FACULTY OF ENGINEERING OF THE UNIVERSITY OF PORTO

Vibrations induced by pile driving: an expedite methodology for prediction and mitigation

Ahmed Moustafa Kamel Mohamed Abouelmaty Omar



Doctoral Program in Civil Engineering

Supervision team:

Dr. Aires Colaço

Prof. Pedro Alves Costa

2024

Vibrations induced by pile driving: an expedite methodology for prediction and mitigation

Ahmed Moustafa Kamel Mohamed Abouelmaty Omar

Doctoral Program in Civil Engineering

2024

Abstract

Deep foundations are widely used to address issues related to weak geotechnical characteristics of the ground. However, the resulting vibrations from the pile-driving process can negatively impact the integrity and stability of nearby structures and cause significant annoyance. Therefore, prediction excessive of vibrations and its mitigation is crucial.

This work aims to develop an experimentally validated surrogate model with a user-friendly interface for prediction of vibrations induced by pile driving. The surrogate model, based on machine learning techniques, offers the advantages of almost zero-time prediction, low computation cost, high accuracy, and easy of use.

The initial step involves contextualizing the topic, briefly explaining the fundamental aspects and the prediction numerical model used to generate the learning dataset. A parametric study identifies the key variables affecting the generation and propagation of vibrations from pile driving. Additionally, an experimental test site near Aveiro, Portugal, is characterized, where vibrations from pile-driving activities were recorded. The FEM-PML numerical model, previously developed, is used to generates a vibration database. This database, created through extensive simulations, represents typical scenarios encountered in urban building foundation projects. Machine learning techniques analyze the relationship between input factors (e.g., hammer and ground properties) and vibration levels. A comparative study among seven different algorithms was also performed.

As a result, machine learning models were developed to predict ground-borne vibrations from pile driving in different soil profiles. The model is also trained to evaluate the impact of using open trench as a mitigation measure for recessive vibration levels. The vibration response can be predicted in both frequency and space domains with acceptable accuracy. Both XGBoost and Artificial Neural Networks effectively address the phenomena under investigation. However, XGBoost demonstrates superiority over other algorithms used in this work, particularly in handling advanced complex geotechnical conditions. This study introduces an integrated tool for quickly predicting, assessing, and mitigating vibrations induced by pile driving processes.

Keywords: Pile driving; Ground-borne vibrations; Numerical modeling; Experimental test site; Artificial Intelligent; Machine learning; ANNs; XGBoost.

Resumo

As soluções de fundações indiretas são amplamente utilizadas em cenários geotécnicos desfavoráveis. A cravação de estacas é uma possibilidade interessante para a sua materialização. No entanto, as vibrações resultantes deste processo construtivo podem impactar negativamente a integridade e estabilidade das estruturas próximas e causar incômodos significativos. Deste modo, a previsão atempada dos níveis de vibração induzidos é crucial. Este trabalho visa desenvolver um modelo expedito de previsão de vibrações induzidas pela cravação de estacas, acompanhado de uma interface gráfica e validado experimentalmente. Este modelo é baseado em técnicas de inteligência artificial, permitindo uma previsão quase instantânea dos níveis de vibração induzidos, tendo ainda associado um custo computacional reduzido, elevada fiabilidade e facilidade de utilização. A fase inicial do desenvolvimento do presente trabalho envolve a contextualização do tema, com a exposição dos aspetos basilares da problemática em estudo, e a descrição do modelo numérico axissimétrico FEM-PML, o qual serve de suporte ao desenvolvimento do modelo de previsão expedito. Através de um extenso estudo paramétrico são identificadas as principais variáveis que afetam a geração e propagação de vibrações provenientes da cravação de estacas. Posteriormente, é apresentado e caraterizado um campo experimental, no qual vibrações induzidas pela cravação de estacas foram registadas. Estes dados são utilizados na validação experimental do modelo numérico. Este mesmo modelo é posteriormente empregue na geração de uma base de dados alargada. Esta base de dados representa cenários típicos frequentemente encontrados em projetos de fundações de edifícios urbanos. A aplicação de técnicas de inteligência artificial à base de dados referida permitiu o desenvolvimento do modelo de previsão expedito. Este modelo é capaz de prever os níveis máximos de vibração à superfície do solo e o respetivo conteúdo em frequência da resposta, utilizando para tal diferentes algoritmos de inteligência artificial, tal como XGBoost e Redes Neurais Artificiais (ANNs). Ambos os algoritmos são capazes de capturar de forma eficaz as principais tendências do fenómeno em investigação. No entanto, o algoritmo XGBoost demonstra superioridade em relação a outros algoritmos, incluindo ANNs, utilizados neste trabalho, particularmente para condições geotécnicas mais complexas. O modelo foi igualmente treinado para avaliar o impacto do uso de trincheiras abertas como medida de mitigação para casos com níveis de vibração excessivos. Este trabalho apresenta uma ferramenta integrada para a previsão, avaliação e mitigação de vibrações induzidas por atividades de cravação de estacas.

Palavras Chave: Cravação de estacas; Vibrações no solo; Modelação numérica; Validação experimental; Inteligência artificial; Aplicação computacional.

Acknowledgements

I express my deepest gratitude to my distinguished supervisor, ***Dr. Aires Colaço***, whose unwavering guidance, steadfast support, and insightful instructions have played a pivotal role in shaping this work. His expertise and mentorship have been indispensable throughout my Ph.D. journey.

My appreciation and acknowledgment should also go to ***Prof. Pedro Alves Costa***, whose invaluable feedback, warm welcome into his laboratory, and unwavering support have significantly contributed to this work. His support and guidance laid a robust foundation for this achievement.

I extend my thanks to my exceptional circle of friends, including ***Yassine Boukili***, ***Omar As-Saadi***, ***Lana Farhat***, ***Georges Hatem***, ***Reihane Shafie*** ,***Chanalisa Salvi*** and ***Anis Mohabeddine***, who have been integral to my journey, shaping my confidant zone, and providing unwavering support.

Special acknowledgment goes to ***Amr Ragab***, ***Motaz Ayiad*** and ***Ahmed Faris***, who consisted of my surrogate family, offering guidance, support, and inclusion during the most challenging times I faced. I am grateful for their generosity and kindness, a constant source of strength throughout my Ph.D. journey.

To my beloved family, whose unwavering love, encouragement, and understanding have been my bedrock, I owe the most profound gratitude. Their sacrifices and constant support have been the driving force behind my academic pursuits. Their belief and love have sustained me through the challenges. Indeed, this achievement is as much theirs as it is mine, and I am profoundly thankful for the solid foundation they have provided throughout my entire life.

This work was financially supported by: Project PTDC/ECI-CON/29634/2017 - POCI-01-0145-FEDER-029634 – VIPIB - funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI); national funds through the FCT/MCTES (PIDDAC), under the project PTDC/ECI-EGC/3352/2021 – IntRAIL – Intelligent prediction tool for ground-borne noise and vibrations induced by railway traffic. The author also acknowledges FCT for the financial support provided through the Individual Grant No. 2022.11639.BD.



Dedication

"To my mother, the embodiment of unconditional love, and to my father, who passed early in this trip, I dedicate this work with profound reverence and gratitude."

"Knowledge is three oceans; whoever realizes the first becomes arrogant,
whoever realizes the second becomes humble,
and the third ocean cannot realize."

Emam Al-shafi'i, 767-820

Contents

1 Introduction **1**
- 1.1 Context . . . 1
- 1.2 Aim and objectives . . . 2
- 1.3 Work description . . . 3

2 Vibration induced by pile driving: a general overview **5**
- 2.1 Generalities . . . 5
- 2.2 Pile driving technology . . . 6
- 2.3 Generation and propagation of seismic waves . . . 9
  - 2.3.1 Soil behavior under driving process . . . 9
  - 2.3.2 Waves propagation in soil . . . 10
  - 2.3.3 Energy transmission during impact pile driving . . . 12
  - 2.3.4 Wave attenuation . . . 13
- 2.4 Prediction and measurement of vibrations induced by pile driving . . . 15
  - 2.4.1 Vibration key indicators . . . 15
  - 2.4.2 Empirical methods . . . 16
  - 2.4.3 Numerical approaches . . . 18
  - 2.4.4 International standards . . . 21
- 2.5 Prediction methods based on machine learning techniques . . . 22
  - 2.5.1 Machine learning techniques . . . 22
  - 2.5.2 Machine learning prediction models . . . 23

3 Vibrations induced by impact pile driving: numerical modeling and phenomenological analysis **27**
- 3.1 Introduction . . . 27
- 3.2 Numerical prediction model . . . 28
  - 3.2.1 General considerations . . . 28
  - 3.2.2 Modelling of the pile-ground system . . . 29
  - 3.2.3 Hammer model . . . 32
  - 3.2.4 Numerical validation . . . 34
- 3.3 Physics of the generation and propagation of ground vibration . . . 37
  - 3.3.1 Reference case . . . 37
  - 3.3.2 The influence of ground properties . . . 40
  - 3.3.3 Pile properties and penetration depth influence . . . 43
  - 3.3.4 Hammer properties influence . . . 45
  - 3.3.5 Results discussion . . . 46
- 3.4 Final conclusions . . . 51

4 Experimental and numerical prediction of ground-borne vibrations induced by impact pile driving 53
4.1 Generalities . . . 53
4.2 Low strain test . . . 55
4.2.1 General overview and objectives . . . 55
4.2.2 Description of the experimental activities and main results . . . 56
4.2.3 Low strain results . . . 57
4.2.4 Numerical evaluation of the ground properties . . . 62
4.3 Vibration induced by impact pile driving . . . 66
4.3.1 Description of the experimental activities . . . 66
4.3.2 Experimental results . . . 66
4.3.3 Experimental validation of the numerical model . . . 71
4.4 Experimental evaluation of the influence of the ground stiffness on ground vibration 73
4.4.1 Vibration levels for two test sites . . . 73
4.5 Numerical study of a mitigation measure . . . 77
4.5.1 General overview of the numerical model . . . 77
4.5.2 Trench solution . . . 79
4.5.3 Continuous distribution of a heavy mass resting on the ground surface next the pile . . . 81
4.6 Final remarks . . . 83

5 Artificial intelligent-based prediction models for vibration induced by pile driving 85
5.1 Introduction . . . 85
5.2 Generation of the learning database . . . 86
5.2.1 Feature selection . . . 86
5.2.2 Simulation technique of the input parameters . . . 88
5.2.3 The execution of the learning database . . . 91
5.3 Data division strategy . . . 93
5.4 Data Standardization and Normalization . . . 94
5.5 Model stopping criteria . . . 95
5.6 Model performance indicator . . . 96
5.7 Machine learning algorithms . . . 97
5.7.1 Artificial Neural Networks, ANNs . . . 97
5.7.2 Multiple linear regression . . . 99
5.7.3 Regression tree . . . 99
5.7.4 Random forest . . . 100
5.7.5 Gaussian Process Regression . . . 101
5.7.6 Support Vector Machine . . . 102
5.7.7 Extreme Gradient Boosting . . . 103
5.8 Application of ANNs model . . . 104
5.8.1 Peak particle velocity prediction . . . 104
5.8.2 Vibration velocity prediction in frequency domain . . . 108
5.9 Comparative study based on different machine learning models . . . 110
5.9.1 Hyper-parameter optimization . . . 110
5.9.2 Machine learning model performance assessment . . . 113
5.10 Machine learning model for soil profile with 3 layers . . . 118
5.10.1 General introduction . . . 118
5.10.2 Learning data set development . . . 119
5.10.3 Development of the machine learning model . . . 119

5.11 Experimental validation 121
5.12 ANNs application in vibration mitigation 123
5.13 Standalone computer application 125
5.13.1 ZBZBAT copyright registration 126
5.13.2 ZBZBAT website 126
5.14 Final conclusions 126

6 Conclusions and future works **129**
6.1 Introduction 129
6.2 Work performed summary 129
6.3 Future work 131

List of Figures

1.1 Ground-borne vibrations induced by pile driving: influence zone. 2
1.2 Work-breakdown structure. 4

2.1 Illustration for driving process influence on the surroundings. 5
2.2 Pile classification and construction technique. 6
2.3 Piles types advantages and drawbacks. 7
2.4 The Scottish steam hammer. 7
2.5 Pile hammering machines: a) diesel hammer, b) vibratory hammer, c) hydraulic hammer. 8
2.6 Hypothetical soil behavior zones near a driven pile [12]. 9
2.7 Waves propagation: a) body waves, b) surface waves. 10
2.8 Horizontal and vertical vibration amplitude of the Rayleigh wave [17]. 12
2.9 Generation of seismic waves during impact driving in homegeneous soil [21]. . . 13
2.10 Predicted vibration levels for class II and class III. 18
2.11 Comparative analysis for the empirical approaches Colaço et al. [53]. 18
2.12 Envelope of the PPV Vs. the distance from the pile due to impact pile driving with a transferred energy from 3.4 to 19.2 kJ, computed by, a) linear analysis; b) nonlinear analysis, and compared with experimental results(dashed-dotted line) [56]. 19
(a) linear analysis 19
(b) nonlinear analysis 19
2.13 Khoubani and Ahmadi [57] model validation. 20
(a) Experimental verification 20
(b) Numerical verification 20
2.14 Sofiste et al. model: numerical validation [60]. 20
2.15 Experimental validation for the numerical model Colaço et al. [4]. 21
2.16 Standards and guidelines limits for PPV. 22
2.17 Machine learning algorithms 23
2.18 Validation for the artificial neural network model. 24
2.19 Validation for the artificial neural network model proposed by Connolly et el. [67] 25

3.1 Vibrations induced by pile driving: representative scheme of the main physics. . . 28
3.2 Representative scheme of the numerical modelling approach. 28
3.3 Pile ground system modeled by an axisymmetric FEM-PML numerical approach. 29
3.4 Stress – strain path during cyclic loading [76]. 31
3.5 Modulus reduction curves for: (a) non-plastic soils and (b) plastic soils [78]. . . . 32
(a) 32
(b) 32

3.6 Two degrees of freedom pile-hammer model [4]. 34
3.7 Free vibration (blue line) and truncated force (red line) applied at the anvil [85]. . 34
3.8 Geometry of an embedded pile in a layered medium. 35
3.9 FE–PML mesh adopted to model the pile-ground system. 35
3.10 Impact hammer force: (a) time history, (b) frequency content. 36
(a) 36
(b) 36
3.11 Numerical validation between FEM-PML model and Masoumi et al. model [55]. 37
3.12 Evolution of the dynamic properties of a plastic soil (PI=50) with shear strain: a) stiffness degradation; b) damping increase [78]. 38
3.13 Pile driving in a homogeneous half-space soil: a) application example; b) pile-ground axisymmetric FE–PML mesh. 38
3.14 PPV vs. distance for the reference case: a) linear analysis; b) equivalent linear analysis. 39
(a) 39
(b) 39
3.15 Velocity versus distance: a) linear analysis; b) equivalent linear analysis. 40
(a) 40
(b) 40
3.16 Vertical velocity of the particles, in the frequency domain, assuming: a) linear analysis; b) equivalent linear analysis. 40
(a) 40
(b) 40
3.17 PPV versus distance from the pile due to impact pile driving for different ground stiffness: a) linear analysis; b) equivalent linear analysis. 41
(a) 41
(b) 41
3.18 PPV versus distance for different soil profiles: a) linear analysis; b) equivalent linear analysis. 42
(a) 42
(b) 42
3.19 PPV versus distance from the pile for different soil damping value: a) linear analysis; b) equivalent linear analysis. 42
(a) 42
(b) 42
3.20 PPV versus distance for different penetration depth: a) linear analysis; b) equivalent linear analysis. 43
(a) 43
(b) 43
3.21 PPV vs. distance for different pile diameter: a) linear analysis; b) equivalent linear. 44
(a) 44
(b) 44
3.22 PPV vs. distance for different pile materials: a) linear analysis; b) equivalent linear. 44
(a) 44
(b) 44
3.23 PPV versus distance for different cushion stiffness: a) linear analysis; b) equivalent linear analysis. 45
(a) 45

(b) . . . 45
3.24 PPV versus distance for different ram mass: a) linear analysis; b) equivalent linear analysis. . . . 46
(a) . . . 46
(b) . . . 46
3.25 PPV versus distance for different drop height values: a) linear analysis; b) equivalent linear analysis. . . . 46
(a) . . . 46
(b) . . . 46
3.26 Parametric based presenting for MAAD evaluation for the linear approach results. 47
3.27 Zone based presenting for MAAD evaluation for the linear approach results. . . . 47
3.28 Parametric based presenting for AUC evaluation for the linear approach results. . 48
3.29 Zone based presenting for AUC evaluation for the linear approach results. 48
3.30 Parametric based presenting for MAAD evaluation for the equivalent linear results. 49
3.31 Zone based presenting for MAAD evaluation for the equivalent linear results. . . 49
3.32 Parametric based presenting for AUC evaluation for the equivalent linear results. 50
3.33 Zone based presenting for AUC evaluation for the equivalent linear results. . . . 50

4.1 Satellite image for the location of the experimental tests [91]. . . . 53
4.2 Aveiro, Portugal: photographic record of the experimental test site. . . . 54
4.3 Delimitation of the experimental test site on geological map (Plan 16 - A, Aveiro) [92]. . . . 54
4.4 SPT equipment in location of borehole S4 [92]. . . . 55
4.5 Schematic setup for the low strain experiment. . . . 56
4.6 Photographic record of low strain test. . . . 57
(a) Vertical and horizontal accelerometers . . . 57
(b) Data receiver . . . 57
4.7 Experimental setup used for low-strain test. . . . 57
4.8 ICP ® impact force: (a) time domain; (b) frequency domain. . . . 58
(a) . . . 58
(b) . . . 58
4.9 Low strain test - Vertical velocity at 8 m: (a) time domain; (b) frequency domain. 59
(a) . . . 59
(b) . . . 59
4.10 Low strain test - Vertical transfer functions at 8 m. . . . 59
4.11 Low strain vertical coherence: a) 2 m; b) 8 m. . . . 60
(a) . . . 60
(b) . . . 60
4.12 Ricker impulse function. . . . 61
(a) Ricker Impulse Time history . . . 61
(b) Ricker Impulse Frequency content . . . 61
4.13 Ricker impulse: particle peak velocities PPV vs. distance from the source. . . . 61
4.14 Different shear waves velocity profiles. . . . 63
4.15 FEM–PML mesh: an axisymmetric approach for pile-ground system. . . . 63
4.16 Comparison between experimental vs. numerical for the studied parameters. . . . 64
4.17 Damping influence on the PPV vs. the distance. . . . 65
4.18 PPV numerical vs. experimental approach. . . . 66
4.19 PPV comparison between the experimental and numerical results- low strain test. 67
(a) Time domain at distance 2 m. . . . 67

(b) Frequency domain at distance 2 m. . . . 67
(c) Time domain at distance 12 m. . . . 67
(d) Frequency domain at distance 12 m. . . . 67
(e) Time domain at distance 32 m. . . . 67
(f) Frequency domain at distance 32 m. . . . 67
4.20 Experimental test: a) Pile driving machine, Junttan PMx25 equipped with the hydraulic impact hammer SHK110-7; b) and c) unidirectional accelerometers placed along the ground surface. . . . 68
4.21 Vibration induced by pile driving activities in different penetration depth and distance from excitation source. . . . 69
(a) Penetration depth 1-2 ms at 2 m. . . . 69
(b) Penetration depth 1-2 ms at 20 m. . . . 69
(c) Penetration depth 3-4 ms at 2 m. . . . 69
(d) Penetration depth 3-4 ms at 20 m. . . . 69
(e) Penetration depth 5-6 ms at 2 m. . . . 69
(f) Penetration depth 5-6 ms at 20 m. . . . 69
4.22 Hammering respond accelerations at distance 8 meter and depth 3-4. . . . 70
4.23 Hammering respond velocities at distance 8 meter and depth 3-4. . . . 70
4.24 Pile driving experiment results: PPV vs. the distance. . . . 71
4.25 Experimental vertical vibration velocities measured at different distances from pile and considering the penetration depth in the interval 5-6 m: a) time-history; b) frequency content (r=8, 16 and 24 meters). . . . 71
(a) . . . 71
(b) . . . 71
4.26 Frequency content of the experimental vertical vibration velocities considering different distance from pile: a) r=8 m; b) r=16 m. . . . 72
(a) . . . 72
(b) . . . 72
4.27 FE–PML mesh adopted to model the pile-ground system. . . . 72
4.28 Experimental vs. numerical results at penetration depth 4-5 m. . . . 73
4.29 Time domain comparison for a point located at: a) 2 m; b) 4 m; c) 12 m; d) 32 m. 74
(a) . . . 74
(b) . . . 74
(c) . . . 74
(d) . . . 74
4.30 Frequency domain comparison for a points at: a) 2 m; b) 4 m. . . . 74
(a) . . . 74
(b) . . . 74
4.31 Frequency domain comparison for a points at: a) 12 m; b) 32 m. . . . 75
(a) . . . 75
(b) . . . 75
4.32 Porto, Portugal: photographic record of the experimental test site [93]. . . . 75
4.33 SPT N-value for INEGI site. . . . 76
4.34 Maximum peak particle velocity versus distance from the pile (circle marker - INEGI test site: Pile 1 - o; Pile 2 - o; Pile 3 - o; Pile 4 - o; cross marker - Aveiro test site: Pile 1 - +; Pile 2 - +; Pile 3 - +; Pile 4 - +) and maximum permitted PPV values for transient vibrations to prevent damage defined in the Portuguese Standard NP2074 (I-sensitive structures; Il-current structures; III-reinforced structures). . . 76

4.35 Frequency content of the experimental vertical vibration velocities considering: a) and c) different distances from pile and for a penetration depth in the interval 5-6 m; d) and b) considering different pile penetration depths and an observation point at 20 m from pile (upper: INEGI test site; down: Aveiro test site). 78
(a) 78
(b) 78
(c) 78
(d) 78
4.36 FEM–PML mesh adopted to model the pile-ground system. 79
4.37 Geometry of the trench solution. 79
4.38 Comparison between experimental and numerical results (before and after the introduction of the trench) - PPV vs distance - for a pile penetration depth of 2 m. . 80
4.39 Norm of the particle velocity at the time step t=75 ms for two different scenarios: a) Case 0; b) Case 2. 81
(a) 81
(b) 81
4.40 Comparison of the soil surface response before and after the construction of the trench in the frequency domain for an observation point located at the ground surface and at a distance from the pile of: a) 8 m; b) 24 m. 81
(a) 81
(b) 81
4.41 Geometry of the heavy mass next to the pile. 82
4.42 Comparison of the soil response before and after the construction of the continuous distribution of a heavy mass resting on the ground surface for an observation point located at the ground surface and at a distance from the pile of 8 m: a) time record; b) frequency content. 83
(a) 83
(b) 83
4.43 Comparison of the PPV vs distance after and before the introduction of a continuous distribution of a heavy mass resting on the ground surface for a pile penetration depth of 2 m (the trench solution results are also presented). 83

5.1 2D-Latin hypercube sampling. 88
5.2 Normalized input parameters of two layer stratified soil learning database. 90
5.3 Normalized input parameters of homogenous soil learning database. 90
5.4 Schematic presentation for FEM-PML numerical model. 91
5.5 Schematic presentation for the modification in FEM-PML numerical model. . . . 92
5.6 The FEM-PML numerical model output, PPV, for the homogenized and two layer stratified soil profile: a) observation point 1; b) observation point 2. 93
(a) 93
(b) 93
5.7 Comparison between the learning data set and subsets in two layer stratified soil scenario: a) pile diameter; b) shear wave velocity. 94
(a) 94
(b) 94
5.8 Human brain neuron cell vs. artificial nerve cell: a) structure of typical human neuron; b) structure of artificial neuron cell. 97
(a) 97
(b) 97

5.9 Artificial Neural Networks and node structure. 98
5.10 Machine learning schemes: a) multiple linear regression; b) regression tree; c) SVM; b) GPR. 101
(a) 101
(b) 101
(c) 101
(d) 101
5.11 Machine learning Tree group [107]. 104
5.12 RMSE for single hidden layer ANN: a) homogeneous soil; b) layered soil. 105
(a) 105
(b) 105
5.13 RMSE for double hidden layer ANN: a) homogeneous soil profile; b) layered soil. 106
(a) 106
(b) 106
5.14 Architecture of ANN model: a) homogeneous soil profile; b) stratified soil. . . . 106
(a) 106
(b) 106
5.15 Database dividing ratio of the ANN: a) homogeneous soil profile; b) layered soil. 107
(a) 107
(b) 107
5.16 Training algorithms for ANN model. 107
5.17 Validation performance of the ANN: a) homogeneous soil profile; b) layered soil. 108
(a) 108
(b) 108
5.18 Prediction vs. numerical values: a) homogenize soil profile; b) stratified soil. . . 108
(a) 108
(b) 108
5.19 Architecture of ANN model for frequency domain: a) homogeneous; b) Layered. 109
(a) 109
(b) 109
5.20 Prediction vs. numerical values for frequency : a) homogeneous; b) layered. . . . 110
(a) 110
(b) 110
5.21 Schematic representation for hyper-parameters tuning during ML process. 111
5.22 ML algorithms improving iteration vs. min. MSE for predicting PPV in homogeneous soil profile: a) regression tree; b) random forest; c) GPR; b) SVM. 112
(a) 112
(b) 112
(c) 112
(d) 112
5.23 ML algorithms improving iteration vs. min. MSE for predicting frequency domain in homogeneous soil profile: a) regression tree; b) random forest. 112
(a) 112
(b) 112
5.24 ML algorithms improving iteration vs. min. MSE for predicting PPV in stratified soil profile: a) regression tree; b) random forest; c) GPR; b) SVM. 113
(a) 113
(b) 113

(c) . . . 113
(d) . . . 113
5.25 Performance of ML models for predicting PPV in homogeneous soil profile: a) multi linear regression; b) random forest; c) GPR; d) SVM; e) regression tree; f) XGBooster. . . . 115
(a) . . . 115
(b) . . . 115
(c) . . . 115
(d) . . . 115
(e) . . . 115
(f) . . . 115
5.26 Performance of ML models for predicting PPV in stratified soil profile: a) multi linear regression; b) random forest; c) GPR; d) SVM; e) regression tree; f) XGBooster. . . . 116
(a) . . . 116
(b) . . . 116
(c) . . . 116
(d) . . . 116
(e) . . . 116
(f) . . . 116
5.27 Performance of ML models for predicting frequency domain in homogenise soil profile: a) multi linear regression; b) random forest; c) regression tree; d) XGBooster. 117
(a) . . . 117
(b) . . . 117
(c) . . . 117
(d) . . . 117
5.28 Performance of ML models for predicting frequency domain in stratified soil profile: a) multi linear regression; b) random forest; c) regression tree; d) XGBooster. 118
(a) . . . 118
(b) . . . 118
(c) . . . 118
(d) . . . 118
5.29 Prediction vs. numerical values for PPV models : a) ANNS; b) XGBooster. . . . 120
(a) . . . 120
(b) . . . 120
5.30 Prediction vs. numerical values for frequency models : a) ANNs; b) XGBoost. . . . 121
(a) . . . 121
(b) . . . 121
5.31 Experimental validation for ANNs and XGBoost PPV model. . . . 122
5.32 Experimental validation for ANNs and XGBoost frequency model at 8 m. . . . 123
5.33 Schematic presentation for driven pile while using open trench. . . . 124
5.34 Predicate vs. numerical values for a soil profile with open-trench vibration mitigation measure: a) homogeneous; b) stratified profile. . . . 125
(a) . . . 125
(b) . . . 125
5.35 ZBZBAT computer software interface. . . . 127
5.36 ZBZBAT website. . . . 127

List of Tables

3.1 Table of geomechanical properties of the assumed profile. 41
3.2 Pile material mechanical properties. 44

4.1 Range of layers thickness in parametric study. 62
4.2 Soil properties from the first phase. 64
4.3 Experimental site soil properties. 65
4.4 Elastodynamic properties of the pile-ground system. 78

5.1 The parameters influence through the space domain on the final response. 87
5.2 The values for the learning database input parameters. 89
5.3 Case scenarios of the dividing ratio for the database of ANN. 106
5.4 Optimised values of hyperparameters for regression models applied. 114
5.5 Comparison of the predictive performance of ML algorithms. 114
5.6 Table comparison between ANNs and XGBooster. 117
5.7 The values for the learning database input parameters for three layer soil profile. . 119
5.8 Table comparison between ANNs and XGBooster for three layer soil profile. . . 120

Chapter 1

Introduction

1.1 Context

In the twenty-one century, the world holds nearly eight billion people and with a growth rate of 80 million annually [1]. With this growth, the concept of big cities has become more demanding. The construction engineering needs to develops and to implements new techniques to meet this demand. The deep foundation technique has the leading in overcoming the problem motivated by weak geotechnical profiles. It can be categorized based on installing techniques. The deep foundation installed in the soil by driving equipment is known as the driven pile. During the last decades, the ground-borne vibrations induced by pile driving has attracted the research community for its significance.

The vibration induced by the pile driving process is a significant phenomenon that needs to be analysed. During the pile driving operations, the vibration levels are generally high, which can induce, in extreme cases, damage to nearby buildings [2], as depicted in Figure 1.1. The prediction of these vibrations is complex and depends on several factors, justifying the use of empirical models as initial step [2, 3]. The empirical models were developed based on a limited number of case studies, justifying the possibility of having a lean prediction for non-similar cases.

Later, researchers presented significant contributions in predicting the ground-borne vibrations induced by pile driving using numerical approaches. Despite the numerical model's accuracy for predicting the vibration levels suffering from complexity, a computationally costly and time-consuming. The author aims, from this work, to present the steps for developing an expedite prediction tool (surrogate model) for the vibration induced by pile driving to overcome the surrounding drawbacks of the previous trials without neglecting to the significance of the accuracy.

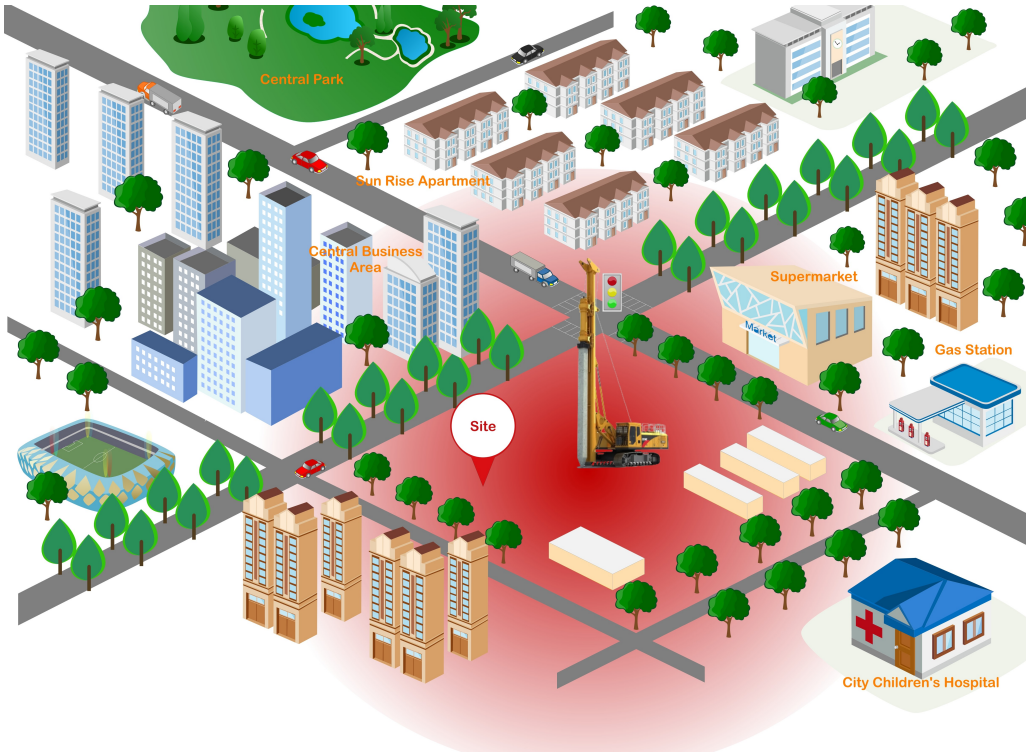


Figure 1.1: Ground-borne vibrations induced by pile driving: influence zone.

1.2 Aim and objectives

This work outlines the steps toward developing an alternative model that aids in predicting vibrations induced by the pile driving process. The surrogate model will merge between a developed numerical model and artificial intelligence techniques. The numerical model used for developing this surrogate model was presented by Colaço et al. [4]. For that purpose initially set, the main objectives of this work can be introduced through the following points:

- *Phenomenological analysis of the problem*: general overview for the main physics of generating and propagation of the ground-borne vibration induced by pile driving.
- *Development and exploration of an experimental test site*: an experimental test site was implemented in Aveiro, Portugal. The experimental work compressed the measurement of ground vibrations induced by pile driving.
- *Development of an expedite prediction model*: development of an expedite prediction model based on machine learning techniques. This model can predict the vibration in spatial and frequency domain in almost zero-time consumption.
- *Development a user friendly interface*: develop of a user friendly interface for the developed model in order to predict the ground borne vibrations induced by pile driving.

1.3 Work description

In Figure 1.2, the work breakdown structure is schematically presented, clarifying the dependency activities. While the following presents, in a summary form, the work accomplished under each chapter.

- *General overview*: The literature review will focus on the general principles related to the generation and propagation of vibrations induced by pile driving; the prediction and measurement of these vibrations; the role of machine learning techniques in the prediction process.

- *Numerical model*: in chapter 3, the numerical model previously developed by Colaço et al. [4] is used for deep understanding for the generation and propagation of the ground-borne vibration induced by pile driving. Through the chapter, a parametric study is presented. The numerical model introduced in this chapter is used to performed the database of the machine learning models.

- *Experimental test site*; in chapter 4, an experimental test site was explored. The chapter begin with general characteristic of the test site, where a low strain test was performed to determine the ground dynamic properties. The results from the pile driving were subjected to a numerical assessment which shows reasonable matching between the experiment and numerical approaches.

- *Surrogate model development*; over chapter 5 a large database will be created through massive numerical simulations that aim to represent the most typical situations that can be faced in this type of problems. The extensive database was used to train machine learning models for the rapid prediction of vibrations caused by the pile driving process. MATLAB® [5] and Python are used to train the model with different machine learning algorithms. During the implementation, the work flow will be in iterative process, starting from model training to validation and testing the model. Furthermore, a comparison study and experimental validation was presented. The chapter finish with presentation of stand-alone computer software with user friendly interface for predicting the ground-borne vibration induced by pile driving process.

- *Considerations and future development*: In chapter 6, general consideration that can be extracted from this works are presented, in addition to presenting the open opportunities for future developments.

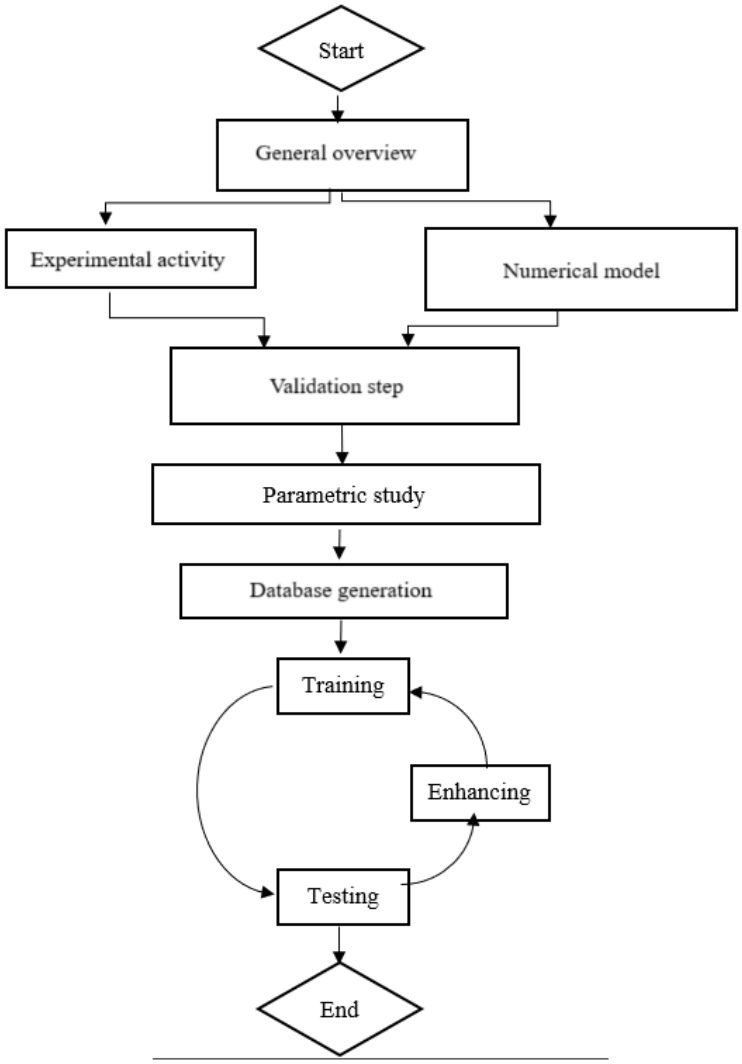


Figure 1.2: Work-breakdown structure.

Chapter 2

Vibration induced by pile driving: a general overview

2.1 Generalities

The building construction activity correlates with human civilization during history, which placed it under continuous development and enhancement to achieve humanity's growing aspirations. The vibrations induced from the construction activities are considered a critical aspect, mainly when the vibration threatens the nearby structures' integrity and stability (see Figure 2.1). The researchers explore during the last decades their practices for profoundly understanding this phenomenon. The practices involve experiments and analyses for the phenomenon under investiga-

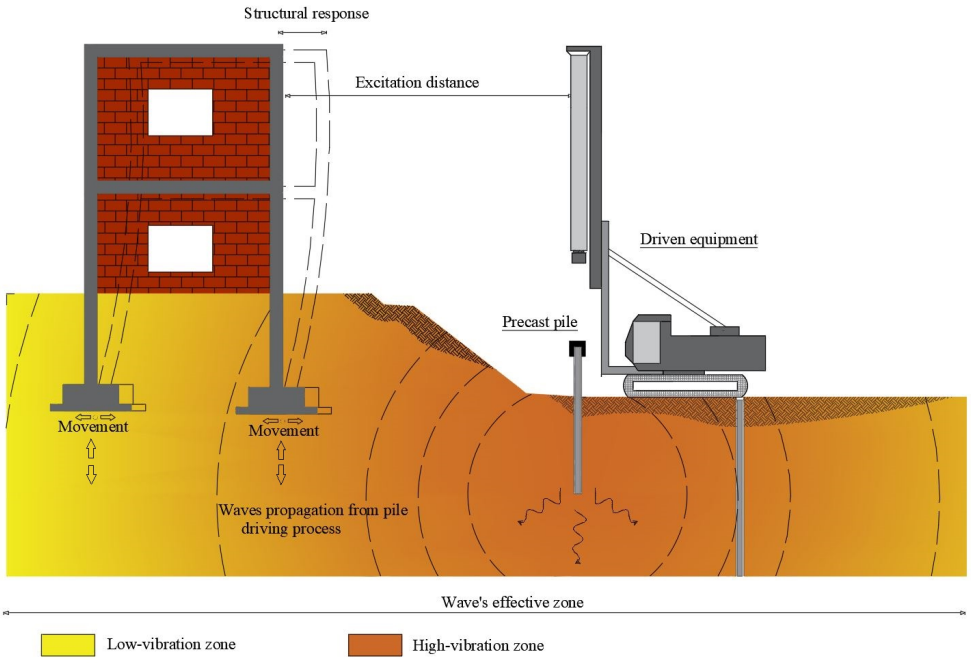


Figure 2.1: Illustration for driving process influence on the surroundings.

tion, which generates different empirical and numerical solutions for prediction of the vibrations induced by the pile driving process.

In order to explore the previous contributions properly in investigating the vibrations induced by the pile driving process, some aspects should be collected and synthesized. The considered aspects include pile driving technology, wave propagation and attenuation, empirical and numerical prediction tools and key indicators, machine learning in predicting the ground-borne vibration, and finally, the international standards for vibration limitation. It should be clear that these aspects will be clarified and evaluated through this chapter.

2.2 Pile driving technology

The pile is a slender structural element, used to overcome the problem motivated by weak geotechnical profiles. The function of the pile is transferring the loads that come from the structure to deep ground layers, involving interaction with the ground through the shaft and the base. Matos Fernandes [6] categorized the piles based on the installation mechanism as schematically illustrate in Figure 2.2.

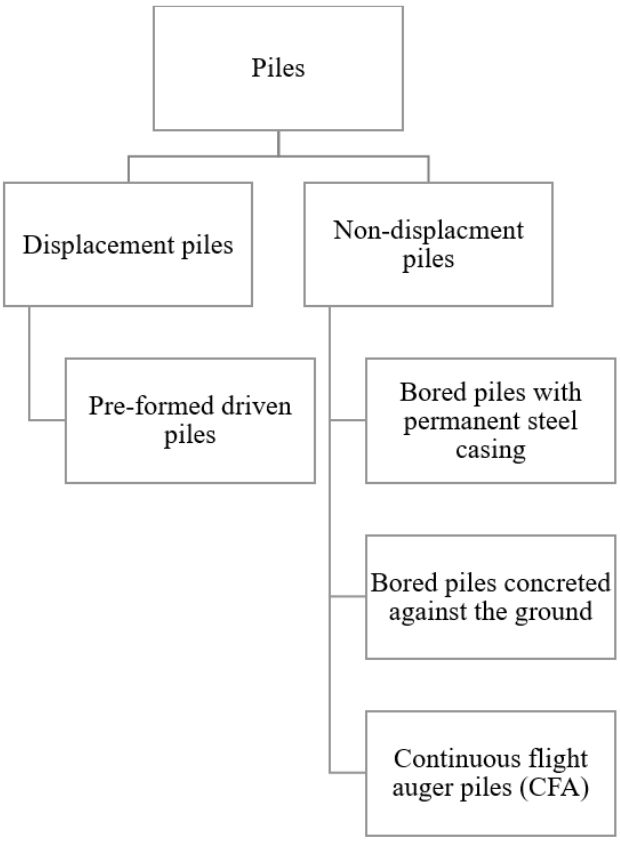


Figure 2.2: Pile classification and construction technique.

Choosing the most proper type of pile should be based on previous knowledge for the advantages and disadvantages for each type. Under this context, Figure 2.3 present the main advantage and drawbacks of each one pile construction technique [6].

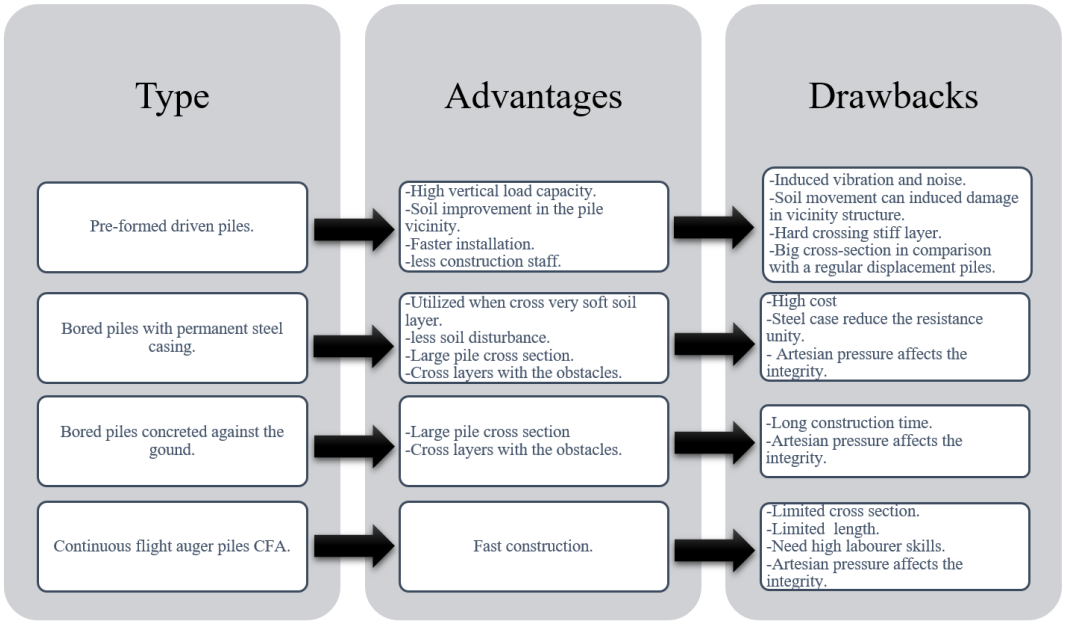


Figure 2.3: Piles types advantages and drawbacks.

For a focused and organized exploration, it is essential to recall that this work concerned with vibration induced by the driven process for precast piles, categorized as pref-formed driven piles, justifying exposing the interesting area elaborately. The detailed history of driven pile foundations could not be tracked along the timeline.

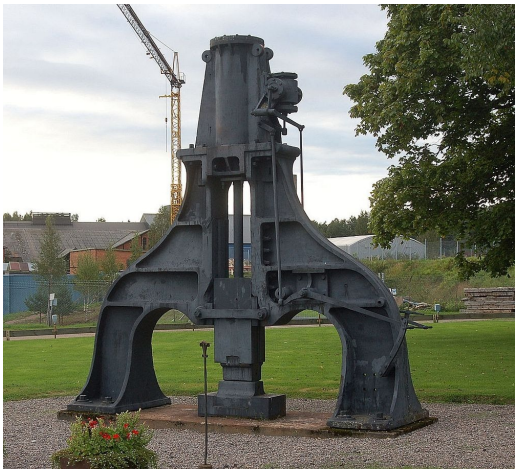


Figure 2.4: The Scottish steam hammer.

However, it is important to point out that Naysmith, in the 1830s, developed the first pile-driving machine, the Scottish steam hammer, using steam pressure (see Figure 2.4). The steamed pile driving hammer was used in the USA between the middle and late 1800s. Concrete-driven piles debuted, as it is now known, in Europe in the 1890s and USA in the 1900s [7]. As the concrete piles became the domain, the Federal Republic of Germany developed pile-hammer diesel power in the 1920s. In the 1940s, the Soviet Union introduced the first vibratory hammer, which became demanding in the USA in the 1960s. Later, primary hydraulic pile hammers were developed in Scandinavia in the 1960s. Throughout the lifetime of the driving hammer, with different types, it was subjected to development to mitigate the impact-side of the driving process on the vicinity and improve the driving efficiency [8]. Generally, three main types of hammer will be present:

I) Diesel hammers are widely used in the foundation industry. The concept of the diesel hammer is similar to a hammer used to drive a nail. The pistons lift the ram to a certain height. When the ram reaches the required height, the operator makes a controlled falling for the ram on the pile head. The compressive stress impacts the pile to make it go deep in the soil. Recently, modern hammers have increased compression pressure (see Figure 2.5.a).

II) Hydraulic hammers are a contemporary form of piling hammers used rather than diesel hammers for driving steel pipe, precast concrete, and timber piles. The advantage is a reduction of machine wear and improved operator comfort for improving reliability and productivity (see Figure 2.5.c).

III) A vibratory hammer is used to drive piles into the soil by vertical vibrations. The soil particles are put into motion so the soil is 'loosened'. The dynamic weight of the hammer can drive the elements into the soil. Vibrating technology is simpler and it also reduces operating time and cost. Recent vibrating hammers are used for piling and extracting of steel sections such sheet pile walls, H-cross section pile, and pipes as well as, cast-in-place piles (see Figure 2.5.b).



(a) (b) (c)

Figure 2.5: Pile hammering machines: a) diesel hammer, b) vibratory hammer, c) hydraulic hammer.

In the last decades, piles with a bearing capacity of 70 tons have been commonly used. Nowadays, 200 tons is more widely used and this number is subjected to increase in following years. Such increment in pile capacity is achived by increasing the cross-section dimension demanding the use of bigger hammers, justifying the possibility of increasing the vibrations levels. The wave propagation should be understood to mitigate any side effects on the vicinity [9].

2.3 Generation and propagation of seismic waves

Vibration is a oscillatory movement of the particles of a body or a medium, commonly resulting from the interaction between the exciting force on the body and the equilibrium resistance. In the present context, the propagating medium of the waves is the ground, being the vibrations generated by the pile driving process. From this context, and to meet the purpose initially set, the concepts of generation, propagation and attenuation of vibrations will be presented in the following.

2.3.1 Soil behavior under driving process

During the driving process of the pile, the soil is subjected to a dynamic strain, where the ground can be divided for three zones. The elastic zone where the shear strain level is below $10^{-3}\%$, with no permanent deformations expected. The elastoplastic zone, where the strain level ranges between $10^{-3}\%$ and $10^{-1}\%$, where some permanent deformations will occur. The plastic zone, where the soil is in a failure condition and is subjected to large strain levels $>10^{-1}\%$ [10, 11]. Figure 2.6 presents these three zones schematically.

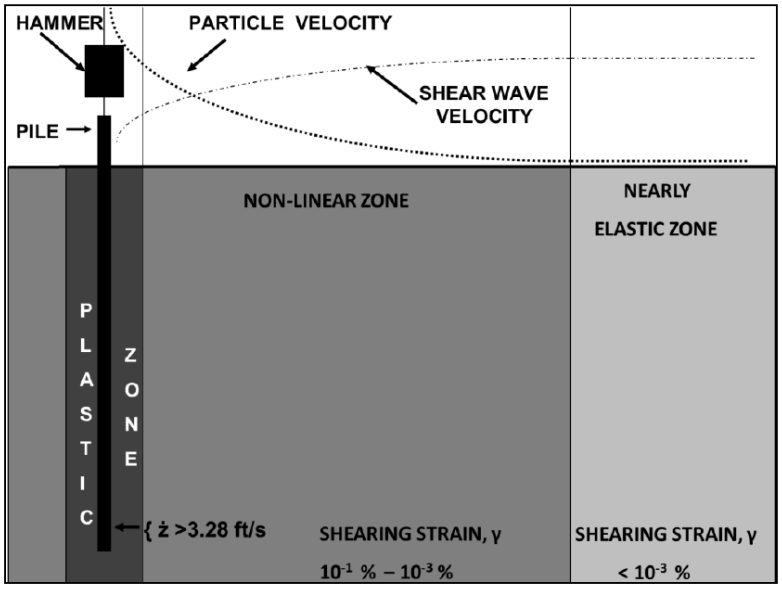


Figure 2.6: Hypothetical soil behavior zones near a driven pile [12].

2.3.2 Waves propagation in soil

During the driving process, the penetration of the pile happens when the hammering force is transmitted to the pile head to overcome the resistance of the ground. While the pile penetrates into the soil, the first waves start propagating induced by the transmitted force from the pile to the soil, known as body waves. The body waves are followed by surface waves on the ground [13]. The body waves consisted of compressional (P) and shear (S) waves, while the surface waves consisted of Rayleigh (R) and The Love (L) waves. The waves are categorized based on the corresponding periodic movement, as illustrated in Figure 2.7.

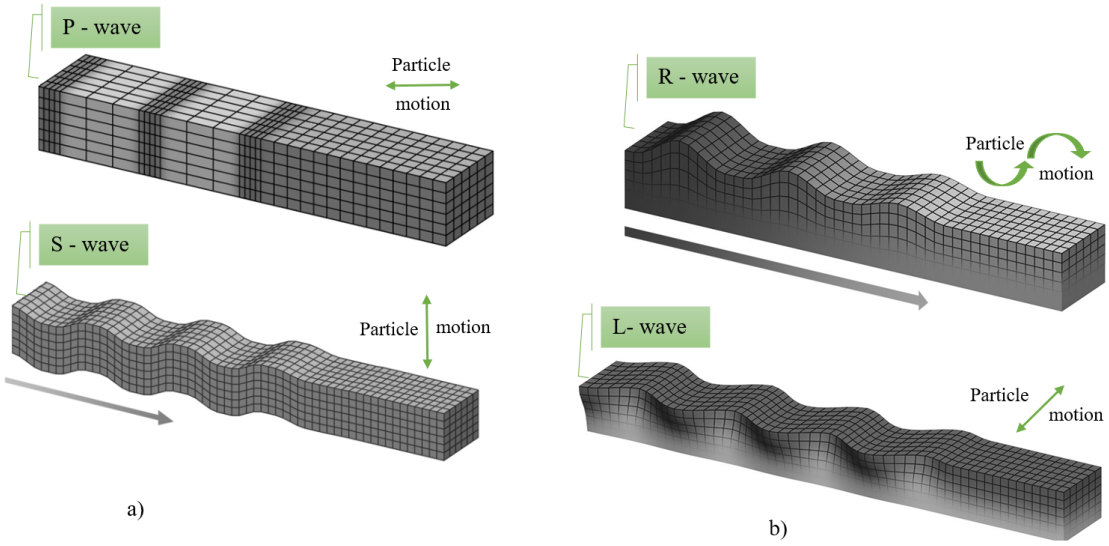


Figure 2.7: Waves propagation: a) body waves, b) surface waves.

2.3.2.1 Compressional waves (P-waves)

also known as longitudinal or primary waves

P-waves induce the particles of the soil moving backward and forward, causing compression and expansion actions, as shown in Figure 2.7a. P-waves do not cause shearing or rotating for the material particles as it passes through. P-wave propagation velocity is affected by the degree of soil saturation as the water and soil do not expand or compressive in the same way. Since P-waves can travel through both solids and fluids, the velocity of P-wave in a saturated soil does not represent the velocity in the soil alone but also the water inside the soil [14]. The P-waves are the fastest wave present in the solid material, and the velocity of P-waves, C_p, can be calculated from the following formula:

$$C_P = \sqrt{\frac{\mu + 2G}{\rho}} \tag{2.1}$$

as,

$$G = \frac{E}{2(1+\nu)} \tag{2.2}$$

$$\mu = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{2.3}$$

where ν is the Poisson's ratio, ρ is the soil density, μ is Lamé's constant, G is the shear modulus, and E is the soil Young's modulus.

2.3.2.2 Shear waves (S-waves)

also known as transverse, distortional or secondary waves

The S-wave induces the vibration of particles perpendicular to the direction of the wave propagation, as shown in Figure 2.7a. In any single plane, S-waves are often divided into two perpendicular components, a vertical component (S_v), and a horizontal component (S_h). The polarized shear wave, C_s, is given by the following equation [14]:

$$C_s = \sqrt{\frac{G}{\rho}} \tag{2.4}$$

2.3.2.3 Rayleigh waves (R-waves)

also known as in-plane surface waves

Rayleigh waves, R-waves, are the most common type of surface wave. R-waves are a combination of refracted and reflected P and S waves. Figure 2.8 shows the Rayleigh wave's horizontal and vertical amplitude as a function of depth, d, Poisson's ratio, υ, and wavelength, λ. From Figure 2.8, it is noticed that the vertical amplitude is greater than the horizontal amplitude, as well the vertical amplitude decreases rapidly with depth. The propagation velocity of a Rayleigh wave is presented through the following formula [15, 16, 17, 18]:

$$C_r \cong 0.9194\, C_s \tag{2.5}$$

2.3.2.4 Love waves (L-waves)

Love waves, L-waves, are the type of surface wave resulting from the interaction of S_H-waves with a soft surface layer, as shown in Figure 2.7b. L-waves are horizontally polarised shear waves with no vertical component [19]. L-waves only exist when a low-velocity layer overlays a higher velocity layer. In a homogenous half-space, L-waves are not produced [20].

Essentially, Love waves consist of S_H-waves reflected within the surface layer. The displacement amplitude of the L-wave varies sinusoidally with depth with exponentially decays [19].

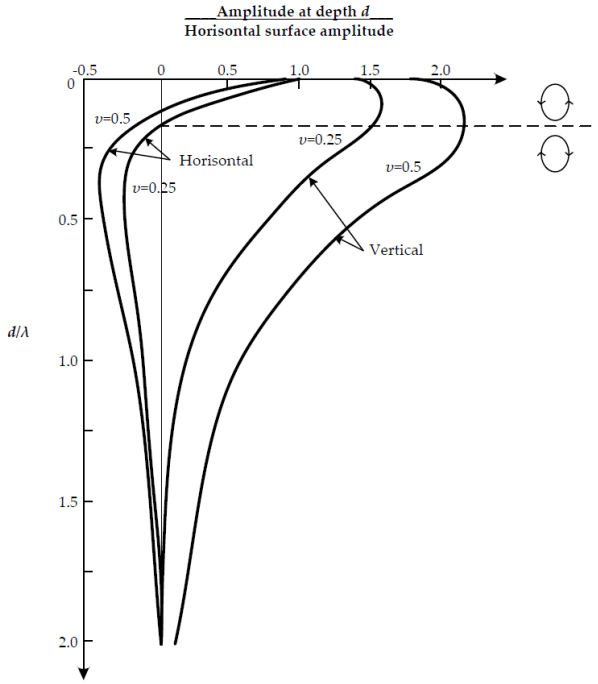


Figure 2.8: Horizontal and vertical vibration amplitude of the Rayleigh wave [17].

The propagation velocity of Love waves is between the S-wave velocity of the upper (softer) and lower ground (stiffer)l layers. The velocity of the Love wave varies with the frequency between an upper and lower limit; hence they are dispersive [19]. The wave propagation velocity for Love waves depends on the frequency.

2.3.3 Energy transmission during impact pile driving

At the tip of the pile, each impact causes a volume displacement, which results in both primary and shear waves traveling outward from the cavity (see Figure 2.9). Both P and S waves travel outward from the tip of the pile on spherical wave fronts. The P wave travels faster than the S wave, so its wave front precedes the shear wave at any given point in the ground. When the P and S waves encounter the surface of the ground, part of their energy is converted to surface waves (Rayleigh waves, R-waves) and part is reflected back into the ground as reflected P- and S waves. These waves that transmit energy to the ground surrounding a pile are potentially damaging to neighboring structures and annoying to people.

The principles of elastic wave propagation near the ground surface provide some insight into the expectations for the relative amplitude of vibrations from pile driving. Rayleigh waves decay more slowly along the ground surface than other seismic waves and consequently transmit damaging energy over greater distances than body waves (see section 2.3.4.1). As the pile is driven deeper, the mechanisms of skin friction and end bearing dominate wave generation and body waves are produced, which decay more rapidly with distance. In addation, preboring is used to minimize

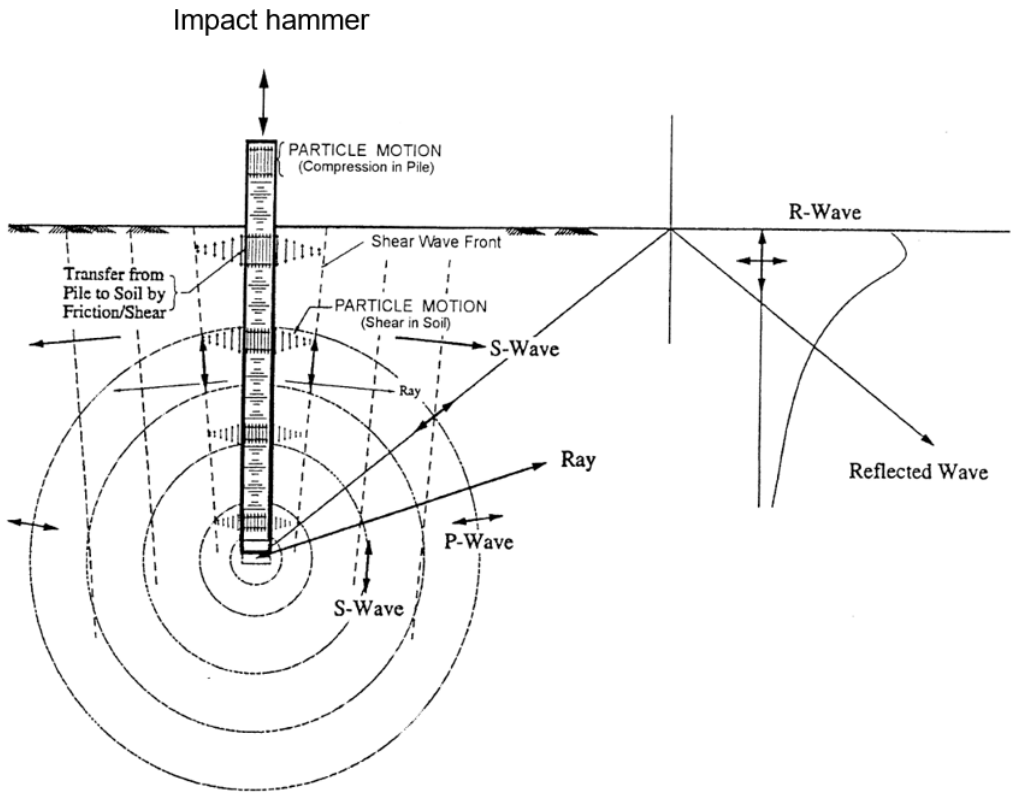


Figure 2.9: Generation of seismic waves during impact driving in homegeneous soil [21].

the transmission of wave energy near the ground surface. However, if the target is underground, such as a pipeline or tunnel, Rayleigh waves generated near the surface will not cause damage, but body waves generated by the pile tip when it is deeper may be causing damages [21].

2.3.4 Wave attenuation

Wave attenuation is defined as the reduction in the wave amplitude during propagation due to geometric and material factors. Geometrical damping derives from the expansion of the wave-fronts, resulting in the spreading of energy over an increasing area. Material damping is a wave propagation resistance due to the soil's dissipation of wave energy [22]. Following both types will be present.

2.3.4.1 Geometrical attenuation

When an impulse is apply at some point on the ground, body waves would be generated and travel into the medium with a hemispherical geometry [14]. When body waves spread out around a

hemispherical wave, the energy is distributed over an area under the following relationship:

$$Energy \propto \frac{1}{r^2} \tag{2.6}$$

where, r is the radius. However, the amplitude is proportional to the square root of the energy per unit area:

$$Amplitude \propto \sqrt{Energy} \propto \frac{1}{r} \tag{2.7}$$

or,

$$Amplitude \propto \frac{1}{r} \tag{2.8}$$

The amplitude of the body waves is proportional to $1/r^2$. Similarly, the amplitude of the Rayleigh waves, which spread out in cylindrical wave-front, is proportional to $1/\sqrt{r}$. Thus the attenuation of the amplitude of the Rayleigh waves is slower than that for the body waves [14].

From the theory of energy conservation, the wave attenuation due to geometric damping can be described with the following expression [21]:

$$A_2 = A_1 \left(\frac{r_1}{r_2}\right)^n \tag{2.9}$$

where A_2 = amplitude of motion at distance r_2 from the source (m) A_1 = amplitude of motion at distance r_1 from the source (m), $n = 1/2$ for R-waves, equal 1 for body waves, and 2 for body waves at the surface.

2.3.4.2 Material attenuation

Soil material dissipated part of the energy of the waves pass through it. The following equation consider the attenuation of surface waves in terms of material attenuation [23]:

$$v = v_1 \sqrt{\frac{d_1}{d}}\ e^{\alpha(d-d_1)} \tag{2.10}$$

where v is the velocity at a distance d measured along the ground surface from the source, and v_1 is the velocity at a reference distance d_1. The value of α is dependent upon the properties of the soil wave attenuation and the following relationship is suggested [24]:

$$\alpha = \frac{2\pi D f}{c} \tag{2.11}$$

where α is the material damping coefficien, D is the material damping (%), f is the vibration frequency, and c is the wave propagation velocity.

2.4 Prediction and measurement of vibrations induced by pile driving

2.4.1 Vibration key indicators

Vibrations is a common physics phenomenon that can be assessed using different approaches and corresponding features. So, it is essential to find out which indicator is the best option to investigate the ground vibrations induced by the pile driving process.

The ground vibrations induced by different sources have been evaluated mostly in terms of peak particle displacement (PPD), peak particle velocity (PPV), and peak particle acceleration (PPA) by different researchers over the last decades [25, 26, 20, 27, 28, 29, 30, 31, 32]. PPD, PPA, and PPV are the maximum displacement, acceleration, and velocity that soil particles experience during a vibration event, respectively. However, several contributions present a better consistency of PPV to the initiation of damage in the structures than PPD and PPA [33, 34, 35, 36, 37], justifying introducing PPV as the most appropriate descriptor for evaluating the ground vibrations. In addition, it should be emphasized that the standards mostly address the ground vibrations based on PPV.

In order to obtain the value of PPV, the vibration generated for a given distance the source is recorded simultaneously in three orthogonal directions of X, Y, and Z as the velocity components (V_R, V_V, and V_T). Indeed, V_R, V_V, and V_T are known as radial (or longitudinal), vertical, and tangential (or transverse) velocity components, and V_R is in the direction of the source. After identifying the three physical quantities, the PPV can be determinant from the following approaches:

1. Peak vector sum (PVS): the most popular approach for PPV determination in scientific literature is the maximum resultant of velocity components that a soil particle experiences at certain time during construction work, as shown in the following:

$$\text{PPV} = \max \sqrt{(V_R(t))^2 + (V_V(t))^2 + (V_T(t))^2} \tag{2.12}$$

   Standards such as Portuguese Standard NP 2074 applied this approach for considering the PPV value.

2. The square root of squares sum (SRSS): PPV is determined in this approach from maximum velocity components from considering the occurrence time, as shown in the following:

$$\text{PPV} = \sqrt{(V_{\text{Rmax}})^2 + (V_{\text{Vmax}})^2 + (V_{\text{Tmax}})^2} \tag{2.13}$$

   From the previous contributions, the simultaneous occurrence of maximum velocity components for a soil particle during construction is far from reality [2], and the calculated PPV would be conservative. Regardless, several researchers have applied this approach to their investigations [38, 39, 40].

3. Peak unidirectional velocity: in this approach, PPV equals the highest recorded value (V_R, V_V, and V_T) during the entire construction time, as shown in the following:

$$\text{PPV} = |V_{V_{max}}, V_{R_{max}}, V_{T_{max}}|_{max} \tag{2.14}$$

   Several standards have used this approach [BS 7385-2:1993 (BSI 1993), DIN 4150-3 (DIN 1999)] and Zhang et al. [41].

4. Peak vertical velocity: the maximum vertical component of velocity obtained from measurements, as shown in the following:

$$\text{PPV} = V_{\text{Vmax}} \tag{2.15}$$

   In this regard, this approach has been widely applied by several researchers [42, 43, 34, 20, 27, 44, 45, 46, 47], as well as standards such as Swiss Standard SN640312 (SNV 1992), and Swedish Standard SS0252 11 (SIS 1999). Hamidi et al. [2] justified the widely using of this approach by several factors such, the better correlation with damage inception, easier to measure or compute and close results to the component resultant by Peak vector sum (PVS).

The previous exposition presents four common approaches proposed for calculating the PPV. The author should emphasize that the fourth approach for considering the PPV value will be used during the present work for vibration assessment. However, it should be pointed out that the PPV does not describe the vibration phenomenon completely. Actually, the frequency content is also considered as a relevant parameter. Therefore, the international standards introduce the PPV side by side with excitation frequency as the most appropriate description for the vibration phenomenon.

2.4.2 Empirical methods

The empirical vibration prediction models were the first trials for the investigators to develop prediction tools for the vibration induced by pile driving. The empirical approaches are based on establishing a set of experimental tests. The results are placed under analysis and studying to find a formula representing the data. As the vibrations are a complex phenomenon, it merges different parameters. The main drawback of the empirical approach is typically the limited number of case studies, which raises concerns about accurately representing the parameters. However, when applied in similar conditions to those that originated them, the simplicity of the application makes them valuable tools for a scope analysis, allowing a quick assessment of the maximum vibration levels.

Wiss in 1967 [48] and Attwell and Farmer in 1973 [42] were the first authors to provide an empirical model for an expeditious estimation of vibration levels due to pile driving. The Peak Particle Velocity (PPV) is estimated considering the energy transmitted by the hammer to the pile,

and the distance between the observation point and excitation source, (see Equation 2.16).

$$v = k\left(\frac{\sqrt{W_0}}{r}\right)^x \tag{2.16}$$

where W_0 is the hammer energy and r is the horizontal distance to the pile. k and x stand for empirical constants. Attwell and Farmer [48] acquired a sufficient and conservative correlation with experimental results. For the sake of enhancement, the investigators proposed different values for k over the last years.

Based on a more extensive database of experimental results, Attwell et al. [49] found that a linear relation in the log-log space is not strictly valid. By using a quadratic regression curve, Attwell [49] suggested an empirical formula for enhancing the representing of the experimental results:

$$\log v = -0.519 + 1.38\log\frac{\sqrt{W_0}}{r} - 0.234\log^2\frac{\sqrt{W_0}}{r} \tag{2.17}$$

In more advanced studies, the authors recommend adopting a one-half standard deviation line for the quadratic regression:

$$\log v = -0.296 + 1.38\log\frac{\sqrt{W_0}}{r} - 0.234\log^2\frac{\sqrt{W_0}}{r} \tag{2.18}$$

The Federal Railroad Administration (FTA) [50] proposed a formula based on the empirical approach to predict the vibration levels caused by general construction activities (see Equation 2.19):

$$v = v_{ref}\left(\frac{r_{ref}}{r}\right)^{1.5} \tag{2.19}$$

where v is the peak particle velocity, r the distance between the source and the receiver, v_{ref} a reference vibration level of the construction equipment and r_{ref} is reference distance of construction equipment. The FTA [50] defines two reference values for vibration levels induced by impact pile driving, The upper range is 38.60 mm/s and 7.60 m for v_{ref} and r_{ref} in respectively, and the typical value for v_{ref} is 16.36 mm/s.

Federal Highway Administration FHWA [51] suggests using the empirical formula presented by Bay [52], which the soil classified between, Class I, Class II and Class III. Class I defines the soil with SPT-N below 5 blows. Soil Class II is addressed with SPT-N values of 5 to 15 blows, and Class III is hard soils with SPT-N values of 15 to 50 blows. The curves associated with these empirical relationships are represented in Figure 2.10 for Class II and III.

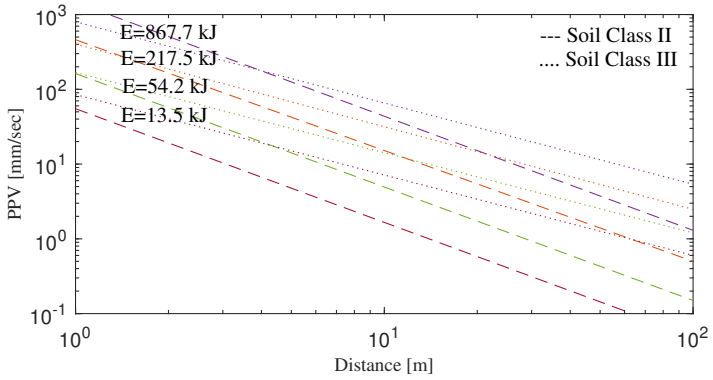


Figure 2.10: Predicted vibration levels for class II and class III.

Colaço et al. [53] presents a comparative analysis of the peak particle velocity levels according to the pile-receiver distance obtained from the previous empirical approaches, and considering rated energy equal to 100 kJ for the empirical models proposed by Attwell and Farmer [48] and Attwell et al. [42], as illustrated in Figure 2.11.

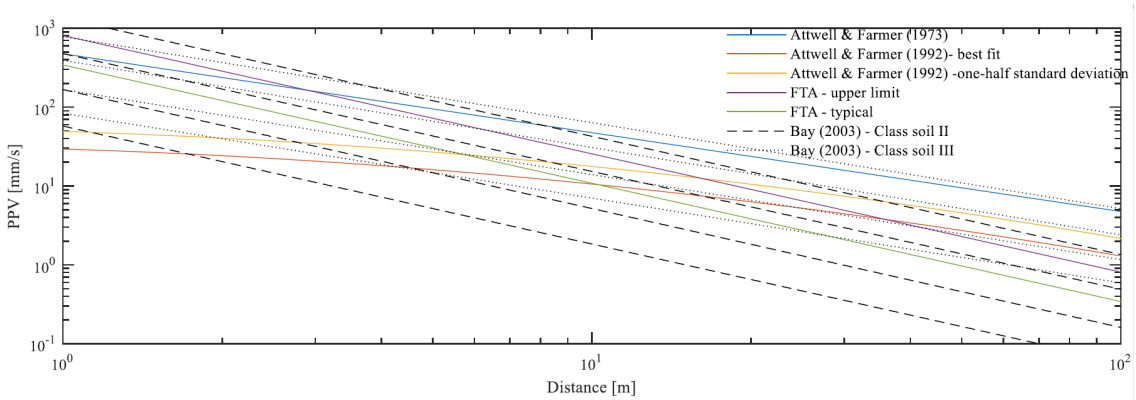


Figure 2.11: Comparative analysis for the empirical approaches Colaço et al. [53].

From the previous presentation, several contributions from the researchers for identifying empirical models to enhance vibration prediction were presented. However, the empirical models presented a high variability, attributed to the phenomenon's complexity and the limiting of the experimental database. These results generate momentum for investigating a numerical approach, looking for predictions enhancement.

2.4.3 Numerical approaches

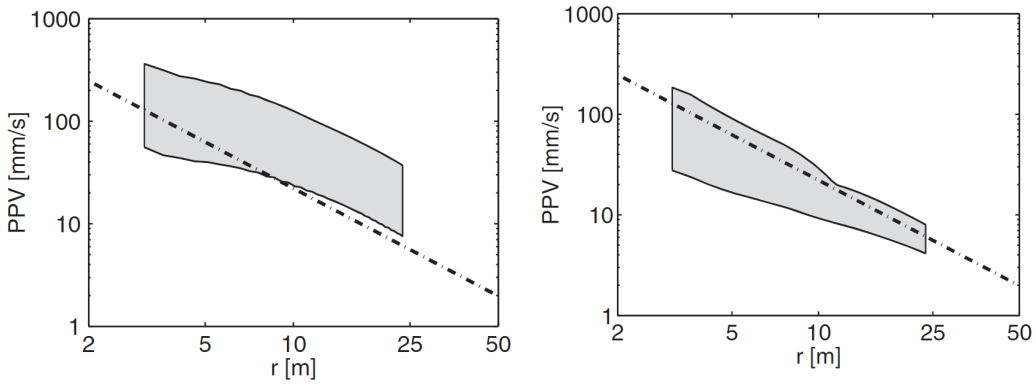
Despite the prevalence of empirical approaches, some numerical methodologies have been proposed in recent years, seeking for a deeper understanding of the physical behavior of the problem

and a more general prediction tool, not restricted by the particular conditions for which the empirical rules were defined. The most relevant numerical modeling should be highlighted in the following.

Madheswaran et al. in 2005 [54] used the finite-element software Plaxis to study ground acceleration time history due to impact pile driving in sandy soil. The pile and the soil were modeled using elastic and elastic-plastic models, respectively. Absorbent boundaries were used at the bottom and side boundaries to avoid Spurious wave reflection. The predicted vertical response was in close agreement with field data, but the predicted radial peak acceleration was more than 20% greater than field data. Later, they used a similar model to study the effect of concrete-filled trenches as a mitigation approach for the excessive peak particle velocity.

Serdarogl et al. [46] in 2010 presented a non-linear model based on finite-element software Abaqus to study ground borne vibrations in cohesive soils under saturated conditions. An artificially damped non-reflecting boundary consisting of several soil layers with different damping ratios was defined at the model's boundary to minimize the reflection of stress waves. However, the model underestimated the measured peak vertical and radial velocities.

Masoumi et al. 2007 [55] introduce a different technique to address the same problem, based on the FEM-BEM (Finite Element Method-Boundary Element Method) coupling. The linear approach overestimates the free field PPVs as observed in reported field experiments during vibratory and impact pile driving at comparable levels of the transferred energy, attributed to the energy dissipation related to plastic deformations in the soil around the pile, (see Figure 2.12a). Therefore, Masoumi et al. 2009 [56] computed the vibration with a non-linear model where the soil is modeled as an isotropic elastic, perfectly plastic solid, which yields according to the Drucker–Prager failure criterion. The result is in much better agreement with experimental records during the driving process, (see Figure 2.12b).

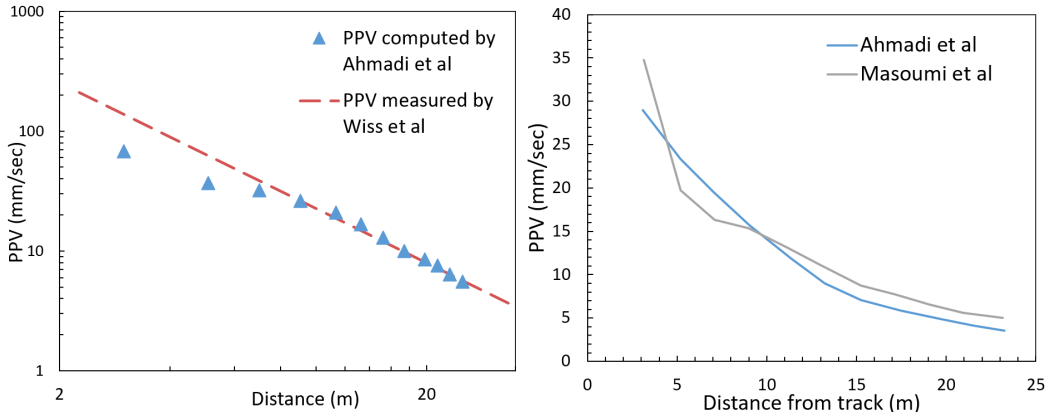


(a) linear analysis (b) nonlinear analysis

Figure 2.12: Envelope of the PPV Vs. the distance from the pile due to impact pile driving with a transferred energy from 3.4 to 19.2 kJ, computed by, a) linear analysis; b) nonlinear analysis, and compared with experimental results(dashed-dotted line) [56].

Khoubani and Ahmadi [57] in 2014 developed an axisymmetric finite-element model that

uses an adaptive meshing algorithm, through the commercial software Abaqus, to simulate full penetration of the pile from the ground surface to the desired depth by applying successive hammer impacts. The model has been verified by comparing the computed particle velocities with those measured in the field and verified numerically with Masoumi et al. 2009 [56], (see Figure 2.13). The results show a good match with the experimental test data at greater distances. However, in the numerical validation, a significant discrepancy is observed between the models at shorter distances from the excitation source.



(a) Experimental verification (b) Numerical verification

Figure 2.13: Khoubani and Ahmadi [57] model validation.

In 2019, Homayoun Rooz and Hamidi [58] employed an axisymmetric FEM model with artificial boundaries, incorporating a gradual increase in soil damping to prevent incident waves from reflecting back into the analysis domain. Subsequently, Sofiste et al. [59] adopted a similar approach for handling artificial boundaries in 2020.

Sofiste et al. [60] in 2021 presented a numerical model developed in the time domain, considering a finite element axisymmetric formulation and an adequate time marching procedure. The numerical results present agreement with the numerical model proposed by Masoumi et al. [56], which indicates the effectiveness of the developed numerical model, (see Figure 2.14).

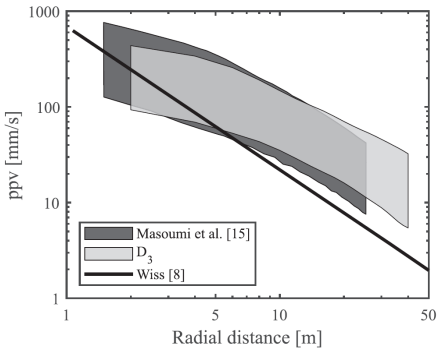


Figure 2.14: Sofiste et al. model: numerical validation [60].

Colaço et al.[4], in 2021, introduced a numerical tool for predicting the vibration level induced by pile driving solved in the frequency domain, and considering the linear and nonlinear behaviour for the soil. The ground's non-linear behavior is a necessary condition for a realistic prediction of vibrations levels induced by pile driving. The model was validated experimentally, and the result present a good agreement between the experimental test and the model based on equivalent linear behaviour, while the linear model present more concerned results, (see Figure 2.15).

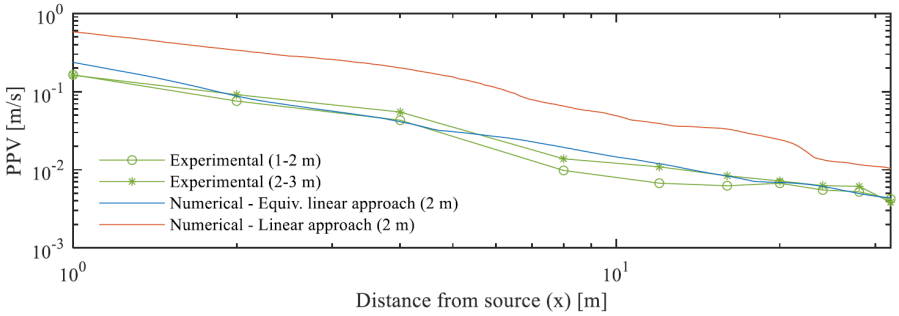


Figure 2.15: Experimental validation for the numerical model Colaço et al. [4].

2.4.4 International standards

Excessive ground-borne vibration can have detrimental effects on the integrity and stability of nearby structures. Due to the varied responses of different structures, it is crucial to establish precise vibration limits. However, setting a Peak Particle Velocity (PPV) limit without considering the frequency of the vibrations raises concerns about the reliability of such limits. There is a well-established link between excitation frequency and structural response [61]. The most important standards and guidelines related to vibration limits are presented below:

- Swedish standard SS 25211 (Vibration and shock – Guidance levels and measuring of vibrations in buildings originating from piling, sheet piling)
- German standard DIN 4150-3 (Vibration in buildings – Part 3: Effects on structures)
- Swiss standard SN 640 312 (Swiss Standard on vibration effects on buildings)
- British standard BS 7385-2 (Evaluation and measurement for vibration in buildings. Part 2: Guide to damage levels from ground-borne vibration)
- Portuguese standard NP 2704 (Evaluation of impulsive vibrations in structures)
- Federal Transit Administration (FTA) – Transit Noise and Vibration Impact Assessment Manual

Figure 2.16 compares three essential standards and guidelines, British standard BS 7385-2, German standard DIN 4150-3, and Portuguese standard NP 2704. The limits of the guidelines

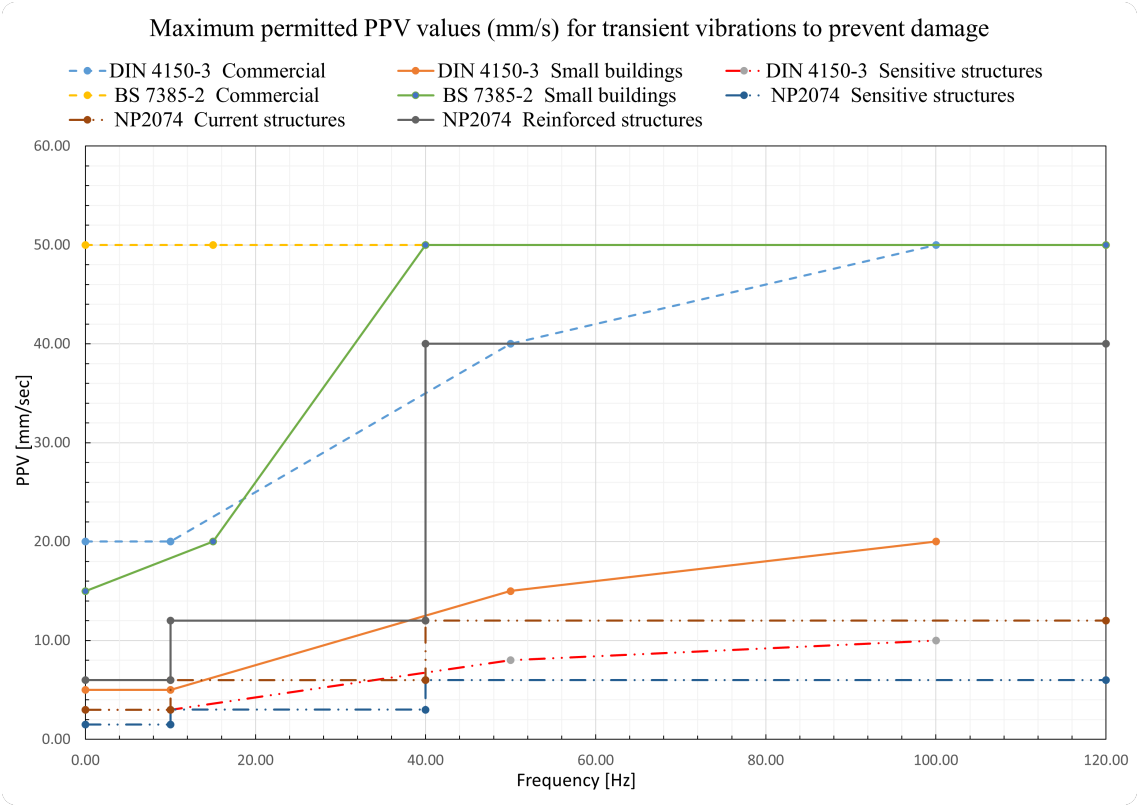


Figure 2.16: Standards and guidelines limits for PPV.

were plotted in the frequency domain. The data trends are more concerned with low-frequency than a high-frequency. It should be mention that different levels of vibrations limits are defined according the structure type.

2.5 Prediction methods based on machine learning techniques

2.5.1 Machine learning techniques

Recently, machine learning applications have become more popular as the term of the big data surfaced. The essentiality of the machine learning application attribute to its ability to discover the data pattern and data prediction. However, the main drawback of the machine learning is lack of transparency. As shown in Figure 2.17, the machine learning can be classified for two categories: supervised and unsupervised techniques.

Every category is classified to a set of subcategories. The subcategories perfectly matching with specific kinds of data type as following:

- ***Supervised machine learning***: The supervised learning algorithm takes the training set and output for training the model to generate reasonable predictions. The supervised learning methods are divided to classification and regression.

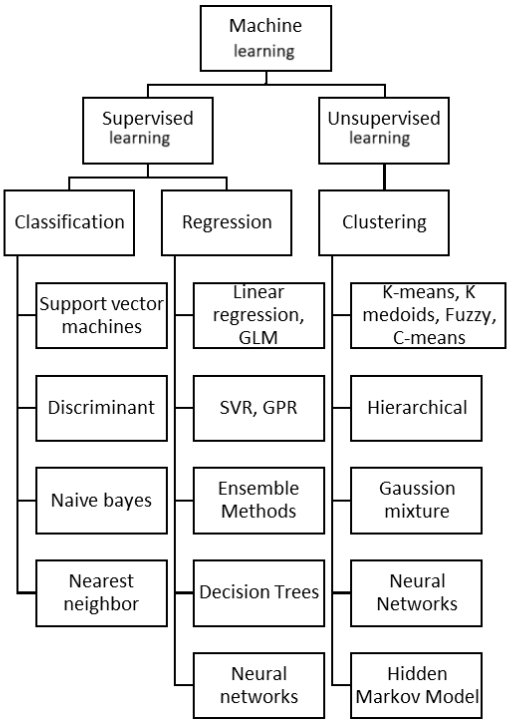


Figure 2.17: Machine learning algorithms

  - ***Classification models*** are trained to classify data. Applications include medical imaging, voice recognition, etc.
  - ***Regression methods*** can predict continuous responses, such as data prediction, and acoustic signal processing, etc.

- ***Unsupervised machine learning***, learning process is executed without output availability.

From Figure 2.17, the algorithms in subcategories are perfectly matching with specific kinds of data distribution. Selecting the most matching algorithm is a long process of data analysis and trials, considering the existing trade-off between the specific characteristics of the algorithms, such as speed of training, memory usage, prediction accuracy, etc.

2.5.2 Machine learning prediction models

In the last few years, researchers endeavored to benefit from machine learning abilities in data pattern detection and prediction. Researchers present various contributions in machine learning applications in complex phenomena, particularly ground-borne vibration prediction. In the following, the essential contributions will be highlighted.

Monjezi et al. [62] in 2010 presented a model based on an artificial neural network (ANN) to predict blast-induced ground vibration in a tunneling project. The application of ANN proved to be more accurate than conventional statistical and empirical methods. The input parameters were the maximum charge per delay, distance from blasting, stemming, and hole depth. Furthermore,

the peak particle velocity was taken as an output. A database consisting of 162 data sets was utilized for training a four-layer feed-forward back-propagation neural network. Twenty data sets were randomly selected as a testing set. For the optimum ANN model, coefficient of determination and mean absolute error between measured and predicted PPV were calculated 0.9493 and 0.002, respectively (see Figure 2.18).

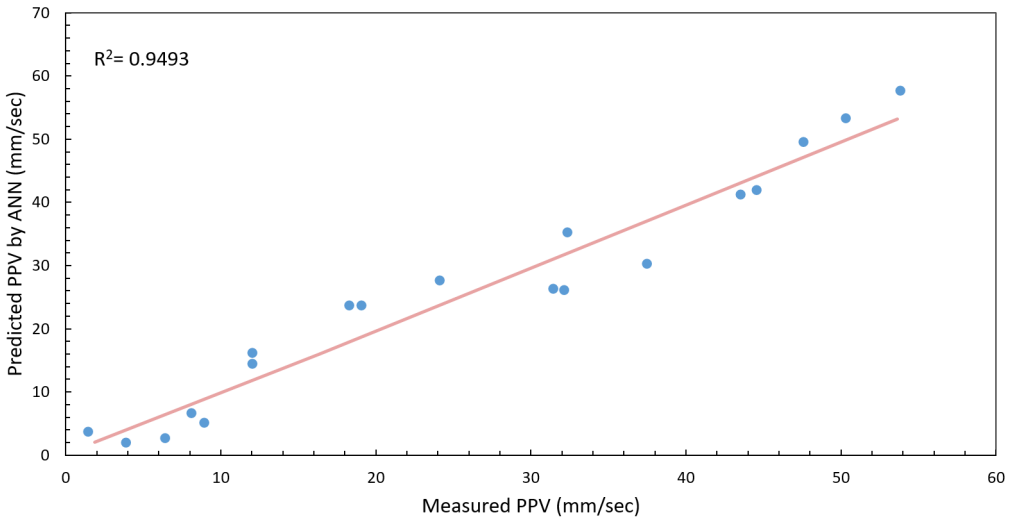


Figure 2.18: Validation for the artificial neural network model.

Singh et al. [63] in 2005 presented a model based on an artificial neural network (ANN) to predict with the ground vibration induced by mines process. Multi-layer neural networks with back-propagation training procedures carry out the necessary inversion process, as well the research concluded that the proposed ANN tool possesses great potential for solving complex problems like ground vibration prediction [64]. Furthermore, ANNs techniques have been used successfully in wave propagation modelling to predict vibration levels from 2D trenches to isolate railway vibration [65], and railway track geometry to vehicle performance [66].

Connolly et al. [67] in 2014 suggested a high-accuracy initial assessment prediction tool for the ground vibration induced by high-speed lines. The advantage of the new approach is including the effect of soil conditions in its calculation and zero run times, thus allowing for the rapid assessment of vibration levels across rail networks. The model development was founded using a fully validated three-dimensional finite element model to generate synthetic vibration records for a wide range of soil types(learning data). These records are analyzed using a machine learning approach to map relationships between soil conditions, train speed, and vibration levels. Connolly [67] suggests using a neural network algorithm (NNA) in the training process due to its non-linear regression ability. Its performance is tested by predicting two independent international vibration metrics on four European high-speed lines, and it is found to have high prediction accuracy (see Figure 2.19).

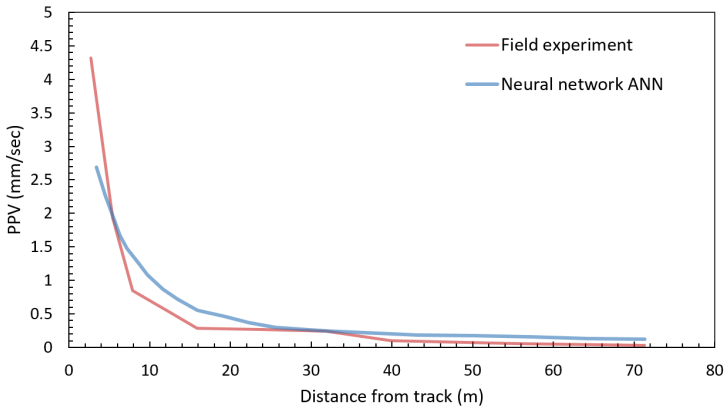


Figure 2.19: Validation for the artificial neural network model proposed by Connolly et el. [67]

Extreme Gradient Boosting, commonly known as XGBoost, was introduced by Chen and Guestrin 2016 [68] as a machine learning algorithm within the Tree group. This algorithm has gained significant attention in both data science competitions and real-world applications due to its outstanding performance in tasks involving classification and regression. The notable success of XGBoost can be attributed to its systematic training process. The training procedure in XGBoost involves iteratively adding decision trees to the model, with each tree aiming to correct errors made by its predecessors. The predictions from each tree are assigned weights based on their contribution to the overall objective function. This unique feature enables XGBoost to effectively address complex phenomena, making it particularly promising for geotechnical applications. However, it is important to note that boosted trees, including those used in XGBoost, tend to over-fit the training dataset. As a result, achieving high accuracy on the training data does not necessarily lead to a good model for unseen data. To mitigate this, it is important to employed the holdout technique to split the data into training, testing, and validation sets. This approach helps ensure that model's performance is robust and generalize-able for new data. Despite its potential, the utilization of XGBoost in geotechnical problem-solving, especially in pile driving scenarios, remains relatively limited. However, a comparative study conducted by Amjad et al. 2022 [69] assessed various machine learning models for predicting pile-bearing capacity. The findings revealed that the XGBoost model outperformed other models, including AdaBoost, Random Forest (RF), Decision Trees (DT), and Support Vector Machines (SVM), showcasing its superior capability in predicting bearing capacity.

Nguyen et al. 2023 [70] introduced a new approach combining Student Psychology Based Optimization (SPBO) with XGBoost to predict the base resistance of super large and long bored-barrette piles in soft soil. The study concluded that this hybrid method enhanced the efficiency of hyper-parameter tuning, improved computational stability, and increased prediction accuracy compared to existing methods. In another innovative application, Kamran in 2023 [71] proposed a hybrid model that integrated the Whale Optimization Algorithm (WOA) with XGBoost to predict the slope of soil stability. The results demonstrated that the WOA-XGBoost hybrid exhibited high reliability, providing a promising solution for accurate slope predictions in soil stability studies.

From the previous expose, the investigators presented a significant contribution in predicting the ground-borne vibration-induced with different sources. However, the scientific community suffering from a lack of applications for machine learning techniques in predicting the ground-borne vibration induced by the pile driving process, which justifies the author's desire to present a significant contribution to the engineering society in this field.

Chapter 3

Vibrations induced by impact pile driving: numerical modeling and phenomenological analysis

3.1 Introduction

Vibrations generated by pile driving activities can disturb people and activities, cause damages to existing equipment and, in extreme scenarios, damage in the surrounding structures. In such situations, the prediction, monitoring and control of vibrations are important issues to reduce or prevent pernicious effects induced by this construction activity. From the literature review, it was found a lack in what concerns to expedite predictions methods, with neither excessive engineering costs nor large computational times, incorporate all subsystems of the problem. From this perspective, this chapter explores the phenomenological analyses of the driving process from the hammer impact to the ground response (see Figure 3.1), as well identifies the main variables that interfere with the ground response in order to allow the development of such a prediction tool in a subsequent step.

For the purpose initially set, an advanced numerical method is applied. The author should emphasize that the numerical methodology was previously developed by the research group and it will be essential for the thesis's main goal: developing a prediction method based on the AI techniques [4]. In short, the advanced numerical model will be used to generate a synthetic vibration database. Then, the relation between input factors and vibration levels is obtained by analyzing the database using the most suitable machine learning techniques. The previous context justifies a deeper exposition of the axisymmetric FEM-PML approach developed by Colaço et al. [4].

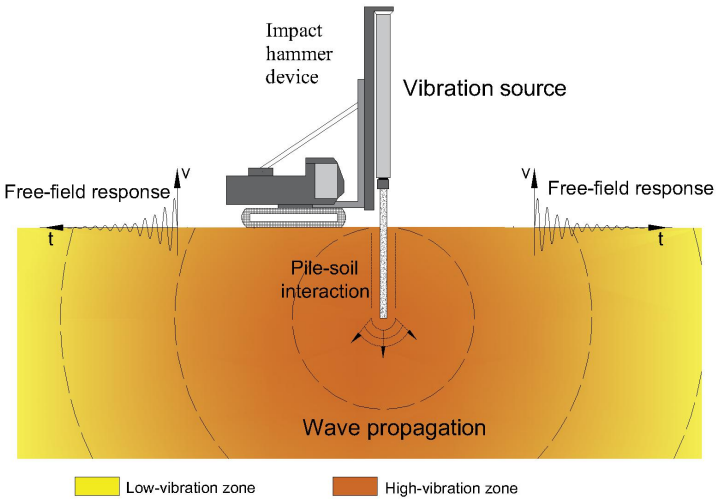


Figure 3.1: Vibrations induced by pile driving: representative scheme of the main physics.

3.2 Numerical prediction model

3.2.1 General considerations

The numerical model developed by Colaço et al. [4] is split into two main modules. The first module comprises the pile-ground system, modelled by a FEM-PML approach in axisymmetric conditions. The second one corresponds to the dynamic simulation of the hammer device. Since the hammer device and the pile have a dynamic interaction, both models are merged, meeting the equilibrium and compatibility requirements. Figure 3.2 shows the interaction scheme of the numerical model.

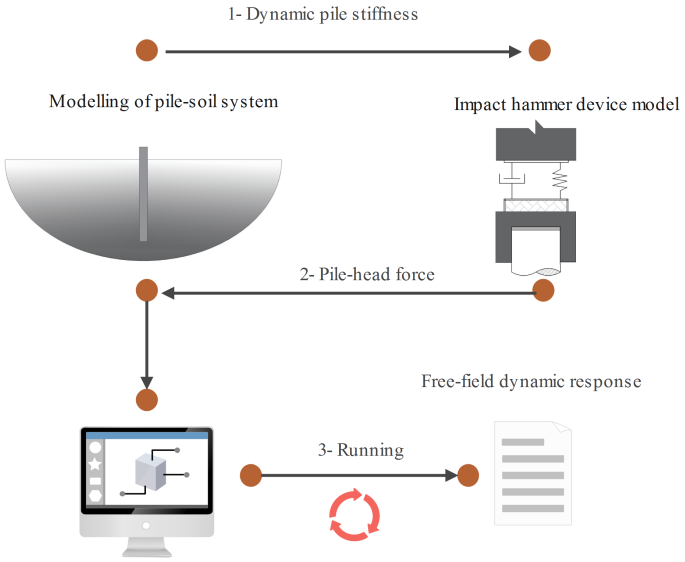


Figure 3.2: Representative scheme of the numerical modelling approach.

3.2.2 Modelling of the pile-ground system

3.2.2.1 General formulation

The numerical modeling of the pile-ground system is based on the coupling between the finite element method and perfectly matched layers in axisymmetric conditions, where a schematic illustration is presented in Figure 3.3. Perfectly matched layers suppress undesirable spurious waves' reflections in infinite media modeled with finite elements. It should be clarified that the modeling approach was implemented in MATLAB®[5]. The equilibrium equation for any point of a three-dimensional domain can be derived based on the principle of virtual work and following the usual steps of the finite element procedure as follows:

$$\int_V \delta\varepsilon^T \sigma dV + \int_V \delta u^T \rho \frac{\partial \mathrm{u_i}(x,t)}{\partial \mathrm{t}^2} dV = \int_S \delta u^T p dS \tag{3.1}$$

where $\delta\varepsilon$ is the virtual strain field, σ is the stress field, δu is the virtual displacement field, u is the displacement field, ρ is the volumetric mass and p is the external loads applied to surface S.

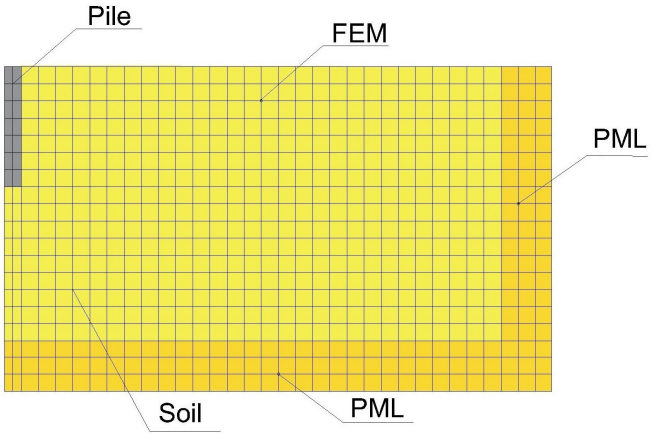


Figure 3.3: Pile ground system modeled by an axisymmetric FEM-PML numerical approach.

The assumption of an axisymmetric model helps for improving the computational effort as it allows to obtain a three-dimensional response to the problem without the need for a 3D discretization. Consequently, the integral volume in Equation 3.1 can be adapted for axisymmetric problems by expressing the differential volume dV as $2\pi rdS$, where r is the radial distance from the axisymmetric axis. Likewise, the differential surface can be converted to the linear form by expressing dS as $2\pi rdL$. According to the principles of the finite element, following will be extract:

$$[K] = 2\pi \int_S B^T DBrdS \tag{3.2}$$

$$[M] = 2\pi \int_S N^T \rho Nrds \tag{3.3}$$

where [K] and [M] are the stiffness and mass matrices, respectively.

In the PML domain, PML elements must allow the attenuation, without reflection, of the wave field that impinges the FEM-PML boundary [72]. For that purpose, a bounded layer consisting of elements with artificial attenuation is introduced [73]. The region replaces the infinite domain with a finite dimension layer, which can absorb the wave field and is also non-reflective. The absorbing condition is accomplished by stretching the coordinates of the PML elements to a complex domain, leading to an artificial increase in the attenuation of the wave that propagates along the PML domain [74]. The real part of the stretching functions is introduced in order to undertake a mesh sprawl so that the thickness of the obtained PML domain can be equivalent to at least one wavelength of the propagating wave, as proposed by Lopes et al. [75]. The natural and modified coordinates are related through the following relationships:

$$\tilde{x} = \int_0^x \lambda_x(x)dx, \quad \tilde{y} = \int_0^y \lambda_y(y)dy \tag{3.4}$$

where λ_x and λ_y are the stretching functions in the x and y directions, respectively. Additional information about this procedure can be found in Lopes et al. [75].

The differential equation of PML domain can be write as the following:

$$[K^*(\omega)] = 2\pi \int_y \int_x B^{*T} D B^* r \lambda_x \lambda_y dxdy \tag{3.5}$$

$$[M^*(\omega)] = 2\pi \int_y \int_x N^T \rho N r \lambda_x \lambda_y dxdy \tag{3.6}$$

where $[B^*]$ is the strain matrix in the stretched domain, $[K^*]$ is the stiffness matrix, and $[M^*]$ is the mass matrix. After the assemblage between the matrices of the PML boundary and the remaining domain, the final solution in frequency domain would be shown as the following:

$$\left\{([K] + [K^*(\omega)]) - \omega^2 ([M] + [M^*(\omega)])\right\} \{u(\omega)\} = \{p(\omega)\} \tag{3.7}$$

Since the simulation is performed in the frequency domain, it is possible to get transfer functions between the response and a unitary loading condition [4].

3.2.2.2 Equivalent linear formulation

The frequency domain model is based on the visco- elastic and linear behavior of the pile-soil system. However, this approach simplifies the actual behavior of the soil, which is highly dependent on the induced strain level, as discussed in Chapter 2. As shown in Figure 3.4, with an increase in the strain level, the stiffness decreases, and the energy dissipation increases for the conditions of asymmetric periodic loading [76, 77].

Pile driving operations can subject the soil to a strain level higher than the allowable elastic limit. Therefore, if large strain values are expected, i.e., more than 10^{-3}, the equivalent linear approach should be integrated into the model, whereby the effects of the non-linear behavior of the soil can be taken into account. Colaço et al. [4] clarifying that the equivalent linear technique integrated into their model involves an iterative process that replaces inelastic behavior by adjusting the elastic material parameters for significant strain levels. The existence of laws describing the decrease in stiffness and the increase in damping as strain levels increase are key features to make this iterative technique [78, 79]. The size of the area within the curve, (see Figure 3.4), represents

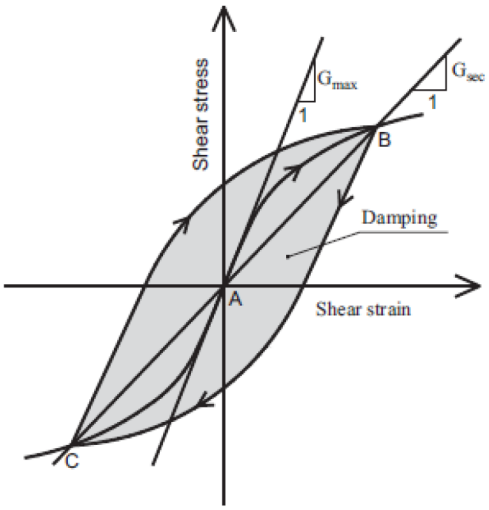
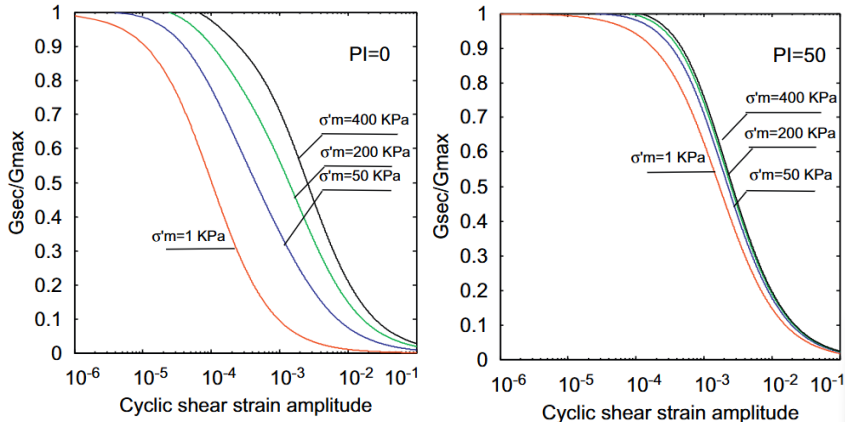


Figure 3.4: Stress – strain path during cyclic loading [76].

the energy density lost in every cycle; the dimension is energy per volume unit (energy density). Both the initial shear modulus and the corresponding shear stress are subjected to decrease with the number of cycles, known as cyclic degradation. The stress-strain relationship and the degradation law depend to a large extent on the soil type. Cohesive soils are less susceptible to cyclic degradation than non-cohesive soil as the particles are more tightly connected to each other. In non-cohesive soils, the particles can rearrange and lose contact during vibrations [80].

For the development of an equivalent linear analysis, two main parameters are needed: (i) the elastic properties for very small strains (usually obtained from the propagation velocity of S and P waves) and (ii) the laws describing stiffness degradation and damping increase with the increase of the strain levels [81]. These laws should be obtained from laboratory tests or, in its absence, from correlations with some physical parameters of the soil. Figure 3.5 shows the global aspect of the typical degradation laws and their dependence on the plasticity index (PI) and mean confining stress

As implied, the definition of the strain level at each finite element is required. In 3D problems, the strain level γ_{eff} is usually defined by the effective octahedral shear strain, see Equation 3.8 as



(a) (b)

Figure 3.5: Modulus reduction curves for: (a) non-plastic soils and (b) plastic soils [78].

proposed by Lysmer et al. [79] and Naggar et al. [82]:

$$\gamma_{eff} = \alpha \frac{1}{3} \sqrt{(\varepsilon_z - \varepsilon_r)^2 + (\varepsilon_r - \varepsilon_\theta)^2 + (\varepsilon_z - \varepsilon_\theta)^2 + 6(\gamma_{zr})^2} \tag{3.8}$$

where α is a parameter in the range between [0.5 - 0.7]. The variables γ_{zr} and ε_i correspond to the strains components of the strain tensor. In the numerical process, and for a discretized medium through an element-by-element procedure, it is possible to determine the induced strain at each element and correct the element properties until a match is obtained with the involved strains. The present model assumes constant properties inside each finite element, meaning the strains in the central point of the element are considered representative of the strains in the whole element domain. For the update of the properties of the elements, a set of linear problems are successively solved until a match between the strain level and the dynamic properties of the soil is reached. The procedure can be summarized as the following:

1. Assume low-strain properties for all elements of the medium.
2. Compute the time history of strains and evaluate the maximum value of shear strain γ^i_{eff}.
3. Use γ^i_{eff} to choose new equivalent linear values,G^{i+1}_{sec} and ξ^{i+1}_{sec} for the next iteration.
4. Repeat steps 2 to 3 until the differences between shear modulus and damping in two successive iterations are below a pre-established tolerance for all finite elements.

In terms of the convergence tolerance, Colaço et al. [4] assumed a value of 3%, since the linear equivalent model corresponds to an approximation to the real problem.

3.2.3 Hammer model

Following the objective initially set, analyze the dynamic response of the ground during pile driving, the time history of the force generated by the impact of the hammer is required. Thus, the

impact hammer device and its interaction with the remaining domain must be simulated to estimate the pile head force. The system shown in Figure 3.6 corresponds to a free-falling weight (ram) that impacts the anvil on the pile head represented as a hammer cushion and a pile cap or helmet. The ram and anvil are linked by a spring that reproduces the stiffness of the cushion, and a dashpot that presents the cushion deforms non-linearly and absorbs energy [83]. This model was previously proposed by Deeks and Randolph [84] and is mathematically described by the following equations:

$$\begin{aligned} &m_r\ddot{u}_r + c_c\left(\dot{u}_r - \dot{u}_a\right) + k_c\left(u_r - u_a\right) = 0 \\ &m_a\ddot{u}_a + c_c\left(\dot{u}_a - \dot{u}_r\right) + k_c\left(u_a - u_r\right) + k_p u_a = 0 \\ &\dot{u}_r(0) = v_0 \end{aligned} \tag{3.9}$$

where m_r is the ram mass; m_a is the anvil mass, and k_c and c_c are the stiffness and damping coefficient of the cushion, respectively; v_0 is the impact velocity of the ram and k_p is the dynamic stiffness of the pile.

When considering the pile–soil dynamic interaction, the dynamic stiffness of the pile can be comparable to the ratio between the force applied at a point on the pile head and its vertical displacement. This force can be represented as a Dirac delta function, which corresponds to a large force distributed over a small time interval. This dynamic stiffness, responsible for the dynamic coupling between the driving device and the pile–ground system, can be computed taking into account the results obtained with the numerical model presented in the previous section. In fact, this is given by the ratio between the unitary load, $P(w)$, applied to the pile head and the dynamic displacement, $u_z(w)$ determined at the same location. Mathematically, this is translated to:

$$k_p(w) = \frac{P(w)}{u_Z(w)} = \frac{1}{u_Z(w)} \tag{3.10}$$

Colaço et al. [4] clarify that by converting Equation 3.9 in the frequency domain, and considering a pseudo-force application whose value is equal to the modulus of the momentum transferred from the ram to the pile ($m_r v_0$), the ram and the anvil displacement values can be computed by solving Equation 3.11:

$$\left(-w^2\begin{bmatrix} m_r & 0 \\ 0 & m_a \end{bmatrix} + iw\begin{bmatrix} c_c & -c_c \\ -c_c & c_c \end{bmatrix} + \begin{bmatrix} k_c & -k_c \\ -k_c & k_c + k_p \end{bmatrix}\right) \times \begin{bmatrix} u_r \\ u_a \end{bmatrix} = \begin{bmatrix} m_r v_0 \\ 0 \end{bmatrix} \tag{3.11}$$

When the ram hits the anvil surface, the soil resists the pile movement and causes a rebound for the ram. Since the linear elastic approach for the anvil was assumed, tractions are generated at the moment of the ram's rebound. From this rebound, a traction force was generated. In order to prevent these tractions and, consequently, the occurrence of negative forces, the force of the anvil surface in the time domain is truncated, as illustrated in Figure 3.7 for a generic case.

The truncated anvil surface force is converted back to the frequency domain, and the Equation 3.11 is recalculated, considering this force instead of the momentum, $m_r v_0$. Therefore, new displacements for the ram (u_r) and the anvil (u_a) are calculated. Afterward, it is possible to compute

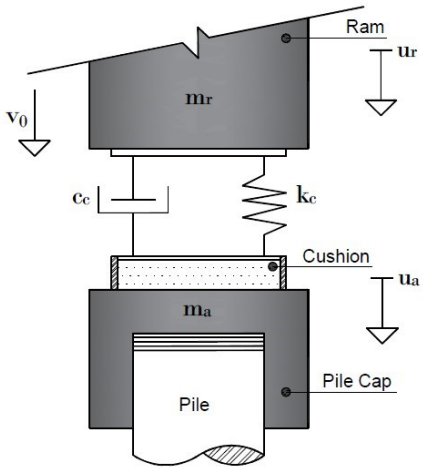


Figure 3.6: Two degrees of freedom pile-hammer model [4].

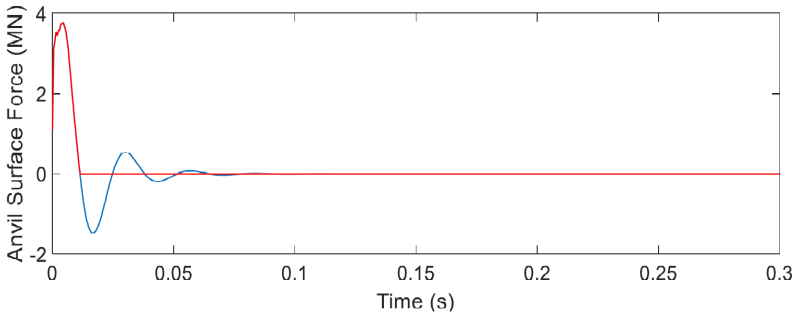


Figure 3.7: Free vibration (blue line) and truncated force (red line) applied at the anvil [85].

the pile head force F_p in the frequency domain from the following equation:

$$F_p(w) = u_a(w) * k_p(w) \tag{3.12}$$

Finally, after computing the load applied to the pile head, the dynamic response of the system is obtained, in the frequency domain, by multiplying the transfer function, computed through the axisymmetric FEM-PML approach, by the loading function F_p. The results in the time domain are obtained by an inverse fourier transform operation.

3.2.4 Numerical validation

Before the exposition of detailed analysis about the generation and propagation of ground-borne vibrations, it is intended to present a simple numerical validation case, taking into account the numerical model proposed by Masoumi et al. [55].

Figure 3.8 expose the properties of soil material.BSP-357 hammer is used to drive a concrete pile with a circular section. The hammer cushion stiffness is $k_c = 1.6 \times 10^9$ N/m. The ram and

anvil masses of the hammer are $m_r = 6860$ kg and $m_a = 850$ kg, respectively. According to Masoumi et al. [55] the pile has a length $L_p = 10$ m, a diameter $d_p = 0.50$ m, a Young's modulus $E_p = 40000$ MPa, a Poisson's ratio $\nu_p = 0.25$ and a density $\rho_p = 2500\ \mathrm{kg/m^3}$. The longitudinal wave velocity of the pile $C_p = 4000$ m/s.

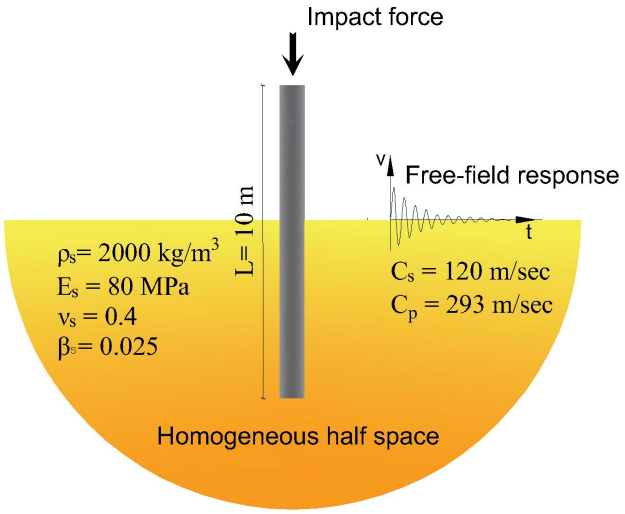


Figure 3.8: Geometry of an embedded pile in a layered medium.

In terms of numerical modeling, the rated energy $E_r = m_r gh$ is assumed to be equal to the kinematic energy $E_t = m_r v_0^2/2$ that is effectively transmitted to the pile, with v_0 the impact velocity.

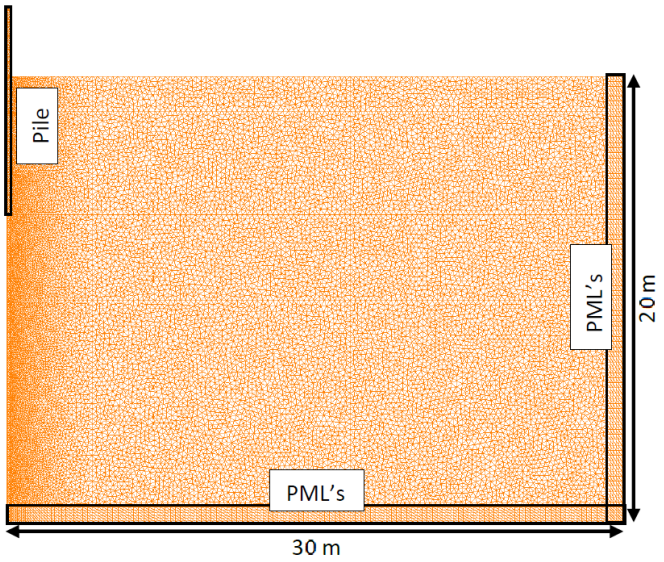
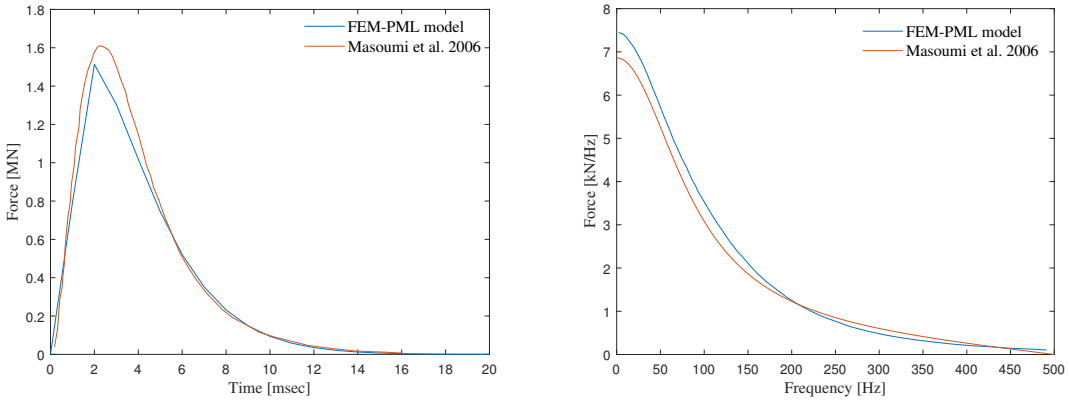


Figure 3.9: FE–PML mesh adopted to model the pile-ground system.

The pile-ground system is discretized by a FE-PML mesh with triangular elements on 6 nodes (69859 nodes total), corresponding to a 30x20 m^2 discretized section. By using the axisymmetric formulation condition, the problem is solved like a two-dimensional problem, and the central axis

of the pile corresponds to the axisymmetric axis. The PML layers are bounding the FEM region, as shown in Figure 3.9.

The force applied to the pile head, considering the procedure explained above, is illustrated in Figure 3.10, being compared with the force presented by Masoumi et al. [55]. The comparison was performed using the corresponding value for the impact velocity, v_0, and the transferred energy, E_t, reported by Masoumi et al. [55] that equal 1 m/s and 3.4 kJ, respectively (see Figure 3.10). However, these values would be changing later as the pile goes deeper. The figures prove the attendance of an acceptable matching between both approaches in the time and frequency domains.



(a) (b)

Figure 3.10: Impact hammer force: (a) time history, (b) frequency content.

Taking into account the applied pile-head force, the ground response is computed and compared with the reference case. It is possible to systematize the extended set of information through a linear plot, where the peak particle vertical velocity of vibration is plotted as a function of the distance for both numerical approaches (see Figure 3.11). The contributions seismic waves are investigated for different penetration depths $e_p = 2$ m and 10 m. The values of the penetration depth were chosen equal to the minimum and maximum penetration depth reported in Masoumi et al. [55] model. For the penetration depths $e_p = 2$ m, an impact velocity equal to $v_0 = 2.4$ m/s, resulting in impact with a transferred energy E_t of about 19.4 kJ. On the other hand, for the penetration depths $e_p = 10$ m, an impact velocity equal to $v_0 = 0.63$ m/s, resulting in a low impact with a rated energy E_r of about 3.4 kJ.

Analysing the results depicted in Figure 3.10, it is possible to identify a very good match between the results provided by the axisymmetric FEM-PML approach proposed by Colaço et al. [4] and the results reported in Masoumi et al. [55]. This initial step allowed to eliminate the practice erroneously factor by the user and enhanced the confidence of the modeled results in the present work.

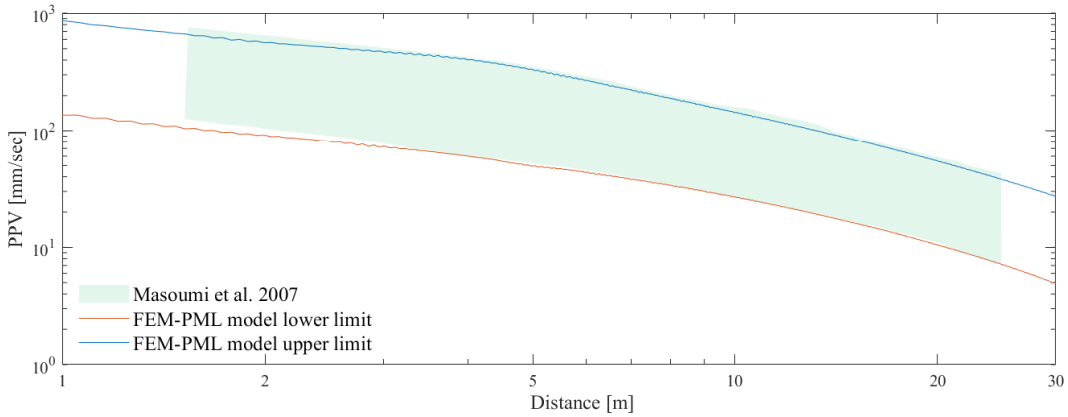


Figure 3.11: Numerical validation between FEM-PML model and Masoumi et al. model [55].

3.3 Physics of the generation and propagation of ground vibration

3.3.1 Reference case

The ground vibration induced by pile driving is a complex phenomenon, as it is affected by several parameters. From this context, the author intends, through this section, to present the influence level of the most relevant parameters for the phenomenon under investigation. These parameters can be grouped into three main groups: propagation field, pile properties, and hammering properties influence. In the following expose, the influence of each parameter correlated to the phenomenon under investigation will be analyzed individually.

As previously discussed, the non-linear behavior of the soil near the pile has a relevant role in the vibration field, and the analysis should take this effect into account. In connection with what was previously presented, an equivalent linear methodology is adopted to take this behavior into consideration. Additionally, in a purely theoretical perspective, the numerical analysis can be performed using an elastic-linear approach, i.e., neglecting the degradation of soil stiffness and the increase in damping according to the strain level. For the sake of presenting a comprehensive investigation, the parametric study is presented in a comparison form between a linear and an equivalent linear analysis for the propagation medium, assuming the degradation curve presented in Figure 3.12.

The objective initially set would be achieved by present the reference case with full detail. From the reference case, a set of cases will be produced based on one change per parameter per case. The comparison will be performed based on grouping all cases subjected to changes in the same parameter. From the author's perspective, the target number for changes in the same parameter is limited by three, as this number of cases will be sufficient to figure out if the parameter has a touchable influence on the ground response or not. In the following, the general properties of the numerical simulation will be exposed.

Regarding the numerical simulation of the pile-ground system, a FEM-PML mesh consisting

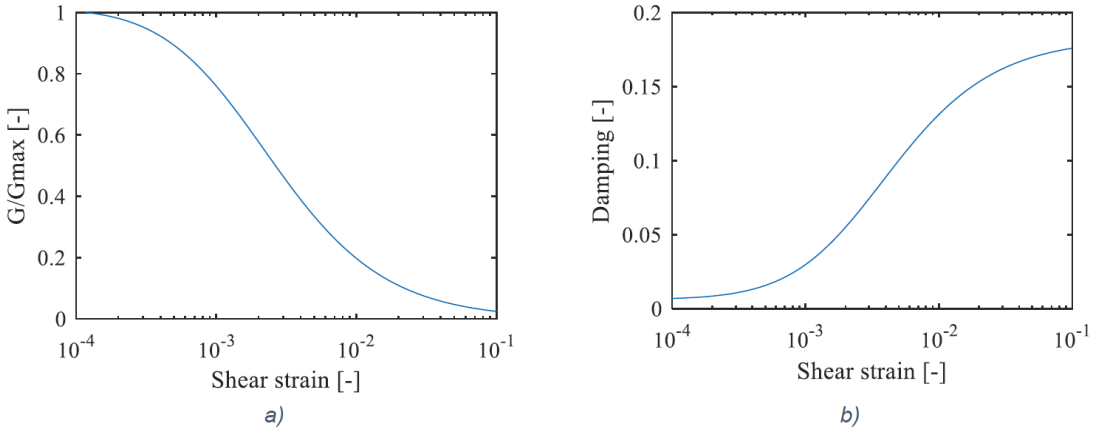


Figure 3.12: Evolution of the dynamic properties of a plastic soil (PI=50) with shear strain: a) stiffness degradation; b) damping increase [78].

of 55796 triangular elements with 6 nodes (total of nodes – 107189) is considered, corresponding to a discretized cross-section of 50x25 m x m. The pile present a Young's modulus, E, of 30 GPa and a mass density, ρ, of 2500 kg/m3. A hysteretic damping factor of 0.01 and a Poisson's ratio equal to 0.15 are also considered for the pile. Figure 3.13 shows an overview of the configuration, properties and mesh adopted for the numerical example.

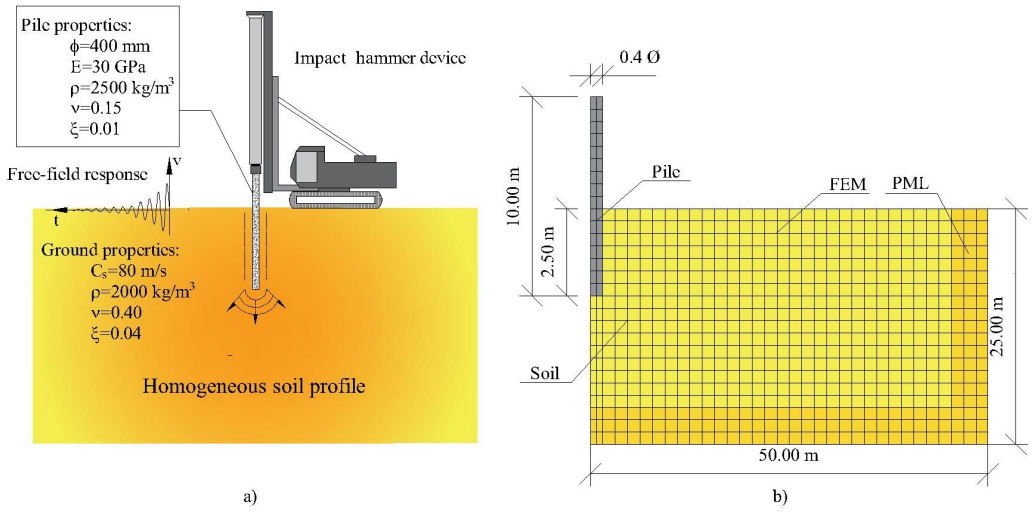


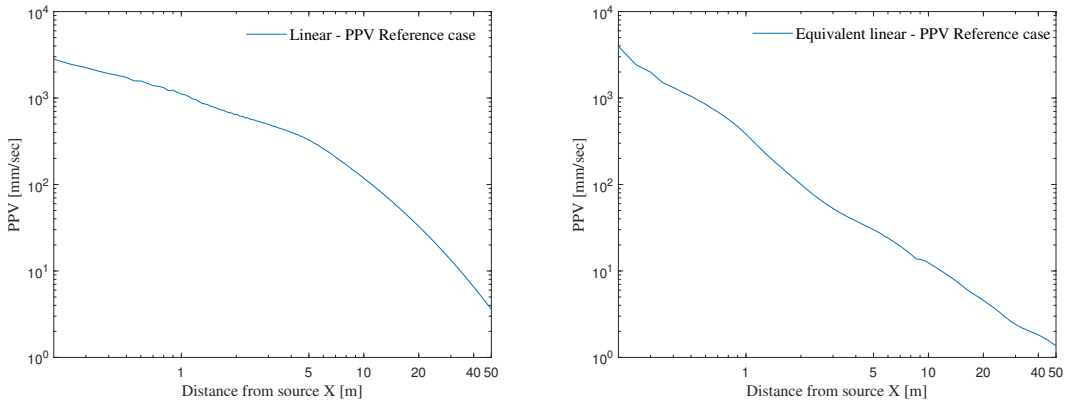
Figure 3.13: Pile driving in a homogeneous half-space soil: a) application example; b) pile-ground axisymmetric FE–PML mesh.

Regarding the driving device, a hydraulic impact hammer with a ram mass, m_r, of 7 tonnes was considered. An anvil mass, m_a, equal to 850 kg was assumed as well. The hammer cushion has a stiffness value of kc = $2x10^6$ kN/m and a damping value of 588 kNs/m.

The transmitted energy $E_t = \eta E_r$ is lower than the rated energy E_r due to energy loss related

to precompression, friction between the ram and the cylinder, preignition and misalignment. The transfer coefficient η for assumed 0.95. The height of fall H is 0.5 m. Figure 3.14 shows the PPV versus distance from the pile.

Figure 3.14 exposes the difference in the ground response between the linear and equivalent linear approach. According to the results, the ground surface presents PPV values higher whenn assumed linear approach. In addition, the result of the linear approach tends to be concave upwards relative to the equivalent linear approach, which tends to be concave downwards. Such differences can be attributed to the effect of the non-linearity of the soil, especially in the zone near the excitation sources.

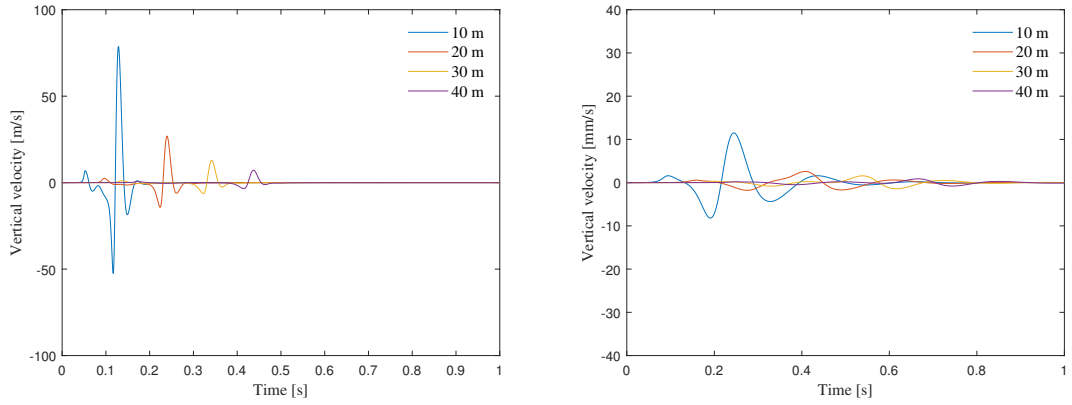


(a) (b)

Figure 3.14: PPV vs. distance for the reference case: a) linear analysis; b) equivalent linear analysis.

For wide exploring for the reference case, Figure 3.15 shows the time series of the vertical velocity of the particles in 4 points located at the ground surface and for horizontal distances to the centre of the pile of 10 m, 20 m, 30 m and 40 m. As expected, the vibration amplitude tends to attenuate as the distance to the pile increases. Due to the geometric attenuation, this effect is observed regardless of the type of analysis performed. However, peak vibration values are higher when a linear elastic behaviour of the system is assumed.

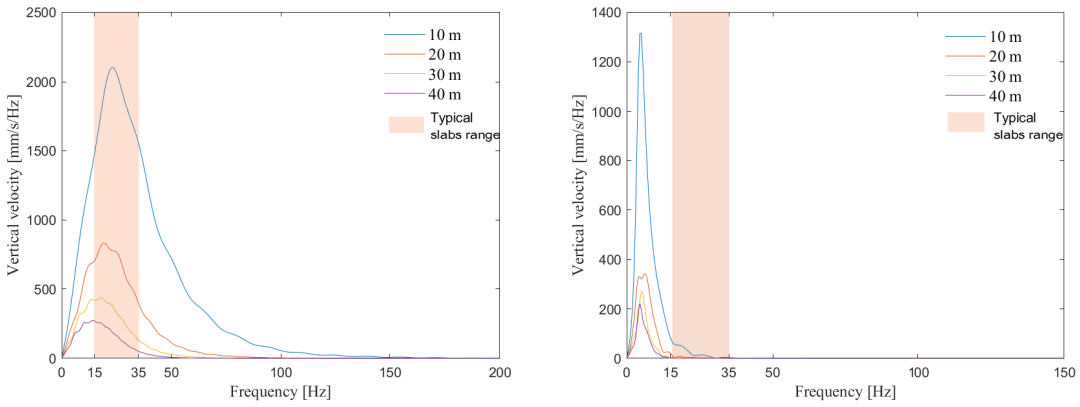
The frequency range observed in the vertical velocity under free-field conditions is quite relevant when dealing with the dynamic response of existing nearby buildings. Since reinforced concrete buildings slabs typically have a resonant frequency ranging from 15 Hz to 35 Hz, therefore in case of attending the matching between the frequency applied from the vibration and the natural frequency of the slab, large displacement can be expected, which affects the integrity and stability of the structure [86, 87]. In this context, the frequency content of the vertical velocity of the particle is plotted in Figure 3.16 for the scenario under analysis. It is possible to see that there is a tendency for a concentration of energy in lower frequency ranges for the equivalent linear approach rather than the linear approach. This concentration of energy is also verified as the distance from the receiver to the pile increases. This is essentially justified by the material damping,



(a) (b)

Figure 3.15: Velocity versus distance: a) linear analysis; b) equivalent linear analysis.

that consider larger when equivalent linear approach is used, that introduces an higher attenuation in the higher frequency range. In the figure the orange band present the range of the a resonant frequency for a typical reinforcement concrete.



(a) (b)

Figure 3.16: Vertical velocity of the particles, in the frequency domain, assuming: a) linear analysis; b) equivalent linear analysis.

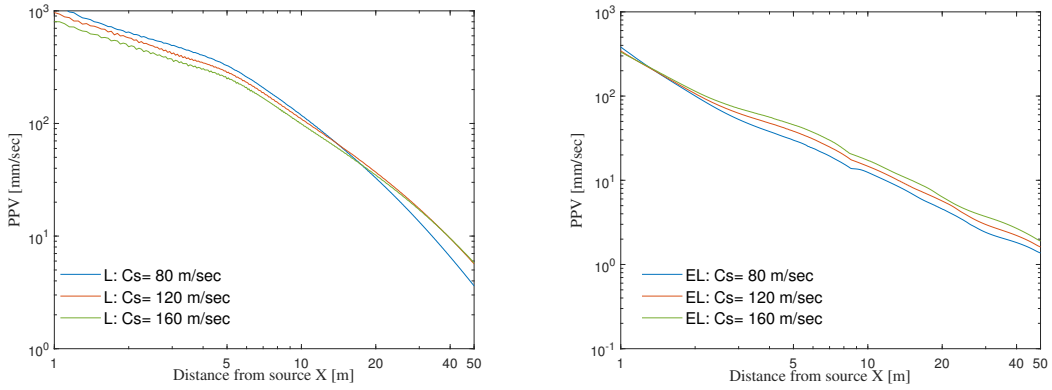
Based on the previous exposition, now, the parametric study will be presented in order to assess the influences of the propagation field, pile properties, and hammer properties on the ground response.

3.3.2 The influence of ground properties

The main objective of this section is to assess the influence of the propagation medium on the ground surface response. The propagation medium is mainly subjected to changes in two dimensions, soil material properties, and soil profile. In the following, the influence of the ground

stiffness, the influence of the ground layering, and influence of the soil damping are exposed.

Following the objectives initially set, in order to test the influence of soil properties two different values of shear wave velocity C_s would be consider: 120 m/sec and 160 m/sec. It is important to mention that the reference case has a shear wave velocity C_s equal to 80 m/sec. Figure 3.17 presents the evolution of the peak particle velocity according to the distance from the pile. As observed, higher vibration levels are associated with stiffer soils. There is only one exception associated with the linear analysis and shorter distances. However, this trend is reversed at large distances.



(a) (b)

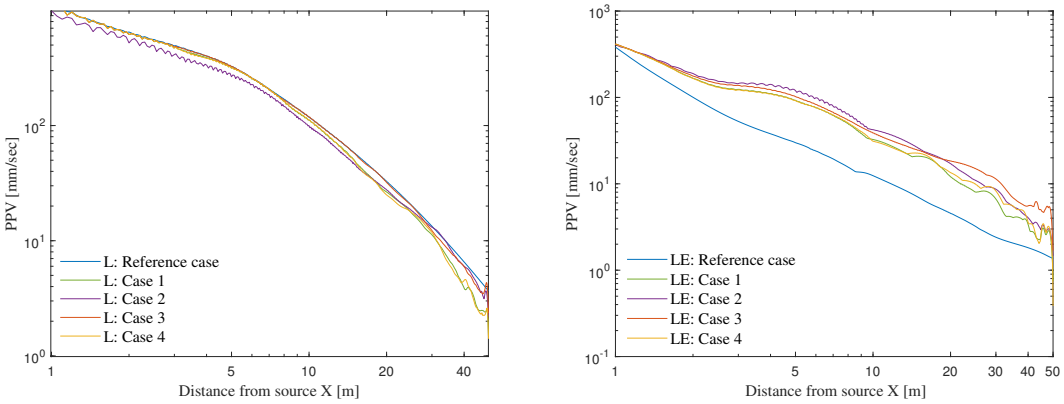
Figure 3.17: PPV versus distance from the pile due to impact pile driving for different ground stiffness: a) linear analysis; b) equivalent linear analysis.

Now, it is essential to expose the influence of stratified soil. For presenting this parameter, the comparison will be established between the reference case, which represents the homogeneous soil, and case 1, case 2, case 3, and case 4, which represents the stratified soil. Table 3.1 exposes the properties of the reference case, case 1, case 2, case 3, and case 4.

Table 3.1: Table of geomechanical properties of the assumed profile.

Item	layer 1		layer 2	
	H_1[m]	C_{s1}[m/sec]	H_2[m]	C_{s2}[m/sec]
Ref. Case	INF.	80	–	–
Case 1	5	80	INF.	300
Case 2	5	120	INF.	300
Case 3	10	80	INF.	300
Case 4	5	80	INF.	500

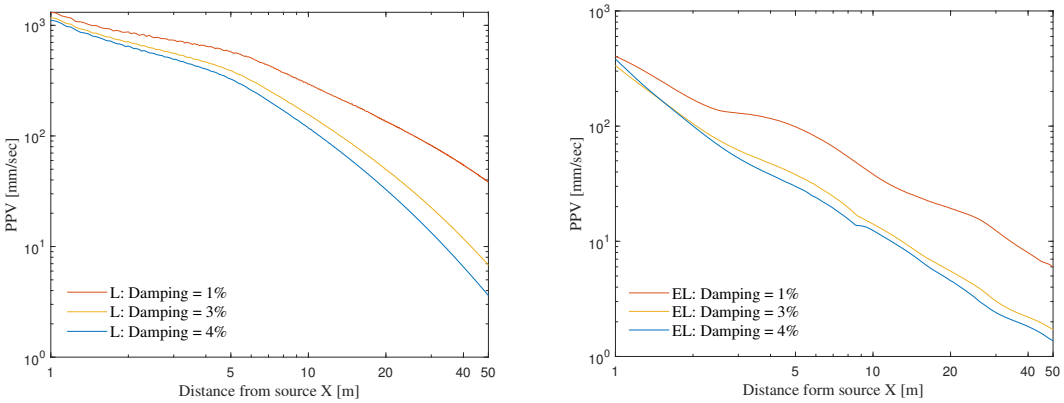
As presented in the Figure 3.18, the stratified soil influence the ground response in a linear and equivalent linear approach. The contrast can be attributed to the contribution of the reflected and refracted waves generated by waves crossing from one layer to another with different stiffness. The influence of changing the height of the layers can be figured out if the result of case 1 is compared with the result of case 3.



(a) (b)

Figure 3.18: PPV versus distance for different soil profiles: a) linear analysis; b) equivalent linear analysis.

Finally, the damping ratio is the final parameter related to the soil properties placed under test. Two scenarios with damping ratio 1% and 3% will be present in comparison with the reference case study, damping ratio 4%, to clarify the effect of the damping on the ground response (see Figure 3.19). From the comparison, it can be concluded that the influence of this parameter is extremely high. From the figure, the PPV decrease whenever the damping ratio increases, especially for higher distances, such that behaviour can be attributed to increasing the energy dissipated by the soil when the damping ratio increase.



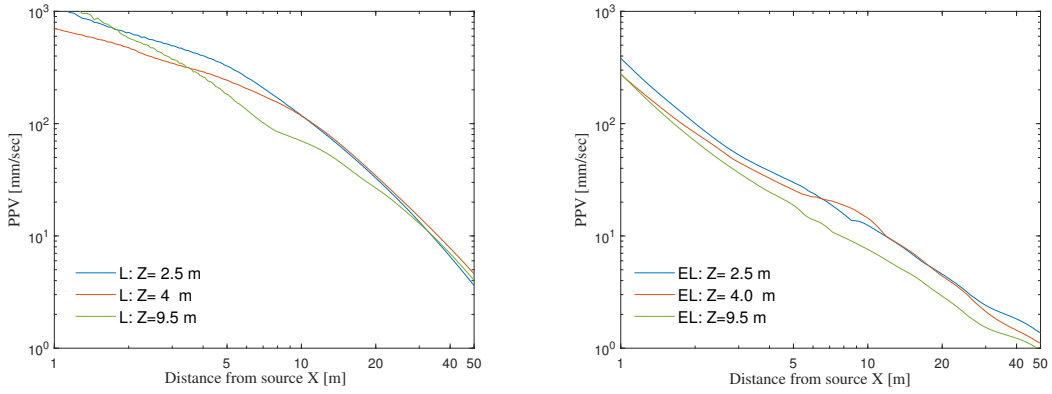
(a) (b)

Figure 3.19: PPV versus distance from the pile for different soil damping value: a) linear analysis; b) equivalent linear analysis.

3.3.3 Pile properties and penetration depth influence

Piles, in this context, can be defined as the transmission medium for the vibrations from the source placed on the pile head to the propagating medium. Based on this definition, the properties of the transmission field essentially have an effect on the vibration waves that travel through it. The pile properties are mainly subjected to changes in two dimensions, material properties, and geometry. In the following, the influence of pile penetration depth, pile diameter, and pile material properties will be addressed.

Regarding the influence of the pile penetration depth (see Figure 3.20), three penetration depth would be considered, 2.5 (reference case), 4 and 9.5 meters. From the figure it can be extract that in the linear approach, minimum penetration depths lead to higher levels of vibration on the short distance. The same behaviour can be extracted from the results provided by an equivalent linear approach. However, the authors should clarify that these comments are valid for a homogeneous ground. When a layered ground is considered, an additional complexity is involved and a detailed analysis is recommended.

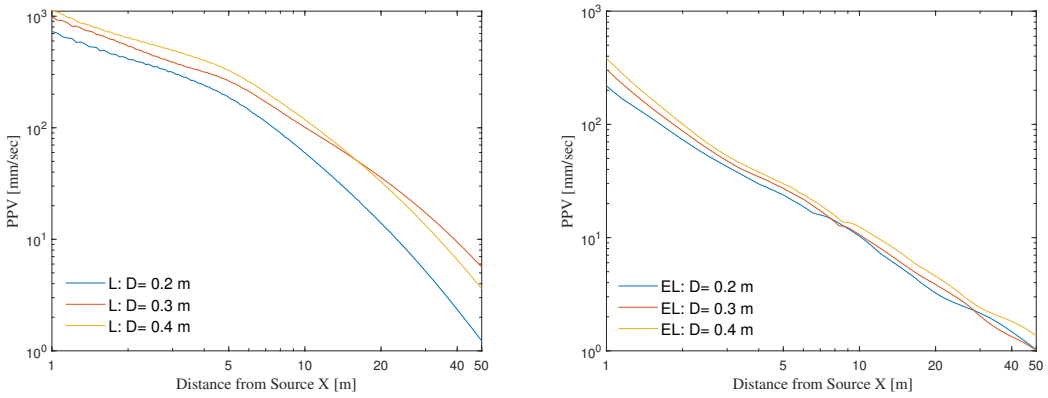


(a) (b)

Figure 3.20: PPV versus distance for different penetration depth: a) linear analysis; b) equivalent linear analysis.

The obtained results regarding the influence of the pile diameter are plotted in Figure 3.21. Three pile diameters were considered: 0.2, 0.3, and 0.4 (reference case) meters. In the linear approach, it is easy to confirm that the changes in pile diameter can lead to a significant effect on the level of vibration. In the equivalent linear approach, the influence does not assume the same importance, however it exist and can be extracted. The general behaviour for this parameter is as the pile diameter decrease the vibrations level decrease as well.

Regarding the influence of the construction material of the pile will be under study. Piles can be fabricated with different materials, i.e., steel, concrete, and wood. As it is recognized, the mechanical properties of the material essentially dominate the levels of the vibration crossing this material. The mechanical properties are exposed in the Table 3.2.



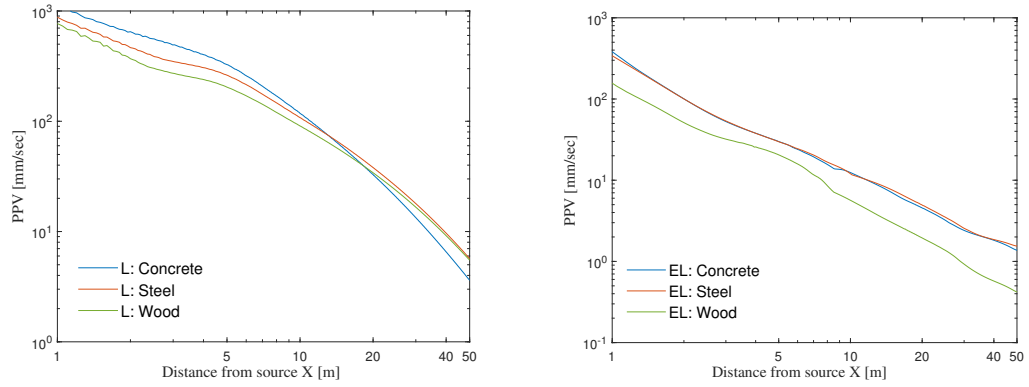
(a) (b)

Figure 3.21: PPV vs. distance for different pile diameter: a) linear analysis; b) equivalent linear.

Table 3.2: Pile material mechanical properties.

Material	E[GPa]	Poisson coefficient	Damping [%]	Density $[Kg/m^3]$
Concrete [ref. case]	30	0.15	0.01	2500
Steel [88]	210	0.28	0.02	7850
Wood [88]	10	0.4	0.07	600

Figure 3.22 present the influence of material properties on the vibration at the ground surface by linear and equivalent linear approaches. In the linear approach, concrete leads to higher levels of vibration and the wood gives the lowest levels. The same clarity is not found in the equivalent linear approach, it is difficult to confirm significant changes between concrete and steel scenarios. However, the wood case scenario present the lowest values in the comparison for PPV. The presented contrast between the different materials can be attributed to the damping ratio, as presented in Table 3.2, the damping ratio assumed to the wood is seven times the damping ratio of concrete [89, 90].



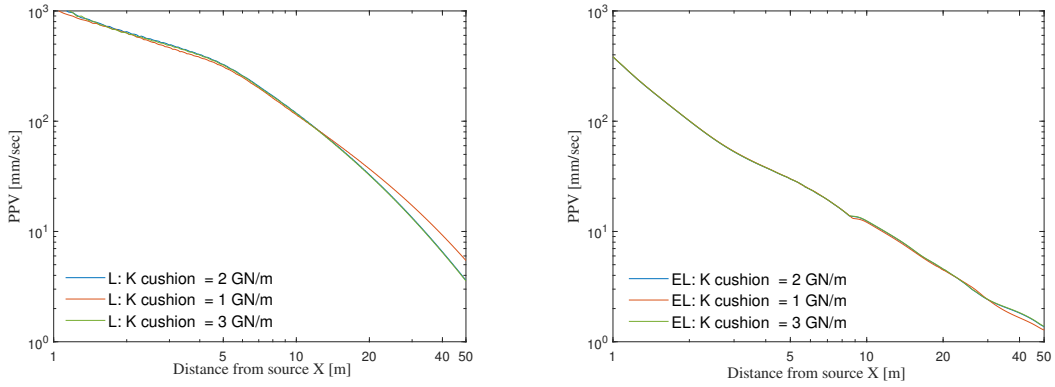
(a) (b)

Figure 3.22: PPV vs. distance for different pile materials: a) linear analysis; b) equivalent linear.

3.3.4 Hammer properties influence

The hammering process has an essential influence on the ground surface response, as it is the source of the vibration in the SSI system. In this context, it is essential to recall Figure 3.6 for a schematic representation of the hammering process. From the figure, it can extract that m_r, k_c, c_c, m_a and v_0 are the controller parameters.

Following the objectives initially set, in order to test the influence of k_c a two different values of cushion stiffness would be consider: 1 GN/m and 3 GN/m. It is important to mention that the reference case has a cushion stiffness k_c equal to 2 GN/m . Figure 3.23 presents the evolution of the peak particle velocity according to the distance from the pile. As can see from the figure, the linear and equivalent linear approaches present the same clarity, it is hard to find significant changes between different scenarios.



(a) (b)

Figure 3.23: PPV versus distance for different cushion stiffness: a) linear analysis; b) equivalent linear analysis.

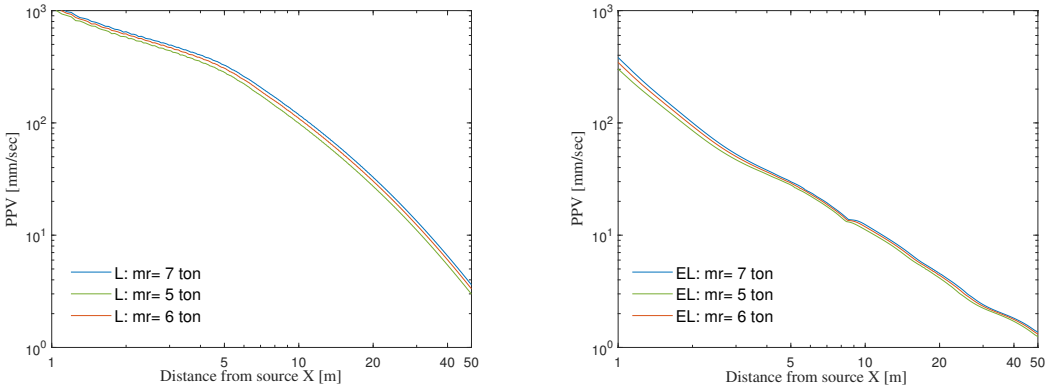
Regarding the influence of the ram mass, m_r, see Figure 3.24. Three different values would be considered, 5, 6, and 7 (reference case) ton. In the linear approach, higher ram mass lead to higher levels of vibration. The same clarity is found in the results provided by an equivalent linear approach.

Regarding the influence of the impact velocity of the ram v_0, this parameter is directly related to the drop height, experimentally measured as presented in Equation 3.13.

$$v_0 = \sqrt{2 * g * h * \alpha} \tag{3.13}$$

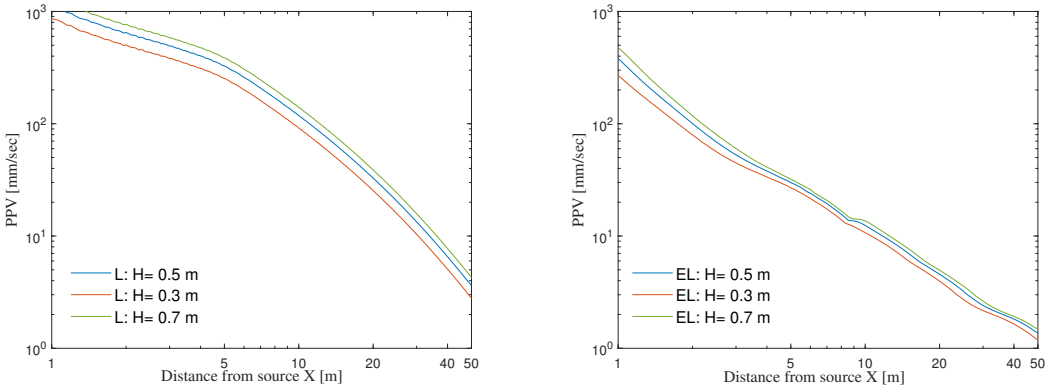
where g is gravity constant, h is the drop height, and α is loss coefficient.

From the present equation, based on the gravity is a constant value, and the loss coefficient depends on the driving equipment. It can be extracted that, the drop height parameter is the most significant one, as it can be easy placed under control. So, Figure 3.25 present the magnitude of ground surface vibrations based on changes on drop height, 0.3, 0.5 (reference case), and 0.7



(a) (b)

Figure 3.24: PPV versus distance for different ram mass: a) linear analysis; b) equivalent linear analysis.



(a) (b)

Figure 3.25: PPV versus distance for different drop height values: a) linear analysis; b) equivalent linear analysis.

meters. As represented in the figure, while the height of fall increases, the vibration level also increases. This observation is valid for the linear and equivalent linear approaches.

3.3.5 Results discussion

According to the previous exposition, a parametric study was presented to expose the influence of the parameters in the phenomenon under study. The parametric study was presented for nine different parameters: shear wave velocity, pile penetration depth, pile diameter, pile material, K cushion, ram mass, drop height, soil profile, and damping ratio. In addition, the previous parameters were assessed through a linear and equivalent linear approach.

In order to get a general reading for the results of the parametric study, two approaches were considered. The first one is to divide the space domain into four zones: zone 1 from zero to 2 m;

zone 2 from 2 to 4 m; zone 3 from 4 to 7 m; and finally zone 4 from 7 to 50 meters. These zones were schematic to expose the parameters influence corresponding to hypothetical soil behavior in plastic, nonlinear and linear zone (see section 2.3.1). The second approach is regarding the assessment method. Two different assessment methods would be present, maximum absolute deviation, MAAD, and area under the curve, AUC. The max absolute deviation method, MAAD, will present the influence of the parameters in every zone based on the max deviation of these parameters from the reference case. In contrast, the method of the area under the curve, AUC, will expose the influence of the parameters per zone based on the area between the curve of the reference case and the curve of the parameter separately.

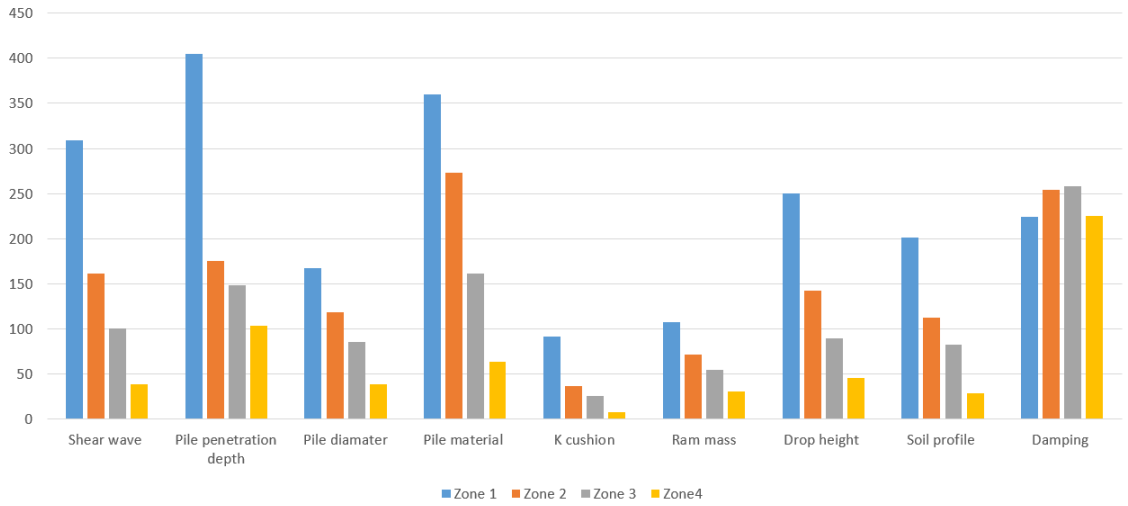


Figure 3.26: Parametric based presenting for MAAD evaluation for the linear approach results.

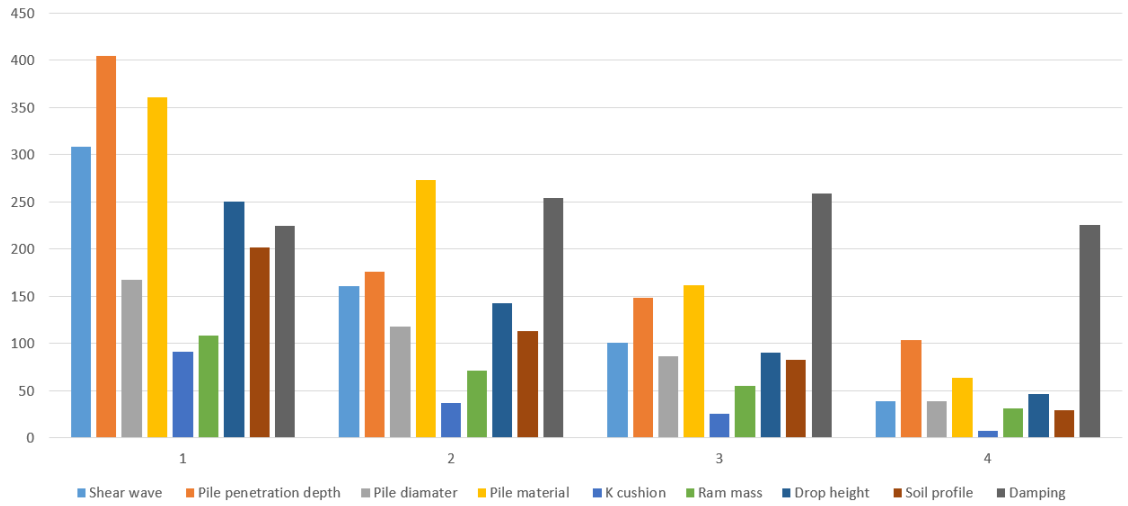


Figure 3.27: Zone based presenting for MAAD evaluation for the linear approach results.

Based on the previous expose, Figure 3.26 presents the parametric analysis for the linear approach model based on the maximum absolute deviation method MAAD. From Figure 3.26, it can extract that the general trending of the influence of the parameters decreases as long the distance increases. However, the damping ratio presented a different behaviour as its influence was roughly equaled in all zones, reflecting the importance of this parameter in the linear approach.

From Figure 3.27 it can be extracted that in zone 4, the damping ratio has the leading influence in the linear approach against the rest of the parameters, while in the first zone, the response of the ground was dominated by six parameters, damping, drop height, pile material, penetration depth, shear wave velocity and soil profile. In addition, it is essential to mention that the stiffness of the cushion has the lowest influence on the ground vibration relative to the rest of the parameters.

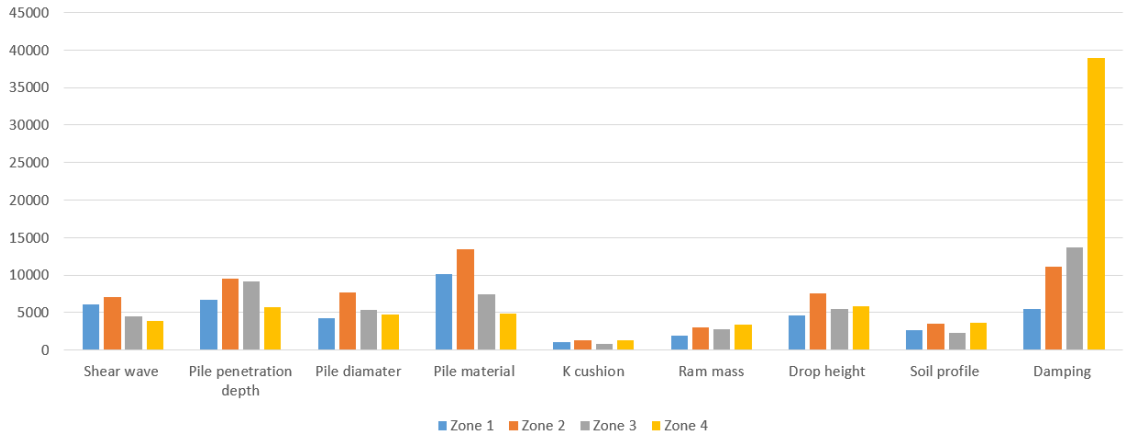


Figure 3.28: Parametric based presenting for AUC evaluation for the linear approach results.

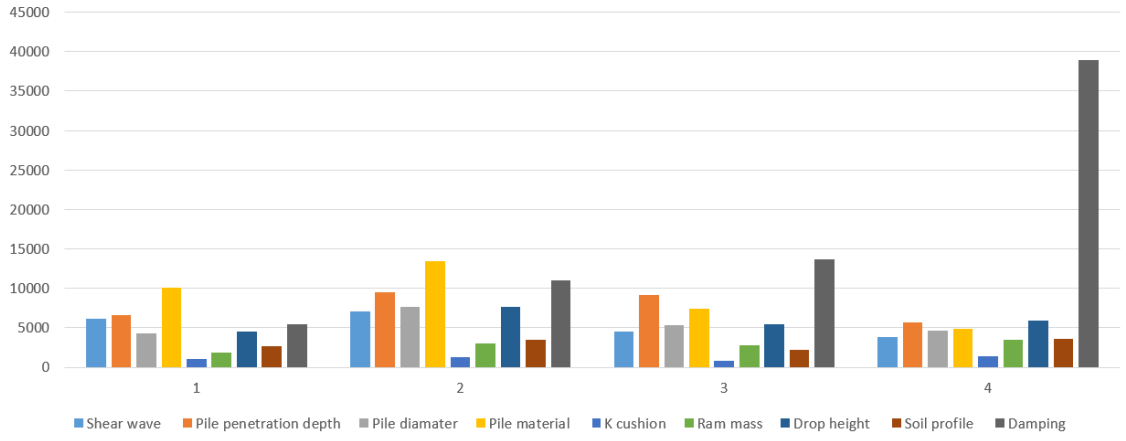


Figure 3.29: Zone based presenting for AUC evaluation for the linear approach results.

Other observations can be extracted from the parametric analysis for the linear approach model based on the area under the curve method AUC. From Figure 3.28 and 3.29, the general trending of the most parameters follow a gamma distribution with peak value located on zone 2. At the

same time, the damping ratio dominates on the ground response. However, the influence of the damping increase as long the distance increase. The significant of the damping in zone 4 can be detected. In addition, in the first zone, the ground response was dominated by six parameters, damping, drop height, pile material, soil profile, penetration depth, and shear wave velocity, as in the MAAD methodology. Finally, the stiffness of the cushion has the lowest contribution on the ground response, as previously mentioned.

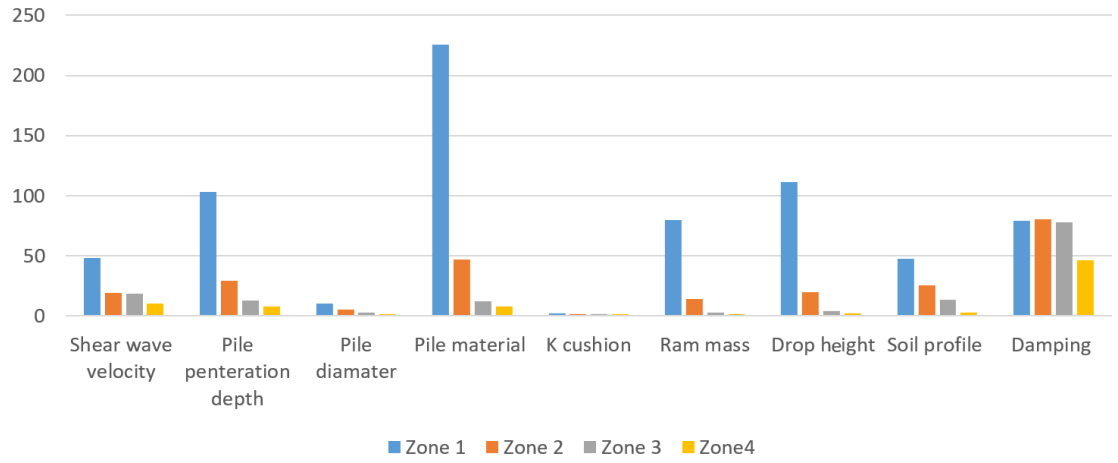


Figure 3.30: Parametric based presenting for MAAD evaluation for the equivalent linear results.

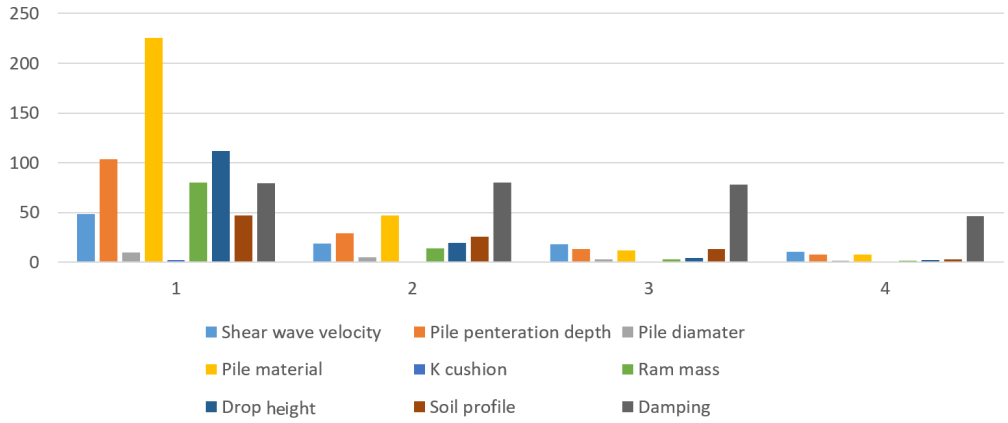


Figure 3.31: Zone based presenting for MAAD evaluation for the equivalent linear results.

Moving on to the second part of the parametric study. From Figures (3.17 - 3.25), the equivalent linear approach and the linear approach were consistent. Figure 3.30 and 3.31 analyze the parametric study modeled by the equivalent linear approach based on maximum absolute deviation. From the figures, it can be extracted that the general trending of the influence of the parameters decrease as the distance increases with an exception for the damping ratio, in recalling for the same behavior in the linear approach. In contrast with the linear approach, the pile material

dominated on the ground surface response. In addition, in the first zone, the ground response was dominated by six parameters, damping ratio, drop height, pile material, penetration depth, soil profile, and shear wave velocity, matching with the linear approach. Finally, the stiffness of the cushion keeps the lowest contribution on the ground response in matching with the linear approach.

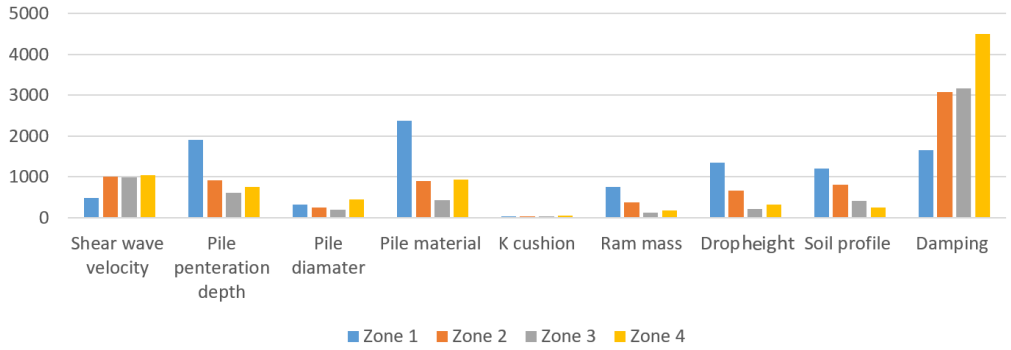


Figure 3.32: Parametric based presenting for AUC evaluation for the equivalent linear results.

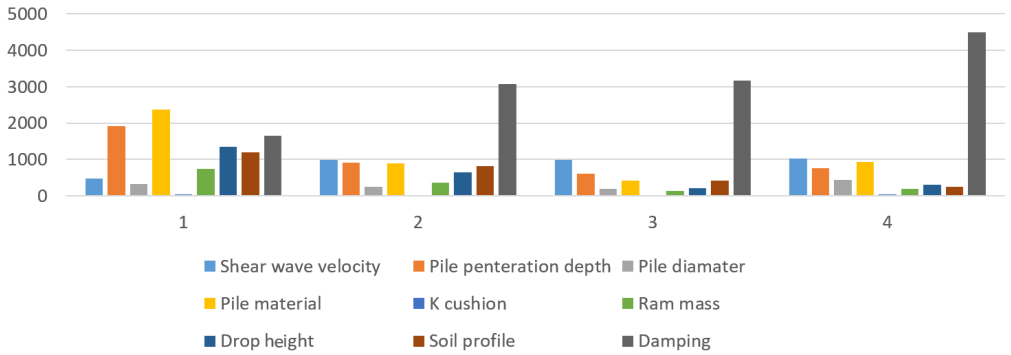


Figure 3.33: Zone based presenting for AUC evaluation for the equivalent linear results.

As the last step, Figure 3.32 and Figure 3.33 analyze the parametric study modeled by the equivalent linear approach based on the area under the curve AUC. From the figures, the damping ratio keeps the domination on the ground response. The stiffness of the cushion keeps the lowest contribution on the ground response. In the first zone, the ground response was dominated by the six parameters, damping, drop height, pile material, penetration depth, soil profile, and shear wave velocity. It is essential to mention that the pile material does not have the same impact on the ground response in contrast to the previous assessment method.

Based on the previous discussion, the parametric study was analyzed from eight different perspectives. It can be extracted that some parameters influence the ground response, while others have less impact. From the analysis, the soil damping ratio contributes significantly on the ground response besides other parameters such as drop height, pile material, penetration depth, shear wave

velocity, soil layering, different ram mass, and pile diameter. The K cushion parameter presents, in eight different comparisons, the lowest influence on the ground surface.

3.4 Final conclusions

Through the context of this chapter, a numerical axisymmetric FEM-PML approach for the prediction of free field vibrations due to impact pile driving was presented. A sub-structuring approach has been used, where the numerical model is split into two main modules: the first one comprises the pile-ground system, and the second one corresponds to the dynamic simulation of the hammer device. The non-linear behavior of the ground is addressed through an equivalent linear methodology, rendering simplicity but reliable methodology.

Given the complexity of the problem under analysis, the proposed numerical approach was placed under a numerical validation process. The numerical validation was performed by exposing the agreement between the presented numerical model (FEM-PML model) and the numerical model reported by Masoumi et al. [55], which presents a satisfying matching.

In the last part of the chapter, a comprehensive parametric study was performed to identify the main variables that interfere with the ground response. The parameters were presented under three main groups, propagation field, pile properties, and hammering properties influence. Two approaches were followed in the parametric study: the first is dividing the space domain into four zones to consider the hypothetical soil behavior, and the second approach is applying two different assessment methods: maximum absolute deviation, MAAD, and area under the curve, AUC. The MAAD and AUC present a similarity in extremes parameters. It can be extracted that the group of the propagation field parameters introduces the most influence on the ground responses. From the analysis, the damping ratio for the soil has the dominating effect on the system's response, while the K cushion exposed a feeble impact. It is also possible to affirm that the hammer and pile properties only significantly affect the ground response for short distances, and a correct estimate of the ground properties will be essential for predicting more considerable distances.

Chapter 4

Experimental and numerical prediction of ground-borne vibrations induced by impact pile driving

4.1 Generalities

The essentiality of an experimental test site lies in realistically demonstrating the physical phenomena. In addition, the connection between experimental and numerical approaches is essential from scientific research perspectives. Thus, the experimental test regarding the vibration induced by the pile driving process and the assessment of the result with the FEM-PML numerical model will be exposed in this chapter. The experiment explanation includes the test proceedings, site characteristics, data treatment, and final results. A site located in south Aveiro, Portugal, was chosen to perform a set of tests. Figure 4.1 shows the location of the test site in the Portuguese territory and Figure 4.2 present a general overview for the site. At the site, a warehouse structure was intended to establish. Therefore, a group of precast piles founded the structure and were installed on-site by the impact driving process.



Figure 4.1: Satellite image for the location of the experimental tests [91].



Figure 4.2: Aveiro, Portugal: photographic record of the experimental test site.

Figure 4.3 shows a plan of the Portugal Geological Map (Plan 16 - A, Aveiro) with the location of the experimental test site. According to this cartography, the following geological formations are found at the test site:

- ***Moderno*** (Aluviões - a): they are represented by sands and silty clays and silt, with shells.
- ***Plio-Plistocénico*** (Dep. De Praias antigas - Q_2^b): These deposits are formed essentially by sand and gravel layers.
- ***Cretácico*** (Arenitos e argilas de Aveiro - C5): they form an upper clay-loam complex, consisting of grayish, greenish, brownish, reddish sandy clays and marbles.



Figure 4.3: Delimitation of the experimental test site on geological map (Plan 16 - A, Aveiro) [92].

According to the geological-geotechnical report that supports the building design, a total of 22 boreholes are covering an area above 27,000 m^2, with SPT tests spaced every 1.50 m in depth. The boreholes have a total length from 6 to 9 meters.



Figure 4.4: SPT equipment in location of borehole S4 [92].

From the SPT tests, it is possible to identify 2 main layers: the first one is alluvium layer with a total depth approximately equal to 3 m, and the second one is a stiffer layer and continuous to the end of the borehole test. In addition, the report shows that a weak backfilling layer exists on the ground surface. This layer contains gravel and leftover plants roots, with brown color and thickness around 1.5 m. As well, the groundwater table was found nearby one-meter depth.

For construction purposes, an embankment layer with an approximate thickness equal to 0.5 m was built after removing the existing backfilling layer. According to the geological-geotechnical report, the embankment material should be gravel or sand belonging to classes A-1 or A-2 of the classification system AASHTO.

From the previous expose, a general overview regarding the site of the experiment was presented. In this chapter, and to obtain the objective initially set, the author presents the low strain analysis; assessment of vibrations induced by pile driving; and experimental validation of the numerical model.

4.2 Low strain test

4.2.1 General overview and objectives

The information provided by the geological-geotechnical report presented before is more qualitative than quantitative for problems involving the propagation of seismic waves. When the numerical modeling of the phenomenon is an objective assumed, it is essential to perform some

geophysics tests, such as Cross-hole or SASW (Spectral Analysis of Surface Waves) tests. However, it was not possible to perform such tests at the present test site. Thus, an alternative process was followed. For that, a low strain test was performed. This test consists of applying an impulse force on the pile head using an instrumented hammer and register the ground response through a set of sensors. The elastodynamic properties of the ground will be then derived by an inverse procedure, comparing the results provided by the numerical model (assuming an initial soil profile) and the experimental data (provided by the low strain test).

This test is concerned with the behavior of the soil under a low strain level. The behavior of soil under low strain levels tends to be linear, and the soil properties will defined through the inverse procedure. The low strain test setup and results will be present in the following.

4.2.2 Description of the experimental activities and main results

For the purpose initially set, a low strain test was performed. This test consists of applying an impulse force on the pile head, previously installed on the ground (section of 270 mm $\times$ 270 mm and penetration depth of 6 m) and using an instrumented hammer (ICP ® Impact Hammer Model No. 086D50). The transient signal was recorded using unidirectional accelerometers placed in a straight line starting from the location of the pile (see Figure 4.5 and Figure 4.6.a).

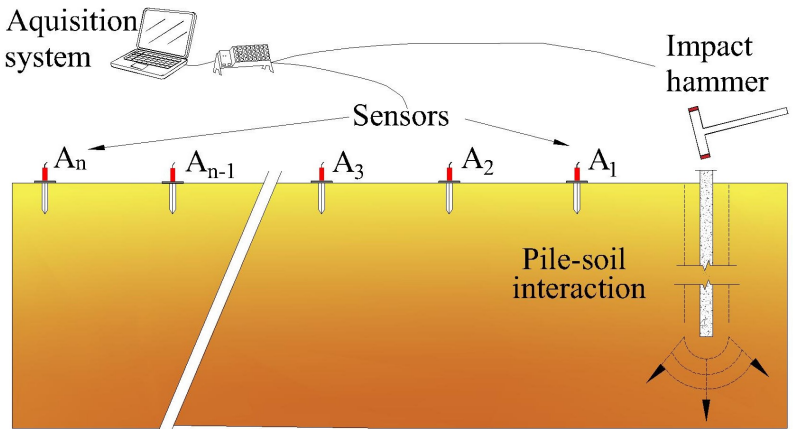


Figure 4.5: Schematic setup for the low strain experiment.

A total of 10 measurement points recorded by 10 accelerometers, with an interval of 4 m was considered (see Figure 4.6.b). An acquisition code developed in LabView was used to record the ground response data and the applied load. A general view of the location of the sensors at the test site can be seen in Figure 4.7. Additionally, the pile used to perform the test is represented by the color blue.

The accelerometers were fixed to cover a distance of 2 m to 36 m from the excitation point, in order to record the ground surface vibration values throughout the space domain.



(a) Vertical and horizontal accelerometers (b) Data receiver

Figure 4.6: Photographic record of low strain test.

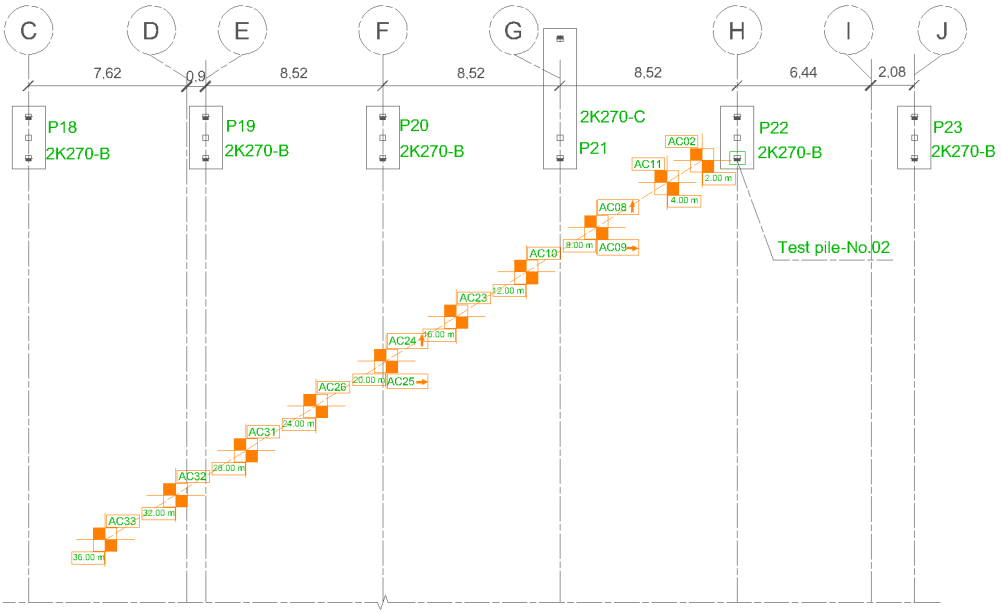
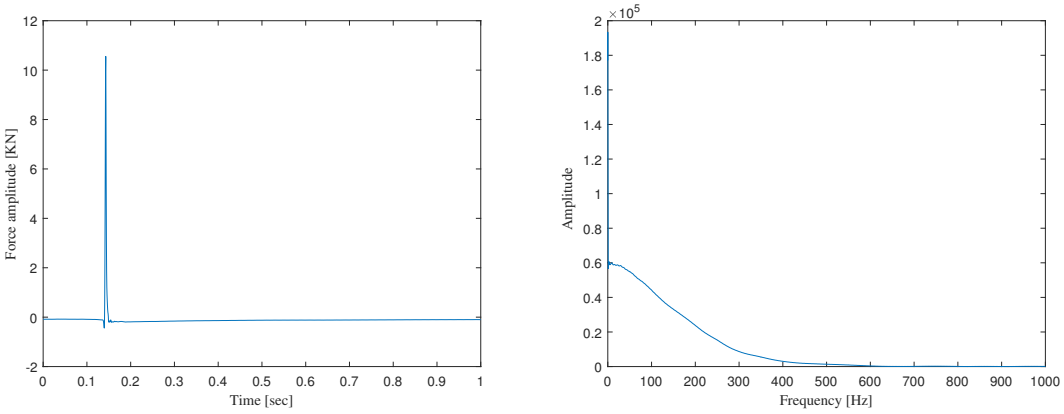


Figure 4.7: Experimental setup used for low-strain test.

4.2.3 Low strain results

From the previous illustration of the experimental procedure, the impact hammer will induce the excitation source and consequent propagation of seismic waves through the ground. In Figure 4.8, the impact force, in time and frequency domain, for a generic impact is presented.



(a) (b)

Figure 4.8: ICP ® impact force: (a) time domain; (b) frequency domain.

While the experimental tests is performing, it can occur that the recorded data are subjected to deviations due to ambient noise from external influences such as surrounding activities, heavy equipment movement, etc. As mentioned before, the applied impacts are at the low strain level, and the ground response will be in the elastic and linear region, so the recorded data values are relatively small. Based on this expose, the recorded data are more likely to be affected by the noise. In order to reduce the influence of the external noise, two different approaches are followed. The first one is to apply several impacts at the pile head to minimize the effect of random noise through the consideration of average registers. The second one is to apply the data treatment proceedings to minimize the effect of noise. The recorded data were subjected to detrending function in the data treatment processes. In addition, Chebyshev type I and Chebyshev type II (low-pass and high pass filter) were applied with cutoff frequency for low pass filter equal to 450 Hz and the high pass filter equal to 3 Hz.

In order to expose the results, Figure 4.9 present the output velocities of observation point in vertical direction. The presented velocities corresponding to 25 applied hammer impacts.

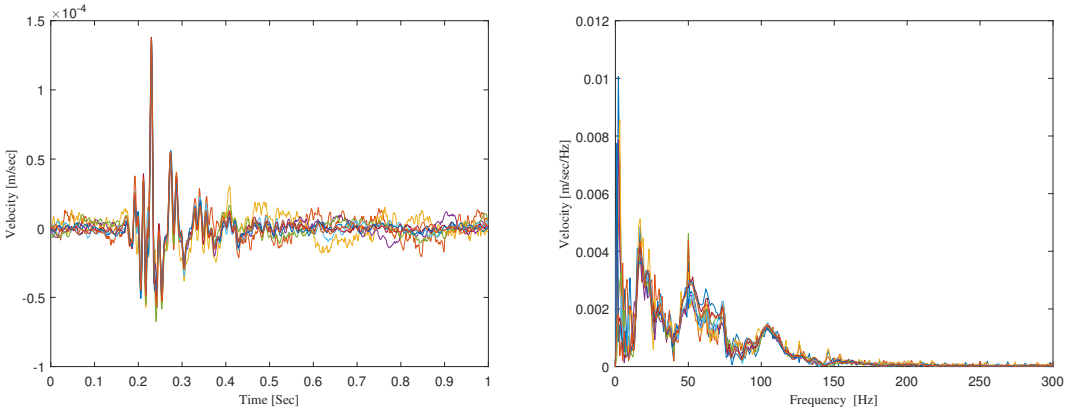
From the recorded results, a transfer function between the free-field response (particle velocity), *channel n*, and the applied impulse, *channel m*, can be computed in the frequency and time domain. The number of impacts (N) was considered. The average transfer function between the receiver and the source, $\widehat{H}_{mn}(\omega, r)$, was computed using:

$$\widehat{H}_{mn}(\omega, r) = \frac{\hat{S}_{nm}}{\hat{S}_{mm}} \tag{4.1}$$

where,

$$\hat{S}_{nm}(\omega) = \frac{1}{N}\sum_{i=1}^{N} \hat{x}_m^i(\omega)\hat{x}_n^{i*}(\omega) \tag{4.2}$$

and ω is the angular frequency, r is the distance between the receiver and the source, $\hat{x}_m^i$ is the



(a) (b)

Figure 4.9: Low strain test - Vertical velocity at 8 m: (a) time domain; (b) frequency domain.

signal in the frequency domain recorded in channel m for impact i and $\hat{x}_n^{i*}$ is the complex conjugate of $\hat{x}_n^i$. Figure 4.10 presents the transfer function in the vertical and horizontal direction at the same observation point in the frequency domain.

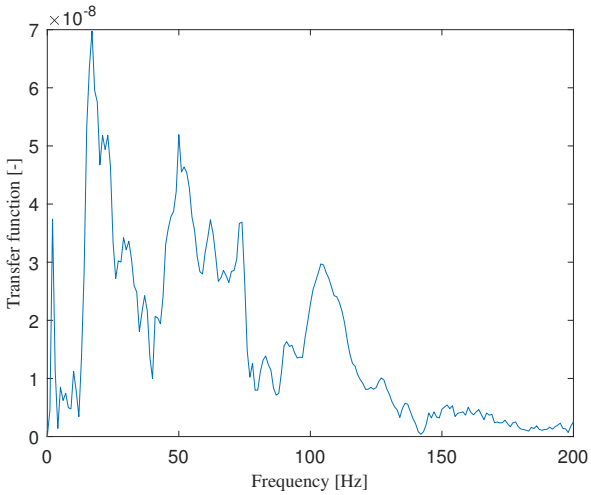


Figure 4.10: Low strain test - Vertical transfer functions at 8 m.

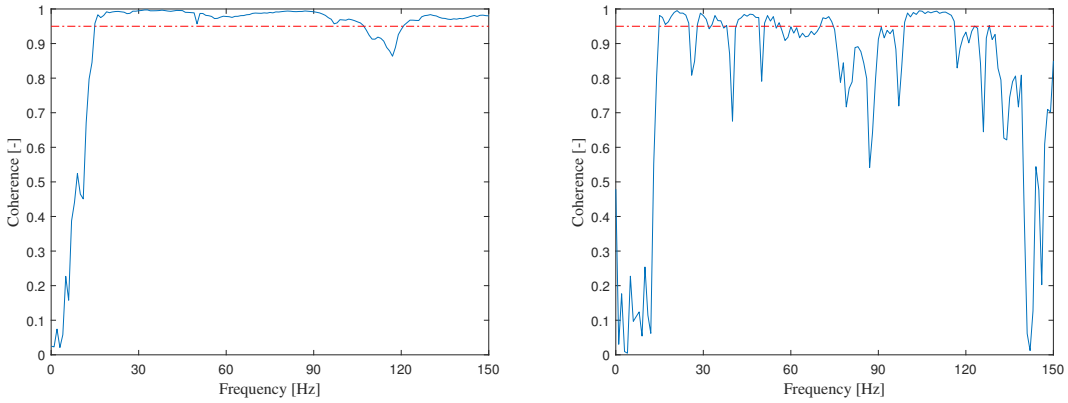
Figure 4.10 present the transfer function in the vertical direction. The vertical transfer function tends to be in decreasing as the frequency increases. This trending was expected as it is in matching with the trending of the velocity in the frequency domain (see Figure 4.9.a).

The coherence is a mathematical term that can be used to assess the relationship between two signals. Commonly it is used to estimate the power of the correlation between input and output for a system, the coherence can be defined as:

$$C_{nm}(f) = \frac{|G_{nm}(f)|^2}{G_{nn}(f) \times G_{mm}(f)} \tag{4.3}$$

where $G_{xy}(f)$ is the cross-spectral density between the free-field response, n, and the applied impulse, m, while $G_{nn}(f)$ and $G_{mm}(f)$ are the autospectral density of n and m signals respectively.

Figure 4.11a presents the function of vertical coherence at 2 meters from the excitation source with presenting a red dotted line for 95% coherence. The figure exposes a sufficient coherence between the applied impact and free-field response. Figure 4.11b represents the vertical coherence in the frequency domain for distances equal to 8 m. The drops in the coherence in the vertical data in distance 2 meters are less than the drops in the coherence in the vertical data in distance 8 meters. These drops happened when a signal for a specific frequency is close to zero, easily a considerable discrepancy between two signals (input and output) would exist, or in other words, drops in the coherence.



(a) (b)

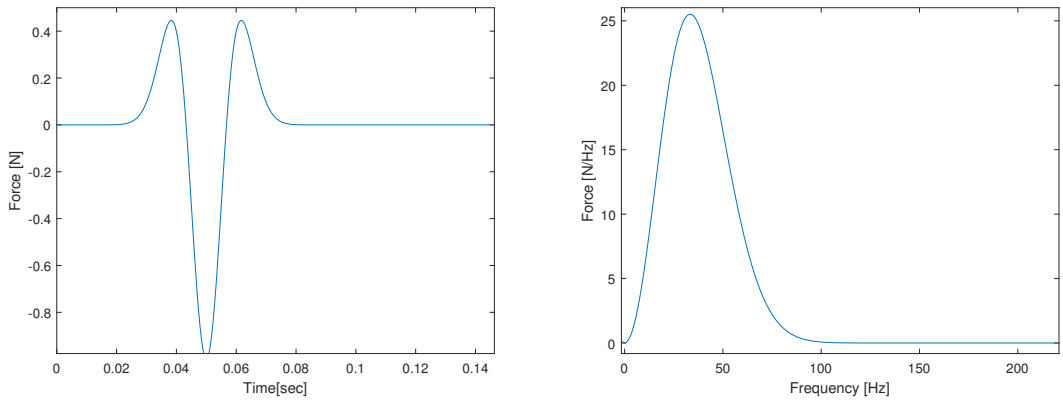
Figure 4.11: Low strain vertical coherence: a) 2 m; b) 8 m.

After computing the transfer functions a Ricker impulse (Equation 4.4) was applied as a loading function in order to get a better visualization of the data in the time domain.

$$R_I(t) = \left[2\left(\pi * \frac{t - t_s}{t_d}\right)^2 - 1\right] e^{-\left(\frac{\pi(t-t_s)}{t_d}\right)^2} \tag{4.4}$$

where, t_s is time shift (t_s=0.05), t_d is the characteristic period (t_d=0.03) and t is the time. Figure 4.12 illustrate Ricker impulse in the time and frequency domain.

From section (2.4), the quantification of the effects of vibration on humans and buildings is usually performed using particle vibration velocity as the primary indicator once there is a correlation between the damage suffered by buildings and the velocity of vibration that the soil experience. From this context, the PPV is an acceptable indicator to characterize the dynamic response of the soil. Figure 4.13 shows the result of applying the Ricker impulse on the transfer



(a) Ricker Impulse Time history

(b) Ricker Impulse Frequency content

Figure 4.12: Ricker impulse function.

functions versus the distance from the excitation point. The author should mention that the result transferred from frequency to time domain, and the peak values of the time history are taken.

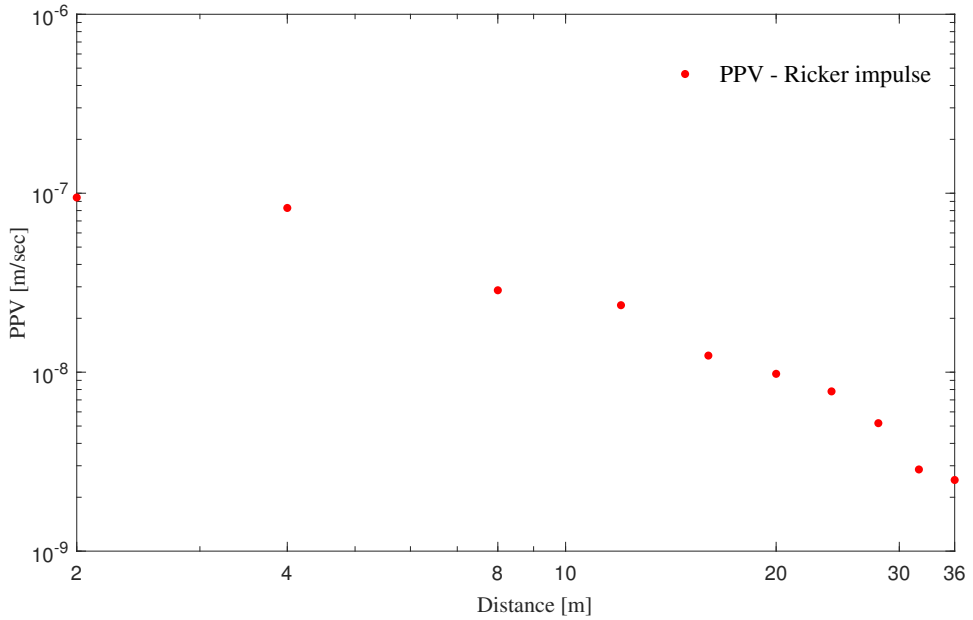


Figure 4.13: Ricker impulse: particle peak velocities PPV vs. distance from the source.

From the previous figure, it could be easily extracted that the value of PPV tends to decrease as the distance from the excitation point increases. That behavior goes in a logical way because of the attendance for the attenuation in the soil material.

4.2.4 Numerical evaluation of the ground properties

As mentioned before, the numerical assessment for the performed test (low strain test) is using to estimate the dynamic ground properties. The numerical evaluation would be performed by comparing the PPV values between a set of combination scenarios for the soil properties and the result presented in Figure 4.13, obtained from the previous section, to find the most matched soil properties. The combination scenarios will be generated based on soil properties rangs, while it will be modeled using the FEM-PML numerical model (see chapter 3).

In connection with exposing the numerical approach, the dynamic soil properties should be involved in a critical way. According to the geological-geotechnical report that supports the building design, the availability of information regarding the dynamic properties for the experiment soil such as shear wave velocity, Poisson's coefficient, damping...etc, are not not available. In addition, the attending for uncertainty regarding the soil geometric profile changes, as just 22 boreholes are covering an area of 27,000 m^2. From the previous context, the elastic properties of the ground were evaluated through a retro-analysis of the ground surface response when excited by an impact load. For that purpose, an extended parametric study was performed in order to get the best fit between experimental and numerical analysis.

According to the geotechnical report, the soil profile has two different stratified layers (alluvium layer and stiffer layer), and there was an additional embankment layer executed on the top of natural ground. Based on the uncertainty of the soil layer thickness, the following range of layers thickness was performed in the parametric study, (see Table 4.1).

Table 4.1: Range of layers thickness in parametric study.

Item	**First layer**	**Second layer**	**Third layer**
Range	0.5-1.5 m	3-5 m	*Infinite*

For the shear wave velocity, the variation is based on the extremes compatible values for the soil material in each layer. For the embankment layer, the shear wave velocity was admitted in the range between 150 - 250 m/sec, for the second layer a range between 60 - 100 m/sec was suggested, and the last layer ranges between 300 - 500 m/sec. In Figure 4.14 the shear waves velocity is presented graphically.

Regarding for Poisson's coefficient, it is assumed equal to 0.49 for the layers under the water table and equal to 0.3 for embankment layer. The mass density 2000 kg/m^3 is for all soil layers.

To clarify, the author should mention that calculating all possible combinations within the given soil properties ranges results in a total of 243 potential soil profiles. However, analyzing this many cases would require excessive computational effort. Therefore, to reduce the computational cost, only the extreme combinations will be considered, reducing the total number of cases to 27.

The numerical transfer function between the response at the ground surface and the dynamic load applied at the pile head is initially calculated by discretizing the pile-ground medium using an FEM-PML mesh. This mesh consists of a total of 47,098 nodes, corresponding to a discretized cross-section of 36x11 meters. As mention before, the problem will be handle in a

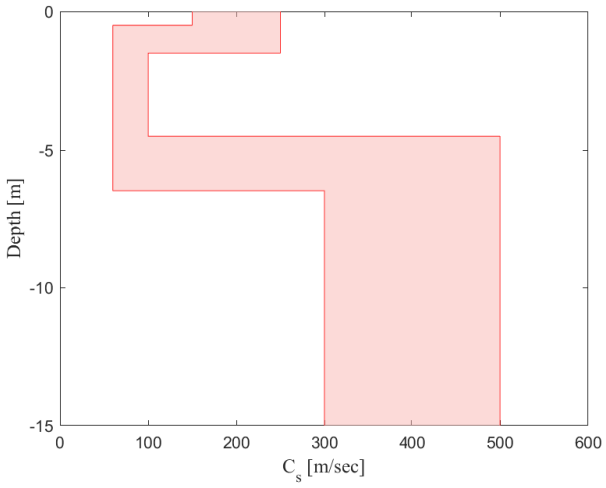


Figure 4.14: Different shear waves velocity profiles.

two-dimensional domain with a pile centre line axis corresponds to the axisymmetric axis. The PML layers are bounding the FEM region, as illustrated in Figure 4.15.

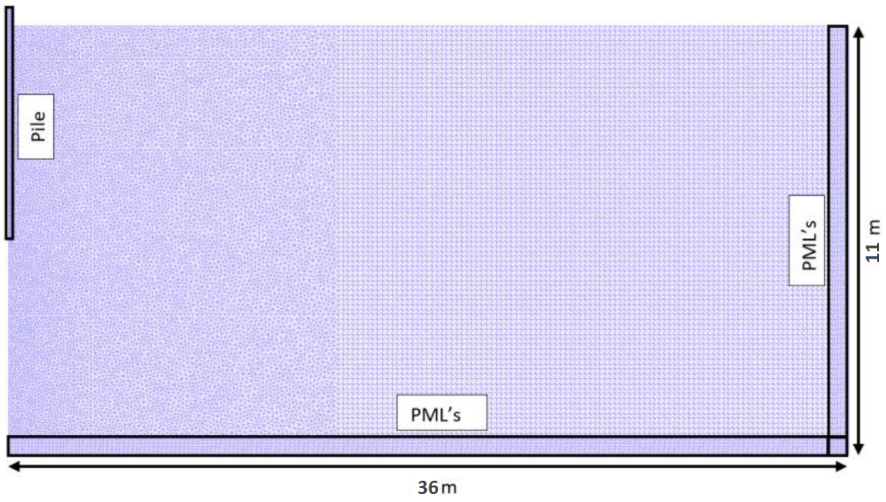


Figure 4.15: FEM–PML mesh: an axisymmetric approach for pile-ground system.

For better visualization for the transfer function, and in order to establish a harmonious comparison between the experimental and the numerical approach, it is mandatory to apply the Ricker impulse, (see Equation 4.4), as a loading function and with the same values of the variables that applied in the experimental approach (see Figure 4.12).

In order to increase the quality of the parametric study, it was performed in two phases. The first phase combines the shear wave velocity, soil layers thickness, and Poisson's ratio. A parametric study was performed to obtain the most matched soil profile properties to the experiment results. As shown in Figure 4.16, envelope provided by the numerical simulations for the particle

peak velocity PPV were presented versus the results of the experimental site test.

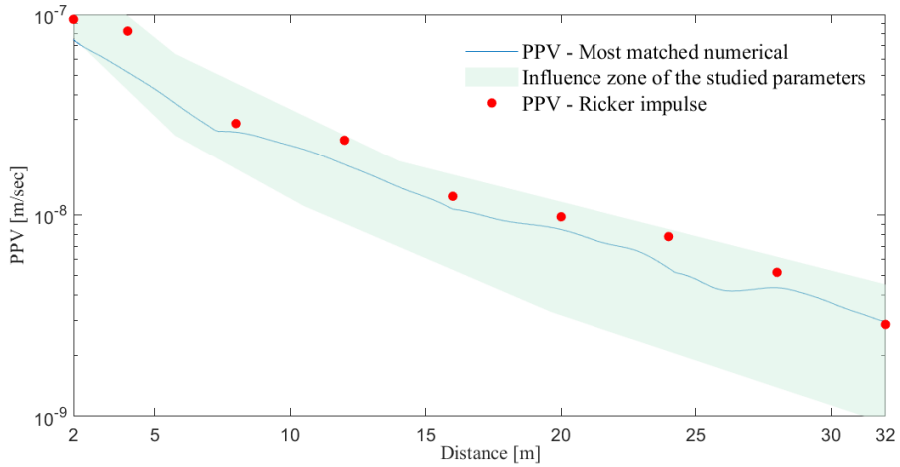


Figure 4.16: Comparison between experimental vs. numerical for the studied parameters.

Based on results from the first phase of combinations, Table 4.2 present the selected properties for the soil from the first phase.

Table 4.2: Soil properties from the first phase.

Layer No.	**Thickness** [m]	C_s [m/sec]	**Poisson Ratio** -	**E** [Mpa]
1	0.5	150	0.3	117
2	3	100	0.49	59.6
3	-	300	0.49	536.4

During the previous analysis, the damping ratio was 3% for all considered scenarios. After getting the most matching properties in the first round, it is necessary to go sequentially to the second phase, which studies the effect of soil damping ratio on the particle peak velocity. Damping is considered one of the essential aspects of a simple evaluation of the dynamic response. So, a wide range of damping ratios was considered, from 1.5% to 4.5%, in the parametric study in order to see the effect of this term on the particle peak velocity and to obtain the proper estimation for the damping ratio in the site, considering the selected soil profile properties from the first phase (see Figure 4.17).

Figure 4.17 present the effect of changing the damping ratio parameter. The green band illustrates the extreme limits for the damping ratio combination. The most matched scenario was plotted versus the experiment result in blue and red colors, respectively.

The study exposed the essentiality of the parameters on the final prediction value. In addition, the parametric study helps to choose the most matching soil properties layer to mitigate the uncertainty about the soil profile. The numerical matched layer properties were tabularized in Table 4.3.

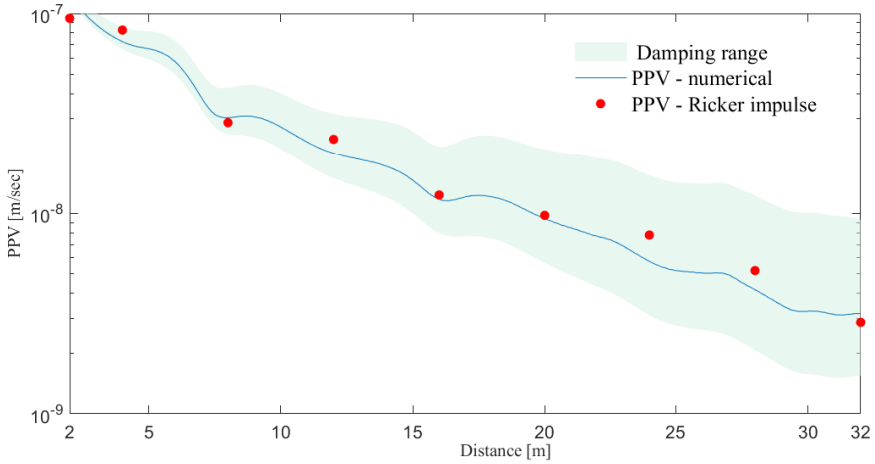


Figure 4.17: Damping influence on the PPV vs. the distance.

Table 4.3: Experimental site soil properties.

Layer No.	**High** [m]	**Damping** [%]	C_s [m/sec]	**Poisson Ratio** -	**E** [Mpa]
1	0.5	3	150	0.3	117
2	3	3	100	0.49	59.6
3	-	3	300	0.49	536.4

From the previous expose, both experimental and numerical approaches are ready for plotting (see Figure 4.18). The figure shows a very good matching between the experimental and numerical approaches. It should be highlighted that this validation comprises distances up to 32 m from the pile. In the large distance domain, effects like soil inhomogeneity are plausible to occur and are not easy to include in numerical analysis, justifying the differences between numerical and experimental results.

For a deeper understanding, the experimental and numerical approaches were compared in the time and frequency domain. The comparison was performed for all recording points to cover the vibration influence zone as long as possible. The covered distance reached up to 32 meters from the centerline of the pile head. The comparison between the experimental and numerical results is illustrated in Figure 4.19. The figures show matching between the numerical and experimental approaches at the short and large distances.

Finally, there is a satisfactory agreement between experimental and numerical results from the previous simulating for the vibrations at the low strain level. However, as expected, the correspondence between them is more significant for the closer observation points to the source. Some discrepancies may occur when the observation point is far from the impact source. For this occurrence, the volume of soil involved in the response can be identified as an additional uncertainty factor in the numerical prediction process. Generally, the previous study gives a good assessment for matching numerical and experimental approaches, and the results can be credible.

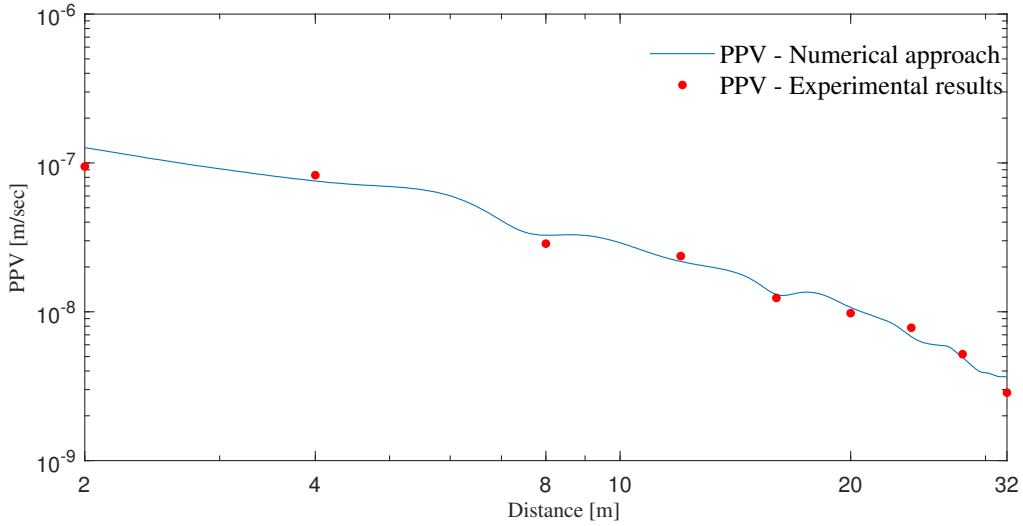


Figure 4.18: PPV numerical vs. experimental approach.

4.3 Vibration induced by impact pile driving

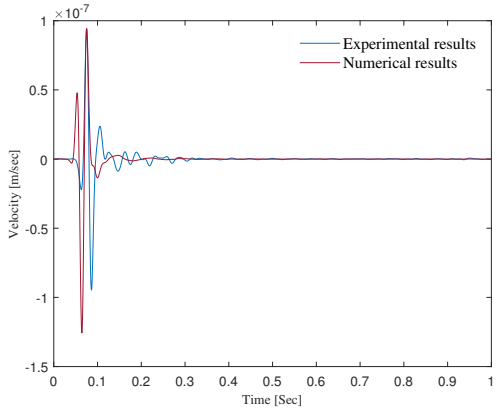
4.3.1 Description of the experimental activities

The vibration register measured for five piles with a cross-section of 270 x 270 mm are presented through the current section. The piles have a total length equal to 6 meters. During the driving operations, the drop height for the hammer is variable, increasing with an increase in the penetration depth (from 15 cm to a maximum of 60 cm). The driving equipment is a Junttan PMx25 equipped with the hydraulic impact hammer SHK110-7. The mass of the hammer ram is equal to 7 tons. The transient signal is recorded at the ground surface for a wide range of points, covering a distance up to 54 meters from the pile and using 12 accelerometers. Figure 4.20 presents photographic records for the driving equipment and the unidirectional accelerometers used to record the ground response.

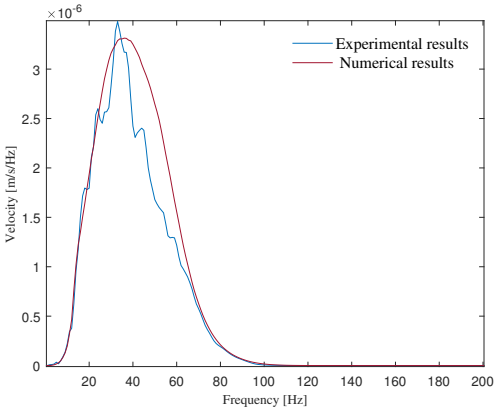
4.3.2 Experimental results

In order to obtain valuable results from the experiment, it is essential to expose the entire proceeding of the data processing. The collected data is a time record of the ground acceleration for each observation point. The ground vibration changes as the pile go deeply into the soil. So, it is essential to go through a splitting process to obtain the acceleration-time recorder for each pile segment. One-meter is the segment length. Figure 4.21 clarifies the difference in the magnitude and frequency for the recorded acceleration, where pile penetration depth and distance from the excitation source change.

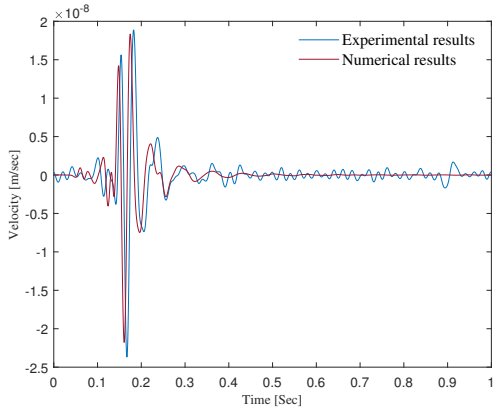
The presented results demonstrate the records of the acceleration for the ground surface in the time domain. From the results recorded at 2 meters, the peak value of acceleration at the



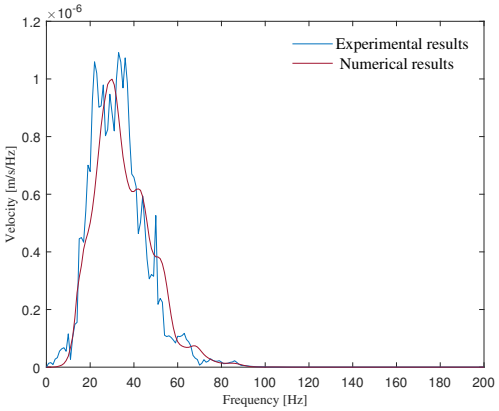
(a) Time domain at distance 2 m.



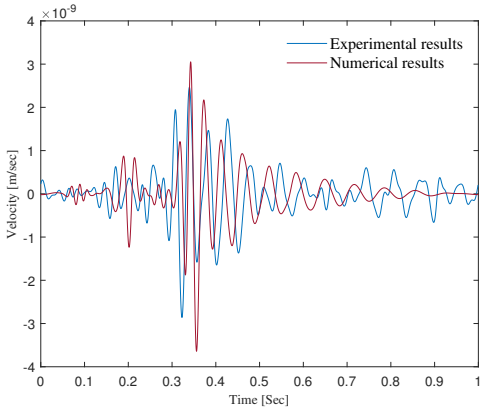
(b) Frequency domain at distance 2 m.



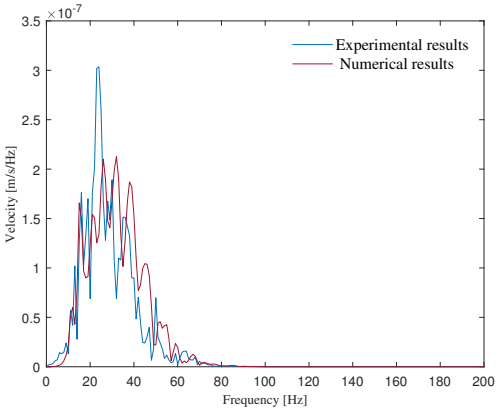
(c) Time domain at distance 12 m.



(d) Frequency domain at distance 12 m.



(e) Time domain at distance 32 m.



(f) Frequency domain at distance 32 m.

Figure 4.19: PPV comparison between the experimental and numerical results- low strain test.

ground surface increases as the pile goes more deeply into the soil induced by the increase in the transmitted energy. However, the same behavior can not detect at 20 meters, justified by the effectiveness of some factors (such as non-linearity, attenuation, layering, etc.) becomes more influential as the distance increase.



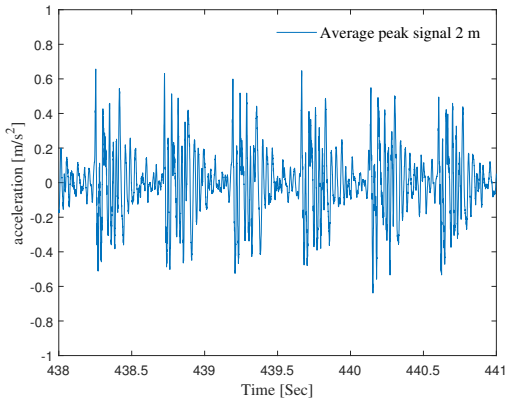
Figure 4.20: Experimental test: a) Pile driving machine, Junttan PMx25 equipped with the hydraulic impact hammer SHK110-7; b) and c) unidirectional accelerometers placed along the ground surface.

As mentioned before, a set of impacts is applied on the pile head from the driving machine. In order to split each impact for each segment, an auto-correlation function was applied to get the impacts in periodic time. In Figure 4.22 all accelerations induced by a hammering process per one segment were plotted together. From the figure, it is possible to verify the prevalence of peak particle acceleration maximums around 0.4 m/s^2 in a vertical direction for 8 meters from the pile.

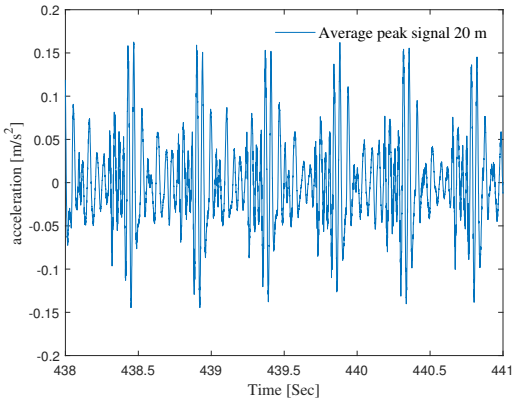
In the previous exposition, the quantification of the effects of vibration on humans and buildings is usually performed using particle vibration velocity as a primary indicator once there is a correlation between the damage suffered by buildings and the velocity of vibration that the soil experience.

In order to obtain particle vibration velocity for the pile driving process, the acceleration register was numerically integrated. Figure 4.23 shows all vertical and horizontal velocities at a distance of 8 meters from the excitation center at a pile penetration depth in the interval 3 to 4 meters. The figure exposes that the maximum velocity at vertical direction is close to 2 mm/sec, while it is 4 mm/sec in the horizontal direction.

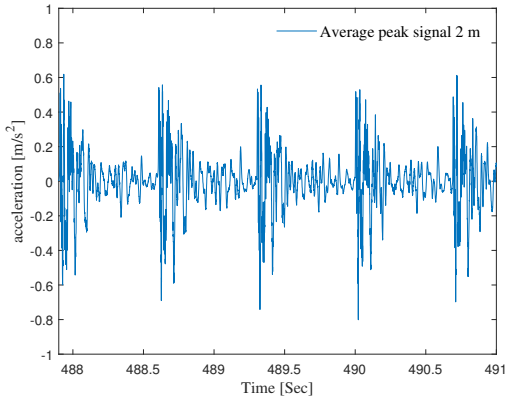
Now, it is possible to systematize the extended set of information through a scatter plot, where the vertical peak particle velocity, PPV, of vibration is plotted as a function of the distance to the vibration source. This result has a high practical value, allowing to identify the maximum peak vibration velocity that can be expected at the level of the building foundations located in the vicinity (if the kinematic compound of the soil structural interaction, SSI, is neglected). In this way, the values evaluated for the set of measurement points and different pile penetration depths are represented through the scatter plot in Figure 4.24. In addition, the maximum permitted PPV values for transient vibrations to prevent damage defined in the Portuguese Standard NP2074 for dominant frequency from 10 to 40 Hz are presented.



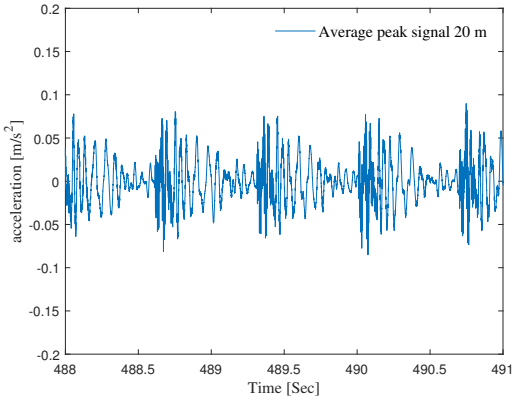
(a) Penetration depth 1-2 ms at 2 m.



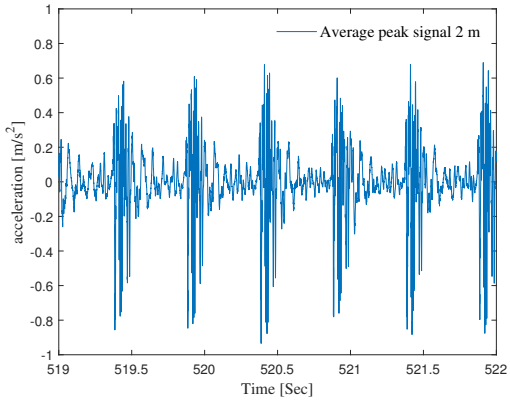
(b) Penetration depth 1-2 ms at 20 m.



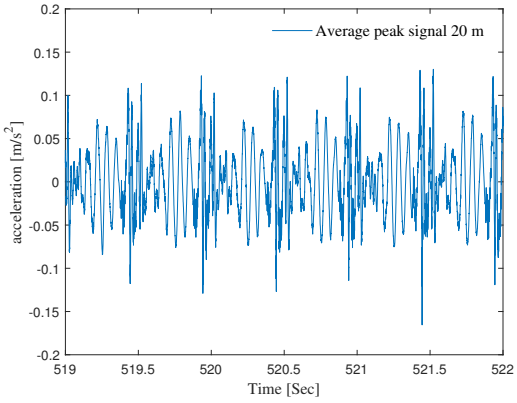
(c) Penetration depth 3-4 ms at 2 m.



(d) Penetration depth 3-4 ms at 20 m.



(e) Penetration depth 5-6 ms at 2 m.



(f) Penetration depth 5-6 ms at 20 m.

Figure 4.21: Vibration induced by pile driving activities in different penetration depth and distance from excitation source.

From the results achieved, it is possible to verify the prevalence of peak particle velocity maximums around 25 mm/s in the vicinity of the pile and 2 mm/s for a distance of 8 meters from the pile. This value is reduced to below 1 mm/s for distances higher than 28 meters.

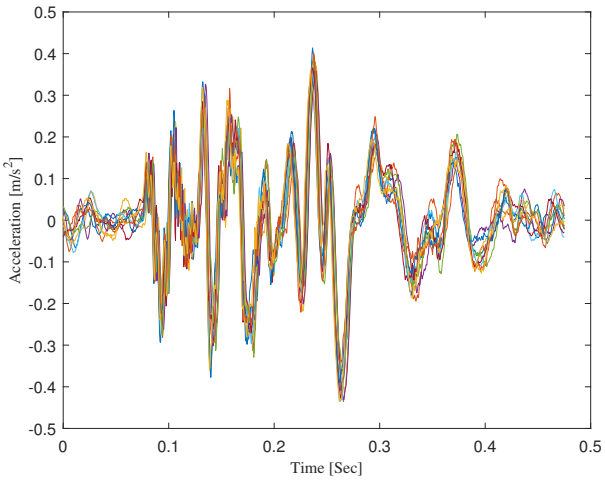


Figure 4.22: Hammering respond accelerations at distance 8 meter and depth 3-4.

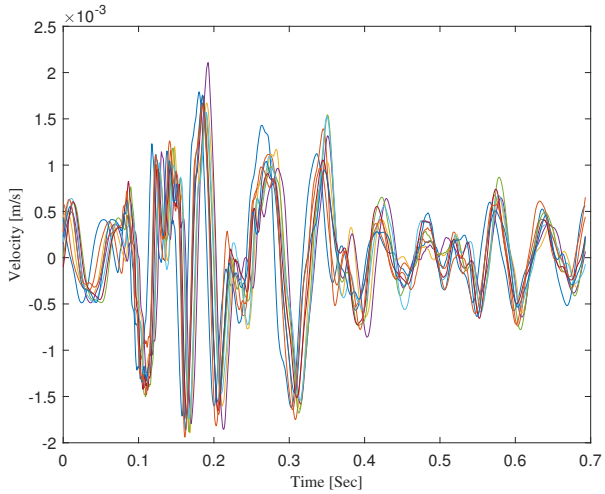


Figure 4.23: Hammering respond velocities at distance 8 meter and depth 3-4.

Despite the interest of this information, the PPV values should be analyzed considering the range of the most relevant frequency content of the response. The standards that address this issue impose limits on the peak value assessed near the foundations of buildings; this limit value is variable, according to the dominant frequency in the velocity spectrum. Thus, and considering the results evaluated during the driving of one of the recorded piles with a cross-section of 270x270 mm, Figure 4.25 shows the frequency content for three points placed at different distances from the pile, namely: 8, 16, and 24 meters. In the figure, the corresponding time-domain records are also presented.

As exposed by the curves illustrated in the previous figure, and in addition to the expected attenuation of peak velocity values with increasing distance, there is an attenuation of the response

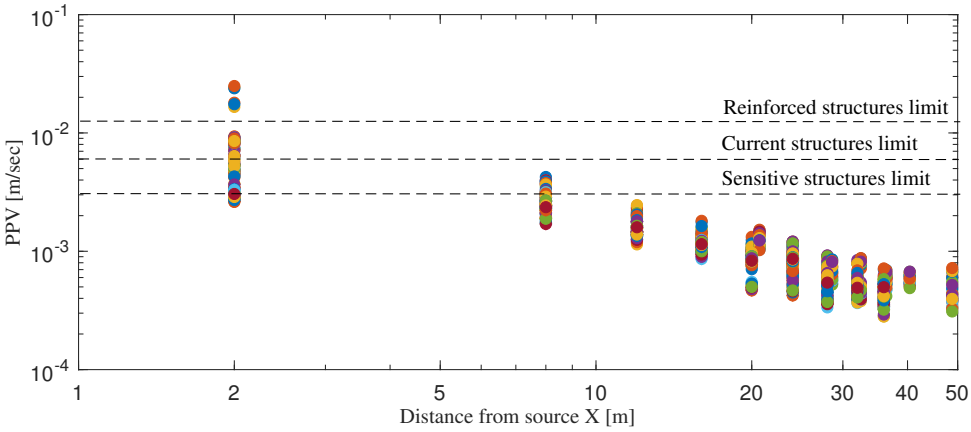
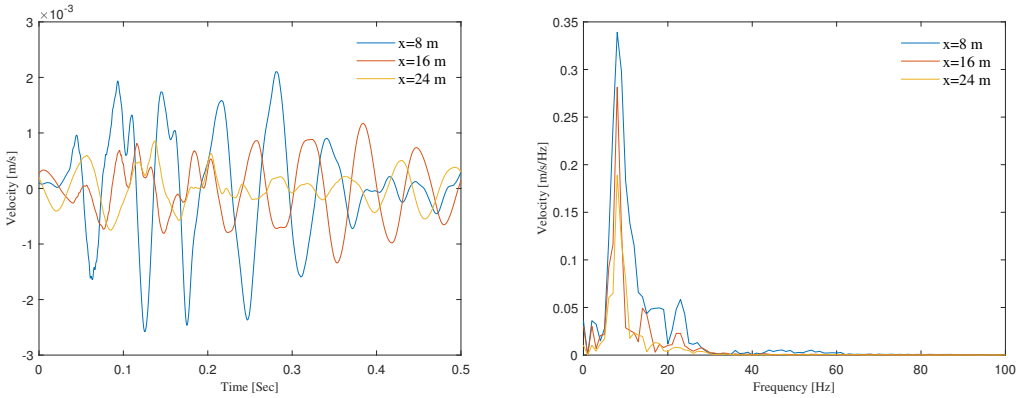


Figure 4.24: Pile driving experiment results: PPV vs. the distance.



(a) (b)

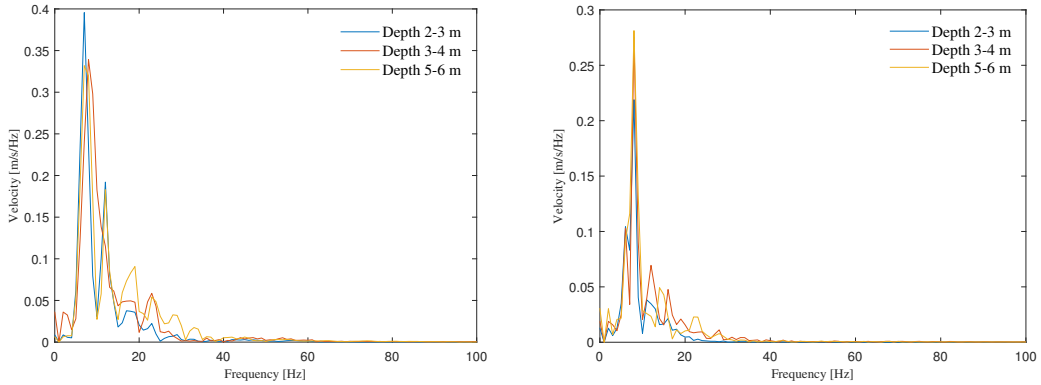
Figure 4.25: Experimental vertical vibration velocities measured at different distances from pile and considering the penetration depth in the interval 5-6 m: a) time-history; b) frequency content (r=8, 16 and 24 meters).

at higher frequencies. This fact derives from the influence of the material damping in the higher frequency content attenuation.

The pile penetration depth is an essential aspect that needs more emphasis corresponding to the evolution of the frequency content with the increasing. Thus, Figure 4.26 presents three curves that correspond to pile penetration depths in the intervals [2-3] m, [4-5] m, and [5-6] m, for distances equal to 8 and 16 meters form excitation source. As shown, no significant variation of the frequency content can be observed with increasing driving depth.

4.3.3 Experimental validation of the numerical model

As initially established, experimental validation of implemented numerical models is one of the main objectives of this chapter. Therefore, the initially defined pile-ground elastic dynamics prop-



(a) (b)

Figure 4.26: Frequency content of the experimental vertical vibration velocities considering different distance from pile: a) r=8 m; b) r=16 m.

erties are used in the numerical simulation of the case study. Considering the limitations of the axisymmetric formulation, which cannot model the actual shape of the pile (rectangular cross-section 270 x 270 mm), it is considered an equivalent circular pile cross-section.

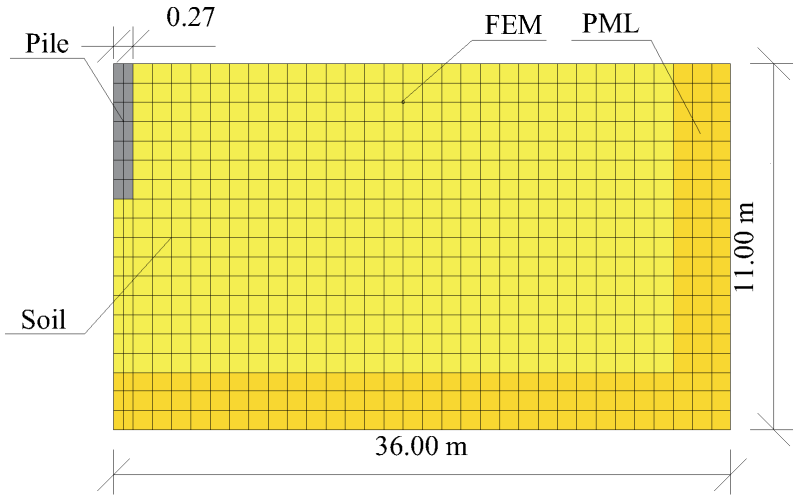


Figure 4.27: FE–PML mesh adopted to model the pile-ground system.

In general, pile-soil media is discretized by a FEM-PML mesh with triangular elements of six nodes (54372 nodes total), corresponding to a 36x11 m^2 discretized section, as presented in Figure 4.27.

In terms of the driving device, the following parameters were adopted: $m_a = 850$ kg, $m_r = 7000$ kg, $k_c = 2 \times 10^6$kN/m and $c_c = 6\text{x}10^5$kNs/m. Regarding the impact velocity of the ram, v_0, a short explanation is necessary: theoretically, the potential energy, $E_p = m_r gh$, is assumed to be equal to the kinematic energy, $E_c = m_r v_0^2/2$. From this relation, it is possible to relate the drop height hammer (experimental data) with the velocity of the ram. However, in practical

applications, it is necessary to take into account the energy loss during the process of hammer fall. Thus, a loss coefficient, α, about 0.75 was considered: $E_c = \alpha E_p$.

To obtain a reasonable representation of the comparison, a penetration depth would be considered in the numerical model equal 4.50 m. Figure 4.28 shows the comparison between experimental results for all piles for interval 4-5 m penetration depth versus the numerical linear and equivalent liner approach.

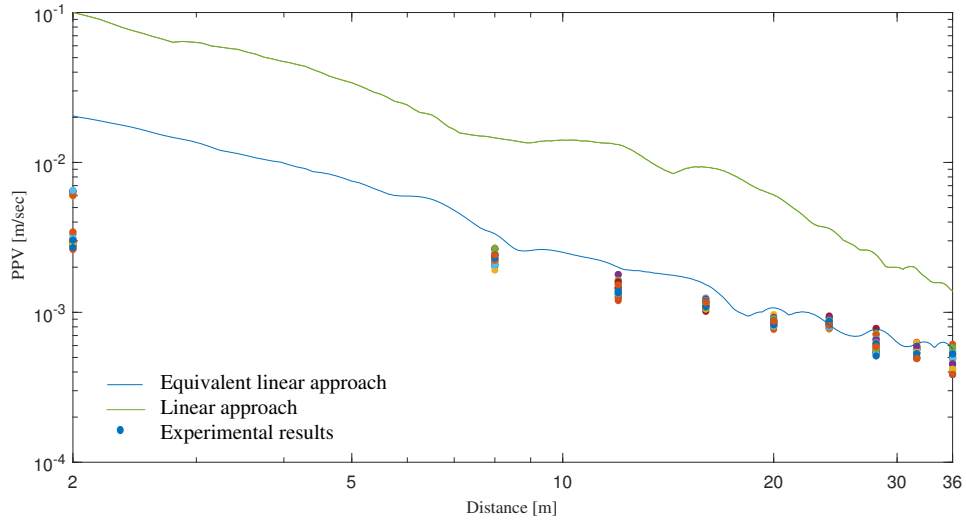


Figure 4.28: Experimental vs. numerical results at penetration depth 4-5 m.

By analyzing the results shown in Figure 4.28, a reasonable match was reached between the experimental data and the numerical results obtained using the equivalent linear approach in the large distance. However, in short distance, the numerical model give more conservative results than experiment. The full-time record shown in Figure 4.29 illustrate a variance between experimental and numerical results. The frequency domain analysis is also essential for a deeper understanding of the problem. Following this goal, Figure 4.30 and Figure 4.31 show a comparison of experimental and numerical results (equivalent linear analysis) in the frequency domain. As illustrated, an acceptable agreement is reached for observation points larger than 2m.

4.4 Experimental evaluation of the influence of the ground stiffness on ground vibration

4.4.1 Vibration levels for two test sites

Additional to the previous experimental results, and with the aim of evaluation the influence of the ground stiffness on the ground vibration level, an additional experiment test site report by Colaço et al. [93] is presented. INEGI test site, is located near the city centre of Porto (Portugal). Figure 4.32 exposes an over view for INEGI site.

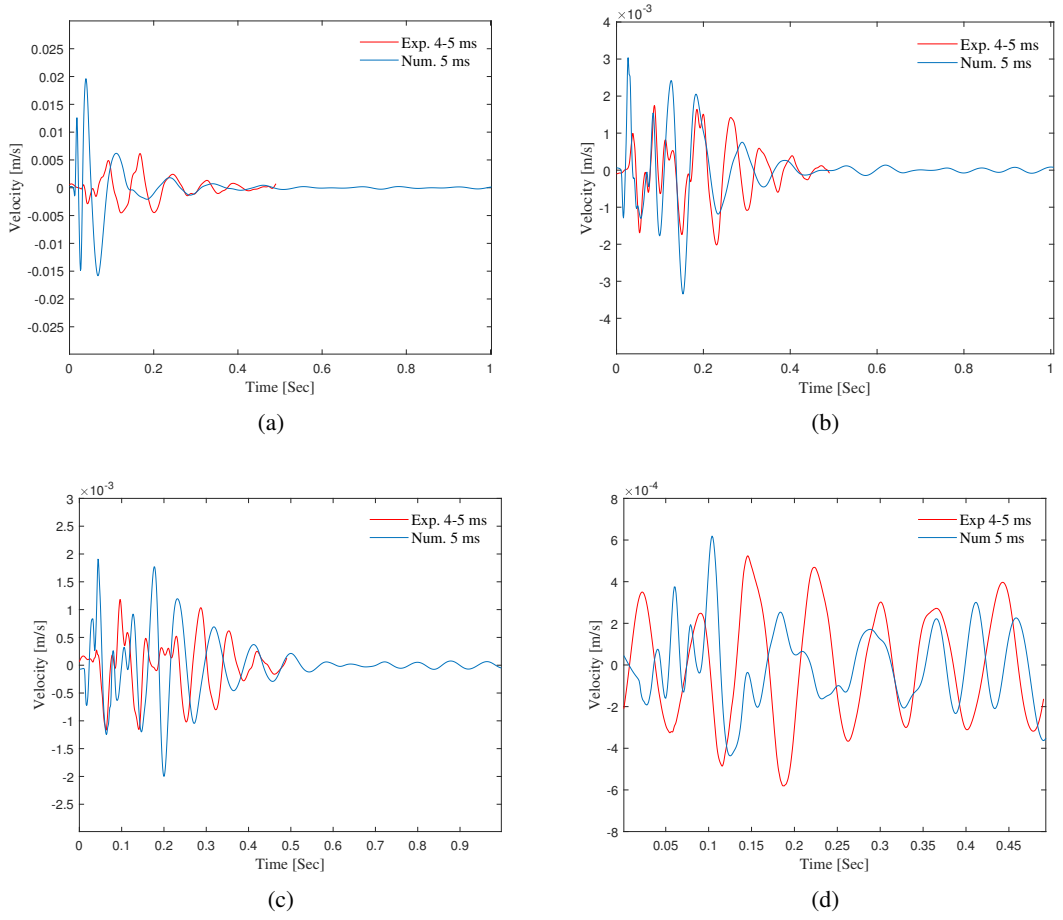


Figure 4.29: Time domain comparison for a point located at: a) 2 m; b) 4 m; c) 12 m; d) 32 m.

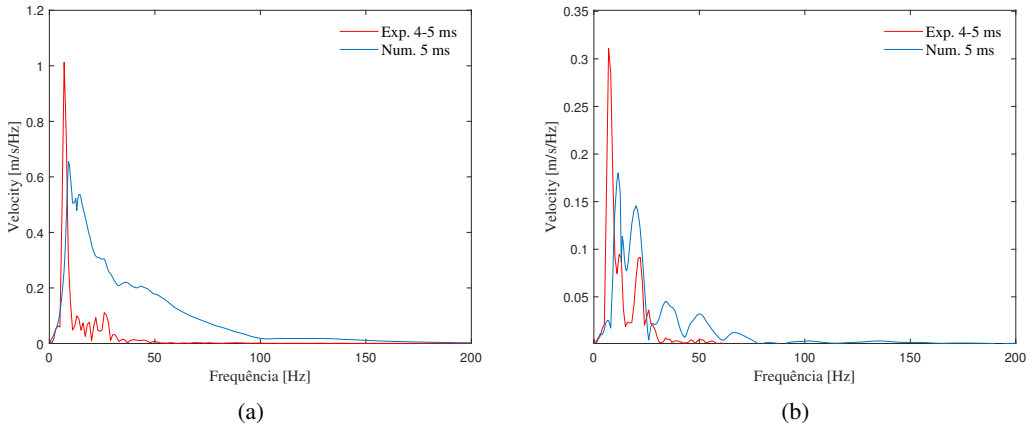
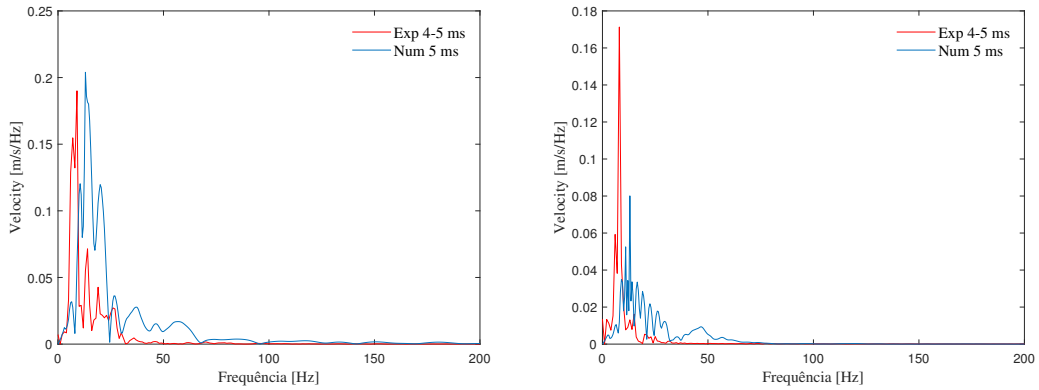


Figure 4.30: Frequency domain comparison for a points at: a) 2 m; b) 4 m.

According to the geological-geotechnical report that supports the building design, a total of seven boreholes were made, with SPT tests spaced at a depth of 1.50 m. Moreover, soil samples were collected to perform a few laboratory tests for the classification and identification of the soil.



(a) (b)

Figure 4.31: Frequency domain comparison for a points at: a) 12 m; b) 32 m.



Figure 4.32: Porto, Portugal: photographic record of the experimental test site [93].

Essentially, the site is geologically formed by a granite residual soil, with the stiffness and strength increasing along the depth. The groundwater table was found approximately 8 m deep (from the base of the excavation). Figure 4.33 presents the SPT N-value for the soil profile.

Concerning the foundations of the buildings, INEGI test site comprises a total of 160 quadrangular piles with two distinct sections of 400 x 400 mm and 350 x 350 mm. These piles are made of precast concrete and they have a total length varying from 8 to 15 meters. The driving equipment corresponds to a Junttan PMx25 equipped with the hydraulic impact hammer SHK110-7. The mass of the hammer ram is equal to 7 tons.

Looking for a systematization of the extended set of information recorded, it should be recall the results from Aveiro site to put in comparison with the INEGI site. Figure 4.34 presents the compilation of the maximum peak particle velocity according to the horizontal distance to the pile. The results from both sites are represented on this figure. It should be clarified that a separation of

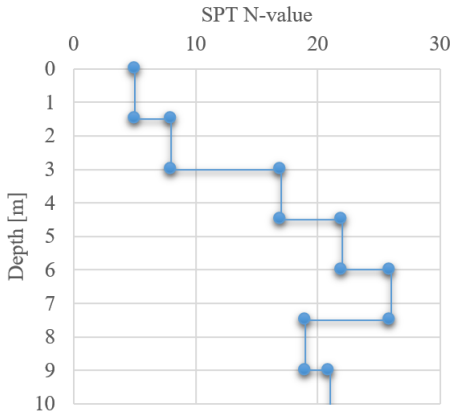


Figure 4.33: SPT N-value for INEGI site.

the impacts was made for each meter of driving, allowing to adopt average values for each driving interval.

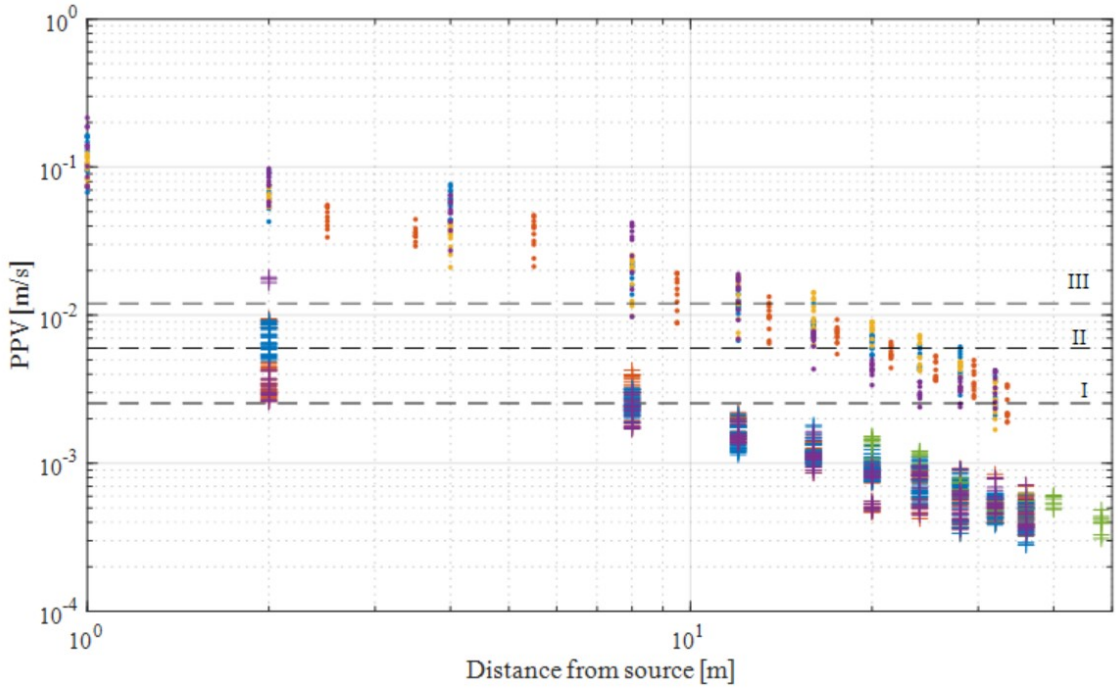


Figure 4.34: Maximum peak particle velocity versus distance from the pile (circle marker - INEGI test site: Pile 1 - o; Pile 2 - o; Pile 3 - o; Pile 4 - o; cross marker - Aveiro test site: Pile 1 - +; Pile 2 - +; Pile 3 - +; Pile 4 - +) and maximum permitted PPV values for transient vibrations to prevent damage defined in the Portuguese Standard NP2074 (I-sensitive structures; Il-current structures; III-reinforced structures).

In a general way, the results from both sites present a typical behaviour characterized by a continuous attenuation of vibration levels as the distance to the pile increases. This effect is expected given the increase of the geometric damping with the distance. On the other hand, it is observed a clear difference on the vibration amplitude levels between both sites, being the

vibration levels recorded at the INEGI test site much higher. Among others, it can be pointed out as the main reason for the differences observed the geotechnical properties of the ground. In fact, Aveiro test site is located in an alluvial deposit, characterized by the presence of soft soils in the first meters. Contrarily, the INEGI test site is geologically formed by a granite residual soil with higher stiffness. Furthermore, the differences in the impact energy for both cases has is contribution over the vibration levels.

Additionally, the frequency content of the soil response is presented. In this kind of phenomenon, the PPV values should be analysed taking into account the most relevant frequency content of the response. As previously expressed, the limits on the peak vibration values imposed by the vast majority of the standards/guidelines are variable according to the dominant frequency in the velocity spectrum. Thus, Figure 4.35 shows the frequency content of the ground considering the driving of two representative piles (one from INEGI test site and another one from Aveiro test site), addressing two variables: i) variation of the distance from the pile (8, 20 and 32 m); ii) variation of the pile penetration depth.

As evidenced by the curves illustrated in the previous figure, and in addition to the expected attenuation of peak velocity values with increasing distance, there is an attenuation of the response at higher frequencies with the increase of the distance between the observation point and the pile. The larger influence of the material damping in the higher frequency content of the response justifies this observation. In what concerns to the evolution of the frequency content with the increase of the pile penetration depth, although with a variable hammer height-of-fall, no significant variation of the frequency content can be observed with increasing driving depth.

4.5 Numerical study of a mitigation measure

4.5.1 General overview of the numerical model

In situations where there is potential to exceed the legally established limits of vibrations, the design of simple and practical mitigation measures can be an alternative to a simple impediment to the use of a given technique, as pile driving. Numerical modelling can be a powerful tool to perform this task. Thus, FEM-PML numerical model, previously developed by Colaço et al. [4], is used.

With the aid of the equivalent linear FEM-PML numerical model, the efficiency of different mitigation measures is analysed. For that, and given the higher vibration levels found at INEGI test site, the present case study adopts its geomechanical characteristics. Laboratorial characterization of the soil indicates a mass density, ρ, close to 1900 Kg/m^3, a hysteretic damping factor, ξ, of 5% is admitted for the first layer and a value of 2.5% for the remaining layers. According to the geotechnical characterization, the elastodynamic properties of the ground are expressed in Table 4.4, where the variable h stands for the thickness of the layers.

As a general overview, the pile-ground medium is discretized by a FEM-PML mesh, corresponding to a discretized cross-section of 35x25 m2. The PML layers are bounding the FEM

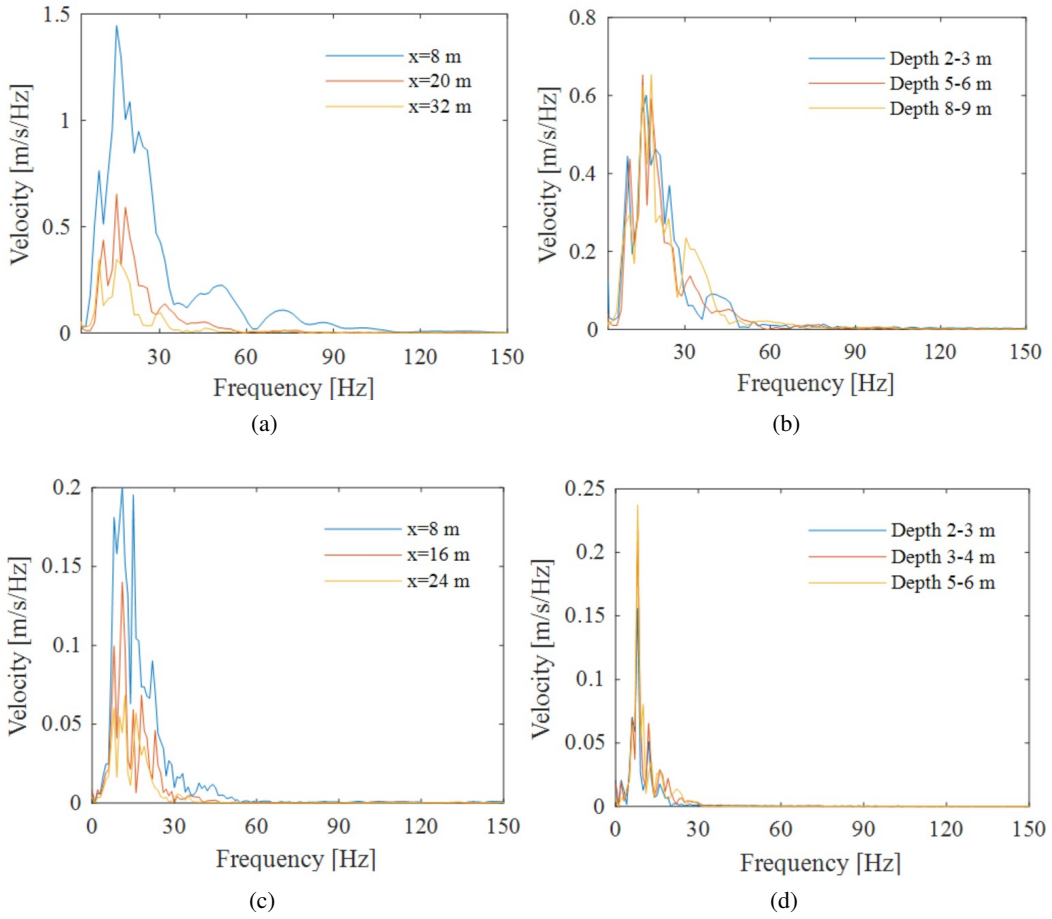


Figure 4.35: Frequency content of the experimental vertical vibration velocities considering: a) and c) different distances from pile and for a penetration depth in the interval 5-6 m; d) and b) considering different pile penetration depths and an observation point at 20 m from pile (upper: INEGI test site; down: Aveiro test site).

Table 4.4: Elastodynamic properties of the pile-ground system.

Element	h[m]	C_s[m/s]	C_p[m/s]	ρ[kg/m3]	ξ[%]
Soil	2	180	330	1900	5
	6	230	415	1900	2.5
	5	330	1900	1900	2.5
	inf	620	1900	1900	2.5
Pile	L = 12 m	2284	3560	2500	1

region, as illustrated in Figure 4.36. Given the limitations of an axisymmetric formulation, in which it is not possible to model the true geometry of the pile (rectangular section 400 x 400 mm), an equivalent circular pile section was considered. The elastodynamic properties of the pile can be found in Table 4.4.

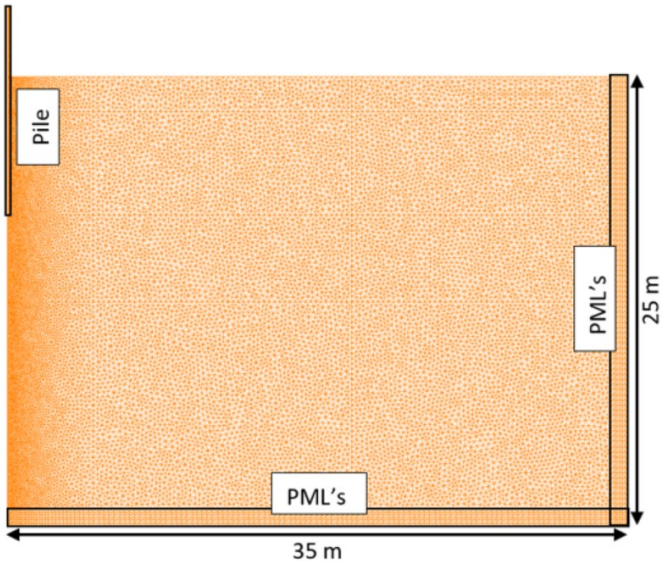


Figure 4.36: FEM–PML mesh adopted to model the pile-ground system.

4.5.2 Trench solution

Aiming to compare the response of the soil free surface before and after the construction of an open trench, two variables were analysed: variation of the trench depth (D) and distance fom the pile to the trench (x). It should be mentioned that, it is intention the design of a measure that can be easyly implemented on the site, with no significative cost for the project. A schematic representation of the trench solution is presented in Figure 4.37.

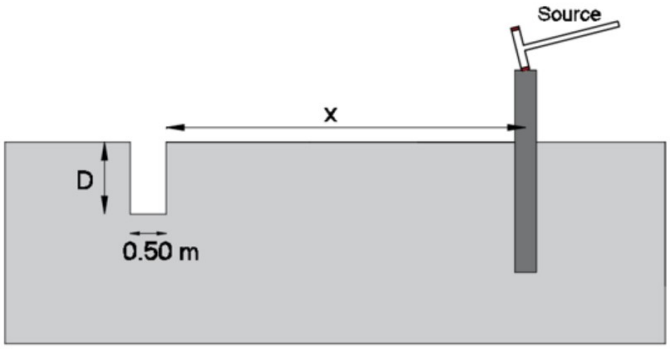


Figure 4.37: Geometry of the trench solution.

Considering the particular case of the pile penetration depth equal to 2 meters, Figure 4.38 presents the experimental results of one of the monitored piles with a cross-section of 400x400 mm. In terms of experimental results, the curve shown in this figure was obtained for pile penetration depths in the interval [1.5-2.5] m, assuming a mean value. In the same figure, the numerical

results provided by equivalent linear approach are also plotted for different cases: Case 0 – Reference case (without trench); Case 1 – Trench (D=1 m; x=5 m); Case 2 – Trench (D=2 m; x=5 m); Case 3 – Trench (D=1 m; x=10 m).

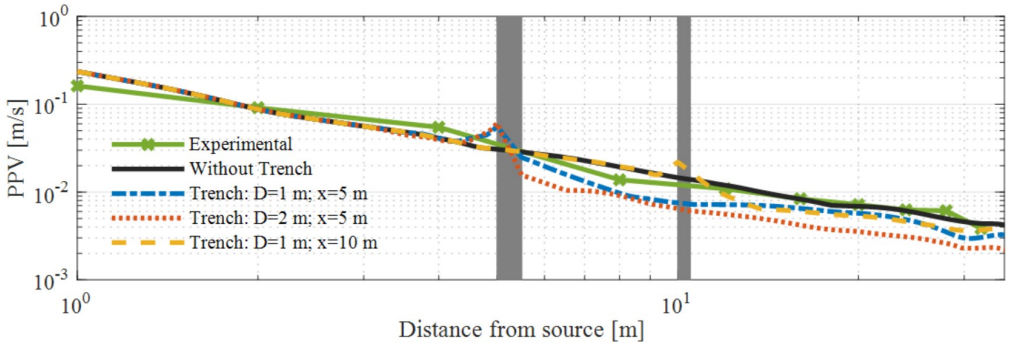
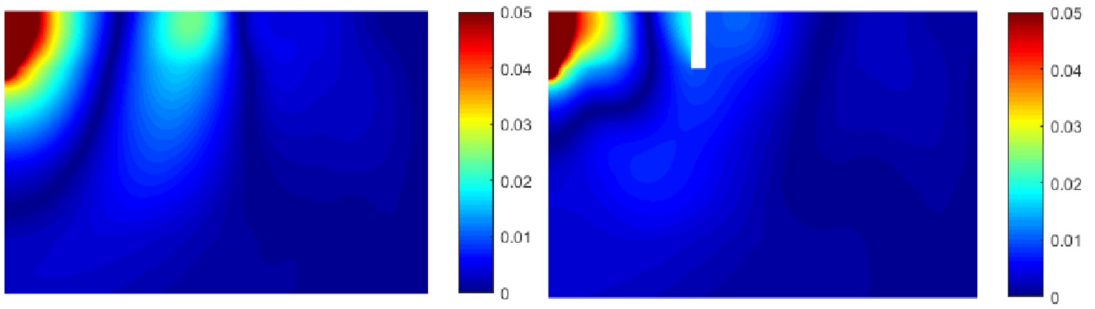


Figure 4.38: Comparison between experimental and numerical results (before and after the introduction of the trench) - PPV vs distance - for a pile penetration depth of 2 m.

Analysing the results depicted in Figure 4.38, it is possible to point out two main conclusions: i) a very good match between the experimental data and the numerical results provided by the equivalent linear approach. The overall behaviour of the system is satisfactorily captured by the numerical model; ii) the introduction of a trench induces a significant attenuation on the PPV. This attenuation appears immediately after the trench and increases with the increase of the trench depth. However, from a practical standpoint, the maximum depth of open trenches is conditioned by the geotechnical properties in order to avoid its walls to collapse. Given the temporary nature of this measure, and taking into account the characteristics of each site, adopting sloped walls for the trench could be a strategy to allow the increase the depth without compromising the trench's performance.

To visualize the effect of the trench on the wave's propagation, Figure 4.39 shows the comparison of the norm of the particle velocity for a given instant temporal (75 milliseconds) for Case 0 (without trench) and Case 2 (D=2 m; x=5 m) under analysis. As can be seen, the presence of the open trench introduces a wave barrier, with a clear perception of the interference in the waves' propagation. As a small note, the amplification of the particle vibration just before the trench, as shown in Figure 4.38, is motivated by the Rayleigh wave reflection on the trench wall, leading to a concentration of energy around this location, as can be seen from Figure 4.39b.

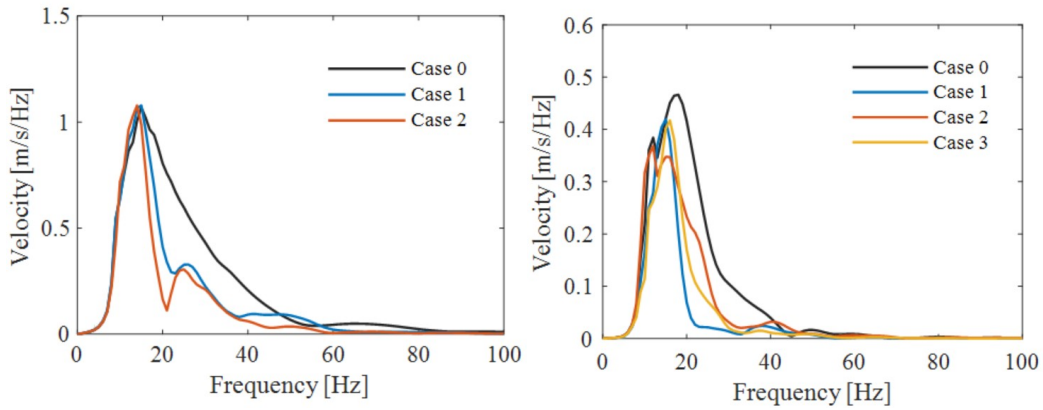
Complementary to the previous analysis, an analysis in the frequency domain is mandatory for a deeper understanding of the problem. In fact, as defined by the different standards, it is not only necessary to take into account the maximum levels of vibration, but also the most relevant frequency range of the response. Following that goal, Figure 4.40 presents the comparison of the soil free surface response before and after the construction of the trench. As can be seen, with the presence of the trench there is just an attenuation of the medium/higher frequencies. This behaviour can be easily explained, since low frequencies are associated with high wavelengths,



(a) (b)

Figure 4.39: Norm of the particle velocity at the time step t=75 ms for two different scenarios: a) Case 0; b) Case 2.

which are insensitive to local discontinuities, such as a trench.



(a) (b)

Figure 4.40: Comparison of the soil surface response before and after the construction of the trench in the frequency domain for an observation point located at the ground surface and at a distance from the pile of: a) 8 m; b) 24 m.

4.5.3 Continuous distribution of a heavy mass resting on the ground surface next the pile

A different mitigation measure with practical application corresponds to a colocation of a continuous distribution of a heavy mass resting on the ground surface next to the pile. This type of mitigation was studied and implemented in other areas involving vibrations, as the vibrations induced by railway traffic [94, 95, 96].

Figure 4.41 gives a schematic view of the study configuration, where a concrete block (mass equal to 2500 kg/m^3), with the geometrical dimensions indicated in Figure 4.41, is placed over

the ground, 5 m away from the pile. To evaluate its efficiency on vibration screening, the calculations are performed with the axisymmetric FEM-PML approach, considering the geo-mechanical properties of INEGI test site.

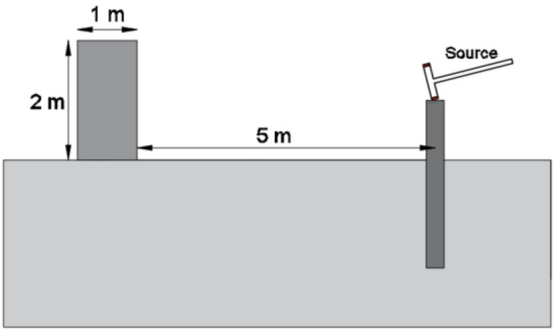


Figure 4.41: Geometry of the heavy mass next to the pile.

To understand the physical mechanisms responsible for the wave impeding effect, the behaviour of the continuous distribution of a heavy mass resting on the ground surface can be resembles that of a mass–spring system. The effectiveness of this kind of mitigation approach is related with the introduction of a new resonance frequency of the system, where the stiffness of the ground and the mass of the wall play a relevant role [95]. Theoretically, at low frequencies, the mass follows the vertical displacement of the soil's surface, and no reduction of the dynamic response is expected. Around the mass-soil resonance frequency, an amplification of the response is expected. Above this frequency, the introduction of the mass on the system can lead to a reduction of the dynamic response due to the inertia of the mass [95]. This theoretical behaviour can be observed from the results exposed in Figure 4.42, concerning to the response of the observation point located 8 m away from the pile (at the free surface), where it is possible to identify a resonant frequency around 16.5 Hz. In a simplified way, this resonance frequency, f_0, can be estimated from $f_0 = \frac{1}{2\pi}\sqrt{\frac{K_{\text{soil}}}{m}}$, where m corresponds to the total mass of the element and K_{soil} stands for the vertical dynamic stiffness of the soil under the continuous mass. The computation of the K_{soil} was performed through the axisymmetric FEM-PML numerical approach. From the calculations performed, a value of f_0 equal to 19 Hz was found. The slight difference found can be justified by the damping, once it was considered the undamped natural frequency through this simplified methodology.

Looking for a general assessment of the efficiency of the mitigation measure, Figure 4.43 presents the comparison, with the reference case, in terms of peak particle velocity (PPV) as a function of the distance from the pile centre. As can be seen, the introduction of the continuous distribution of a heavy mass resting on the ground surface next to the pile can leave to a decrease of the maximum vibration levels experimented. However, and comparatively, to the trench solution, the results are more modest, especially when considered deeper trench, as can be seen in Figure 4.43.

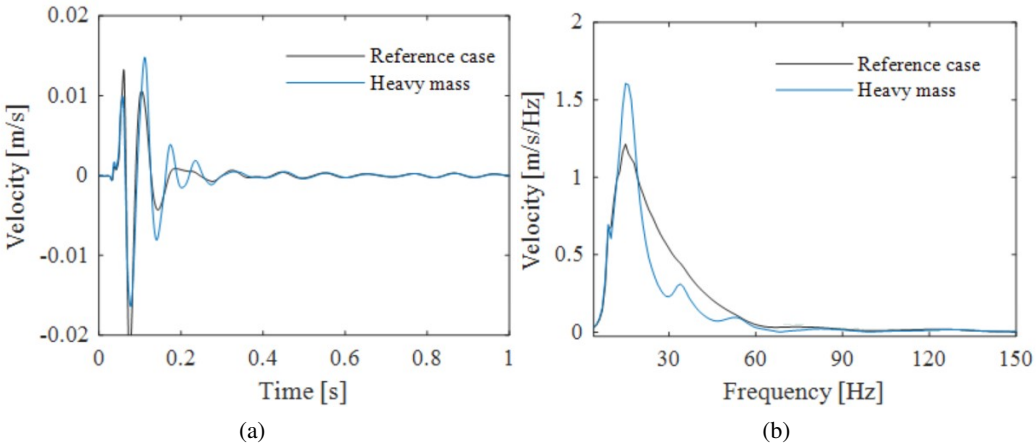


Figure 4.42: Comparison of the soil response before and after the construction of the continuous distribution of a heavy mass resting on the ground surface for an observation point located at the ground surface and at a distance from the pile of 8 m: a) time record; b) frequency content.

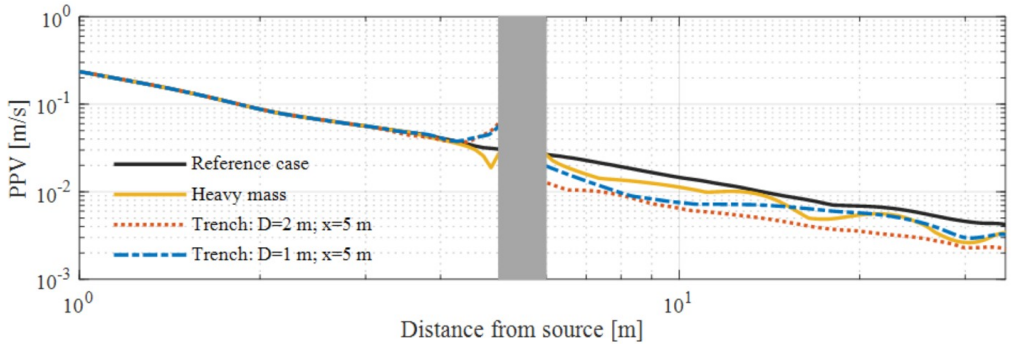


Figure 4.43: Comparison of the PPV vs distance after and before the introduction of a continuous distribution of a heavy mass resting on the ground surface for a pile penetration depth of 2 m (the trench solution results are also presented).

4.6 Final remarks

This chapter introduces the experimental part of this work by exposing the performed tests on the site and a numerical assessment for the experimental results. The experiment of site located in Aveiro, Portugal, was introduced where a pile-driving machine was used for installing a group of precast piles to founded a warehouse structure. From the exposition of the general characterization of the soil, it can extract that the nature of the information addressed by the geological-geotechnical report regarding the soil profile is nearby a qualitative description than quantitative for problem involving propagation of vibration. Thus low strain test was used to quantification the dynamic properties of ground.

The second part of this chapter is regarding the vibration induced by the pile driving process, and its numerical assessment. The experiment setup was exposed, where 10 accelerometers were

utilized to cover the influence zone up to 50 meters from the excitation source. The recorded results were subjected to a numerical treatment process to quantify the vibration level. The numerical assessment for the vibration was performed, where a reasonable matching between the experimental and numerical results was presented.

Chapter 5

Artificial intelligent-based prediction models for vibration induced by pile driving

5.1 Introduction

Artificial intelligence, AI, is one of the best-known characteristics of the twenty-first century. It consists of a set of technologies and algorithms that, in co-operate, help humans to complete complex tasks with swiftness and effectiveness. Artificial intelligence refers to the general ability of computers to emulate human thought. When merging with computer abilities, this emulation helps predict the behavior of complex phenomena and perform tasks in real-time environments. This context justifies the momentum in finding new applications of artificial intelligence in various scientific fields, particularly geotechnical engineering.

Machine learning, ML, is one of the well-known techniques of Artificial Intelligence that prove successful in dealing with engineering problems. The literature review reveals that ML algorithms, particularly ANNs and XGBoost, have been used successfully in addressing the phenomena driven by complex geotechnical problems such as, pile bearing capacity prediction, modeling soil behavior, site characterization, settlement of structures, slope stability, design of tunnels and underground openings, liquefaction, soil permeability, and hydraulic conductivity, soil compaction, soil swelling and classification of soils [97] (see Section 2.5.2).

ML are rigidly connected with the term big-data. This connection is an essential advantage of the ML. The algorithms of ML can deal with and process the big-data much better and faster than human brain. As the big-data express complex phenomena deeply, the algorithms can catch the data trending and make accurate predictions more than human brain. In addition to advantages such, time saving, human error reduction, all-time available, unbiased decisions, fast performance, etc. However, from another perspective, non-systematically developed of data-base and model training would make the resulted model mislead or over-fit, in extreme cases. Furthermore, the

generated model has drawbacks such as high technical skills, unemployment, lack of creative solutions, etc.

In this chapter, the author intends to introduce the application of ML in predicting ground-borne vibration induced by pile driving. This chapter elaborates the entire procedure of generating a learning data set, including parameters, simulation techniques, data set size, etc. The chapter introduces a model development based on Artificial Neural Networks, ANNs. Later, the ANNs model was compared with other popular machine-learning algorithms such Multiple Linear Regression, Regression Tree, Random Forest, Gaussian Process Regression, Support Vector Machine, and Extreme Gradient Boosting, XGBoost. These algorithms are trained on MATLAB and Python to generate surrogate models. The most competitive models, ANNs and XGBoost, have been subject to training on an advanced complex database corresponding to a three-layer-stratified soil profile. After that, the generated models are subject to experimental validation. A user-friendly interface was developed and presented also through this chapter.

5.2 Generation of the learning database

5.2.1 Feature selection

The physics exposition and the numerical modeling performed in the previous chapter show that the phenomenon under study has a significant complexity. The high level of complexity can be attributed to the considerable number of variables. To systematically progress in the creation of a learning database, it is crucial to define and position the parameters surrounding the phenomenon under investigation. The essentiality of this investigation arises from the existence of the diversity in the contribution between the parameters on the final response of the ground-borne vibration.This initial step refines the conceptual framework of the learning database by identifying key parameters to retain and those that can be excluded. Eliminating some parameters from the input of the learning database helps to reduce the complexity level and computational cost and helps the algorithm to find the data trends. The eliminated parameters would be chosen based on two features, i.e., parameter influence level on the ground-borne and the prediction scope of the artificial model.

For the purpose initially set, the author performed in section 3.3 a parametric study using the FEM-PML numerical model. The influence of the parameters were present in linear and equivalent linear approaches in the space domain. From the numerical assessment of the experimental test site in Aveiro, Portugal (section 4.3.3), the equivalent linear approach presents more matching than the linear approach, driven by the fact of considering the material properties degradation in the equivalent linear approach. The previous expose justifying consider the equivalent linear approach in generating the database. Therefore, it is essential to explore the results of the parametric study under the equivalent linear approach.

The author in section 3.3.5 presents a bar-chart for the influence of parameters in an equivalent linear approach in the space domain; see Figure 3.30 to 3.33. However, it is essential to go deeply into reading these results in order to generate a systematic design of the learning database.

Table 5.1: The parameters influence through the space domain on the final response.

Item	Change [%]					
	Input change	Zone 1	Zone 2	Zone 3	Zone 4	Total
Shear wave velocity	50	−1.80538	15.86908	27.81146	22.47675	64.35192
	100	0.033925	30.79745	52.57571	44.5308	127.9379
Pile penetration depth	60	−18.2468	−11.9193	−9.19106	5.837055	−33.5201
	280	−28.9089	−31.2048	−38.71	−36.9534	−135.777
Pile diameter	−50	−21.3898	−15.6501	−17.8028	−15.9505	−70.7932
	−25	−3.90519	−5.78291	−7.60826	−14.283	−31.5794
Pile material	N/A	−4.13717	0.93408	0.871582	5.148066	2.816556
	N/A	−54.1365	−41.1957	−33.1619	−54.7527	−183.247
K cushion	−50	0.486624	0.719751	0.437885	−2.0341	−0.38984
	50	0.487421	0.315738	0.196559	0.531746	1.531463
Ram mass	−28.57	−17.9903	−11.7892	−7.21491	−8.36786	−45.3622
	−14.29	−8.25947	−5.32443	−3.16745	−3.57143	−20.3228
Drop height	−40	−25.0421	−16.4033	−10.311	−12.0892	−63.8457
	40	22.03507	14.24117	7.226243	8.331118	51.8336
Soil profile	N/A	−11.9794	−8.55886	−4.52477	5.318307	−19.7447
	N/A	−17.756	−29.3543	−28.1558	−1.24249	−76.5086
Damping	−75	35.7773	129.8507	229.5192	271.4646	666.6119
	−25	−3.71533	14.49283	26.63508	20.04531	57.45789

The result of the parametric study, presented in section 3.3, were passed over numerical processing in order to get a general reading for the final results, see Table 5.1. The table introduces the percentage of change in each parameter and the corresponding in the space domain. An essential reading that can be extracted from the table is the contrast in the influence between damping and cushion stiffness. When the damping ratio of the soil profile decreased by 75% and 25%, the final ground-borne vibration response increased by 666.6% and 57.45%, respectively. In contrast, the final response changes only by -0.38% and 1.53% when the cushion stiffness changes up and down by 50%. Such a contrast between the parameters helps to figure out the general contribution of each parameter to the final ground vibration response. The author should note that specific parameters, such as pile material and soil profile, cannot be quantified in terms of percentage change. For instance, in the case of pile material parameters, the material may vary between steel, concrete, and wood. Similarly, regarding the soil profile, the geometric of the soil profile changes from homogeneous to stratified soil profile, and such changes are not easily expressed in numerical terms.

To achieve the purpose initially set, the author would eliminate other parameters based on the main scope of the machine learning model. The machine learning model considers the most common scenarios in the construction site. Therefore, the pile material parameter will be eliminated as the model will predict the vibration induced by driving a concrete pile. The material properties of the concrete pile are assumed as Young's modulus $E_p = 30000$ MPa, a Poisson's ratio $v_p = 0.15$ and a density $\rho_p = 2500\ \mathrm{kg/m^3}$.

Regarding the hammering properties for the driving process, as previously mentioned, the cushion stiffness, k_c, will be excluded from the learning database and be adjusted on $2.0 \times 10^9\ \mathrm{N/m}$

over the scope of this study. Furthermore, for practical reasons, the ram and the anvil masses (m_r and m_a) of the hammer are assumed as 7000 kg and 850 kg, respectively.

From the previous, excluding some parameters from the learning data set helps to enhance the training process and avoid misleading scenarios. The eliminated parameters were chosen based on the lowest contribution or out-of-interest area for the prediction model.

5.2.2 Simulation technique of the input parameters

5.2.2.1 Latin hypercube sampling

From the previous context, pile diameter, penetration depth, hammer drop height, soil properties, and soil profile would be the input parameters in the learning database. Presenting these parameters in the learning database should be done equally and systemically to avoid misleading, overestimating, or underestimating scenarios. For this purpose, the Latin hypercube sampling approach would be used to generate the input parameters in the learning database.

Latin hypercube sampling is a stratified sampling technique where the range of possible values of each random input variable is partitioned into “strata", and a value of each stratum is randomly selected as a representative value. These values are then combined so that each representative value is considered once and only once in the simulation process. The objective of this technique is to reduce the number of simulations needed to obtain reasonable results, see Figure 5.1.

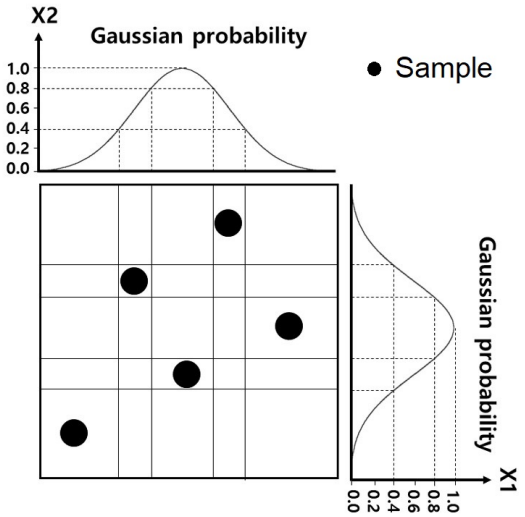


Figure 5.1: 2D-Latin hypercube sampling.

The basic steps in Latin hypercube sampling are as follows:

- Partition the range of each variable X_i into N-intervals. The partitioning should be done so that the probability of a value of X_i occurring in each variable interval is $1/N$.
- For each X_i variable and each of its intervals, randomly selected a representative value for the interval.

- Select N such that each representative value appears once and only once in these combinations, using the following procedure:

  1. To obtain the first combination, randomly select one of the representatives values for each of the input variables;
  2. To obtain the second combination, randomly select one of the $N-1$ remaining representative values of each random variable;
  3. Continues this process until getting N combinations.

The author should clarify that the previous processing was written on a Matlab script file in order to generate input parameters for the learning set. It should note that this process is valid for a variable with a normal distribution.

5.2.2.2 Schematics of learning database input

From the previous exposition, the simulating techniques conducted in generating the parameters of the learning database, i.e., Latin hypercube, were presented and written on a Matlab script file. Early in the chapter, the input parameters in the learning base were allocated from analyzing the results of the parametric study and the scope of the machine learning model. Two learning databases would be conducted in parallel, i.e., homogeneous and two layer stratified soil databases. This separation was performed to avoid misleading scenarios and enhance the ability of the machine learning model to catch the data trending. The author should mention that the initial learning databases were set to 100 cases for homogeneous models and 100 for two layer stratified soil. Table 5.2 present the values for the learning database input parameters.

Table 5.2: The values for the learning database input parameters.

Item	Unite	Homogeneous soil			Two layer soil		
		Max	Min	Distribution	Max	Min	Distribution
Pile diameter	m	0.5	0.2	uniform	0.5	0.2	uniform
Pile penetration depth	m	15	0.5	uniform	15	0.5	uniform
hammer drop height	m	1	0.3	normal	1	0.3	normal
Soil first layer							
Shear wave velocity	m/sec	200	60	uniform	200	60	uniform
Poisson's ratio		0.49	0.25	normal	0.49	0.25	normal
Damping ratio	%	0.08	0.02	normal	0.08	0.02	normal
layer thickness	m	$Inf.$			15	4	normal
Soil second layer							
Shear wave velocity	m/sec	N/A			800	200	uniform
Poisson's ratio		N/A			0.49	0.25	normal
Damping ratio	%	N/A			0.04	0.02	normal
layer thickness	m	N/A			$Inf.$		

In the stratified soil learning database, the input parameters are generated based on the soil profile with two soil layers. The based layer (second soil layer) has a shear wave velocity, C_s, higher than the first layer, which simulates the most common cases in the sites where pile driving

is applied. The distribution of the generated parameters, normally distributed, is present in Figure 5.2. The figure shows the normal distribution for the input parameters generated using Latin hypercube simulation techniques.

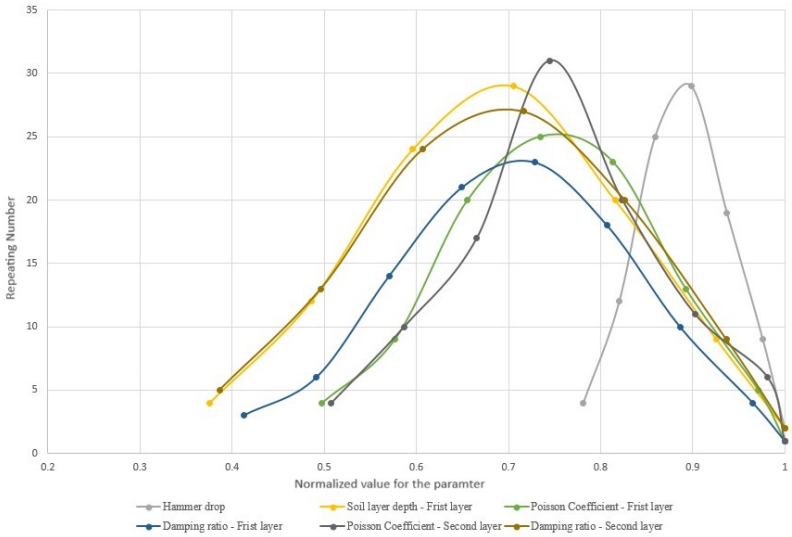


Figure 5.2: Normalized input parameters of two layer stratified soil learning database.

In contrast, the homogenous model presented the input parameters that simulated a soil profile with a single type of soil layer. Therefore, the soil geometric parameter can not be applicable as the soil is homogenized. Figure 5.3 presents the normal distribution of the 100 cases for a homogenous soil model.

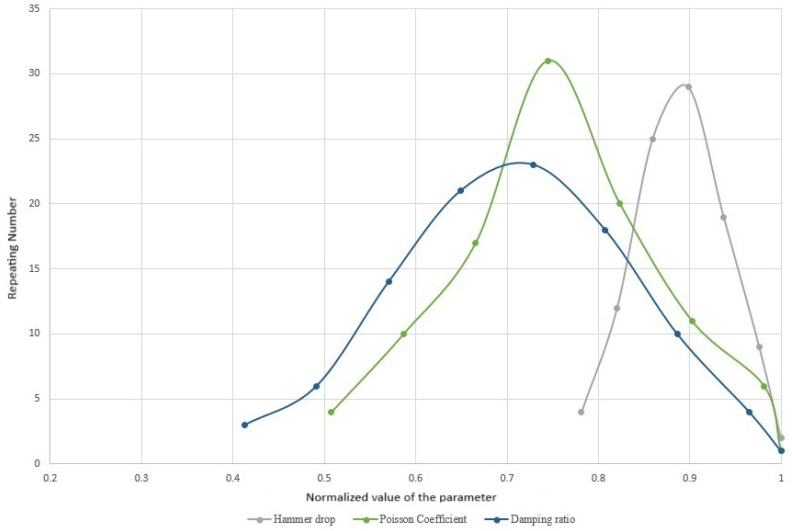


Figure 5.3: Normalized input parameters of homogenous soil learning database.

5.2.3 The execution of the learning database

5.2.3.1 FEM-PML Numerical model

This section is designed to introduce the proceedings of generating the output of the learning database. As introduced in Chapter 3, the FEM-PML numerical model is used to generate the particle peak velocity, PPV, on the ground surface in the space and frequency domains. The generated PPV from the numerical model would be the output of the learning database. Therefore, it is essential to model the input parameters of the learning data set to get the corresponding particle peak velocities. The numerical model validation was presented in Section 3.2.4 to eliminate erroneous practices. However, the current version of the numerical model should be enhanced in order to be more compatible with modeling extensive-size databases. The importance of this contribution arises from the minimizing of the time-consuming and human intervention as exposed in the following.

The current version of the FEM-PML numerical model needs manual interaction in 3 phases (see Figure 5.4). The first interaction for the model will be on the Gmsh program to generate the geometric file for the generation of the finite element mesh of the soil-pile system. The second phase of human interaction is placed in assigning the soil-pile material properties on the geometric file generated by the Gmsh program in the previous step. The third phase is for identifying the hammering condition on the numerical model. All of this proceeding is conducted together in FEM-PML code and running on the computer to get the results.

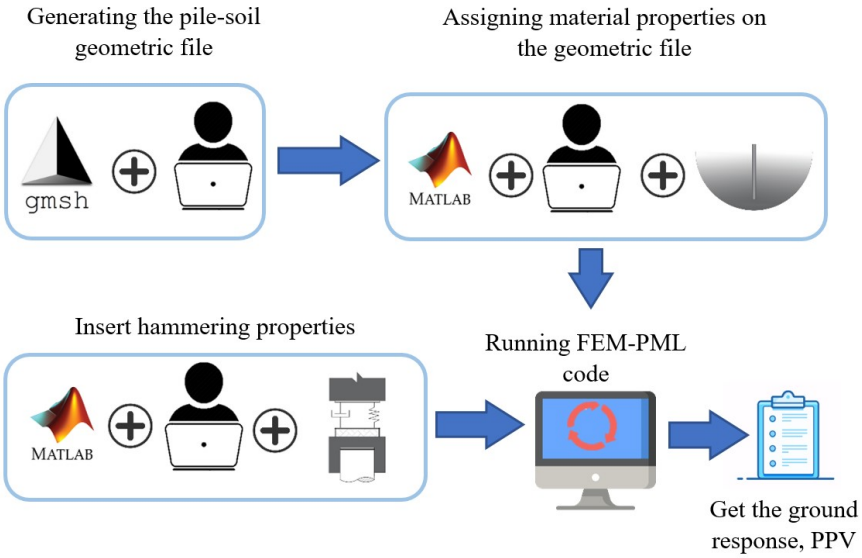


Figure 5.4: Schematic presentation for FEM-PML numerical model.

Performing 200 case scenarios with this work layout would waste time and increase the probability of occurring input mistake scenario from human interaction. Therefore, it is essential to work on removing manual interaction as possible and converting it to automatic entire processing.

Figure 5.5 presents the modification applied to the FEM-PML numerical model to minimize human interaction. A general input user interface on a spreadsheet is generated. Matlab script file is written to call the Gmsh program and automatically generate the pile-soil geometric file based on the data implemented in the spreadsheet. Insert the hammering properties is converted to be fully automatic based on previous data in the spreadsheet. Then FEM-PML numerical model is run to get the final ground response. The author should mention that in this development, manual interaction occurs in the initial step, in particular at the stage of input parameters entry. Such development reduce the human interaction and enhancing output reliability.

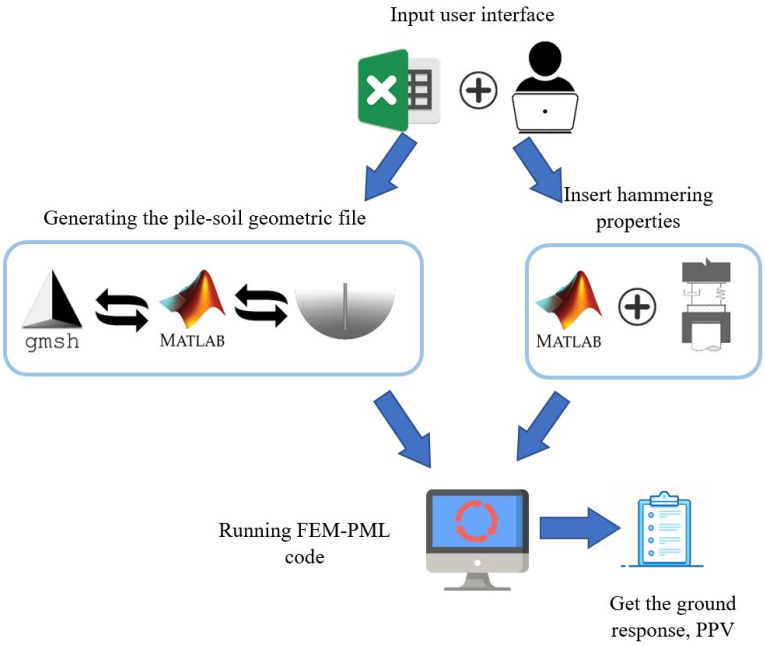


Figure 5.5: Schematic presentation for the modification in FEM-PML numerical model.

5.2.3.2 Schematics of learning database output

From the previous, the FEM-PML numerical model was adjusted to be compatible with automatically modeling 200 case scenarios of pile-soil profiles. These 200 cases form the machine learning database and need more than 300 hr of computing on high performance unit, HPU. As mentioned before, the numerical model output gives the PPV values in the space domain, and the frequency content. Therefore, it is essential to consider the distance from the pile center as an input parameter in the learning database. For each case scenario, 14 distance points were chosen to be considered in the database. These distances are subject to change for a few centimeters, more or less, based on the mesh size for each scenario. Therefore, each case has sub-cases with the same input parameters but changes in the measuring point. From the previous context, it can be concluded that the actual size of the learning database is 1400 cases for the homogenized soil profile and 1400 for the two layer stratified soil profile. Figure 5.6 presents the scatter plot for the

PPV in the space domain for the homogenized and two layer stratified soil profile resulting from the numerical model in two observation points.

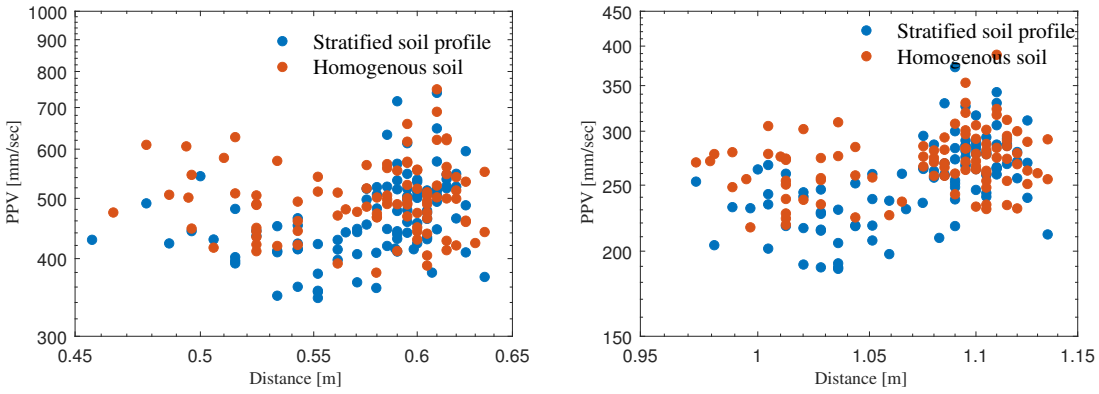


Figure 5.6: The FEM-PML numerical model output, PPV, for the homogenized and two layer stratified soil profile: a) observation point 1; b) observation point 2.

5.3 Data division strategy

From the previous context, generating the learning database (inputs and outputs) performed. Combining the generated learning database with the machine learning algorithms is the next milestone in the proceeding of generating the model. However, perform training with the current layout of the learning database increases the probability of going through a misleading scenario. The literature review proves that the learning database should be divided into subsets. Every subset has its utilizing, characterizing, and influence on the final output results, as present in the following.

To divide the dataset of machine learning, typically split it into three sets: a training set, a validation set and a test set. The training set is used to train a model, the validation set is used to evaluate the model during training and the test set is used to evaluate the model after it has been trained. There are various methods for dividing a dataset into these three sets, which the most popular approaches are mentioned in the following:

- Random sampling: randomly selecting samples from the dataset to include in each set;
- Stratified sampling: ensuring that the proportion of different classes in each set is the same as the proportion in the original dataset;
- Cross-validation: dividing the dataset into "folds", and then training and evaluating the model on each fold.

As mentioned before, it is essential to split data carefully, as the quality of the training and test sets can significantly impact the machine learning model's performance. Therefore, it is essential to ensure that the different sets are representative of the overall dataset. For example, if the dataset is

imbalanced, it is important to ensure that the training, validation, and test sets are also imbalanced. This strategy of data division is confirmed by Masters et al. (1993) [98], as it is mentioned that "if our training set is not representative of the data on which the network will be tested, we will be wasting our time."

As presented before, a systemic approach for building the learning dataset was followed. This approach helped to comprehend the nature of the data in the learning dataset. In connection with the abovementioned purpose, data division strategy, the author builds an algorithm for dividing data. The algorithm works on splitting the learning dataset into three subsets. The generated subsets hold the same layout and arithmetical characteristics, i.e., mean and standard deviation. Shahin et al. (2004) [99] state that a promising result was obtained when 20% of the data was used for validation and the remaining data were divided into 70% for training and 30% for testing. However, the algorithm was built to connect with the training model to get a loop of model training with different ratios for subsets. This proceeding enhances the opportunity to get a promising machine-learning model. Figure 5.7 compares the learning set with the subsets generated from the data-dividing algorithm developed by the author. The Comparison proves a harmony between all sets. Furthermore, the author should clarify that the ratio of the subsets (training, testing, and validation) presented in the figure are 64%, 14%, and 22%, respectively, from the learning database.

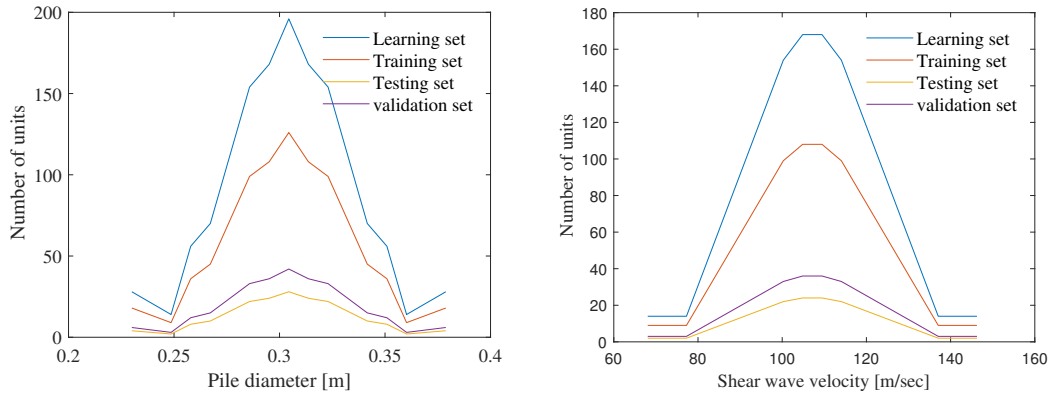


Figure 5.7: Comparison between the learning data set and subsets in two layer stratified soil scenario: a) pile diameter; b) shear wave velocity.

5.4 Data Standardization and Normalization

The previous expose has divided the learning set into subsets, i.e., training, testing and validation. It is essential to introduce the data inside these subsets to the training algorithm in a suitable form. This procedure in machine learning is known as data pre-processing. Data pre-processing ensures all variables receive equal attention during the training process, as incorporating input parameters with different distributions can lead to a model showing bias towards features with larger values

and variance. Furthermore, pre-processing enhances the speed of the learning process. Input parameters scaling concludes mainly in two forms, i.e., standardization and normalization. However, some transformation forms for data scaling are addressed in some types of data distribution, e.g., log, exponential, etc.

Normalization typically means re-scaling the input data into a range between 0 to 1. In some cases, normalization helps for obtaining the optimal results in machine learning. The normalization function is presented in the following equation:

$$\hat{X} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{5.1}$$

Standardization typically means re-scaling data with a mean of 0 and a standard deviation of 1. Standardization function is presented in the following equation:

$$\hat{X} = \frac{X - \mu}{\sigma} \tag{5.2}$$

As a result, standardization is ideal for data that fits a standard or Gaussian distribution. It is also superior when handling data with outliers, as it is more resistant to extreme values. On the other hand, normalization is the alternative when the distribution of the data is not defined.

Finally, the author should emphasize that the testing data should not be considered when determining the parameters for a scaling operation. The parameters used in any scaling should always be determined by the training data alone. While the testing data is also transformed, its values are only scaled based on the parameters derived from the training data.

5.5 Model stopping criteria

After dividing the learning data set into its subsets, the model training phase starts. As exposed before, the model uses the training, testing and validation sets to develop the artificial model. The number of iterations over the entire dataset, epochs, controls the accuracy level of the artificial model. If the number of epochs is low for the model to get the general trend of the data, the model outputs will not be accurate enough. On the other hand, if the epochs number is relatively high, the resulting model would be memorized rather than a generalized model. From this context, precisely identifying the model-stopping criteria is essential. Therefore, several approaches can be used as stopping criteria for training the machine learning model. Some standard methods include:

- Early stopping: This involves training the model until the performance on a validation set starts to degrade, and then stopping the training at that point. This can prevent overfitting, as the model is stopped before it has a chance to learn patterns in the training data that do not generalize to the validation data.

- Fixed number of epochs: Another approach is to train the model for a fixed number of epochs (iterations over the entire dataset) and then evaluate its performance. This can be

useful if have a sense of how long the model should take to converge based on past experience or if have a specific time budget for training the model.

- Performance threshold: it can also specify a performance threshold that the model must meet or exceed in order to be considered for deployment. If the model does not meet this threshold after a certain number of epochs, the training process can be stopped.

- Manual intervention: In some cases, to stop training the model manually when it has reached an acceptable level of performance. This can be based on visual inspection of the training and validation loss or accuracy, or on other performance metrics such as precision, recall, or F1 score.

- Automatic convergence detection: Some machine learning libraries include algorithms for automatically detecting convergence of the model during training. These algorithms can be used to stop the training process when the model has converged.

  The Automatic convergence detection is used in training the model under this work.

5.6 Model performance indicator

Model evaluation is the process of assessing the performance of a machine-learning model on a specific task. In other words, once the training phase of the model has been successfully accomplished, the performance of the trained model should be validated. The purpose of the model validation phase is to ensure that the model has the ability to generalize within limits set by the training data, rather than simply having memorized it. There are many ways to evaluate the generated model. The appropriate evaluation method depends on the type of model and the goals of the modeling process. Some common evaluation methods for machine learning models include:

- Training and validation accuracy: one way to evaluate a model is to measure its accuracy on the training and validation data. If the model has high accuracy on both the training and validation data, it is likely to generalize well to new, unseen data.

- Training and validation loss: another way to evaluate a model is to measure its loss on the training data and the validation data. If the model has low loss on both the training and validation data, it is likely to be performing well.

- Confusion matrix: a confusion matrix is a table that is used to describe the performance of a classification model on a set of test data. The rows of the matrix correspond to the true classes of the test examples and the columns correspond to the predicted classes.

- Cross-validation: cross-validation is a method for evaluating the performance of a model by training it on a subset of the data and evaluating it on the remaining data. This can be done using k-fold cross-validation, where the data is split into k folds, and the model is trained and evaluated k times, each time using a different fold as the evaluation set.

Furthermore, the root means squared error, RMSE, the mean absolute error, MAE, and the coefficient of determination, R^2, are the main criteria that are often used to evaluate the prediction performance of the machine learning model dealing with regression problems. The RMSE is the most popular measure of error and has the advantage that large errors receive much greater attention than small errors [100]. In contrast with RMSE, MAE eliminates the emphasis given to large errors. Both RMSE and MAE are desirable when the evaluated output data are smooth or continuous [101]. While, the coefficient of determination is a measure that is used to determine the relative correlation and the goodness of fit, R^2, between the predicted and observed data. The following guide for values of R^2 between 0.0 and 1.0 [101]:

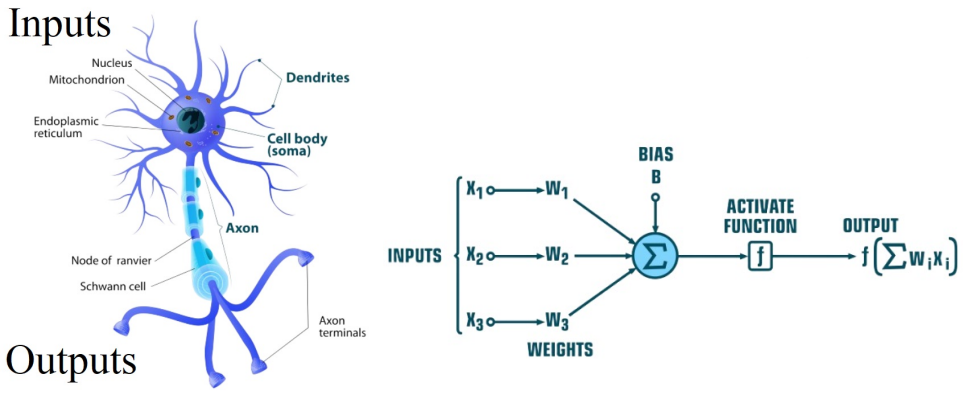
- $R^2 \geq 0.8$ strong correlation exists between two sets of variables;
- $0.2 < R^2 < 0.8$ correlation exists between the two sets of variables;
- and $R^2 \leq 0.2$ weak correlation exists between the two sets of variables.

5.7 Machine learning algorithms

5.7.1 Artificial Neural Networks, ANNs

5.7.1.1 General overview

Artificial Neural Networks, ANNs, is one of machine learning techniques that endeavors to simulate the function of the human brain and nervous system, see Figure 5.8. ANNs learn from data-sets presented to it to capture the relationships among the data, even if the underlying relationships are unknown or the physical meaning is difficult to explain. ANNs are adequate for modeling complex phenomena such as geotechnical engineering problems.



(a) (b)

Figure 5.8: Human brain neuron cell vs. artificial nerve cell: a) structure of typical human neuron; b) structure of artificial neuron cell.

From the previous expose, this modeling capability and learning from the experience of data-sets give ANNs superiority over most traditional modeling methods.

5.7.1.2 Theory of Artificial Neural Networks

Artificial Neural Networks, ANNs, consist of several artificial neurons known as processing elements (PEs), nodes, or units. ANNs consists of multi-layer perceptrons, MLPs. In MLPs, artificial neuron nodes/units are usually arranged in layers: an input layer, an output layer, and one or more intermediate layers called hidden layers, see Figure 5.9.

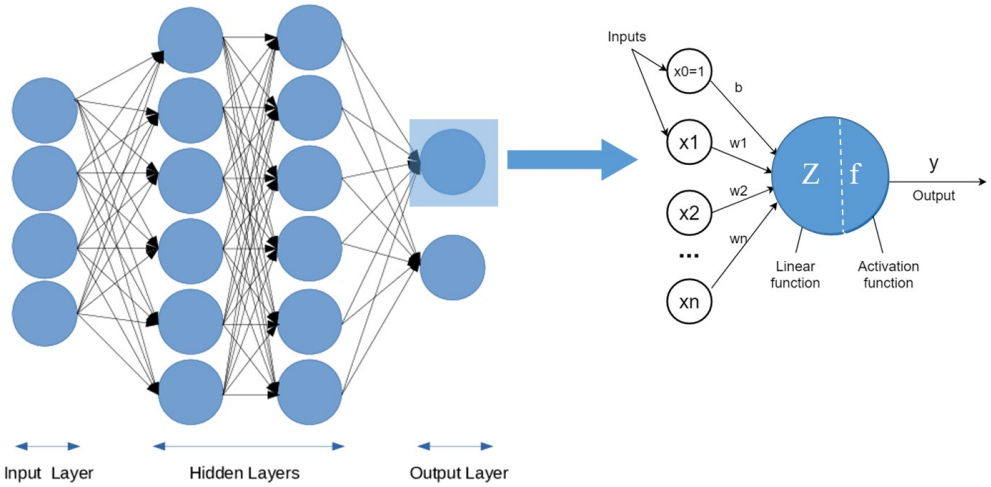


Figure 5.9: Artificial Neural Networks and node structure.

Each node in a specific layer is connected, fully or partially, to many other nodes via weighted connections. The scalar weights determine the strength of the connections between interconnected neurons. A zero weight refers to an absence of the connection between two neurons and a negative weight refers to a prohibitive relationship. Every node receives from other nodes its weighted inputs. These weights are assembled and a bias unit or threshold is added or subtracted. The bias unit is used to scale the input to a useful range to improve the convergence properties of the neural network. The results are passed through an activation function such as logistic sigmoid or hyperbolic tangent, etc., in order to produce the output of the processing element. The previous expose is presented mathematically in the following equation for node j:

$$\begin{aligned} Z_j &= b_j + \sum_{i=1}^{n} w_{ji} x_i \\ y_j &= f(I_j) \end{aligned} \tag{5.3}$$

where Z_j is the activation level of node j; w_{ji} is the connection weight between nodes j and i; x_i is the input from node $i, i = 0, 1, \ldots, n$; b_j is the bias or threshold for node j; y_j is the output of node j; and f. the activation function.

The information propagation in MLPs initiates at the input layer where the input data is presented. The inputs are weighted and received by each node in the next layer. The weighted inputs are assembled and passed through an activation function to produce the node output, which is also

weighted and passed to nodes in the next layer. The previous learning rule was considered until getting the network's final output. The network adjusts its weights by forward and backward feeding until it finds a set of weights that will produce the input-output mapping with the minimum possible error. The process mentioned above is known as learning or training. The author should mention that the previous learning process corresponds to supervised feedback neural network, which is popularly used in geotechnical engineering. Furthermore, the supervised feedback neural network will be used in this work, justifying drive the spotlight on this specific type of ANNs.

In the ANNs model, the degree of non-linearity between input and output can be easily changed by changing the activation function and network architecture. The network architecture consists of the number of hidden layers and nodes, as exposed before. The type of non-linearity can be changed by varying the number of hidden layers and the number of nodes in each layer. It is essential to emphasize that the most compatible network architecture would be achieved through a set of trial-and-error alterations. Additionally, ANNs models can be upgraded from uni-variate to multi-variate by increasing the number of input nodes. Such an advantage gives the pioneering to the ANNs in the geotechnical engineering applications [97].

5.7.2 Multiple linear regression

Multiple Linear Regression is a regression technique that postulates a linear association between the regression function and a set of p input features (x_1, x_2, ..., x_p). Linear regression models have proven to be valuable tools in various engineering applications because of their straightforwardness and intuitive interpretability. In certain situations, such as those involving sparse or limited training data, multiple linear regression can even outperform more intricate nonlinear machine learning models [102], see Figure 5.10.a. The essence of multiple linear regression can be summarized as follows:

$$\widehat{y_i} = \beta_0 + \sum_{j=1}^{p} x_j \beta_j \tag{5.4}$$

where $\widehat{y_i}$ is the estimated target output, β_0 is the intercept, β_j is the feature's coefficient for x_j, p is the number of input features. The coefficients β_j are generally estimated using a training dataset $(x_1, y_1), (x_2, y_1), \ldots, (x_n, y_n)$ by means of a least squares approach. Given a number of features (p), each x_i can be written as a vector $(x_{i1}, x_{i2}, \ldots, x_{ip})^T$. A vector of coefficients $(\beta_0, \beta_1, \ldots, \beta_p)^T$ can be found to minimise the sum of squared residuals such that:

$$RSS = \sum_{i=1}^{n} \left(y_i - \beta_0 + \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 \tag{5.5}$$

5.7.3 Regression tree

The Regression Tree algorithm, as described in reference [103], is specifically designed for addressing regression problems [104]. This approach involves the iterative construction of a decision

tree, depicted in Figure 5.10.b, by generating a series of splitting rules from the top of the tree to the bottom [103]. This resultant structure is commonly referred to as a decision tree. The initial step involves partitioning the entire dataset into m distinct, non-overlapping regions denoted as "terminal nodes" or "leaves" and labeled as $(R_1, R_2, \ldots, R_m)$.

Within this context, the points at which the predictor space undergoes subdivision are designated as "internal nodes," while the connecting segments between these nodes are termed "branches" [102, 104]. The algorithm's core functionality lies in its automatic selection of suitable internal nodes for splitting input features, which ultimately defines the tree's shape.

In each of these regions R_m, characterized by its terminal node, predictions are established based on the output target y_i. This prediction is computed as the mean of y_i values within the respective region. Subsequently, the algorithm identifies the set of regions $(R_1, R_2, \ldots, R_m)$ that collectively minimize the Residual Sum of Squares (RSS) function, denoted as $f(x)$ in order to optimize the tree's structure and predictive performance.

$$f(x) = \sum_{m=1}^{M} \sum_{i \in R_m} (y_i - \widehat{y}_{R_m})^2 \tag{5.6}$$

where y_{R_m} is the mean of the target output observed in the training data in m. A greedy approach is adopted to split the input features j based on the lowest RSS. A splitting point j and s can be defined as the pair of half-planes:

$$R_1(j,s) = (X \mid X_j \leq s), R_2(j,s) = (X \mid X_j > s) \tag{5.7}$$

and the value of j and s are determined to solve

$$\sum_{i:x_i \in R_1(j,s)} (y_i - \widehat{y}_{R_1})^2 + \sum_{i:x_i \in R_2(j,s)} (y_i - \widehat{y}_{R_2})^2 \tag{5.8}$$

where $\widehat{y}_{R_1}$ and $\widehat{y}_{R_2}$ is the mean target of the training data in $R_1(j,s)$ and $R_2(j,s)$, respectively. The process is repeated for all regions until a specific criterion is attained. However, Regression Tree models can be overfitted if the tree is very large while a small tree might result in underfitting [102, 104].

5.7.4 Random forest

Random Forest, as outlined in reference [105], stands as an ensemble of Decision Trees. What distinguishes this method is its adaptation of the bagging approach, which serves to reduce the inter-tree correlation [103]. The pivotal departure lies in the manner of selecting feature subsets during the tree-growing process, a process referred to as "tree decorrelation." This particular approach leads to a significant reduction in variance among the constituent trees, enhancing the model's reliability [104].

To elaborate, the decision tree algorithm is repeatedly trained on various, randomly chosen, of subsets derived from the training data. The ultimate prediction from this approach is derived

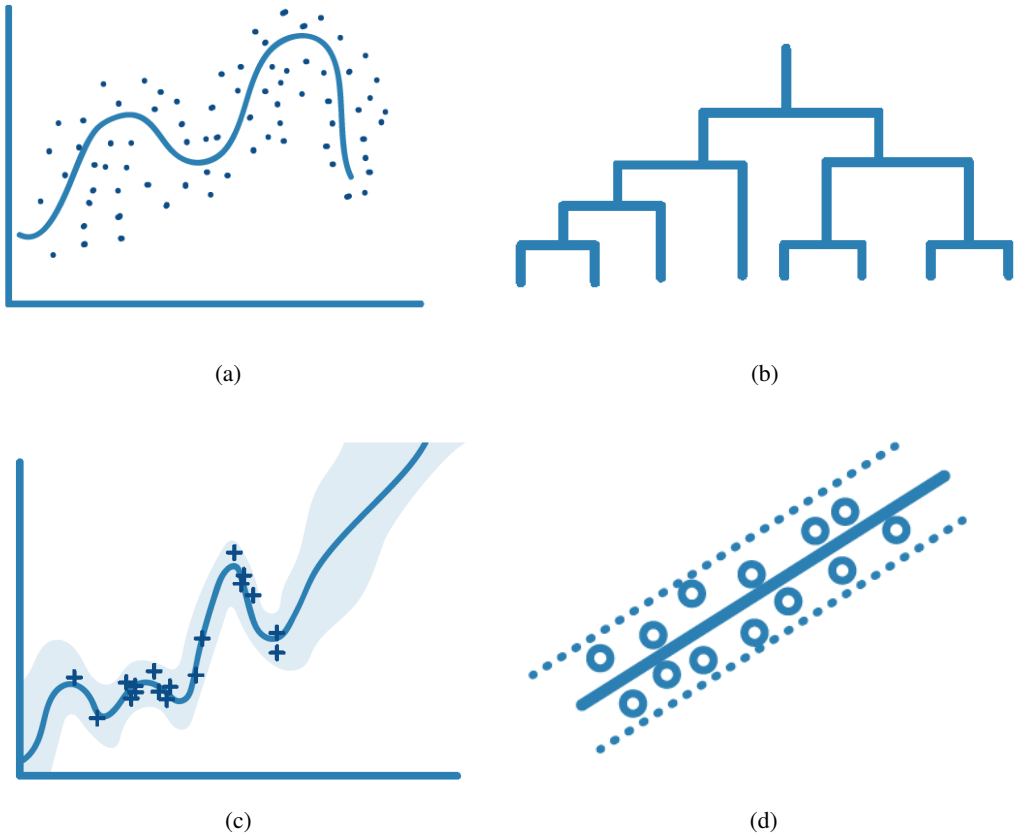


Figure 5.10: Machine learning schemes: a) multiple linear regression; b) regression tree; c) SVM; b) GPR.

by averaging the predictions generated from multiple decision trees, as depicted in Figure 5.10.b. This is mathematically expressed in Equation 5.9

$$\widehat{f}_{RF}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{5.9}$$

where $\widehat{f}_{RF}(x)$ signifies the final prediction derived from the averaging of the outputs of the individual trees T_b, and B denotes the total count of Decision Trees.

Due to the inclusion of numerous independent trees in the prediction process, Random Forest exhibits a notable advantage over the standalone Decision Tree in terms of model generalization. Nonetheless, it is important to note that constructing a large number of trees and integrating their results incurs a substantial computational cost.

5.7.5 Gaussian Process Regression

Gaussian Process Regression (GPR) stands as a powerful and versatile tool within the realm of machine learning and statistical modeling, see Figure 5.10.c. At its core, GPR is built upon the

principles of Gaussian processes, representing a collection of random variables with a joint Gaussian distribution. This distinctive feature allows GPR to capture not only the mean behavior of a system but also the relationships and uncertainties inherent in the data. By defining a kernel function that quantifies the similarity between input points, GPR crafts a predictive framework that excels in scenarios where a probabilistic understanding of predictions is crucial. The Bayesian nature of GPR further enhances its appeal, treating model parameters, such as those within the kernel function, as random variables and allowing for a principled incorporation of prior knowledge [106].

The application of GPR extends across diverse domains, ranging from finance and geostatistics to optimization problems and robotics. One of its notable advantages is the ability to provide not only point estimates but entire probability distributions over potential functions, offering a nuanced perspective on the uncertainties associated with predictions. The selection of an appropriate kernel function plays a pivotal role in shaping the GPR model, allowing practitioners to tailor the analysis to the characteristics of the underlying data. However, it's essential to note that the computational complexity of GPR, particularly inverting the covariance matrix, may pose challenges for large datasets. In response, various approximations and sparse GP techniques have been developed to address scalability concerns.

In practical terms, GPR finds application in scenarios where a nuanced understanding of predictions and uncertainties is paramount. Whether modeling complex physical phenomena, or optimizing processes with limited data, GPR's probabilistic framework and flexibility make it a valuable asset. Researchers and practitioners alike leverage GPR to gain insights into the underlying patterns of their data, especially in situations where capturing uncertainty is as crucial as making accurate predictions. The elegance of GPR lies in its ability to balance precision with a rich representation of uncertainty, making it a cornerstone in the toolkit of machine learning and statistical modeling [106].

5.7.6 Support Vector Machine

Support Vector Machine (SVM) [105] is typically employed for classification tasks. However, SVM can also be effectively used to tackle regression problems, giving rise to Support Vector Regression (SVR). SVR operates with a characteristic mapping function:

$$f(x) = \sum_{m=1}^{M} w_m \varphi_m(x) + b \tag{5.10}$$

In this equation, M signifies the count of mapping functions, w_m denotes the weights associated with these functions, facilitating the transformation from the original to high-dimensional feature spaces. Additionally, φ_m represents the vectors of input features (x), and b corresponds to the bias parameter.

The primary objective of SVR is to identify a hyperplane such that a majority of the training data points fall within a specified margin or tolerance level (ε). Here, ε serves as a hyperparameter governing the maximum permissible error margin between the target value and the fitted

hyperplane. Consequently, the algorithm endeavors to minimize the coefficients while imposing constraints on the error term, as illustrated in Figure 5.10.d. The error function subject to optimization in SVR can be expressed as:

$$C\sum_{i=1}^{n}(\xi_i+\widehat{\xi}_i)+\frac{1}{2}\|w\|^2 \text{ subject to } \begin{cases} \left(\sum_{m=1}^{M} w_m\varphi_m(x_{ij})+b\right)-y_i \leq \varepsilon+\widehat{\xi}_i \\ y_i-\left(\sum_{m=1}^{M} w_m\varphi_m(x_{ij})+b\right) \leq \varepsilon+\widehat{\xi}_i \end{cases} \tag{5.11}$$

In this equation, ξ_i and $\widehat{\xi}_i$ represent the slack variables pertaining to data points, where i ranges from 1 to n, with n representing the total number of training data points. Furthermore, w denotes a vector in a high-dimensional space, y_i is the target output value, x_i signifies the feature vector associated with y_i, and C stands as a hyperparameter indicative of regularization levels, governing the allowance for training data points to deviate from the error margin ε.

The use of various mapping functions can enhance the precision of the prediction model, with kernel functions being a popular choice to capture nonlinear relationships [105]. In this context, the radial basis function kernel is frequently employed. Moreover, SVR hyperparameters, such as C and ε, play a vital role in balancing the bias-variance trade-off.

5.7.7 Extreme Gradient Boosting

Extreme Gradient Boosting, XGBoost, is a powerful and widely-used machine learning algorithm that has gained popularity in various data science competitions and real-world applications due to its exceptional performance in both classification and regression tasks. This algorithm is advanced form from the Tree group, see Figure 5.11. It is an ensemble learning method based on the gradient boosting framework, known for its scalability, speed, and ability to handle complex, high-dimensional datasets.

To understand the mathematics behind XGBoost, it's essential to grasp its objective function, which is optimized during the training process. The objective function in XGBoost combines a loss term (typically a mean squared error for regression or a log loss for classification) and a regularization term to prevent overfitting.

Here is a simplified representation of the XGBoost objective function for regression:

$$L(\theta)=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\Omega(f) \tag{5.12}$$

where,

- $L(\theta)$ is the total objective to minimize.
- n is the number of training samples.
- y_i is the true target value for the i-th sample.
- $\hat{y}_i$ is the predicted value for the i-th sample.

- $\Omega(f)$ is the regularization term, which includes both L1 and L2 regularization.

The training process in XGBoost involves adding decision trees to the model iteratively, with each tree attempting to correct the errors made by the previous ones. The predictions from each tree are weighted based on their contribution to the objective function.

In summary, XGBoost is a powerful machine learning algorithm known for its scalability, speed, and accuracy. Its success in various data science competitions and practical applications has made it a popular choice for building predictive models. By understanding its mathematical foundations, practitioners can effectively use and customize XGBoost for their specific machine learning tasks. The author should mention that this XGBooster algorithm is not built in Matlab software and it will be training on Python with TensorFlow library and PyCharm environment [68].



Figure 5.11: Machine learning Tree group [107].

5.8 Application of ANNs model

5.8.1 Peak particle velocity prediction

As mentioned before, building a machine learning model based on a neural network should be done systemically to get a promising result. The neural network's architecture is one of the most critical aspects of the neural network that needs to be defined well. The importance of this aspect arises from the fact of being the degree of non-linearity between input, and output controls the complexity level of the neural network architecture. For this purpose, a set of trying and assessment proceedings would be performed in order to define the most proper neural network architecture layout between the input and the output for the phenomenon placed under investigation.

As previously exposed, during the proceeding of generating the machine learning database, the homogeneous soil profile database was separated from the stratified soil profile database to enhance the neural network algorithm's ability to find the data-trending. Therefore, the author intends to highlight that the subsequent proceeding for generating the neural model would be applied to each database individually. Defining the characters of the architecture of the neural networks, i.e., the number of layers and nodes, in addition to dividing ratios of the database and training algorithm, will be introduced in the following.

The first step in identifying the neural network's architecture is determining the number of nodes in the first layer. This goal can be achieved by determining the value of the RMSE for a

single hidden layer neural network with nodes ranging from 1 to 20. This process is repeated 50 times to get each network's average RMSE, $\bar{R}$.

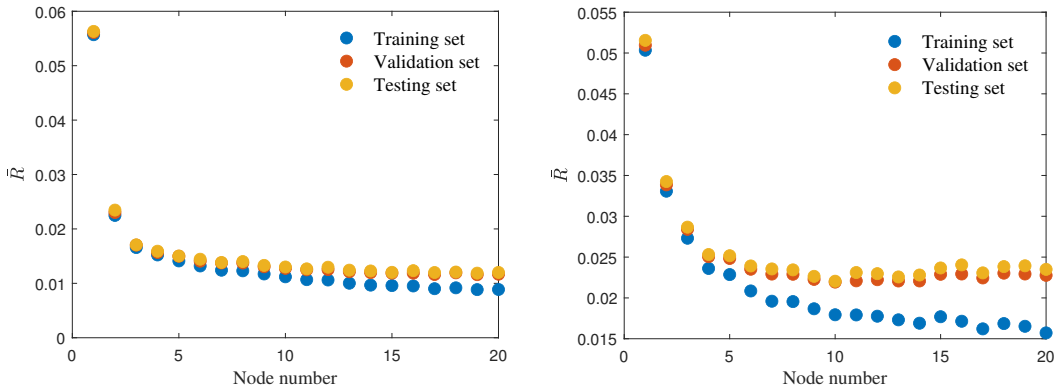


Figure 5.12: RMSE for single hidden layer ANN: a) homogeneous soil; b) layered soil.

Figure 5.12 presents the value of $\bar{R}$ for each number of the node for the single hidden layer network. The figure exposed that as the number of nodes increases, the model goes through an overfitting scenario as the model becomes more fitting for the training set than the validation and testing set. Such behavior can be noticed from the contrast between the values of $\bar{R}$ of the three sets at the high node numbers. The author should carefully choose the proper number of nodes that achieve the min $\bar{R}$ and avoid the overfitting scenario. From the figure, the number of nodes in the first hidden layer in homogeneous and layered soil profiles is 3 and 6, respectively.

After determining the total number of nodes in the first hidden layers in ANN, the focus should be to evaluate the network $\bar{R}$ after adding the second hidden layer. In this iteration, the node number ranges from 1 to 20, with the repeating times equal 50. Figure 5.13 shows the value of $\bar{R}$ for the networks. The figure does not indicate any significant improvement in the values of $\bar{R}$, justifying that adding a second layer for ANN is unavailing.

From the previous exposure, the architecture of the ANN was defined from a set of trying iterations. The primary purpose of these iterations is to avoid insufficient accuracy due to improper ANN architecture. Figure 5.14 presents the final architecture of the neural networks used in the next phase of building the ANN model.

The second phase in building the ANN model is identifying the ratio of the database. As mentioned before, the database should be divided into learning, testing, and validation sets. The dividing ratio affects the ANN model output, as mentioned before. In this context, an assessment for different ratios would be performed. For this purpose, seven cases scenario of the data dividing ratio is presents. The case scenarios are mentioned in Table 5.3.

Figure 5.15 presents the $\bar{R}$ for the case scenarios. The figure presents a contrast between the case scenarios. From the figure, case scenario number five was chosen to model the neural network of the homogeneous and layered soil profile.



Figure 5.13: RMSE for double hidden layer ANN: a) homogeneous soil profile; b) layered soil.

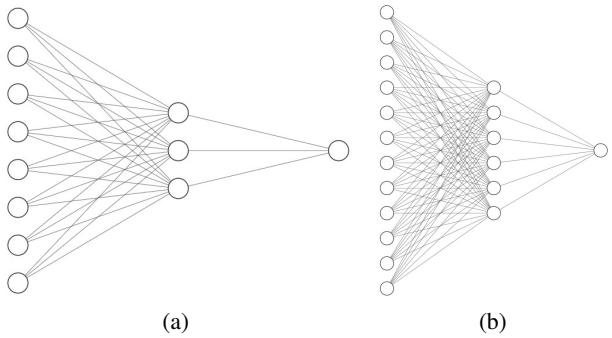


Figure 5.14: Architecture of ANN model: a) homogeneous soil profile; b) stratified soil.

Table 5.3: Case scenarios of the dividing ratio for the database of ANN.

Set	Unite	Case scenario Number						
		1	2	3	4	5	6	7
Training	%	40	50	56	60	70	45	40
Validation	%	30	20	20	30	20	25	25
Testing	%	30	30	24	20	10	30	35

In the last phase of building the ANN model, 50 iterations would be performed to determine the most proper algorithm. Figure 5.16 presents the $\bar{R}$ values of the different algorithms. From the figure, it is possible to choose algorithm "Levenberg-Marquardt" as the training algorithm in both models.

From the previous, two ANN models corresponding to predicting the PPV induced by the pile driving process were developed. Figure 5.17 presents the MSE versus the number of epochs. From the figure, the best validation performance for the stratified soil profile model is 0.0006437 at epoch 21, while for homogeneous soil profile MSE is 0.00029849 at epoch 200.

Finally, Figure 5.18 presents the regression graph between the numerical results provided by FEM-PML numerical model and the predicted results by the developed ANNs model for valida-

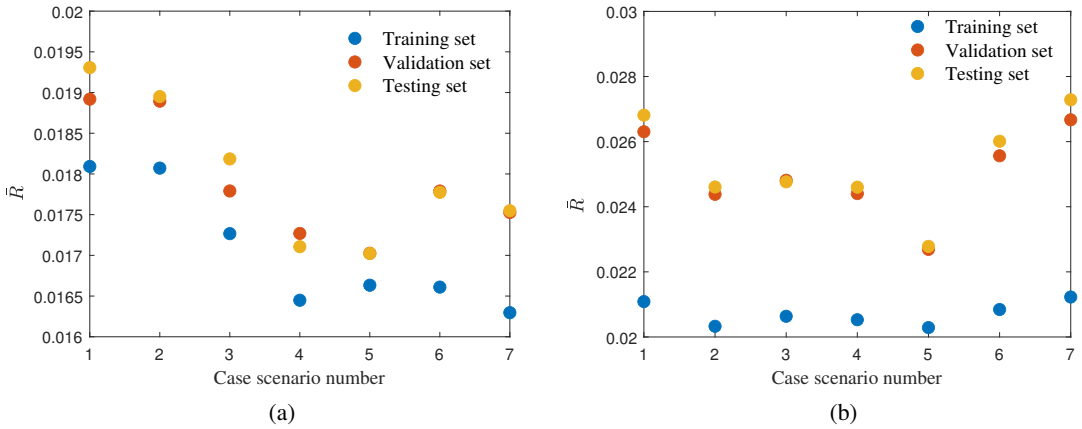


Figure 5.15: Database dividing ratio of the ANN: a) homogeneous soil profile; b) layered soil.

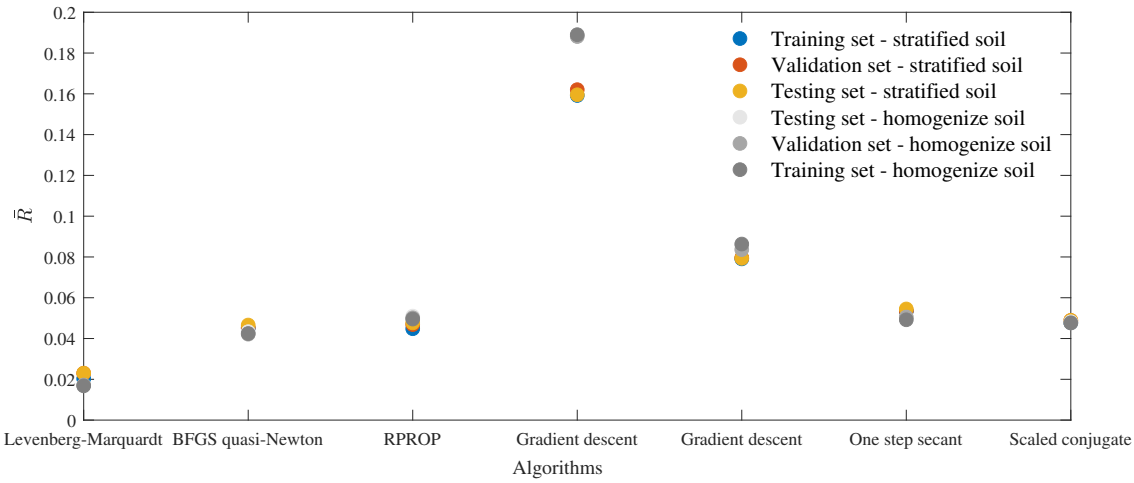
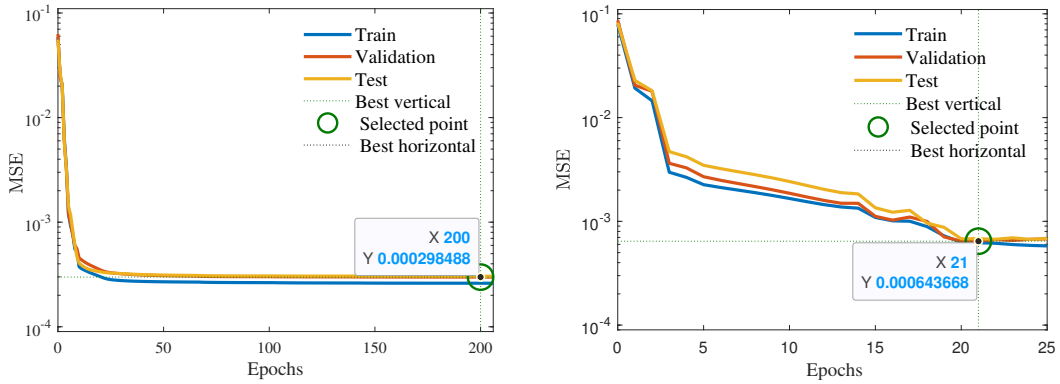


Figure 5.16: Training algorithms for ANN model.

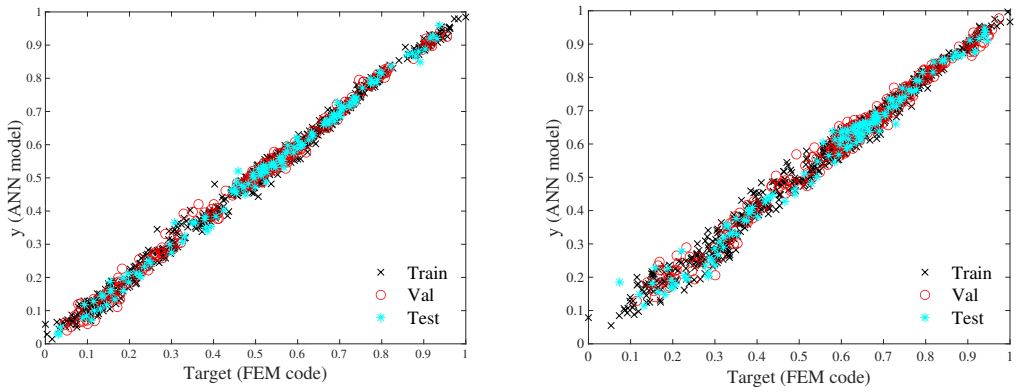
tion, training, and testing sets. The graphs present an acceptable matching between the numerical values and predicted values in homogeneous and layered soil profiles. This matching is driven by a systemic build for the machine learning model, from choosing the problem input parameters to selecting the most proper features for the ANNs algorithm.

As presented early in this work, the vibration induced by the pile driving process can not be only addressed in the space domain. But also, the vibration in the frequency domain gives a more important information of the phenomenon under investigation. This, it is important to expose the ANNs' ability to understand the phenomenon in the frequency domain, which is more complex process. The complexity is driven from increasing the input parameters of the database as exposes in the next section.



(a) (b)

Figure 5.17: Validation performance of the ANN: a) homogeneous soil profile; b) layered soil.



(a) (b)

Figure 5.18: Prediction vs. numerical values: a) homogenize soil profile; b) stratified soil.

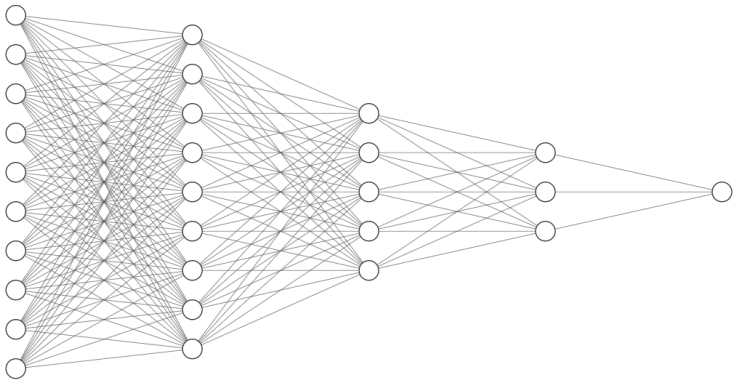
5.8.2 Vibration velocity prediction in frequency domain

As mentioned before, evaluating ground-borne vibration requires consideration of both partial peak velocity and the frequency content. It is essential to recall that the international standards, as explained in section 2.4.4, establish limits for partial peak velocity based on the frequency domain. Given the purpose of enhancing the accuracy of the artificial neural networks model assessment discussed in this work, it is crucial to construct a machine-learning model to predict the response in the frequency domain. Such contribution will increase the reliability of the ANNs model assessment and add more value to the present work.

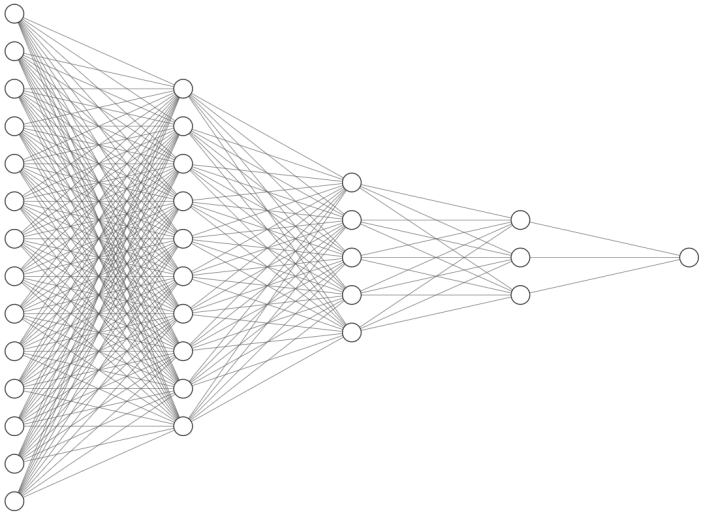
The author must emphasize that developing the ANNs model for predicting the frequency domain follows the same systematic strategy as exposed previously in developing PPV models, which were explained in full detail in Section 5.8.1. Notably, the complexity level of the generated model increased as the learning database for the frequency domain requires two additional inputs, frequency and space domain.

As a result, every case scenario was collected for 400 points in the frequency domain corresponding to rangs from 1 to 200 Hz. Therefore, the total number of cases in the database is 560,000. Such large and unbiased database would be used to train the ANNs model, as it presents in the following.

As mentioned before, the first step in identifying the neural network's architecture is determining the number of nodes and the number of hidden layers. The dimensions of the architecture control, to a large extent, the model's output. Therefore, a set of case scenarios was performed and evaluated. These trials resulted in the final architecture for the ANNs for homogeneous and two-layer soil profile models, see Figure 5.19. The figure illustrates that the network of the homogeneous profile consists of 3 hidden internal layers with 10:9:5:3:1, while 15:10:5:3:1 with three internal hidden layers for the stratified soil profile model. This advanced architecture makes the learning process more challenging and complex. Therefore, the proper term used to point out the learning process in this advanced form is the Deep Learning Process.



(a)

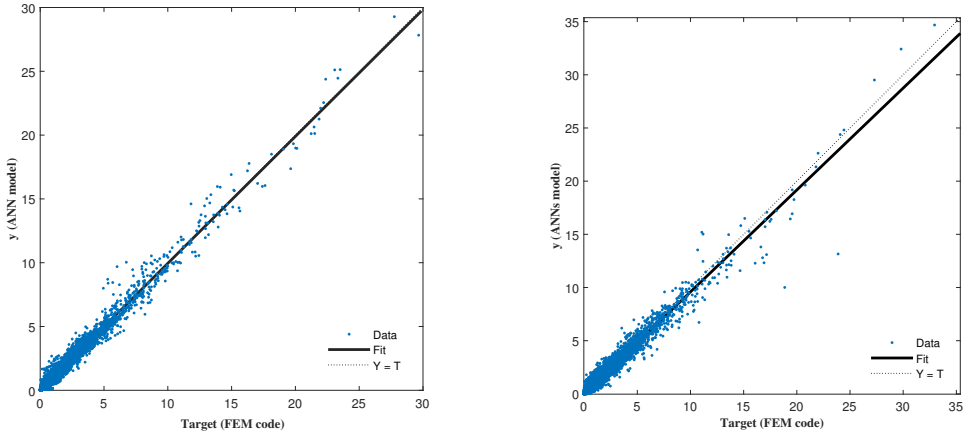


(b)

Figure 5.19: Architecture of ANN model for frequency domain: a) homogeneous; b) Layered.

In the second part, a set of trials and assessments was used to reach for the final properties

of the training model. The dividing ratio for the database is determined to be 40% for training, 30% for validation, and 30% for testing sets. The assessment also revealed that the "Levenberg-Marquardt" algorithm could be used as the training algorithm for both models. Both models were trained, and the stopping criteria were adjusted to be 1000 epoch or min MSE, which comes first. The training resulted in a model with R^2 equal 0.992 and RMSE equal 0.082 for homogeneous soil profile, while R^2 equal 0.985 to and RMSE equal 0.1 for two-layer layered soil profile. Figure 5.20 illustrates the regression graph, which compares the numerical results provided by the FEM-PML numerical model with the predicted results generated by the developed ANN model for the validation, training, and testing sets. The graphs demonstrate an acceptable matching level between the numerical and predicted values for homogeneous and layered soil profiles.



(a) (b)

Figure 5.20: Prediction vs. numerical values for frequency : a) homogeneous; b) layered.

5.9 Comparative study based on different machine learning models

5.9.1 Hyper-parameter optimization

Hyper-parameter optimization, often referred to as hyper-parameter tuning, is a crucial step in machine learning (ML) model development. It involves the process of finding the best combination of hyper-parameters for a given ML algorithm to maximize its performance on a specific task or dataset. Hyper-parameters are settings or configurations that are not learned from the data but must be specified before training a machine learning model. Examples of hyper-parameters include the learning rate in a neural network, the depth of a decision tree in a random forest, or the number of clusters in a k-means clustering algorithm [108].

The goal of hyperparameter optimization is to search through a predefined hyperparameter space efficiently and effectively to find the hyperparameters that result in the best model performance, (see Figure 5.21). This process typically involves the following steps:

1. Define a Hyperparameter Search Space: define the range or values that each hyperparameter can take. This space defines the scope of the search.

2. Select an Optimization Strategy: There are various strategies for searching the hyperparameter space, including grid search, random search, and more advanced techniques like Bayesian optimization and genetic algorithms. The choice of strategy depends on factors such as the computational resources available and the size of the search space.

3. Evaluation of Models: For each set of hyperparameters, you train and evaluate the model using a suitable evaluation metric (e.g., accuracy, F1 score, mean squared error, etc.) on a validation dataset or through techniques like cross-validation.

4. Iterative Search: repeat the process of training and evaluating models for different hyperparameter combinations, iteratively refining the search based on the results.

5. Select Best Hyperparameters: After the search is complete, choose the set of hyperparameters that resulted in the best model performance on the validation data.

6. Test on a Holdout Dataset: Finally, test the model with the selected hyperparameters on a separate, unseen test dataset to assess its generalization performance.

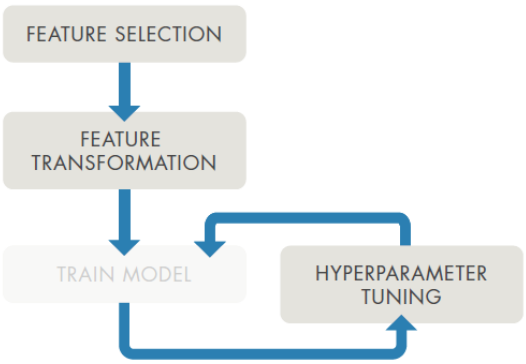


Figure 5.21: Schematic representation for hyper-parameters tuning during ML process.

After delving into the process of hyperparameter optimization, the subsequent phase involves execution. As explored earlier, algorithms dealing with extensive databases and undergoing iterative optimization justify the computational costs incurred. To undertake this task, a High-Performance Computing Unit (HPCU) is employed. The performance of the hyperparameter tuning process is monitored and depicted in Figures 5.22-5.24. These figures illustrate the iteration number versus mean squared error, and the optimized results are squared. The optimized answers are tabulated and presented in Table 5.4 for various soil profiles and in both the PPV and frequency domains.

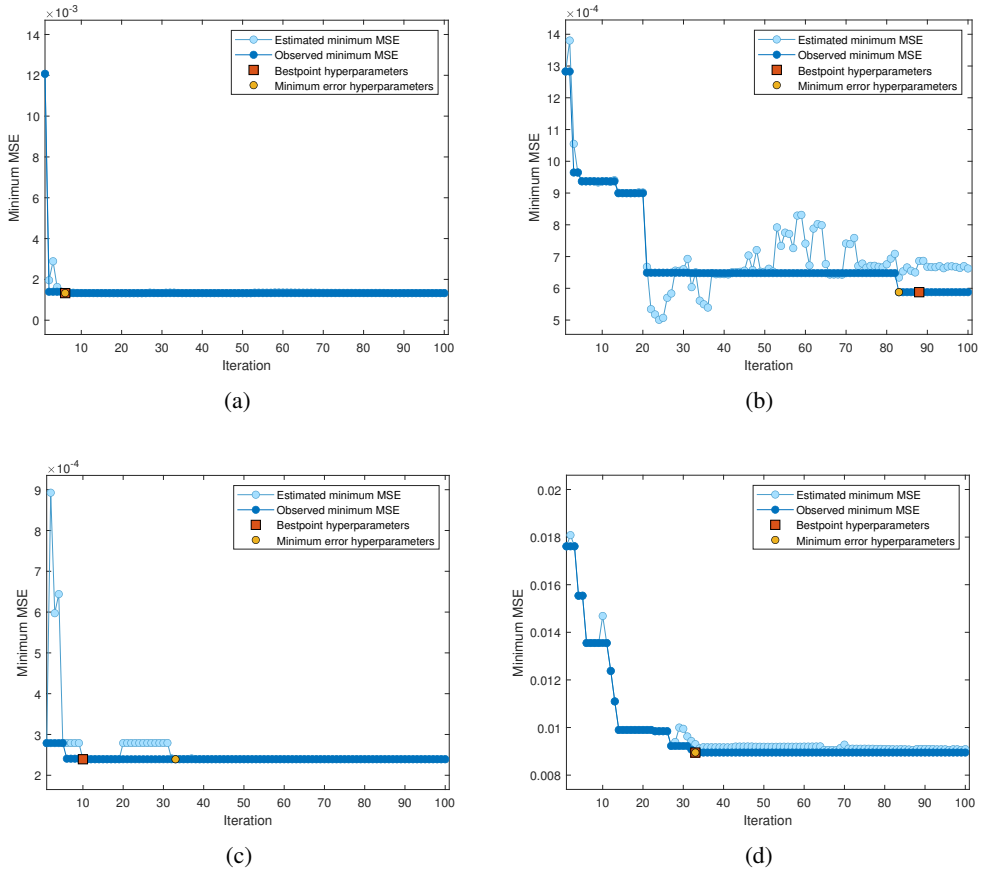


Figure 5.22: ML algorithms improving iteration vs. min. MSE for predicting PPV in homogeneous soil profile: a) regression tree; b) random forest; c) GPR; b) SVM.

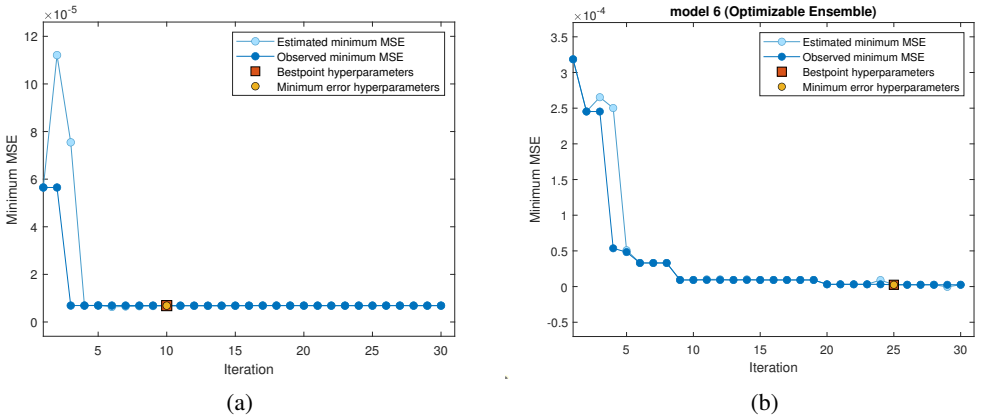


Figure 5.23: ML algorithms improving iteration vs. min. MSE for predicting frequency domain in homogeneous soil profile: a) regression tree; b) random forest.

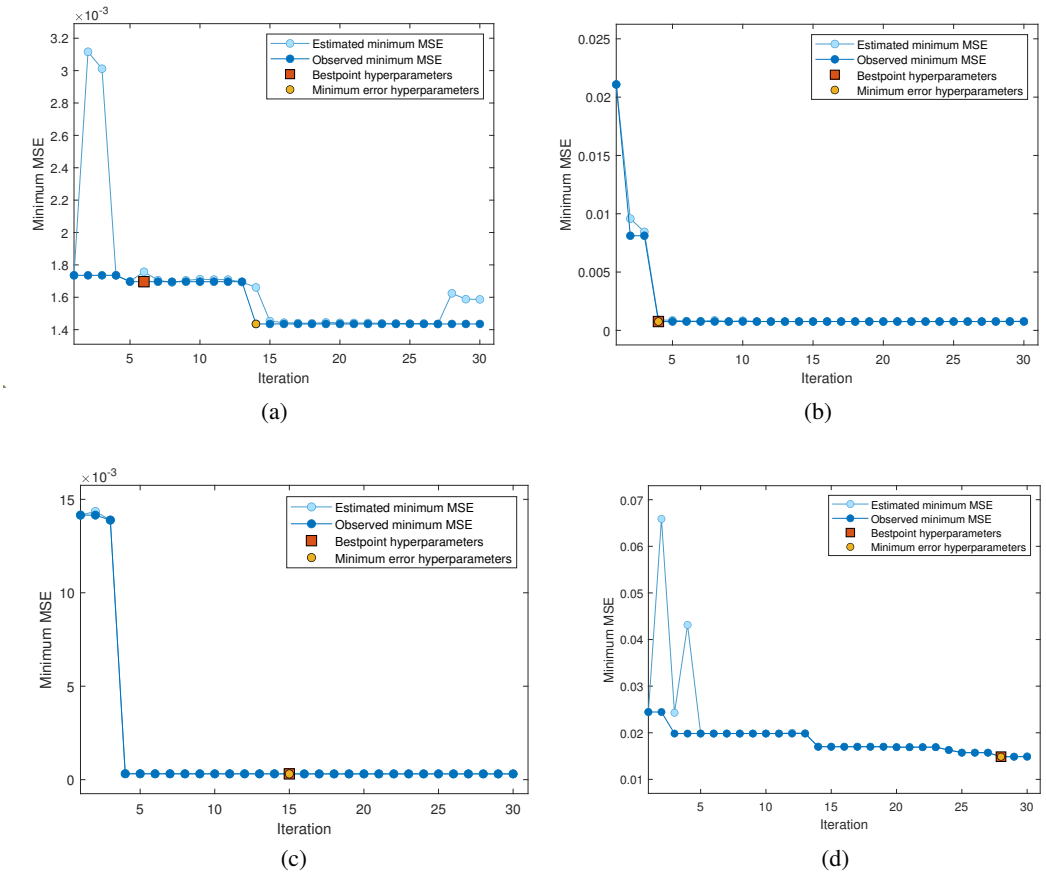


Figure 5.24: ML algorithms improving iteration vs. min. MSE for predicting PPV in stratified soil profile: a) regression tree; b) random forest; c) GPR; b) SVM.

5.9.2 Machine learning model performance assessment

As mentioned earlier in this chapter, the assessment of prediction performance for the developed machine learning models employs two key performance metrics: i) the coefficient of determination, R^2, and ii) the root mean square error, RMSE. Widely embraced in machine learning performance studies, these indicators are computed using the 'unseen' test dataset, as discussed earlier. Table 5.5 presents the values of the aforementioned metrics for both the training and test datasets, encompassing all machine learning algorithms used. The corresponding predictions versus numerical results are illustrated in Figure 5.25 - 5.28.

Overall, all models exhibit comparable predictive performance between the training and test sets, suggesting they are not prone to overfitting and are likely to generalize well to unseen data. With the exception of multiple regression and SVM, all other machine learning algorithms demonstrate high predictive power, as evidenced by the elevated R^2 values and low RMSE values presented in Table 5.5. Among the models considered, multiple regression stands out as the least effective, yielding a test set R^2 value ranging from 0.27 to 0.55. This performance is notably lower compared to analogous studies in vibration engineering, such as those referenced in [108].

Table 5.4: Optimised values of hyperparameters for regression models applied.

Training model	Hyperpartmer proprties			
	homogeneous profile		Two layer stratified profile	
	PPV	Frequency	PPV	Frequency
Multiple regression	Not applicable	Not applicable	Not applicable	Not applicable
Regression Tree	Minimum leaf size = 4	Minimum leaf size = 3	Minimum leaf size = 8	Minimum leaf size = 3
Tree forest	Ensemble method LSBoost	Ensemble method Bag	Ensemble method LSBoost	Ensemble method Bag
	Number of learners = 53	Number of learners = 31	Number of learners = 305	Number of learners = 54
	Learning rate = 0.6408	Learning rate = 0.6501	Learning rate = 0.0423	Learning rate = 0.6408
	Minimum leaf size = 4	Minimum leaf size = 1	Minimum leaf size = 1	Minimum leaf size = 1
	No. of predictors to sample = 7	No. of predictors to sample = 9	No. of predictors to sample = 2	No.of predictors to sample = 10
GPR	Sigma = 1.2156	Not applicable	Sigma = 1.3258	Not applicable
	Basis function is 0		Basis function is 0	
	Kernel function is Nonisotropic 5/2		Kernel function is Rational Quadratic	
SVM	Box const. = 22.74	Not applicable	Box const. = 0.0013	Not applicable
	Epsilon = 0.0724		Epsilon = 0.0205	
	Kernel function is Cubic		Kernel function is quadratic	

Table 5.5: Comparison of the predictive performance of ML algorithms.

Database name	Accuracy (Normalized RMSE $\times 10^{-3}/R^2$)					
	LR	*RT*	RF	GPR	SVM	XGBooster
Hom.PPV.Valid	120/0.27	36/0.93	27/0.96	15/0.99	0.094/0.55	0.3/0.98
Hom.PPV.Test	113/0.29	40/0.91	21/0.97	7/1.00	0.091/0.54	0.5/0.97
Hom.Freq.Valid	16/0.33	2.6/0.98	1.5/0.99	*N*/*A*	*N*/*A*	0.003/0.99
Hom.Freq.Test	16.5/0.32	2.7/0.98	1.8/0.99	*N*/*A*	*N*/*A*	0.004/0.99
Stra.PPV.Valid	126.2/0.242	43.5/0.91	30.4/0.96	17.5/0.99	121/0.29	0.9/0.96
Stra.PPV.Test	139.8/0.26	39.3/0.94	41/0.94	18.4/0.99	137.7/0.28	0.7/0.97
Stra.Freq.Valid	11.6/0.54	2.8/0.97	1.7/0.99	*N*/*A*	*N*/*A*	0.007/0.97
Stra.Freq.Test	11/0.55	2.4/0.98	1.7/0.99	*N*/*A*	*N*/*A*	0.007/0.97

A conspicuous disparity between predictions and numerical results is evident in Figure 5.25.a/d to Figure 5.28.a/d for Multiple Regression and Support Vector Machine. This incongruence suggests that traditional algorithms may struggle to capture the full spectrum of drift responses observed in the phenomenon under investigation. This limitation can be attributed, at least in part, to their inability to describe highly nonlinear processes and relationships between input features and drift demands. Conversely, all other ML models showcase comparatively higher prediction performance, with R^2 values exceeding 0.90 on both training and test sets across all algorithms and data sets. Nevertheless, the GPR and SVM algorithms in the frequency domain struggle to capture the data trends within the learning dataset. This behavior can be attributed to their limited

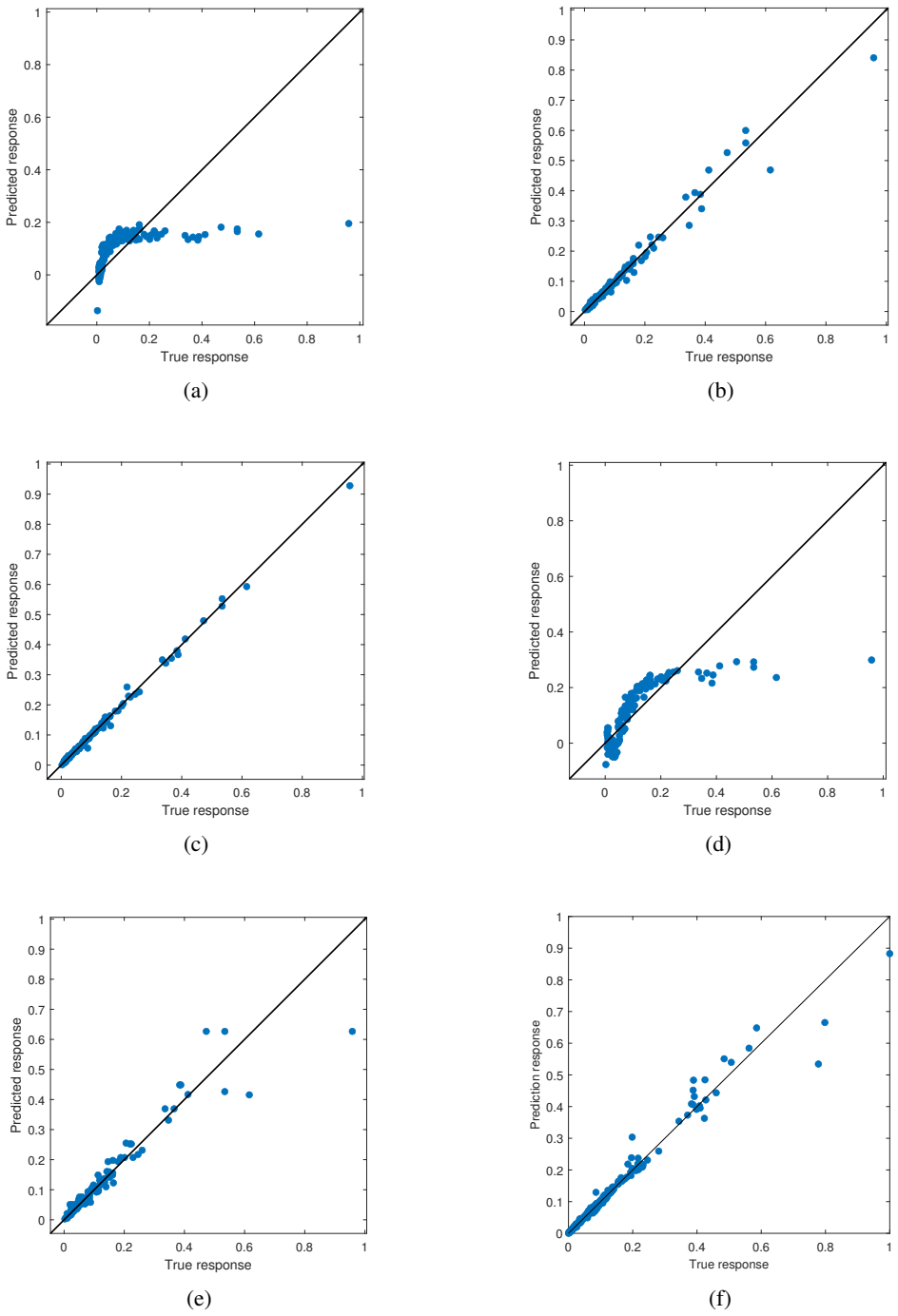


Figure 5.25: Performance of ML models for predicting PPV in homogeneous soil profile: a) multi linear regression; b) random forest; c) GPR; d) SVM; e) regression tree; f) XGBooster.

ability to handle big data, as evidenced in the frequency domain where they fall short, particularly when faced with up to 0.5 million cases and an escalating level of complexity.

In the context of a comparative study, an in-depth analysis was conducted, and the results presented in Table 5.5, revealing the superior performance of the XGBooster algorithm, see Table

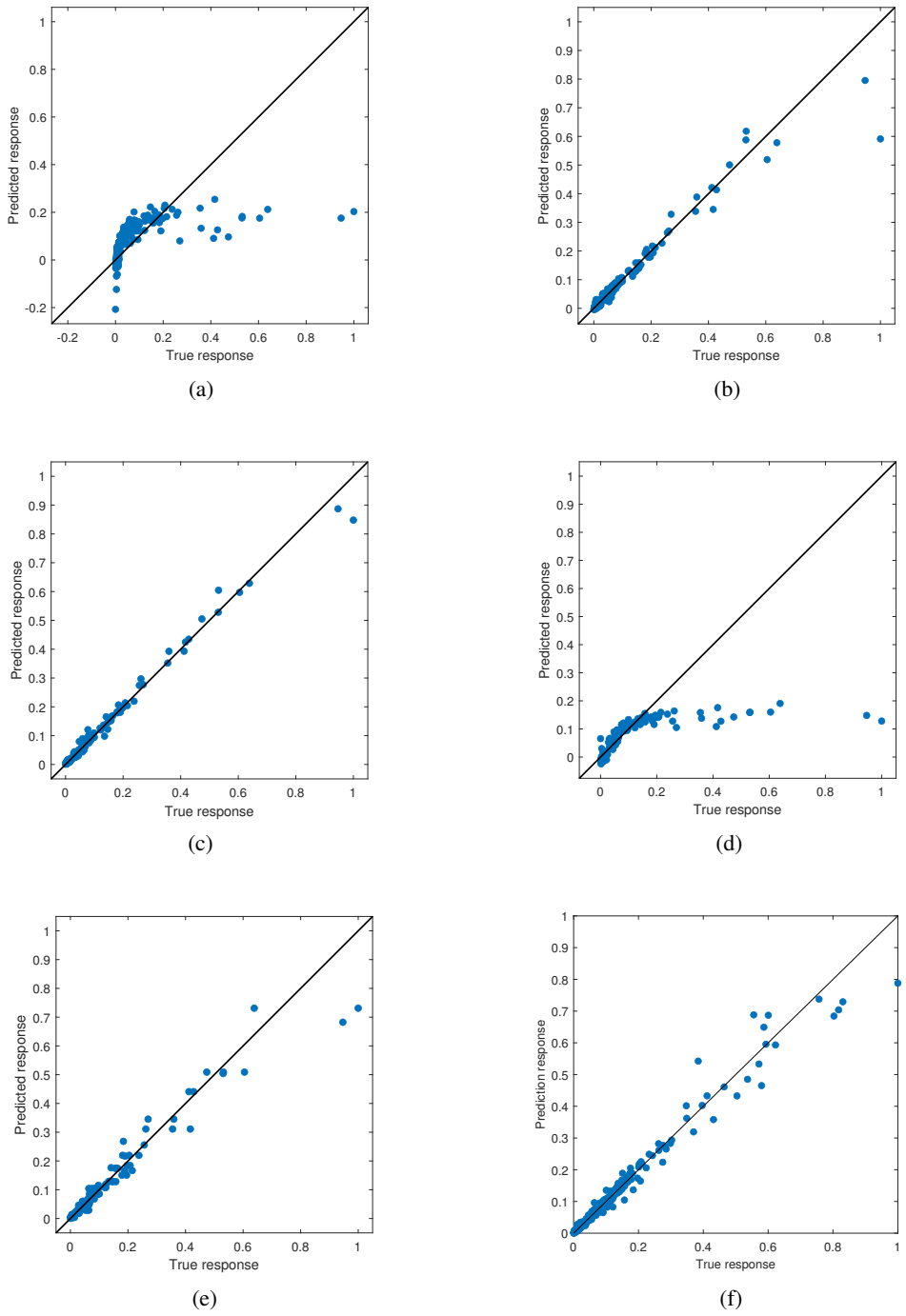


Figure 5.26: Performance of ML models for predicting PPV in stratified soil profile: a) multi linear regression; b) random forest; c) GPR; d) SVM; e) regression tree; f) XGBooster.

5.6. Remarkably, XGBooster exhibited the highest R^2 score, indicative of its exceptional ability to explain the variance in the data, coupled with the lowest Root Mean Square Error, signifying its precision in predicting the observed values. This notable achievement is attributed to XGBooster's inherent capability to adeptly navigate high levels of complexity within the dataset. The

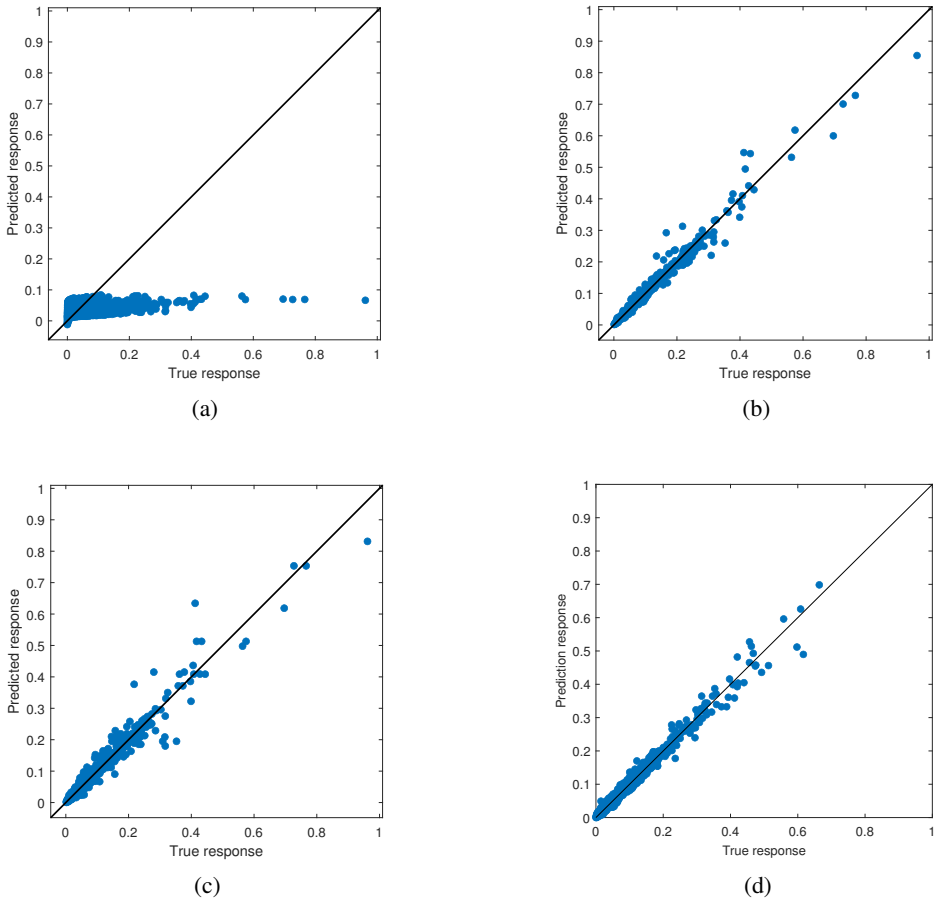


Figure 5.27: Performance of ML models for predicting frequency domain in homogenise soil profile: a) multi linear regression; b) random forest; c) regression tree; d) XGBooster.

Table 5.6: Table comparison between ANNs and XGBooster.

Training model	Accuracy (Normalized RMSE $\times 10^{-3}/R^2$) - Test Set			
	homogeneous profile		Two layer stratified profile	
	PPV	Frequancy	PPV	Frequancy
ANNs	0.028/0.996	2.7/0.992	0.042/0.994	2.8/0.985
XGBooster	0.5/0.97	0.004/0.99	0.7/0.97	0.007/0.97

algorithm's resilience in handling intricate patterns and relationships in the data positions it as a robust contender for predictive modeling, especially ANNs. Encouraged by these promising outcomes, the author would delve further into the predictive modeling arena by comparing the efficacy of Artificial Neural Networks and XGBooster in tackling the high-complexity prediction challenge. Specifically, this study seeks to unravel the intricate phenomenon of predicting vibrations induced by pile processes within a soil profile comprising three distinct soil layers, adding depth and specificity to the predictive modeling exploration, as showen it the next section.

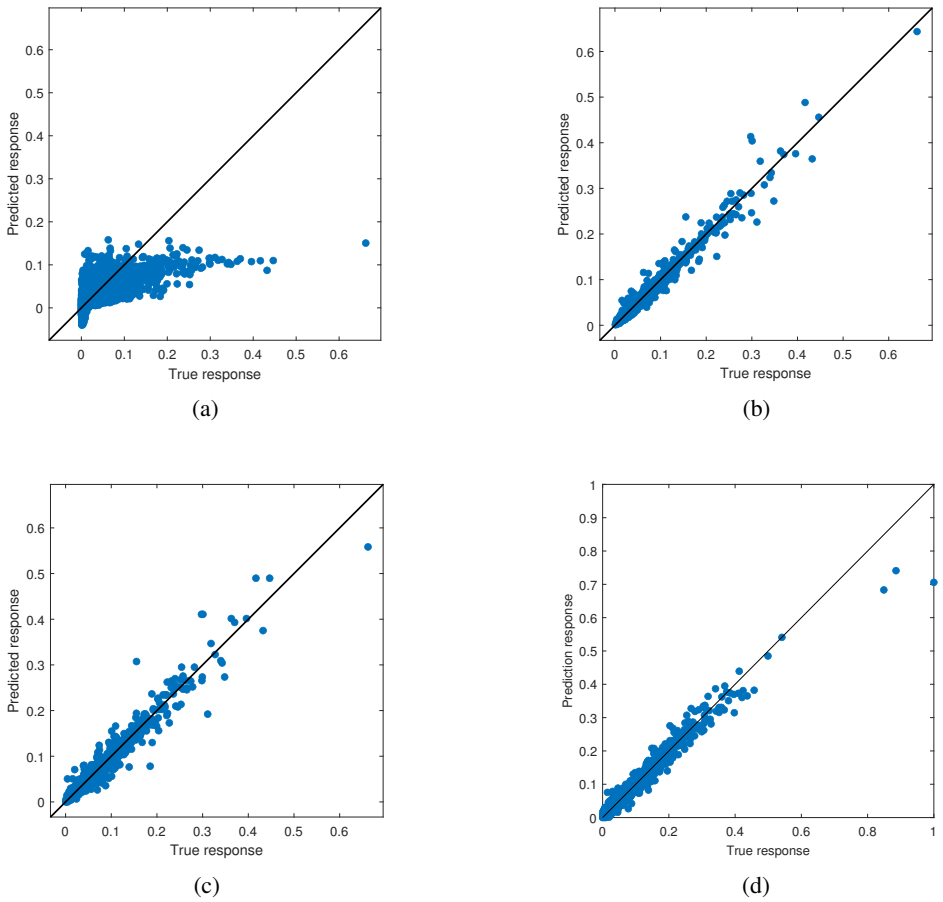


Figure 5.28: Performance of ML models for predicting frequency domain in stratified soil profile: a) multi linear regression; b) random forest; c) regression tree; d) XGBooster.

5.10 Machine learning model for soil profile with 3 layers

5.10.1 General introduction

As previously exposed, a comprehensive study was conducted involving seven distinct algorithms. The findings of this study revealed that Artificial Neural Networks, ANNs, and the XGBoost algorithm exhibited the highest predictive performance. Encouraged by these results, the author intends to delve deeper into investigating the efficacy of Artificial Neural Networks and XGBoost by undertaking more complex geotechnical problems.

The challenge entails predicting the vibrations induced by the pile driving process for a soil profile consisting of three distinct soil layers. The inclusion of these layers elevates the complexity of the phenomenon. To navigate through this process, a corresponding database must be generated. The properties of this database are elucidated in the following section.

In addition, the development of the machine learning model and a comparison between ANNs and XGBoost in terms of revealing data trends and performance will be presented. Finally, the developed models will be experimentally validated using the results obtained earlier in this work.

5.10.2 Learning data set development

As discussed earlier, the development of the learning dataset is a crucial phase in creating a machine learning model. The learning database should accurately reflect the realities of a construction site. While it is essential to focus on common occurrences, emphasizing rare events excessively in the database can lead to a misleading model. Therefore, the database must cover a broad spectrum and include a representation of various possible scenarios to generate a reliable machine learning model.

The databases intended for generation should adhere to the same features and prediction scope outlined in Section 5.2.1. Table 5.7 presents the parameters of the learning dataset for a stratified soil profile consisting of three layers. Additionally, the parameters related to pile properties and the hammering process, such as pile diameter (ranging from 0.2 to 0.5 meters), pile penetration depth (ranging from 0.5 to 15 meters), drop height (ranging from 0.3 to 1 meter), and hammer efficiency (ranging from 0.6 to 0.95). These parameters encompass the most common scenarios encountered in the field. One hundred cases were generated. To accomplish this task, a High-Performance Computing Unit, HPCU, is employed. The HPCU performs this task in over 500 hours.

Table 5.7: The values for the learning database input parameters for three layer soil profile.

Item	Unite	Distribution	1^{st} layer soil		2^{nd} layer soil		3^{th} layer soil	
			Max.	Min.	Max.	Min.	Max.	Min.
Shear wave velocity	m/sec	uniform	200	60	800	200	800	600
Poisson's ratio	-	normal	0.49	0.25	0.49	0.25	0.3	
Damping ratio	%	normal	0.02	0.08	0.02	0.04	0.02	
layer thickness	m	normal	5	0.5	6	3	*Inf.*	

As the complexity of the phenomena increases, the algorithms require more cases to identify data trends. Therefore, for each case, the author collects 20 points (covering a length of 50 meters) for the PPV dataset and 600 points for each point in the frequency domain. As a result, the total number of cases for PPV dataset is 2000, and for frequencies dataset is 1.2 million cases. These sets will be used for training the machine learning models, namely XGBoost and ANNs. The results will contribute to understanding the performance of these algorithms in dealing with advanced levels of complex phenomena.

5.10.3 Development of the machine learning model

For the initially established purpose, a systematic approach was employed to develop, implement, and prepare extensive learning datasets for the subsequent training process. The training process will be conducted using MATLAB and Python for Artificial Neural Networks (ANNs) and XGBoost algorithms, respectively. Given the substantial size of the learning dataset, comprising dimensions of 2000×16 for Peak Particle Velocity and $1.2E6 \times 18$ for frequency domain data, the utilization of a High-Performance Computing Unit, HPCU, is imperative in the training process. The training process underwent hyper-parameter tuning to optimize the model's performance.

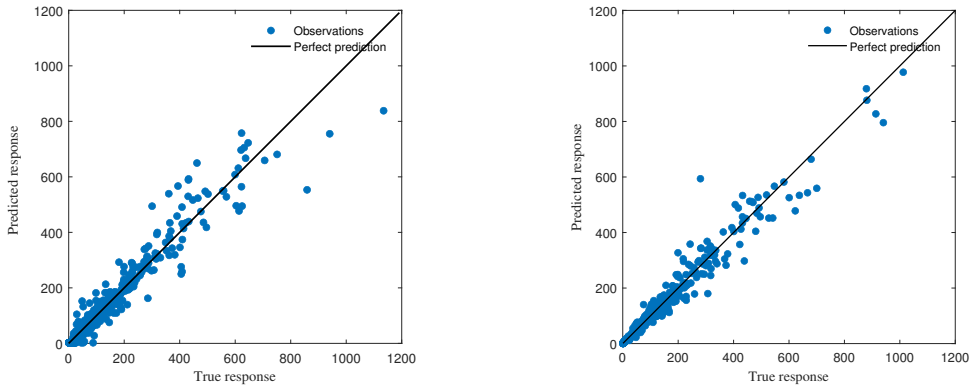
Model performance was assessed using Root Mean Squared Error, RMSE, and the coefficient of determination, R^2. In the subsequent sections, the trained models are introduced from various perspectives.

The learning set for Peak Particle Velocity underwent an extensive training process employing both Artificial Neural Networks, ANNs, and the XGBoost algorithm. The architecture of the ANNs comprises three hidden layers, with the number of nodes configured as 153:1:4 for the respective layers, reflecting the complexity required for investigating the phenomenon.

The training times for ANNs and XGBoost were recorded at 520 and 330 seconds, respectively, on the High-Performance Computing Unit. Both models demonstrated a prediction rate exceeding 5000 observations per second. The learning dataset was partitioned into training, validation, and test sets in a 60:25:15 ratio for both algorithms. The normalized Root Mean Squared Error (RMSE) and R^2 for the PPV model are presented in Table 5.8. The table shows a notable contrast between the results of the validation sets and the test set conclude that both models effectively avoided overfitting scenarios. Figure 5.29 presents the performance of the machine learning models in a scatter presentation.

Table 5.8: Table comparison between ANNs and XGBooster for three layer soil profile.

Training model	Accuracy (Normalized RMSE $\times 10^{-3}/R^2$)			
	Testing set		Validation set	
	PPV	Frequency	PPV	Frequency
ANNs	48/0.92	3.8/0.88	42/0.93	3.6/0.88
XGBooster	0.034/0.96	0.004/0.99	0.039/0.93	0.0131/0.95



(a) (b)

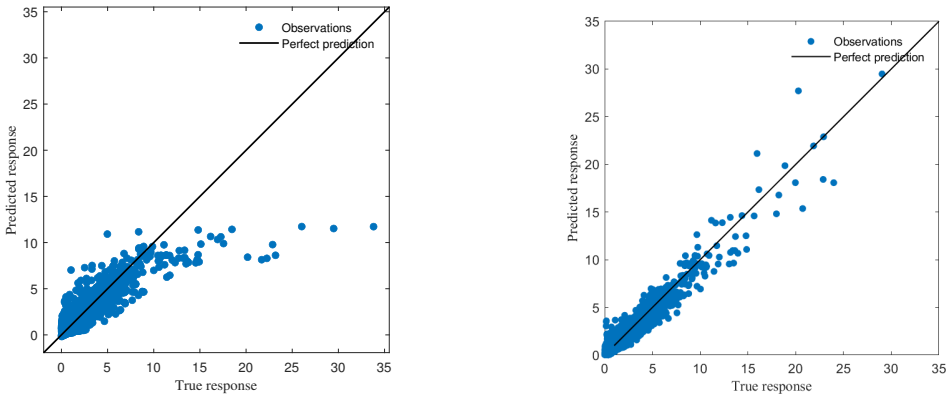
Figure 5.29: Prediction vs. numerical values for PPV models : a) ANNS; b) XGBooster.

The dataset for velocity in the frequency domain underwent training on both Artificial Neural Networks (ANNs) and the XGBoost algorithm. This dataset is considerably more complex than the PPV dataset, as discerning data trends for a matrix consisting of $1.2E6 \times 18$ cases poses a challenging task. The architecture of the ANNs comprises two hidden layers, with the number of nodes configured as 88:14 for the respective layers. Training times for ANNs and XGBoost were

recorded at 3.5E5 (4 days) and 900 seconds, respectively, on the High-Performance Computing Unit. Both models demonstrated a prediction rate exceeding 3000 observations per second.

The learning dataset was partitioned into training, validation, and test sets in a 55:25:20 ratio for both algorithms. Normalized Root Mean Squared Error (RMSE) and R^2 for the PPV model are presented in Table 5.8. The table reveals a notable contrast between the results of the validation sets and the test set, concluding that both models effectively avoided overfitting scenarios. Additionally, the R^2 of the ANNs is below 0.9, indicating that ANNs may encounter misleading scenarios as the complexity level increases. From another perspective, Figure 5.30 illustrates the performance of the ANNs model and XGBoost in a scatter graph.

As depicted in the figure, ANN models exhibit misleading results in the high amplitude values, which are not observed in the XGBoost algorithms. This conclusion indicates a limited ability for the ANNs model to avoid overestimated or underestimated scenarios for high amplitude in the frequency domain.



(a) (b)

Figure 5.30: Prediction vs. numerical values for frequency models : a) ANNs; b) XGBoost.

The preceding exploration highlighted the limited ability of the ANNs algorithms to comprehend data trends when faced with an increased level of complexity. Conversely, XGBooster demonstrated acceptable performance in handling high complexity levels. To further assess both the ANNs model and XGBooster, they underwent experimental validation. This validation aims to understand how reliable these models are for practical use, as elucidated in the following section.

5.11 Experimental validation

The previous exploration introduced machine learning models based on two different algorithms: ANNs and XGBooster. A comparative study was conducted to understand the differences, strengths, and weaknesses of each model. This section provides the experimental assessment for these models. The importance of this assessment arises from understanding how reliable these models can be as tools for rapid prediction of vibrations induced by the pile driving process, which is the

main purpose outlined in this chapter. The following presents the exploration of experimental validation.

The author will recap some relevant information corresponding to the experiment test site. Non-intrusive geophysical tests were performed at the test site: for the ground characteristics, a Seismic Refraction Testes and SASW tests (Spectral Analysis of Surface Waves) were performed. It was assumed that the damping ratio of the soil is 0.025 and that the soil has a mass density close to 1900 kg/m3.

Concerning the foundations of the building, a total of 160 quadrangular piles were driven with two distinct sections of 400 × 400 mm and 350 × 350 mm. These piles are made of precast concrete, with Young modulus, E, around 30 GPa and mass density of 2500 kg/m3. A hysteretic damping factor equal to 0.01 is considered for the pile and the value of 0.15 is assumed for the Poisson's ratio. The piles have a total length ranging from 8 to 15 meters. The driving equipment corresponds to a Junttan PM25 equipped with the SHK110-7 hydraulic impact hammer with a drop height of 0.4 meters. The mass of the hammer ram is equal to 7 tons. Based on this exposition, the authors should highlight that all test site parameters fall within the limits of the learning database.

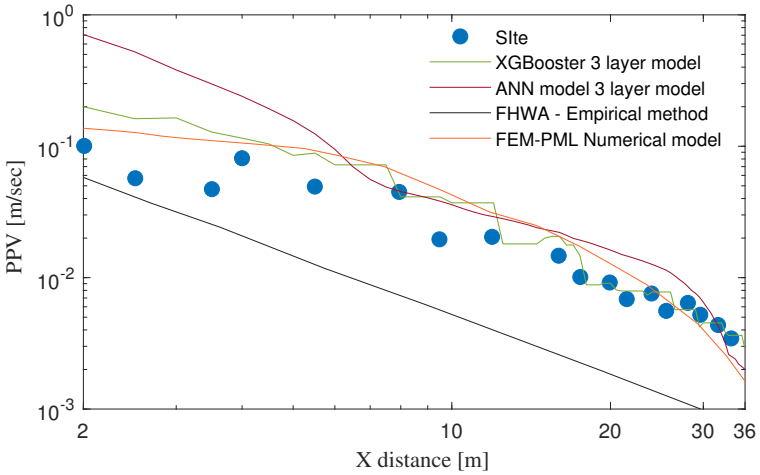


Figure 5.31: Experimental validation for ANNs and XGBoost PPV model.

Figure 5.31 illustrates the experimental validation of the machine learning models and the empirical method recommend by the FHWA in predicting PPV values in the space domain. It is possible to visualize a strong correlation between the experimental results and the predictions from both the numerical model and the XGBoost algorithm across all evaluated distances. However, it highlights an overestimation by the ANNs model in predicting vibrations induced by pile driving close to the source of excitation, and the underestimation of the empirical method. The over-prediction of the PPV for X $\leq$ 5m by the ANN algorithm can be attributed to its interpretation of the data trend, where it builds a single model to describe the entire database. Since the spatial domain is large, the number of collected points at short distances is relatively small compared to the rest of the space domain, leading to a larger representation in the linear zone and a smaller one

in the nonlinear zone (closer to the pile). This imbalance in data representation affects the ANN's ability to accurately predict near the pile, causing it to overestimate the PPV. In contrast, XGBoost builds sub-models that predict each group of data separately, justifying its superior performance in the zone near the pile. While it might be challenging to find structural locations affected by excessive vibration in this area, the XGBoost algorithm exhibits a notable match with the experimental site. This is because XGBoost works by focusing on smaller sections of data, allowing it to capture the localized behavior more effectively. Moreover, the accuracy metrics used in the analysis are somewhat misleading due to the structure of the database. The sparse representation of data points near the pile reduces the impact of this critical area on the overall accuracy metrics. This justifies why the accuracy metrics in Table 5.8 appear strong for both algorithms, despite potential local inaccuracies observed in ANNs' algorithm.

Regarding the frequency domain, the previous assessment suggests that the XGBoost results are superior and more closely aligned with the experimental site compared to the results derived from the ANNs, especially in terms of amplitude levels at 20 m from the excitation source. Figure 5.32 presents the experimental validations for the frequency domain models, confirming these findings. The ANNs model struggles to predict the velocity in the frequency domain, both in terms of amplitude and in the dominant frequency range. In contrast, the XGBoost shows an acceptable match with the experimental results, as depicted in Figure 5.32.

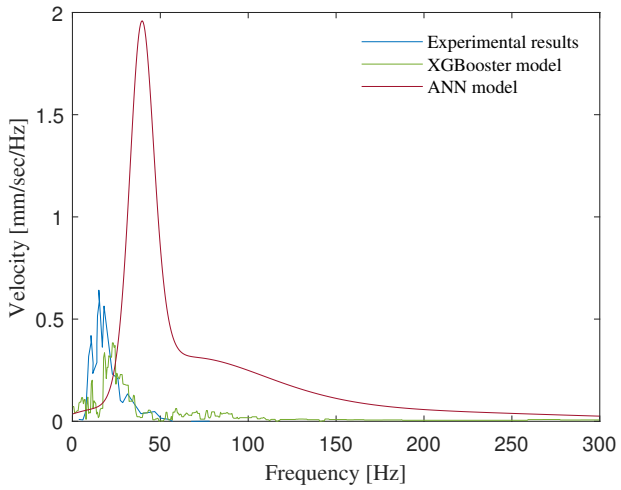


Figure 5.32: Experimental validation for ANNs and XGBoost frequency model at 8 m.

5.12 ANNs application in vibration mitigation

The process of pile driving can induce ground-borne vibrations that, in extreme cases, pose a threat to the stability and integrity of nearby structures. Therefore, it is important to predict the vibration levels at an early stage and, if necessary, employ vibration mitigation techniques to control them.

These techniques aim to ensure that the vibration level remains within acceptable limits as defined by international standards.

In this regard, the author would like to draw attention to a comparative study between two vibration mitigation techniques, namely the use of an open trench and a heavy mass placed on the ground surface (refer to Section 4.5). The study found that an open trench is more effective in reducing vibration levels than placing a heavy mass on the ground surface. This result justifies the use of artificial neural networks, ANNs, to test their ability in vibration mitigation prediction. This section outlines the process of creating a surrogate model based on ANNs for predicting ground surface vibrations while using an open trench as a vibration mitigation solution.

As previously mentioned, the first step in developing the ANNs model is to generate the learning database using the FEM-PML model. It is important to note that the input database in Table 5.2 will be used, and a total of 100 cases will be generated for homogeneous soil and 100 cases for stratified soil (two layers). However, for this ANNs model, there will be a new addition of three inputs that correspond to the trench dimensions and position. These new inputs are Htrench, Wtrench, and Xtrench, as shown in Figure 5.33. The values for these two new inputs are uniformly distributed within different ranges. Htrench have values of 0.5, 1.0, 1.5, and 2.0 meters, while Xtrench has values ranging from 5.0 to 10.0. The process for developing this ANNs model will be the same as before, with the only difference add of these new inputs.

From another perspective, adding this mitigation methodology, open trench, to the phenomena under study pushes the level of the phenomenon complexity up. Therefore, the author increases the number of observation points in the space domain to 370 points that cover 50 meters from the center of the driving pile. Such an increase reach the learning database 37,000 cases. This increase enhances the ability of the ANNs to find the data trending.

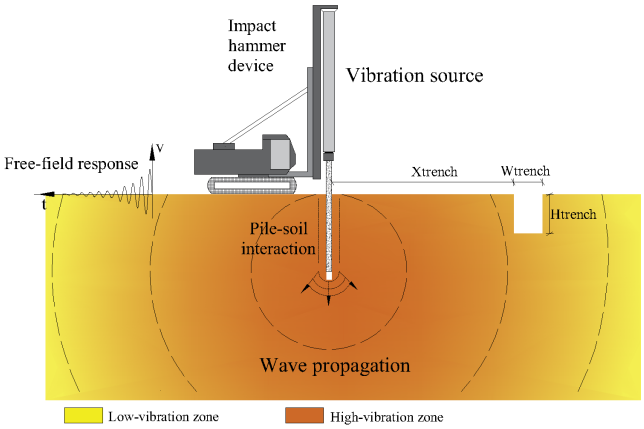
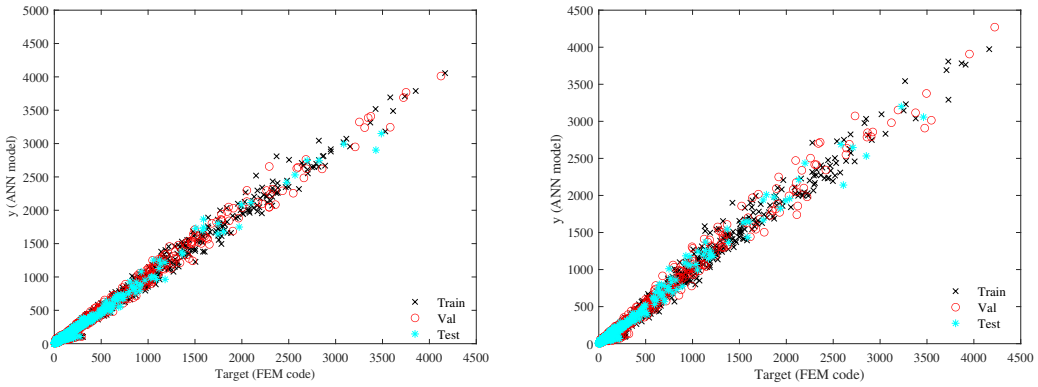


Figure 5.33: Schematic presentation for driven pile while using open trench.

In developing the architecture of the ANN model, a systematic strategy was employed. As

a result, it was determined that a network with a single layer containing three neurons would be used for a homogenous soil profile, while a single layer with seven neurons would be used for a stratified soil profile. After assessing the situation, a dividing ratio of 60% for training, 30% for validation, and 10% for testing sets were deemed appropriate. Additionally, the "Levenberg-Marquardt" algorithm was found to be suitable as the training algorithm for both models. Figure 5.34 present a comparison between the ANN model and FEM-PML model response in both profiles. The Figure presents an acceptable matching and refers to remarkable addressing for the main features of the problem.



(a) (b)

Figure 5.34: Predicate vs. numerical values for a soil profile with open-trench vibration mitigation measure: a) homogeneous; b) stratified profile.

5.13 Standalone computer application

To facilitate user interaction with the previous machine learning algorithms, a user-friendly graphical interface has been developed in Python, integrating ANNs and XGBoost models in a standalone program. This program combines ease of use with a powerful prediction capability, delivering results in real-time. The software program, named ZBZBAT, features a simple and intuitive interface (see Figure 5.35), enabling users to input soil profile data and hammer-pile parameters easily. It also allows users to visualize predictions generated by the ANNs and XGBoost models.

The interface of ZBZBAT includes four main input modules:

1. **General Input Parameters:** this section includes three subsections detailing key properties of the pile-driving process:
    - **Pile Properties:** pile diameter and penetration depth.
    - **Hammering Properties:** Drop height.
    - **Soil Profile:** Shear wave velocity (Cs), Poisson's coefficient, damping ratio, and layer thickness.

2. **Frequency Domain Response:** this module introduces an additional input: the location of the evaluation point to compute the frequency response, corresponding to the distance between the pile and the evaluation point.

3. **Standard Limits:** the tool features three standard limits:

   - Portuguese Standard NP2074
   - German Standard DIN 4150-3
   - British Standard BS 7385-2

   For each standard, inputs include the standard name, structure type, and dominant frequency range.

4. **Mitigation Measure (Open Trench):** this module requires inputs for trench parameters:

   - Trench Depth: options of 0.5, 1, 1.5, and 2 meters.
   - Distance to the Pile: ranging from 5 to 10 meters.

Regarding the output information, there are two graph plot zones providing visual representation of the results:

1. PPV (Peak Particle Velocity) values as a function of the distance from the pile;

2. Frequency spectrum for chosen evaluation point.

5.13.1 ZBZBAT copyright registration

ZBZBAT is registered as Intellectual Property (IP) under registration number 1701/2024 on 14/08/2024, awarded by IGAC. IGAC is a public entity directly accountable to the government official responsible for Culture. In Portugal, IGAC specializes in copyright and related rights.

5.13.2 ZBZBAT website

A website is developed for ZBZBAT in order to make it easy for engineering community to reach out the software (Figure 5.36). In the website, all relevant publications that stand behind ZBZBAT are listed. As well, introduction for the software is presented and a link for download. The website link is : https://sites.google.com/view/zbzbat/home

5.14 Final conclusions

The aim of this chapter is to provide a clear and comprehensive explanation of the process for creating a surrogate model using machine learning techniques, which can be used to rapidly predict the vibration levels induced by pile driving. An extensive database was created through massive

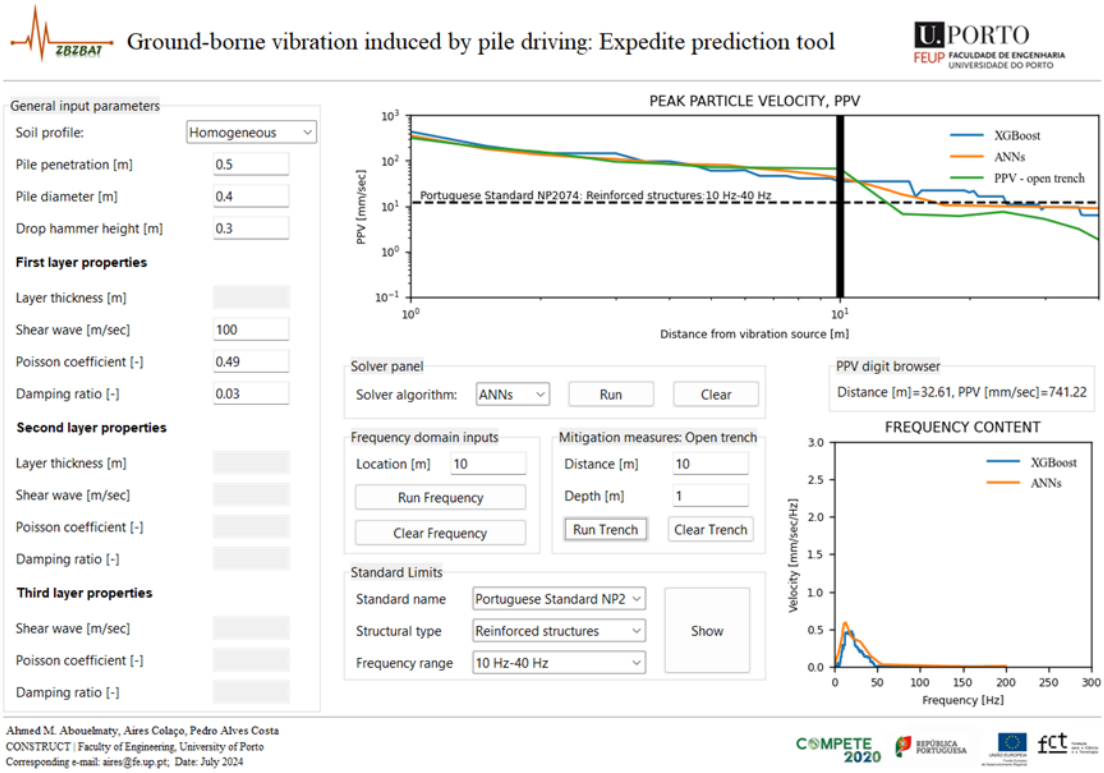


Figure 5.35: ZBZBAT computer software interface.

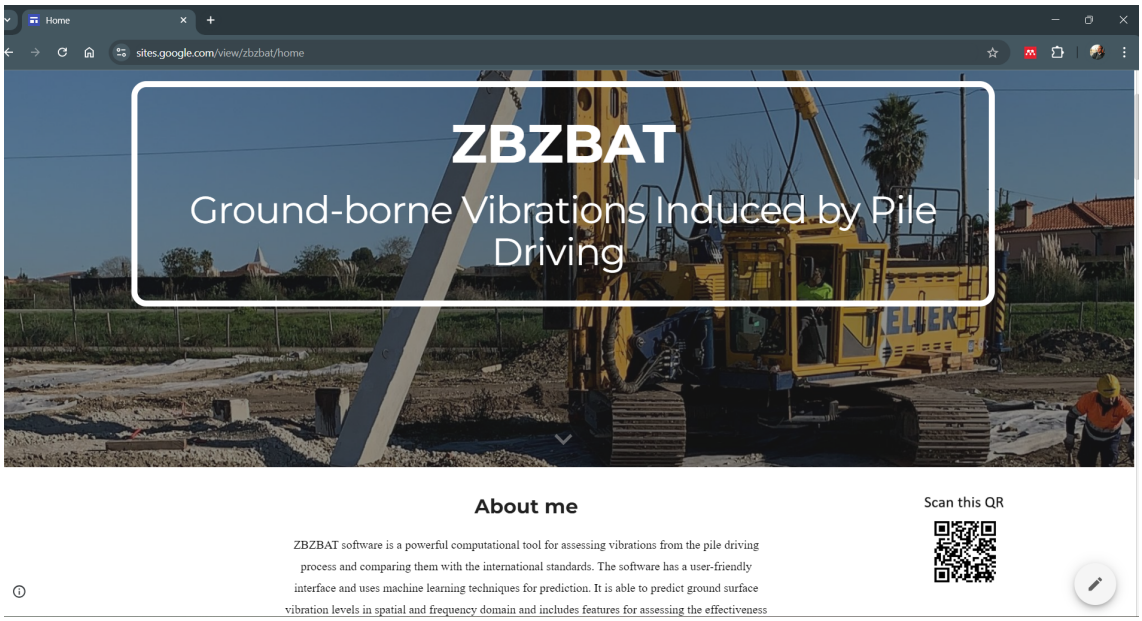


Figure 5.36: ZBZBAT website.

numerical simulations that aim to represent the most typical situations that can be faced in this type of problem. This data was provided by applying the equivalent linear FEM-PML approach pre-

sented previously. The resulting data from the numerical model was passed through a data mining process to generate the learning database for the machine learning model. The model was trained on seven different algorithms systematically to identify the algorithms that could catch the data trends. The model was trained to predict the peak particle velocity and the velocity in the frequency domain. The study showed that artificial neural networks and XGboost are accepted matches in the address phenomenon. However, in increasing the complexity level, XGboost demonstrated a greater understanding of the data trend than ANNs.

Moreover, this chapter details the steps taken to create an artificial intelligent model capable of predicting vibration on the ground surface when using an open trench as a vibration mitigation measures. The model was trained using ANNs and demonstrated a high level of agreement with the numerical approach.

To make the generated models accessible and easy to use, the author developed ZBZBAT, a stand-alone computer software that incorporates the best generated models after assessment into a single user-friendly interface with straightforward data presentation.

Finally, an experimental validation process was conducted for the generated models. The experimental results reveal that XGBoost demonstrated a better alignment with the site results compared to artificial neural networks (ANNs). This discrepancy in performance is attributed to a nuanced understanding of the algorithms for capturing data trends.

While ANNs established a comprehensive relationship between all input variables and the output, XGBoost partitioned the database into sub-models. This distinctive feature empowered XGBoost to discern and account for both linear and non-linear effects within the space domain, thus enhancing its overall performance.

Chapter 6

Conclusions and future works

6.1 Introduction

Ground-borne vibrations induced by the pile driving process are a significant phenomenon that needs to be place under investigation. These vibrations, in extreme cases, can threaten the integrity and stability of nearby structures. The present work outlines the steps taken to develop a surrogate model for the expedited prediction of ground-borne vibrations induced by the pile driving process. Simulating these vibrations can help identify potential risks to surrounding utilities and structures, which is particularly crucial in densely populated urban areas where buildings are closely spaced. This chapter will present a summary of the work performed and future works.

6.2 Work performed summary

An expanded literature review has been elaborated in five main aspects: the generality of aspects inherent to the generation and propagation of vibrations induced by pile driving; prediction and measurement of vibrations; and the contribution of the machine learning techniques in the prediction process; key indicators; and international standards. The literature review was synthesized to stand on the previous contributions for the research corresponding to the phenomenon under investigation. The literature review exposes the complexity of the phenomenon under investigation, which depends on several factors. Many of them are difficult to incorporate into a rational model, justifying the challenge of developing a prediction tool for the corresponding behavior in cooperation with the artificial intelligence application. Despite the high complexity of the problem, the current state of knowledge allows conceptual interpretation for the phenomenon, which helped develop empirical and numerical models for predicting the ground-borne vibrations. The information collected in the state-of-the-art made it possible to identify an opportunity to develop surrogate models based on a cooperation between numerical models and artificial intelligence techniques. The models aimed to predict the ground-borne vibration induced from several excitation sources,

e.g., blasting, high-speed trains, and highways. However, there is an absence of a valuable contribution in predicting the vibration induced by the pile driving process, which considers the research gap for the present work.

The linear and equivalent linear axisymmetric FEM-PML approach is presented, developed by the research group, for the prediction of free field vibrations due to impact pile driving. The numerical model is essential for the thesis's main goal, developing a prediction method based on the AI technique, where the numerical model will be applied to analyze the main aspects of the pile driving process, identify fundamental variables, and generate the database. For the purpose initially set, comprehending, practicing, and applying the numerical model are essential steps during the exposition of this work. The numerical model was subjected for a numerical evaluation with the numerical model reported by Masoumi et al. [55]. From the numerical validation, there is good matching between models. In addition, the author intended to present the main behavior of the hammer-pile-ground system from the parameters perspective. For that purpose, a parametric study was performed for nine different parameters to examine their influence on the ground response. The influence of the parameters were placed under study using eight different types of comparative study to identify the main variables that interfere with the ground response.

A comprehensive experimental test site was implemented in Aveiro, Portugal, to understand the phenomenon deeply. The performed experiments were grouped into two components: i) geotechnical investigation; ii) measurement of vibrations induced by pile driving. From the low stain test and combined with FEM-PML numerical tool, low strain test was performed to identify the dynamic soil characteristic and the soil profile. The collected data from the driving process was numerically treated, presented graphically, and assessed by the Portuguese Standard NP2074. Experimental validation was performed using the numerical prediction tool. A good agreement was observed between the experimental data and the prediction performed by the equivalent linear numerical model.

A study on vibration mitigation techniques was undertaken, focusing on a comparative analysis between two widely used methods: open trench and placing heavyweight on the ground surface. The study used the FEM-PML numerical model. It was concluded that both approaches are effective in preventing ground surface vibrations. However, the open trench technique demonstrated greater influence in attenuating vibrations on the ground surface.

Exploring the Development of a Machine Learning-Based Model to accelerate vibration predictions induced by pile driving. An extensive database was established through extensive numerical simulations, designed to capture the most typical situations encountered in this type of problem. This data was generated by employing the equivalent linear axisymmetric FEM-PML approach as previously presented. The resulting data from the numerical model underwent a data mining process to create a learning database for the machine learning model. The model underwent training with seven different algorithms systematically, aiming to identify those capable of capturing the data trends. The model was specifically trained to predict both the peak particle velocity and velocity in the frequency domain.

The study revealed that artificial neural networks and extreme boosts are suitable matches for

addressing the phenomenon. However, as the complexity level increased, XGBoost demonstrated a greater understanding of the data trends than artificial neural networks. The machine learning model was further trained on additional databases to predict ground vibrations induced by the pile driving process when employing an open trench as a mitigation measure. The generated model was then implemented as a standalone computer application named 'ZBZBAT' and register it as a copyright. Finally, an experimental validation process was conducted to assess the effectiveness of the generated models. As well, a website is developed for the ZBZBAT software.

6.3 Future work

The research presented in this thesis and the accompanying papers have focused on the efficacy of machine learning models in capturing vibrations induced by the pile driving process, taking into account soil characteristics, pile geometry, and hammering characteristics. Conclusions drawn from this research, along with identified areas for further exploration, have led to several suggestions for future studies.

- ***FEM-PML improving opportunity***: there is potential to expand the Finite Element Method with Perfectly Matched Layer (FEM-PML) to incorporate other techniques of pile driving as vibratory pile driving. This enhancement would contribute to a more comprehensive understanding of the numerical model.
- ***Vibration in nearby building***: an extension of this study could encompass the examination of vibrations affecting nearby structures. By broadening the scope to include the impact on adjacent structures, the research could provide valuable insights into the broader implications of pile driving activities on the surrounding environment.
- ***Offshore application***: the findings of this study hold promise for application in offshore monopile driving, where the study of vibrations on the seabed can offer insights into their impact on marine life. This extension of the study could prove invaluable in understanding and mitigating potential environmental effects associated with offshore construction activities.

In summary, these suggestions for further research aim to build upon the current findings and advance the understanding of machine learning applications in predicting vibrations induced by pile driving, with implications for offshore construction and structural dynamics in diverse settings.

Bibliography

[1] M. Roser and H. Ritchie. World population growth. *Our World in Data*, 2013.

[2] A. Hamidi, A. Rooz, and M. Pourjenabi. Allowable distance from impact pile driving to prevent structural damage considering limits in different standards. *Practice Periodical on Structural Design and Construction*, 23(1):04017029, 2018.

[3] J. Dungca, D. Acosta, M. Juego, and I S Sanchez, H.and Sanchez. The propagation behavior of pile-driving-induced vibration done on soil at varying distances and its effects on existing structures. *International Journal*, 10(21):1877–1883, 2016.

[4] A. Colaço, P. Costa, C. Parente, and A. Cardoso. Ground-borne noise and vibrations in buildings induced by pile driving: An integrated approach. *Applied Acoustics*, 179:108059, 2021.

[5] MATLAB. *version 23.2.0 (R2023a)*. The MathWorks Inc., Natick, Massachusetts, 2023.

[6] M. Fernandes. *Analysis and design of geotechnical structures*. Taylor and Francis Group, 1 edition, 2020.

[7] M. Xiao. *Geotechnical engineering design*. John Wiley and Sons, 2015.

[8] FHWA. Implementation of AASHTO LRFD Design Specifications for Driven Piles State DOT LRFD Design Guidance for Driven Piles. Technical report, Federal Highway Administration, Nov 2013.

[9] B. Grand. Types of piles: their characteristics and general use. *Highway Research Record*, 1(333), 1970.

[10] A. Gkrizi, A. Athanasopoulos, and R. Woods. Understanding the energy transfer mechanism in the near field of impact driven piles. *10th International Conference on Stress Wave Theory and Testing Methods for Deep Foundations*, 2018.

[11] K. Massarsch. Effects of vibratory compaction. In *International Conference on Vibratory Pile Driving and Deep Soil Compaction, Keynote Lecture*, pages 33–42, 2002.

[12] A. Zekkos, R. Woods, A. Grizi, et al. Effect of pile-driving induced vibrations on nearby structures and other assets. Technical report, Michigan. Dept. of Transportation. Office of Research and Best Practices, 2013.

[13] K. Massarsch. Keynote lecture: Static and dynamic soil displacements caused by pile driving. In *Proceedings of the Fourth International Conference on the Application of Stress-Wave Theory to Piles, The Hague, The Netherlands*, pages 15–24, 1992.

[14] M. Das. *Fundamentals of soil dynamics*. Vieweg, 1983.

[15] P. Auning, C. Lomnitz, F. Wuttke, and R. Saragoni. Prograde rayleigh-wave motion in the valley of mexico. *Geofísica internacional*, 45(3):149–162, 2006.

[16] T. Tanimoto and L. Rivera. Prograde rayleigh wave particle motion. *Geophysical Journal International*, 162(2):399–405, 2005.

[17] F. Deckner. *Ground vibrations due to pile and sheet pile driving: influencing factors, predictions and measurements*. PhD thesis, KTH Royal Institute of Technology, 2013.

[18] A. Bodare. Kompendium œ jord-och bergdynamik. *KTH, Stockholm*, 1996.

[19] S. Kramer. *Geotechnical earthquake engineering*. Pearson Education India, 1996.

[20] G. Athanasopoulos and P. Pelekis. Ground vibrations from sheetpile driving in urban environment: measurements, analysis and effects on buildings and occupants. *Soil dynamics and earthquake engineering*, 19(5):371–387, 2000.

[21] R. Woods. *Dynamic effects of pile installations on adjacent structures*, volume 253. Transportation Research Board, 1997.

[22] C. Ramshaw. *Computation of ground waves from pile driving and their effects on structures*. PhD thesis, Durham University, 2002.

[23] G. Bomitz. About the propagation of the ground vibrations generated by large-piston machines into the depth. *Springer, B*, 1931.

[24] W. Haupt. *Soil dynamics: basics and application*. Vieweg, 1986.

[25] A. Uromeihy. *Ground vibration measurements with special reference to pile driving*. PhD thesis, Durham University, 1990.

[26] M. Mabsout and J. Tassoulas. A finite element model for the simulation of pile driving. *International Journal for numerical methods in Engineering*, 37(2):257–278, 1994.

[27] T. Thandavamoorthy. Piling in fine and medium sand—a case study of ground and pile vibration. *Soil Dynamics and Earthquake Engineering*, 24(4):295–304, 2004.

[28] B. Elebi and Vural. Field experiments on wave propagation and vibration isolation by using wave barriers. *Soil dynamics and earthquake engineering*, 29(5):824–833, 2009.

[29] S. Ekanayake, D. Liyanapathirana, and C. Leo. Influence zone around a closed-ended pile during vibratory driving. *Soil Dynamics and Earthquake Engineering*, 53:26–36, 2013.

[30] A. Longinow. Discussion of "tolerable limits of construction vibrations. *Practice Periodical on Structural Design and Construction*, 21(4):07016003, 2016.

[31] M. Pourjenabia and A. Hamidi. Numerical modeling of dynamic compaction process in dry sands considering critical distance from adjacent structures. *Structural engineering and mechanics: An international journal*, 56(1):49–56, 2015.

[32] M. Rezaei and A. Hamidi, A.and Rooz. Investigation of peak particle velocity variations during impact pile driving process. *Civil Engineering Infrastructures Journal*, 49(1):59–69, 2016.

[33] BM. New. Ground vibration caused by civil engineering works. Technical report, Sciences Engineering Medicine, 1986.

[34] K. Massarsch and B. Broms. Damage criteria for small amplitude ground vibrations. *International conferences on recent advances in geotechnical earthquake engineering and soil symanic*, 1991.

[35] S. Jones. Transportation and construction-induced vibration guidance manual. *Sacramento, Calif: Calif Department of Transport, Noise, Vibration, and Hazardous Waste Management Office*, 2004.

[36] M. Svinkin. Minimizing construction vibration effects. *Practice periodical on structural design and construction*, 9(2):108–115, 2004.

[37] E. Abdel-Rasoul and M. Mohamed. Measurement and evaluation of blasting ground vibrations and airblasts at the limestone quarries of assiut cement company (cemex). *J. Eng. Sci. Assiut Univ*, 34(4):1293–1309, 2006.

[38] J. Head and F. Jardine. *Ground-borne vibrations arising from piling*. Number 142 in 142. Public Works Research Institute, Japan, 1992.

[39] D. Rockhill, M. Bolton, and D. White. Ground-borne vibrations due to press-in piling operations. In *BGA International Conference on Foundations: Innovations, observations, design and practice: Proceedings of the international conference organised by British Geotechnical Association and held in Dundee, Scotland on 2–5th September 2003*, pages 743–755. Thomas Telford Publishing, 2003.

[40] D. Hiller and V. Hope. Groundborne vibration generated by mechanized construction activities. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 131(4):223–232, 1998.

[41] J. O'Connor, W. O'Brien, and J. Choi. Industrial project execution planning: Modularization versus stick-built. *Practice periodical on structural design and construction*, 21(1):04015014, 2016.

[42] P. Attewell and I. Farmer. Modern Piling: Part Two-Attenuation of Ground Vibrations from Pile Driving. *Ground engineering*, 1973.

[43] D. Martin. Ground vibrations from impact pile driving during road construction. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 1980.

[44] A. Ahlqvist and E. Enggren. *Impact on surrounding environment from vibro driven sheet piles*. American Society of Civil Engineers, 2006.

[45] V. Whenham, L. Areias, F. Rocher-Lacoste, D. Vié, M. Bourdouxhe, and A. Holeyman. Full scale sheet pile vibro-driving tests. In *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering (Volumes 1, 2, 3 and 4)*, pages 1354–1357. IOS Press, 2009.

[46] M. Serdaroglu. *Nonlinear analysis of pile driving and ground vibrations in saturated cohesive soils using the finite element method*. PhD thesis, University of Iowa, 2010.

[47] B. Schümann and J. Grabe. Fe-based modelling of pile driving in saturated soils. In *Proceedings of the 8th international conference on structural dynamics, EURODYN*, pages 894–900, 2011.

[48] J. Wiss. Damage effects of pile driving vibration. *Highway Research Record*, 141(155), 1967.

[49] P. Attewell, A. Selby, and L. O'Donnell. Tables and graphs for the estimation of ground vibration from driven piling operations. *Geotechnical and Geological Engineering*, 10(1):61–85, 1992.

[50] A. Quagliata, M. Ahearn, E. Boeker, C. Roof, L. Meister, H. Singleton, et al. Transit noise and vibration impact assessment manual [2018]. Technical report, United States. Federal Transit Administration. Office of Planning and ..., 2018.

[51] P. Hannigan, F. Rausche, L. Garland, B. Robinson, M. Becker, and H. Shelsta. *Design and construction of driven pile foundations*. US Department of Transportation, Federal Highway Administration, 2016.

[52] J. Bay. A summary of the research on pile driving vibrations. *Proceeding Of 7th Pile Driving Contractors Association Annual Winter Roundtable, Atlanta, GA*, 2003.

[53] A. Colaço, M. Ferreira, and P. Costa. Empirical, experimental and numerical prediction of ground-borne vibrations induced by impact pile driving. *Vibration*, 5(1):80–95, 2022.

[54] C. Madheswaran, R. Sundaravadivelu, A. Boominathan, and K. Natarajan. Response of ground during pile driving. *Journal of the Institution of Engineers(India), Part CV, Civil Engineering Division*, 86:22–27, 2005.

[55] H. Masoumi, G. Degrande, and G. Lombaert. Prediction of free field vibrations due to pile driving using a dynamic soil–structure interaction formulation. *Soil Dynamics and Earthquake Engineering*, 27(2):126–143, 2007.

[56] H. Masoumi, S. François, and G. Degrande. A non-linear coupled finite element–boundary element model for the prediction of vibrations due to vibratory and impact pile driving. *International Journal for Numerical and Analytical Methods in Geomechanics*, 33(2):245–274, 2009.

[57] A. Khoubani and M. Ahmadi. Numerical study of ground vibration due to impact pile driving. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 167(1):28–39, 2014.

[58] A. Rooz and A. Hamidi. A numerical model for continuous impact pile driving using ale adaptive mesh method. *Soil Dynamics and Earthquake Engineering*, 118:134–143, 2019.

[59] T. Sofiste, L. Godinho, P. Costa, and D. Soares. An effective time domain numerical model for the prediction of ground-borne vibrations induced by pile driving. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 261, pages 5442–5452. Institute of Noise Control Engineering, 2020.

[60] T. Sofiste, L. Godinho, P. Costa, D. Soares, and A. Colaço. Numerical modelling for prediction of ground-borne vibrations induced by pile driving. *Engineering Structures*, 242:112533, 2021.

[61] Rix. J. *Dynamic Soil-Structure Interaction*. Geosyntec Consultants, Inc. Atlanta, Georgia, USA, 1 edition, 2015.

[62] M. Monjezi, M. Ghafurikalajahi, and A. Bahrami. Prediction of blast-induced ground vibration using artificial neural networks. *Tunnelling and Underground Space Technology*, 26(1):46–50, 2011.

[63] T. Singh and V. Singh. An intelligent approach to prediction and control ground vibration in mines. *Geotechnical and Geological Engineering*, 23(3):249–262, 2005.

[64] S. Nazari, S. Saljooghi, M. Shahbazi, and H. Akbari. Application of artificial neural networks in solving inversion problem of surface wave method on pavements. In *Second international conference on engineering optimization. Lisbon, Portugal*, pages 1–8, 2010.

[65] C. Hung and S. Ni. Using multiple neural networks to estimate the screening effect of surface waves by in-filled trenches. *Computers and Geotechnics*, 34(5):397–409, 2007.

[66] D. Li, A. Meddah, K. Hass, and S. Kalay. Relating track geometry to vehicle performance using neural network approach. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 220(3):273–281, 2006.

[67] D. Connolly, G. Kouroussis, A. Giannopoulos, O. Verlinden, P. Woodward, and M. Forde. Assessment of railway vibrations using an efficient scoping model. *Soil Dynamics and Earthquake Engineering*, 58:37–47, 2014.

[68] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[69] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński, and U. Amjad. Prediction of pile bearing capacity using xgboost algorithm: modeling and performance evaluation. *Applied Sciences*, 12(4):2126, 2022.

[70] T. Nguyen, D. Ly, T. Huynh, and T. Nguyen. Soft computing for determining base resistance of super-long piles in soft soil: A coupled spbo-xgboost approach. *Computers and Geotechnics*, 162:105707, 2023.

[71] M. Kamran. Predictive modeling of slope stability using hybrid machine learning approaches: Pca-gmm based woa-xgboost. *Computers and Geotechnics*, 2023.

[72] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.

[73] S. Johnson. Notes on perfectly matched layers, online mit course notes, 2007.

[74] P. Lopes, P. Costa, M. Ferraz, R. Calçada, and A. Cardoso. Numerical modeling of vibrations induced by railway traffic in tunnels: From the source to the nearby buildings. *Soil Dynamics and Earthquake Engineering*, 61:269–285, 2014.

[75] P. Lopes, P. Costa, R. Calçada, and A. Cardoso. Numerical modeling of vibrations induced in tunnels: a 2.5 d fem-pml approach. *Traffic induced environmental vibrations and controls: theory and application*, pages 133–166, 2013.

[76] B. Hardin and V. Drnevich. Shear modulus and damping in soils: measurement and parameter effects. *Journal of the soil mechanics and foundations division*, 98(6):603–624, 1972.

[77] M. Vucetic and R. Dobry. Effect of soil plasticity on cyclic response. *Journal of geotechnical engineering*, 117(1):89–107, 1991.

[78] Isao Ishibashi and Xinjian Zhang. Unified dynamic shear moduli and damping ratios of sand and clay. *Soils and foundations*, 33(1):182–191, 1993.

[79] J. Lysmer, T. Udaka, C. Tsai, and H. Seed. Flush-a computer program for approximate 3-d analysis of soil-structure interaction problems. Technical report, California University, 1975.

[80] A. Holeyman and C. Legrand. Soil modeling for pile vibratory driving. In *International Conference on Design and Construction of Deep Foundations, Orlando, Florida*, volume 2, pages 1165–1178, 1994.

[81] Pedro Alves Costa, Rui Calçada, António Silva Cardoso, and Anders Bodare. Influence of soil non-linearity on the dynamic response of high-speed railway tracks. *Soil Dynamics and Earthquake Engineering*, 30(4):221–235, 2010.

[82] A. Halabian and M. El Naggar. Effect of non-linear soil–structure interaction on seismic response of tall slender structures. *Soil Dynamics and Earthquake Engineering*, 22(8):639–658, 2002.

[83] H. Masoumi and G. Degrande. Numerical modeling of free field vibrations due to pile driving using a dynamic soil-structure interaction formulation. *Journal of Computational and Applied Mathematics*, 215(2):503–511, 2008.

[84] A. Deeks and M. Randolph. Analytical modelling of hammer impact for pile driving. *International Journal for Numerical and Analytical Methods in Geomechanics*, 17(5):279–302, 1993.

[85] Aires Colaço, Pedro Alves Costa, Cecília Parente, and Ahmed M Abouelmaty. Vibrations induced by a low dynamic loading on a driven pile: Numerical prediction and experimental validation. *Vibration*, 5(4):829–845, 2022.

[86] P. Lopes, J. Ruiz, P. Costa, L. Rodríguez, and A. Cardoso. Vibrations inside buildings due to subway railway traffic. experimental validation of a comprehensive prediction model. *Science of the Total Environment*, 568:1333–1343, 2016.

[87] A. Colaço, P. Costa, P. Amado-Mendes, and L. Godinho. Prediction of vibrations and reradiated noise due to railway traffic: a comprehensive hybrid model based on a finite element method and method of fundamental solutions approach. *Journal of Vibration and Acoustics*, 139(6), 2017.

[88] H. Aksoy, M. Gor, and E. İnal. A new design chart for estimating friction angle between soil and pile materials. *Geomechanics and Engineering*, 10(3):315–324, 2016.

[89] Falk and Itani. Dynamic characteristics of wood and gypsum diaphragms. *Journal of structural Engineering*, 113(6):1357–1370, 1987.

[90] Kudu, S. Ucak, and Bayraktar. Estimation of damping ratios of steel structures by operational modal analysis method. *Journal of Constructional Steel Research*, 112:61–68, 2015.

[91] Google limited liability company (LLC). Google maps, location: https://goo.gl/maps/a3cbija9e2xa7kmp6. *Google Maps directions for driving from Ingolstadt, Germany, to Geneva, Switzerland*, 2020.

[92] Civilria. Relatório Geotécnico Implantação de Edifícios Oliveirinha, Aveiro. Technical report, Gabinete de serviços Geotecnicos Lda., June 2020.

[93] A Colaço, AM Abouelmaty, and P Alves Costa. Ground-borne vibrations induced by impact pile driving: experimental assessment and mitigation measures. *Earthquake Engineering and Engineering Vibration*, 22(1):105–115, 2023.

[94] V. Krylov. Control of traffic-induced ground vibrations by placing heavy masses on the ground surface. *Journal of low frequency noise, vibration and active control*, 26(4):311–321, 2007.

[95] A. Dijckmans, P. Coulier, J. Jiang, M. Toward, D. Thompson, G. Degrande, and G. Lombaert. Mitigation of railway induced ground vibration by heavy masses next to the track. *Soil Dynamics and Earthquake Engineering*, 75:158–170, 2015.

[96] V. Krylov. Scattering of rayleigh waves by heavy masses as method of protection against traffic-induced ground vibrations. In *Environmental Vibrations: Prediction, Monitoring, Mitigation and Evaluation (ISEV 2005)*, pages 393–398. CRC Press, 2021.

[97] A. Shahin, B. Jaksa, and R. Maier. State of the art of artificial neural networks in geotechnical engineering. *Electronic Journal of Geotechnical Engineering*, 8(1):1–26, 2008.

[98] T. Masters. *Practical neural network recipes in C++*. Morgan Kaufmann, 1993.

[99] A. Shahin, H. Maier, and M. Jaksa. Data division for developing neural networks applied to geotechnical engineering. *journal of Computing in Civil Engineering*, 18(2):105–114, 2004.

[100] R. Hecht-Nielsen. On the algebraic structure of feedforward network weight spaces. In *Advanced Neural Computers*, pages 129–135. Elsevier, 1990.

[101] M. Twomey, A. Smith, et al. Validation and verification. *Artificial neural networks for civil engineers: Fundamentals and applications*, pages 44–64, 1997.

[102] T. Hastie, R. Tibshirani, J. Friedman, and H. Friedm. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[103] J. Quinlan. Induction of decision trees. *Machine learning*, 1:81–106, 1986.

[104] G. James, D. Witten, T. Hastie, R. Tibshirani, et al. *An introduction to statistical learning*, volume 112. Springer, 2013.

[105] L. Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

[106] C. Rasmussen and C. Williams. *Gaussian processes for machine learning*, volume 1. Springer, 2006.

[107] XG-Boost in Machine learning. Enjoy Algorithms. https://www.enjoyalgorithms.com/blog/xg-boost-algorithm-in-ml, 2023. [Online; accessed 11-Nov.-2023].

[108] E. Junda, C. Málaga-Chuquitaype, and K. Chawgien. Interpretable machine learning models for the estimation of seismic drifts in clt buildings. *Journal of Building Engineering*, 70:106365, 2023.