has been conducted in which predictions from the 1D model are compared with the original training data. This comparison indicates that the 1D model provides a close fit to the calibration analyses; the PISA metrics η, ρ (Equation (4.20)) comparing the chalk model predictions with the PLAXIS 3D calculations for all the calibration analyses are listed in Table 8.7.

Table 8.7: Values of accuracy metric η and ratio metric ρ for the piles' ultimate response.

Pile	η_{ult}	ρ_{ult}	Pile	η_{ult}	ρ_{ult}
CST1	0.994	0.997	**CST11**	0.990	0.974
CST2	0.989	1.009	**CST12**	0.998	0.995
CST3	0.964	0.964	**CST13**	0.994	0.992
CST4	0.986	0.977	**CST14**	0.954	0.954
CST5	0.956	0.952	**CST15**	0.991	0.998
CST6	0.976	0.976	**CST16**	0.983	0.976
CST7	0.985	0.983	**CST17**	0.994	0.994
CST8	0.984	0.979	**CST18**	0.985	0.976
CST9	0.931	0.925	**CST19**	0.978	0.971
CST10	0.915	0.921	**CST20**	0.960	0.954
Average	0.973	0.975	**CoV(%)**	-	2.067

Comparisons between predictions from the 1D model and PLAXIS 3D with field data for the ALPACA+, ALPACA, Wind Support field test piles (considered to be unseen cases) are shown in Figure 8.11. The 1D model ('chalk model') is able to deliver a close match with the field data and the PLAXIS 3D data in all cases. The PISA metrics η, ρ (Equation (4.20)) comparing the chalk model predictions with the PLAXIS 3D calculations for the four unseen cases are listed in Table 8.8. The average value of η is 0.9855; this indicates a good match

with the 3D FEA data. Furthermore, the ρ ratio metric data indicate that there is no significant overall bias.

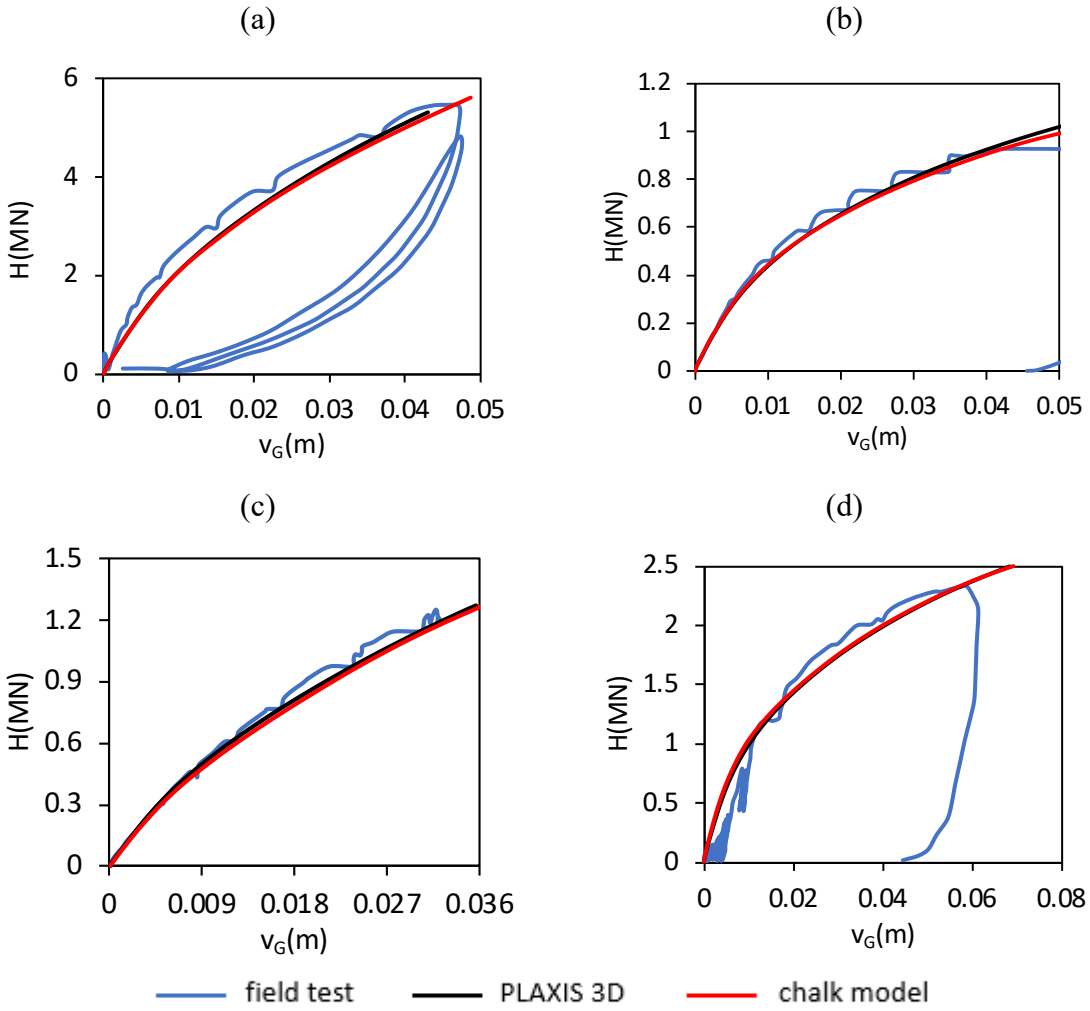


Figure 8.11: Comparison between predictions from the chalk model (shown in red), PLAXIS 3D results (in black) and field test data (in blue): (a) ALPACA+ (b) ALPACA ($L = 3.05$m) (c) ALPACA (L=10.16m) (d) Wind Support.

Table 8.8: Values of accuracy metric η and ratio metric ρ for the piles' ultimate response.

Pile	η_{ult}	ρ_{ult}
A	0.981	0.977
B	0.985	0.974
C	0.984	0.972
D	0.992	0.997
Average	0.9855	0.980
CoV(%)	-	0.91

8.7 Refining the chalk model for different thicknesses for the de-structured chalk zones

A set of 10 additional calibration analyses employing the piles (CST21 to CST30) in Table 8.10 were implemented to improve the chalk model's accuracy for design cases corresponding to different de-structured chalk zones thickness. The refined 1D model can be used for piles with L/D ranging from 4.5 to 6.5 with a normalised wall thickness t_w/D between 0.02 to 0.06 and a thickness of remoulded chalk zone in the range $t_1 = 0.6t_W - 1.5t_W$ and thickness of fractured chalk zone in the range $t_2 = 2t_W - 5t_W$. The FEA performed to develop this model are detailed in Appendix A (refer to Table A.10).

Table 8.9: Exploratory analyses for the ALPACA+, ALPACA and Wind Support piles.

Case	t_1/t_w	t_2/t_w
1	1.25	3
2	0.9	4

To assess the refined chalk model a set of additional exploratory calculations have been conducted for the piles A, B, D (Table 8.1) with different values of t_1 and t_2. The values of t_1 and t_2 employed are specified in Table 8.9. Both values of t_1/t_w and t_2/t_w are unseen, although they lie within the bounds of the parameter space.

Table 8.10: Calibration piles employed for the refined chalk model.

Pile	$D(m)$	$L(m)$	$t_w(mm)$	$h(m)$	L/D	t_w/D	t_1/t_w	t_2/t_w
CST21	0.4	1.8	24	1	4.5	0.06	0.6	5
CST22	0.4	2.6	24	1	6.5	0.06	0.6	5
CST23	0.4	1.8	24	1	4.5	0.06	1.05	3.5
CST24	0.4	2.6	24	1	6.5	0.06	1.05	3.5
CST25	0.9	5.85	45	1	6.5	0.05	1.05	3.5
CST26	0.9	4.05	45	1	4.5	0.05	1.05	3.5
CST27	0.9	5.85	18	1	6.5	0.02	0.6	2
CST28	0.9	4.05	18	1	4.5	0.02	0.6	2
CST29	0.9	5.85	18	1	6.5	0.02	1.05	3.5
CST30	0.9	4.05	18	1	4.5	0.02	1.05	3.5

Comparisons between the response computed with the 1D model and the corresponding 3D FE results are shown in Figure 8.12. The 1D model successfully reproduces the 3D FE data in all cases. This suggests that by calibrating correctly a chalk model using the possible minimum, average and maximum de-structured zone thicknesses observed after the piles' installation procedure, the 1D model will be able to provide high-accuracy predictions for arbitrary values of t_1/t_w and t_2/t_w within the calibrated range.

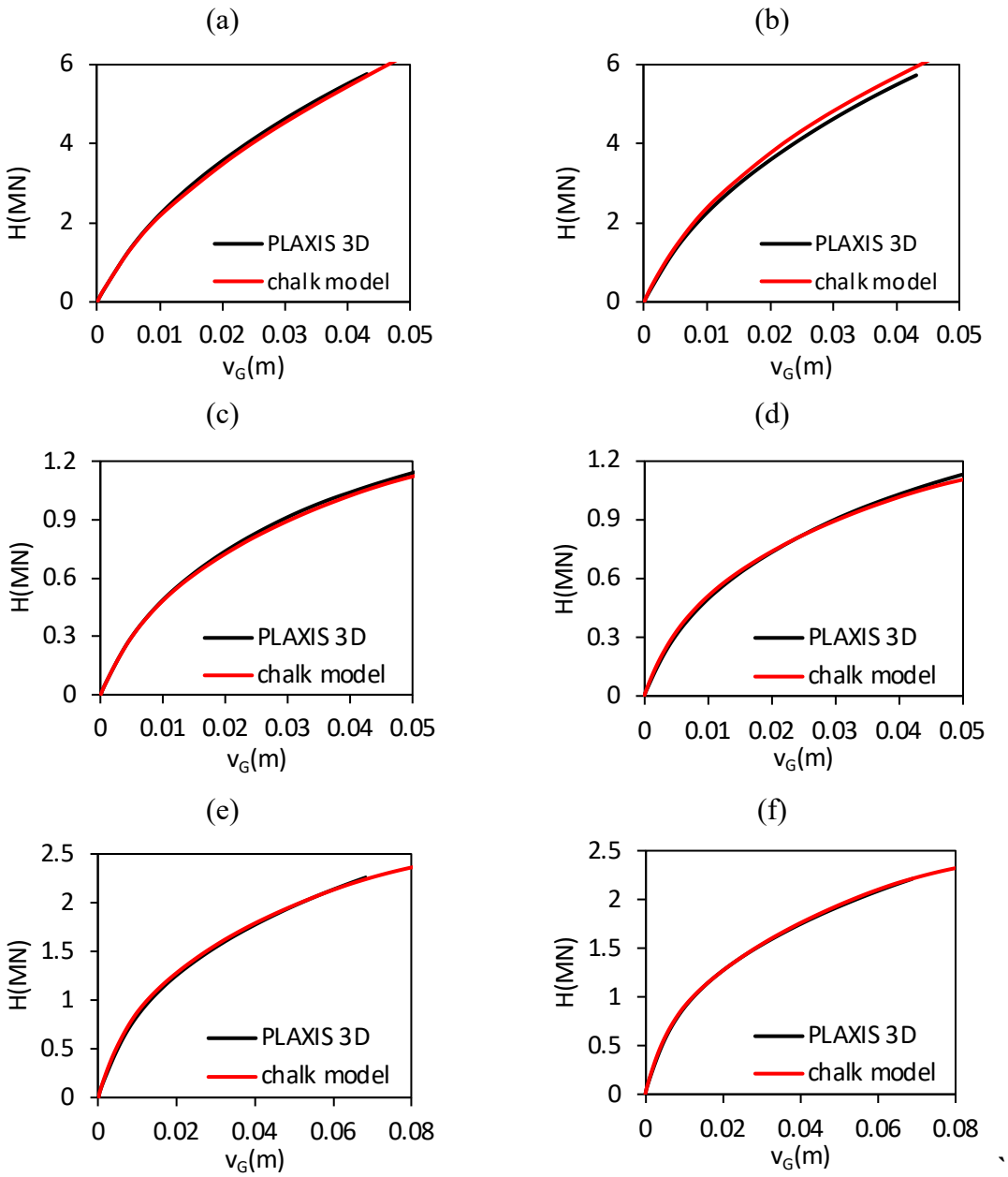


Figure 8.12: Comparison between predictions from the chalk 1D model (shown in red), PLAXIS 3D results (in black) where Case 1 (Table 8.9) is presented on the left graphs and Case 2 (Table 8.9) on the right graphs: (a, b) ALPACA+ (c, d) ALPACA ($L = 3.05m$), (e, f) Wind Support.

8.8 Summary

This chapter introduces a framework for generating a chalk 1D model using a data-driven method. A chalk 1D model was developed for the ALPACA, ALPACA+ and Wind Support field test piles embedded in the St. Nicholas chalk site.

The pile installation process de-structures the chalk around the pile. Three different distinct zones with different strength and stiffness characteristics were simulated in the 3D model to capture the pile installation effects. In this study, the Mohr-Coulomb model was used to simulate the brittle behaviour of intact chalk employing a conservative calibration for the c', ϕ' parameters. The FEA results were validated with measured field test data. It was found that for the field test piles geometry, the soil plug (local damage or remoulded plug model) produce very similar results – for the large diameter offshore monopile geometries it is expected that the local damage model would be a more realistic assumption (increasing the magnitude of the base soil reaction curves). The thickness of the remoulded and fractured chalk zones outside the pile wall (which depend on the pile wall thickness) have a significant effect on the pile's lateral capacity. This effect increases the complexity of the chalk 1D model compared to the 1D model's developed for soils; additional normalised variables (t_W/D, t_1/t_W, t_2/t_W) are required to develop a practical 1D model. The additional variables introduced make the PISA design model unsuitable to generate a 1D model for chalk, since there is no straightforward normalisation approach to map the numerical soil reaction curves. The chalk 1D model developed in this thesis using the data-driven method is capable of computing in good accuracy the field test piles' lateral response for different thicknesses of the de-structured chalk zones. The method is readily extended to full-scale monopile configurations. The UCS or the dry density of the intact chalk material can be employed as an additional training input to create a general chalk 1D model that will consider a range of chalk profiles of different unconfined compression strength and stiffness.

Chapter 9

9 Conclusions

9.1 Introduction

This thesis explores the lateral response of offshore monopiles subjected to monotonic loading conditions. The main aim is to develop a data-driven method to generate 1D models capable of computing the pile's backbone curve and induced bending moments. Data-driven 1D design models employ the GPR ML model to formulate the soil reaction curve data and come as an alternative to the PISA design model improving the generality, accuracy and applicability of the generated 1D models. The soil reaction curves used to calibrate the 1D models were generated by numerical analyses performed using the PLAXIS 3D software. This chapter describes the key contributions and possible directions of future work.

9.2 Key Contributions

Data-driven design method

This thesis outlines the key stages involved in processing and normalising soil reaction curves data, making them suitable for ML implementation to calibrate data-driven 1D design models. The data-driven method developed in this thesis focuses on laterally loaded monopiles but can also be adapted for other geotechnical applications, including axially loaded monopiles, laterally loaded suction caissons, anchor design for floating OWTs, and

retaining walls. The data-driven method – using appropriate training inputs – can be employed to develop 1D models for clay, sand, chalk or layered soils. Both the FNN regression and the GPR ML models were found to describe well the knot point parameters data. For the relatively small database employed in the data-driven 1D design models (a few thousand knot points), the GPR model outperformed the FNN in terms of accuracy for unseen data and interpretability. Additionally, the GPR model estimates the uncertainty of the predicted knot point parameters through its confidence intervals.

The data-driven 1D design models employ the GPR model together with the eight-knot point piecewise spline function with the PCHIP interpolation scheme; they calculate in high fidelity the numerical soil reaction curves at a local level. Therefore, they do not require any optimisation stages as it is required for the PISA design model. Moreover, the data-driven 1D design models consistently predict with a high accuracy the pile's load displacement and moment rotation backbone curves and the distribution of bending moment along the pile. Finally, using the GRP model, the data-driven 1D design models can provide reliable predictions for design cases that do not fall significantly away from the 1D model's calibration space – this type of analysis is not recommended with the PISA design model.

3D FEA for laterally loaded monopiles using built-in PLAXIS constitutive models

Numerical analyses were performed for piles embedded in clay, sand and chalk sites. All analyses were performed employing widely-used built-in constitutive models available in PLAXIS 3D. The model performance was validated through: (i) previous field test data, (ii) the bespoke FEA undertaken in PISA and ALPACA projects. The results indicated that the adopted constitutive models can capture in relatively good accuracy the pile's monotonic lateral behaviour.

Soil reaction curves extraction

A bespoke code was developed to extract the soil reaction curves from the PLAXIS 3D analyses. The procedure followed was the same as employed in the original PISA project and adopted later in the PLAXIS Monopile Designer software (Panagoulias et al., 2018). PLAXIS Monopile Designer software is a very useful tool to extract the numerical soil reaction curves for offshore pile geometries. However, it is not suitable for the smaller geometry field test piles, since it provides approximately one soil reaction curve per 1m.

Implementation of the machine learning predictions in the PISA 1D FE framework

New 1D models were developed in the OxPile software. These 1D models are suitable for clay, sand and chalk profiles. They use the PCHIP spline interpolation scheme in MATLAB to calculate the forces values in the entire displacement domain, generating a suitable approach to accurately reproduce the soil reaction curves of the knot point parametrised soil reaction curves and calculate the pile's lateral response. The PCHIP interpolation together with the eight-knot points provide a much richer modelling framework of the soil reaction curves than the PISA conic function. The disadvantage of using more parameters to define the soil reaction curves is conveniently dealt with the machine leaning models employed. The selected eight knot points (e.g. predefined displacements values) were found to provide a very good fit to the soil reaction curves. Additional knot points can be employed (e.g. ten-knot points introducing $\bar{y}_9$ at $\bar{x} = 0.15$ and $\bar{y}_{10}$ at $\bar{x} = 0.2$ in Table 4.4) if the designer wants to generate a 1D model to predict the pile response at larger displacements ($v_G = 0.2D$).

Data-driven 1D design model for clay soils with arbitrary strength and stiffness profiles

A new general clay 1D model was developed for offshore glacial clay till sites with arbitrary stress history varying from a normally consolidated clay to a highly overconsolidated clay. This type of generalised model cannot be developed using the PISA design model where a

site-specific calibration is required for every different clay soil profile. The general clay 1D model was evaluated through various design cases concerned with pile geometries and clay profiles falling within or outside the calibration space. Finally, it was shown that the 1D model (calibrated for offshore pile geometries) can successfully scale down its predictions and calculate accurately the response of the smaller diameter PISA Cowden field test piles.

Layered soil 1D models

The original PISA concept for layered soils was to use the PISA design model to calibrate 1D models for homogenous soil profiles and then to employ these 1D models for calculations concerned with layered soils. This approach was found to work well for several layering configurations. However, it is not suitable for all layering configurations (e.g. it provides bad results when a soft clay is adjacent to a dense sand). In these cases, there is no established method on how to incorporate the interactions between adjacent layers in the developed 1D models. In this thesis a new approach was proposed – to generate data-driven layered soil models calibrated by soil reaction curves that are extracted from analyses performed in layered soils. The data-driven layered soil models can account for spatial variation of the thickness and the soil properties of each soil layer across the site. This approach is applicable to all layered soil profiles regardless the number of soil layers and does not have any restrictions in the context of the soil strength and stiffness employed for each soil layer.

The data-driven layered soil models become more practical for design when a site-specific calibration for each layering configuration (site-specific layered soil models) is employed. In this approach, various layering configurations are identified across the OWF and the wind turbine sites are grouped into these areas. Relatively few calibration analyses are required to form the 1D model for each layering configuration depending on: (i) the number of piles forming the calibration space, (ii) the variation of the thickness of each layer,

(iii) the variation of soil properties (i.e. strength and stiffness) of each layer. Each 1D model can be easily extended with a small number of additional calibration analyses to account for extra thin soil layers that may be found embedded within the soil matrix.

Chalk 1D model

The additional variables introduced in the problem due to de-structured chalk zones make the PISA design model unsuitable to generate a 1D model for chalk. A new framework to develop a chalk 1D model using the data-driven method is presented. The developed chalk 1D model is capable of capturing the different remoulded and fractured zone thickness generated by the pile's installation effect. The new chalk 1D model was tested and validated through various design examples including the ALPACA, ALPACA+ and Wind Support field test piles embedded in the St. Nicholas site. The method is readily extended: (i) to full-scale monopile configurations, (ii) to create a general chalk 1D model that will consider a range of chalk profiles of different unconfined compression strength and stiffness.

9.3 Integrating data-driven method in design procedure

One of the biggest advantages of the PISA design models is that once the design engineer acquires the PISA 1D FE framework (e.g. OxPile software), they can easily code in the calibrated 1D models by simply adding the parameters of the conic function for each calibrated 1D model. However, this task is not as straightforward using the data-driven method due to the black-box nature of the machine learning models. This section provides insight into the fundamental steps involved in applying the data-driven method in design. The devised methodology for generating a data-driven 1D design model is illustrated in Figure 9.1. Furthermore, an example of a snippet code employing the GPR model to train a data-driven 1D design model is presented in Appendix B.

Store the data: A well-structured and readable approach is needed to store the data. In this project, the knot point data along with the training inputs are stored in SQLite format (e.g. use SQLite (DB4S) open source tool) – other storage protocols could be also used. Table 9.1 lists an example form of the SQLite database used to generate a general clay 1D model.

Train the ML models: The SQLite database is loaded using a Python script (e.g. use Python Jupyter notebook to run Python code). The Scikit-learn Python library is used to: (i) split the data to train and test data, (ii) train the GPR models with the selected kernel (e.g. Matérn), (iii) use the MSE and R^2 regression metrics to assess the performance of the GPR models. A GPR model (32 in total) is trained for each knot point parameter for $\bar{p}$, $\bar{m}$, $\bar{H}_B$, $\bar{M}_B$. All the trained models are saved using the Joblib library. The joblib file can be distributed to other design engineers or can be used later to predict the knot point parameters from the trained GPR models.

Table 9.1: Example format of the SQLite database and CSV file used to generate a general clay 1D model. In the first column 1 is used to define the training inputs and knot points for p, m and 2 is used to define the training inputs and knot points for H_B, M_B. In the last columns $f_1 \dots f_8$ is used for the knot point parameters of p, H_B and $m_1 \dots m_8$ for the knot point parameters of m, M_B.

	Training inputs (Soil - Pile geometry characteristics)				Output (Knot point data)			
	z/L	L/D	G_o/s_u	σ'_v/s_u	$f_1 \dots$	f_8	$m_1 \dots$	m_8
1 or 2	…	…	…	…	…	…	…	…

Obtain the ML models predictions: The joblib file including all the information about the trained ML models (e.g. for the GCCM) is loaded using a python script. The ML models are used to predict the knot point parameters for the selected engineering inputs (e.g. z/L, L/D, G_o/s_u, σ'_v/s_u); these inputs can be inserted in the Python script using an excel or a text file format. The GPR models' predictions (e.g. the depth variation of $\bar{p}, \bar{m}$ knot points along the pile and the point predictions of $\bar{H}_B$, $\bar{M}_B$ knot points) can be plotted using a python script to check the confidence interval of the GPR predictions (i.e. estimating the uncertainty).

Subsequently, the GPR predictions (knot points) are saved in a CSV file. The format of the CSV file for a general clay 1D model is listed in Table 9.1.

Import ML results into OxPile: The CSV file imports the knot point predictions in OxPile. 1D models developed in OxPile (suitable for clay, sand and chalk profiles) use the PCHIP interpolation scheme to reproduce the soil reaction curves in the entire displacement domain. OxPile calculates the pile's backbone curve and induced bending moments (similar to the PISA design models).

Checking process: The user should go over a checking process to validate a data-driven 1D design model: (i) for a set of trained inputs – plot the depth variation of the knot points predicted by the GPR models with the training data to see if the ML model describes well the data, (ii) for a set of unseen inputs – read the joblib file, export the GPR predictions in a CSV file and use OxPile to calculate the pile response. Then, conduct some independent 3D calculations to compare and validate the data-driven 1D design model.

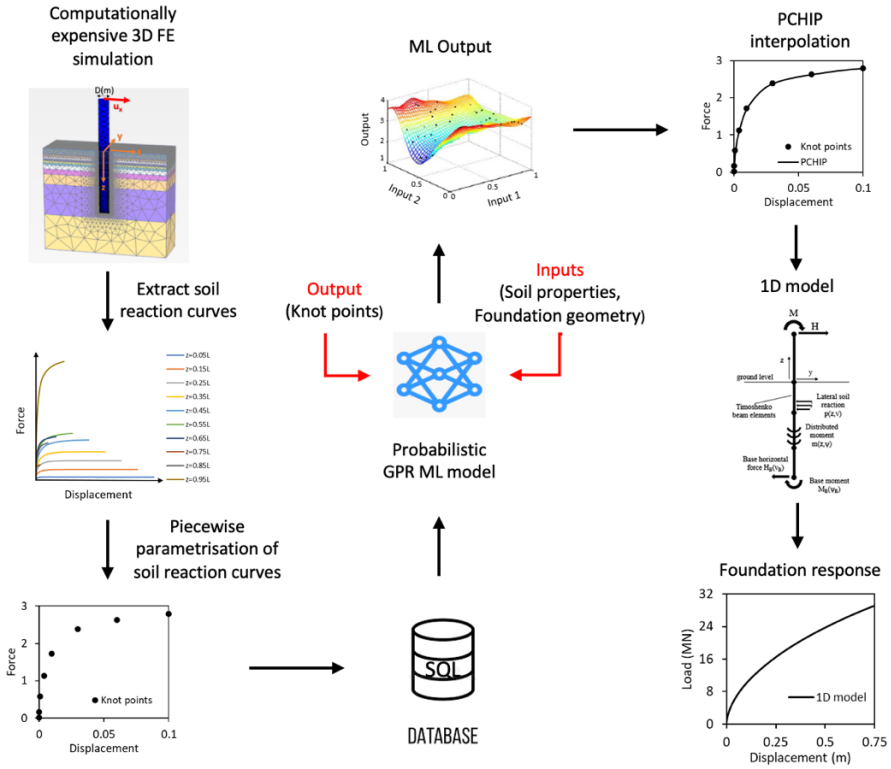


Figure 9.1: Devised methodology for generating a data-driven 1D design model.

9.4 Future Work

Data-driven chalk model

The data-driven chalk model presented in this thesis has limited applicability to offshore design cases. Further work needs to be carried to develop a general chalk 1D model for chalk profiles and pile geometries related to offshore conditions. The same methodology as presented in Chapter 8 can be used for this scope.

Layered soil profiles involving chalk layers

The current formation of the data-driven layered soil models was developed for layered soils that include clay or sand layers. However, in some offshore sites all clay, sand and chalk layers may occur. Therefore, the data-driven layered soil models need to be extended to account for possible chalk layers. Separately the accuracy of the independent layer method should be re-evaluated in the case of layered soils involving chalk layers.

Confidence of the data-driven 1D design models' predictions

In general, it is not advised to use the data-driven 1D design models without a validation 3D FEA for calculations outside the 1D model's calibration space. However, as it was shown through various design examples, the data-driven 1D design models have the ability to provide reliable predictions for design cases falling outside the calibration space. This generates questions about to what extent should a data-driven 1D design model be used for calculations outside the 1D model's calibration space. An advantage of the GPR model is that it provides the standard deviation of their mean predictions quantifying in some extent the uncertainty of the predicted soil reaction curves. However, a more straightforward reliability measure should be developed to estimate more precisely the overall pile's load-displacement performance uncertainty.

Compatibility of the data-driven 1D design models with design approaches corresponding to cyclic loading such as HARM

The data-driven 1D design models can provide calculations appropriate for the ULS and FLS design (monotonic loading conditions). The design of the OWT monopiles is also governed by the SLS design (cyclic loading conditions). During the PICASO project, the HARM model (Houlsby et al., 2017; Abadie et al., 2019; Abadie et al., 2023) has been promoted for complex cyclic loading calculations. The HARM can be calibrated using the pile's backbone curve or the monotonic soil reaction curves. A series of design analyses are required to be performed to ensure that the produced results of the data-driven method can be easily employed by HARM.

9.5 Summary

This thesis brings insights into using artificial intelligence to generate high fidelity data-driven 1D design models to rapidly calculate the OWT pile's backbone curve and the distribution of bending moment along the pile. The developed data-driven 1D design models are successfully applied to monopiles embedded in various soil profiles found in the North Sea involving clay, sand and chalk layers or layered soils with spatial variability of the layer thickness and the soil profile properties (e.g. strength and stiffness). A relatively small number of calibration analyses is required to train and develop the data-driven 1D design models. Overall, this project forms a part of a wider effort by Oxford University to enhance the PISA project outcomes through the PICASO project and to improve the design procedure employed for the monopiles used as foundation to offshore wind turbines.

References

Abadie, C.N., Houlsby, G.T. and Byrne, B.W. (2019). A method for calibration of the Hyperplastic Accelerated Ratcheting Model (HARM). *Computers and Geotechnics*, 112, pp.370-385.

Abadie, C.N., Beuckelaers, W.J., Byrne, B.W., Houlsby, G.T., Burd, H.J. and McAdam, R. (2023). Modeling Lifetime Performance of Monopile Foundations for Offshore Wind Applications. *Journal of Geotechnical and Geoenvironmental Engineering*, 149(8), p.04023051.

Abdel-Rahman, K. and Achmus, M. (2005). Finite element modelling of horizontally loaded monopile foundations for offshore wind energy converters in Germany. DOI: 10.1201/NOE0415390637.ch38.

Achmus, M., Thieken, K., Saathoff, J.E., Terceros, M. and Albiker, J. (2019). Un-and reloading stiffness of monopile foundations in sand. *Applied Ocean Research*, 84, pp.62-73.

Aghakouchak, A. (2015). *Advanced laboratory studies to explore theaxial cyclic behaviour of driven piles*. PhD thesis, Imperial College London, London, UK.

Alvarez-Borges, F., Madhusudhan, B.N. and Richards, D.J. (2020). Mechanical behaviour of low–medium density destructured White Chalk. *Géotechnique Letters*, 10(2), pp.360-366.

API (2011). *Geotechnical and Foundation Design Considerations.* IS0 19901-4:2003. Washington.

Arany, L., Bhattacharya, S., Macdonald, J. and Hogan, S.J. (2017). Design of monopiles for offshore wind turbines in 10 steps. *Soil Dynamics and Earthquake Engineering*, 92, pp.126-152.

AWG (Academic Work Group) (2016). *Confidential report by the PISA Academic Work Group to the PISA industrial partners*, Rev. E, Document No. 2595138, PISA Final Report. Oxford, UK.

AWG (Academic Work Group) (2018). *Confidential report by the PISA2 Academic Work Group to the PISA2 industrial partners*, Rev. E, Document No. 3107410, PISA2 Final Report. Oxford, UK.

Bauer, M., van der Wilk, M. and Rasmussen, C.E. (2016). Understanding probabilistic sparse Gaussian process approximations. *Advances in neural information processing systems*, 29.

Baghbani, A., Choudhury, T., Costa, S. and Reiner, J. (2022). Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Science Reviews*, *228*, p.103991.

Been, K. and Jefferies, M.G. (1985). A state parameter for sands. *Géotechnique*, 35(2), pp.99-112.

Beuckelaers, W.J.A.P. (2017). *Numerical modelling of laterally loaded piles for offshore wind turbines*. DPhil thesis, University of Oxford.

Beuckelaers, W.J.A.P., Burd, H., Houlsby, G., McAdam, R. and Byrne, B. (2016). A generalised Winkler model for the hysteretic and ratcheting behaviour of monopiles in clay and sand. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics.* ISFOG 2021b.

Benz, T. (2007). *Small-Strain Stiffness of Soils and It`s Numerical Consequences*. PhD Thesis, University of Stuttgart—Institute of Geotechnical Engineering, Stuttgart, Germany.

Bialowas, G.A. and Diambra, A. (2018). Time and stress dependent strength and stiffness of reconstituted chalk. *In Engineering in Chalk: Proceedings of the Chalk 2018 Conference* (pp. 503-508). ICE Publishing.

Bialowas, G.A., Diambra, A. and Nash, D.F. (2018). Stress and time-dependent properties of crushed chalk. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 171(6), pp.530-544.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Bolton, M.D. (1986). The strength and dilatancy of sands. *Geotechnique*, 36(1), pp.65-78.

Boulton, G. and Dobbie, K. (1993). Consolidation of sediments by glaciers: Relations between sediment geotechnics, soft-bed glacier dynamics and subglacial ground-water flow. *Journal of Glaciology*, 39(131), 26-44. doi:10.3189/S0022143000015690.

Bowden, A.J., Spink, T.W. and Mortimore, R.N. (2002). The engineering description of chalk: its strength, hardness and density. *Quarterly Journal of Engineering Geology and Hydrogeology*, 35(4), pp.355-361.

Brinkgreve, R.B.J., Engin, E. and Engin, H.K. (2010). Validation of empirical formulas to derive model parameters for sands. *Numerical methods in geotechnical engineering*, 137, p.142.

Brinkgreve, R.B.J., Kumarswamy, S. and Swolfs, W.M. (2018). *Plaxis manual 2018*. Plaxis bv Delft, the Netherlands.

Brosse, A., Kamal, R. H., Jardine, R. J. and Coop, M. R. (2017). The shear stiffness characteristics of four Eocene- to-Jurassic UK stiff clays. *Géotechnique*, 67 (3), pp. 242–259.

Buckley, R.M., Jardine, R.J., Kontoe, S., Parker, D. and Schroeder, F.C. (2018). Ageing and cyclic behaviour of axially loaded piles driven in chalk. *Géotechnique*, 68(2), pp.146-161.

Buckley, R.M., McAdam, R.A., Byrne, B.W., Doherty, J.P., Jardine, R.J., Kontoe, S. and Randolph, M.F. (2020a). Optimization of impact pile driving using optical fiber Bragg-grating measurements. *Journal of Geotechnical and Geoenvironmental Engineering*, 146(9), p.04020082.

Buckley R., Byrne B., McAdam R. and Schranz F. (2020b). Overview of ALPACA progress & update on ALPACA+. Oxford, UK, University of Oxford.

Buckley, R.M., Byrne, B.W., Martin, S.C., McAdam, R.A., Sheil, B.B., Aghakouchak, A. and Lindeboom, R. (2021). Large diameter pile testing for offshore wind applications with a focus on cyclic lateral loading and rate effects. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics.* ISFOG (pp. 1403-1412).

Bui, T.D., Moayedi, H., Gör, M., Jaafari, A., and Foong, L.K. (2019). Predicting slope stability failure through machine learning paradigms. *ISPRS International Journal of Geo-Information* 8, no. 9 (2019): 395.

Bundy, S.P. (2013). *Geotechnical properties of chalk putties*. PhD Thesis, University of Portsmouth.

Burd, H.J., Taborda, D.M.G., Zdravkovic, L., Abadie, C.N., Byrne, B.W., Houlsby, G.T., Gavin, K.G., Igoe, D.J.P., Jardine, R.J., Martin, C.M., McAdam, R. A., Pedro, A.M.G. and Potts, D. (2020a). PISA design model for monopiles for offshore wind turbines: application to a marine sand. *Géotechnique* 70, No. 11, 1048–1066.

Burd, H.J., Abadie, C.N., Byrne, B.W., Houlsby, G.T., Martin, C.M., McAdam, R.A., Jardine, R.J., Pedro, A.M., Potts, D.M., Taborda, D.M. and Zdravković, L. (2020b). Application of the PISA design model to monopiles embedded in layered soils. *Géotechnique*, 70(11), pp.1067-1082.

Burt, D.R., Rasmussen, C.E. and Van Der Wilk, M. (2020). Convergence of sparse variational inference in Gaussian processes regression. *The Journal of Machine Learning Research*, 21(1), pp.5120-5182.

Byrne, B.W., Burd, H.J., Zdravkovic, L., Abadie, C.N., Houlsby, G.T., Jardine, R.J., Martin, C.M., McAdam, R.A., Pacheco Andrade, M., Pedro, A.M. and Potts, D.M. (2019). PISA design methods for offshore wind turbine monopiles. *Offshore Technology Conference* (p. D011S005R002).

Byrne, B. W., McAdam, R. A., Burd, H.J., Beuckelaers, W. J. A., Gavin, K. G., Houlsby, G.T., Igoe, D. J. P., Jardine, R. J., Martin, C.M., Wood, A. M., Potts, D. M., Gretlund, J. S., Taborda, D. M. G., and Zdravković L. (2020a). Monotonic laterally loaded pile testing in a stiff glacial clay till at Cowden. *Géotechnique* 70, No. 11, 970–985.

Byrne, B. W., Houlsby, G. T., Burd, H. J., Gavin, K. G., Igoe, D. J. P., Jardine, R. J., Martin, C. M., McAdam, R. A., Potts, D. M., Taborda, D. M. G. and Zdravkovic, L. (2020b). PISA design model

for monopiles for offshore wind turbines: Application to a stiff glacial clay till. *Géotechnique* 70, No. 11, 1030–1047.

Byrne, B. W. (2021a). Lateral Pile Design for Offshore Wind Turbines. In Piling 2020: Proceedings of the Piling 2020 Conference (pp. 13-33). ICE Publishing.

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. (2021b). 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 2021b.

Byrne, B.W. (2021c). Documentation provided for predictions of the PICASO tests. Confidential Report.

Carotenuto, P., Meyer, V.M., Strøm, P.J., Cabarkapa, Z., St John, H. and Jardine, R. (2018). Installation and axial capacity of the Sheringham Shoal offshore wind farm monopiles–a case history. In *Engineering in chalk: proceedings of the chalk 2018 conference* (pp. 117-122). ICE Publishing.

Castillo-Fuentes, Zdravković, L. and Potts, D. M. (2023). On the calibration and application of the NorSand model. *10th European Conference on Numerical Methods in Geotechnical Engineering*. NUMGE 2023**.**

Ciavaglia, F., Carey, J. and Diambra, A., (2017a). Time-dependent uplift capacity of driven piles in low to medium density chalk. *Géotechnique Letters*, 7(1), pp.90-96.

Ciavaglia, F., Carey, J. and Diambra, A., (2017b). Monotonic and cyclic lateral tests on driven piles in Chalk. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 170(4), pp.353-366.

Clarke, B. G. (2018). The engineering properties of glacial tills. *Geotechnical Research*, 5(4), 262-277.

Cox, W.R., Reese, L.C. and Grubbs, B.R. (1974). Field testing of laterally loaded piles in sand. In *Offshore Technology Conference*. OnePetro.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4), 303-314.

Diaz, E. and Spagnoli, G. (2024). A super-learner machine learning model for a global prediction of compression index in clays. *Applied Clay Science*, *249*, p.107239.

Doughty, L.J., Buckley, R.M. and Jardine, R.J. (2018). Investigating the effect of ageing on the behaviour of chalk putty. In *Engineering in chalk: proceedings of the chalk 2018 conference* (pp. 695-701). ICE Publishing.

DNVGL (2018). *Support structures for wind turbines*. Tech. rep. DNVGL-ST-0126. Oslo: Det Norske Veritas.

Doherty, P. and Gavin, K. (2012). Laterally loaded monopile design for offshore wind farms. *Proceedings of the institution of civil engineers-energy*, 165(1), pp.7-17.

Duroska, P., Avgerinos, V., Andrade, M.P., Andersen, K.W. and Lingaard, M.W. (2020). Comparison of predicted pile load-displacement response using PISA API p-y to static laterally loaded onshore test piles in sand at Cuxhaven. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics* (ISFOG 2021).

Ebid, A.M., 2021. 35 Years of (AI) in geotechnical engineering: state of the art. Geotechnical and Geological Engineering, 39(2), pp.637-690.

Ferrero, J. A., Ruiz López, A., Taborda, D. M. G. and Brasile, S. (2023). Applying the observational method to a deep braced excavation using an artificial neural network. *10th European Conference on Numerical Methods in Geotechnical Engineering*. Proceedings 10th NUMGE 2023**.**

Forrester, A., Sobester, A. and Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Fritsch, F.N. and Carlson, R.E. (1980). Monotone Piecewise Cubic Interpolation, SIAM. *Journal on Numerical Analysis*, 17, pp. 238-246.

Gareau, L. F., Molenkamp, F., Sharma, J. and Hegtermans, M. H. (2004). Engineering geology of glaciated soils. In *Advances in geotechnical engineering: The Skempton conference: Proceedings of a three day conference on advances in geotechnical engineering, organised by the Institution of Civil Engineers and held at the Royal Geographical Society*, London, UK, on 29–31 March 2004 (pp. 1280-1291). Thomas Telford Publishing.

Georgiadis, M. (1983). Development of p–y curves for layered soils. In *Geotechnical practice in offshore engineering* (ed. S. G. Wright), vol. 1, pp. 536–545. New York, NY, USA: American Society of Civil Engineers.

Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning* (Vol. 1). Cambridge: MIT press.

Grammatikopoulou, A., Schroeder, F. C., Pedone, G., Brosse, A., Sørensen, T., Taborda, D. M. G. and Potts, D. M. (2021). 3D finite element analysis of monopiles and its application in offshore wind farm design. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics.* ISFOG.

Grimstad, G., Andersen, L. and Jostad, H. P. (2012). NGI-ADP: Anisotropic shear strength model for clay. *International Journal for Numerical and Analytical Methods in Geomechanics*. 36(4):483 – 497. DOI: 10.1002/nag.1016

GWEC (2022). Global Offshore Wind Report 2022. https://gwec.net/global-offshore-wind-report-2022/.

Ritchie, H., Roser, M. and Rosado, P. (2020). CO_2 and Greenhouse Gas Emissions. Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/co2-and-greenhouse-gas-emissions' [Online Resource].

Hardin, B. O. (1978). The nature of stress–strain behaviour for soils. *In Earthquake engineering and soil dynamics – proceedings of the ASCE Geotechnical Engineering Division specialty conference*, June 19–21, Pasadena, CA, USA, vol. 1, pp. 3–90. New York, NY, USA: American Society of Civil Engineers.

Hartigan, J.A. and Wong, M.A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1), pp.100-108.

Heaton, J. (2008). *Introduction to neural networks with Java*. Heaton Research, Inc.

Hensman, J., Fusi, N. and Lawrence, N.D. (2013). Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*.

Hipp, R.D. (2020). *SQLite*, Available at: https://www.sqlite.org/index.html.

Houlsby, G.T. (2016). Interactions in offshore foundation design. *Géotechnique*, 66, pp.791-825.

Houlsby, G.T., Abadie, C.N., Beuckelaers, W.J.A.P. and Byrne, B.W. (2017). A model for nonlinear hysteretic and ratcheting behaviour. *International Journal of Solids and Structures*, 120, pp.67-80.

Hu, Q., Han, F., Prezzi, M., Salgado, R. and Zhao, M. (2022). Finite-Element Analysis of the Lateral Load Response of Monopiles in Layered Sand. *Journal of Geotechnical and Geoenvironmental Engineering*, 148(4), p.04022001.

Jaky, J. (1944). The coefficient of earth pressure at rest. J. Soc. Hung. Arch. Eng. 1944, 78, 355–388.

Jardine, R.J., Buckley, R.M., Kontoe, S., Barbosa, P. and Schroeder, F.C., (2018). Behaviour of piles driven in chalk. *In Proceedings of the conference "Engineering in Chalk"*, London, UK: Institution of Civil Engineers (ICE) Publishing, 33–51.

Jefferies, M. and Been, K. (2015). *Soil liquefaction: a critical state approach*. CRC press.

Jong, S.C., Ong, D.E.L. and Oh, E. (2021). State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. *Tunnelling and Underground Space Technology*, 113, p.103946.

Igoe D. J. P. and Jalilvand S. (2020). 3D finite element modelling of monopiles in sand validated against large scale field tests. ISFOG 2021. Texas.

Iwan W.D. (1966). A distributed-element model for hysteresis and its steady-state dynamic response.

Kallehave, D., Byrne, B.W., LeBlanc Thilsted, C. and Mikkelsen, K.K. (2015). Optimization of monopiles for offshore wind turbines. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2035), p.20140100.

Kamas, I., Burd, H. J., Byrne, B. W. and Suryasentana, S. K. (2023). Generation of a general Cowden clay model (GCCM) using a data-driven method'. *9th International SUT OSIG Conference 2023, Session 17 on 'Pile Design II',* London, United Kingdom, 12/09/23 - 14/09/23.

Kardani, N., Zhou, A., Nazem, M. and Shen, S.L. (2020). Estimation of bearing capacity of piles in cohesionless soil using optimised machine learning approaches. *Geotechnical and Geological Engineering*, 38(2), pp.2271-2291.

Le, T. M. H., Eiksund, G. R., Strøm, P. J., and Saue, M. (2014). Geological and geotechnical characterisation for offshore wind turbine foundations: A case study of the Sheringham Shoal wind farm. *Engineering geology*, 177, 40-53.

LeBlanc, C., Houlsby, G.T. and Byrne, B.W. (2010). Response of stiff piles in sand to long-term cyclic lateral loading. *Géotechnique*, 60(2), pp.79-90.

Li, X. S. & Wang, Y. (1998). Linear representation of steady-state line for sand. *J. Geotech. Geoenviron. Engng 124*, No. 12, 1215–1217.

Liu, T., Ushev, E. R., and Jardine, R. J. (2020a). Anisotropic stiffness and shear strength characteristics of a stiff glacial till. *Journal of Geotechnical and Geoenvir. Engineering*, 146(12).

Liu, H., Ong, Y.S., Shen, X. and Cai, J. (2020b). When Gaussian process meets big data: A review of scalable GPs. *IEEE transactions on neural networks and learning systems*, 31, pp.4405-4423.

Liu, H., Diambra, A., Abell, J.A. and Pisanò, F. (2020c). Memory-enhanced plasticity modeling of sand behavior under undrained cyclic loading. *Journal of Geotechnical and Geoenvironmental Engineering*, 146(11), p.04020122.

Liu, D., Lin, P., Zhao, C., and Qiu, J. (2021). Mapping horizontal displacement of soil nail walls using machine learning approaches. *Acta Geotechnica* 16, no. 12 (2021): 4027-4044.

Liu, H.Y. and Kaynia, A.M. (2022a). Monopile responses to monotonic and cyclic loading in undrained sand using 3D FE with SANISAND-MSu. *Water Science and Engineering*, 15(1), pp.69-77.

Liu, T., Ahmadi-Naghadeh, R., Vinck, K., Jardine, R.J., Kontoe, S., Buckley, R.M. and Byrne, B.W. (2022b). An experimental investigation into the behaviour of destructured chalk under cyclic loading. *Géotechnique*, pp.1-13.

Loukidis, D. and Salgado, R. (2009). Modeling sand response using two-surface plasticity. *Computers and Geotechnics*, 36(1-2), pp.166-186.

Lutenegger, A. J. (1995). Geotechnical behavior of overconsolidated surficial clay crusts. *Transportation research record*, 1479, 61.

Manceau, S., McLean, R., Sia, A. and Soares, M. (2019). Application of the findings of the PISA Joint Industry Project in the design of monopile foundations for a North Sea wind farm. In *Offshore Technology Conference*. OnePetro.

Martin, S. C., B. W. Byrne, R. M. Buckley, K. Wu, R. A. McAdam, Aghakouchak A., Turner S, Cameron A. and Lindeboom R. (2023). Medium-scale field testing of laterally loaded monopiles in glacial till at Cowden. *Paper presented at the 9th International SUT OSIG Conference 2023, Session 6 on 'Lateral Loaded Piles'*, London, United Kingdom, 12/09/23 - 14/09/23.

Matthews, M.C. and Clayton, C.R.I. (2004). Large diameter plate tests on weathered in-situ Chalk. *Quarterly journal of engineering geology and hydrogeology*, 37(1), pp.61-72.

Matthews, A.G.D.G., Van Der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., Le, P., Ghahramani, Z. and Hensman, J., 2017. GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40), pp.1-6.

Matlock, H. (1970). Correlation for design of laterally loaded piles in soft clay. In *Offshore technology conference*. OnePetro.

Mayall, R. O. (2019). *Monopile response to scour and scour protection*. DEng thesis, University of Oxford.

Mayne, P. W., and Kulhawy, F. H. (1982). Ko-OCR relationships in soil. *Journal of the Geotechnical Engineering Division*, 108(6), 851-872.

McAdam, R.A., Byrne, B.W., Houlsby, G.T., Beuckelaers, W.J., Burd, H.J., Gavin, K.G., Igoe, D.J., Jardine, R.J., Martin, C.M., Muir Wood, A. and Potts, D.M., (2020). Monotonic laterally loaded pile testing in a dense marine sand at Dunkirk. *Géotechnique*, 70(11), pp.986-998.

McGann, C.R., Arduino, P. and Mackenzie-Helnwein, P. (2011). Applicability of conventional py relations to the analysis of piles in laterally spreading soil. *Journal of Geotechnical and Geoenvironmental engineering*, 137(6), pp.557-567.

Minga, E. and Burd, H.J. (2019). *Validation of PLAXIS MoDeTo based on the Cowden Till PISA Field Tests*. Oxford University: Oxford, UK.

Minga, E., and Burd, H.J. (2019b). *Validation of the PLAXIS MoDeTo 1D Model for Dense Sand*. Oxford University: Oxford, UK.

Moler, C.B. (2004). *Numerical computing with MATLAB*. Society for Industrial and Applied Mathematics.

Momeni, E., Dowlatshahi, M.B., Omidinasab, F., Maizir, H. and Armaghani, D.J. (2020). Gaussian process regression technique to estimate the pile bearing capacity. *Arabian Journal for Science and Engineering*, 45, pp.8255-8267.

Mortimore, R.N. (2012). Making sense of Chalk: a total-rock approach to its Engineering Geology. *Geological Society of London*.

Murff, J. D. and Hamilton, J. M. (1993b). P-ultimate for undrained analysis of laterally loaded piles. *Journal of Geotechnical Engineering*. American Society of Civil Engineers, 119(1), pp. 91–107. doi: 10.1061/(ASCE)0733-9410(1993)119:1(91).

Murphy, G., Igoe, D., Doherty, P., and Gavin, K. (2019). 3D FEM approach for laterally loaded monopile design. *ELSEVIER Computers and Geotechnics*, Vol. 100, pp. 76–83.

Nguyen V.C., Keawsawasvong, S., Nguyen, D.K. and Lai, V.Q. (2023a). Machine learning regression approach for analysis of bearing capacity of conical foundations in heterogenous and anisotropic clays. *Neural Computing and Applications*, 35(5), pp.3955-3976.

Nguyen, D.K., Nguyen, T.P., Ngamkhanong, C., Keawsawasvong, S. and Nguyen, T.K. (2023b). Prediction of uplift resistance of circular anchors in anisotropic clays using MLR, ANN, and MARS. *Applied Ocean Research*, 136, p.103584.

Nguyen, T., Duy-Khuong, L., Shiau, J. and Nguyen P.D. (2024). Optimizing load-displacement prediction for bored piles with the 3mSOS algorithm and neural networks. *Ocean Engineering* 304 (2024): 117758.

Nikishkov (2010). Three-dimensional Isoparametric Elements. *In: Programming Finite Elements in Java™. Springer*, London. https://doi.org/10.1007/978-1-84882-972-5_12.

Obozinski, G., Taskar, B., and Jordan, M. (2006). *Multi-task feature selection*. Technical report, Department of Statistics, University of California, Berkeley.

O'Neill, M.W. and Murchison, J.M. (1983). *An evaluation of py relationships in sands*. University of Houston.

Page, A.M., Grimstad, G., Eiksund, G.R. and Jostad, H.P. (2018). A macro-element pile foundation model for integrated analyses of monopile-based offshore wind turbines. *Ocean Engineering*, 167, pp.23-35.

Panagoulias, S., Brinkgreve, R.B.J., and Zampich, L. (2018). *PLAXIS MoDeTo manual*.

Pedone, G., Kontoe, S., Zdravković, L., Jardine, R.J., Vinck, K. and Liu, T. (2023). Numerical modelling of laterally loaded piles driven in low-to-medium density fractured chalk. *Computers and Geotechnics*, 156, p.105252.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, pp.2825-2830.

Peralta, P., Ballard, J.C., Rattley, M. and Erbrich, C.E. (2017). Dynamic and cyclic pile-soil response curves for monopile design. *Proc. 8th Int. Conf. Offshore Site Inv. Geo.* London, pp. 1054-1061.

Phoon, K.K. and Zhang, W. (2023). Future of machine learning in geotechnics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 17, no. 1 (2023): 7-22.

Pirrone, A. I. and Taborda, D. M. G. (2023). An application of machine learning to the back analysis of monopile response. In *Offshore Site Investigation Geotechnics 9th International Conference Proceeding* (Vol. 2028, No. 2035, pp. 2028-2035). Society for Underwater Technology.

Potts, D. M. and Zdravković, L. T. (1999). *Finite element analysis in geotechnical engineering: Theory*. London: Thomas Telford.

Powell, J. J. M. and Butcher, A. P. (2003). Characterisation of a glacial clay till at Cowden, Humberside. In *Characterisation and engineering properties of natural soils*, pp. 983–1020. Lisse, the Netherlands: Swets and Zeitlinger.

Qiu Z., Byrne B. W. and Martin C. M. (2023). Multidirectional cyclic simple shear testing of an overconsolidated glacial till. *Paper presented at the 9th International SUT OSIG Conference 2023, Session 15 on 'Seabed Stabilisation & Advanced Laboratory Testing'*, London, United Kingdom, 12/09/23 - 14/09/23.

Quinonero-Candela, J. and Rasmussen, C.E. (2005). A unifying view of sparse approximate Gaussian process regression. *The Journal of Machine Learning Research*, *6*, pp.1939-1959.

Randolph, M. and Gourvenec, S. (2011). *Offshore geotechnical engineering*. CRC Press.

Randolph, M. F. (2018). Potential Damage to Steel Pipe Piles During Installation. *IPA News Letter* 3.1, pp. 3–10.

Rasmussen C.K., and Williams, C.E. (2006). *Gaussian processes for machine learning* (Vol. 2, No. 3, p. 4). Cambridge, MA: MIT press.

Reese, L.C., Cox, W.R. and Koop, F.D. (1975). Field testing and analysis of laterally loaded piles on stiff clay. In *Offshore technology conference* (pp. OTC-2312). OTC.

Reese, L.C., Allen, J.D. and Hargrove, J.Q. (1981). Laterally loaded piles in layered soils. In *Proceedings of the tenth international conference on soil mechanics and foundations engineering*, Stockholm (Vol. 2, pp. 819-822).

Reese, L.C. (1997). Analysis of laterally loaded piles in weak rock. *Journal of Geotechnical and Geoenvironmental engineering*, 123(11), pp.1010-1017.

Richards, I.A., Byrne, B.W. and Houlsby, G.T. (2020). Monopile rotation under complex cyclic lateral loading in sand. *Géotechnique*, 70(10), pp.916-930.

Rossi, S., Heinonen, M., Bonilla, E., Shen, Z. and Filippone, M. (2021). Sparse Gaussian processes revisited: Bayesian approaches to inducing-variable approximations. In *International Conference on Artificial Intelligence and Statistics* (pp. 1837-1845). PMLR.

Rudolph, C., Bienen, B. and Grabe, J., (2014). Effect of variation of the loading direction on the displacement accumulation of large-diameter piles under cyclic lateral loading in sand. *Canadian geotechnical journal*, 51(10), pp.1196-1206.

Samui, P. (2008). Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Computers and Geotechnics*, 35(3), pp.419-427.

Sauer, E.K., Egeland, A.K., Christiansen, E.A. (1993). Compression characteristics and index properties of tills and intertill clays in southern Saskatchewan. Canada, *Canadian Geotechnical Journal,* Vol. 30, pp. 257-275.

Schanz, T. (1999). Formulation and verification of the Hardening-Soil Model. *RBJ Brinkgreve, Beyond 2000 in Computational Geotechnics*, pp.281-290.

Schofield, A.N. and Wroth, P. (1968). *Critical state soil mechanics*. London: McGraw-hill.

Segurado, J. and LLorca, J. (2004). A new three-dimensional interface finite element to simulate fracture in composites. *International journal of solids and structures*, 41(11-12), pp.2977-2993.

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

Shuttle, D. and Jefferies, M. (2010). Norsand: description, calibration, validation and applications. *Retrieved June*, *10*, p.2019.

Siegl, K., Dührkop, J. and Spill, S. (2020). A design framework for monopiles in sand based on large-scale pile load tests and FE analysis.

Simonin, L.E., Houlsby, G.T. and Byrne, B.W. (2022). HySand: A New Constitutive Model for Sand Under Cyclic Loading. In *International Conference of the International Association for Computer Methods and Advances in Geomechanics* (pp. 142-149). Cham: Springer International Publishing.

Snelson, E. and Ghahramani, Z. (2006). Sparse Gaussian Processes using Pseudo-inputs. *In Advances in Neural Information Processing Systems* (NIPS) (pp. 1257-1264).

Stuyts, B. and Suryasentana, S. (2023). Applications of data science in offshore geotechnical engineering: state of practice and future perspectives. In *Session on 'AI applications' in 9th OSIG conference*.

Suryasentana, S.K. and Lehane, B.M. (2014). Numerical derivation of CPT-based p–y curves for piles in sand. *Géotechnique*, 64(3), pp.186-194.

Suryasentana, S. K., Burd, H. J. and Byrne, B. W. (2019). *Data-Driven Approach for Monopile Foundation Design: Preliminary Study*. Confidential report for the PICASO project WP1 to Ørsted.

Suryasentana, S. K., Burd, H. J., Byrne, B. W. (2020). *Data-Driven Approach for Monopile Foundation Design*. Confidential report for the PICASO project WP1 to Ørsted.

Suryasentana, S. K., Burd, H. J., Byrne, B. W. (2021). *Data-Driven Approach for Monopile Foundation Design*. Confidential report for the PICASO project WP1 to Ørsted..

Suryasentana, S. K., Burd, H. J., Byrne, B. W., Aghakouchak, A. and Sørensen, T. (2022). Comparison of Machine Learning Models in a Data-Driven Approach for Scalable and Adaptive Design of Laterally-Loaded Monopile Foundations. *ISFOG.*

Suryasentana, S. K. and Sheil, B.B. (2023). Demystifying the connections between Gaussian Process regression and kriging theories. In *Session on 'AI applications' in 9th OSIG conference*.

Taborda, D. M. G., Zdravkovic, L., Kontoe, S. and Potts, D. M. (2014). Computational study on the modification of a bounding surface plasticity model for sands. *Computers and Geotechnics*, Vol. 59, pp. 145–160.

Taborda, D. M. G., Potts, D. M. and Zdravkovic, L. (2016). On the assessment of energy dissipation through hysteresis in finite element analysis. *Computers and Geotechnics*, Vol. 71, pp. 180-194.

Taborda, D. M. G., Zdravkovic´, L., Potts, D. M., Burd, H. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Jardine, R. J., Liu, T., Martin, C. M. and McAdam, R. A. (2020). Finite element modelling of laterally loaded piles in a dense marine sand at Dunkirk. *Géotechnique*, https://doi.org/10.1680/jgeot.18.PISA.006.

Tejada, G.I. and Antolin, P. (2022). Use of machine learning for unraveling hidden correlations between particle size distributions and the mechanical behavior of granular materials. *Acta Geotechnica*, 17(4), pp.1443-1461.

Thieken, K., Achmus, M. and Lemke, K. (2015). A new static p-y approach for piles with arbitrary dimensions in sand. *geotechnik*, 38(4), pp.267-288.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), pp.267-288.

Titsias, M. (2009). Variational Learning of Inducing Variables in Sparse Gaussian Processes. In Proceedings of the 12th International Conference on Artificial Intelligence and Statistics (AISTATS) (pp. 567-574).

Tsiampousi, A., Zdravkovic, L. and Potts, D. M. (2013). A new Hvorslev surface for critical state type unsaturated and saturated constitutive models. *Computers and Geotechnics*. Vol 48.

United Nations (2015). *Framework Convention on Climate Change - Paris Agreement*. Tech. rep.

Van der Wilk, M. (2019). *Sparse Gaussian process approximations and applications.* Doctoral dissertation, University of Cambridge.

Van Eekelen, H. A. M. (1980). Isotropic yield surfaces in three dimensions for use in soil mechanics. *Int. J. Numer. Analyt. Methods Geomech*. 4, No. 1, 89–101.

Vinck, K., Liu, T., Jardine, R.J., Kontoe, S., Ahmadi-Naghadeh, R., Buckley, R.M., Byrne, B.W., Lawrence, J.A., McAdam, R.A. and Schranz, F. (2022). Advanced in situ and laboratory characterisation of the ALPACA chalk research site. *Géotechnique*, pp.1-15.

Wang, L., Lai, Y., Hong, Y. and Mašín, D. (2020). A unified lateral soil reaction model for monopiles in soft clay considering various length-to-diameter (L/D) ratios. *Ocean Engineering*, 212, p.107492.

Wang, H., Lehane, B.M., Bransby, M.F., Wang, L.Z., Hong, Y. and Askarinejad, A. (2023). Lateral behavior of monopiles in sand under monotonic loading: Insights and a new simple design model. *Ocean Engineering*, 277, p.114334.

Whyte, S.A., Burd, H.J., Martin, C.M., and Rattley, M.J. (2020). Formulation and implementation of a practical multi-surface soil plasticity model. *Computers and Geotechnics*. Vol. 117, January 2020, 103092.

Williams, C. and Seeger, M., (2000). Using the Nyström method to speed up kernel machines. *Advances in neural information processing systems*, 13.

Wood, A.M., Mackenzie, B., Burbury, D., Rattley, M., Clayton, C.R.I., Mygind, M., Andersen, K.W., Thilsted, C.L. and Liingaard, M.A., (2015). Design of large diameter monopiles in chalk atWestermost Rough offshore wind farm. In *Frontiers in Offshore Geotechnics III: Proceedings of the 3rd International Symposium on Frontiers in Offshore Geotechnics (ISFOG 2015)* (Vol. 1, pp. 723-728). Taylor & Francis Books Ltd.

Wu, X., Hu, Y., Li, Y., Yang, J., Duan, L., Wang, T., Adcock, T., Jiang, Z., Gao, Z., Lin, Z. and Borthwick, A. (2019). Foundations of offshore wind turbines: A review. *Renewable and Sustainable Energy Reviews*, 104, pp.379-393.

Yang, Z. and Jeremic, B. (2005). Study of soil layering effects on lateral loading behaviour of piles. *J. Geotech. Geoenviron. Engng (ASCE)* 131, No. 6, 762–770.

Yin, X., Wang, H., Pisanò, F., Gavin, K., Askarinejad, A. and Zhou, H., (2023). Deep learning-based design model for suction caissons on clay. *Ocean Engineering*, 286, p.115542.

Zdravkovic, L., Jardine, R. J., Taborda, D. M. G., Abadias, D., Burd, H. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Igoe, D. J. P., Liu, T., Martin, C. M., McAdam, R. A., Muir Wood, A., Potts, D. M., Skov Gretlund, J. and Ushev, E. (2020a). Ground charaterisation for PISA pile testing and analysis. *Géotechnique* 70, No.11, 999–1013.

Zdravković, L., Taborda, D. M. G., Potts, D. M., Abadias, D., Burd, H. J., Byrne, B. W., Gavin, K., Houlsby, G. T., Jardine, R. J., Martin, C. M., McAdam, R. A. and Ushev, E. (2020b). Finite element modelling of laterally loaded piles in a stiff glacial clay till at Cowden. *Géotechnique* 70, No. 11, 945–960.

Zdravković, L., Taborda, D. M. G. and Potts, D. M. (2021). Effect of interface conditions on the response of laterally loaded monopiles in sand. In *Frontiers in Offshore Geotechnics IV: Proceedings of the 4th International Symposium on Frontiers in Offshore Geotechnics.* ISFOG.

Zhang, Y., Andersen, K. H. and Tedesco, G. (2016). Ultimate bearing capacity of laterally loaded piles in clay ¬- Some practical considerations. *Marine Structures*, Vol. 50, pp. 260-275.

Zhang, Y. and Andersen, K.H., (2017). Scaling of lateral pile py response in clay from laboratory stress-strain curves. *Marine Structures*, 53, pp.124.

Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y. and Ding, X. (2021a). Application of deep learning algorithms in geotechnical engineering: a short critical review. *Artificial Intelligence Review*, pp.1-41.

Zhang, R., Wu, C., Goh, A.T., Böhlke, T. and Zhang, W. (2021b). Estimation of diaphragm wall deflections for deep braced excavation in anisotropic clays using ensemble learning. *Geoscience Frontiers*, 12(1), pp.365-373.

Appendix A

The following appendix includes the finite element analyses (FEA) performed for each trained Gaussian process regression model presented in this thesis. All the analyses were conducted using PLAXIS 3D software. Table A.1 and Table A.2 present the FEA used to generate the GCCM and GDSM data-driven 1D design models discussed in Chapter 5. Table A.3 to Table A.6 show the FEA used to develop the layered soil data-driven 1D design models described in Chapter 6. Table A.7 to Table A.9 illustrate the FEA conducted for the development of layered soil models intended for practical design, as detailed in Chapter 7. Finally, Table A.10 presents the FEA performed to develop the chalk model for the Kent St. Nicholas field tests, as explained in Chapter 8.

Figure A.1 to Figure A.5 illustrate the profiles of K_o, G_o, s_u and σ'_v for the layered soil sites used in the calibration FEA performed to develop the layered soil models in Chapter 6. Figure A.6 and Figure A.7 show the profiles of K_o, G_o, s_u and σ'_v for the layered soil sites used in the calibration FEA performed to develop the layered soil models intended for practical design, as detailed in Chapter 7. As previously mentioned in this thesis, the properties of the sand layers are based on the PISA representative offshore homogeneous sand site presented in Burd et al. (2020a). The properties of the glacial till soil are based on the PISA representative offshore glacial clay till site (Byrne et al., 2020b). The stress history of each clay layer is determined using the procedure presented in Section 5.2.1. The properties of the Bothkennar clay soil profile are the same as those of the PISA representative offshore Bothkennar clay site (AWG, 2018).

Table A.1: Finite element calibration analyses performed to develop the GCCM presented in Section 5.2. A total of 36 FEA were conducted.

Soil Profiles (refer to Figure 5.4)	**Piles** (refer to Table 5.1)
PISA representative offshore glacial clay till site NC representative offshore glacial clay till site	C1, C2, C4, C5, C6, C8, C9, C10, C11
heavily OC representative offshore glacial clay till site corresponding to $d_{ice} = 100m$.	C1, C2, C4, C5, C6, C8, C9, C10, C11
slightly OC representative offshore glacial clay till site corresponding to $d_{ice} = 25m$.	C1, C2, C4, C5, C6, C8, C9, C10, C11
NC representative offshore glacial clay till site	C1, C2, C4, C5, C6, C8, C9, C10, C11

Table A.2: Finite element calibration analyses performed to develop the GDSM V2 presented in Section 5.3. A total of 36 FEA were conducted.

Soil Profiles (refer to Figure 5.18)	**Piles** (refer to Table 5.1)
PISA representative offshore homogenous sand site with a relative density of $D_R = 45\%$	C1, C2, C4, C5, C6, C8, C9, C10, C11
PISA representative offshore homogenous sand site with a relative density of $D_R = 60\%$	C1, C2, C4, C5, C6, C8, C9, C10, C11
PISA representative offshore homogenous sand site with a relative density of $D_R = 75\%$	C1, C2, C4, C5, C6, C8, C9, C10, C11
PISA representative offshore homogenous sand site with a relative density of $D_R = 90\%$	C1, C2, C4, C5, C6, C8, C9, C10, C11

Table A.3: Finite element calibration analyses performed to develop the general two-layered soil model presented in Section 6.6.1. A total of 36 new FEA were conducted. Additionally, the general two-layered soil model was trained using soil reaction curves extracted from FEA conducted on reference homogeneous soil profiles, specifically Bothkennar clay and Dunkirk sand with $D_R = 90\%$. This contributed another 18 FEA. In total, knot point data from 54 FEA were used to train the GPR models.

Soil Profiles (refer to Figure 6.20)	**Top layer thickness**	**Piles** (refer to Table 6.5 & Table 5.1)
two-layered soil profile with a Bothkennar clay layer located over a Dunkirk sand layer with $D_R = 90\%$	$z_{L1} = 0.2L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
two-layered soil profile with a Bothkennar clay layer located over a Dunkirk sand layer with $D_R = 90\%$	$z_{L1} = 0.5L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
two-layered soil profile with a Bothkennar clay layer located over a Dunkirk sand layer with $D_R = 90\%$	$z_{L1} = 0.8L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
Bothkennar clay	$z_{L1} = 1L$	C1, C2, C4, C5, C6, C8, C9, C10, C11
Dunkirk sand with $D_R = 90\%$	$z_{L1} = 1L$	C1, C2, C4, C5, C6, C8, C9, C10, C11

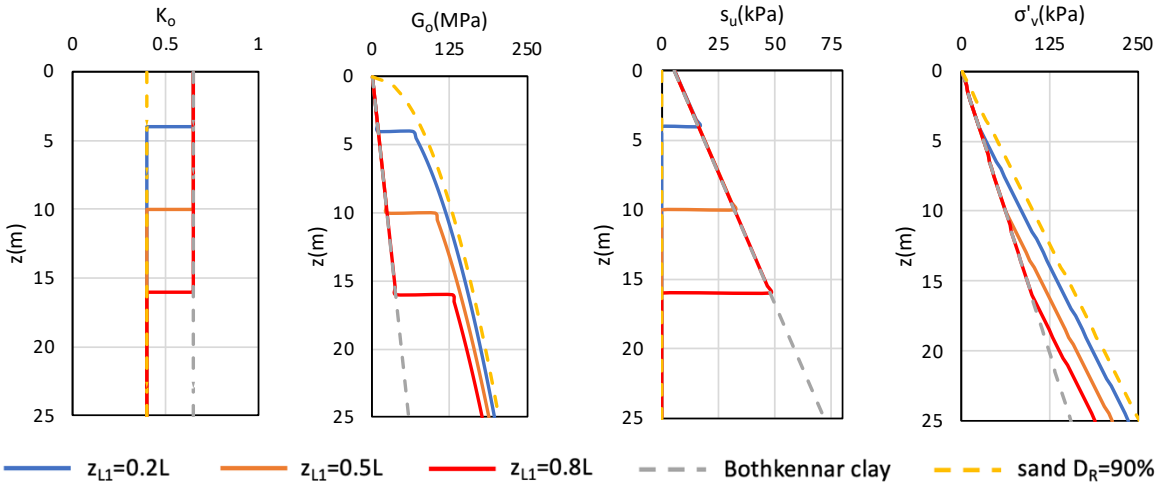


Figure A.1: Example profiles of K_o, G_o, s_u and σ'_v used in the calibration FEA for pile LC1 (see Table 6.5). These profiles are utilised to develop the general two-layered soil model described in Section 6.6.1. The values of K_o, G_o, s_u and σ'_v are determined based on the z_{L1} value specific to each layered soil profile, and therefore vary with different pile geometries.

Table A.4: Finite element calibration analyses performed to develop the general three-layered soil model presented in Section 6.6.2. A total of 72 new FEA were conducted. Additionally, the general three-layered soil model was trained using soil reaction curves extracted from the FE calibration analyses presented in Table A.3. In total, knot point data from 126 FEA were used to train the GPR models.

Soil Profiles (refer to Figure 6.23)	**Top layer thickness** (Bothkennar clay)	**Embedded Layer thickness** (Dunkirk sand)	**Piles** (refer to Table 6.5)
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.2L$	$z_{L2} = 0.25L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.2L$	$z_{L2} = 0.1L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.6L$	$z_{L2} = 0.25L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.6L$	$z_{L2} = 0.1L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.2L$	$z_{L2} = 0.6L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12
embedded Dunkirk sand layer of $D_R = 90\%$ within a Bothkennar clay matrix	$z_{L1} = 0.3L$	$z_{L2} = 0.4L$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12

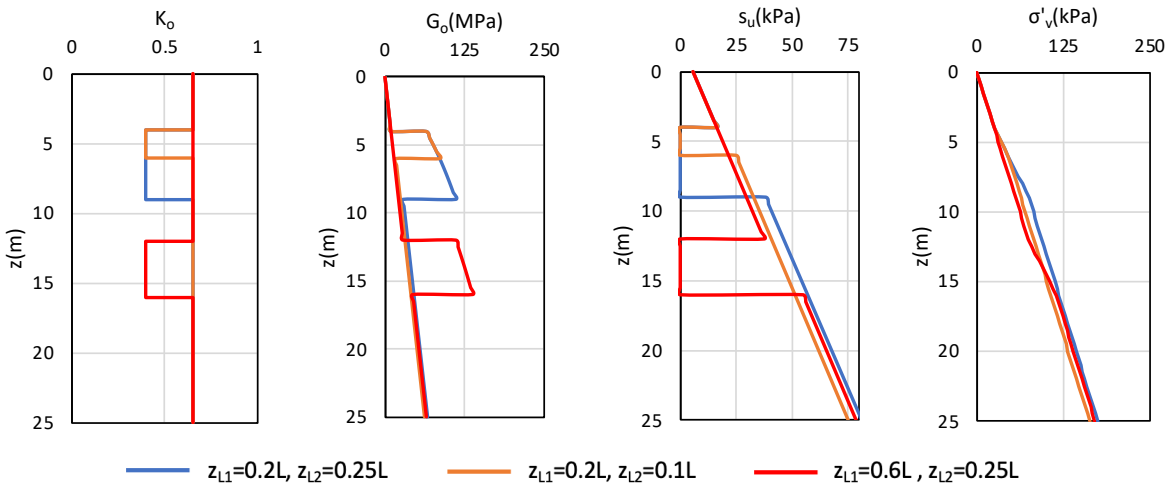


Figure A.2: Example profiles of K_o, G_o, s_u and σ'_v used in the calibration FEA for pile LC1 (see Table 6.5). These profiles are utilised to develop the general three-layered soil model described in Section 6.6.2. The values of K_o, G_o, s_u and σ'_v are determined based on the z_{L1} and z_{L2} values specific to each layered soil profile, and therefore vary with different pile geometries. The soil profiles illustrated in this figure correspond to the top three-layered soil profiles presented in Table A.4.

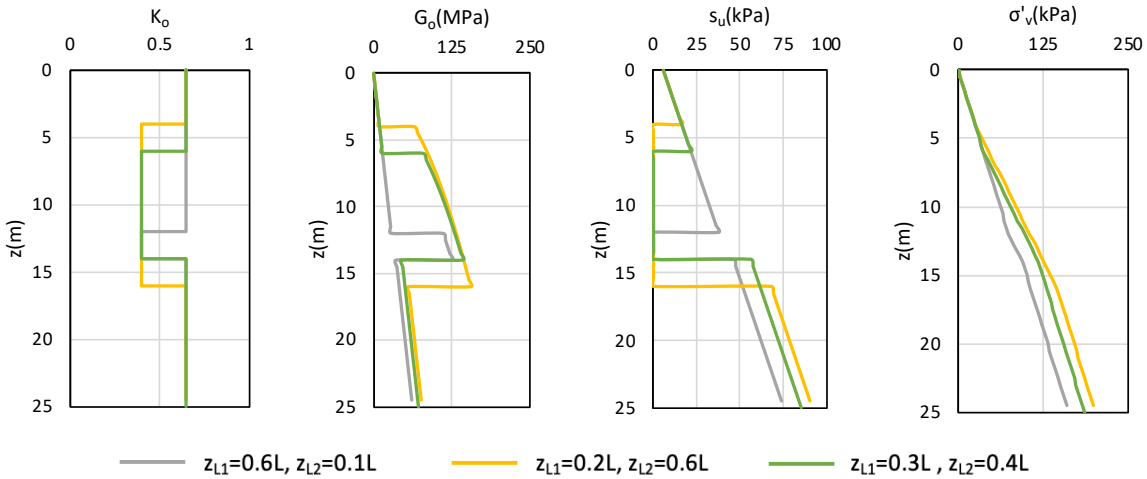


Figure A.3: Example profiles of K_o, G_o, s_u and σ'_v used in the calibration FEA for pile LC1 (see Table 6.5). These profiles are utilised to develop the general three-layered soil model described in Section 6.6.2. The values of K_o, G_o, s_u and σ'_v are determined based on the z_{L1} and z_{L2} values specific to each layered soil profile, and therefore vary with different pile geometries. The soil profiles illustrated in this figure correspond to the bottom three-layered soil profiles presented in Table A.4.

Table A.5: Finite element calibration analyses performed to develop the site-specific layered soil model at a single location presented in Section 6.7.1. A total of 10 FEA were conducted.

Soil Profiles (refer to Figure 6.26, Figure A.1)	**Piles** (refer to Table 6.8)
conceptual multi-layer system combining Bothkennar clay and Dunkirk sand with $D_R = 90\%$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10

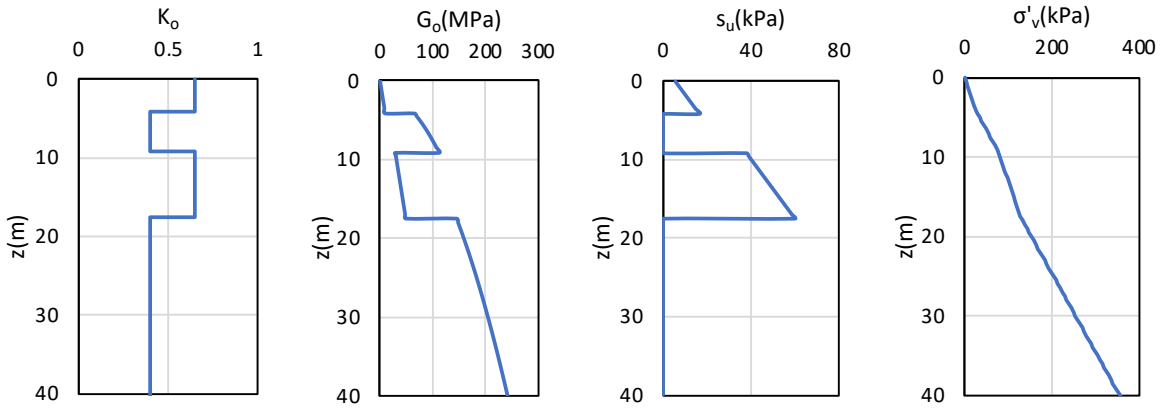


Figure A.4: Profiles of K_o, G_o, s_u and σ'_v adopted for the calibration FEA performed to develop the site-specific layered soil model at a single location, as presented in Section 6.7.1.

Table A.6: Finite element calibration analyses performed to develop the site-specific layered soil model accounting for spatial variations presented in Section 6.7.2. A total of 20 FEA were conducted.

Soil Profiles (refer to Figure 6.20)	**Layer thickness**	**Piles** (refer to Table 6.8)
conceptual multi-layer system combining Bothkennar clay and Dunkirk sand with $D_R = 90\%$	$z_{L1} = 3.2\text{m}$, $z_{L2} = 4\text{m}$, $z_{L3} = 6.8\text{m}$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10
conceptual multi-layer system combining Bothkennar clay and Dunkirk sand with $D_R = 90\%$	$z_{L1} = 4.8\text{m}$, $z_{L2} = 6\text{m}$, $z_{L3} = 10.2\text{m}$	LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10

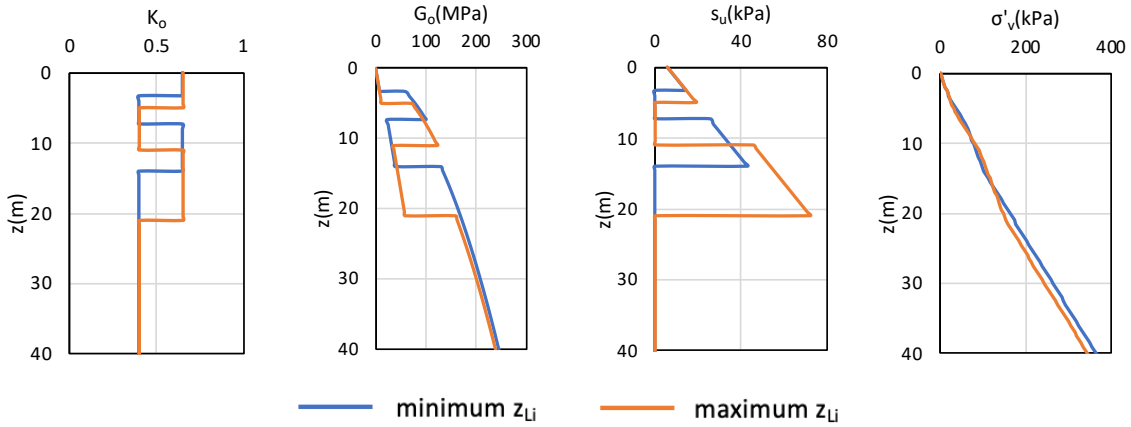


Figure A.5: Profiles of K_o, G_o, s_u and σ'_v adopted for the calibration FEA performed to develop the site-specific layered soil model accounting for spatial variations, as presented in Section 6.7.2.

Table A.7: Finite element calibration analyses performed to develop the layered soil model 'Model V1' for *Configuration I* presented in Section 7.5. A total of 16 FEA were conducted. In all cases, Layer 4 consists of the PISA representative Cowden clay soil profile.

Layered Configuration (refer to Figure 7.1(a))	**Strength and stiffness properties of each layer** (refer to Figure 7.2)	**Layer thickness**	**Piles** (refer to Table 7.1)
four-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$	$z_{L1} = 0.5\text{m}$ $z_{L2} = 4.5\text{m}$ $z_{L3} = 15\text{m}$	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$	$z_{L1} = 0.5\text{m}$ $z_{L2} = 4.5\text{m}$ $z_{L3} = 15\text{m}$	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$	$z_{L1} = 3.7\text{m}$ $z_{L2} = 12.7\text{m}$ $z_{L3} = 30.6\text{m}$	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$	$z_{L1} = 3.7\text{m}$ $z_{L2} = 12.7\text{m}$ $z_{L3} = 30.6\text{m}$	CP1, CP2, CP3, CP4

Table A.8: Finite element calibration analyses performed to develop the layered soil model 'Model V2' for *Configuration I* presented in Section 7.5. A total of 24 FEA were conducted. In all cases, Layer 4 consists of the PISA representative Cowden clay soil profile.

Layered Configuration (refer to Figure 7.1(a))	**Strength and stiffness properties of each layer** (refer to Figure 7.2)	**Layer thickness**	**Piles** (refer to Table 7.1)
four-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$	$z_{L1} = 0.5$m $z_{L2} = 4.5$m $z_{L3} = 15$m	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$	$z_{L1} = 0.5$m $z_{L2} = 4.5$m $z_{L3} = 15$m	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$	$z_{L1} = 2.1$m $z_{L2} = 8.6$m $z_{L3} = 22.8$m	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$	$z_{L1} = 2.1$m $z_{L2} = 8.6$m $z_{L3} = 22.8$m	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$	$z_{L1} = 3.7$m $z_{L2} = 12.7$m $z_{L3} = 30.6$m	CP1, CP2, CP3, CP4
four-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$	$z_{L1} = 3.7$m $z_{L2} = 12.7$m $z_{L3} = 30.6$m	CP1, CP2, CP3, CP4

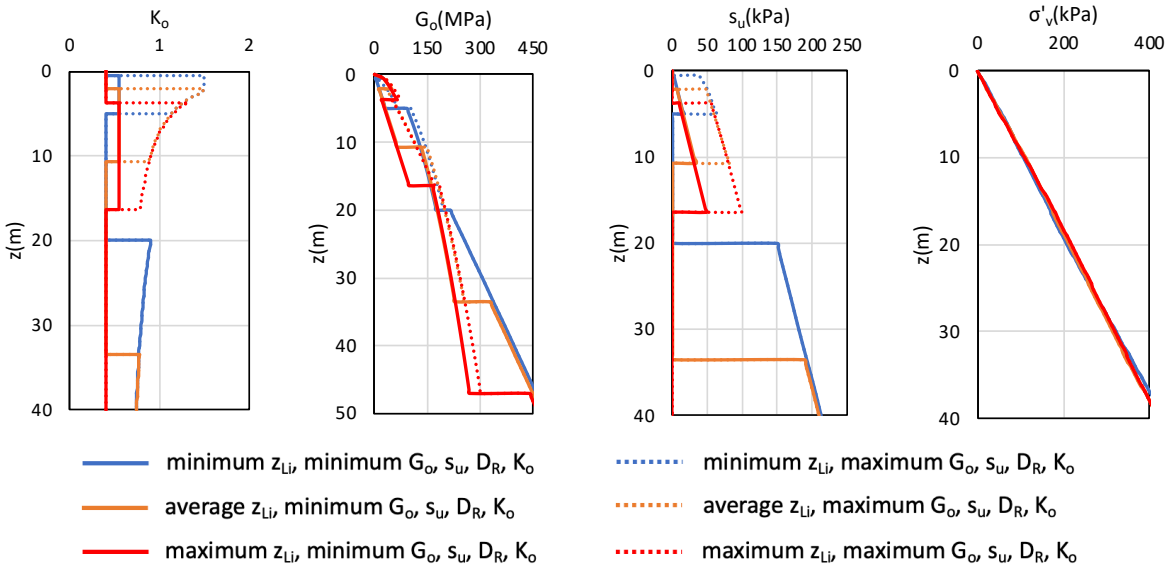


Figure A.6: Profiles of K_o, G_o, s_u and σ'_v adopted for the calibration FEA performed to develop the layered soil models 'Model V1' and 'Model V2' for *Configuration I*, as presented in Section 7.5.

Table A.9: Additional finite element calibration analyses performed to update the layered soil model 'Model V2' to account for *Configuration II* presented in Section 7.5. This contributed another 8 FEA, hence a total of 32 FEA were used to train the GPR models for *Configuration II*.

Layered Configuration (refer to Figure 7.1(b))	**Strength and stiffness properties of each layer** (refer to Figure 7.2)	**Layer thickness**	**Piles** (refer to Table 7.1)
six-layer soil profile	Layer 1: $D_R = 45\%$ Layer 2: $d_{ice} = 0m$ Layer 3: $D_R = 80\%$ Layer 4: $d_{ice} = 75m$ Layer 5: $D_R = 80\%$	$z_{L1} = 0.5$m $z_{L2} = 4.5$m $z_{L3} = 1.2$m $z_{L4} = 9$m $z_{L5} = 24.8$m	CP1, CP2, CP3, CP4
six-layer soil profile	Layer 1: $D_R = 75\%$ Layer 2: $d_{ice} = 25m$ Layer 3: $D_R = 100\%$ Layer 4: $d_{ice} = 100m$ Layer 5: $D_R = 100\%$	$z_{L1} = 0.5$m $z_{L2} = 4.5$m $z_{L3} = 1.2$m $z_{L4} = 9$m $z_{L5} = 24.8$m	CP1, CP2, CP3, CP4

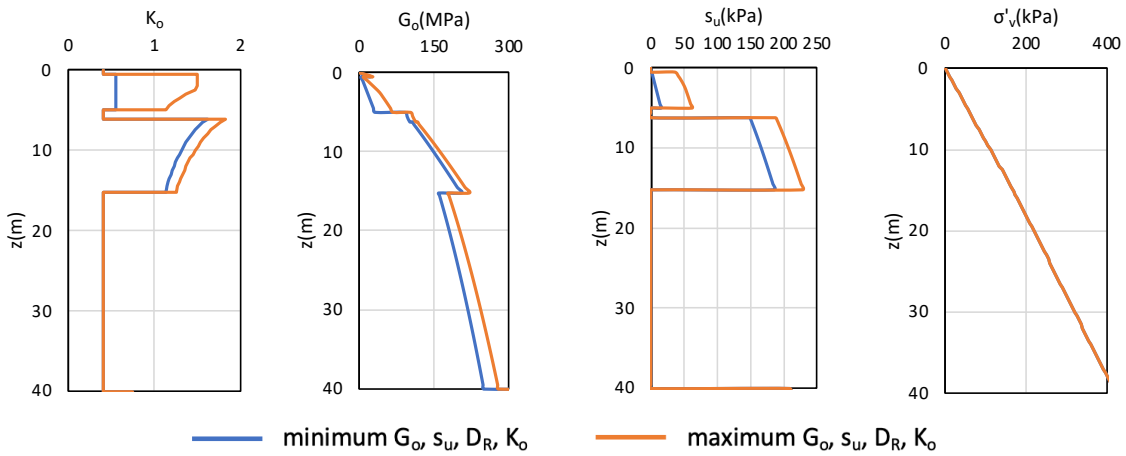


Figure A.7: Profiles of K_o, G_o, s_u and σ'_v adopted for the calibration FEA performed to update the layered soil model 'Model V2' to account for *Configuration II*, as presented in Section 7.5.

Table A.10: Finite element calibration analyses performed to develop the chalk model presented in Chapter 8. A total of 30 FEA were conducted.

Soil Profiles (refer to Figure 8.3)	**Piles** (refer to Table 8.5 & Table 8.10)
Kent St. Nicholas site	CST1, CST2, CST3, CST4, CST5, CST6, CST7, CST8, CST9, CST10, CST11, CST12, CST13, CST14, CST15, CST16, CST17, CST18, CST19, CST20, CST21, CST22, CST23, CST24, CST25, CST26, CST27, CST28, CST29, CST30

Appendix B

The following appendix includes an example of a snippet code employing the Gaussian process regression model for the knot points of the distributed lateral load reaction, $\bar{p}$, distributed moment reaction, $\bar{m}$, base horizontal force reaction, $\bar{H}_B$, and base moment reaction, $\bar{M}_B$, knot points to create the generalised Cowden clay model (GCCM) presented in Section 5.2. The code includes the development of an SQLite database containing all the required data for the GCCM machine learning model inputs and outputs. The knot point data are extracted and determined for each soil reaction curve following the procedure presented in Section 4.5. For simplicity, the following code demonstrates adding data (knot point data, soil profile, and pile geometry properties) from an Excel spreadsheet to the SQLite database, though various methods can be used to achieve this. The data in the code are categorized as relevant to the distributed lateral load reaction, $\bar{p}$, and distributed moment reaction, $\bar{m}$, ('SoilReactionType'=1), and relevant to the base soil reactions, $\bar{H}_B$ and $\bar{M}_B$ ('SoilReactionType'=2). In total, 36 GPR models are trained for each data-driven 1D design model. All the trained GPR models are saved in a joblib file. To use the trained models for prediction, one needs to load the saved models from the joblib file. By providing the necessary inputs (either seen or unseen), one can obtain the GPR model's predictions and the confidence intervals, as shown in the code. Finally, the GPR model predictions, along with the inputs used, are saved in a CSV file.

Code in jupyter notebook (python):

Create Sqlite database

```
import sqlite3
# Enter the path to SQLite database file here
database_name = 'path.db'
#create the table
create_table_sql = """
   CREATE TABLE IF NOT EXISTS soil_rection_data (
      pile_num INTEGER NOT NULL,
      length FLOAT, diameter FLOAT,
      z FLOAT, SoilReactionType INTEGER NOT NULL,
      sv FLOAT, p FLOAT, ko FLOAT, OCR FLOAT, Go FLOAT, su FLOAT,
      soil_unit TEXT, soil_type TEXT,
      f1 FLOAT, f2 FLOAT, f3 FLOAT, f4 FLOAT, f5 FLOAT, f6 FLOAT, f7 FLOAT, f8 FLOAT,
      m1 FLOAT, m2 FLOAT, m3 FLOAT, m4 FLOAT, m5 FLOAT, m6 FLOAT, m7 FLOAT, m8
FLOAT
   );
"""
conn = sqlite3.connect(database_name)
cursor = conn.cursor()
cursor.execute(create_table_sql)
conn.commit()
conn.close()
print("Table 'soil_properties' has been created successfully in", database_name)

#Insert data from an excel spreadsheet to Sqlite table
import sqlite3
import pandas as pd
# Enter the path to excel file here
excel_file = 'path.xlsx'
def read_excel_and_clean(excel_file):
   try:
      df = pd.read_excel(excel_file, header=None)
```

```
        df_cleaned = df.dropna(how='all').reset_index(drop=True)
        df_cleaned.columns = df_cleaned.iloc[0]
        df_cleaned = df_cleaned.drop(0)
        df_cleaned = df_cleaned.dropna(axis=1, how='all')
        df_cleaned = df_cleaned.reset_index(drop=True)
        return df_cleaned
    except Exception as e:
        print(f"Error reading Excel file: {e}")
        return None
if __name__ == "__main__":
    cleaned_df = read_excel_and_clean(excel_file)

# Enter the relative or absolute path to SQLite database file here
database_name = 'path.db'

# Function to insert data into the Soil_data table
def insert_data(data):
    try:
        # Connect to the SQLite database
        conn = sqlite3.connect(database_name)
        cursor = conn.cursor()
        insert_query = """
            INSERT INTO soil_rection_data (pile_num, length, diameter, z, SoilReactionType, sv, p,
                                           ko, OCR, Go, su, "soil_unit", "soil_type", f1, f2, f3,
                                           f4, f5, f6, f7, f8, m1, m2, m3, m4, m5, m6, m7, m8)
            VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)
        """
        data_tuples = [tuple(x) for x in data.to_numpy()]
        cursor.executemany(insert_query, data_tuples)
        conn.commit()
        print("Data inserted successfully into Soil_data table")
    except sqlite3.Error as error:
        print(f"Failed to insert data into SQLite table: {error}")
    finally:
```

```
        if conn:
            conn.close()
if __name__ == "__main__":
    # Insert data from DataFrame into the database
    insert_data(cleaned_df)
```

Load Data from Sqlite database

```
#load Sqlite table data to a dataframe
import sqlite3
import pandas as pd
def load_table_to_dataframe(database_name, table_name):
    try:
        # Connect to the SQLite database
        conn = sqlite3.connect(database_name)
        # Construct the SQL query to select all data from the table
        query = f"SELECT * FROM {table_name};"
        # Use pandas to read the SQL query into a DataFrame
        df = pd.read_sql_query(query, conn)
        print(f"Data from table '{table_name}' loaded successfully into DataFrame.")
        return df
    except sqlite3.Error as error:
        print(f"Error while loading data from table '{table_name}':", error)
        return None
    finally:
        if conn:
            conn.close()
# Enter the path to SQLite database file here
database_name = 'path.db'
table_name = 'soil_rection_data'
df = load_table_to_dataframe(database_name, table_name)

#Generate the synthetic features (inputs) from the soil properties and pile geometries
def norm_inputs(df):
```

```
    df['z_L'] = df['z']/df['length']
    df['L_D'] = df['length']/df['diameter']
    df.loc[:, 'su_sigv'] = df.loc[:, 'su']/df.loc[:, 'sv']
    df.loc[:, 'Ir'] = df.loc[:, 'Go']/df.loc[:, 'su']
    df.loc[:, 'Isig'] = df.loc[:, 'Go']/df.loc[:, 'sv']
    df.loc[:, 'Irsig'] = df['Ir'] # by default for clay
    df.loc[:, 'Sj'] = 1/df['su_sigv'] # by default for clay
    return df
def clean_training_data(df):
    dist = 0.01
    df = df.drop(df[(df.z_L < dist) & (df.SoilReactionType == 1)].index) # avoid data near surface
    return df
def preprocess(df):
    df = norm_inputs(df)
    # remove outliers in training data
    df = clean_training_data(df)
    return df
df = preprocess(df) # Process the data
# Print the status of the loaded data
print(f'Loaded {df.shape[0]} soil reactions training data')
```

Train the 36 GPR models and save them to a single joblib file

```
from sklearn.model_selection import train_test_split
from sklearn.gaussian_process import GaussianProcessRegressor
from sklearn.gaussian_process.kernels import Matern, WhiteKernel, ConstantKernel
from sklearn.metrics import mean_squared_error, r2_score
import joblib
import pandas as pd

# Define the target outputs
outputs = ['f1', 'f2', 'f3', 'f4', 'f5', 'f6', 'f7', 'f8',
           'm1', 'm2', 'm3', 'm4', 'm5', 'm6', 'm7', 'm8']
# Initialize a dictionary to store the results and models
```

```
results = {}
models = {}

# Function to calculate scores
def score(Y_true, Y_pred):
    mse = mean_squared_error(Y_true, Y_pred)
    r2 = r2_score(Y_true, Y_pred)
    return mse, r2

# Loop over each soil reaction type
for soilreactiontype in [1, 2]:
    print(f"Soil Reaction Type {soilreactiontype}:")
    # Define inputs based on soil reaction type
    if soilreactiontype == 1:
        inputs = ['z_L', 'L_D', 'Irsig', 'Sj']
        # Define the kernel with length scale bounds and constant term
        kernel = ConstantKernel(constant_value=1.0, constant_value_bounds=(1e-5, 1e3)) \
                * Matern(length_scale=[1.0] * len(inputs),
                         length_scale_bounds=[(1e-1, 1e2), (5e-1, 1e2), (50, 1e3), (0.1, 1e3)],
                         nu=2.5) \
                + WhiteKernel(noise_level=1e-5, noise_level_bounds=(1e-6, 1e4))
    elif soilreactiontype == 2:
        inputs = ['L_D', 'Irsig', 'Sj']
        # Define the kernel with length scale bounds and constant term
        kernel = ConstantKernel(constant_value=1.0, constant_value_bounds=(1e-5, 1e3)) \
                * Matern(length_scale=[1.0] * len(inputs),
                         length_scale_bounds=[(5e-1, 1e2), (50, 1e3), (0.1, 1e3)],
                         nu=2.5) \
                + WhiteKernel(noise_level=1e-5, noise_level_bounds=(1e-6, 1e4))
    else:
        raise ValueError(f"Unexpected soil reaction type: {soilreactiontype}")

    sdf = df[df.SoilReactionType == soilreactiontype]
    # Loop over each output variable
```

```
for output in outputs:
    print(f"Training model for {output}...")
    Y = sdf[output]
    X = sdf[inputs]
    Xtrain, Xtest, Ytrain, Ytest = train_test_split(X, Y, test_size=0.2, random_state=42)

    # Create the GaussianProcessRegressor object
    gpr = GaussianProcessRegressor(kernel=kernel, n_restarts_optimizer=10, normalize_y=False)
    # Fit to the training data
    gpr.fit(Xtrain, Ytrain)
    # Print the optimized kernel
    print(f"Optimized kernel for {output} (Soil Reaction Type {soilreactiontype}):", gpr.kernel_)

    # Evaluate metrics
    Ytrain_pred = gpr.predict(Xtrain)
    Ytest_pred = gpr.predict(Xtest)
    trainscores = score(Ytrain, Ytrain_pred)
    testscores = score(Ytest, Ytest_pred)

    # Store the results
    if output not in results:
        results[output] = {}
    results[output][soilreactiontype] = {
        'optimized_kernel': gpr.kernel_,
        'train_mse': trainscores[0],
        'test_mse': testscores[0],
        'train_r2': trainscores[1],
        'test_r2': testscores[1]
    }
    # Store the model
    if output not in models:
        models[output] = {}
    models[output][soilreactiontype] = gpr
    print(f'MSE (Train) for {output}: {trainscores[0]:.2f}')
```

```
        print(f'MSE (Test) for {output}: {testscores[0]:.2f}')
        print(f'R2 (Train) for {output}: {trainscores[1]:.2f}')
        print(f'R2 (Test) for {output}: {testscores[1]:.2f}')
        print('-' * 50)

# Save all models in a joblib file
joblib.dump(models, 'gpr_models.joblib')
print('Saved all models in gpr_models.joblib')
```

Load the 36 GPR models

```
import joblib
# Define the target outputs
outputs = ['f1', 'f2', 'f3', 'f4', 'f5', 'f6', 'f7', 'f8',
           'm1', 'm2', 'm3', 'm4', 'm5', 'm6', 'm7', 'm8']
# Load models
loaded_models = joblib.load('gpr_models.joblib')
#Accessing and printing optimized kernel for each output
for output in outputs:
    print(f"Loaded model for {output}")
    models = loaded_models[output]
    for soilreactiontype, model in models.items():
        print(f"Soil Reaction Type {soilreactiontype}:")
        print(f"Optimized kernel for {output}: {model.kernel_}")
```

Determine the design case inputs (pile geometry, soil properties)

```
import pandas as pd
import numpy as np

#Insert the soil profile data of the design case from an excel file
excel_file = 'path.xlsx'
dfsp = pd.read_excel(excel_file)
dfsp_cleaned = dfsp.dropna(how='all').reset_index(drop=True)
dfsp_cleaned.columns = dfsp_cleaned.iloc[0]
```

```
dfsp_cleaned = dfsp_cleaned.drop(0)
dfsp_cleaned = dfsp_cleaned.dropna(axis=1, how='all')
dfsp = dfsp_cleaned
numeric_columns = ['su', 'sv', 'Go', 'z']
for col in numeric_columns:
    dfsp[col] = pd.to_numeric(dfsp[col], errors='coerce')
dfsp = dfsp.dropna()
dfsp = dfsp[dfsp['z'] != 0]

# Calculate synthetic features (inputs)
dfsp['su_sigv'] = dfsp['su'] / dfsp['sv']
dfsp['Ir'] = dfsp['Go'] / dfsp['su']
dfsp['Isig'] = dfsp['Go'] / dfsp['sv']
dfsp['Irsig'] = dfsp['Ir']  # by default for clay
dfsp['Sj'] = 1 / dfsp['su_sigv']  # by default for clay
# Convert dataframe columns to numpy arrays
Irsig_sp = dfsp['Irsig'].to_numpy()
Sj_sp = dfsp['Sj'].to_numpy()
z = dfsp['z'].to_numpy()

# Provide the soil_unit of the design case
soil_unit = 'ISL 100m Cowclay'

# Values to provide for test pile
test_L_D = 2 # L/D ratio of design pile
test_D = 10 # Diameter of design pile
num = 100  # number of points for z/L test

# Generate test points for p, m predictions
z_L_pred = np.linspace(0.01, 1, num).reshape(num, 1)
L_D_pred = np.ones(num) * test_L_D
L_D_pred = L_D_pred.reshape(-1, 1)
z_pred = (z_L_pred * test_L_D * test_D)
Irsig_pred = np.interp(z_pred, z, Irsig_sp)
```

```
Sj_pred = np.interp(z_pred, z, Sj_sp)
Irsig_pred = Irsig_pred.reshape(-1, 1)
Sj_pred = Sj_pred.reshape(-1, 1)

#For p,m knot points: X=[z/L   L/D   Go/su.  sv/su]
Xpred_1 = np.concatenate((z_L_pred, L_D_pred, Irsig_pred, Sj_pred), axis=1) #inputs for prediction
#For Hb, Mb knot points: X=[L/D   Go/su.  sv/su]
Xpred_2 = Xpred_1[-1, 1:].reshape(1, -1)
#load the knot point data to plot them along predictions if the design pile is embedded
#in a soil profile that was used in the training of the GPR models
database_name = 'path.db'
table_name = 'soil_rection_data'
df = load_table_to_dataframe(database_name, table_name)
df = preprocess(df)
```

Obtain the GPR predictions for the design case – plot GPR results and confidence interval – extract GPR results to csv file

```
import joblib
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Load models from the .joblib file
loaded_models = joblib.load('gpr_models.joblib')

# Create a dictionary to hold the results
results = {
    'soilreactiontype': [],'z_L': [],'L_D': [],'Isig': [],'Sj': [],
    'output': [],'prediction': [],'lower_bound': [],'upper_bound': []
}

# Loop over each soil reaction type
for soilreactiontype in [1, 2]:
```

```
print(f"Soil Reaction Type {soilreactiontype}:")
sdf = df[df.SoilReactionType == soilreactiontype]
# Select the appropriate Xpred based on soil reaction type
if soilreactiontype == 1:
    Xpred = Xpred_1
elif soilreactiontype == 2:
    Xpred = Xpred_2
else:
    raise ValueError(f"Unexpected soil reaction type: {soilreactiontype}")

# Loop over each loaded model and make predictions
for output in outputs:
    gpr = loaded_models[output][soilreactiontype]
    print(f"Loaded model for {output} (Soil Reaction Type {soilreactiontype}):")
    print(f"Optimized kernel: {gpr.kernel_}")

    # GPR predictions
    Ypred, Ystd = gpr.predict(Xpred, return_std=True)
    # Calculate confidence intervals
    lower_bound = Ypred - 1.96 * Ystd
    upper_bound = Ypred + 1.96 * Ystd

    for i in range(len(Xpred)):
        results['soilreactiontype'].append(soilreactiontype)
        if soilreactiontype == 1:
            results['z_L'].append(Xpred[i, 0])
            results['L_D'].append(Xpred[i, 1])
            results['Isig'].append(Xpred[i, 2])
            results['Sj'].append(Xpred[i, 3])
        elif soilreactiontype == 2:
            results['z_L'].append(np.nan)
            results['L_D'].append(Xpred[i, 0])
            results['Isig'].append(Xpred[i, 1])
            results['Sj'].append(Xpred[i, 2])
```

```
        results['output'].append(output)
        results['prediction'].append(Ypred[i])
        results['lower_bound'].append(lower_bound[i])
        results['upper_bound'].append(upper_bound[i])


    # Plot based on soilreactiontype
    plt.figure(figsize=(6, 4))
    if soilreactiontype == 1:
        # Plot the predictions with confidence interval
        plt.plot(Xpred[:, 0], Ypred, 'k', label='Predicted')
        plt.fill_between(Xpred[:, 0], Ypred - 1.96 * Ystd, Ypred + 1.96 * Ystd, alpha=0.2, label='95% Confidence Interval')
        # Scatter plot based on conditions
        if soil_unit in df['soil_unit'].values:
            df_filtered = df[(df['soil_unit'] == soil_unit) & (df['SoilReactionType'] == 1)]
            plt.scatter(df_filtered.loc[df_filtered['L_D'] == 2, 'z_L'], df_filtered.loc[df_filtered['L_D'] == 2, output],
                        c='red', marker='o', label='L/D=2')
            plt.scatter(df_filtered.loc[df_filtered['L_D'] == 6, 'z_L'], df_filtered.loc[df_filtered['L_D'] == 6, output],
                        c='blue', marker='o', label='L/D=6')
        else:
            print(f"No soil reactions found for the soil profile {soil_unit}.")
        plt.title(f'Gaussian Process Regression for {output} (Soil Reaction Type {soilreactiontype})')
        plt.xlabel('z_L')
        plt.ylabel(output)


    elif soilreactiontype == 2:
        # Single point plot with confidence interval
        plt.errorbar(Xpred[:, 0], Ypred, yerr=1.96 * Ystd, capsize=6, fmt='o', color='black', label='Predicted')
        # Scatter plot based on conditions
        if soil_unit in df['soil_unit'].values:
            df_filtered = df[(df['soil_unit'] == soil_unit) & (df['SoilReactionType'] == 2)]
            plt.scatter(df_filtered.loc[df_filtered['L_D'] == 2, 'L_D'], df_filtered.loc[df_filtered['L_D'] == 2, output],
```

```
                    c='red', marker='o', label='L/D=2')
            plt.scatter(df_filtered.loc[df_filtered['L_D'] == 6, 'L_D'], df_filtered.loc[df_filtered['L_D'] == 6, output],
                    c='blue', marker='o', label='L/D=6')
        else:
            print(f"No soil reactions found for the soil profile {soil_unit}.")
        plt.title(f'Gaussian Process Regression for {output} (Soil Reaction Type {soilreactiontype})')
        plt.xlabel('L_D')
        plt.ylabel(output)
    plt.legend()
    plt.show()

# Convert the results dictionary to a DataFrame
results_df = pd.DataFrame(results)
# Save the results to a CSV file
results_df.to_csv('gpr_predictions.csv', index=False)
```