U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

COWI

PROBABILISTIC ANALYSIS IN THE DESIGN OF EMBANKMENTS OVER SOFT SOIL

IVAN OLIVEIRA DE ALMEIDA

Dissertation submitted for partial requirements attendance for the degree on

MASTER ON CIVIL ENGINEERING — GEOTECHNICAL SPECIALIZATION

Supervisor: Professor Doutor José Manuel Mota Couto Marques

Co- Supervisor: Professora Doutora Sara Rios da Rocha e Silva

Co- Supervisor: Professor Doutor Ole Hededal

2020, JUNE

MESTRADO INTEGRADO EM ENGENHARIA CIVIL 2019/2020

DEPARTAMENTO DE ENGENHARIA CIVIL

Tel. +351-22-508 1901

Fax +351-22-508 1446

✉ miec@fe.up.pt

Editado por

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Rua Dr. Roberto Frias

4200-465 PORTO

Portugal

Tel. +351-22-508 1400

Fax +351-22-508 1440

✉ feup@fe.up.pt

http://www.fe.up.pt

À Nossa

If you get tired, learn to rest not to quit

Banksy

ACKNOWLEDGEMENTS

The path towards this dissertation has been labyrinthine, every day with new challenges. Fortunately, I have had the privilege of being guided by Engineer Dimitris Fotiadis, to whom I owe all the patience, wisdom and spectacular mentoring skills and without whom this thesis would not pass the form of a mere vision of an enthusiastic student. All my gratitude goes to you. My sincere gratefulness to Professor Ole Hededal for making this fantastic opportunity possible. Another word of appreciation goes to my co-supervisor Professor Sara Rios, for her handy suggestions on how to improve this work, giving a significant contribution to shaping it in the final product. I would like to thank to Professor José Couto Marques for trusting me and allowing me the opportunity to develop the dissertation abroad. I cannot also disregard the opportunity to acknowledge the care of José Calejo, for his friendship and for always being ready to help me along the journey.

My appreciation also extends to the staff of COWI A/S. To their host spirit and inclusion notwithstanding all the adverse conditions. For sharing with me their work-ethic, the know-how exchange, forming a unique work environment.

I wish to acknowledge the support of my friends, who were always there to fill in the pauses with laughter and companionship. To Daniel, for being my side by side comrade in this struggling ride. To all the friends I made throughout life.

The final word goes to my family, mother and brother, to whom I cannot thank enough. They have put up with me in the most stressful situations, supporting me and rooting for me even when I could not be as present as I would wish. You are the best. Thank you for remembering me that I am loved. Behind every great play, there is a whole set of fundamental people that one can easily forget, and this dissertation is no exception. Thank you all for the fantastic support and for always believing this would be possible.

Ivan Almeida

June 2020

ABSTRACT

Reclamation projects are of increasing importance at present due to many reasons, from the rise of world population to environmental concerns. Optimization and innovation in project design procedures are a daily basis in Engineers life. Applications of probabilistic methods in geotechnical engineering have increased remarkably in recent years, in order to complement the observational method that has been used for many years up to the present. Uncertainty and lack of information is a routine in Geotechnical Engineering, which creates a challenge to Geotechnical Engineers. Probabilistic methods in the design can be a helpful tool to complement current design methodologies in order to quantify the level of risk involved. The probabilistic methods are required when combining random variables is essential, which cannot be managed over analytical methods.

The most critical factor in reclamation design and construction is the estimation of the expected settlements over time, which includes total settlements and the rate at which they will develop. Statistically defining the soil parameters that support soil settlements predictions is a comprehensive answer to understand how conservative the design can be made. This study is focused on probabilistic approaches in land reclamation projects. A practical case was studied, a port expansion, where the adopted soil improvement technique was preloading with prefabricated vertical drains.

In this work, the variables numerical modelling process takes in concern simulation methods, such as Monte Carlo simulation and Latin Hypercube sampling in order to confront them to find the appropriate method to implement.

The vital objectives of the present study rely on the definition of soil parameters with the best probability distribution, based on the measured values obtained in the field and further implementation of the most efficient simulation method, Latin Hypercube sampling with 1000 samples. The soil parameters appropriate to the numerical modelling in this case study rely on the coefficient of consolidation and the compression index due to their importance to characterize the rate and magnitude of settlement in soils. These procedures gather the requirements needed to obtain the remaining settlements with their corresponding probability to occur. The definition of the data probability density function with EasyFit and the adopted tests allowed to fit the best-detailed distribution for each variable, which in this particular case correspond to the Johnson SB and Dagum distributions for the soil parameters, coefficient of consolidation and the compression index, respectively. These procedures also provide a starting point on the implementation of statistical approaches in the design, which can be optimized to use in real-time processing.

KEYWORDS: land reclamation, soil improvement, probability distributions, statistical approaches, simulation methods, Monte Carlo simulation, Latin Hypercube sampling, probabilistic design, numerical modelling, settlement analysis

RESUMO

Nos dias que correm, os projetos de reclamação de áreas são cada vez mais importantes devido a inúmeras razões, desde o aumento da população mundial até às preocupações ambientais. A otimização e inovação nos métodos de desenvolvimento de projetos são regulares na vida de um Engenheiro. As aplicações de métodos probabilísticos em engenharia geotécnica aumentaram notavelmente nos últimos anos, de modo a complementar o método observacional que tem sido utilizado durante muitos anos até ao presente. A incerteza e a falta de informação são uma rotina em Engenharia Geotécnica, criando assim um desafio aos Engenheiros Geotécnicos. Os métodos probabilísticos no desenvolvimento de projetos podem ser uma ferramenta útil para complementar as metodologias de projeto atuais de forma a quantificar o nível de risco envolvido. Os métodos probabilísticos são necessários quando se combinam variáveis aleatórias, que não podem ser resolvidas por métodos analíticos.

O fator mais crítico na conceção e construção de projetos de reclamação de solos é a estimativa do valor dos assentamentos do solo que ocorrem ao longo do tempo, que inclui o total de assentamentos e a taxa a que estes se desenvolvem. Definir estatisticamente os parâmetros do solo que suportam as previsões de assentamentos dos solos é uma resposta abrangente para entender o quão conservador o projeto em causa é desenvolvido. Este estudo centra-se em abordagens probabilísticas em projetos de reclamação de terras. Aplicada a um caso prático, uma expansão portuária, onde as técnicas de melhoria do solo adotadas são a pré-carga com drenos verticais pré-fabricados.

Neste trabalho, o processo de modelação numérica de variáveis tem em conta métodos de simulação, tais como, "Monte Carlo simulation" e "Latin Hypercube sampling", a fim de as comparar, de forma a entender o método adequado a implementar.

Os principais objetivos do presente estudo baseiam-se na definição dos parâmetros do solo com a correspondente distribuição de probabilidade com base nos valores medidos obtidos por ensaios de campo e na implementação posterior do método de simulação mais eficiente, "Latin Hypercube sampling" com 1000 amostras. Os parâmetros do solo adequados à modelação numérica neste caso de estudo baseiam-se no coeficiente de consolidação e no índice de compressão devido à sua importância para caracterizar a taxa e a magnitude dos assentamentos nos solos. Estes procedimentos reúnem os requisitos necessários para obter os restantes assentamentos com a sua correspondente probabilidade de ocorrência. A definição da função de densidade de probabilidade de parâmetros com o recurso ao software EasyFit, e os testes adotados permite interpretar a melhor distribuição de probabilidade a atribuir a cada variável, que neste caso em particular corresponde às distribuições de probabilidade, distribuição Johnson SB e distribuição Dagum para os parâmetros do solo, coeficiente de consolidação e índice de compressão, respetivamente. Estes procedimentos fornecem também um ponto de partida para a implementação de abordagens estatísticas no desenvolvimento de projetos, que podem ser otimizadas de forma a serem implementadas no seu desenvolvimento ao longo de todo o processo construtivo.

PALAVRAS-CHAVE: reclamação de áreas, melhoria de solo, distribuição de probabilidades, abordagens estatísticas, métodos de simulação, Monte Carlo simulation, Latin Hypercube sampling, design probabilístico, modelação numérica, análise de assentamentos

ABSTRAKT

Projekter for udvidelse af grundareal vigtige i forbindelse med voksende verdensbefolkning og af miljøhensyn. Optimering og innovation af projektering er en daglig opgave for ingeniørerne. Anvendelse af sandsynlighedsmetoder i geoteknik har for nylig haft en betydelig udvikling. Usikkerhed og manglende information er en del af geoteknik, som skaber udfordringer for geotekniker. Sandsynlighedsmetoder er et hjælpsomt værktøj for at kvantificere risici i projektering. Disse metoder er væsentlige for at kombinere aleatoriske variabler, som ikke kan håndteres ved brug af analytiske metoder.

Den kritiske faktor i designet og udførelse af grundareal udvidelse projekter er estimering af sætninger langs tiden, inkluderende totale sætninger og sætningsudvikling. Ved at definere statistiske jordparametre for sætningsvurdering giver en omfattende oversigt af designs konservatisme. Dette projekt fokuserer på sandsynligheds metoder anvendt i grundareal udvidelse projekter.

I dette projekt, arbejdsmetoden for variablers numeriske modellering tager i betragtning simulering metoder, som Monte Carlo simulering og Latin Hypercube, som sammenlignes for at finde den meste passende metode.

Undersøgelsens formål er bestemmelsen af jordparametre med den bedste sandsynlighedsfordeling, baseret på værdier målt på marken og med implementering af den meste simuleringsmetode, Latin Hypercube med 1000 prøver. Jordparametre anvendt i den numeriske modellering er konsolideringsfaktor og dekadehældning, da disse er vigtige for at beregne sætningsudvikling og størrelsesorden. Disse procedurer indeholder de krav for at få de resterende sætninger og tilsvarende forekomst sandsynlighed. Data for sandsynlighedstæthedsfunktion er defineret med EasyFit og tilpassende tester angiver den bedste detaljeret fordeling for hvert variabel, som i denne undersøgelse svarer til Johnson SB og Dagum fordeling for henholdsvis konsolideringsfaktor og dekades hældning. Disse metoder giver et startpunkt for implementering af statistiske tilgang i projektering, som kan bruges for optimering i real tid behandling.

NØGLEORD: Grundareal udvidelse, jordstabilisering, sandsynlighedsfordeling, statistiske tilgang, simulering metoder, Monte Carlo simulering, Latin Hypercube metode, statistisk design, numerisk modellering, sætningsanalyser

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

In this section, a list of symbols, abbreviations and acronyms used in the realisation of this dissertations is presented. Each symbol is usually associated with a specific subject, and it is easy to distinguish its meaning, given the context in which it is presented.

Latin Alphabet

$A \cap B$	Intersection of event A and event B
$A \cup B$	Reunion of event A and event B
c	Cohesion
C_c	Compressibility index
C_{ec}	Compression index
c_h	Horizontal coefficient of consolidation
C_i	Consequences of each event
C_k	Permeability index
C_r	Recompression index
c_v	Consolidation coefficient
D	Kolmogorov-Smirnov statistic
d_e	Equivalent diameter of a unit PVD influence zone
d_s	Equivalent diameter of the disturbed zone
d_w	Equivalent diameter of PVD
d_{sz}	Equivalent diameter of the smear zone
E_i	Range of events
E_{oed}	Oedometric modulus
F	Safety factor
F_d	Effect of drain spacing, soil disturbance and well resistance
G	Limit state
H	The more significant distance that water needs to travel to a direct drained boundary
h_{dr}	Maximum vertical drain path
h_{sur}	Surcharge height
i	Hydraulic Gradient
J	Percolation Force
k	Soil permeability coefficient

k_h	Horizontal permeability of the soil
k_r	Smear zone permeability
k_u	Permeability in the undisturbed zone
k_v	Vertical permeability coefficient
k_{hi}	Initial horizontal permeability of the soil
L	Length for one-way drainage
l_m	PVD maximum discharge length
M	Margin of safety
m_v	Volumetric compressibility
ø'	Internal Friction Angle
P_f	Probability of Failure
P_i	Probability of each event to occur
P(A)	Probability that A is true.
P(A|B)	Probability of event A given event B
Q	Load
qc	Cone penetration resistance
q_w	Discharge capacity of PVD
Q_1	25th Percentile
Q_3	75th Percentile
R	Resistance
r	Range of a set of data
r_e	Radius of the influence zone of a PVD
r_s	Smear zone radius
r_t	Transition zone radius
Rf	Friction ratio
S	Area of the sample
s_{target}	Target for primary consolidation settlement
S_∞	Predicted long-term primary consolidation settlement
T	Time factor
t	Time
T_v	The dimension time factor
t_{max}	Preloading time
u_e	Excess pore pressure

U_h	Horizontal consolidation
U_v	Vertical consolidation
v	Flow-velocity
$x_{0,5}$	Midpoint of a sample
(x_1, x_2)	Correlated variables
(y_1, y_2)	Gaussian random variables
w_l	Liquid Limit
w_p	Plastic Limit
w_s	Shrinkage Limit
Z	Monitoring of settlements information data
z	Soil Depth
Z_d	Depth factor

Greek Alphabet

$\bar{A}$	Complement of event A
α	Significance level
β	Reliability index
$\Delta\sigma'_v$	Additional effective vertical stress
γ_w	Water Impulse
μ	Mean
ρ	Correlation coefficient
ρ_c	Consolidation settlement
ρ_i	Immediate settlement
ρ_s	Secondary compression settlement.
ρ_t	Total soil settlement
σ	Standard deviation
σ'_p	Pre-consolidation stress tension
σ'_{v0}	Initial vertical stress
σ^2	Variance
Φ	Cumulative standard normal distribution function
Φ^{-1}	Inverse cumulative standard normal distribution function

Acronyms

®	Trademark
A^2	Anderson-Darling test
CDF	Cumulative density function
COV	Coefficient of variation
CPT	Cone penetration test
Diff	The probability difference graph
ECDF	Empirical cumulative distribution function
FEUP	Faculty of Engineering University of Porto
Georef	Geotechnical engineering boundary value
GIS	Geographic information systems
GOF	Goodness of fit test
ICDF	Inverse cumulative distribution function
IQR	Interquartile range
InSAR	Differential interferometric synthetic aperture radar
LHS	Latin Hypercube Sampling
MC	Monte Carlo Simulation
OCR	Overconsolidation ratio
PDF	Probability density function
PMF	Probability mass function
P-P plot	Probability-Probability plot
PVC	Polyvinyl chloride
PVD	Prefabricated Vertical Drains
Q-Q plot	Quantile-Quantile plot
RF	Random field
RMS	Root-mean-square
SAA	ShapeArray
SPT	Standard penetration test
SRV	Single random variable

1 INTRODUCTION

1.1 BACKGROUND

Reclamation projects are increasingly important around the world. The main reasons are the lack of available space on land in densely populated coastal areas, climate change protection and waterfront development. The most critical factor in reclamation design and construction is the estimation of the expected settlements over time, which includes total settlements and the rate at which they will develop. These factors have a significant impact on the construction cost, on the time-schedule of construction as well as on the earliest possible operational start of the completed project. The settlements depend on the processes of consolidation and creep, for which several analytical and numerical methods of analysis exist that may be used in the design and planning phase. During construction, a geotechnical monitoring program may be implemented, which monitors on-going settlements and thereby allows for prediction of the final settlements and their time development. Currently, the geotechnical analysis and design of reclamation projects are mostly done assuming some representative input soil parameters together with an engineer's estimate of a reasonable range within which they can vary. Based on this input, the analysis models supply a range of expected outcomes, which are then evaluated, and appropriate design modifications are made. However, this approach has the following drawbacks:

a) The selected soil parameters input range is subjective and does not represent a realistic statistical distribution based on geotechnical survey data.
b) There is not an assigned probability to the different analysis outcomes, which is misleading about the importance of the different outcomes.

1.2 CONTEXT AND OBJECTIVES

Applications of probabilistic methods in geotechnical engineering have increased remarkably in recent years. Geotechnical engineers and geologists deal with materials whose properties and spatial distribution are poorly known and with problems in which loads and resistances are often coupled. Thus, a somewhat different philosophical approach is necessary. Historically, the geotechnical profession has dealt with uncertainty on essential projects by using the observational approach; this is entirely compatible with reliability-based methods (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). In order to quantify the uncertainties of the parameters, statistical approaches are going to be more common in the future. The probabilistic methods are required when combining random variables is essential, which cannot be managed using analytical methods. That way, the expected results of this work can be divided into the following points:

- Compilation of a rational literature review of the relevant methods;
- Knowledge acquisition of consolidation and creep settlements;
- Statistics familiarity applied to geotechnical engineering;
- Awareness about simulation methods;
- Statistical description of soil properties;
- Parameters numerical modelling;
- Probabilistic settlement analysis interpretation;
- Evaluation of different probabilistic methods and conclusions on their strengths and limitations for settlement analysis in reclamation design.

The proposed study will have three goals:

1) Determine appropriate probabilistic methods which address the design input uncertainty.
2) Distribution fit criteria and requirements demanded.
3) Settlement analysis based on probability distributions.

1.3 DISSERTATION OUTLINE

The present dissertation is structured into seven chapters.

The present chapter, chapter 1, contains a general outline of the thesis scope. It comprehends a brief Background of reclamation projects and its importance, covering the Context and Objectives of this dissertation.

Chapter 2 and 3, are focused in a literature review divided into main subjects focused on the main topics of this work. Chapter 2 begins with the mathematically necessary reviews of Probability Statistics. It comprises the concept of Uncertainty and Risk in Geotechnical Engineering, as well as the management of risk assessment. It also presents the Statistical Approaches considered in the study.

Chapter 3 starts with a presentation of the main issues related to Reclamation of Coastal Areas, and Soil Improvement methods. Geotechnical approaches for Settlement Analysis section allows the reader to comprehend the theoretical engineering background needed to interpret the procedures implemented. The final subject, Construction Monitoring, allows a perception of the bases of instrumentation and monitoring used in real case studies.

Subsequently, in chapter 4, the Case Study of this dissertation starts to be introduced, describing the main features of this port expansion project. It includes the description of Site Investigation, field and laboratory tests made. The Geotechnical Conditions present on the construction field leading to the Geotechnical Parameters Evaluation. A brief explanation is introduced to how the land reclamation constructions projects are performed. Lastly, the Monitoring Program adopted in this specific project is presented.

Chapter 5 is focused were the Numerical modelling towards design parameters contemplating all the followed steps. The preliminary approaches made throughout a Statistical Description of Soil Properties gave the necessary background to improve the Parameters Numerical Modelling. Combining the approaches done, with data analysis and simulation application, EasyFit, it was possible to obtain the final Modelling Variables. In order to make some improvements to the modelling variables, Modelling Optimization was implemented, taking into account two different simulation methods. Throughout the examination of the Simulation Methods Distinctions, it was possible to identify the best method to implement in the modelling procedures, so the final and Adopted Modelling Variables is accomplished.

Chapter 6 is reserved for the presentation of the final settlement results obtained. The outcomes result from considering the modelling parameters evaluated in the 5th chapter. The drains and soil characteristics used in the settlement design are stated. A discussion of the analysis and the Remaining Settlements results obtained is also presented.

Chapter 7 summarizes the main outputs as well as conclusions obtained. There are many other methods and different concepts that can be considered to improve the obtained results. In the section Future Developments Perspectives, some research subjects that can be further introduced are presented and discussed, where they represent the continuation of the covered topics.

2
Statistics Applied to Geotechnical Engineering

2.1 PROBABILITY STATISTICS

2.1.1 VARIABLES DEFINITIONS

In statistical studies, there are two types of random variables whose definitions it is essential to distinguish (Oliveira, 2010):

1. Discrete variables – their possible values form a finite range of values;
2. Continuous variables – their values do not constitute a limited range of values, it consists of a variety of values.

A histogram is the most typical form to represent continuous data graphically, as shown in Figure 2.1. Histograms are visually similar to bar charts, but this type of graphical display is used to represent descriptive data. The histogram splits data into fixed intervals, called bins. Composed by different height bars, each bin of a histogram contains the number of occurrences in the data set that is contained within that bin. The width of each bin is arbitrary but must be uniform and visually comfortable. In a histogram, it is the area of the bar that express how often each bin occurs (Oliveira, 2010). Due to that, it is possible to create a frequency distribution by dividing each vertical bar by the total number of measurements. This gives the relative frequency of observed values in each interval as a decimal fraction. The cumulative frequency distribution represented by a solid curve in Figure 2.1, is the sum of all frequencies per bin, existing in a frequency distribution.

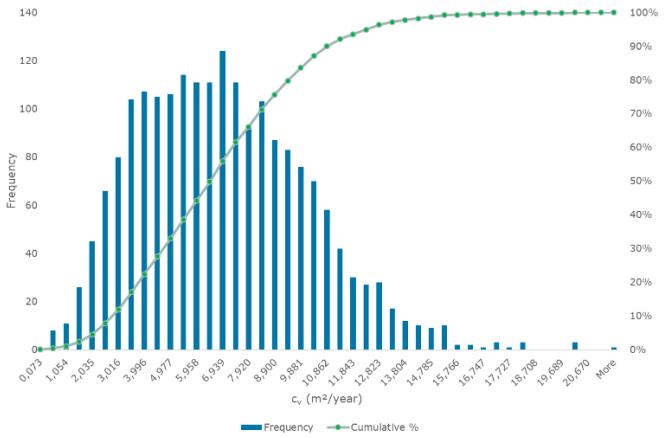

Figure 2.1 Example of a histogram

The histogram also gives visually notions about summary statistics like central tendency, the dispersion, scale variation, and if there is a point of concentration, it gives the mode value. It also shows if the variable is symmetric or asymmetric.

2.1.2 DESCRIPTIVE STATISTICS

Descriptive statistics provides information about the data in our data set. Due to that, it is possible by just visualising the data to see the differences between them. There are two main types of statistics evaluation (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003):

- Central tendency;
- Dispersion.

Central tendency measures are classically known as the mean, median, and mode (Oliveira, 2010). The mean is the arithmetic average of a set of data usually denoted as μ. To get the mean of a data set, sum up all the numbers in the data set, and then divide that total by the number of digits in the data set. The median is the number which splits the data set into two equal intervals. The median represents the midpoint number of a set of ordered data, usually denoted as $x_{0.5}$. The mode is the most common value that occurs in our data set, or in other words, the value that has a higher frequency.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.1}$$

$$F_x(x_{0.5}) = 0.5 \tag{2.2}$$

Conventional measures of dispersion are the standard deviation, range, and inner quartiles of the frequency distribution (Oliveira, 2010).

The variance, typically denoted as σ^2 in statistics, is a measurement of the spread between numbers in a data set. Variance measures how distant each number in the set is from the mean.

$$\mathrm{Var(x)} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2.3}$$

The standard deviation is a statistical parameter that measures the dispersion of a dataset relative to its mean and his calculation as the square root of the variance. If the data points are further from the mean, there is a higher deviation within the data set; thus, the more spread out the data, the higher the standard deviation. The calculation is represented as the square root of the variance.

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \tag{2.4}$$

The coefficient of variation (COV) of a data set is defined as the standard deviation divided by the mean. When needed to compare different samples from different studies, the sample with the higher value of cov means that the sample has more variation in comparison to its mean.

$$Cov_x = \frac{\sigma_x}{\bar{X}} \tag{2.5}$$

The range of a set of data denoted *r*, is the difference between the largest and smallest values,

$$r_x = |x_{max} - x_{min}| \tag{2.6}$$

An interquartile range is a measure of where the majority of the values are.

$$Q_1(x_{0.25}) = 0.25 \tag{2.7}$$

$$Q_3(x_{0.75}) = 0.75 \tag{2.8}$$

An interquartile range is simple to understand with the box-plot of Figure 2.2. A box plot is a graphical method to recap a data set by visualising the minimum value, the 25th percentile (Q_1), median, mean, 75th percentile (Q_3), the maximum value, and potential outliers.

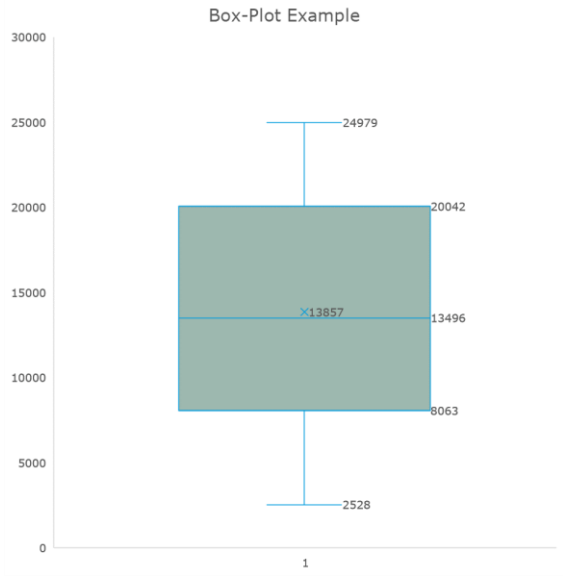

Figure 2.2 How to read a box-plot

2.1.3 PROBABILITY THEORY

Probability theory is a branch of mathematics linked by the analysis of a random event. The probability is the chance of an event to happen in a random experiment (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). In order to quantify the probability of an event, its value is between zero and one, where zero means that the event will not occur and one means that it is certain the occurrence of that event, or in percentage 0% chance to happen and 100% sure it will happen. The outcomes of a random experiment are denominated in a sample space, denoted by S in Figure 2.3, where all the possible values to occur are contained. While with sample points, we are in handling with discrete variables as illustrated in Figure 2.4. By that, it is possible to accomplish that the sum of the corresponding probabilities of each set is 1, or 100% chance that will occur. When we have two independent events (A and B), as in Figure 2.3, the probability that one of them occurs, does not affect the likelihood of the other to occur simultaneously. The product between them indicates the possibility of both events to happen.

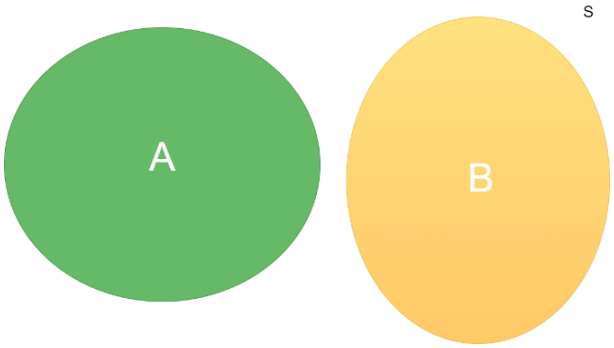

Figure 2.3 Independent events in S sample space

2.1.3.1 Random Events

The mathematical theory of probability deals with experiments and their outcomes. An experiment is a random process generating specific and a priori unknown results or outcomes (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). In Figure 2.4, there are 90 discrete sample points of a sample space S. There are two different and independent events denominated as event A and event B. Event A holds 12 sample points, and event B holds 6 sample points within S. The complement of A is denoted as $\bar{A}$, which contains all the other sample points that are not part of event A, thus 78 sample points. The reunion of 2 different events (A and B) is the collection of sample points contained in both events mathematically symbolised as $A \cup B$ which result in an event shaped by the sum of both sample points, in this specific case 18 sample points. On the other hand, the intersection between two events mathematically symbolised as $A \cap B$ is the collection of sample points contained in both A and B, exemplified in Figure 2.5, by the Venn diagram. The term mutually exclusive means the events in a study cannot occur at the same time, like in Figure 2.4. Though collectively exhaustive means that the events combined represent everything that can happen, which in probability theory such events probabilities sum to 1, in percentage, 100%.

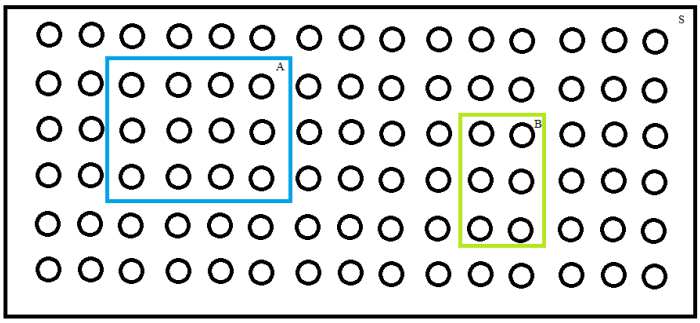

Figure 2.4 Sample space showing sample points and an event A and event B

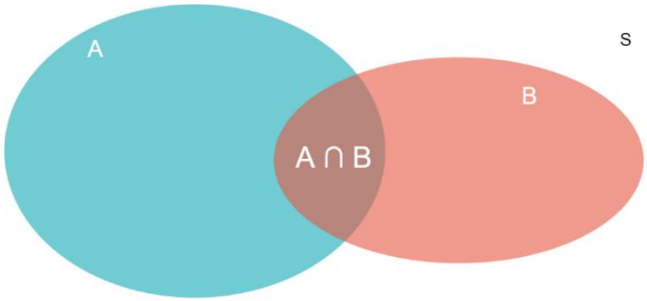

Figure 2.5 Venn diagram of events A and B in sample space S

2.1.3.2 Conditional Probability, Dependence, and Independence

Conditional probability enables to understand if an event can be changed by the assumption or evidence from an additional event that already occurred. For example, the probability assigned to high values of clay undrained shear strength is presumably changing by knowing the pre-consolidation pressure of such fine soils. From empirical observation and theory, high values of undrained strength have an association with high pre-consolidation demands, and conversely. Thus, the probability of high undrained strength should rise if we learn that the pre-consolidation pressure is high (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). Mathematically, the likelihood of an event A given event B is usually represented by P(A|B).

$$\mathrm{P(A|B)} = \frac{\mathrm{P(A \cap B)}}{P(B)} \tag{2.9}$$

As shown before, when two independent events occur, the product between them indicates the possibility of both events to happen. Adding the concept of conditional probability (Figure 2.5), it is possible to manage the next equation:

$$\mathrm{P(A\ and\ B)} = \mathrm{P(A)} * \mathrm{P(B|A)} = \mathrm{P(B)} * \mathrm{P(A|B)} \tag{2.10}$$

2.1.4 RANDOM VARIABLES

Once there are two different types of variables, there is a mathematical function to define either of them, named probability distribution. These functions designate the probability over a sample space of obtaining the probable values that a random variable can assume, that is, their outcomes.

2.1.4.1 Probability Mass Functions

Whenever our sample space is discrete, the discrete probability distribution, p(x) is named as probability mass function (PMF). The probability that x can take a specific value is referred to p(x).

$$P(X = x) = p(x) \tag{2.11}$$

Where p(x) cannot be a negative value for all real x, once the outcome is the sum of the PMF over the sample space, the final result must be 1. Figure 2.6 is representing the probability of each number on a dice, visually representing a probability mass function.

$$\sum_{i=1}^{n} \mathrm{P}(x_i) \;=\; 1 \tag{2.12}$$

$$x_i > 0 \tag{2.13}$$

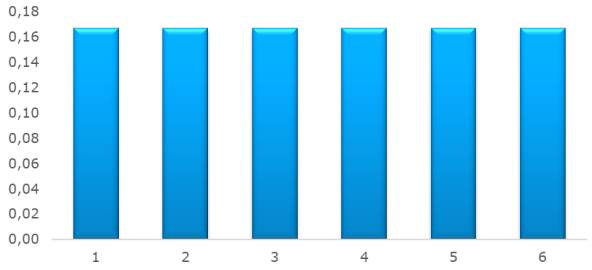

Figure 2.6 Probability mass functions dice probability

2.1.4.2 Probability Density Functions

Once there are many values for the outcome, the mathematical definition is based on a density function, named as probability density function (PDF), exemplified in Figure 2.7. Similarly, to PMF, this function needs to respect some conditions. In Figure 2.1, PDF is noticeably defined as the area between two points. Due to that, the probability of x to be between two points a and b, is P(x), analytically expressed:

$$P(a \leq x \leq b) = \int_a^b f(x)dx \tag{2.14}$$

Once this is a probability approach, the sum of p(x) over all possible values of x is 1.

$$\int_{-\infty}^{+\infty} fx\,(x_i)dx = 1 \tag{2.15}$$

$$x_i > 0 \tag{2.16}$$

The cumulative density function (CDF) shows the probability of an outcome of X to be more or less equal to a particular value. An example of the normal cumulative distribution function is visible in Figure 2.8.

$$P(x \leq x_i) = \int_{-\infty}^{x_i} f_X(x)dx \tag{2.17}$$

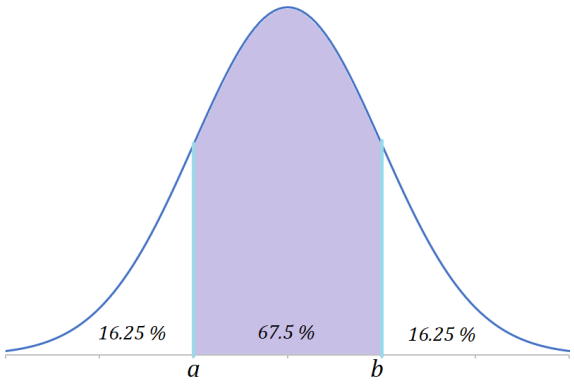

Figure 2.7 Normal probability density function example

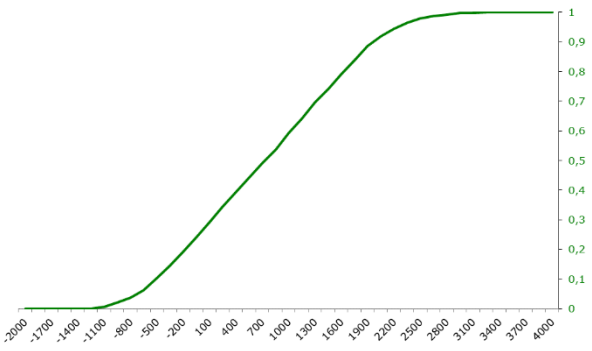

Figure 2.8 Normal cumulative distribution function example

2.1.5 PROBABILITY DISTRIBUTIONS

A statistical distribution is a practical tool to measure uncertainty, for discrete or continuous variables. A probability distribution is a mathematical function which gives the probabilities of something to occur in a range of numerous outcomes. When talking about distributions, an important parameter to be able to define the distribution in the study are the measures of shape. Firstly, the definitions of measures of shapes are presented. Secondly, the distributions which represent the continuous variables in the study are discussed.

2.1.5.1 Measures of Shape

There are two different measures of shape (Jain, 2018):

- Skewness;
- Kurtosis.

Skewness measures the lack of symmetry in the data distribution. It is very convenient to understand and define normal distributions, and to evaluate if a descriptive data corresponds to a normal distribution or not. Skewness represents in which tail the bulk of values are. It allows having insight if the values are either above or below the mean. There are two types of skewness, positive and negative, visible in Figure 2.9. When skewness is zero, this means the distribution is symmetrical, equivalent to normal distribution.

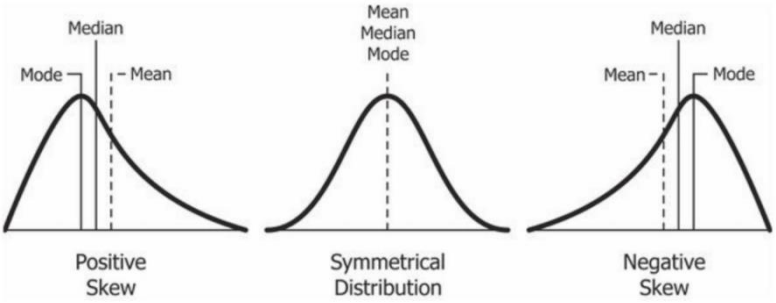

Figure 2.9 Skewness of frequency distribution (Jain, 2018)

Kurtosis is related with the tails of the distribution. It describes the extreme values in either tail. It provides an understanding of the presence of outliers in the case study. Outliers are due to many reasons, even human error by introducing wrong data values. The lower the kurtosis indicates that data has light tails and outstanding to that, no outliers. On the other hand, high values of kurtosis show that the data has outliers. Kurtosis can be defined in three types of different shapes, visible in Figure 2.10.

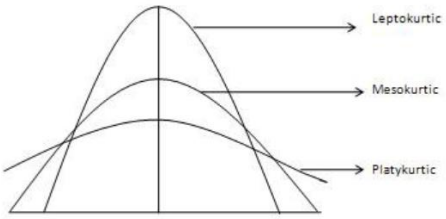

Figure 2.10 Kurtosis of frequency distribution (Jain, 2018)

Kurtosis standard value for a normal distribution is three. The most similar to the normal distribution is the middle one, mesokurtic, the one with the value three. Leptokurtic has a value higher than three, meaning that the distribution shape is with more massive tails, and the value of peak is elevated. Having distributions with these characteristics means that the data set has outliers. Platykurtic have kurtosis values smaller than three and have different shape and definition than Leptokurtic, which means, lower peak, shorter distribution, and light-tailed distribution resulting in no outliers.

2.1.5.2 Uniform Continuous Distribution

This is the most straightforward continuous distribution, continuous in a range [a, b]. In Figure 2.11, the density functions are represented. Series 1 is in the range of [-5,4], and series 2 [-7,6]. This distribution has a constant probability, and its definition is:

$$\mathrm{f}(x) = \begin{cases} \dfrac{1}{b-a}\,, a < x < b \\ 0\,, otherwise \end{cases} \tag{2.18}$$

Abbreviated by:

$$X \sim \mathrm{U}(a, b) \tag{2.19}$$

Where the variable X is a discrete uniform distribution with the parameters a and b.

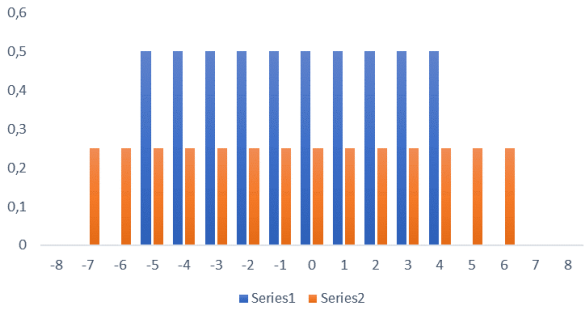

Figure 2.11 Uniform continuous distribution

Limiting a = 0 and b = 1, the resultant uniform continuous distribution U(0,1) is named a Standard Uniform Distribution.

2.1.5.3 Normal Distribution

The normal distribution is a probability function that describes how the values of a variable are distributed. It is a symmetric distribution where most of the observations cluster around the central peak and the probabilities for values further away from the mean taper off equally in both directions (Frost, 2020). A normal distribution is a continuous distribution with the shape of a bell curve, visible in Figure 2.12 that can assume different types of flatness or positions in the horizontal axis. A normal distribution has skewness values close to zero and kurtosis values close to three. With the same values is represented in Figure 2.8, the normal cumulative distribution function.

If a random variable *X* follows a normal distribution, it is characterised by two parameters, μ and σ^2, which represent the mean and variance. The mean defines the peak of the case in the study. Through the visualization of different graphs, if the mean value shifts, the curve on the horizontal axis also shifts either right or left.

The normal distribution can be defined by the following equation:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} * e^{(-\frac{(x-\mu)^2}{2\sigma^2})} \tag{2.20}$$

Abbreviated by:

$$X \sim \mathcal{N}(\mu, \sigma^2) \tag{2.21}$$

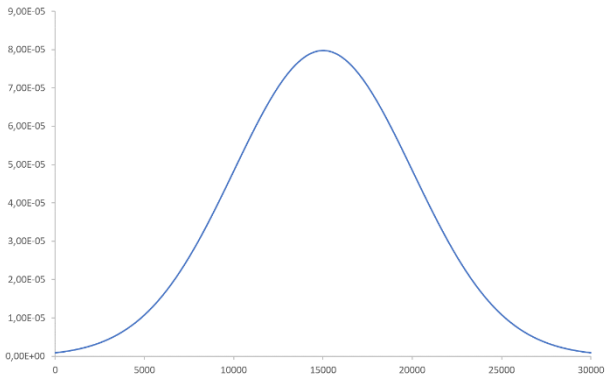

Figure 2.12 Normal distribution

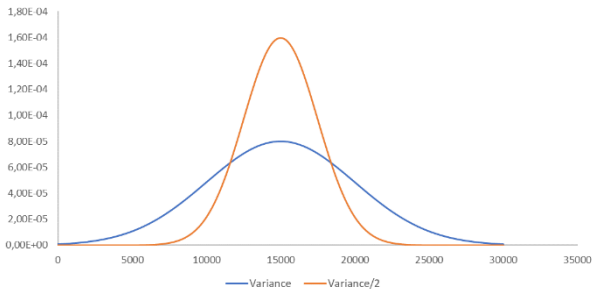

Figure 2.13 Normal distribution with different variance

Normal distributions have some distinctive properties (Frost, 2020):

- The normal distribution is symmetric at the centre;
- The mean, median and mode have the same value;
- Half of the sample is minor then the mean and the other half is superior to the mean;
- The total area under the curve is 1;
- Linear combinations applied to random independent variables that follow the normal distribution result in a random variable with a normal distribution.

Thus, when the subject is a probability, the area under the curve means 100% probability to occur, where the left side of the curve has 50% chance to happen, and the other side also has 50% chance to occur.

Standardisation is a statistical method comprehended to study data provided by normal distributions with different means and variances, rearranged on a standard scale to be able to compare them (Frost, 2020). This method is used when the software is not available to use. The final formula is called Z-scores, which result in:

$$\mathrm{Z} = \frac{X - \mu}{\sigma} \tag{2.22}$$

Z-scores follows a normal distribution where the mean value results in 0 and the standard deviation in 1.

$$X \sim \mathcal{N}(0,1) \tag{2.23}$$

Finding areas under the Z-scores curve, Figure 2.14, it is possible to calculate by the Standard Normal distribution table.

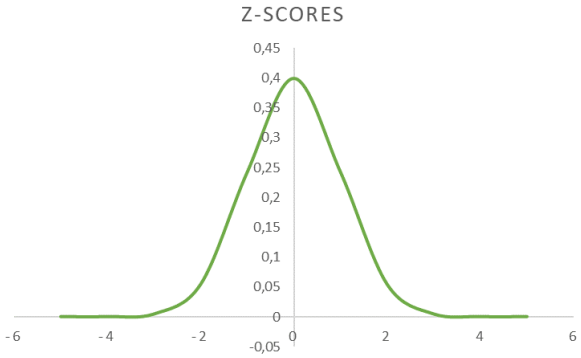

Figure 2.14 Standardization normal distribution

The standard normal distribution and its cumulative density function (CDF), are commonly used in the statistical world, represented by a Greek symbol for shorthand notation Φ. So, standard normal distribution CDF is commonly expressed as Φ(z) instead of F(z), analytically represented through the following expression:

$$\Phi(t) = \frac{1}{2\pi}\int_{-\infty}^{t} e^{-\frac{z^2}{2}}\,dz \tag{2.24}$$

Integration to the point *t* is processed to give the cumulative density function below *t* for the standard normal distribution.

2.1.5.4 Lognormal Distribution

The lognormal distribution is a continuous distribution in which the logarithm of a variable has a normal distribution. Lognormal distribution plays an essential role in probabilistic design because negative values of engineering phenomena are sometimes physically impossible. Typical uses of lognormal distribution are found in descriptions of fatigue failure, failure rates, and other phenomena involving an extensive range of data (Chang, 2015). A random variable is lognormally distributed if its logarithm is normally distributed.

The lognormal distribution of a random variable *X* with expected value μ_X and standard deviation σ_X is denoted $\mathrm{Ln}(\mu_X, \sigma_X)$ and is defined as:

$$F_x(x) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2}(\frac{\ln(x)-\mu_y}{\sigma_y})^2}, 0 < x < \propto \tag{2.25}$$

in which $f_X(x)$ is the PDF of the random variable *X*, and:

$$\sigma_y = \sqrt{\ln((\frac{\sigma_x}{\mu_x})^2 + 1)} \tag{2.26}$$

$$\mu_y = \ln(\mu_x) - \frac{1}{2}\sigma_y^2 \tag{2.27}$$

The last equations represent the standard deviation and mean value for the normal distribution variable $y=\ln(x)$. Lognormal distribution functions are shown in Figure 2.15. The cumulative distribution function of a lognormal distribution results in the next equation:

$$F_x(x) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2}(\frac{\ln(x)-\mu_y}{\sigma_y})^2}, 0 < x < \propto \tag{2.28}$$

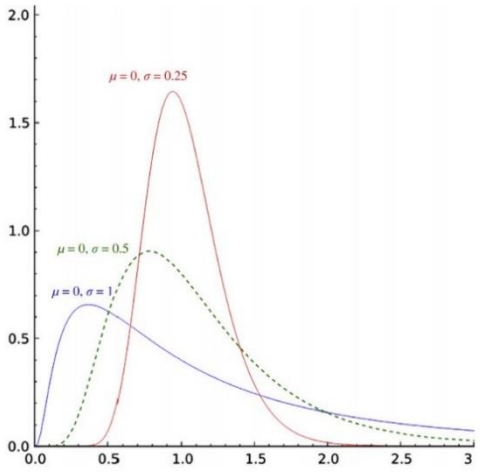

Figure 2.15 Lognormal distribution with different mean and standard deviations values (Chang, 2015)

The normal distribution and the lognormal distribution are unlike in different facts. The apparent difference is in its shape. Opposed to the normal distribution, lognormal is an asymmetrical distribution. The lognormal distribution is right-skewed once all values are positive.

2.1.5.5 Dagum Distribution

The Dagum distribution was proposed by Camilo Dagum, as an alternative model to the lognormal distribution since Dagum was not satisfied with how the lognormal distribution handles with heavy-tailed data (Glen, Dagum Distribution: Definition, CDF & PDF, 2020). Dagum distribution often accomplishes improved results when applied to empirical data.

The domain of Dagum distribution is truncated from below by a continuous location parameter γ. The four-parameter Dagum distribution PDF (Figure 2.16) and CDF is analytically defined as (Continuous Distributions, u.d.):

$$\mathrm{f}(x) = \frac{\alpha k (\frac{x-\gamma}{\beta})^{\alpha k-1}}{\beta(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha})^{k+1}} \tag{2.29}$$

$$\mathrm{F}(x) = \left(1+\left(\frac{x-\gamma}{\beta}\right)^{-\alpha}\right)^{-k} \tag{2.30}$$

where:

k - continuous shape parameter (k> 0);

α - continuous shape parameter (α> 0);

β - continuous scale parameter (β> 0);

γ – continuous location parameter.

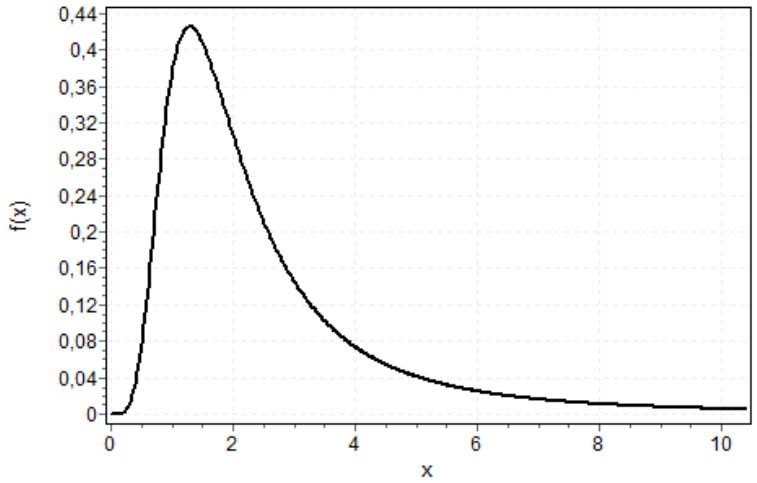

Figure 2.16 Dagum distribution

2.1.5.6 Beta Distribution

Beta distributions have two free parameters α_1 and α_2 (Weisstein, u.d.).

Beta distribution PDF (Figure 2.17) and CDF is analytically represented by the following equations (Continuous Distributions, u.d.):

$$f(x) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1 - 1}(b-x)^{\alpha_2 - 1}}{(b-a)^{\alpha_1 + \alpha_2 - 1}} \tag{2.31}$$

$$F(x) = I_z(\alpha_1, \alpha_2) \tag{2.32}$$

$$z \equiv \frac{x-a}{b-a} \tag{2.33}$$

where:

α_1 - continuous shape parameter ($\alpha_1 > 0$);

α_2 - continuous shape parameter ($\alpha_2 > 0$);

a, b – continuous boundary parameters (a> b);

I_z – regularized incomplete beta function.

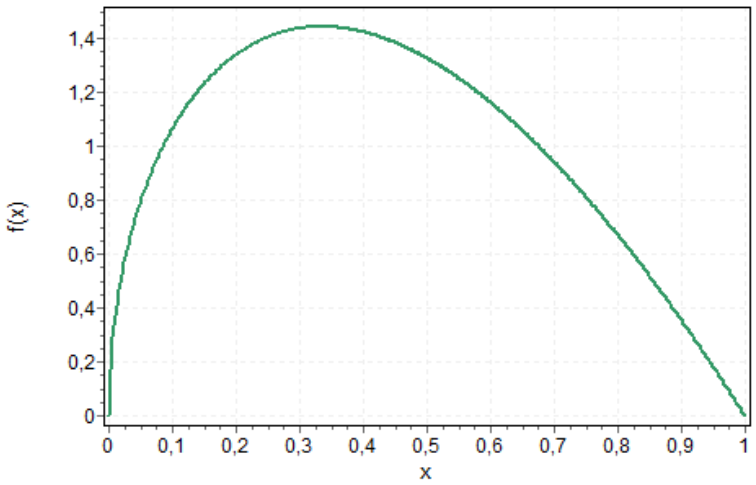

Figure 2.17 Beta distribution

2.1.5.7 Johnson SB Distribution

Johnson SB distribution is named due to Johnson system which contains only four distributions: the normal distribution, the lognormal distribution, the SB distribution (which models bounded distributions), and the SU distribution (which models unbounded distributions) (Wicklin, 2020).

Johnson SB distribution PDF (Figure 2.18) and CDF is analytically represented by the following equations (Continuous Distributions, u.d.):

$$f(x) = \frac{\delta}{\lambda\sqrt{2\pi}z(1-z)} \exp\left(-\frac{1}{2}\left(\gamma + \delta \ln\left(\frac{z}{1-z}\right)\right)^2\right) \tag{2.34}$$

$$\mathrm{F}(x) = \Phi(\gamma + \delta \ln\left(\frac{z}{1-z}\right)) \tag{2.35}$$

$$\mathrm{z} \equiv \frac{x-\zeta}{\lambda} \tag{2.36}$$

where:

γ – continuous scale parameter;

δ – continuous shape parameter (δ> 0);

λ – continuous scale parameter (λ> 0);

ζ – continuous location parameter;

Φ – Laplace Integral.

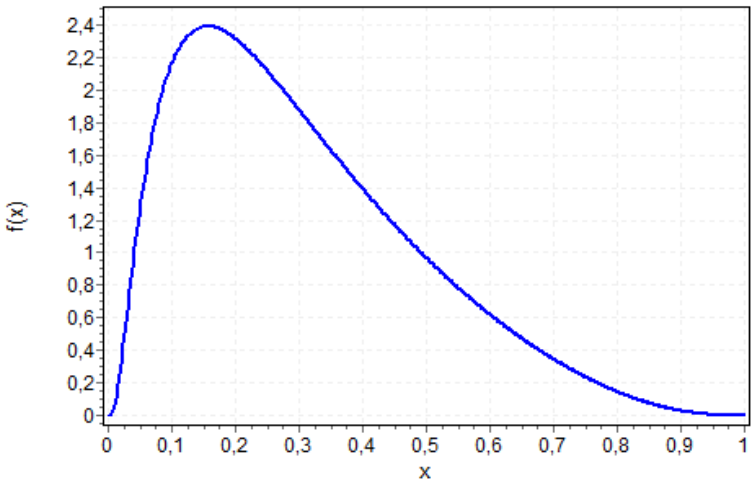

Figure 2.18 Johnson SB distribution

2.1.5.8 Generalized Extreme Value Distribution

The generalized extreme value distribution is often used to model the smallest or largest value among a broad set of independent, identically distributed random values representing measurements or observations. The generalized extreme value combines three simpler distributions into a single form, allowing a continuous range of possible shapes that includes all three of the simpler distributions (Generalized Extreme Value Distribution, u.d.).

Generalized extreme value distribution PDF (Figure 2.19) and CDF is analytically defined as (Continuous Distributions, u.d.):

$$\mathrm{f}(x) = \begin{cases} \frac{1}{\sigma}\exp\left(-(1+kz)^{\frac{-1}{k}}(1+kz)^{-1-\frac{1}{k}}\right), k \neq 0 \\ \frac{1}{\sigma}\exp(-z-\exp(-z)), k = 0 \end{cases} \tag{2.37}$$

$$F(x) = \begin{cases} \exp\left(-(1+kz)^{\frac{-1}{k}}\right), k \neq 0 \\ \exp(-\exp(-z)), k = 0 \end{cases} \quad (2.38)$$

$$z = \frac{x-\mu}{\sigma} \quad (2.39)$$

where:

k - continuous shape parameter;

σ - continuous scale parameter (σ> 0);

μ - continuous location parameter.

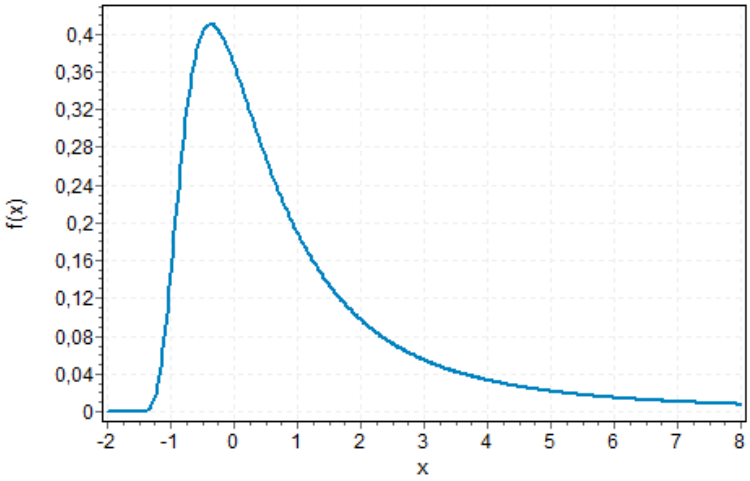

Figure 2.19 General extreme value distribution

2.1.5.9 Truncated Distribution

When applying probability statistics to the physical world, the range of values has to be restricted. A truncated distribution is a result of cutting off range from a distribution. Taking into account a specific domain of a distribution, x-values, there are several types of truncated distributions (Glen, Truncated Distribution / Truncated Normal Distribution, 2016):

- Truncated from above: high values of the domain are cut off, so the range is from negative infinity to some maximum value x [-∞, x_{max}];
- Truncated from below: low values of x are cut off, so the range is from some minimum value of x to positive infinity [x_{min}, +∞];
- Double truncation: both the low values and x values are cut off {x_{min}, x_{max}};
- If values range from negative infinity to infinity {-∞, +∞}, there is no truncation.

Truncation happens when datasets have values that are outside of a usual range.

When the percentage of cutter range is relatively low, it is possible to assume that the parameters that define the distribution remain equivalent.

2.1.6 PROBABILITY INTEGRAL TRANSFORM

In probability theory, the probability integral transform, also known as the universality of the uniform, relates to the result that data values that are modelled as being random variables from any given continuous distribution can be converted to random variables having a standard uniform distribution (Dodge, 2006). Random variables can be generated using a CDF and the standard uniform distribution, named *X*, in this example case, so, $X \sim Unif\ (0,1)$. This theorem says that plugging *X* into the inverse CDF, a new random variable will be distributed according to the original CDF.

Assuming that a random variable *X* follows a continuous distribution, whose CDF is F_X then a random variable *Y*, a standard uniform distribution is defined as:

$$Y = F_X(X) \tag{2.40}$$

By this definition, this method can be demonstrated through the next rearrangements (Dodge, 2006):

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ F_Y(y) &= P(F_X(X) \leq y) \\ F_Y(y) &= P(X \leq F_X^{-1}(y)) \\ F_Y(y) &= F_X(F_X^{-1}(y)) \\ F_Y(y) &= y \end{aligned} \tag{2.41}$$

Through this demonstration, it is possible to consider that F_Y is the CDF of a *Unif (0,1)* random variable. Therefore, *Y* has a uniform distribution on the interval [0,1].

This method is a careful process to create random numbers through any continuous distribution having their distribution properties in concern.

2.1.7 PROBABILITY PLOTS

Probability plots are used to have a graphical comparison of data defined through different distributions. These comparisons can involve three different circumstances (Lewinson, 2019):

- Two empirical sets;
- One empirical and one theoretical set;
- Two theoretical sets.

The comparison between one empirical and one theoretical set is the typical analysed plot. The empirical data corresponds to measured data obtained, and it is compared with a specific probability distribution

that corresponds to the theoretical set. The different type of plots that can be analysed is presented in the following section.

2.1.7.1 Probability-Probability Plot

The probability-probability plot, P-P plot, is a picturing that plots CDFs of two different distributions, the empirical CDF and the theoretical CDF, to be able to compare them visually. It is used to determine how well a specific distribution fits the observed data. This plot will be approximately linear if the specified theoretical distribution is the correct model. This graph can also be used to determine whether the data follow or not the theoretical distribution (Learn More About EasyFit, 2004-2010). In Figure 2.20, an example of a P-P plot is presented.

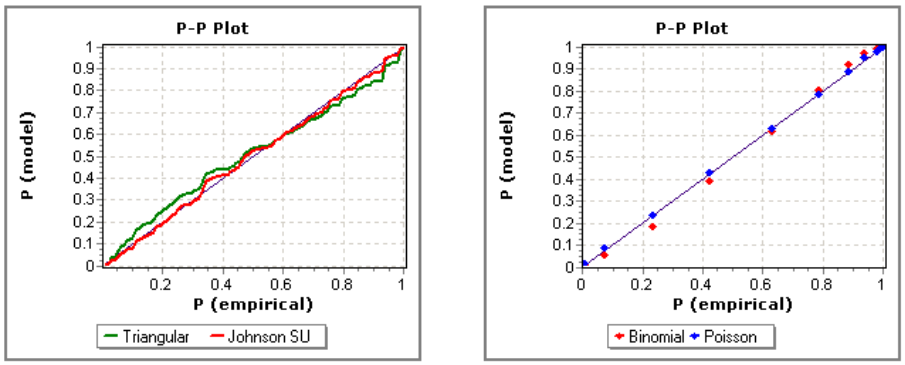

Figure 2.20 Probability-probability plot example (Learn More About EasyFit, 2004-2010)

Some leading information on P-P plots must be mentioned (Lewinson, 2019):

- The interpretation of the points on the plot considers two distributions and a point of evaluation (any value). The point on the plot indicates what percentage of data lies at or below the point in both distributions (as per definition of the CDF);
- To be able to compare the distributions, the points lie on a 45-degree line ($x=y$). In case they deviate, the distributions differ;
- P-P plots can be used to evaluate the skewness of a distribution visually;
- P-P plots are most useful when comparing probability distributions that have a nearby or equal location.

2.1.7.2 Quantile-Quantile Plot

The Quantile-quantile plot, Q-Q plot, has the same basis as P-P plot, to compare distributions by plotting their quantiles against each other. Q-Q plot is a graph of the input, measured data values, plotted against the theoretical distribution quantiles. Both axes of this graph are in units of the input data set. This graph is produced by plotting the observed data values x_i against the X-axis, and the following values against the Y-axis (Learn More About EasyFit, 2004-2010):

$$F^{-1}(F_n(x_i) - \frac{0.5}{n}) \quad (2.42)$$

where:

$F^{-1}(x)$ – Inverse cumulative distribution function (ICDF);

$F_n(x)$ – Empirical CDF;

n – Sample size.

The Q-Q plot will be approximately linear if the specified theoretical distribution is the correct model.

Some leading information on Q-Q plots must be mentioned (Lewinson, 2019):

- Interpretation of the points on the plot: a point on the chart corresponds to a particular quantile coming from both distributions;
- On a Q-Q plot, the reference line is dependent on the location and scale parameters of the theoretical distribution. The intercept and slope are equal to the location and scale parameters respectively;
- A linear pattern in the points indicates that the given family of distributions reasonably describes the empirical data distribution;
- Q-Q plot gets excellent resolution at the tails of the distribution but worse in the centre (where probability density is high);
- Q-Q plots do not require specifying the location and scale parameters of the theoretical distribution, because the theoretical quantiles are computed from a standard distribution within the specified family.

2.1.7.3 Probability Difference Graph

The probability difference graph (Diff) is a plot of the difference between the empirical CDF and the theoretical CDF, and it is represented by the following equation:

$$Diff\ (x) = F_n(x) - F(x) \quad (2.43)$$

This graph can be used to determine how well the theoretical distribution fits the observed data and compare the goodness of fit of several fitted distributions. It is displayed as a continuous curve or a scatterplot for continuous distributions and a collection of vertical lines (at each integer x) for discrete distributions (Learn More About EasyFit, 2004-2010). An example of Diff graph is presented in Figure 2.21.

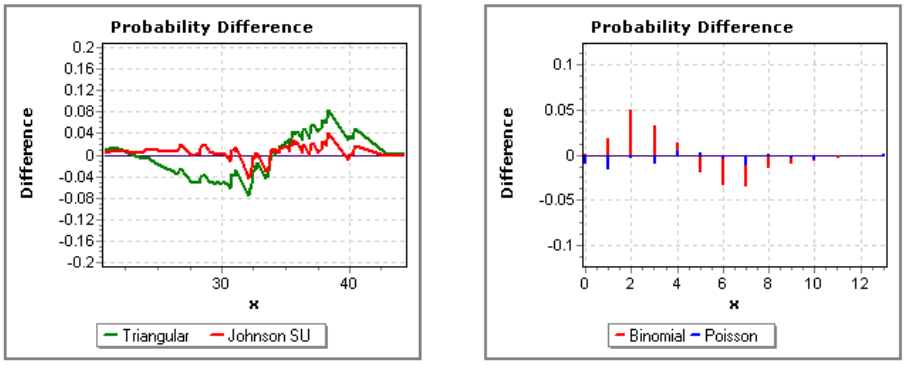

Figure 2.21 Probability difference graph (Learn More About EasyFit, 2004-2010)

2.1.8 GOODNESS FIT TEST

The goodness of fit (GOF) test measures the compatibility of a random sample with a theoretical probability distribution function. This test shows how well the selected distribution fits the data (Learn More About EasyFit, 2004-2010).

The most common GOF tests considered are the following:

- Kolmogorov-Smirnov;
- Anderson-Darling;
- Chi-Squared.

To this case study, the GOF test considered was the Kolmogorov-Smirnov and Anderson-Darling due to the type of data in the study.

2.1.8.1 Kolmogorov-Smirnov Test

This test is used to decide if a sample comes from a theorised continuous distribution. It is based on the empirical cumulative distribution function (ECDF). Assuming a random sample [x_1, x_n] from some continuous distribution with CDF $F(x)$, the empirical CDF is denoted by (Learn More About EasyFit, 2004-2010):

$$F_n(x) = \frac{1}{n}[Number\ of\ observations \leq x] \tag{2.44}$$

The Kolmogorov-Smirnov statistic (D) is based on the most significant vertical difference between $F(x)$ and $F_n(x)$. It is defined as the following:

$$D_n = sup_x |\, F_n(x) - \ F(x)| \tag{2.45}$$

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, D, is higher than the critical value obtained from a table (Learn More About EasyFit, 2004-2010).

2.1.8.2 Anderson-Darling Test

The Anderson-Darling procedure is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function (Learn More About EasyFit, 2004-2010). This test is named as A^2, automatically considered in the software EasyFit, and is defined by the following formula:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln F(X_i) + \ln(1 - F(X_{n-i+1}))] \quad (2.46)$$

The hypothesis regarding the distributional form is rejected at the chosen significance level if the test statistic, A^2, is higher than the critical value obtained from a table. The fixed values of α (0.01, 0.05) are generally used to evaluate the null hypothesis (H_0) at various significance levels. A value of 0.05 is typically used for most applications. However, in some critical industries, a lower α value may be applied.

In general, the critical values of the Anderson-Darling test statistic depend on the specific distribution being tested. However, tables of critical values for many distributions (except the most widely used ones) are not easy to find (Learn More About EasyFit, 2004-2010).

2.1.9 CORRELATION COEFFICIENT

The correlation coefficient is a statistical parameter to measure how strong is a relationship between two variables. There are different types of correlations coefficient, such as Pearson's correlation. The correlation coefficient can assume values within -1 to 1 with the following meanings (Stephanie G. , u.d.):

- A correlation coefficient of 1 means that for every positive increase in one variable, there is a positive increase in a fixed proportion in the other;
- A correlation coefficient of -1 means that for every positive increase in one variable, there is a negative decrease of a fixed proportion in the other;
- Zero means that for every increase, there is neither a positive or negative increase. The two variables are not related.

The absolute value of the correlation coefficient gives us the relationship strength. The larger the number, the stronger the relationship.

The correlation coefficient can be analytical determined through the following formula:

$$\rho_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (2.47)$$

where:

ρ_{xy} - correlation coefficient of the linear relationship between the variables x and y;

x_i – values of the variable x in the sample;

$\bar{x}$ - mean of the values of the variable x;

y_i – values of the variable y in the sample;

$\bar{y}$ - mean of the values of the variable y.

2.1.10 Confidence Intervals

When talking about probabilistic values, those values need to be conceivable to the case study. The confidence interval tells how confident you are in your results (Confidence Interval: How to Find a Confidence Interval: The Easy Way!, 2020). Confidence intervals express a reasonable range of plausible values to reduce the uncertainty related to the values that every sample can outcome. Typically, confidence intervals are used with a margin of error that represent the belief of the represented values. Confidence intervals are combined with confidence levels. Confidence levels are expressed as a percentage, for example, a 95% confidence level, like in Figure 2.22. For this confidence level, when an experiment is frequently repeated, 95% of the time the outcomes will match the results obtained from an analysis.

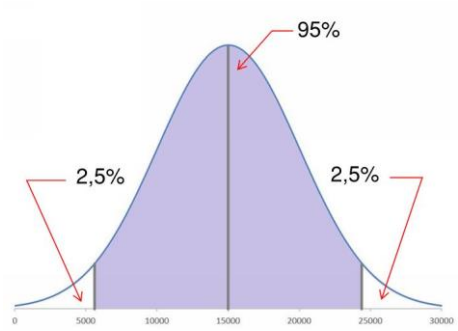

Figure 2.22 Confidence level of 95%

2.2 Uncertainty and Risk in Geotechnical Engineering

2.2.1 Historical Development in Geotechnical Design

The world is imperfectly knowable. When an engineer is performing a design, his base concern is safety. The main issue in geotechnical engineering is the ambiguity and uncertainties related to the geological material and the variable state of nature, to progress to geotechnical design. In the past, engineers like Casagrande (Casagrande, 1965) and Peck (Peck, 1969) developed a method to counteract these uncertainties, named the observational method. Engineers used to create the design based on conservative values of loads and materials properties, even if sometimes it was not possible to build in

reality, since often the expected values fell outside the design range. In recent times engineers are applying the results of reliability theory to geotechnical engineering. Even though engineers are conscious about the fact that uncertainty is not excluded and that parameters have to be examined. By with this approach, engineers can provide a way of quantifying those uncertainties and handling them consistently.

2.2.2 RELIABILITY ANALYSIS

Since geotechnical engineering is not an exact science, reliability analysis is even more needed. Reliability analyses allow engineers to picture and estimate the variable parameter definitions that geotechnical engineering faces. Reliability approaches do not substitute conventional methods used but complement them, providing more consolidated information on the parameters and uncertainties involved. Reliability analysis is the only type of analysis that can provide to the designer a perception of the essential risk level associated with what is being studied.

2.2.3 STATISTICS VS. PROBABILITY

The approach of evaluating risk and reliability is intimately associated with probability and statistics. However, these concepts have individual definitions. Probability can be thought of as an algebra – a set of results derived by rigorous reasoning from a set of hypotheses. Statistics deals with the description of the observed world. That they often arrive at similar statements and conclusions does not change the fact that the primary reasoning processes of the two disciplines are different (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003).

2.2.4 UNCERTAINTY

Uncertainty is related to a lack of knowledge. It is connected with the unidentified, through multiple variable conditions of what is being interpreted. Uncertainty guides us to something that is not sure that will happen, or that has the chance to happen - due to that, uncertainty is predictable. In geotechnical engineering, the term uncertainty is a daily basis since it is present in all site investigations. The primary uncertainty sources in geotechnical design are due to the definition of the primary variables combined with models that determine the variability of the secondary variables. Talking about uncertainty in geotechnics divides us into two different topics, aleatory and epistemic uncertainties, schematically represented in Figure 2.23.

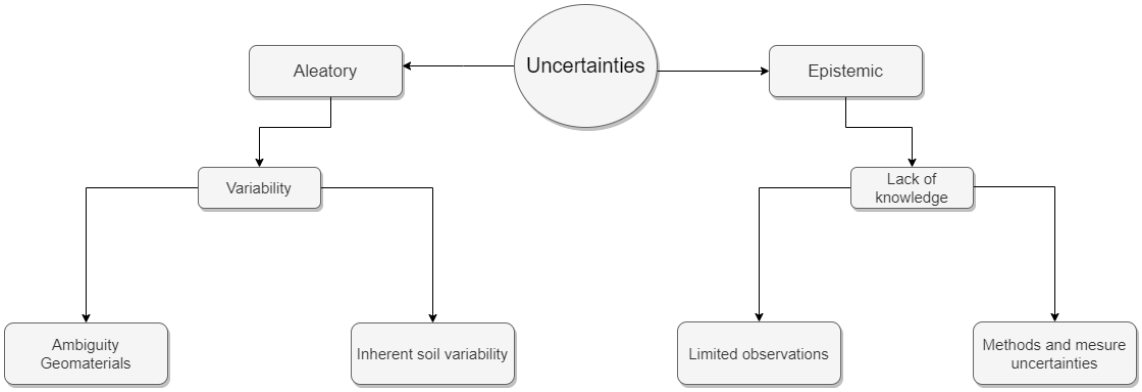

Figure 2.23 Types of uncertainties

Variability is a natural process that outcomes from the geological developments which are continuously modifying. One can define variability over time for singularities that take place at a single location (temporal variability), or as variability over space for singularities that take place at different locations but at a single time (spatial variability), or as variability over both time and space (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). Epistemic uncertainties can be due to the equipment or failures from the operator, data handling and transcription errors, uncertainty on the actions (permanent, variable, accidental) estimation. The lack of data, or subjective information, or even incompatible representativeness of observations can be due to time and space restrictions. On the decision model, uncertainty is introduced when the field or laboratory measurements are transformed into soil/rock design parameters using empirical or other correlation models (Phoon, Prakoso, Wang, & Ching, 2016), resulting in the inability of a model or design method to characterize a parameter or a range of parameters accurately. Thus, this happens when the approaches used in the design are not well-matched with real behaviour. In geotechnical engineering, uncertainty is always linked with risk. Engineering practice is based on risk analysis modelled by uncertainty, in order to diminish the allied risk.

2.2.5 RISK ANALYSIS

Risk is associated with all topics of natural life, including Engineering design. If certain events occur, they will create devastating consequences. Risk analysis is the ability to predict events that have not happened yet. The main goal in engineering is to reduce risk to the society and environment, assuming that it is possible to determine the probability of the event to occur and to quantify the magnitude or cost of the consequences associated with that occurrence. The combination of an uncertain event with the corresponding adverse consequences is what determines the risk (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003):

$$Risk = \sum_{i} p_i * c_i \tag{2.48}$$

where:

p_i – Probability of each event to occur;

c_i – Consequences of each event.

However, this definition does not always match reality since different occurrences can lead to different failures and different time sequences that result in different consequences.

2.2.5.1 Margin of Safety

Civil engineering is always related to risk and a margin of safety, in other words, acceptable risk. The margin of safety is associated with failure probability, which allows engineers to interpret if the associated risk is acceptable or not. Most of the engineering design is based on safety factors (*F*), which correspond to a parameter defined by the ratio between resistance (*R*) and the load (*Q*). The margin of safety (*M*), defined by engineers, is the difference between resistance and the load.

$$F = \frac{Resistance}{Load} \tag{2.49}$$

$$M = Resistance - Load \tag{2.50}$$

Our critical value is when M = 0, this means, once M > 0, engineers assume that the structure is safe, otherwise, when M ≤ 0, the structure is out of the safe zone.

2.2.5.2 Reliability Index and Probability of Failure

By defining the probability distribution parameters *R* and *Q*, represented in Figure 2.24, it is possible to determine also the mean (μ) and the variance (σ^2) of M, by the approaches presented in previous sections. A reliability index (β) expresses the distance of the mean margin of safety from its critical value in units of standard deviation. This index corresponds to the M mean divided by its variance.

$$\beta = \frac{\mu_M}{\sigma_M} \tag{2.51}$$

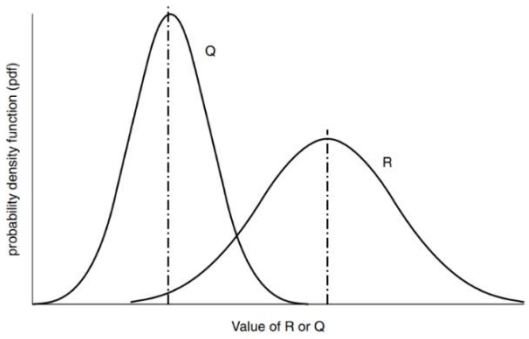

Figure 2.24 Probability densities for typical resistance (R) and load (Q) (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003)

When considering the associated uncertainties, there is also uncertainty in the safety margin. The failure probability, P_f, is the zone under the probability distribution of M where $M \leq 0$, as represented in Figure 2.25. Note that the area under the curve (a) and to the left of the axis corresponds to the probability of failure identified in (b).

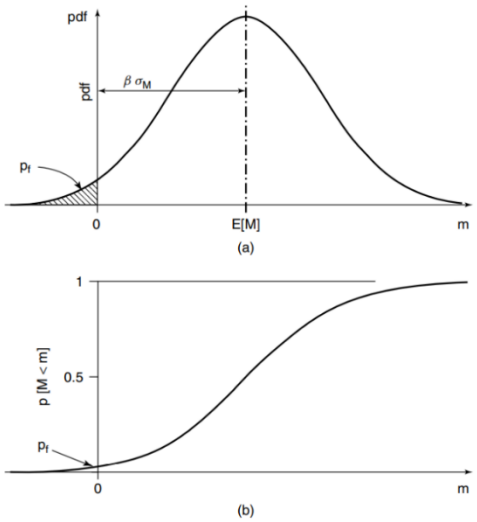

Figure 2.25 Probability density (a) and cumulative probability (b) for margin (M). (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003)

2.2.5.3 Risk Assessment and Management

Once it is not possible to have a zero-failure probability, engineers need to know until when the risk is tolerable in order to proceed to the expected design. Risk management is the process of identifying, analysing and assessing risks to enable informed decisions on accepting, treating and controlling risks to minimise them (Lacasse S. , et al., 2019). The risk that can result in damage is denoted as danger or hazard. Minor risks involve high costs because structures need to be reinforced to reduce the failure probability. It is called risk mitigation, as represented in Figure 2.26, defined as the process of deciding and implementing actions to counteract the identified risks. Some of those risks can be treated with monitoring and periodic review, otherwise, advanced risks require more advanced measures. Subsequently, engineering judgment based on probabilistic analysis has to decide the best answer to counter the risk. Risk management has been formalised into a framework by ISO 31000 (2018) (ISO/TC, 2018).

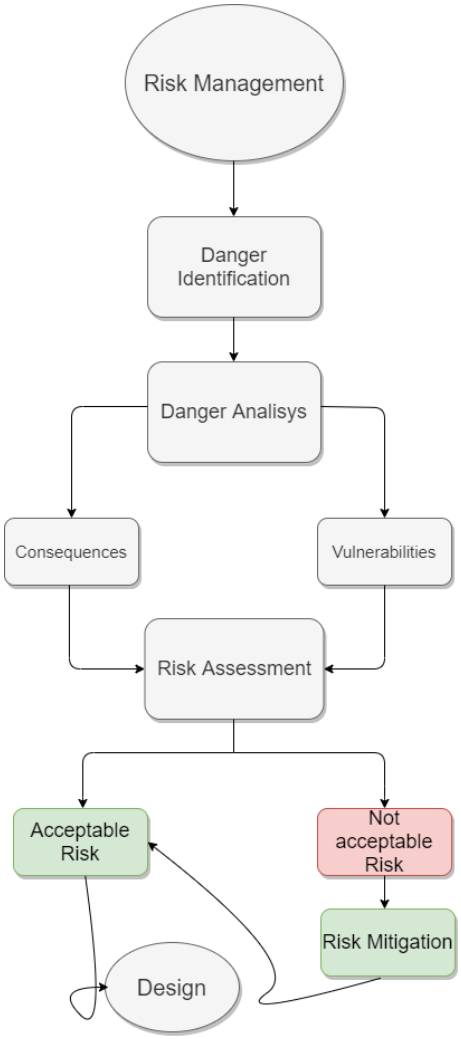

Figure 2.26 Risk management

Nowadays, there are several methods to do the risk assessment, which can be divided into two different types:

1. Qualitative methods;
2. Quantitative methods.

Since there are dozens of methods, only the most common of each type will be briefly presented.

The most common tool in the group of qualitative methods is the "traffic-light" matrix, represented in Figure 2.27. The qualitative matrices can be very useful, mainly when assessed through the consensus of several individuals with different expertise. Over the years, the 5x5 matrix has gained more popularity than the original 3x3 matrix (Lacasse S. , Nadim, Boylan, Liu, & Choi, 2019).

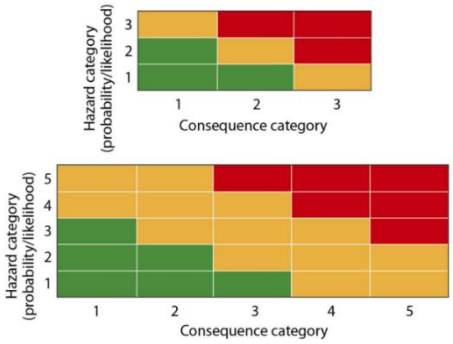

Figure 2.27 Qualitative risk assessment with 3x3 and 5x5 matrix: Hazard categories 1 to 3 or 1 to 5 from very low to very high hazards (Lacasse S. , Nadim, Boylan, Liu, & Choi, 2019)

The matrix contains three different colours, related to varying rates of risk. The low risk is represented by the green colour, the orange is the intermediate-risk, while the red colour represents the higher associated risk. Primarily, these qualitative estimates allow the reader to understand whether or not more detailed analyses are needed if red and orange colours are the primary colours present in the matrix. It is possible to implement this method in a macro-operated Excel sheet.

When the subject is the quantitative method, parameters like materials or load properties are described by a probability distribution. With all of them described as statistical parameters associated in more typical cases, the analyst must usually employ a technique that yields an approximation to the real value of the reliability index and the probability of failure. Several methods are available, each having advantages and disadvantages (Baecher & Christian, Reliability and Statistics in Geotechnical Engineering, 2003). These quantitative methods consist of:

- The First Order Second Moment (FOSM);
- The Second Order Second Moment (SOSM);
- The Point Estimate method (Rosenblueth);
- Event tree analysis;

Additionally, there are more sophisticated methods that will be detailed further on:

- Monte-Carlo Simulation;
- Latin Hypercube Simulation;
- Bayesian Approaches.

2.2.5.4 Acceptable Risk

The demarcation between acceptable and unacceptable risk is usually a gradual transition. In Figure 2.28, the zone for F between 10^{-3} and 10^{-4} and one to eight mortalities seems to belong to two categories. If a risk estimate should fall in that zone, the most severe action (red line) should be applied (Lacasse S. , et al., 2019).

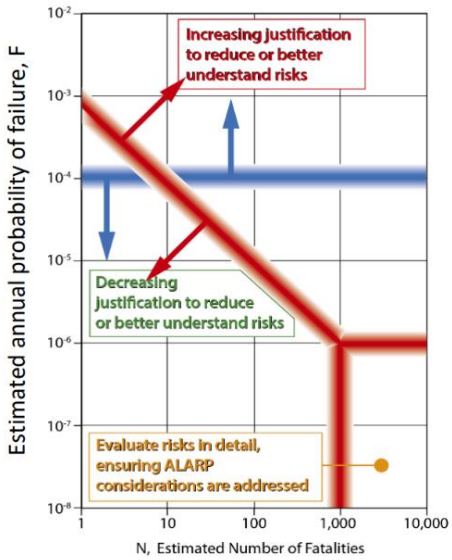

Figure 2.28 US Bureau of reclamation 2011 guidelines (Lacasse S. , et al., 2019)

2.2.6 WHY PROBABILISTIC APPROACHES

Different approaches are used in engineering, even though all of them have gaps. Engineering judgment is a requisite to every decision made to evaluate and to make decisions for the case study. Deterministic analytical models should be used when material properties, failure modes (mechanisms and geometries) and forces are known with reasonably high accuracy. Probabilistic analyses should be used when the uncertainty in parameters may govern the results of the investigations (Lacasse & Nadim, 2007). As seen before, the deterministic method is the right approach but not a complete one. Most of the time, deterministic approaches are based on safety factors or tabled values that can result in oversized designs. Thus, more time and costs involved. Probabilistic methods do not replace deterministic analysis, though they complement them. With engineering judgment to evaluate and study, the combination of both approaches will result in more conscious and valid interpretations for every situation. Spatial variability of ground properties and other geotechnical uncertainties may be modelled probabilistically using either a single random variable (SRV) or a series of spatially distributed and correlated random variables, a random field (RF). A scheme comparing both methods can be observed in Figure 2.29.

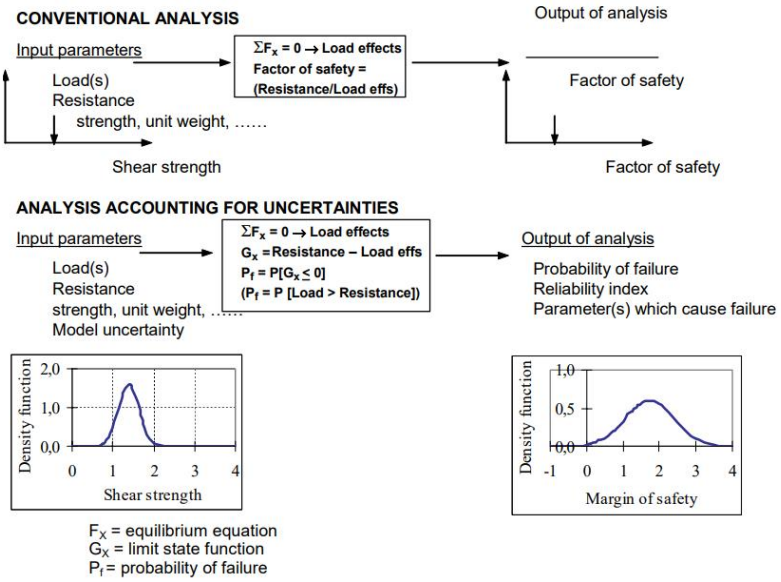

Figure 2.29 Comparison of deterministic and probabilistic analysis (Lacasse & Nadim, Risk and Reliability in Geotechnical Engineering, 1998)

Probabilistic analysis is similar to sensitivity analysis but more rational with additional rigour, and accuracy.

2.2.7 PROBABILITY INTERPRETATION

Although probability and statistics are well-developed subjects, there are philosophically different schools of probability, the most prominent being the frequentist and Bayesian schools (Wang, Zhao, & Cao, 2019). The introduction of probabilistic and risk analysis in practical activities implies, among other things, the need to clarify the interpretation of the concept of probability.

2.2.7.1 Frequentist Theory

The frequentist approach does not allow a subjective interpretation of the probability. The uncertainties of knowledge cannot be treated in the same way as the inherent and measurable ones. Those who support this point of view are obliged to use other approaches to deal with these uncertainties (not probabilistic, according to their interpretation), such as the standard tools, like confidence intervals, safety factors, safety margins and conservative estimates, or others. Frequentist theory can be summarized in the following points:

- The statistical frequency of a value is the relationship between the number of times that a value appears and the total number of observations;
- Only observable and countable events can be considered within the domain of this probability theory;
- The probability assessment should only be based on enough data or unambiguous theoretical arguments;
- This interpretation only finds adequate justification in a stationary world, where the amount of statistical or theoretical evidence is considerable;

- This interpretation does not fit the field of civil engineering applications since in almost all cases, data is very scarce and often very generic.

2.2.7.2 Bayesian Statistics

Was named after its inventor, the 18th-century Presbyterian minister Thomas Bayes. Bayes' theorem is a method for calculating the validity of beliefs (hypotheses, claims, propositions) based on the best available evidence (observations, data, information) (John, 2016). The bayesian approach is used for the treatment of unknown parameters. Bayes' theorem is based on a rearrangement of the conditional probability equation to get more accurate values of something to occur. Thus, Bayesian inference is frequently known as Bayesian updating. The more confidence in the evidence, the more precise is the final value. In geotechnical practice, the degree of belief may be interpreted as a reinforcement of engineering judgment, to get the typical engineering practice value. Due to this, it is possible to conclude that most geotechnical engineers are intuitive Bayesians whether they know it or not. Through that rearrangement the basic mathematical formula for the Bayes' theorem (2.9) results in the next equation:

$$P(A|\mathrm{B}) = \frac{\mathrm{P}(\mathrm{B}|\mathrm{A}) * \mathrm{P}(\mathrm{A})}{P(B)} \tag{2.52}$$

where:

P – Probability;

A – Belief;

B – Evidence;

P(A) – Probability that A is true;

P(B) – Probability that B is true;

P(A|B) – Probability of A if B is true;

P(B|A) – Probability of B if A is true.

2.3 STATISTICAL APPROACHES FOR RISK ASSESSMENT

2.3.1 COMPUTATIONAL STATISTICS

Computers and developed software are a powerful tool in terms of creating, and processing more values, equations and exhaustive statistical methods in less time than was ever possible for humans to process in a suitable time. The connection between statistics and computer science is the definition of the term computational statistics. Resampling methods, like the Monte Carlo and Latin Hypercube methods, are intense methods that computers are useful to deal with. These methods require the generation of random numbers running multiple analyses of simulated datasets. Three crucial areas of statistical computing are random number generation, numerical linear algebra, and optimisation (Gentle, 2010). A model in an Excel spreadsheet, having a certain number of input parameters and a few equations that use those inputs to give a set of outputs (Figure 2.30), is usually deterministic, meaning that the same results are obtained no matter how many times it is re-calculated (Wittwer, Jon, 2004). On the contrary, in the resampling methods indicated above, each time the method is recalculated, the result may be different.

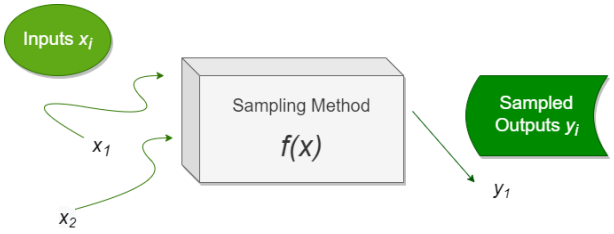

Figure 2.30 Deterministic model with a set of input variables to a set of output variables

2.3.2 SAMPLING IN STATISTICS

Data sets can have much information. Sometimes those data sets are too heavy to process in statistics in terms of time and costs. Samples are part of the data set, visible in Figure 2.31. When having a good-sized sample, and analysing it, it is possible to have expectations of the data set behaviour, processing the information in a more accessible way. There are different types of techniques for sampling in elementary statistics. The one to focus in this study is random sampling. With random sampling, each object does not have an equal chance of being chosen.

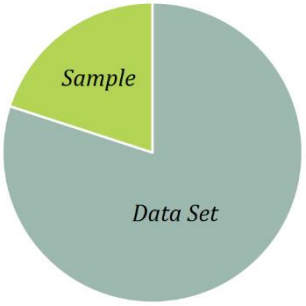

Figure 2.31 Sampling in statistics

2.3.3 GENERATING RANDOM VARIABLES

As presented before, with Probability Integral Transform, it is possible to generate random numbers respecting probability distribution properties. When F is a cumulative distribution function, and it is truncated at values $t_0 < t_1$, then random values can be obtained as (Dodge, 2006):

$$Y = F^{-1}(X) \tag{2.53}$$

where X is uniformly distributed in the interval $[a, b] = [F(t_0), F(t_1)]$.

Starting from data that is defined through a specific distribution, using distribution values into its CDF, the correspondent data points are distributed in the standard uniform distribution.

Whenever two random variables are correlated, the random produced samples must also be correlated, mainly if they are intended to be used in further calculations. The Cholesky decomposition is commonly used for simulating systems with correlated variables. Cholesky decomposition reduces a symmetric matrix into a lower-triangular matrix which, when multiplied by its transpose produces the original

symmetric matrix (Pistilli, 2019). In practice, when using, for example, Monte Carlo (MC) method, first, both sets of samples must first be created randomly. Then by the application of a formula, one of the parameters must be affected by the correlation factor, and further MC method is applied.

To generate two correlated variables, x_1 and x_2, a correlation coefficient, ρ, must be calculated. First, the two variables must generate two uncorrelated random variables, y_1 and y_2, through Monte Carlo simulation. Assuming the obtained correlation coefficient, given by equation 2.47, the correlated variables can be designed through the resulting calculation:

$$x_2 = \rho y_1 + \sqrt{1-\rho^2} y_2 \tag{2.54}$$

2.3.4 SIMULATION METHODS

Statistics can deal with the lack of data or theoretical background since with some samples, it is possible to predict with reasonable error what is happening in the study case, or export the data wanted. Due to simulation methods and computer software, that process is eased. Some of the advantages related to simulation methods can be described as (Gentle, 2010):

- The use of generated random numbers instead of collected data;
- Faster information processing;
- Analyse the behaviour of complex processes;
- When evaluating proper data, methods give reliable outcomes.

Simulation methods work as virtual laboratories where hypotheses about observed problems can be tested and processed through iterations as rework loops that are almost impossible for human mental analysis.

Probabilistic approaches work with a combination of random variables, based on a model, which contains numerous input parameters. This type of model comprises significant mathematical or arithmetic processing where deterministic ways must be sufficient to solve them. Combining random variables where the mathematical or arithmetic ways are not applied, the technique to adopt is over probabilistic approaches.

Even though reliable outcomes are expected, operating analyses need to hold all the requirements that the case study demands. The following section presents the simulation methods considered in this particular work. They characterize the most straightforward and powerful methods, with no restrictions associated.

2.3.5 MONTE CARLO

Monte Carlo simulation (MC) is a method for an iterative evaluation of a deterministic model, mainly using sets of random numbers as input. This method is often used when the model is complex, nonlinear, or involves more than just a couple of uncertain parameters (Wittwer, Jon, 2004). Whenever Monte Carlo method is needed, there must be a large number of samples to have a good representation of probability distribution, (having lower samples will not cover all the case in study). As each sample is entirely random, every time this method starts, new sample points are generated without considering those previously made, this means the Monte Carlo is a nonmemory method. Monte Carlo is one of the

methods for analysing uncertainties, to predict specific parameters based on the probability of them to occur. With the samples number obtained, it is possible to find a type of distribution that matches them. The methodology behind Monte Carlo is considerably useful when a specific equation cannot solve the case study. This method consists of the following points:

- Know the probability density functions of all the random variables;
- Randomly select those random variables, assigning a value for each variable;
- With the range of values obtained, it is calculated the comparable value to variable data;
- Continually repeating the process, it is possible to achieve a range of values to the variable from whom it can be defined as its probability density function, and then draw a conclusion.

The data generated from the simulation can be represented as probability distributions (or histograms) or converted to error bars, reliability predictions, tolerance zones, and confidence intervals (Wittwer, Jon, 2004). Figure 2.32 presents the basic principle behind the Monte Carlo simulation.

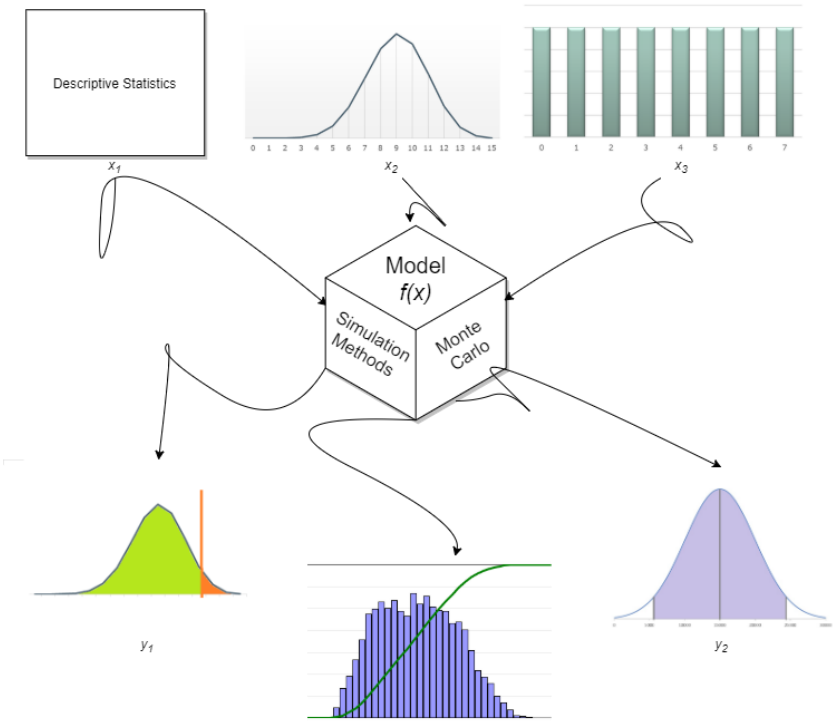

Figure 2.32 Monte Carlo basic principle

After input treatment and analysis, a PDF for the available data can be defined. Afterwards, MC is implemented, generating the needed number of samples through the procedure mentioned in section 2.3.3 adapted to the input variable characteristics for each input parameter. Then, the model is processed for each sample. These generated samples can define a complete PDF variable, the one to use in further calculations.

As said before, to get better results when using the Monte Carlo method, a large number of samples is needed. In order to reduce the number of samples required, stratified sampling techniques such as Latin Hypercube Sampling (LHS) were developed, as discussed in the next section.

2.3.6 LATIN HYPERCUBE SAMPLING (LHS)

The primary purpose behind LHS is to accurately recreate the probability distribution with fewer samples compared to the Monte Carlo approach. The key to this method is the stratification of the input probability distribution. Stratification means to divide the cumulative curve into equal intervals. Then a sample from each stratification is randomly selected. When a sample is selected from a stratification, this stratification is not sampled from again. In LHS, the number of stratifications of the cumulative distribution is equal to the number of iterations performed. In Figure 2.33, LHS is exemplified considering just one variable, representing a one-dimension sample. In this illustration four random intervals are represented, where per stratification only one sample is taken.

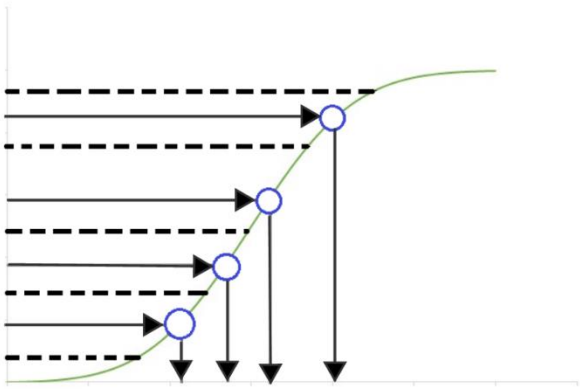

Figure 2.33 One-dimension sampling stratification (LHS)

When using the Latin Hypercube technique to sample from multiple variables, it is essential to maintain independence between variables. The values sampled for one variable need to be independent of those sampled for another. This independence is maintained by randomly selecting the interval to draw a sample for each variable (Natvig, 2005). In the context of statistical sampling, a square grid containing a sample position is a Latin square if there is only one sample in each row and each column. Visible in Figure 2.34 is a distinction between a normal random situation and two LHS dimensions. Once in LHS, it is only possible to select a sample from each stratification. This method can be characterised as a method with memory.

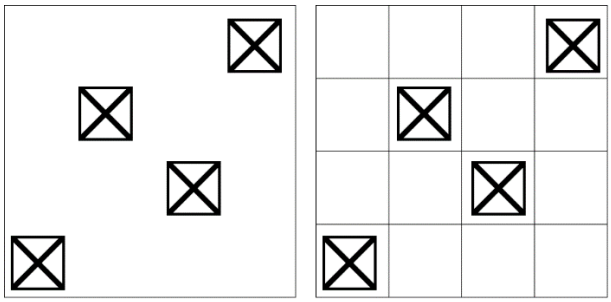

Figure 2.34 Random sampling and two-dimension (LHS)

The main advantage of LHS over MC is the fact that LHS has memory, meaning fewer samples needed than the memoryless method MC.

2.3.6.1 Stratification Development

In LHS, the generation of random variables follows the same adopted in MC, contemplating the procedure mentioned in section 2.3.3. The main difference between these two simulation methods procedures is the stratification process characteristic of LHS. For the stratified sampling to be accomplished, the CDF of the normal distribution is divided into segments, as visually described in Figure 2.33. A probability is randomly selected within each segment using a uniform distribution (section 2.1.6) and then plotted to the correct characteristic value of the actual variable distribution. A simulation with 500 iterations would fragment the probability into 500 segments, each representing 0.2% of the total distribution. For the first sector, a number would be chosen between 0.0% and 0.2%. For the next sector, a number would be chosen between 0.2% and 0.4% and so forth. This number would be used to calculate the actual variable value based upon its distribution (Fenniak, 2004). The following equation denotes the random value selected in each stratification:

$$P(X \leq x) = n \tag{2.55}$$

The probability is solved for *x*, where *n* represents the random point nominated in the sector. This process is different for each distribution, but it is just the reverse process of the probability function.

3
Land reclamation

3.1 Reclamation of Coastal Areas

Many densely populated cities are built inland, even in soft ground. Due to human activity with continuing development and redevelopment of urban and coastal areas, most of the possible construction sites are entirely occupied. New strategies need to be in concern to continue cities development. The expansion of land to the sea by creating new territory is a development that started its implementation since the last century. This procedure is called land reclamation. Artificial islands are an example of land reclamation, like the worldwide known Dubai's artificial islands, pictured in Figure 3.1. Another example is shown in Figure 3.2, representing Yangshan Port in Hangzhou Bay south of Shanghai, China, during its construction phase.

Figure 3.1 Palm Jumeirah (left) and Palm Deira (right) with the world and the universe archipelagos (Karlhuber, 2008)

Figure 3.2 Land reclamation of Yangshan deep-water port (YH, 2007)

Frequently engineering and environment are opposed. Land reclamation is interdependent, this means, engineering will only be complete if the environment does not get compromised. In the process of conceiving land reclamation projects, some points on the environment need to be considered, such as:

- Geographical constraints;
- Environmental impact;
- Marine ecology;
- Operational efficiency;
- Infrastructure development nearby.

3.1.1 SOIL IMPROVEMENT

The simplest methods of land reclamation are called drained reclamation, consisting of filling the area with natural materials. These natural materials can have different sources, such as construction sites or excavations for stations and tunnels. With these approaches, it is possible to recycle and reuse waste soils extracted due to construction. The drained reclamation method commonly used in Hong Kong involves putting layers of earth and sand on the sea bed, creating pressure to consolidate them and extract water from within layers through weep-holes.

On the other hand, there are soil improvement methods such as the following:

- Deep cement mixing;
- Stone columns;
- Prefabricated vertical drains (PVD).

The deep cement mixing method is widely implemented in Japan, as well as in the continents of Asia, Europe and America. Deep cement mixing slowly injects and blends cement into the soft mud of the sea bed, stiffening it into the form of cement columns able to sustain the reclamation load.

Stone columns helps to improve load-bearing capacity mitigating the effect of soil settlements. They allow shallow foundations on ground that would usually require deep foundations. In structures with an evenly spread load, the columns are installed in a regular grid pattern. Their diameter is generally between 40 and 120 cm. The technique consists of a vibrating probe subjected to a combination of vibration, self-weight and a pull-down force which pushes the probe into the ground until it reaches the

required layer. Coarse aggregate composed of various sizes is incorporated into the ground. The tool is raised either from the top of the bore area or directly to the base of the probe through an adjacent tube. The aggregate is compacted by vibration as the tool is raised. This technique offers some advantages, such as:

- Permeability properties of the aggregate speeds up consolidation of the ground;
- Reduce total and differential settlement;
- A practical solution to overcome liquefaction.

Prefabricated vertical drains, also designed for the acceleration of soft soil consolidation settlement and compaction, consist of a specially designed synthetic core which conducts maximum water flow along the length of the drain (Rujikiatkamjorn, 2005). Also called "wick drains" or "strip drains", PVDs base is fully wrapped in a durable geotextile filter jacket that has very high permeability and filtration properties. Both the core and the geotextile filter jacket have significant strength and durability, visible in Figure 3.3. PVD installation involves attaching the end of the PVD to an anchor plate on the soil. Further, PVD is installed into the ground using a hollow mandrel that protects the PVD material, Figure 3.4. The mandrel is then hydraulically pushed into the field to the specific spaces where PVD is going to be implemented. Once the PVD is at the correct depth and pressure, the mandrel is withdrawn back into the mast, leaving the PVD in place. After installation, as illustrated in Figure 3.5, the PVD position is checked by a specific measuring and recording technology. After all PVD installation is complete, a new layer of loading must be applied to the site. This layer, combined with the PVD, reduces settlement time by reducing consolidation time, as also increasing soil bearing capacity.

Figure 3.3 Typical types of PVD (a) Colbond drain (b) Mebra drain (Rujikiatkamjorn, 2005)

Figure 3.4 Typical installation (Rujikiatkamjorn, 2005)

Figure 3.5 Horizontal drains in the transverse and longitudinal direction (Rujikiatkamjorn, 2005)

These methods have rapid and efficient installation processes, providing a more cost-effective, safer and faster time frame.

3.2 GEOTECHNICAL APPROACHES FOR SETTLEMENT ANALYSIS

3.2.1 SOIL MECHANICS REVIEW

"In engineering practice, difficulties with soils are almost exclusively due not to soils themselves but to water contained in their voids. On a planet without any water there would have been no need for Soil Mechanics" Terzaghi, Karl

Water seeps through the soil. This occurrence is denoted as percolation, where water travels through soil particles pores. This phenomenon is dependent on soil permeability. Permeability is the facility of soil pores to allow water passage through them. Those pores are mainly composed of air. Once those pores become only occupied by water, this means the soil is saturated. Knowing that soil is a layered

structure, the determination of permeability allows to understand how water flows through the different layers in order to improve the soil towards the intended construction goal.

The law that governs water movement in soils is named Darcy's law, based on Darcy's experiment.

Darcy concluded that the water flow results in the next equation (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006):

$$Q = k\frac{h_1 - h_2}{L}S = k * i * S \tag{3.1}$$

where:

k – Soil permeability coefficient;

i – Hydraulic gradient;

S – Area of the sample.

Considering that all soil layers are homogeneous, it is possible to calculate flow velocity by the rearrangement of the above equation, resulting in (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006):

$$v = \frac{Q}{S} = k * i \tag{3.2}$$

It is essential to notice that velocity takes different values along soil routes due to variations in the dimension of the pores present in all the different layers that constitute the soil.

Water flow creates forces between soil particles. This force is denoted as percolation force represented by the effect that water applies to each soil volume unit due to water movement. Considering the water unit weight (γ_w) and making the difference between hydrodynamic and hydrostatic forces, it is possible to quantify percolation force (J) obtaining the following equation (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006):

$$J = i * \gamma_w \tag{3.3}$$

The hydrostatic case is when soil is submerged being only submitted to water buoyancy. The hydrodynamic case is when water buoyancy and percolation forces act at the same time.

The coefficient of permeability is denoted as *k* and has an extensive range of possible values, varying eight to nine orders of magnitude. Soil grain size distribution is the most relevant factor for permeability. However, the soil void ratio, the mineralogical composition (mainly for clays) and the soil degree of saturation may also influence (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006). Due to laboratory tests or *in situ* surveys, it is possible to evaluate the coefficient of permeability. In clays, the water movement is dependent on the electrochemical forces between particles, and therefore

clay activity affects permeability, reducing it when the activity is high. The permeability coefficient for clays generally ranges from 10^{-7} to 10^{-9} m/s, which practically represents an impervious soil.

Clays generally show a volume increase when wet, and when dried, the volume decreases, creating many cracks (Ural, 2018). For that reason, several parameters are needed to understand the clay behaviour, such as:

- Atterberg's limits;
- Permeability (Hydraulic conductivity);
- Swelling-shrinkage;
- Undrained shear strength;
- Consolidation rate and settlement.

Atterberg's limits are the relationship between soil particles and water and how soil react for variable water contents. Those limits are visible in Figure 3.6. When water content increases, clay changes from solid-state, to a semi-solid state, to a plastic state and then turning into a liquid state. For high values of water content, the mixture water-soil behaves like a liquid. When regularly and gradually reducing water content, at some point clay starts to be mouldable, preserving the form that was designed for. When continuing water reduction soil starts to behave like a semi-solid, which means soil separates himself in fragments when trying to get a particular shape. Keeping water reduction, there is one point where volume does not vary. The points that separate these previously mentioned conditions are visible in Figure 3.6, which are called liquid limit (w_l), plastic limit (w_p) and Shrinkage limit (w_s), respectively.

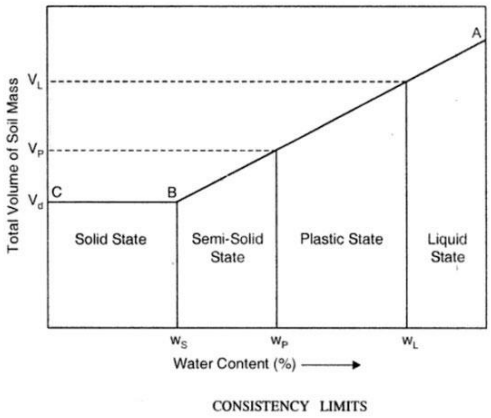

Figure 3.6 Definition of Atterberg's limits (Ural, 2018)

The soil shrinkage is defined as the specific volume change of soil relative to its water content. This process occurs mainly due to clay swelling properties (Haines, 1923). It can be measured in soils with more than 10% of clay content (Boivin, Garnier, & Vauclin, 2006) and shows a typical S-shape, visible in Figure 3.7.

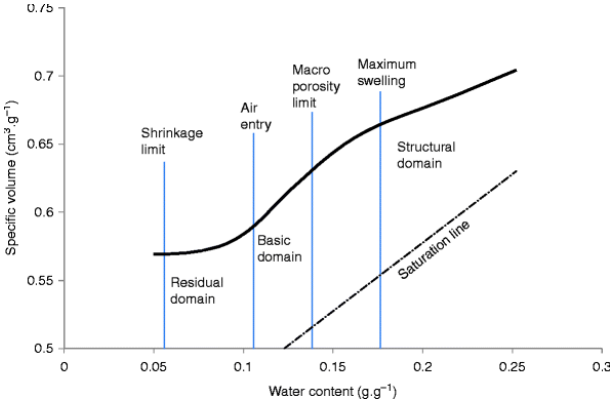

Figure 3.7 Shrinkage and swelling phenomena in soils (Boivin P. , 2011)

3.2.2 SOIL SETTLEMENTS

Compressibility is the volumetric deformation suffered by soils. Volume reduction does not mean a reduction in the volume of solids, it is the reduction of soil voids. Different soils have different reactions which are schematically represented in Figure 3.8.

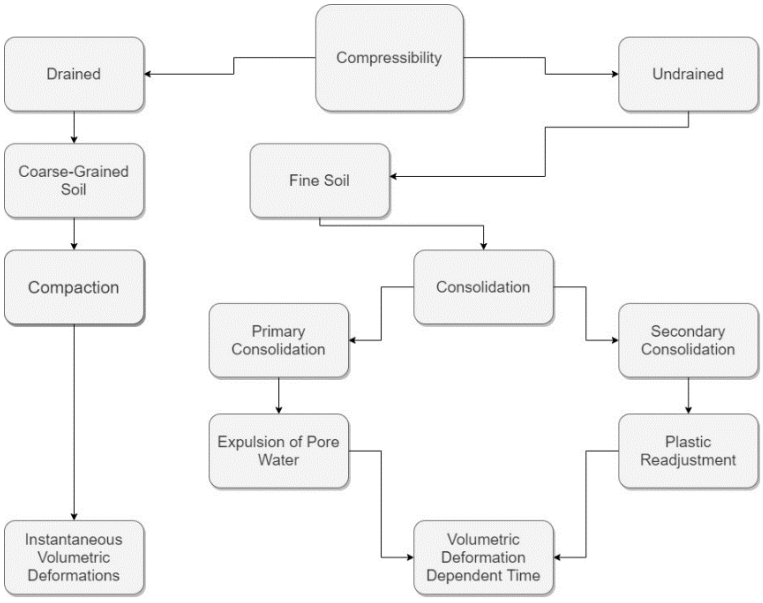

Figure 3.8 Soils compressibility

Coarse soils, such as sand or gravel, are soils with high permeability. When subjected to loads, volumetric deformations occur instantly due to fast dissipation on pore water pressures and volume reduction of these pores. This process is called compaction, where settlements on soil surface occur almost immediately and stay constant over the years. These are typically called drained loadings.

Fine soils, like silts or clays, are soils with very low permeability, making water expulsion generally very slow. For this reason, volumetric deformations, which cause settlements on the soil surface, can take extended periods. These are called undrained loadings.

The total soil settlement can be described as the sum of three settlements parameters.

$$\rho_t = \rho_i + \rho_c + \rho_s \tag{3.4}$$

where:

ρ_t – Total soil settlement;

ρ_i – Immediate settlement;

ρ_c – Primary consolidation settlement;

ρ_s – Secondary consolidation settlement.

Immediate settlement is not time-dependent, this means, when it occurs, it is immediate. This settlement occurs due to shear stresses with constant volume while loads are being applied to the soil. Primary consolidation settlement and secondary consolidation settlement are time-dependent. Fine saturated soils have an increase of pore water pressure when loaded. The dissipation of those excess pore pressure is the cause of volume changes leading to the primary consolidation settlement. Secondary consolidation is due to the effect of time-dependent stress-strain behaviour of soil structural viscosity (Rujikiatkamjorn, 2005). Two different studies show different approaches to secondary consolidation. The first approach assumes that the secondary consolidation occurs after the end of primary consolidation (Mesri & Choi, 1985). The other study defends that creep behaviour occurs during the entire primary consolidation process (Tatsuoka, 2002).

Since total settlement occurs over time, the time-settlement curve can be represented, as shown in Figure 3.9.

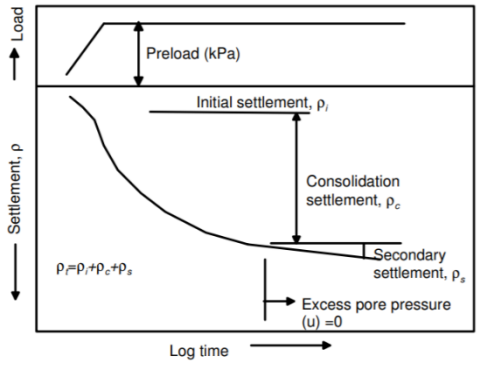

Figure 3.9 Typical components of settlements (Rujikiatkamjorn, 2005)

At a given depth, the relationship between the average void ratio and the logarithm of average effective stress has different approaches for different consolidation phases. In Figure 3.10 is illustrated an over-consolidated soil submitted to an additional effective vertical stress ($\Delta\sigma'_v$), which, accumulating the initial vertical stress (σ'_{v0}), exceeds the pre-consolidation stress (σ'_p). Two consolidation phases may be identified in Figure 3.10:

- Between σ'_{v0} and $\sigma'_{p,}$ the soil suffers compressibility in stresses that it has already experienced, represented by the recompression section whose slope is given by the recompression index (C_r);
- Between σ'_{p} and $\sigma'_{v0} + \Delta\sigma'_{v}$, the soil is experiencing virgin consolidation, characterised by the compression index (C_c).

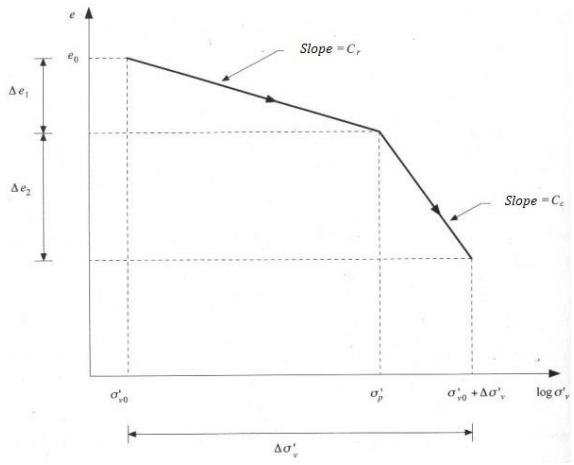

Figure 3.10 Void ratio variation due to additional vertical stress (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006)

The void ratio variation associated with the compressibility in the two branches results in (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006):

$$\Delta e_1 = -C_r log \frac{\sigma'_p}{\sigma'_{vo}} \tag{3.5}$$

$$\Delta e_2 = -C_c log \frac{\sigma'_{vo} + \Delta\sigma'_v}{\sigma'_p} \tag{3.6}$$

In radial drainage, the horizontal soil permeability decreases with the average void ratio. The relationship between these two parameters results in the following equation and visible in Figure 3.11.

$$e = e_o + C_k \log\left(\frac{k_h}{k_{hi}}\right) \tag{3.7}$$

The C_c is determined from consolidation tests. The C_{ec} is a simplification of the compression index Cc considering the initial void ratio of the soil in the study. C_{ec} is calculated considering the following formula:

$$c_{ec} = \frac{c_c}{(1 + e_o)} \tag{3.8}$$

where:

e_0 is the initial void ratio.

The permeability index (C_k) is generally independent of stress history.

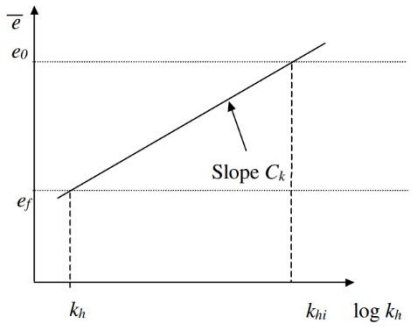

Figure 3.11 Semi-log permeability-void ratio relationship (Rujikiatkamjorn, 2005)

Consolidation settlements are generally the dominant component of the total settlement calculation in normally consolidated fine-grained soils. Its magnitude is significant, and it takes a long time to stabilize even after the construction is completed.

The consolidation rate depends on soil properties, namely:

- Soil compressibility and permeability;
- The initial distribution of excess pore pressure;
- Drainage of soil boundary conditions.

During the consolidation process, it is possible to conclude that pore pressure is decreasing, effective stress is increasing, and settlement is also increasing.

The original assumptions in Terzaghi consolidation theory are (Fernandes, Mecânica dos Solos Introdução à Engenharia Geotécnica Volume 2, 2011):

- Soil is homogenous and isotropic;
- Soil is saturated;
- Water and solids are incompressible;
- Darcy's law applies;
- The flow and consolidation are one dimensional in the vertical direction;
- Continuity on soil parameters;
- Compressibility behaviour.

Defining the consolidation coefficient (c_v) as the material property that quantifies how fast a soil will consolidate, the next equation was obtained:

$$\frac{\partial u_e}{\partial t} = c_v \frac{\partial^2 u_e}{\partial z^2} \tag{3.9}$$

$$c_v = \frac{k}{m_v \gamma_w} \tag{3.10}$$

where:

u_e – Excess pore pressure;

m_v – Volumetric compressibility;

k – Soil permeability coefficient;

t – Time;

z – Soil depth.

This deduction is valid when the soil layers have only one draining boundary which does not happen in every case scenario. To consider other scenarios for single or double drainage boundary conditions in alternative to the magnitudes of *t* and *z* it is healthier to work with other sizes, directly proportional but dimensionless resulting in the following expressions (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006):

$$Z_d = \frac{z}{H} \tag{3.11}$$

$$T = \frac{c_v}{H^2} \tag{3.12}$$

where:

Z_d – Depth factor;

T – Time factor;

H – The more considerable distance that water needs to travel to a direct drained boundary.

With these rearrangements, equation 3.9 results in:

$$\frac{\partial u_e}{\partial t} = \frac{\partial^2 u_e}{\partial Z^2} \tag{3.13}$$

Through Figure 3.12, it is possible to obtain the different parameters, considering that initial excess pore pressure is constant over layer depth.

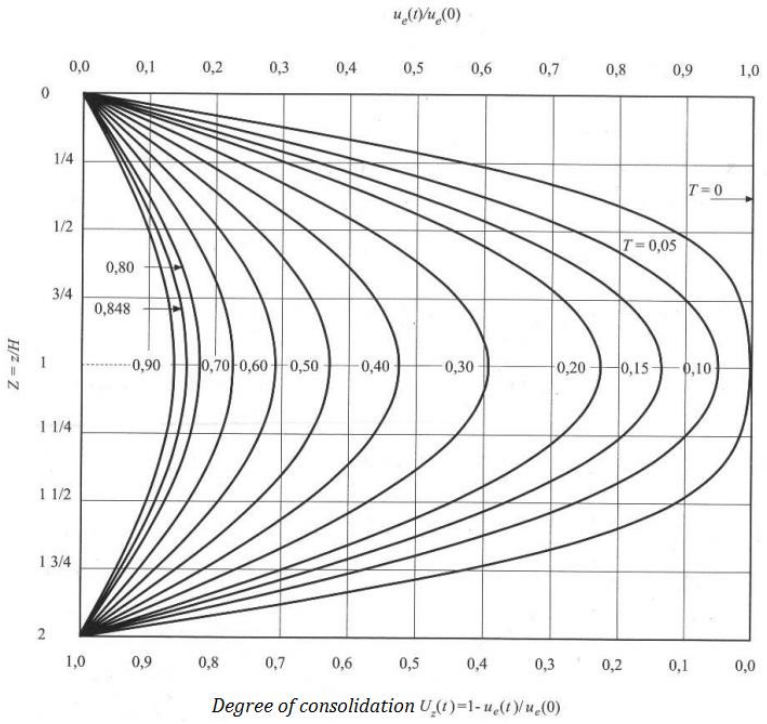

Figure 3.12 Solution of consolidation equation (Fernandes, Mecânica dos Solos Conceitos e Princípios Fundamentais, 2006)

The consolidation of saturated soft clay is a procedure that can take more than twenty-five years to be fully accomplished. Thus, there are methods and technologies to mitigate the long-time process into several years or months according to each situation.

3.2.3 EVALUATION OF CONSOLIDATION PARAMETERS

3.2.3.1 Oedometric Modulus

The Oedometric modulus (E_{oed}) is calculated either directly from oedometer tests, or using the following correlation with qc measured from CPT:

$$E_{oed} = \alpha * qc \tag{3.14}$$

The coefficient α is given in the following table presented in Table 3.1.

Table 3.1 Coefficient α for determining E_{oed} from CPT qc (Building on soft soils: design and construction of earth structures both on and into highly compressible subsoils of low bearing capacity, 1996)

Cone resistance (MPa)	Water content (%)	Type of soil	α coefficient	
			Mechanical cone	Electrical cone
< 0.7		Clay with low plasticity	3.0-8.0	3.7-10
0.7-2.0			2.0-5.0	2.5-6.3
> 2.0			1-2.5	1.25-3.0
1.2-2.0		Silt with low plasticity	3.0-6.0	3.5-7.5
> 2.0			1.0-3.0	1.25-3.7
< 2.0		Clay and silt with high plasticity	2.0-6.0	2.5-7.5
< 1.2		Organic silt	2.0-8.0	2.5-10
< 0.7	50-100	Peat and organic clay	1.5-4.0	
	100-200		1.0-1.5	
	> 200		0.4-1.0	
2.0-3.0		Gravel	2.0-4.0	
> 3.0			1.5-3.0	
< 5.0		Sand	2	
> 10.0			1.5	

3.2.3.2 Coefficient of Consolidation

The coefficient of consolidation, c_v, can be determined directly from consolidation tests. However, this method usually underestimates the coefficient's value since it does not capture the large-scale soil structure such as fissures and the permeability anisotropy due to the depositional structure of clays. Besides, the laboratory-derived values are very sensitive to sample disturbances. Consequently, the real large-scale value of c_v might be 10 to 100 times larger than the one derived from consolidation tests.

Due to the limitations mentioned above of laboratory-derived c_v values, the laboratory results are crosschecked with estimates coming from other sources, such as correlations with liquid limit, visible in Figure 3.13.

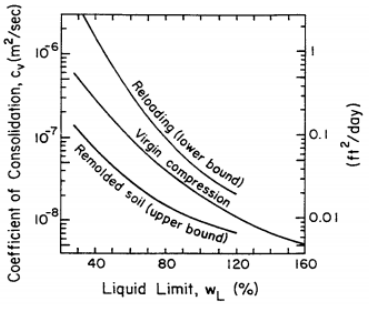

Figure 3.13 Coefficient of consolidation as a function of liquid limit (Kulhawy & Mayne, 1990)

The value of c_v can be determined from vertical soil permeability, k_v, as given by the following relationship:

$$c_v = \frac{k_v * E_{oed}}{\gamma_w} \tag{3.15}$$

The permeability can be calculated by in-situ or laboratory tests or estimated from correlations with soil classification, Table 3.2.

Table 3.2 Vertical permeability as a function soil type (Burt, 2007)

Soil type	Description	USC-symbol	Permeability (m/s)
Gravels	Well graded	GW	10^{-3} to 10^{-1}
	Poorly graded	GP	10^{-2} to 10
	Silty	GM	10^{-7} to 10^{-5}
	Clayey	GC	10^{-8} to 10^{-6}
Sands	Well graded	SW	10^{-5} to 10^{-3}
	Poorly graded	SP	10^{-4} to 10^{-2}
	Silty	SM	10^{-7} to 10^{-5}
	Clayey	SC	10^{-8} to 10^{-6}
Inorganic silts	Low plasticity	ML	10^{-9} to 10^{-7}
	High plasticity	MH	10^{-9} to 10^{-7}
Inorganic clays	Low plasticity	CL	10^{-9} to 10^{-7}
	High plasticity	CH	10^{-10} to 10^{-8}
Organic	with silts/clays of low plasticity	OL	10^{-8} to 10^{-6}
	with silts/clays of high plasticity	OH	10^{-7} to 10^{-5}
Peat	Highly organic soils	Pt	10^{-6} to 10^{-4}

3.2.4 METHODS TO ACCELERATE CONSOLIDATION

Since consolidation settlements are delayed in time, this is a very problematic issue in construction sites. In some constructions methods to accelerate consolidation must be implemented, and this leads to additional advantages, schematically observable in Figure 3.14. For that reason, these methods are sometimes considered in the soil improvement category. The methods covered in this work are preloading, combined with prefabricated vertical drains.

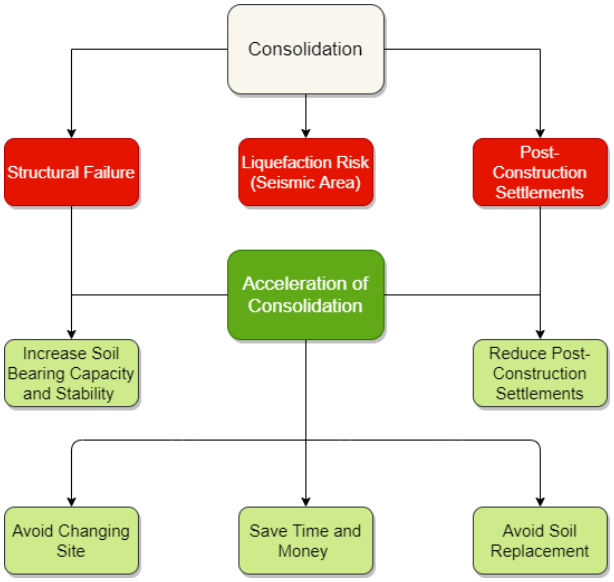

Figure 3.14 Advantages of acceleration of consolidation

The typical acceleration of consolidation methods used in land reclamation projects are the following:

- Stone columns
- Prefabricated vertical drains
- Preloading
- Vacuum preloading

In this work, the preloading combined with PVDs will be detailed below focusing on the design procedures.

3.2.4.1 Preloading with Prefabricated Vertical Drains

Preloading application is a classic and widespread method used to boost consolidation settlement. Preloading is the application of an overload on the soil to accelerate primary consolidation process (Chu, Varaksin, Ultich, & Mengé, 2009). While the application of preloading alone may considerably reduce the consolidation time in some soil deposits, in some other cases (especially if the fine-grained layers are very thick and with very low permeability), these techniques need to be associated to other methods. Prefabricated vertical drains (PVD) associated with preloading accelerates soil radial drainage and consolidation by decreasing the drainage path in the radial direction, as visible in Figure 3.15.

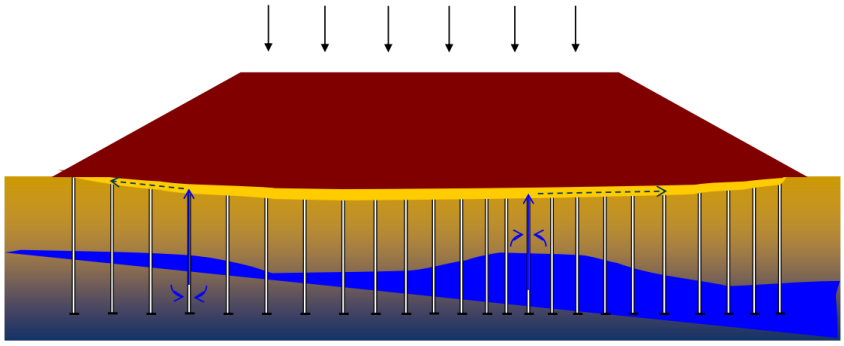

Figure 3.15 Preloading with vertical drains (Chu, Varaksin, Ultich, & Mengé, 2009)

In reclamation projects, the consolidated soil is on the sea bed (thus completely saturated) but still it generates excess pore pressures above the hydrostatic pressure due to soil placed above that is dissipated during consolidation. So, the preloading with vertical drains will improve the flow of water from the sea bed to the soil above balancing the pressure until the hydrostatic pressures are re-established in all points.

As previously shown, PVDs have a rectangle shape. The available theories of radial consolidation have been derived for drains having a circular shape. To be able to apply the theories to the design considering PVDs, the equivalent diameter of the PVDs shape must be defined (Barron, 1948). The adaption process is visible in Figure 3.16. PVDs rectangle shape is defined through a width, *b*, and a drain thickness, *t*. The main focus is to obtain the equivalent diameter of the drain, d_w, and the drain influence diameter, d_e.

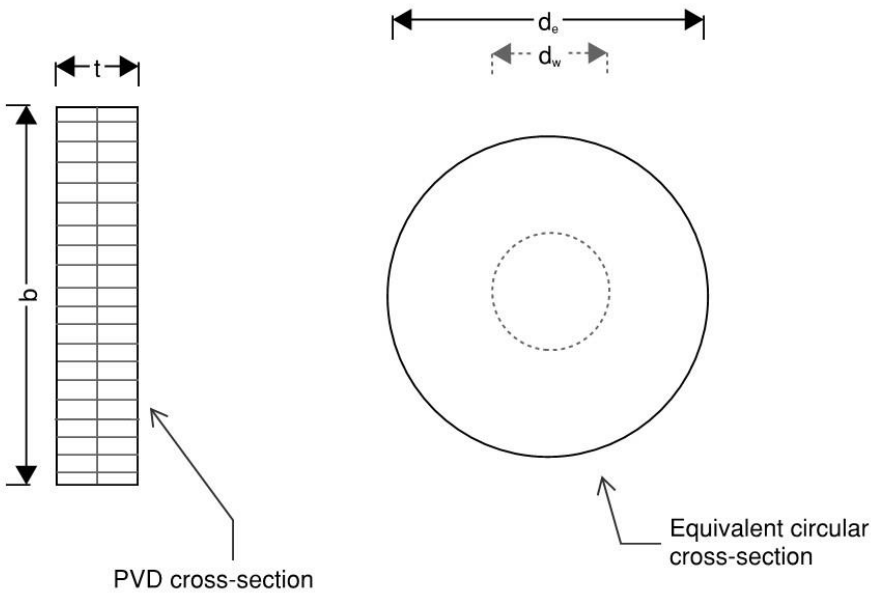

Figure 3.16 PVD cross-section

Hansbo (1979) used a finite-element analysis and concluded that the equivalent diameter of a drain is a result of the following equation:

$$d_w = \frac{2(b+t)}{\pi} \tag{3.16}$$

The drain influence zone, d_e, is a function of drain spacing, d_s. The typical drains grid geometry is in square or triangular patterns, observable in Figure 3.17. Square pattern layouts have greater ease and control in the field. However, the triangular patterns are preferred to provide more uniform consolidation between drains (Holtz, 1987).

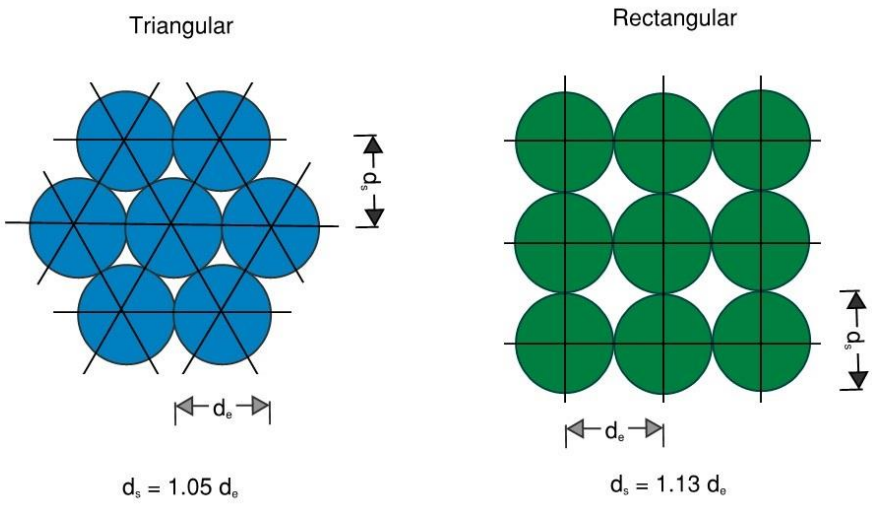

Figure 3.17 Drains grid geometry

The drain and its zone of influence, named as the unit cell, are assumed to be circular in plan with radius r_c for the unit cell (Figure 3.18). The disturbed zone is divided into two distinct parts (Basu, Basu, & Prezzi, 2013):

- Smear zone with a radius r_s;
- Transition zone with an outer radius r_t.

A smear zone is the disturbed zone after PVDs installation, namely when the mandrel is pushed through the clay displacing the soil material. Its definition is crucial because of its different hydraulic conductivity properties, namely the reduced lateral permeability in the smear zone.

Two parameters are essential to describe the smear effect, specially, the diameter of the smear zone (d_s) and the permeability ratio (k_h/k_r), represented through, the value in the undisturbed zone (k_h) over the smear zone (k_r). Both the diameter of the smear zone and its permeability are difficult to quantify in laboratory tests, and, so far, there is no comprehensive or standard method of measuring them. The

extent of the smear zone and its permeability vary with the installation procedure, size, and shape of the mandrel (Indraratna, Sathananthan, Bamunawita, & Balasubramaniam, 2015).

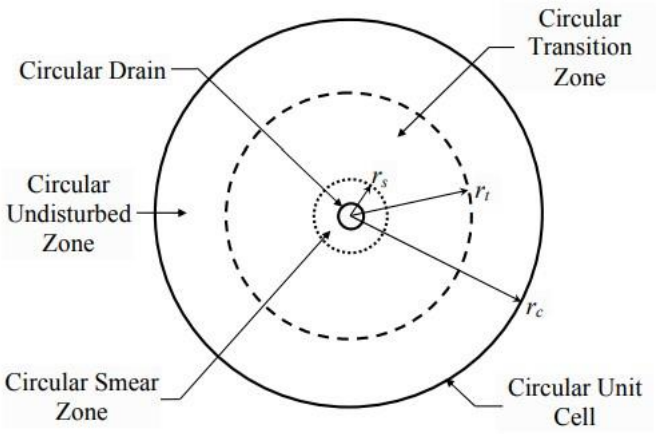

Figure 3.18 Unit cell with disturbed zone (Basu, Basu, & Prezzi, 2013)

Well resistance refers to the finite permeability of the vertical drain concerning the soil. Head loss occurs when water flows along the drain and delays radial consolidation. Well resistance is controlled not only by the discharge capacity of the drain q_w, but also by the permeability of the soil k_h, the maximum discharge length l_m, and any geometric deficiencies (bending, kinks, etc.) on the drains (Indraratna, Sathananthan, Bamunawita, & Balasubramaniam, 2015).

Analytical solutions already developed for consolidation of ground improved with vertical drains invariably contain the “unit cell” model, as represented in Figure 3.19.

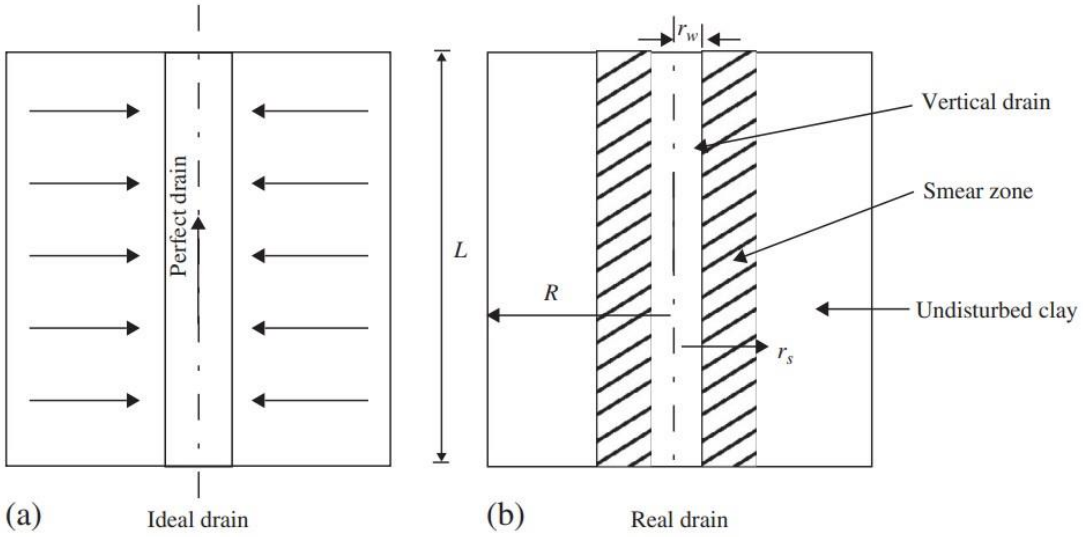

Figure 3.19 Unit-cell model of a drain surrounding by soil cylinder (Indraratna, Sathananthan, Bamunawita, & Balasubramaniam, 2015)

Through the installation of a PVD mesh in a loaded fine soil layer, consolidation runs vertically and horizontally towards the drains. That spatially averaged degree of consolidation is expressed in function of time, *t* (Carrillo, 1942) defined by Carillo (1948) by the following equation:

$$U(t) = 1 - [1 - U_v(t)][1 - U_h(t)] \tag{3.17}$$

Hansbo (1981) presented an approximate solution for vertical drain based on the equal strain by considering both smear and well resistance, a complement of Barron (1948) research. The average horizontal or radial consolidation rate is defined by the next equation:

$$\bar{U}_h(t) = 1 - \exp\left(-\frac{8c_h t}{d_e^2 F}\right) \tag{3.18}$$

$$F = \ln\left(\frac{d_e}{d_r}\right) - 0.75 + \pi\left(\frac{2}{3}\right) l^2 \frac{k_h}{q_w} + \frac{k_h}{k_r} \ln\left(\frac{d_r}{d_w}\right) \tag{3.19}$$

where:

C_h – Horizontal coefficient of consolidation;

d_e – Radius of the influence zone of a PVD;

F – Effect of drain spacing, soil disturbance and well resistance.

Vertical consolidation is based on a time factor, T_v. The close relationships between T_v and U_v are the following:

$$Uv \leq 60\ \%\ (Tv \leq 0.471) \tag{3.20}$$

$$U_v = 100 * \sqrt{\frac{4T_v}{\pi}} \tag{3.21}$$

$$Uv > 60\ \%\ (Tv > 0.471) \tag{3.22}$$

$$U_v = 100 * 10^{\frac{1.781 - T_v}{0.933}} \tag{3.23}$$

3.2.5 PROBABILISTIC DESIGN AND MONITORING OF SETTLEMENT FOR RECLAMATIONS

3.2.5.1 Probabilistic Design Procedure

To have a probabilistic design, primarily, all the geotechnical parameters must be characterised probabilistically. After characterisation, a distribution of the primary consolidation settlement, named *S* in Figure 3.20, is determined. Secondly, s_{target} is determined such that post-completion primary compression occurs only with the predefined target probability p_{FT} for this serviceability limit state (Akbas & Kulhawy, 2009). Thirdly, the degree of consolidation at t_{max} is assessed probabilistically for some initial PVD design. Fourthly, for a range of surcharge heights, h_{sur}, the corresponding settlement and OCR at t_{max} are assessed probabilistically (Spross, Prästings, & Larsson, Probabilistic Evaluation of Settlement Monitoring with the Observational Method during Construction of Embankments on Clay, 2019). The initial surcharge height is then a decision to the engineer, predicted on the calculated probabilities respecting the following two design conditions:

$$P(S_{t_{max}}^{sur} \geq s_{target}) \geq p_{acc} \tag{3.24}$$

$$P(OCR_{t_{max}}^{sur} \geq 1.1) \geq p_{acc} \tag{3.25}$$

Where $S^{sur}{}_{tmax}$ is uncertain predictions of the settlement, $OCR^{sur}{}_{tmax}$ is the OCR at the end of the preloading time, and p_{acc} is the acceptable probability with which the design criteria are satisfied (Spross, Prästings, & Larsson, Probabilistic Evaluation of Settlement Monitoring with the Observational Method during Construction of Embankments on Clay, 2019). In Figure 3.20, these principles are graphically represented. Essentially, the supposition of the initial surcharge height has a critical impact since the savings of selecting a low surcharge cannot occur when having to raise the surcharge during the preloading.

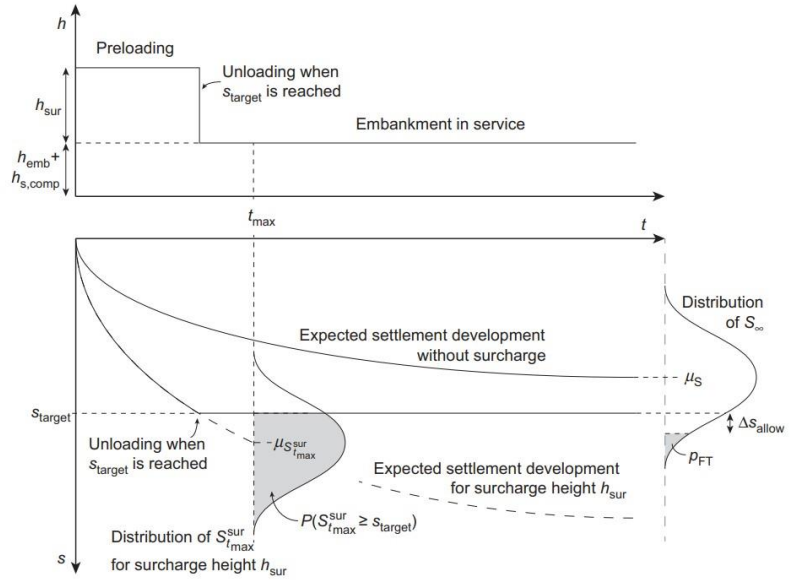

Figure 3.20 Conceptual idea of the design procedure. Top: embankment height plotted against time. Bottom: developed settlement plotted against time. The S_∞ is used to determine the s_{target} value. To ensure that ΔS_{allow} is only exceeded with p_{FT}, a surcharge height h_{sur} is selected so that the s_{target} value and OCR = 1.1 are attained within t_{max} with acceptable probability (Spross & Larsson, Probabilistic Observational Method for Design of Surcharges on Vertical Drains, 2019)

3.2.5.2 Updating predicted settlements with Bayesian statistics by monitoring

Bayesian view on statistics can accurately improve outcomes quality. In Figure 3.20 example, the new prediction regarding the settlement at t_{max} acquired from the measurements can be used to update the prior mean, $\mu'_{s,tmax}$, and variance, $\sigma'^2_{s,tmax}$, of S'^{sur}_{tmax} into a posterior distribution of the expected settlement by the end of the preloading, named S''^{sur}_{tmax}. To that procedure, some conventions must be considered (Spross, Prästings, & Larsson, Probabilistic Evaluation of Settlement Monitoring with the Observational Method during Construction of Embankments on Clay, 2019):

1. Prior measurements are normally distributed.
2. The difference between the estimated and the real error variance in the regression is negligible.

Regarding these assumptions, S''^{sur}_{tmax} can be described in terms of mean and corresponding variance (Alfredo & Tang, 2007):

$$\mu''_{s,t_{max}} = E\left(S^{sur}_{t_{max}} \middle| Z\right) = \frac{s_{t_{max}}\sigma'^2_{s,t_{max}} + \mu'_{s,t_{max}}(\sigma^2_{s,t_{max}}/n)}{\sigma'^2_{s,t_{max}} + (\sigma^2_{s,t_{max}}/n)} \tag{3.26}$$

$$\sigma''^2_{s,t_{max}} = \frac{\sigma'^2_{s,t_{max}}(\sigma^2_{s,t_{max}}/n)}{\sigma'^2_{s,t_{max}} + (\sigma^2_{s,t_{max}}/n)} \tag{3.27}$$

Monte Carlo can be used when the prior measurements do not follow a normal distribution.

With Bayesian statistics, the measured settlements are used to update the prior prediction of the final settlement. These approaches reduce uncertainty related to the first settlement calculated at the design phase. Based on the result of Bayesian updating, through the monitoring observations, a prepared contingency action to increase the surcharge load can be put into operation in due time (Spross, Prästings, & Larsson, Probabilistic Evaluation of Settlement Monitoring with the Observational Method during Construction of Embankments on Clay, 2019), so the target settlement is reached in the preloading time. Bayesian allows making earlier adjustments, through the real-time interpretation ground conditions reducing time and costs to the desired target.

3.3 CONSTRUCTION MONITORING

Uncertainty is a daily basis in construction, even more in geotechnical engineering where the soil is heterogeneous. Construction monitoring is an accurate and helpful technique of controlling and examining the quality of a construction project and can anticipate difficulties before it occurs. In situ monitoring can optimize construction methods or just confirm design expectations. In some cases, real-time monitoring systems can evaluate each circumstance. To this case study, some relevant monitoring systems need to be in concern, such as:

- Extensometers;
- Piezometers;
- Inclinometers;
- Remote Sensing.

3.3.1 EXTENSOMETERS

Extensometers are helpful to measure vertical displacements in soil. A Polyvinyl chloride (PVC) tube is introduced in a borehole where the extensometers are installed. In Figure 3.21, the installation is represented and in Figure 3.22, the measuring scheme is presented. In Step 1, a borehole is drilled in the ground. In Step 2, the anchors are inserted. While on step 3, the anchors are expanded to lock in with the surrounding soil. In Step 4, Figure 3.22, the joint rock moves and the distance between the two anchors changes. The evolution of length between the two anchors will be recorded in the data logger (Rajapakse, 2016).

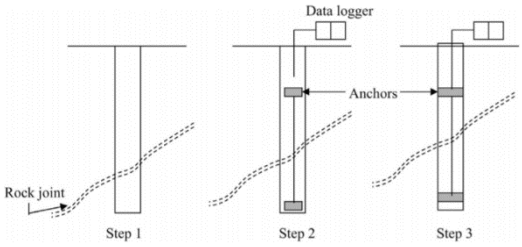

Figure 3.21 Extensometer installation procedure (Rajapakse, 2016)

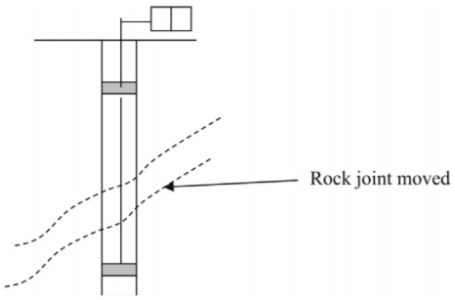

Figure 3.22 Extensometer in action (Rajapakse, 2016)

3.3.1.1 Magnetic Extensometer

The magnet extensometer system is a type of extensometer. It is composed by a probe, a graduated cable, a tape reel with built-in light and buzzer, and some magnets positioned along the length of an access pipe. The magnets are coupled to the surrounding soil, and they move up or down as heave or settlement occurs. Readings are obtained by driving the probe through the access pipe to find the depth of the magnets. When the probe enters a magnetic field, a reed switch closes, activating the light and buzzer. The operator then refers to the graduations on the cable and notes the depth of the magnet. When the access pipe is anchored in the stable ground, the depth of each magnet is referenced to a datum magnet, which is fixed to the bottom of the access pipe. If the bottom of the access pipe is not in the stable ground, the depths of the magnets must be referenced to the top of the pipe, which is optically surveyed before readings are taken (Magnet, 2002).

The magnetic extensometers components are the following, visible in Figure 3.23:

- Access Pipe;
- Datum Magnet;
- Spider Magnet;
- Plate Magnet.

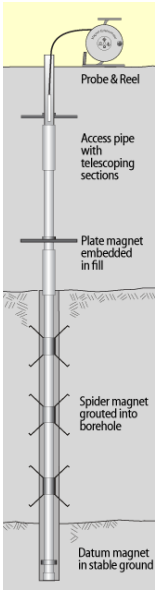

Figure 3.23 Magnetic extensometer (Magnet, 2002)

3.3.2 PIEZOMETERS

3.3.2.1 Standpipe

Pore pressure is an important parameter that affects soil behaviour when subjected to a different type of stresses. This parameter can be measured through piezometers. The principal bases of piezometers operation are elementary:

- The porous structure is placed on the layer point where the measure is needed;
- Soil water flows through the porous structure and enters a compartment (typically a tube);
- The height reached by water in the pipe represents the pore pressure in that specific soil point.

The simplest type is called standpipes. The pressure of the groundwater pushes water into and up the standpipe as represented in Figure 3.24. The level of water inside the standpipe is equivalent to the pore water pressure in the ground at the elevation of the porous filter (Piezometers, 2019).

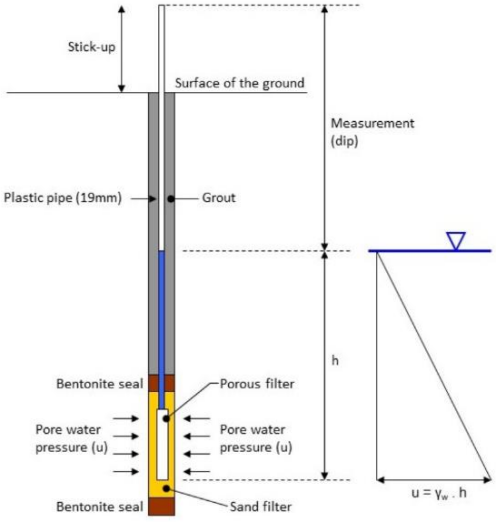

Figure 3.24 Standpipe piezometer principle (Piezometers, 2019)

3.3.2.2 Vibrating Wire

Electric piezometers are typically used in fine soils with low permeability, schematically represented in Figure 3.25. Electric piezometers consist of a deflecting diaphragm and a porous filter separated by a small reservoir of water. Deflections of the diaphragm are detected using vibrating wire and are converted to an equivalent pressure using a suitable calibration. The piezometer is inserted into a borehole. The annulus between the porous filter and the borehole is filled with either sand or cement/bentonite grout. Water from the ground forces its way into the reservoir. It causes the diaphragm to deflect until the pressure inside the reservoir is the same as the pore water pressure in the ground at the elevation of the porous filter (Piezometers, 2019). These types of piezometers can measure lower pressures than atmospheric pressure. By isolating the transducer with impermeable material, allows measuring pore pressure at different heights, as also use more piezometers in the same borehole.

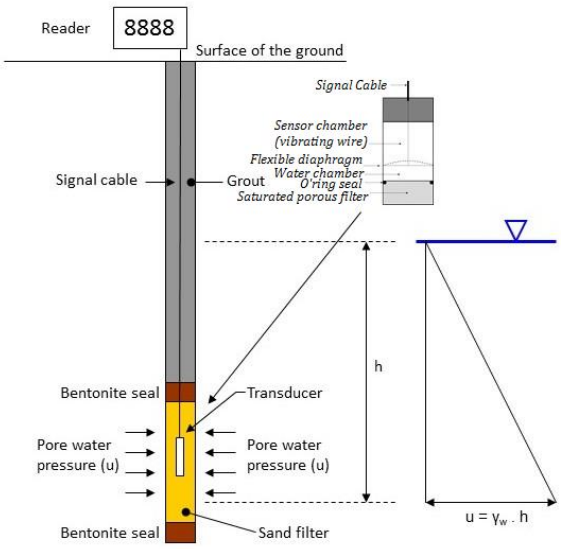

Figure 3.25 Vibrating wire piezometer principle (Piezometers, 2019)

3.3.3 INCLINOMETERS

Inclinometers are used in slope stability monitoring by measuring horizontal displacements over various points. A slope inclinometer is a wheeled instrument with a probe. A fixed casing is installed and fixed in a borehole so it would only move if soil also moves. The inclinometer is inserted into the borehole. It sends signals to the data logger about the verticality of the borehole, picturing an outline of the borehole. When the slope does not suffer any movement, the outcome data is a vertical line, as shown in Figure 3.26. When slope has moved, the inclinometer shows on the vertical line the deformation that occurred in the borehole, represented in Figure 3.27.

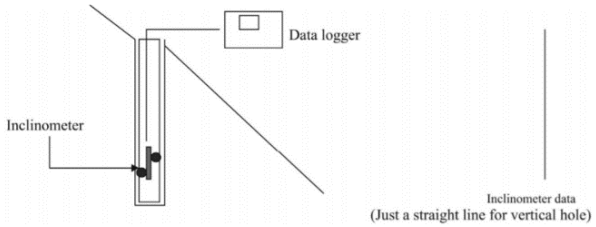

Figure 3.26 Inclinometer with slope stable (Rajapakse, 2016)

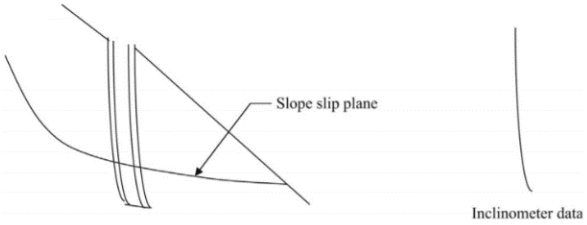

Figure 3.27 Inclinometer with slope movement (Rajapakse, 2016)

When the inclinometer is inclined as shown in Figure 3.28, by an angle α, the pendulum is closer to detector *A* than to detector *B*. Hence detector *A* would record a higher reading than detector *B*. This electrical potential difference can be utilised to obtain the angle of inclination (Rajapakse, 2016). The point where the angle inclination is located is one representative point of the slope failure.

Through the installation of various inclinometers at different depths and distances like represented in Figure 3.29, it is possible to predict slope failure surface by measuring slope movements that occur before the slope failure.

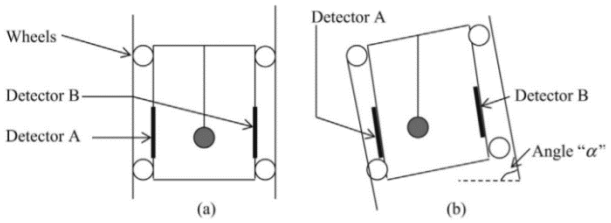

Figure 3.28 Working of the inclinometer. (a) Vertical inclinometer, (b) inclined inclinometers (Rajapakse, 2016)

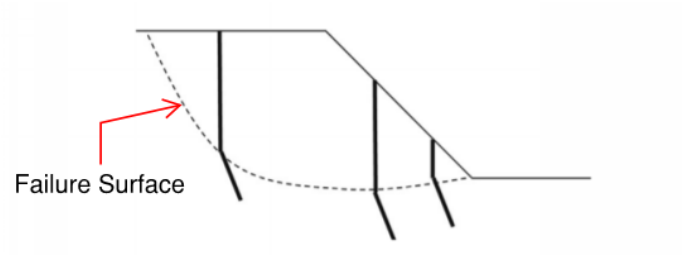

Figure 3.29 Inclinometers slope failure surface (Rajapakse, 2016)

3.3.3.1 Shape Array (SAA)

Shape arrays are advanced inclinometers used for monitoring settlement, movement and vibration. A shape array is a chain of rigid segments connected by flexible joints, visible in Figure 3.30. The joints are designed to resist twist but allow the segments to tilt in any direction. Shape arrays can be installed in a variety of ways (Shape accel arrays, u.d.):

- Installed in existing inclinometer casings to automate readings;
- Inserted into a small diameter pipe and installed directly into boreholes;
- Fixed to tunnel linings or concrete slabs.

Each segment is instrumented with three orthogonally mounted tilt sensors and a microprocessor. The microprocessor calculates the XYZ position of the segment based on the length of the segment and measurements from the tilt sensors. The overall shape of the array is found by cumulating the XYZ positions of the connected segments. Initial measurements serve as a baseline. Subsequent measurements, obtained at regular intervals, are compared to the baseline. Changes indicate that deformation has occurred and reveal the direction and magnitude of deformation. Shape array measurements are forwarded to automated monitoring platform which processes the measurements, checks for alarms, and posts plots and plans views on the project website fulltime.

Figure 3.30 Shape array inclinometer (Shape accel arrays, u.d.)

3.3.4 REMOTE SENSING

Nowadays, more methods are available to visualise, control and plan construction sites. Geographic information systems (GIS) can gather, store, manage and analyse spatial or geographic data. With satellites or drones, the process becomes even more accessible. Remote sensing can map all the area in the study and collect primary data from the surface of the earth. The combination of these sources and methods allows engineers to decide more accurately where to build, the conditions of the surrounding area or to control different conditions such as ground settlements. When the main subject is soil improvement, to measure and monitor deformations on earth surface with reliable accuracy, differential interferometric synthetic aperture radar (InSAR) is consistent and straightforward practice. Detecting changes in the position of the Earth's surface requires two radar images of a selected area taken from approximately the same position in space but at two different times. By bouncing signals from a radar satellite off the ground in successive orbits and looking at the differences between the images, InSAR can detect small differences in the distance between its position and the ground as the land surface moves. Interferometry is based on processing the pair of images to map out the differences in the reflected signals over the area. The amount and pattern of deformation in an interferogram are shown by using the range of colours in the spectrum from red to violet. Figure 3.31 shows an interferogram documenting subsidence in the Santa Clara Valley of California (upper image) and a shaded-relief map (lower image) correlating colour bands with the deformation pattern (Helz, 2005).

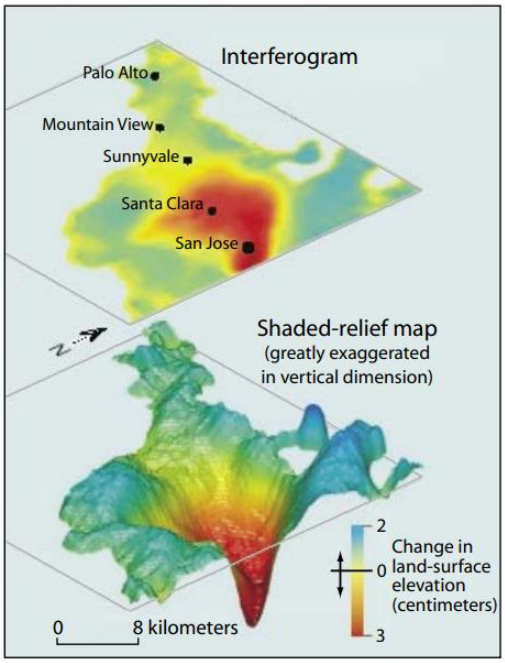

Figure 3.31 InSAR Interferogram and shaded-relief map illustration (Helz, 2005)

The main benefits of InSAR for monitoring stand through the following facts:

- Reliability;
- Simplicity;
- Low-cost;
- Measure large areas;
- Weather independent;
- Daylight independent.

4
Case Study Description

4.1 PORT EXPANSION

The present case study consists in the expansion of a marine port, in a country situated in Southeast Europe. The port is situated in an advantageous geographical location. This port is one of the essential ports in Southeast Europe, serving millions of inhabitants of its international mainland.

This project aims to collect soil from land to fil into the sea space in order to expand the pier area so that the ships can dock, and more containers can be accommodated.

For soil reclamation, it is imperative to know the ground conditions up to twenty or thirty meters of depth. The primary issue in a reclamation project is to know how big the settlements will be and how long they will take to develop. Knowing remaining settlements is crucial to recognize how soon the port can operate.

4.2 SITE INVESTIGATION

The data from five geotechnical surveys performed *in situ* have been evaluated. Four of them contain sampling boreholes from geotechnical surveys executed in 1970, 1972, 1997 and 2015. A fifth survey that only contains cone penetration tests (CPTs) was executed in 2000 as part of construction at that time. These five surveys include in total 56 rotary continuous sampling boreholes and 11 electrical cone CPTs with SPTs.

An additional survey was executed in September 2019, as part of the design scope. It included four additional rotary continuous sampling boreholes, comprising three offshore and one on land. Also, six CPT with mechanical cone were performed.

4.3 GEOTECHNICAL CONDITIONS

4.3.1 GEOLOGY

The major part of the city area extending from the port to the south-eastern waterfront suburbs is located on Quaternary-Neogene formations. Recent deposits of this particular region, covering the Neogene formation, are Holocene clays, sands and pebbles and they have formed an extensive plain.

The Holocene sediments from deltaic terrains of four principal rivers caused an approximately 80 km shift of primordial coastline by progressively delivering sediments and shallowing waters. The project region is formed from the continuous deposition of deltaic sediments, resulting in a complex stratigraphy. The stratigraphy of the plain in the area between the four rivers corresponds to new

complex laterally variable delta formed in the last few thousand years, in a more than 600 meters deep Quaternary graben, cut in Neogene.

The project area has high seismic activity during recorded history. The described geological situation creates a deltaic environment where sedimentation comes from two sources - river and sea. That has an impact on its stratigraphy (e.g. faults or irregular layers/lenses) and sensitivity (e.g. liquefaction) of local soils. Deltas form interbedding layers of finer, cohesive material and coarser material, further interbedded with marine deposits. That is reflected in the top stratigraphy of the Gulf, which is composed of mainly fine, very soft to firm clayey and silty material. The bedrock (metamorphic gneiss, epigneiss and green schists) is located at an immense depth, which was not reached by any of the new and existing investigations. Based on the available geological and geophysical studies for the broader project area it is estimated that very soft to firm sand/silty clays reach up to 100-150 m below the seabed.

4.3.2 STRATIGRAPHY

Based on the previous and supplementary investigations, the following four main layers have been determined:

1. Fill
   Appears in the project area due to reclamation works performed for the construction of the Pier during the last 40 years. The fill is composed of a mixture of all type of soils starting from fat clays and ending in coarse gravels/pebbles. The fill layer is up to 20 m thick in the Pier project area.
2. Very Soft Silt
   This layer starts at the seabed, and it is composed of very soft to incredibly soft deltaic deposits. Black or very dark grey SILTs and clayey SILTs, locally with very silty CLAYs were encountered within this layer. The thickness of this layer in the project area varies from 2.3 to 9.5 meters. *In situ* tests performed in this layer present extremely low to very low strength parameters, where Standard penetration test (SPT) N-values are between zero and four blows per 300mm. In contrast, no cone penetration resistance (qc) was recorded by CPT tests, due to very low resistance of this layer.
3. Sand/Gravel Lenses
   Alluvial delta deposits characterized by clayey, gravelly either well-graded sands and clayey, sandy either well-graded gravels. A typical example of delta deposits in more energetic river seasons. The SPT N-values in these lenses varies between 14 and >50.
4. Predominantly Fine-Grained Delta/Marine Deposits
   This layer is represented by various mixtures of fine- and coarse-grained soils such as very soft to very stiff silty, sandy and locally gravelly clay, sandy to very sandy silt and clayey to very clayey, silty to very silty sand and gravel. This layer contains locally both marine fossils and clay with high plasticity, interbedded locally by the poorly graded coarse-grained deposits of layer "Sand/Gravel Lenses" The sedimentation environment is a combination of alluvial and marine soils and brackish waters. It extends to depths below seabed larger than 60 m, which corroborates with the broader area geology. SPT N-values generally vary between 4 and 36 within the fine-grained deposits but reach up to >50 within the local occurrence of coarse-grained soil.

4.4 CONSTRUCTION SEQUENCE

The first step is dredging. Construction on soft soil, also named mud, is associated with many problems due to the low shear strength and high compressibility of these soils. In this case, dredging means removing the mud from the sea bed, more precisely layer 2. An example of a dredging operation is visible in Figure 4.1.

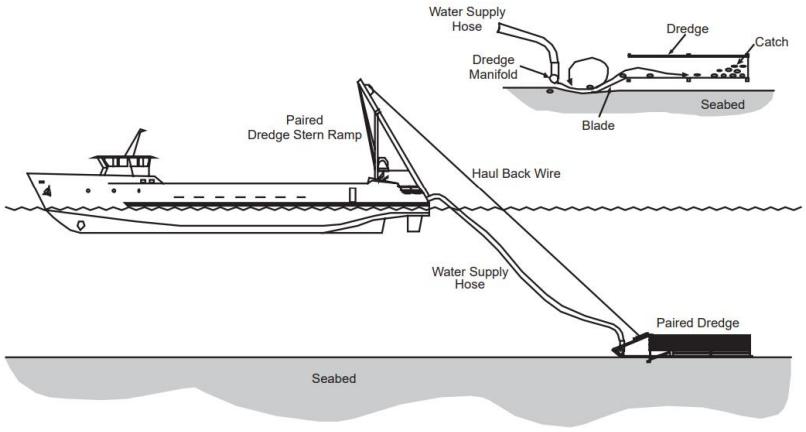

Figure 4.1 Example of hydraulic clam dredging gear and method (Gilkinson, 2003)

The second step is to place an approximately one-meter thick layer of sand at the bottom of the sea. Then PVDs are driven and placed as deep as required in the design over the sand layer, which goes around 20 to 25 meters. By visualizing Figure 4.2, is observable an example of PVDs installation in offshore constructions. Beyond PVDs installation, an additional 1 m gravel layer is required over the sand layer. With this extra layer, the sand cannot mix with the soil placed above and will not adulterate it. Now, the conditions to start the preloading are fulfilled. The soils are placed in barges, that are continuously dumping pure soil into the sea. An example of a barge transporting soil is perceptible in Figure 4.3. One of the techniques adopted to deposit soil on the land reclamation site can be through doors located at the bottom of the barge: when the doors open the soil can fall to the bottom of the sea. Accuracy is achieved by manoeuvring the barge precisely above the designated location.

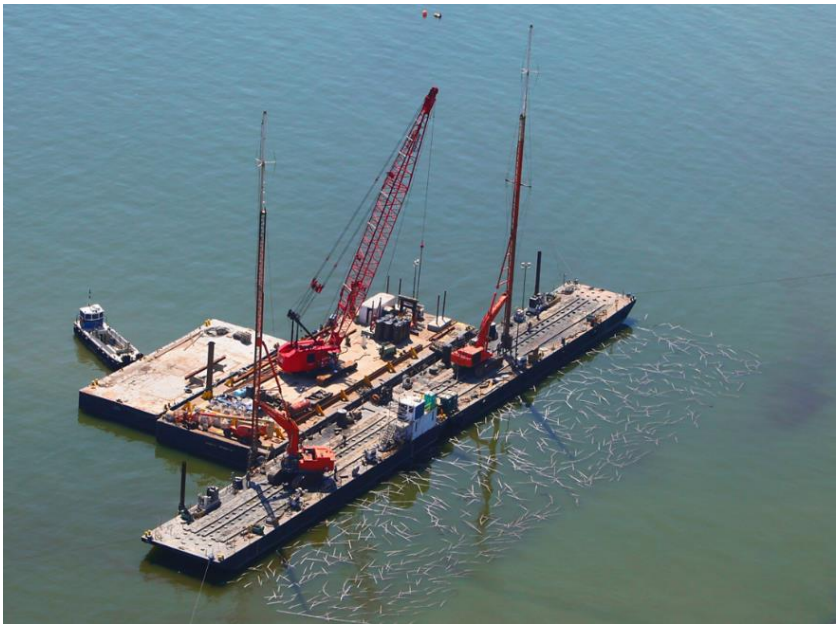

Figure 4.2 Offshore PVDs installation (Jimmy, 2012)

Figure 4.3 Barge transporting soil for land reclamation (Wikimedia Commons, 2011)

The preloading is divided into phases. It is done gradually since the *in situ* soil is too soft initially. If the preloading is performed in one single time, the probability of failure is higher due to the *in situ* soil poor characteristics. By performing the preloading gradually, the soil can consolidate and improve their properties after a specific time. After the preloading phase is complete, the soil is vibro compacted, and then the next layer of preloading is added. When this consolidation is achieved, the second layer of preloading can be placed and so on, until the design elevation is reached. Figure 4.4 shows an example of different soil layers being pumped into the land reclamation site in a port expansion in Sri Lanka.

Figure 4.4 Pumping soil into reclamation area (China – Sri Lanka jointly build “Shining Pearl of Indian Ocean”, 2019)

The main purpose of this work is to find the time that the preloading needs to be removed. On chapter 6, the settlement graphs will be presented where it is possible to identify the time for removing the preloading. The typical soil height in these projects is around 2 to 3 meters above the water. After the *in situ* soil is consolidated, the preloading layers placed above the intended height can be removed since just the layers placed at the requested height are going to be placed permanently. A visible perception of the preloading process is represented in Figure 4.5.

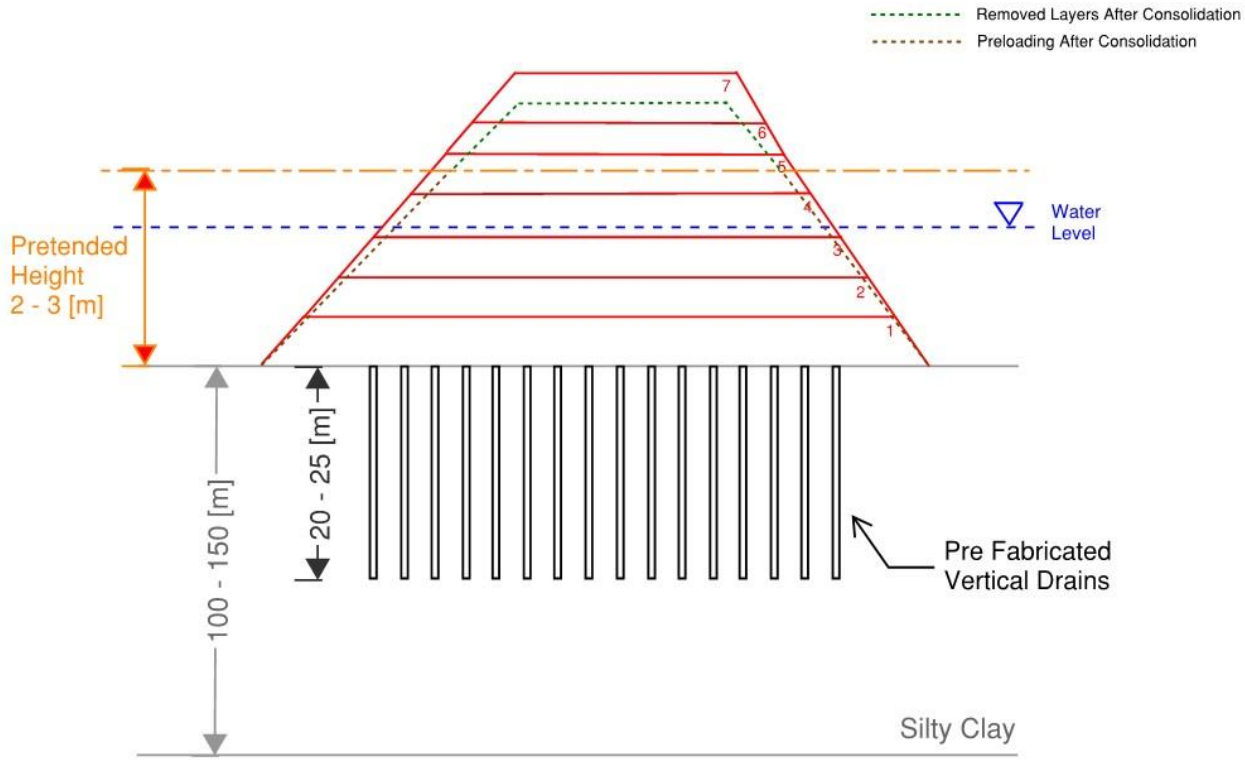

Figure 4.5 Preloading description

At this point, some soil adjacent to the end of the embankment by the sea is removed so a prefabricated caisson wall can be implemented. Caisson wall is a prefabricated reinforced concrete structure and stays there permanently guaranteeing soil stability. The empty space between the wall and the structure is filled with soil, and vibro compacted, so the design height and conditions for the embankment are now respected on all the site. All the pavement is now complete, so the construction can start operating.

4.5 GEOTECHNICAL PARAMETERS EVALUATION

The geotechnical design parameters were derived by analysing a total of 56 rotary continuous sampling boreholes with SPT as well as data from 11 electrical CPTs tests and six CPTs with mechanical cone gathered from the five site investigations mentioned in section 4.2.

The appropriate parameters evaluation to this case study were also performed through laboratory tests, more precisely the Oedometric test. This test allows measuring soil consolidation properties, necessarily the Oedometric modulus and the coefficient of consolidation in the same test, through the analyses of the soil collected in the boreholes. The data presented in this chapter was provided already treated and processed, so the laboratory results interpretation was not contemplated in this work.

Therefore, the values indicated in Appendix 7.2A.1 resulted from the merged interpretation of laboratory and in situ tests. The descriptive statistics of each parameter is presented below in the present section.

4.5.1 IN SITU TESTS

4.5.1.1 Standard Penetration Test

The variation of N-SPT in depth is presented in Figure 4.6, based on the surveys mentioned in section 4.2. The interpretation of this evaluation clarifies the soil stratification according to the penetration resistance recognised.

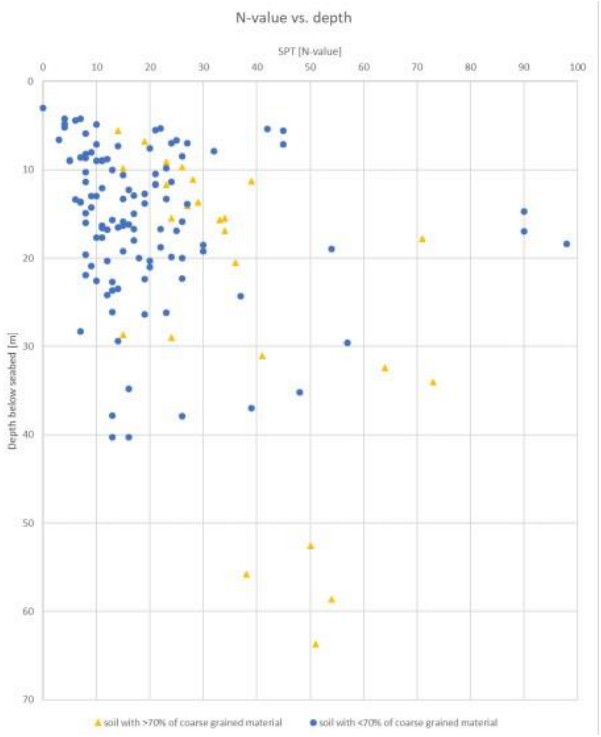

Figure 4.6 N-SPT vs depth

In all surveys, the N-SPT value shows a tendency to increase with depth, which is consistent with the profile of a normally consolidated clay. However, there is a wide scatter of the values, which can be attributed to the randomly appearing coarse-grained soil lenses (layer 3). It is considered that N-SPT values more significant than 20 generally indicate coarse-grained soils. The scatter may also be due to the different surveys which may not be entirely comparable due to different equipment and operators. This is even more pronounced in the SPTs performed in the 1970 and 1972 surveys, which may not fully conform to the test standards used in the more recent surveys.

4.5.1.2 Cone Penetration Tests

Six CPTs with mechanical cone have been performed during the supplementary survey and their combined results of qc and friction ratio, *Rf*, are presented in Figure 4.7, in terms of qc (left graph) and Rf (right graph) vs depth below the seabed. The older eleven CPT results cannot be presented, as their numerical results are not available, and they were also performed with an electrical cone which is not directly comparable with the mechanical cone of the new CPTs.

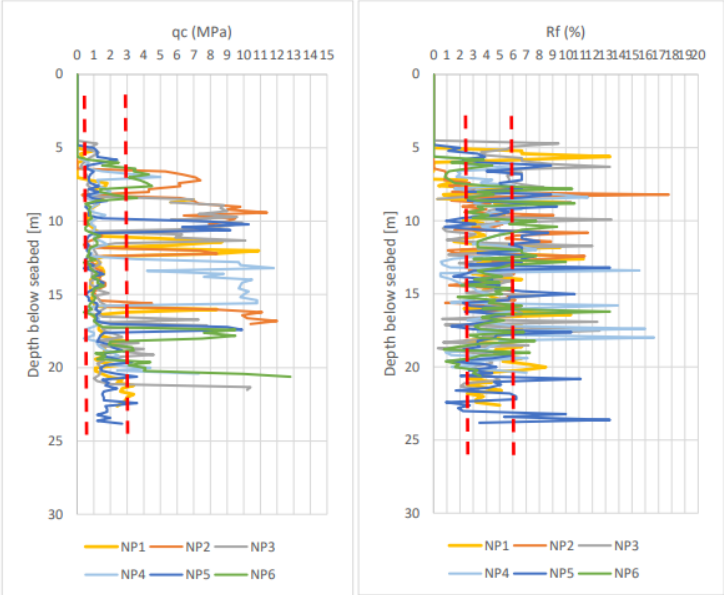

Figure 4.7 qc (left graph) and Rf (right graph) vs depth below the seabed. Dashed red lines indicate a range of values corresponding to layer (4) (COWI A/S, 2020)

The qc and Rf values in Figure 4.7 indicate the existence of two groups of values:

- A group of values with a friction ratio above 2.5 and up to 6.0, corresponds to predominantly fine-grained material (layer 4). This has a corresponding range of qc between 0.70 MPa and 3.00 MPa, which shows an increase with depth;
- A group of values with a friction ratio below 2.5 and higher than 1.0, matches to predominantly coarse-grained material (layer 3). This has a corresponding range of qc from 6.50 MPa and up to 10 MPa or CPT refusal. This soil appears in 2-4 m thick layers (lenses) randomly interbedded with the principal (layer 4), without any clear trend with depth.

4.5.1.3 Soil Physical Properties

The physical properties' variation with depth is presented in Figure 4.8. These results were obtained from laboratory tests (granulometry tests), and SPTs field tests mentioned in section 4.5.1.1.

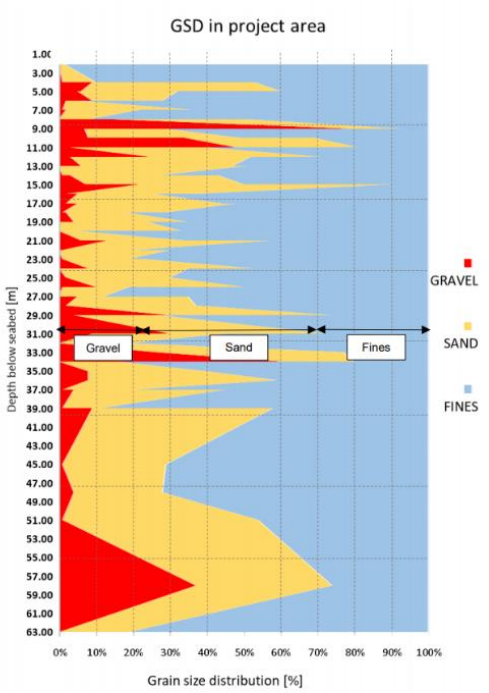

Figure 4.8 Grain size distribution vs depth below seabed (COWI A/S, 2020)

From Figure 4.8, it is concluded that the project geology is mainly composed by fine-grained soils (with fines content of 60-70%) interbedded with lenses of coarse-grained soils having fines content of less than 40%.

4.5.2 DEFORMATION PARAMETERS

4.5.2.1 Oedometric Modulus

The variation of E_{oed} vs depth is presented in Figure 4.9. The points indicated in the graph are derived from lab tests and from correlations with qc, as presented in section 3.2.3.1, applying a correlation factor (α) of 4, Table 3.1.

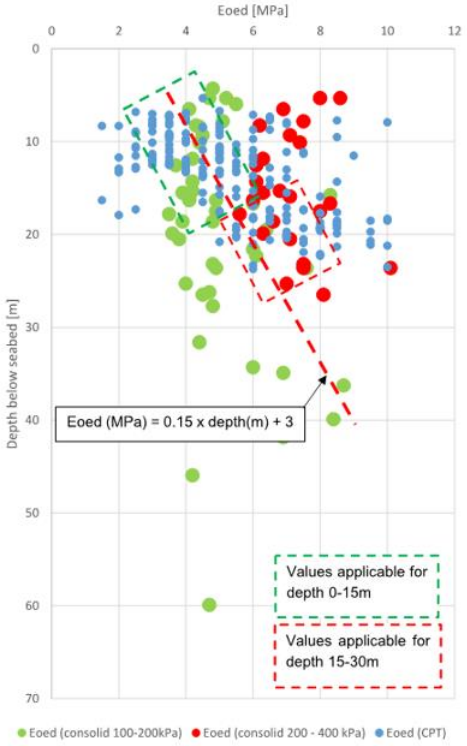

Figure 4.9 Constrained modulus (E_{oed}) vs depth below the seabed. Values corresponding to coarse-grained layers are not shown (COWI A/S, 2020)

For this work, the Oedometric modulus values considered in the analysis were the green points in Figure 4.9, corresponding to depths up to 15 m adjusted with the blue points, obtained in the CPTs investigations. The values were considered fitted from an engineering point a view. In Table 4.1 are charted the descriptive statistics of the considered values included in Appendix A (Table A. 1) to the E_{oed} (100-200 kPa).

Table 4.1 Oedometric measures considered in the study

Constrained E modulus values (MPa)	
Mean	5,08
St. Deviation	1,28
Min	3,5
Max	8,7

4.5.2.2 Compression Index (C_c, C_{ec})

The variation of compression index, C_c, and C_{ec} vs depth is presented in Figure 4.10, based on oedometer tests. Since C_c and C_{ec} are correlated over the void ration equation 3.8, it is possible to obtain both parameters from consolidation tests results. C_{ec} is the parameter considered in the study, once it is a standardisation of C_c, allowing to ease the calculation process since it already includes the soil void ratio.

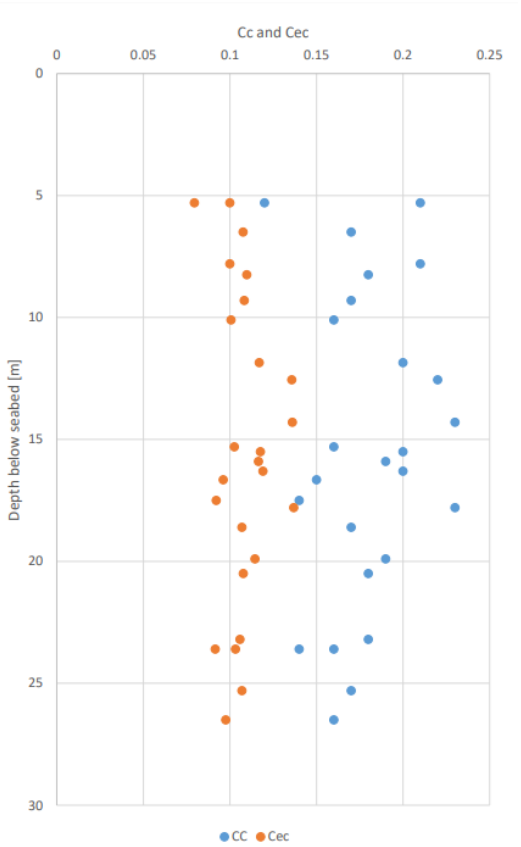

Figure 4.10 Compression Index vs depth below seabed (COWI A/S, 2020)

From the above diagram, a representative value of C_{ec} = 0.11 is selected.

In Table 4.2 is tabulated the descriptive statistics of the C_{ec} values considered in the case study.

Table 4.2 Compression Index measures considered in the study

Compression index Cec (-)	
Mean	0.1102
St. Deviation	0.0267
Min	0.0700
Max	0.2400

4.5.2.3 Coefficient of Consolidation

The coefficient of consolidation vs depth is presented in Figure 4.11. These values are determined directly from consolidation tests.

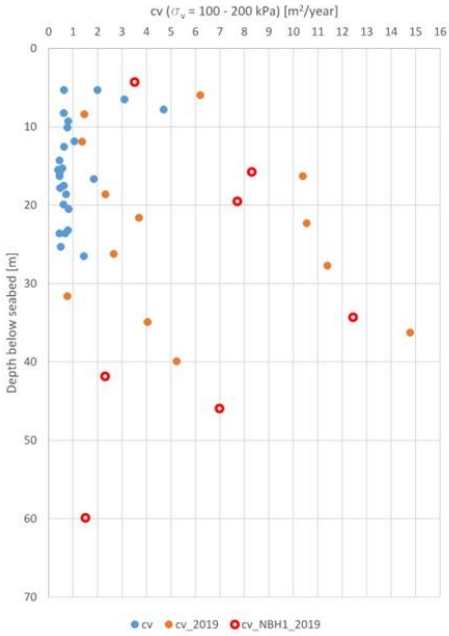

Figure 4.11 Coefficient of consolidation vs depth bellow seabed (COWI A/S, 2020)

The representative value from this graph is approximately 0.7 m²/year. However, for the reasons explained in section 3.2.3.2, the c_v values calculated from oedometer tests underestimate the c_v value.

In Figure 4.11 are presented c_v values derived from three different sources. The blue and orange points outcome from the surveys mention in section 4.2, and the red points outcomes from the samples taken from the boreholes and processed through the oedometric test. To this case study, the values adopted were between the range [0.3, 15] since the blue points results are not compatible with the engineering point of view values requiring in this way a restriction of values.

In Table 4.3 are organised the descriptive statistics of the considered values in the study.

Table 4.3 Coefficient of consolidation measures considered in the study

Coefficient of consolidation (m²/year)	
Mean	3.47
St. Deviation	3.98
Min	0.39
Max	14.78

4.6 MONITORING PROGRAM

Construction monitoring is an accurate and helpful technique of controlling and examining the quality of a construction project providing real-time measurements.

The design cannot account for the long-term settlements of the deeper non-improved layer, the layer starting bellow the PVDs. However, it is vital to have an estimate during the monitoring of the long-

term settlement of this deeper unimproved layer since this is critical for the future operation and maintenance cost of the port.

Typically, in this type of projects, it is not possible to have measurements immediately after the start of preloading construction, since it is not conceivable to install the instruments before the preloading reaches the water surface. Thus, before the preloading reaches the water surface, it is not possible to know the settlements or pore pressures which represent critical information for predicting the future development of the settlements.

The instruments used in this case, are placed in three sections in four different locations. The number of monitoring instruments is the following, whose detailed description was made in section 3.3:

- 12 Extensometers;
- 12 Vibrating Wire piezometer;
- 12 Standpipe piezometer;
- 8 Shape Array (SAA).

All three sections contain the instruments mentioned above. Per section, there is one of each instrument. Inclinometers are just placed in the two sections close to the slope since its primary function is to evaluate slope stability.

5
Numerical modelling towards design parameters

5.1 Statistical Description of Soil Properties

5.1.1 Introduction

Geotechnical design based in probabilistic and risk analysis consists of the assignment to each soil parameter a statistical distribution instead of assuming characteristic values, corresponding to averages or increased costs by safety factors. Soil parameters have a significant influence on the geotechnical analysis, especially when the estimation of settlement and consolidation time is in consideration, and thus when the best approaches to speed up consolidation need to be defined. Statistical and probabilistic methods can quantify the uncertainties and interpret them in a balanced and reliable way. When using mathematical description in soil parameters, those parameters are interpreted as random variables for statistical language. At this point, the goal is to assign and justify each parameter a probabilistic model.

To describe these random variables, primarily the measured data obtained must be described so the Generating Random Variables can be realized. The data set comprises both laboratory tests and in situ surveys. Secondly, through the application of Simulation Methods, a specific distribution is attributed regarding the compatibility between the descriptive statistics of the data set and the attributed distribution. The description of some terms such as mean, standard deviation, higher value and the lowest value is calculated considering the boundaries defined by the engineering point of view, where standard deviation represents the uncertainty associated with each parameter. Once the distribution is characterised and confirmed, the soil parameter can be defined through a probabilistic model of a random variable, process schematically represented in Figure 5.1.

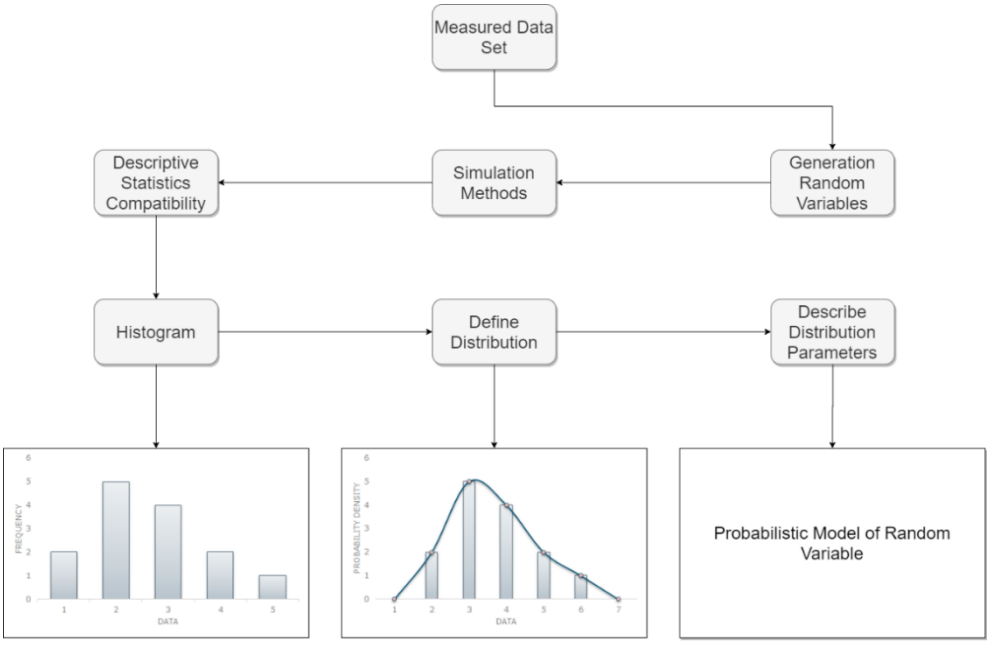

Figure 5.1 Pictorial representation for the probabilistic modelling of a random variable

The main goal is to find the correct input to the calculation of the settlement, either for design modelling and for the analytical approach. The inputs could be the coefficient of vertical permeability or the coefficient of consolidation. In order to progress with the parameters modelling, some procedures need to be concerned to conclude a correct design with Monte Carlo simulation (MC). Here are presented preliminary considerations to accomplish the correct input parameters and the correct operation of MC. At this point, two different approaches are regarded. In order to implement the Monte Carlo simulation, 2000 samples were randomly generated, assuming that the parameters follow a normal distribution function. The next statistical description of soil proprieties is based on the previous results showed in Chapter 4, namely section 4.5.2.

The initial step, presented in the current section, for the numerical modelling procedure, was based on the normal distribution due to its simplicity. These preliminary studies allow to diagnose and understand the statistical characteristics of the parameters, the modelling process and restrictions. At this point, there are three variables considered, c_v, E_{oed} and k_v. These variables concern the input parameters to the settlement analysis required design purposes. Their management and performance in this step were crucial to recognize the improvements that need to be concerned. After their numerical definition using the normal distribution, 2000 random samples were generated through Monte Carlo simulation method.

Based on the drawn conclusions from the initial step, it was clear that the normal distribution has some limitations to define the parameters. As a consequence, the truncated normal distribution and the truncated lognormal distribution were adopted. After a more realistic numerical modelling was processed, the 2000 random samples were generated as described in section 2.3.3, which also confers a more realistic approach. Afterwards, MC was also implemented featuring the correlation of the parameters, and the final 2000 random samples were generated. Up until now, the match among measured data and the probabilistic model assigned to each parameter was based on descriptive statistics compatibility and distribution shape.

Although the second approach is supplementary realistic, due to the numerous existing continuous distributions and the restrictions of the physical parameters, the consideration of those distributions is

crucial. In this step, the software used was EasyFit, which is a powerful software that also takes into consideration continuous empirical distributions that match better for the variables in study. At this stage, different probability distributions were tested for each parameter, and the one that matches the requirements needed was implemented. Figures for each parameter are presented, taking into account the distributions adopted in the second approach and the ones that best fit the EasyFit software in order to compare them visually. MC was also the simulation procedure adopted where the outcome of the random samples comes from the software EasyFit.

In order to complement the modelling process, it was implemented a different simulation method. Hereupon in order to compare the outcome results with the previous ones obtained, distinguish their differences, and select the one that accomplishes best results to the variables modelling.

The elapse of these steps were vital to the correct definition of the parameters, to understand different improvements that need to be concerned, and to be able to conceive the modelling processes and techniques that need to be adopted thoroughly.

All the steps previously mentioned are carefully described in the succeeding sections.

5.1.2 FIRST APPROACH

The target at this approach is to obtain a coefficient of consolidation and compare it with the measured values. This approach comprises three stages:

1. Create a probability density function (PDF) for the coefficient of vertical permeability
    - Assuming E_{oed} = 4000 kPa;
    - c_v measured values, Table 4.3;
    - Considering equation 3.15.
2. Create PDF for E_{oed}
    - E_{oed} measured values, Table 4.1.
3. Create PDF for c_v
    - Considering equation 3.15.

5.1.2.1 PDF for the coefficient of vertical permeability

Taking into account the c_v values presented in Table 4.3 assuming a constant oedometer modulus of 4000 kPa and using equation 3.15, a sampling simulation using the Monte Carlo method was created to obtain the PDF of the permeability coefficient (k_v) presented in Figure 5.2. Since this is a statistical approach, a geotechnical engineering boundary, *Georef*, needs to be considered. Once the values of k_v cannot be physically lower than or equal to zero, the k_v geotechnical reference minimum acceptable value is 0.0001 m/year. Due to that, all the values below the reference minimum need to be removed, truncated from below. This procedure results in a truncated normal distribution. For that reason, the actual PDF for the coefficient of vertical permeability is presented in Figure 5.3.

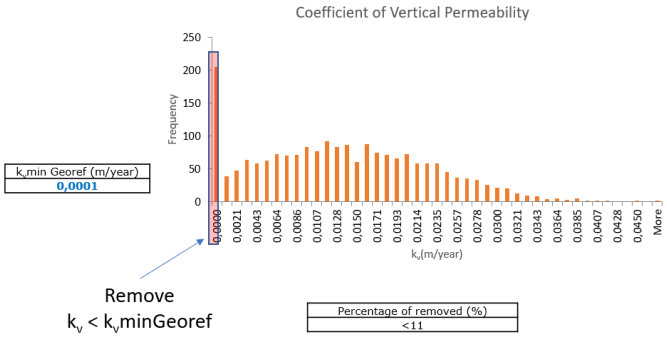

Figure 5.2 Coefficient of vertical permeability PDF before truncation (first approach)

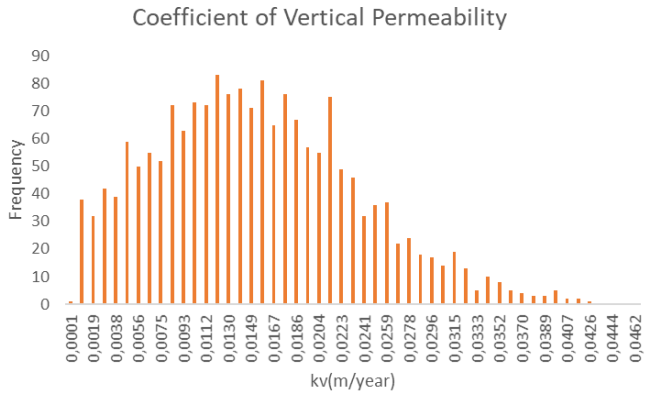

Figure 5.3 Coefficient of vertical permeability PDF (first approach)

The normal distribution that matches Figure 5.3 is right-skewed since the majority of values are concentrated close to the minimum value, and the sample has more variation around its mean.

5.1.2.2 Oedometric Modulus PDF

Through the tabulated values in Table 4.1, implementing MC, the Oedometric modulus PDF results in Figure 5.4.

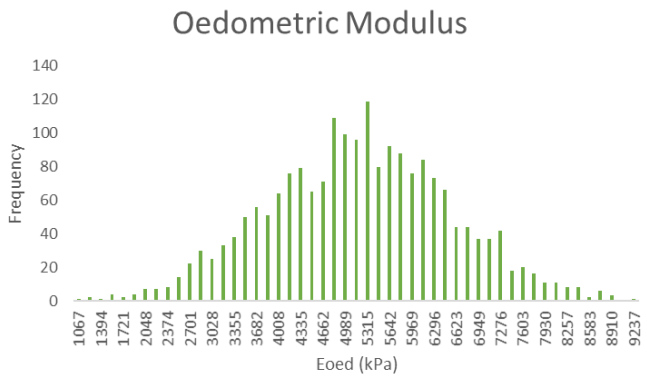

Figure 5.4 Oedometric modulus PDF (first approach)

This parameter ties a typical normal distribution, symmetric at the centre, establishing that most of the observations gather around the central peak and the probabilities for values further away from the mean taper off equally in both directions.

5.1.2.3 Coefficient of Consolidation PDF

Considering the previously mentioned designs obtained with Monte Carlo simulation (for the oedometer modulus and the permeability coefficient) and applying them in equation 3.15, using a water unit weight of 9.81 kPa/m^3, through the MC, the obtained coefficient of consolidation PDF is presented in Figure 5.5.

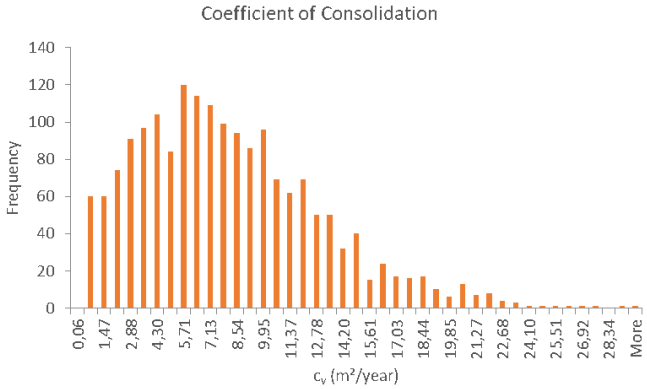

Figure 5.5 Coefficient of consolidation PDF (first approach)

5.1.3 Second Approach

The goal of this approach is to obtain a coefficient of vertical permeability and compare it with the one obtained in the first approach. This approach is separated into three phases:

1. Create PDF for c_v
    - c_v measured values, Table 4.3.
2. Create PDF for E_{oed}

- E_{oed} measured values, Table 4.1.

3. Create PDF for k_v
   - Through the rearrangement of 3.15 equation.

5.1.3.1 Coefficient of Consolidation PDF

The same approach made for the calculation of the coefficient of vertical permeability PDF, on section 5.1.2.1, needs to be considered in this step. After applying the MC, the c_v PDF is obtained and presented in Figure 5.6. The statistical approach needs to be adapted to a geotechnical reference value, once the values of c_v cannot be lower than or equal to zero. The contemplated reference value is 0.5 m²/year, resulting in discarding all the values below the referenced value, truncated from below. So, the c_v PDF with geotechnical boundaries outcome in the truncated normal distribution represented in Figure 5.7.

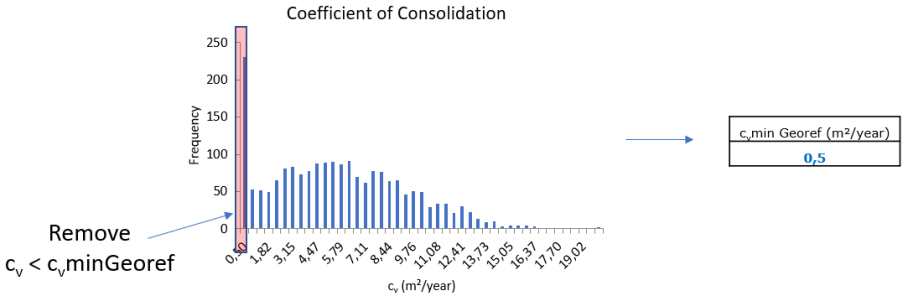

Figure 5.6 Coefficient of consolidation PDF before truncation (second approach)

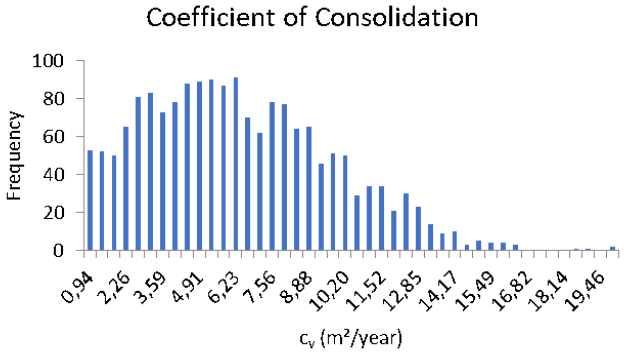

Figure 5.7 Coefficient of consolidation PDF (second approach)

5.1.3.2 Oedometric Modulus PDF

The Oedometric Modulus follows the same process detailed in section 5.1.2.2. In Figure 5.8, the E_{oed} PDF is represented. The values are different from Figure 5.4 to Figure 5.8. This occurs since MC always samples different random values. When running MC, obtaining different values, but roughly the same PDF shape, incomes that the method is processing correctly, so a correct approach is being considered. So, the conclusion took in section 5.1.2.2 are equivalent in this phase.

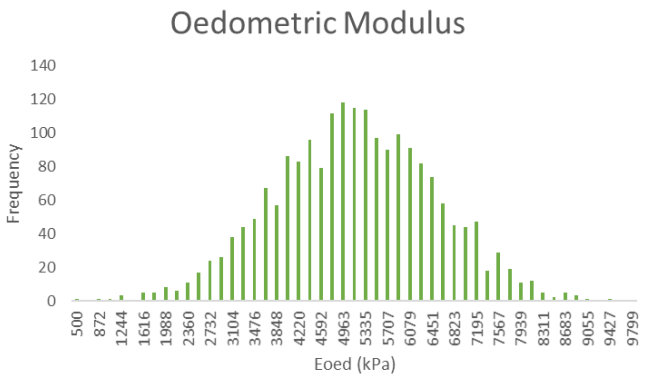

Figure 5.8 Oedometric modulus PDF (second approach)

5.1.3.3 Coefficient of Vertical Permeability PDF

The calculation of k_v PDF follows the same scheme of section 5.1.2.3. Regarding the two earlier mentioned PDFs and considering the water unit weight of 9.81 kPa/m^3 the calculation process to k_v through MC outcomes in Figure 5.9, through the rearrangement of 3.15 equation.

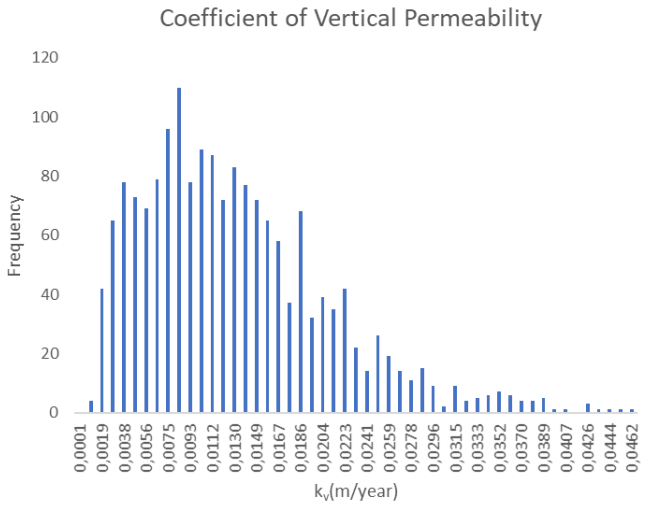

Figure 5.9 Coefficient of vertical permeability PDF (second approach)

5.1.4 RESULTS ANALYSIS

Figure 5.10 contemplates the PDFs comparison between the laboratory measures obtained in Figure 4.11 and Figure 5.5 over the MC.

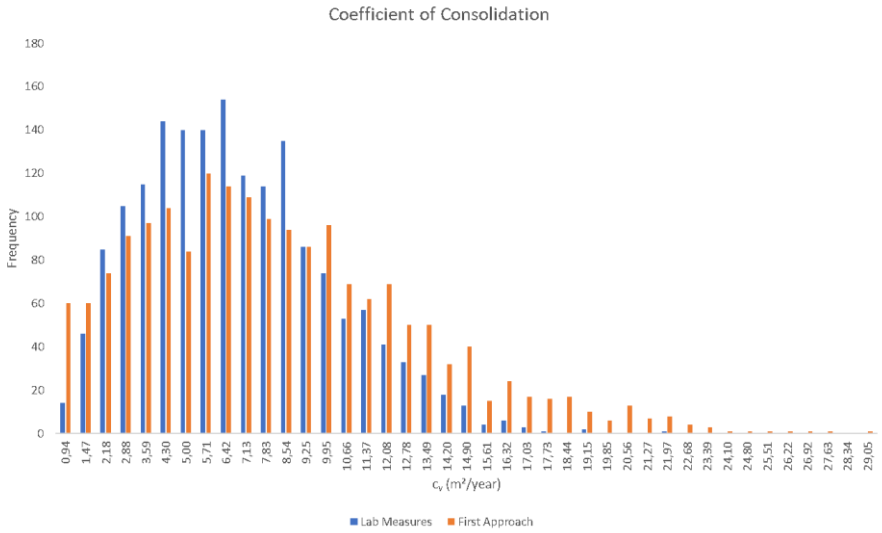

Figure 5.10 Coefficient of consolidation PDF comparison

It is visible some differences in the PDF shapes. The first approach PDF shows a steep right tail. This happens since the statistical approach does not contemplate a maximum value resulting in a more spread shape. To have a true definition of reality, the c_v PDF considered must be the one with the values obtained in the laboratory measures.

Figure 5.11 represents the comparison among the k_v, calculated in the two different approaches.

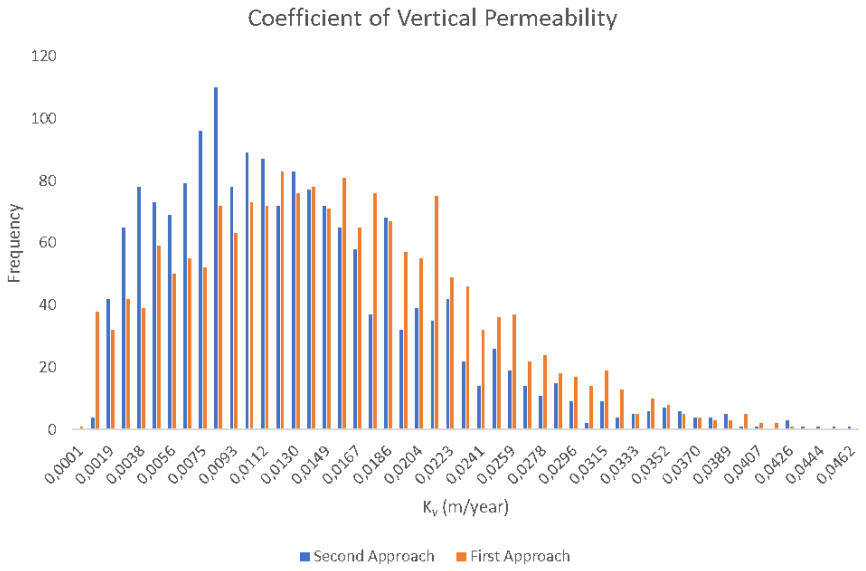

Figure 5.11 Coefficient of vertical permeability PDF comparison

Due to the parameters in concern, the k_v PDF should be a right-skewed distribution. It is visible when seeing the input variables PDFs. The first approach PDF in Figure 5.11 outcomes from assumption values. Through this assumption, outcomes a spread distribution since the engineering boundaries were

not considered in the input parameters. This consents to conclude that the best approach is the second, once the input parameters are the real ones obtained in measured surveys contemplating the real conditions preventing from statistical errors.

In sum, the parameters that should be numerically modelled must be the coefficient of consolidation and the oedometric modulus using the measured values obtained for each parameter. For the calculation of settlements in Plaxis®, the input parameter must be k_v, calculated in the same way as in the second approach. For the analytical approach, the input parameters must be the modelled, namely the coefficient of consolidation and the oedometric modulus.

These analyses allow us to achieve some conclusions. The statistical distributions to define the parameters must be truncated distributions to respect the engineering boundaries. The distributions that typically fit the best are the normal distribution for parameters that are symmetrical and lognormal distributions for the parameters that have a majority of values concentrated in one particular length.

5.2 PARAMETERS NUMERICAL MODELLING

The numerical modelling process needs to have essential considerations to be able to match statistical approaches with physical realism. In Figure 5.1 is represented the procedures to define a probabilistic model of a random variable assumed. With the considerations taken from the previous section, Figure 5.12 schematically represents the adjustments assumed in the modelling process and their proposal at this phase.

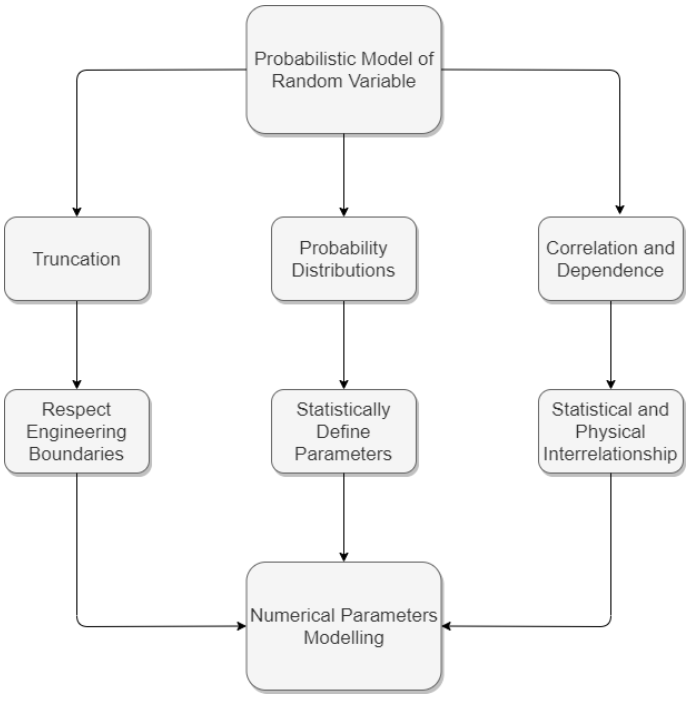

Figure 5.12 Numerical modelling flowchart

In this stage, parameters such as the coefficient of consolidation must be implemented in the modelling software considering the lognormal distribution with lower and upper bounds. Statistically, this is called truncated lognormal distribution. By performing this, it is possible to compare the normal distributions

modelled in the previous section, with this approach, to find the best to use. The advantage of the lognormal truncated is that there is no need to remove any values, so, all values sampled are positive, and there is no stacking of values at the zero value, as presented before. Also, the c_v is usually better modelled by a lognormal distribution, once the consolidation tests typically show many values around the peak and then a long tail to the right. For the above reasons, if the truncated lognormal gives equally good estimates of the mean and standard deviation when compared with the normal distribution, then the modelling process will keep the lognormal distribution.

Parameters that present a more symmetric distribution, with tails on both sides of the peak, the normal distribution is a satisfactory approach. Parameters like oedometric modulus must be defined through normal distribution because the measured E_{oed} values present a more symmetric distribution.

In reality, the interdependence between physical parameters is significant. This phenomenon happens because those properties are physically correlated, although their measures are done independently during the laboratory test. The correlation coefficient is determined following the procedure of section 2.1.9. Our measured values, c_v and E_{oed} or $C_{ec,}$ is an example of correlated parameters. Whenever two random variables are correlated, the randomly produced samples are also correlated, mainly if they are intended to be used together in further calculations. In practice, both sets of samples must be first created, and then the expression 2.54 may be applied to one of these sets. From a mathematical point of view, it is correct to use the correlation coefficient to any of the two variables. In this specific case, the affected variable must be E_{oed} or C_{ec}, and c_v sampled values keep unchanged. Parameters like E_{oed} and C_{ec} indicate how much the soil in a study is going to settle, while c_v describes to us how fast the settlement is going to occur. Once E_{oed} and C_{ec} are diversified throughout all the soil dimensions, these parameters must be affected by the correlation with the c_v data to get values nearer to physical reality. Since the correlation between variables is independent of the distribution that defines each parameter, their correlation must also be considered in the modelling process.

5.2.1 IMPLEMENTATION DETAILS

Microsoft Excel has a well-known and durable privations. The use of explicit and straightforward formulas eases the Excel documentation and computations designs. Modelling random variables considering normal distributions decreases the quantity of testing that needs to be completed in order to found reliable results and maximize the chance to understand how the software works appropriately. Modelling considering a Lognormal distribution is simple considering the cumulative distribution function (Φ) and inverse cumulative distribution function (Φ^{-1}) of the standard normal distribution. Considering the expression presented in section 2.1.5.4, the modelling results in the following functions:

$$F(x) = \Phi\left(\frac{Log(x) * \mu}{\sigma}\right) \tag{5.1}$$

$$F^{-1}(q) = exp(\mu + \sigma\Phi^{-1}(q)) \tag{5.2}$$

In Excel along with other numerical software programming languages, a uniform random variable *X* in an interval [a, b] is generated by obtaining a random variate *Y* from the interval [0, 1] resulting in:

$$X = a + (b - a)Y \tag{5.3}$$

The modelling application form for normal distributions meets the same procedures mentioned above, considering the standard normal distribution, adapting their properties as explained in section 2.1.5.3.

5.2.2 MODELLING VARIABLES

The focus of this step is to manage the core variables essential to the estimation of settlements in time. After the numerical modelling procedure, the sampled values obtained for each variable are going to be the inputs values to the design.

5.2.2.1 Coefficient of Consolidation

As stated before, the coefficient of consolidation is statistically modelled as a truncated lognormal distribution. According to the Implementation Details presented and the data collection of the laboratory results, the c_v modelling results in the following PDF, presented in Figure 5.13, where the sampling data and the measured data frequencies are represented.

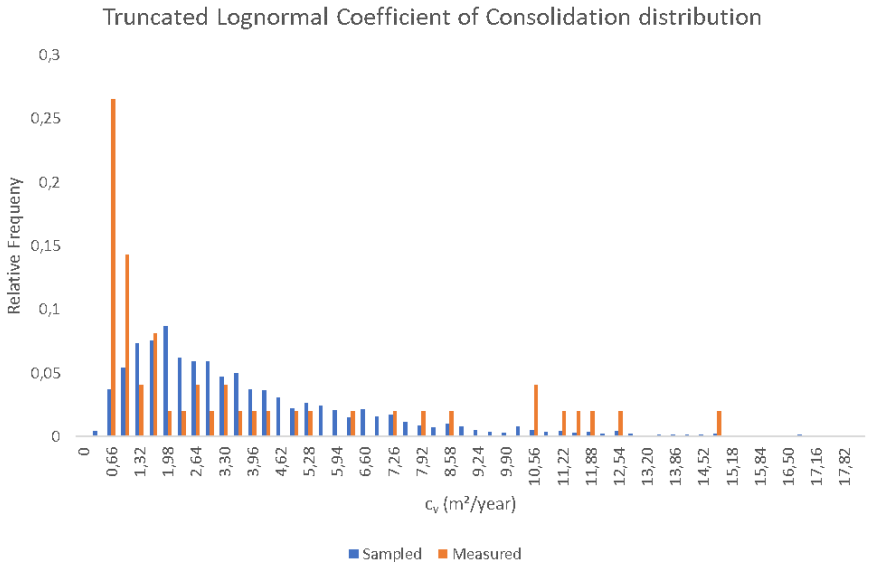

Figure 5.13 Coefficient of consolidation PDF

5.2.2.2 Oedometric Modulus

The E_{oed} modelling procedure is done considering the normal distribution since the data obtained in the laboratory results expresses symmetric distribution. The initial step for the E_{oed} modelling process is the

definition of the parameter only considering the laboratory results, resulting in the PDF represented in Figure 5.14.

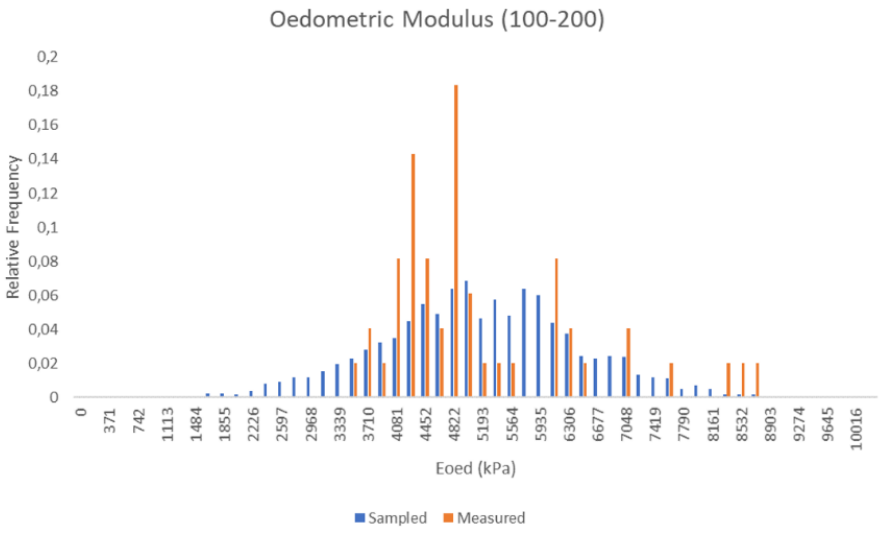

Figure 5.14 Oedometric modulus PDF

This parameter needs to have in concern the correlation with the coefficient of consolidation. This factor is based on both measured data outcoming the value 0.426 (Table 5.1). Applying the factor to the sampled data obtained, accomplish the PDF presented in Figure 5.15.

Table 5.1 Coefficient of consolidation and oedometric modulus (100-200 kPa) correlation

	c_v (m²/year)	*E_{oed} (MPa)_100-200 kPa*
c_v (m²/year)	1	
E_{oed} (MPa)_100-200kPa	0.426	1

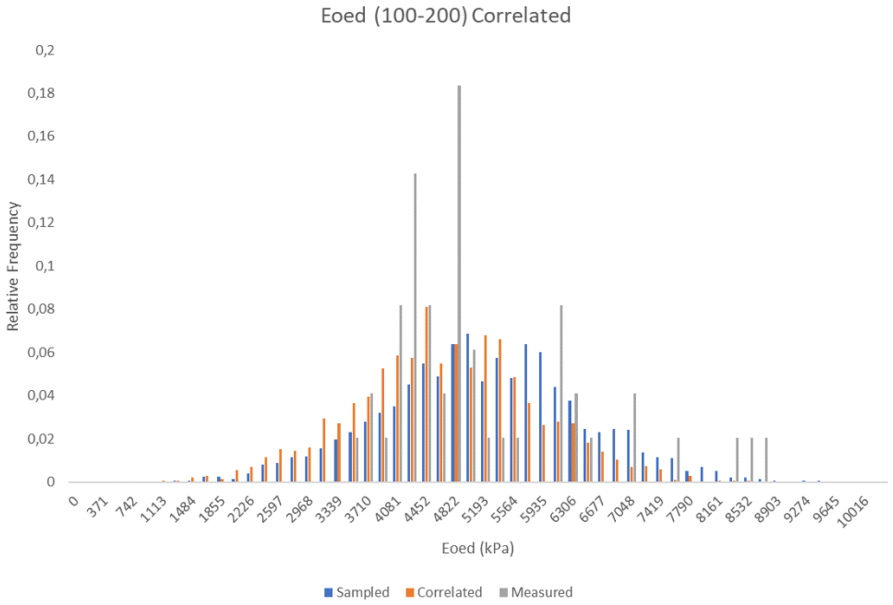

Figure 5.15 Oedometric modulus correlated PDF

It is visible that the data moves slightly left. This move is a result of a mathematical approach since it does not have any influence on the parameters physical meaning.

5.2.2.3 Compression Index

Another pertinent parameter for the calculation of the settlement is the compression index, C_{ec}. The approach adopted in this modelling stage is the normal distribution since the parameter shows a symmetric distribution. It is modelled the C_{ec} instead if C_c since C_c modelling process needs to have in apprehension also the soil void ratio, which results in an increase of information for the software to process. It is also notable from the correlation obtained in Table 5.2 that c_v and C_{ec} do not have any correlation between them.

Table 5.2 Coefficient of consolidation and compression index correlation

	c_v (m²/year)	*C_{ec} (-)*
c_v (m²/year)	1	
C_{ec} (-)	-0.0139	1

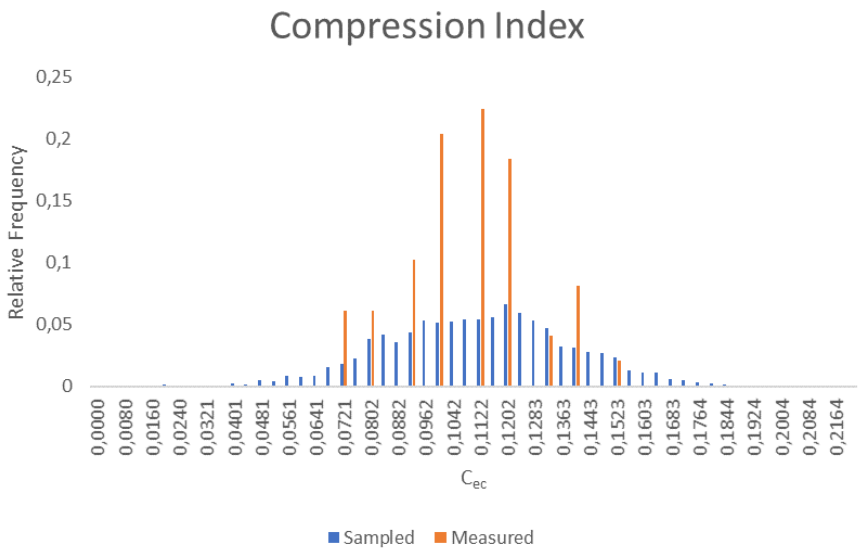

Figure 5.16 Compression index PDF

5.3 ALTERNATIVE MODELLING DEVELOPMENT

There are many probability distributions to define the desired variables. Previously, it was considered just two different distributions, based on the preliminary examinations done. These two distributions can define the variables in the study. However, perchance, it is not taking into account all the characteristics needed to represent the physical reality of the variable. In this section, to counter and to complement the previous results, a data analysis and simulation application named EasyFit is considered.

5.3.1 EASYFIT DATA ANALYSIS AND SIMULATION

EasyFit is a data analysis and simulation application allowing to fit probability distributions to sample data, select the best model, and apply the analysis results to make better decisions. EasyFit can be used as a stand-alone Windows application or with Microsoft Excel and other third-party Excel-based simulation tools. EasyFit combines classical statistical analysis methods and innovative data analysis techniques. This application has some key features (Learn More About EasyFit, 2004-2010):

- Support more than 50 continuous and discrete distributions;
- Calculate descriptive statistics;
- Dominant automated data fitting mode;
- Interactive graphs;
- The goodness of fit tests;
- Random number generation.

The development to find the distribution that fits the best for each variable is made considering the following topics:

- Distribution shape analyses comparing with the measured data;
- Probability-Probability Plot;
- Quantile-Quantile Plot;
- Probability Difference Graph;
- Kolmogorov-Smirnov Test;
- Anderson-Darling Test;
- Descriptive Statistics when compared both outcomes (measured and theoretical).

The fitness tests are ranked as follows, according to their importance:

1. The visual fit of PDF with frequencies histogram and CDF with cumulative histogram. Both the shape and the extremes (left tail, right tail) must fit as pleasant as possible;
2. Small difference (e.g. up to 0.10-0.12) in the Probability Difference Graph of EasyFit;
3. References from the literature review that this distribution is appropriate for the specific physical variable (e.g. soil strength parameters are Normal, grain size distribution is lognormal, rock strength are Weibull or lognormal, and so on);
4. Probability-Probability Plot and Quantile-Quantile Plot tests;
5. Goodness of fit statistics, such as the Kolmogorov-Smirnov Test and Anderson-Darling Test.

However, before checking all the test, all the considered data should be correct from a geotechnical point of view. We should never try to fit a complicated distribution to some weird-shaped data unless we first make sure that the weird shape is real and not due to erroneous data. If the data are correct, then they should fit popular distributions, i.e. Normal, Lognormal, Exponential, Beta, Weibull.

5.3.2 MODELLING VARIABLES

At this point are known the core variables to have in attendance. In this modelling phase, there are going to be in concern two of those variables, c_v and C_{ec} once they are the considered inputs for the analytical settlement calculation. The distribution that fits the best for each parameter is based on the Probability Plots tests and the Goodness Fit Test, Anderson-Darling Test, described in section 2.1.8.2.

5.3.2.1 Coefficient of Consolidation

The truncated lognormal distribution obtained in the previous section was considered, and the distributions that fit the best through the mentioned tests were adopted. Their differences are visible in Figure 5.17, where is also evident the distribution that matches better the measured results.

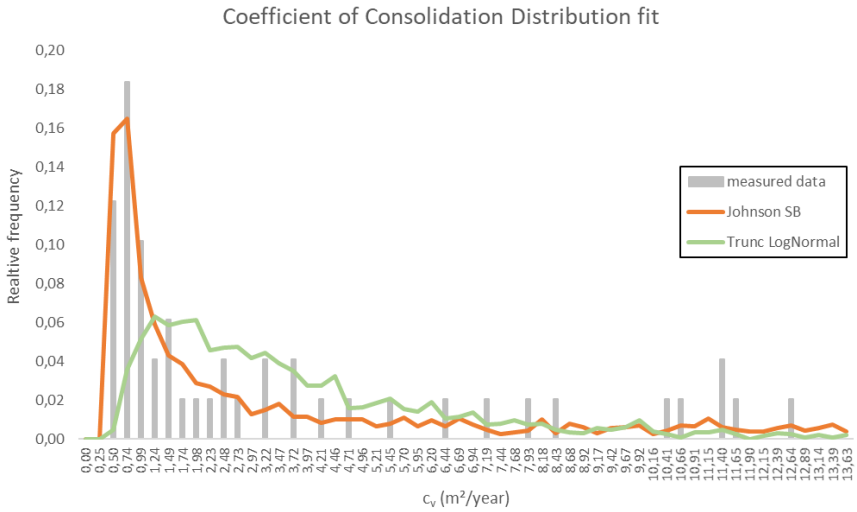

Figure 5.17 Coefficient of consolidation distribution fit

It is visible the limitations associated with the truncated lognormal distribution when compared with the measured data. The distribution adopted for the c_v numerical modelling is Johnson SB once it is the distribution that fulfils the requirements mentioned in the previous section.

5.3.2.2 Compression Index

For the C_{ec} parameter (Figure 5.18), three more continuous distributions were considered beyond the normal distribution. The normal distribution, noticeable in blue, is based on the same procedures adopted in the previous sections with the random numbers obtained in section 2.3.3. The remaining distributions studied are the Dagum distribution (yellow), the Beta distribution (brown) and the Lognormal distribution (green).

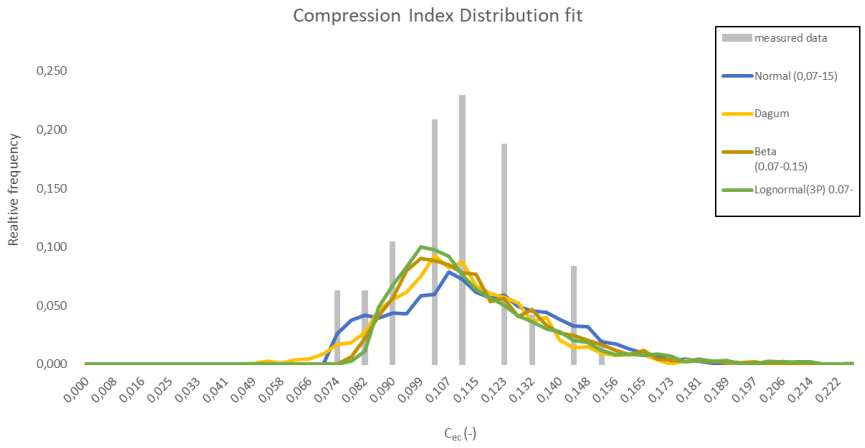

Figure 5.18 Compression index distribution fit

The remaining distributions and sampled values were attained through the software EasyFit. The numerical modelling distribution adopted for this parameter is the Dagum distribution (yellow line), which represent the distribution that fulfils the requirements tests mentioned at the beginning of the section 5.3.1. Proper distributions were also considered in the study as the Beta distribution and Lognormal distribution. However, as shown in Figure 5.16, it does not contemplate the values below 0.078, which in reality, they happen as shown in the measured data.

5.4 MODELLING OPTIMIZATION

Since the parameters modelling is processed, a few optimizations can be implemented. Refining the statistical approach was already performed with the commercial software, EasyFit. Appealing to Monte Carlo simulation, which is a memoryless method, the use of significant samples values to obtain reliable results is needed. In order to reduce the number of samples, the optimization process is through the implementation of Latin Hypercube method.

5.4.1 LATIN HYPERCUBE SAMPLING APPLICATION

The modelling procedures previously adopted were founded through the Monte Carlo simulation by the generation of 2000 samples for each variable. At this point, modelling optimization leads through the implementation of Latin Hypercube sampling. This implementation consents a better response from the software, once the number of inputs introduced is less and extra reliable. The first step for LHS is to decide the number of samples. The simulation method adjusts the number of intervals based on the number of samples introduced implementing them on the variable CDF. Each interval has a probability associated, comprehending multiple random values. After defining the intervals due to his memory ability, LHS stratifies the intervals in order to obtain a sample value for each stratification producing a more consistent and precise sample value. Figure 5.19 diagrammatically represent the procedures adopted to implement the LHS.

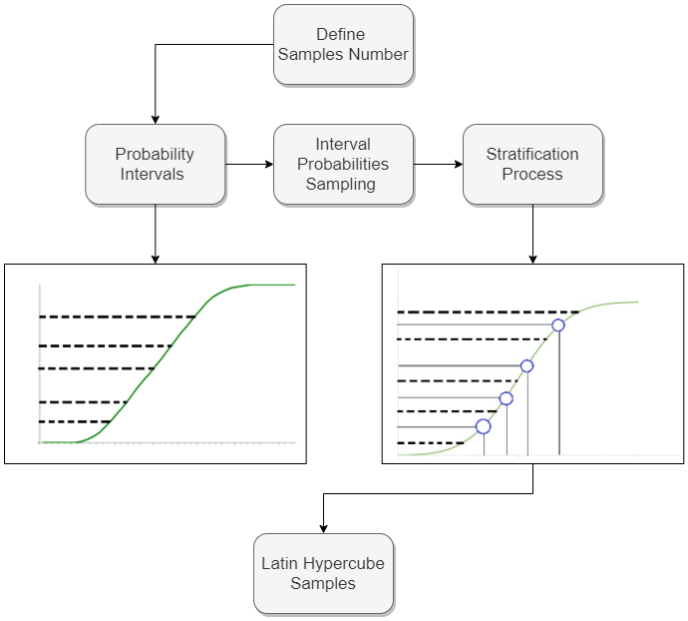

Figure 5.19 LHS recreation procedure

5.4.2 SIMULATION METHODS DISTINCTIONS

Monte Carlo simulation and Latin Hypercube sampling have the same functionality but oppose in the data processing. The main differences between the methods were already stated, highlighting the MC lack of speed and memory, in contrast to LHS. This section presents an observable comparison between the methods taking into account the two variables modelled in the previous section. The adopted process consists of the simulation of 500, 1000 and 2000 sampled values for each method. Afterwards, it is discussed the performance of both methods and their distinctions.

5.4.2.1 Monte Carlo simulation

For variables modelling, considering the simplest simulation methods, MC is the method to implement. MC allows to evaluate uncertainties in a model, and it can be executed with parameters that require significant complexities in an uncomplicated way. The outcomes obtained with MC are also simple to comprehend and to interpret. Figure 5.20 and Figure 5.21 present the MC simulation application for two distinct variables with a different number of samples being tested.

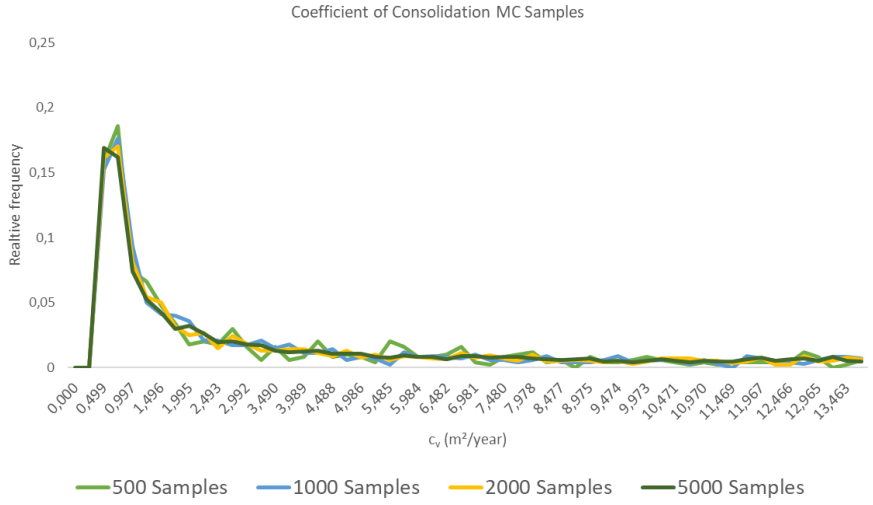

Figure 5.20 Coefficient of consolidation MC samples contrast

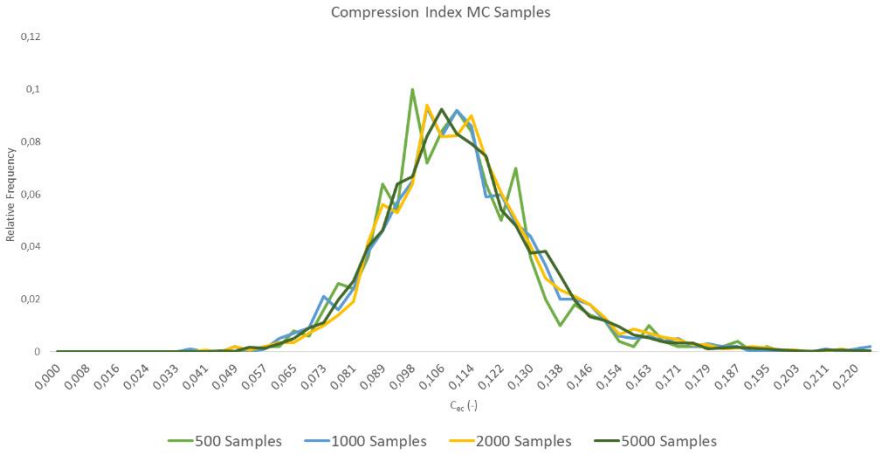

Figure 5.21 Compression index MC samples contrast

It is visible the disparity between the generated random number considering a different number of samples. It is also perceptible that the generated numbers did not stabilize not even for the 5000 samples since it is visible peaks and discontinuities throughout the sampled line. Hence, MC demands a significant number of samples, which demands an extra computational complexity.

5.4.2.2 Latin Hypercube sampling

Based on the same principles of MC, LHS is a more accurate method since it is an outcome of stratified sampling. Figure 5.22 and Figure 5.23 present the LHS simulation application, the same two distinct variables previously analysed with a different number of samples considered.

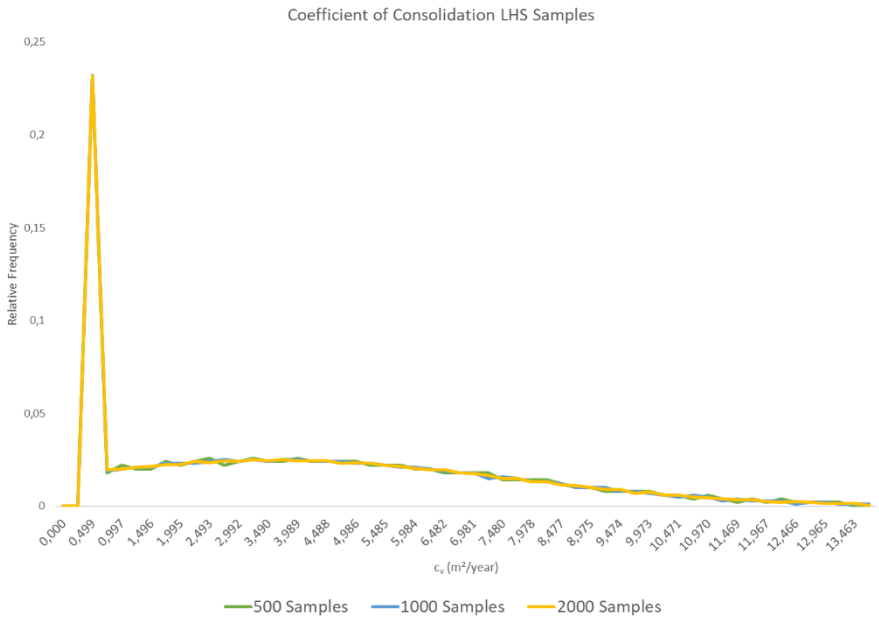

Figure 5.22 Coefficient of consolidation LHS samples contrast

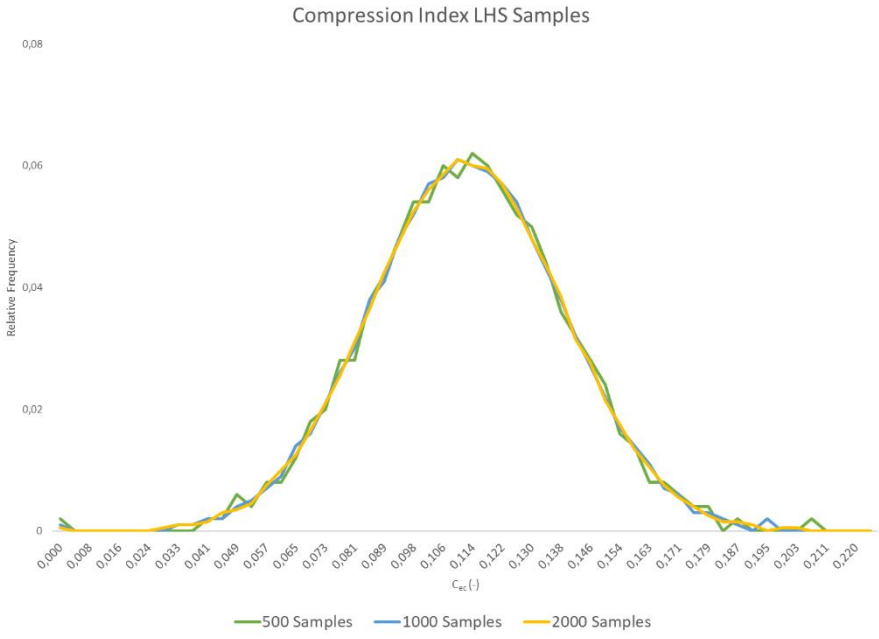

Figure 5.23 Compression index LHS samples contrast

The output results are clearly more stabilized when compared with the MC approaches. LHS shows some disparity when considering 500 samples, but it is noticeable that stabilizes when 1000 samples are tested. Even though LHS is not as simple as MC to variables modelling, it reduces the number of samples needed and gives more reliable results for fewer samples values also with less computational complexity.

5.4.2.3 Final Remarks

For both methods, it is evident to conclude that by increasing sample sizes, the results get more accurate. The range of samples values is smaller in LHS, but even though the curves almost coincide, a fact that does not occur when applying MC. LHS can achieve better performance using the same number of samples when compared to MC and can achieve the same outcome using a fewer number of samples. The difference in terms of computational complexity is significant. Fewer numbers of samples result in less computational complexity, which is only possible using LHS. This analysis allows concluding that an optimization that can be executed on the variables modelling process is to implement LHS with 1000 samples since it was evident its reliability through the comparison between the previous figures displayed.

5.5 ADOPTED MODELLING VARIABLES

Founded on all the assumptions and steps that were previously showed and discussed along with this chapter, who played a crucial role to forward the final variables modelling, Figure 5.24 and Figure 5.25 represent the final variables modelling that is going to be considered in the analysis of the case study. This modelling is a result of the LHS simulation method, considering 1000 samples.

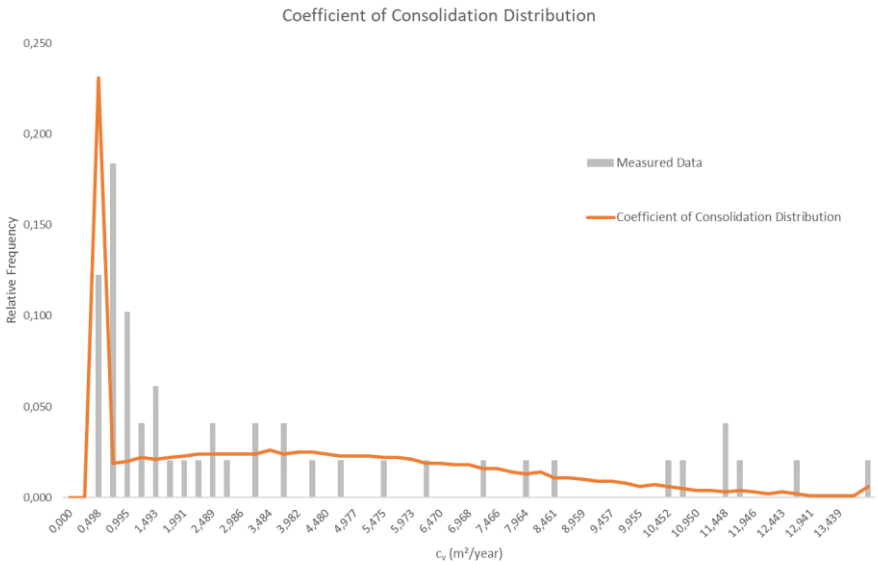

Figure 5.24 Coefficient of consolidation final distribution

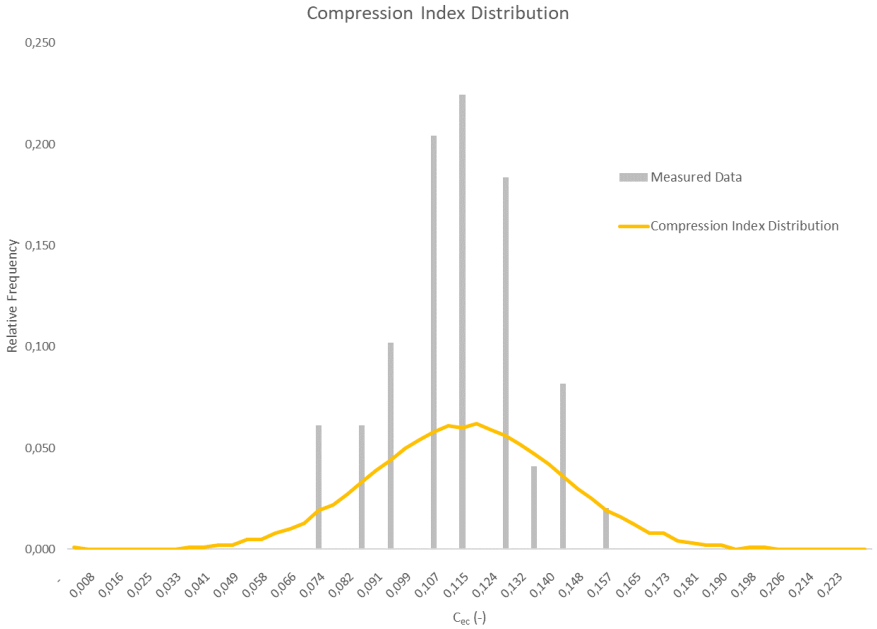

Figure 5.25 Compression index final distribution

The coefficient of consolidation and the compression index final distributions that define the mentioned variables are the Johnson SB and Dagum distributions, respectively.

6
Settlements Analysis

6.1 SETTLEMENT DESIGN

Settlement design is not part of this project. However, its understanding is crucial as it concerns the final aim of any geotechnical design of a reclamation. The modelling variables obtained in the previous chapter are the inputs to the settlement design, representing the innovative characteristic in the design development conversely to the usual deterministic approach. The proposed improvement consists, instead of having a single settlement versus time curve, on the possibility to obtain a range of settlements curves with their associated probability to occur, providing a more detailed and sensible design. This section presents the settlements calculated following the methodology presented in section 3.2.4, applying the modelling variables developed in the previous chapter.

The drain characteristics and the soil properties used in the design are presented in Table 6.1 and Table 6.2, respectively. The preloading height is estimated to go up to 10 meters above the seabed, separated in different stages due to the weak bearing conditions of the seabed layer. Combining the preloading with the PVDs technique, the required settlement is achieved after a specific time. At this point, part of the preloading is removed to the final pretended level of soil above the water, which in this particular case is 3 meters. That particular point (i.e., the required settlement at a specific time) is defined based on the remaining settlement curves discussed on the next section.

Table 6.1 Settlement design drains characteristics

Drains drid geometry		Drain characteristics	
Shape	Rectangular	Equivalent diameter d_w (m)	0,07
Drain spacing (m)	1.75	Max. drain flow (m^3/s)	1,00E-05
Drain lenght (m)	20	Smear zone ratio	1.5

Table 6.2 Settlement design soil characteristics

Soil Characteristics	
Layer	Silty Clay
Layer thickness (m)	100
Coefficient of consolidation (m²/year)	Modelling Variable
Compression index	Modelling Variable
Soil to smear ratio	2
Permeability ratio	3

6.2 REMAINING SETTLEMENT

The evaluation of the remaining settlement, that is the settlement that will occur after removing the preloading, will indicate how much the final platform will settle and for how long. In a port expansion project, this is decisive to identify the time to start operating. In this particular case, it is the customer who will decide on how conservative the design is going to be, in order to have the safety conditions to start operating. Since soil settlements can last for more than 20 years, the analysis is made based on the remaining settlements. The purpose of the building structure and its functionality after the end of construction, namely its sensitivity to additional settlements, will condition the decisions. In this reclamation case, the decision will consider the following points:

- The remaining settlement, as well as the time to remove the preloading and start the construction phase;
- The need for implementing more or diverse methods to improve consolidation speeding up settlement rate;
- Avoid future costs due to unexpected settlements;
- Avoid unplanned constructions damages.

The importance of this study is to show the impact on the settlements curves when different input variables values are considered. In this practical case, the variable values correspond to the soil parameters coefficient of consolidation and the compression index due. Varying these parameters, the soil settlement also vary resulting in different soil settlement curves.

The engineer model adopted to calculate the soil settlements is the same adopted in previous projects and designs with deterministic approaches. The only distinction to the previous designs is the implementation of the probabilistic methods developed and detailed in the previous chapter.

Based on COWI A/S inhouse Excel spreadsheets adopted for settlements designs according to the methodology reviewed in section 3.2.4, using the modelled variables as input, the remaining settlements curves are the ones presented in Figure 6.1 for different probability values. In addition the curve that represents the deterministic analysis is also presented for comparison (dashed line)

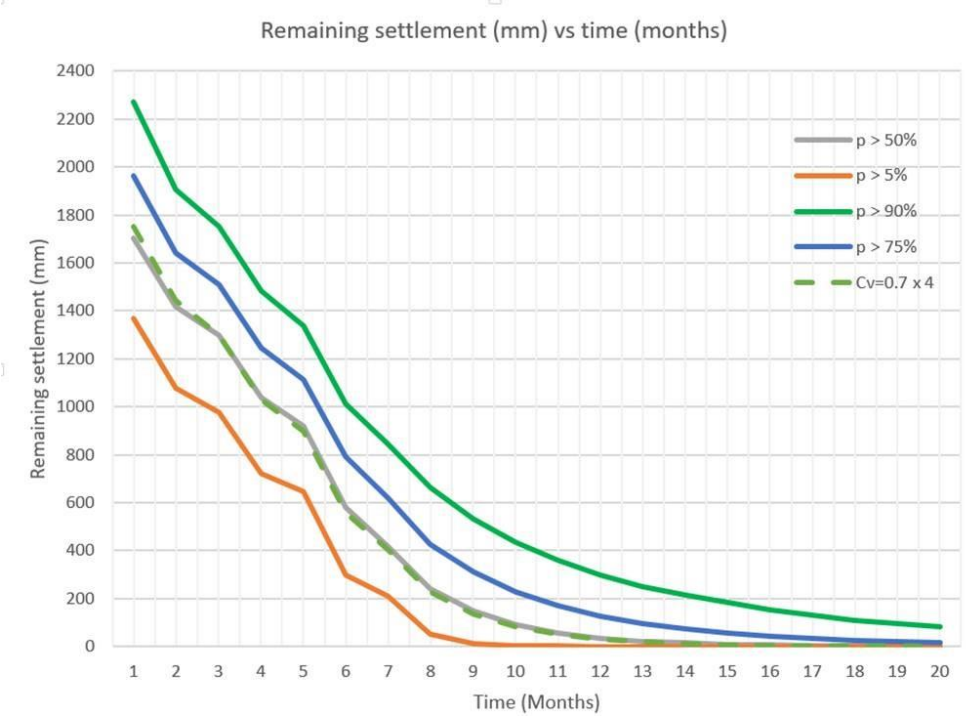

Figure 6.1 Remaining settlements distribution

Over the analysis of the obtained curves for each probability, calculating a probability to each settlement curve it is possible to assemble them a distribution.

The design made with the deterministic approach, considering a single value for each input parameter overlaps the 50% probability remaining settlement curve. This curve is represented as a cross-check of the adopted approach. In addition, it physically indicates that the design can be much more conservative, especially in land reclamation projects where the time to reach the final settlement is crucial.

Depending on how much risk the client wants to accept, one of the curves that concern the remaining settlement is selected. In this practical case, the risk is associated with the construction that is going to be built on top of the reclamation area.

Probabilistic methods allow to have a more rigorous risk analysis. The risk associated with the failure of the structures and the materials is well measured when probability is correlated to each case, since the soil settlement is probabilistically defined. So, it is possible to measure the structures behaviour in order to quantify the risk of failure associated. The client decides how soon wants to start constructing being informed about the associated risk, which in this case, concerns the soil settlement that will still occur after preloading removal.

However, it is essential to cross-check the obtained distribution with its physical meaning. In this case, the distribution was compared with the deterministic curve settlement in order to validate it. It should be highlighted that the deterministic method, considers the safety factors adopted in the Eurocode.

7
Conclusions and Future Developments

7.1 CONCLUSIONS

Probabilistic approaches represent an excellent approach to combine random variables whose analytical methods cannot accomplish. The most important feature is the fact that the introduced data should be correct from an engineering point of view. Once probabilistic approaches complement the sensitive analysis that engineers accomplish over the years, a critical thought and literature reviews must be used for validation.

According to the case in study, the method to implement must be appropriate to the case under discussion. Monte Carlo simulation is the simplest method since it does not report any restrictions. However, for a specific case, one of the quantitative methods mention in section 2.2.5.3, can be selected to solve the problem more accurately.

The distribution fit for the introduced random variables must also count with an engineering judgment. If from the literature review, it is concluded that some parameters always follow a specific probability distribution, and the outcome from the analysis does not match, the engineering judgment is crucial to find a solution.

The present case study highlights the need to run a Monte Carlo simulation for a different number of samples (e.g. 500, 1000, 5000) in order to check the convergence of the outcome result. This step is vital for a qualitative analysis where the error propagation was not yet confirmed. Visualising the different trials it is possible to conclude that the method is well implemented.

The contrast between Monte Carlo simulation and Latin Hypercube sampling was evident. Latin Hypercube sampling can achieve better performance using the same number of samples when compared to the Monte Carlo simulation. It can achieve the same outcome using a fewer number of samples. The probability distributions that fitted the best for each parameter and fulfilled all the requirements needed were the Johnson SB distribution and Dagum distribution, for the coefficient of consolidation and the compression index, respectively.

The settlement curve also has to be cross-checked with their physical meaning, namely with the theoretical curve settlement in order to validate it, detecting disparities or dissimilarities that can correspond to a not realistic outcome result.

7.2 FUTURE DEVELOPMENTS PERSPECTIVES

This case study was based on a numerical resolution. Through the consideration of c_v and C_{ec} descriptive statistics obtained in the laboratory measures and in-situ tests, implementing simulation methods, it was

possible to conceive a numerical response. These types of analysis have an error associated. In statistics, the error propagation, also named uncertainty propagation, is a verification that must be estimated. When dealing with uncertainty, getting further calculations involves an increase in the error propagation. Nowadays were already elaborated tools to evaluate error propagation based on formulas and methods that can be applied in this case. The implementation of these error estimation is an added value binding when applying statistical approaches. The next step for the continuity of this study is to achieve the error estimation value in order to support the analysis made.

Likewise, the addition of more input variables can complement the settlement results. Modelling other soils parameters, introducing them as inputs in the settlement designs, improves the outcome results, being more precise and explicit in the matter in question. Two examples of extra input variables that shall be of concern in the future are the consolidation layer thickness and the soil unit weight.

During real-time construction, Bayesian statistical approaches are a powerful tool to implement. As soon as the construction starts being executed, the instruments in the field start to register measures. Apart from comparing the expected results obtained with the real data, considering the real measures, introducing them in the analysis as a prior, it is possible to back-calculate the soil parameters. Applying to Bayesian, it is possible to see how the measurements go with time in reality when compared to the expected. Moreover considering the measured data as prior, it is also possible to further define a new settlement curve adopted with the real data. An introduction to this approach is contemplated in section 3.2.5.2.

Another approach that can be conceived to crosscheck the numerical modulation or to solve the case study is based on analytical and theoretical approaches. All the parameters formulas are obtained by manipulating all the formulas with the variables PDF. Considering that not every variable has a PDF defined, is humanly exhaustive and almost inconceivable in proper time. Appealing to MATLAB and introducing all the formulas and PDFs parameters, it is possible to do mathematical calculations to obtain a formula of the settlement as a function of time. This approach is valuable to evaluate the cross-check and differences between the numerical and the analytical approach. The main advantage of using a numerical approach is the fact that it is adaptable to almost every practical case. The analytical approach is not so flexible, since the formulas and conditions are continually changing, which means that implementing analytical approaches in this specific case, would not provide useful conclusions in a reasonable time.

This case study can represent a starting point for the implementation of statistical approaches in Geotechnical engineering strategies.

REFERENCES

Administration, S. T. (2013). *Swedish Transport Administration's Technical Requirements on Geotechnical Engineering Structures.*

Akbas, S., & Kulhawy, F. (2009). Reliability-Based Design Approach for Differential Settlement of Footings on Cohesionless Soils. doi:10.1061/(ASCE)GT.1943-5606.0000127

Alfredo, H.-S. A., & Tang, W. H. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering.* Hoboken, NJ: John Wiley Publishers.

Baecher, G. B. (2017). Bayesian Thinking in Geotechnics.

Baecher, G. B., & Christian, J. (2003). *Reliability and Statistics in Geotechnical Engineering.*

Barron, R. A. (1948). *The equivalent diameter of the bandshaped drain* (Vol. 113). Trans. ASCE.

Basu, D., Basu, P., & Prezzi, M. (2013). *A Rational Approach to the Design of Vertical Drains Considering Soil Disturbance.* doi:10.1061/9780784412770.037

Bergado, D., Miura, Balasubramaniam, N., & Arumugam. (1998). PVD improvement of soft Bangkok clay with combined vacuum and reduced sand embankment preloading. *29*. Retrieved from https://pdfs.semanticscholar.org/89b6/0eab6886f394b180fb4678cd4c9389a44e0f.pdf?_ga=2.17642681.1892793600.1587974105-1547314848.1587974105

Boivin, P. (2011). Shrinkage and Swelling Phenomena in Soils. In *Encyclopedia of Agrophysics.* Springer Netherlands. doi:10.1007/978-90-481-3585-1_139

Boivin, P., Garnier, P., & Vauclin, M. (2006). Modeling the soil shrinkage and water retention curves with the same equations. doi:10.2136/sssaj2005.0218

Building on soft soils: design and construction of earth structures both on and into highly compressible subsoils of low bearing capacity. (1996). Rotterdam, Netherlands; Brookfield, VT, USA: A.A. Balkema.

BULK FREIGHT. (n.d.). (West Coast Tug) Retrieved from http://westcoasttug.ca/bulk-freight

Burt, G. L. (2007). Rock classification. In *Handbook of Geotechnical Investigation and Design Tables.* London: CRC Press. doi:10.1201/b16520-4

Carrillo, N. (1942). Simple Two and Three Dimensional Case in the Theory of Consolidation of Soils. *Journal of Mathematics and Physics, 21*(1-4). doi:10.1002/sapm19422111

Casagrande, A. (1965). The role of the 'calculated risk' in earthwork and foundation engineering. *91*.

Chang, K.-H. (2015). Reliability Analysis. In *e-Design* (pp. 523-595). Boston: Academic Press. doi:https://doi.org/10.1016/B978-0-12-382038-9.00010-7

China – Sri Lanka jointly build "Shining Pearl of Indian Ocean". (2019, January 24). Retrieved from https://cdn.shortpixel.ai/client/to_webp,q_lossless,ret_img,w_898/https://www.beltandroad.news/wp-content/uploads/2019/01/Colombo-Port-City.jpg

Chu, J., Varaksin, S., Ultich, K., & Mengé, P. (2009). *Construction Processes.* 17th International Conference on Soil Mechanics and Geotechnical Engineering. Retrieved from https://www.tc211.be/wp-content/uploads/2018/06/TC211_State_of_the_Art_Report_2009.pdf

Confidence Interval: How to Find a Confidence Interval: The Easy Way! (2020). Retrieved from Statistics How To: https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/confidence-interval

Continuous Distributions. (n.d.). (MathWave Technologies) Retrieved from Mathwave data analysis & simulation: http://www.mathwave.com/help/easyfit/html/analyses/distributions/_continuous.html

COWI A/S. (2020). *Geotechnical Interpretive Report.* Denmark.

Dodge, Y. (2006). *The Oxford Dictionary of Statistical Terms* (6 ed.). Oxford University Press.

Donald, P. C. (1988). Geotechnical Engineering Principles and Practices. Upper Saddle River, NJ, USA: Prentice Hall.

Fenniak, M. (2004, April 30). *Latin Hypercube Sampling*. Retrieved from Mathieu Fenniak: https://mathieu.fenniak.net/latin-hypercube-sampling/

Fernandes, M. d. (2006). *Mecânica dos Solos Conceitos e Princípios Fundamentais* (Vol. 1). Porto: FEUPEdicoes.

Fernandes, M. d. (2011). *Mecânica dos Solos Introdução à Engenharia Geotécnica Volume 2.* Porto: FEUPedicões.

Frost, J. (2020). *Normal Distribution in Statistics*. Retrieved from Statistics By Jim: https://statisticsbyjim.com/basics/normal-distribution/

Generalized Extreme Value Distribution. (n.d.). Retrieved from MathWorks: https://se.mathworks.com/help/stats/generalized-extreme-value-distribution.html

Gentle, J. (2010). *Computational Statistics.* Virginia: Elsevier. doi:https://doi.org/10.1016/B978-0-08-044894-7.01316-6

Gilkinson, K. D. (2003). Immediate and longer-term impacts of hydraulic clam dredging on an offshore sandy seabed: effects on physical habitat and processes of recovery. *Continental Shelf Research, 23*(14-15), 1315-1336. doi:10.1016/s0278-4343(03)00123-7

Glen, S. (2016, May 17). *Truncated Distribution / Truncated Normal Distribution*. Retrieved from Statistics How To: Elementary Statistics for the rest of us!: https://www.statisticshowto.com/truncated-normal-distribution/

Glen, S. (2020, May 26). *Dagum Distribution: Definition, CDF & PDF*. Retrieved from Statistics How To: https://www.statisticshowto.com/dagum-distribution/

Haines, W. B. (1923). The volume-changes associated with variations of water content in soil. *13*(3), 296–311. doi:doi.org/10.1017/S0021859600003580

Hansbo, S., Jamiolkowski, M., & Kok, L. (1981). Consolidation by vertical drains. *31*(1), 45-66. doi:10.1680/geot.1981.31.1.45

Helz, R. L. (2005). *Monitoring Ground Deformation from Space.* U.S. Department of the Interior. Reston: U.S. Geological Survey. Retrieved from https://pubs.usgs.gov/fs/2005/3025/2005-3025.pdf

Holtz, R. D. (1987). Preloading with prefabricated vertical strip drains. *Geotextiles and Geomembranes, 6*(1), 109-131. doi:https://doi.org/10.1016/0266-1144(87)90061-6

How would you describe the difference between silt and clay when doing an experiment? (2013, September 7). (eNotes Editorial) Retrieved 2020, from eNotes: https://www.enotes.com/homework-help/how-would-you-decribe-diffrence-betwwen-silt-clay-452365

Indraratna, B., Sathananthan, I., Bamunawita, C., & Balasubramaniam, A. S. (2015). Chapter 3 - Theoretical and Numerical Perspectives and Field Observations for the Design and Performance Evaluation of Embankments Constructed on Soft Marine Clay. In B. Indraratna, J. Chu, & C. Rujikiatkamjorn (Eds.), *Ground Improvement Case Histories* (pp. 83-122). Butterworth-Heinemann. doi:https://doi.org/10.1016/B978-0-08-100192-9.00003-X

ISO/TC. (2018). *ISO 31000 (2018). Risk Management – Guidelines. Technical Committee* (2 ed.). Technical Committee ISO/TC 262. Retrieved from https://www.iso.org/standard/65694.html

Jain, D. (2018, August 23). *Skew and Kurtosis: 2 Important Statistics terms you need to know in Data Science*. Retrieved from codeburst.io: https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis-388fef94eeaa

Jimmy, D. (2012, January 16). *APE News*. (APE Mid-Atlantic (VA))

John, H. (2016, January 4). *Scientific American*. Retrieved from https://blogs.scientificamerican.com/cross-check/bayes-s-theorem-what-s-the-big-deal/

Karlhuber, T. (2008, February 7). Retrieved from https://upload.wikimedia.org/wikipedia/commons/2/24/The_universe.jpg

Kulhawy, H. F., & Mayne, W. P. (1990). *Manual on Estimating Soil Properties for Foundation Design.* Cornell University, Geotechnical Engineering Group, New York. Retrieved from https://www.geoengineer.org/storage/publication/20489/publication_file/2745/EL-6800.pdf

Lacasse, S., & Nadim, F. (1998). Risk and Reliability in Geotechnical Engineering. Retrieved from https://scholarsmine.mst.edu/icchge/4icchge/4icchge-session00/11

Lacasse, S., & Nadim, F. (2007). Probabilistic geotechnical analyses for offshore facilities. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 1*(1), 21-42. doi:10.1080/17499510701204224

Lacasse, S., Nadim, F., Boylan, N., Liu, Z., & Choi, Y. (2019, January). Risk Assessment and Management for Geotechnical Design of Offshore. *Offshore Technology Conference.* doi:10.4043/29398-MS

Lacasse, S., Nadim, F., Liu, Z., Eidsvig, U., Le, T., & Lin, C. (2019). Risk assessment and dams – Recent developments and applications. doi:10.32075/17ECSMGE-2019-1110

Learn More About EasyFit. (2004-2010). (MathWave Technologies) Retrieved from MathWave data analysis & simulation: http://www.mathwave.com/articles/goodness_of_fit.html#ks

Lewinson, E. (2019, April 16). *Explaining probability plots*. Retrieved from Towards Data Science: https://towardsdatascience.com/explaining-probability-plots-9e5c5d304703

(2002). *Magnet.* Washington: Slope Indicator Company. Retrieved from https://durhamgeo.com/pdf/manuals/magnet-extensometer.pdf

Mesri, G., & Choi, Y. K. (1985, April). Settlement analysis of embankments on soft clays. *111*(4), 441-464. doi:https://doi.org/10.1061/(ASCE)0733-9410(1985)111:4(441)

Natvig, B. (2005, February 9). *Appendix A: Sampling Methods* . University of Oslo. Retrieved from University of Oslo: https://www.uio.no/studier/emner/matnat/math/STK4400/v05/undervisningsmateriale/Sampling%20methods.pdf

Oliveira, P. M. (2010). Estatística.

Peck, R. B. (1969). Advantages and Limitations of the Observational Method in Applied Soil Mechanics. *Géotechnique*, *19*(2), 171-187. doi:10.1680/geot.1969.19.2.171

Phoon, K.-K., Prakoso, W., Wang, Y., & Ching, J. (2016). Chapter 3 Uncertainty representation of geotechnical design parameters. doi:10.1201/9781315364179-4

Piezometers. (2019). Retrieved from Geotechnical Observations: http://www.geo-observations.com/piezometers

Pistilli, T. (2019, May 23). *Behind The Models: Cholesky Decomposition*. Retrieved from Towards Data Science: https://towardsdatascience.com/behind-the-models-cholesky-decomposition-b61ef17a65fb

Qian, J. H., Zhao, W. B., Cheung, Y. K., & Lee, P. K. (1992). The theory and practice of vacuum preloading. *Computers and Geotechnics, 13*(2), 103-118. doi:https://doi.org/10.1016/0266-352X(92)90027-Q

Rajapakse, R. (2016). Geotechnical instrumentation. In *Geotechnical Engineering Calculations and Rules of Thumb.* Elsevier. doi:https://doi.org/10.1016/B978-0-12-804698-2.00027-1

Rujikiatkamjorn, C. (2005). Analytical and Numerical Modelling of Soft Clay Foundation Improvement Via Prefabricated Vertical Drains and Vacuum Preloading.

Shape accel arrays. (n.d.). (GEO-Instruments) Retrieved from https://www.geo-instruments.co.uk/expertise/technologies/shape-accel-arrays

Spross, J., & Larsson, S. (2019). Probabilistic Observational Method for Design of Surcharges on Vertical Drains. *0*(0), 1-13. doi:10.1680/jgeot.19.P.053

Spross, J., Prästings, A., & Larsson, S. (2019). *Probabilistic Evaluation of Settlement Monitoring with the Observational Method during Construction of Embankments on Clay.* doi:10.3850/978-981-11-2725-0-IS3-3-cd

Stephanie. (2016, July 15). *Triangular Distribution / Triangle Distribution: Definition*. Retrieved from Statistics How To: https://www.statisticshowto.com/triangular-distribution/

Stephanie, G. (n.d.). *Correlation Coefficient: Simple Definition, Formula, Easy Steps*. Retrieved from Statistics How To: Elementary Statistics for the rest of us!: https://www.statisticshowto.com/probability-and-statistics/correlation-coefficient-formula/

Sun, J., & Zhao, Z. (2013, December). Stability Charts for Homogenous Soil Slopes. *139*(12). doi:https://doi.org/10.1061/(ASCE)GT.1943-5606.0000938

Tatsuoka, C. (2002, July 30). Data analytic methods for latent partially ordered classification models. doi:https://doi.org/10.1111/1467-9876.00272

Ural, N. (2018). The Importance of Clay in Geotechnical Engineering. In *Current Topics in the Utilization of Clay in Industrial and Medical Applications.* doi:10.5772/intechopen.75817

Wang, Y., Zhao, T., & Cao, Z. (2019). *Bayesian Perspective on Ground Property Variability for Geotechnical Practice*. doi:10.3850/978-981-11-2725-0-key5-cd

Weisstein, E. W. (n.d.). *Beta Distribution*. Retrieved from MathWorld--A Wolfram Web Resource: https://mathworld.wolfram.com/BetaDistribution.html

Wicklin, R. (2020, January 20). *The Johnson SB distribution*. Retrieved from BlogsSas: https://blogs.sas.com/content/iml/2020/01/20/johnson-sb-distribution.html

Wikimedia Commons. (2011, July 12). Retrieved from https://upload.wikimedia.org/wikipedia/commons/thumb/6/60/Barge_%C3%A0_charbon.jpg/800px-Barge_%C3%A0_charbon.jpg

Wittwer, Jon. (2004). *Monte Carlo Simulation Basics*. Retrieved from Vertex42: https://www.vertex42.com/ExcelArticles/mc/MonteCarloSimulation.html

YH. (2007, January 10). Retrieved from https://upload.wikimedia.org/wikipedia/commons/6/63/Yangshan_Deepwater_Port.jpg

APPENDIX

Appendix A: Available Data

A.1 Considered Data

Table A. 2 Considered data

Samples	c_v (m²/year)	E_{oed} (MPa)_100-200kPa	C_{ec} (-)	E_{oed} (MPa)_200-400kPa
1	0.39	3.90	0.12	6.30
2	0.71	3.90	0.11	6.60
3	0.68	7.60	0.09	10.10
4	0.77	4.40	0.10	7.40
5	0.47	3.50	0.14	5.60
6	0.80	4.50	0.11	7.10
7	0.45	4.20	0.14	6.10
8	0.50	4.00	0.11	7.00
9	0.63	5.20	0.10	8.00
10	0.56	4.20	0.10	6.80
11	0.61	3.60	0.11	6.30
12	3.10	4.10	0.11	6.90
13	0.46	4.10	0.12	7.10
14	0.79	4.80	0.11	7.50
15	0.82	3.80	0.11	7.10
16	0.44	4.90	0.10	7.50
17	0.62	4.30	0.11	6.20
18	0.63	3.70	0.14	6.10
19	0.62	4.90	0.09	8.00
20	1.44	4.50	0.10	8.10
21	2.00	4.70	0.08	8.60
22	1.05	4.20	0.12	6.30
23	1.85	6.00	0.10	8.30
24	4.70	5.10	0.10	7.50
25	0.45	4.10	0.12	6.00
26	3.52	4.80	0.11	7.20
27	8.31	8.30	0.09	11.30
28	7.72	6.40	0.07	10.70
29	1.45	6.00	0.09	8.10
30	2.31	6.90	0.08	7,00
31	6.99	4.20	0.09	7.70
32	1.51	4.70	0.10	8.70
33	1.46	4.40	0.13	6.40
34	1.37	6.30	0.15	5.70
35	2.32	4.80	0.12	7.20
36	3.70	6.00	0.07	9.90
37	2.66	4.70	0.10	7.50
38	0.77	4.40	0.11	7.30
39	4.05	6.90	0.11	10.20
40	5.24	8.40	0.10	11.00
41	6.20	5.50	0.12	6.80
42	11.21	3.90	0.12	5.80
43	10.39	4.90	0.24	6.50
44	10.55	6.10	0.08	10.00
45	14.78	8.70	0.07	11.80
46	2.99	4.80	0.13	6.10
47	11.62	4.70	0.14	5.40
48	11.40	4.80	0.12	6.10
49	1.19	6.30	0.12	8.00