Metamodelling of the load-displacement response of offshore piles in sand

Alessio Mentani, Laura Govoni, Franck Bourrier, Riccardo Zabatta

Computers and Geotechnics 159 (2023) 105490

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Metamodelling of the load-displacement response of offshore piles in sand

Alessio Mentani [a,*], Laura Govoni [a], Franck Bourrier [b], Riccardo Zabatta [a]

[a] DICAM, Dept. of Civil, Chemical, Environmental and Materials Engineering, Univ. of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy
[b] Univ. Grenoble Alpes, INRAE, IGE, CNRS, IRD, Grenoble INP, IGE, Grenoble 38000, France

ARTICLE INFO

Keywords:
Offshore piles
Sand
Metamodel
Polynomial Chaos Expansion

ABSTRACT

The paper illustrates the development of metamodels of the response of steel piles driven in sand and subjected to pull-out. The metamodels are created for the prediction of the pile tensile capacity and secant stiffness. They were developed using the results of finite element analyses, which made use of finite element models of robustness assessed employing a selection of available data from large-scale model pile tests. Four hundred finite element analyses allowed for the calibration of very accurate metamodels, which were also demonstrated to closely track the outputs of the experimental results. Once calibrated, the metamodels can be used independently from the finite element models they stemmed from. The outcomes of the study show that metamodels of piles response can yield very accurate results within a wide and realistic range of soil-pile configuration, avoiding the laborious implementation and computational cost which underpins the use of finite element models. As the use of metamodels in this context is new, the paper relies on particularly simplified problem, but the procedure could be extended to accommodate modelling features of higher complexity.

1. Introduction

Advanced numerical models now provide a reliable support to the prediction of offshore foundations behaviour, particularly aiding with understanding and integrating model tests data. The activities connected with the development, validation and use of these numerical models are particularly complex and time-consuming and, as such, essentially remain a research prerogative. Simplifications are then generally required to take the results of these studies to the design practice, for instance through the extrapolation of load-transfer curves or the development of macro-element models (e.g., Jeanjean et al. 2017; Burd et al. 2020; Page et al. 2019).

A further option to store, make accessible and ready-to-use, the results of elaborated numerical studies and to support design activities is now offered by metamodeling techniques. Metamodels replace more complex models in the calculation of output data from a combination of inputs. They maintain the original model accuracy but are much simpler to use and computationally more efficient (Teixeira et al. 2021). They are built on the data of a series of simulations conducted with the complex model, starting from a set of input variables, suitably sampled over their domain and combined (Baudin et al. 2017). In general, the number of simulations required to obtain accurate metamodels is relatively small (Wang et al. 2021), therefore the use of sophisticated and computationally expensive models remains sustainable. The advantage is that the resulting metamodel can be employed as surrogate of the complex model it stemmed from, for any possible combinations of input data.

Available metamodeling techniques are generally grouped into classification (e.g., Support Vector Machine) and regression methods (e.g., Kriging, Polynomial Chaos Expansion) and their use is rather well-established in geomechanics (e.g., Kang et al. 2015; Toe et al., 2018; Soubra et al. 2019; Lambert et al. 2021; van den Eijnden et al. 2021). The application to model a foundation response was explored by Sudret (2008), who addressed the classical bearing capacity problem of a footing resting on a layered soil using the Polynomial Chaos Expansion (PCE) technique.

The PCE was first introduced by Wiener (1938) to represent explicitly the response of a mechanical system whose input parameters are modelled by random variables (Blatman and Sudret, 2008). The PCE can handle several input data over a large interval of variation. Thus, it is particularly adequate to describe comprehensively the load displacement response of foundations, which in fact features various types of input including the geometry, the loading conditions and the soil properties; and output variables, such as the capacity and displacements

* Corresponding author.
E-mail addresses: alessio.mentani2@unibo.it (A. Mentani), l.govoni@unibo.it (L. Govoni), franck.bourrier@inrae.fr (F. Bourrier), riccardo.zabatta2@unibo.it (R. Zabatta).

https://doi.org/10.1016/j.compgeo.2023.105490
Received 19 September 2022; Received in revised form 29 March 2023; Accepted 20 April 2023
Available online 4 May 2023

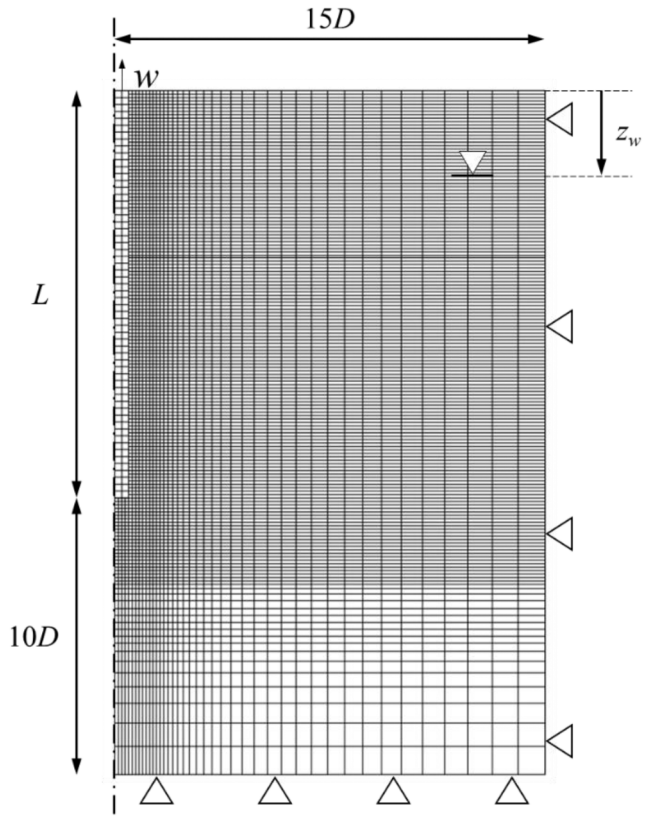

Fig. 1. Mesh and dimension of the pile 2D FE model with boundary conditions.

at failure.

The approach can be employed in a deterministic way, combining different output variables to draw a complete picture of the foundation response for a given set of inputs, a procedure first explored by Mentani et al. (2022). However, its best application is in the context of statistical analyses as metamodels allow to rationally deal with the variability and uncertainties of material and methods, which are intrinsic of any foundation problem and are particularly crucial in the context of offshore wind turbine design (Houlsby, 2016). This trend is well documented by the upsurge in the research interest on statistical approaches in offshore geomechanics, for instance in the field of pile design (Stuyts, 2020; Cai et al. 2021).

Applying metamodelling techniques in the field of offshore foundations has great potential, but is new, therefore the aim of this paper is mostly procedural. The general procedure to build a metamodel of a mechanical model was introduced, with specific reference to the PCE technique applied to Finite Element (FE) models. The simple problem of an offshore pile subjected to a drained and fully plugged monotonic pull-out from a Mohr-Coulomb soil was considered as proof of concept of the procedure. Offshore piles, subjected to axial load, were selected also due to the availability of several, high quality data of model tests, which were collected over the last two decades for the development of design procedures for piles supporting offshore platforms (Yang et al. 2017). The large-scale experiments were used to both assess the FE modelling strategy to be emulated, and to demonstrate the prediction's quality of the calibrated PCE.

Several studies have shown the suitability of FE models (Broere and van Tol, 2006; Said et al. 2009; Han et al. 2017) to capture the essential features of the pile axial response. The choice of considering tensile conditions is instead underpinned by the fact that they are typical of foundation for floating structures and results of the study would be therefore relevant to the upcoming deployment of offshore wind in deeper water.

The paper is organised as follows, the details of the FE approach are first introduced and the resulting FE models are used to reproduce the response of selected experimental results, to assess its robustness and identify reliable output variables. Details of the PCE technique are given and the procedure to train a metamodel is introduced and applied to the assessed FE model. The evolution in the accuracy of the PCE to varying the size of the training dataset was then explored. Finally, the most accurate PCE metamodels are tested to predict the experimental load–displacement response as observed in the selected model pile tests. This showed a satisfying agreement that confirmed the potential of the proposed approach.

2. Fe modelling

2.1. Description of the approach

The approach addresses a pile of diameter D, length L and wall thickness t, subjected to a drained and fully plugged pull-out, following a fully cored installation process in a sand deposit of constant relative density D_r. The FE software Abaqus (Dassault Systèmes, 2020) was employed to implement the FE models. Four-nodes, bilinear and axisymmetric elements were used (CAX4R). The model extends laterally for 15 diameters, where horizontal displacements are restrained, while a vertical boundary condition is applied at the lower model bound, located 10D below the pile base. The mesh is built so that the smaller elements, included within 1D from the pile outer limits, had dimension of 1/8D. The mesh of the deposit then becomes coarser moving out from the pile shaft with larger elements having the size as the pile radius. The geometry and mesh are illustrated in Fig. 1. Since linear elements are known to perform poorly in situations where the gradients of stresses and strains are substantial and the use of better-quality elements was out of the procedural vocation of the paper, a mesh sensitivity study was carried out to assess that no effects on the results occurred by further reducing the element size across the pile shaft, with this is ascribable to the linearity of the selected interface model.

The piles were modelled as deformable bodies obeying a linear elastic response. A uniform cross-section was prescribed to the pile models, for the results to be representative of open-ended piles which are installed fully-cored and experience a fully plugged pull-out, as typical of offshore piles. The soil is prescribed to be linear elastic, failing according to the Mohr-Coulomb criterion. Data required for the soil model implementation are the soil density, ρ_{soil}, and the elastic (E_{soil}, ν_{soil}) and plastic (Φ'_p, Ψ) parameters. The earth pressure coefficient (K_{ini}) is also needed to set the initial stress state.

2.2. Implementation of the FE models and identification of input parameters

A numerical strategy to define the properties of the FE model was identified. The strategy defines a rule to build the model by relating some input parameters. The FE model can be built for any input combination, but the strategy avoids unrealistic conditions (e.g., low stiffness for very dense sand; $\delta > \Phi'_{cv}$; etc.).

The soil properties can be linked to an artificial cone resistance, q_c - which is associated to the deposit's constant relative density - through the use of CPT-based empirical correlations. An expression for q_c as function of D_r, suitable for uniform, normally consolidated sand deposits is that of Jamiolkowski et al. (2003)

$$q_c = 20 p_a \left(\sigma'_{v0}/p_a\right)^{0.5} \exp(2.86 D_r) \quad (1)$$

where σ'_{v0} is the vertical effective in situ stress, p_a is the atmospheric pressure. Equation (1) can be in turn used to calculate the soil elasticity, according to the functional form proposed by Robertson (2009)

$$E_{soil} = 0.015\left(10^{0.55 I_c + 1.68}\right)\left(q_c - \sigma'_{v0}\right) \quad (2)$$

Table 1
Input variables and ranges.

Input type	Input variable	Range
Geometry	Diameter, D [mm]	250–1000
	Slenderness ratio, L/D [-]	10–70
	Thickness ratio, D/t [-]	10–70
Soil property	Relative density, D_r [%]	40–100
	Soil behaviour type index, I_c [-]	1.31–2.05
	Critical state friction angle, ϕ'_{cv} [°]	28–40
Water level	Parameter for water level location, λ_w [-]	0.0–1.0
Interface	Parameter to define interface friction angle,λ_δ [-]	0.70–0.95
Earth pressure coefficient	Coefficient to define initial soil stress state, μ [-]	0.03–0.07

where I_c is the soil behaviour type index. Similarly, the equation is applied to calculate the soil peak resistance, following the equation of Kulhawy and Mayne (1990)

$$\varphi'_p = 17.6 + 11log\left(\frac{q_c}{(p_a/\sigma'_{v0})^{0.5}}\right) \quad (3)$$

Equation (3) in turn allows to estimate the evolution of the dilation angle (Ψ) with depth, assuming $\Psi' = \phi'_p - \phi'_{cv}$, being ϕ'_{cv} the soil constant volume friction angle. To account for the effects of pile driving on subsequent load stages in the wished-in-place model, the equation by Randolph (1994) was used

$$\sigma'_{ri} = K_{min} + \left(0.01\frac{q_c}{\sigma'_{v0}} - K_{min}\right)exp\left(-\mu\frac{L-z}{D}\right) \quad (4)$$

where σ'_{ri} is the post-installation radial stress along the pile shaft, K_{min} is the active earth pressure coefficient and μ is a dimensionless variable, ranging between 0.03 and 0.07 (Randolph, 2003). The equation particularly suits the problem examined, as it accounts for the effects of friction fatigue and geometrical effects induced in the soil by pile driving. Based on eq. (4), the earth pressure coefficient is simply derived as $K_{ini} = \sigma'_{ri}/\sigma'_{v0}$and applied along the radial direction.

The interface response is taken rigid up to failure, governed by the interface friction angle, δ, according to the Coulomb criterion. The interface friction angle is the critical state's, assuming that peak has been yet mobilised during the installation process (Lehane et al. 2005). The interface friction angle is prescribed to be always smaller than the constant volume soil friction angle (Tovar-Valencia et al. 2018) and expressed as a fraction of it ($\delta = \lambda_\delta \cdot \phi'_{cv}$). The ground water level is function of the model depth ($\lambda_w \cdot (L + 10D)$). This model feature was introduced to extend the procedure to model piles tested onshore.

According to the FE modelling strategy, when failure at the interface is reached, relative motion between the pile and the soil can occur (i.e., slipping), differently, the soil remains attached to the pile (i.e., sticking). Due to the lower interface resistance, compared to that available in the soil, the pile is easily extracted, with small plastic deformations and no dilation taking place. As a result, the model response is expected to be essentially linear elastic up to failure.

The implementation of the approach requires nine input variables which can be easily traced back to standard site and laboratory test data, allowing for the development of simple site-specific FE models. All input variables are listed in Table 1 and grouped according to the features they describe. Three variables define the pile geometry D, L/D, D/t, three the soil mechanical properties D_r, I_c, ϕ'_{cv}, one the interface, λ_δ, one the post-installation earth pressure coefficient μ, one the ground water levelλ_w.

2.3. FE modelling of the response of large-scale tests

Data of eight tensile tests, concerning steel open-ended piles driven in sand, were selected among those now available in literature, to understand the capacity of the FE approach, to reproduce the experimentally observed behaviour and to identify, as consequence, a reliable set of output indicators. Five tests were carried out in onshore test sites and yet included in the Yang et al. (2015) database, which was created for the development of CPT-based design methods for piles supporting oil platforms. The other three are large-scale laboratory tests, belonging to an experimental programme, recently performed to investigate the response of piles for jacket supporting offshore wind turbines (Schmoor et al. 2018).

The list of experiments is given in Table 2, where they are also assigned a label, hereinafter used for identification purposes. The essential information on the experiments were retrieved in the main bibliographic references, also included in Table 2, and concern the test results, described by the tensile load–displacement curve as well as the tests details, required to implement the FE models. These are the nine input variables introduced in Section 2.2. Data about pile dimensions, ground water levels and interface angles, were directly accessible. For two test sites only (Gavin et al. 2013 and Kolk et al. 2005a), the constant volume friction angle was missing, and was, therefore, guessed, in agreement with the soil type and ensuring δsmaller than ϕ'_{cv}, which was instead known. The relative density and the soil behaviour type index were indirectly inferred from the local CPT data applying Eq. (1) and the definition of I_c, respectively. The values were averaged along the pile shaft to comply with the FE approach requirement for a uniform soil profile. For the empirical parameter μ, the two extreme values were considered to bound all possible solutions.

2.4. Performance of the FE models and identification of behavioural indicators

The input values are all shown in Table 2 and were used to generate the FE models and run the analyses which were small strain and carried out applying an upward 0.1D displacement to the pile head.

The results of the simulations are displayed in Fig. 2 in terms of normalised load, $\overline{V}$, and displacement, $\overline{w}$, in continuous black lines. Two curves are shown for each of the tests, resulting from the use of the two μ threshold values. As expected, the numerical response is roughly bi-linear, a result which is qualitatively similar to that attainable with current design standards, which implement linear load-transfer curves (ISO, 2016). Despite the simplifications introduced, the results are in good and consistent agreement with the experiments, which were also

Table 2
Pile geometries, site information and soil properties of the experimental database.

ID	Reference	D [mm]	L/D [-]	D/t [m]	λ_w [-]	D_r [%] *	I_c [-] *	ϕ'_{cv} [°]	λ_δ [-]	μ [-]
J-06	Jardine et al. 2006	457	42.23	33.85	0.16	76.5	1.69	32.0	0.84	0.03
G-13	Gavin et al. 2013	340	20.59	24.29	0	88.9	1.67	37.9§	0.95	0.07
K-05a2	Kolk et al. 2005a (2)	763	50.72	21.43	0.021	78.5	1.99	32.6§	0.95	
K-05a3	Kolk et al. 2005a (3)	763	61.60	21.43	0.018	85.7	1.84	32.6§	0.95	
K-05b	Kolk et al. 2005b	763	61.21	21.43	0.018	83.8	1.98	32.6§	0.95	
S-18P1	Schmoor et al 2018 (P1)	273	20.88	54.60	0.059	73.6	1.40	28.7	0.84	
S-18P3	Schmoor et al 2018 (P3)	356	16.01	56.51	0.053	73.6	1.40	28.7	0.84	
S-18P4	Schmoor et al 2018 (P4)	356	18.82	56.51	0.048	73.6	1.41	28.7	0.84	

*Average values following interpretation of CPT data; §Assumed data, not available in the Reference work.

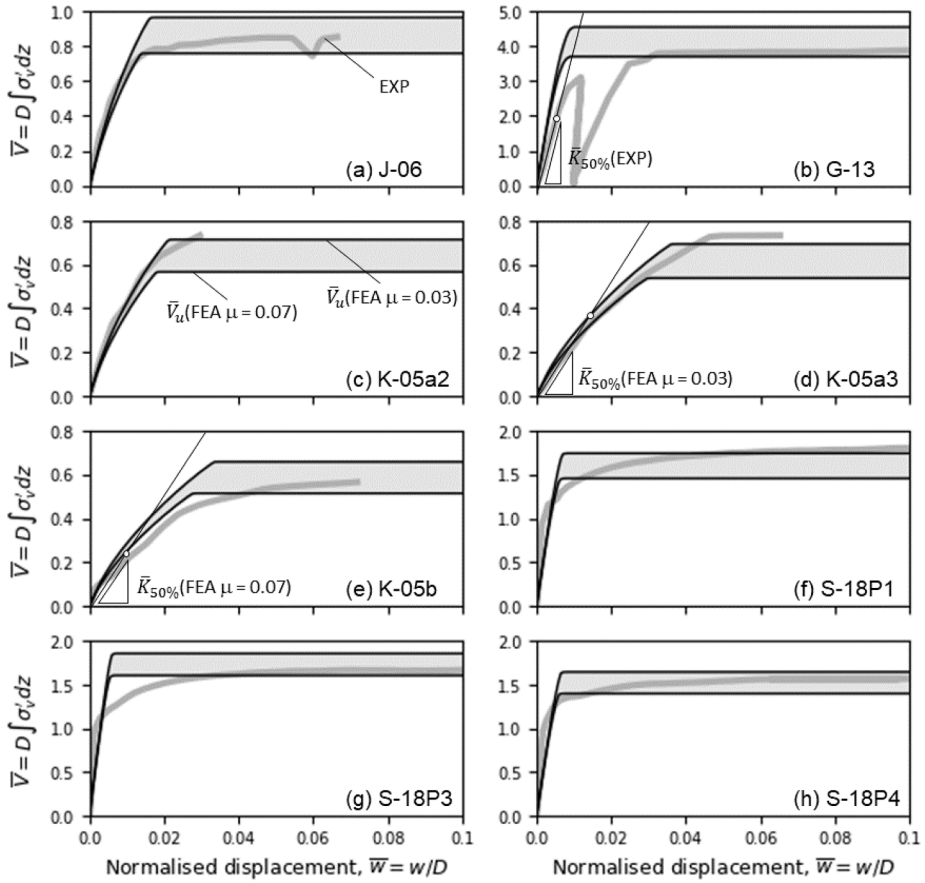

Fig. 2. Comparison between the experimental results and the FE outcomes of the pile database.

included in Fig. 2 to allow for a direct comparison. Among the numerical curves, two values were identified as behavioural indicators: the normalised pull-out capacity, $\overline{V}_u$, and the secant stiffness measured at a service load level corresponding to the 50% of the mobilised bearing capacity, $\overline{K}_{50\%}$. The latter was chosen within the typical pile working range as proposed by Lehane et al. (2020).

An average relative error of the pile capacity equal to 11.45% (standard deviation, σ, of 7.25%) was measured for the eight experiments when considering the FE model predictions with the two extreme values of μ coefficients. On the selected data, it was observed that an average error of 7.2% (σ = 5.2%) could be obtained using μ = 0.05 as also suggested by Randolph (2003).

Predictions of $\overline{K}_{50\%}$ yielded an average error of 39.91%, on the whole dataset (σ = 21.75%), with the larger value observed for pile G-13. Another method to measure the model accuracy is to compute the normalised difference between the measured and calculated head pile displacements (i.e., $(w_{FE} - w_{EXP})/D$) corresponding to the 50% of the mobilised capacity. This returns a maximum value of 0.65% of the pile diameter for pile G-13. A prediction error similar to those obtained using the most recent advances in terms of load-transfer curves methods (Lehane et al. 2020).

Based on this comparison, and reminding here the procedural vocation of the study, the normalised pull-out capacity, $\overline{V}_u$ and secant stiffness, $\overline{K}_{50\%}$ were considered to be behavioural predictors of sufficient and consistent reliability and consequently taken as the approach output indicators.

3. Fe-based metamodelling

3.1. PCE details

The metamodelling study was built in the Python programming language, using the open source package OpenTURNS (Baudin et al. 2017) and the PCE metamodel was used as detailed in the following.

In general, a mechanical system, like the FE model, has n uncertain and independent input parameters that can be represented by the vector $\times = \{x_1, \ldots, x_n\}$. The model function, $f(\boldsymbol{x})$, which returns the evaluations of the mechanical model as a set of m output quantities in the vector $\boldsymbol{y} = \{y_1, \ldots, y_m\}$, is known only for some input combinations by running the simulation, but has no explicit expression. A metamodel is an analytical function, $g(\boldsymbol{x})$, that is able to approximate the model function.

The PCE metamodelling technique consists in building the approximation function by expanding the original model function onto a finite-dimensional basis made of orthogonal polynomial sequences of the Hilbert space. The PCE function is built as a linear combination of selected multivariate orthonormal polynomial basis, $\psi_k(\boldsymbol{x})$, and their corresponding coefficients, α_k, as represented by

$$y = f(x) \cong \hat{y} = g(x) = \sum_{k \in K} \alpha_k \Psi_k(T(x)) \tag{5}$$

Where $K \in \mathbb{N}$ is the number of terms used in the expansion and T is an isoprobabilistic transform. This transformation rescales the input variables into common distribution forms (e.g., uniform, normal, exponential). This process governs the choice of the family of polynomial basis to be used, which are associated to a specific distribution type in a Hilbert space (Sudret, 2008). The polynomial basis obeys orthogonality and normality rules, that is their inner product is $\langle \Psi_a, \Psi_b \rangle = \delta_{ab}$, where δ_{ab} is the Kronecker delta function. Once an appropriate set of basis (e.g.,

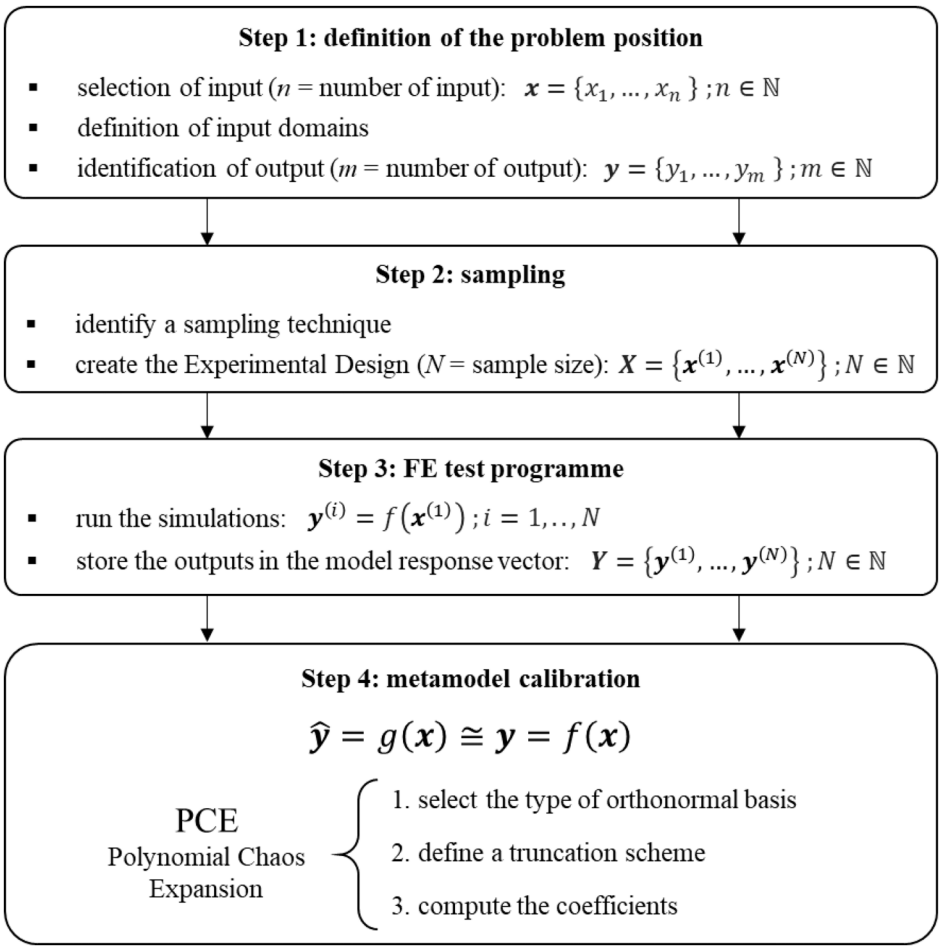

Fig. 3. Procedure for generating a metamodel.

Legendre, Hermite, Krawtchouk) is selected according to the transformed input distribution form, the problem reduces to the adoption of a truncation scheme for determining the expansion size (i.e., K in eq.5) of the approximation function, followed by the computation of the coefficients.

In this study, a standard truncation scheme was adopted that means all the polynomials up to a maximum degree, p, are used in the expansion (Le Gratiet et al. 2017). This means that the size of truncation of the set of polynomial bases equals to:

$$K = \text{card}\mathscr{A}^{n,p} = \frac{(n+p)!}{n!p!} \tag{6}$$

Therefore, the maximum polynomial degree adopted in the expansion leads to the number of coefficients to be computed.

For computing the coefficients, the common regression approach was adopted. It means that a set of coefficients is determined which minimises the mean square error between the original model observations and the PCE metamodel predictions given a set of input and output pairs (i.e., $\boldsymbol{X} = \{\boldsymbol{x}^{(i)}, i = 1, \cdots, N\}$ and $\boldsymbol{Y} = \{\boldsymbol{x}^{(i)}, i = 1, \cdots, N\}$) as

$$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_K) = \underset{\alpha \in \mathbb{R}^K}{\text{argmin}} \mathbb{E}\left[(\text{f}\ (\boldsymbol{X}) - \text{g}\ (\boldsymbol{X}))^2\right] \tag{7}$$

The data derives from a parametric test programme of the mechanical system being surrogated as further detailed in the following Section. In the equation, the expectation operator is the empirical mean over the size of the parametric sample.

3.2. Procedure to train a PCE metamodel

The procedure to calibrate a metamodel of a mechanical system consists of four steps, as also illustrated in the chart of Fig. 3. The procedure applies to all metamodelling techniques and it is further detailed here with reference to the PCE metamodel.

Step 1: definition of the problem position

According to the problem under investigation, the problem position is defined by three choices. First, a finite number, n, of input variables must be selected. The inputs are parameters required to generate the mechanical model being surrogated and that influencing its response. There is no limit to the number of input variables to be selected, but a large number of parameters would increase the unknown PCE coefficients to be computed in eq.7, also according to eq.6. Together with the selection of the input, their range of variations is to be defined in this phase. The PCE can handle large variable domains, but care should be taken when identifying their ranges, as the PCE function cannot extrapolate outside the calibration domain.

Last, the output variables computed by the mechanical model for what the metamodel wants to be trained must be identified at this stage. For each output, a set of polynomial basis coefficients is computed for the PCE, therefore their number, m, might be unlimited.

Step 2: sampling

Training a metamodel requires a sufficiently large number of input and output combinations, generated through a suitably designed test programme. This relies on using an adequate sampling technique to allow for a rational combination of the inputs. Different sampling methods can be used: from traditional random Monte Carlo, to space-filling methods like the Latin Hypercube Sampling (LHS, McKay et al.

1979), rather than quasi-random sequence sampling like Halton or Sobol.

The input pairs of the generated sample are collected in a vector $\boldsymbol{X} = \{\boldsymbol{x}^{(i)}, i = 1, \cdots, N\}$ named the Experimental Design (ED) of size N. In the PCE, the sample size is relevant for computing the α_k coefficients as it must be larger than the truncation size, K.

Step 3: test programme

Given the sample, the test programme follows which is carried out with the original mechanical model. The simulations are run for each pairs of the ED and the results, for the output identified in the problem position, are stored in the so-called model response vector $\boldsymbol{Y} = \{\boldsymbol{y}^{(i)}, i = 1, \cdots, N\}$ that has the same size of $\boldsymbol{X}$.

Step 4: metamodel calibration

Once an ED and its outcomes are stored in respective vectors ($\boldsymbol{X}$ and $\boldsymbol{Y}$), the calibration phase of a metamodel consists on determining the analytical function, g, that better approximates the original computational model, f. As described in Section 3.1 and detailed in Fig. 3, for the PCE this reduces to three choices. First, the family of orthonormal basis must be selected. This choice derives by the distribution for of the input variables of the ED. Then, a truncation scheme is identified and the standard one was used in this study. This, together with the size of the ED governs the maximum degree, p, that can be used for the polynomial basis as further detailed in the following Section. The last step is the computation of the coefficients, α_k, which are computed through eq.7.

3.3. Application to the FE pile model

The metamodel procedure was applied to the case study of the offshore pile subjected to tensile loading described in Section 2. The FE model was the mechanical model. This was characterised in Section 2.2 by $n = 9$ input parameters, as summarised in Table 1, which were the selected input variables of the problem position. As for the output, the two variables ($\overline{V}_u$, $\overline{K}_{50\%}$) identified as the system behavioural indicators in Section 2.3, were selected. Thus, the output vector $\boldsymbol{y}$ had size $m = 2$.

The problem position was finally set by defining the domain of the input variables. Each range of variation was defined to have a realistic and comprehensive picture of the pile response and to comply with the available experimental evidences (Yang et al. 2015). As for the pile geometry, diameters were allowed to assume values within the range of available large-scale model tests, with thresholds for slenderness and thickness ratios selected to include both wind and oil and gas applications. Soil conditions were intended to encompass medium to very dense sand, according to the available evidences from onshore and laboratory pile tests to exploited offshore sites. Realistic values were prescribed to the constant volume soil and interface angles. The ground water level was set to accommodate both offshore and onshore conditions. The soil behaviour index and the earth pressure coefficient were restrained to values defined in literature (Robertson, 2009; Randolph, 2003). Threshold values for each parameter range are shown in Table 1.

The following step was the generation of the ED. The LHS was used to sample and combine the input variables. The LHS is the most suitable techniques when variables have an equal probability of occurrence over their domain of variation. It is a stratified or space-filling sampling type that generates a sample of size N by partitioning the range of each input variable into N non-overlapping equal probability intervals. The sample is created by randomly selecting one value from each interval. The method was used since it builds a random sample, while allowing to cover the full domain of variation of the input variables providing a uniform distribution for each parameter. Thanks to this characteristic, the trained PCE metamodel is expected to have a uniform accuracy within the whole domain of each input data. Further, the LHS sample size can be increased without losing its property, but the augmented size must be multiple of N (Sallaberry et al. 2008). Following the procedure, four ED were created of increasing size from $N = 50$, according to the simple geometric series of ratio two (i.e., $N = 100$, 200, 400). The samples will be referred as calibration samples hereinafter.

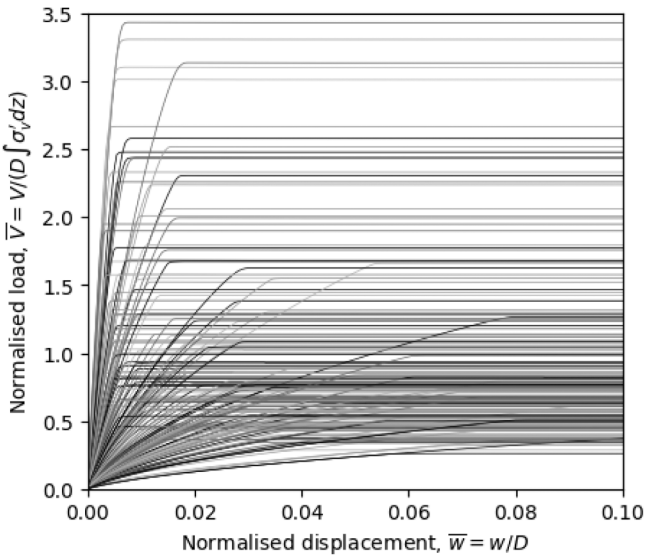

Fig. 4. Results of the FE test programme for the sample of size $N = 200$.

The calibration samples were used as input parameters for the following FE test programme, which involved a total of 400 simulations. As per Section 2.3, simulations were small strain and carried out applying an upward 0.1D displacement to the pile head. The results of the first 200 analyses (corresponding to the calibration sample of size n = 200) are shown in Fig. 4, in terms of normalised load, $\overline{V}$, and displacement, $\overline{w}$. Despite the applied normalisation procedure and simplified FE modelling approach, the variability of the outcomes is significant, highlighting the strong non-linear response of the model function.

All the information, input and output, of the four calibration samples were stored in respective vectors, $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$, with $i = (50; 100; 200; 400)$. A uniform distribution of the nine input variables was assumed when generating the samples with the LHS technique. The distribution of the input variable over the domain of variation drives the choice of the orthonormal basis of polynomials. Being in a Hilbert space, a uniform distribution requires the use of the Legendre family of polynomials.

Since a standard truncation scheme is adopted in this study, a thumb rule of $N \approx 2 \bullet K$ was used for the problem to be well-posed as also recommended by Le Gratiet et al. 2017. Then, according to eq.6 the maximum degree of the polynomial basis p was established for each PCE. Other truncation schemes like the hyperbolic or sparse (Blatman and Sudret, 2008) can be adopted to reduce the number of multivariate basis considered in the truncation, thus increasing the maximum degree of the univariate basis. However, the accuracy obtained with the standard scheme showed to be already satisfying for emulating the pile FE model.

The stored input and output data were finally used to determine the best fit polynomial coefficients, α_k, of the polynomial basis according to eq. (7). At the end of the procedure, four PCEs were created PCE_{50}, PCE_{100}, PCE_{200} and PCE_{400} using data of calibration samples of size 50, 100, 200 and 400.

3.4. Accuracy of the metamodel

Once trained, the accuracy of the metamodel has to be assessed. This is done by comparing the observations of the original mechanical model, $\boldsymbol{Y}$, with the metamodel predictions, $\widehat{\boldsymbol{Y}}$, obtained for a new ED. This sample can also be generated following a different sampling strategy (i.e., sampling technique and distribution form of the inputs). Other methods, like the traditional cross-validation techniques can be employed to evaluate the PCE accuracy. They allow to assess the met-

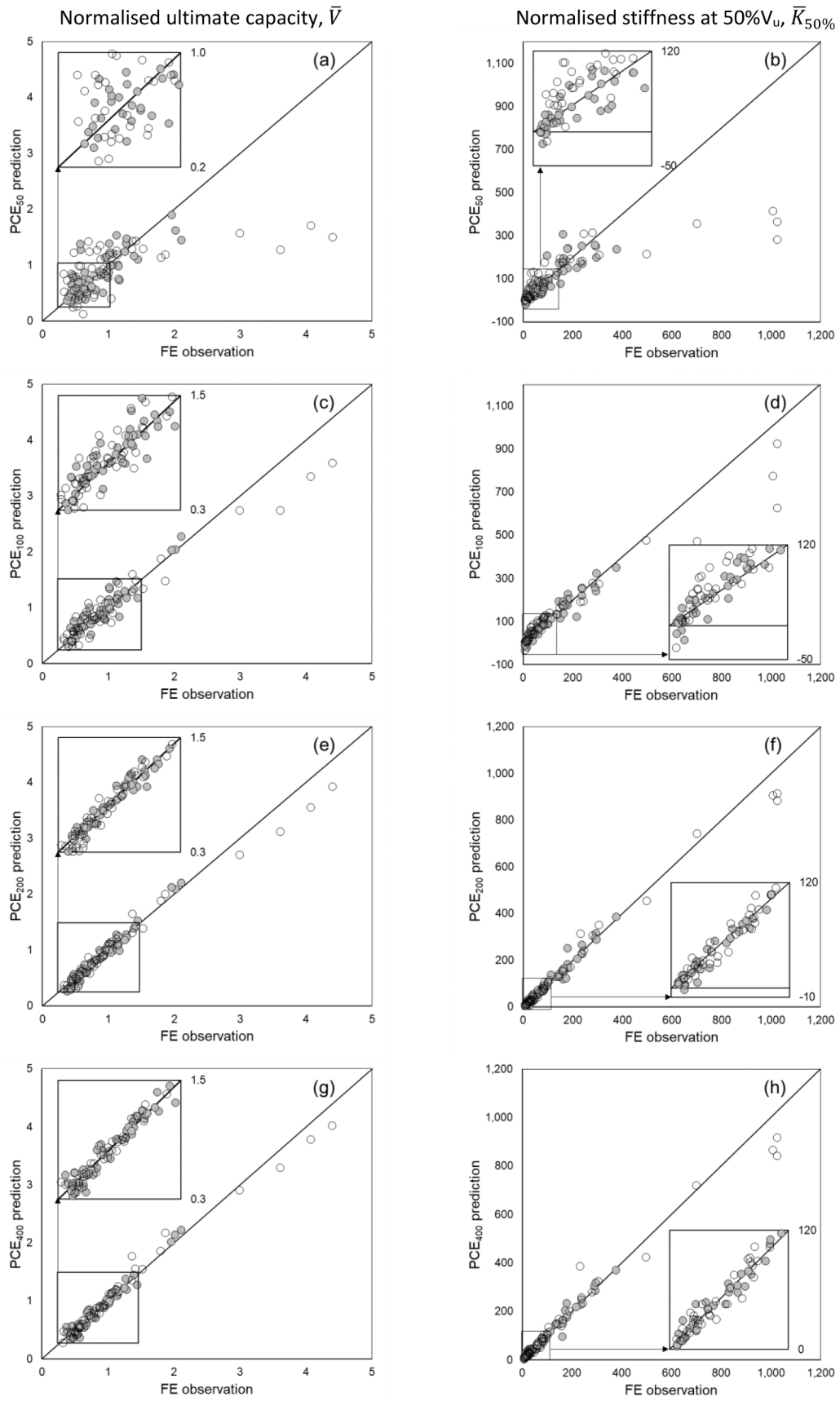

Fig. 5. PCE_n metamodels predictions Vs FE model observations for the normalised tensile capacity (left side) and for the normalised stiffness at a load rate of 50% the ultimate capacity (right side): (a) and (b) for PCE_{50}; (c) and (d) for PCE_{100}; (e) and (f) for PCE_{200}; (g) and (h) for PCE_{400}. The grey shaded circles referred to the additional 50 data of the augmented sample.

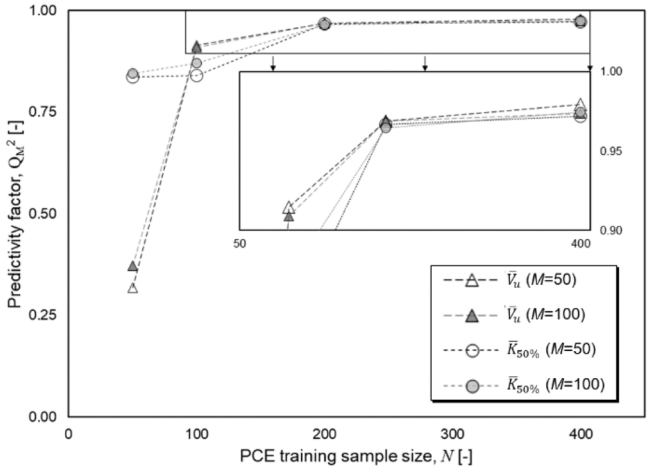

Fig. 6. Variation of the predictivity factors, Q^2, for the two selected outcomes as function of the training sample size, N, and of the validation sample size, M.

amodel without requiring a new testing programme, thus reducing the time–cost, but the PCE is then iteratively tested over the same ED. In this study the first option was followed also to investigate if there was any effect of the validation sample size on the PCE accuracy.

The performance of the PCE can be measured by analysing the residuals or the relative errors of its predictions with respect to the FE results, but a common estimator of regression metamodel's accuracy is the predictivity capacity factor (Blatman and Sudret, 2008)

$$Q_M{}^2 = 1 - \frac{\sum_{i=1}^{M}\left[\boldsymbol{y}^{(i)} - g\left(\boldsymbol{x}^{(i)}\right)\right]^2}{M \bullet Var(\boldsymbol{Y})} \tag{8}$$

where M is the validation sample size and $Var(\boldsymbol{Y})$ denotes the variance of the model response vector.

4. PCE assessment and exploitation

4.1. Accuracy

The PCEs accuracy was measured by comparing with two new ED that will be referred to as validation samples. They were built with the LHS technique as in the calibration procedure, with size $M = 50$, then augmented to $M = 100$. The reason of building two validation samples rather than one, was to assess possible effects of variation of parameter M on the accuracy measurements. The validation samples were used as input parameters for a new FE test programme involving a total of 100 simulations to obtain the model response vectors, $\boldsymbol{Y}_i = \{\overline{V}_u{}^{(i)}, \overline{K}_{50\%}{}^{(i)}; i = 50; 100\}$. The four developed PCEs were also applied to the two validation samples and their predictions were similarly stored in $\widehat{\boldsymbol{Y}}_i$. The normalised pull-out capacity, $\overline{V}_u$, and secant stiffness, $\overline{K}_{50\%}$, obtained with the FE and the corresponding PCEs predictions are compared in Fig. 5. The figure explores the effects of the size of the calibration sample as follows: Fig. 5a–b shows the prediction with PCE_{50}, Fig. 5c–d those with PCE_{100}, Fig. 5e–f those with PCE_{200} and Fig. 5g–h those with PCE_{400}. Data resulted from the smallest calibration sample ($M = 50$) are given in white circles, while the further grey data points refer to the other validation sample ($M = 100$).

Observing that best results would have data points falling close to the diagonal, an overall improvement is clearly observed on both output

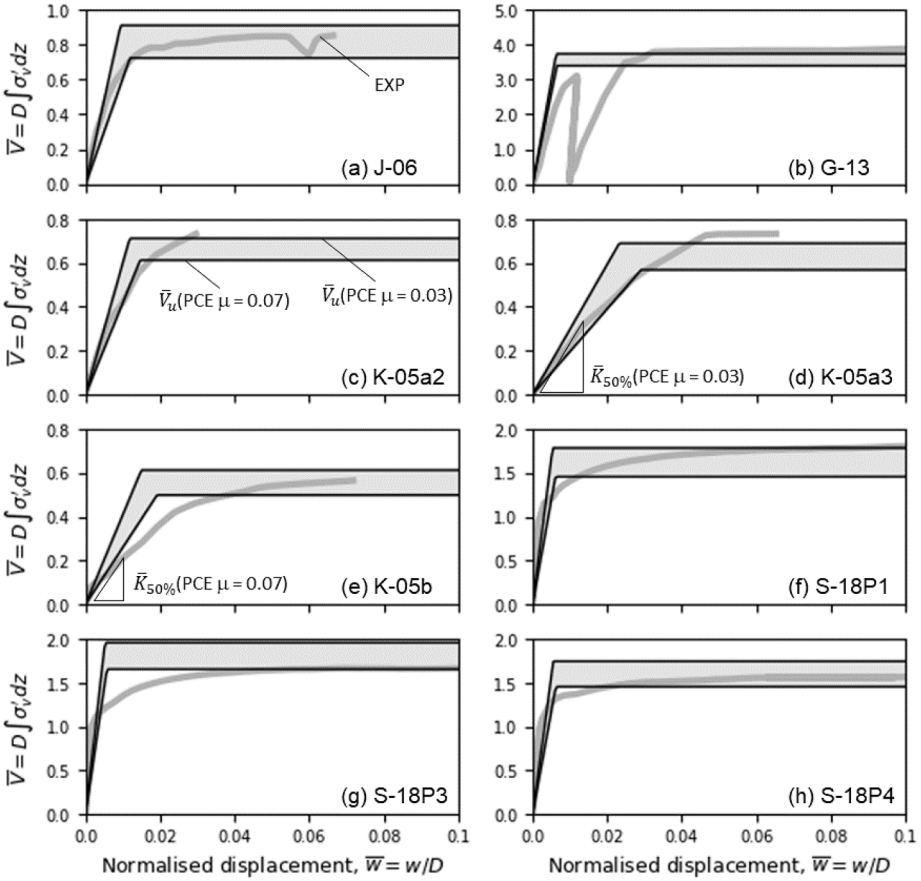

Fig. 7. Comparison between the experimental results and the PCE_{400} predictions of the pile database.

Table 3
Prediction errors of the PCE_{200} and PCE_{400} with respect to the FEA predictions of the experimental database.

PCE prediction	μ coefficient	J-06	G-13	K-05a2	K-05a3	K-05b	S-18P1	S-18P3	S-18P4
V_u [-]	0.03	3.99%*	25.59%*	0.46%*	6.29%	3.22%	8.94%*	1.05%*	2.46%*
PCE_{200}	0.07	5.78%	17.69%*	4.92%*	7.53%	1.81%	1.08%	7.29%	8.09%
V_u [-]	0.03	5.65%*	17.84%*	0.18%*	0.62%*	6.95%*	2.42%	5.63%	6.23%
PCE_{400}	0.07	4.65%*	8.19%*	8.13%	5.67%	3.04%*	0.05%	3.24%	4.16%
$K_{50\%}$ [-]	0.03	6.40%	17.01%	43.03%	13.69%	37.31%	6.70%	4.03%*	4.81%*
PCE_{200}	0.07	13.53%	1.29%	24.28%	4.87%*	60.31%	11.78%*	17.35%*	3.01%
$K_{50\%}$ [-]	0.03	20.29%	9.86%	6.05%	10.55%*	60.00%	3.17%*	7.56%*	2.07%
PCE_{400}	0.07	3.26%	3.62%*	13.04%	5.92%*	49.33%	4.46%*	9.15%*	5.00%*

* Negative values.

variables with increasing the sample size, starting from a rather disperse set of data obtained with PCE_{50}. A small enhancement using PCE_{100} is shown, with large residuals still on $\overline{V}_u$ and a few negative predictions on small values of $\overline{K}_{50\%}$.

The agreement between FE results and metamodel predictions significantly improved using the PCE_{200}. The measured average relative error of the two output for the largest validation sample ($M = 100$) is 8.90% ($\sigma = 8.37\%$) and 16.38% ($\sigma = 18.45\%$) for $\overline{V}_u$ and $\overline{K}_{50\%}$, respectively. Few negative predictions of the secant stiffness were observed for PCEs built with small calibration samples, as more clearly visible in the figure boxes where the graphs were scaled up to allow for a focus close to the origin of the axis. No negative predictions resulted from the use of PCE_{400}, with an average error reducing to 12.61% ($\sigma = 14.60\%$) but no further improvement on $\overline{V}_u$ (i.e., average error 9.10%, $\sigma = 8.90\%$).

The predictivity capacity factors were computed according to eq. (8) and the results were reported in Fig. 6. The metamodel accuracy for the two selected output increases rapidly with the calibration sample size N to stabilise at values equal or greater than 97% for both $\overline{V}_u$ and $\overline{K}_{50\%}$ for N above 200, a trend which showed to be rather independent from the size of the validation sample, M.

4.2. Use of the Pce to model large-scale pile tests

The PCE_{200} and the PCE_{400} were used to investigate the approach potential to predict the response of the model piles introduced in Section 2. To the aim, the input data of Table 2 were used in the PCEs algorithm. In Fig. 7 the prediction of the two outputs $\overline{V}_u$ than $\overline{K}_{50\%}$ obtained using PCE_{400} were combined into a bi-linear shape and presented along the experimental load–displacement curves for a qualitative comparison. The overall agreement is satisfactory. Consistently with the results of the FE analyses illustrated in Fig. 2, the PCE offers a better approximation on $\overline{V}_u$ than on $\overline{K}_{50\%}$.

The relative errors of the PCEs predictions of $\overline{V}_u$ and $\overline{K}_{50\%}$ with respect to the FE observations of the two output are reported in Table 3 for the eight model pile tests. This allowed to explore the PCEs performance with respect to its original computational model on this new dataset. The average relative error on $\overline{V}_u$ decreases from 6.64% ($\sigma = 6.6\%$) using PCE_{200} to 5.17% ($\sigma = 4.26\%$) for PCE_{400}. On $\overline{K}_{50\%}$ the decreasing trend is also apparent, from 16.84% ($\sigma = 16.72\%$) to 13.33% ($\sigma = 16.88\%$) passing from PCE_{200} to PCE_{400}. The larger discrepancies observed for piles G-13 and K-05b, are ascribable to the simultaneous presence of some input parameters (i.e., I_c; λ_w and λ_δ) close to their range thresholds.

These quantities are in line with the performance indicators calculated in Section 4.1, confirming that accurate PCEs can be yet obtained with a set of 200 simulation, with a slight increase in the overall performance using 400 simulations. Results also show that a PCE, calibrated on a set of FE analyses, is expected to maintain the calibration accuracy when applied to any other input combinations. Such PCE can be then used with confidence as a predictive tool, provided that input parameters belong to the calibration domain, with care taken when they are sampled close the boundaries.

5. Concluding remarks

The paper has described the process to develop FE-based PCE metamodels of the response of offshore piles in sand and has explored their potential as predictive tools. Tensile loading conditions were considered so that the results of the study are relevant to the deployment of offshore wind in intermediate and deep water, as such condition is typical of foundations of jackets and floating structures.

Due to the procedural vocation of the paper, the approach was kept very simple in all phases, but allowed for the entire procedure to be displayed and key points addressed which concern:

- the definition of the variables describing the soil-pile system;
- the identification of a rational procedure to turn the input parameters into FE model parameters;
- a thorough assessment of the FE model ability to reproduce the experimentally observed response, with the identification of FE model shortcomings which may affect the results and therefore the metamodeling activities.
- the identification of key reliable indicators of the pile response based on the accuracy with which the FE model is able to capture them;
- the choice of the most adequate metamodelling technique and sampling method;
- the quantification of metamodel accuracy.

The soil-pile system was described via nine input variables, which were related to FE model parameters using well-established empirical and analytical correlations. Two FE model outputs were identified as sufficiently reliable indicators of the experimental pile response and PCE metamodels were developed to approximate such response. Their accuracy was quantified higher than 97% using a minimum number of 200 FE simulations.

A similar accuracy was observed when the metamodels were applied to the experimental dataset, providing enough confidence in the use of the developed metamodels as predictive tools.

The paper also showed the crucial role played by field and laboratory pile tests. The FE model was in fact shown able to reproduce some of the features (i.e., $\overline{V}_u$, $\overline{K}_{50\%}$) observed in eight large-scale pile tests. These features were in turn accurately approximated by the metamodel, whose aim was to make them accessible and exploitable in a simpler way than that offered by the original physical and numerical FE models.

Significant experimental efforts have characterised the offshore foundation research in the last decades, first in the oil and gas then in the wind sector and the available data could be now encapsulated in powerful metamodels, while supporting their development and assessment.

The procedure, which was showed here in the essential step, could be further extended to accommodate modelling features of higher complexity, optimising the selection of the input variables and can be employed to predict other behavioural aspects. Metamodels can be also exploited to support statistical analyses, particularly suitable to address

the variability and uncertainties which characterise any foundation problem, particularly in the offshore environment.

CRediT authorship contribution statement

Alessio Mentani: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Funding acquisition. **Laura Govoni:** Conceptualization, Methodology, Supervision, Writing – original draft, Project administration. **Franck Bourrier:** Conceptualization, Supervision. **Riccardo Zabatta:** Formal analysis, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

As for the funding research framework all data will be open access

Acknowledgement

This work forms part of the activities of the project SEAFLOWER, which has received funding from the European Union's Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No 891826.

References

Baudin, M., Dutfoy, A., Iooss, B., Popelin, A.L., 2017. OpenTURNS: an industrial software for uncertainty quantification in simulation. In: Ghanem, R., Higdon, D., Owhadi, H. (Eds.), Handbook of Uncertainty Quantification. Springer, Cham. https://doi.org/10.1007/978-3-319-12385-1_64.

Blatman, G., Sudret, B., 2008. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. Comptes Rendus Mécanique. 336 (6), 518–523.

Broere, W., van Tol, A.F., 2006. Modelling the bearing capacity of displacement piles in sand. Proc. Inst. Civil Eng. – Geotech. Eng. 159 (3), 195–206.

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., Potts, D.M., 2020. PISA design model for monopiles for offshore wind turbines: application to a marine sand. Géotechnique 70 (11), 1048–1066.

Cai, Y., Bransby, F., Gaudin, C., Uzielli, M., 2021. A framework for the design of vertically loaded piles in spatially variable soil. Comput. Geotech. 134, 104140.

Gavin, K.G., Igoe, D.J.P., Kirwan, L., 2013. The effect of ageing on the axial capacity of piles in sand. Proceedings of the Institution of Civil Engineers - Geotechnical Engineering 166 (2), 122–130.

Han, F., Salgado, R., Prezzi, M., Lim, J., 2017. Shaft and base resistance of non-displacement piles in sand. Comput. Geotech. 83, 184–197.

Houlsby, G.T., 2016. Interactions in offshore foundation design. Géotechnique 66 (10), 791–825.

ISO, 2016. Petroleum and natural gas industries – specific requirements for offshore structures. Part 4: Geotechnical and foundation design considerations. Geneva, Switzerland.

Jamiolkowski, M.B., Lo Presti, D.F.C., Manassero, M., 2003. Evaluation of relative density and shear strength of sands from CPT and DMT. Soil behaviour and soft ground constructions. ASCE, Geotechnical Special Publication 119, 201–238.

Jardine, R.J., Standing, J.R., Chow, F.C., 2006. Some observations of the effects of time on the capacity of piles driven in sand. Géotechnique 56 (4), 227–244.

Jeanjean, P., Zhang, Y., Zakeri, A., Andersen, K.H., Gilbert, R., Senanayake, A., 2017. A framework for monotonic p-y curves in clays. In: Keynote Lecture, OSIG SUT Conference, 2017, London.

Kang, F., Han, S., Salgado, R., Li, J., 2015. System probabilistic stability analysis of soil slopes using Gaussian Process regression with Latin Hypercube Sampling. Comput. Geotech. 63, 13–25.

Kolk, H.J., Baaijens, A.E., Vergobi, P., 2005b. Results of axial load tests on pipe piles in very dense sands: the EURIPIDES JIP. In: Proc. Int. Symp. on Frontiers in Offshore Geomechanics, ISFOG. Taylor & Francis, London, pp. 661–667.

Kolk, H.J., Baaijens, A.E., Sender, M., 2005a. Design criteria for pipe piles in silica sands. In: Proc. Int. Symp. on Frontiers in Offshore Geotechnics. Taylor & Francis, London. pp. 711–716.

Kulhawy, F.H., Mayne, P.W., 1990. Manual on estimating soil properties for foundation design. Electric Power Research Institute EL-6800, Project 1493–6. Electric Power Research Institute, Palo Alto, Calif.

Lambert, S., Toe, D., Mentani, A., Bourrier, F., 2021. A meta-model-based procedure for quantifying the on-site efficiency of rockfall barriers. Rock Mech. Rock Eng. 54, 487–500.

Le Gratiet, L., Marelli, S., Sudret, B., 2017. Metamodel-based sensitivity analysis: Polynomial Chaos Expansions and Gaussian Processes. In: Ghanem, R., Higdon, D., Owhadi, H. (Eds.), Handbook of Uncertainty Quantification. Springer, Cham. https://doi.org/10.1007/978-3-319-12385-1_38.

Lehane, B.M., Schneider, J.A., Xu, X., 2005. The UWA-05 method for prediction of axial capacity of driven piles in sand. In: Proc. of Int. Symp. on Frontiers in Offshore Geomechanics ISFOG, Taylor & Francis, London, pp. 683–689.

Lehane, B.M., Li, L., Bittar, E.J., 2020. Cone penetration test-based load-transfer formulations for driven piles in sand. Géotech. Lett. 10 (4), 568–574.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 21 (2), 239–245.

Mentani, A., Govoni, L., Bourrier, F., 2022. The use of CPT based metamodels to predict the performance of offshore anchor piles. In: Proc. of 5th Int. Symposium on Cone Penetration Testing (CPT'22), Bologna, Italy. pp. 1016–1022. doi:10.1201/9781003308829.

Page, A.M., Grimstad, G., Eiksund, G.R., Jostad, H.P., 2019. A macro-element model for multidirectional cyclic lateral loading of monopiles in clay. Comput. Geotech. 106, 314–326.

Randolph, M.F., 2003. Science and empiricism in pile foundation design. Géotechnique 53 (10), 847–875.

Randolph, M.F., 1994. Design methods for pile groups and piled rafts. In: Proc. 13th Int. Conf. Soil Mech. Found. Engng, New Delhi. Vol. 5, pp. 61–82.

Robertson, P.K., 2009. Interpretation of cone penetration tests – a unified approach. Can. Geotech. J. 46, 1337–1355.

Said, I., De Gennaro, V., Frank, R., 2009. Axisymmetric finite element analysis of pile loading tests. Comput. Geotech. 36, 6–19.

Sallaberry, C.J., Helton, J.C., Hora, S.C., 2008. Extension of Latin hypercube samples with correlated variables. Reliability Eng. Syst. Safety 93 (7), 1047–1059.

Schmoor, K.A., Achmus, M., Foglia, A., Wefer, M., 2018. Reliability of design approaches for axially loaded offshore piles and its consequences with respect to the North Sea. J. Rock Mech. Geotech. 10 (6), 1112–1121.

Soubra, A.H., Al-Bittar, T., Thajeel, J., Ahmed, A., 2019. Probabilistic analysis of strip footings resting on spatially varying soils using Kriging metamodeling and importance sampling. Comput. Geotech. 114, 103107.

Stuyts, B., 2020. Data science applications in geo-intelligence. In: Proc. 4th Int. Symposium on Frontiers in Offshore Geotechnics, ISFOG-04, Austin, Texas, pp. 120–149.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. 93, 964–979.

Systèmes, D., 2020. Abaqus Analysis User's Manual. SIMULIA, Providence, RI, USA.

Teixeira, R., Nogal, M., O'Connor, A., 2021. Adaptive approaches in metamodel-based reliability analysis: a review. Struct. Saf. 89, 102019.

Toe, D., Mentani, A., Govoni, L., Bourrier, F., Gottardi, G., Lambert, S., 2018. Introducing meta-models for a more efficient hazard mitigation strategy with rockfall protection barriers. Rock Mech Rock Eng. 51, 1097–1109. https://doi.org/10.1007/s00603-017-1394-9.

Tovar-Valencia, R.D., Galvis-Castro, A., Salgado, R., Prezzi, M., 2018. Effect of surface roughness on the shaft resistance of displacement model piles in sand. J. Geotech. Geoenviron. Eng. 144 (3), 04017120.

van den Eijnden, A.P., Schweckendiek, T., Hicks, M.A., 2021. Metamodelling for geotechnical reliability analysis with noisy and incomplete models. Georisk: Assessment Manage. Risk Eng. Syst. Geohazards. https://doi.org/10.1080/17499518.2021.1952611.

Wang, Z.Z., Xiao, C., Goh, S.H., Deng, M.-X., 2021. Metamodel-based reliability analysis in spatially variable soils using convolutional neural networks. J. Geotech. Geoenviron. Eng. 147 (3), 04021003.

Wiener, N., 1938. The homogeneous chaos. Amer. J. Math. 60 (4), 897–936.

Yang, Z.X., Jardine, R.J., Guo, W.B., Chow, F., 2015. A new and openly accessible database of tests on piles driven in sands. Géotech. Lett. 5, 12–20.

Yang, Z.X., Guo, W.B., Jardine, R.J., Chow, F., 2017. Design method reliability assessment from an extended database of axial load tests on piles driven in sand. Can. Geotech. J. 54, 59–74.