PROBABILISTIC FRAMEWORK FOR THE SOIL MODELLING IN THE SOUND PROPAGATION DURING IMPACT PILE DRIVING

by

GIORGIO CAUMO

Master Thesis

MASTER OF SCIENCE IN CIVIL ENGINEERING

Master track: Structural Engineering

Delft University of Technology

Student n°: 5130905

Committee:	Dr. ir. Apostolos Tsouvalas	TU Delft
	Dr. ir. Oswaldo Morales-Nápoles	TU Delft
	PhD. Yaxi Peng	TU Delft

January 18, 2022

Giorgio Caumo: *Probabilistic framework for the soil modelling in the sound propagation during impact pile driving*
January 2022

ABSTRACT

In the last decades, renewable energies gained an increasing interest due to the environmental awareness of people, especially in the developed countries. Moreover sustainable resources represent a long-term investment full of possibilities of use.
In the Netherlands a technology which is getting more and more popular to produce green energy is represented by offshore wind turbines (OWT). However, during the installation of the structural elements for these systems, the risk of noise pollution and animal harm is an issue that has to be considered.
Previous works developed models to predict the noise emission and propagation during the construction phase, however the uncertainty related to the environmental properties has not been yet fully investigated. Since the model characteristics are uncertain, so will be the prediction of the noise.
This thesis aims to fill this gap, investigating in the underwater soil property uncertainties and the resulting variation in sound predictions. The main goal of this work is to settle a sounding methodology to model the soil characteristics and interpret the sound levels.

In the first part of this work, the soil uncertainties are treated.
A framework on how to use measurements from cone penetration tests (CPT) and obtain mechanical and dynamic soil features is presented.
By means of statistical approaches, the procedure to define the optimal depth for different homogeneous layers (given the software used for the noise prediction) is described. Another topic dealt with is the definition of proper characteristic distributions and the choice of the optimal one representing the available empirical measurements.
Finally a procedure to generate random samples for the analyses that will follow is shown.
An important feature presented is the use of the correlation between the properties to define copulas. The samples then are not completely random and independent, but instead combinations that are more likely to appear are obtained.

In the second part, the results of several analyses are presented.
The steps to treat the sound levels and obtain the probability density (and cumulative) distributions are discussed. These results will help in estimating the probability of exceeding a particular defined sound level. With this information, additional measures and precautions, as noise barriers, may be adopted in the installation of the pile to prevent exceeding the threshold.

The correlation between soil properties and obtained sound levels is investigated, in order to highlight if there are soil properties that greatly affect the outcomes.
The insight obtained may help in determining which particular features need a careful estimation, both by more accurate measurements or new techniques. Another benefit related to the detection of parameters not affecting significantly the sound levels, is the reduction of simulations necessary to cover enough combinations. That is because if a property can be neglected in the generation of samples and be taken as a fixed value, less combinations are needed to be considered.
Finally, a comparison between the obtained estimations and empirical measurements in the North Sea is made, to test the validity of the framework proposed.

The joy of life comes from our encounters with new experiences, and hence there is no greater joy than to have an endlessly changing horizon, for each day to have a new and different sun.

Happiness is only real when shared.

— Alex Supertramp - Christopher McCandless

ACKNOWLEDGMENTS

Before starting, I would like to take a small really personal space.
It is only at the end of a journey, once you finally stop and with a free mind, that you can really elaborate what has happened.
New places, new persons, new experiences, hard work, (more, hopefully) accomplishments, (less) failures. It's part of life.
In all this chaos, something always remains there as a lighthouse during a tempest, while other special things happen unexpectedly changing someone's life forever.
Here I would like to thank who did that for me.

Laura, Alberto, Aldo, I never really told you how much you affected, helped, supported, motivated me. Within ups and downs, between arguments and amazing moments, you were always there, ready to unconditionally give the love a family should always spread.
Although I've always wanted to be independent and never liked to ask for help, all the steps I took were possible because of you.
I am here because of you.
Aurelio, Ruggero, Roberta, Bruna, I owe to each of you a part of who I am now, the things that I value, and the memories of a past and beautiful time.

Apostolos, thank you. I never really had a particular passion in a topic before being your student, everything was mostly just a duty. All of your questions made me realize what I needed to improve, what I still didn't understand and forced me to have clearer in mind what to look for. I learnt a lot from you.
Oswaldo, thank you. It was after talking with you and with your suggestions that I really appreciated and understood the importance of statistics and probability. I also answered some of the questions that, to me, are really of value.
Yaxi, thank you. You had to deal with me more than the others! Every single problem, error, question, idea, you were there to help and give your knowledge. Most of this work was possible also because of you.

Daniela, there is not enough space here to really tell how precious you are to me, as a friend, as a study partner, as a dancing teacher! You probably saw my worst, and yet you have remained giving me some of the most valuable help I needed and an amazing friendship!
Menelaos, I really value the time spent together, the discussions, the jokes, the never happened holidays in Greece or in Italy. There will be time for that!
Mirco, although the different studies we had, although the distance and the few time spent you always gave me something to laugh, an advice, or something to feel a bit less lonely.
Akis, Hari, Karim, *The Boys*, I had such a great time with all of you, studying, laughing, discussing and getting advices, but mostly getting hurt seeing you put hot sauce on pizza!
Carolina, Andrés, the *Latinos*, thank you for all the days spent playing football, training, enjoying the cold beach and playing cards, is thank to you if I got through some hard times.
Cristina, I'll never thank you enough for all the help and support you gave me lately. You made me find the strength I needed to complete this journey!

Thanks to TU Delft, for the university environment, the challenges (the rain and wind too), the opportunities. All these were some things that I think not only improved me as a future engineer, but as a person too.

I have never really considered how the persons surrounding us can so much affect our paths.
As I said, it is only now that I realize how lucky I have been, meeting amazing people and giving value to every moment, even silly ones.
Each of you left a mark.
A chapter has now finished, another begins, and I wish to be lucky enough to remain surrounded by amazing people as you are.

Giorgio Caumo
January 2022

CONTENTS

Part I

INTRODUCTION AND BACKGROUND

1 INTRODUCTION

1.1 SUMMARY

The following section presents a brief overview of the content of each chapter.
In Chapter 1 an introduction of the topic, of the motivation and an overview on the sound level limitations in different countries is presented.
Chapter 2 addresses the available methods to model the sound generation and propagation in marine environments. Some statistical methods to deal with uncertainties are also discussed. A brief discussion is made upon the advantages and disadvantages of each one.
In Chapter 3 the adopted sound generation and propagation solver is described. The software, called *SILENCE*, is based on a semi-analytical model. This has better performance in the accuracy and time efficiency while solving the analyses compared to other existing tools.
Chapter 4 opens the discussion on the estimation of the soil properties from raw cone penetration test (CPT) results. Through empirical relations the main soil characteristics are obtained. The uncertainties related to these measurements are also addressed.
In Chapter 5 the main statistical approaches used in this work are presented. These consist of the definition of soil layer's depths, the determination of various distributions types and the choice of the best representing the empirical data.
In Chapter 6 the main model set-up and the sampling procedure is outlined. Some assumptions and remarks are then discussed.
Chapter 7 deals with the results of the sound propagation analyses. The statistical interpretation is outlined and the sound distributions are obtained. The correlation between soil properties and sound levels is investigated. At last, a comparison between sound measurements performed in an offshore site and the simulation predictions is analyzed.
Chapter 8 summarizes the whole framework presented, discussing the limitations or hypotheses made, highlighting the results obtained and the advantages of the framework.
In Chapter 9 the conclusions of this work are drawn.
Finally some recommendations for future works or approaches which can improve the framework presented are described.

1.2 PROBLEM STATEMENT AND MOTIVATION

In the last decade climate change and human footprint have become major issues and people, companies, and governments are now more than aware of the consequences these factors can lead to.
A solution to this problem could be the use of renewable energies to preserve the environment and improve the conditions of people living around polluted areas. The use of green alternatives looks appealing also from an economic point of view, since it may be a long term investment to self produce the required energy for growth and well-being.
There are, however, some limitations when using sustainable resources, mainly regarding their efficiency. A common solution to overcome this issue is the installation of extensive systems, which however often require space that could be used for other purposes, for example farming.

This thesis focuses on one particular system to produce renewable energy, that is, *offshore wind turbines* (OWT), which increased their popularity due to various advantages compared to the inland alternatives. These installations, for example, do not occupy land, and not having the spatial limitation it is possible to build larger machineries. Moreover they benefit from steadier and stronger winds, thus increasing the energy production. Studies estimate that the same turbine excited by ocean winds can produce up to 50% more energy than onshore ones [1].
During the construction phase another issue arises, regarding the noise emitted and propagating in water, air and soil during the installation. It is a major concern, since it can harm marine wildlife, affecting their behaviour and/or causing damages to their hearing organs [2, 3, 4].

Governments of different nations have then decided to establish rules which limit the construction time to specific seasons, or thresholds to the noise that can be generated during installation [5]. It is believed this may help the harm to animals.
Unfortunately, the *real* noise levels that will be generated are of difficult prediction.
Existing numerical models are used to predict the sound levels which would be produced during installation. However to do so an accurate description of the geometry and material properties is required. Especially the latter ones are highly variable in the large domains which need to be considered, often tens of kilometers.
Approximations are then made, but even if the sound prediction may be performed correctly, it could only apply to a specific, and most probably not realistic, environment case.
To the author's knowledge, the various approaches available have mainly focused on the accurate numerical solution. The sensitivity of the noise

levels due to the uncertainties of environmental parameters has not yet been fully investigated [6].

This thesis aims to fill this gap, by building a probabilistic framework addressing the uncertainties in the soil properties. By employing the noise prediction model for offshore pile driving, the sound distributions can be obtained which are then used to define the probability of exceeding a certain threshold. Another objective is to identify correlations between specific soil features and the resulting sound levels.

From available *in-situ* tests material properties of the soil can be defined. However, to deal with the uncertainty in the soil parameters, due to measurement error and spatial variation of the marine environment, ranges of data sets are used and the probability distributions are introduced. Particular attention is also given to the dependence between different properties, since it is of importance to know when some combination of them may be more likely to appear.

Through several simulations of noise propagation, reproducing different soil stiffness and wave speed scenario, resulting sound levels are expected. The following step is the definition of a sound level distribution, which is used to estimate, depending on the specific limitation or criteria that needs to be met, the probability of exceeding a particular value.

Among the purposes of this work, the identification of the dominant soil parameters will help practitioners in focusing their efforts on an accurate measurement of soil properties. Newer and more accurate techniques are expected to be developed in the future.

It is also believed that the probabilistic framework here presented will be of use for upcoming investigations. It allows improvements and is sufficiently flexible to fulfill the needs of the user.

Once the sound level probability is obtained, the choice of the installation location, as well as the use of noise mitigation methods [7] can be optimized.

Another aspect to take into account is the possibility of reducing the overall amount of simulations performed during the projecting phase to obtain the necessary information.

1.3 OVERVIEW ON PILE DRIVING AND SOUND PROPAGATION

Different structural typologies of piles for offshore purposes exist. One of the most common for OWT consists of a monopile, obtained by welding several steel tubes: the dimensions can reach $8 - 9\,m$ of diameter and $70 - 80\,m$ of length, the thickness is generally around $90 - 150\,mm$.
The most used techniques to drive the pile in the soil, sometimes even tens of meters into the ground, is by the use of hydraulic impact hammer or large vibratory systems. New studies are investigating on the effectiveness of the latter method to produce lower sound pressure levels.
In general, considering the impact hammer, each blow delivered to the top of the pile makes it gradually penetrate the soil; the complete installation requires hundreds or thousands of blows, and an enormous amount of energy is involved.
Each blow is then followed by a high noise pressure level, which propagates through different *channels*. The main propagation medium is the water column, where Mach cones travelling at the compressional speed in the water ($c_p = 1500\,m/s$) generate around the pile. The second, but equally important, is represented by the seabed interface, where Scholte surface waves occur. As will be discussed further, they represent an intense noise propagation path which can travel many kilometers from the source origin.

Figure 1.1: Example of offshore platform for the pile installation, image downloaded from the internet

1.4 OVERVIEW ON SOUND INTENSITY LIMITATIONS

In this section a brief overview on the limitations considered in this thesis is outlined.
The main referenced national guideline chosen is the German one [5] which, summarizing the most important aspects requires, at 750 m from the pile driving location:

- a Sound Exposure Level (SEL) lower than 160 bB re $1\mu Pa^2$
- a Sound Pressure Level (SPL/L_{peak}) lower than 190 bB re $1\mu Pa$

The choice of relying on the German legislation is made since, at the moment, is the only one which introduces objective and measurable thresholds, while other countries do not rely on quantitative limits [8, 9]. In the Netherlands, for example, the only limitation regards a seasonal restriction (1st of January - 1st of July), without sound thresholds. In the United Kingdom pile driving has to be monitored by Marine Mammal Observer (MMO) and Passive Acoustic Monitoring (PAM) during the whole process.
To the author's knowledge, it is not clear if these limitations would completely prevent animal harm, yet these thresholds and periodic prohibitions would for sure prevent excessive damage to marine wildlife.
Stressing out the difference that appears between the referenced sound pressure levels (SPL) in air and water (eq 1.1), and to put the regulations in perspective to some sound pressure levels found in literature [10], table (1.1) present some common activities sound levels.

Activity	SPL in air [dB]	SPL in water [dB]
Threshold of hearing	0	26
Whisper	20	46
Normal conversation	60	86
Painful to human hear	130	156
Jet engine	145	171
Whale	139	165

Table 1.1: Some sound pressure levels for common activities

$$SEL[dB] = 10\log_{10}\left(\frac{1}{T}\int_0^T \frac{P^2}{P_{ref}^2}dt\right) \quad L_{peak}[dB] = 20\log_{10}\left(\frac{|P|}{P_{ref}}\right) \tag{1.1}$$

Where the P_{ref} in air and water are chosen differently: respectively $P_{ref,air} = 20\,\mu Pa$ at $1\,m$, $P_{ref,water} = 1\,\mu Pa$ at $1\,m$ from the sound source.

In general the sound conversion from air to water is obtained adding $+26\,dB$.

Looking at table(1.1) now the German sound limitations have clearer meanings, having thresholds really close to normal sounds in marine environments and that, even at the peak, should not provoke excessive harm to the wildlife.

1.5 PROJECT DESCRIPTION

In this section, the overview of the thesis is presented together with the main research questions involved.
As introduced before, the main problem related to pile driving in terms of environmental impact is the high noise emission during the installation. Since it may affect marine wildlife and damage their hearing organs precaution measures had been taken by governments.
To verify if the threshold will be met, a correct sound propagation estimation is necessary. This requires, other than an accurate solver method, also a realistic definition of the environment where the installation will take place.
This thesis mainly deals with the environmental feature uncertainties, primarily related to the soil characteristics. Since the model input are uncertain, even the sound level output will be. The aim is to investigate on the effects of this variability.

1.5.1 *Research questions and objectives*

This thesis aims to answer one main question:
Is it possible to predict the probability of exceeding a dangerous sound level, in order to take precautions and prevent animal harm?
Along with the main question, this work is intended to respond to some *sub-questions* related to the several topics analyzed:

- With empirical measurements (CPT) how is it possible to account for the variability of the marine environment?
- Since the input for the solver must be set, how is it possible to generate samples that account for correlation between the different properties?
- Once obtained the raw sound level measurements, how can they be treated?
- Analyzing the soil characteristics and the sound level output, is there any significant correlation?

To answer these questions, the following objectives are to be achieved:

- Define a flexible framework to treat the empirical recordings, which then allows to describe soil stratification. This will be helpful in determining the input soil parameters for the existing models with horizontal stratified ground layers.
- Since the soil is represented in a simplified manner, to account for the uncertainties of the 'real' model, a large amount of simulations is needed. To generate the input combinations of each simulation, a statistical framework is defined, accounting also for the existing correlation between soil features.
- A statistical/probabilistic approach is used to interpret the sound level outcomes and define probability distributions. With these information it is possible to estimate the probability of required outcomes and decide whether precaution measures are needed to be taken.
- Finding if certain properties are highly affecting the sound propagation may then help highlighting the features that need an accurate evaluation. Moreover future projects may use these information to design accordingly to reduce the noise emissions.

1.5.2 *Methodology*

A brief overview of the steps followed in this project is now outlined.
From available cone penetration test (CPT) recordings, the dynamic and mechanical soil properties are obtained. The results are statistically interpreted and given the need to define homogeneous layers the optimal depth for each of them is estimated. For each layer is then found the optimal distribution matching with the empirical data.
By the introduction of copulas, a high number of soil features combination accounting for the existing internal correlation is generated. Using these as input for several sound propagation simulations, the sound exposure and peak levels are obtained.
Again, the results are statistically interpreted and the optimal sound distributions fitting the data are found. Relating the outcomes with the simulation input the correlation is estimated.
A final comparison with experimental measurements is conducted to verify the validity of the method.

2

STATE OF THE ART AND AVAILABLE TECHNIQUES

In this chapter, an overview on the sound propagation phenomenon in marine environment, as well as some of the most common methods used for its evaluation will be discussed.
Some of the most common probabilistic approaches used to deal with uncertainty will also be presented.
In addition, the known advantages and limitation of each of them will be brought to attention, so that it will be clear the reason behind the choice of the adopted model described in ch(3), and the procedures in ch(4,5).

2.1 UNDERWATER SOUND PROPAGATION OVERVIEW

Sound propagation is a well known phenomenon, which requires the solution of differential equations by defining potentials for the required field.
One of the most common expression is the three-dimensional *wave equation for displacement potential* [11]:

$$\nabla^2\psi - \frac{1}{c_p^2}\frac{\partial^2\psi}{\partial t^2} = f(\mathbf{r},t) \tag{2.1}$$

Where c_p is the compressional wave speed in the water, the displacement potential ψ is defined as

$$\mathbf{u} = \nabla\psi$$

While $f(\mathbf{r},t)$ represents a forcing source, function of space ($\mathbf{r}(x,y,z)$) and time.
The acoustic pressure is related to the displacement potential through the following expression:

$$p = -K\nabla^2\psi \tag{2.2}$$

With K being the water bulk modulus: $K = \rho c^2$, so that then the acoustic pressure can be found as

$$p = -\rho\frac{\partial^2\psi}{\partial t^2}$$

A frequently used approach of solution is by using the Fourier transform on eq(2.1) obtaining the so called *Helmholtz equation*, in which appears the wavenumber expression $k_p = \frac{\omega}{c_p}$:

$$[\nabla^2 + k_p^2(\mathbf{r},\omega)]\tilde{\psi}(\mathbf{r},\omega) = \tilde{f}(\mathbf{r},\omega) \tag{2.3}$$

Where the quantities with the tilde represent the frequency domain expressions after the application of the Fourier transform

$$\begin{cases} x(t) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \tilde{x}(\omega) e^{i\omega t} d\omega \\ \tilde{x}(\omega) = \displaystyle\int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt \end{cases} \tag{2.4}$$

The solution of the problem requires the definition of appropriate boundary conditions and to find an adequate particular solution other than the Hankel functions. The whole solution process will not be presented since it has been exhaustively described in the referenced works.
Instead, an overview of the available techniques to solve eq(2.3) will be discussed.

2.2 NUMERICAL MODELS

A class of solution method often used to deal with differential equations, not only related to wave propagation, relies on discretization of the problem and numerical computation.
If on one hand each numerical model description and solution is relatively easy, and they allow to also consider properties variation in the domain considered, on the other hand some critical limitations arises.
First, the choice of the discretization dimensions is of crucial importance because it can not only affect the accuracy of the solution, but also its stability and convergence to the "real" one.
However the sound propagation problems often requires the definition of large domains, and the requirements to satisfy the stability and accuracy of the solution would make the solution computationally expensive. Considering that one of the most common techniques to deal with the uncertainties in properties is by solving several problems, numerical models do not represent the best alternative.
Two of the most widely used numerical methods will now be briefly presented, without entering in the details of the coupled problem and various versions of each one. For a deeper understanding of the results that can be obtained, as well as some stochastic finite difference method works (not directly related to sound propagation), the reader can refer to [1, 12, 13, 14, 15, 16]. A third method, more focused on sound propagation far from the source, will be then presented.

2.2.1 *Finite difference methods (FDM)*

One of the oldest and widely used numerical method is the *finite difference method*, due to its simplicity in the definition of the necessary

equations and the solution, as well as its versatility for various different problems.
In this section an overview of the method, through an example will be shown.

Recalling eq(2.1), this time defined for pressure p as done in [11], and for simplicity referring to a 2-D problem, the wave equation for a point source of strength $S(t)$ located at (x_s, z_s) is:

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} - \frac{1}{c_p^2}\frac{\partial^2 p}{\partial t^2} = S(t)\delta(x - x_s)\delta(z - z_s) \tag{2.5}$$

Where δ is the Dirac delta function: $\delta(0) = 1, \delta(\neq 0) = 0$.
The discretization and consequently the expression of (2.5) follows as presented in [12]:

$$\begin{aligned} & x_i = i\,\Delta x,\ i = 0..N_x,\ t_i = i\,\Delta t,\ i = 0..N_t \\ & \frac{\partial^2}{\partial t^2} p(x_i, z_j, t_n) \sim \frac{p_{i,j}^{n+i} - 2p_{i,j}^n + p_{i,j}^{n-1}}{\Delta t^2} \\ & \frac{\partial^2}{\partial x^2} p(x_i, z_j, t_n) \sim \frac{p_{i+1,j}^n - 2p_{i,j}^n + p_{i-1,j}^n}{\Delta x^2} \\ & \frac{\partial^2}{\partial z^2} p(x_i, z_j, t_n) \sim \frac{p_{i,j+1}^n - 2p_{i,j}^n + p_{i,j-1}^n}{\Delta z^2} \end{aligned} \tag{2.6}$$

Which leads to:

$$\begin{aligned} \frac{p_{i,j}^{n+i} - 2p_{i,j}^n + p_{i,j}^{n-1}}{c_p^2 \Delta t^2} = & \left(\frac{p_{i+1,j}^n - 2p_{i,j}^n + p_{i-1,j}^n}{\Delta x^2} + \frac{p_{i,j+1}^n - 2p_{i,j}^n + p_{i,j-1}^n}{\Delta z^2} \right) + \\ & + S(t_i)\delta(x_i - x_s)\delta(z_i - z_s) \end{aligned} \tag{2.7}$$

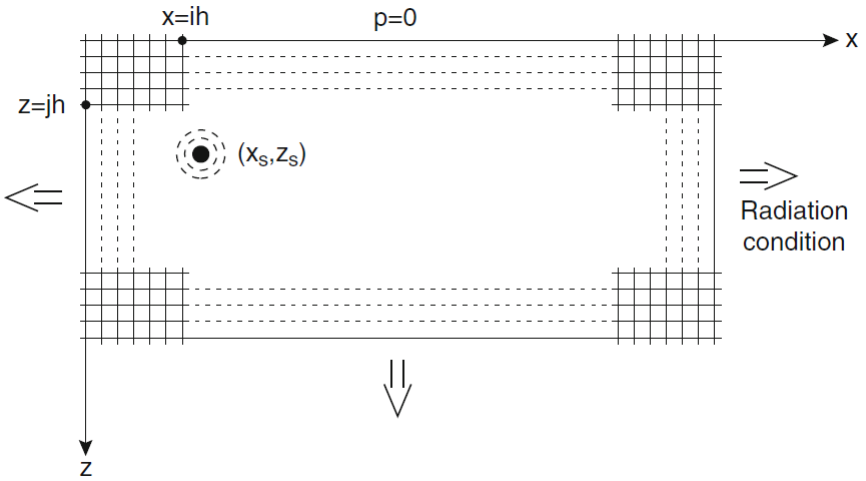

Figure 2.1: FDM visual discretization for the point source problem

In fig(2.1) a visual interpretation of the discretized domain is presented, while in fig(2.2) an example of solution is shown.
Although the mesh size can be reduced in order to try to minimize the error with the real solution, this arises the problem of excessive computation time.

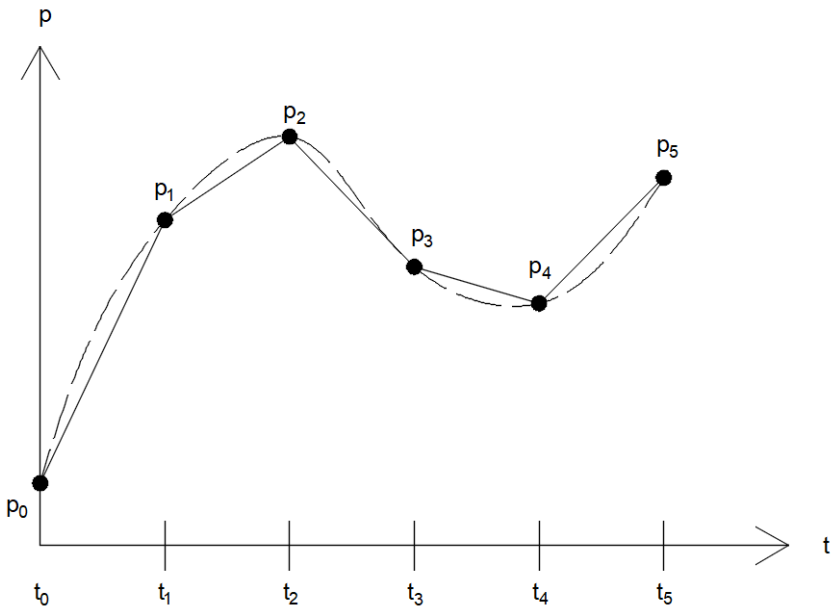

Figure 2.2: Discretized solution for $p(x, z, t)$, the step size needs to be small enough to accurately approximate the real solution (dashed line)

Other than the accuracy of the solution, another critical requirement of the discretization is to lead to stable results. For complex problems (as the sound propagation in marine environment) it is not so easy to determine the convergence and stability of the algorithm used, and is not sufficient to simply chose a really small step size, because then the computational time would become excessive. Moreover, the description of the boundary conditions, as well as reflection between different surfaces, can be troublesome.

2.2.2 *Finite element methods (FEM)*

Another method relying on the discretization and numerical solution is the *finite element method*.
Unlikely FDM, that as said discretizes the differential equations at local level, FEM discretizes the problem on a global scale, so that the definition of what is called (for structural elements) *stiffness matrix*, is independent from boundary conditions.
This makes FEM more versatile for complex geometries or complex boundary conditions, that can be problematic for FDM.

An example of FEM discretization is shown in fig(2.3).
Another feature of the FEM is the possibility to define different order of shape functions, still maintaining the compatibility between elements: these interpolates the coefficients at the nodes, and can be chosen to fulfill requirements on the type of displacement or stresses expected, see the example in fig(2.4).

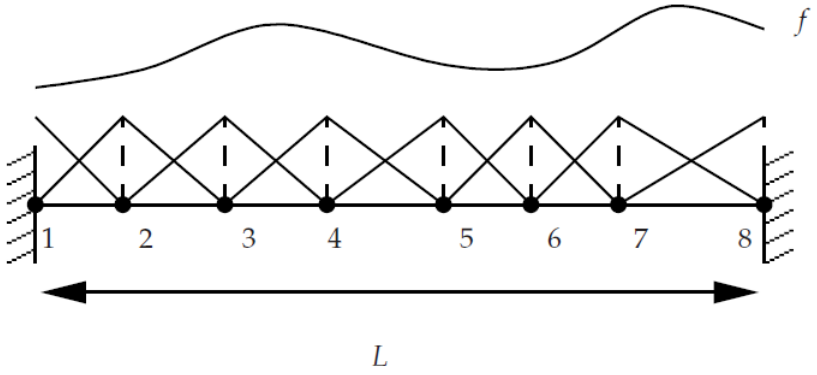

Figure 2.3: Example of FEM discretization, the displacement field is defined as connection of single element shape functions

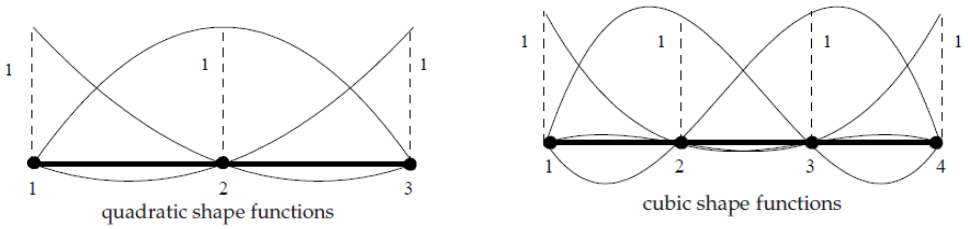

Figure 2.4: Different types of shape functions that can be used in a model

On the other hand, more than FDM, FEM can be problematic for large domains. Due to the excessive number of elements that need to be defined for a fine mesh the computational time required to solve several simulations would be unbearable.
However, due to their versatility and their possibility to take into account also phenomena like reflection of waves between different materials, FE are commonly used in the near field (i. e. near the noise source). This allows to define the necessary number of elements only in a limited region of the domain. For the far field (distant from the noise source) the FEM results are coupled to parabolic equations (PE): with the assumption that the reflection has already been attenuated, it is possible to simplify the problem to simple wave propagation.
An example of the use of this coupling is presented in [1, 17].

Although FE are a valuable method to approach the coupled problem, their high demand of computational power and time still do not make it the best option for determining the solution of underwater noise propagation.

2.2.3 *Parabolic equations (PE)*

When dealing with sound propagation problems, the domain that needs to be considered usually reaches kilometers away from the sound source, and that is the main reason why FD and FE models alone are not a convenient choice, because as mentioned would require a too demanding computational time.
It is then of practical use the definition of a *near-field* close to the sound source, where the problem is modelled using either FD or FE to account for the interaction between pile, water and soil. A *far-field* distant from the source is then defined, in which the main focus regards the sound propagation, and the main numerical scheme used is the *parabolic equations*.
These equations mainly focus on propagation in a medium, using the results from FD/FE as matching and initial conditions. Although their versatility and simplicity of implementation, there are some issues which can lead to significant errors in the prediction of the solution.

First of all, as mentioned in [11], the definition of PE makes them sensitive to the propagation direction, leading to inaccurate results for steep angles: situations like this may occur especially during reflection at the seabed interface, especially if the bottom is inclined. The errors can then increase particularly in shallow waters environments.
In recent years some PE schemes have been implemented such that can deal with steep propagation angles and still provide accurate results, although these require more computation effort.

Another issue encountered while using PE regards the energy conservation during propagation, with energy loss or gain depending on the bottom slope, which again result in prediction errors.
For seek of completeness, the general expression for a 3-D PE propagation as given in [11] will be shown:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 p}{\partial \varphi^2}+\frac{\partial^2 p}{\partial z^2}+\frac{\omega^2}{c_p^2(r,\varphi,z)}p=0 \tag{2.8}$$

In the case of φ-coordinate independence, the equation reduces to the standard 2-D Helmholtz equation:

$$\frac{\partial^2 p}{\partial r^2}+\frac{1}{r}\frac{\partial p}{\partial r}+\frac{\partial^2 p}{\partial z^2}+k_0^2 n^2 p=0 \tag{2.9}$$

Where $p(r,z)$ is the acoustic pressure, $k_0 = \omega/c_0$ is a reference wavenumber, and $n(r,z) = c_0/c_p(r,z)$ is the index of refraction.
The trial solution is in the form of a cylindrical outgoing wave:

$$p(r,z) = \psi(r,z)H_0^{(1)}(k_0 r) \tag{2.10}$$

The Hankel function that satisfies the Bessel differential equation is often replaced by the asymptotic relation $k_0 r \gg 1$, so that it becomes:

$$H_0^{(1)}(k_0 r) \sim \sqrt{\frac{2}{\pi k_0 r}} e^{i(k_0 r - \pi/4)}$$

Once substituting the trial solution (2.10) into eq(2.9) and again assuming the farfield condition that $k_0 r \gg 1$, we obtain the simplified elliptic wave equation:

$$\frac{\partial^2 \psi}{\partial r^2} + 2ik_0 \frac{\partial \psi}{\partial r} + \frac{\partial^2 \psi}{\partial z^2} + k_0^2(n^2 - 1)\psi = 0 \tag{2.11}$$

Finally is introduced the *paraxial approximation* for small angles, expressed as:

$$\frac{\partial^2 \psi}{\partial r^2} \ll 2ik_0 \frac{\partial \psi}{\partial r}$$

So that the standard parabolic equation becomes:

$$2ik_0 \frac{\partial \psi}{\partial r} + \frac{\partial^2 \psi}{\partial z^2} + k_0^2(n^2 - 1)\psi = 0 \tag{2.12}$$

Which can be solved either by use of FD of FE schemes once the boundary conditions are defined.

2.3 (SEMI) ANALYTICAL MODELS

A different class of solution techniques rely on analytical expressions, but making use of numerical solvers to determine the unknowns of the problem; therefore they are often called *semi-analytical methods*.
Among the advantages these methods have, there's the accuracy, due to the possibility to avoid discretization and use analytical expressions, and computation speed, compared to numerical models.
One of the major limitation of the semi-analytical approaches regards the assumptions made on the domain geometry. Often symmetry is used to simplify the expressions and allow a faster computation. It then follows the impossibility of defining continuously varying properties in space, so that the only way to approximate these variations is through definition of different layers.

In the next chapter the adopted model, which falls into the semi-analytical class, will be described. It has been exhaustively explained in [18, 19], where specifically three different approaches to model the soil have been discussed. The advantages and disadvantages of each one will be then considered.

2.4 PROBABILISTIC APPROACHES DEALING WITH UNCERTAINTY

Many physical problems are usually solved relying on models that somehow simplify the *real* world, or that are valid under some assumptions (often met). If these assumptions are met, then the solution can be determined.
However, either the assumptions not always halt, or there is a high uncertainty in setting the model inputs. It is generally the case when material properties are involved, which are never a completely deterministic feature.
To deal with variability and errors that may occur, in recent years probabilistic and statistical approaches gained a lot of popularity in many different fields. Their advantages consist in the relative simplicity with which they can account for uncertainty; moreover the concept of low or high probability of an outcome is generally of strong impact.
In this section some of the most common statistical techniques will be presented.

2.4.1 *Monte Carlo analysis*

One of the most common method used to solve complex problems, in which several variables are involved, is the Monte Carlo analysis.
The main concept of it is to randomly generate the input of the model, which should be deterministic, and obtain the results. Repeating the solution several times it is possible to obtain each time different outcomes, which however can highlight trends or insight on the sensitivity with respect to certain inputs.
One of the reason behind its popularity is that it doesn't require any change in the deterministic model. Generally only the distributions of the input features are required from which samples will be generated.
A drawback of the method is however the excessive computation time required to obtain accurate results: the more are the trials the better.
To partially solve or improve the method, convergence criteria are usually set. For example if doubling the samples the results vary less than a threshold 'ε', the method may be considered to have reached convergence.
Another technique is the *importance sampling* [20]: instead of randomly

generate from the real distributions, new ones are defined such that they may investigate on critical combinations. Usually this allows to reduce the overall number of samples to obtain the same results.

2.4.2 *Multi-dimensional random fields*

Uncertainty in physical properties is a component that, especially in laboratory tests, is often solved by the definition of distributions of the latter. One thing that needs to be remembered is that all the samples can be compared between each other having generally same volume, shape, or other.
In reality, however, most of the time not only the physical uncertainty is present, but also *spatial variability*.
When elements define a domain, and their properties can vary in space (and time) then they define a *random field* [21]. Depending on the type of correlation between the properties, different random fields can be defined, one for example, the Gaussian field.
Fairly recent studies are involved in the definition of fields that can predict the spatial variation of several soil properties based on just a few empirical test. Then by mean of *spatial correlation matrices* 'connect' the tests from different places and define the behaviour in-between.
Imagine for example that in a domain big as a football pitch, CPT have been performed at the corners so that the properties are known in those locations. Through the spatial correlation matrix it is possible to define the correlation of each point of the pitch with the surrounding elements. Inserting then the input measurements of the corners, the in-between properties are then evaluated, accounting for the strength of correlation with the neighbouring points.
A spatial correlation matrix element $C_{i,j}$, of a point located at position x^*, y^* in space, can be defined as shown in eq(2.13).

$$C_{i,j} = \exp\left(-\frac{\sqrt{(x_i - x^*)^2 + (y_j - y^*)^2}}{L_c}\right) \tag{2.13}$$

Where L_c is the so called *correlation length*, and estimates the spatial intensity of the correlation. A higher correlation length means that a point is highly related to points even quite distant from it.
In fig(2.5, 2.6) an example of correlation matrices for two different points, with different correlation lengths, are shown.
In the first case, a point located at $x = 15, y = 10$, with a correlation length of $L_c = 10\,m$ is presented. It can be seen how a high correlation ($r > 0.7$) is spread for $10 \leq x \leq 20, 5 \leq y \leq 15$.
For the second case the length is reduced to $L_c = 5\,m$, meaning that the specific point is only related to neighboring elements, and not highly

affected by points distant to it.
This approach, however, is applicable only when at least a few measurements in different positions are available. Moreover, the results may be more useful in models which allow to consider varying properties.

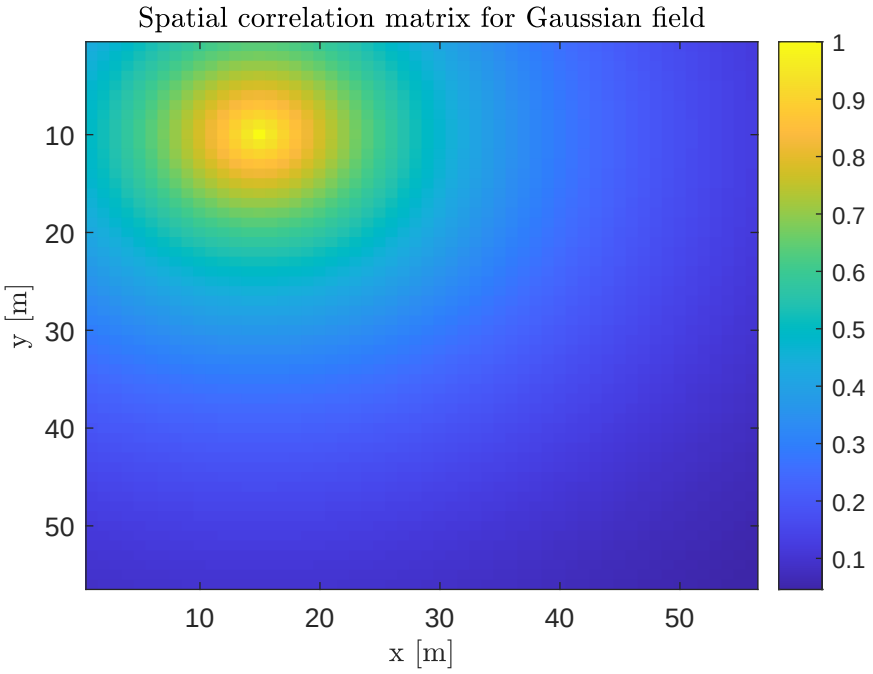

Figure 2.5: Spatial correlation for a point located at $x = 15, y = 10\,m$, $L_c = 10\,m$

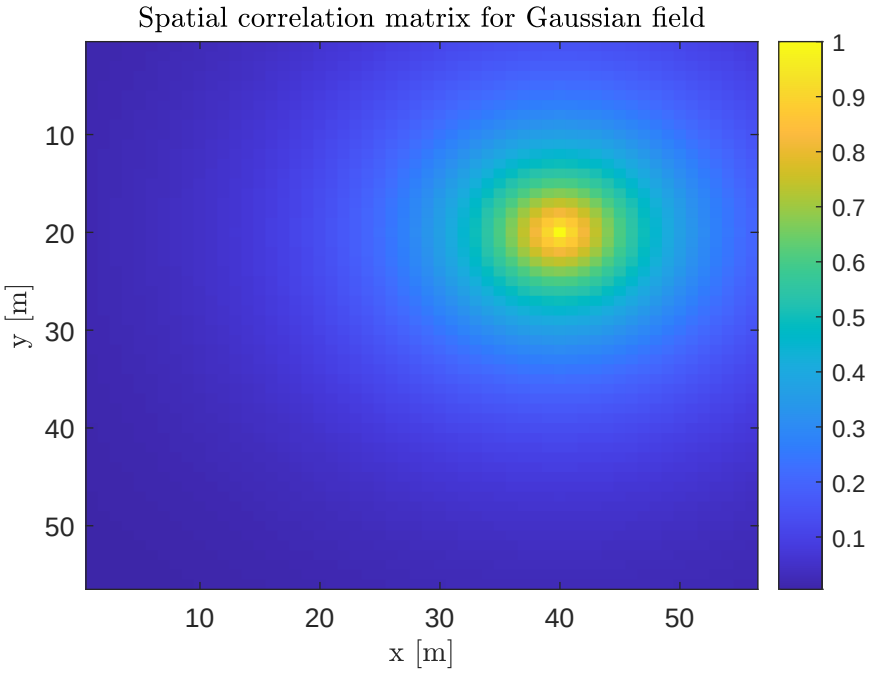

Figure 2.6: Spatial correlation for a point located at $x = 40, y = 20\,m$, $L_c = 5\,m$

3

ADOPTED NOISE PREDICTION MODEL

In the following chapter, the *Coupled problem* model that is used throughout the work, and implemented in the software *Silence*, is presented.
The model is the result of previous works which described the problem in analytical terms and investigated in the most accurate description of the soil medium [18, 19, 22]. In the conclusive works the soil, treated as an elastic medium, accurately accounts for surface (Scholte) waves which have been found to be, after the Mach cones in the water column, the second most important sound propagation path.
In the works which followed [23] the numerical implementation is addressed, and this chapter mainly reports the main information needed to have a clear picture of the features of the approach and its advantages and limitations, which are also present in [7, 24].
The model belongs to the semi-analytical class, and as highlighted in the previous chapter, its advantages reside in the computational speed with which the sound propagation results can be obtained. On the other hand, a disadvantage remains the limitation to simple and constant geometries.
The main aim of this thesis is to obtain sufficient sound outcomes to define quite accurate sound distributions. Another goal is the identification of the dominant parameters which highly affect the propagation and intensity of sound. Therefore, the faster computation time is considered to be more important than an accurate geometry description. Once the critical features of the problem are found, an accurate model is easier to be defined.

3.1 MODEL OVERVIEW

In fig(3.1) a schematization of the model is shown: a pile, surrounded by a fluid-soil domain, excited by an external force which can be related to the hammer blow.
The pile is modelled and described by shell equations, including shear deformation and rotational inertia; the constants $E, \nu, R, \rho, 2h$ stand for the complex modulus of elasticity of the pile, the Poisson's ratio, the mid-surface radius, the density and the thickness of the shell. A coordinate system is defined at the top of the pile such that it occupies the domain $0 \leq z \leq L$ and $0 \leq r \leq R$.
The fluid occupies the region $z_0 \leq z \leq z_1$ and $r > R$, it is modelled as a three dimensional inviscid compressible medium with a pressure

release boundary describing the sea surface. The soil, occupying the region $z_1 < z < H > L$ and $r > R$ is modelled as a three dimensional elastic continuum; at $z = H$ a rigid boundary is defined.
In the next sections the governing equations are presented.

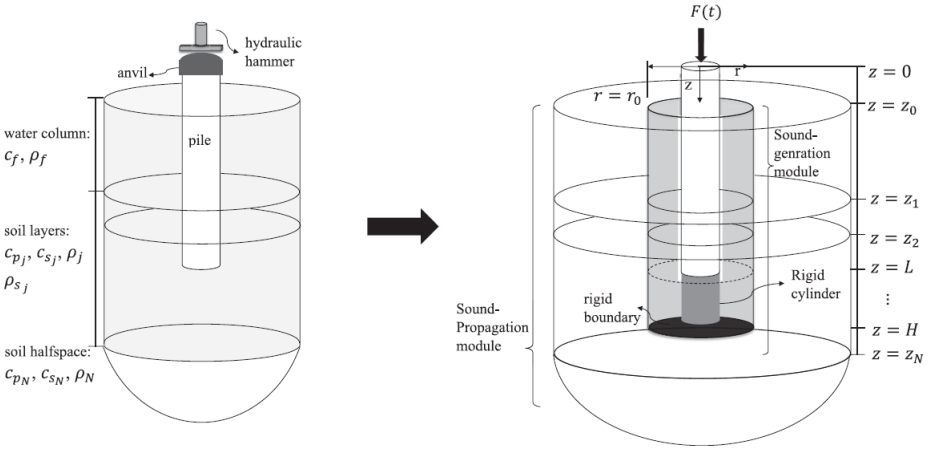

Figure 3.1: Model schematization. The distance $r = r_0$ is where the coupling between near and far-field happens [23].

3.2 GOVERNING EQUATIONS

In this section the governing equations that define the problem are described. Each component of the model is solved and at the end a coupling is performed to obtain the final solution.
The case considered is the one of a cylindrical symmetric case, without asymmetric forcing input.

3.2.1 *Shell vibration*

The equation that governs the pile under forcing is the following:

$$\mathbf{Lu} + \mathbf{I\ddot{u}} = [H(z-z_2) - H(z-L)]\mathbf{t}_s - [H(z-z_1) - H(z-z_2)]\mathbf{p}_e + \mathbf{f}_e, 0 \leq z \leq L \tag{3.1}$$

Where $\mathbf{u} = [u_z(z,t), u_r(z,t)]^T$ is the displacement vector in vertical and radial direction, **L, I** are respectively the stiffness and modified inertia operators of the shell (pile), H is the *Heaviside* function, $\mathbf{t}_s, \mathbf{p}_e$ are the stress reaction of the soil surrounding the shell, and the fluid pressure outside the pile, finally $\mathbf{f}_e = [f_{rz}(z,t), f_{rr}(z,t)]^T$ is the external force vector.
We apply the Fourier transforms (2.4) and we obtain the frequency domain expression of the shell motion.

$$\mathbf{L\tilde{u}} + \mathbf{\tilde{I}\tilde{u}} = [H(z-z_2) - H(z-L)]\tilde{\mathbf{t}}_s - [H(z-z_1) - H(z-z_2)]\tilde{\mathbf{p}}_e + \tilde{\mathbf{f}}_e$$

(3.2)

3.2.2 *Fluid domain*

$$\nabla^2\phi_f(r,z,t) - \frac{1}{c_f^2}\frac{\partial^2\phi_f(r,z,t)}{\partial t^2} = 0, \quad z_0 < z < z_1, \quad R < r < \infty \tag{3.3}$$

Where c_f is the speed of waves in the fluid.
The motion of an inviscid compressible fluid can be described by a scalar velocity potential $\tilde{\phi}_f(r,z,\omega)$ by the frequency domain Helmholtz equation:

$$\nabla^2\tilde{\phi}_f(r,z,\omega) + \frac{\omega^2}{c_f^2}\tilde{\phi}_f(r,z,\omega) = 0 \tag{3.4}$$

By separation of variables and considering the radiation condition at infinity $\lim_{r\to\infty}\tilde{\phi}_f = 0$ a solution is found as:

$$\tilde{\phi}_f(r,z,\omega) = H_0^{(2)}(k_{\phi,f}\,r)\left(A_1\exp(-\alpha_f z) + A_2\exp(\alpha_f z)\right) \tag{3.5}$$

Where $k_{\phi,f}$ is a separation constant, $\alpha_f = \sqrt{k_{\phi,f}^2 - \omega^2/c_f^2}$, A_1, A_2 are unknown complex constants.
The velocity and pressure potential are then:

$$\begin{aligned} &\tilde{\mathbf{v}}(r,z,\omega) = \nabla\tilde{\phi}_f(r,z,\omega) \qquad with\ \nabla = \frac{\partial}{\partial r}\hat{r} + \frac{\partial}{\partial z}\hat{z} \\ &\tilde{p}_f(r,z,\omega) = -i\omega\rho_f\tilde{\phi}_f(r,z,\omega) \end{aligned} \tag{3.6}$$

3.2.3 *Soil domain*

The equations that govern the wave propagation in an undamped soil medium, through displacement potentials, are the following, valid in the region $z_1 < z < H$ and $R < r < \infty$:

$$\begin{aligned} &\nabla^2\tilde{\phi}_s(r,z,t) - \frac{1}{c_L^2}\frac{\partial^2\phi_s(r,z,t)}{\partial t^2} = 0 \\ &\nabla^2\psi_s(r,z,t) - \frac{\psi_s(r,z,t)}{r^2} - \frac{1}{c_T^2}\frac{\partial^2\psi_s(r,z,t)}{\partial t^2} = 0 \end{aligned} \tag{3.7}$$

Similarly, the displacement potentials that describe wave propagation in a layered soil medium in frequency domain are:

$$\begin{aligned} &\nabla^2\tilde{\phi}_s(r,z,\omega) + \frac{\omega^2}{c_L^2}\tilde{\phi}_s(r,z,\omega) = 0 \\ &\nabla^2\tilde{\psi}_s(r,z,\omega) - \frac{\tilde{\psi}_s(r,z,\omega)}{r^2} + \frac{\omega^2}{c_T^2}\tilde{\psi}_s(r,z,\omega) = 0 \end{aligned} \tag{3.8}$$

With $c_L^2 = \frac{\lambda_s + 2\mu_s}{\rho_s}$, $c_T^2 = \frac{\mu_s}{\rho_s}$.

The displacement potentials and stress relations in the frequency domain read:

$$\begin{aligned}
\tilde{u}_{s,r} &= \frac{\partial \tilde{\phi}_s}{\partial r} - \frac{\partial \tilde{\psi}_s}{\partial z} \\
\tilde{u}_{s,z} &= \frac{\partial \tilde{\phi}_s}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r\tilde{\psi}_s) \\
\tilde{\sigma}_{s,rr} &= \lambda_s \left(\frac{\partial \tilde{u}_{s,r}}{\partial r} + \frac{\tilde{u}_{s,r}}{r} + \frac{\partial \tilde{u}_{s,z}}{\partial z} \right) + 2\mu_s \frac{\partial \tilde{u}_{s,r}}{\partial r} \\
\tilde{\sigma}_{s,zz} &= \lambda_s \left(\frac{\partial \tilde{u}_{s,r}}{\partial r} + \frac{\tilde{u}_{s,r}}{r} + \frac{\partial \tilde{u}_{s,z}}{\partial z} \right) + 2\mu_s \frac{\partial \tilde{u}_{s,z}}{\partial r} \\
\tilde{\sigma}_{s,zr} &= \mu_s \left(\frac{\partial \tilde{u}_{s,r}}{\partial z} + \frac{\partial \tilde{u}_{s,z}}{\partial r} \right)
\end{aligned} \tag{3.9}$$

Finally, the expression for the two potential reads as:

$$\begin{aligned}
\tilde{\phi}_s &= H_0^{(2)}(k_{\phi,s}\, r)(A_3 \exp(-\alpha_s z) + A_4 \exp(\alpha_s z)) \\
\tilde{\psi}_s &= H_1^{(2)}(k_{\psi,s}\, r)(A_5 \exp(-\beta_s z) + A_6 \exp(\beta_s z))
\end{aligned} \tag{3.10}$$

With $\alpha_s = \sqrt{k_{\phi,s}^2 - \omega^2/c_L^2}; \beta_s = \sqrt{k_{\psi,s}^2 - \omega^2/c_T^2}$ being the wavenumbers; the coefficients $A_3..A_6$ are the unknown complex coefficients of the soil region.
The introduction of the appropriate boundary conditions allows to define a system in which, for each wavenumber, is possible to determine the amplitude coefficients of the modes. The complete definition and solution process, as mentioned, is described in [19] and is not repeated here.

3.3 IMPORTANT PROPERTIES AND INSIGHT ON THE MODEL

The adopted model schematization and solution process has been chosen on the basis of some important advantages compared to alternative techniques. In this section they are highlighted, as well as some limitations that still remain.

3.3.1 *Completeness and accuracy*

The conclusions obtained at the end of previous works [18, 19, 24, 23] show how the semi-analytical model results in an accurate solution method, taking into account many physical phenomena involved in the problem.
In particular, different ways of modelling the soil have been investigated

[19], treating it as a *dashpot-spring system*, as an equivalent *acoustic fluid* and finally as a *three-dimensional continuum*.
The tests made to validate the models showed how the dashpot-spring system, although accurate, requires particular attention in the choice of the equivalent spring stiffness and damping coefficients, which makes it of difficult application in regions in which the soil properties, as said, are not well specified.
The use of the equivalent acoustic fluid represent a valid alternative in terms of computational speed, but shows to be accurate only when the following conditions apply [19]:

1. The energy is released in the fuid region and only the radiated field in the fluid domain is required.

2. The source and/or receiver are not positioned close to the seabed surface. At low frequencies and for sources and receivers close to (or in) the sea bottom, shear wave and interface modes of propagation may dominate.

3. The dynamics of the soil region do not influence the radiating source itself (dynamic response of the pile in this case).

Unfortunately, in the coupled problem these conditions are not sufficiently met. Therefore, the use of an equivalent acoustic fluid to represent the soil should be used with the maximum attention, otherwise would produce inaccurate results.
The final and most complete model considered, the three-dimensional continuum, has produced the most accurate results, compared also with empirical recordings. It has showed to be a computationally efficient option, when compared to FDM or FEM alternatives [1, 25, 26], but most importantly a complete model, enabling to consider the many physical phenomena occurring in the sound propagation.
Although the primary sound travelling path is the fluid, where Mach cones radiates from the pile, it has resulted that a second route is the seabed interface. There, the generation of interface Scholte waves represents a significant phenomenon of energy and sound pressure propagation.
The model has also shown (regarding the near-field) to be the most accurate when considering the evanescent spectrum of modes (imaginary and complex modes) [19, 23]. These have resulted to determine the accuracy in the near-field, and usually it is required to determine the necessary amount of which need to be accounted for, since these modes are infinite. In practice it has been demonstrated that the account for at least 100 of them results in an error of less than 0.5% compared to a solution that account for more [19].

3.3.2 *Limitations of the model*

While the model has proved to be accurate and computationally efficient, some limitations still remain related to the geometry representation and the properties assigned to the materials, first of all the soil.
The mentioned model can only represent a constant geometry, not describing the changes that a real seabed would have. Another, and probably more crucial lack, is the inability to account for space varying soil properties, which would also occur in a real environment. This problem, however, would be in common also with the alternative FDM and FEM models, since it is really difficult to account for all the variations of characteristics in a domain that expands, usually, for kilometers from the pile.
The feature that somehow allows to consider variations in the soil is the introduction of horizontally stratified layers. Since a greater number of them increases the computation time, up to three layers can be defined in the current version.
As is discussed more in detail in the next chapter, the empirical results that are possible to obtain only provide a local description for the soil. The main purpose of this thesis is, from the obtained recordings, to determine probabilistic distributions for the properties and investigate their influence on the sound level and its propagation.
Once the estimated sound levels are obtained, and statistically interpreted, the correlation between soil input and sound output can be investigated.
These information may help in estimating more carefully the dominant characteristics, developing also newer techniques.

Part II

STATISTICAL SOIL MODELLING AND SAMPLE GENERATION

4

SOIL CHARACTERIZATION

The method presented in ch.(3), implemented in *Silence*, is able to solve the problem of sound propagation accounting for fluid-pile-soil interaction when all the material properties and the (simple) geometry are known.

In order to solve the vibro-acoustic problem it is necessary to assign the input parameters to the materials involved. However, many uncertainties appear particularly regarding their dynamic properties, due to the high spatial variation marine environments can have. The lack and/or difficulties of performing geotechnical investigations in these environments makes it difficult to accurately define characteristics such as shear or compressional velocity, density, Poisson's ratio and other features.

One of the main practice used nowadays to obtain soil information is through *Cone penetration test* (CPT): one of the least invasive and yet accurate techniques which measures the shear friction and tip resistance applied on the instrument's tip during penetration in the ground.

The 'needle' is driven up to several tens of meters below the seabed, and the stress measurements are recorded.

In this chapter, relying on experimental results obtained from an *offshore wind farm* (OWF), a method to estimate properties ranges is presented, which later leads to the possibility of assuming a soil conformation.

The measurements obtained are then related to other properties through empirical relations [27, 28, 29], and by use of statistical relations, a probabilistic description of the soil is made.

Hopefully the method further presented will be helpful in situations where only numerical results, and not material samples (which in marine environments are hardly obtained), are available.

4.1 CPT AND EMPIRICAL SOIL PROPERTIES

As mentioned, the domain considered in the vibro-acoustic problem usually spans kilometers from the pile. It is then difficult to model the geometry of the marine environment. Moreover usually it is possible to rely on just a few measurements which cannot predict how the geometry and properties may vary in radial and vertical direction, the presence of local inhomogeneities, instrumental errors.

It is necessary to highlight that then, even if accurate, all these information could not be introduced in the model used (*Silcence*). The model

adopts a simple and homogeneous geometry, which means that the properties assigned do not vary in space. The possibility to define multiple layers tends to mitigate this simplification, although as soon as the number of layers increases, also the computation time does.
The strategy used in this thesis is to investigate in the definition of proper distributions for the parameters, so that various scenario are considered. Although still using a simple geometry, the aim is to find which and which parameter's combinations are highly affecting the sound levels.

The CPT provides measurements and variation along the depth for shear friction S_f and tip resistance q_t of the soil, which then through empirical relations can be related to other mechanical and dynamic characteristics.

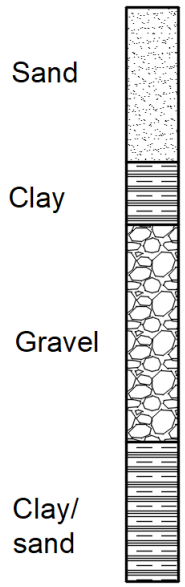

Figure 4.1: Possible estimation of the soil stratification

From the recordings, it is also possible to estimate the type of materials which compose the stratification, as shown in fig(4.1), although the only way to verify the prediction would be the extraction of samples. In these environments, however, this is hardly done due to its difficulty: the samples should be collected often tens of meters below the water, and then other tens in the soil. Additionally, this would also change the *in-situ* conditions compared to the laboratory ones, since the real environment conditions would not be reproduced (pressure, water content...). This would then affect the determination of the dynamic properties.
The prediction is then, often, made based upon similar environments which provided similar results, or based on manuals which contain ranges for each property [30].
In this work, although its high importance (especially from a design and

structural point of view), the correct determination of the soil strata is not investigated. The focus is instead put on the definition of accurate property ranges, allowing to investigate on various combination cases. Doing so, from a dynamic and acoustic point of view, it is believed that this allows to consider the most critical sound propagation outcomes.

4.2 FROM RECORDINGS TO DISCRETE VALUES

The continuous set of recordings from a CPT can't be directly used, therefore a discretization of the data is necessary: from the available measurements the area underneath the curve of measurements is evaluated at every $0.5\,m, 1\,m$, averaging over the step used. The discrete recordings are shown in fig(4.2), they extend up to $55\,m$ below seabed, and the ones obtained every $0.5\,m$ show more discontinuities. Since the recordings may be affected by instrument sensitivity and may be highly affected by local inhomogeneities, the measurements obtained averaging every $1\,m$ are considered. The reason is that they tend to show a less oscillatory behaviour, which could refer to a more gradual change in values.
One last remark is the choice of neglecting CPT recordings in the first two meters below seabed. This has been done for two main reasons.
First, the recordings in the first meters of the soil are almost unreadable, therefore the errors which may be introduced are extremely high. The shear velocity estimation in these first meters indeed presented unrealistically high values, since it is expected to have a low shear speed in a completely saturated material. Another study (in a different region and conditions) to confront the values with [31] also confirmed that the values obtained were excessively high and unrealistic. Second, the extreme change between fluid and soil is generally a source of errors for the instrument itself, therefore even an accurate reading should not be trusted completely. As is discussed later, for the first saturated layer, of a depth assumed of $2\,m$, *ad-hoc* properties have been defined, that rely on empirical experiments and physical behaviour.

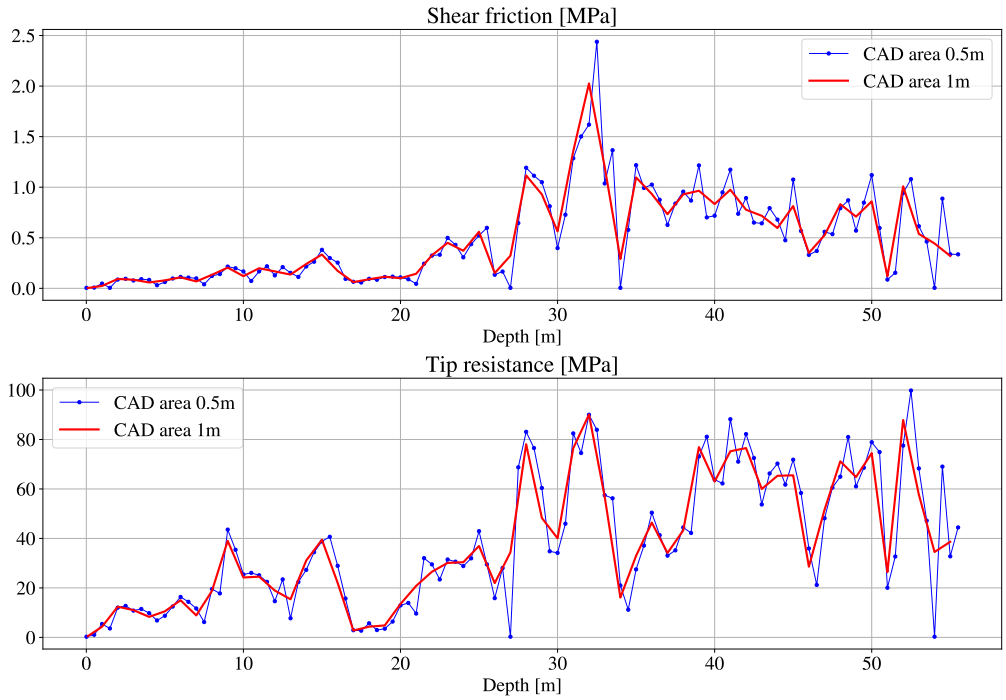

Figure 4.2: Shear friction S_f and tip resistance q_t discrete measurements from area averaging

4.2.1 *Shear velocity*

Once the stress components are known, it is possible to estimate (except for the first 2 m) the shear velocity applying one the empirical relations collected in [28], in this case the most general which applies to most of the marine soil types is:

$$V_s = (10.1 \cdot \log(q_t[kPa]) - 11.4)^{1.67} \cdot \left(\frac{S_f}{q_t} \cdot 100\right)^{0.3} \tag{4.1}$$

To verify how sensitive the empirical relation is, and how both measurement or reading differences may affect the values, on the shear friction and tip resistance of the CPT results has been applied an error of $\pm 10\%$. In appendix (A.1) the sensitivity for the three main soil characteristic is investigated. The numerical results for the shear velocity produced values which were different for less that 5% even in the most extreme case, with respect to the so called *undisturbed CPT values*, which are the ones obtained directly from the CPT. Since the distribution later defined covers a wider range of possible values, these have been obtained without applying any error to the measurements. It also demonstrate that although more empirical data might improve the distribution definition, the statistical interpretation helps in dealing with uncertainties and spatial variability.

In fig(4.3) the shear velocity estimation from eq(4.1)are shown. Although evaluated, the results in the first 2 m are disregarded.

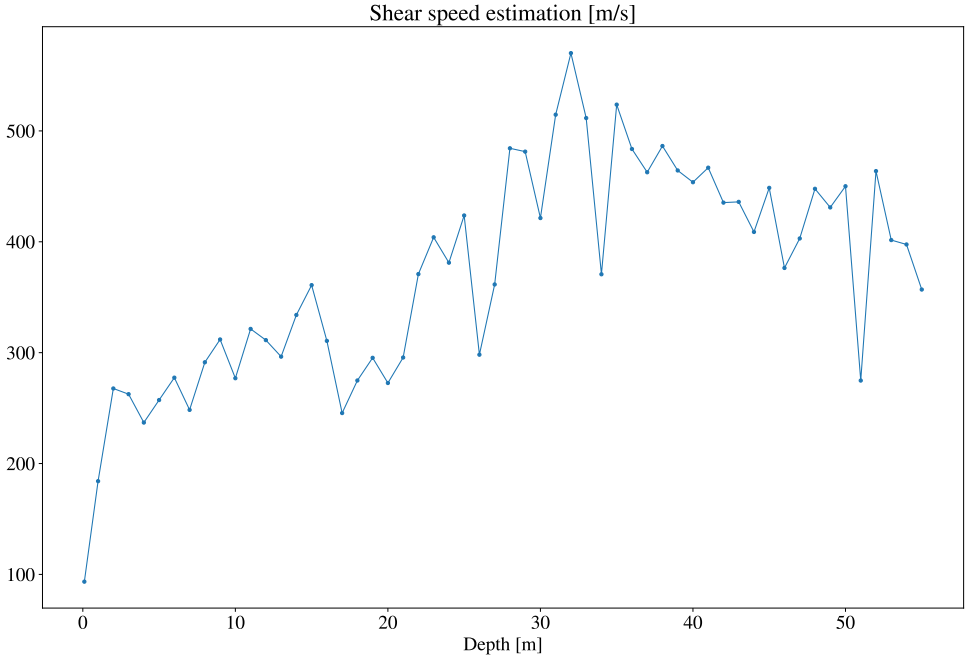

Figure 4.3: Shear velocity estimation from CPT measurements

4.2.2 *Density*

Density, as for the shear-wave velocity, can be obtained from empirical relations [28]:

$$\gamma_r[kN/m^3] = 8.32 \cdot \log(V_s) - 1.61 \cdot \log(z) \tag{4.2}$$

Applying a $\pm 10\%$ to the CPT recordings, as has been done for the shear speed case, allows to verify how sensitive to the measurements eq(4.2) is: the estimation do not differ for more than 1%. As said before, then, the recordings can be used, and having more accurate ones would only improve the results.
The density estimation are plotted in fig(4.4) and seem in line with experimental results and common properties value, as referred in [30].

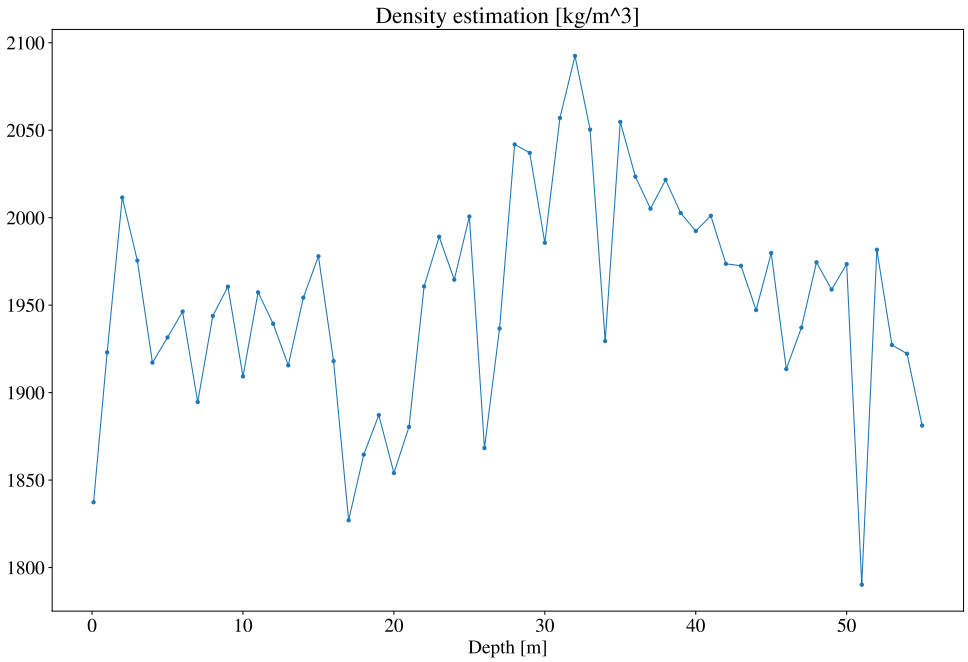

Figure 4.4: Density estimation from CPT results

4.2.3 *Compressional velocity*

The shear speed and the density properties can be obtained directly from the CPT measurements. Another fundamental soil feature is the compressional velocity.
There exist many empirical relations to derive the compressional speed both from the density or the shear speed [29]. Unfortunately, they are based on investigations made on the mainland and not in marine environments. However, two main empirical relations are compared, and based on measurements performed in marine environments [31] a mean between the two is chosen.
The first relation matches the compressional speed to the shear one (expressed in m/s):

$$V_p[\frac{km}{s}] = 0.9409 + 2.0947V_s - 0.8206V_s^2 + 0.2683V_s^3 - 0.0251V_s^4 \quad (4.3)$$

The second relates the compressional speed to the density (expressed in g/cm^3), although it is valid for relatively high soil densities ($\sim 2000 - 4500\, kg/m^3$):

$$V_p[\frac{km}{s}] = 39.128\rho - 63.064\rho^2 + 37.083\rho^3 - 9.1819\rho^4 + 0.8228\rho^5 \quad (4.4)$$

The results are shown in fig(4.5), where also the mean between the approaches is presented.

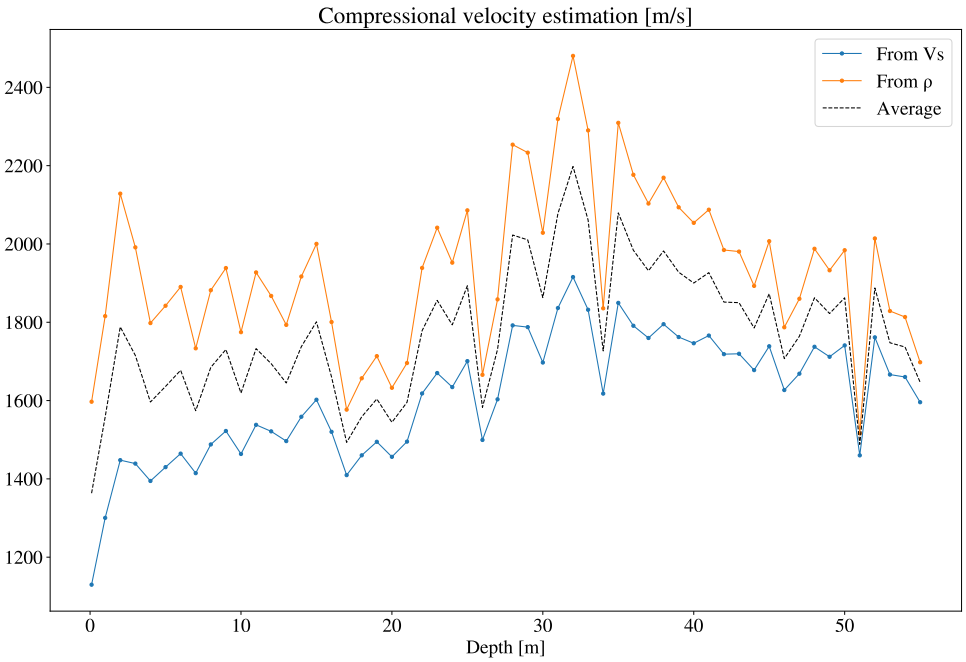

Figure 4.5: Compressional velocity estimation from shear speed and density, and their mean

The main concern regarded the velocity in the first few meters of depth. Since the soil is assumed to be highly saturated, it is expected that the compressional velocity is similar to the one in the water ($c_p = 1500\,m/s$). It is then expected an increase in compressional speed, since at lower depths, due to the compression the soil undergoes by the soil above, the density will be higher. A denser material allows compressional waves to travel faster.
The estimated results in fig(4.5) show these behaviours.

4.2.4 *Poisson's ratio*

The Poisson's ratio is an important material characteristic. Unfortunately, it is still not known how to measure it from experiments, and it is then necessary to derive it once the shear and compressional velocity are known.
From the well known definition of compressional velocity, where $\kappa = \sqrt{\frac{E}{3(1-2\nu)}}$ is the Bulk modulus, G is the shear modulus and ρ the density:

$$V_p = \sqrt{\frac{\kappa + \frac{4}{3}G}{\rho}}$$

And introducing the expression of $V_s = \sqrt{\frac{G}{\rho}}$, rearranging the equation a relation for the Poisson's ratio is found as [32]:

$$\nu = \frac{V_p^2 - 2V_s^2}{2(V_p^2 - V_s^2)} \tag{4.5}$$

It has to be mentioned that the Poisson's ratio is highly sensitive to changes in either the compressional or shear velocity, therefore a high numerical accuracy should be taken into account. Another aspect that has to be considered is the high nonlinearity of the relationship. It does not make then much sense to define a Poisson's ratio distribution.
Since the features which may be measured are, as shown, shear speed, density, and (partly) compressional velocity, these three are the ones later generated from appropriate distributions. The Poisson's ratio, and the Young's modulus introduced in the next section are then derived once the main three characteristics are defined.

4.2.5 *Young's modulus*

Finally, Young's modulus is estimated. The relation with the shear velocity and Poisson's ratio is the well known [32]:

$$E = 2(1+\nu)G = 2(1+\nu)\rho V_s^2 \tag{4.6}$$

As for the Poisson's ratio, Young's modulus is not directly measured, and for the analysis that follow it is determined without the involvement of distributions.

5

STATISTICAL ANALYSIS OF THE SOIL

The adopted model, as mentioned before, can't describe spatially varying properties. The CPT recordings, and the obtained soil properties, cannot be directly inserted as input in the software.
It is then necessary to define appropriate soil layers. To do so, analyzing the estimated properties, similarities along the depth are searched for.
One of the first questions that arises is how to establish the depth of each layer, and moreover how to define proper distributions that represent the obtained data.
As mentioned, it has been preferred to disregard the recordings for the first two meters since a specific saturated layer is defined. The rest of the recordings are investigated so that an additional (2nd) layer or two additional (3 layers in total) can be defined.
In this section is presented the main procedure adopted to define the depth of the layers through *ANOVA*, the definition of the distributions using *maximum-likelihood estimation*. Finally the choice of the best fitting one by mean of the *AIC* criterion.
The correlation between the variables **within** the layer is evaluated, while it is assumed that each strata is independent from the others. This has been done to simplify the samples generation.
Due to the lack of available data, an approach based on *Gaussian random fields* [21] was not possible. Future works are investigating on how to treat large spatially varying domains and how to extrapolate information from the recordings.
The method presented further aims to be a solid approach even for future works that need to deal with limited discretized recordings and uncertainty.

5.1 SOIL LAYERS DEFINITION

When using homogeneous soil layers it is of importance to know how many sufficiently represent the 'real' environment. Usually a few are sufficient and a trade-off between number of layers and computational time is performed.
Based on the number of strata that one needs to define, it is possible to group the recordings in regions which show similar traits. In statistical words, the division of the layers is done by grouping data that show similar mean and minimum standard deviation.
An approach that analyzes if two or more set of data have different

means, and therefore belong to different groups, is the *analysis of variance*.

5.1.1 *ANOVA*

The analysis of variance (ANOVA) refers to a technique to verify whether two or more groups have different means. Depending on the amount of independent variables involved in the various samples, ANOVA can be single, double or triple factor. The aim of the analysis is to tell if at least one of the groups considered presents a high difference with respect to the others. The limitation consists in the fact that cannot tell which group/s differ from the others, but since in the extreme case here we deal with 2 groups (layers), this won't be a problem.
Another limitation is that it does not tell if elements of one group may belong to another. Therefore, for the determination of the 'optimal' layer's depth an iterative approach is necessary.
In the case followed, both shear velocity, compressional speed and density have been analyzed to determine the depths of the soil stratifications.
In principle, the method wants to confront the "null" hypothesis, H_0, that **all** means of the groups are equal, against the "alternative" hypothesis H_a that **at least** one group mean is different from the others (5.1).

$$\begin{cases} H_0 : \mu_1 = \mu_2 = ...\mu_n \\ H_a : \text{at least one mean differs from the others} \end{cases} \tag{5.1}$$

The calculations involved regards the mean the quadratic sum, and the "degrees of freedom" (dof).
If we define:

$$\begin{aligned} &a\,, b : \text{the 'a' elements within the 'b' groups (i. e. layers)} \\ &N : \text{total number of elements, } a \cdot b \text{ when all the layers} \\ &\quad \text{have same number of elements} \\ &X_{ij} : \text{the i-th observation of the j-th layer} \\ &\overline{X}_j : \text{the mean of the j-th layer} \\ &\overline{\overline{X}} : \text{the mean of all layers} \end{aligned} \tag{5.2}$$

The sum of the squares (SS) can be divided in two parts: one referring to the model variability, and one due to casual error:

$$\begin{aligned} &\sum_{i=1}^{a}\sum_{j=1}^{b}(X_{ij} - \overline{\overline{X}})^2 = \sum_{j=1}^{b} a(\overline{X}_j - \overline{\overline{X}})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(X_{ij} - \overline{X}_j)^2 \\ &SS(\text{Total}) = SS(\text{Model}) + SS(\text{Error}) \end{aligned} \tag{5.3}$$

Each sum of squares has a particular number of dof assigned as summarized in table(5.1).

	dof
SS(Total)	$N-1$
SS(Model)	$b-1$
SS(Error)	$N-b$

Table 5.1: Degree of freedom for the different squared sum components

Once the dof are known, it is possible to evaluate the F ratio eq(5.4). This ratio, that has a distribution, represents which of the two variance components is governing. If the ratio is close to $F = 1$ the model error and casualties have similar importance, then it is highly possible that the means are all equal (null hypothesis). Values higher than 1 increase instead the probability that the alternative hypothesis is true.
In other words, if the error due to the model is higher than the casual error, then the model has been defined in the wrong way and the assumption of all means are equal does not halts.

$$F = \frac{\frac{SS(Model)}{b-1}}{\frac{SS(Error)}{N-b}} \tag{5.4}$$

Once the F ratio is obtained, particular tables allow to evaluate the probability of the null hypothesis, that is then called *p-value*.
In this case, the *p-value* was obtained automatically through Excel.
From the shear speed estimation, fig(4.3), it is visually possible to recognize two different regions, since there are values sensibly different. That is, however, not always possible, and is therefore necessary to perform the analysis. An ANOVA has been carried out comparing varying each time the depths of the soil layers, considering all the possible combinations.
In table(5.2) are presented only the three best cases as example.
In Excel, other than the F value, F_{crit} is also estimated: it represents the F value which would correspond to a $p - value = 5\%$. If $F > F_{crit}$, then the null hypothesis can be rejected.
From the results obtained, the null assumption (all means equal, therefore a single layer) is hardly possible, therefore the existence of 2 different layers below the first is highly probable. This does not mean, however, that the use of a model with only 2 layers would be 'wrong', however wouldn't represent adequately the empirical measurements.
The analysis has been performed considering also density and compressional velocity, with the same results.

	μ	σ	F	F_{crit}	$p-value$
Layer 2 [3-20]	288.1	33.00	71.71	4.03	$2.71 \cdot 10^{-11}$
Layer 3 [21-55]	427.5	65.47			
Layer 2 [3-21]	288.5	32.12	86.67	4.03	$1.40 \cdot 10^{-12}$
Layer 3 [22-55]	431.377	62.24			
Layer 2 [3-22]	292.6	36.29	84.19	4.03	$2.23 \cdot 10^{-12}$
Layer 3 [23-55]	433.2	62.27			

Table 5.2: ANOVA results for the three best cases, analyzing shear speed

The choice for the depth is made upon the minimum *p-value* obtained. Given the *ANOVA* results, it has been preferred to use for all the simulation in the software a model with three layers. For the particular case considered, the soil depths (starting from seabed level) are the following:

$$z_{L,3}[m] = \begin{bmatrix} 0-2 \\ 3-21 \\ 22-55 \end{bmatrix}$$

In the software the third layer has been extended up to a depth of $110\,m$ even if no recordings are available at that depth. This has been done because it is necessary to define a sufficient soil depth below the pile bottom such that the boundary conditions result as a rigid bottom.

5.2 PROBABILITY DISTRIBUTIONS DEFINITION

Once the layers have been defined, the following step involves the definition of different distributions. In this section a practical framework for determining the optimal parameters for the distributions considered (*Maximum Likelihood Estimator*) [33] is presented. These are then compared, and the one best representing the available data is chosen using *AIC* criterion.

5.2.1 *Maximum Likelihood estimators*

When data have to be described by a certain distribution type, the one giving the highest probability density function to the collected ones has to be preferred. To do so, the coefficients of the distributions are obtained through maximum likelihood estimation (MLE).
The likelihood is a function of the distribution coefficients $\theta = \theta_1 ... \theta_n$

and represents the joint probability of the n recordings, and is obtained by multiplying their pdf, eq(5.5).

$$\mathcal{L}(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) \tag{5.5}$$

The distributions considered and their MLE is discussed hereafter.

5.2.1.1 *Normal distribution*

One of the simplest distribution to construct is the normal, which probability and cumulative density functions are shown in eq(5.6).

$$\begin{cases} f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \\ F(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds \end{cases} \tag{5.6}$$

The MLE of the Normal distribution are not shown since it is a well known procedure, but the results are given. The coefficients of the distribution (mean μ^* and standard deviation σ^*) that best describe n samples are obtained using the well known eq(5.7):

$$\begin{cases} \mu^* = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i \\ \sigma^* = \sqrt{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \mu^*)^2} \end{cases} \tag{5.7}$$

5.2.1.2 *Gumbel distribution*

Another distribution investigated is the Gumbel [34].
The probability density function and the cumulative distribution function (pdf-cdf) are shown below in eq(5.8):

$$\begin{cases} f(x) = \dfrac{1}{\beta}\exp(-z - e^{-z}); \qquad z = \dfrac{x-\mu}{\beta} \\ F(x) = \exp(-e^{-z}); \qquad F^{-1}(p) = \mu - \beta \cdot \ln(-\ln(p)) \end{cases} \tag{5.8}$$

Where μ is the mode, β is a scaling parameter, p refers to the probability. The MLE obtained elevating the pdf to the n^{th} power is:

$$\begin{aligned} \mathcal{L}(\mu,\beta|x) &= \prod_{i=1}^{n} \frac{1}{\beta}\exp(-z_i - e^{-z_i}) = \left(\frac{1}{\beta}\right)^n \prod_{i=1}^{n} e^{-(\frac{x_i-\mu}{\beta})} \cdot e^{-e^{-(\frac{x-\mu}{\beta})}} \\ \mathcal{L}(\mu,\beta|x) &= \left(\frac{1}{\beta}\right)^n \cdot e^{-\sum_{i=1}^{n}\frac{x_i-\mu}{\beta}} \cdot e^{-\sum_{i=1}^{n} e^{-(\frac{x_i-\mu}{\beta})}} \end{aligned} \tag{5.9}$$

It is common practice to use the log-likelihood, so that then it is easier to take the derivatives with respect to the necessary coefficients.

$$\ln(\mathcal{L}(\mu,\beta|x)) = -n\ln(\beta) - \left(\sum_{i=1}^{n}\frac{x_i}{\beta}\right) + \frac{n\mu}{\beta} - e^{\frac{\mu}{\beta}}\cdot\left(\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}\right) \tag{5.10}$$

To maximize the fitting with the available values, the derivatives are set to zero. It should be verified that also the second derivative has to be positive, but it won't be demonstrated here. The two derivatives read as follow:

$$\begin{aligned}
&\frac{\partial\ln(\mathcal{L})}{\partial\mu} = \frac{n}{\beta} - \frac{e^{\frac{\mu}{\beta}}}{\beta}\cdot\left(\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}\right) = 0\\
&\frac{\partial\ln(\mathcal{L})}{\partial\beta} = -\frac{n}{\beta} + \sum_{i=1}^{n}\frac{x_i-\mu}{\beta^2} - e^{\frac{\mu}{\beta}}\left[\frac{\mu}{\beta^2}\cdot\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}} + \frac{1}{\beta^2}\left(\sum_{i=1}^{n}x_i\cdot e^{-\frac{x_i}{\beta}}\right)\right] = 0
\end{aligned} \tag{5.11}$$

The mode coefficient can then be estimated through:

$$\hat{\mu} = -\beta\ln\left(\frac{1}{n}\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}\right) \tag{5.12}$$

For the scaling parameter it is instead necessary the use of an iterative scheme (commonly the Newton-Raphson). From an initial guess, it corrects the value at each step; usually a tolerance is set such that if two consecutive iterations differ less than the tolerance, the process stops. The general expression is shown in eq(5.13).

$$\beta_{i+1} = \beta_i - \frac{g(\beta_i)}{g'(\beta_i)} \tag{5.13}$$

In general, $g(\beta)$ must tend to zero, and it is the case, since we can define it as the derivative with respect to the scaling parameter:

$$\frac{\partial\ln(\mathcal{L})}{\partial\beta} = g(\beta) = \beta - \bar{x} + \frac{\sum_{i=1}^{n}x_i e^{-\frac{x_i}{\beta}}}{\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}} = 0 \tag{5.14}$$

The first derivative simply follows:

$$g'(\beta) = 1 + \frac{\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}\cdot\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^2\cdot e^{\frac{x_i}{\beta}} - \sum_{i=1}^{n}x_i e^{-\frac{x_i}{\beta}}\cdot\sum_{i=1}^{n}\frac{x_i}{\beta^2}e^{-\frac{x_i}{\beta}}}{\left(\sum_{i=1}^{n}e^{-\frac{x_i}{\beta}}\right)^2} \tag{5.15}$$

An inition β_0 is assumed, and after that the iteration process starts, and the coefficients can be determined.

5.2.1.3 *Lognormal distribution*

The Lognormal distribution is usually defined when the logarithm of the values are normally distributed: if X is lognormal, $Y = \ln(X)$ is a normal distribution.
In appendix(A.2) a clarification is made upon the coefficients of this distribution. The actual parameters necessary to describe the distribution are the ones of the equivalent normal distribution.
The probability density function and cumulative are defined as:

$$\begin{cases} f(x) = \dfrac{1}{x\sigma_{LN}\sqrt{2\pi}} \cdot e^{-\frac{(\ln(x)-\mu_{LN})^2}{2\sigma_{LN}^2}}; \\ F(x) = \dfrac{1}{\sigma_{LN}\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \frac{1}{s} e^{-\frac{(\ln(s)-\mu_{LN})^2}{2\sigma_{LN}^2}} ds \end{cases} \tag{5.16}$$

The procedure for the determination of the coefficients μ^*_{LN}, σ^*_{LN} of the equivalent normal distribution is:

$$\begin{cases} \mu^*_{LN} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \ln(x_i) \\ \sigma^*_{LN} = \sqrt{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} (\ln(x_i) - \mu^*)^2} \end{cases} \tag{5.17}$$

In fig(5.1,5.2, 5.3) a visual representation of the distributions obtained with the mentioned equations is shown.

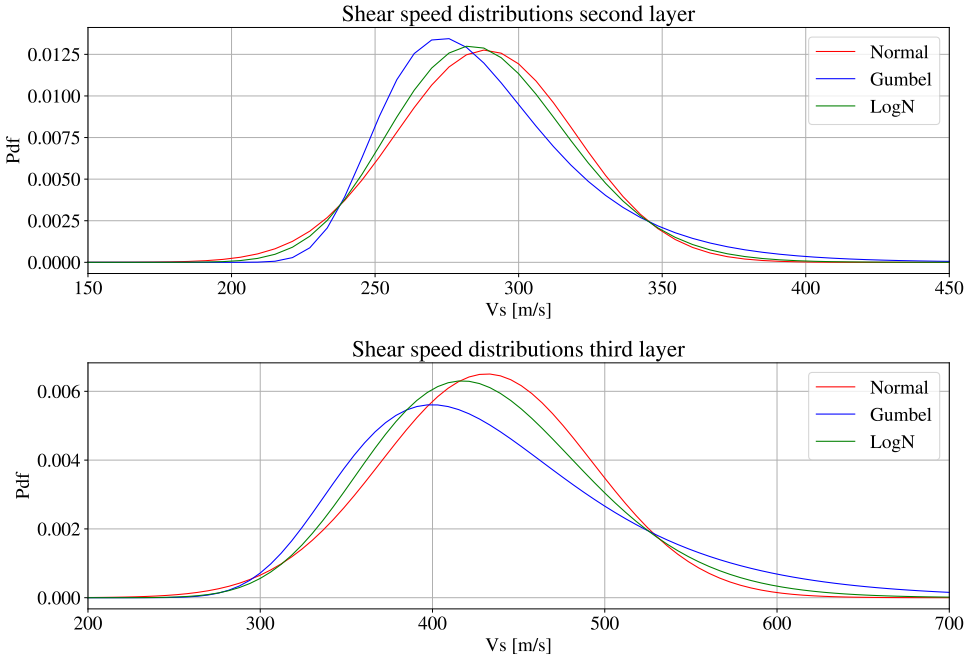

Figure 5.1: Shear velocity distribution comparison in second and third layer

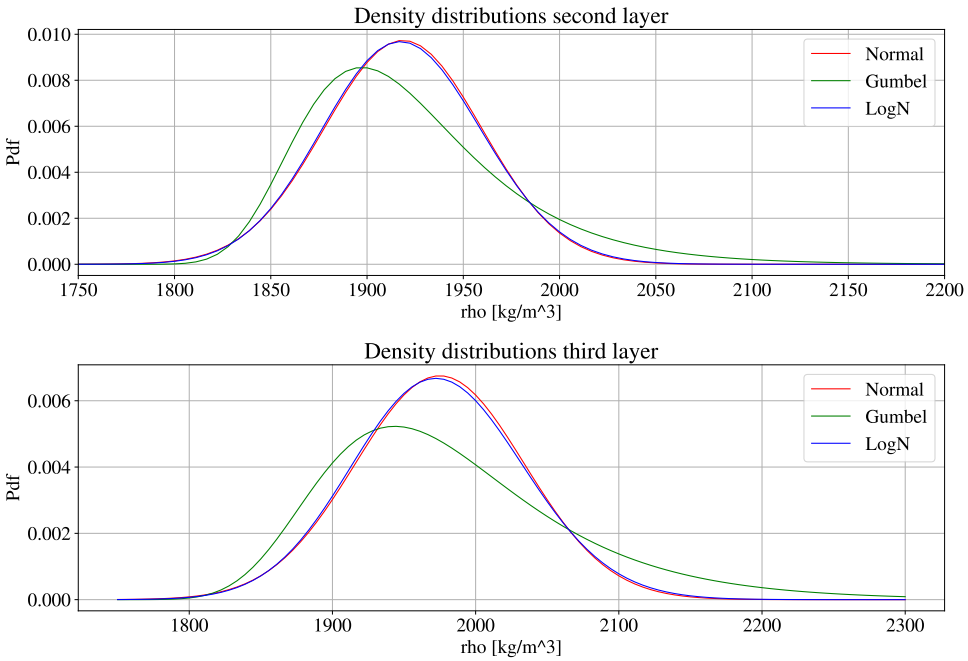

Figure 5.2: Density distribution comparison in second and third layer

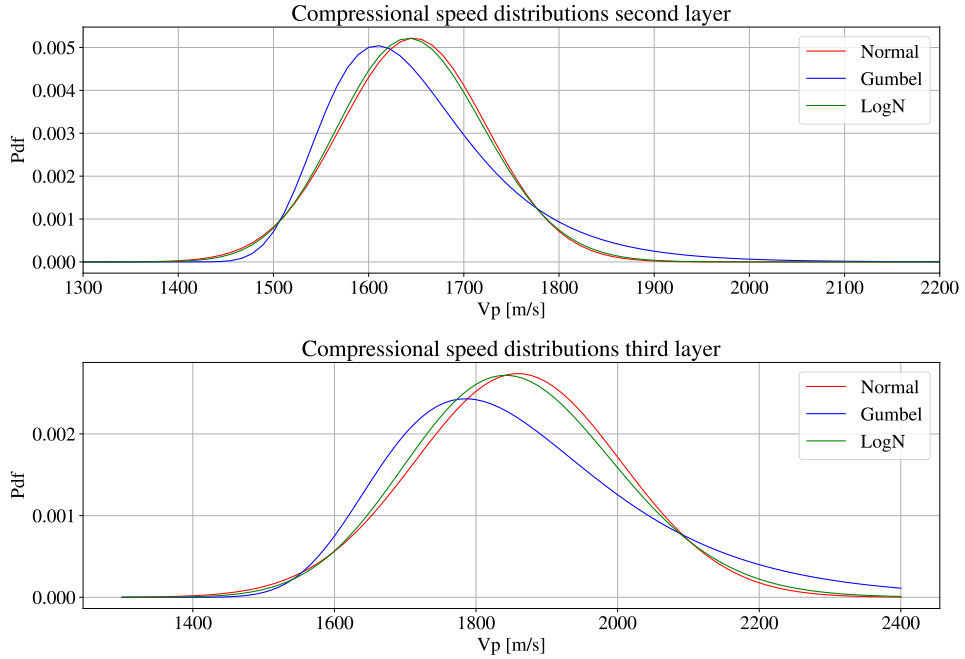

Figure 5.3: Compressional velocity distribution comparison in second and third layer

5.2.2 *Akaike Information Criterion*

Once more distribution types are available, one of the most important question to be answered is: *which of these distributions is the best for my*

particular case?
To answer this, the *Akaike Information criterion* (AIC) helps in determining the quality of the compared models. It estimates their error as information loss, and performs a trade-off between their goodness of fit (represented by the likelihood) and their simplicity (intended as the number of coefficients used by the model, e. g. mean and std, denoted by k).
The equation to evaluate the model is the following:

$$AIC = 2 \cdot k - 2 \cdot \ln(\mathcal{L}(\theta|x)) \tag{5.18}$$

Where $\mathcal{L}(\theta|x)$ is the likelihood estimator defined in eq(5.5).
Once obtained the minimum AIC_{min}, the probability that still a different i^{th} distribution could minimize the error is estimated through:

$$P[\text{i-th distr. min. error instead of } AIC_{min}] = e^{\frac{AIC_{min} - AIC_i}{2}} \tag{5.19}$$

A more clear interpretation of this probability is: *how likely is this distribution to still describe accurately the samples*.
In table(5.3) are summarized the results from the AIC performed comparing different scenarios. The recordings belonging to the first $2\,m$ depth are neglected. Although the recording obtained averaging the CPT every $0.5\,m$ have not been used for determining the distributions, the AIC obtained from a greater data sample shows the purpose of the criterion.
A comparison is made between the case of only two soil layers ($0-2; 3-55$) and the three layers model ($0-2; 3-21; 22-55$). As can be seen, the use of a total of 3 layers highly improve the description of the soil.
A few conclusions, based on the results, can be drawn:

- The normal distribution appears to be the best distribution type among the considered, except for the compressional velocity when only few recordings are considered. It is then the distribution type used to define the properties samples in all cases.
- The lognormal distribution may still well describe the data in the case of a single layer, especially referring to shear velocity.
- The Gumbel distribution is, for the particular case, the one less representing the data. However it has to be highlighted that is always necessary to verify all the distributions. It may be possible that a distribution performing poorly in a case may result the best choice for another (e. g. , when choosing 2 layers and the shear speed measurements with step $0.5\,m$, the Gumbel results as a better choice than the lognormal; but still not better than the normal).

	V_s		ρ		V_p	
	AIC	**P**	**AIC**	**P**	**AIC**	**P**
2 Layers						
53 recordings						
Normal	627	-	587	-	693	$8 \cdot 10^{-15}$
Gumbel	629	0.368	622	$2.5 \cdot 10^{-8}$	628	-
LogN	627	1	588	1	692	$1 \cdot 10^{-14}$
106 recordings						
Normal	1285	-	1302	-	1401	0.607
Gumbel	1313	$8 \cdot 10^{-7}$	1399	$8 \cdot 10^{-22}$	1408	0.018
LogN	1319	$4 \cdot 10^{-8}$	1323	$3 \cdot 10^{-5}$	1400	-
3 Layers						
53 recordings						
Normal	565	-	572	-	658	-
Gumbel	574	0.011	591	$7 \cdot 10^{-4}$	667	0.011
LogN	567	0.368	574	0.369	659	0.607
106 recordings						
Normal	1201	-	1268	-	1335	-
Gumbel	1250	$2 \cdot 10^{-11}$	1340	$2 \cdot 10^{-16}$	1368	$7 \cdot 10^{-8}$
LogN	1256	$1 \cdot 10^{-12}$	1285	$2 \cdot 10^{-4}$	1348	0.001

Table 5.3: AIC evaluation and comparison of the available distributions

It has to be noticed how, when the model assumes three layers instead of two, the AIC becomes lower. This means that the Likelihood is higher, and therefore a better representation of the actual data available. This is another reason, as mentioned before, to use for the simulations three layers.

5.2.3 *Q-Q plots*

Another visual approach to verify if the assumption of a distribution may apply to the available data is the use of quantile plots (Q-Q).
On the horizontal axis, for example, the standard normal ($\mu = 0, \sigma = 1$) observation are drawn, while on the vertical axis the sorted empirical recording are paired. If the recordings match the theoretical distribution defined, then the choice made can be taken as a sufficiently accurate representation of the data.
In fig(5.4,5.5) the QQ plots for all the properties taken in consideration are presented, for the second and third soil layer. Considering the limited amount of data, the normal distribution still seems to be the appropriate choice.
A few values seem to fall off from the theoretical estimation, especially on the extremes. These exceptions may be related to instrumental error or local and sudden change in the material, however the trend seems to be the one of a normal distribution.

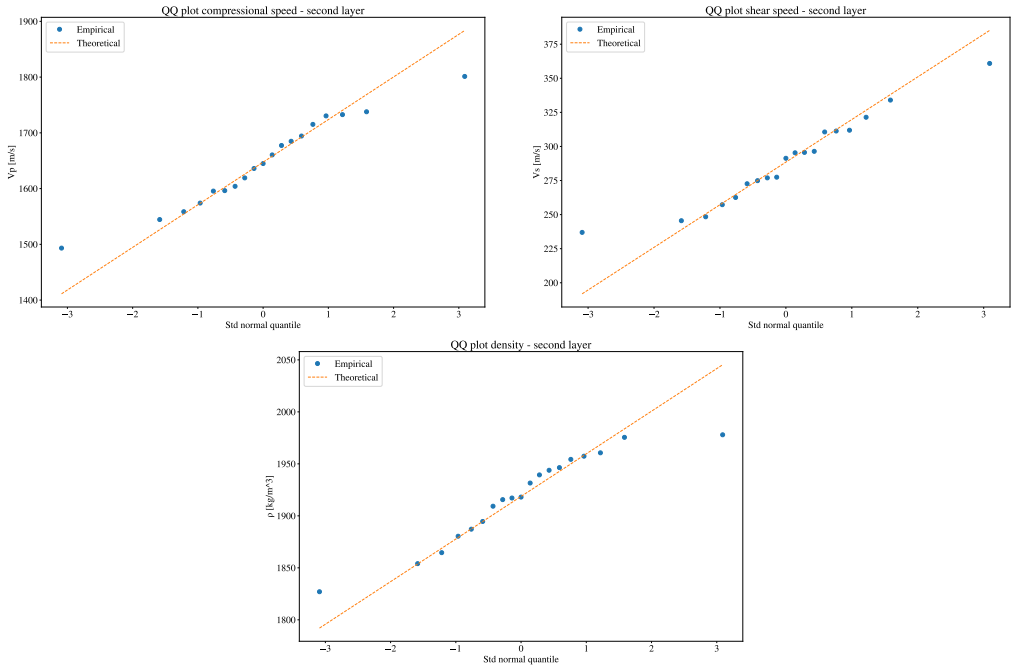

Figure 5.4: QQ plots for compressional/shear speed and density recordings against theoretical normal distribution for Layer 2

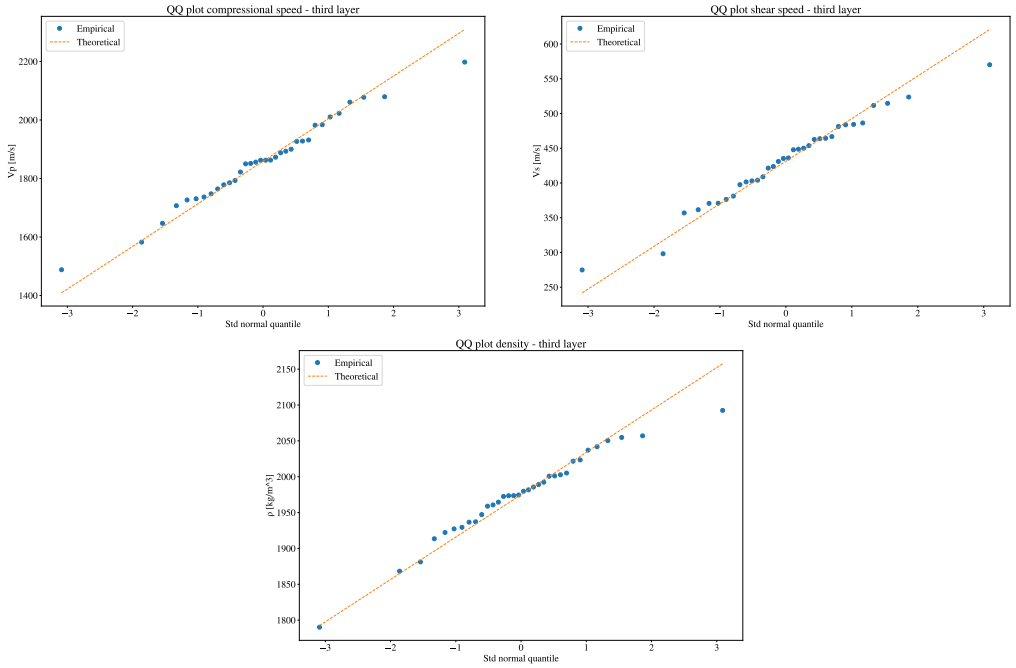

Figure 5.5: QQ plots for compressional/shear speed and density recordings against theoretical normal distribution for Layer 3

5.3 CORRELATION

Being all the properties defined since here derived from the same measurements, they should show a high level of correlation. Moreover, it is physically expected that at lower depths due to higher pressure, density and wave velocities increase.
Considering also the use of copula that will be used for the generation of the samples, it is now introduced the *Spearman's correlation coefficient* [35].
The choice of using the Spearman's rank coefficient instead of the Pearson's one is now briefly explained. The latter correlation index assesses only linear dependences, therefore measures if the real values do have a constant trend of appearence (may it be positive or negative correlation). The Spearman's instead addresses more in general monotonic correlation, therefore simply investigate whether the occurrence of high values of one variable are paired with high or low values of another.
To evaluate the rank correlation between two variables (X, Y), the individual values have to be ranked. A rank of a value in a sample of n elements is simply a number $\in (1..n)$ that expresses the position of that value in the group. One obtain then n pairs of ranks, denoted as (R_i, S_i),

where n is the size of the records for both variables.

$$r = \frac{12}{n(n+1)(n-1)} \sum_{i=1}^{n} R_i S_i - 3\frac{n+1}{n-1} \in [-1, 1] \tag{5.20}$$

In the case examined, three correlation sub-matrices have been defined, and then assembled in the global one.
This has been done since to obtain the correlation all the data samples (in this case, the empirical estimation of V_p, V_s, ρ in each layer) must have the same length. Since each layer has a different amount of recordings, it is only possible to evaluate the correlation within a single layer.
The first sub-matrix refers to the first soil layer. Given the limited amount of available recordings, it has been preferred to consider its properties (V_p, V_s, ρ) independent among each other. Another reason to assume the independence between the properties is that, as mentioned, in the first soil meters high instrumental errors may occur.
The second and third sub-matrices refer respectively to the second and third layers. In this case the rank correlation between the variables has been evaluated through eq(5.20).
The correlation matrix for the first soil layer, which accounts for independence, and with the following variable order V_p, V_s, ρ, is:

$$C_{3L,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.21}$$

The first column elements represent respectively the correlation $V_p - V_p$, $V_p - V_s$, $V_p - \rho$.
The correlation matrices for the second and third layer are respectively:

$$C_{3L,2} = \begin{bmatrix} 1 & 0.749 & 0.951 \\ 0.749 & 1 & 0.544 \\ 0.951 & 0.544 & 1 \end{bmatrix} \quad C_{3L,3} = \begin{bmatrix} 1 & 0.964 & 0.989 \\ 0.964 & 1 & 0.930 \\ 0.989 & 0.930 & 1 \end{bmatrix} \tag{5.22}$$

The complete correlation matrix, assuming that the layers are independent among each other, is obtained assembling the sub-matrices:

$$C_{V_p,V_s,\rho} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0.749 & 0.951 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0.749 & 1 & 0.544 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0.951 & 0.544 & 1 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & 0 & 0 & 1 & 0.964 & 0.989 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0.964 & 1 & 0.930 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0.989 & 0.930 & 1 \end{bmatrix} \tag{5.23}$$

In fig(5.6) a visual representation is shown, where it is also clearer each correlation element of the matrix.

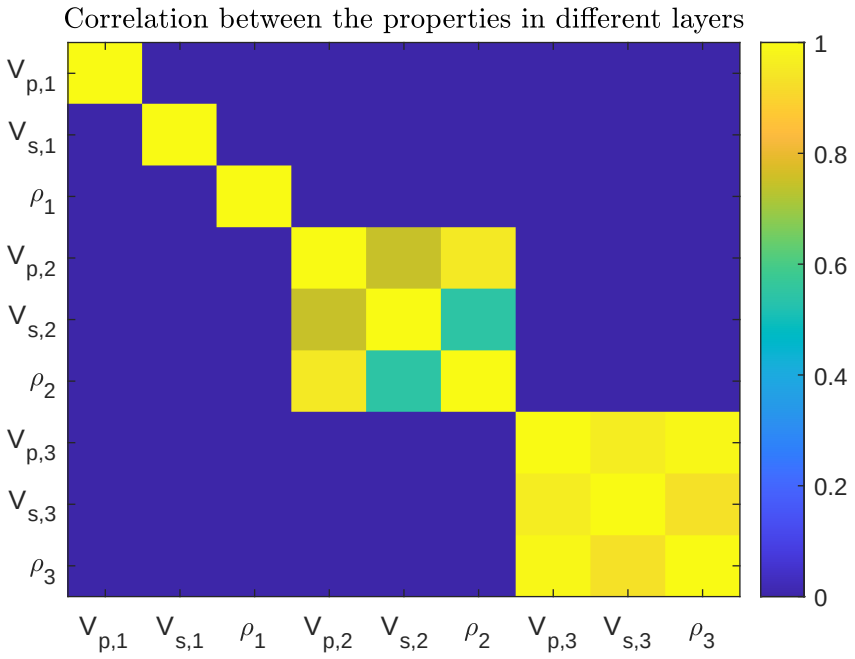

Figure 5.6: Correlation matrix for soil properties in all three layers. Each layer is independent from the others.

As expected, there is a high correlation between the properties, especially at greater depth. In the third layer the high correlation coefficient would allow to obtain the different variables in an almost deterministic way.
The assumptions of independence between the layers, given the limited

amount of data, cannot be verified. However, a few physical considerations to support them can be outlined.
As said, the first few meters in the soil represent a region with really specific properties: compressional velocity comparable with the one in the water and low shear speed. Given the limited extension of the saturated first layer, and the extremely variable water content and Poisson's ratio, these two properties may vary significantly and without any clear dependence. Regarding the density, although in general on the seabed a material with low one is present, no constraint are known on the possible values. It is not unrealistic to imagine that different environments present similar wave velocities but way different densities. The assumptions for the first soil layer can therefore be considered valid.
Regarding the independence among each layer, it has to be considered that the whole differentiation in layers establishes that they have to be different. Since the gradual variation in depth has been substituted by separate regions it is expected each one will have different values.
Moreover, estimating the correlation between the layers may be valid only in a specific location but significantly vary for another measurement, even in the same domain under consideration.

5.4 FINAL SOIL MODEL

In this section the complete distributions overview is be presented. The definition of the first soil layer is also introduced, with the assumptions behind its realization.
A visual representation of the multivariate distribution, accounting for the correlation is shown, although in the next chapter the concept of copula is introduced. It is the main tool to generate the combination samples.

5.4.1 *First soil layer*

Until now the first soil layer has not yet been defined, since the high uncertainty in the measurements for the meters immediately below seabed generated unrealistic values.
In this section the final soil parameter distributions defined is presented. The first soil layer has been chosen to have a depth of 2 meters. This because it is assumed that in this region, highly saturated, the compressional velocity has to be similar to the one in the water. Another physical feature, demonstrated by experiments [31], is the low shear velocity. In the water there are no shear waves that propagate and it is therefore assumed that, due to the saturation of the soil, the shear speed are significantly lower compared to the sediments at greater depths. Fi-

nally, due to the high water content, even the density is believed to be affected by the water content.
Each feature is defined independently based on previous researches and experimental values [27, 29, 31, 32]. This is in accordance with the independence of the properties in the first soil stratum.
In general a mean value and an upper and lower bound (corresponding to the $95 - 5^{th}$ percentiles) have been defined for each feature. These values have been taken considering, as mentioned, previous empirical results. Once enough percentiles have been defined, it is possible to estimate a distribution with a required amount of coefficients. To maintain coherence with the second and third layer, even for the first one the distribution chosen is a normal one.
In the case of a normal distribution, two percentiles are required. Denoting P_1, P_2 the probability corresponding to two set values X_1, X_2, being μ, σ the unknown coefficients of the normal distribution, and x_1, x_2 the distances of the two values from the mean, then eq(5.24) can be set:

$$\begin{cases} \Phi^{*-1}(P_1 < 0.5) = \mu - x_1 \cdot \sigma = X_1 \\ \Phi^{*-1}(P_2 > 0.5) = \mu + x_2 \cdot \sigma = X_2 \end{cases} \tag{5.24}$$

It seems that the equation has a surplus of unknowns, but the distances x_1, x_2 can be found using the properties of the standard normal distribution. Replacing $\mu^* = 0, \sigma^* = 1$, the distances can be found.

$$\begin{cases} \Phi^{*-1}(P_1 < 0.5) = -x_1 \\ \Phi^{*-1}(P_2 > 0.5) = x_2 \end{cases} \tag{5.25}$$

Which then can be replaced in eq(5.24):

$$\begin{cases} X_2 - X_1 = (x_2 + x_1)\sigma \\ \mu = X_1 + x1 \cdot \sigma \end{cases} \tag{5.26}$$

In table(5.4) are summarized the coefficients to define *normal distributions* in the three layers. While the ones for the first soil layer are simply assumptions, the ones for the second and third have been obtained following the procedure discussed in this chapter.

For clarity, in fig(5.7) are presented the shear speed distributions defined for each of the three layers. The same has been done for density and compressional velocity.

	V_p [m/s]		V_s [m/s]		ρ [kg/m^3]	
	μ	σ	μ	σ	μ	σ
Layer 1 [0-2]	1500	36.5	60	18.2	1600	91.2
Layer 2 [3-21]	1647	76.4	288.5	31.3	1919	41.0
Layer 3 [22-55]	1859	145.6	431.4	61.3	1975	59.1

Table 5.4: Estimated distribution coefficients for the soil stratifications, i. e. mean values and standard deviation

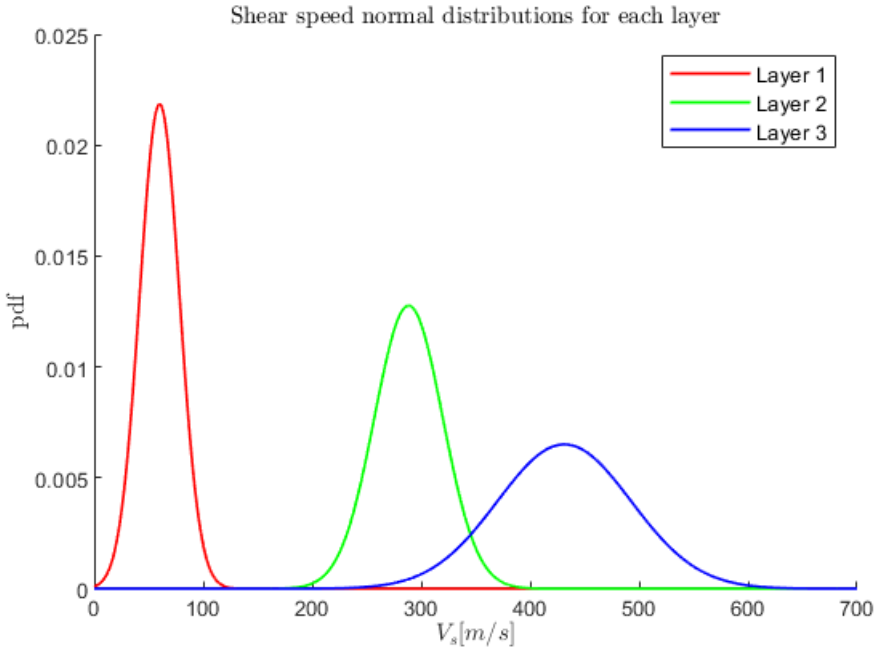

Figure 5.7: Shear speed distributions for each layer

5.4.2 *Multivariate distributions*

With the coefficients obtained, summarized in table(5.4), and this time considering the correlation found in eq(5.23), it is possible to define a multivariate distribution [36].

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r_{XY}^2}} \cdot \exp\left(-\frac{(\frac{x-\mu_X}{\sigma_X})^2 - \frac{2r_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + (\frac{y-\mu_Y}{\sigma_Y})^2}{2(1-r_{XY}^2)}\right) \tag{5.27}$$

Where $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ are respectively the mean and standard deviations of the (in this case) two variables X, Y. The Spearman's correlation between the variables is instead denoted bt r_{XY}, as in eq(5.20).

An example of a multivariate (bivariate) distribution is presented in fig(5.8,5.9): an independent case and one where correlation is present. In fig(5.10,5.11) the three complete pdf is shown.

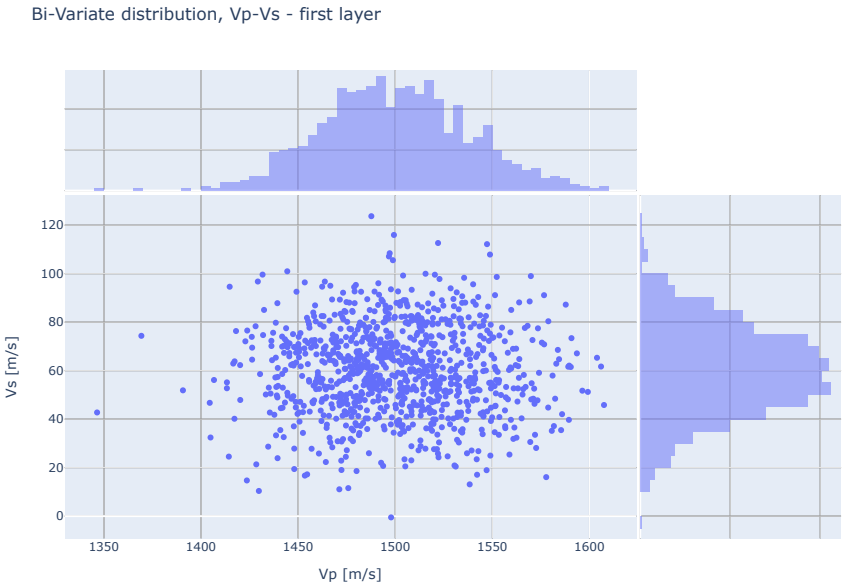

Figure 5.8: Bivariate distribution between compressional and shear speed. Being the first layer, there is no dependence between the properties (r=0).

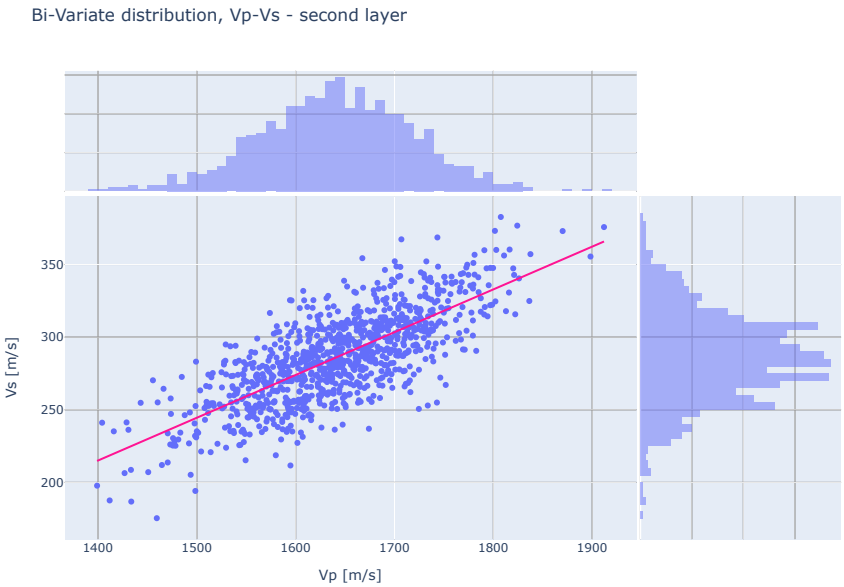

Figure 5.9: Bivariate distribution between compressional and shear speed. In the second layer there is significant dependence (r=0.749).

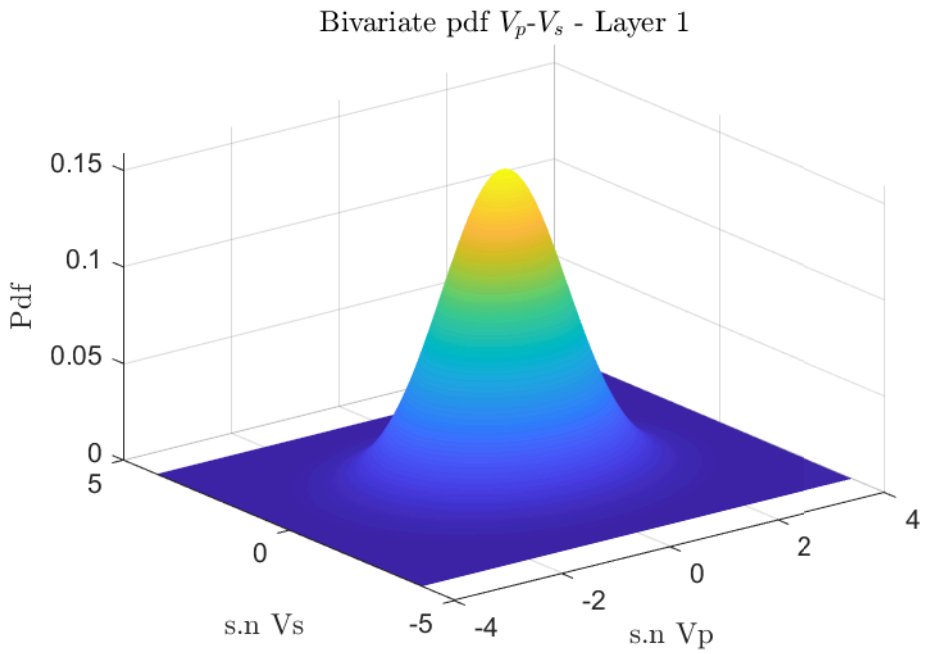

Figure 5.10: Bivariate pdf between compressional and shear speed, first layer.

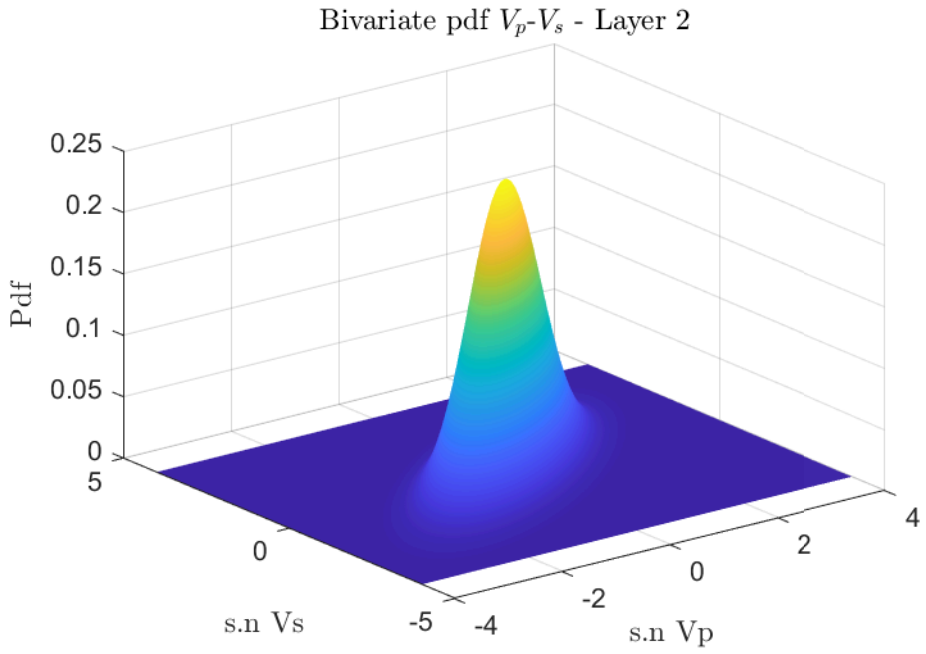

Figure 5.11: Bivariate pdf between compressional and shear speed, second layer.

6

MODEL SET-UP AND SAMPLES

In the previous chapter, the procedure to determine the distributions of the properties involved in the sound prediction model has been presented. In the last section it has been introduced the correlation between the variables. In this chapter, that information is used to generate the samples as input for the software *Silence*.
In this chapter the main settings for the simulations are presented: this includes the pile characteristics and the input blow energy.
As mentioned before, the high number of characteristics involved and the complexity of the problem makes it difficult to predict in advance which features are the dominant ones. In the next part this problem is addressed.
The compressional and shear speed, and density are obtained through a random sampling from a defined copula.
The idea behind numerous simulations is to investigate on multiple possible scenarios, and *a posteriori* try to identify the governing parameters through correlation.
The concept is not so different from a *Monte Carlo* analysis. However, in the latter case, the number of samples considered is at least of the order of $n = 10^5$, and it has also to be considered the convergence of the solution.
Due to the high computation time required for the sound propagation analysis, a limited amount of them was possible to perform ($n = 309$). Moreover, the main objective is to present an approach that can be later improved with a higher number of samples.

6.1 SETUP OF THE MODEL

In this section the main features of the pile model are shown. Usually, the pile characteristics are defined by the offshore company, therefore they do not represent a major concern in terms of uncertainty.
The input blow and the energy definition, instead, is generally highly uncertain, however the expression of a realistic hammer strike is taken from literature [23]. Since the main interest is focused on the soil properties, the input blow energy remains constant and is assumed deterministic.

6.1.1 *Pile features*

The pile dimensions and properties refer to a realistic case of a pile driven in the North Sea. All the simulations and results reported in this thesis rely on the values summarized in table(6.1). Therefore the pile can be considered as a *case-study*.
The origin of the depth coordinate is taken at the top surface of the pile.

Young's modulus	E	$2.1\,10^{11}\,N/m^2$
Poisson's ratio	ν	0.28
Density	ρ	$7850\,kg/m^3$
Shell thickness	t	$0.09\,m$
Pile diameter	D	$8\,m$
Pile length	L	$77\,m$
Water surface	z_0	$1.4\,m$
Sea bottom	z_1	$41.4\,m$
End of 1^{st} soil layer	z_2	$43.4\,m$
End of 2^{nd} soil layer	z_3	$64.4\,m$
End of 3^{rd} soil layer	z_4	$110\,m$

Table 6.1: Pile geometry and physical characteristics

The Lamé parameters can be obtained from the Bulk modulus κ (related to Young's modulus E and Poisson's ratio ν):

$$\begin{cases} \kappa = \dfrac{E}{3(1-2\nu)} \\ \kappa = \lambda + \dfrac{2}{3}\mu \end{cases} \tag{6.1}$$

Where λ is the first Lamé parameter while μ is the second.
The damping in the soil is defined by the percentage of the imaginary part, with respect to the real, for each of the Lamé coefficients. For example $Im(\lambda) = 0.01$ specifies that the imaginary part of the first Lamé parameter is 1% of the real part.
In the simulations performed, the imaginary part of the coefficients have been defined as presented in table (6.2).

	$Im(\lambda)$	$Im(\mu)$
First soil layer	0.003	0.015
Second soil layer	0.004	0.015
Third soil layer	0.004	0.018

Table 6.2: Pile geometry and physical characteristics

The definition of the Lamé damping coefficients is of difficult determination and an investigation on their estimation could be investigated by future works.

6.1.2 *Input force and blow energy*

The input blow force is defined in such a way that it represents a realistic hammer strike [23]. It is usually not possible to predict accurately the time and frequency domain of the strike. Moreover, it is even more complicated to anticipate the energy that is transferred in the pile by the strike. This is due to the fact that each energy input is dependent from the specific conformation of the soil in that location, therefore different scenarios should each time estimate the energy introduced. This topic is later discussed.
The analytical expression for the hammer blow is:

$$F(t)[MN] = \begin{cases} F_A \ \sin(F_B(t-t_0))e^{-F_C(t-t_0)}, & t_0 < t < t_1 \\ 0, & t < t_0 \ or \ t > t_1 \end{cases} \tag{6.2}$$

Where the coefficients are taken as reported in tab(6.3). In fig(6.1) eq(6.2) is plotted in the frequency and time domain.

t_0	t_1	F_A	F_B	F_C
$0.001s$	$0.05s$	503	149	150

Table 6.3: Coefficient values for the definition of the force blow input

6.1.3 *Frequency and time steps*

Two important settings that need to be defined carefully are the frequency and time steps. A step too large could lead to unrealistic results or to not describe accurately the response of the system. On the other hand, a step too small would affect the computation time making it unfeasible for running many simulations.

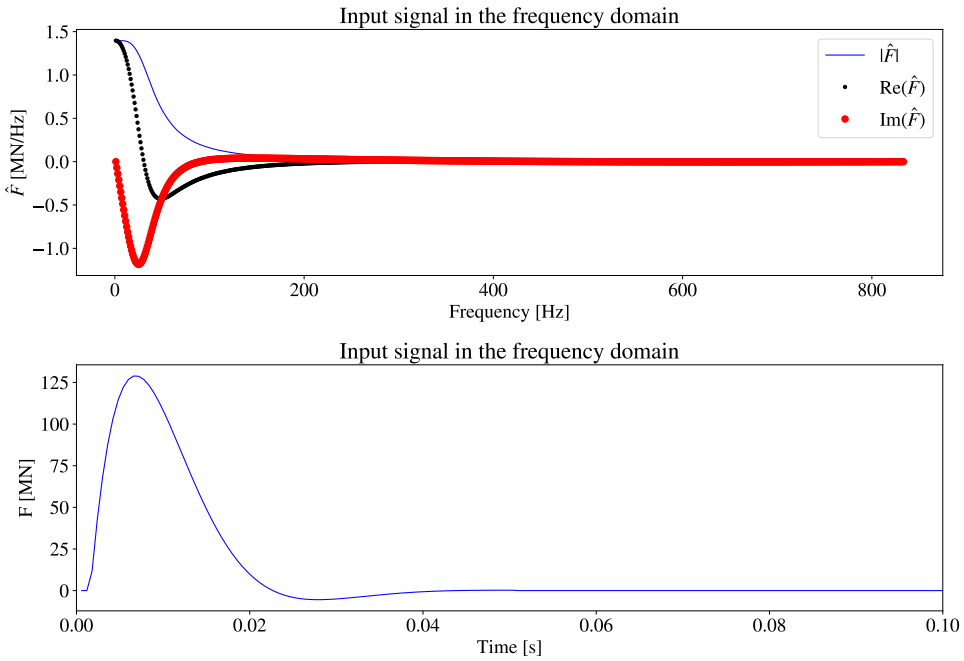

Figure 6.1: Frequency and time component of the input blow force

In the model used, it is first defined the time step, their number, and the total time that has to be considered.

$$dt = 0.0006\,s; \quad n_t = 1600; \quad T_{tot} = n_t \cdot dt = 0.96\,s$$

The sampling and Nyquist frequencies are defined as:

$$f_s = \frac{1}{dt} = 1666.67\,Hz; \quad f_N = \frac{f_s}{2} = 833.33\,Hz$$

Being the frequency range considering an interval between $f \in [-f_N, f_N]$, the number of frequency steps is taken as half the one of the time steps, therefore $n_f = 800$.
It is important to point out that the frequency spectrum needs to be sufficiently broad to consider sufficient modes of the pile and consider the frequency content of the input blow, other than a sufficient number of propagation modes. As stated in [19] there is a high frequency response for frequencies in the range $0.5 f_R \leq f \leq 0.8 f_R$, where f_R is the ring frequency of the pile, defined as:

$$f_R = \frac{1}{2\pi R}\sqrt{\frac{E}{\rho(1-\nu^2)}} \tag{6.3}$$

Where R, E, ρ, ν correspond respectively to the radius, Young's modulus, density and Poisson's ratio of the pile structure.
Inserting the pile properties from table(6.1) in eq(6.3), we obtain $f_R =$

214.37 Hz. Since $f_N \gg f_R$, and from the input blow in fig(6.2) there is no significant frequency content for $f > 200\,Hz$, it is believed that the frequency window is sufficiently broad.

6.1.4 *Mesh receivers*

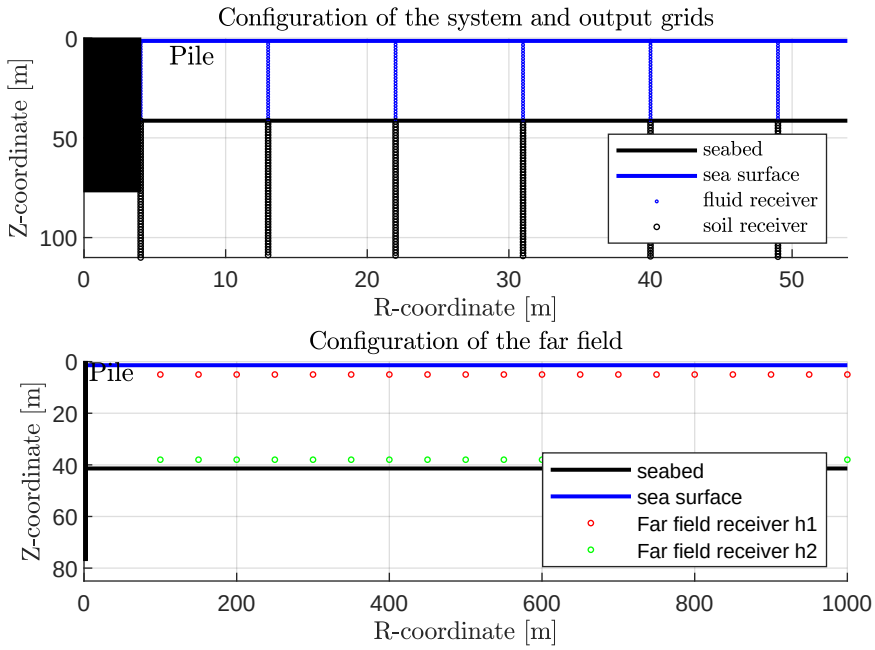

Figure 6.2: Mesh definition for the near and far field

Similarly to what happens in a FEM model, a space discretization is required to select the places where the results are needed to be evaluated. In fig(6.2) the near and far-field receivers are shown.
In the near field, due to the necessary accurate representation of the coupling between pile-water-soil, more 'station points' are needed. In the far field, instead, since the main interest is of the sound propagation, less receivers can be defined. Two depths for the receivers have been defined, one close to seabed (2 m above it) and one almost at the sea surface.
All analyses share the same receivers definition, summarized in table(6.4).

Notice that the far-field domain is defined $r \in [100 - 1000]\,m$, and the total time of the analysis considered, $T_{tot} = 0.96\,s$, allows a compressional wave in the water to travel $\Delta = T_{tot} \cdot c_p = 1440\,m$. Therefore the results obtained describe the effects of the waves travelling in the water. Regarding the shear speed in the soil instead, only the near field accounts for it in the estimation of the sound levels (being in the first soil

Near field		Far field	
Radial mesh size	9 m	Radial mesh size	50 m
Number of receivers	6	Number of receivers	19
Vertical mesh size	0.2 m	Radial distance of first receiver	100 m
		Water depth of top receivers	2 m
		Water depth of bottom receivers	38 m

Table 6.4: Mesh definition for near and far field. The water depth of the receivers in the far field is relative to sea surface

layer on average $V_s = 60\,m/s$, $\Delta = T_{tot} \cdot V_s = 57\,m$).
A more complete analysis should consider a time window sufficiently broad to consider also the arrive (and therefore effects) of the Scholte waves.

6.2 NEAR FIELD APPROXIMATIONS

The software used for the simulation (*SILENCE*) is divided in two parts: one related to the sound generation and one to the sound propagation. Moreover, the two parts respectively relate to the near field and the far field, i. e. close and far from the pile.
In order to reduce the computation time, the focus has been put exclusively on the sound propagation, i. e. the far field results. The near field results have been instead considered constant, obtained from a single analysis.
In this section the main sound generation results and reasons for approximations are presented.

6.2.1 *Energy in the pile*

When the hammer hits the top of the pile, there is an intense transfer of energy that travels through the shell structure and ultimately gets transferred in the soil. The compressional waves travelling from top to bottom then reflects and starts travelling back towards the top. While the wave travels the energy is radiated from the pile to the surrounding media in contact with it, where it dissipates.
It has been demonstrated [19] that the majority of the energy inserted in the pile transfers in the soil. Only a limited amount ($\sim 10 - 14\%$, [37]) directly radiates in the water, being the primary noise propagation path. The majority of the energy inserted in the system is therefore radiated in the soil and in the proximity of the pile. Therefore the main effects of the energy with the sound levels in the far field are related to the

energy directly radiated in the water.
Although the hammer strike always has the same force input, the energy that get inserted in the system depends on the soil configuration. This may have significant effects in the sound pressure levels.
It has been found that in general stiffer soils tend to quickly radiate more energy in the media, consequently producing higher sound levels.
Although this conclusion has been already verified, what has not yet been completely addressed is that different soil configurations introduce a different energy level.
A few sound generation analyses have been performed drastically varying the soil properties combinations. Combinations of high and low percentiles have been used, spanning from extremely soft to stiff soil conformations, with high or low wave speeds and different densities. The aim was to highlight if the soil configuration highly affects the energy entry, which later affect the sound propagation.
The results of $n = 11$ analyses are shown in fig(6.3), presenting the energy profile from the top to the pile tip. Highlighted are the lowest and highest results, and the mean value. These results span from a minimum of $E_n = 1750\,kJ$ to a maximum of $E_n = 3000\,kJ$.
Each profile, as expected, is different, meaning that the same input force (eq 6.2) introduces in the system different energy levels, which also radiates differently in the water and soil.
From the profiles is not yet clear how the energy is affected by the soil configuration and by which feature. A correlation study has been carried out analyzing whether relations exist between soil parameters and energy introduced in the system.
The energy level considered refers to the result at $z = 1.4\,m$ from the pile head, which correspond to the sea surface and therefore is assumed the energy introduced in the water column.
In fig(6.4) the correlation matrix is presented, while in appendix(A.5) the numerical values are reported.

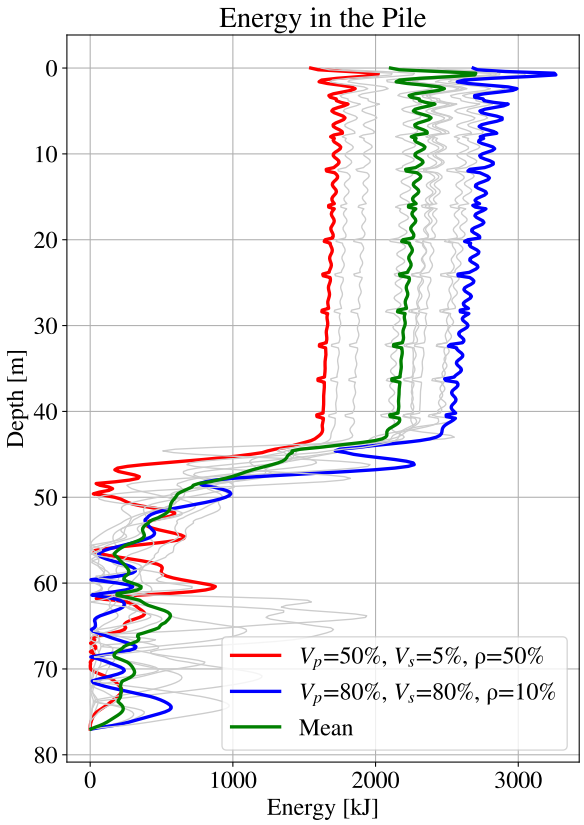

Figure 6.3: Different energy profiles in the pile based on different soil configurations. The energies refer to the final blows when the pile is almost completely embedded in the soil.

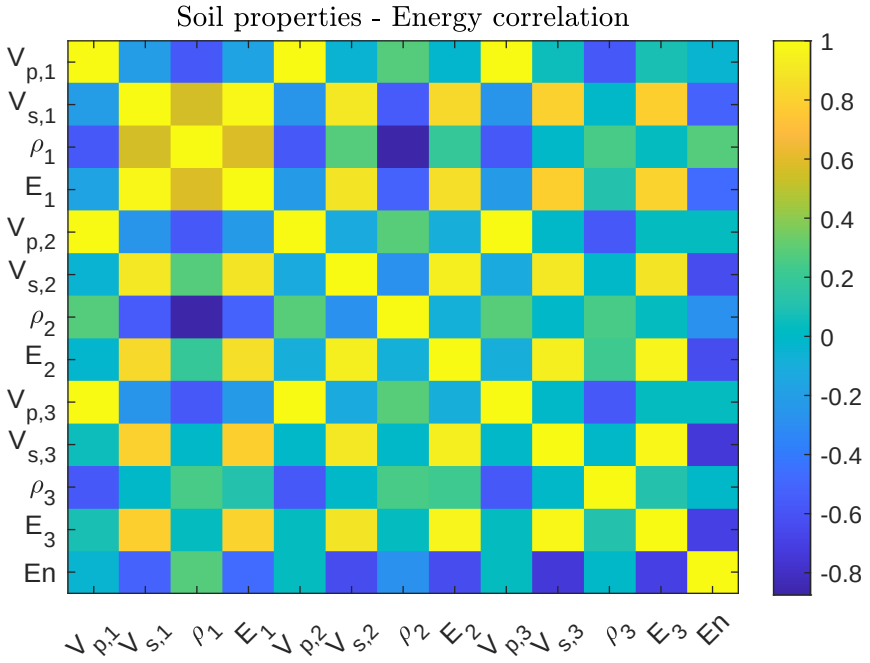

Figure 6.4: Energy correlation with soil characteristic.

It is now interesting to analyze the last column of the matrix, presenting the energy correlation with the soil features.
A clear result is the high negative correlation between Young's modulus and energy inserted in the system ($r \sim -0.8$). This means that soft soils introduce more energy compared to stiffer ones. This however can't provide conclusions on the expected sound levels. If on one hand softer soils show that a greater energy is introduced compared to stiffer ones, the latter radiate more energy in the media. Both these aspects could then affect the sound levels.
Although for each case a sound generation analysis should be required, to reduce the computational time a single sound generation analysis has been taken as reference. This particular case is obtained setting all the soil properties in the layers to the average values, summarized in table(5.4). In fig(6.5) the energy introduced in the system and the mean value of the different analyses previously performed are plotted. The high similarities allow to consider the average configuration as a good approximation to perform all the sound propagation.
The reference energy that results from a blow defined with eq(6.1) and the average conformation is estimated having a value of $E_{Amp=1} = 2500kJ$.

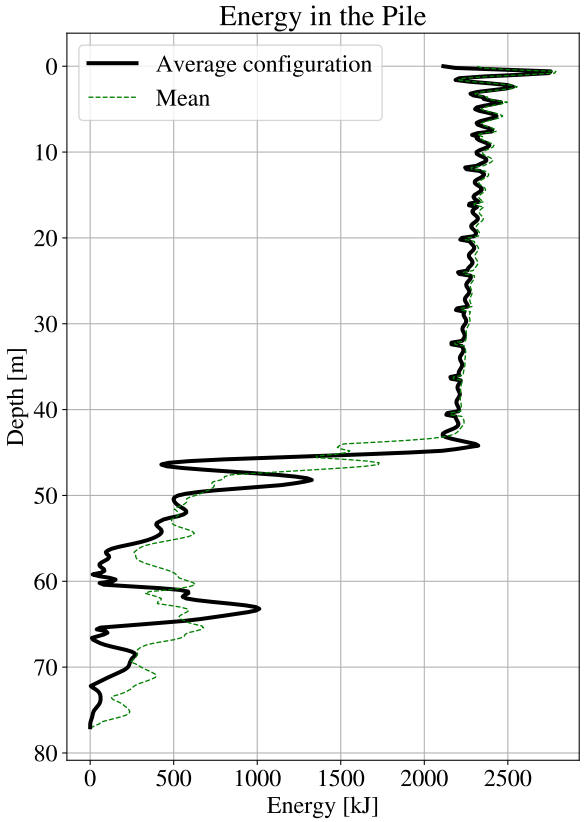

Figure 6.5: Energy in the pile for average soil conformation (black line) compared with the mean value of several simulations (green line).

It is clear that the most of the energy disperse in the soil region, as can be noticed from the abrupt change in the profile ($z = 1.4 + 40\,m$).
During the construction, the impact of the hammer can be reduced if critical sound levels are reached. A lower force introduces less energy in the system. It is now presented how the scaling of the energy could be carried out, although in all the results the amplification factor is always considered equal to $Amp = 1$.
For a particular soil configuration, to obtain the new energy input E^*, the amplification coefficient Amp^* can be determined using:

$$Amp^* = \sqrt{\frac{E^*}{E_{Amp=1}}} \tag{6.4}$$

Where $E_{Amp=1}$ is the initial energy obtained from the sound generation analysis.
Then, the sound levels obtained in the case of $Amp = 1$ can be reduced by a quantity δ, without the need of performing another analysis. The decibel quantity related to a change of amplitude is:

$$\delta[dB] = 20\log_{10}(Amp^*) \tag{6.5}$$

So that the sound level results can be modified accounting for the change in energy input by:

$$SEL^* = SEL_{Amp=1} + \delta[dB] \qquad L_p^* = L_{p,Amp=1} + \delta[dB] \tag{6.6}$$

6.2.2 *Near-field sound levels*

Another approximation deriving from the use of a single case for the near-field regards the sound levels. The results obtained from the several sound generation analyses are presented in fig(6.6,6.7). Highlighted are the results from the extreme energy (lowest and highest) entries presented in the previous section. It has to be noticed that the sound levels do not follow the energy trends: the highest energy introduced in the system doesn't produce the highest sound level as well as the opposite. The sound levels at the furthest extreme of the near field do differ for a maximum of $\sim 8\,dB$, which correspond to a relative difference of $\sim 4\%$. The average configuration results present profiles that are similar to the others, which have soil conformations extremely different.
Although a complete analysis, therefore sound generation and propagation with the appropriate soil properties should be carried out, the use of a single reference case can be considered a valid approximation.

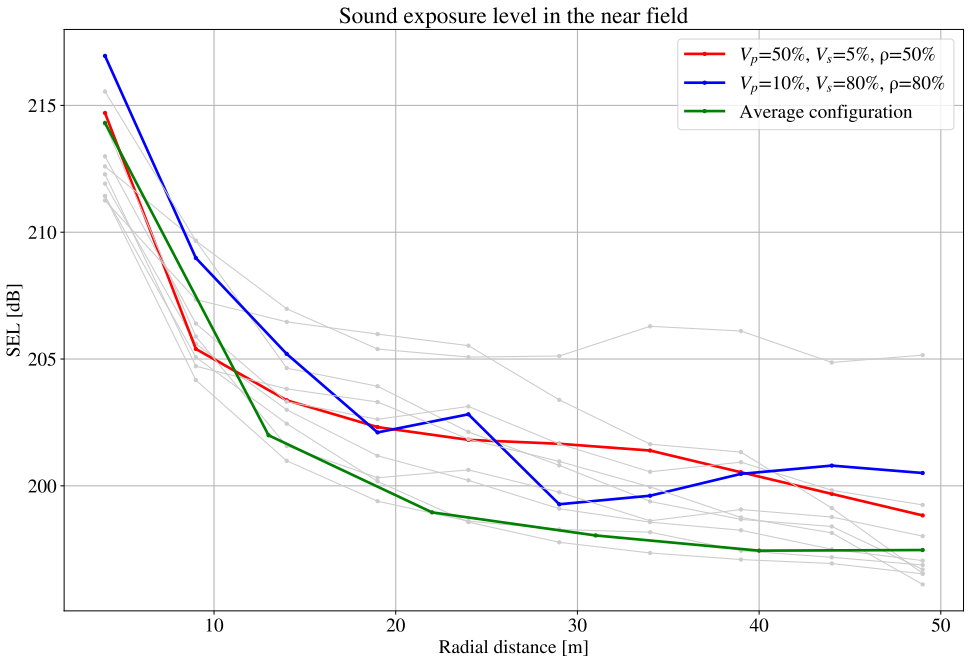

Figure 6.6: Sound exposure levels in the near-field. Highlighted are the profiles obtained with the lowest (red), highest (blue) and average (green) energy input.

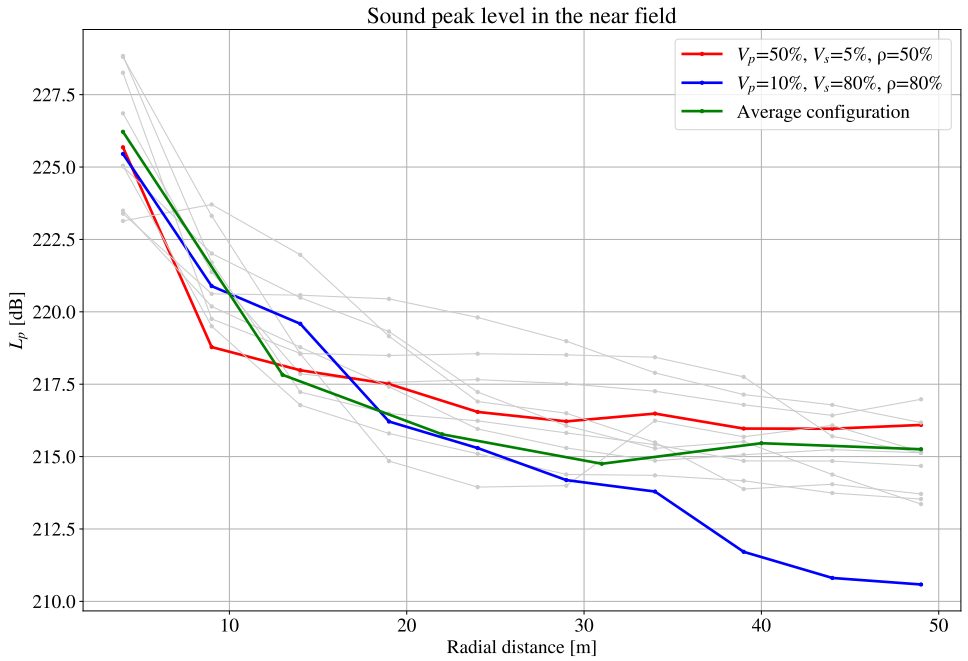

Figure 6.7: Sound peak levels in the near-field. Highlighted are the profiles obtained with the lowest (red), highest (blue) and average (green) energy input.

In fig(6.8) are shown the sound levels for the near field, derived with the average configuration and that is assumed constant for all the sound propagation analyses.

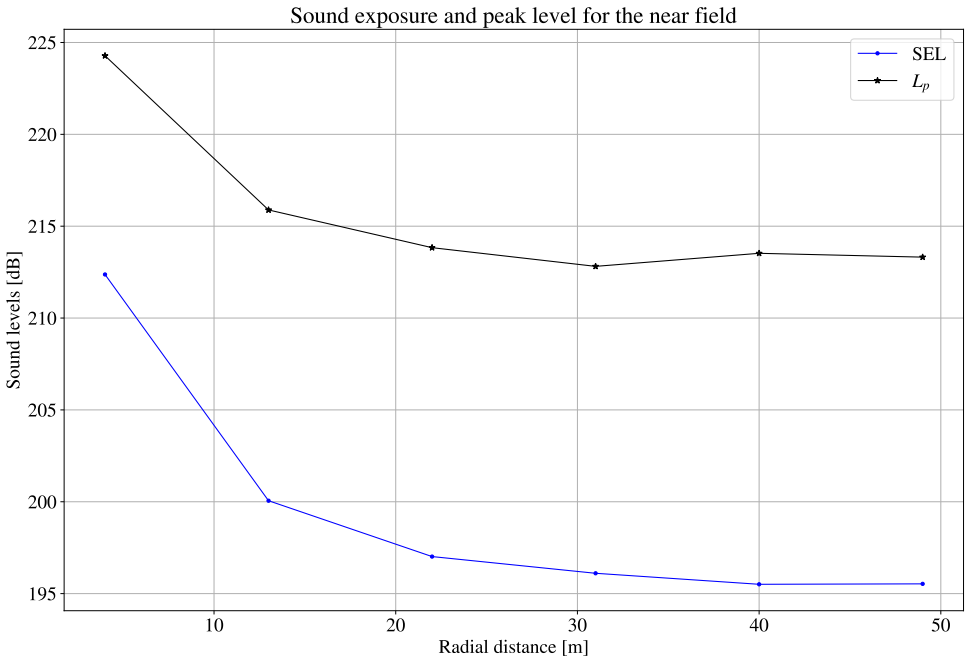

Figure 6.8: SEL and L_{peak} in the near field obtained with the average soil properties. These levels remain constant for the far-field analyses.

6.2.3 *Sampling and combinations*

As mentioned in the previous chapters the soil definition is far from certain, and through statistical approaches it was possible to define distributions to describe it (ch. 5). The main soil properties believed to influence on the sound propagation are wave compressional, shear velocity and density. These features are also the ones easier to obtain from empirical relations (ch. 4) [29].
The approach followed for the generation of the samples made use of *copulas* [36].
A copula is a multivariate cumulative function, where the marginal probability distribution of each variable is uniform in the domain $\in [0-1]$.
A *Gaussian* copula between two variables X_1, X_2, having cumulative distribution F such that $F(X_1) = u, F(X_2) = v, (u, v) \in (0, 1)$ and correlation parameter r, is defined as:

$$C_r(F(X_1), F(X_2)) = C_r(u, v) = \Phi_r(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{6.7}$$

Where Φ_r is the multivariate Gaussian cumulative, Φ^{-1} is the normal inverse cumulative distribution (for the single variable).

Copulas are preferred to multivariate distributions since they only represent the dependence between the variables, and do not need the marginals to be known. The choice of the best distribution for each variable is an independent step, but while the marginals can change, the copula, therefore the dependence, does not.
The realization of a copula is common practice in many softwares. In this thesis the copula have been defined using *Matlab*, and once given the correlation matrix of the soil characteristics, random cumulative samples have been obtained.
In the particular case of this work, all distributions have been taken as normal. Therefore, once having generated from a copula $u_1, u_2..u_n$, $u_i \in [0-1]$, where n is the number of variables, the inverse normal returns the 'real' value sample $X_1..X_n$.

$$X_i = \Phi^{-1}(u_i)$$

In fig(6.9,6.10) the pdf and cumulative copula are shown, for the case of independent and dependent variables.
In fig(6.11-6.16) are instead presented the random samples of V_p, V_s, ρ and the corresponding combinations of E, ν, ρ as input for the software *Silence*, obtained through the relations described in ch(4).
It is clear the independent pattern in the first soil layer, while for the deeper strata the correlation coefficients are the ones defined in eq(5.23).

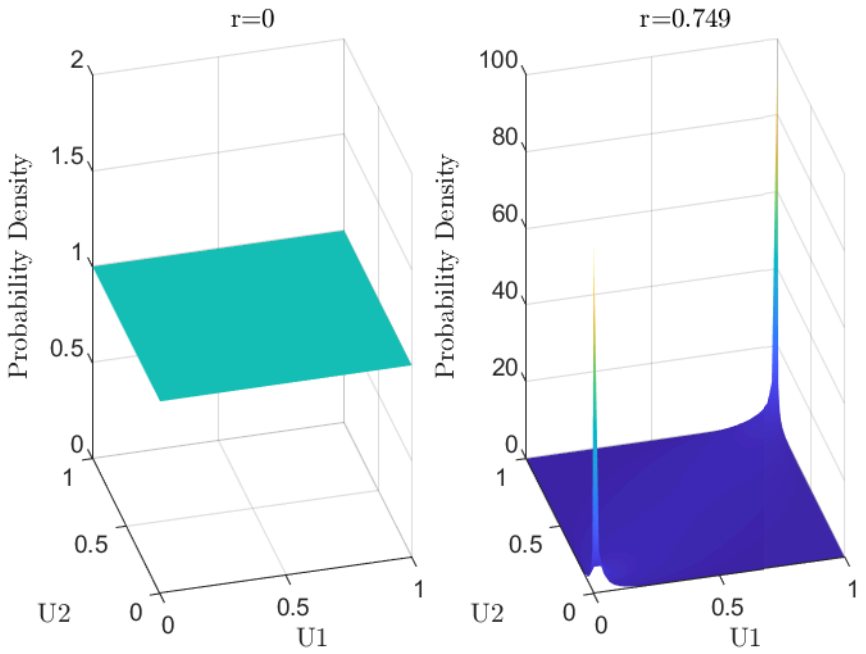

Figure 6.9: Copula probability density function between two variables, independent (r=0) and dependent (r=0.749) examples

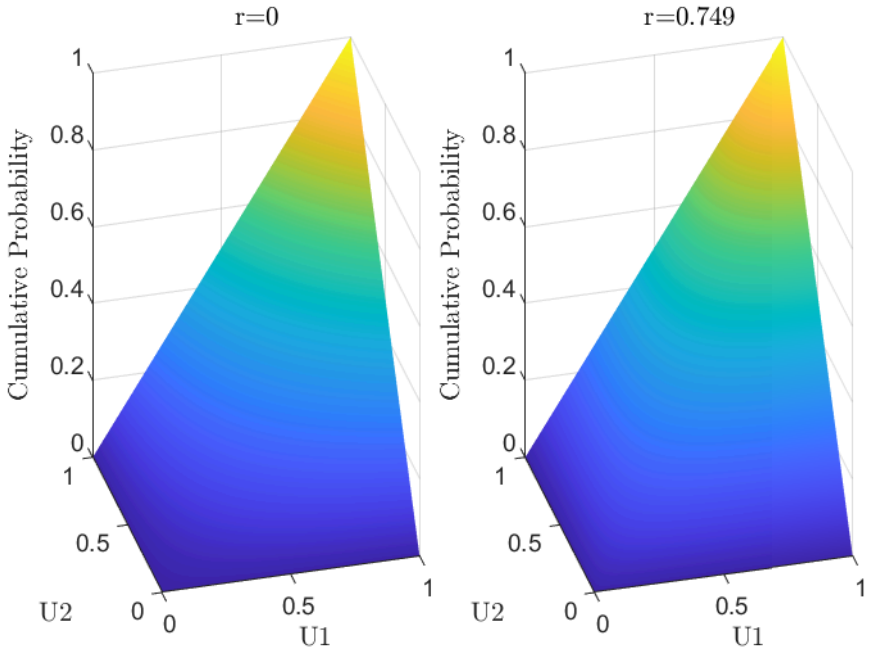

Figure 6.10: Copula cumulative function between two variables, independent (r=0) and dependent (r=0.749) examples

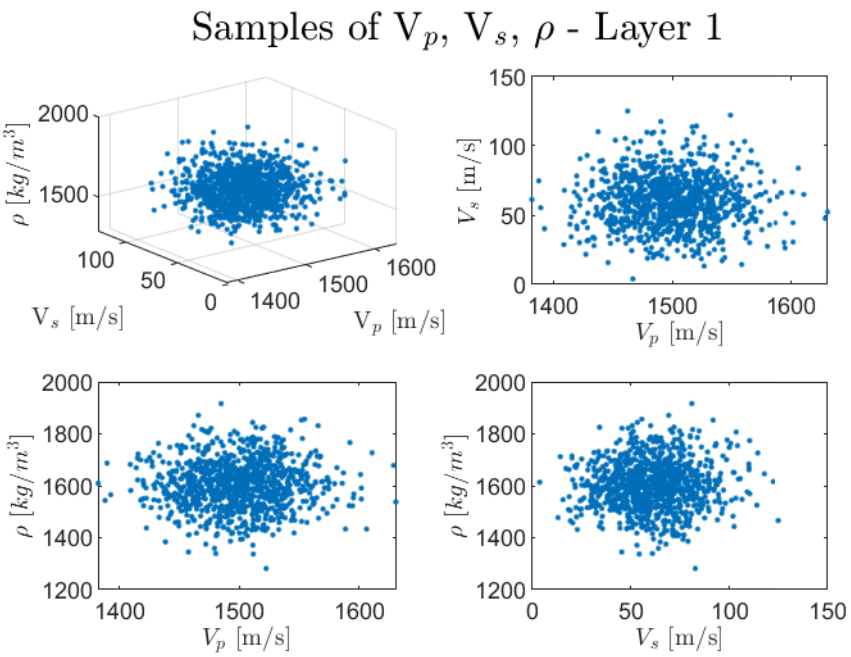

Figure 6.11: Compressional, shear velocity and density random samples 1st layer

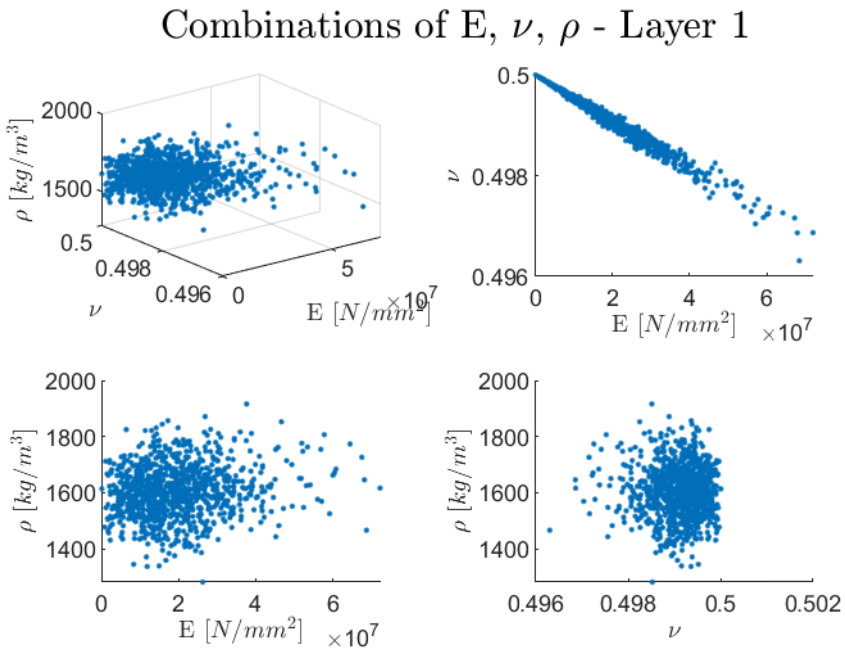

Figure 6.12: Young's modulus, Poisson's ratio and density input for the simulations, 1st layer

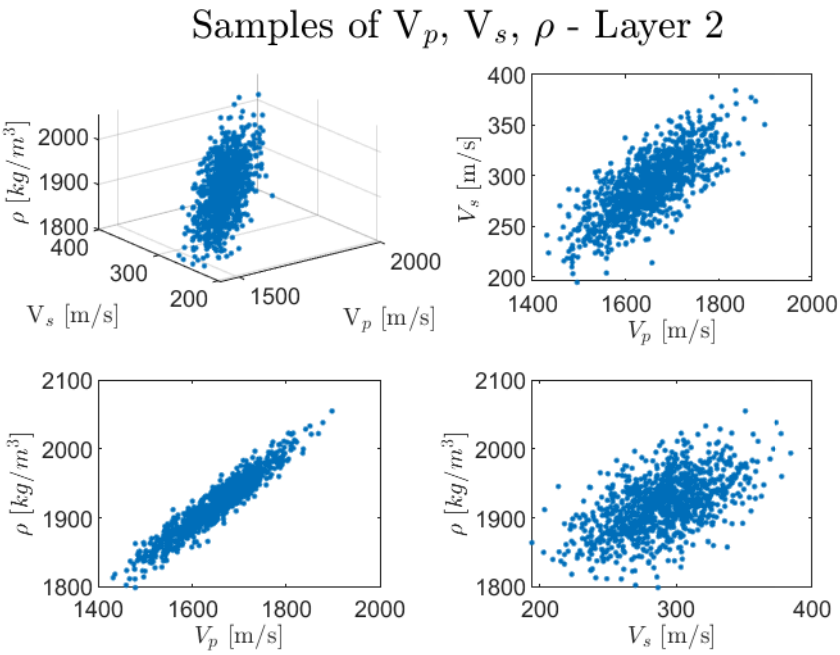

Figure 6.13: Compressional, shear velocity and density random samples 2nd layer

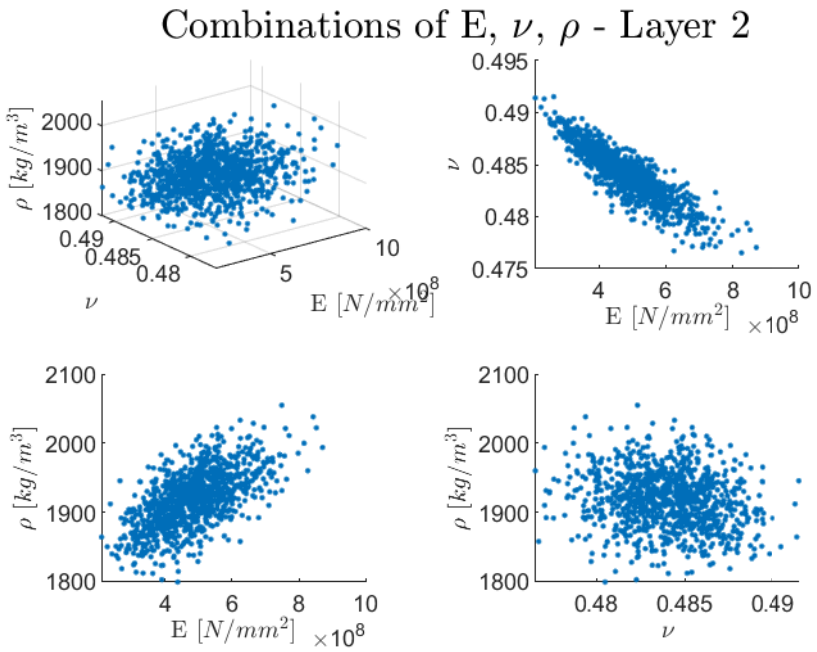

Figure 6.14: Young's modulus, Poisson's ratio and density input for the simulations, 2nd layer

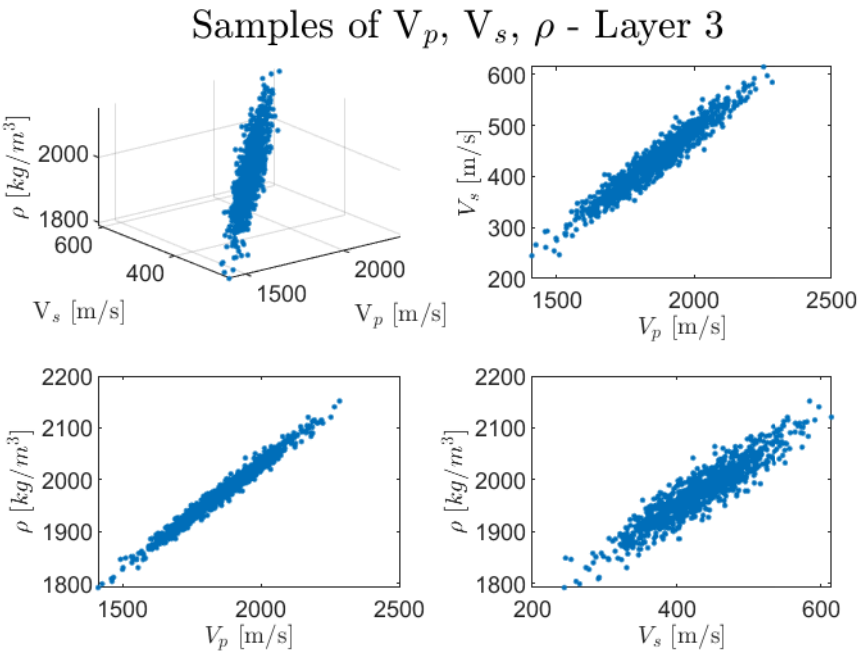

Figure 6.15: Compressional, shear velocity and density random samples 3rd layer

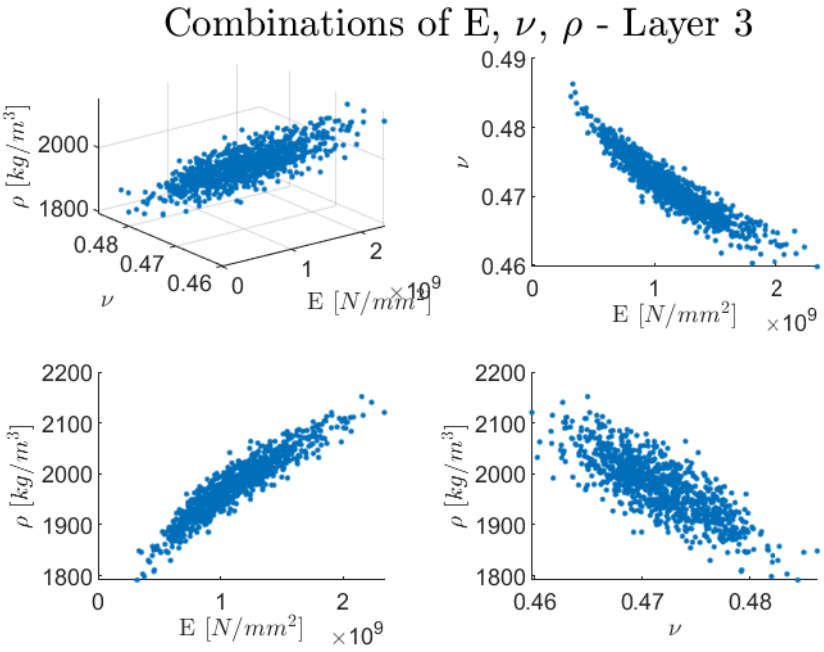

Figure 6.16: Young's modulus, Poisson's ratio and density input for the simulations, 3rd layer

6.3 FINAL REMARKS

Before presenting the results of the simulations, a few remarks need to be pointed out, considering also possible future improvements.
The first limitation encountered is the definition of a maximum of three soil layers. Although this is already assumed to provide sufficient accordance to the measurements, a higher number of soil strata could result in more accurate results.
Moreover, as mentioned, the three layers have been assumed independent between each other for simplicity. Although physical considerations do provide solid ground for the assumptions, future studies with sufficient data could finally verify them.
Another topic not completely investigated is the sensitivity of the results to the number of frequency/time steps used. A few trials have been made by using $N_f = 500, 600, 700, 800$, and no significant differences have been found. However, in order to have the most complete and accurate response, it has been preferred to use, as mentioned, $N_f = 800, N_t = 1600$. Increasing the number of frequency steps could help in improving the accuracy once reading the frequency response spectrum, avoiding to mismatch modes or critical frequencies. On the other hand, the time required for the analyses is affected.
Only one input blow force has been considered, defined analytically in eq(6.2). It is known that the frequency domain of the blow force can impact the overall response of the system. However, the main interest in this work is to highlight the influence of the soil parameters.
A single energy input has been defined as a constant for all the analyses. It is discussed later that this assumption may under/overestimate the sound levels, since the soil properties for the sound generation do not match the ones in the sound propagation. Moreover the energy introduced in the system is different. However, it has been shown how even extremely different soil configuration do produce similar sound levels.
The mesh definition adopted is believed to not be a variable affecting the sound levels. However, since a reference sound generation result has been used for all the sound propagation analyses, the near and far field should be defined the closest possible.
Having defined all the necessary input for the analysis, in the next chapter the results and their interpretation are discussed.

Part III

SIMULATION RESULTS

7

NOISE PREDICTION OUTCOMES

In this chapter, the main results of the simulations regarding the sound levels are discussed.
As done for the soil properties, the best fitting distribution is searched. From these results, the correlation between the input and output of the analyses is investigated.
The main results presented refer to sound levels at a radial distance of $r = 750\,m$ and at $+2\,m$ above seabed, therefore close to the ground.

7.1 STATISTICAL INTERPRETATION OF SOUND LEVELS

In this section the sound exposure and peak levels obtained from the sound propagation part of the simulations is presented. Of great importance are the sound levels at a distance $r = 750\,m$ from the pile and $2\,m$ above the ground (recall the German sound threshold).
Particular attention is given to the use of the sound level results, since it is not possible to directly investigate on them: it is necessary to analyze the pressure results (A.4).
As done for the soil, a few distributions are compared to estimate the one best fitting the results obtained. This process is done for all the defined radial distances of the far-field. Once found the optimal distribution it is possible to estimate a probability of exceeding a certain sound level. After that one can determine the correlation between the input soil properties and the sound level outcomes.

7.1.1 *Average and standard deviation of sound levels*

It has been demonstrated [38, 39] that defining distribution coefficients directly using the sound levels (in dB) may often underestimate their average and variation.
In order to statistically analyze the sound measurements, it is necessary to obtain back the real pressure quantities. The following steps, presented in the mentioned works, easily allow to fulfill the purpose.
We recall eq(1.1), that is the analytical expression for the SEL and L_{peak} (abbreviated now as L_p).
We define now the logarithm arguments of SEL, L_p as:

$$\frac{1}{T}\int_0^T \frac{P^2}{P_0^2}dt = \alpha; \qquad \max\left[20\log_{10}\left(\frac{|P|}{P_0}\right)\right] = \beta \tag{7.1}$$

So that eq(1.1) becomes:

$$SEL(\alpha) = 10 \cdot \log_{10}(\alpha) \qquad L_p(\beta) = 20 \cdot \log_{10}(\beta) \tag{7.2}$$

From the simulations performed, several sound levels SEL_i, $L_{p,i}$ have been obtained. Taking the inverse function, we obtain the integral (or max) of the Pressure quantities α_i, β_i as:

$$\alpha_i = 10^{SEL_i/10}; \quad \beta_i = 10^{L_{p,i}/20} \tag{7.3}$$

It is now possible to apply the maximum likelihood to determine the coefficients for the distributions of α, β (in appendix A.4 the other distributions are treated).
In the case of a Gaussian distribution the coefficients are:

$$\begin{cases} \hat{\alpha} = \dfrac{\sum_i^n \alpha_i}{n}; \\ \sigma(\alpha) = \sqrt{\dfrac{1}{n}\sum_i^n (\alpha_i - \hat{\alpha})^2} \end{cases} \qquad \begin{cases} \hat{\beta} = \dfrac{\sum_i^n \beta_i}{n}; \\ \sigma(\beta) = \sqrt{\dfrac{1}{n}\sum_i^n (\beta_i - \hat{\beta})^2} \end{cases} \tag{7.4}$$

Let's consider now a continuous function and its Taylor expansion in a specific point x_0:

$$g(x) \sim g(x_0) + g'(x)(x - x_0) \tag{7.5}$$

Assuming now x_0 being the mean, it can be proven that the standard deviation of a normal distribution is equal to:

$$\sigma(g(x)) = |g'(x_0)|\sigma(x) \tag{7.6}$$

This is useful because indeed we need to determine the standard deviation of a composed function: $\sigma(SEL(\alpha))$, $\sigma(L_p(\beta))$.
Being the first derivatives

$$\frac{d}{d\alpha}\left(10\log_{10}(\alpha)\right) = \frac{10}{\ln(10)\alpha} \qquad \frac{d}{d\beta}\left(20\log_{10}(\beta)\right) = \frac{20}{\ln(10)\beta} \tag{7.7}$$

So that finally the expected value (mean) and standard deviation of the required sound levels become:

$$\begin{cases} E[SEL(\alpha)] = 10\log_{10}(\hat{\alpha}) \\ \sigma(SEL(\alpha)) = \dfrac{10}{\ln(10)} \cdot \dfrac{\sigma(\alpha)}{\hat{\alpha}} \end{cases} \qquad \begin{cases} E[L_p(\beta)] = 20\log_{10}(\hat{\beta}) \\ \sigma(L_p(\beta)) = \dfrac{20}{\ln(10)} \cdot \dfrac{\sigma(\beta)}{\hat{\beta}} \end{cases} \tag{7.8}$$

This procedure has been followed for all the radial distances defined in the far-field (fig 6.2, table6.4). As an example in table(7.1) are summarized the coefficients of the distribution at a radial distance $r = 750\,m$, 2 m above seabed.

	SEL		L_p	
Normal	$\mu = 188.4$	$\sigma = 6.9$	$\mu = 196.7$	$\sigma = 3.03$
Gumbel	$\mu = 185.8$	$\beta = 4.39$	$\mu = 195.3$	$\beta = 2.54$
LogN	$\mu_{LN} = 5.23$	$\sigma_{LN} = 0.024$	$\mu_{LN} = 5.28$	$\sigma_{LN} = 0.014$
	$\mu = 187.8$	$\sigma = 4.54$	$\mu = 196.6$	$\sigma = 2.78$

Table 7.1: Coefficients of the considered distributions for SEL and L_p, $r = 750\,m$

Fig(7.1,7.2) show the near field results (from 6.8), all the $n = 309$ far field estimations (grey lines) and the mean value for a Normal, a Gumbel and a Lognormal distribution. All the analysis refer to an energy input of $E = 2500kJ$ ($Amp = 1$).
A main consideration has to be made on the far field results: at $r = 100\,m$ there is a broad range (even $20\,dB$) of sound levels. In sec(6.2.2) has been presented the assumption of relying on sound levels in the near field obtained from a single analysis. This has allowed to perform only the sound propagation part of the problem, reducing significantly the necessary computation time. However, considering also that each soil configuration would have a different energy entry in the system and a different sound profile in the near field, performing the complete sound generation **and** propagation may reduce the spread of results.
As shown in fig(6.3), if only the average configuration is considered, it is possible that in many cases the energy in the pile has been underestimated, while in other overestimated. For example let's consider the lowest energy entry of the graph, $E_{min} = 1750\,kJ$. Other than match the near and far-field sound levels corresponding to different soil properties, the energy entry has been exaggerated to $E = 2500\,kJ$.
Using eq(6.4,6.5), this would mean that the sound levels obtained from the simulations have been overestimated by:

$$Amp^* = \sqrt{\frac{2500}{1750}} = 1.195 \rightarrow \delta[dB] = 20\log_{10}(Amp^*) = 1.55\,dB$$

The results from the case in which the energy entry was $3000\,kJ$ have been instead underestimated by a quantity:

$$Amp^* = \sqrt{\frac{2500}{3000}} = 0.91 \rightarrow |\delta[dB]| = 20\log_{10}(Amp^*) = 0.8\,dB$$

Other than this needed remark, the sound profiles show that indeed the soil configuration plays a fundamental role in their levels. The samples generated then allowed to consider several possible scenarios, and by statistically treating the sound levels, their probability distributions are found.

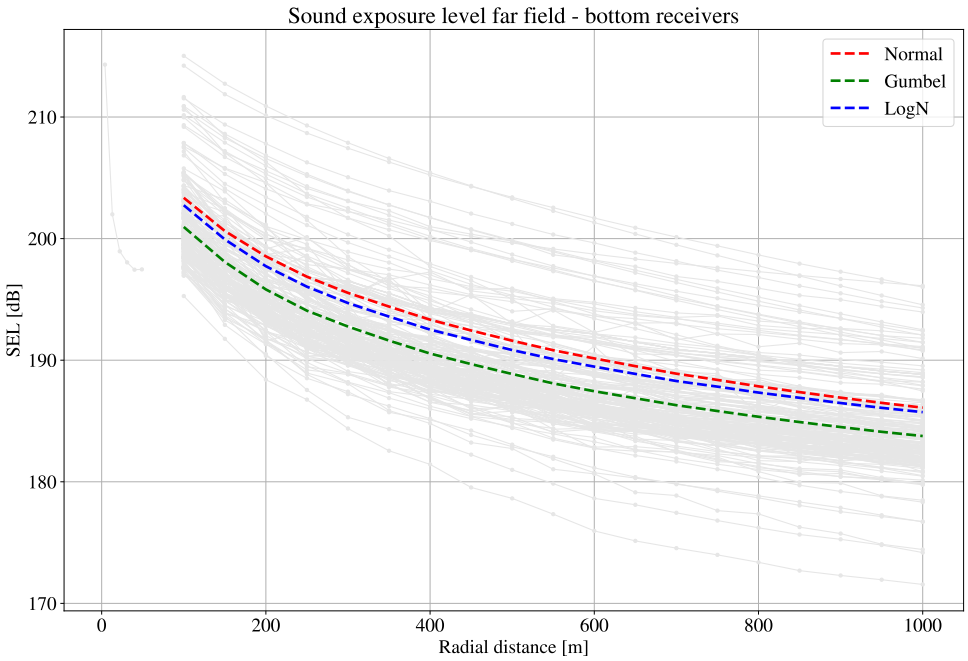

Figure 7.1: Sound exposure level outcomes and mean values for different distribution types. Energy input $E = 2500kJ$

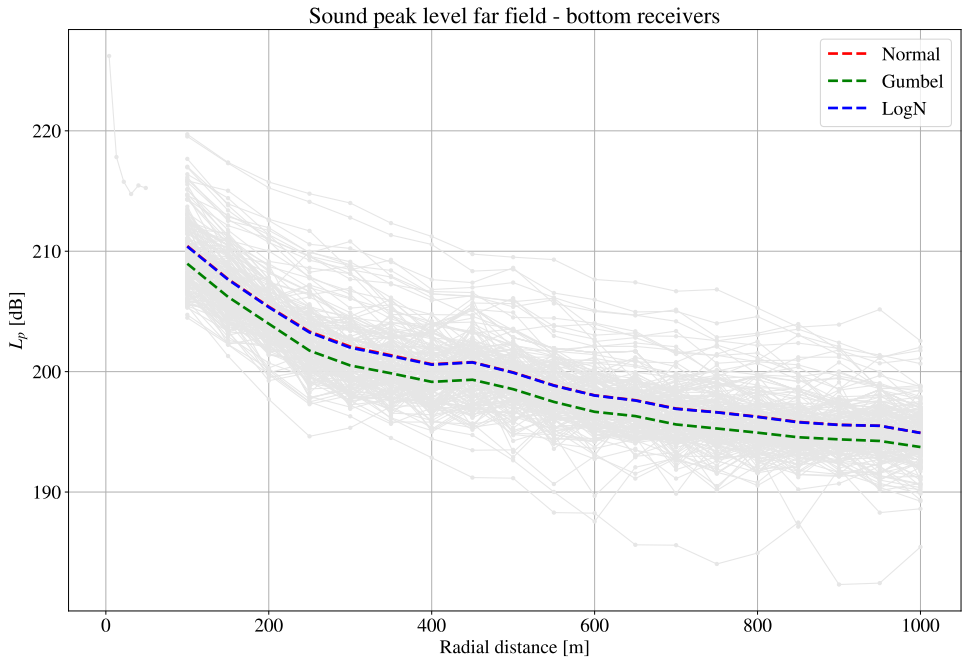

Figure 7.2: Sound peak level outcomes and mean values for different distribution types. Energy input $E = 2500kJ$

7.1.2 *Choice of the distribution*

Fig(7.1,7.2) show the near field results (from 6.8), the far field estimations and the mean values for each radial distance for three different distributions.
As done in ch(5), once the different distributions are obtained, the AIC helps in determining the optimal one.
In table(7.2) the results for the sound exposure and peak levels are presented. The quantity P refers to the probability of the related distribution to still agree with the measurements, compared with the optimal distribution.

	SEL		L_p	
	AIC	**P**	**AIC**	**P**
Normal	1265	~ 0	1012	0.14
Gumbel	1173	0.25	1143	~ 0
LogN	1170	-	1009	-

Table 7.2: AIC evaluation and comparison of the available distributions

From the analysis it can be determined that the Lognormal best describes both SEL, L_p.
In sect(7.2.1) a more detailed overview on the quality of fitting will be described.

7.2 SOUND LEVEL PROBABILITY DENSITY FUNCTIONS

One of the most important results that can now be defined is the probability density function. From the coefficients determined in the previous section (table 7.1) the continuous distributions can be drawn.
In fig(7.3-7.4) all three probability density functions considered are presented.
Moreover, the empirical results from the simulations are presented as a discrete probability function: each column represents the probability of a certain sound level interval over all the samples considered, $P(x_1 \leq X \leq x_2)$.
The use of continuous distribution over discrete, empirical ones, can bring several advantages.

- First, they allow to represent information also on results that have not been recorded. In the figures it can be clearly seen how the distributions give a probability density also to values below and above the measurements.

- Once the deterministic coefficients have been obtained from the measured samples the continuous distributions can be defined. The coefficients are less sensitive to the variation of the samples compared to the discrete distributions.

- Continuous distributions easily allow to define the cumulative functions, which then enables to estimate cumulative or exceedance probabilities. This calculation is performed using analytical expressions, and generally only requires a few coefficients of the distribution. In the discrete case it is often necessary to sum each interval probability, and as mentioned there is no real information on values outside the domain of measurements.

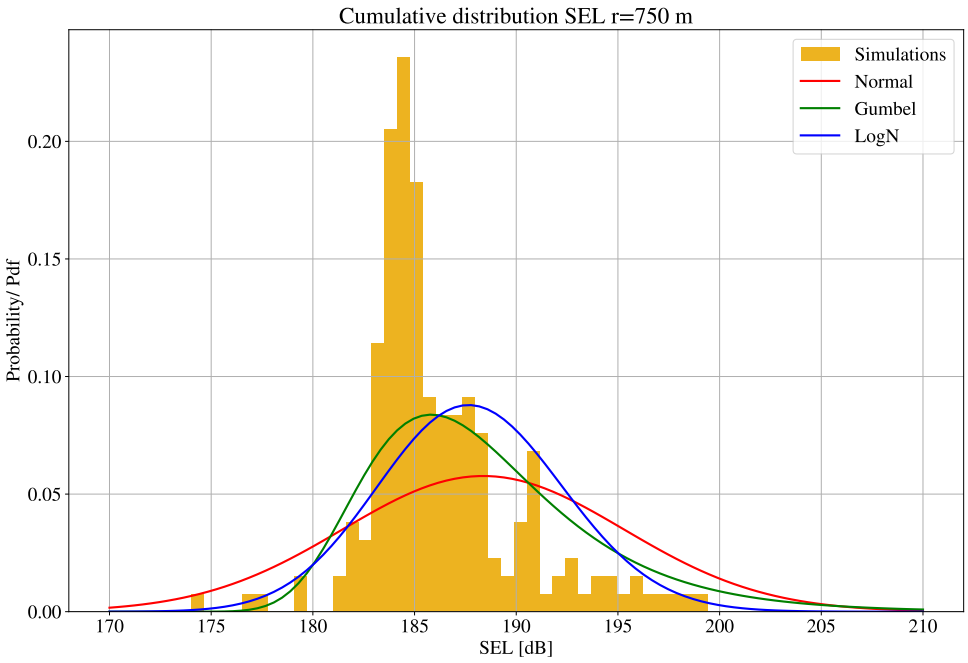

Figure 7.3: Simulation probability and pdfs of SEL

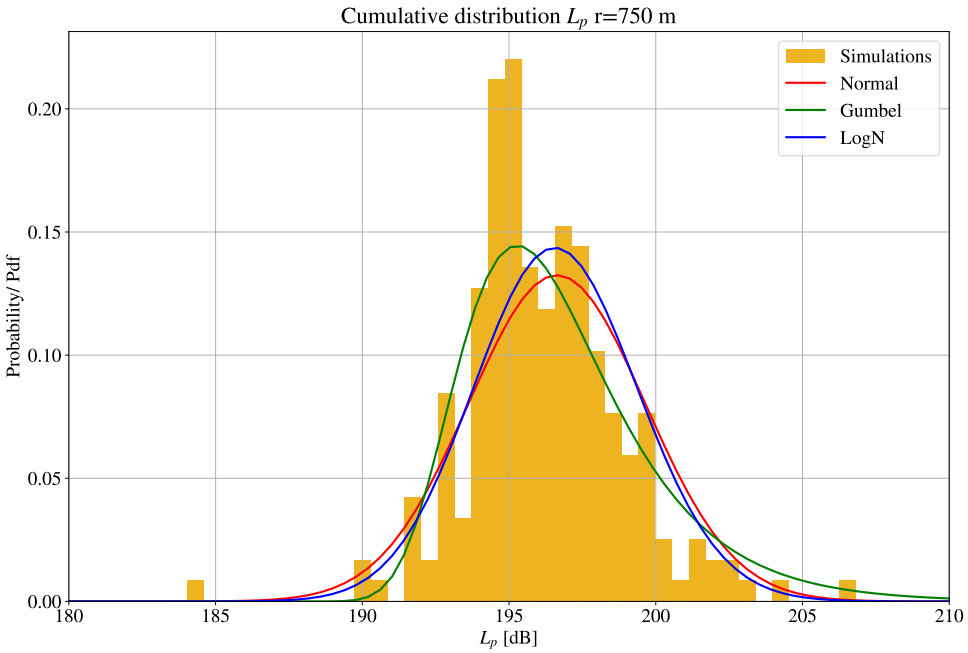

Figure 7.4: Simulation probability and pdfs of L_p

In fig(7.5,7.6) the cumulative distributions, and the empirical one, are presented. The results from the AIC evaluations are now clearer, since the Lognormal distributions show their matching with the empirical data.

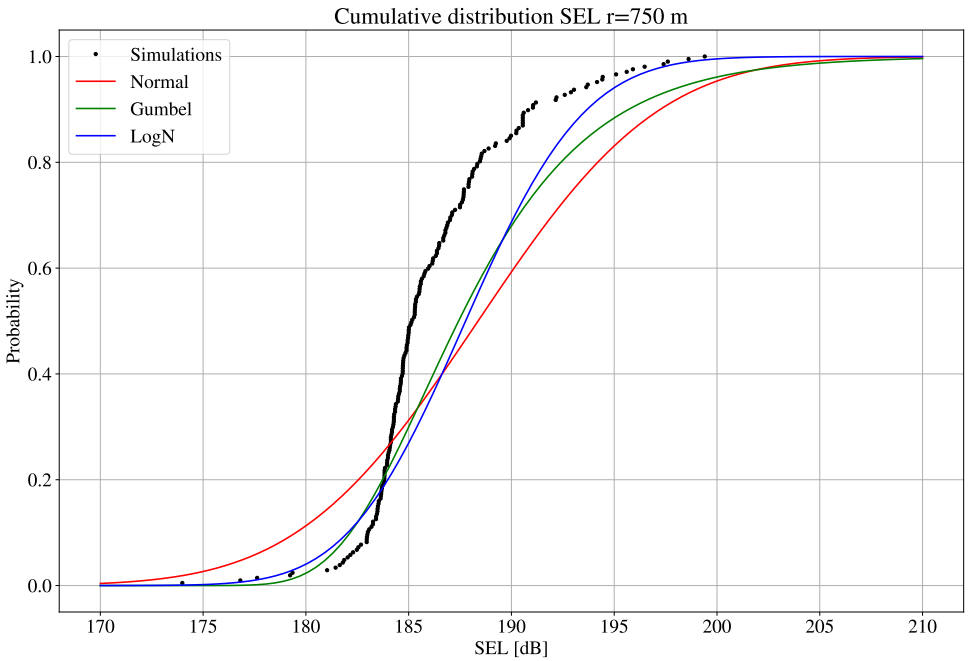

Figure 7.5: Simulation and estimated cumulative distributions of SEL

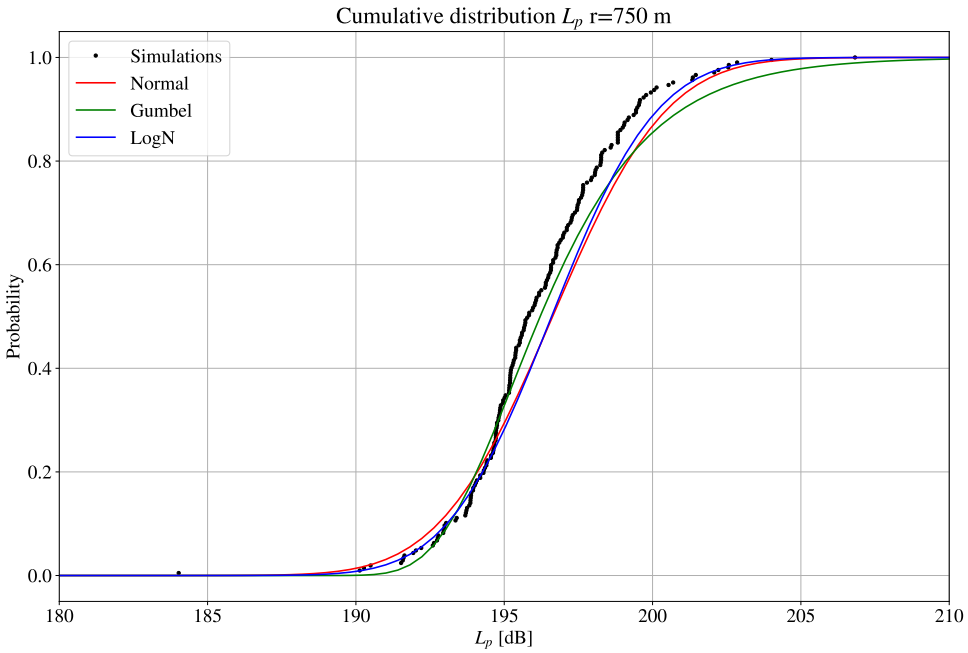

Figure 7.6: Simulation and estimated cumulative distributions of L_p

Recalling the remark previously made on the use of a single energy entry for all the analysis, and that by performing the complete analysis the sound levels are expected to produce different profiles.

7.2.1 *Goodness of fit*

From the AIC analysis it resulted that the Lognormal distribution is the one best fiting the simulation sound levels. However from fig(7.3-7.6) it is difficult to clearly see whether the choice indeed well represents the data. Especially for the SEL it seems a different distribution could be used although in this work the main concepts are presented.
A QQ-plot, however, can help in judging if the choice of the distribution really represents the obtained data. In fig(7.7-7.12) the QQ plots are presented for the SEL and L_p, for the three distributions considered. The Lognormal QQ-plots are indeed the ones that show how the choice made was the most appropriate.
While the L_p estimations show a good agreement with the Lognormal distribution, in the case of the SEL ones there are still some inconsistencies regarding the middle values. A different distribution type could, as said, improve the representation of the results.

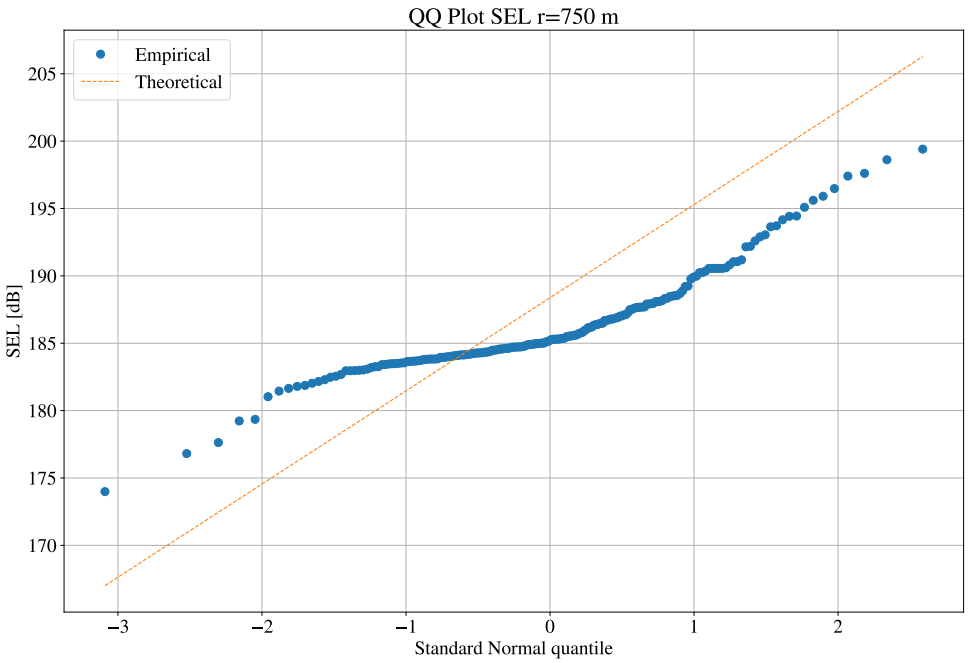

Figure 7.7: QQ-plots for SEL. The horizontal axis is the theoretical standard normal distribution $\mu = 0, \sigma = 1$.

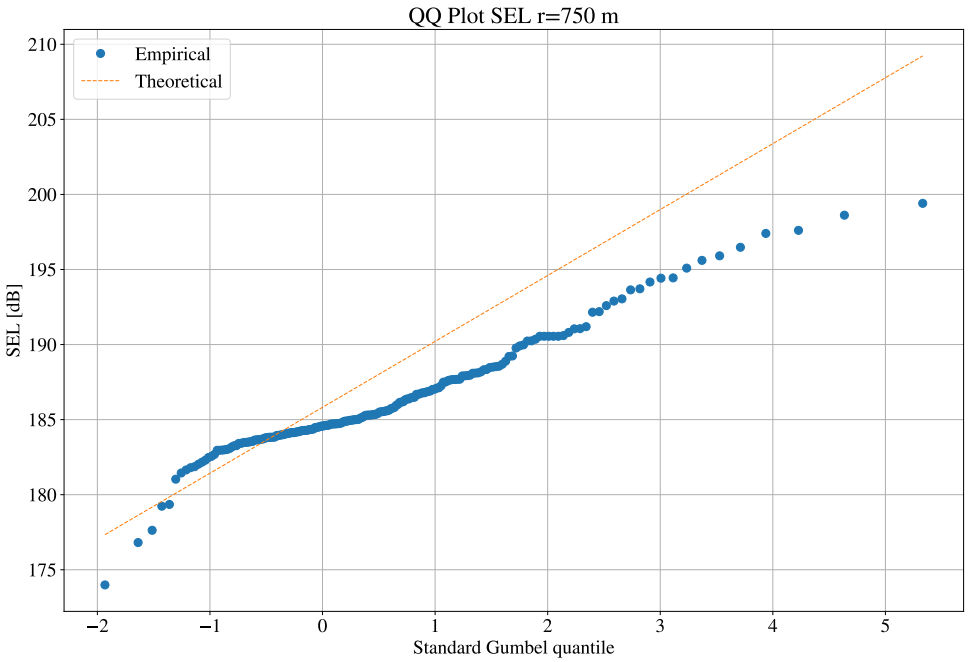

Figure 7.8: QQ-plots for SEL. The horizontal axis is the theoretical standard Gumbel distribution $\mu = 0, \beta = 1$.

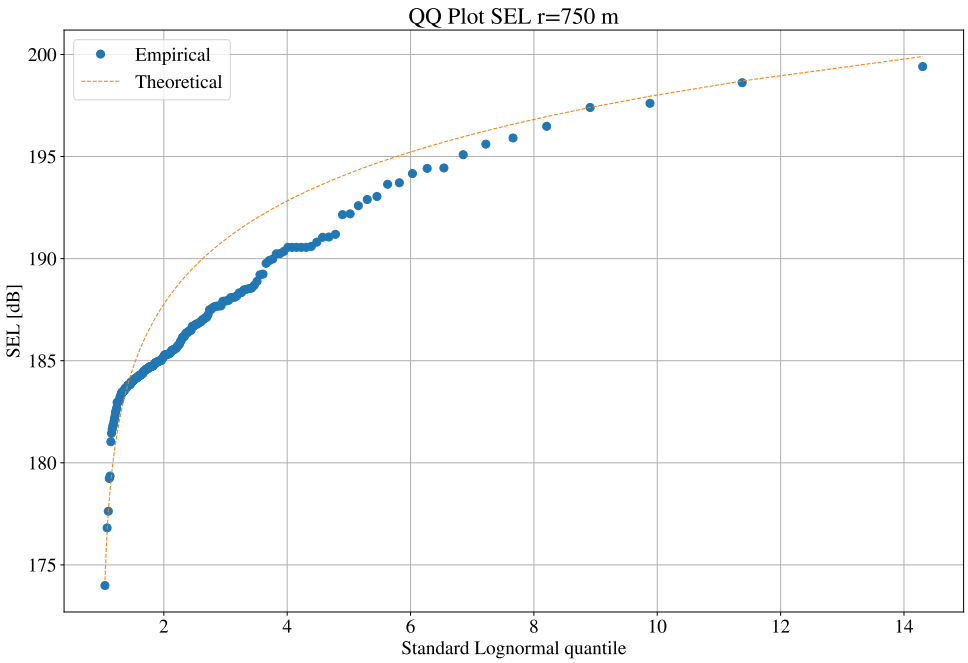

Figure 7.9: QQ-plots for SEL. The horizontal axis is the theoretical standard lognormal distribution $\mu_{LN} = 0, \sigma_{LN} = 1$.

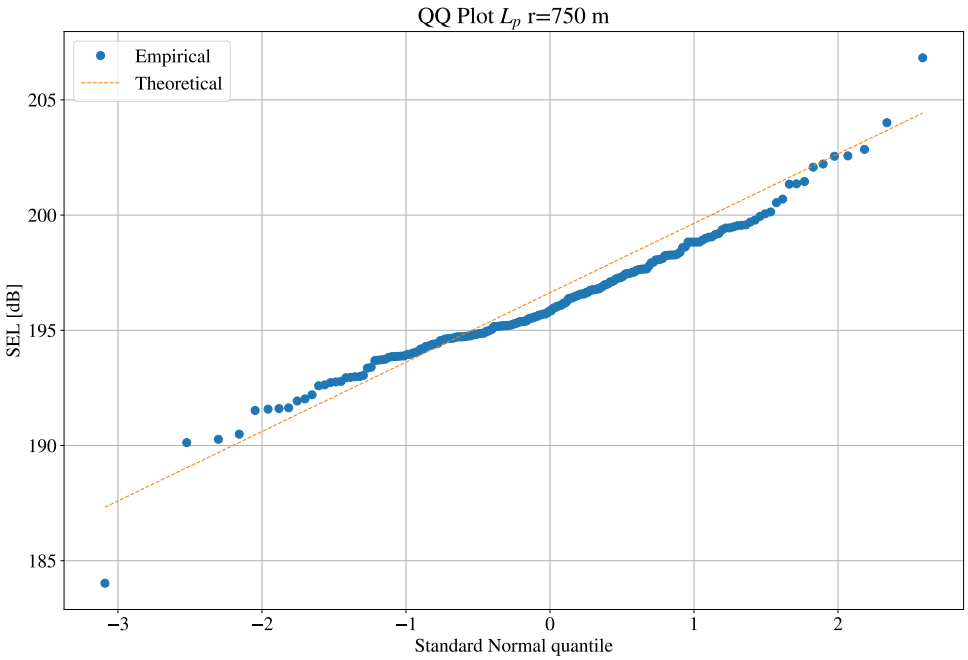

Figure 7.10: QQ-plots for L_p. The horizontal axis is the theoretical standard normal distribution $\mu = 0, \sigma = 1$.

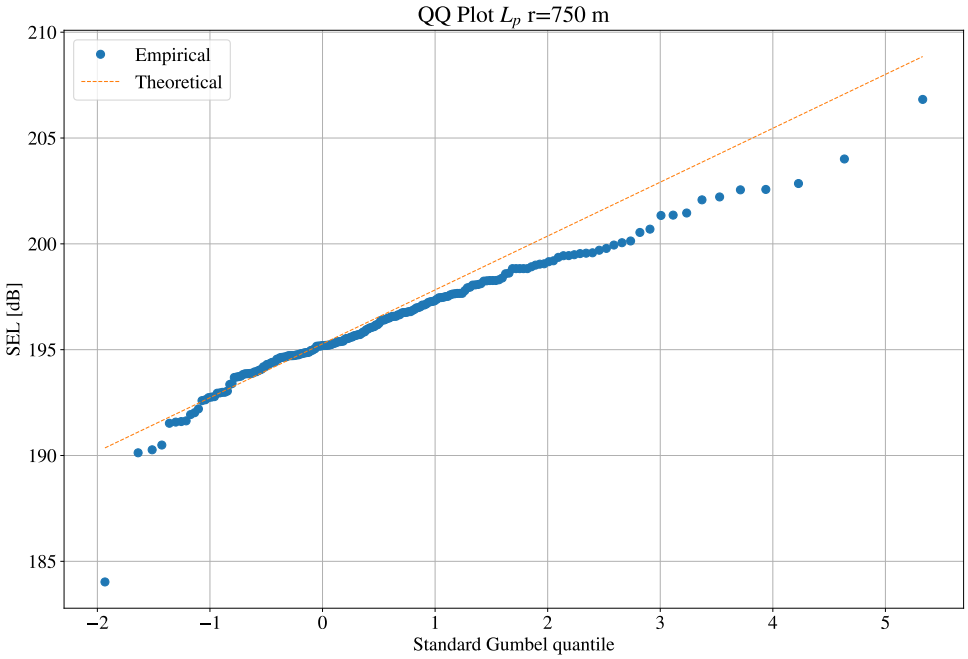

Figure 7.11: QQ-plots for L_p. The horizontal axis is the theoretical standard Gumbel distribution $\mu = 0, \beta = 1$.

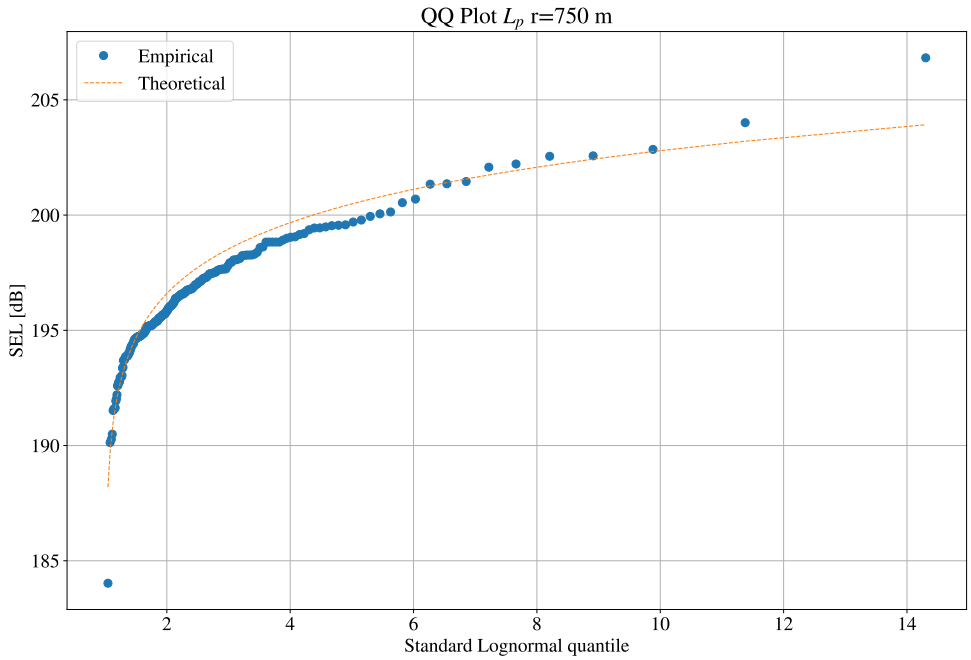

Figure 7.12: QQ-plots for L_p. The horizontal axis is the theoretical standard lognormal distribution $\mu_{LN} = 0, \sigma_{LN} = 1$.

7.2.2 *Use of the distribution for determining probabilities*

From the results obtained, it is now possible to estimate either the cumulative or the exceedance probability of a certain value.
In a realistic case a sound level could be set as threshold, and the probability of exceeding it could be a good indicator of which and how many measures should be taken. For example, the introduction of noise barriers or the reduction in the energy force.
To demonstrate the application that could be made of the sound distributions, an example will now be presented.
Four cases will be considered, namely:

- No sound barriers applied, energy input $E = 2500$ kJ.
- Introduction of sound barrier (-20 dB), energy input $E = 2500$ kJ.
- Sound barrier, reduction of energy input $E^* = 2000$ kJ.
- Sound barrier, energy input $E^{**} = 1600$ kJ.

The sound reduction due to the noise insulation system has been taken considering a well performing barrier. At the moment there exist different sound barrier types and in general the average decibel reduction, verified by experiments, is about $\delta = -15$ dB [19]. New techniques are in development, therefore it has been assumed that in the nearest future

an average reduction of about $\delta = -20$ dB will be easy to achieve.
In table(7.3,7.4) are summarized some relevant probabilities that could be estimated with the defined distribution (Lognormal). The estimations obtained at $r = 750$ m and 2 m above seabed are considered. Other than the sound barrier reduction, the decibel decrease due to different energy input is considered (eq 6.5).
For the sound exposure and peak levels it has been evaluated the probability of exceeding the German regulations that need to be met during pile driving, however different threshold can be used.
The results presented are valid for the specific environment related to the CPT results. From those, it was possible to define the layer's depth, property distributions and combinations, and obtain the simulation results which have their own distribution as well.
Even though, the framework proposed is intended to adjust to different environments, it can be modified in terms of distribution types considered and improved with the use of more empirical recordings.

SEL, $r = 750\,m$	μ_{LN}, μ	σ_{LN}, σ	$P(X > 160\,dB)$
Raw measurements	5.235	0.024	~ 1
$E = 2500\,kJ$	187.8	4.54	
Sound barrier (-20dB)	5.122	0.027	0.960
$E = 2500\,kJ$	167.8	4.54	
Sound barrier (-20dB)	5.116	0.027	0.936
$E = 2000\,kJ$, $\delta = -0.969$dB	166.9	4.54	
Sound barrier (-20dB)	5.113	0.027	0.918
$E = 1750\,kJ$, $\delta = -1.549$dB	166.3	4.54	

Table 7.3: SEL distr. coefficients and exceedance probability for different cases

L_p, $r = 750\,m$	μ_{LN}, μ	σ_{LN}, σ	$P(X > 190\,dB)$
Raw measurements	5.281	0.014	0.992
$E = 2500\,kJ$	196.6	2.78	
Sound barrier (-20dB)	5.174	0.016	$1.6 \cdot 10^{-6}$
$E = 2500\,kJ$	176.6	2.78	
Sound barrier (-20dB)	5.168	0.016	$3.3 \cdot 10^{-7}$
$E = 2000\,kJ$, $\delta = -0.969$dB	175.6	2.77	
Sound barrier (-20dB)	5.165	0.016	$1.1 \cdot 10^{-7}$
$E = 1750\,kJ$, $\delta = -1.549$dB	175.0	2.78	

Table 7.4: L_p distr. coefficients and exceedance probability for different cases

The results presented show the importance of sound barriers and reduction in impact force (which then releases in the pile less energy). Especially for the sound peak level, from an almost certain exceedance of the sound limit, the introduction of a sound barrier would significantly prevent this from happening.
Regarding the SEL, it appears that in this particular case it is the governing sound level that requires more attention. Even applying a sound barrier and significantly reducing the hammer strike, the probability of not meeting the requirements is still high.
A few considerations must be brought to attention.
First of all, as mentioned before, the simplification relying on a single energy in the pile may have under/over-estimated the sound levels. Also, the coupling of the near and far-field may have affected the results, since in the two steps different soil configurations has been used. Consequently, also the probability distributions have been influenced, and if the whole analyses would have been carried out different means and std values could have been obtained.
The second aspect to consider regards the large amount of scenarios considered and the possibility to overcome instrumental and reading errors from CPT, as well as to account for spatial variability of the soil. It has been shown in sec(4.2) and appendix(A.1) how even considerable measurement errors/differences would not significantly affect the soil properties definition.
This means that unless abrupt or extremely different environments are encountered, the definition of feature distributions sufficiently accounts for the spatial changes of the soil. This is a great benefit, since although the soil is not entirely known, even with a limited amount of data it is possible to replicate a certain extent of variability.
Focusing now on the results obtained, still stressing out that some simplifications have been made, it appears that meeting the German regulations on SEL would hardly be possible. It looks that the North Sea environment may be a difficult one where to install monopiles and more sound reduction measures should be applied. For example if a sound barrier allowing a reduction of $\delta = -30\,dB$ was available, then the probability of exceeding the threshold of $SEL_{LIM} = 160\,dB$ would be $P(X > 160) \sim 30\%$.
This project has shown how, starting from a limited amount of empirical measurements is possible to reproduce the variability of the soil and obtain how likely the environment under consideration may produce certain sound levels. These results can definitely be of use during project stage. Improvements and more accuracy can of course be obtained, but the overall framework can be adapted to many different environments and conditions.

7.3 CORRELATION SOIL PROPERTIES - SOUND LEVELS

Another objective of this thesis was to obtain insight upon which soil characteristics affect affects the most the sound levels.
Confronting all the generated samples input with the estimated noise outputs should help in highlighting possible trends. As done previously only for the soil features and for the energy, the Spearman's rank correlation (eq 5.20) is evaluated, looking for dependence between certain soil characteristic and sound levels.
All the generated soil properties and sound exposure/peak levels are given a rank value. Given a sample of n elements, the values belonging are given a rank $\in (1, ..n)$ defining their position in the ensemble.
The correlation matrix is shown in fig(7.13), the details are instead in appendix(A.14).

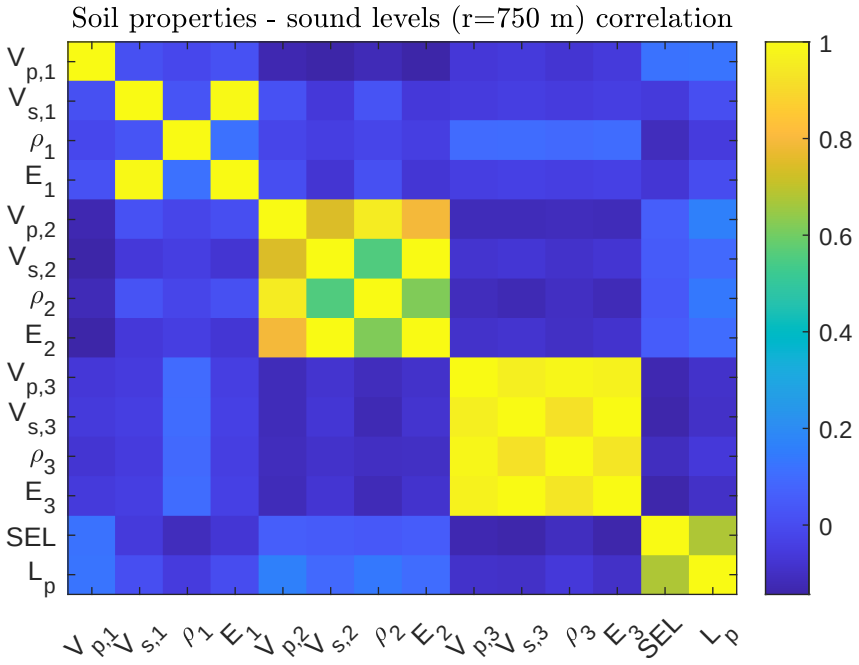

Figure 7.13: Correlation matrix relating input soil properties and sound levels output

The results obtained do not show strong correlations between the soil characteristics and the resulted sound levels. A significant correlation corresponds to values above $r = \pm 0.3 - 0.4$ [33].
Analyzing them more in detail, however, it appears that especially the compressional wave velocity in the first two layers has a positive correlation with the sound levels. This can be explained considering that surface waves tend to quickly fade with depth, while they still have an intense energy at the seabed. Fast wave speed in the top soil region may allow positive interference between the waves in the soil and the ones

in the water, resulting in higher sound levels.
Two main reasons may explain that.
The first is that, given the simplification of relying on a single sound generation analysis (a single energy in the pile) which as said may have affected the sound levels, their dependence with the soil properties may have been affected too. It has also to be considered that the coupling between near and far field matches different soil configurations, which may introduce errors in the sound propagation module. If the sound levels had been under or overestimated, then the calculated correlation may produce results not entirely correct or may not discover the real relation.
The second is the type of correlation that is being evaluated. The Spearman correlation investigates on monotonic relations: if a variable increases it returns how the others either increase or decrease. In such a complex multi-dimensional problem like sound propagation, looking only for monotonic trends may not be sufficient enough. Different and more advanced types of correlations (*pair-copulas* or *vine copulas* [40, 41, 42]) may detect relations between the input and output.
Another reason to use more advanced techniques is that one has to remember that null correlation **does not** equal independence between the variables.

7.4 COMPARISON WITH EXPERIMENTAL SOUND MEASUREMENTS

A comparison of the obtained sound levels is now performed against experimental measurements obtained by an OWT company in the same location as the CPT recordings.
The settings used for the simulations (table 6.1) are close to the configuration of the monopile considered in the experiment [23]. Only the final embedment of the pile was not matched, and in the model the third layer expands until a depth $z = 110 - 41.4 = 68.6\,m$ although the CPT measurements provide information up to a depth $z = 55\,m$.
The exact force excitation is approximated by eq(6.1) [23]. As is discussed further, these uncertainties and other main differences are expected to create some differences in the sound level comparisons.
The empirical readings have been obtained at a radial distance $r = 750\,m$ from the piles, 2 m above seabed, like the simulation estimations.
The measurements have been obtained during the whole installation process, which means that the hammer blows have been modified during the process. However, confidence intervals have been obtained considering the latest stage of the installation, in which the energy derived by the strike remained almost constant around $E \sim 1600$ kJ (fig 7.14).

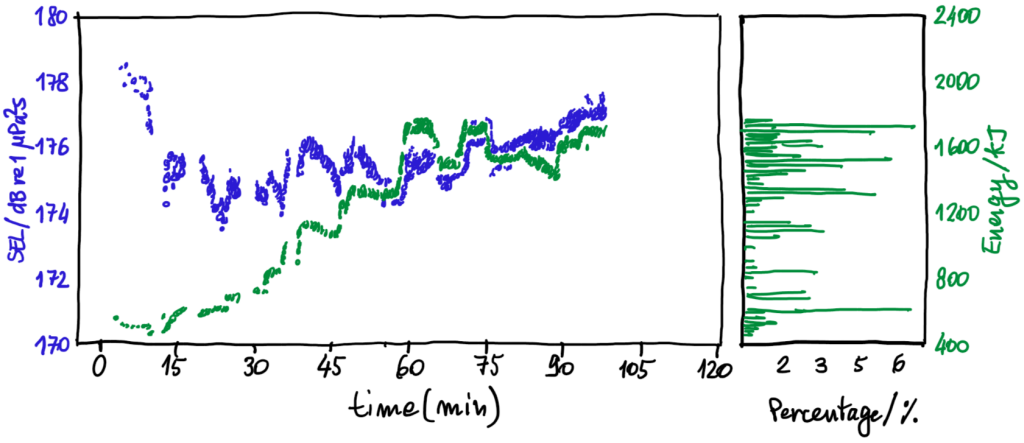

Figure 7.14: Sketch of the experimental measurements for SEL and energy input during construction phase. The sketch is made only for clarity purposes.

The available information regard the lower bound, such that 5% of the measurements stand below it, a mean value and an upper bound, such that only 5% of the measurements are above it. Four stations with same radial distance ($r = 750\,m$) were used each providing the mentioned quantities.
To account for measurement error, a standard uncertainty of $\pm 2\,dB$ is generally applied. This uncertainty is applied on the upper bound: relying on the value that accounts for 95% of all the measurements, it is believed to correctly define the measurement ranges.
In table(7.5) are summarized the percentiles and the standard error.

	5^{th}	μ	95^{th}	Error
SEL				$\pm 2\,dB$
Location 1	$175\,dB$	$176\,dB$	$177\,dB$	
Location 2	$175\,dB$	$177\,dB$	$178\,dB$	
Location 3	$175\,dB$	$177\,dB$	$178\,dB$	
Location 4	$175\,dB$	$177\,dB$	$178\,dB$	
L_p				$\pm 2\,dB$
Location 1	-	$200\,dB$	-	
Location 2	-	$201\,dB$	-	
Location 3	-	$201\,dB$	-	
Location 4	-	$202\,dB$	-	

Table 7.5: Mean, upper and lower bounds from empirical measurements

In order to compare the experimental results with the *SILENCE* ones, the latter need to be scaled to a proper energy entry. Since the raw estimations refer to a case where $Amp = 1, E = 2500\,kJ$, by mean of eq(6.4,6.5) the values can be adjusted.

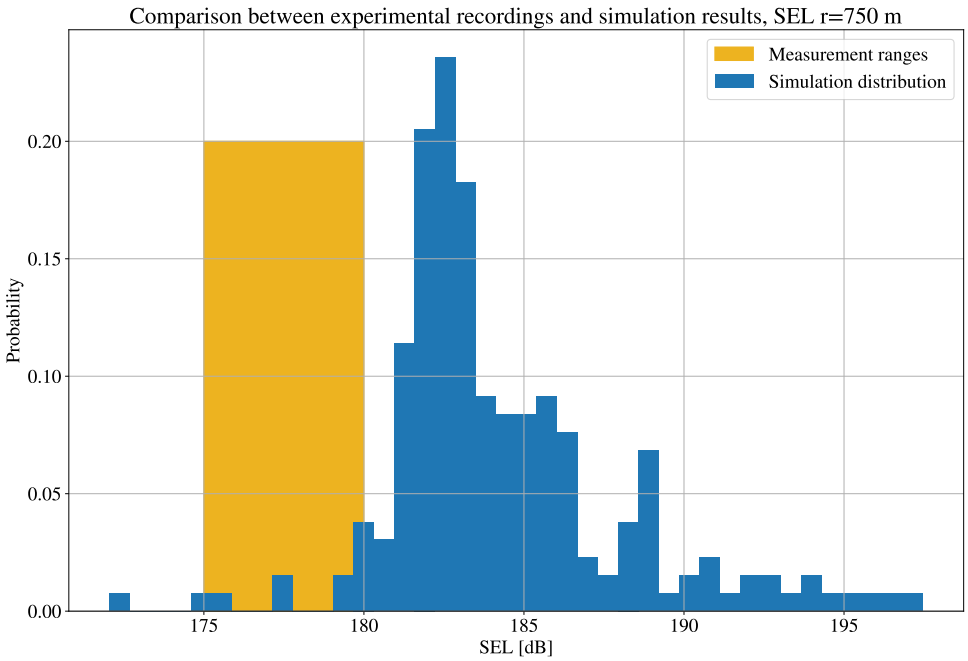

Figure 7.15: Experimental ranges and SILENCE results, SEL, $r = 750\,m$, 2m above seabed. No sound barriers considered, energy input $E^* = 1600$ kJ, $Amp^* = 0.8$.

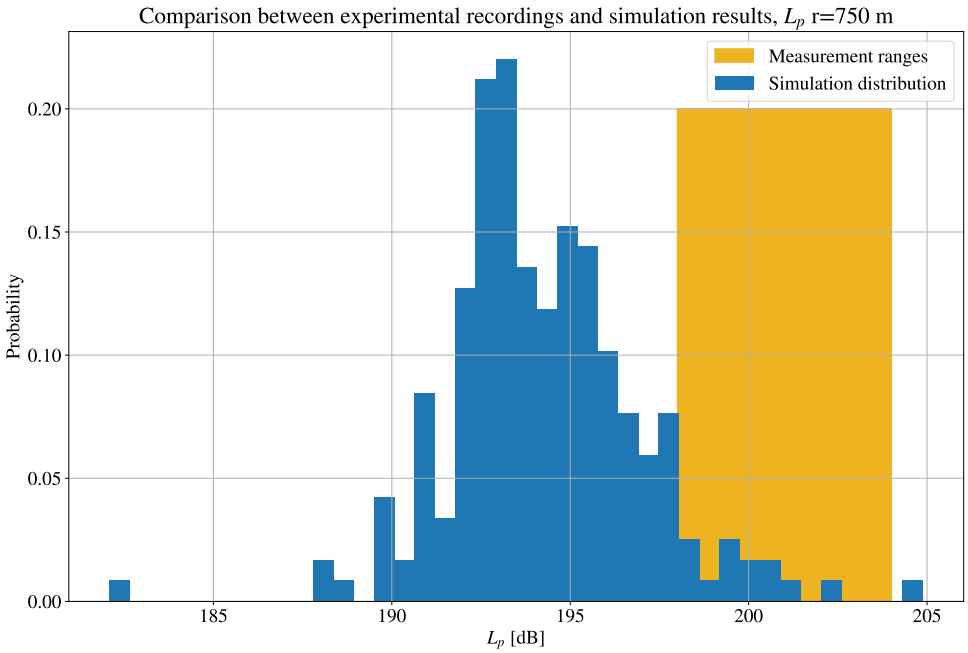

Figure 7.16: Experimental ranges and SILENCE results, L_p, $r = 750\,m$, 2m above seabed. No sound barriers considered, energy input $E^* = 1600$ kJ, $Amp^* = 0.8$.

The new amplitude is therefore $Amp^* = 0.8$ and a decibel reduction $\delta[dB] = -1.938\,dB$ is applied. The new coefficients of the distribution can be obtained applying this correction to all the simulation results, and re-evaluating table(7.1).
In fig(7.15,7.16) the empirical sound ranges ($175 - 180\,dB$ for SEL, $198 - 204\,dB$ for L_p) and the simulation distributions are presented.
The comparisons show some noticeable differences.
The majority of the simulation SEL estimations stand above the experimental sound range. Considering that the range has been defined relying on the 95^{th} percentile, it seems the simulations produced values overestimating the 'real' ones.
On the other hand, the peak levels stand below the experimental interval showing a great accordance: 95% of the simulation results are below the experimental upper bound.
There are several reasons why the empirical 95% of recording are exceeded. As mentioned the not complete match of experimental conditions and the simulation settings, for example the input force, may have affected the comparison. Another explanation is the simplification made on relying on a single sound generation analysis, which may have affected the sound levels. By under or overestimating the energy introduced in the system, and with an approximated coupling between near and far field, the simulation distribution has been affected.
Finally, the experimental results are performed in a single and real envi-

ronment. All the simulations performed can not model exactly the soil properties or the geometry of the seabed. The cases considered are all simplifications, and sometimes investigate on completely different configurations, far from reality.
To summarize, the methodology presented in this thesis focuses on quantifying and deal with the uncertainties involved in the soil characteristics. The comparison between the measured data and the simulated sound levels show that complete accordance has not been reached.
A larger set of samples, the introduction of the correct correlation between the soil layers and the complete simulation steps are suggested for future studies to achieve more accurate prediction of the sound levels.

Part IV

CONCLUSIONS, RECOMMENDATIONS AND FUTURE WORKS

8

DISCUSSION AND CONCLUSIONS

In this chapter a summary of the complete process followed in this thesis is discussed.
The methodology steps for describing the soil and the treatment of the results obtained from the simulations is analyzed.
The suitability and some of the main advantages are pointed out, together with some limitations or improvements that could be adopted.

8.1 SOIL MODELLING

In ch(4,5) the main methodology for dealing with experimental measurements of the soil had been treated.
The approach resulted to be highly flexible referring to its adaptability to the necessary needs and showed the capacity to overcome to a certain extent the lack of data or uncertainty.
Now the steps followed are discussed more in detail.

8.1.1 *CPT measurements*

The first action described was the acquisition of the CPT measurements. A simple technique of reading the area underneath the graphs of the measurements has been used, averaging over the length of the step size defined.
All the steps which followed took into account the possible reading errors, other than the instrumental errors which could not be avoided. It has been shown that even a significant difference in measurements would not drastically change the soil properties and the distributions. This is a major achievement, since it means that possible unavoidable errors would not affect the modelling. Moreover even the real changes in the environment, given that are not significant ($\pm$10%), may still be considered by the definition of the probability distributions.
Once the soil features have been obtained, the values have been compared with different experimental works which verified the reliability of the process.
A higher number of CPT measurements would improve the accuracy of the distribution definition, as well as give insights on the variation of the environment. As mentioned, however, unless the changes would be abrupt (which shouldn't happen in the region involved in the sound

propagation problem) the definition of the distributions and division in layers showed its value in considering minor changes due to the natural variation of the environment.

8.1.2 *Layer's depth*

A procedure to obtain the optimal layer's depth has been investigated. In the particular case of this work, only two extra layers (below the saturated one) have been found, considering also the limitation given by the software. However, in the case it was possible to define more layers the same analysis of variance could be performed, comparing more groups. In this sense, the method results flexible and can adapt to different needs.
A larger number of measurements would improve the definition of a greater set of layers, however in general no more than three/four are necessary for a sufficiently accurate description of the soil.
It has also to be considered that to overcome the approximation using homogeneous layers, a different solution method would be required. In the case of a FEM model, techniques such as spatial correlation could be introduced.
Considering however the high computational time required by these numerical methods, the ANOVA can be considered an efficient technique.

8.1.3 *Soil properties and distributions*

With the available CPT measurements, three fundamental soil properties have been estimated. To do so, many manuals and previous works have been referenced [27, 29, 31, 32], and the empirical relations leading to the most reliable results have been used [30, 43].
These estimations could only improve given more accurate measurements, or once better empirical relations could be found, therefore the steps presented can be considered sufficiently reliable. An investigation upon the sensitivity of these empirical relations has been performed, showing that even significant measurements errors would not produce the same effects on the estimations.
A limited amount of distributions have been considered, however these proved to highly represent the available data. More types could be investigated to increase the fitting, but that was not the scope of this work.
Some assumptions have been made on the correlation between properties and between soil layers. Although it was not possible to verify them, some physical considerations has been used, which agree with the assumed model. Future works may investigate on the correct definition of the correlation between the layers, to improve also the copula distri-

butions.
Since the scope was the definition of a workflow applicable to a model defining horizontal homogeneous layers, and investigating on ways to generate samples for the analyses, the work presented resulted in an efficient approach.

8.2 SOUND LEVEL RESULTS AND STATISTICAL INTERPRETATION

In ch(7), the simulation results have been presented and statistically analyzed. The main steps followed will now be commented, and some advantages of the model highlighted.

8.2.1 *Sampling of input soil properties*

Once the correlation between soil properties had been defined, it was then possible to generate samples from a copula. The advantages of using this technique reside in the versatility of this distribution, which can easily be modified depending on the available data. Moreover, the use of correlation is a major improvement compared with the simply independent random generation. On one hand, as shown, different layers may be considered independent, most probably representing different materials; moreover the dependence that could be found can only apply in a specific location. On the other hand instead, within the layer, the properties are in general correlated: higher density results in higher compressional velocity and shear speed as well.
During the sampling, many different scenarios has been considered, however the single properties do account for relations that would be present even in the real environment.

8.2.2 *Model assumptions*

Given the results from the soil definition, three layers have been used throughout the whole process. The use of only two layers has not been investigated since, as said, three layers showed a better representation of the available data.
A single time and frequency signature has been used to represent the input force of the hammer strike. Although representing a realistic blow, an investigation of the effects of the frequency content on the sound level outputs could help in highlighting critical excitations. On the other hand could also provide blow configurations helping to reduce the noise emission.
Due to the computational time necessary to perform the complete analysis, a main simplification has been made on the sound generation part.

A single simulation, using the average soil properties for all layers, has been used to estimate the energy entry. This energy has then been assumed constant for all the sound propagation problem.
Although still providing reliable results, it is believed this main simplification increased the variation in the sound levels for the far field. That is because each soil configuration radiates differently the energy introduced by the hammer blow, varying the intensity of the pile oscillation and the direct radiation both in the water and in the soil. By using only one energy level, in many cases this may have under or overestimated the proper energy for the particular soil combination, which then may have resulted in lower or higher sound levels.
This approximation can be improved given a sufficient computational time is available. This way more accurate sound levels may be obtained and possibly more insight on the correlation as well.

8.2.3 *Sound level results and distributions*

The sound level estimations from the simulations have then been presented and investigated. Particular attention has been given to the sound results at a radial distance $r = 750\,m$, 2 m above seabed.
The correct procedure to treat the sound levels has been discussed, since treating directly the decibel quantities would produce errors in the estimation of the distribution coefficients.
As done for the soil, a few common distributions have been compared to find the one that best describes the sound estimations. An investigation on the fitting accuracy has been discussed. It is stressed out again that the simplification of using a single sound generation analysis may have affected the results.
A larger number of simulations could improve the determination of a proper sound distribution and improve the information on the most affecting soil parameters. However the number of cases performed ($n = 309$) is sufficiently high to obtain enough insight and verify the validity of the methodology proposed.

8.2.4 *Probability estimations*

The obtained probability/cumulative distribution functions allowed then to evaluate cumulative/exceedance probabilities. An example of the applicability of this method has been presented. Although the exact definition of the soil was not possible, it is believed that by simulating several scenarios this variability could still be investigated. By reproducing different environments it is possible to obtain information on the most probable outcomes.

This result is believed to be highly important to determine whether actions to reduce the sound levels must be taken, such as sound barriers and/or reduction in the blow force.

8.2.5 *Correlation input-output*

The correlation matrix between soil properties input and sound levels output have been evaluated. No significant dependencies have been highlighted, and a reason could be due to the simplification made in the sound generation analysis. Performing the complete analysis may produce better results.
The highest correlation had been found in the top and second layer, and this could be related to the fact that surface waves tend to fade with depth. Soil configuration that allow pressure waves to positively interfere with the waves in the water column may be the most critical.
Given the complexity of the task, however, more advanced techniques such as pair or vine copulas may highlight more insights. Future works may also introduce more parameters to investigate into to obtain more information.

8.2.6 *Experimental measurements comparison*

A final comparison with experimental sound recordings, performed by an OWT company in the North Sea, has been carried out.
These measurements refer to the sound propagating in the case no sound barriers are applied during the construction phase. The campaign has collected the data at a distance $r = 750\,m$ from the pile, $2\,m$ above seabed, therefore they can be confronted with the simulation predictions.
The *SILENCE* estimations have been scaled in such a way to have the same energy input as the one obtained from measurements. Although scaled, the energy simplification in the simulation still persists.
Differences are expected since the information on the pile dimensions, water depths and penetration are not available.
The software results show indeed some differences with the experimental ones, especially for the sound exposure levels (SEL). By performing the complete analysis, it is believed that the results from the software would better agree with the empirical data. Moreover if the exact settings of the pile were known, the results would better agree.
Another, even greater reason, is that the experimental results refer to a single soil configuration, which can not be completely represented by the available model. The simulations reproduce several cases which

may significantly differ from the actual situation, but still provide insights on the most probable outcomes.
To correctly compare the two quantities, a more complete model for the soil, together with improved measurements and estimation of the soil properties, should be used.
The framework presented, however, showed how to obtain important information and is still a valuable tool to take into account during the project phase.

9

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORKS

In this chapter the conclusions of this project are discussed, verifying whether the initial questions have been addressed and answered and if the objectives have been reached. Finally possible approaches and future works which could improve the presented framework are outlined.

9.1 CONCLUSIONS

The main focus of this thesis is the investigation on the uncertainties in the soil modelling and the probabilistic quantification of their effects in the sound propagation. The ultimate goal is to find correlation between the soil parameters and the resulted noise levels, in order to provide more reliable noise predictions to prevent an harmful impact on marine mammals.
This work has been divided in two main parts: the uncertainties in the soil and its correlation in the sound emissions.
The overall framework presented is composed by several statistical and probabilistic intermediate steps. The methodology here presented aims to overcome the limitation of previous works [6] on the soil properties uncertainty and their strong effects on the sound propagation.
To answer the main research question of this project, the following conclusions and discussions are presented below.

1. *With empirical measurements (CPT) how is it possible to account for the variability of the marine environment?*
   In ch.(4) have been presented the steps to treat available experimental records. It would be possible to implement these data in a numerical model, such as FEM, but at the moment is not feasible to measure the complete domain of interest to reproduce it in all the details. Approximations are therefore needed starting from *local* information.
   The existing relations to obtain the soil features have been described. Moreover, a sensitivity to measurement or reading errors has been investigated. It has been demonstrated that even significant errors ($\pm 10\%$ on both the stress data) would not produce the same effects on the properties estimation. This is a major achievement: not only it has been proven that in general minor errors are easily overcame, but the natural fluctuation of the environment is

also took into account. It is normal that measurements in different locations will present different results but in the domain considered it is believed the changes would not be abrupt. Comparing different CPT measurements and statistically interpreting them would improve the definition of the distributions. These would still account for the real variability of the environment, and moreover when sampling different scenarios are investigated. Even if an approximation it is believed it would produce reliable results. The aim to define a method that can easily adapt to different cases is then achieved.

2. *Since the input for the solver must be set, how is it possible to generate samples that account for correlation between the different properties?*
   In ch.(5-6), the main statistical modelling of the soil has been analyzed.
   Two main assumption have been used, regarding the correlation between the properties in the same and in different layers. Although the hypothesis have not been verified, some physical considerations have been adopted to still consider the model valid.
   The first soil layer, being in direct contact with the water column and therefore partially or fully saturated, has been assumed having particular property ranges, each independent from the others. The choice has been made relying on previous experimental works [27, 29, 31, 32] and, as said, physical considerations. Due to the high water content in the first meters of the soil, the compressional velocity is expected to have values very similar to the sound speed in the sea. For the same reason, since there are no shear waves travelling in the water, it is expected that the shear speed in the top soil is generally low. A direct dependence between the two features is however not possible to define in advance. Usually the recordings in this specific region are the most affected by measurement errors and oscillations given the sudden change from the fluid environment. Moreover the Poisson's ratio, which is a property not yet possible to directly estimate, given its variability would allow different combination of the shear and compressional velocity, so that considering them independent results in a valid assumption. Regarding the density, although in general the seabed is less dense than the materials below, different values may be experienced. Again due to the variability of the Poisson's ratio, no direct correlation between the characteristics can be defined in advance, for this limited and particular region.
   Considering the lower layers, when the water content is expected to be smaller, it is then possible to obtain the correlation between the different properties. It has been shown that high correlation

exists.
The second assumption regards the independence of the features belonging to different layers. The whole model is intended to separate the measurements in layers, which have significantly different values and distributions. Finding correlation between different layers, other than difficult due to the limited amount of data, is also extremely specific. The results one would obtain would hardly apply to the whole domain in consideration. The hypothesis of independence results again an effective choice.
With the results obtained, the definition of the copula has been presented, which is a distribution that can easily adapt to different marginal choices. The generation of samples from it then account for the correlation between the properties, producing combinations that are more likely to appear.
It is believed that the question has been sufficiently answered, although a verification of the assumptions could give more value to the framework.

3. *Once obtained the raw sound level measurements, how can they be treated?*
In ch.(7) the sound propagation results have been presented and discussed.
A main issue has been highlighted, due to a simplification performed during the model set-up (ch 6). A single sound generation analysis has been performed, which resulted in a single energy entry in the pile. An analysis on a few different sound generation results has been carried out. It resulted that the average energy of even extremely different environments would correspond to the one obtained by setting all the features to their average value. That is the main reason why all the performed simulations rely on this estimation.
Since, however, each simulation would result in a different energy introduced in the system, the simplification has affected the sound level results. It is believed that the complete analysis would produce more accurate results. Considering this simplification and the minor differences in settings from the empirical measurements case, as shown in the last section of ch.(7), the results provide valuable results.
Nevertheless, the framework presented shows how to correctly statistically interpret the sound levels. Even for this case, the steps allow the method to quickly adapt to different data.
One of the main achievements of this project is the definition of sound probability density functions. Although previous works obtained similar outcomes, they have not intensively investigated on the definition of appropriate input for the simulations, rely-

ing on simple standard cases. The main value of the distributions defined in this work is that they are derived from combinations which account for relations between the parameters. Investigating in several configurations, even if not exactly describing the real environment, provides insight on the most likely outcomes in terms of sound levels. Therefore it answers to how troublesome the environment under consideration might be in terms of expected sound levels.

4. *Analyzing the soil characteristics and the sound level output, is there any significant correlation?*
   One of the last objectives aimed to obtain insights on the relations between particular soil features and the obtained sound levels. The Spearman's correlation matrix presented however did not provide significant ($> 0.3 - 0.4$ [33]) information. The major dependence found regards the compressional velocity in the two top layers ($r \sim 0.16$). It is believed that fast compressional speed in the soil, which is still comparable with the one in the water, may allow positive interference in the pressure levels during the propagation, resulting in higher sound levels. The negative correlation between Young's modulus and sound levels ($r = -0.14$) may be referred to the experienced higher energy introduced in the system when soft soils are considered.
   Some considerations are now made on why no great correlations have been found.
   Firstly, as mentioned before, the simplification made on the energy entry has then affected the sound level outcomes. Since the level used may have under or overestimated the correct energy that the specific combination would have introduced, the sound levels may have been under or overestimated as well. Then the correct correlation between the soil features and resulting sound levels may have been compromised.
   A second reason for not finding sufficient insights is due to the type of correlation that has been estimated. The Spearman's one estimates the monotonic relation between the variables. The sound generation and propagation is a multi-dimensional problem with a significant amount of variables. Finding direct relations is therefore difficult with the mentioned test, and more advanced ones should be used for future works.

Given these conclusions, the main research question is now addressed.
Is it possible to predict the probability of exceeding a dangerous sound level, in order to take precautions and prevent animal harm?
Yes. The uncertainty in the marine environment and the consequent variability in the sound outcomes has been addressed. With the frame-

work presented it is possible to account for the inevitable variation of the soil properties. Although some approximations are necessary, performing several simulations reproducing different environments, still conforming to the real one, provides the most likely outcomes. These results can provide insights on how troublesome an environment can be based on the most likely sound levels obtained. The comparison with experimental measurements, given the approximations made, still show a good agreement and the efficacy of the the method.
With the probability functions then defined, it is possible to estimate, in the environment under consideration, how critical the sound levels may be.
Although some assumptions still need to be verified and some simplifications may be corrected, the framework presented in this thesis is believed to provide a simple, flexible, and yet accurate tool that can help evaluating the risks of excessive sound levels in the required environment.

9.2 RECOMMENDATION AND FUTURE WORKS

The work presented in this thesis achieved most of the aimed objectives and answered most of the required questions.
There are, as already mentioned throughout the document, some improvements or new techniques which can improve the method.
In this section the main ones are brought to attention.

9.2.1 *Stochastic profile composition*

A considerable improvement of the presented framework would be obtained introducing the stochastic profile composition [44]. In general the soil properties increase with depth, that is due to the higher compression force the soil undergoes from the above sediments. Therefore in the measurements it is possible to highlight an increasing trend in the values. By removing this trend and analyzing the filtered recording better estimations for the feature distributions may be obtained, as well as a more accurate correlation estimation.

9.2.2 *Inter-layer correlation*

In this project it has been assumed that only the properties within the stratum are correlated, while each layer is independent from the others. Physical considerations have been used to sill consider reliable the hypothesis.
Given the limited amount of available data, it was not possible to ver-

ify the assumptions. In the future more and more accurate could allow to overcome this limitation. With a more complete correlation of the properties, also the sampling generation would benefit from it.

9.2.3 *Spatial correlation*

As mentioned in ch.(2), recent studies on the definition of multi dimensional random fields are developing relations to describe spatially varying domains.
A random field of particular interest given the many applications that can have is the *Gaussian random field*. By defining the spatial correlation of the properties it would be possible to define the whole domain starting from measurements obtained at different locations. Given the large extension of the marine environment that needs to be considered in the sound propagation problem, this techniques is believed to help in describing the soil. Other than improving the definition of homogeneous layers, could also be applied in FEM to account for varying soil characteristics.

9.2.4 *Complete simulation results*

The simplification applied on the energy entry in the sound generation module may have affected the estimation of the sound levels. Although it is believed that the ones obtained are still reliable, for future works is recommended to perform the complete analysis each time. This would improve the estimation of the probability density functions and the correlation between the features.

9.2.5 *Advanced correlation estimation*

One of the last objectives of this thesis was to identify the critical soil features that mostly affect the sound levels. The adopted test may be inadequate to find relations in a such complex problem as the sound propagation one.
It is believed that more advanced correlation estimators would help in detecting critical dependencies. The use of *pair-copulas* or *vine copulas* may help to approach high-dimensional problems (i. e. many variables) and investigate on their correlation.

9.2.6 *More features considered*

In this project, the main focus had been posed on the soil characteristics, while other model input have been maintained constant or assumptions

have been made on their values. In future investigations more parameters may be analyzed to calculate their effects on the sound levels, as for example different time and frequency signatures for the hammer blow.

BIBLIOGRAPHY

[1] Huikwan Kim et al. "Long range propagation modeling of offshore wind turbine noise using finite element and parabolic equation models." In: *The Journal of the Acoustical Society of America* 131 (4 Apr. 2012), pp. 3392–3392. ISSN: 0001-4966. DOI: 10.1121/1.4708798.

[2] Madsen PT et al. "Wind turbine underwater noise and marine mammals: implications of current knowledge and data needs." In: *Marine Ecol-Progress Ser* (2006).

[3] Slabbekoorn H et al. "A noisy spring: the impact of globally rising underwater sound levels on fish." In: *Trends Ecol Evol* (2010).

[4] Jefferson TA, Hung SK, and Wursig B. "Protecting small cetaceans from coastal development: impact assessment and mitigation experience in Hong Kong." In: *Trends Ecol Evol* (2009).

[5] Andreas Müller et al. *Assessment of Effects of Offshore Wind Energy Facilities on the Marine Environment "Assessment approaches for underwater sound monitoring associated with offshore approval procedures, maritime spatial planning and the marine strategy framework directive-BeMo" Technical Report Monitoring and assessment of the effectiveness of noise mitigation measures to reduce impact by percussive pile driving in German waters*. Federal Maritime and Hydrographic Agency (BSH), 2019.

[6] Tristan Lippert and Otto von Estorff. "The significance of parameter uncertainties for the prediction of offshore pile driving noise." In: *The Journal of the Acoustical Society of America* 136 (5 Nov. 2014), pp. 2463–2471. ISSN: 0001-4966. DOI: 10.1121/1.4896458.

[7] Apostolos Tsouvalas. "Underwater noise emission due to offshore pile installation: A review." In: *Energies* 13 (12 June 2020). ISSN: 19961073. DOI: 10.3390/en13123037.

[8] Pieter Meulendijk-de Mol, Lies van Nieuwerburgh, and Audrey van Mastrigt. *Underwater Noise Social Cost Benefit Analysis*. Royal HaskoningDHV, June 2015. URL: www.royalhaskoningdhv.com.

[9] Maarten Jaspers Faijer and Eric Arends. "(Inter)national discussion on underwater noise; regulation of underwater noise caused by pile driving in the Netherlands." In: *TKI Wind Op Zee* (2015).

[10] A. R. Collins. Miscellaneous Technical Articles. URL: https://www.arc.id.au/SoundLevels.html.

[11] William A. Kuperman et al. *Computational Ocean Acoustics*. 2011. URL: http://www.springer.com/series/3754.

[12] Hans Petter Langtangen. "Finite difference methods for wave motion." In: (2016).

[13] Daniel R. Wilkes, Tim P. Gourlay, and Alexander N. Gavrilov. "Numerical Modeling of Radiated Sound for Impact Pile Driving in Offshore Environments." In: (2015).

[14] Andrew Papanicolaou. "Introduction to Stochastic Differential Equations (SDEs) for Finance." In: (Apr. 2015). URL: http://arxiv.org/abs/1504.05309.

[15] Simo Särkkä and Arno Solin. "Applied Stochastic Differential Equations." In: (2019).

[16] Pep M Mascaró Monserrat. *Stochastic Differential Equations and Applications*. Departament de Matemátiques i Informática, 2017.

[17] Jonas von Pein et al. "A hybrid model for the 3D computation of pile driving noise." In: (2020).

[18] A. Tsouvalas and A. V. Metrikine. "A three-dimensional vibroacoustic model for the prediction of underwater noise from offshore pile driving." In: *Journal of Sound and Vibration* 333 (8 Apr. 2014), pp. 2283–2311. ISSN: 0022460X. DOI: 10.1016/j.jsv.2013.11.045.

[19] Apostolos Tsouvalas. "Underwater noise generated by offshore pile driving." In: (2015). URL: www.gildeprint.nl.

[20] Adolphus Lye, Alice Cicirello, and Edoardo Patelli. "Sampling methods for solving Bayesian model updating problems: A tutorial." In: *Mechanical Systems and Signal Processing* 159 (Oct. 2021). ISSN: 10961216. DOI: 10.1016/j.ymssp.2021.107760.

[21] Dionissioss T. Hristopulos. *Advances in Geographic Information Science - Random Fields for Spatial Data Modeling*. Techincal University of Crete, 2020. URL: http://www.springer.com/series/7712.

[22] Apostolos Tsouvalas and Andrei V. Metrikine. "Structure-borne wave radiation by impact and vibratory piling in offshore installations: From sound prediction to auditory damage." In: *Journal of Marine Science and Engineering* 4 (3 2016). ISSN: 20771312. DOI: 10.3390/jmse4030044.

[23] Yaxi Peng et al. "A fast computational model for near- and far-field noise prediction due to offshore pile driving." In: *The Journal of the Acoustical Society of America* 149 (3 Mar. 2021). ISSN: 0001-4966. DOI: 10.1121/10.0003752.

[24] A. Tsouvalas, Y. Peng, and A. Metrikine. "Underwater Noise Generated By Offshore Pile Driving: A Pile-Soil-Water Vibroacoustic Model Based On A Mode Matching Method." In: (2019).

[25] Moritz B. Fricke and Raimund Rolfes. "Towards a complete physically based forecast model for underwater noise related to impact pile driving." In: *The Journal of the Acoustical Society of America* 137 (3 Mar. 2015), pp. 1564–1575. ISSN: 0001-4966. DOI: 10.1121/1.4908241.

[26] Qingpeng Deng et al. "Modelling of offshore pile driving noise using a semi-analytical variational formulation." In: *Applied Acoustics* 104 (Mar. 2016), pp. 85–100. ISSN: 1872910X. DOI: 10.1016/j.apacoust.2015.11.002.

[27] Hao Yang et al. "Elastic modulus calculation model of a soil-rock mixture at normal or freezing temperature based on micromechanics approach." In: *Advances in Materials Science and Engineering* 2015 (Feb. 2015). ISSN: 16878442. DOI: 10.1155/2015/576080.

[28] Paul W. Mayne. *A Synthesis of Highway Practice NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM NCHRP SYNTHESIS 368*. URL: www.TRB.org.

[29] Thomas M. Brocher. "Empirical relations between elastic wavespeeds and density in the Earth's crust." In: *Bulletin of the Seismological Society of America* 95 (6 Dec. 2005), pp. 2081–2092. ISSN: 00371106. DOI: 10.1785/0120050077.

[30] Burt G Look. *Handbook of Geotechnical Investigation and Design Tables*. 2007.

[31] Kai Schumann et al. "P and S wave velocity measurements of water-rich sediments from the Nankai Trough, Japan." In: *Journal of Geophysical Research: Solid Earth* 119 (2 2014), pp. 787–805. ISSN: 21699356. DOI: 10.1002/2013JB010290.

[32] Wojciech Sas et al. "Dynamic characterization of cohesive material based on wave velocity measurements." In: *Applied Sciences (Switzerland)* 6 (2 2016). ISSN: 20763417. DOI: 10.3390/app6020049.

[33] John A. Rice. *Mathematical statistics and data analysis*. Thomson, Brooks, and Cole, 2007. ISBN: 0534399428.

[34] Smail Mahdi and Myrtene Cenac. *Estimating parameters of Gumbel distribution using the methods of moments, probability weighted moments and Maximum Likelihood*. Universidad de Costa Rica, May 2004.

[35] Christian Genest and Anne-Catherine Favre. "Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask." In: *JOURNAL OF HYDROLOGIC ENGINEERING* (July 2007). DOI: 10.1061/ASCE1084-0699200712:4347.

[36] J. K. Vrijling and P. H. A. J. M. Van Gelder. *CTWA5310 - PROBABILISTIC DESIGN IN HIDRAULIC ENGINEERING*. TU Delft, 2002.

[37] Jorrit Van Rhijn. *Reduction of underwater piling noise An optimization of the impact force to reduce underwater noise during the installation of a large sized monopile*. Technical University of Delft, 2017.

[38] Frode Haukland Gunnar Taraldsen Truls Berge. "Uncertainty of decibel levels." In: *The Journal of the Acoustical Society of America* 138 (3 Sept. 2015), EL264–EL269. ISSN: 0001-4966. DOI: 10.1121/1.4929619.

[39] Mikko Kylliäinen. *Standard Deviations in Field Measurements of Impact Sound Insulation*. 2004.

[40] Kjersti Aas et al. "Pair-copula constructions of multiple dependence." In: *Insurance: Mathematics and Economics* 44 (2 Apr. 2009), pp. 182–198. ISSN: 01676687. DOI: 10.1016/j.insmatheco.2007.02.001.

[41] Johannes ; Stübinger, Benedikt ; Mangold, and Christopher Krauss. *Statistical arbitrage with vine copulas*. University of Erlangen-Nurnberg, Department of Statistics and Econometrics, Oct. 2016. URL: http://hdl.handle.net/10419/147450www.econstor.eu.

[42] Ruolan Yu et al. "A vine copula-based modeling for identification of multivariate water pollution risk in an interconnected river system network." In: *Water (Switzerland)* 12 (10 Oct. 2020). ISSN: 20734441. DOI: 10.3390/w12102741.

[43] P.W. Mayne F.H. Kulhawy. "Manual on estimating soil properties for foundation design." In: (2007).

[44] J M Sinke. *A probabilistic approach to sheet pile driveability predictions by vibro hammers*. Delft University of Technology, Apr. 2020. URL: http://repository.tudelft.nl/..

Part V

APPENDIX

A

APPENDIX

A.1 SOIL PROPERTIES SENSITIVITY TO CPT MEASUREMENTS

In this section the sensitivity of the soil properties to changes in the CPT measurements is performed.
Due to possible instrumental error, as well as reading errors, it has been assumed that an error of $\pm 10\%$ could occur, although it is unlikely such a magnitude.
The recordings obtained directly, without applying any error, are labelled as *raw measurements*. Then the error is applied on the CPT stresses, and the properties derived with the new values are obtained.
Finally two extreme cases are considered, where the lower bound of one CPT measurement is paired with the upper bound of the other. The soil features such obtained are the ones the most different from the original values. However, even in these extreme cases, the changes in values on the characteristics is not that significant.
This is a valuable result. Not only it verifies that even a significant error in the measurement would slightly affect the soil properties but that then natural variation of the properties in the environment could still accurately be described by a limited amount of CPT measurements.
In table(A.1) the results are performed from the measurements obtained at a depth $z = 15\,m$ below seabed.

	$S_f[MPa]$	$q_t[MPa]$	$V_s[m/s]$		$\rho[kg/m^3]$		$V_p[m/s]$	
Raw	0.3341	39.44	360.90		1978	$(z = 15)$	1801	
-10%	0.3007	35.50	352.98		1970		1780	
$+10\%$	0.3676	43.39	368.12		1985		1820	
$-10\%S_f, +10\%q_t$			346.61	96.0%	1963	99.2%	1763	97.9%
$+10\%S_f, -10\%q_t$			374.88	103.9%	1992	100.7%	1838	102.1%

Table A.1: Sensitivity study on the effects of errors in CPT measurements on soil features.

In fig(A.1-A.3) the complete profiles are presented to demonstrate the effects of possible measurement errors.

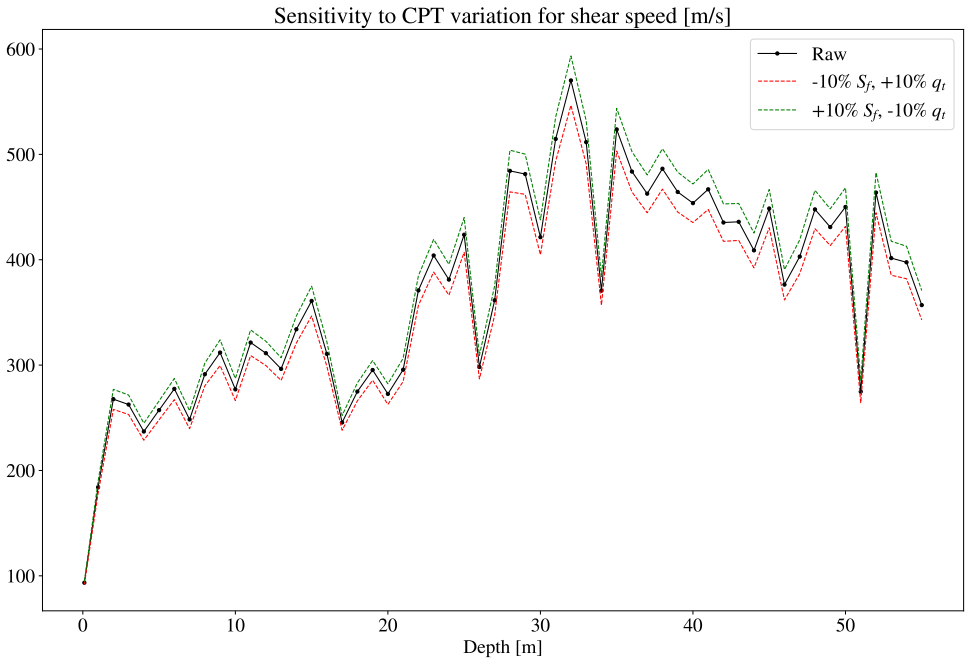

Figure A.1: Shear speed profile, showing the variation due to possible CPT measurement errors

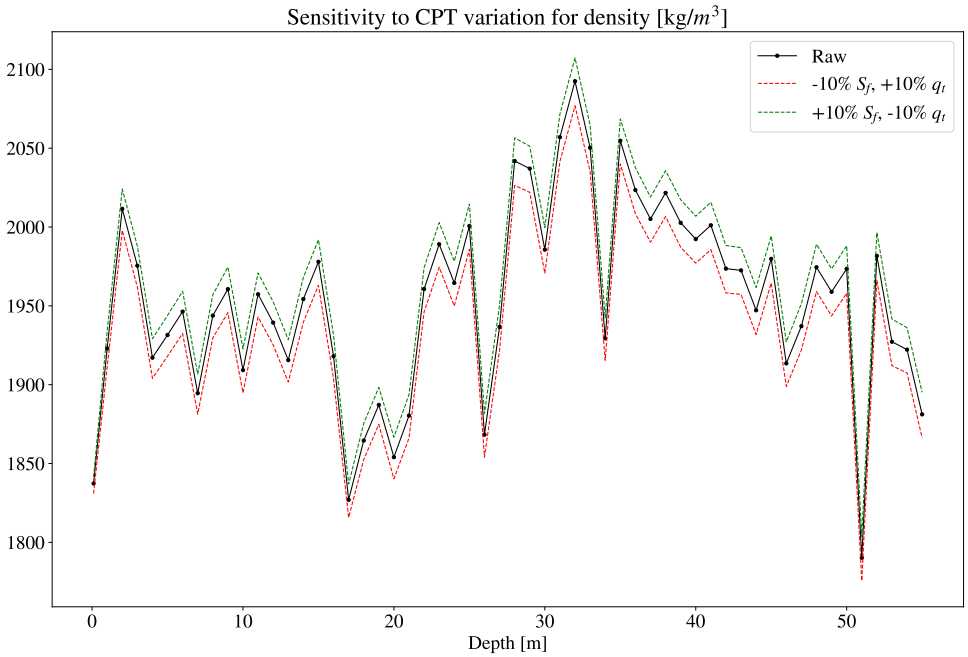

Figure A.2: Density profile, showing the variation due to possible CPT measurement errors

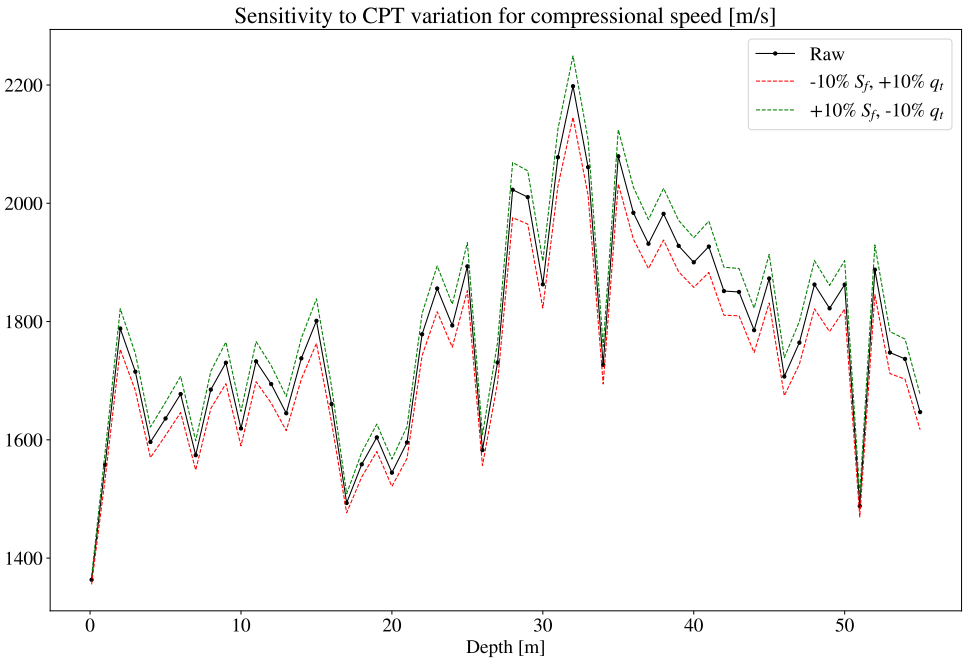

Figure A.3: Compressional speed profile, showing the variation due to possible CPT measurement errors

A.1.1 *Partial derivatives analysis*

Another approach to investigate how the variability of the CPT measurements reflects on the property estimation is to evaluate the variance of the equation.
If we assume a distribution for the CPT, with a certain mean and std, we can evaluate the corresponding mean and std of the soil property. This way we have an estimation of the variability of the equation to the input variables.
Let's assume that the CPT measurements have a coefficient of variation (c.o.v) of 10%:

$$c.o.v = \frac{\sigma}{\mu}$$

Where μ, σ respectively represent the mean and standard deviation of the CPT.
If we want to estimate the mean and std of eq(4.1), for the mean we simply evaluate inserting the average values of S_f and q_t, while for the std we need to evaluate the partial derivatives.

$$\begin{cases} \mu_{V_s} = (10.1 \cdot \log(\mu_{q_t}[kPa]) - 11.4)^{1.67} \cdot \left(\frac{\mu_{S_f}}{\mu_{q_t}} \cdot 100\right)^{0.3} \\ \sigma_{V_s} = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial V_s}{\partial X_i} \cdot \sigma_{X_i}\right)^2} = \sqrt{\left(\frac{\partial V_s}{\partial q_t} \cdot \sigma_{q_t}\right)^2 + \left(\frac{\partial V_s}{\partial S_f} \cdot \sigma_{S_f}\right)^2} \end{cases} \quad (A.1)$$

Where X_i are the variables contained in the equation, in this case S_f and q_t.
If, as said, we consider a $c.o.v = 10\%$ for the CPT data, the resulting coefficient of variation for the shear speed is instead $c.o.v = 3\%$, which is three times smaller. For the density, it is even smaller: $c.o.v = 0.05\%$. This result confirm again that even a considerable difference in CPT measurements would not highly vary the soil property estimation.

A.2 LOGNORMAL DISTRIBUTION

A clarification is here presented in order to avoid committing mistakes when using the Lognormal distribution.
As mentioned, the logarithm of the same distribution follows a normal one. Therefore it is possible to define four coefficients: the mean and standard deviation of the equivalent normal distribution, and the effective mean and standard deviation of the lognormal.
The probability density function and cumulative are defined as mentioned in the work like:

$$\begin{cases} f(x) = \dfrac{1}{x\sigma_{LN}\sqrt{2\pi}} \cdot e^{-\frac{(\ln(x)-\mu_{LN})^2}{2\sigma_{LN}^2}}; \\ F(x) = \dfrac{1}{\sigma_{LN}\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \frac{1}{s} e^{-\frac{(\ln(s)-\mu_{LN})^2}{2\sigma_{LN}^2}} ds \end{cases} \tag{A.2}$$

Where μ_{LN}, σ_{LN} are the equivalent mean and std of the normal distribution, which however in many softwares are the necessary coefficients used to define the lognormal. Once the lognormal is defined, it has its own mean and standard deviation, denoted as μ, σ.
Here will be presented the relation that exists between the coefficients of the equivalent normal and the ones of the lognormal.
If the available coefficients are the equivalent normal distribution ones, then the lognormal mean and std can be obtained through:

$$\begin{cases} \mu = e^{(\mu_{LN}+0.5\sigma_{LN}^2)} \\ \sigma = \mu\sqrt{e^{\sigma_{LN}^2}-1} \end{cases} \tag{A.3}$$

If instead given are the mean and std of the lognormal, μ, σ, the equivalent normal distribution parameters become:

$$\begin{cases} \mu_{LN} = \ln(\mu) - 0.5\sigma_{LN}^2 \\ \sigma_{LN} = \sqrt{\ln\left(1+\left(\dfrac{\sigma}{\mu}\right)^2\right)} \end{cases} \tag{A.4}$$

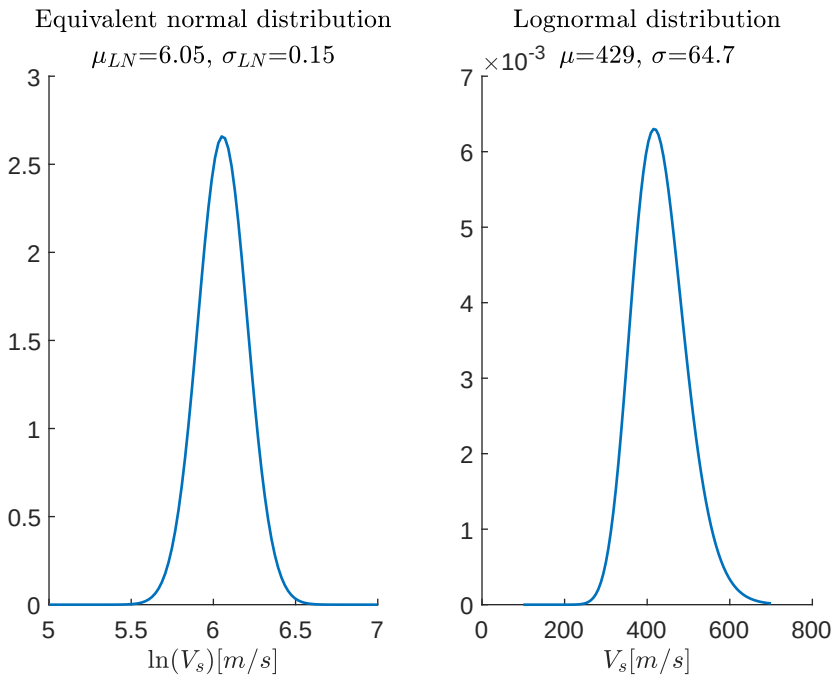

Figure A.4: Equivalent normal distribution and Lognormal distribution, with respective coefficients

In fig(A.4) the equivalent normal distribution, obtained from the logarithm of the data, and the lognormal distribution with the actual values is presented.

A.3 ENERGY CORRELATION MATRIX

In eq(A.5) is presented the correlation matrix obtained analyzing the soil properties and the related energy inserted in the system.

$$C_{Sp,En} = \begin{bmatrix}
1 & -0.20 & -0.56 & -0.17 & 0.35 & 0.99 & -0.05 & 0.28 & -0.03 & 0.34 & 0.99 & 0.05 & -0.56 & 0.08 & 0.32 & -0.04 \\
-0.20 & 1 & 0.56 & 0.98 & -0.98 & -0.25 & 0.90 & -0.56 & 0.84 & -0.89 & -0.25 & 0.80 & 0 & 0.79 & -0.79 & -0.52 \\
-0.56 & 0.56 & 1 & 0.58 & -0.58 & -0.57 & 0.28 & -0.88 & 0.19 & -0.33 & -0.57 & 0 & 0.25 & 0.03 & -0.14 & 0.28 \\
-0.17 & 0.98 & 0.58 & 1 & -0.96 & -0.22 & 0.89 & -0.52 & 0.87 & -0.87 & -0.22 & 0.79 & 0.11 & 0.81 & -0.77 & -0.48 \\
0.35 & -0.98 & -0.58 & -0.96 & 1 & 0.41 & -0.89 & 0.52 & -0.83 & 0.93 & 0.41 & -0.79 & -0.11 & -0.77 & 0.84 & 0.54 \\
0.99 & -0.25 & -0.57 & -0.22 & 0.41 & 1 & -0.13 & 0.28 & -0.09 & 0.41 & 1 & 0 & -0.57 & 0.03 & 0.38 & 0.03 \\
-0.05 & 0.90 & 0.28 & 0.89 & -0.89 & -0.13 & 1 & -0.28 & 0.94 & -0.94 & -0.13 & 0.90 & 0 & 0.89 & -0.86 & -0.64 \\
0.28 & -0.56 & -0.88 & -0.52 & 0.52 & 0.28 & -0.28 & 1 & -0.08 & 0.22 & 0.28 & 0 & 0.25 & 0.03 & 0 & -0.28 \\
-0.03 & 0.84 & 0.19 & 0.87 & -0.83 & -0.09 & 0.94 & -0.08 & 1 & -0.92 & -0.09 & 0.94 & 0.22 & 0.96 & -0.89 & -0.64 \\
0.34 & -0.89 & -0.33 & -0.87 & 0.93 & 0.41 & -0.94 & 0.22 & -0.92 & 1 & 0.41 & -0.89 & -0.22 & -0.87 & 0.96 & 0.67 \\
0.99 & -0.25 & -0.57 & -0.22 & 0.41 & 1 & -0.13 & 0.28 & -0.09 & 0.41 & 1 & 0 & -0.57 & 0.03 & 0.38 & 0.03 \\
0.05 & 0.80 & 0 & 0.79 & -0.79 & 0 & 0.90 & 0 & 0.94 & -0.89 & 0 & 1 & 0 & 0.98 & -0.91 & -0.74 \\
-0.56 & 0 & 0.25 & 0.11 & -0.11 & -0.57 & 0 & 0.25 & 0.22 & -0.22 & -0.57 & 0 & 1 & 0.11 & -0.28 & 0 \\
0.08 & 0.79 & 0.03 & 0.81 & -0.77 & 0.03 & 0.89 & 0.03 & 0.96 & -0.87 & 0.03 & 0.98 & 0.11 & 1 & -0.89 & -0.70 \\
0.32 & -0.79 & -0.14 & -0.77 & 0.84 & 0.38 & -0.86 & 0 & -0.89 & 0.96 & 0.38 & -0.91 & -0.28 & -0.89 & 1 & 0.78 \\
-0.04 & -0.52 & 0.28 & -0.48 & 0.54 & 0.03 & -0.64 & -0.28 & -0.64 & 0.67 & 0.03 & -0.74 & 0 & -0.70 & 0.78 & 1
\end{bmatrix} \tag{A.5}$$

A.4 SOUND LEVEL MEAN AND STANDARD DEVIATION

In this section an example of the correct statistical interpretation of the sound levels and the errors that would arise treating directly the decibel quantities is addressed.
Let's assume we obtained four sound exposure level quantities:

$$SEL = \begin{bmatrix} 181.5 \\ 179.8 \\ 184.2 \\ 182.7 \end{bmatrix}$$

Estimating directly directly the mean and standard deviation we would obtain:

$$\hat{SEL} = 182.05\,dB \qquad \sigma_{SEL} = 1.61\,dB$$

Using instead the approach presented in section(7.1.1) we first need to evaluate the following quantities α:

$$\alpha_i = 10^{SEL_i/10} \rightarrow \alpha = \begin{bmatrix} 141.25 \\ 95.50 \\ 263.03 \\ 186.21 \end{bmatrix} \cdot 10^{16} \tag{A.6}$$

Since the high values involved, it is more practical to divide the quantities in this case by a factor 10^{16}. Obtaining now the average value and std for α:

$$\hat{\alpha} = 171.5 \qquad \sigma_\alpha = 61.815 \tag{A.7}$$

Applying now eq(7.8), multiplying back the quantity to its original dimension, we obtain the correct estimation of the mean and standard deviation, which are:

$$\begin{cases} E[SEL(\alpha)] = 10\log_{10}(\hat{\alpha} \cdot 10^{16}) = 10\log_{10}(171.5 \cdot 10^{16}) = 182.34\,dB \\ \sigma[SEL(\alpha)] = \dfrac{10}{ln(10)} \cdot \dfrac{\sigma_\alpha}{\hat{\alpha}} = \dfrac{10}{ln(10)} \dfrac{61.815}{171.5} = 1.57\,dB \end{cases} \tag{A.8}$$

As we can see, the coefficients obtained are different from the ones calculated directly with the decibel quantities. Therefore, to avoid under/overestimation fo the sound levels, the real quantities need to be obtained before the use of the maximum likelihood estimator.
For seek of completeness, the coefficients for the Gumbel and Lognormal will now be obtained.

A.4.1 *Gumbel distribution*

By using the approach described in ch(5) on the α quantities the Gumbel distribution parameters can be estimated. An initial β_0 is required to start the Newton iteration scheme and a value quite far from the final is taken, to show the efficiency of it.

$$\begin{cases} \beta_0 = 100 \\ g(\beta_0) = 67.6 \\ g'(\beta_0) = 1.25 \end{cases} \quad \begin{cases} \beta_1 = \beta_0 - \dfrac{g(\beta_0)}{g'(\beta_0)} = 46.15 \\ g(\beta_1) = -7.65 \\ g'(\beta_1) = 1.58 \end{cases} \quad \begin{cases} \beta_2 = 50.97 \\ \mu = 141.8 \end{cases} \tag{A.9}$$

If we continue with the iterations, the β value would stabilize at 51.04, $\mu = 141.8$. Then the Gumbel coefficients for the SEL quantity are:

$$\begin{cases} \mu_{SEL} = 10\log_{10}(\mu \cdot 10^{16}) = 10\log_{10}(141.8 \cdot 10^{16}) = 181.52\,dB \\ \beta_{SEL} = \dfrac{10}{ln(10)} \cdot \dfrac{\beta}{\mu} = \dfrac{10}{ln(10)} \cdot \dfrac{\beta}{\mu} == \dfrac{10}{ln(10)} \dfrac{51.04}{141.8} = 1.56\,dB \end{cases} \tag{A.10}$$

A.4.2 *Lognormal distribution*

The same procedure can be applied to the Lognormal distribution.
A few more steps are required, first the logarithm of the initial α quantities are required:

$$\frac{\alpha}{10^{16}} = \begin{bmatrix} 141.25 \\ 95.50 \\ 263.03 \\ 186.21 \end{bmatrix} \rightarrow \ln(\alpha) = \begin{bmatrix} 4.95 \\ 4.56 \\ 5.57 \\ 5.23 \end{bmatrix} \tag{A.11}$$

The equivalent normal coefficients, and then the average and standard deviation of the lognormal distribution of α are:

$$\begin{cases} \hat{\alpha}_{LN} = 5.077 \rightarrow \mu_\alpha = exp(\hat{\alpha}_{LN} + 0.5\sigma^2_{LN,\alpha}) = 171.78 \\ \sigma_{LN,\alpha} = 0.371 \rightarrow \sigma_\alpha = \mu_\alpha \sqrt{exp(\sigma^2_{LN,\alpha}) - 1} = 1.67 \end{cases} \tag{A.12}$$

Now to transform back to sound levels, first the lognormal parameters are obtained, and then the equivalent normal ones.

$$\begin{cases} \mu_{SEL} = 10\log_{10}(\mu_\alpha \cdot 10^{16}) = 182.35\,dB, \sigma_{SEL} = \dfrac{10}{\ln(10)} \dfrac{\sigma_\alpha}{\mu_\alpha} = 1.67\,dB \\ \mu_{LN,SEL} = \ln(\mu_{SEL}) - 0.5\sigma^2_{LN,SEL} = 5.206, \sigma_{LN,SEL} = \sqrt{\ln\left(1 + \left(\dfrac{\sigma_\alpha}{\mu_\alpha}\right)^2\right)} = 0.009 \end{cases} \tag{A.13}$$

A.5 CORRELATION MATRIX SOIL PROPERTIES - SOUND LEVEL

In eq(A.14) is shown the evaluated correlation matrix between soil properties input and sound levels from the *SILENCE* simulations.

$$C_{Sp,Sl} = \begin{bmatrix} 1 & 0.02 & -0.01 & 0.02 & -0.13 & -0.14 & -0.12 & -0.14 & -0.07 & -0.06 & -0.08 & -0.06 & 0.12 & 0.13 \\ 0.02 & 1 & 0.03 & 0.99 & 0.02 & -0.06 & 0.02 & -0.06 & -0.05 & -0.04 & -0.05 & -0.04 & -0.06 & 0.01 \\ -0.01 & 0.03 & 1 & 0.12 & -0.02 & -0.04 & -0.02 & -0.04 & 0.10 & 0.11 & 0.10 & 0.10 & -0.11 & -0.05 \\ 0.02 & 0.99 & 0.12 & 1 & 0.01 & -0.07 & 0.01 & -0.07 & -0.04 & -0.03 & -0.04 & -0.03 & -0.07 & 0 \\ -0.13 & 0.02 & -0.02 & 0.01 & 1 & 0.74 & 0.95 & 0.79 & -0.11 & -0.11 & -0.11 & -0.12 & 0.06 & 0.16 \\ -0.14 & -0.06 & -0.04 & -0.07 & 0.74 & 1 & 0.55 & 1.00 & -0.08 & -0.07 & -0.09 & -0.07 & 0.05 & 0.10 \\ -0.12 & 0.02 & -0.02 & 0.01 & 0.95 & 0.55 & 1 & 0.62 & -0.11 & -0.12 & -0.10 & -0.12 & 0.04 & 0.14 \\ -0.14 & -0.06 & -0.04 & -0.07 & 0.79 & 1.00 & 0.62 & 1 & -0.09 & -0.08 & -0.10 & -0.08 & 0.05 & 0.10 \\ -0.07 & -0.05 & 0.10 & -0.04 & -0.11 & -0.08 & -0.11 & -0.09 & 1 & 0.96 & 0.99 & 0.97 & -0.13 & -0.09 \\ -0.06 & -0.04 & 0.11 & -0.03 & -0.11 & -0.07 & -0.12 & -0.08 & 0.96 & 1 & 0.92 & 1 & -0.14 & -0.09 \\ -0.08 & -0.05 & 0.10 & -0.04 & -0.11 & -0.09 & -0.10 & -0.10 & 0.99 & 0.92 & 1 & 0.94 & -0.10 & -0.06 \\ -0.06 & -0.04 & 0.10 & -0.03 & -0.12 & -0.07 & -0.12 & -0.08 & 0.97 & 1 & 0.94 & 1 & -0.14 & -0.09 \\ 0.12 & -0.06 & -0.11 & -0.07 & 0.06 & 0.05 & 0.04 & 0.05 & -0.13 & -0.14 & -0.10 & -0.14 & 1 & 0.68 \\ 0.13 & 0.01 & -0.05 & 0 & 0.16 & 0.10 & 0.14 & 0.10 & -0.09 & -0.09 & -0.06 & -0.09 & 0.68 & 1 \end{bmatrix} \tag{A.14}$$

A.6 RELATIONS BETWEEN INPUT FEATURES AND SOUND RESULTS

Below are presented the various properties samples paired with the respective sound levels. The linear correlation is highlighted.

Figure A.5: Distribution of the compressional velocity samples and the obtained SEL/L_p - 1st layer

Figure A.6: Distribution of the shear velocity samples and the obtained SEL/L_p - 1st layer

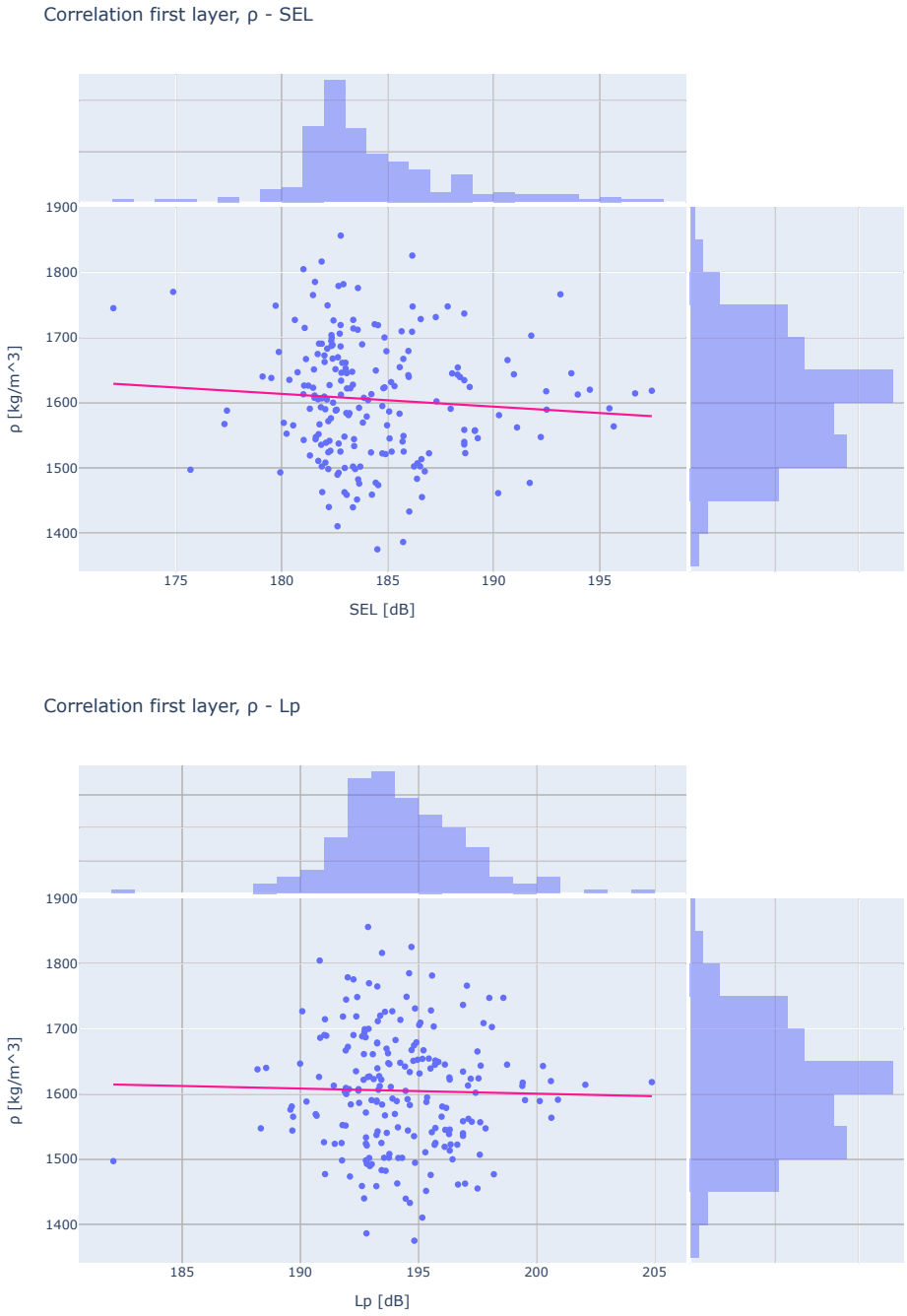

Figure A.7: Distribution of the density samples and the obtained SEL/L_p - 1st layer

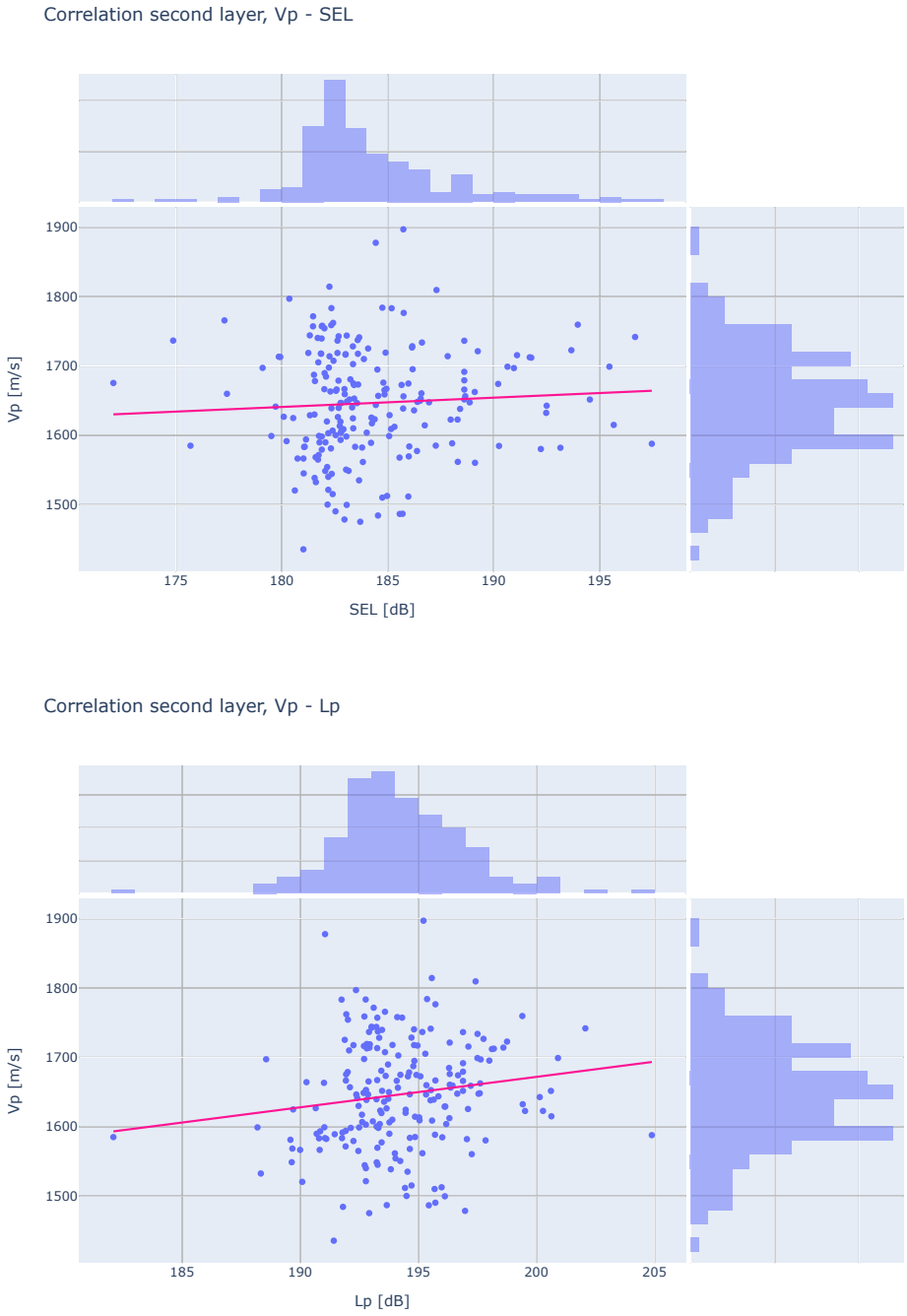

Figure A.8: Distribution of the compressional velocity samples and the obtained SEL/L_p - 2nd layer

Figure A.9: Distribution of the shear velocity samples and the obtained SEL/L_p - 2nd layer

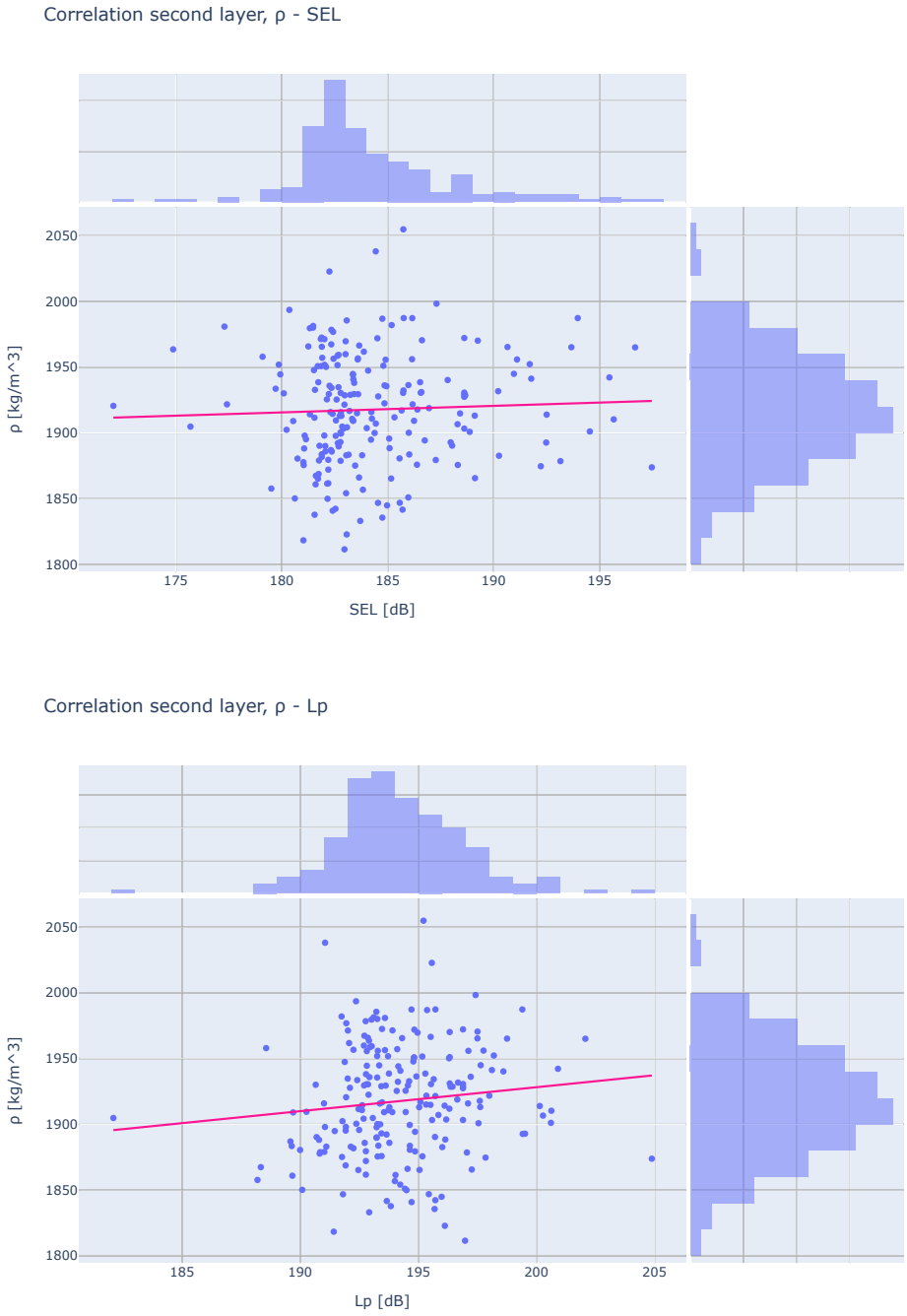

Figure A.10: Distribution of the density samples and the obtained SEL/L_p - 2nd layer

Figure A.11: Distribution of the compressional velocity samples and the obtained SEL/L_p - 3rd layer

Figure A.12: Distribution of the shear velocity samples and the obtained SEL/L_p - 3rd layer

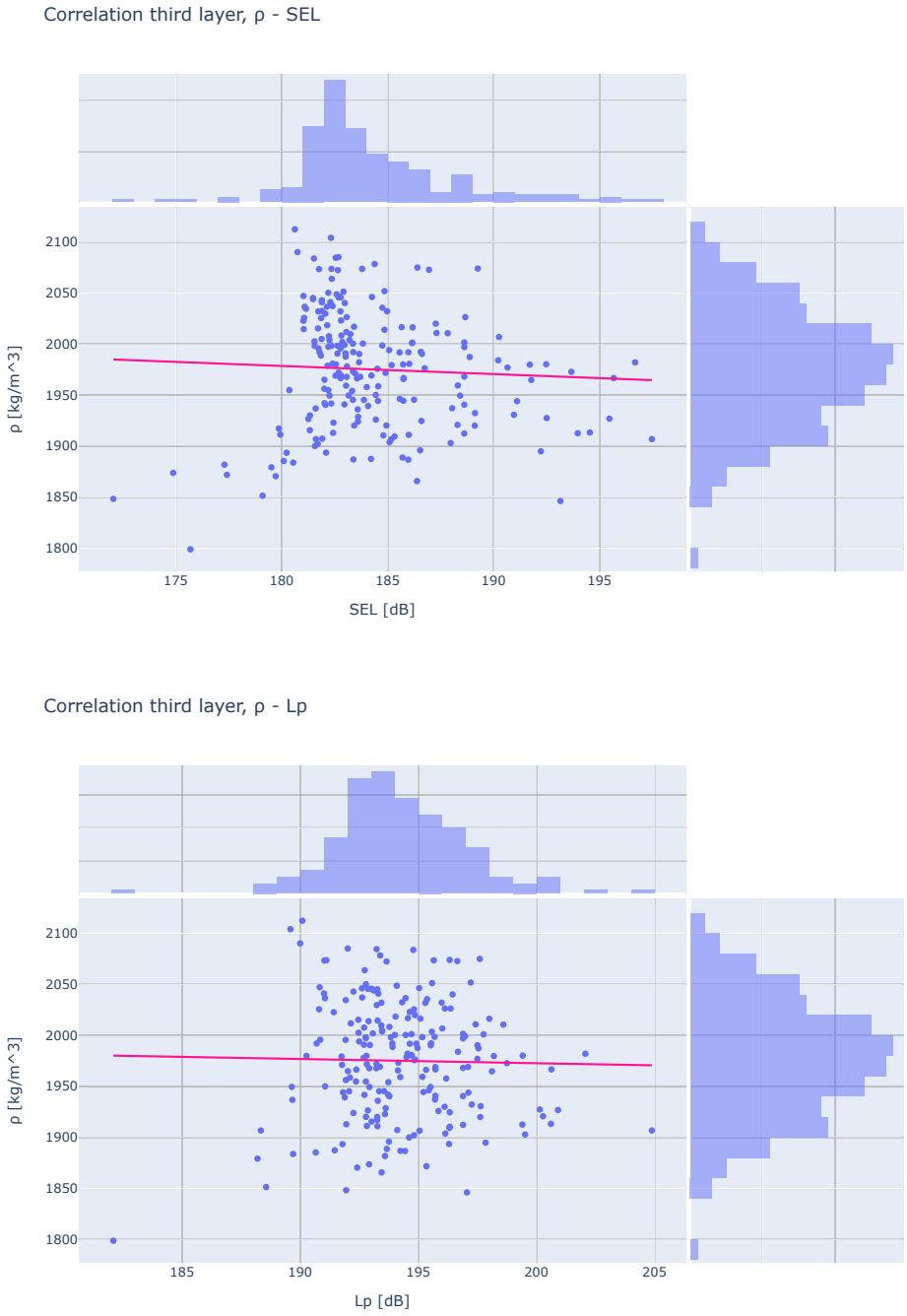

Figure A.13: Distribution of the density samples and the obtained SEL/L_p - 3rd layer