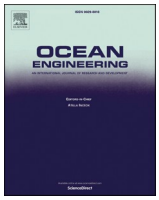
Ocean Engineering 284 (2023) 115095

Contents lists available at ScienceDirect

Ocean Engineering

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



# Integrated and data Science-informed seabed characterisation for optimised foundation design

M.P. O'Neill [a,*], A.L. Osuchowski [a], Y. Cai [a], M.F. Bransby [a], P.G. Watson [a], C. Gaudin [a], J. Doherty [a], E. Dalgaard [b], R. Ross [c]

[a] ARC ITRH TIDE, Oceans Graduate School, The University of Western Australia, Perth, Australia
[b] SolidGround, Copenhagen, Denmark
[c] Qeye Labs, Perth, Australia

ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

ABSTRACT

Unlike traditional offshore oil and gas projects which generally involve a limited number of structures and are confined to a relatively small seabed footprint, renewable energy projects such as offshore wind farms can comprise many tens or even hundreds of structures dispersed across an extensive area. Driven by the need for offshore wind to reduce costs while also minimising project risk, there is currently a strong focus within the offshore industry on identifying new ways to extract the full benefit from all available site investigation data and carry this benefit through to engineering design. This paper provides a demonstration of two such approaches that utilise geo-data from offshore Western Australia, potentially representative of conditions of future offshore wind farms; the first is an integrated approach combining geophysical and geotechnical data through a seismic inversion process, while the second encompasses statistical analysis of geotechnical cone penetrometer and soil strength test data combined with a Bayesian compressive sampling-based spatial interpolation method. The demonstrations yielded useful findings about the methodologies and associated input requirements. It is envisaged this work will lead to the development of an efficient integrated framework for interpreting geo-data that will inform future offshore site investigation and geotechnical design practice.

## 1. Introduction

With any offshore fixed or floating wind development, uncertainty associated with defining the seabed geotechnical conditions can lead to significant project cost and risk. The layout of infrastructure, and subsequent assessment of foundations, anchoring systems and cables, requires geotechnical knowledge of the seabed that may not exist at specific locations, and at best is likely to be sparse along the route of individual power cables used in a wind farm development. To deal with this, the traditional design approach involves selection of conservative 'design lines' representing key soil properties, which can result in foundations and anchors that may be larger than required, potentially incurring increased cost of fabrication, transportation and installation. Conversely, unconservative design lines may expose the infrastructure to installation and operational (in-place) risks. Design lines are typically selected by combining all available geotechnical, geophysical and geological information, with engineering judgement and experience also playing a key role. While the level of installation or operational reliability achieved using such approaches is difficult to quantify, the view of the authors is that the multiple conservatisms inherent within traditional design processes normally combine to give over-safe (and hence over-expensive) design outcomes.

A subject of ongoing discussion within the offshore industry is the possibility to accept a lower target level of reliability for offshore wind developments (relative to oil and gas installations) due to reduced consequences of failure – as these are less likely to lead to loss of life or significant damage to the natural environment. Other considerations for design include the typically large spatial extent of offshore wind developments, and the need to continually reduce the cost of wind energy to achieve parity with other energy sources. Hence, there is a strong and ever-growing need to develop and implement new and smarter approaches to offshore infrastructure design.

Recent advancement of statistical methods to handle large volumes of data, combined with increases in computational power (at low

* Corresponding author.
*E-mail address:* michael.oneill@uwa.edu.au (M.P. O'Neill).

https://doi.org/10.1016/j.oceaneng.2023.115095
Received 28 March 2023; Received in revised form 1 June 2023; Accepted 10 June 2023

relative cost) to perform repetitive complex calculations, provide the opportunity for rigorous examination of uncertainty and the quantitative evaluation of different approaches to aid geotechnical design. In this context, potential developmental steps include:

1. Establishing (and validating) industry-implementable methods that quantify point-to-point uncertainty in geotechnical properties at individual sites. Such approaches will need to consider the proximity of the location of interest to nearby geotechnical information, compared to the scale of fluctuation of the seabed properties. The integrated use of geophysical data to 'bridge' between sparse geotechnical data is likely to be beneficial in this regard.
2. Establishing relationships between the level of uncertainty in key geotechnical data and the associated consequences for foundation design in terms of installation, in-place capacity and serviceability (settlement, tilt), as well as for optimum decision making in terms of cost and/or residual project risk. This may involve more informed selection of design lines and/or partial factors in deterministic (e.g. LRFD) design, or direct use of statistical design methods.
3. Developing a framework to optimally locate geotechnical site investigations (SIs) based on relationships between cost (risk) and benefit (value). The aim of this would be to accommodate maximum flexibility of infrastructure layout, whilst also ensuring the highest value data are acquired:
   - For new ('green-field') developments, such a framework would inform the selection of SI locations relative to a preliminary field layout, but with the goal of ensuring value from the acquired data is retained as the field layout matures (and changes) through the development lifecycle.
   - For expansion or life-extension of existing ('brown-field') developments, such a framework can inform decisions on whether additional SI is required to 'fill gaps', or whether the in-hand data are sufficient to ensure uncertainty (risk) is at acceptable levels for design.

Development of SI planning tools may also create opportunities to modify (in real time) field campaigns via real-time analysis of recently-acquired data and 'live' assessment of remaining uncertainty at infrastructure locations.

This paper relates a body of interconnected work to investigate existing approaches and identify and refine new approaches in the use of geo-data. It is anticipated that the work will involve development of new mathematical models, as well as the application of statistical approaches that are gaining in popularity in onshore geotechnical engineering and now being extended to offshore applications. It is noted that most of the current statistical approaches typically use geotechnical data only (for example, cone tip resistance profiles and/or point measurements of geotechnical properties from sampling) to make predictions of point-to-point geotechnical properties and their uncertainty. It is believed that the integration of geophysical data into these statistical approaches will lead to a reduction of residual uncertainty, particularly in seabeds that have non-stationary statistical properties (i.e. have layers of material with different properties). The overall objective of this research is to develop an integrated workflow as illustrated on Fig. 1, with key aspects being the utilisation of geophysical and geotechnical input data from different sources, the integrated analysis and modelling of these data, and the implementation of the analysis output directly into a reliability-based foundation assessment.

To illustrate the above, this paper presents selections of the workflow as follows:

- We first review quantitative approaches to incorporate geophysical and geotechnical data for the purposes of site characterisation ('A2' on Fig. 1). Case Study 1 introduces seismic inversion, using data collected at a carbonate site offshore Western Australia – which was selected to investigate the suitability of these approaches in characterising the types of seabed conditions expected to be encountered in upcoming offshore wind farm developments on the Australian continental shelf (and elsewhere in the world where carbonate sediments are prevalent such as the regions shown in Watson et al., 2019).

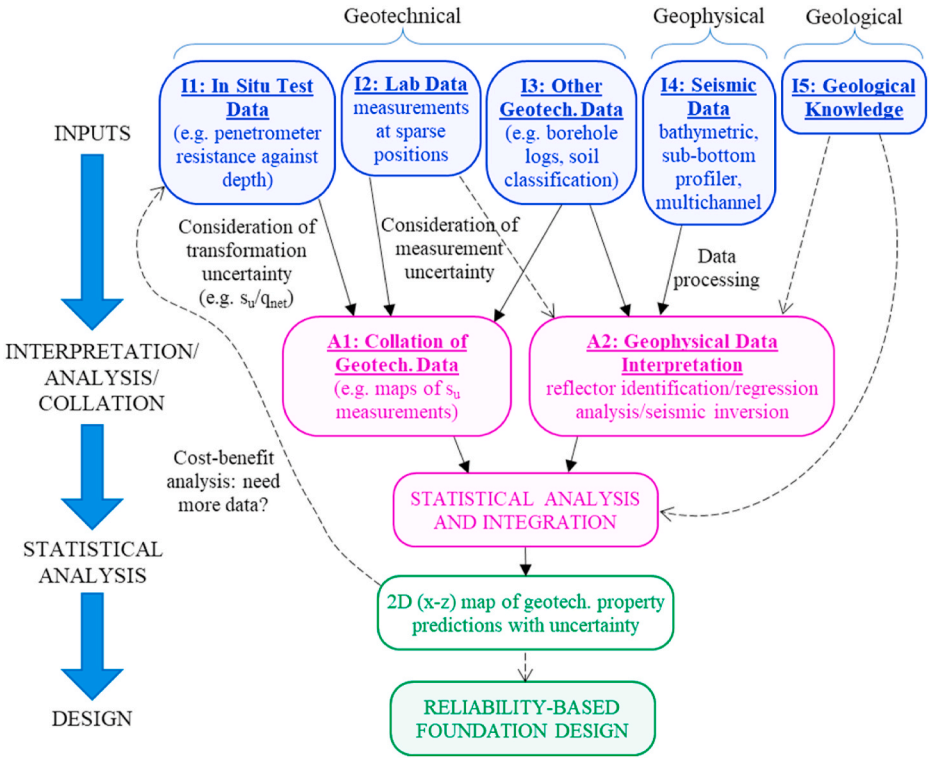



**Fig. 1.** Schematic showing integrated workflow combining seismic and geotechnical data to facilitate reliability-based foundation design.

- We then demonstrate the application of statistical methods to quantify geotechnical data at 'sampled' locations, and to extrapolate beyond this data to predict properties (and quantify their uncertainty) at 'unsampled' locations. Case Study 2 uses cone penetrometer test (CPT) and laboratory data to explore the left hand branch of Fig. 1 (without using geophysical or geological information), in order to explore how geotechnical-only statistical approaches fare with complex data, including consideration of the transformation uncertainty from net cone tip resistance ($q_{net}$) to undrained shear strength ($s_u$) – which is known to be particularly challenging for carbonate soil conditions.

We then conclude by summarising the key messages from the two case studies, providing implications for design practice, and making suggestions for future work.

## 2. Maximising value from data: quantitative analysis of geophysical data for site characterisation

### 2.1. Marine seismic reflection overview

Marine seismic reflection data are typically acquired in the early stage of an offshore project as part of the geophysical survey, providing initial insight into subsurface conditions over large areas. Other geophysical data sets typically collected include single/multibeam echosounder and side-scan sonar data, which provide information primarily about the seafloor, including bathymetry and surface features. In seismic reflection surveys, the acoustic energy is of sufficient amplitude and frequency such that it can penetrate through the seafloor and deeper into the subsurface, thus providing information to a target investigation depth.

Seismic data sets with differing vertical resolution (defined as the thinnest layer which can be interpreted in the data) can be collected during a seismic reflection survey. There is a trade-off between vertical resolution and depth of penetration; higher frequency acoustic energy will result in higher vertical resolution data but the energy attenuates more quickly and thus will have limited penetration, whereas lower frequency energy will penetrate deeper but produces lower resolution data. Typically, complimentary data sets meant to capture both high vertical resolution and sufficient penetration depth are collected during a marine seismic reflection survey. One data set may provide decimetre-scale vertical resolution in the shallow subsurface (e.g. seafloor to 50 m depth) commonly collected through the use of a sub-bottom profiler (SBP) system, also referred to as acoustic profiling. A second data set may provide lower (metre-scale) vertical resolution but with greater penetration (e.g. to around 200 m depth), often referred to as ultra-high resolution seismic (UHRS) data. SBP data are commonly acquired through the use of a single source-receiver arrangement (termed 'single channel'), which may be hull-mounted on a survey vessel or placed onboard a tow device or autonomous underwater vehicle. In many cases, SBP systems use a swept frequency source such as a Chirp device (e.g. Schock, 1989). In contrast, UHRS data are commonly collected using a multichannel streamer with an air gun (or similar) source arrangement. The streamer trails from the survey vessel providing horizontal offsets between the source and trailing receivers, in an arrangement more closely resembling surveys performed for hydrocarbon prospecting. Surveys may be designed such that two-dimensional (2D) or three-dimensional (3D) data are collected. Although 3D surveys are now commonplace in hydrocarbon exploration, they are rarely performed for shallow seabed investigation. Instead, a pattern of closely-spaced 2D 'main' lines with additional intersecting 'tie-in/cross' lines are performed across the survey area; interpolation between these adjacent 2D lines then provides a way of achieving 'pseudo-3D' (also termed '2.5D') coverage.

During a seismic reflection survey, sound waves (P waves) are produced by the source and propagate outwards as a spherical wavefront. The downward traveling portion of the wavefront propagates to the seafloor, whereupon part of the acoustic energy is reflected and part is transmitted deeper into the subsurface; some energy is also lost to scattering and attenuation. The reflection at the seafloor is produced because the propagating P wave encounters a contrast in acoustic impedance ($Z$) – the product of P-wave velocity ($V_P$) and bulk density ($\rho_b$). The values of acoustic impedance on either side of a reflecting interface affect the nature (i.e. amplitude, phase and frequency) of the resulting reflection. For the normally incident (vertically propagating) portion of the P wave, the amplitude of the reflection is determined by the reflection coefficient ($RC$) of the interface, calculated using Equations (2)–(1) below (Dix, 1952):

$$RC = \frac{Z_2 - Z_1}{Z_2 + Z_1} = \frac{\rho_{b2} V_{P2} - \rho_{b1} V_{P1}}{\rho_{b2} V_{P2} + \rho_{b1} V_{P1}} \quad \text{(Eq. 2-1)}$$

where $Z_1$, $\rho_{b1}$, and $V_{P1}$ are acoustic impedance, bulk density and P-wave velocity of the upper layer, respectively, and $Z_2$, $\rho_{b2}$, and $V_{P2}$ are those of the lower layer, as illustrated on Fig. 2. When transitioning from a higher to lower acoustic impedance layer, as from Sediment layer 1 to Sediment layer 2 on Fig. 2, a negative reflection coefficient is produced and a trough is seen in the seismic trace.

The reflected energy is recorded using hydrophone receivers which measure changes in seawater pressure (compressions and rarefactions) as P waves pass by – these measurements being equated to seismic amplitude. The recordings of seismic amplitude are sampled in time with a single recording referred to as a seismic ('wiggle') trace, like the one shown on Fig. 2. A collection of seismic traces acquired over an area of seabed then makes up a seismic data set. In the case of 2D seismic data, the series of seismic traces are collected along a transect, and these can be plotted together, either as a series of 'wiggle' traces or by using a colour mapping scheme, thereby creating a cross-sectional image of the subsurface called a seismic section. Fig. 3a provides an example of a 2D seismic section which can be conceptualised as a spreadsheet containing seismic amplitude values, constructed of many rows and columns, with each cell in the spreadsheet (i.e. pixel in the image) representative of a single seismic amplitude measurement. In a multichannel seismic survey, the data include multiple traces imaging the same point on the seabed but collected at different offsets between source and receiver (referred to as multifold coverage) and may be analysed in either a pre-stack or post-stack condition. In a pre-stack data set the individual traces are retained, and in a post-stack data set the traces are averaged together in a multi-step process called stacking to create a single trace for each imaged point on the seabed (Telford et al., 1990; Barclay et al., 2008). This typically improves the signal-to-noise ratio but removes the variation in amplitude with offset present in the pre-stack traces. These and other data processing techniques (e.g. those listed on Fig. 4, Step 1) are applied to the seismic data to progress them from their initial form – essentially the results of a field experiment measuring the interaction of sound with the surrounding environment – to a seismic section that is a spatially accurate representation of subsurface geological structure.

Although seismic amplitude is commonly displayed on a seismic section, other seismic 'attributes' such as instantaneous frequency or phase may also be computed and displayed for analysis purposes (Taner et al., 1979). The vertical axis of a seismic section is commonly presented as two-way travel time (TWT) which is the time required for acoustic energy to leave the source, propagate to a reflecting interface, then return back to the receiver(s). In order to convert from TWT to measured depth, the P-wave velocities of the water column and underlying sediment and rock layers must be known or estimated. This assemblage of different P-wave velocities across the seismic section is referred to as the velocity model. The time-to-depth conversion using the velocity model brings the seismic data set into the depth domain, where it can be compared with geotechnical data, which by nature are measured in depth.

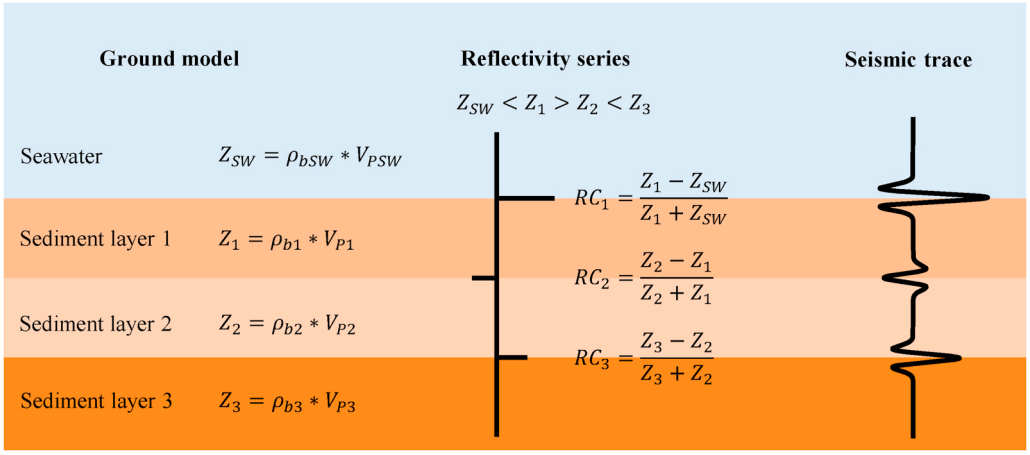
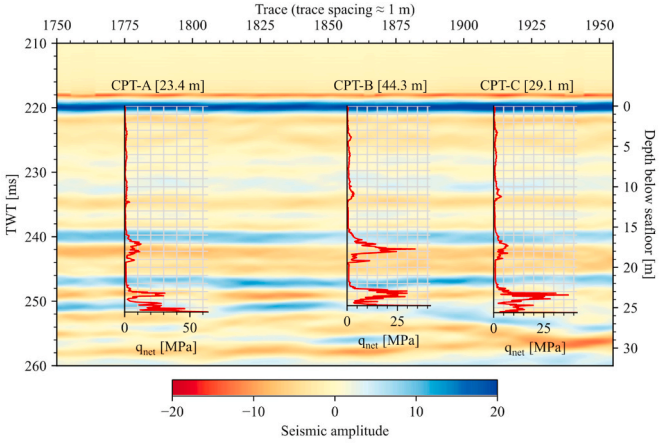
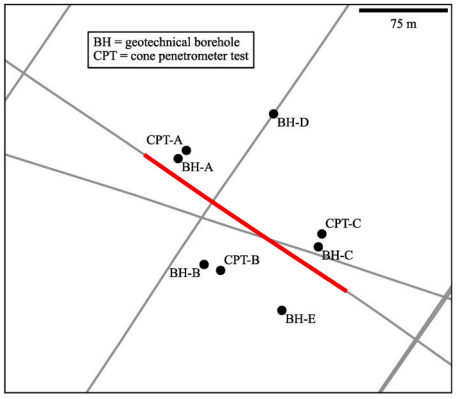



**Fig. 2.** Acoustic properties of sediment layers and idealised (noise-free, zero phase) seismic trace.




(a): Section view showing seismic amplitude data and $q_{net}$ profiles for CPTs A, B and C (numbers in square brackets indicate lateral offset distance between CPT and seismic line).




(b): Plan view showing 2D seismic (grey) lines and geotechnical locations, red line indicates the segment of seismic line shown in profile in (a).

**Fig. 3.** Case Study 1 site.

### 2.2. Qualitative interpretation to quantitative analysis

For the purposes of geotechnical site characterisation, initial site assessment and preliminary development of a ground model is often based on qualitative interpretation of the seismic reflection data. This visual interpretation may be aided by shallow, widely-spaced ground truth sampling performed at the time of the geophysical survey. If time permits, the initial ground model can be used to help plan the geotechnical investigation, by targeting investigation sites at areas exhibiting complex seabed conditions and by aiming to capture geotechnical information at each relevant subsurface unit interpreted in the seismic data. Once the geotechnical survey has been completed, the newly available in situ and laboratory geotechnical data are then used to assign index and mechanical parameters to each geotechnical unit. At this stage, the seismic and geotechnical data can be brought together, using simultaneous inspection of both data sets to revisit the geotechnical unit interpretations, and interpolate/extrapolate them across the 2D sections covered by the seismic survey (Birchall, 2012; Rattley et al., 2017; Thomas, 2017). The geotechnical data may also be used to refine the geophysical interpretation, providing a means for evaluating and updating the seismostratigraphic and velocity models. These processes are naturally iterative, and benefit from close communication between geophysicists, geologists, geotechnical engineers and other geo-practitioners, whose definitions and methods for describing the subsurface may be fundamentally different and are complementary in nature. A combined understanding and appreciation for both seismic and geotechnical data, and an effort to harmonise the two, allows for greater overall insights to be gained regarding subsurface conditions at a site.

The interpretive process described above uses the seismic data in a largely qualitative manner, based on visual inspection of the data rather than quantitative assessment. A variety of quantitative approaches are presented in the literature, some of which are gaining traction in terms of use in industry. Characterisation of the seafloor and shallow seabed is one such research area, where various methods involving analysis of the amplitude, phase and frequency of shallow reflections are used to predict seafloor sediment type and associated physical properties (e.g. Schock et al., 1989; Leblanc et al., 1992; Panda et al., 1994; Bull et al., 1998; Stevenson et al., 2002; Lyu et al., 2021), noting that the Chirp SBP system was designed primarily with this purpose in mind (Schock, 1989). Another category of quantitative seismic analysis which extends beyond the seafloor and deeper into the subsurface is known as seismic inversion. Seismic inversion is a broad term encompassing different types of analyses, but in general means to use a set of observations (seismic data) to estimate the causal factors (subsurface conditions) that produced them. The output of the seismic inversion process then is a numerical model representing some physical property of the subsurface (Barclay et al., 2008). Originally developed for hydrocarbon prospecting, seismic inversion strategies have more recently been adapted to suit shallow, high frequency subsurface data. As Table 2 in Vardy et al. (2017) indicates, there is a substantial body of published work in this area, with seismic inversion analyses generally divided into the categories of reflectivity, attenuation and full waveform inversion. The most recent studies, largely falling under the reflectivity inversion

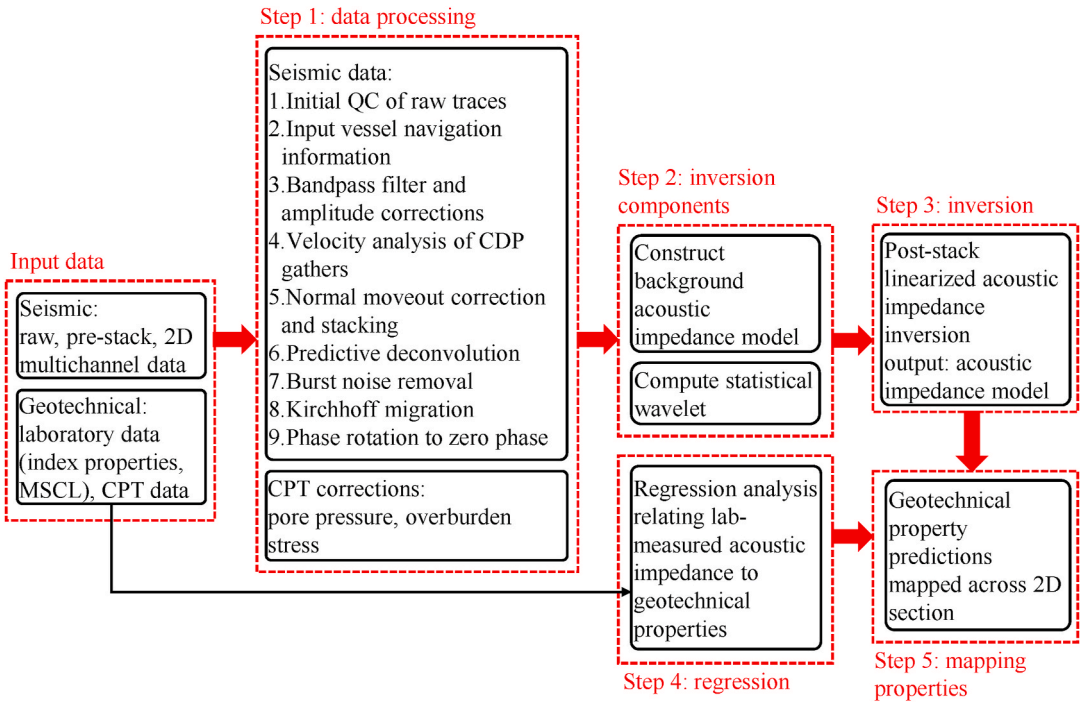



**Fig. 4.** Seismic inversion workflow developed for this study.

category, venture further from the traditional geophysical realm and have begun tailoring seismic inversion outputs specifically to geotechnical characterisation and design (e.g. Vardy, 2015; Vardy et al., 2015; Forsberg et al., 2017; Vardy et al., 2018; Sauvin et al., 2019; Chen et al., 2020; Chen et al., 2021; Karkov et al., 2022). The post-stack acoustic impedance workflow of Vardy (2015) uses a stochastic genetic algorithm optimisation technique which allows for the assignment of probability distributions to inverted estimates of acoustic impedance and empirically correlated geotechnical properties. Using pre-stack seismic data, Karkov et al. (2022) implement an amplitude-versus-offset inversion approach to invert for bulk density, bulk modulus and shear modulus, ultimately relating these inverted parameters with CPT data to produce predictions of tip resistance, sleeve friction and pore pressure. In some cases the outcome of the inversion process is fed directly into foundation calculations, for example using estimates of maximum shear modulus ($G_{max}$) produced through seismic inversion as the primary input for lateral pile response analysis, such as the Pile Soil Analysis (PISA) methodology (Byrne et al., 2017).

Much of the recent work has focused in and around the North Sea region, and for bottom-founded offshore wind farm sites where water depths are modest. There are limited published examples of these approaches being implemented for the shallow carbonate environments on the continental shelf of Australia, or for carbonate regions in other parts of the world. Carbonate sediments, consisting primarily of the skeletal remains of marine organisms, have been shown to exhibit substantial differences in both their physical characteristics and engineering behaviour compared to their siliceous counterparts (Semple, 1988; Sharma and Joer, 2015; Watson et al., 2019). Post-depositional cementation may also occur to varying degrees (i.e. light cementation between sediment grains to continuous, rock-like layers) which can have a large impact on geotechnical engineering design. These and other factors make engineering in these environments uniquely challenging.

With the offshore wind industry currently in an early development stage in Australia, it is anticipated that site investigation work followed by the design and construction of infrastructure will occur in coming years. The ability to implement improved quantitative site characterisation strategies in these (carbonate) environments holds high potential value for these projects. This provides the impetus for the work presented in this section, which demonstrates the application (and current limitations) of a quantitative seismic inversion workflow for geotechnical characterisation of carbonate sites on the Australian continental shelf.

### 2.3. *Case study 1 – Geophysical data overview and summary of subsurface conditions*

The work presented in this section uses a data set from the northwestern continental shelf of Australia, in 80–200 m water depth. Two 2D seismic surveys were performed at the site – Chirp SBP and multichannel UHRS. This study utilises the multichannel data which were acquired using a small volume air gun source and 24-channel, 46-m-long hydrophone streamer. Main lines from the survey were spaced approximately 200 m apart, with a collection of cross lines interspersed, providing roughly uniform coverage over the 8 × 15 km survey area (note the survey footprint is similar to that of an offshore wind farm). The frequency bandwidth of the multichannel seismic data is approximately 50–500 Hz. Seabed sampling was also performed as part of the geophysical survey, including shallow box and piston core sampling, with the latter penetrating to a maximum depth of 5 m below seafloor. Ultrasonic (230–500 kHz) P-wave velocity ($V_p$) measurements were made on the piston core samples using multi-sensor core logger (MSCL) equipment. A subsequent geotechnical investigation included more extensive sampling within a smaller region of the overall site, along with in situ CPT and T-bar testing. A majority of the investigations were made within the upper 50 m below seafloor, with one sampling borehole and one CPT extending to 250 m depth. A laboratory programme was then conducted, including classification and intact element testing of a selection of the recovered samples. All data produced during each stage of the investigation were provided for this research.

For Case Study 1, a selection of the available geophysical and geotechnical data were used, specifically a subset of a single 2D multichannel seismic line and adjacent geotechnical data. The selected seismic section, plotted together with profiles of CPT net cone resistance ($q_{net}$), is shown on Fig. 3a. A plan view showing geotechnical CPT and borehole locations in relation to the seismic line is provided on Fig. 3b.

Per Fig. 3a, the section of seabed considered in this study comprises a

succession of largely horizontal to moderately subhorizontal layers, which comprise carbonate sediments ranging from muds to sands with carbonate ($CO_3$) content typically greater than 90%. Layers of carbonate silty sand to sandy silt, with $q_{net}$ ranging from near zero to approximately 4 MPa, are present from seafloor to a depth of approximately 15 m. A localised increase in $q_{net}$ can be seen at a depth of approximately 11 m; this lens is reported in the geotechnical core logs as being very weakly to moderately cemented and containing fine gravel-sized shells, bioclasts, lithified tube worm burrows and calcarenite nodules. A more substantially cemented layer is evident from approximately 15-19 m depth, with recovered samples containing gravel-sized nodules of weakly to moderately cemented calcarenite, and with $q_{net}$ ranging from approximately 13-35 MPa. This layer is underlain by a soft carbonate sandy mud layer approximately 4 m in thickness. Below the sandy mud layer is a highly variable zone consisting of interlayered regions of uncemented and cemented sediment (calcarenite). This variable stratigraphy is highlighted by the erratic nature of the $q_{net}$ measurements ranging as high as 65 MPa, but also dropping as low as 2 MPa. The thickness of this variably-cemented stratum is approximately 10 m, the base of which was the end depth for this study.

When considering the seismic amplitude data on Fig. 3a, the seismic reflections are largely consistent with observations from the geotechnical data. The strongest (highest seismic amplitude) reflection is observed at the seafloor, followed by another continuous reflection at approximately 16 m depth, situated in near proximity to the top of the first substantially cemented zone. A series of more complex, discontinuous reflection events are present in the variably-cemented region across the bottom of the section.

#### 2.3.1. Seismic inversion workflow

In order to expand upon the initial characterisation described above, a quantitative seismic inversion workflow was developed and applied to the seismic section shown on Fig. 3a. A diagram outlining the overall workflow is presented on Fig. 4, and the individual components (steps) are described in further detail below.

**Step 1**. Data Processing

Raw seismic data collected in the field are generally not suitable for direct use in inversion, thus requiring an initial processing stage. There are numerous data processing techniques, and depending on the intended end use of the data (e.g. a qualitative or quantitative analysis) these techniques are selected from and applied – typically as a process of sequential steps. In the case of quantitative seismic inversion, care must be taken to preserve the original amplitude and frequency content of the signal. For this study, the initial seismic data processing steps are listed on Fig. 4, a majority of which were performed by the data acquisition contractor.

The geotechnical data from the field also require processing, for example applying pore pressure and overburden corrections to the measured cone resistance ($q_c$) in order to compute $q_{net}$ (see Section 3.1 for more detail). This is particularly important if the ultimate goal of the inversion workflow is to produce synthetic CPT parameters, which is an area of ongoing research by the authors but is not presented as part of this study.

**Step 2**. Inversion Components

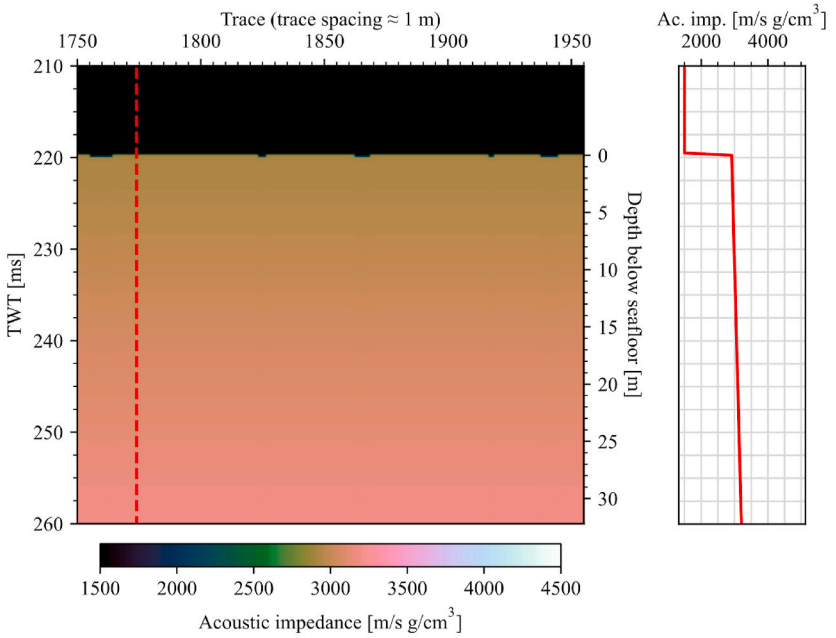
i. Background Acoustic Impedance Model

Inverting for acoustic impedance using seismic data alone will generate only a 'relative' acoustic impedance model – in which relative changes in acoustic impedance are determined across each seismic data point. While useful in a qualitative sense, and of value in sharpening the subsurface stratigraphy model, this does not lead to estimates of actual (absolute) values of acoustic impedance. To obtain these, it is necessary to develop a background acoustic impedance model, also referred to as a low frequency model (as the seismic data do not contain low frequency trends). When the background acoustic impedance model is coupled with relative changes in acoustic impedance produced from the inversion process, absolute values are obtained (Barclay et al., 2008). To enable this, $V_P$ and bulk density values, and therefore acoustic impedance, must be estimated across the full area of seismic data being inverted. In some cases they may be measured in situ, for example through PS logging (e.g. Nauroy et al., 1998), although this is rare for projects on the Australian continental shelf as it is difficult to maintain the necessary open borehole in soft near-surface sediments. For this study, the background acoustic impedance model was developed using laboratory MSCL measurements of $V_P$ and attenuated gamma density (an analogue of bulk density) made on intact shallow piston core samples. Due to the shallow depth of the piston cores (ranging from 1 to 5 m), the trend in acoustic impedance observed in the shallow zone was extrapolated down to the bottom depth of the study area. The background acoustic impedance model developed for this study is shown on Fig. 5a. At the time of writing, the authors are making similar geoacoustic measurements using triaxial bender element testing of intact samples recovered from greater depth, subjected to stresses representative of in situ conditions. The bender elements are used to measure $V_P$ within the triaxial specimen, which is then used to compute acoustic impedance. The results of this study will allow for future refinement of the acoustic impedance background model.

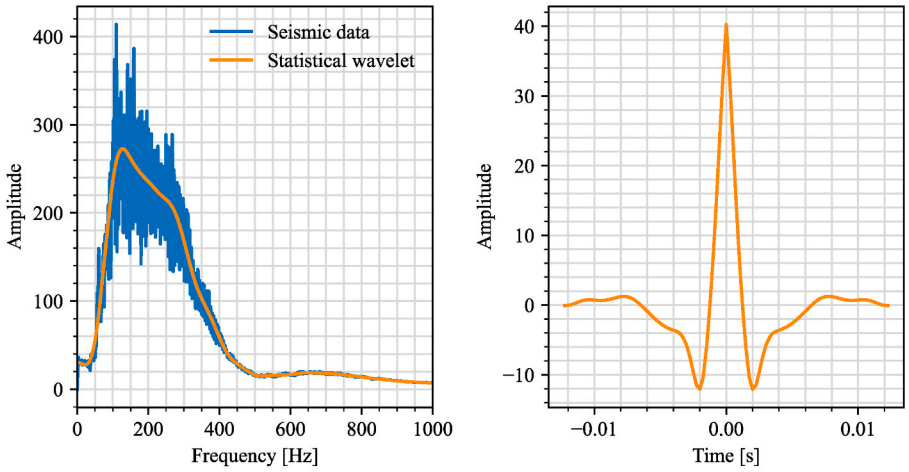
The term 'wavelet' refers to the acoustic signature (i.e. amplitude time series) of the P wave produced by the source used in the seismic survey. In some cases this may be measured (verified) during the survey by suspending a hydrophone receiver in the water column beneath the source. As this was not the case for the data set under study, the wavelet was computed from the seismic data (therefore termed a statistical wavelet) using a Fourier analysis procedure. The wavelet was made to be zero phase, meaning the central positive peak of the wavelet in the time domain represents the position of a reflector (using the positive polarity convention in which an increase in acoustic impedance across a reflecting interface is represented as a peak, like shown on Fig. 2). The computed statistical wavelet is shown in both the frequency and time domains on Fig. 5b. The amplitude spectrum of the wavelet is shown superimposed on that of the seismic data, demonstrating the consistency in frequency content between the two. An amplitude scaling was applied to the wavelet by first calculating the reflection coefficient of the seafloor using piston core MSCL and seawater acoustic impedance measurements in the vicinity of the seismic section included in the study. Using this information and Equations (2)–(1), a reflection coefficient of approximately 0.3 was estimated for the seafloor. The wavelet amplitude was then scaled such that the product of the reflection coefficient and wavelet amplitude was equivalent to the amplitude of the seafloor reflection contained in the trace nearest the ground truth measurements. The wavelet was assumed to be stationary, meaning it was not changed with lateral or vertical position across the section of seismic data being inverted.

**Step 3**. Inversion

The inversion computations were performed using the pylops.avo.poststack.PoststackInversion module from the open source PyLops library (Ravasi and Vasconcelos, 2020). This module implements a linear operator based on the convolutional forward model, in which the seismic wavelet is convolved with the reflectivity series calculated from the acoustic impedance model (the relationship between acoustic impedance ground model, reflectivity series, and resulting seismic trace is illustrated on Fig. 2). An iterative process is employed until a sufficient match is achieved between the input seismic data and the (synthetic) seismic section generated through application of the forward model to the estimated acoustic impedance model solution. The background acoustic impedance model and wavelet, generated separately, are provided as inputs to the PyLops module in order to run the inversion computations. As post-stack data were used in this study, each seismic trace is representative of the response of a normally incident P wave, and




(a): Background acoustic impedance model. 2D representation is shown on the left and a 1D representation (extracted from the single trace location identified by the red dashed line) is shown on the right.




(b): Amplitude spectra for the seismic data and statistical wavelet (left) and statistical wavelet shown in the time domain (right) prior to scaling based on nearby ground truth data.

**Fig. 5.** Components used in the seismic inversion process.

ii. Wavelet

amplitude-versus-offset variations were not considered. The output of the post-stack acoustic impedance inversion process is shown on Fig. 6, with the CPT $q_{net}$ profiles superimposed. Estimated absolute acoustic impedance values within the subsurface range from approximately 2200 m/s g/cm$^3$ to 4000 m/s g/cm$^3$. Subtle layer variations can be seen in the top 15 m of the profile, while the zones containing cemented sediments from 15 to 19 m, and from 23 m to the bottom of the profile, are apparent as zones of higher acoustic impedance, consistent with increases in the CPT $q_{net}$ measurements. The sandy mud layer from approximately 19-23 m is also evident, showing low acoustic impedance values.

**Step 4.** Regression

Although acoustic impedance is a physical property, it is not a parameter commonly used in geotechnical characterisation and design, and hence it is necessary (using the seismic inversion methodology presented here) to relate acoustic impedance to more common geotechnical engineering properties using empirical relationships. An example regression analysis performed for this study, which incorporates laboratory data from the study site as well as carbonate sediment data retrieved from the literature (Shumway, 1960; Johnson et al., 1977; Hamilton and Bachman, 1982; Richardson et al., 1997; Richardson and Briggs, 2004) is shown on Fig. 7a and b. In order to capture the spread in data points across different values of acoustic impedance, low estimate (LE), best estimate (BE) and high estimate (HE) trends were established, with the LE to HE interval representative of the percentile range between the 10th and 90th percentiles. For the bulk

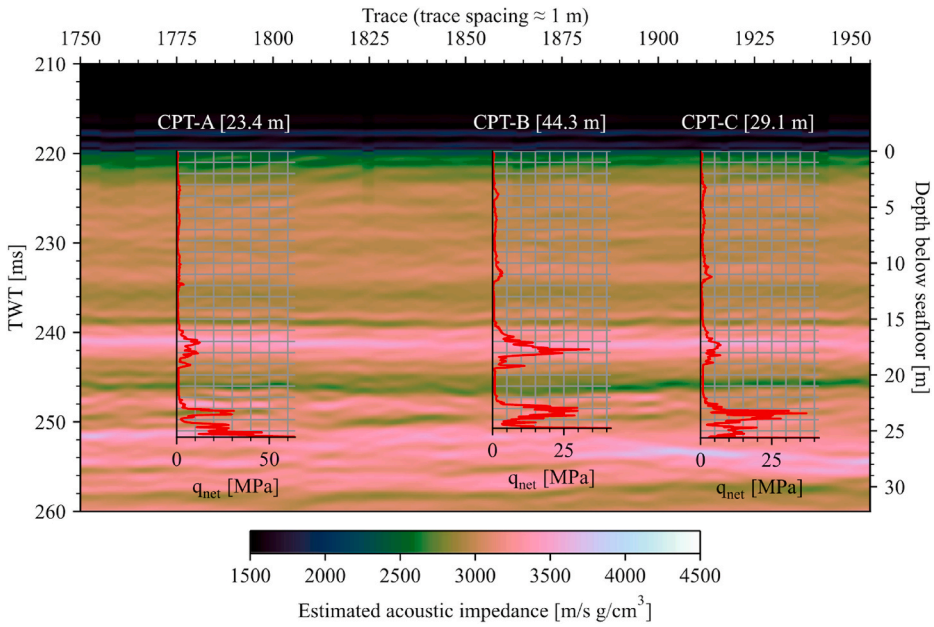
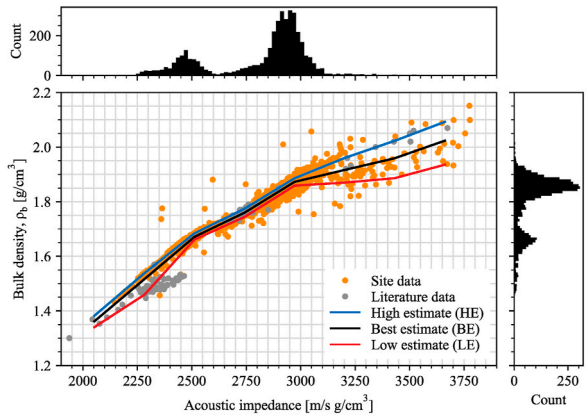
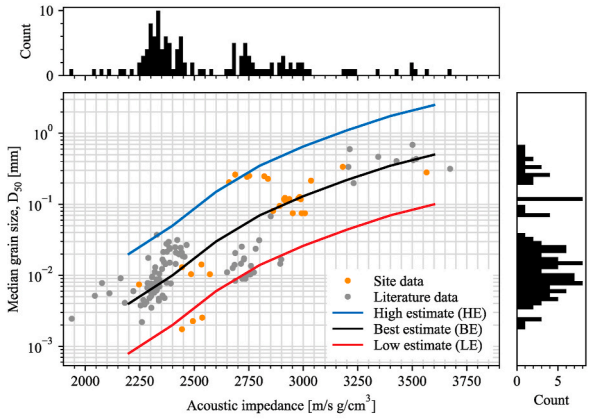



**Fig. 6.** Output of post-stack acoustic impedance inversion with CPT $q_{net}$ profiles superimposed.




(a): Relationship between acoustic impedance and bulk density.




(b): Relationship between acoustic impedance and median grain size (note data sets from the literature report mean grain size).

**Fig. 7.** Correlations between geophysical and geotechnical data developed in this study. Histograms are included to show number and distribution of data points.

density-acoustic impedance relationship shown on Fig. 7a, LE, BE and HE trends were computed using a statistical approach. Eight uniformly-sized acoustic impedance ranges ('bins') were first defined and within each bin a least squares linear interpolation was performed to establish BE lines. A search algorithm was then used to identify lines (constrained to the same slopes as the BE lines) at which 10% and 90% of the data points fell below. The centre points of these lines (contained within each bin) were then connected to produce the LE, BE and HE lines shown on Fig. 7a. Because of the lower number of data points for the median grain size-acoustic impedance relationship shown on Fig. 7b, the statistical approach described above was not implemented, instead assigning the LE, BE and HE lines manually using a visual approach.

Fig. 7a shows a strong relationship as bulk density is a component of acoustic impedance and thus the two parameters are closely linked. As can be seen on Fig. 7b, there is greater spread in the median grain size data, reflected in the greater distance between the LE, BE and HE lines. Although a relationship is apparent, the spread in the data is an indicator of the weaker physical linkage between acoustic impedance and median grain size. For example, a single grain size distribution parameter such as median grain size may be similar for two distinctly different sediment types having different bulk density and $V_P$ values, and therefore differing acoustic impedances.

**Step 5.** Mapping Properties

Using the correlations shown on Fig. 7a and b, predictions of bulk density and median grain size (including LE, BE and HE values) can now be mapped across the subsurface section by converting inverted acoustic impedance to values of bulk density and median grain size. To assess the validity of these predictions, ground truth data from the five geotechnical boreholes in the immediate study area (locations shown on Fig. 3b) were compared with inversion results from the trace location closest to each borehole (with lateral offset distances ranging from approximately 14 m–91 m). Best estimate predictions and ground truth data are presented on Fig. 8. The gray shaded regions in the plots represent the low to high estimate range based on the correlations presented on Fig. 7a and b.

### 2.3.2. *Discussion of results and recommendations for further improvement*

In reviewing the results presented on Fig. 8, the predictions output from the seismic inversion workflow are in reasonable agreement with the ground truth data points over some depth ranges, while in others

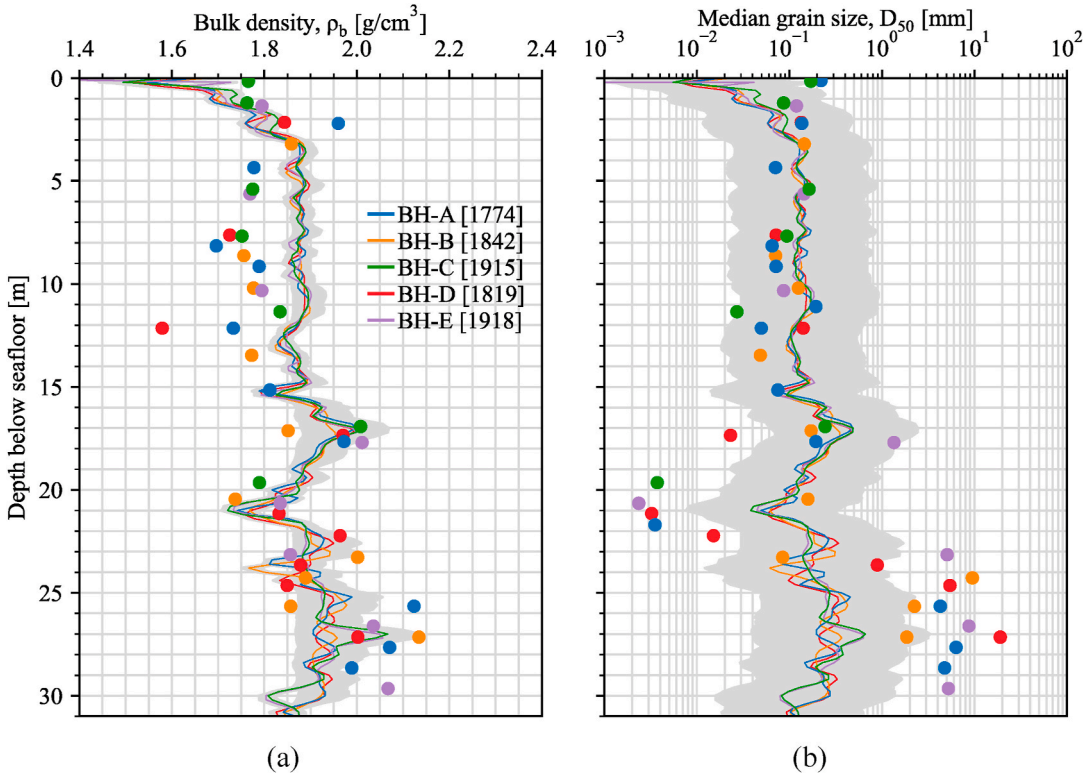



**Fig. 8.** Comparison of predictions from geophysical inversion and borehole sample data. The trace location in closest proximity to each geotechnical borehole is shown in the legend.

there is significant discrepancy. The workflow appears to over predict bulk density in the top 4–14 m, while under predicting in some regions of the variably-cemented zone from 23 m to the bottom of the profile. Similar variability is seen in the median grain size results, although the greater uncertainty in the correlation between acoustic impedance and median grain size is captured via the larger shaded region spanning the low to high estimate predictions.

Keeping in mind the offset distances between the inverted seismic traces and their closest geotechnical boreholes from which the ground truth data were acquired, some of the inconsistency between inversion output and ground truth data may be attributed to the variability of geotechnical conditions over (relatively) short lateral distances, this being especially true in the variably-cemented regions. Further explanation lies in the selected background acoustic impedance model (shown on Fig. 5a) which does not encode any of the geological layering evident at the site (as described in Section 2.3). Through this study, it was observed that the mathematical process underlying the inversion does not deviate substantially from the background model and thus an accurate background model is key to achieving good results. It is proposed to update the background model based on the first pass results, and then iterate until an acceptable correlation between inversion output and ground truth data is observed – and this work is ongoing. This and other aspects of the seismic inversion workflow are under continued development. Considering the case study as a whole, the insights gained by employing this type of integrated quantitative workflow are promising and with further improvement, the spatial mapping of geotechnical property predictions across 2D sections can be more confidently implemented.

## 3. Maximising value from data: statistical analysis of geotechnical data for foundation design

Direct geotechnical measurement of the seabed is limited to specific locations that are sampled or probed. These locations are typically determined early in a project lifecycle and may not necessarily be aligned with or spatially optimal relative to the final location of infrastructure, such as wind turbine foundations or anchors. Engineering judgement is then needed to manage variability in ground conditions between (and sometimes beyond) the investigated locations. Such judgement can (as discussed in the previous section) be supported by qualitative interpretation of geophysical information, including through reflector picking to define individual soils units – but spatial interpolation carries a level of uncertainty that needs to be reflected in design. Additionally, measurement uncertainty (associated with the equipment/tools used to acquire the data) and interpretation (or transformation) uncertainty (where a fundamental soil parameter is inferred but not directly measured) need consideration.

The previous section introduced an approach to address some of the above uncertainties through direct use of geophysical data. The following section introduces the application of statistical methods to maximise value from geotechnical data for use in foundation design. This study considers a data set comprising field CPT data from a different carbonate soil site offshore north-western Australia, along with soil strength data obtained from laboratory tests.

### 3.1. Case study 2 – Geotechnical data overview and summary of subsurface conditions

Fig. 9 presents a layout of the site, showing the locations of five CPTs (CPT-A to CPT-E) that comprise the data set. These were performed along the indicated survey line at horizontal ($x$) spacings of 34–38 m, with a total horizontal distance of 142 m separating CPT-A from CPT-E. Based on geotechnical laboratory testing performed on samples from the site, the sediment comprises high plasticity carbonate muddy silt and silty mud, with localised thin layers of carbonate silty/muddy sand.

Profiles of the sampled (measured) cone resistance ($q_c$) and pore pressure ($u_2$) versus depth below the seabed ($z$) are presented on Fig. 10, together with profiles of net cone resistance ($q_{net}$) and excess pore

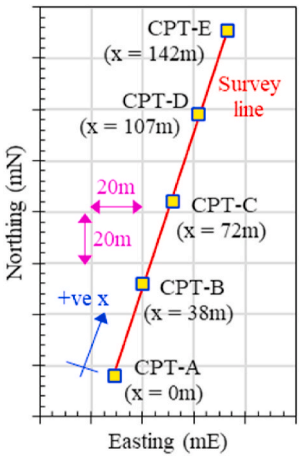



**Fig. 9.** Site layout showing CPT locations.

pressure ratio ($B_q$) calculated using:

$$q_{net} = q_c + u_2(1-\alpha) - \sigma_{vo} \quad \text{(Eq. 3-1)}$$

$$B_q = (u_2 - \gamma_w z) / q_{net} \quad \text{(Eq. 3-2)}$$

and where $\alpha$ is the cone tip area ratio, $\sigma_{vo}$ is the total in situ vertical stress and $\gamma_w$ is the unit weight of seawater. The CPT results in the upper 8 m (or so) are relatively uniform, with low but steadily increasing $q_{net}$ (<0.5 MPa) and relatively high $B_q$ (0.2–0.6) values indicating silt/mud. At greater depth the CPT profiles are more variable, with moderate sharp increases in $q_{net}$ combined with reductions in $B_q$ indicating higher proportions of sandy material.

### 3.1.1. Relationship between CPT response and undrained shear strength

For fine grained soils, geotechnical design of offshore infrastructure is typically undertaken using the undrained soil response. In this case, the sediment strength is often taken as the undrained shear strength ($s_u$), which may be estimated directly from CPT data as:

$$s_u = q_{net}/N_{kt} \quad \text{(Eq. 3-3)}$$

where $N_{kt}$ is the cone factor. Selecting an appropriate value of $N_{kt}$ can be difficult. It is commonly evaluated via calibration against laboratory strength tests undertaken on undisturbed soil samples, but can also be (broadly) estimated using other site-specific geotechnical data (e.g. soil classification properties) combined with previous experience in similar soil types. For regional (i.e. North West Shelf, Australia) soft silts and muds that generally exhibit an undrained response, $N_{kt}$ values typically lie in the range of 9–15; while for sandier soils that exhibit a drained to partially drained response, $N_{kt}$ typically exceeds 30. For simplicity, or where the investigated sediments are considered to be relatively homogenous, a uniform 'best fit' or average value of $N_{kt}$ is often adopted for design purposes. However, for a non-uniform stratigraphy like that shown on Fig. 10, a constant $N_{kt}$ is unable to adequately capture the changing response (and the corresponding variation in $s_u$). Fig. 11 illustrates this, showing results from 30 simple shear tests performed on samples obtained in the vicinity of the offshore site CPT locations (at various depths). The pairing of measured laboratory and CPT data is presented on Fig. 11a (as green dots) in terms of an inferred $N_{kt}$ (determined using Equations (3)–(3)) versus $q_{net}$, and on Fig. 11b in terms of the same value of $N_{kt}$ versus $B_q$. Fig. 11b clearly shows the expected trend of decreasing $N_{kt}$ with increasing $B_q$, while direct comparison of the field-measured $q_{net}$ and laboratory-measured $s_u$ data implies a 'uniform' best fit $N_{kt}$ equal to 12.35, indicated on Fig. 11 as a blue line.

Mayne and Peuchen (2018, 2022) undertook a review of more than 400 laboratory strength tests paired with CPT data from more than 60 different intact clayey soils (ranging from soft to hard). The authors propose a generalised trend linking $N_{kt}$ with $B_q$ described as:

$$N_{kt} = 10.5 - 4.6\ ln(B_q + 0.1) \text{ with } B_q > -0.1 \quad \text{(Eq. 3–4)}$$

The range covered by the Mayne-Peuchen relationship was approximately $B_q = -0.05$ to 1.25. While the soils included in this study differ from the high carbonate content soils included in the current case study, this relationship is included on Fig. 11b and provides an improved representation of the measured data – while still not capturing the full range of $N_{kt}$ values.

In the present work a new 'variable' $N_{kt}$ relationship was investigated which incorporates the in situ vertical effective stress ($\sigma'_{vo}$) along with the measured $q_{net}$ and $B_q$ values, described as:

$$N_{kt} = a(B_q)^b \left(\frac{q_{net}}{p_{atm}}\right)^c \left(\frac{\sigma'_{vo}}{p_{atm}}\right)^d + e \leq 40 \text{ with } B_q \geq 0.01 \quad \text{(Eq. 3–5)}$$

where $p_{atm}$ is atmospheric pressure = 100 kPa, and $a$, $b$, $c$, $d$ and $e$ are

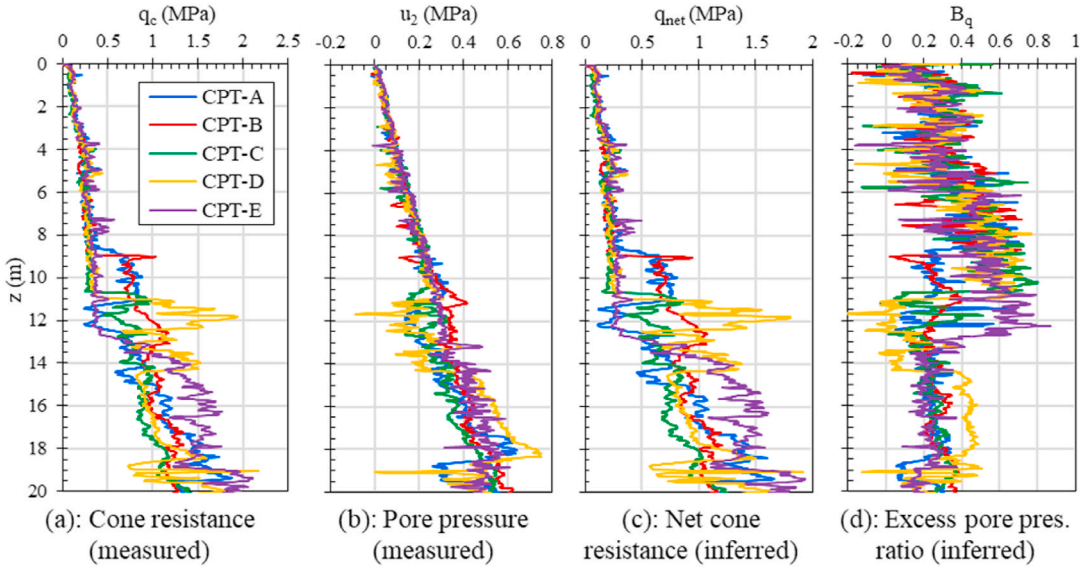



**Fig. 10.** Measured CPT profiles.

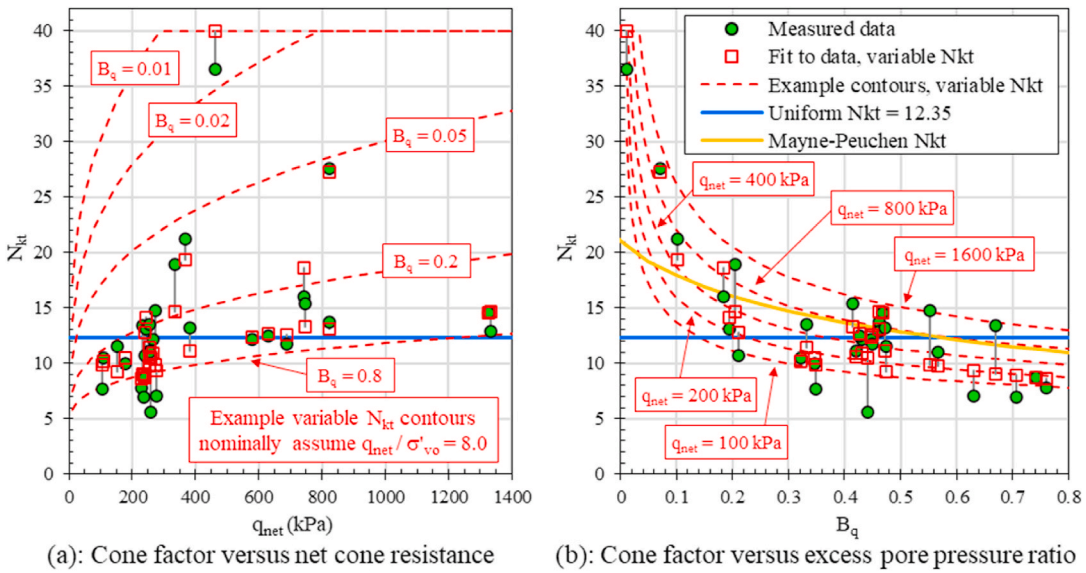



**Fig. 11.** Examination of relationship between offshore site CPT response and undrained shear strength.

fitting parameters determined via calibration against the measured CPT ($q_{net}$, $B_q$) and laboratory ($s_u$) data to provide the lowest mean squared error. The resulting estimates of $N_{kt}$ using this relationship are included on Fig. 11 as red squares (using fitting parameters $a = 2.793$, $b = -0.425$, $c = 0.435$, $d = -0.142$ and $e = 3.698$), and it can be seen that the relationship provides a marked improvement in predicting $N_{kt}$. For information, Fig. 11 also shows example indicative contours illustrating the variable $N_{kt}$ relationship for various values of $q_{net}$ and $B_q$ and nominally assuming an average representative value of $q_{net}/\sigma'_{vo} = 8.0$ calculated from the data.

The appropriateness of adopting uniform and variable $N_{kt}$ relationships to estimate undrained shear strength from CPT data at the current site is summarised on Fig. 12, which shows frequency-based histograms of the ratio of measured $s_u$ (obtained from each laboratory strength test) versus estimated $s_u$ (obtained using the uniform and variable $N_{kt}$ relationships discussed above). Also included are corresponding natural log-normal distribution fits to the histograms. The variability encapsulated by the log-normal fits may be expressed in terms of the coefficient of variation (COV, calculated as $[\exp(\sigma^2) - 1]^{0.5}$ where $\sigma^2$ is the distribution variance) – noting in this case that lower COV values indicate better overall ability of the relationship in predicting $s_u$. The variable $N_{kt}$ relationship (Equations (3)–(5)) provides the best overall performance, and confirms that a significant portion of the variability in measured-versus-estimated $s_u$ associated with the uniform $N_{kt}$ may be accounted for through consideration of $B_q$ and $\sigma'_{vo}$.

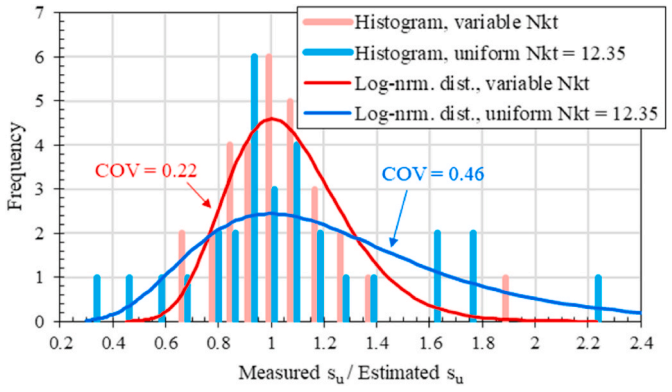



**Fig. 12.** Histogram of measured $s_u$ divided by cone-estimated $s_u$ values at offshore site using variable and uniform $N_{kt}$ relationships.

Application of the variable $N_{kt}$ relationship (Equations (3)–(5)) and a uniform $N_{kt} = 12.35$ results in the undrained shear strength profiles shown on Fig. 13a and b, respectively. The profiles calculated assuming the uniform $N_{kt}$ show a significantly larger $s_u$ range between CPTs, particularly at depths beyond approximately 9 m. Given that use of a variable $N_{kt}$ model is expected to better reflect the 'true' seabed undrained shear strength properties, the comparison between Fig. 13a and b suggest that much of the variability of the cone resistance profile is due to different drainage conditions occurring during cone penetration – rather than variability in undrained shear strength of the in situ material. This is important, since for a foundation design where drainage path lengths, coefficient of consolidation ($c_v$) values and loading times reveal that the soil response in undrained, it is the undrained shear strength that is relevant for design.

### 3.1.2. *Conventional use of spatial CPT data for geotechnical design*

In traditional offshore geotechnical engineering practice, assessment of CPT data generally involves the determination of simplified design profiles that are representative of the wider data set. Importantly, these simple profiles are often assumed to apply across the full survey area and are independent of horizontal location. An example is presented on Fig. 14a, which shows $s_u$ profiles determined from CPT data using the variable $N_{kt}$ relationship presented above, together with a simple piecewise-linear 'design-mean' profile (in black) generated based on natural log values of $s_u$. Natural log values were adopted since log-normal distributions are generally considered to be a good statistical model for geological (and geotechnical) properties (Lacasse, 1994; Griffiths et al., 2009) and their use ensures non-negative probabilistic estimates (which is appropriate for soil strength).

Also shown are proposed design 'low estimate' (LE) and 'high estimate' (HE) $s_u$ profiles for the site Which capture the likely range in $s_u$ at any given depth (and at any given position horizontally). In this instance the design LE and HE profiles nominally represent the 10th and 90th quantiles of the data respectively and were generated using quantile regression (Uzielli et al., 2019). It is noted that use of the uniform $N_{kt}$ approach gives a higher LE design line, which is unconservative when used for foundation sizing to achieve in-place capacity, emphasising the importance of correct transformation from $q_{net}$ to $s_u$. In addition, the data show a far wider spread of design lines when using a constant value of $N_{kt}$.

The conclusion can be drawn that although adoption of location-independent bounding profiles does simplify the design process, their

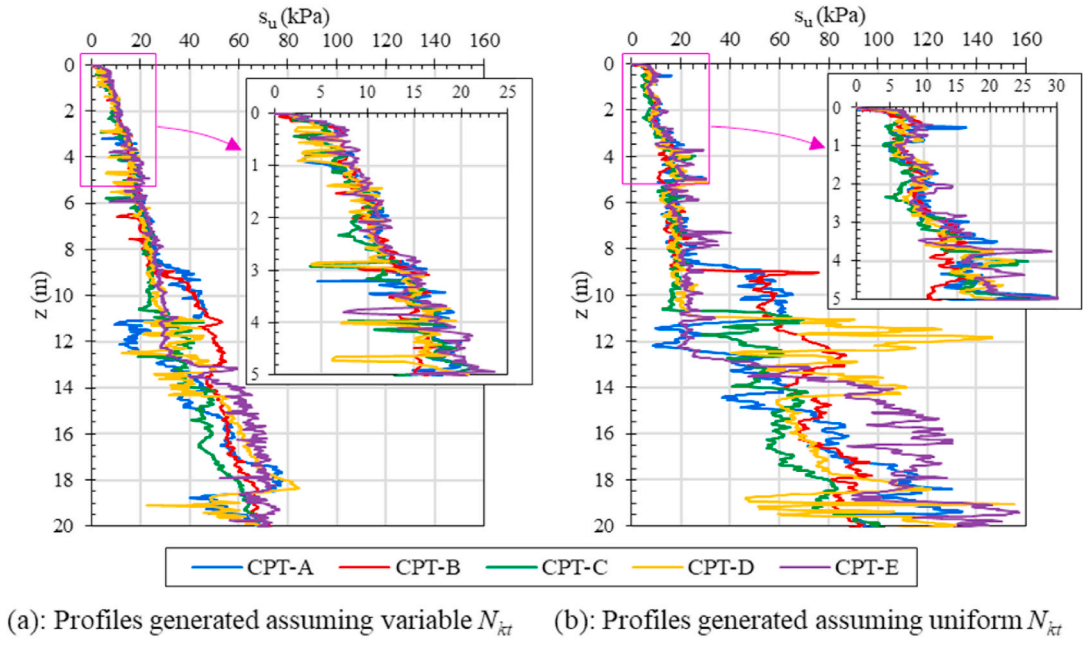



**Fig. 13.** CPT-inferred undrained shear strength profiles at the offshore site.

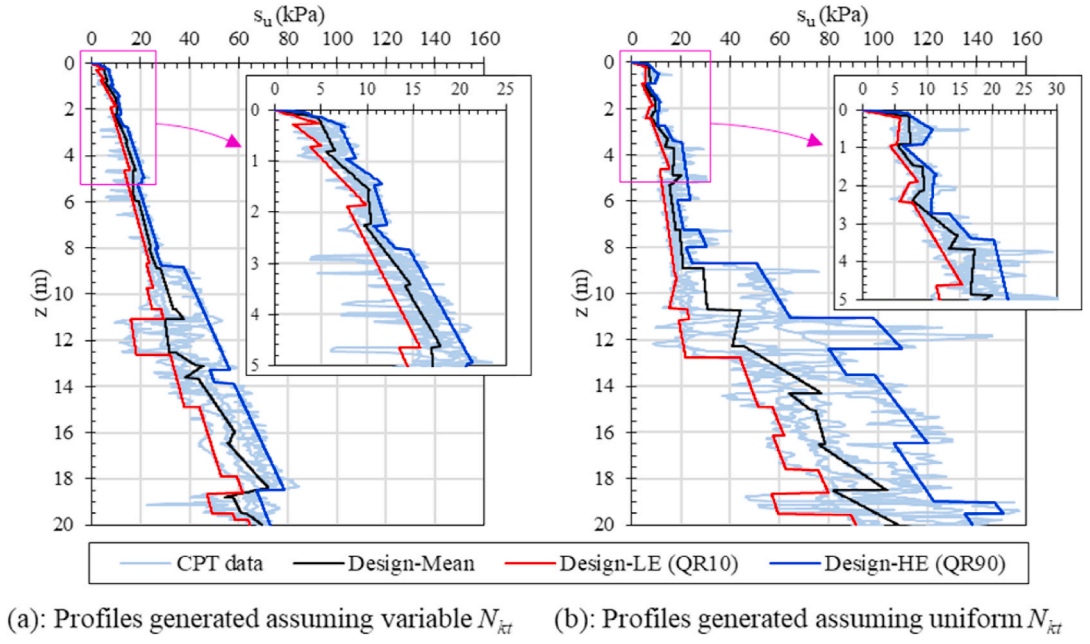



**Fig. 14.** CPT-based design undrained shear strength profiles at the offshore site.

use is likely to result in oversized (overly conservative or risk-averse) foundations that in turn leads to increased (and unnecessary) material, fabrication, transportation and installation costs. Hence, there is strong impetus to investigate and develop new, robust approaches for analysing spatially-varying data to determine location-specific soil strength (with uncertainty estimates) for use in geotechnical design.

### 3.1.3. *Updated approaches for interpolation of CPT data*

Ongoing improvement in computational capability (stemming from improvements and increased availability of both hardware and software) has led to increased interest in the application of data-centric methods for analysis of measured geotechnical data, with recent publications describing approaches for the formal treatment of uncertainty and the associated spatial interpolation (and extrapolation) of data via statistical and analytical approaches. These approaches aim to predict the most likely (mean) spatial distributions of geotechnical properties from point-to-point (and their uncertainty) based on a given volume of input geotechnical data. Examples of these approaches include:

- Kriging (Zhang et al., 2011; Cai et al., 2019; Hu et al., 2020; Rahman et al., 2021);
- Bayesian compressive sampling (Wang and Zhao, 2017; Wang et al., 2020, 2021);
- Bayesian compressive sampling combined with Markov Chain-Monte Carlo simulation (Zhao et al., 2020);
- Conditioned random field methods (Cai et al., 2019);

- Ensemble radial basis function networks (Shi and Wang, 2021a);
- Multiple point statistics (Shi and Wang, 2021a, 2021b);
- Approaches utilising iterative convolution extreme gradient boosting (Shi and Wang, 2021c);
- Direction-dependant coupled Markov Chain models (Li et al., 2019);
- Bayesian compressive sensing-based response surface method (Li and Wang, 2021);
- Inverse distance weighted interpolation (Rahman et al., 2021);
- Natural neighbour models (Rahman et al., 2021); and
- Artificial neural networks (Sauvin et al., 2019; Chen et al., 2021).

In the following sections, we explore the use of Bayesian compressive sampling (BCS) with Markov Chain-Monte Carlo simulation as applied to the current data set, in order to estimate the most likely cone tip resistance ($q_c$), pore pressure ($u_2$) and therefore undrained shear strength ($s_u$) values at unsampled locations, as well as the uncertainty in these estimated values.

### 3.1.4. BCS assessment of CPT data

The BCS method is suited to interpolating and extrapolating sparse geotechnical data, as it directly incorporates uncertainty in the estimation of geotechnical properties at unsampled (untested) locations and can accommodate non-stationary data – here, non-stationary signifies non-uniform soil stratigraphies, comprising multiple soil 'units' with spatially varying statistical properties. Additional details on the application of the BCS method to field CPT data are provided in O'Neill et al. (2022a, b).

The analyses undertaken extends horizontally from $x$ = 0–142 m, vertically from $z$ = 0–20 m, and is based on horizontal and vertical spatial increments of $\Delta x = 1.0$ m and $\Delta z = 0.02$ m. The BCS analysis of $q_c$ adopts natural log values of the input data, while the $u_2$ analysis adopted standard input values, principally because $u_2$ can take on negative values.

Each BCS analysis undertaken in the study comprised a total of 50,500 iterations. Initialisation of the analysis required that results from the first 500 iterations be disregarded. Subsequently, and to ensure statistical independence, results from every 100th iteration (from the remaining 50,000 iterations) were saved, while those from the other iterations were also disregarded.

In addition, it was assumed the measured (sampled) CPT data included a component of measurement uncertainty. To account for this uncertainty and the resulting variation in estimated $q_c$ and $u_2$ vales at unsampled locations, at the commencement of each 100 iteration block a set of random normally distributed measurement uncertainty values was added to the input $q_c$ or $u_2$ data in accordance with the recommendations outlined in Peuchen and Terwindt (2015). Hence, each set of saved BCS analysis results reflected a different (independent) estimate of CPT measurement uncertainty. The resulting BCS analysis output 500 independent sets of estimated cone resistance and pore pressure across the full 142 m by 20 m $x$-$z$ cross-section. The BCS-estimated $q_c$ and $u_2$ sets were then used with Equations (3)–(1) and Equations (3)–(2) to calculate 500 corresponding sets of estimated $q_{net}$ and $B_q$, and the results are discussed below.

Fig. 15a presents a $z$-$x$ heat map showing the mean $q_{net}$ across the survey line calculated from the 500 independently estimated $q_{net}$ sets. The locations of the five (sampled) CPTs are indicated on the map by vertical dashed lines. Fig. 15b presents profiles of the mean, 10th and 90th percentile (P10–P90) $q_{net}$ (also calculated from the 500 estimated

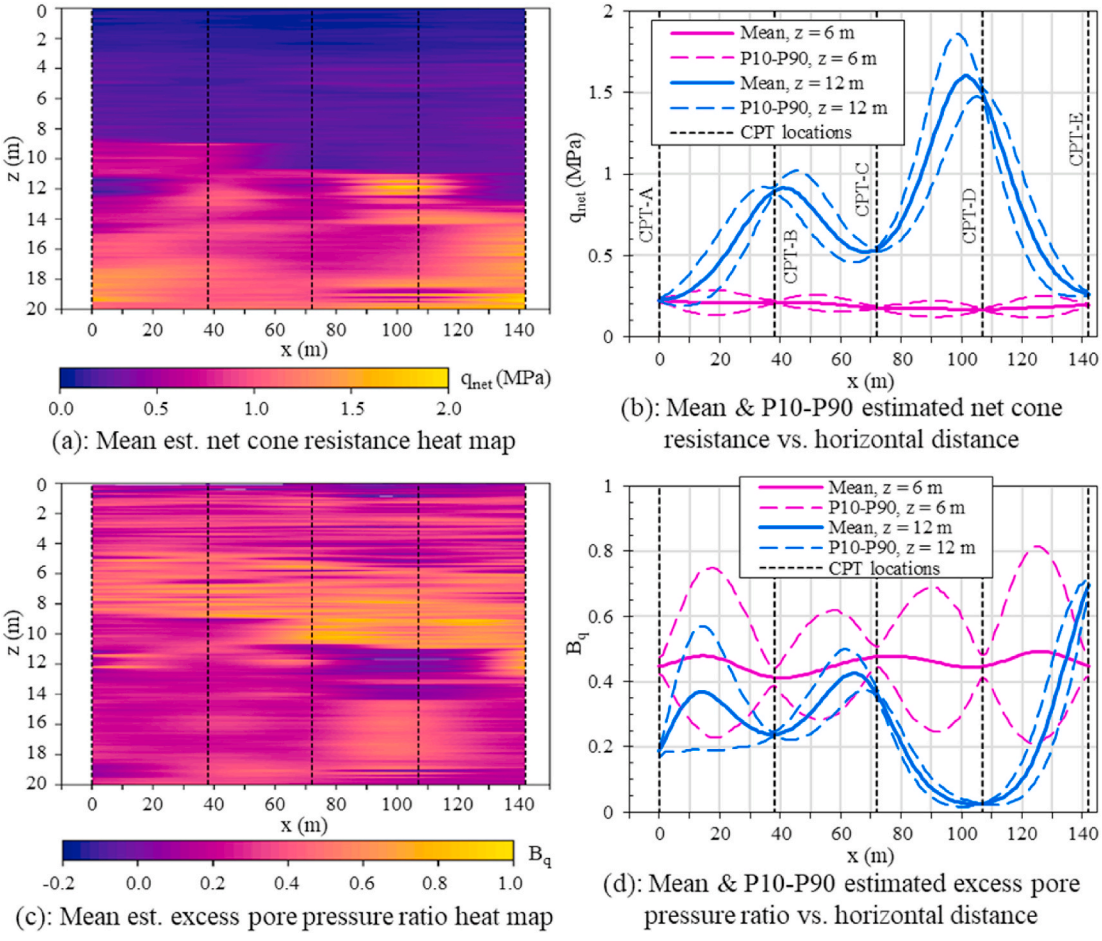



**Fig. 15.** BCS results – estimated net cone resistance and excess pore pressure ratio.

$q_{net}$ sets) versus horizontal distance along the survey line at nominal depths below seafloor of 6 m and 12 m. The profiles featured on Fig. 15b are useful in demonstrating the functionality of the BCS approach – at each CPT location $q_{net}$ is 'known', such that the range in estimated $q_{net}$ implied by the P10–P90 profiles is low (the small variability at the location of each CPT originates from the inclusion of CPT measurement uncertainty in the analyses); however as the horizontal distance from the nearest CPT location increases the P10–P90 range for $q_{net}$ widens, reflecting increasing spatial variability in the BCS estimates.

A similar heat map showing the mean of the BCS-estimated excess pore pressure ratio across the survey line is presented on Fig. 15c, while profiles of the mean and P10–P90 of the estimated excess pore pressure ratio versus x at depths of 6 m and 12 m are presented on Fig. 15d. It can again be seen that the range of estimates is low at the CPTs and high midway between the CPTs.

For each pair of independently estimated $q_{net}$ and $B_q$ profiles obtained from the BCS analyses, a corresponding estimate of the undrained shear strength across the survey line was calculated using the variable $N_{kt}$ fit described by Equations (3)–(5). Uncertainty associated with the variable $N_{kt}$ relationship was included by multiplying each set of estimated $s_u$ values by an 'uncertainty factor' equal to a random value selected from the log-normal distribution fit to the frequency curve of

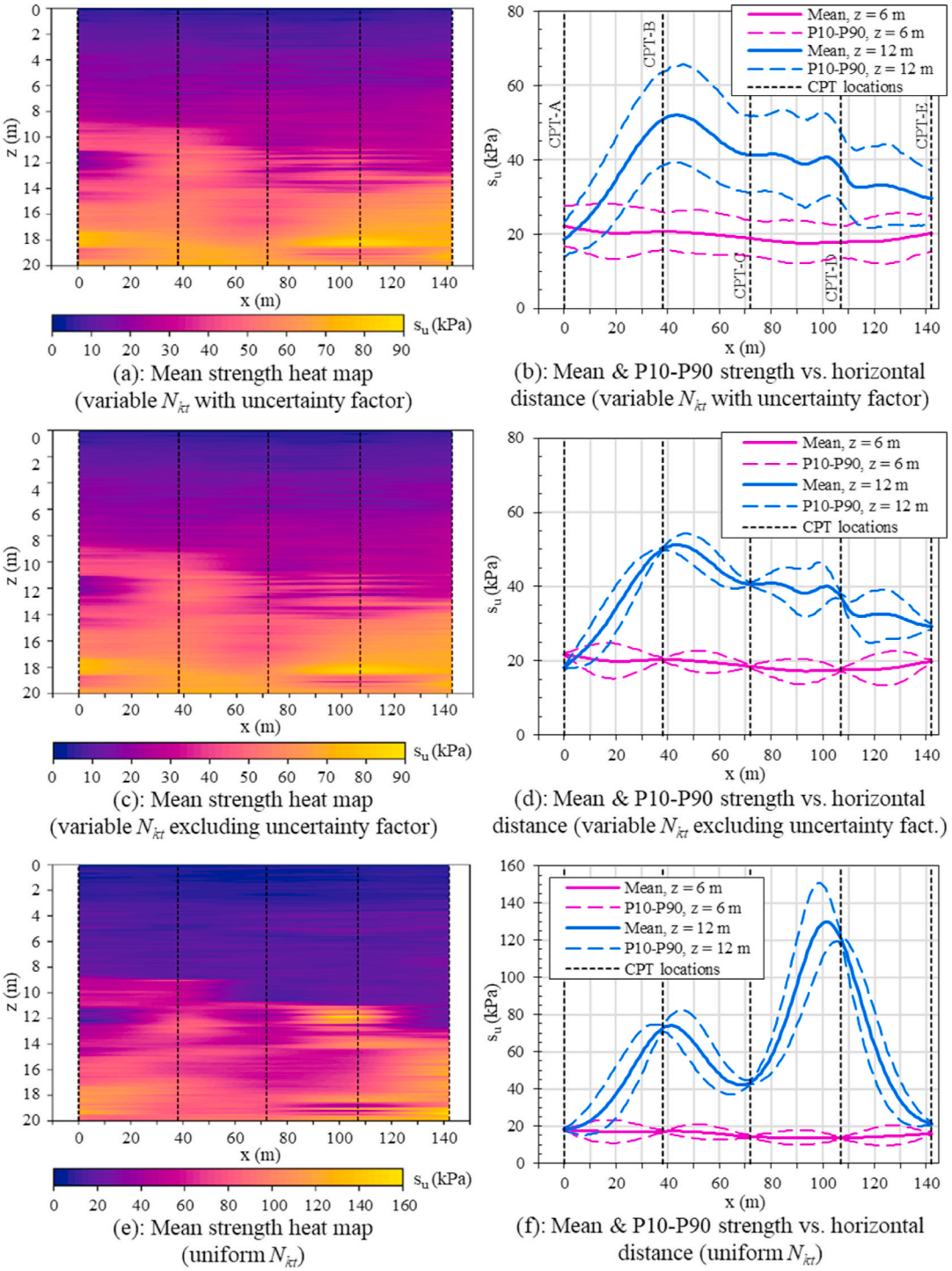



**Fig. 16.** BCS analysis results – estimated undrained shear strength.

[Measured $s_u$/Estimated $s_u$] as presented on Fig. 12 for the variable $N_{kt}$ relationship.

Fig. 16a shows a heat map of the resulting mean BCS-estimated undrained shear strength across the survey line, while Fig. 16b shows corresponding profiles of mean and P10–P90 estimated $s_u$ versus $x$ at depths of 6 m and 12 m. In this case the range of uncertainty in $s_u$ inferred by the P10–P90 profiles at the (known) CPT locations is only slightly less than the range mid-way between the CPTs – which is due to the estimation uncertainty associated with $N_{kt}$, and which in this study is independent of location and is statistically greater than the spatially-driven estimation uncertainty associated with $q_{net}$ and $B_q$.

Note that these findings may be a feature of carbonate soil conditions where small changes in grading can produce significant changes in drainage condition (and therefore tip resistance) during cone penetrometer tests. However, such transformational uncertainty will still feature for 'conventional' soils and should be accounted for rigorously in design. It is also worth noting that if predictions of $s_u$ were made at locations further away from CPT locations, it would be expected the spatial uncertainty in $q_{net}$ and $B_q$ would become more dominant than at locations nearer to CPT locations. Whether the spatial uncertainty in $q_{net}$ and $B_q$ would dominate over the uncertainty in $N_{kt}$ would depend on various factors.

In order to ascertain the relative impact of uncertainty associated with the variable $N_{kt}$ relationship, the estimation of $s_u$ using the BCS-derived $q_{net}$ and $B_q$ sets was repeated using the variable $N_{kt}$ fit described by Equations (3)–(5), but excluding the $N_{kt}$ uncertainty factor. The results are presented on Fig. 16c and d, and while the mean BCS-estimated $s_u$ appears unaffected by the inclusion of the uncertainty factor, this is not the case for P10-90 values – and it can be concluded that uncertainty stemming from $N_{kt}$ is considerably higher than the spatial uncertainty inherent in the CPT data.

Similar plots are presented using a uniform $N_{kt} = 12.35$ on Fig. 16e and f. Here the difference in estimated $s_u$ at 12 m depth is particularly noticeable – due primarily to the presence of lower $B_q$ soil detected across most of the survey line at 12 m depth, which leads to higher $N_{kt}$ values via the variable $N_{kt}$ relationship.

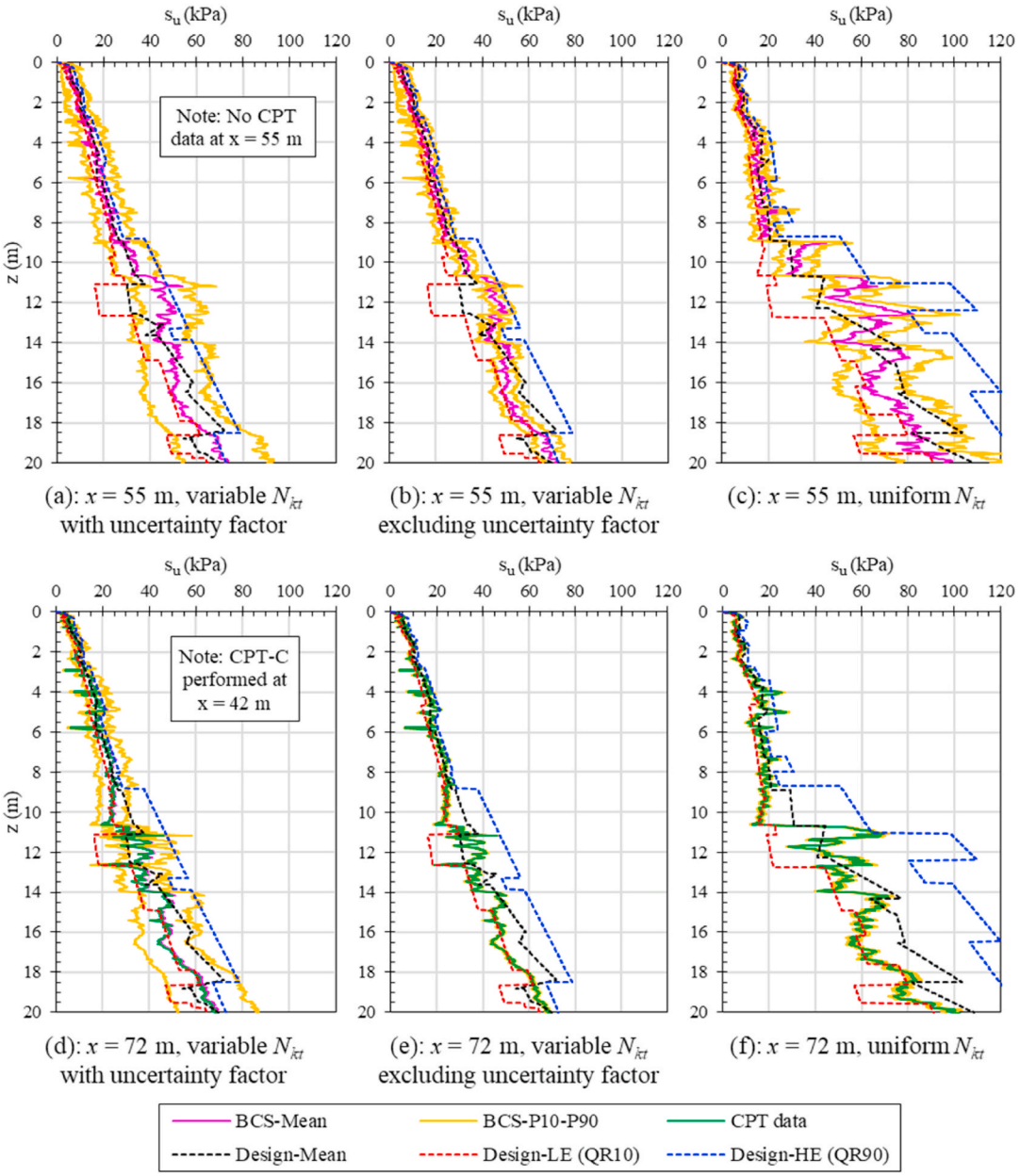



**Fig. 17.** Comparison between offshore site CPT data, BCS analysis results and design undrained shear strength profiles.

Additional examples of using BCS analysis are presented on Fig. 17 showing profiles of BCS-estimated mean and P10–P90 $s_u$ versus depth midway between CPT-B and CPT-C ($x = 55$ m) and at the location of CPT-C ($x = 72$ m), using all of the measured data in the analyses. Separate plots are shown considering the variable $N_{kt}$ relationship with and without the uncertainty factor, and using a uniform $N_{kt}$. For comparison, the $s_u$ profiles determined directly from CPT-C are included on Fig. 17d–f – noting that these profiles are consistent with the trends inferred by the $s_u$ versus $x$ results presented on Fig. 16. At the CPT-C location ($x = 72$ m) the BCS-estimated mean $s_u$ profiles are nearly identical to their measured counterparts. Furthermore, and with the $N_{kt}$ uncertainty factor excluded, the small (but discernible) range in the P10–P90 values at CPT-C (Fig. 17e and f) is due to the inclusion of CPT measurement uncertainty.

Finally, Fig. 17 also shows the respective design mean, LE (QR10) and HE (QR90) $s_u$ profiles previously shown on Fig. 14 for this site – and it is seen that in general, the location specific BCS-estimated P10–P90 profiles reflect a narrower estimated $s_u$ range than the site QR10-QR90 profiles.

#### 3.1.5. Concluding remarks on case study 2

Aside from the advantage in formally quantifying and incorporating spatial uncertainty in the estimation of parameter values relative to sampled locations, use of the BCS method to interpolate CPT data also allows for and facilitates the statistical assessment of foundation capacity (and other geotechnical design attributes).

## 4. Discussion and implications for design

Section 2 and Section 3 discussed the forms of geotechnical and geophysical inputs (identified as 'I1' through 'I4' on Fig. 1) and explored (a) how geophysical information may be used to quantify seabed properties through seismic inversion; and (b) how geotechnical data alone can be used to generate estimates of geotechnical properties ($s_u$ in this case) and their uncertainty at any point. This has been done for two carbonate sediment sites on the Australian continental shelf which may have properties broadly similar to areas for future wind farm development in Australia, and globally where carbonate sediments may be encountered. As indicated on Fig. 1, a further step in the development of this methodology will be to fully integrate the geophysical and geotechnical data such that the statistical analysis encompasses all of the data (and specific uncertainties) of each input type – in order to obtain information for the design of offshore wind farms in carbonate seabeds.

While conducting the work for the two cases studies, the following general findings were revealed:

- When planning a seismic inversion workflow, a thorough assessment of the seismic and geotechnical survey data should be made, the details of which largely determine what is possible in terms of output from the analysis. If the seismic data are of poor quality, for example containing a low signal-to-noise ratio, or are not properly migrated such that reflections are incorrectly positioned in the subsurface, these deficiencies will also be present in the inversion output. Time and effort spent to ensure input data is of high quality will result in better output from the seismic inversion process. As the output of a seismic inversion workflow is a numerical model of the subsurface representing some physical property, the model should be re-evaluated and updated if necessary as more information becomes available.
- The inversion method used in Case Study 1 is deterministic, meaning it produces the best solution for a given data set. While some attempt was made to quantify uncertainty via the acoustic impedance to geotechnical property regression analysis, no attempt was made to quantify uncertainty associated with the geophysical data and the seismic inversion process itself. This is ongoing work and will be required when quantitatively integrating geophysical and geotechnical data in future statistical approaches.
- For the carbonate soils studied, it was found that uncertainty associated with the transformation of cone penetrometer data to undrained shear strength was more important than (and in fact dominated) the uncertainty associated with the point-to-point profiles of cone resistance. This may be unique to carbonate soil conditions (because of drainage transitions during cone penetrometer testing) but should be included appropriately in future work, alongside considerations of spatial variability.

## 5. Conclusions

The accelerating push to transition from fossil-based to renewable energy has led to growing interest in the development of approaches that maximise the value of geotechnical and geophysical data to support design. Unlike offshore oil and gas projects which typically involve a limited number of structures, renewable energy projects (e.g. offshore wind farms) may comprise many tens or even hundreds of structures, spread across an extensive area. Coupled with a need to reduce cost in order for wind energy to be competitive with other energy sources, while not unnecessarily increasing project risk, there is a focus on identifying new ways to extract the full benefit from all available data.

This paper demonstrated two such approaches that utilise geo-data from offshore Western Australia, potentially representative of conditions of future offshore wind farms. Firstly, an integrated approach combining geophysical and geotechnical data through a seismic inversion process was presented, followed by a statistical analysis of geotechnical CPT and soil strength data combined with BCS-based spatial interpolation. Useful findings regarding each methodology and their respective input requirements were revealed, potentially feeding future site investigation practice and the development of integrated interpretation approaches.

It is envisaged these approaches form a starting point for further research that will develop and examine other (new) approaches and ultimately lead to a framework for efficiently integrating the analysis of geophysical and geotechnical data, and implementing the analysis output directly into foundation design.

## Credit author statement

We (M.P. O'Neill, A.L. Osuchowski, Y. Cai, M.F. Bransby, P.G. Watson, C. Gaudin, J. Doherty, E. Dalgaard and R. Ross), the co-authors of the manuscript for submission to Ocean Engineering titled "*Integrated and Data Science-Informed Seabed Characterisation for Optimised Foundation Design*", declare that all co-authors contributed equally to the writing and preparation of the manuscript. Regards, Michael O'Neill.

## 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

The data that has been used is confidential.

## Acknowledgements

This research is part supported by the ARC ITRH for Transforming energy Infrastructure through Digital Engineering (TIDE, http://TIDE.edu.au) which is led by The University of Western Australia (UWA), delivered with The University of Wollongong and a number of Australian and international research partners, and funded by the Australian Research Council, INPEX Operations Australia, Shell Australia,

Woodside Energy, Fugro Australia Marine, Wood Group Kenny Australia, RPS Group, Bureau Veritas and Lloyd's Register Global Technology (grant No. IH200100009). The fourth author holds the Fugro Chair in Geotechnics at UWA, whose support is gratefully acknowledged. The fifth author leads the Shell Chair in Offshore Engineering research team at UWA, which is supported by Shell Australia.

The seismic inversion workflow presented in Section 2 was performed using a combination of open source Python libraries and internally-developed software. Open source Python libraries used included segyio for loading of seismic data (SEG-Y) files, pandas for loading of tabulated data (CSV) files, NumPy and SciPy for digital signal processing and regression analysis, PyLops (Ravasi and Vasconcelos, 2020) for seismic inversion computations, and Matplotlib (Hunter, 2007) for creation of figures. The creators and maintainers of these libraries and the Python scripting language are gratefully acknowledged.

## References

Barclay, F., Bruun, A., Rasmussen, K.B., Camara Alfaro, J., Cooke, A., Cooke, D., Salter, D., Godfrey, R., Lowden, D., McHugo, S., Özdemir, H., Pickering, S., Gonzalez Pineda, F., Herwanger, J., Volterrani, S., Murineddu, A., Rasmussen, A., Roberts, R., 2008. Seismic inversion: reading between the lines. Oilfield Rev. 42–63.

Birchall, R., 2012. The Need to Integrate Geophysical with Geotechnical Data to Aid Pile Design and Installation: a Case Study of the Sheringham Shoal Offshore Wind Farm. Proc. 7th Int. Conf. Offshore Site Investigation and Geotechnics, London, UK, pp. 667–673.

Bull, J.M., Quinn, R., Dix, J.K., 1998. Reflection coefficient calculation from marine high resolution seismic reflection (chirp) data and application to an archaeological case study. Mar. Geophys. Res. 20, 1–11.

Byrne, B.W., McAdam, R.A., Burd, H.J., Houlsby, G.T., Martin, C.M., Beuckelaers, W.J.A. P., Zdravkovic, L., Taborda, D.M.G., Potts, D.M., Jardine, R.J., Ushev, E., Liu, T., Abadias, D., Gavin, K., Igoe, D., Doherty, P., Skov Gretlund, J., Pacheco Andrade, M., Muir Wood, A., Schroeder, F.C., Turner, S., Plummer, M.A.L., 2017. PISA: New Design Methods for Offshore Wind Turbine Monopiles. *Proc. 8th Int. Conf. Offshore Site Investigation and Geotechnics*, London, UK, pp. 142–161.

Cai, Y., Li, J., Li, X., Li, D., Zhang, L., 2019. Estimating soil resistance at unsampled locations based on limited CPT data. Bull. Eng. Geol. Environ. 78, 3637–3648. https://doi.org/10.1007/s10064-018-1318-2.

Chen, J., Sang, L., Willems, B., Newlin, J., Ruden, O., Sarker, R., Hu, S., 2020. A machine learning based quantitative integration of geotechnical and geophysical data for deepwater site characterization. Proc. 4th Int. Symp. Frontiers in Offshore Geotechnics 653–663. Austin, USA.

Chen, J., Vissanga, M., Shen, Y., Hu, S., Beal, E., Newlin, J., 2021. Machine learning-based digital integration of geotechnical and ultrahigh-frequency geophysical data for offshore site characterizations. J. Geotech. Geoenviron. Eng. 147 (12), 1–14. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002702.

Dix, C.H., 1952. Seismic Prospecting for Oil. Harper, New York, NY, p. 414pp.

Forsberg, C.F., Lunne, T., Vanneste, M., James, L., Tjelta, T.I., Barwise, A., Duffy, C., 2017. Synthetic CPTs from Intelligent Ground Models Based on the Integration of Geology, Geotechnics and Geophysics as a Tool for Conceptual Foundation Design and Soil Investigation Planning. Proc. 8th Int. Conf. Offshore Site Investigation and Geotechnics, London, UK, pp. 1254–1259.

Griffiths, D.V., Huang, J., Fenton, G.A., 2009. Influence of spatial variability on slope reliability using 2-D random fields. J. Geotech. Geoenviron. Eng. 135 (10), 1367–1378.

Hamilton, E.L., Bachman, R.T., 1982. Sound velocity and related properties of marine sediments. J. Acoust. Soc. Am. 72 (6), 1891–1904.

Hu, Y., Wang, Y., Zhao, T., Phoon, K.-K., 2020. Bayesian supervised learning of site-specific geotechnical spatial variability from sparse measurements. ASCE-ASME J. Risk. Uncertainty. Eng. Sys. Part A.: Civ. Eng. 6 (2), 1–12. https://doi.org/10.1061/AJRUA6.0001059.

Hunter, J.D., 2007. Matplotlib: a 2D graphics environment. Comput. Sci. Eng. 9 (3), 90–95.

Johnson, T.C., Hamilton, E.L., Berger, W.H., 1977. Physical properties of calcareous ooze: control by dissolution at depth. Mar. Geol. 24, 259–277.

Karkov, K.H., Dalgaard, E., Diaz, A.T., Duarte, H., Hansen, H.J., Hviid, S., Høegh van Gilse, N.C., Krogh, L., Kuppens, S., Salaun, G., 2022. Case Study: AVO Inversion and Processing of Ultra-high Resolution Seismic for a Windfarm Application. *Proc. 83rd EAGE Annual Conf. & Exhib.*, Madrid, Spain, pp. 1–4.

Lacasse, S., 1994. Reliability & Probabilistic Methods. Proc. 13th Int. Conf. Soil Mechanics and Foundation Engineering, New Delhi, India, pp. 225–227.

LeBlanc, L.R., Panda, S., Schock, S.G., 1992. Sonar attenuation modelling for classification of marine sediments. J. Acoust. Soc. Am. 91 (1), 116–126.

Li, J., Cai, Y., Li, X., Zhang, L., 2019. Simulating realistic geological stratigraphy using direction-dependent coupled Markov chain model. Comput. Geotech. 115, 1–9. https://doi.org/10.1016/j.compgeo.2019.103147.

Li, P., Wang, Y., 2021. Development of an efficient response surface method for highly nonlinear systems from sparse sampling data using Bayesian compressive sensing. ASCE-ASME J. Risk. Uncertainty. Eng. Sys. Part A.: Civ. Eng. 7 (4), 1–14. https://doi.org/10.1061/AJRUA6.0001155.

Lyu, C., Park, J., Carlos Santamarina, J., 2021. Depth-dependent seabed properties: geoacoustic assessment. J. Geotech. Geoenviron. Eng. 147 (1), 1–13.

Mayne, P.W., Peuchen, J., 2018. Evaluation of CPTU Nkt Cone Factor for Undrained Strength of Clays. Proc. 4th Int. Symp. Cone *Penetration Testing*, Delft, Netherlands, pp. 423–429.

Mayne, P.W., Peuchen, J., 2022. Undrained Shear Strength of Clays from Piezocone Tests: A Database Approach. *Proc. 5th Int. Symp. Cone Penetration Testing*, Bologna, Italy, pp. 546–551. https://doi.org/10.1201/9781003308829.

Nauroy, J.-F., Dubois, J.-C., Colliat, J.-L., Kervadec, J.-P., Meunier, J., 1998. The GEOSIS method for integrating VHR seismic and geotechnical data in offshore site investigations. Proc. Int. Conf. Offshore Site Investig. Foundation Behav. 175–198.

O'Neill, M.P., Bransby, M.F., Doherty, J., Watson, P., 2022a. Spatial Interpolation of Sparse PCPT Data to Optimise Infrastructure Design. *Proc. 5th Int. Symp. Cone Penetration Testing*, Bologna, Italy, pp. 1023–1028.

O'Neill, M.P., Bransby, M.F., Watson, P., 2022b. The Effectiveness of Spatial Interpolation of Sparse PCPT Data to Optimise Offshore Design. Proc. 8th Int. Symp. Geotechnical Safety and Risk (ISGSR 2022), Newcastle, Australia.

Panda, S., LeBlanc, L., Schock, S., 1994. Sediment classification based on impedance and attenuation estimation. J. Acoust. Soc. Am. 96 (5), 3022–3035.

Peuchen, J., Terwindt, J., 2015. Measurement Uncertainty of Offshore Cone Penetration Tests. Proc. 3rd Int. Symp. Frontiers in Offshore Geotechnics, Oslo, Norway, pp. 1209–1214. https://www.proquest.com/docview/2115977051.

Rahman, M.H., Abu-Farsakh, A., Jafari, N., 2021. Generation and evaluation of synthetic cone penetration test data using various spatial interpolation techniques. Can. Geotech. J. 58, 224–237. https://doi.org/10.1139/cgj-2019-0745.

Rattley, M., Salisbury, R., Carrington, T., Erbrich, C., Li, G., 2017. Marine Site Characterisation and its Role in Wind Turbine Geotechnical Engineering. Proc. of TC 209 Workshop, 19th Int. Conf. Soil Mechanics and Geotechnical Engineering, Seoul, Korea, pp. 21–34.

Ravasi, M., Vasconcelos, I., 2020. PyLops—a linear-operator Python library for scalable algebra and optimization. SoftwareX 11, 1–9. https://doi.org/10.1016/j.softx.2019.100361.

Richardson, M.D., Lavoie, D.L., Briggs, K.B., 1997. Geoacoustic and physical properties of carbonate sediments of the Lower Florida Keys. Geo Mar. Lett. 17, 316–324.

Richardson, M.D., Briggs, K.B., 2004. Empirical Predictions of Seafloor Properties Based on Remotely Measured Sediment Impedance. Proc. High Frequency Ocean Acoustics Conf., LaJolla, USA, p. 10.

Sauvin, G., Vanneste, M., Vardy, M.E., Klinkvort, R.T., Forsberg, C.F., 2019. Machine Learning and Quantitative Ground Models for Improving Offshore Wind Characterization. Proc. Offshore Technology Conf., Houston, USA, pp. 1–17. OCT-29351-MS.

Schock, S., 1989. The Chirp Sonar—A High Resolution, Quantitative Subbottom Profiler. Doctoral thesis, University of Rhode Island.

Schock, S., LeBlanc, L., Mayer, L., 1989. Chirp subottom profiler for quantitative sediment analysis. Geophysics 54 (4), 445–450.

Semple, R.M., 1988. The Mechanical Properties of Carbonate Soils. *Engineering For Calcareous Sediments*. Jewell & Khorshid, pp. 807–836.

Sharma, S.S., Joer, H.A., 2015. Some Characteristics of Carbonate Sediments from North West Shelf, Western Australia. Proc. 3rd Int. *Symp. Frontiers in Offshore Geotechnics*, London, UK, pp. 1109–1114.

Shi, C., Wang, Y., 2021a. Non-parametric machine learning methods for interpolation of spatially varying non-stationary and non-Gaussian geotechnical properties. Geosci. Front. 12, 339–350. https://doi.org/10.1016/j.gsf.2020.01.011.

Shi, C., Wang, Y., 2021b. Nonparametric and data-driven interpolation of subsurface soil stratigraphy from limited data using multiple point statistics. Can. Geotech. J. 58 (2), 261–280. https://doi.org/10.1139/cgj-2019-0843.

Shi, C., Wang, Y., 2021c. Development of subsurface geological cross-section from limited site-specific boreholes and prior geological knowledge using iterative convolution XGBoost. J. Geotech. Geoenviron. Eng. 147 (9), 1–17. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002583.

Shumway, G., 1960. Sound speed and absorption studies of marine sediments by a resonance method, Part 1. Geophysics 25 (2), 451–467.

Stevenson, I.R., McCann, C., Runciman, P.B., 2002. An attenuation-based sediment classification technique using Chirp sub-bottom profiler data and laboratory acoustic analysis. Mar. Geophys. Res. 23, 277–298.

Taner, M.T., Koehler, F., Sheriff, R.E., 1979. Complex seismic trace analysis. Geophysics 44 (6), 1041–1063.

Telford, W.M., Geldart, L.P., Sheriff, R.E., 1990. Applied Geophysics, second ed. Cambridge University Press.

Thomas, S., 2017. A Phased and Integrated Data Interpretation Approach to Site Characterisation. Proc. 8th Int. Conf. Offshore Site Investigation and Geotechnics, London, UK, pp. 71–87.

Uzielli, M., Zei, M., Cassidy, M.J., 2019. Probabilistic Assignment of Design Undrained Shear Strength Using Quantile Regression. Proc. 7th Int. Symp. Geotechnical Safety and Risk, Taipei, Taiwan, pp. 188–193.

Vardy, M.E., 2015. Deriving shallow-water sediment properties using poststack acoustic impedance inversion. Near Surf. Geophys. 13 (2), 143–154.

Vardy, M.E., Vanneste, M., Henstock, T.J., Morgan, E., Pinson, L.J.W., 2015. Can high-resolution marine geophysical data be inverted for soil properties? Proc. Inst. Acoustics. 37 (1), 149–156.

Vardy, M.E., Vanneste, M., Henstock, T.J., Clare, M.A., Forsberg, C.F., Provenzano, G., 2017. State-of-the-art remote characterization of shallow marine sediments: the road to a fully integrated solution. Near Surf. Geophys. 15, 387–402.

Vardy, M.E., Clare, M.A., Vanneste, M., Forsberg, C.F., Dix, J.K., 2018. Seismic Inversion for Site Characterization: when, where and Why Should We Use it? Proc. Offshore Technology Conf., Houston, USA, pp. 1–9.

Wang, Y., Zhao, T., 2017. Statistical interpretation of soil property profiles from sparse data using Bayesian compressive sampling. Geotechnique 67 (6), 523–536. https://doi.org/10.1680/jgeot.16.P.143.

Wang, Y., Hu, Y., Zhao, T., 2020. Cone penetration test (CPT)-based subsurface soil classification and zonation in two-dimensional vertical cross section using Bayesian compressive sampling. Can. Geotech. J. 57, 947–958. https://doi.org/10.1139/cgj-2019-0131.

Wang, Y., Hu, Y., Phoon, K.K., 2021. Non-parametric modelling and simulation of spatiotemporally varying geo-data. Georisk 16 (1), 1–21. https://doi.org/10.1080/17499518.2021.1971258.

Watson, P., Bransby, F., Delimi, Z.L., Erbrich, C., Finnie, I., Krisdani, H., Meecham, C., O'Neill, M., Randolph, M., Rattley, M., Silva, M., Stevens, B., Thomas, S., Westgate, Z., 2019. Foundation Design in Offshore Carbonate Sediments – Building on Knowledge to Address Future Challenges. Proc. XVI Pan-American Conf. Soil Mechanics and Geotechnical Engineering, Cancun, Mexico, p. 35.

Zhang, J., Zhang, L.M., Tang, W.H., 2011. Kriging numerical models for geotechnical reliability analysis. Soils Found. 51 (6), 1169–1177.

Zhao, T., Xu, L., Wang, Y., 2020. Fast non-parametric simulation of 2D multilayer cone penetrometer test (CPT) data without pre-stratification using Markov chain Monte Carlo simulation. Eng. Geol. 27 https://doi.org/10.1016/j.enggeo.2020.105670.